Volume 9 Number 5 1981 Nucleic Acids Research

# Locations of three repetitive sequence families found in BALB/c adult β-globin clones

Nancy L.Haigwood*, Carolyn L.Jahn, Clyde A.Hutchison III and Marshall H.Edgell

Department of Bacteriology and Immunology, and Curriculum in Genetics, The University of North Carolina, Chapel Hill, NC 27514, USA

Received 24 November 1980

## ABSTRACT

Three different repeat sequences have been mapped within the cloned EcoRI fragments that contain the adult β-globin genes from the BALB/c ($Hdd^d$) mouse. One sequence, "a", occurs 1.5-2 kb 3' to the β-major gene. A second, "b", is found 4kb 5' and 7.5kb 3' to the β-minor gene. The 14kb EcoRI fragment bearing the β-minor gene carries at least one additional repetitive element, "c". Probing a BALB/c DNA library with each repeat has demonstrated that these sequences are moderately to highly repetitive and are extensively interspersed with each other throughout the genome. In addition, repeats "a" and "b" are preferentially found in satellite and main-band DNA, respectively. The occurence of these repeats elsewhere in the β-globin cluster was demonstrated by probing the non-adult globin clones with each repeat. The arrangement of these repeats around the non-adult genes is 5'-"b"-"b"-εy-βh1-βh2-"c"-βh3- 3'. Probing the C57BL/10 ($Hbb^s$) adult gene clones with these repeats demonstrated that the distribution of these sequences in the adult region of these two haplotypes is essentially the same.

## INTRODUCTION

Inbred strains of mice are polymorphic at the β-globin locus (Hbb, We have been comparing the DNA sequence organization of the β-globin gene complexes from two such strains, BALB/c ($Hbb^d$) and C57BL/10 ($Hbb^s$), in order to determine the source of this genetic variation (2,3,4). The embryonic and adult β-globin genes are linked within a 65kb gene cluster that has been isolated as clones in phage lambda (3,5). As part of the characterization of the M. musculus β-globin haplotypes we have been interested in the distribution of non-globin sequences repeated within the β-globin cluster and in the genome. Sequences repeated within the rabbit β-globin cluster have been mapped in detail (6). A sequence on the human β-globin cluster has been found 5' to the Gγ and δ genes and 3' to the β gene, and this sequence is transcribed by RNA polymerase III in vitro (7). Heteroduplex analysis of genomic clones of mouse α- and β-major globin genes (8) revealed a common sequence about 165 base pairs long situated 1.5 to 2 kilobases (kb)

© IRL Press Limited, 1 Falconberg Court, London W1V 5FG, U.K.

to the 3' side of the coding regions. No such repetitive elements shared by β-major and β-minor have been detected by heteroduplex analyses. (9). This report describes the distribution of several additional such sequences interspersed amongst the adult, embryonic, and pseudoglobin genes in the $\underline{Hbb}^d$ and $\underline{Hbb}^s$ haplotypes.

## MATERIALS AND METHODS

### DNA Samples and Clones

Mouse DNA was prepared from livers of male BALB/c or C57BL/10 mice obtained from the Jackson Laboratory (2). Genomic DNA was prepared for 2-dimensional (2) analysis by digestion with endonuclease EcoRI and chromatography on RPC-5, followed by analytical electrophoresis of the fractions as described previously (2).

Mouse satellite and main band DNA's were a gift of S. Weaver. They were prepared from high molecular weight mouse liver DNA made according to Gross-Bellard et al. (10) and fractionated on a single silver cesium sulfate gradient (11) without prior shearing.

DNA was prepared from bacteriophage as described (12). Recombinant bacteriophage MβG2 (β-major EcoRI clone, 11), MβG3 (β-minor EcoRI clone, 5), αM9 (BALB/c a genomic globin EcoRI clone, (8) and a BALB/c genomic ribosomal clone (13) were obtained from Philip Leder, NIH. Recombinant globin clones CE14, CE17, and CE18 were previously isolated from an EcoRI partial-digestion library (gift of Mark Davis, Cal. Tech.) (3). Recombinant globin clones BA3 and BA4 were previously isolated from a C57BL/10 partial digestion EcoRI library constructed by S. Weaver, UNC (Weaver et al., manuscript in preparation). DNA from BA3 and BA4 was a gift of S. Weaver.

### Restriction Digestious, Gel Analyses, DNA Transfers

Restriction endonuclease digestious, gel analyses, and southern transfers were performed as previously described (2). DNA from phage plaques was transferred to nitrocellulose according to the method of Benton and Davis (14).

### Preparation of Probes and Hybridization

DNA fragments were prepared for use as probes by electrophoretic elution of endonuclease digested fragments from polyacrylamide gels (2) or by preparative electrophoresis (15). Fragments were assayed for purity by nick translation (below) and hybridization to Southern blots of appropriate digests of the starting material. All of the probes used here except the 14kb EcoRI fragment showed less than 2% hybridization to bands other than

self, and in most cases no cross hybridization was detected. We found that preparative electrophoresis yielded purer fragments than electrophoretic elution. Nick translation of probes and pretreatment of filters for hybridization was carried out as previously described (3). Hybridizations were carried out in 6XSSC (1XSSC=0.15M NaCL, 0.015M Na citrate), 0.5% SDS, 1X Denhardts (16) with cold DNA's and polyribonucleotides: salmon sperm DNA, E. coli DNA, yeast tRNA (Sigma), poly (rA) and poly (rC) (3).

The temperature of incubation and wash conditions varied according to the stringency desired. High stringency hybridizations were incubated at 70°C, and the filters were washed in 0.1XSSC at 52°C or in 6XSSC at 70°C with 4, 30 minute washes of 250 ml each. Low stringency hybridizations were incubated at 48°C., and filters were washed in 6XSSC at 48°C. The period of time of incubation was varied according to the length of the hybridizing sequence in the probe. All reactions were carried out to a Cot (17) sufficient to hybridize 80% of the sequences of the complexity of the probe. Filters were air dried and exposed to X-ray film (3).

Recombinant DNA materials were prepared under conditions of P3-EK2 containment according to the NIH guidelines for recombinant DNA Research which were in effect at that time. The containment for these materials have since been reduced to P1-EK2.

Data Processing

Autoradiographic data were analyzed by assigning an intensity value (I) to each hybridizing fragment by hand matching to a set of hybridized intensity standards. The accuracy of these assigments was assessed by comparing I values from an autoradiogram of an intact filter to scintillation counter data from the slices of the filter. Of the 10 hybridizing bands analyzed, all agreed easily within a factor of 1.5. Intensities were corrected for the amount of DNA analyzed and the efficiency of transfer to Southern blots. Intensities were converted to equivalent nucleotide pairs, or the extent of homology of perfectly matched sequences, by normalizing to an internal hybridization standard of known homology to the probe. The values used were: adult globin cDNA clone x β-major or β-minor = 550 nucleotides (9,18); α-genomic x β -major = 150 nucleotides (8); β-major x β-minor = 1400 nucleotides (18). All experiments were performed at low and high stringency. Numerical values represent the mean of at least two independent experiments at each stringency level.

RESULTS

Sequences in the adult clones repeated in genomic DNA

The presence of repeat sequences was first demonstrated by probing Southern blots of two-dimensional (2D) genomic spreads (RPC-5 chromatography followed by analytical gel electrophoresis) of mouse DNA with the two adult β-globin clones (β-major and β-minor) which showed that each contained sequences widely dispersed throughout the genome. In order to determine the location of these repetitive sequences within the adult clones, the clones were digested with various restriction enzymes and the fragments were isolated and used to probe the genome.

Subfractionation of the Eco RI fragment bearing the β-major gene with Xba I (map in Fig. 1) and use of the two resulting subfragments as probes to the genome indicated that all of the repetitive sequences in the 7kb EcoRI fragment were located 3' to the coding region (3 to 7 kb on Fig. 1). To localize these repeat sequences more precisely, the Hae III sites in the cloned β-major fragment were mapped (Fig. 1) and purified Hae III fragments were used individually to probe 2D spreads of genomic DNA. All the Hae III fragments with the exception of K, L, and M were used to probe spreads. The

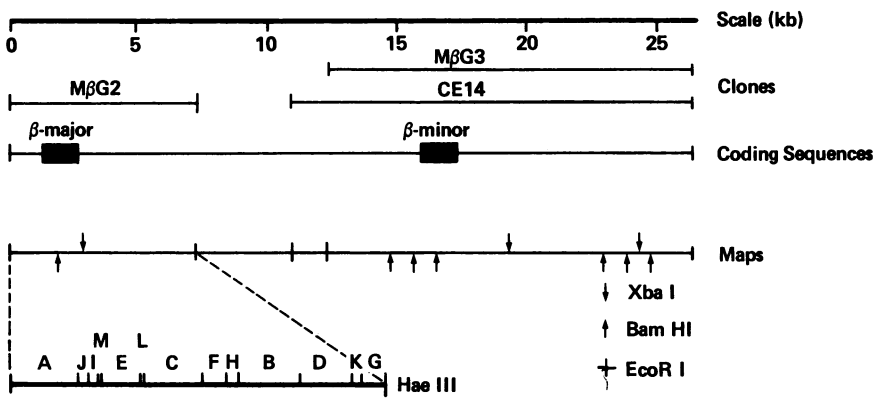


Figure 1. Map of the adult β-globin region from BALB/c. The adult globin clones used in this study are shown above a map of the region. MβG2 contains a single 7.3 kb Eco RI fragment; MβG3 contains a single 14 kb Eco RI fragment; CE14 contains an additional 1.35 kb Eco RI fragment plus the 14 kb fragment. The black boxes along the central line delineate the extent of coding sequences for the β-major and β-minor genes, where the direction of transcription is from left to right. Maps of the relevant restriction endonuclease cleavage sites are presented below, with an expanded Hae III map of the 7.3 kb β-major Eco RI fragment. Xba I and Bam HI maps of MβG2 and MβG3 were obtained from published material (5).

only Hae III fragment showing significant homology to genomic fragments in either BALB/c or C57BL/10 other than the genomic β-major fragment was Hae III fragment B (an example is shown in figure 2).

The β-minor clone was divided into subfragments in a similar fashion (see Fig. 1 for map). However, in this case each of the three XbaI and EcoRI subfragments and the 5' proximal Bam HI subfragments (14.6 to 16.5 kb on Fig. 1) hybridized to many genomic fragments (data not shown). The 1.35 kb fragment (11 to 12.3 kb on Fig. 1) linked 5' to the 14 kb Eco RI fragment bearing the β-minor gene also hybridized extensively to genomic fragments

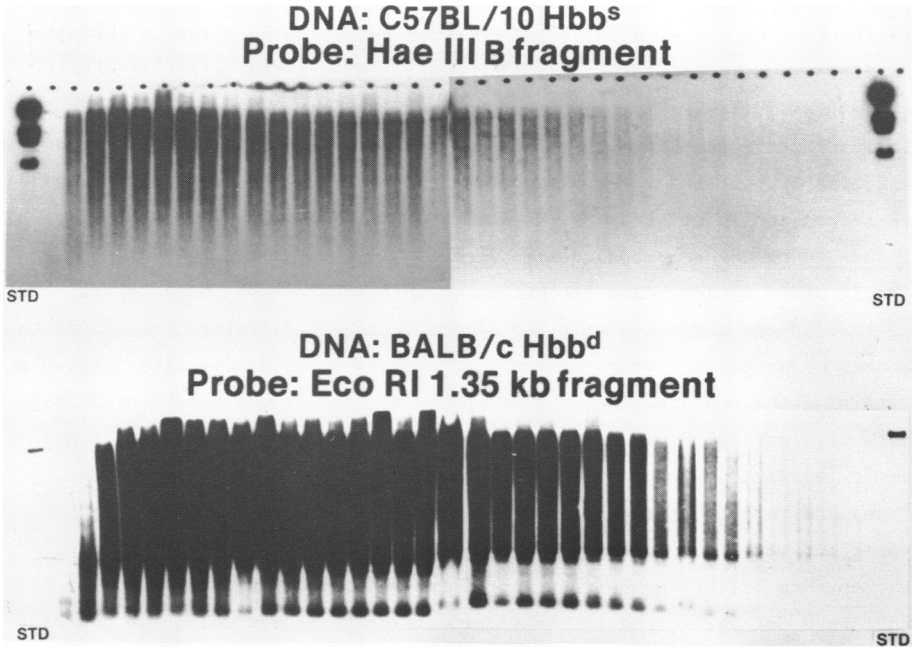


Figure 2. Identification of repeated sequences in subfragments or adjacent fragments of BALB/c adult β-globin clones. Aliquots of an RPC-5 chromatograph of an Eco RI digest of C57BL/10 DNA were applied to a horizontal 1% agarose gel 20 cm x 20 cm with two rows of sample wells. In the photograph the two rows of samples are rearranged so that RPC-5 fractions are continuous from left to right. The samples were electrophoresed, stained and photographed, transferred to nitrocellulose filters and hybridized at 70°C with the $^{32}$P nick translated globin fragments indicated. The Hae III-B fragment used in this experiment was purified from electrophoretically eluted β-major 7.3 kb fragment after subsequent digestion with Hae III and electrophoretic elution. The 1.35 kb Eco RI fragment was purified by electrophoretic elution from a gel of an Eco RI digest of CE14 DNA.

(Fig. 2). Shared homology between the β-major and β-minor clones is confined to Hae III fragments bearing coding region, as shown in cross hybridization experiments (data not shown.) Therefore these β-major and β-minor-linked repetitive elements represent at least two distinct repeat families. We chose two repetitive subfragments for further study: the Hae III B fragment from β-major (repeat "a") and the 1.35 kb EcoRI fragment from β-minor (repeat "b).

Hybridization to Satellite and Main band DNAs

As part of the investigation into the distribution throughout the genome of the repetitive elements present in the adult region of the β-globin cluster, genomic DNA was fractionated into satellite and main band preparations on a $Ag^{++}Cs_2SO_4$ gradient. The satellite preparation cut with Hae III gave the predicted ladder of fragments with the repeat length of 240 np (19). The β-major subfragment Hae III B probe hybridized preferentially with a subset of the Eco RI digested satellite bands which had a mean size of 7 kb (Fig. 3). In contrast, the 1.35 kb Eco RI fragment probe hybridized poorly if at all with satellite DNA (Fig. 3).

Dispersion of the repeat sequences in the genome

In order to determine the extent to which members of these repeat families in the adult β-globin Eco RI fragments are dispersed throughout the genome, we probed a library of Eco RI fragments in Charon 4A with fragments containing the repeats. A DNA library constructed from size-selected fragments of a partial Eco RI digest of BALB/c DNA (20) was plated out at a density sufficient to produce isolated plaques (ca. 120-180 plaques per plate) and replica filter lifts were made. Sets of filters were then hybridized with four probes (the β-major Eco RI fragment, the β-major subfragment Hae III B, the β-minor Eco RI fragment, and the 1.35 kb Eco RI fragment) to a Cot sufficient to anneal repetitive sequences only. We had previously shown that a β-globin cDNA probe detected embryonic clones in this library at the expected frequency, although only one of the adult β-globin clones was found, and its frequency of occurrence was several-fold higher than expected (3). The β-major probe hybridized to 93% of the plaques bearing inserts, subfragment Hae III B to 85%, and β-minor and 1.35 kb to 100%. These percentages have been corrected to represent the fraction of insert-bearing plaques by subtracting from the total average number of plaques the 25% shown to carry no inserts (20). This correction seems reasonable since the same 25% of the plaques in our experiments did not hybridize to any probe.

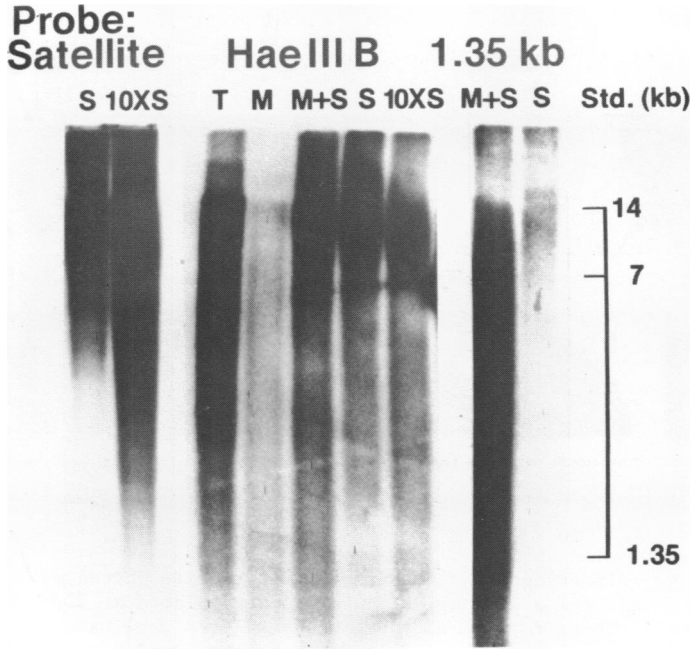


Figure 3. Satellite and main band DNA hybridization studies. Mouse density satellite DNA (S) and main band DNA (M) from the same density gradient and total mouse DNA (T) were digested to completion with Eco RI and samples were applied to 1% agarose slab gels, electrophoresed, and transferred to nitrocellulose filters. Digests of MβG2 and CE14 with Eco RI were electrophoresed in adjacent channels and transferred as size standards for hybridization (data not shown). Identical sets of filters were hybridized with the nick translated DNAs indicated at the top of the figure. S is 0.55 µg of satellite DNA; 4 x S is 2.2 µg; T is 5.5 µg of total mouse DNA; M is 5 µg of main band DNA; M + S is 5 µg of main band DNA plus 0.55 µg of satellite DNA. The Hae III-B fragment was purified from a Hae III digest of the β-major fragment purified by preparative electrophoresis; the 1.35 kb Eco RI fragment was purified as described in Figure 2.

In contrast to other probes, the 1.35 kb probe exhibited two distinct classes of hybridization intensity: a very intense class, 36%, and a markedly less intense class, 54% (Fig. 4). Variability in hybridization intensities could result from different numbers of copies of a well matched sequence in a probe or from a variation in the degree of mismatch to a given probe, or from a combination of the two. Control plaques of known homology to the probes exhibited the expected differences in hybridization intensities and set the limits of detectability.

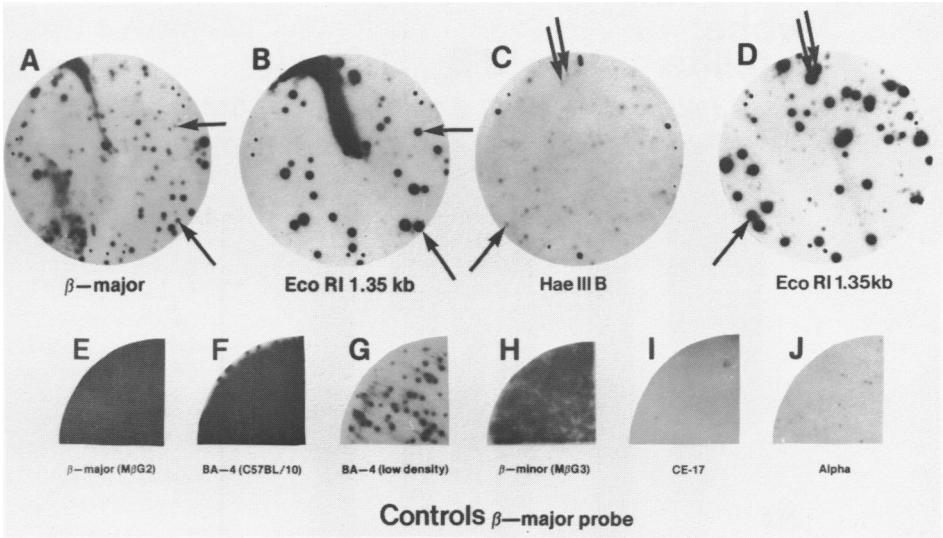


Figure 4. Dispersion and interspersion of repeats in the mouse genome. Recombinant phage from a library of mouse DNA were plated at 120-180 pfu per dish and sets of three replica nitrocellulose filters were made and marked for alignment. Mixtures of these filters were hybridized with the β-major 7.3 kb Eco RI fragment (A), the β-major Hae III-B subfragment (C), the 1.35 kb Eco RI fragment adjacent to β-minor (B and D), and the β-minor 14 kb Eco RI fragment (data not shown). After hybridization, autoradiograms were compared to the original plates and scored individually for percent positive plaques. Some replica filters were used to test reproducibility of transfer of DNA, and no differences were seen in the three lifts tested. To compare hybridization between two probes, autoradiograms were overlaid using alignment marks. Two examples of plaques which hybridize to both the β-major Eco RI fragment and the 1.35 kb Eco RI fragment are shown in replica filters A and B (arrows). Three examples of plaques which hybridize to β-major Hae III-B and the 1.35 kb fragment are denoted in replica filters C and D (arrows). Filter lifts were also made from characterized globin phage from the BALB/c adult β-globin region (MβG2, MβG3, CE14), the BALB/c embryonic region (CE17, CE18), the C57BL/10 adult region (BA3, BA4), an alpha globin clone (αM9), a nonrecombinant phage (λgtWES.B), and a blank filter. These filters were included in all hybridizations to monitor phage vector contamination, which was never detected, and to test for globin homologies. Six of the ten standards which were hybridized and exposed for autoradiography with filter A are shown in E-J. Probes Hae III-B and the 1.35 kb fragment were prepared as described in Figure 2. The 7.3 kb β-major fragment was purified by preparative electrophoresis; the 14 kb β-minor fragment was prepared by electrophoretic elution.

Homologies of 165 nucleotides (the homology 3' to the coding region shared by the β- and α-globin clones) was barely detected in these experiments (Fig. 4). Homologies of 550 nucleotides (the extent of globin coding

region homology) were easily detected as seen in the hybridization of the β-major probe to the β-minor plaques (Fig. 4H). These controls also demonstrated consistency in plaque size and signal which was relatively independent of the plating density (Fig. 4, F and G). The intensity standards in hybridizations with the 1.35 kb probe indicated that some of the characterized globin phage hybridized at the weaker of the two intensity levels seen with this probe (data not shown).

Sequences in the Adult Clones Repeated Within the Beta Cluster

The 14 kb Eco RI fragment bearing β-minor purified from MβG3 (which does not contain the 1.35 kb fragment) hybridized strongly to the 1.35 kb fragment (Fig. 5). In the reciprocal hybridization, the 1.35 kb fragment hybridized strongly to the 2 kb 3' terminal fragment of an Eco RI plus Xba I digest of the β-minor 14 kb Eco RI fragment (Fig. 5). This demonstrates that an element in the 1.35 kb fragment is repeated within the adult region of the BALB/c cluster, both 5' and 3' to the gene, and the approximate location of these sequences is known.

Five other structures homologous to the β-globin cDNA make up the characterized portion of the BALB/c β-globin cluster (3). The 32 kb stetch of DNA containing these structures will be called for identification purposes the "embryonic" region since the only known protein coded for by these structures is the embryonic globin, εy. In order to test the embryonic region for sequences shared with the adult region, four fragments purified from the adult region were used as probes to Southern blots of restriction digests of clones from the embryonic region. The probes were the β-major Eco RI fragment, the β-major subfragment Hae III B, the 1.35 kb Eco RI fragment and the β-minor Eco RI fragment. A β-globin cDNA clone was also used to probe the embryonic region to monitor homology due to coding region sequences. Figure 6 shows representative Southern blots with a schematic summary of the hybridization data below a map of the region.

Hybridization with these probes detected three embryonic region Eco RI fragments (W,X, and S) with non-globin coding homology to the adult probe (Fig. 6). Fragments W and X share sequences with both the β-minor and the 1.35 kb fragment. Fragment S shares sequences with β-major, β-major subfragment Hae III B, and β-minor. Fragment V shares sequences with β-minor at low and high stringency conditions and with the β-globin cDNA at low stringency. The β-minor hybridization to V does not appear to be predominantly due to the coding sequences in β-minor because of (a) the stringency differentiation, (b) the known sequence divergence between the adult and

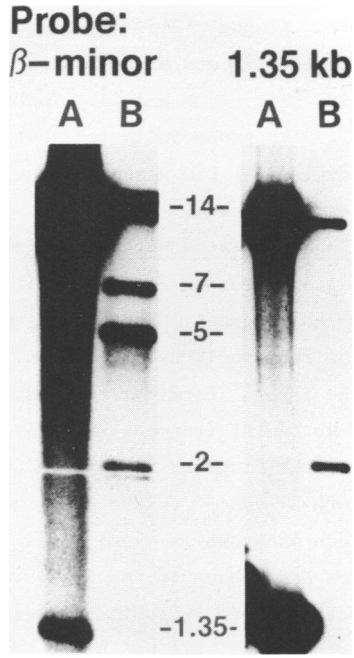


Figure 5. Identification of a repeat sequence 5' and 3' to β-minor. DNA from recombinant globin clone CE14 was digested with Eco RI, yielding the 14 kb and 1.35 kb insert fragments shown in the lanes marked A. MβG3 was digested with Xba I + Eco RI to yield the 7 kb, 5 kb, and 2 kb products of complete digestion plus a partial digestion product of 14 kb shown in the lanes marked B. Sizes of the fragments (in kb) are indicated in the center of the two panels. The DNA was electrophoresed on 1% agarose slab gels, transferred to nitrocellulose filters, and hybridized with the nick translated 14 kb or 1.35 kb fragments. The 14 kb β-minor fragment was purified by electrophoretic elution from a gel containing a digest of MβG3, which does not contain the 1.35 kb fragment; the 1.35 kb fragment was purified from CE14 by electrophoretic elution. The filters were cut for hybridization; the two pieces are evident in the β-minor samples.

embryonic coding sequences, and (c) β-minor hybridization to fragment V at a level ten-fold greater than to any other fragment bearing coding sequences in the embryonic region (Fig. 6). The repeat element shared by V and β-minor is called "c".

Extent of Homology Between Repeats

Under probe-saturating conditions two features of the repeat sequence being probed determine the intensity of hybridization: the length of the homologous sequence and the degree of homology. Hybridization intensities

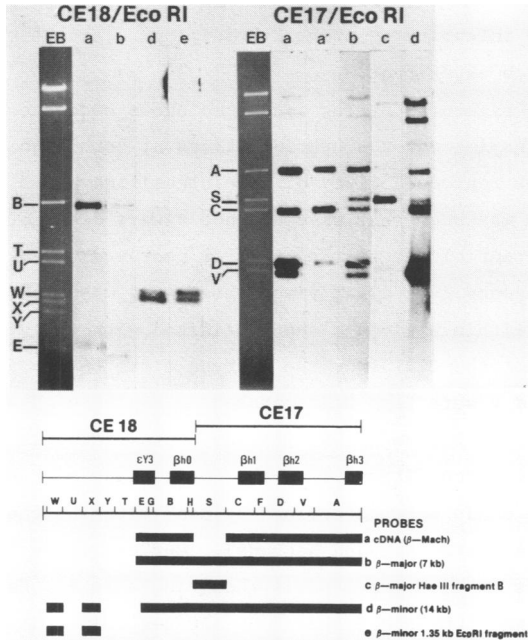


Figure 6. Hybridization of adult probes to embryonic region clones. Clones CE17 and CE18 span a 32 kb region diagrammed schematically in the lower portion of the figure. The β-globin coding regions are delineated with labeled black boxes above an Eco RI map of this region derived previously (3). CE17 and CE18 DNAs were digested with Eco RI, and identical portions of these digests were applied to 1% agarose gels, electrophoresed, stained with ethidium bromide and photographed. Examples are in the lanes marked EB. The two largest unlabeled bands in the photographs are the arms of the vector, Charon 4A. The insert bands have been labeled according to the Eco RI map shown. Fragments F, G, and H are not shown here. The DNA on the gels was transferred to nitrocellulose and hybridized with the probes listed in the lower portion of the figure. The lanes adjacent to the ethidium bromide stains are representative blots from these experiments, keyed by letter to indicate the probe used. Intensities of hybridization are not comparable except within individual lanes. A different gel was used for each probe, and variations in fragment migration have been corrected in the reproduction so that fragments are aligned. Not all of the results can be seen in the exposures shown here. The qualitative results are summarized in the schematic diagram below the Eco RI map, where the presence of a bar below an Eco RI fragment indicates that the probe hybridizes to it. Apparent hybridization of cDNA to fragment T in CE18 (lane a) is an artifact of filter cutting. Lane a' in CE17 is a blot hybridized at the higher stringency level. The β-globin cDNA fragment was purified by electrophorectic elution of the Hha I fragment containing the insert. The 7.3 kb Eco RI fragment bearing the β-major coding region was purified by preparative electrophoresis. This purified fragment was digested with Hae III, and the B fragment was purified by electrophoretic elution. The 14 kb and 1.35 kb Eco RI fragments were purified as described in Figure 5.

were converted to "equivalent nucleotides" (length assuming perfect homology) by using the intensity of probe hybridizing to itself as the normalizing standard in each experiment.

In principle fragments bearing less homologous sequences should be identifiable by carrying out the hybridizations at both high and low stringency. However, the reproducibility of our intensities normalized to equivalent nucleotides was such that we could not always discriminate between fragments in different clones bearing well or poorly matched sequences. Variation in stringency gave clear visual assays of mismatch in some comparisons of fragments within a given clone hybridized with the same probe. For example, analysis of the CE17 from the embryonic region demonstrated that the fragments D and V were less well matched to β-globin cDNA than are fragments A and C. In Figure 6, DNA in lane a under CE17 is hybridized under lower stringency conditions than the DNA in lane a'. Hybridization intensities of the bands were determined by hand matching the bands to intensity standards (Materials and Methods). Our estimates for the sizes of repetitive elements in each fragment is shown in Table 1. Estimates from these hybridizations of homology shared by the β-globin cDNA probe and the embryonic and adult region coding regions are included in Table I as examples of the reproducibility and accuracy of the method of estimation.

Comparison of the Adult Regions From BALB/c and C57BL/10

C57B1/10 mice contains two adult β-globin genes, β-s, and β-t, detected by cDNA hybridization to genomic DNA (2, S. Weaver, personal communication). This adult region from C57B1/10 has been isolated from a C57B1/10 library (Weaver, et al., manuscript in preparation. Results of hybridizations to Southern blots of digests of two clones from this region, BA3 and BA4, with the four probes from the adult region of BALB/c and β-globin cDNA demonstrated that repeats "a" and "b" are found in an arrangement around βs and βt that is identical to their arrangement with respect to β-major and β-minor (as summarized in Figure 7).

The β-major subfragment Hae III B hybridized very well to the β-s clone on the 3' side of the β-s gene and minimally if at all to the 5' side of β-s or to the β-t clone. The alpha globin clone also hybridized preferentially to the β-s clone: 100 equivalent np; which is the same extent to which the alpha clone hybridized to β-major (data not shown). The β-minor linked 1.35 kb Eco RI fragment hybridized to at least two locations in the β-t clone and minimally if at all to the β-s clone.

Hybridization of β-major to the β-s clone is 6 - 7 fold greater as

TABLE I. ESTIMATES OF LENGTHS OF REPEAT SEQUENCES AND β-GLOBIN HOMOLOGIES IN BALB/c

| REPEAT/β-GENE | "b" | "b" | εy | βh0 | "a" | βh1 | βh2 | "c" | βh3 | βmaj | "a" | "b" | βmin | "c" | "b" |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Fragment(s) | Eco W | Eco X | Eco E+G | Eco B+H | Eco S | Eco C+F | Eco D+V | Eco V | Eco A | Eco 7.3 | Hae 1.3 | Eco 1.35 | Eco 14 | Eco 14 | Xba 2 |
| Low stringency | 300 | 300 | nd | 19 | 18 | 38 | 53 | 1870 | 30 | 550* | 1300* | 1350* | nd | nd | 1700 |
| High stringency | 300 | 300 | nd | 16 | 0 | 38 | 4 | 1130 | 18 | 550* | 1300* | 1360* | 581 | nd | 1700 |

*standardizing values

The top row lists the repetitive sequence elements and the β-globin coding sequences in the order that they have been mapped in the 60 kb of the cloned BALB/c DNA analyzed. The second row indicates the fragment(s) from which measurements were obtained to estimate the homology. The values for equivalent nucleotide pairs of homology (length assuming perfect match) were obtained by normalizing the intensity of hybridization to an internal hybridization standard of known homology to the probe, as described in Materials and Methods. The standards are indicated by "*" in the Table: the 1.35 kb fragment ("b") is 1350 in equivalent nucleotides; the 1.3 kb Hae III B fragment ("a") is 1300; the "c" repeat was standardized using the 14kb Eco RI fragment hybridization to itself. Numbers of equivalent nucleotide pairs of homology to cDNA which were standardized using β-major X cDNA hybridization= 550 nucleotide pairs. nd=not determined

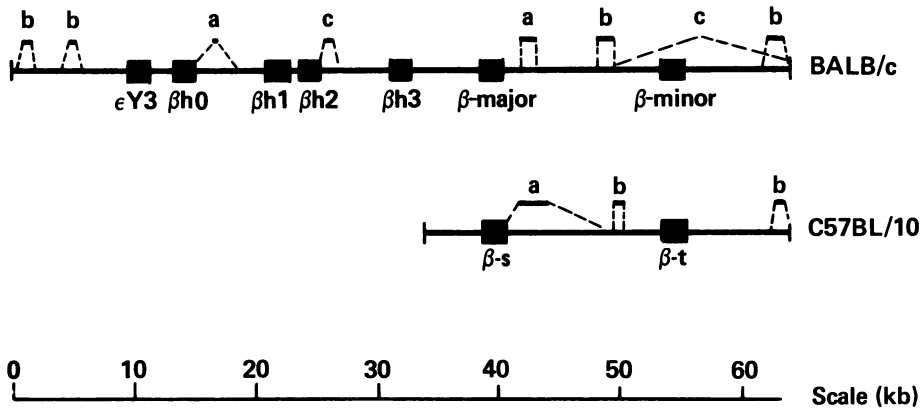


Figure 7. Summary of repeat element location in the globin clusters. β-globin genes are delineated as black boxes in the BALB/c (top line) and C57Bl/10 (middle line) regions of the cluster studied here. Known locations of homology to repeat elements "a", "b", and "c" from the adult region of BALB/c are marked with the identifying letters. The estimated sizes (in equivalent nucleotides) are represented by the length of the bars below the letters identifying the repeats. The regions bounded by the dotted lines are the areas in which the various homologies are contained.

measured in equivalent nucleotides than to the clone bearing β-t, where hybidization is limited to fragments bearing coding regions. A qualitatively similar result was obtained with 2D genomic spreads of C57BL/10 DNA probed with the β-major probe (data not shown). In addition, BA4 phage hybridized nearly as well to the β-major probe as did MβG2 phage (Fig. 4, E and F), while BA3 phage hybridized to the extent that MβG3 (β-minor) phage hybridized (data not shown). In analyses of cloned fragments, hybridization of β-minor to the β-t clone is two fold greater than to the β-s clone. β-minor hybridization to β-t spans 14.5 kb of the 15.9 kb clone (11 adjacent fragments); homology between β-minor and β-s is confined to the two fragments which bear coding sequences.

## DISCUSSION

The DNA fragments bearing the adult globin genes carry several middle to highly repetitive elements which are repetitive both within the genome and the 65 kb β-globin cluster. This is a common feature of eucaryotic gene organization which has been seen for example in Drosophilia (21), sea urchin (22), rabbit (6), and human (23). Twenty-five percent of the middle repeti-

tive sequence elements are estimated to be clustered while the remaining seventy-five percent are dispersed throughout the "unique" sequences in the eucaryotic genome (24). The studies presented here were designed (a) to identify the location of some of these elements in the adult region of the β-globin cluster at a scale sufficient to direct subsequent sequence analyses; (b) to determine the minimum number of repeat elements present in the adult β-globin region; and (c) to test other regions of the globin cluster and the genome for the presence of these elements.

Subfractionation of the adult globin clones and use of these fragments as probes demonstrated the presence of at least three non-homologous sequences which are moderately to highly repetitive in the genome, one ("a"), near β-major and at least two ("b" and "c") near β-minor. One of these ("b") is located both 5' and 3' to the β-minor coding sequence. None of these sequences hybridized to a mouse ribosomal DNA clone (13) bearing 18S and 28S sequences (data not shown).

We interpret the preferential hybridization of the β-major Hae III B fragment to satellite DNA to be indicative of a high proportion of (A+T)-rich sequences near the β-major gene which causes it to copurify with satellite DNA. The (A+T)-rich sequences probably reside in Hae III B, since the only region within the Eco RI fragment bearing β-major which hybridizes to the satellite DNA is the Hae III B fragment (data not shown.) It seems unlikely that the simple repetitive element in mouse satellite DNA is present in the Hae III B fragment since the pattern of hybridization of Hae III B to satellite DNA is different from satellite DNA hybridization to itself (Fig. 3). Hence Hae III B cannot contain a sequence present in the bulk of satellite DNA. The repetitive element "b" from the 1.35 kb fragment hybridized preferentially to main band DNA. Repeat elements in other systems have also been shown to hybridize specifically to main band (25).

We estimate that sequences homologous to "a" and "b" are dispersed to approximately $2X10^5$ sites in the genome and extensively interspersed since sequences homologous to both probes occur at least once in essentially all the clones of a DNA library of Eco RI fragments (10-20 kb insert size). If these probes each contain a single repeat family, then the members of these families are not over-represented in the β-globin cluster, and it seems unlikely that these elements could serve a globin-specific function. Of course, the complexity of these two repeat-containing probes is such that globin-specific sequences could be a subset within the repetitive elements.

At least three adult region repeat element probes have homology to

fragments in the embryonic region, and portions of at least two of these repeat elements ("a" and "b") are also present in the C57BL/10 adult region in approximately the same locations as in BALB/c (Fig. 7). The $\underline{Hbb}^{s}$ mouse β-globin cluster has been shown by cross blotting experiments to be very homologous to the $\underline{Hbb}^{d}$ cluster (Weaver et al., manuscript in preparation). The greater hybridization observed between the β-major and β-s and between β-minor and β-t suggests that β-s is embedded in an environment homologous to β-major and that β-t is embedded in an environment homologous to beta-minor.

The extent of homology estimated between β-major Hae III and the β-s clone suggests that the entire Hae III fragment is conserved, while only a very short region of homology is seen between Hae III B and the embryonic fragment S. Homology to "a" may indeed be different in fragment S and in the β-s clone. The extents of homology between 1.35 kb probe and fragments in the embryonic and adult regions are sufficiently similar to postulate that they are the same repeat element. Sequence analysis in embryonic region Eco RI fragment X and in the 1.35 kb Eco RI fragment containing repeat "b" has confirmed the homology detected by the DNA hybridization data presented here. (Voliva et al., manuscript in preparation).

The embryonic and adult β-globin genes of the mouse are temporally controlled as well as coordinately and, in at least two haplotypes ($\underline{Hbb}^{d}$ and $\underline{Hbb}^{p}$), differentially expressed. The identification of sequences common to the embryonic and adult β-globin regions, for example, may be useful to direct biological studies of these sequences and their possible role in gene expression. One postulated function of moderately repetitive sequences is as posttranscriptional controlling elements (26), since many moderately repetitive sequences are transcribed into nuclear RNA (27,28), and RNA complementary to one family of sequences, the human AluI family, has been found hydrogen bonded to nuclear RNA (29). Given the complexity of eukaryotic RNA production, it is reasonable to consider regulatory functions for shared sequences in locations other than 5' to the start of the message. Examples of such regulatory elements are the central region of 55 RNA (30,31) and an element in human β-globins 3' to the γ genes which affects expression of γ genes in the HPFH (32). Numerous roles can be postulated for functional sequences which would fall within the middle repetitive class e.g. roles specific to the maintenance, structure and replication of the chromosome. These sequences may of course be serving no current function.

The beta cluster is commonly thought to have arisen from a series of

gene duplications. If the repetitive sequence elements repeated in the β-globin cluster were present in the ancestral sequence which was duplicated and maintained, one might expect to see some regularity or order in the distribution of those sequences which would correlate with the order of the globin homologus sequences. The sequence elements examined in this study do not appear to be correlated with the globin sequences in any regular fashion. Therefore the process of gene duplication alone is not sufficient to explain the distribution of these repetitive elements in the β-globin cluster.

ACKNOWLEDGEMENTS

This work was supported by grants to M.H.E. and C.A.H. III from the U.S.P.H.S. (AI08998 and GM 21313). C.L.J. also received support from a training grant to the U.N.C. curriculum in Genetics (GM07121).

We thank Steven Weaver for generous gifts of DNA and for mapping information for the adult globin clones from C57BL/10 and Dan DiMaio and John Morrow for critical reading of the manuscript.

*Current address: Microbiology Department, The Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

REFERENCES

1. Stern, R.H., Russell, E.S. and Taylor, B.A. (1976) Biochem. Genet. 14, 373.
2. Weaver, S., Haigwood, N.L., Hutchison, C.A. III and Edgell, M.H. (179) Proc. Natl. Acad. Sci. 76, 1385.
3. Jahn, C.L., Hutchison, C.A. III, Phillips, S.J., Weaver, S., Haigwood, N.L., Voliva, C.F. and Edgell, M.H. (1980) Cell 21, 159.
4. Edgell, M.H., Weaver, S., Jahn, C.L., Padgett, R.W., Phillips, S.J., Voliva, C.F., Comer, M.B., Hardies, S.C., Haigwood, N.L., Langley, C.H., Racine, R.R. and Hutchison, C.A. III. (1981) Cellular and Molecular Regulation of Hemoglobin Switching, ed. Stammatoyannopolus and Nienhaus, Grune & Stratten publishers.
5. Tiemeier, D.C., Tilghman, S.M., Polsky, F.I., Seidman, J.G., Leder, A., Edgell, M.H. and Leder, P. (1978) Cell 14, 237.
6. Shen, C.-K.J. and Maniatis, T. (1980) Cell 19, 379.
7. Duncan, C., Biro, P.A., Choudary, P.V., Elder, T.T., Wang, R.R.C. Forget, B.G., deRiel, J.K. and Weissman, S.M. (1979) Proc. Natl. Aca. Sci 76, 5095.
8. Leder, A. Miller, H.J., Hamer, D.A., Seidman, J.G., Norman, B., Sullivan, M. and Leder, P. (178) Proc. Natl. Aca. Sci. 75, 6187.
9. Konkel, D.A., Tilghman, S.M. and Leder, P. (1978) Cell 15, 1125.
10. Gross-Bellard, M., Oudet, P. and Chambon, P. (1973) Eur. J. Biochem. 36, 32.

11. Corneo, G., Ginelli, E., Soave, C. and Bernardi, G. (1968) Biochemistry 7, 4373.
12. Enquist, L., Tiemeier, D., Leder, P., Weisberg, R. and Sternberg, N. (1976) Nature 259, 596.
13. Tiemeier, D.C., Tilghman, S.M. and Leder, P. (1977) Gene 2, 173.
14. Benton, W.D. and Davis, R.W. (1977) Science 196, 180.
15. Polsky, F., Edgell, M.H., Seidman, J.B. and Leder, P. (1978) Anal. Biochem 87, 397.
16. Denhardt, D. (1966) Biochem. Biophys. Res. Commun. 23, 641.
17. Britten, R.J. and Kohne, D.E. (1968) Science 161, 529.
18. Konkel, D.A., Maizel, J.V. Jr., and Leder, P. (1979) Cell 18, 865.
19. Schmookler-Reis, R.J. and Biro, P.A. (1978) J. Mol. Biol. 121, 357.
20. Davis, M.M., Calame, K., Early, P.W. Livant, D.L., Toho, R., Weissman, I.L., and Hood, L. Nature 283, 733.
21. Wensink, P.C. (1977) Cold Spring Harbor Symp. Quant. Biol 42, 1033.
22. Lee, A.S., Britten, R.J. and Davidson, E.H. (1977) Cold Spring Harbor Symp. Quant. Biol 42, 1065.
23. Rubin, C.M., Houck, C.M., Deininger, P.L, Friedmann, T. and Schmid, C.W. (1980). Nature 284, 372-374.
24. Davidson, E.H., Hough, B.R., Amenson, C.S. and Britten, R.J. (1973) J. Mol. Biol. 77, 1.
25. Manuelidis, L. (1980) Nucleic Acids Rs. 8, 3247.
26. Davidson, E.H. and Britten, R.J. (1979) Science 204, 1052.
27. Ginelli, E., DiLernia, R. and Corneo, G. (1977). Chromosoma 61, 215.
28. Costantini, R.D., Scheller, R.H., Britten, R.J. and Davidson, E.H. (1978) Cell 15, 173.
29. Jelinek, W.R., Toomey, T.P., Leinwand, L., Duncan, C.H., Biro, P.A., Choudary, P.V., Weissman, S.M., Rubin, C.M., Houck, C.M., Deininger, P.L. and Schmid, C.W. (1980) Proc. Natl. Acad. Sci. USA 77, 1398-1402.
30. Sakonju, S., Bogenhagen, D.F. and Brown, D.D. (1980) Cell 19, 13.
31. Bogenhagen, D.F., Sakonju, S. and Brown, D.D. (1980) Cell 19, 27.
32. Fritsch, E.F., Lawn, R.M. and Maniatis, T. (1979) Nature 279, 598.