Regional localization on the human X of DNA segments cloned from flow sorted chromosomes

Louis M.Kunkel, Umadevi Tantravahi, Michael Eisenhard and Samuel A.Latt

Genetics Division, Children's Hospital Medical Center, 300 Longwood Avenue, Boston, MA 02115, USA

Received 14 December 1981; Revised and Accepted 5 February 1982

ABSTRACT

Fluorescence activated sorting of chromosomes from 49,XXXXY human lymphoblasts has been used to obtain DNA enriched for the human X. This DNA was cloned in λ phage Charon 21A to obtain a library of approximately 60,000 pfu. Phage inserts free of human highly repeated DNA sequences were localized to different regions of the human X by two independent hybridization analyses. The first utilized comparative hybridization to rodent-human hybrid cell DNA samples containing all or known portions of the human X, while the second was based on hybridization dosage to DNA samples from human cell lines differing in the number of X chromosomes or X chromosome segments. Of five unique sequence inserts tested, three were X chromosome specific and were localized to regions Xpter- $\rightarrow$ Xcen, Xq1- $\rightarrow$ Xq22 and Xq24- $\rightarrow$ Xqter, respectively. The library presented here represents a highly enriched source of human X chromosome-specific DNA sequences.

INTRODUCTION

The human X chromosome contains nearly 2×10^5 kb of DNA, corresponding approximately to 200 recombination map units (1). Thus far, approximately 100 human X-linked genetic loci have been identified (2). Most of these loci can undergo genetic inactivation (3), which, in eutherian mammals, achieves X chromosome dosage compensation. Also, the location of genes on the X chromosome is subject to evolutionary conservation (4). A molecular analysis of the mammalian X chromosome, using recombinant DNA technology, might thus provide information about X-inactivation and gene linkage. The latter information might be of diagnostic use in studying the segregation of X-linked disorders, such as hemophilia and Duchenne's muscular dystrophy. Recombinant DNA libraries that are highly enriched for X chromosome DNA should facilitate such efforts.

A variety of approaches have been employed to isolate cloned X-specific DNA. For example, human cell lines containing four X chromosomes have been used (5) as a source of DNA for a plasmid library, from which clones specific for the human X were isolated. Bruns et al (6) employed an alternative approach, in which a cloned DNA library from a human-hamster cell hybrid, that had retained the human X chromosome but little human autosomal chromatin, was screened for human DNA inserts by probing (7) for human-specific repeated sequences. Human-mouse hybrid cells were also used by Schmeckpepper et al to identify repeated DNA segments that were present both on the human X and on human autosomes (8). Another approach for obtaining DNA enriched for the X chromosome is fluorescence activated flow sorting (9,10,11). This methodology has previously been employed in our laboratory to clone DNA enriched for the mouse Cattanach (12) translocation, which has a piece of chromosome number 7 inserted into the X, forming the X(7) chromosome (13).

The present report describes experiments, using fluorescence activated sorting, directed at cloning DNA enriched for the human X chromosome. For this purpose metaphase chromosomes from a 49,XXXXY lymphoblast cell line were isolated (14), stained with 33258 Hoechst and sorted according to fluorescence intensity. DNA was isolated from a chromosome fraction enriched for the X and cloned in λ phage Charon 21A, a Hind III vector (15). Cloned DNA segments were localized to the human X chromosome in experiments employing both hybridization to DNA samples from human-rodent hybrid cells and quantitative hybridization to DNA from human cells differing in the number of X chromosomes or X chromosomal segments. These two screening methods permit independent subchromosomal localization of X-specific fragments.

While the present study was in progress, a report describing an independent human X chromosome library in the phage λ gt Wes 1B, an EcoRI vector, was published (16).

MATERIALS AND METHODS**1) Cell lines****a. Human**

Five human cell lines were employed: a 46,XY lymphoblast line established at Children's Hospital Medical Center; CCRF-CEM, a 46,XX, 9p⁻ lymphoid line (17); a 46,X,i(Xq) lymphoblast line (18) in which all cells examined contained an i(Xq) as well as a normal X; a 46,X,dic(X) (Xpter→Xq24::Xq24→Xpter) (19) lymphoblast line, containing, along with a normal X, the dic(X)(q24) chromosome in more than 98% of the metaphases examined, established by Dr. Herbert Lazarus, Sidney Farber Cancer Institute; and a human 49,XXXXY lymphoblast line (GM1202) obtained from "The Human Genetic Mutant Cell Repository" at the Institute for Medical Research, Camden, New Jersey.

b. Human-rodent hybrids

One (hybrid #2) was formed from fusion between 46,XY human lymphocytes and Chinese hamster (YH21) cells. This hybrid retained one intact human X per YH21 genome, plus an unidentified human autosomal fragment in 20% of the cells; it lacked human autosomal enzyme markers (G.P. Bruns, unpublished data). The other hybrids were products of fusion, between human lymphocytes with a reciprocal X-autosome translocation and rodent cells, that had retained the translocation products containing different segments of the human X plus a few autosomes (that varied from hybrid to hybrid) (G.P. Bruns, unpublished data). One was a hybrid (hybrid #1) of 46,X,t(X;19) lymphocytes and mouse (RAG) cells which retained the X-19 translocation product which includes Xq1→Xqter. Another (hybrid #3) was a hybrid between 46,X,t(X;13) human lymphocytes and mouse (RAG) cells which retained the t(X;13) chromosome which includes Xq22→Xqter (19; G.P. Bruns, unpublished data). The last (hybrid #4) was a hybrid of 46,X,t(X;19) (19) lymphocytes and Chinese hamster (E36) cells, which retained the X-19 translocation product which includes Xq24→Xqter.

The chromosome constitution of all cell lines was ascertained by Q-banding (20), late replication banding (18,19), and, in the case of cell hybrids, isozyme markers (G. Bruns, data not shown here; hybrids described elsewhere (29)).

2) Chromosome isolation

Metaphase chromosomes were isolated by a method derived from procedures by Carrano et al (9) and Otto et al (14). Lymphoblast cell lines (49,XXXXY, GM1202 or 46,XY) were diluted to approximately 10^5 cells per ml of medium and allowed to divide for 48 hours. Colcemid was added (0.01 ug/ml) for the final 12 to 16 hours and cells (as much as 40% mitotic) were collected by centrifugation (500 x g) and resuspended in 1 ml (2×10^7 cells) of hypotonic buffer (5 mM $MgCl_2$, 7.5 mM Tris-HCl, pH 7.5). After 10 minutes incubation at between 18° and 22°, Triton X-100 was added to a final concentration of 0.5%. During another 10 minutes incubation, the cells were periodically disrupted by mild vortexing. Two volumes of isotonic buffer (25 mM $MgCl_2$, 0.1 M NaCl, 15 mM Tris-HCl, pH 7.5) were added to the resulting suspension, and the suspension was left at 4° for 1 hour. The chromosome suspension was separated from the settled nuclear pellet and stained with 33258 Hoechst (final concentration 4 ug/ml).

3) Flow cytometry and sorting

The techniques for chromosome fluorescence analysis and chromosome sorting have been described (13). Sorting was done with a Becton-Dickenson FACS II instrument equipped with a Spectra Physics Model 164-05 argon ion laser with an output of about 100 mW in the ultraviolet (351-364 nm). The analysis rate for preparative sorting experiments was approximately 3000 particles per second. Between 4-6% of the particles, exhibiting fluorescence in the interval expected for the human X chromosome, were sorted into tubes maintained at 4°. The sorted chromosomes were stored at -20° until use in DNA isolation.

4) DNA preparation

a) From flow sorted chromosomes

A pool of 6×10^6 chromosomes, obtained after approximately 20 hours of sorting, was thawed and centrifuged at 12,000 x g for 15'. The chromosome pellet was resuspended in 0.4 ml 0.075 M NaCl, 0.024 M EDTA, 10 mM Tris, pH 7.5 and, when dispersed, the solution was made 1% SDS, 100 ug/ml proteinase K and 50 ug/ml tRNA (Miles). The tRNA had previously been tested to insure that it did not inhibit enzyme reactions or phage

packaging. The mixture containing DNA was then extracted twice with phenol, buffered to pH 8 with 0.1 M Tris-HCl, and then twice with chloroform. After adding Na acetate to the solution to a concentration of 0.1 M, the DNA was precipitated with ethanol (2.5 volumes) at -80° , and the precipitate collected by centrifugation. The DNA was resuspended in 10 mM Tris-HCl, pH 7.5, 1 mM EDTA and stored frozen at -20° . Yields were approximately 0.1 ug DNA per 10^6 chromosomes.

b) From cells or tissues

Nuclei were prepared from placenta and from cells grown in culture by homogenation in 0.32 M sucrose, 10 mM Tris-HCl pH 7.6, 5 mM $MgCl_2$ and 1% Triton-X 100 and collected by centrifugation. DNA was prepared from these nuclei as previously described (21).

c) From lambda phage

Large plates of confluent phage lysate were prepared, and the phage were diffused into 15 ml of 0.1 M NaCl, 10 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$. The phage solution was then placed into a centrifuge tube, and a few drops of chloroform were added. Following a 10 minute $10,000 \times g$ centrifugation to remove bacterial debris, the phage suspension was layered onto a CsCl step gradient (1.4 g/cc and 1.6 g/cc) in an SW27.1 nitrocellulose tube and banded at the 1.4 g/cc: 1.6 g/cc interphase by centrifugation at 22,000 RPM for 2.5 hours. The phage band was collected in approximately 0.5 ml and diluted to 4 ml in 10 mM Tris-HCl 5 mM EDTA. Phage particles were disrupted by addition of SDS to 0.5% and heated at 65° for 10'. After addition of potassium acetate to 0.5 M, the solution was placed on ice for 20'. The capsid protein: detergent precipitate was removed by centrifugation and the phage DNA solution dialyzed in 10 mM Tris-HCl pH 7.5, 30 mM NaCl, 1 mM EDTA. Following dialysis, the solution was extracted once with phenol (pH adjusted to 8.0) and then concentrated with butanol to approximately 0.4 ml. Following one extraction with chloroform, the solution was adjusted to 0.1 M Na acetate and the DNA was precipitated by addition of 2.5 volumes of ethanol. DNA was collected by centrifugation, and the pellet was washed in 90% ethanol and subsequently dried. The DNA was resuspended in 10 mM Tris-HCl pH 7.5

and stored frozen. Yields were approximately 10 ug of DNA per single 150 mm plate confluent lysate.

5) Cloning of DNA fragments and plaque screening

Charon 21A (15) was chosen as the vector for chromosomal DNA cloning. Between 80 and 90% of human DNA, when cleaved with Hind III, falls within the acceptance range of this phage (0 to 9 kb). Approximately 0.1 ug of DNA from sorted chromosomes was cleaved with Hind III and then ligated (T4 DNA ligase, New England Biolabs) with 0.05 ug phage DNA (Hind III cut and subsequently treated with calf intestinal alkaline phosphatase (Boehringer) to limit self ligation). The Charon 21A DNA used here yielded 20 to 30 times more viable phage when Hind III digested genomic DNA was added to a ligation and packaging reaction than when this DNA was absent. The ligated DNA fragments were packaged following the procedure of Blattner et al (15). The 60,000 resultant phage were amplified on lawns of LE392 bacteria (15), purified by centrifugation in a CsCl gradient, and stored at 4°.

Screening of the library for phage containing repetitive human DNA was accomplished by plating a low density of phage on lawns of bacteria and subsequent transfer of plaques to nitrocellulose filters by the method of Benton and Davis (22). Filters were probed with ³²P-labeled DNA in 10% dextran sulfate as previously described (13).

Enzyme reactions

Restriction endonuclease digestions were carried out as described by the manufacturers (New England Biolabs or Boehringer Mannheim). Genomic DNA samples used for electrophoresis and blotting (23) were cleaved preparatively by either EcoRI or Hind III restriction endonucleases in the following manner: approximately 100 ug of DNA was incubated with 50 units of enzyme for 1 hour in buffer described by the manufacturer at 37°. At the end of one hour another 50 units was added for an additional hour. The reaction was monitored by gel electrophoresis for completion of digest; all digests contained prominent human satellite DNA bands detectable in the ethidium bromide-stained gel. Reactions were terminated by the addition of proteinase K (100 ug/ml) for 30' at 37° and subsequent

phenol extraction. After two phenol extractions and two chloroform extractions, the DNA was precipitated with ethanol. The precipitate was resuspended in 100 μ l of 10 mM Tris. The restriction enzyme digestion buffer plus 50 units more enzyme was added, and incubation proceeded another hour at 37°. Reactions were terminated by brief heating at 60° and stored frozen.

The preparative restriction endonuclease digests were judged to be complete by the absence of partial digest products upon Southern hybridization (23) analyses with multiple radio-labeled probes, including globin in the case of the Hind III digests. An aliquot of the same endonuclease digested samples were used for each Southern hybridization (23) presented here as well as for other analyses not presented.

Radiolabeling of DNA was accomplished by either the nick translation activity of *E. coli* DNA polymerase I (24; Boehringer Mannheim), or by T4 DNA polymerase (Bethesda Research Laboratories; 25). The labeling of inserts from the phage was accomplished in the following manner. Pools of 2 or 3 phage DNA samples were cut to completion with Hind III. All resultant fragments were labeled using T4 DNA polymerase as described by O'Farrell (25). The exonuclease reaction proceeded for 10 minutes, and the subsequent polymerizing reaction for 15 minutes. Following removal of unincorporated nucleotides by seive column, the radiolabeled DNA fragments were separated by electrophoresis in 0.6% low melting point agarose (Bethesda Research Laboratories) gels. Autoradiography of the separated fragments in the gel allowed identification of desired inserts. Inserts were excised from the agarose gel, the agarose was melted, and the ³²P-labeled DNA recovered by phenol extraction. Following ethanol precipitation, the radioactive samples were boiled and applied to the nitrocellulose blots (23) of genomic DNA.

The specific activity of radiolabeled products were approximately 2×10^8 dpm/ μ g when utilizing T4 DNA polymerase (25) and 5×10^8 dpm/ μ g with *E. coli* DNA polymerase I (24).

6) Gel electrophoresis and hybridization

For most experiments, three μ g of DNA was loaded into a

horizontal gel slot in 30 μ l of 30% glycerol, 10 mM Tris HCl, 5 mM EDTA, 0.1% SDS and 0.01% bromophenol blue. DNA concentrations were determined by a fluorometric assay using the dye 4', 6-diamidino-2-phenylindole (DAPI) (26,27). Restriction endonuclease digested DNA from various sources was separated by gel electrophoresis on horizontal agarose gels (0.8% agarose in 2 mM EDTA, 90 mM Tris-borate, pH 8). The DNA fragments from gels were transferred to nitrocellulose filters according to Southern (23), using 20X SSC (3.0 M NaCl, 0.3 M Na citrate, pH 7). The filters were prehybridized overnight at 42° in 4X SSC, 30 mM Tris-HCl, pH 7, 50% formamide, 10X Denhardt's (28) solution, 0.1% SDS, 1 mM EDTA with 100 μ g/ml of yeast tRNA and 25 μ g/ml of denatured salmon sperm DNA. A second prehybridization was carried out for 2-4 hours at 42° in 50% formamide, 10X Denhardt's (28) solution, 4X SSC, 50 mM sodium phosphate buffer (pH 6.8), 0.1% SDS, 1 mM EDTA, 8% dextran sulfate, 1% glycine, 100 μ g/ml yeast tRNA and 25 μ g/ml denatured salmon sperm DNA. Hybridization was carried out overnight at 42° in the second prehybridization buffer, typically using 1×10^6 cpm of 32 P radiolabeled phage insert DNA per 5 ml. The filters were washed 5-6 times (5 minutes each) at room temperature in 0.1X SSC, 0.1% SDS, 5-6 times (1 hour total) in the same buffer at 50°C, and then twice at 65° (5 minutes each). Alternatively, those filters containing Hind III digests were prehybridized in 4X SSC, 2X Denhardt's, 0.1% SDS, 1 mM EDTA, 100 μ g/ml of yeast tRNA, 25 μ g/ml of salmon sperm DNA, 20 mM sodium phosphate buffer pH 7.0, 5% dextran sulfate at 68° for 6-8 hours. Hybridizations were carried out in the same buffer overnight at 68°. The filters were washed at room temperature in 0.1X SSC and 0.1% SDS 5-6 times (5 minutes each) then transferred to a solution containing 2X SSC, 0.1% SDS, 20 mM sodium phosphate, 0.08% sodium pyrophosphate at 68°. Filters were washed in the above solution for 4-5 times (10 minutes each wash). The filters were transferred to a 1:1 diluted solution of the above and washed for 10 minutes at 68°. Finally the filters were rinsed in 1X SSC at room temperature. The filters were dried and autoradiographed with Kodak X-omat film against an intensifying screen at -80° for 2-7 days.

RESULTS**Flow cytometry of the human X chromosome**

Carrano et al (9) analyzed human chromosomes, stained with 33258 Hoechst, by fluorescence activated flow sorting. The fluorescent intensity of such chromosomes depends on the amount and emission efficiency of DNA bound dye. Larger chromosomes, with more DNA, generally fluoresce brighter than smaller ones with less DNA. Shown in Figure 1A is a flow histogram of chromosomes from a lymphoblast line with the karyotype 46,XY. The arrow in figure 1A identifies the peak presumed to contain the X plus chromosome #7, based on identification of a similar histogram presented by Gray et al (10). Presented in Figure 1B is a flow histogram of chromosomes from a cell line exhibiting the karyotype 49,XXXXY. The peak containing X and 7 chromosomes (denoted by arrow) in the histogram of 46,XY has increased more than two fold in height and is assumed to be caused by the addi-

Figure 1. Flow histograms of 46,XY and 49,XXXXY lymphoblast metaphase chromosomes stained with 33258 Hoechst. (A) 46,XY human chromosomes. Shown is a reproduction of the histogram displayed by the ND-100 multichannel analyzer of a Becton-Dickenson FACS II sorter. Data were accumulated to approximately 500 fluorescent objects in the tallest, central peak. The arrow is over the peak, containing the X and chromosome #7. (B) 49,XXXXY lymphoblast chromosomes stained with 32258 Hoechst. Data were accumulated to approximately 1000 objects in the central tallest peak. The arrow is over the X and 7 peak and the bar represents the fluorescent intensity range actually sorted and collected.

tion of 3 X chromosomes. The area of fluorescent intensity defined by the bars in Figure 1B was sorted for DNA isolation. Judging from the relatively increased number of the chromosomes in this range of fluorescence, the sorted material was estimated to contain approximately 30 to 60% X chromosomes. Because of the very high condensation of the lymphoblast chromosomes caused by prolonged exposure to colcemid, precise cytogenetic identification was difficult. DNA was isolated from the sorted chromosomes, digested to completion with Hind III restriction endonuclease, ligated with Charon 21A (15) DNA (see methods), and the resultant 60,000 phage amplified and stored.

Characterization of the human X library

A four step procedure was chosen. First, the human origin of the library of sorted chromosome DNA was established. Second, inserts from the library which are free of highly repeated sequences were identified. Third, the X-specificity of an insert was established. Fourth, various X chromosome sequences were localized to subchromosomal regions. Both the regional localization and the establishment of X-specificity were accomplished by two independent tests. One employed hybridization dosage to DNA from human cell lines containing supernumerary (5) or structurally abnormal X chromosomes. The other utilized correlative hybridization to DNA from human-rodent cell hybrids containing various pieces of human X chromosomes in a rodent DNA background (5,6,7,16,29). The use of two different restriction enzymes with these samples aids confirmation of X chromosome specificity, by indicating hybridization dosage with fragments of different sizes.

To demonstrate the presence of human DNA in the constructed phage library, aliquots of the amplified library were plated on bacterial lawns and the resultant plaques of phage diffused onto nitrocellulose discs (22). Nick translated (24) ^{32}P -labeled human DNA was hybridized to the filter containing a replica of the plaques. Shown in Figure 2 is one such hybridization. More than 70% of plaques on the filter contained inserts homologous to ^{32}P -labeled human DNA. Conditions of hybridization were such that only human repeated sequences would hybridize (7). The negative plaques were presumed to represent phage containing

Figure 2. Screening (20) of plaques from an X enriched phage library for repeated human DNA sequences. ³²P-labeled human DNA was hybridized to a replica filter of a plate of 56 phage plaques from the enriched library (see methods for details). More than 70% of plaques showed positive hybridization.

inserts of human DNA free of highly repeated sequences.

Plaques which were presumptively identified as negative for repeated sequences were picked and amplified. The goal was to prepare inserts of human origin free of repeated sequences. CsCl purified phage were prepared from each of a dozen plaques and the DNA isolated from each. The purified DNA was cut with Hind III and the digest products separated by electrophoresis on agarose gels. Ethidium bromide staining of the gel revealed inserts of DNA ranging in size between 0.5 and 7 kb, with a mean size of approximately 4 kb. The DNA in the gel was then transferred to nitrocellulose filters (23) and human ³²P-labeled DNA hybridized to the filter. Hybridization to Southern blots appeared to be more sensitive than plaque screening for detecting repetitive DNA, since some of the inserts scored negative by plaque screening appeared positive on subsequent blot (23) analysis. Five phage DNA samples isolated as described above, plus two others, chosen at random from the library, containing inserts negative for repeated sequences, were mixed in three pools. The Hind III restriction fragments from these pools were labeled by the T4 DNA polymerase labeling procedure described by O'Farrell (25) and recovered as described (methods).

Shown in Figures 3A and B are the hybridization patterns of two ³²P-labeled presumably unique sequence inserts from the sorted X chromosome library with test DNA samples cleaved with

Figure 3. Screening of cloned DNA inserts for X chromosome specificity. Three μ g of Eco RI digested genomic DNA samples from cells differing in human X content were separated by electrophoresis in 0.8% agarose gels and transferred to nitrocellulose (see methods). The DNA samples used on the gel were the following: lane 1, 46,XY placental DNA; lane 2, 49,XXXXY lymphoblast DNA; lane 3, mouse-human hybrid cell DNA (hybrid #1, see methods) with all of the human X chromosomes long arm; lane 4, 32 P-end labeled (T4 DNA polymerase), Hind III cut λ DNA; lane 5, hamster-human hybrid cell DNA (hybrid #2) with an intact X chromosome; lane 6, mouse-human hybrid cell DNA (hybrid #3) containing the distal half of the X chromosome long arm. A schematic representation of the human X with the portion present in the cell from which the DNA is isolated is depicted above the lanes (shaded). (A) Autoradiograph of the hybridization of a 1.7 kb Hind III fragment, end labeled with 32 P using T4 DNA polymerase to EcoRI digested DNA samples immobilized in nitrocellulose (Methods). The patterns observed are consistent with localization of this fragment to the human X. (B) Autoradiograph of the hybridization of a 32 P end-labeled 2.0 kb Hind III insert to aliquots of the same immobilized samples as in (A). This second DNA fragment does not appear to be localized on the human X.

EcoRI. In panel A is the hybridization pattern of a phage insert (released by Hind III) of 1.7 kb size, which appears to be derived from the short arm of the X chromosome. There is a significantly greater hybridization to a 6.6 kb EcoRI fragment of DNA from cells with 4 X chromosomes than to DNA from cells with only 1 X. The other lanes, except #5 (containing an intact X chromosome) were negative for this 6.6 kb EcoRI fragment. The

negative lanes contained DNA which included X chromosome representation limited to the long arm. The additional hybridization, most apparent in the human lanes, is assumed to be either homology with sequences not found on the X chromosome or the presence of a small amount of repeated DNA still present in the insert sequence. Based on both the dosage (lanes 1 and 2) and the selective hybridization pattern (to lane 5, but not to lanes 3 and 6), this insert is probably derived from the short arm of the X chromosome.

A second insert shown in Figure 3B, when hybridized to equal aliquots of the same samples of DNA cleaved with EcoRI, appears not to be localized to the human X chromosome. The amount of the 3.6 kb EcoRI fragment is not increased in the DNA from 49,XXXXY cells with 4 X chromosomes and is not present in DNA from a hybrid cell with an intact human X chromosome.

To confirm that the insert tested in Figure 3A was located in the short arm of the X, the 1.7 kb Hind III fragment was ³²P-labeled and hybridized to Hind III-digested samples of DNA from human cells with different X chromosome compositions (Figure 4). A distinct difference in hybridization intensity was observed between DNA from cells with one X (lane 1) and DNA from cells with two X chromosomes (lanes 2 and 3) for a 1.7 kb DNA fragment. The fourth lane contains DNA from cells with the karyotype 46,X,i(Xq), which should contain one copy of the X chromosome short arm and 3 copies of the X chromosome long arm. The hybridization observed was similar in extent to that seen with DNA from cells with only 1 X chromosome (lane 1) and distinctly less than that observed with DNA from cells with two X chromosomes (lanes 2 and 3). Greater hybridization was observed in lane 5 (involving DNA from cells containing three copies of all of the X except the distal part of the long arm) than in DNA from cells containing two X chromosomes (shown in lanes 2 and 3). Hybridization of a ³²P-labeled globin probe to the same amounts of these samples yielded less variation in hybridization intensity for all samples than that observed here for DNA with one X chromosome versus that with two X chromosomes (data not shown). The results presented in Figure 4 with Hind III genomic digests and in Figure 3A with EcoRI genomic DNA

Figure 4. Regional localization of a cloned 1.7 kb Hind III insert within the human X chromosome. DNA from human cells differing in the number of X chromosomes or containing a structurally abnormal X was digested with Hind III and treated as described in Figure 3. The top of each lane is labeled with a schematic representation of the X chromosomes present in these human cell lines. Autoradiography was for 1 day. Each lane contained 4 ug of Hind III digested DNA: lane 1, 46,XY lymphoblast DNA; lane 2, 46,XX lymphoblast DNA; lane 3, 46,XX 9p lymphoid DNA; lane 4, 46,X,i(Xq) lymphoblast DNA; lane 5, 46,X,dic(X)(q24) lymphoblast DNA; lane 6, 49,XXXXY lymphoblast DNA. Consistent with the results of Figure 3A, this 1.7 kb DNA fragment appears to be localized to the short arm of the human X.

digests substantiate an X chromosome short arm localization for this 1.7 Hind III insert from the X library.

Of seven inserts used as probes against the DNA digest test panels shown in Figures 3 and 4, two were clearly from human chromosomes other than the X. Two other inserts were slightly repeated, and chromosomal localization was not possible. Three inserts were clearly from the X chromosome. The one tested in Figure 3A and 4 is from the X chromosome short arm, while two others are from the long arm of the X chromosome.

One of these latter cloned sequences can presumptively be localized to the distal part of the X chromosome long arm. This 5.5 kb Hind III insert, when tested against EcoRI digests, hybri-

dizes much more strongly with two fragments (4.3 and 2.7 kb, respectively) in the 4X-containing DNA than with the 1X DNA (Figure 5A, lanes 1 and 2). These two restriction fragments are also found in all four of the rodent-human hybrid cell DNA samples, including the DNA sample in lane 7, Figure 5A. This last DNA sample contains only that section of the X-chromosome long arm, which is distal to band Xq24. The possible distal long arm

Figure 5. Localization of a cloned 5.5 kb Hind III insert to the distal part of the long arm of the human X chromosome. (A) 32 P-labeled 5.5 kb insert was prepared with T4 DNA polymerase. Aliquots of EcoRI digested DNA (3 ug) of the same samples from the cell panel described in the caption to Figure 3A, plus DNA from a hybrid cell containing the segment of the human X distal to band q24 were loaded as follows; lane 1, 46,XY placenta DNA; lane 2, 49,XXXXY DNA; lane 3, hybrid #1 DNA; lane 4, 32 P- λ DNA cleaved with Hind III; lane 5, hybrid #2 DNA; lane 6, hybrid #3 DNA; lane 7, hybrid #4 DNA (see Methods for hybrid descriptions). (B) Electrophoresis of aliquots of Hind III-digested human DNA (4 ug) samples (see methods). The 32 P-labeled DNA used as a probe was the 5.5 kb Hind III insert subcloned in pBR 322 and labeled by nick translation (24; methods). The samples are in the following order; lane 1, human 46,XY DNA; lane 2, human 46,XX DNA; lane 3 human 46,XX 9p⁻ DNA; lane 4, human 46,X,i(Xq) DNA; lane 5, human 46,X,dic(X)(Xq24) DNA; lane 6, 49,XXXXY DNA. The results of both autoradiograms are constitent with localization of this 5.5 kb insert to the segment of the human X between Xq24 and Xqter.

localization with the EcoRI digests (Figure 5A) was confirmed by similar analysis using Hind III digested samples (Figure 5B). A 5.5 kb Hind III fragment of DNA from the 46,X,i(Xq) cells (lane 4, Figure 5B), the same size as the cloned fragment, hybridizes to a greater amount than does the same fragment from DNA samples from cells with two X-chromosomes. This 46,X,i(Xq) cell line has 3 copies of the human X chromosome long arm and one of the X chromosome short arm. Furthermore, equivalent intensity of hybridization is seen with the 46,X,dic(X)(q24) (lane 5) and 46,XY DNA (lane 1). The translocation is such that there are 3 copies of the majority of the X, but a single dose of Xq24->Xqter. The cloned fragment must be from the area deleted in this translocation chromosome, confirming the distal long arm localization based on results (Figure 5A) using EcoRI-digested rodent-human hybrid DNA.

A third X-specific insert from the library gave more complicated hybridization patterns. Hybridization of this 1.3 kb Hind III fragment to EcoRI genomic DNA digests (Figure 6A) yielded 3 fragments (2.0, 1.2 and 0.5 kb) in the 49,XXXXY DNA (lane 1). The 1.2 kb fragment is present in 46,XY DNA (lane 2), where it exhibits approximately one quarter the hybridization intensity as that seen with 49,XXXXY DNA (lane 1). Hybridization to 2.0 kb and 0.5 kb fragments was very faint, requiring long exposures for detection in 46,XY DNA, in which an additional 2.5 kb fragment was observed. This 2.5 kb fragment presumably represents a lack or reduction of cleavage at a single EcoRI site linking the 0.5 and 2 kb fragments. Further addition of EcoRI enzyme to the previously cleaved 46,XY DNA sample did not change the hybridization intensity with this 2.5 kb fragment.

The X-chromosome specificity of this 1.3 kb Hind III insert is confirmed by analysis of its hybridization to the DNA samples from hybrid cells (lanes 3, 4, 5 and 6, Figure 6A). All three EcoRI fragments (2.0, 1.2 and 0.5 kb) are present in DNA from hybrid cells containing the entire human X (lane 4) or most of the X chromosome long arm (lane 3), but they are absent from DNA from cells containing only distal segments of the X chromosome long arm (lanes 5 and 6). On the basis of these data, this cloned 1.3 kb Hind III insert is tentatively localized to a

Figure 6. Localization and characterization of a cloned 1.3 kb Hind III DNA fragment. The methodology and hybrid X chromosome representations are as in the previous figures. In both autoradiograms (A) and (B), the 1.3 kb Hind III fragment hybridized was 32 P-labeled using T4 DNA polymerase. (A) Aliquots of EcoRI digests of DNA (3 ug) were as follows; lane 1, 49,XXXXY lymphoblast DNA; lane 2, 46,XY placenta DNA; lane 3, hybrid #1 DNA; lane 4, hybrid #2 DNA; lane 5, hybrid #3 DNA; lane 6, hybrid #4 DNA; lane 7, 32 P-labeled Hind III digested λ DNA. (B) Hind III digested DNA was from the following samples: lane 1, 46,XY lymphoblast DNA; lane 2, 46,XX lymphoblast DNA; lane 3, 46,XX 9p⁻ lymphoid DNA; lane 4, 46,X,i(Xq) lymphoblast DNA; lane 5, 46,X,dic(X)(Xq24) lymphoblast DNA; lane 6, 49,XXXXY lymphoblast DNA. The hybridization patterns observed localize this fragment to the proximal part of the long arm of the human X.

segment of the X chromosome between bands Xq1 and Xq22. Also present in both hybrid cell DNA samples (lane 3 and 4) is the 2.5 kb fragment. The intensity of hybridization with this 2.5 kb fragment is not diminished by prior digestion with additional EcoRI endonuclease.

Localization of this 1.3 kb Hind III insert to the X chromosome long arm is confirmed by the results of hybridization to Hind III digests of DNA from cell lines with structurally abnormal X chromosomes (Figure 6B). Hybridization to DNA samples with 3 copies of X chromosome long arm (lanes 4 and 5)

yielded intensities of a 1.3 kb fragment intermediate between 46,XX DNA (lanes 2 and 3), with two X chromosomes, and 49,XXXXY DNA (lane 6), with four X chromosomes. This intensity is approximately that expected for 3 doses of the 1.3 kb Hind III fragment and correlates well with the 3 copies of X chromosome long arm present in those cell lines. By these criteria, the 1.3 kb Hind III insert is located between Xql and Xq24. No variant restriction fragment sizes were observed with Hind III cut genomic DNA samples.

DISCUSSION

The experiments described demonstrate that DNA sequences specific for one chromosome type, in this case the human X-chromosome, can be isolated from a lambda phage library constructed from flow sorted chromosomes. Based on flow histogram comparisons as shown in figures 1A and 1B, between 30 and 60% of the chromosomes isolated here were from the X chromosome. Phage with inserts lacking repeated DNA were selected and the X chromosome localization was made for 3 of the 5 inserts examined that were free of repeated sequences. Analysis of additional repeat sequence-negative phage from this library is in process. Thus far, the yield of X-specific fragments is within the range expected from flow histogram data. The X chromosome amounts to approximately 5% of the human haploid autosomal chromatin mass (9). The X-library presented here, then, represents nearly a 10 fold enrichment for X chromosomal DNA, and it provides a ready source of X chromosome specific DNA sequences.

The X-enriched library described in the present report and cloned in the Hind III vector Charon 21A should complement the EcoRI library, cloned in λ gt Wes, described by Davies et al (16). The test panel of genomic DNA samples utilized in the present study (Figures 3, 4, 5 and 6) affords a ready analysis of inserts from this library and others (16,29). The removal of highly repeated sequences is essential to identify specific X chromosomal location, since many (7), although not necessarily all (30) repeated sequences are located on multiple chromosomes. The uniform X-specificity of the cloned inserts presented here is substantiated by the fact that all three inserts, when

hybridized against Hind III digested DNA, yielded single fragments of the same size as the cloned fragment and by the correlation of the hybridization intensity of these fragments with the number of X chromosomes present in each DNA sample. Further studies are in progress to quantitate this hybridization dosage.

The detail with which unique sequence inserts from the library have been mapped extends to a subregional localization on the X chromosome. Two independent methods are employed for this purpose. One relies on human-rodent hybrids retaining different parts of the human X; the second employs human cell lines with structurally rearranged X chromosomes. The first approach offers more variety, since a wealth of hybrid cell lines is available, but is subject to uncertainty due to cellular heterogeneity and chromosome rearrangement in these cells. The second is free of these problems but requires accurate detection of two or three-fold differences in hybridization extent. This last capability has already been demonstrated in our previous work with the mouse X (13) and by initial screening of segments for human X specificity (Figures 4, 5 and 6) using DNA from cell lines containing supernumerary X chromosomes. Importantly, the two methods yield consistent results, serving for mutual validation. By utilizing both hybridization dosage and hybridization with DNA from human-rodent somatic cells, three sequences from our library have been assigned to X chromosome. They are, Xpter->Xqcen (Figure 3A and 4), Xq1->Xq22 (Figure 6), and Xq24->Xqter (Figure 5).

Future sequences can be localized by methods presented here to at least 4 sections of the X chromosome. This subdivision places such sequences on the average within a maximum of 50-100 recombination map units from known X-linked loci. Groups of sequences from the same section of the X chromosome might thus ultimately serve as markers in restriction fragment length polymorphism (31) studies and linkage analysis. Possibly relevant to such studies is one insert from the library presented here (Figure 6). Hybridization of a 1.3 kb Hind III insert to EcoRI digested genomic DNA yields a fragment (2.5 kb) in 46,XY placental DNA and two cell hybrid DNA samples. This

new restriction fragment is assumed to be caused by the loss or reduction of cleavage at a single EcoRI site. This is presumably not due to enzymatic digestion that is artifactually incomplete, for this degree of incompleteness was not observed in these DNA samples when other inserts were hybridized to them. Studies are in progress to determine if this variation is caused by local inability of EcoRI to cleave this site completely due to possible methylation or if this represents a base substitution in a single EcoRI site. The use of EcoRI cleaved DNA isolated from members of individual pedigrees should establish the possible genetic nature of this variation and the possible genetic linkage of this variation with other X chromosome markers. To facilitate linkage analysis and subsequent screening of phage libraries, several of the X-linked DNA sequences described in this report have already been subcloned in plasmid pBR 322 (32).

A chromosomal library is an excellent source of genomic clones for identifying expressed X specific sequences. Towards this goal of isolating this subset of transcribed sequences, pools of cDNA clones derived from Hela cell mRNA have been prepared free of repeated DNA transcripts (33). Screening of the library with these cDNA pools has been initiated and some positive phage identified. Characterization of these phage is in progress.

The approach described, in which chromosome-specific sequences are obtained from DNA libraries enriched by flow sorting, at present emphasizes those fragments lacking repeated sequences. It thus complements the approach of Bruns et al (6, 29), in which species-specific repeated DNA serves as the marker necessary for identification of clones containing X chromosome segments. Sequences isolated by either method can be further subclassified according to their genetic expression or evolutionary conservation. Sets of X-specific sequences so identified are now being isolated for utilization in more detailed mapping of the X, for linkage analysis, and for a molecular analysis of the process of X chromosome inactivation.

ACKNOWLEDGEMENTS

This research was supported by grants from the National Institutes of Health (GM21121 and HD 40807) and the National Foundation March of Dimes (1-353). Lou Kunkel is a fellow of the Muscular Dystrophy Association.

We thank Dr. G.A.P. Bruns for generous provision of hybrid cell lines and unpublished data, and together with Drs. D.M. Kurnit and P. Gerald for many useful discussions during the course of this work, Dr. H. Lazarus for help in establishing lymphoid cell lines, Mr. A. Lojewski for the initial assistance in chromosome sorting and Ms. J. Johnson for expert preparation of this manuscript.

REFERENCES

1. Vogel, F., and Motulsky, A.G. (1979) Human Genetics, Problems and Approaches. Springer Verlag, Berlin, Heidelberg, New York.
2. McKusick, V.A. (1978) Mendelian Inheritance in Man. The Johns Hopkins University Press, Baltimore and London.
3. Lyon, M.F. (1961) Nature 190, 372-373.
4. Ohno, S. (1967) Sex Chromosomes and Sex Linked Genes. Springer, Heidelberg.
5. Wolf, S.F., Marenzi, C.E., Migeon, B.R. (1980) Cell 21, 95-102.
6. Bruns, G.A.P., Gusella, J.F., Keys, C., Housman, D., and Gerald, P.S. (1981) Pediatric Research 15, 599 (abstract #704).
7. Gusella, J.F., Keys, C., Vassanyi-Breiner, A., Kao, F-T, Jones, C., Puck, T.T., and Housman, D. (1980) Proc. Natl. Acad. Sci. USA 77, 2829-2833.
8. Schmeckpepper, B.J., Willard, H.F., and Smith, K. (1981) Nucleic Acids Res. 9, 1853-1872.
9. Carrano, A.V., Gray, J.W., Langlois, R.G., Burkhardt-Schultz, K. and Van Dilla, M.A. (1979) Proc. Natl. Acad. Sci. USA 76, 1382-1385.
10. Gray, J.W., Langlois, R.G., Carrano, A.V., Van Dilla, M.A. (1979) Chromosoma 73, 9-27.
11. Disteché, C.M., Carrano, A.V., Ashworth, L.K., Burkhardt-Schultz, K., Latt, S.A. (1981) Cytogenet. Cell Genet. 29, 189-197.
12. Cattanach, B.M. (1961) Z. Vererbungslehre 92, 165-182.
13. Disteché, C.M., Kunkel, L.M., Lojewski, A., Orkin, S.H., Eisenhard, M., Sahar, E., Travis, B. and Latt, S.A. (in press) Cytometry.
14. Otto, F.J., Oldiges, H., Gohde, W., Barlogie, B., Schumann, J. (1980) Cytogenet. Cell Genet. 27, 52-56.
15. Blattner, F.R., Williams, B.G., Blechl, A.E., Denniston-Thompson, K., Faber, H.E., Furlong, L.A., Grunvald, D.J.,

- Kiefer, D.O., Moore, D.D., Schumm, J.W., Sheldon, E.L., Smithies, D. (1977) *Science* 196, 161-169.
16. Davies, K.E., Young B.D., Elles, R.G., Hill, M.E., and Williamson, R. (1981) *Nature* 293, 374-374.
17. Krishan, A., Raychaudhuri, R., and Flowers, A. (1969) *J. Natl. Cancer Inst.* 43, 1203-1214.
18. Latt, S.A., Barell, E.F., Dougherty, C.P. and Lazarus, H. (1981) 3, 171-181.
19. Latt, S.A., Willard, H.F., and Gerald, P.S. (1976) *Chromosoma (Berl.)* 57, 135-153.
20. Caspersson, T., Lindsten, J., Lomakka, G., Moller, A., and Zech, L. (1972) *Int. Rev. Exp. Pathol.* 11, 1-72.
21. Kunkel, L.M., Smith, K.D., Boyer, S.H., Borgaonkar, D.S., Wachtel, S.S., Miller, O.J., Breg, W.R., Jones, H.W. Jr. and Rary, J.M. (1977) *Proc. Natl. Acad. Sci. USA* 74, 1245-1249.
22. Benton, W.D. and Davis, R.W. (1977) *Science* 196, 180-182.
23. Southern, E.M. (1975) *J. Mol. Biol.* 98, 503-517.
24. Rigby, P.W., Dieckman, M., Rhodes, C. and Berg, P. (1977) *J. Mol. Biol.* 113, 237-251.
25. O'Farrell, P. (1981) *Focus* 3, 1-3.
26. Brunk, C.F., Jones, K.C. and James, T.W. (1977) *Analyt. Biochem.* 83, 252-257.
27. Kapuscinski, J. and Skoczylas, B. (1977) *Analyt. Biochem.* 83, 252-257.
28. Denhardt, D.T. (1966) *Biochem. Biophys. Res. Commun.* 23, 641-646.
29. Bruns, G.A.P., Gusella, J.E., Keys, C., Leary, A.C., Housman, D. and Gerald, P.S., submitted.
30. Willard, H.F., Smith, K.D. and Goss, S.J. (1981) *Amer. J. Hum. Genet.* 33, 59A.
31. Botstein, D., White, R.L., Skolnick, M. and Davis, R.W. (1980) *Am. J. Hum. Genet.* 32, 314-331.
32. Bolivar, F., Rodriguez, R.C., Green, P.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W., Crosa, J.H. and Falkow, S. (1977) *Gene* 2, 95-113.
33. Kurnit, D.M., Kunkel, L.M., and Wentworth, B.M., Villa-Komaroff, L. (1981) *Am. J. Hum. Genet.* 33, 47A.