Construction and analysis of DNA sequence libraries from flow-sorted chromosomes: practical and theoretical considerations

Jeffrey K.Griffith^{1*}, L.Scott Cram¹, Brian D.Crawford¹, Paul J.Jackson¹, James Schilling², Robert T.Schimke², Ronald A.Walters¹, Mark E.Wilder¹ and James H.Jett¹⁺

¹University of California, Los Alamos National Laboratory, Life Sciences Division, Los Alamos, NM 87545, and ²Stanford University, Department of Biochemistry, Stanford, California, USA

Received 12 December 1983; Accepted 30 March 1984

ABSTRACT

We describe the construction and analysis of recombinant DNA libraries representative of chromosomes 1 and 2 of Chinese hamster (Cricetus griseus). Propidium-iodide stained chromosomes were purified by flow cytometric analysis and sorting, and EcoRI digests of purified DNA were cloned into the bacteriophage vector Charon 4A. These libraries contain DNA complementary to 63% and 69% of nick-translated DNA derived from flow-purified chromosomes 1 and 2, respectively. However, sequences complementary to only 24% and 35% of a total Chinese hamster genomic DNA tracer were hybridized in parallel renaturation experiments. The chromosome 2 library contained DNA sequences encoding dihydrofolate reductase (dhfr), a gene previously mapped to Chinese hamster chromosome 2. No sequences complementary to dhfr were found in the library constructed from chromosome 1 DNA. These analyses are discussed with regard to the current limitations and future strategies for the construction of chromosome-specific DNA sequence libraries of high purity and completeness.

INTRODUCTION

Recombinant DNA technology has greatly expanded our knowledge of the structure, organization, and expression of DNA sequences in mammalian genomes, and promises to play important new roles in mammalian pedigree analysis, cytogenetics, and gene mapping. In this context, the cloning of DNA sequences representative of specific chromosomes would facilitate investigations relating to: i) gene mapping by combined somatic cell and molecular genetic techniques (1), ii) linkage and pedigree analyses by the study of DNA restriction fragment length polymorphisms (1,2), iii) specific chromosome alterations associated with neoplastic and inborn genetic disorders (3), iv) the comparative study of DNA sequence organization among individual chromosomes within a given species, and v) inter-species genetic comparisons, such as the study of linkage group homologies and rearrangements during evolution (1,4,5,6).

Previous approaches toward the cloning of chromosome-specific DNA sequences have relied upon competitive hybridization or species-specific DNA

sequences (e.g., Refs. 7-10). The utility of this approach is limited, however, in that a complete library of pure chromosome-specific sequences cannot be generated, (a fact attributable to both the kinetics observed for the reassociation of mammalian DNA (11), and the sequence homology which exists between rodent and human DNA)(12).

A direct approach to the purification of chromosome-specific DNA sequences consists of physical isolation of the desired chromosome by fluorescence-activated sorting in flow (13,14). Improvements in the preparation of metaphase chromosome samples, coupled with advances in flow instrumentation, have improved the resolution and lowered background contamination such that highly purified chromosome fractions can be obtained by sorting (15,16,17). Although flow cytometry alone is very limited in providing large quantities of chromosomes for biochemical or cytogenetic studies, when combined with recombinant DNA technology it is a practicable tool. For example, flow sorting techniques have already been used to isolate DNA to clone, in bacteriophage vectors, recombinant DNA libraries enriched for the human and mouse (Cattanach) X-chromosomes (18,20), human chromosomes 21 and 22 (21), and an abnormal human chromosome number 1 (22).

The utility of chromosome-specific libraries relies upon their purity and completeness. The preparative purification of individual chromosomes by flow sorting requires: 1) that properties such as the DNA content of each chromosome be measured with sufficient precision to permit unambiguous resolution (23); 2) that a sufficiently large number of chromosomes be analyzed in a reasonable time; and 3) that at a practicable analysis rate which does not compromise resolution, the chromosomes of interest be sorted reliably from the rest of the population. In this paper, we address each of these issues in the context of the construction of recombinant DNA libraries representative of chromosomes 1 and 2, respectively, of the Chinese hamster (Cricetulus griseus). In addition, we report biochemical and molecular genetic characterizations designed to examine the purity and completeness of these libraries. The results and interpretations of these investigations have significance with regard to future studies relying upon chromosome-specific DNA sequences for genetic and biochemical investigations.

MATERIALS AND METHODS

1. Fluorescence Activated Chromosome Analysis and Sorting

a. Cell Cultures. The Chinese hamster cell line, M3-1 (24), was obtained from A. Carrano of the Lawrence Livermore Laboratory; this cell line has been

cultured in our laboratory for several years, during which time a few rearrangements have occurred among some of the small metacentric and acrocentric chromosomes. Whole Chinese hamster embryo (WCHE) cell cultures were isolated at Los Alamos. The flow and G-banded karyotypes of the WCHE-5 line are described in detail elsewhere (25). All cells were maintained in exponential growth in medium (α MEM, GIBCO, Grand Island, NY) supplemented with 10% (vol/vol) fetal bovine serum (Reheis, Kankakee, IL).

b. Chromosome Analysis and Sorting Using FACS II Instrumentation.

Chromosomes were isolated for flow analysis using a modified hypotonic lysis and propidium iodide staining procedure (26), detailed previously (25).

Chromosome analysis and sorting were accomplished with a modified Becton-Dickinson FACS II. The modifications were restricted to the electronic components of the system. Preamplified signals from the photomultiplier tube were AC coupled into a Los Alamos designed amplifier/integrator (27). The amplifier and integrator outputs were digitized with 8 bit accuracy and histograms generated in a Digital Equipment Corporation PDP 11/23 using the LACEL hardware (28) and software (29). Sorting windows and coincidence requirements were set in the LACEL sort module. The resulting sort left or sort right commands were issued within 5 μ -sec by the sort module (28) when signals from the desired chromosomes were detected. The prompt sort signals were passed to a sort delay and PZT drive module (30) which issues the properly delayed and phased droplet charging pulse.

c. Calculation of Sort Purity and Yields. The three major determinants of purity in sorting are the quality of the preparation (in terms of debris and chromosome aggregation), the quality of the measurement which serves as the basis for the sorting decision, and the accidental overlap (within a sort time window) of the desired chromosome with undesired chromosomes. Figure 1b (see Results) shows a typical distribution of integrated fluorescence signals obtained from Chinese hamster M3-1 chromosomes at an analysis rate of approximately 1000 events per second. Figure 1b also shows the placement of the left and right sort windows. The contamination attributable to debris and clumped chromosomes is small, as evidenced by the number of events detected in channels between the chromosome peaks. With the resolution obtained, chromosomes 1 and 2 were well separated, and cross contamination between the two regions was negligible. Given the instrumental operating conditions, the accidental contamination of the sorted fractions can be calculated (31). Under the typical conditions used (analysis rate of 1000 chromosomes per second, sorting 10% of the population, sorting two drops, a

droplet formation rate of 40,000 per second, and no overlap rejection), the expected purity was greater than 95%.

2. DNA Purification

a. From Flow-Sorted Chromosomes. Sorts of approximately 10^6 chromosomes (approximately 1 μ g) of each type were centrifuged at 30,000 rpm at 4°C for 30' in an SW-41 rotor. The chromosome pellets were resuspended in 0.5 ml of 150 mM NaCl, 10 mM Tris-HCl, 10 mM EDTA, pH 8.0. This solution was adjusted to 0.5% SDS, 100 μ g/ml proteinase K, and 20 μ g/ml tRNA. (The addition of tRNA did not inhibit restriction endonuclease or ligation reactions, or phage packaging.) After incubation at 37°C for 1-2 hours, DNA was extracted twice with redistilled phenol:chloroform:isoamyl alcohol (25:24:1 by volume), then twice with chloroform:isoamyl alcohol (24:1). Sodium acetate was added to 0.3 M, and the DNA was precipitated by the addition of 2.5 volumes of ethanol and stored at -20°C prior to use. A total of 10^7 chromosomes were sorted and used for DNA isolation for these studies.

b. From Cells. Chinese hamster (CHO) cells were grown in spinner cell culture, and the DNA purified from isolated nuclei by established methods (32).

c. From Bacteriophage. Large (150 mm) plates of confluent phage lysate were prepared, and phage resuspended in 0.1 M NaCl, 10 mM Tris-HCl, pH 7.5, 10 mM MgCl₂. The phage were then purified according to Yamamoto *et al.* (33) and banded in CsCl step gradients (1.45, 1.50, 1.70 g/cc) at the 1.45-1.5 g/cc interphase by centrifugation in an SW50.1 rotor at 35,000 rpm for 1 hour. The phage bands of individual gradients were collected, pooled, and dialysed against 0.1 M NaCl, 10 mM Tris-HCl, pH 7.5, 10 mM MgCl₂. Phage particles were disrupted by bringing this solution to 20 mM in EDTA, 50 μ g/ml proteinase K, and 0.5% SDS, and incubation at 65°C for 1 hour. The DNA was then extracted with phenol and chloroform, adjusted to 0.3 M Na acetate, and precipitated with 2.5 volumes of ethanol as described above. The DNA was collected by centrifugation and resuspended in 10 mM Tris-HCl, 1 mM EDTA, pH 7.6. Yield was approximately 10 μ g DNA per plate of phage lysate.

3. Molecular Cloning of DNA Representative of Chinese Hamster Chromosomes 1 or 2.

DNA was purified from flow-sorted chromosome 1 or 2, and 1-2 μ g of DNA from each chromosome were digested with the restriction endonuclease EcoRI (BRL, Bethesda, MD). The small mass of DNA used precluded size-selection of EcoRI fragments prior to cloning; however, by dilution of the enzyme, and monitoring DNA digestion in control experiments using small masses of total genomic DNA, an attempt was made to achieve partial EcoRI digestion of

flow-sorted material. After agarose gel electrophoresis and ethidium bromide staining, inspection of aliquots (1/10) of the restriction endonuclease digestion reactions revealed nearly complete digestion, however, of the purified chromosomal DNA. Charon 4A (34) was selected as the cloning vector (the acceptance range of this phage is 7.1–20.1 kilobase pairs for EcoRI fragments cloned, by substitution, using purified vector arms). After ligation (T4 DNA ligase, New England Biolabs) with Charon 4A arms at a mass ratio of approximately 2:3 (insert:vector), DNA was packaged in vitro according to modifications by Enquist and Sternberg using the NS428 and NS433 E. coli lysogens for the preparation of packaging extracts (35). The resulting recombinant phage were amplified on lawns of DP50 SupF bacteria to minimize the under-representation of slower replicating phage (36), and stored as phage lysates at 4°C.

4. Screening of the Phage Libraries.

Screening of the phage library for phage containing chromosome-specific gene sequences was performed by plating phage on lawns of bacteria, and subsequent transfer of plaques to nitrocellulose filters by the method of Benton and Davis (37). Filters were hybridized to pDHFR-11, a ³²P-labeled cDNA recombinant clone of mouse dihydrofolate reductase (dhfr), by methods described previously (38). Plasmid pDHFR-11 is a pBR322 derivative which contains a 1500 bp cDNA insert corresponding to nearly full-length mouse dhfr mRNA (38).

5. Nucleic Acid Reassociation Kinetic Analyses in Solution.

Radioactive tracer DNA (specific activity $\geq 2 \times 10^6$ cpm/ μ g DNA) complementary to either total genomic, or flow-sorted chromosome 1 or chromosome 2 DNA was prepared by nick translation (39) using ³H-labeled nucleotides. Driver DNA was prepared from CHO cell nuclei, or the recombinant chromosome 1 and chromosome 2 libraries, as described above, and sheared to a single strand length of approximately 500 nucleotides. Approximately 1–2 ng of the desired tracer was mixed either with 6–60 μ g of sheared genomic DNA or 6 μ g of the sheared chromosome-specific library DNAs, and the reassociation kinetics as a function of DNA concentration (C_0), time (t) and equivalent salt concentration (E) assayed by hydroxylapatite chromatography (32).

RESULTS

1. Chinese Hamster Flow Karyotype Analysis and Chromosome Sorting.

Chromosome analysis and sorting were performed using a cell sorter

Figure 1. Representative flow karyotypes for (a) WCHE-5 cells and (b) M3-1 Chinese hamster cells. Samples were prepared and analyzed using a cell sorter equipped with a FACS II optical bench, as described in Methods. Distributions were obtained at an analysis rate of 1000-2000 chromosomes/second. Sort windows for chromosomes 1 and 2 are noted. Peak assignments, (explained in reference[25]) are, from right to left, chromosomes 1, 2, 3, X, 4, 5, 6, 7, 8, 9, 10 of *Cricetulus griseus*. Chromosomes in the aneuploid M3-1 karyotype which were used for sorting are noted (homologues 1A and 1B, 2).

equipped with a Becton-Dickenson (Sunnyvale, CA) FACS II optical bench. Figure 1a shows a representative flow karyotype for euploid whole Chinese hamster embryo (WCHE) cells recorded at an analysis rate of 1000 events/second, reflecting a coefficient of variation (C.V.) of 1.5-2.0% for chromosomes 1 or 2. Flow karyotype information obtained with this and other instrumentation at Los Alamos has been used, in conjunction with G-banding and cytogenetic analyses, to identify each peak in the Chinese hamster flow karyotype (25,40,41). Similar resolution of chromosomes is shown in Figure 1b, a representative distribution for chromosomes from the aneuploid Chinese hamster fibroblast line, M3-1. At an analysis rate of 1000-2000 events/sec, the two homologues of chromosomes 1 were partially resolved in M3-1 cells, whereas chromosomes 1 and 2 were fully resolved from each other and all other chromosomes in the M3-1 karyotype. Debris present in each of the sort windows (Figure 1b) was less than 5%. This observation, coupled with the ease of culturing this rapidly-growing established cell line, encouraged the use of M3-1 cells as starting material for flow sorting and recombinant DNA cloning experiments.

The relative fluorescence of chromosomes 1 and 2 indicates that they contain 23.8% and 19.6%, respectively, of the total genomic DNA mass of Cricetulus griseus. The relatively large DNA content (approximately 10^{-12} g/chromosome) of these chromosomes, coupled with their excellent resolution in flow, suggested that quantities of DNA sufficient for cloning could be isolated in a reasonable time and with good purity. At an analysis rate of 1,000 chromosomes/second, 1 μ g of DNA representing each of these chromosomes could be collected after 3-4 hours of continuous sorting, with a purity of 95% or greater (see Methods). Furthermore, larger chromosomes such as 1 and 2 would be expected to contain fewer contaminating chromosome fragments, which are the consequences of chromosome damage during sample preparation.

2. Construction and Analysis of Recombinant Bacteriophage Libraries

DNA was purified from flow-sorted chromosomes, digested with the restriction endonuclease EcoRI, and cloned in the lambdaphage vector Charon 4A, as described in the Methods. Amplified stocks derived from the resulting recombinant phage (6×10^4 pfu for chromosome 1; 9×10^4 pfu for chromosome 2) were used to examine the purity and completeness of these libraries were then examined by two experimental approaches.

a) Nucleic Acid Reassociation Kinetics. DNA was extracted and purified from amplified phage representing chromosome-specific "libraries", sheared to a single-stranded length of approximately 500 nucleotides, and mixed with either tracer amounts of radiolabeled total Chinese hamster genomic DNA, or

TABLE I

Nucleic Acid Reassociation Kinetic Analyses to Estimate Sequence Representation in Recombinant Libraries Constructed From Chromosomes 1 or 2 of Chinese Hamster

Source of DNA Tracer	Total Genome	Percent Hybridization to Driver DNA Isolated From:	
		Chromosome 1 Library	Chromosome 2 Library
Total Genomic DNA	98	24	35
Chromosome 1 (sorted)	92	63	48
Chromosome 2 (sorted)	94	28	69

1-2 ng of ^3H -labeled total genomic DNA, chromosome-1 DNA, or chromosome-2 DNA (specific activities $> 2 \times 10^5$ cpm/ μg) were mixed with 6 μg of unlabeled 500 nucleotide-long chromosome 1 or chromosome 2 "library" DNA, or 60 μg of unlabeled 500 nucleotide-long total genomic DNA in 1.5 M NaCl, 10 mM Tris-HCl (pH 7.5) 2 mM Na_4EDTA , and 0.2% SDS. The samples were denatured by heating for 2 minutes at 100°C and incubated at 72°C to an ECot of 5,000, then diluted to 1.0 ml with 0.05 M phosphate buffer (PB) and loaded onto 1 g hydroxylapatite columns at 60°C . The single-stranded and double-stranded DNA fractions were eluted at 60°C with 0.12 M PB and 0.4 M PB, respectively. The double-stranded fraction was normalized to each tracer's maximum reactivity (97%, 90% and 89% for total genomic DNA, chromosome 1 DNA, and chromosome-2 DNA respectively). See Materials and Methods for additional details.

chromosome-specific tracer DNA (prepared from purified chromosome 1 or 2 DNA obtained by flow sorting) as described in the Methods. The kinetics and extent of DNA reassociation were measured by hydroxylapatite chromatography, as described in the Methods and the legends to Table I and Figure 2.

In control experiments, each of the three DNA tracers had reactivities in excess of 90% when annealed with total genomic DNA (Table I, and Figure 2). If the two chromosomal DNA populations isolated by flow-sorting were mutually discrete, the homologous chromosomal DNA libraries and tracers would be expected to reassociate to a much greater extent than parallel heterologous reactions. This prediction was borne out by the results presented in Table I. In homologous reassociations between DNA from chromosome-specific libraries and chromosome-specific tracers, 63% and 69% of the two tracers were in double-stranded molecules when the reactions were terminated at an ECot of 5000. Based on an average insert size of 15 kb in the Charon 4A vector, and by comparison to the kinetics of hybridization of single copy hamster DNA sequences with total genomic hamster DNA (Figure 2b), we estimate that all repetitive DNA sequences, and approximately 80% of all single copy sequences should have annealed at this value of ECot. In parallel reactions designed to measure reassociation between DNA from chromosome-specific libraries and their heterologous tracers, only 28% and 48% of the two tracers

Figure 2. Reassociation kinetics of total genomic or chromosome-specific DNA.

(A) 1-2 ng of ^3H -labeled total genomic DNA ($\blacktriangle$), chromosome 1 DNA ($\bullet$) or chromosome 2 DNA ($\circ$) were mixed with 6-60 μg of unlabeled 500 nucleotide-long total genomic DNA, denatured, and incubated to the indicated value of ECot. The single-stranded and double-stranded DNA fractions were resolved on hydroxylapatite and normalized to tracer reactivity as described in the legend to Table I.

(B) Single-copy Chinese hamster genomic DNA was prepared from the total genomic DNA described in (A) above by self-reassociation to an ECot of 10^3 and isolation of the single-stranded DNA fraction by hydroxylapatite chromatography as described in Table I. 1-2 ng of this single-copy tracer was mixed with 20 μg of the unlabeled 500 nucleotide long total genomic DNA, denatured at 100°C , and the rate and extent of hybridization measured as in "A" above. See Materials and Methods for additional details.

were recovered as double stranded DNA at an ECot of 5,000. Furthermore, only 24% and 35% of the total genomic DNA tracer reassociated with the two chromosomal libraries (Table I). Collectively, these results demonstrate: 1) that sequences complementary to 65-75% of the total Chinese hamster genomic DNA mass are not present in the respective chromosome specific libraries; 2) that different, mutually discrete DNA sequences are contained in

flow-purified chromosome 1 and chromosome 2 DNA; and 3) that these libraries contain DNA complementary to 63% and 69% of the DNA contained in chromosomes one and two respectively. We conclude that each of the chromosomal DNA fractions obtained by flow sorting is enriched for specific sets of DNA sequences. However, these recombinant libraries are probably lacking certain DNA sequences. This issue is addressed in more detail in the discussion.

b) Screening for Recombinant Phage Containing Dihydrofolate Reductase (dhfr) Sequences. The Chinese hamster dihydrofolate reductase (dhfr) gene has been mapped to chromosome 2 by somatic cell genetic techniques (42). This gene, which spans a minimum of 26 kb of Chinese hamster genomic DNA, contains six exons as defined by hybridization to mouse and Chinese hamster dhfr cDNA probes (43). Upon complete restriction of Chinese hamster DNA with EcoRI, 13 discrete fragments of the dhfr gene are generated, seven of which encompass exon-containing regions homologous to the mouse pDHFR-11 probe (detailed in reference 43). These dhfr-encoding fragments are all below the 20.1 kb upper size limit for EcoRI fragments cloned into the Charon 4a vector. Assuming that the two sorted chromosomal DNA populations are distinct, there should be dhfr DNA sequences in the recombinant library derived from chromosome 2, but none in the recombinant bacteriophage constructed from chromosome 1. Phage plaques representing approximately two chromosome equivalents of each library ($\sim 10^5$ total pfu each) were probed with dhfr cDNA from the plasmid pDHFR-11 (38) by the method of Benton and Davis (37). Twelve hybridizing phage plaques were detected upon an initial screening of the chromosome 2 DNA library. Subsequent analysis of several of these plaque isolates confirmed that they contained sequences complementary to the dhfr probe. In contrast to these findings, no dhfr-complementary clones were detected among an equal number of phage examined from the recombinant library constructed from chromosome 1. This result demonstrates that phage containing dhfr sequences are absent or infrequent in the recombinant library constructed from chromosome 1 DNA, indicating that the purity of the contamination of the sorted chromosome 1 and chromosome 2 fractions was as expected from flow analysis (discussed above).

3. Analysis of DNA Sequence Representation in Chromosomes 1 and 2 and Their Respective Chromosome-Specific Libraries.

The discreteness of the two flow-sorted chromosomal DNA populations, as evidenced by the investigation of the recombinant libraries described above, indicated that DNA tracers constructed from these two chromosomes could also be used to compare the mass distribution of repetitive and single copy DNA

sequences in chromosome 1 and 2, in comparison to total genomic DNA. Flow-sorted DNA, purified from each chromosome, was nick-translated, stripped of snap-back sequence by hybridization to very low values of Cot , mixed with a 10^3 - 10^4 fold excess of sheared, total genomic Chinese hamster DNA, and reassociated to different values of $ECot$. As shown in Figure 2a, the reassociation kinetics of the two chromosome-specific DNA tracers were indistinguishable from that of the total genomic DNA tracer. Comparison to the reassociation kinetics of a single-copy DNA tracer prepared from total genomic DNA (Figure 2b), indicated that 40-45% of each of the three tracer DNA's represented repetitive sequences, whereas 55-60% represented single copy DNA sequences. Thus, the proportion of single-copy and repeated DNA sequences in these two chromosomes is indistinguishable from that in total genomic Chinese hamster DNA.

DISCUSSION

Flow sorting has been used previously to obtain recombinant DNA libraries enriched for specific human chromosomes (summarized in the introduction). In each of the reports published to date, strategies have been used which are similar to that which we employed here for the cloning of chromosomes 1 and 2 of Chinese hamster. Specifically, after flow-sorting a given chromosomal fraction, purified DNA has been digested to completion and cloned into a single bacteriophage vector of choice. The recombinant phage thus produced have been extremely useful in providing a source enriched for chromosome-specific DNA sequence probes, and have facilitated regional mapping of such DNA sequences on chromosomes (18-22, 45). The utility of these recombinant libraries for other studies may be limited, however, by inadequate documentation of their purity and completeness. In this context, the results of our studies may be useful both in suggesting a strategy for assessing the purity and completeness of chromosome-specific recombinant libraries, as well as in recommending a strategy to be adopted in the construction of such libraries at the outset.

In this paper, we have demonstrated that after propidium iodide staining, the resolution of the Chinese hamster karyotype in flow is sufficiently precise to permit sorting of chromosomes with high purity. This resolution can be compared with flow karyotypes obtained by single parameter analysis for human cells (16), wherein certain chromosome groups cannot be resolved even at low flow rates. Human chromosomes stained with Hoechst 33258 and chromomycin A3 have been resolved into 18 fractions by sequential excitation with two lasers, at an analysis rate of 1000 chromosomes/second (23).

Pending other instrumental and technical advances, however, the yield of chromosomes obtained with high purity by flow sorting will be low. Such low yields make other (perhaps preferred) cloning approaches, such as those involving physical shearing or partial restriction endonuclease cleavage and subsequent size selection of DNA (46), or cosmid cloning of large DNA fragments (47) ill-advised at best. The studies reported in this paper indicate that the use of a single cloning vector may be inadequate to provide complete representation of the DNA sequences on a chromosome. This results from the acceptance range of the cloning vector, coupled with the difficulty in controlling partial restriction enzyme digestion, and the spacing of the restriction enzyme recognition sequence. For example, large DNA fragments (47) ill-advised at best. The studies reported in this paper indicate that the use of a single cloning vector may be inadequate to provide complete representation of the DNA sequences on a chromosome. This results from the acceptance range of the cloning vector, coupled with the difficulty in controlling partial restriction enzyme digestion, and the spacing of the restriction enzyme recognition sequence. For example, if the spacing for a given recognition sequence is very long (>20 kb) or organized into tandem short repeats, certain genomic sequences might be missed in cloning. This problem perhaps can be obviated by constructing separate chromosome-specific libraries from total restriction digests of DNA, i.e., by using combinations of vectors with different acceptance ranges, and by using vectors which accept different restriction endonuclease cleavage fragments. Our findings thus underscore the need to construct chromosome-specific DNA sequence libraries in more than one recombinant vector.

Aside from the more practical considerations highlighted by this study, information of a more basic nature also has been gained. First, the utility of chromosome-specific libraries in confirming the mapping of a specific gene (that for dhfr) has been demonstrated. In addition to this confirmatory mapping (42) of dhfr in Chinese hamster, an apparent enrichment of this sequence in the chromosome 2-specific library was demonstrated. Such sequence enrichment may be a useful future application of chromosome-specific libraries, particularly for cloning strategies designed to isolate functional genes apart from their related pseudogenes, which may be chromosomally dispersed (48). Furthermore, the availability of cloned chromosome-specific DNA sequence libraries from the Chinese hamster enables molecular genetic approaches to the study of linkage group conservation or rearrangement during evolution. For example, by somatic cell genetic techniques, Stallings and

Siciliano have reported chromosomal mapping of genetic loci in the Chinese hamster (6). These mapping studies have revealed a conservation of certain autosomal linkage groups among Chinese hamster, mouse, and human species. Specifically, synteny for the genetic loci AK2, PGD, PGM1, and EN01 has been conserved in these three species (6). These loci map to chromosome 2 of the Chinese hamster (6). While potentially significant, such conservation may simply reflect close linkage of loci which have not become chromosomally dispersed. Our library representative of chromosome 2 may be useful for molecular genetic studies designed to address this and related questions.

Second, these studies also used flow-sorted chromosomes for a purpose distinct from recombinant cloning. The radiolabeling of DNA from sorted chromosomes enabled nucleic acid reassociation kinetic analyses designed to quantitate sequence representation on specific chromosomes. These experiments demonstrated that single copy and repeated sequences are distributed, on a mass basis, similarly within chromosomes 1 and 2 and the genome of *Cricetulus griseus* as a whole. Not all genomic repeated DNA sequences were confirmed to be present, however, in the recombinant libraries constructed from these chromosomes. From our data we can speculate on the representation of specific DNA sequences in the two chromosomes, and recombinant libraries derived from them. Given that 40-45% of total genomic Chinese hamster DNA reacts with kinetics characteristic of repetitive sequences (Figure 2), and considering that chromosomes 1 or 2 each contains approximately 20% of the total genomic DNA (Figure 1), if each chromosome contains members of all repetitive DNA sequences in the genome, each should bear sequences homologous to 51-57% of the total genomic tracer ($40-45\% + [55-60\% \times 0.2]$). Data presented in Table 1 show, however, that recombinant DNA libraries constructed from chromosomes 1 and 2 hybridized only 24% and 35%, respectively, of the total genomic tracer. Considering that the heterologous reassociations reported in Table 1 demonstrated part of this hybridization to be attributable to the reassociation of single copy DNA sequences, it is clear that the extent of reaction of each chromosome library driver DNA with the total Chinese hamster genomic tracer DNA was substantially lower than would be expected for the reaction of repetitive DNA sequences alone. If we assume a given repetitive sequence present at 1000 copies per genome to be distributed uniformly on a mass basis, such a sequence would be reiterated 200 times on each of chromosomes 1 and 2. Given the statistical representation of DNA sequences we achieved in our cloning, as well as the demonstrated presence of at least one-third of the single copy

sequences in each of the recombinant libraries, the exclusion of certain genomic repetitive sequences would appear to be non-random. Whether the exclusion of these sequences is merely attributable to experimental artifact(s) of recombinant cloning (such as those we have emphasized above), or alternatively, is suggestive of Chinese hamster genomic repetitive sequences truly lacking in these chromosomes, remains to be determined. Clearly, the existence of chromosome-specific repetitive sequences, several of which have been documented (7,48,50), could have significance with regard to the structural and functional organization of the genome. Chromosome-specific DNA sequence libraries of high purity and completeness should aid in addressing this question.

ACKNOWLEDGEMENTS

We thank John L. Hanners, and Judith Buckingham for technical assistance. The editorial assistance of Monica Fink in preparation of the manuscript is also gratefully acknowledged. This work was supported in part by the Los Alamos National Laboratory, the Department of Energy, and the NIH Flow Cytometry National Resource (Grant #RR01315).

*Present address: Department of Cell Biology, University of New Mexico, Cancer Research and Treatment Center, Albuquerque, New Mexico

+To whom correspondence should be addressed

REFERENCES

1. Shows, T. B., Sakaguchi, A. Y., and Naylor, S. L. (1982) in: Advances in Human Genetics, ed. Harris H., and Hirschhorn, K., Vol. 12 pp. 341-451, Plenum, NY.
2. Botstein, D., White R. L., Skolnick, M., and Davis, R. W. (1980) *Am. J. Hum. Genet.* 32:314-331.
3. McKusick, V. A. (1978) Mendelian Inheritance in Man: Catalogs of Autosomal Dominant, Autosomal Recessive, and X-linked Phenotypes. Johns Hopkins Univ. Press, Baltimore, MD.
4. Dutrillaux, B. (1979) *Hum. Genet.* 48:251-314.
5. Nash, W. G. and O'Brien, S. J. (1982) *Proc. Natl. Acad. Sci. U.S.A.* 79:6631-6635.
6. Stallings, R. L. and Siciliano, M. J. (1981) *Somatic Cell Genetics* 7:683-698.
7. Kunkel, L. M., Smith, K. D., Boyer, S. H., Borgaonkar, D. S., Wachtel, S. S., Miller, O. J., Breg, W. R., Jones, H. W., and Rary, J. M. (1977) *Proc. Natl. Acad. Sci., U.S.A.* 74:1245-1249.
8. Olsen, A. S., McBride, D. W., and Otey, M. C. (1980) *Biochemistry* 19:2419-2428.

9. Schmeckpeper, B. J., Smith, K. D., Dorman, B. P., Ruddle, F. H., and Talbot, C. C. (1979) *Proc. Natl. Acad. Sci., U.S.A.* 76:6525-6528.
10. Gusella, J. F., Keys, C., Vassanyi-Breiner, A., Kao, F. T., Jones, C., Puck, T. T., and Housman, D. (1980) *Proc. Natl. Acad. Sci. U.S.A.* 77:2829-2833.
11. Britten, R. J. and Kohne, D. E. (1968) *Science* 16:529-540.
12. Rice, N. (1974) *Carnegie Inst. Wash. Yearbook* 73:1094-1098.
13. Carrano, A. V., Gray, J. W., Langlois, R. G., Burkhardt-Schultz, K. R., Van Dilla, M. A. (1979) *Proc. Natl. Acad. Sci. U.S.A.* 76:1382-1384.
14. Gray, J. W., Carrano, A. V., Steinmetz, L. L., Van Dilla, M. A., Moore, D. H., Mayall, B. H., and Mendelsohn, M. L. (1975) *Proc. Natl. Acad. Sci. U.S.A.* 72:1231-1234.
15. Lebo, R. V. and Bastian, A. M. (1982) *Cytometry* 3:213-219.
16. Young, B. D., Ferguson-Smith, M. A., Sillar, R., and Boyd, E. (1981) *Proc. Natl. Acad. Sci. U.S.A.* 78:7727-7731.
17. Cram, L. S., Arndt-Jovin, D. J., Grimwade, B. S., and Jovin, T. M. (1979) *J. Histochem. Cytochem.* 27:445-453.
18. Davies, K. E., Young, B. D., Elles, R. G., Hill, M. E., and Williamson, R. (1981) *Nature* 293:374-378.
19. Kenkel, L. M., Tantravahi, U., Eisenhard, M., and Latt, S. A. (1982) *Nuc. Acids Res.* 10:1557-1578.
20. Distèche, C. M., Kunkel, L. M., Lojewski, A., Orkin, S. H., Eisenhard, M., Sahar, E., Travis, B., and Latt, S. A. (1982) *Cytometry* 2:282-286.
21. Krumlauf, R., Jeanpierre, M., and Young, B. (1982) *Proc. Natl. Acad. Sci. U.S.A.* 79:2971-2975.
22. Latt, S. A., Alt, F. W., Schreck, R. R., Kanda, N., and Baltimore, D. (1982) in: *Gene Amplification*, ed. R. T. Schimke, Cold Spring Harbor Laboratory, NY. pp. 283-289.
23. Gray, J. W., Langlois, R. G., Carrano, A. V., Burkhardt-Schultz, K., and Van Dilla, M. A. (1979) *Chromosome* 73:9-27.
24. Gray, J. W., Carrano, A. V., Steinmetz, L. L., VanDilla, M. A., Moore, D. H., Mayall, B. G., and Mendelsohn, M. L. (1975) *Proc. Natl. Acad. Sci. U.S.A.* 72:1231-1234.
25. Cram, L. S., Bartholdi, M. F., Ray, F. A., Travis, G. L., and Kraemer, P. L. (1983) *Cancer Research* 43, 4828-4837.
26. Yu, L. C., Aten, J., Gray, J., and Carrano, A. V. (1981) *Nature* 293:154-155.
27. Steinkamp, J. A., Fulwyler, M. J., Coulter, J. R., Heibert, R. D., Harvey, J. L., and Mullaney, P. F. (1973) *Rev. Sci. Instrum.* 44:1301-1310.
28. Heibert, R. D., Jett, J. H., and Salzman, G. C. (1981) *Cytometry* 1:337-341.
29. Salzman, G. C., Wilkins, S. F., and Whitfill, J. A. (1981) *Cytometry* 1:325-336.
30. Heibert, R. D. and Sweet, R. G. in: *Methods of Flow Cytometry* (1983) Academic Press, NY (in press).
31. McCutcheon, M. J. and Miller, R. G. (1982) *Cytometry* 2:219-225.
32. Graham, D. E., Britten, R. J., and Neufeld, B. R., (1974) *Methods in Enzym.* 29E, 246.
33. Yamamoto, K. R., Alberts, B. M., Benzinger, R., Lawhorne, L., and Treiber, G. (1970) *Virology* 40:734-739.
34. Williams, B. and Blattner, F. (1979) *J. Virol* 29:555-561.
35. Enquist, L. and Sternberg, N. (1979) in: *Methods in Enzymology*, ed. R. Wu, Vol. 68 pp. 281-299 (Academic Press, NY).
36. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning*, Appendix C, Cold Spring Harbor Press, NY, pp. 504-506.

37. Benton, W. D. and Davis, R. W. (1977) *Science* 196:180-182.
38. Nunberg, J. H., Kaufman, R. J., Chang, A. C. Y., Cohen, S. N., and Schimke, R. T. (1980) *Cell* 19:355-364.
39. Rigby, P. W.J., Dieckmann, M., Rhodes, C., and Berg, P. (1977) *J. Mol. Biol.* 113:237-251.
40. Bartholdi, M. F., Sinclair, D. C., and Cram, L. S. (1983) *Cytometry*, 3, 395-401.
41. Ray, M. and Mohandas, T. (1976) *Cytogenetics Cell Genetics* 16:83-91.
42. Roberts, M. and Ruddie, F. H. (1980) *Exp. Cell. Res.* 127:47-54.
43. Carothers, A. M., Urlaub, G., Ellis, N., and Chasin, L. A. (1983) *Nuc. Acids Res.* 11:1997-2012.
44. Nunberg, J. H., Kaufman, R. J., Schimke, R. T., Urlaub, G. A., and Chasin, L. A. (1978) *Proc. Natl. Acad. Sci. U.S.A.* 75:5553-5556.
45. Lebo, R. V. (1982) *Cytometry* 3:145-154.
46. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning, Cold Spring Harbor Press, NY, pp. 270-294.
47. *ibid*, pp. 295-307.
48. Sharp, P. (1983) *Nature* 301:471-472.
49. Yang, T. P., Hansen, S. K., Oishi, K. K., Ryder, O. A., and Hamkalo, B. (1982) *Proc. Natl. Acad. Sci. U.S.A.* 79:6593-6597.
50. Willard, H. F., Smith, K. D., and Sutherland, J. (1983) *Nuc. Acids. Res.* 11:2017-2033.