Haematology and Blood Transfusion Vol. 28
Modern Trends in Human Leukemia V
Edited by Neth, Gallo, Greaves, Moore, Winkler

Construction and Characterization of Chromosomal DNA Libraries*

B. D. Young, M. Jeanpierre, M. H. Goyns, G. D. Stewart, T. Elliot, and R. Krumlauf

A. Summary

The construction of an extensive human genetic linkage map will require the generation of large numbers of DNA probes specific for single chromosomes. Recombinant DNA libraries representing chromosomes 22 and 21 have been constructed with a view to studying the specific rearrangements of chromosome 22 observed in chronic myeloid leukaemia and Burkitt's lymphoma and also the association of chromosome 21 with Down's syndrome. This was accomplished by sorting about 2×10^6 copies of chromosome 22 and 21 by flow cytometry and insertion of the DNA into the vector *λgt WESλB*. Twenty clones selected at random from the chromosome 22 library hybridised to *Eco*RI-digested human DNA, and five of these hybridised to single bands identical in size to the phage inserts. Altogether six single-copy sequences and a clone coding for an 8S RNA isolated by screening the chromosome 22 library for expressed sequences were characterised in detail. Hybridisation of all seven clones to a panel of sorted chromosomes and hybrid cell lines confirmed the assignment of the sequences to chromosome 22. The sequences were localised to regions of chromosome 22 by hybridisation to translocated chromosomes sorted from a cell line having a balanced translocation *t(17;22)(p13;q11)* and to hybrid cell lines containing the various portions of another translocation *t(X;22)(q13;q112)*. Five clones reside on the long arm of chromosome 22 between *q112* and *qter*, while two clones and an 18S rRNA gene isolated from the chromosome 22 library reside between *pter* and *q112*. In situ hybridisation has further mapped one of these clones to the region *q13-qter*. An identical approach was used for chromosome 21 and three out of five single copy clones have been identified as specific to this chromosome by hybridisation to DNA from a cell hybrid containing only human chromosome 21. In summary, it has been shown that this approach has yielded DNA libraries of high purity based on chromosomes 21 and 22. This method has the advantage of being direct and applicable to nearly all human chromosomes and will be important in the molecular analyses of human genetic disease.

B. Introduction

The availability of cloned genes has rapidly expanded our knowledge of the fine structure, organisation and expression of sequences in the human genome, and provided insight into the molecular basis of diseases such as the thalassemia syndromes. In addition, the application of recombinant DNA technology to human genetics has played an important role in attempting to bridge the gap in resolving power between DNA sequence analysis and classical cytological approaches to mapping and gene linkage. Recent studies have shown that highly polymorphic loci can be identified from human genomic libraries (see review [1]). Using such probes the construction of

* This work was supported by the Medical Research Council, Cancer Research Campaign and the Leukemia Research Fund

a fine structure linkage map would be particularly useful in cases where the chromosomal location of the genetic lesion associated with a disease is known. In this context a general method for rapidly isolating a wide spectrum of DNA probes for specific human chromosomes would be an invaluable aid to human gene mapping. An essential prerequisite for such an approach is the high resolution fractionation of human chromosomes.

There have been many attempts to fractionate metaphase chromosomes using centrifugation [2], counter-current distribution [3], and 1 g sedimentation [4]. While such attempts have achieved some enrichment of different size classes of chromosomes, the similar size of many human chromosomes has prevented the separation of pure fractions of single chromosomes. The construction of human-rodent cell hybrids has allowed the purification of single human chromosomes using selectable markers [5]. Although this approach has the disadvantage of the rodent cell DNA background, it has been used to obtain clones from particular chromosomes [6]. However, due to the presence of a larger number of clones containing rodent cell DNA, this approach is useful for obtaining only small number of clones. In order to obtain a larger number of clones from a particular chromosome, it is necessary to purify the chromosomes of interest in sufficient quantity to allow direct DNA cloning. We have already reported the sorting of the human chromosomes 22 and 21 by flow cytometry in sufficient quantity to allow the construction of DNA libraries [7]. In this report we extend these observations to include the mapping of more clones on both chromosomes 22 and 21. Chromosomes 22 and 21 were chosen for this study in view of the specific rearrangements of 22 which occur in chronic myeloid leukaemia and the association of 21 with Down's syndrome.

C. Results

I. Sorting and Cloning of Chromosomes 21 and 22

The strategy we have adopted for obtaining DNA sequences from human chromosomes 21 and 22 involves directly sorting these chromosomes on a fluorescence-activated cell sorter (FACS II). A high-resolution flow analysis of a suspension of metaphase chromosomes prepared [8] from the lym-

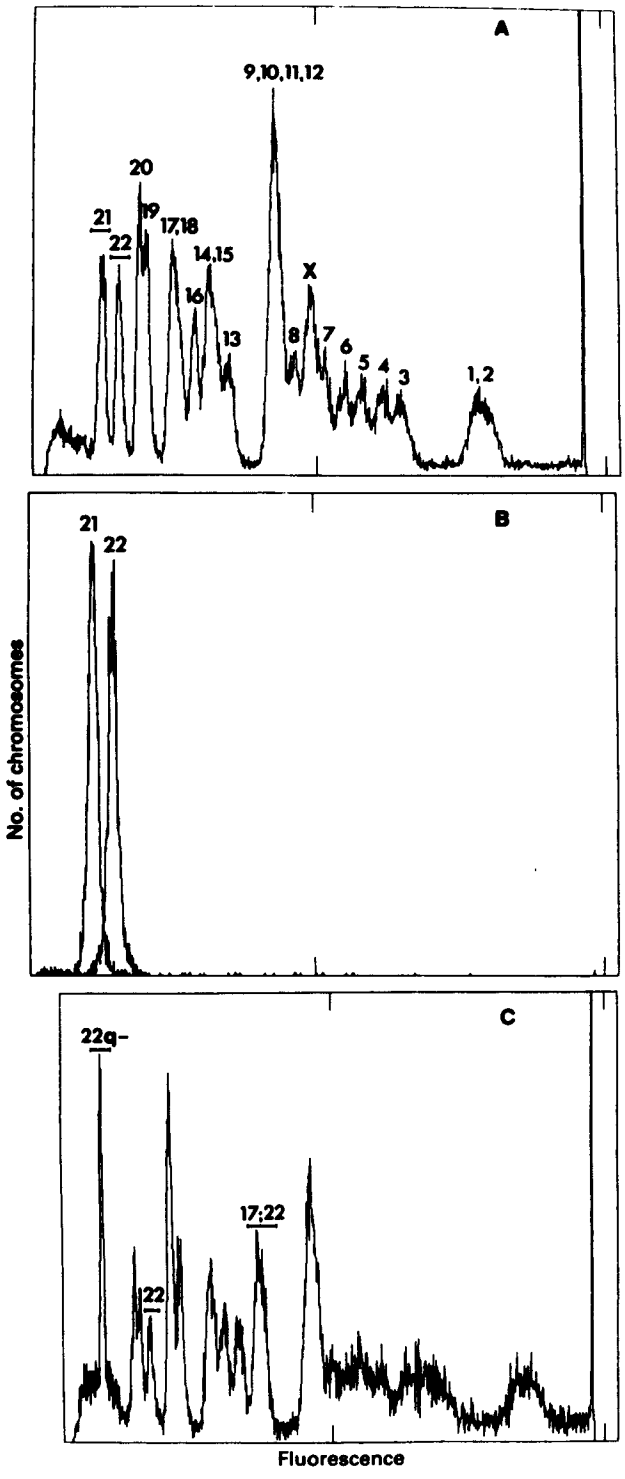

Fig. 1A–C. Flow karyotypes of human chromosomes. *Numbers* indicate the chromosomes contained in each peak, and the *horizontal bars* indicate the fluorescent windows used to sort chromosomes. **A** GM1416: 48,XXXX lymphoblastoid cells. **B** Fractions of chromosomes 21 and 22 sorted from **A**. **C** GM3197: 46,XX cells bearing the translocation *t(17;22)(p13;q11)*

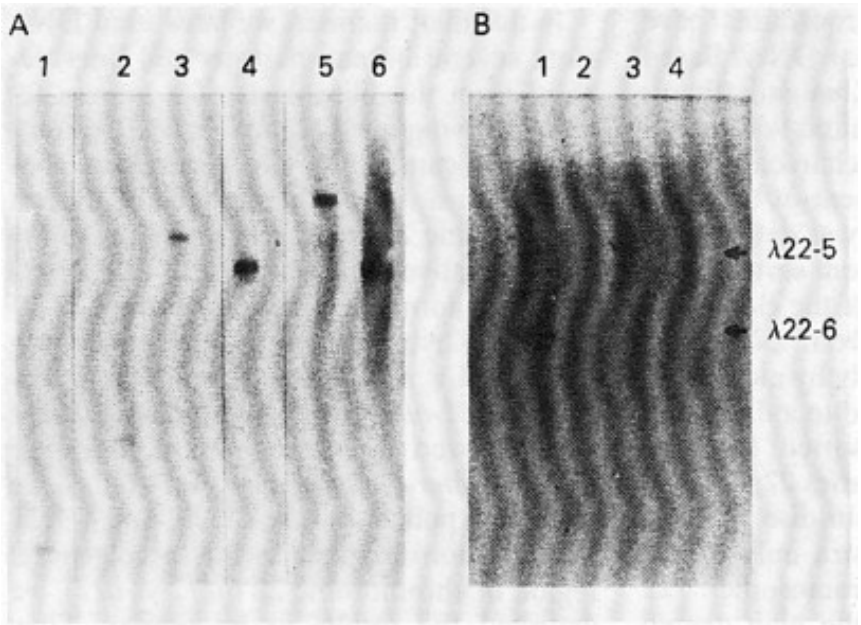

Fig. 2A, B. **A** Hybridisation of single-copy clones. Nitrocellulose strips containing 20 μg of GM1416 DNA digested with *Eco*RI were hybridised with probes prepared from various recombinant phages. *Lanes: 1, λ22-1; 2, λ22-2; 3, λ22-3; 4, λ22-4E; 5, λ22-5; 6, λ22-6.* **B** Assignment of probes to chromosome 22. Clones *λ22-5* and *λ22-6* were labelled separately, mixed, and hybridised to *Eco*RI-digested DNA from sorted fractions containing 1×10^6 copies of various chromosomes. *Lanes: 1,* chromosome 22; *2,* chromosome 21; *3,* chromosomes 1–8; *4,* chromosomes 9–20

phoblast cell line GM 1416: 48, XXXX, generates the flow karyotype shown in Fig. 1A. The assignment of chromosomes to each of the distinct peaks was based on their DNA contents [9] and on the results of a study which established correlations between flow karyotypes and conventional G- and C-banding karyotypes using abnormal human chromosomes and heterochromatic polymorphisms [10]. The peaks containing chromosomes 21 and 22 were individually sorted, and a sample of each was reanalysed on the FACS II to assess the efficiency of sorting (Fig. 1B). Each fluorescence profile showed a sharp single peak, suggesting the sorted fractions were pure.

DNA was extracted from each fraction containing 2×10^6 chromosomes (about 90–100 ng DNA) and was used to construct recombinant libraries in *λgt WES λB* by a modification of the method described by Maniatis et al. [11]. The DNA was completely digested with *Eco*RI, due to the difficulty in controlling digestion of such small amounts. A yield of 3×10^5 unamplified recombinant phage was obtained for each chromosome, which should be sufficient to ensure that the majority of sequences were represented in the libraries. The inserts of all phage examined to date range in size between 2.5 and 14 kilobase pairs. A background of parental phage (about 10%) was estimated by ligation and packaging of cloning arms alone.

II. Isolation and Assignment of Single-Copy Sequences to Chromosome 22

Single-copy sequences required for linkage analysis were initially identified by isolating phage DNA from 20 recombinants selected at random. The DNAs were labelled with P^{32} by nick translation [12] and hybridised to nitrocellulose strips containing *Eco*RI-digested total DNA from GM 1416 cells. All the probes hybridised to the filters, indicating that they contained cloned human sequences. The hybridisation from 13 probes produced smears, indicating that they contained sequences repeated with varying frequencies in the human genome [13]. In two cases, complex multiple bands were observed. However, the remaining clones hybridised to discrete DNA fragments identical in size with their human DNA inserts (Fig. 2A). This verifies that

the sorting and cloning procedures have not altered the chromosomal DNA. Based on these hybridisation data, we initially selected these five recombinants as single-copy probes for further characterisation and mapping. The assignment of these sequences to chromosome 22 was determined by hybridisation to DNA from sorted chromosomes. Results for two of the clones are shown in Fig. 2B. DNA from *λ22-5* and *λ22-6* was labelled separately by nick translation, mixed and hybridised to a filter containing DNA from four sorted fractions corresponding to chromosomes 22, 21, 1–8 and 9–20. Both probes hybridise to fragments of the appropriate size only in the fraction containing chromosome 22 (Fig. 2B). This illustrated the high degree of purity of the sorted chromosomes, and allowed us to assign the five clones to chromosome 22. The presence of these sequences on the chromosome has been independently confirmed by hybridisation to a series of hybrid cell lines. The probes hybridise to only those lines containing human chromosome 22 (data not shown).

III. Localisation of Clones on Chromosome 22

We have utilised three approaches to localise the cloned sequences to specific regions of chromosome 22, based on methods used to map globin genes [14] and non-coding sequences on chromosome 11 [6]. First, we have sorted the two respective portions of one homologue of chromosome 22 from the translocated cell line GM 3197: 46, XX, *t(17;22)(p13;q11)*. The flow karyotype of this line showed that the balanced translocation produced derivative chromosomes that were sufficiently different in size from the normal homologue that they could be identified and sorted on the FACS II (Fig. 1C). By hybridising each of the probes to Southern blots of DNA from the two sorted portions of chromosome 22, the sequences were localised to a region of the chromosome. For example, clone *λ22-1* hybridises only to total DNA and the *22q-* derivative placing it between *pter-q11* (Fig. 3A); while clones *λ22-5* and *λ22-6* are detected only in the total and 17:22 derivative fractions, mapping them to *q11-qter* (Fig. 3B).

In a similar manner, we have used DNA from several human-hamster cell lines derived from the fibroblasts of a human female carrying a *t(X;22)(q13;q112)* reciprocal translocation [14]. The hybrid lines contain either a normal homologue of human chromosome 22 or a segregant of one of the two derivative chromosomes (*X/22, 22/X*) free of its normal counterpart. Probes *λ22-5* and *λ22-6* hybridised to the lane containing *X/22* DNA not *22/X*, placing them between *q112-qter* (Fig. 4A, B). Clone *λ22-1* again showed the reverse pattern, localising it to the *pter-q112* region (Fig. 4C). Since the breakpoints of the *X;22* and *17;22* translocations mapped in the same general region of chromosome 22 (*q112* vs *q11*), we expected the results to be similar. They were identical for each of the five probes, confirming the map assignments made using the sorted translocations. In situ hybridisation [15] was also used to map clone *λ22-5* on chromosome 22. The hybridisation of this clone to normal human chromosomes is presented in Fig. 5A as a histogram of grain count versus chromosome number and it can be seen that there is a distinct peak over chromosome 22. This clone was also hybridised against the chromosomes of a cell line which bore the translocation *t(11;22)(q25;q13)* and it can be seen in Fig. 5B that there is a peak of hybridisation over both the normal chromosome 22 and over the chromosome 11/22 but not over 22/11. Thus it can be concluded that clone *λ22-5* lies distal to the breakpoint in *q13* on chromosome 22. Figure 7 summarises the results and relative map positions of these clones on chromosome 22.

IV. Expressed Sequences

We have screened the chromosome 22 library for coding sequences. A cDNA prepared from polyadenylated RNA isolated from the white cells of a patient with chronic myeloid leukaemia was used to probe an aliquot of the phage. Recombinants giving strong positive signals with this probe were picked for further analysis. A Northern blot of one of these clones (*λ22-4E*) against polyadenylated RNA from several types of human cells shows that it is expressed in many cell types, and suggests that it codes

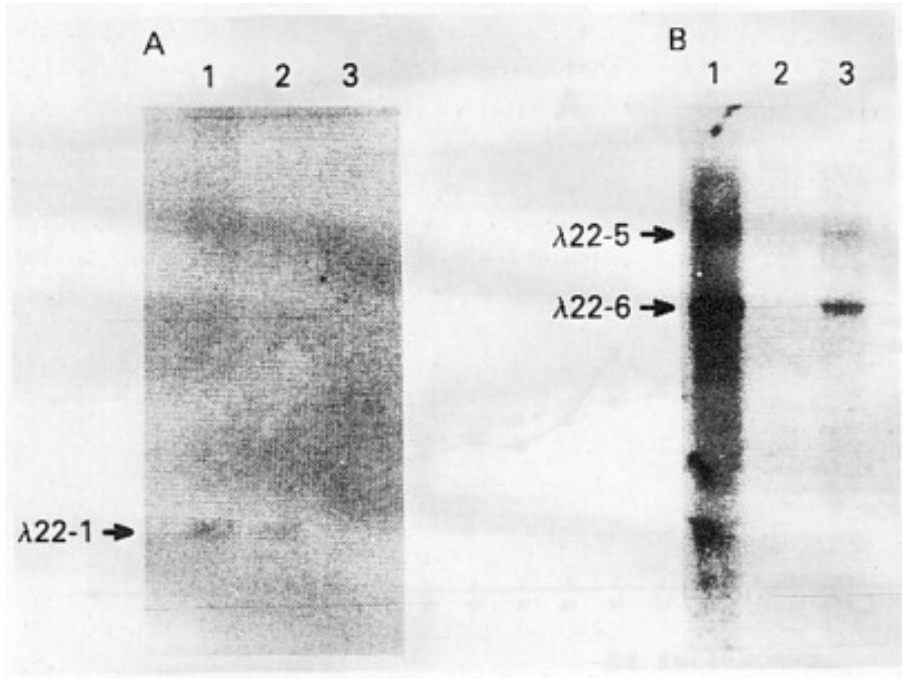

Fig. 3A, B. Mapping probes with sorted translocations. Cell line GM3197 beariing the translocation *t(17;22)(p13;q11)* was used for sorting 7×10^5 copies of the two portions [*22q-*(*lane 2*) and *17; 22*(*lane 3*)] involved in the translocation, and the normal homologue of chromosome 22 (*lane 1*) as indicated in Fig. 10. The filter containing *Eco*RI-digested DNA from these fractions was first hybridised with labelled phage DNA from clone *22-1* **A**, then melted and rehybridised with mixed probes from clones *λ22-5* and *λ22-6* **B**

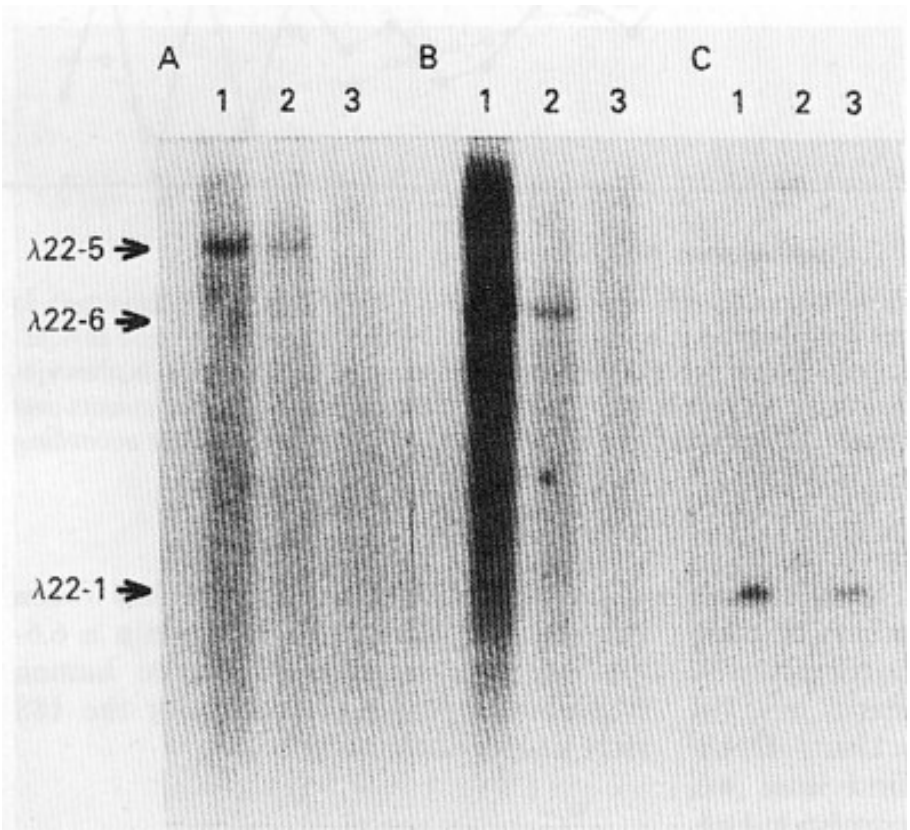

Fig. 4A–C. Mapping probes with hybrid cell lines containing portions of chromosome 22 resulting from the translocation *t(X;22)(p13;q112)*. Filters containing *Eco*RI digests (20 μg/lane) of human DNA (*lane 1*), and DNA from hybrid lines S and H (*lanes 2 and 3* respectively) carrying the *X/22* and *22/X* portions of the translocation were hybridised to the nick translated phage *λ22-5* **A**, *λ22-6* **B** and *λ22-1* **C**

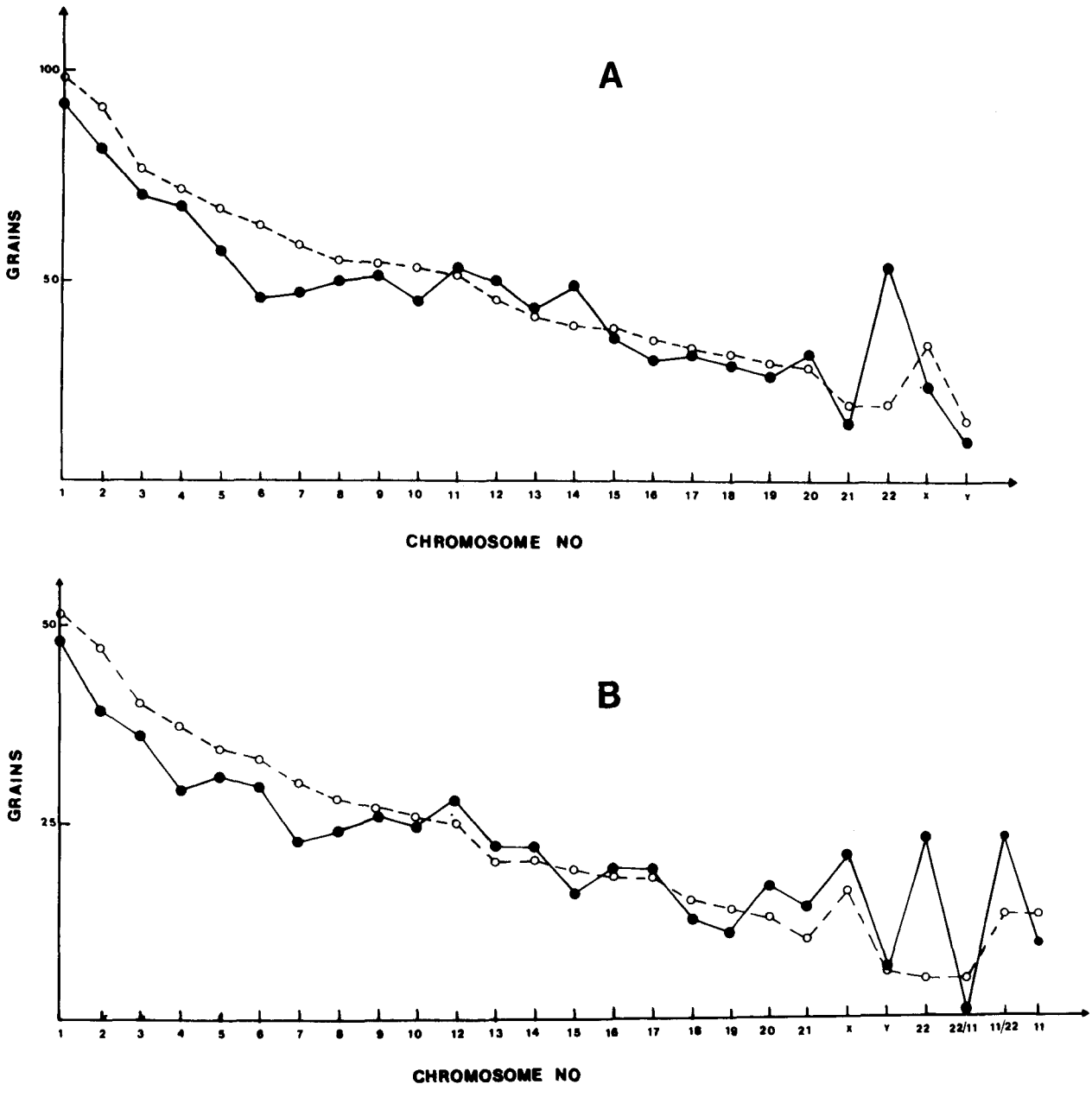

Fig. 5 A, B. Results of in situ hybridisation of nick translated *λ22-5* phage DNA (10^7 dpm/μg) to metaphase chromosomes of normal lymphocytes **A** and to those of a lymphoblastoid cell line **B** bearing the translocation *t(11;22)(q25;q13)*. Grains were counted over 39 normal metaphase lymphocytes **A** and over 15 translocation-bearing lymphocytes **B.** The *solid line* represents actual grain counts and the *broken line* represents the expected values if the grains had been randomly distributed according to chromosome length

for an 8S mRNA (Fig. 6 B). The probe hybridises to a variety of fragments in total human DNA, besides the one corresponding to its insert, suggesting that is may be encoded by multiple genes. Using DNA from the X/22 somatic hybrid lines we have mapped the sequence corresponding to the *λ22-4E* insert to the *q112-qter* region of chromomosome 22 (Fig. 6 A). Finally, chromosome 22 has a nucleolar organiser and we have identified several 18S and 28S rRNA clones using a ribosomal probe. One clone in particular, 22-18S, has been characterised and found to contain a 6.6-kilobase pair fragment from the human ribosomal repeat unit coding for the 18S rRNA (data not shown).

V. Localisation of Clones on Chromosome 21

In a manner similar to that described above for the chromosome 22 library, 96 clones

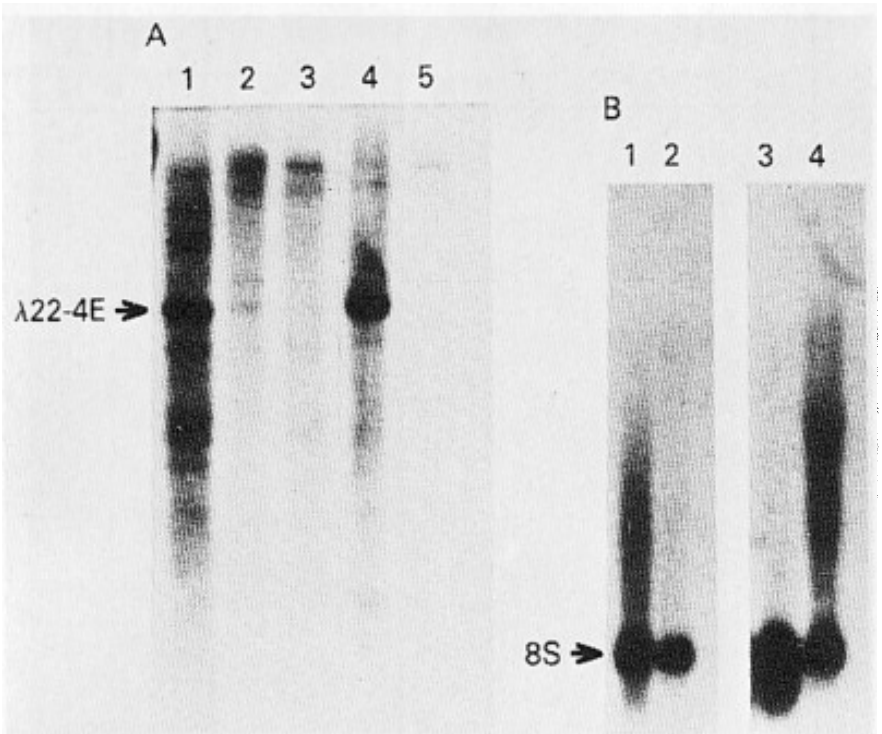

Fig. 6 A, B. Characterisation of an expressed clone. **A** Clone *λ22-4E* was mapped by hybridisation to a filter with *Eco*RI digests (20 μg/lane) of human DNA (*lane 1*), Chinese hamster DNA (*lane 5*) and DNA from the hybrid cell lines M, H and S containing various portions of chromosome 22 [*22, 22/X (lane 2), 22/X (lane 3), X/22 (lane 4)*]. **B** The labelled phage, *λ22-4E,* was also hybridised to a filter containing 3 μg of total polyadenylated RNA isolated from a patient with chronic granulocytic leukaemia (*lane 1*), a patient with chronic lymphatic leukaemia (*lane 2*), cell line K562 (*lane 3*) and the MRC-5 human fibroblast line (*lane 4*)

selected at random were screened with nick translated human DNA. Twenty clones did not hybridise, indicating that they may have had single copy inserts. Six of these twenty clones were found to have single copy inserts and hybridised to *Eco*RI-digested human and mouse DNA and also to DNA from a rodent-human cell hybrid which contained only chromosome 21. As shown in Fig. 7, three of these clones (B3, H8 and D4) hybridised to the cell hybrid DNA, thus localising them to chromosome 21. Of the other three clones, two do not appear to be on chromosome 21 (A6 and B4) and one apparently does not contain human DNA.

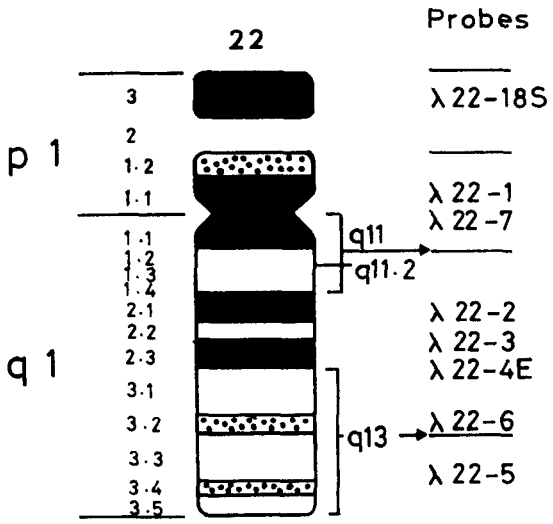

Fig. 7. Summary of mapping results on chromosome 22. Schematic representation of chromosome 22 showing the breakpoint positions of the *X/22 (q112), 17/22 (q11)* and *11/22 (q13)* translocations used for mapping as described in the text. The relative positions of the clones are indicated.

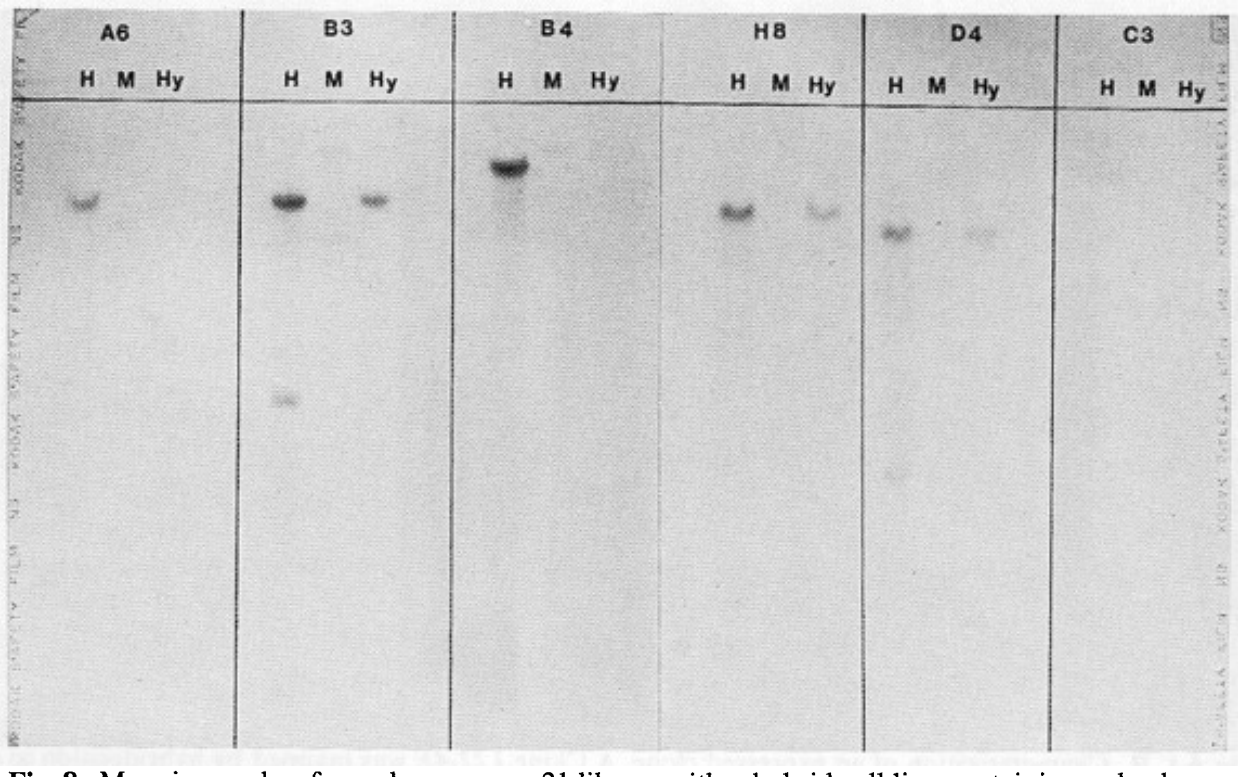

Fig. 8. Mapping probes from chromosome 21 library with a hybrid cell line containing only chromosome 21. Filters containing *Eco*RI digests of human DNA (*H*), mouse parent cell line BW5147 (*M*) and the hybrid cell line Thy B133R (*Hy*) were hybridised to six different nick translated phages

D. Discussion

This study on chromosomes 21 and 22 shows that by using the fluorescence-activated cell sorter to isolate purified fractions of specific human metaphase chromosomes, it is possible to construct DNA libraries representing single human chromosomes. The construction of a chromosome 22 library as a total *Eco*RI digest has enabled us to rapidly isolate and localise single-copy sequences on the chromosome. Analysis of the chromosome 22 library so far indicates that the background of parental phage is low, since 50 recombinants selected at random have all been found to contain human DNA inserts. Of these, ten single-copy clones have been identified and nine have been shown to be on chromosome 22 and one on chromosome 21. The map positions of eight of these sequences are summarised in Fig. 7. This illustrates the purity of chromosomes obtained with the sorting procedure, and suggests that the majority of sequences in this library are derived from chromosome 22.

This approach is simple and direct, offering several advantages over methods utilising total genomic libraries [16] or libraries prepared from human-rodent hybrid cell lines [6]. Relatively small numbers of phage are usually handled, since prescreening steps to identify human clones among a rodent background are not necessary. The level of purity of the sorted chromosomes ensures that most of the clones are derived from the same chromosome. Repetitive sequences can be isolated and characterised regardless of their reiteration frequency or cross-species hybridisation. Sequences on chromosomes that exhibit hybrid instability or lack selectable enzyme markers may also be isolated, since hybrid cell lines do not need to be constructed. Furthermore, since this library consists of sequences representing only 1.5% of the genome, relatively little phage (10,000) needs to be screened to isolate genes located on the chromosome. Of particular interest are the λ-immunoglobulin light chain genes, located on chromosome 22 [17], which are possibly involved in the translocation between chromosomes 8 and 22 in a variant form of Burkitt's lymphoma [18]. A second 22 library prepared following complete digestion with *Bam*HI will allow us to chromosome walk from any sequences of interest identified from the first library.

9. Carrano AV, Gray JW, Langlois RG, Burkhart-Schultz K, Van Dilla MA (1979) Proc Natl Acad Sci USA 76:1382–1384
10. Young BD, Ferguson-Smith MA, Sillar RS, Boyd E (1981) Proc Natl Acad Sci USA 78:7727–7731
11. Maniatis T, Hardison RC, Lowry E, Lauer J, O'Connell C, Quon D, Sim GK, Efstradiatis A (1978) Cell 15:687–701
12. Rigby RWJ, Diekmann M, Rhodes C, Berg P (1977) J Mol Biol 113:237–254
13. Schmidt CW, Deininger PL (1975) Cell 6:345–358
14. Lebo RV, Carrano AV, Burkhart-Schultz K, Dozy AM, Yu LC, Kan YW (1979) Proc Natl Acad Sci USA 76:5804–5808
15. Malcolm S, Williamson R, Boyd E, Ferguson-Smith MA (1977) Cytogenet Cell Genet 19:256–261
16. Wolf SF, Mareni C, Migeon B (1980) Cell 21:95–102
17. Erikson J, Martinis J, Croce CM (1981) Nature 294:173–175
18. Klein G (1981) Nature 294:313–318
19. Davies KE, Young BD, Elles RG, Hill ME, Williamson R (1981) Nature 293:374–376
20. Kunkel LM, Tantravahi U, Eisenhard M, Latt SA (1982) Nucl Acids Res 10:1557–1578
21. Disteche CM, Kunkel LM, Lojewski A, Orkin SH, Eisenhard M, Sahar E, Travis B, Latt SA (1982) Cytometry 2:282–286
22. Gray JW, Langlois RG, Carrano AV, Burkhart-Schultz K, Van Dillen MA (1979) Chromosoma 73:9–27
23. Rowley JD (1973) Nature 243:290–293