A 3.5 genome equivalent multi access YAC library: construction, characterisation, screening and storage

R.Anand, J.H.Riley, R.Butler, J.C.Smith and A.F.Markham
ICI Diagnostics, Gadbrook Park, Northwich, Cheshire CW9 7RA, UK

Received February 16, 1990; Revised and Accepted March 20, 1990

ABSTRACT

The construction of a yeast artificial chromosome (YAC) primary gridded library of 35,000 clones from human lymphoblastoid (48,XXXX) cell line DNA is described. The average YAC size is ~350kb representing a >3.5 times coverage of the genome. The library is stored at −70°C as gridded clones on nylon filters impregnated with 20% glycerol and as glycerol suspensions of individual clones in microtitre plates providing a prolonged multi-user potential. To date we have used 14 single copy probes to screen this library by colony hybridisation as well as PCR and have isolated between 1 and 5 YAC clones for every probe.

INTRODUCTION

Until very recently, the analysis of complex genomes had been severely restricted by the size of DNA fragments that could be fractionated (~50kb on agarose gels) and cloned (~40kb in cosmids). A large void existed between the resolving power of these techniques and conventional genetic analysis with resolution in centi Morgans (cM) which in molecular distance roughly equals a megabase (Mb). A major advance towards bridging this gap was made in 1984 by Schwartz and Cantor when they described the technique of pulsed field gel electrophoresis (PFGE) which was capable of resolving several hundred kilobase fragments of DNA. This technique has since been significantly improved and is now known to be capable of resolving several megabase fragments of DNA [for reviews see: 1, 2]. Another advance towards the construction of long range maps was the concept of jumping and linking libraries [3, 4]. However, the inability to move from long range maps to the detailed analysis of any long stretch of DNA was still a restraining factor. This was resolved following the description of an elegant vector system capable of cloning several hundred kilobase fragments of DNA as linear artificial chromosomes in yeast [5]. There were some technical problems associated with this cloning system. Firstly, the yeast cell wall needed to be partially dissolved to enable the artificial chromosome construct to enter the cell during transformation. This meant that following transformation, the cells had to be allowed to recover by plating them in a high concentration of supporting top agar. As a consequence, most recombinants grew below the surface of the agar making conventional colony lift and hybridisation screening impossible. One solution to this problem was to allow the recombinants to recover and grow, scrape off the top agar containing them, make an emulsion and then plate out on agar or on a filter as an amplified library. However, due to large variations in colony sizes, there are likely to be serious problems with representation. For this reason, primary libraries need to be constructed as gridded arrays of recombinants such that all clones are equally represented. A second drawback was that although the system was capable of replicating large fragments of DNA, manipulating such large DNA fragments in the liquid phase prior to transformation and keeping them intact up to the transformation stage proved to be very difficult. The result has been that most libraries have an insert size of 100 to 200kb [6−11].

It is advantageous to have large inserts for analysing complex genomes as well as for cloning large genes. To this end we had described a method of removing small fragments by PFGE fractionation in order to achieve a larger average YAC size [12]. We have now used this method to construct a primary gridded library with an average size of 350kb. Preliminary studies suggest a good overall representation of the 48,XXXX genome since all 14 probes tested to date have yielded the relevant YAC clones.

MATERIALS AND METHODS

The artificial chromosome cloning vector pYAC4 which has a single EcoRI cloning site and Saccharomyces cerevisiae AB1380 (MATaψ+, ura3, trp1, ade2-1, can1-100, lys2-1, his5) host strain were used [5]. Agarase was from Calbiochem, lyticase was from Sigma and Zymolase-20T was from ICN Biomedicals. Restriction enzymes and calf alkaline phosphatase were from Boehringer Mannheim and T4 DNA ligase was from Amersham International. All reagents for media preparation were from Difco laboratories and Sigma. The vacuum dialysis apparatus and UH100/75 ultra thimbles were from Schleicher and Schüell.

GM1416 (48,XXXX) lymphoblastoid cells from the National Institute of General Medical Sciences human genetic mutant cell repository (Camden, New Jersey, USA) were cultured and karyotyped to confirm their chromosomal content. High molecular weight DNA in 0.5% low gelling temperature agarose plugs was prepared at 1.5×10^7 cells/ml [13, 14]. Each full plug contained ~1.5×10^6 cells equivalent to ~10μg DNA.

Yeast cells were grown without selection in YPD [15] or with selection in synthetic medium (SD) containing 0.7% yeast nitrogen base without amino acids, 2% glucose and 5.5mg/100ml of adenine and tyrosine, pH7.0. For single (-ura) selection, 7ml of 20% casamino acids solution and 2ml of 1% tryptophan was added to every 100ml whereas for double selection (-ura, -trp),

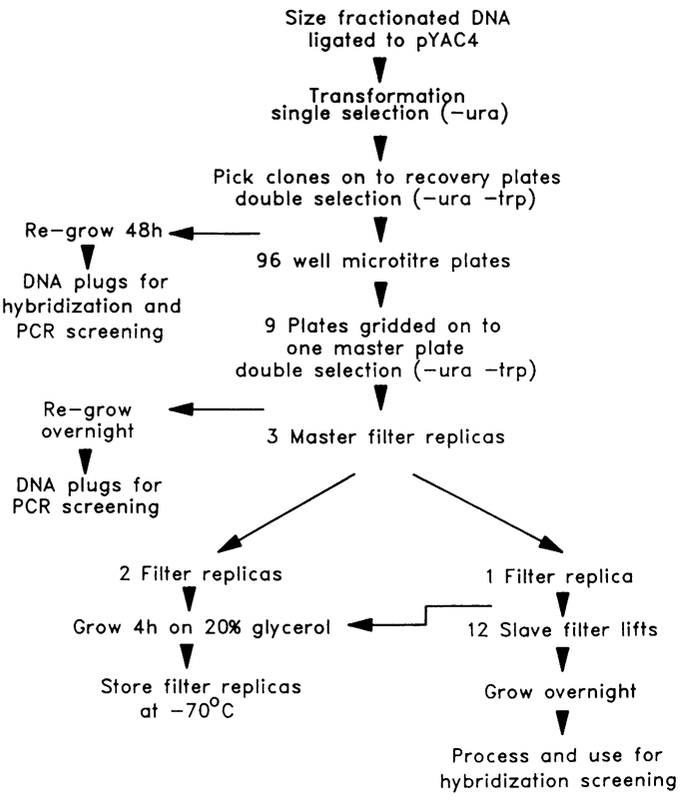


Figure 1. A summary of the steps involved in the construction of a primary yeast artificial chromosome library.

only 7ml of 20% casamino acids/100ml was added. Recovery plates were made from 2% bacto agar in double selection medium. Plates for transformation were made from 2% bacto agar in single (-ura) selection medium containing 1M sorbitol. Top agar for plating out the recombinants was 3% bacto agar in single (-ura) selection medium containing 1M sorbitol.

Yeast chromosomal DNA

High molecular weight DNA was prepared in agarose plugs for PFGE analysis as well as all other procedures. Essentially, yeast cells grown with or without single or double selection in liquid culture or on agar recovery or master plates were harvested. They were washed once in 50mM EDTA and resuspended to $3-5\times10^8$ cells/ml in yeast resuspension buffer (YRB; 1.2M sorbitol, 20mM EDTA, 10mM Tris-HCl pH7.5 and 14mM β-mercaptoethanol) containing 20 units/ml lyticase. Cells were incubated at 37°C and generally ~80% were spheroplasts after 1h. This was monitored by mixing a small aliquot of cells with 10% SDS on a microscope slide and examination under a phase contrast microscope. Spheroplasts appear as dark ghosts compared to bright and shiny intact yeast cells. Cells were gently mixed 1:1 with a 1% solution of LGT agarose in YRB at 37°C and transferred to a plug mould to set [14]. Plugs were transferred to 2−3 volumes of yeast lysis buffer (YLB; 0.1M EDTA, 10mM Tris-HCl pH8.0 and 1% Lithium dodecyl sulphate) and gently shaken at room temperature for 1−2h. The solution was replaced with 10 volumes of YLB and incubated at 40−50°C overnight. Plugs were rinsed once in TE and stored in YLB at room temperature. The DNA was essentially intact and stable by PFGE analysis. Partial or complete restriction digests could be carried out using standard protocols. These plugs could not be stored

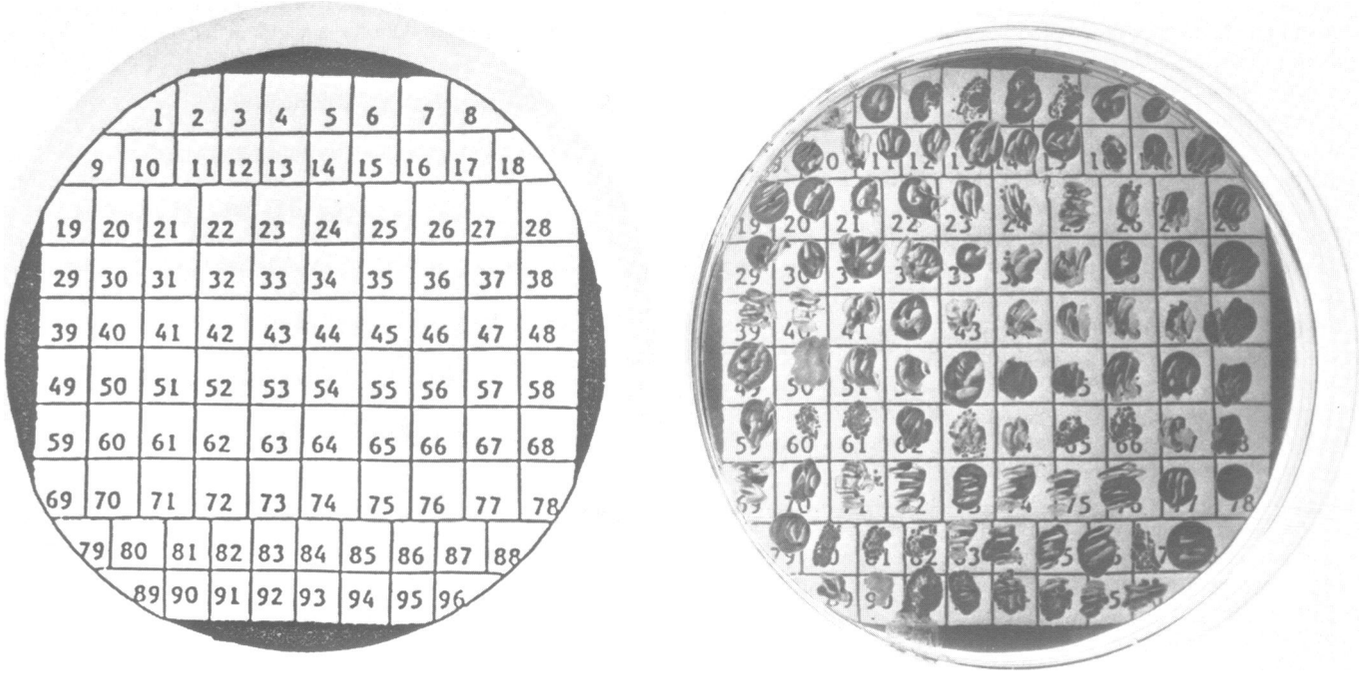


Figure 2. A 96 position grid used with 9cm circular recovery plates, and a double selection (-ura, -trp) recovery plate showing the YAC clones recovered from a primary transformation plate. Two of the clones (No. 50 and 90) were 'white' presumably because there was no interruption of the vector's SUP4 gene responsible for the red colour. These were therefore regarded as non-recombinants.

at 4°C because of possible precipitation of lithium salts. For storage at 4°C the plugs were first equilibrated in 0.1M EDTA, pH9.0.

Fractionation and ligation

These were essentially as described elsewhere [12]. Briefly, EcoRI partially restricted GM1416 cell line DNA was size fractionated using the 'Waltzer' PFGE system [16]. DNA fragments >200kb were recovered, purified and ligated to EcoRI/BamHI restricted and dephosphorylated pYAC4 vector DNA in a 1:1 (w/w) ratio (>20 fold molar excess of vector). The ligated DNA was repurified and concentrated ready for transformation. This DNA was stable at 4°C for several months.

Transformation

AB1380 cells were streaked out on a YPD agar plate and grown at 30°C for 24−48h. A single colony was used to inoculate a YPD culture and grown overnight at 30°C. The cells were harvested at an OD_{600} of 0.7−0.8. The cells were made competent for the uptake of DNA essentially as described by Burgers and Percival [17]. The level of spheroplasting was monitored as described above. Approximately 1μg of ligated DNA in 30μl TE was used with 700μl spheroplasts (from 17.5ml of yeast culture). PEG 6000 was used in all transformations and the cells from each 30μl DNA transformation were plated out on two 9cm diameter plates under single (-ura) selection. Plates were incubated at 30°C for 3−4 days after which yeast colonies were clearly visible. Uncut pYAC4 DNA was used as a control to monitor transformation efficiency.

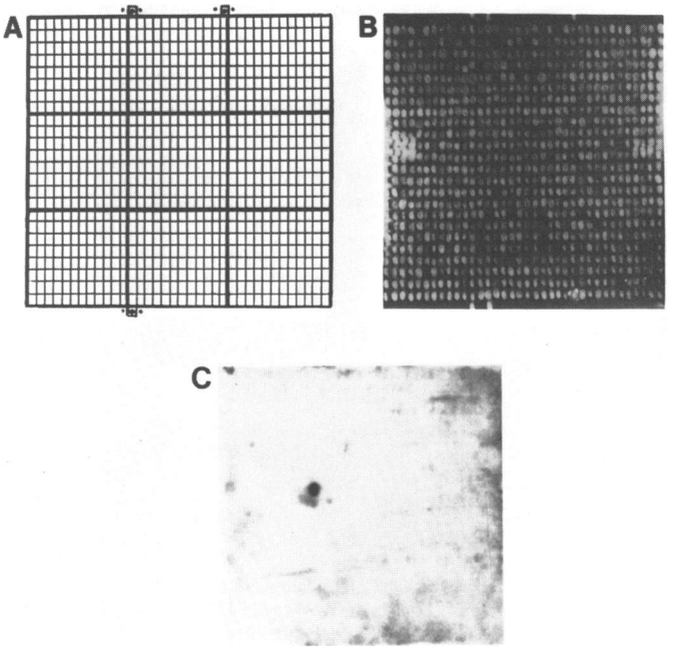


Figure 3. A 9×96 position asymmetric grid (A) used with 10×10cm double selection (-ura, -trp) master plates. A colony lift of the gridded array of 864 recombinant clones on a 9.5×9.5cm master filter (B) after overnight growth at 30°C. An autoradiogram (C) showing positive hybridisation of a single copy sequence probe to a slave filter. There is sufficient background hybridisation to accurately localise the positively hybridising clone amongst the 864 recombinants.

RESULTS

The transformation efficiencies using the same batch of ligated DNA varied from 200−1000 clones per μg. Increasing the DNA concentration or additional carrier DNA did not cause a consistent improvement and therefore, the variation is likely to arise either from within the colonies used to start the AB1380 cultures or in the processing steps to make the cells competent.

All stages in the construction of this primary library are outlined in Figure 1. Clones from the primary transformation plates were individually picked on to double selection (-ura, -trp) recovery plates. Each 9cm circular recovery plate had a 96 position grid attached to its bottom surface so that 96 individual clones could be streaked on each plate. These were grown at 30°C over 36−48h so that there were large numbers of cells for each clone (Figure 2). This also allowed for clear expression of the red colour, and based on this we estimate that > 95% of all colonies were recombinant and contained YACs. Cells from each clone expressing the red colour were scraped off using the broad flat end of sterile wooden toothpicks and transferred to the wells of microtitre plates containing 100μl SD + 20% glycerol. To ensure that each recovery plate contained cells from all 96 recombinant clones in the corresponding microtitre plate, any non-recombinant (white) clones were replaced by recombinant (red) clones from back up recovery plates. This was done by transferring cells to the blank wells of the microtitre plate as well as streaking some cells on to the blank areas of the recovery plate. The recovery plate was regrown at 30°C for 48h, cells harvested and DNA plugs made as described in the Methods section. These plugs contained chromosomal DNA from 96 recombinant clones.

Recombinant clone stocks in microtitre plates were used to grid master plates prior to being frozen and stored at −70°C. Master plates were 10×10cm recovery agar (-ura, -trp) plates with a 9×96 asymmetric grid (Figure 3) stuck on the bottom of the plate. Clones were spotted in the small rectangles on the grid such that the position of a clone on each of the 40 master plates corresponded to a glycerol stock in a microtitre plate. The 9 microtitre plate positions on each master grid were labelled as A−I and therefore any clone could be identified as to which individual microtitre plate it came from. Microtitre plates are numbered as rows A−H and columns 1 to 12 and therefore once any positive clone was identified, it could be numbered (eg. Master 10, Microtitre D, Row F and Column 6 = 10DF6). Master plates were grown overnight and then master filter lifts were made on to 9.5×9.5cm Hybond-N nylon filter squares. This was done by pre-wetting boil washed and autoclaved filters on 10×10cm recovery plates and then placing them on the master plate. The filters were left on the plate for 2−4 min, carefully peeled off and placed on recovery plates with the colony side up. Three master lifts were made from each master plate and then the cells from the master plate were harvested and made into DNA plugs as described in the Methods section. The three master filters were grown on recovery plates overnight. Two of these were grown for another 4h on recovery plates containing 20% glycerol and then sealed in 10×10cm dry petri dishes and stored frozen at −70°C. The third master filter was used to make 12 replicas as described by Coulson et al. [18]. The only difference was that boil washed Nytran nylon membranes were used for all replicas as this gave the most reproducible hybridisation signals. The replica filters were grown for 36−48h at 30°C and processed as described [18] except that we used 50 units/ml lyticase instead of Zymolase.

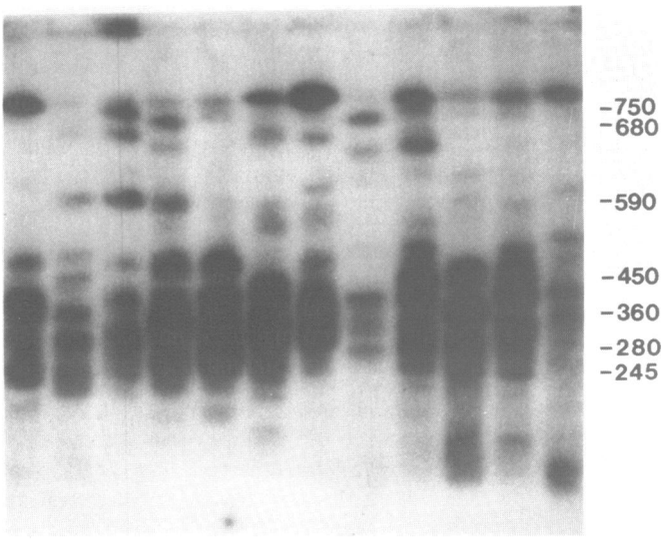


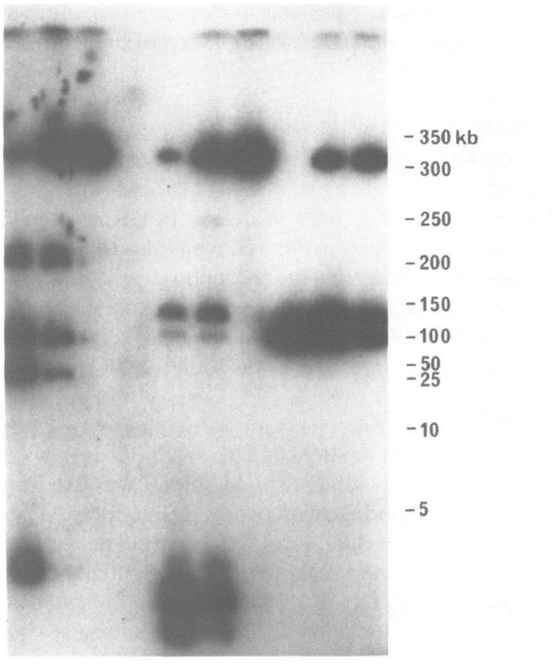
Figure 4. Twelve recovery plate DNA plugs representing 1152 (12×96) recombinant clones were fractionated on a 'Waltzer' PFGE apparatus and a blot of this gel was hybridised to ^{32}P-dCTP labelled pBR322. All YACs hybridise to the probe because of sequence homology to the vector pYAC4, demonstrating the scatter of sizes of the artificial chromosomes. The peak size distribution is between 250−450kb.

YAC Size

To estimate the average YAC size of the library, 100 random clones were picked, grown up and made into plugs. These were analysed by PFGE and blots hybridised to ^{32}P-dCTP labelled total human DNA [19] to identify all the individual YACs. The average size was thus estimated to be ~350kb. The distribution of YAC size in this library was also analysed by PFGE of 12 recovery plate plug preparations, each representing 96 clones. A blot of this gel representing 12×96 = 1152 recombinant clones was hybridised to ^{32}P-dCTP labelled pBR322 to show all artificial chromosomes. Figure 4 shows that the peak size distribution is mainly between 250−450kb with an overall range of 100 to >800kb. The same hybridisation pattern was observed when the filter was stripped and reprobed with ^{32}P-dCTP labelled total human DNA.

Screening

We have to date screened this library for 14 different single copy sequences. Some of these were RFLP markers linked to diseases like Friedreich Ataxia and Cystic Fibrosis whereas others were cDNA/genomic probes from small genes like aldose reductase to the very large ~2.3Mb gene for Duchenne muscular dystrophy. Screening was performed by hybridisation of all 40 master filter replicas in duplicate. Three to five probes were labelled by random priming [20] using 100ng of each probe and ^{32}P-dCTP to achieve a final specific activity of >4×10^{8}dpm/μg. Pooled probes were denatured, reassociated with total human DNA to Cot 10 [21] and used for hybridisation in 150ml hybridisation solution (0.5M sodium phosphate pH7.5, 5×Denhardt's, 1% SDS) at 65°C for 48−72h. Filters were washed down to 0.5×SSC at 65°C and autoradiographed for 2−4 days at −70°C with Dupont Cronex lightning plus intensifying screens. Duplicate positives were usually localised to one of the nine microtitre plates comprising a master. Sometimes however, a clear positive could be identified against a faint background which helped to localise the exact position of the clone within a microtitre plate (Figure 3).



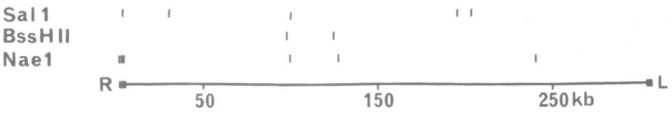


Figure 5. Restriction site mapping of a YAC. Restriction enzymes BssHII, NaeI and SalI were used at three different concentrations to achieve a range of partial digests of a 310kb YAC. These digests were fractionated by PFGE and a blot of this gel probed with a 1.69kb PvuII/BamHI fragment of pBR322, a probe specific for the right hand end (R) of YACs. The autoradiogram shows restriction sites that can be mapped from R. A NaeI site which is 245kb from R is very faint and only shows up on a long exposure. However, this site is clearly visible when mapping from the left hand end (not shown). All sizes are in kilobase pairs and these were used to construct a map of the YAC as shown. The left (L) and right (R) vector arms are marked.

An alternative to the above hybridisation procedure was to screen the library using the polymerase chain reaction [22−24]. We used appropriate oligonucleotide primers (20−30mers) designed to amplify 200−500bp of specific loci from total human genomic DNA. Individual plugs from the 40 master plates were used in PCR amplifications to identify the master which contained the clone of interest. DNA plugs from the 9 microtitre plates comprising that master were then used for PCR amplifications to identify the microtitre plate containing the positive clone.

Once a microtitre plate had been identified as containing the positive clone, it was used to inoculate a fresh microtitre plate containing 100μl per well of double selection SD medium. A 96 prong well replication device (Cat. No. WR081, Denley Instruments Ltd, Billinghurst, West Sussex RH14 9EY) was used for these inoculations. The inoculated plate was grown overnight at 30°C in an orbital incubator. The well replicator was then used to transfer ~10μl of cells in SD to Nytran filters. These were grown on double selection SD plates for 24−48h before processing [18]. The individual positive clones were identified by hybridisation and streak purified. DNA plugs were made for further analysis. Alternatively, DNA plugs from all those microtitre plates containing positive clones were fractionated by

PFGE and blotted on to nylon filters. The blots could be probed sequentially with each one of the original mixture of probes at a time to identify those plates containing positives as well as the YAC sizes of the individual positive clones. This is particularly useful when several positive clones are identified by a single probe. In such cases, at least initially, only that microtitre plate containing the largest YAC need be replicated and have the positive clone purified.

Characterisation & Mapping

All purified clones were analysed by PFGE on 1.5% agarose gels to determine the YAC sizes as well as the number of YACs in each clone. Of the 30 positively hybridising clones analysed so far, 27 had only one artificial chromosome. The average size of YAC in these clones was ~360kb. To obtain a restriction map of the individual clones, plugs were cut into 1/3 for each digest. Up to 20 of these cut plugs were washed in 50ml of TE overnight followed by 50ml TE for 1h at 4°C. Plugs were equilibrated in 1.5ml restriction buffer for 1h at 4°C followed by equilibration in 100μl buffer + enzyme for 30min at 4°C and then incubation at the appropriate temperature for the enzyme for 30min. Enzyme concentrations were varied from 0.1 to 10 units/100μl to achieve a range of partial digests. Following fractionation by PFGE, the gels were blotted on to nylon membranes and sequentially probed with pBR322 probes specific for the left and right ends of the YACs. pBR322 restricted with PvuII and BamHI results in a 2.67kb fragment which hybridises specifically to the left (centromere and telomere) end and a 1.69kb fragment which hybridises specifically to the right (Ura and telomere) end. An example of the hybridisation pattern and the resultant map is shown in Figure 5.

We have so far mapped 15 YACs from 5 different loci. Of these, only one YAC from the D7S8(J3.11) region had a restriction map which did not correspond to the maps obtained from the other 3 overlapping clones from the same region. The nature of rearrangements in this clone have not been characterised since there were other clones covering the same region. In all cases where multiple clones were isolated using a single probe, the YACs did not show any clustering of start or end points suggesting that there was no restriction site bias in the cloning procedure. Finally, the largest region covered by clones (contig) isolated from a single probe was ~800kb.

While this manuscript was in preparation, a method of plating transformed spheroplasts in 2% sodium alginate was described [8]. This method circumvents the problem of recombinants growing inside the top plating agar. The average YAC size of this 30,000 clone library is only 100kb with a range of 70–150kb and only one positive clone has so far been isolated after screening with six different probes. We are currently investigating the plating efficiencies of transformants containing large YACs in sodium alginate.

DISCUSSION

Mapping and eventual sequencing of the human genome is now a real possibility. There are already several PFGE maps covering many megabases of DNA at various loci. However, it is clear that there are serious discrepancies between maps constructed in different laboratories using different apparatus, cell lines and DNA concentrations. Although some of these discrepancies are due to the inherent problem of variations in DNA methylation, there are other more technical problems as well. Cloned DNA can yield much more accurate and detailed maps and therefore it would seem prudent to obtain maps using cloned DNA, compare these to the genomic maps and then arrive at a consensus. Furthermore, cloned DNA lends itself to various manipulations including subcloning single copy regions, searching for coding sequences and eventually sequencing. However, since there are differences between DNA obtained from different sources, it might be helpful to initially use only a limited number of gridded YAC libraries. This should minimise the effects of different starting DNA used to construct the libraries thus making it feasible to arrive at a consensus map.

The representation of the library described herein seems adequate since all probes tested so far have yielded YAC clones. The large average insert size allows for fast walking along the genome from linked markers to gene loci, possibly generating new polymorphic markers in the process. Large YACs are also useful in linking markers several hundred kilobases apart as well as in cloning large genes. Restriction site maps of positive clones can be constructed using YAC end-specific pBR322 probes. Based on these maps, insert ends can be identified and isolated for subsequent walks. Insert ends can also sometimes be PCR amplified using vector end specific and Alu consensus primers [25]. These ends can then be used as probes for chromosomal localisation by hybridisation to blots of chromosome specific hybrid panels or by *in situ* hybridisation. Localisation of both ends of a YAC to a single chromosome reduces the probability of the insert having arisen as a result of an anomalous ligation event between two non-contiguous genomic fragments. We are currently using these and other strategies to isolate insert ends and construct large YAC contigs.

ACKNOWLEDGEMENTS

We would like to thank R. Anwar, D. Black, L. Brown, P. Elvin, R. Finniear, L. Fletcher, A Graham, K. Hart, P. Hedge, B. Hopkins, D. Jenner, K. Murray, D. Ogilvie, W. Potts, S. Powell and G. Slynn for their invaluable help in manually picking clones from primary to recovery plates. We would also like to thank P. Heath for karyotyping the GM1416 cell line, S. Newton for the illustrations and D. Moore for typing this manuscript.

REFERENCES

1. Anand, R. (1986). *Trends in Genetics*, **2**, 278–283.
2. Dawkins, H.J.S. (1989). *J. Chromatography* **492**, 615–639.
3. Collins, F.S. and Weissman, S.M. (1984). *Proc. Natl. Acad. Sci. USA* **81**, 6812–6816
4. Poustka, A. and Lehrach, H. (1986). *Trends in Genetics* **2**, 174–179.
5. Burke, D.T., Carle, G.F. and Olson, M.V. (1987). *Science* **236**, 806–812.
6. Carle, G.F., Burke, D.T., Olson, M.V., Little, R. and Schlessinger, D. (1987). *Cytogenet. Cell Genet.* **46**, 590.
7. Guzman, P. and Ecker, J.R. (1988). *Nucleic Acids Res.* **16**, 11091–11105.
8. Traver, C.N., Klapholz, S., Hyman, R.W. and Davis, R.W. (1989). *Proc. Natl. Acad. Sci. USA* **86**, 5898–5902.
9. Monaco, A.P., Larin, Z., Bates, G. and Lehrach, H. (1989). *Cytogenet. Cell Genet.*, in press.
10. van Ommen, G.J.B., den Dunnen, J.T., Verkerk, J.M.H., Kievits, T., Bakker, E., Blonden, L.A.J., Breuning, M.H., Driessen, M.S., Fishbeck, K.H., Grootscholten, P.M., Oostra, B.A., Skraastad, M.I., Wapenaar, M.C. and Pearson, P.L. (1989). 'Genome mapping and sequencing,' Abstract No. 120, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
11. Brownstein, B.H., Silverman, G.A., Little, R.D., Burke, D.T., Korsmeyer, S.J., Schlessinger, D. and Olson, M.V. (1989). *Science* **244**, 1348–1351.
12. Anand, R., Villasante, A. and Tyler-Smith, C. (1989). *Nucleic Acids Res.* **17**, 3425–3433.
13. Schwartz, D.C. and Cantor, C.R. (1984). *Cell* **37**, 67–75.

14. Anand, R. and Southern, E.M. (1990). Pulsed field gel electrophoresis. In 'Gel electrophoresis of nucleic acids: a practical approach' (D. Rickwood and B.D. Hames, Eds.), in press, IRL Press, Oxford and Washington DC.
15. Lohr, D. (1988). Isolation of yeast nuclei and chromatin for studies of transcription-related processes. In 'Yeast: a practical approach' (I. Campbell and J.H. Duffus, Eds.), pp. 125–145, IRL Press, Oxford and Washington DC.
16. Southern, E.M., Anand, R., Brown, W.R.A. and Fletcher, D.S. (1987). *Nucleic Acids Res.* **15**, 5925–5943.
17. Burgers, P.M.J. and Percival, K.J. (1987). *Anal. Biochem.* **163**, 391–397.
18. Coulson, A., Waterston, R., Kiff, J., Sulston, J. and Kohara, Y. (1988). *Nature* (London) **335**, 184–186.
19. Southern, E.M. (1975). *J. Mol. Biol.* **98**, 503–517.
20. Feinberg, A.P. and Vogelstein, B. (1983). *Analyt. Biochem.* **132**, 6–13.
21. Sealey, P.G., Whittaker, P.A. and Southern, E.M. (1985). *Nucleic Acids Res.* **13**, 1905–1922. 22. Kleppe, K., Ohtsuka, E., Kleppe, R., Molineux, I. and Khorana, H.G. (1971). *J. Mol. Biol.* **56**, 341–361.
23. Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988). *Science* **239**, 487–491.
24. Green, E.D. and Olson, M.V. (1989). 'Genome mapping and sequencing,' Abstract No. 75, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
25. Nelson, D.L., Ledbetter, S.A., Corbo, L., Victoria, M.F., Ramirez-Solis, R., Webster, T.D., Ledbetter, D.H. and Caskey, C.T. (1989). *Proc. Natl. Acad. Sci. USA* **86**, 6686–6690.