Volume 13 Number 6 1985 Nucleic Acids Research

Removal of repeated sequences from hybridisation probes

Paul G.Sealey, Paul A.Whittaker and Edwin M.Southern

MRC Mammalian Genome Unit, King's Buildings, West Mains Road, Edinburgh, EH9 3JT, UK

Received 21 December 1984; Revised and Accepted 22 February 1985

ABSTRACT

Pre-reassociation of human clone probes, containing dispersed highly repeated sequences, (e.g. Alu and KpnI families), with a large excess of sonicated total human DNA allows signal from single and low copy number components to be detected in transfer hybridisations. The signal from non-dispersed repeated sequences is reduced to single copy levels. The procedure, which is simple and quick, is illustrated using model combinations of well characterised cloned probes, and is applied to a sample of randomly chosen cosmid clones. A theoretical assessment is presented which may be useful to those wishing to use this procedure.

INTRODUCTION

Repeated sequences are found dispersed throughout eucaryotic genomes (1). In humans, repeated sequences are found at intervals of a few thousand base pairs throughout at least 80% of the genome; much of the remaining 20% consists of tandem arrays of repeats (2).

All phage and cosmid vectors used in the construction of libraries of DNA have a lower size limit that is well in excess of the repeated sequence interspersion distance, consequently most clones contain at least one representative of a family of highly repeated sequences (3).

The presence of such repetitive elements limit the uses of clones as hybridisation probes to gel transfers of genomic DNA, where a smear of repeat hybridisation may obscure bands produced by low copy components. Lengthy subcloning procedures are often used to free clones of repeated sequences (see, for example, (4)).

This problem has been partly overcome by carrying out hybridisation at high stringency (5); sequence divergence

amongst members of a sequence family reduces the number of repeats capable of forming stable hybrids. Unfortunately this approach suffers from the disadvantage that the stringent hybridisation conditions used reduce the rate of hybridisation of the low copy sequences and lower the signal. Furthermore all repeat families are not highly diverged.

Another approach has been to add a large excess of competitor DNA (eg. sheared total human DNA) to the probe (4,6). Contrary to the expectation that competition would increase the ratio of low to high copy signal results have been reported as poor or irreproducible (6), although some workers have found it effective (4). Competition has been suggested as the cause of failure to detect poly dCdA repeats in eucaryotic DNA when DNA, from a different species, but also containing poly dCdA repeats, such as salmon sperm or calf thymus DNA, was used as carrier during the hybridisation (7).

A third approach has been to reassociate the clone with a large excess of competitor DNA, either in solution, or bound to diazonium cellulose, to a final Cot value just less than the Cot $\frac{1}{2}$ of the sequence of interest (6). The portion remaining single stranded was then separated from the reassociated fragments. The clone was presumed to have been stripped of its repeats, and the remaining DNA was then used as a low copy probe. The method was successfully used, but had the disadvantages that the separations could be time consuming and that a portion of the low copy DNA was lost along with the repeated DNA.

In this paper we describe a simplification of this procedure in which DNA hybridization probes are reassociated with a large excess of total genomic DNA, without removal of double stranded portions prior to hybridisation. DNA clones, representative of highly repeated dispersed human sequence families, have been used as probes in combination with a range of low copy number DNA probes; different degrees of pre-reassociation of the probe with total genomic DNA have been tested for their effectiveness in revealing the low copy sequence signal. A quick and simple routine has been devised which is illustrated using randomly chosen human cosmid clones as probes.

MATERIALS AND METHODS

Cloned DNA probes

The following were used:

BLUR 8 — a plasmid containing a member of the human Alu repeat family with a sequence close to the consensus Alu sequence (300bp insert in pBR322) (8).

pBK(1.5)54 — a plasmid containing a representative of the human Kpn repeat family (1.5kb insert in pBK) (9).

pBS22(X) — a plasmid containing a single copy sequence mapping to the human X chromosome (1.1kb insert in pBR328)(B.Smith, pers.comm.).

pH δ — a plasmid containing the human δ-globin gene (2.3kb insert in pBR322)(10).

pRBα1 — a plasmid containing the human α 1-globin gene (3.9kb insert in pBR322)(11).

PS2(21) — a bacteriophage clone, isolated from a sorted human chromosome 21 library provided by B.D.Young (8.9kb insert in λgtwes). It has both low copy number and repeated components; the low copy component maps to human chromosome 21, (data not shown).

Human cosmid — probes were isolated at random from an amplified cosmid library constructed using the cosmid vector pJB8, (12), cleaved at its BamHl site, and insert DNA, (30-45kb, MbOl partial digest), prepared from a cell line established with lymphocytes from a female human carrying an X:21 balanced translocation. The Y chromosome derived pJB8 clone, clF2, which was a gift from W.R.A.Brown, was prepared using DNA from a mouse/human hybrid cell line.

Isolation of probe DNA.

Plasmids and phage were prepared on a large scale using standard procedures, (13). The phage insert was prepared by Eco Rl digestion, (see below), followed by agarose gel electrophoresis, excision of the portion of the gel containing the desired fragment, and electroelution of the DNA. The DNA was

freed of gel contaminants, and concentrated, by passage through an Elutip-d, (Schleicher & Schuell). Cosmid DNA was prepared from 5ml cultures by the alkaline lysis procedure, (14), as modified by Ish-Horowicz & Burke, (12).

Human, mouse-human & mouse cell lines.

The following were used:

TIL1	an XXXXY human line.
THY B1-33/12	a mouse/human hybrid line in which the only human chromosome present is the X, based on both karyotype analysis & enzyme analysis.
THY B1-33-6-1	as above, chromosome 21 only.
HORL9X	as above, X chromosome only, separate origin.
ALR58r	as above, chromosomes 1,3,5,6,7,9,11,12,14,16, & 21 are present.
BW5147	a mouse only line used as a control with the hybrids above.

All the cell lines described above were supplied by C.Bostock & E.Clarke, with the exception of ALR58r, which was a gift from S.Christie.

Preparation of genomic DNA.

DNA from placental tissue of known sex was prepared by homogenisation in 0.3M Sodium trichloroacetic acid, 2mM EDTA, pH 7.0, followed by lysis in 1% Sodium dodecyl sulphate (SDS) and purification using standard methods. DNA from cultured cells was prepared by lysis in urea/SDS, (15), followed by purification using standard methods.

Restriction enzyme digestion.

All digestions were carried out using EcoR1, (Boehringer-Mannheim), in high salt buffer, (13), at 37^{0}C,in the presence of 100µg/ml gelatin & 4mM spermidine, for periods of 1 to 16 hours. The reaction was stopped and the sample prepared for electrophoresis in the usual way (16).

Gel electrophoresis and transfer.

All digested DNAs were fractionated on 1% agarose gels at a voltage gradient of 1V/cm for 16 hours using phosphate buffer, (16). DNA transfer was carried out using a standard procedure, (16), onto either nitrocellulose, (Schleicher & Schuell BA85), or Zeta-Probe, (Bio-Rad).

Sonication.

Human placental DNA (5-10mg/ml, dissolved in 10-20 ml of 10mM Tris-Cl, pH 8.0, 1mM EDTA, (TE)),in a 50ml plastic beaker, in an ice bath, was subjected to 15x 5sec. sonic bursts (15sec. intervals) using a Dawe sonic converter fitted with a 12mm tip. A sample was analysed by alkaline agarose gel electrophoresis; the bulk of the fragments were between 300 & 500 bases long. The DNA was extracted with phenol/chloroform & then chloroform alone, ethanol precipitated and redissolved in TE, to give a concentration of greater than 20mg/ml.

Nick translation.

These were carried out using a scaled down version of a standard procedure (13), altered as follows. 100ng of DNA, in a final volume of 10 µl was incubated in the presence 10µCi of each of 1 or 2 α-^{32}P labelled nucleotides (Amersham), 1 unit of DNA polymerase 1 (Amersham), 0.2nmole each unlabelled nucleotide, 0.1ng DNA'se 1 at 16^{o}C for 30 min. The reaction was stopped by the addition of 40µl TE & 50µl phenol/chloroform. After extraction 10µg E.coli tRNA and Ammonium acetate to 0.3M were added and the probe was precipitated with ethanol. The pellet was briefly dried in vacuo and the labelled DNA was dissolved in 10µl TE. Average incorporations were 30% giving a final DNA specific activity of 10-100µCi/µg. Between 10 & 100ng of each probe was used per hybridisation. Probe mixtures were made after nick translation and before pre-reassociation.

Pre-reassociation.

The basic procedure is:

1. Mix probe DNA
   driver DNA[a]
   25µl 20x SSC[b]
   and TE to 100µl in a 1.5ml microfuge tube.
2. Denature in boiling water 10'.
3. Cool in ice water 1'.
4. Incubate 68°C for the required time.[c]
5. Add into prewarmed hybridisation mix and apply to filter after mixing well.
6. Hybridise etc., using normal procedures.

Note - for the experiments described in this paper:

a) 1mg sonicated human DNA; [driver DNA] = 10mg/ml.

b) 1x SSC is 0.15M NaCl, 0.015M NaCitrate, pH7.0; so the final $[Na^+]$ = 0.825M.

c) The incubation time was 1, 10 or 100 minutes. Cot values, based on the total human driver concentration of 10mg/ml were calculated using the formula (17): Cot = driver concentration (mg/ml) x time in minutes. This includes a correction factor of 6.2 to give the equivalent Cot achieved in 0.12M phosphate buffer (17).

Transfer hybridisation, washing, and autoradiography.

Filters where pre-hybridised in 10ml of pre-hybridisation mix:
5x SET
5x Denhardt's solution
0.1% SDS
0.1% NaPPi(Sodium pyrophosphate)
100 µg/ml denatured sonicated salmon sperm DNA,
at 68 °C for 16-24 hours. The filter was drained shortly before hybridisation. 20x SET is 3M NaCl, 0.4M Tris-Cl, pH 7.8, 20mM EDTA. Hybridisation was carried out in 10ml of a fresh solution of the same composition plus 10% dextran sulphate, at 68 °C for 16 hours. Pre- & hybridisation mixes were degassed and prewarmed before use. Post-hybridisation washes were carried out at 55 °C as follows, 1hr in 4x SSC then 0.5h each in 3x, 1x and 0.1x SSC. All solutions contained 0.1% SDS and 0.1% NaPPi.

In some experiments N-lauroyl sarcosine was used as an alternative to SDS. Filters were air dried and autoradiographed, with intensification in a standard procedure (13). "Short exposure" means 0-8 hours, "normal" 8-72 hours, "long" 3-14 days; for the purposes of comparing results any more precise statement of exposure would be of little value, given the variation in the amount &/or specific activity of particular probes between experiments. Heparin (18), 500µg/ml, was used in one case, (Fig. 4), as a substitute for both Denhardt's solution & salmon sperm DNA.

RESULTS

Hybridisation to gel transfers.

Pre-reassociation conditions were optimised using probes

made from mixtures of human DNA clones representative of highly repeated and low copy DNA sequences. These were pre-reassociated in the presence of a large excess of sonicated total human DNA, to driver Cot values from 10 to 1000, and then used as probes to gel transfers of restriction enzyme digested human, and mouse/human hybrid DNA's.

Elimination of signal from Alu (500,000 copy) repeats

The first experiment (Fig.1) used a δ-globin probe, pHδ, in combination with an Alu clone, BLUR 8. pHδ bridges the δ-globin coding sequence but excludes the 5' repeated sequence associated with the β-globin gene cluster (19). It hybridises to four fragments from the β-globin cluster found in Eco R1 digested human DNA; two from the δ-globin genes, 2.25 and 1.75kb in size, and two from the β-globin genes, 5.2 and 3.6 kb in size. Alu sequences occur about 0.5x - $1x10^6$ times in the human genome (20), and are dispersed throughout the whole size range of human Eco R1 fragments. BLUR 8 hybridises as a smear down the whole track when hybridised to Eco R1 fragments. If the two probes are mixed and used without pre-reassociation (Fig. 1A) the Alu smear masks the single copy signals from the δ-globin probe; short autoradiographic exposure just allows the 2.25 and 5.2kb bands to be seen (Fig 1B). Pre-reassociation of the probe to Cot 100 in the presence of total human DNA (Fig. 1C) dramatically reduces the background, and the four globin fragments are clearly visible. Also revealed is a larger fragment in track 1 only, which may be Y chromosome specific, as well as other low intensity bands, possibly corresponding to pseudogenes. The hybrid cell track, (Fig. 1D3), exhibits a complex band pattern made up from both mouse and human globin fragments. The β-globin cluster is allocated to human chromosome 11, which has been identified in this hybrid.

An α-globin probe, pRBα1, containing an integral Alu family repeat (19), which hybridises to a 22.5kb EcoR1 fragment (22), was also used, together with the single copy X probe, pBS22X, against similar panels. Without pre-reassociation, (Fig. 2A), an intense smear obscures the X specific band (5.9kb); the 22.5kb globin band is just visible in low exposure, (Fig. 2B). Pre-reassociation to Cot 10, 100, or 1000, (Fig. 2C-E), allows both

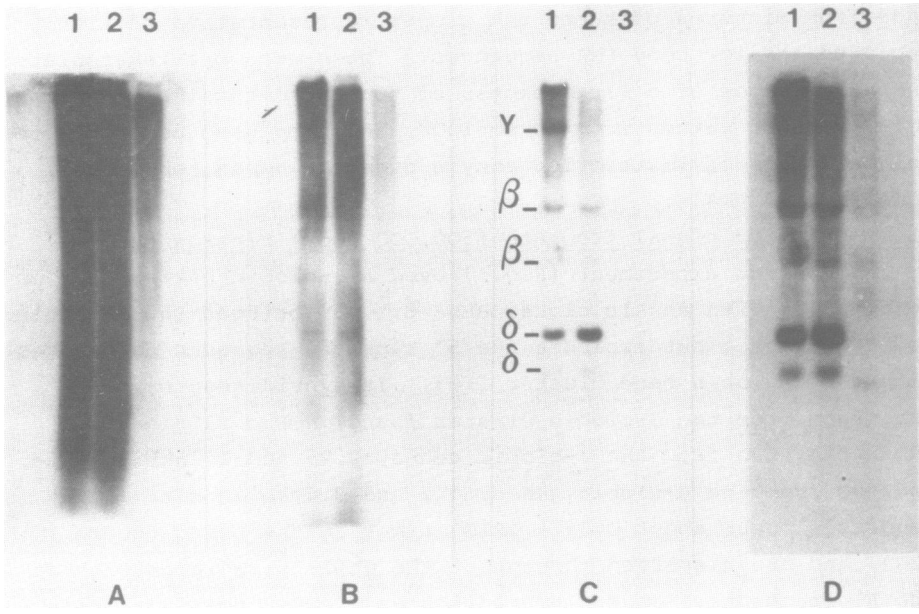

Figure 1. Driven pre-reassociation as a means of reducing Alu repeat hybridisation. DNA panel, track 1, human male, track 2, human female, track 3, ALR58r, all 4μg/track, digested with EcoRl, see Fig. 3E. Approximately 10ng of each of BLUR 8 & pHδ, nick-translated as described in materials & methods, was used as a probe to each of the panels. A & B, probes not pre-reassociated, normal & short autoradiographic exposures, respectively. C & D, probes pre-reassociated to Cot 100, normal autoradiographic exposure in both cases, but the prints have been made with different exposures to show different features. For further explanation, see text.

globin and X bands to be seen clearly against a low background. The background diminishes as the Cot increases; between 10 and 100 the single copy signal is reduced by about 2/3, a larger reduction occurs between 100 & 1000.

Elimination of signal from Kpnl repeats (50,000 copy)

A Kpnl probe, pBK(1.5)54, was mixed with pBS22(X), and with a chromosome 21-specific probe, PS2(21), which contains both a single copy sequence and a dispersed repeat of unknown class. This Kpnl probe has been shown to correspond to about 1% of the human genome (9), and as expected produces a strong and overwhelming background smear when the mixed probes were directly hybridised with panels as used in the previous

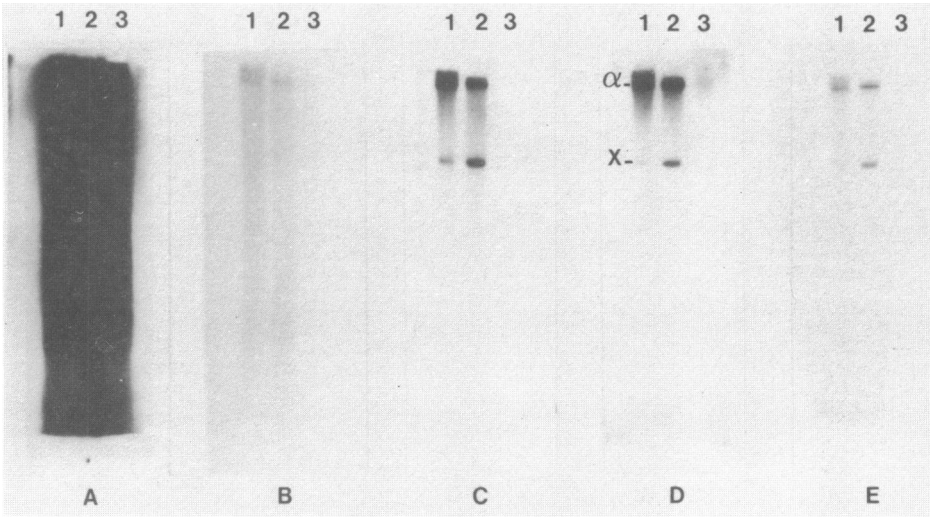

Figure 2. Driven pre-reassociation reduces hybridisation of an integral Alu repeat. The panels were the same as used in Fig. 1. The probe mixture per panel was pRBα1, 10ng, plus pBS22(X) 10ng. A & B, probes not pre-reassociated, normal and low autoradiographic exposure. C,D & E, probes pre-reassociated to Cot 10, 100 & 1000, respectively, normal exposure. For further explanation, see text.

experiment, (Fig.3A). A low autoradiographic exposure shows a 3.3kb band, (Fig.3B), presumably corresponding to the KpnI-containing multi-copy sequence described by Manuelidis & Biro, (21). Pre-reassociation of the probe to Cot 100, (Fig.3C), reveals both X-chromosome (5.9kb) and 21-chromosome (9.6kb) specific bands. The 3.3kb repeat signal has been reduced down to single copy levels. The X probe signal increases with X chromosome dosage. None of these bands is present in the hybrid track, which does not contain the appropriate chromosomes. Pre-reassociation of the probes to Cot 1000, (Fig. 3D), substantially reduces signal strength; at this Cot value most of the single copy sequences will be in duplex form before the filter is exposed to them, (see discussion and appendix).

Use of whole cosmid probes

The usefulness of the pre-reassociation procedure was further investigated in experiments using whole cosmid clones as probes (Figs. 4,5,6). One of the cosmids used, c1F2, has a human

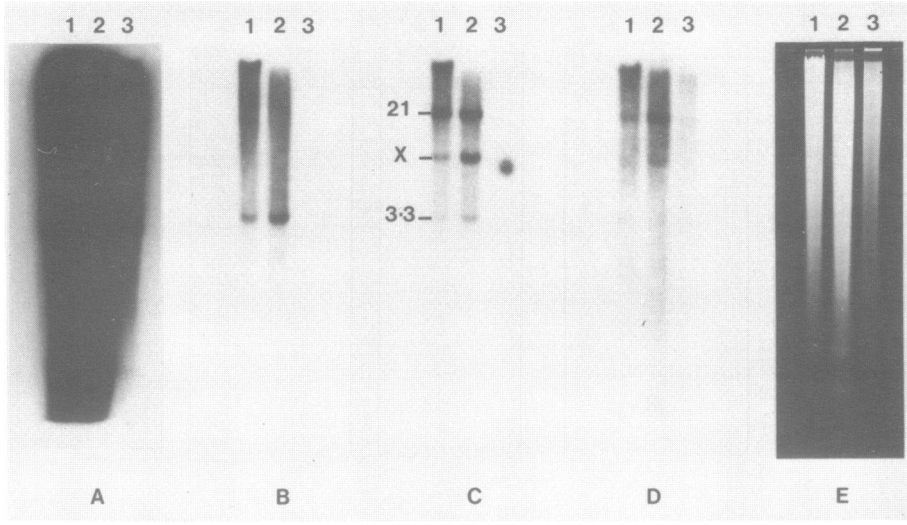

Figure 3. Driven pre-reassociation as a means of reducing Kpn1 repeat hybridisation. The DNA panel is the same as in Fig.1, E is a photograph of one of four panels run per gel, (ethidium bromide stained); all the panels used in each experiment were taken from the same gel. The probe mixture per panel was pBK(1.5)54, 40ng, pBS22(X), 65ng, & PS2(21), purified insert, 40ng, nick translated as before. A & B, probes not pre-reassociated, normal and low autoradiographic exposure. C, pre-reassociated to Cot100, normal exposure. D, pre-reassociated to Cot 1000, normal exposure. For further explanation, see text.

Y chromosome derived insert (Fig. 6); the rest were picked at random from a total human cosmid library. All the cosmids used were constructed by inserting size selected DNA from MboI partial digests into the BamHI site of pJB8, therefore two of the fragments produced by EcoRI digestion of a probe cosmid will hybridise to larger sized EcoRI restriction fragments in genomic DNA. These vector:insert junction fragments would be expected to give a low signal for their size. Possible examples of this are seen with all of the cosmids used.

As expected, when used as probes without pre-reassociation, all of the cosmids gave a strong hybridisation smear down the whole track (eg. see Figs. 4A & 6A); the smear is strongest in the human tracks, and reduced in the hybrid and mouse tracks. All of the cosmid probes gave band patterns when used as a Cot 100 pre-reassociated probes. The size of the

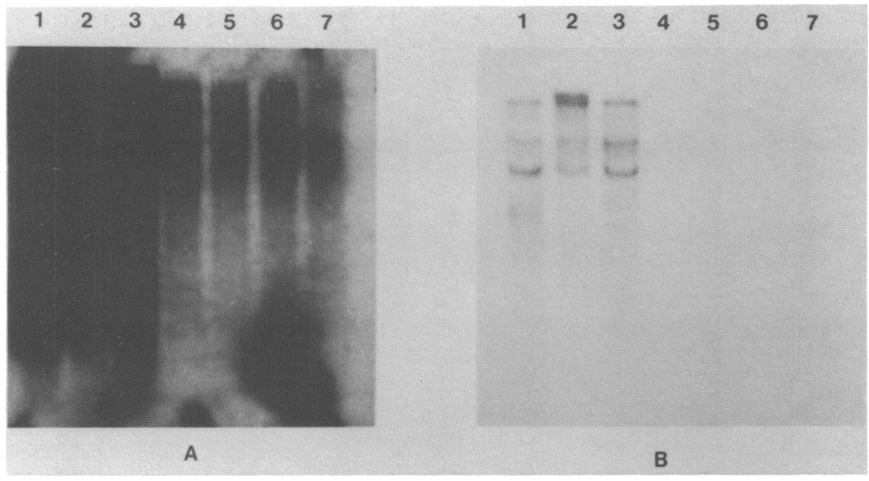

Figure 4. The application of Cot100 pre-reassociation to a randomly chosen cosmid probe. The DNA panel was as described in the legend to Fig. 5. Nick translated cosmid DNA, 50ng per filter, was used as a probe against each of 2 panels, without pre-reassociation (A) and with pre-reassociation to Cot100 (B). Normal exposure. For further details, see text.

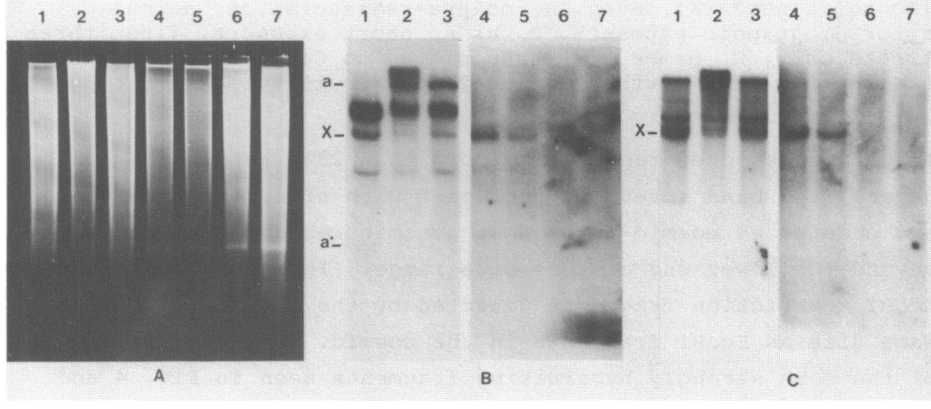

Figure 5. Random cosmid probes. The DNAs in the panel (A) were as follows: track 1, Til 1, 2, human male, 3, human female, 4, ThyB1-33/12, 5, Horl 9X, 6, Thy B1-33-6-1, 7, BW5147, all 4µg/track, all EcoR1 digested. Different, randomly chosen cosmid probes, 100ng, were used as separate probes against panels B & C; in both cases the probe mix contained pBS22(X), 65ng, and was pre-reassociated to Cot 100. Autoradiographic exposure was normal; tracks 4 - 7 have been printed using a longer exposure than for tracks 1 - 3. For further explanation, see text.

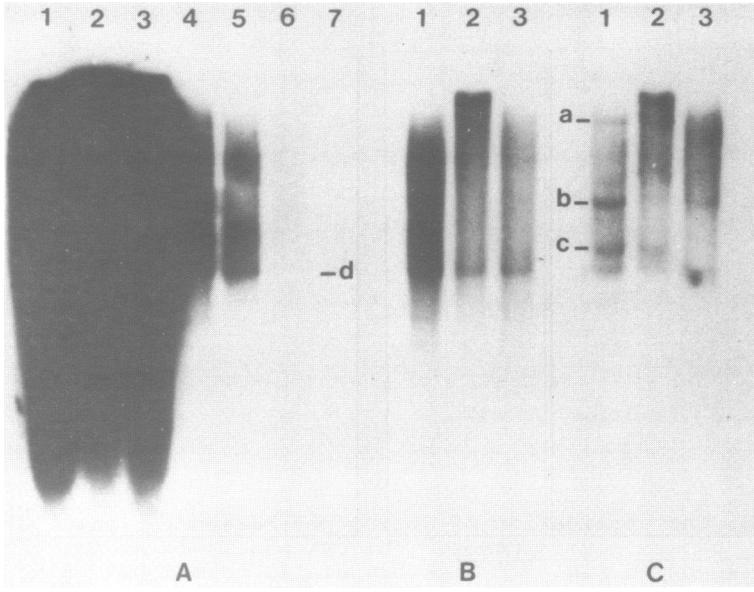

Figure 6. The effect of pre-reassociation on the hybridisation of a Y chromosome derived cosmid probe. DNA panel as described for Fig.5. The probe mixture contained cIF2, 100 ng, & pBS22(X), 65ng. A, no pre-reassociation, normal autoradiographic exposure. B, as A, short exposure, first three tracks only. C, probe pre-reassociated to Cot 100, first three tracks only. For further explanation, see text.

fragments detected ranges from more than 23kb to 2kb; in general the band intensity decreases with size, and thus there are undetected cosmid-homologous genomic restriction fragments beyond the lower end of this size range. In most cases genomic EcoRI restriction fragments detected by the cosmid are of the same size as EcoRI fragments in the cosmid. This applies to all of the more strongly hybridising fragments seen in Fig. 4 and 5B. Non-corresponding hybridising bands may be homologues of EcoRI/MboI insert/vector junction fragments (see above), or non-cognate fragments such as were detected by the δ globin probe. In Fig. 5B a polymorphism is also revealed (band, a in tracks 2 and 3, and a' in track 1). The possibility that this result is produced by incomplete digestion of the DNA in tracks 2 and 3 is unlikely because the same DNA digests were used for all the panels in Figs. 4, 5 and 6 and no other similar result is

obtained with the other cosmid probes.

As an internal control, the single copy X probe pBS22(X) was included in the hybridisations shown in Figs. 5 and 6 and the appropriate dose related signal (5.9kb) is detected in both human and hybrid tracks. In the cases of the probes used in Figs. 5C and 6 the analysis is more difficult because there are substantial differences between the fragments generated by EcoRI digestion of the cosmid and those detected in the transfers. These may be the result of cloning artefacts. In the case of Fig. 5C this could be due to rearrangement during cloning or the inclusion of multiple non-contiguous MboI fragments within the insert and hence novel EcoRI restriction fragments. In the case of Fig. 6 the sum total of the sizes of bands produced by EcoRI digestion of cosmid c1F2 is below 25kb, indicating a substantial deletion of the insert. The large fragment (a), only detected in the Y chromosome DNA containing tracks (1 and 2) is not contained in the cosmid. This could be a deleted fragment, or a junction fragment, but it should be noted that the well characterised δ globin probe also detects fragments much larger than itself, see Fig. 1. Band (b) is produced by control probe pBS22(X). Band (c) corresponds to a fragment contained within c1F2. Band (d) is detected in the absence of pre-reassociation in total human DNA tracks (1, 2 and 3) and X chromosome containing hybrid tracks; a corresponding band is weakly detected by pre-reassociated probes in the human tracks. The sequence may be a conserved repeat found on both X and Y chromosomes. If present in equal numbers on both chromosomes the expected dosage ratios across tracks 1, 2 and 3 would be 5:2:2 and the signal obtained is consistent with this. Clearly the complex patterns detected following pre-reassociation of cosmid probes will require more careful analysis than those detected with smaller probes.

DISCUSSION

A procedure has been devised to supress the hybridisation of repeated sequences and this permits the analysis of low copy DNA components within large cloned DNA fragments from the human genome. The technique has been tested with combinations of well

characterised repeated and low copy sequences which are present on separate cloned DNA fragments (Figs. 1 & 3) or on the same cloned DNA fragment (Figs. 2,4,5 & 6).

The results show the value of the procedure but it is worth comparing them with theoretical expectations (see appendix). These suggest that pre-reassociation to driver Cot 100 should be expected to reduce the background hybridisation signal of highly repeated dispersed sequences to practically zero. For all repeated sequences, the final signal strength should be similar across the copy number range. These theoretical predictions are fairly consistent with the results presented in this paper: Alu & Kpnl repeat signals were practically abolished (Figs.1 & 3), although there is still a low residual signal, & in two cases signal from repeat bands was substantially reduced (Figs.3 & 6). It should be noted that repeated sequences such as ribosomal or satellite DNA which produce specific band patterns in genomic blots may give a hybridisation signal resembling low copy DNA after pre-reassociation of the probe; unlike low copy sequences, such repeated sequences may give strong band patterns that are visible when probe that has not been pre-reassociated is used (Figs. 3 and 6).

Pre-reassociation to Cot 1000 causes a large reduction in single copy signal, and has little to recommend it as a standard procedure. However, using the method described, there is probably not much to choose between pre-reassociation to Cot 10 or Cot 100. Probe pre-reassociated to Cot 10 will continue to reassociate in the hybridisation mix, (diluted 100x, & in the presence of 10% dextran sulphate), and can be expected to reach Cot 100 in about an hour.

For this reason it would be difficult to do the control experiment which does not allow any pre-reassociation to take place. It is important to mix the components well in the prewarmed hybridisation mix before they are applied to the filter. Even a few seconds under such conditions would allow substantial pre-reassociation of high repeats to take place. Thus it is not possible to be completely sure that the success of the procedure is solely due to pre-reassociation as opposed to simple competition. Experiments using single stranded probe

and competitor could answer this question. It would be difficult to find a single stranded alternative to total human DNA. This difficulty of controlling the amount of pre-reassociation may explain the variable results that have been reported by those seeking to use simple competition as a means of suppressing repeat hybridisation.

The use of sonicated driver DNA has the consequence that partially paired duplexes will be formed between molecules that are not completely overlapping and this may be expected to cause some networking with a reduction in the effectiveness of the procedure. This might be reduced by using driver that has been produced by digestion with an appropriate restriction enzyme. We have successfully used restricted vector DNA to suppress vector:vector reassociation in colony hybridisations. However, in the case of the type of transfer hybridisations reported in this paper, the convenience of using a single, cheaply produced preparation of driver DNA with a variety of probes constructed in different ways is clear. The properties of human dispersed repeats (Alu, Kpnl) are such that even with appropriately restricted driver non-cognate hybrids would still be formed.

The procedure described is simple and rapid, and allows the detection of low copy signal, in transfer hybridisations, which would otherwise have been obscured. Using the technique we have observed signal from DNA sequences which varies with chromosome dosage, and, in one case, fragments displaying a restriction fragment length polymorphism. Such observations indicate that the procedure may be of use to those searching for markers linked to loci associated with genetic disease. It should be possible to construct experimental protocols for a range of applications using the methods and results described in this paper.

APPENDIX

Theoretical assessment of pre-reassociation strategy

To determine the best pre-reassociation strategy for reducing unwanted signal from a repeat or from a vector component, whilst preserving the signal from low copy components we start with the assumption that signal from the low copy

component should be kept to a maximum. Let us say that loss of half of the normal signal is the maximum acceptable. We define $t\frac{1}{2}$ as the time of half reassociation of the clone sequence in the pre-reassociation reaction and in the absence of driver. For any given set of conditions $t\frac{1}{2}$ may be calculated from the known reassociation rate of the E.coli genome. The kinetic complexity of this genome equals its genome size, as all sequences are present at the same molar concentration, and this is also the case with most cloned probes. Thus:

$$\frac{\text{Cot}\,\frac{1}{2}\ (\text{probe})}{\text{Cot}\,\frac{1}{2}\ (\text{E.coli})} = \frac{\text{L(probe)}}{4.2 \times 10^{6}}$$

Where Co is the initial DNA concentration and L(probe) is the probe length in base pairs. The $\text{Cot}\frac{1}{2}$ for E.coli is 4 mole sec/litre, in 0.12M phosphate buffer (corresponding to $[Na^+]$ = 0.18M) at 60°C. Thus:

$$\text{Cot}\,\tfrac{1}{2}\ (\text{probe}) = L \times 10^{-6}$$

For example a 50kb probe at a concentration of 100ng/100 µl has $\text{Cot}\frac{1}{2}$ 0.05 mole sec/litre, and hence the maximum permissible time of pre-reassociation $t\frac{1}{2}$ = 5 hours.

To drive specific components further into the double strand state we have to increase their molar concentration. For example, to drive vector into duplex requires the addition of more vector DNA. The appropriate amount can be calculated from the equation:

$$Cd = Co\ (n-1)$$

Where Co is the probe molar concentration, Cd is the molar concentration of the driven component and 1/n is the fraction remaining single stranded. An approximation is to say that any reduction in a particular probe component requires a proportionate increase in its molar concentration, so that:

$$\frac{Co}{Cd} = \frac{1}{n}$$

For example, adding a 100 fold excess of vector, reduces vector signal to about 1% of the uncompeted level.

If the driver DNA is complex - e.g. whole eukaryote genome DNA, the molar concentration of any repeated component is proportional to its copy number (r). The molar contribution (Cd)

of a driver component can be calculated from:

$$Cd = \frac{\text{copy number x driver concentration}}{\text{DNA content of haploid genome}}$$

For each molar concentration the signal reduction will be the same as it would be in the single component driver situation. This relationship remains valid as long as the amount of the complex DNA used as driver is not large enough to significantly affect the molar concentration of the low copy probe component.

The amount of the remaining single stranded probe that will hybridise to gel transfers of genomic DNA, (H), may be assumed to be proportional to that sequence's copy number in the probed genome. If this is the same as the driver genome, then:

$$H \propto r$$

The final signal strength (S), relative to probe that has not been pre-reassociated is the product of these factors:

$$S \propto r \frac{Co}{Cd}$$

in other words the final relative signal is proportional to the relative molar concentrations of the components (in the probe) times the copy number (in the target DNA). This equation may be used to calculate the amount of driver DNA required to produce a desired signal reduction.

To optimise the signal, and minimise the effects of unwanted components the following parameters can be adjusted for pre-reassociation. $t_{\frac{1}{2}}$ is proportional to reaction volume and probe complexity and inversely proportional to concentration. Changing the temperature and/or salt concentration will alter the rate of reassociation and possibly allow more convenient timing. Different stringencies of reassociation may alter the copy number of different genome components.

For the experiments using total human DNA described in this paper a value of 48 minutes is obtained for the $t_{\frac{1}{2}}$ of the undriven probe (assuming 50kb probe as before, but with $[Na^+] = 0.825M$); this is equivalent to a driver Cot value of 480. All repeat classes of human DNA are reduced to the same relative signal strength of 6.9, using a value of 3.4pg for the DNA content of the haploid human genome.

ACKNOWLEDGEMENTS

Preliminary experiments were carried out by Ruggiero Caizzi, Bao Yong-de and Juliet Honeycombe. Thanks are due to Chris Bostock for several stimulating discussions; to Sandy Bruce for photographic assistance and to Anne Deane for typing the manuscript. Paul Whittaker is supported by the Muscular Dystrophy Group of Great Britain.

REFERENCES

1. Britten,R.J., & Kohne,D.E. (1968) Science 161, 529-540.
2. Schmid,C.W. & Deininger,P.L. (1975) Cell 6, 345-358.
3. Maniatis,T., Hardison,R.C., Lacy,E., Lauer,J., O'Connell,C., Quon,D., Sim,G.K., & Efstratiadis,A. (1978) Cell 15, 687-701.
4. Weiss,E.H., Golden,L., Fahrner,K., Mellor,A.L., Devlin,J.J., Bullman,H., Tiddens,H., Bud,H., & Flavell,R.A. (1984) Nature 310, 650-655.
5. Fisher,J.H., Gusella,J.F., & Scoggin,C.H. (1984) Proc. Natl. Acad. Sci. USA 81, 520-524.
6. Brison,O., Ardeshir,F., & Stark,G.R. (1982) Mol. Cell Biol. 2, 578-587
7. Hamada,H., Petrino,M.G., & Kakunaga,T. (1982) Proc. Natl. Acad. Sci. USA 79, 6465-6469.
8. Jelinek,W.R., Toomey,T.P., Leinwand,L., Duncan,C.H., Biro,P.A., Choudary,P.V., Weissman,S.M., Rubin,C.M., Houck,C.M., Deininger, P.L., & Schmid,C.W. (1980) Proc. Natl. Acad. Sci. USA 77, 1398-1402.
9. Shafit-Zagardo,B., Maio,J.J., & Brown,F.L. (1982) Nucleic Acids Res. 10, 3175-3193.
10. Fritsch,E.F., Lawn,R.M., & Maniatis,T. (1980) Cell 19, 959-972.
11. Lauer,J., Shen,C-K.S., & Maniatis,T. (1980) Cell 20, 119-130.
12. Ish-Horowicz,D., & Burke,J.F. (1981) Nucleic Acid Res. 9, 2989-2998.
13. Maniatis,T., Fritsch,E.F., & Sambrook,J. (1982) "Molecular Cloning: A Laboratory Manual" Cold Spring Harbour Laboratory.
14. Birnboim,H.C., & Doly,J. (1979) Nucleic Acids Res. 7, 1513-1523.
15. Ross,J. (1976) J. Mol. Biol. 106, 403-420.
16. Sealey,P.G., & Southern,E.M. (1982) in "Gel Electrophoresis of Nucleic Acids: A Practical Approach" Rickwood.D., & Hames.B.D., Eds, pp. 39-76, IRL Press, Oxford & Washington.
17. Britten,R.J., & Kohne,D.E. (1967) Carnegie Inst. Wash. Yearbook 66, 73
18. Singh,L & Jones,K.W. (1984) Nucleic Acids Res. 12, 5627-5638.
19. Fritsch,E.P., Shen,C.K.J., Lawn,R.M., & Maniatis,T. (1980) Cold Spring Harbour Symp. Quant. Biol. 65, 761-775.
20. Rinehart,F.P., Ritch,T.G., Deininger,P.L., & Schmid,P.W. (1981) Biochemistry 20, 3003-3010
21. Manuelidis,L., & Biro,P. (1982) Nucleic Acids Res. 10, 3221-3229.
22. Orkin,S.H. (1978) Proc. Nat. Acad. Sci. USA 75, 5950-5954.