Chapter 2

GENETIC AND PHYSICAL SELECTION OF EUKARYOTIC GENES CLONED IN *E. COLI*

RONALD W. DAVIS, MARJORIE THOMAS, JOHN R. CAMERON, PETER PHILIPPSEN, RICHARD KRAMER, THOMAS ST. JOHN, KEVIN STRUHL and JILL FERGUSON

Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305, U.S.A.

Introduction

Cloning eukaryotic DNA in E. coli with λgt

Techniques have been recently developed for the construction and cloning of viable molecular hybrids between modified bacteriophage λ DNAs and any foreign DNA [1,2,3,4]. As shown in fig. 1, one of these modifications of λ DNA involves the genetic removal of two *Eco*RI restriction sites (sites 4 and 5), a genetic deletion of non-essential DNA in the right arm, and the biochemical deletion of the λ *Eco*RI C fragment [2,3]. Some of the unique advantages of this modified λ, called λgt, as a vector for foreign DNA are: 1) a segment of DNA must be inserted between the right and left end fragments in order to have sufficient DNA to make a viable virus particle. 2) The DNA can be modified for K restriction so that it can infect any K12 strain of *E. coli* which is sensitive to infection by λ. 3) The phage can grow lytically with active anti-terminated λ promotors which will probably transcribe the foreign DNA. 4) The DNA can be integrated into the bacterial chromosome with the λ promotors repressed. 5) The strands of λ DNA can be readily separated. 6) Deletions that occur in the inserted DNA can be readily selected.

Large pools of hybrids containing many different fragments of DNA from any given genome can be generated. Fig. 2 shows the construction of hybrids with *Saccharomyces cerevisiae* (yeast) DNA. The vector used is λgt-λB, which carries the λ *Eco*RI B fragment for the initial propagation of the vector [3]. This fragment can be removed by preparative agarose gel electrophoresis prior to hybrid construction; however, its removal is not essential, as its reinsertion

The Molecular Biology of the Mammalian Genetic Apparatus: edited by P. Ts'o

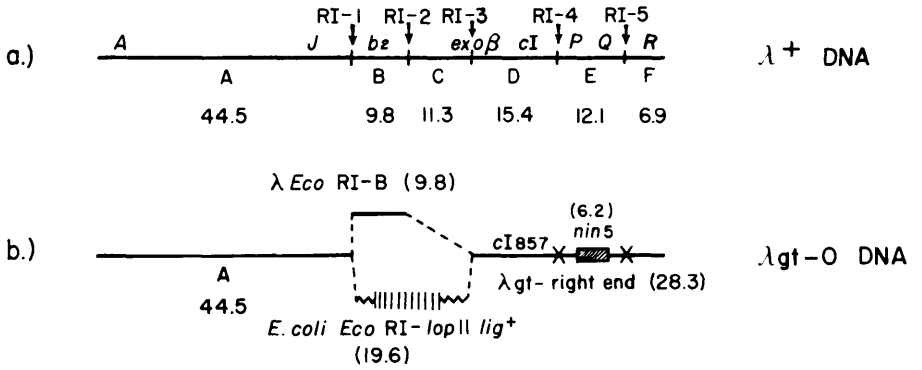

Fig. 1. a) Structure of bacteriophage λ DNA. The five *Eco*RI sites are labeled RI-1 through RI-5. A number of the λ gene positions are labeled A through R. The λ exonuclease gene is at *Eco*RI site 3. The *Eco*RI fragments are lettered sequentially A through F. The size of these fragments is given in fractional units of whole λ DNA which contains about 46,500 base pairs for a mass of 30.8×10^6 daltons. b) Structure of λgt-λB DNA. The two extreme right *Eco*RI sites have been removed by mutation and genetic selection. A non-essential region, *nin5*, has been genetically deleted. The *Eco*RI-C fragment between RI-2 and RI-3 has been biochemically deleted by *Eco*RI endonuclease cleavage and rejoining. This fragment contains the attachment site and genes necessary for the establishment of stable lysogens. The CI857 mutation renders the λ repressor temperature sensitive. Also shown is the hybrid which contains the *E. coli* DNA sequence coding for DNA ligase. This gene also carries the *lop*11 mutation which causes an overproduction of DNA ligase.

after cleavage can be competed with *Eco*RI cleaved yeast DNA [2]. The biochemically constructed hybrid DNA is used to infect calcium treated *E. coli* cells [5] that lack a restriction modification system. The plaques that are produced are combined to produce a hybrid pool. The hybrids formed with yeast DNA are designated λgt-Sc, with either an isolation number or a descriptive term following the Sc. We have prepared three independent λgt/Sc pools, each consisting of about 10,000 pooled plaques from the DNA infection.

EcoRI and HindIII restriction spectra of Saccharomyces cerevisae DNA

The variety of DNA sequences that are found in yeast DNA and in the cloned λgt-Sc DNA pools are readily visualized by separating the restriction endonuclease generated DNA fragments by agarose gel electrophoresis [6,7,8,9]. The display of DNA fragments, visualized by fluorescence from bound ethidium bromide, is termed a restriction spectrum (fig. 3). The *Hind*III and *Eco*RI restriction spectra of *S. cerevisiae* DNA are shown in fig. 3a and c. Fig. 3b shows the *Hind*III, *Eco*RI double digestion restriction spectrum. The arrows indicate faint bands that probably represent unique single copy sequences. The darker bands result from the coincidence of a number of different single copy sequences or from one sequence found in multiple copies. The *Eco*RI restriction spectrum

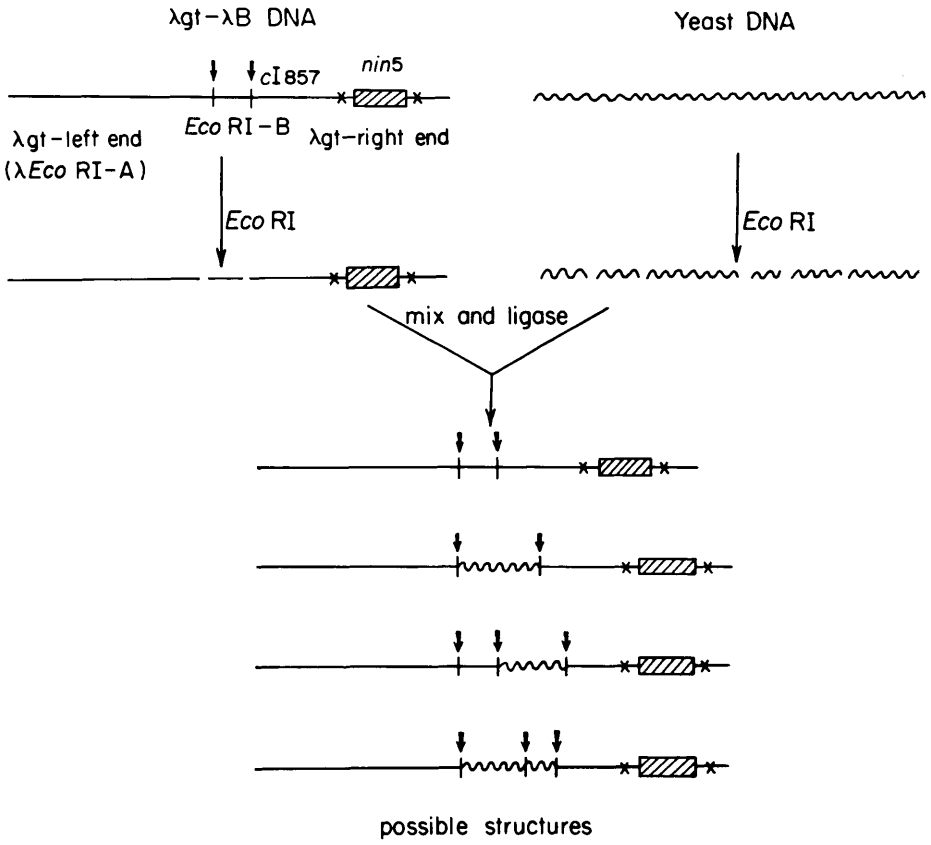

Fig. 2. Construction of hybrids. λgt-λB and yeast (*Saccharomyces cerevisiae*) DNA's were separately cleaved to completion with *Eco*RI endonuclease (partial cleavage products of yeast DNA were used in some cases). They were then mixed at equal concentrations of vector and yeast DNA for covalent joining of the *Eco*RI cohesive ends by *E. coli* DNA ligase. This was carried out at 10°C for 18 h. Four possible resulting structures of viable phage DNA are shown.

of the λgt-Sc DNA pool is very similar to that of whole yeast DNA, with the exception of the λ *Eco*RI B fragment which is found in the pool. This serves as evidence that most of the *Eco*RI DNA fragments from the yeast genome can be cloned in *E. coli*.

Genetic selection

Lytic selection

Direct genetic selections can be applied to the K modified λgt-Sc pools. One of the simplest questions to ask is whether the inserted yeast DNA fragments

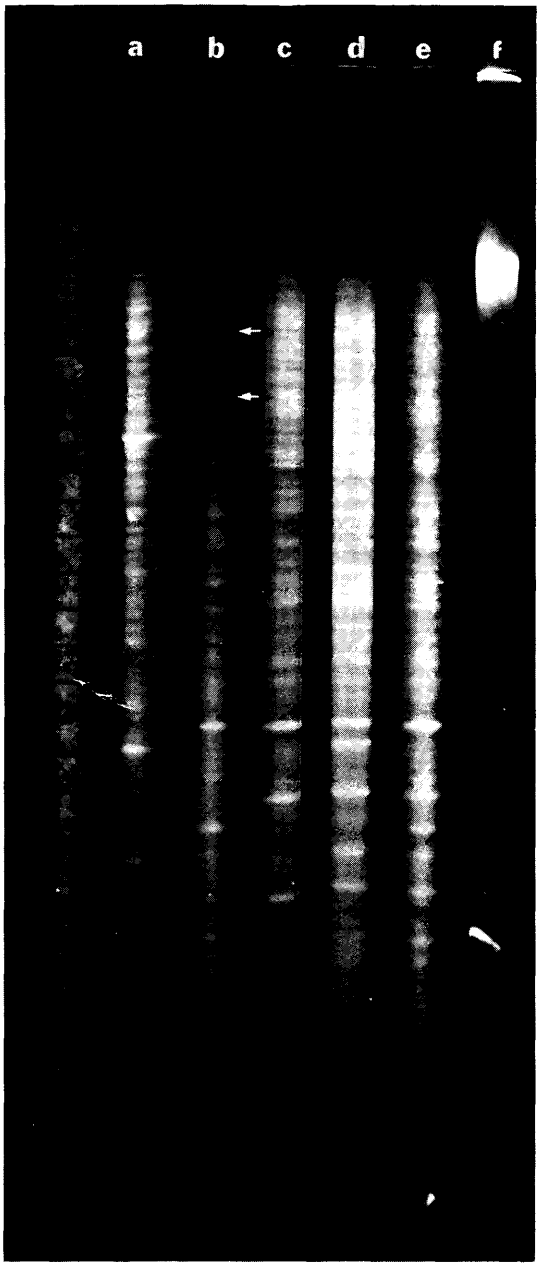

Fig. 3. Restriction spectra of yeast DNA. Yeast DNA was digested to completion with *Hin*dIII and/or *Eco*RI endonuclease in 50 mM Tris–HCl, pH 7.5, 10 mM $MgCl_2$, 50 mM NaCl at 37° C. Aliquots were loaded on 20 cm-long cylindrical 0.7% agarose gels. Electrophoresis was performed at 1.5 v/cm for 30 h. a) 0.8 μg DNA cleaved with *Hin*dIII. b) 0.8 μg DNA cleaved with *Hin*dIII and *Eco*RI. c) 0.8 μg DNA cleaved with *Eco*RI. d) Mixture of samples a and c. 2) Mixture of samples b and c. f) 0.4 μg DNA incubated without enzyme.

affect λ growth or whether they are completely neutral. The simplest way to answer this question is to continuously passage the phage pool and follow the restriction spectra as a function of the number of rounds of infection. It is quite clear that after only a few rounds of infection, a few DNA bands of the restriction spectra become enriched, while most DNA bands become diminished [10]. After 30 sequential infections, about 95% of the λgt-Sc hybrids in the pool contain one yeast DNA fragment of about 8 kb. Heteroduplex analysis shows that this fragment is inserted in λgt in either direction with approximately equal frequency (λgt-Sc1000 and λgt-Sc1000^{I}). Since there is no polar effect of this insertion, transcription from the λ promotors is probably not germane in its selection. It is not clear if this yeast fragment aids λ growth or if it is the least detrimental of the yeast fragments to λ growth. In any event, it grows better than λgt-λB, but not as well as wild type λ^{+}, the burst of λgt-Sc1000 being about one-half that of λ^{+}. Therefore, it was of interest to determine if a yeast DNA fragment could be found that would aid λ growth in a more restrictive host. This is easily accomplished by plating on a DNA polymerase I deficient cell (*pol*A *amber* mutant), since the λ exonuclease gene has been deleted in all the λgt-Sc phage, and λ*exo*$^{-}$ phage will not plaque on a DNA polymerase I deficient cell. Therefore, if the yeast hybrid pool contains phage which can complement the mutation in the DNA polymerase gene or the λ exonuclease gene, then these phage should give plaques on a *pol*A lawn. The *Eco*RI restriction spectrum of the pool, examined after three successive infections of a DNA polymerase I deficient cell, shows about 30 different sizes of fragments [10]. A number of these have been cloned and shown to have different DNA sequences by heteroduplex analysis. The yeast fragment that confers the best growth advantage to a λ *exo*$^{-}$ phage on a *pol*A *amber* cell was selected by 20 successive infections. This phage, called λgt-Sc1001, gives a burst equal to that of a λ *exo*$^{+}$ phage on this *pol*A cell. It contains two fragments of 2.2 and 1.6 kb. Unexpectedly, it gives a lower phage burst on infection of a wild type cell than on a DNA polymerase I deficient cell, indicating that there is an adverse interaction between this phage or phage product and wild type DNA polymerase I levels in a cell. If the inserted DNA fragment codes for a diffusible product, this would be the first evidence of functional genetic expression of eukaryotic DNA in a prokaryotic cell. A *trans* complementation test was conducted to determine whether the presence of the foreign DNA fragment can help a genetically distinguishable λ *exo*$^{-}$ phage during coinfection. λgt-Sc1001 clearly acts in *trans*, while a number of other hybrids do not show a marked *trans* effect [10]. It would appear that λgt-Sc1001 complements the λ exonuclease defect rather than the polymerase defect since this phage can also grow on a DNA ligase ts cell, which also does not plate *exo*$^{-}$ phage.

The selections thus far discussed have focused on phage functions. A method for the direct selection of a phage which supplies a function essential to cell growth is also needed. However, even if the infecting phage carries a genetic

element which allows the cell to grow, what method will determine this if the cell dies from the infection? A plaque cannot be easily seen since the uninfected cells cannot grow, and therefore, cell growth around the plaque cannot be observed. The infected cell should support phage growth if it carries into the cell a function essential for cell growth, and the infected cells will lyse and release the DNA of the cell. This released DNA is easily localized by spraying the plate with ethidium bromide and then viewing with ultraviolet light. The fluorescent spots indicate the plaques from phage which perform an essential function for the cell. An example is shown in fig. 4 of a *λtrp* (which carried the *E. coli* tryptophan operon) infecting an *E. coli trp* cell lawn on a minimal plate. One of the advantages of this type of selection is that the inserted DNA is likely to be transcribed from one of the λ antiterminated promotors, thus allowing the genetic selection of DNA sequences which do not contain their own promotors.

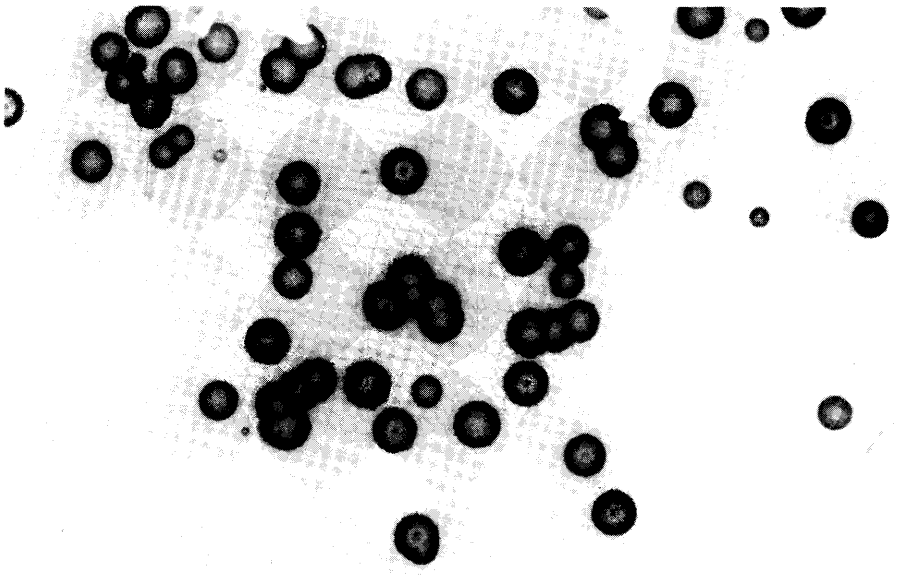

Fig. 4. Visualization of plaques without visible lawns by staining with ethidium bromide. Cells of the tryptophan auxotroph, *trp* A33, were grown overnight in maltose minimal medium supplemented with tryptophan (*trp*). The culture was washed twice with 10 mM $MgSO_4$ and starved for *trp* for 30 min. Approximately 10^8 cells were infected with λgt-λB (10^7) and *λptrp* ABCD (500) and plated on glucose minimal plates at 37 C. After 2 days, plates were sprayed with 0.1 mg/ml ethidium bromide and visualized by fluorescence with ultraviolet light.

Lysogenic selection

Since λ is capable of integrating into the bacterial chromosome to form a stable lysogen, direct genetic selections can also be performed without killing the host. The establishment of lysogeny with the integration defective λgt-Sc phage is achieved by coinfection with an integration helper phage. We have isolated λgt-Sc phage which, when integrated into the chromosome of a auxotroph, allows this bacterium to grow in the absence of histidine [11]. The cell used is a non-reverting *his* B mutant of *E. coli* which lacks the enzyme imidazole glycerol phosphate (IGP) dehydratase. The *his* B complementation is dependent upon the presence of a λgt hybrid since all *his*$^+$ colonies (200) contain a λ prophage, and removal of the prophage from the chromosome by curing results in the original *his* B$^-$ phenotype. The phage (λgt-Sc2601) which is responsible for the *his* B complementation has been cloned and contains a single inserted yeast *Eco*RI DNA fragment of 10.3 kb. This same *Eco*RI DNA fragment has been isolated from another independently generated λgt-Sc phage pool using an identical selection procedure. If λgt-Sc2601 is used in the selection with an integration helper, the frequency of obtaining *his*$^+$ colonies is increased by 10^4 compared to using the hybrid pool with an integration helper. The inserted DNA fragment in λgt-Sc2601 is not a contaminant because labeled RNA made from it hybridizes to the same size *Eco*RI fragment of a different yeast DNA preparation. Therefore, we conclude that there is functional genetic expression of eukaryotic DNA in *E. coli* and that transcription is most probably initiated within the yeast DNA sequence since all known λ promotors are repressed or have been deleted.

Physical selection

Physical screening with complementary RNA

Clearly, not all genes can be obtained by genetic selection. For example, we are presently attempting to isolate the same yeast DNA fragment in λgt-Sc2601 from yeast mutants with no IGP dehydratase activity. A simple screening method has been developed in which ^{32}P RNA is hybridized to denatured DNA in plaques [12,13]. This involves forming plaques on bacterial lawns grown on nitrocellulose filters, and alkaline-denaturing the phage and the DNA to locally fix the single stranded DNA on the filter [13]. After washing the filter and drying, labeled RNA is added and hybridized. The filters are then washed. After RNase treatment, plaques containing DNA sequences complementary to the RNA are localized by autoradiography. An example is shown (fig. 5) using total and 5.8S yeast ribosomal RNA hybridized to randomly selected plaques from the hybrid

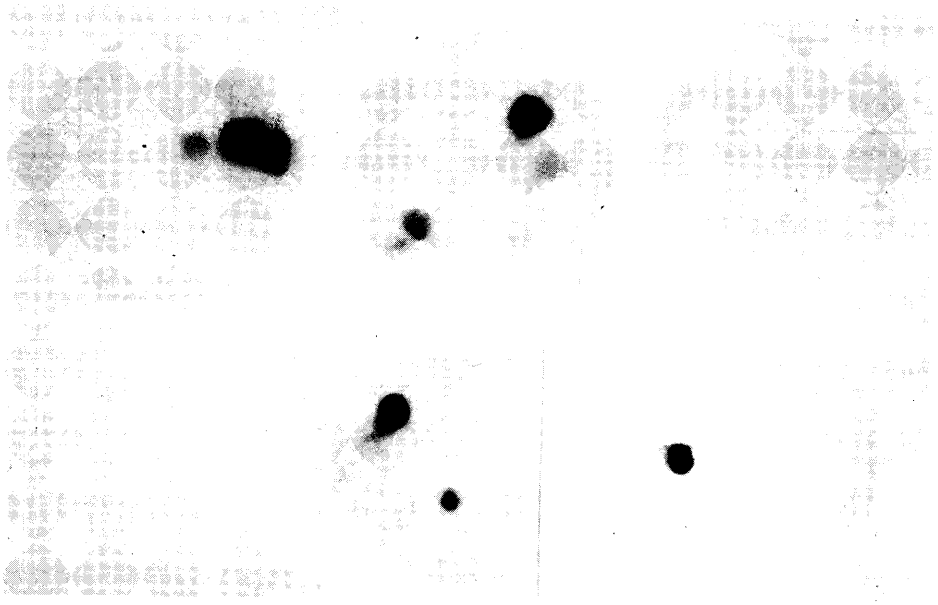

Fig. 5. Autoradiograph of in situ filter hybridization. ^{32}P-labeled RNA was extracted from yeast ribosomes and approximately 5×10^5 cpm applied to each nitrocellulose filter with denatured DNA from phage plaques fixed in situ. After hybridization and washing, the filters were placed against X-ray film to determine the location of phage which contained sequences homologous to rRNA (dark spots). The outline of the filters is visible due to a low level of non-specific binding of the RNA.

pool. Recently, we have found that tRNA can also be used for this screening procedure.

Physical selection with complementary RNA

Using the physical screening method, one can readily find a sequence in the hybrid pool at a frequency of about 10^{-3}, which is the approximate frequency of finding a single copy sequence in yeast. If a hybrid pool is constructed with DNA from higher organisms, the frequency of finding a single copy sequence drops to 10^{-6}. Because of this low frequency, a direct physical selection is desired. We are developing a selection procedure which is based on the observation [14] that RNA can be hybridized to double stranded DNA. A micrograph of such a molecule is shown in fig. 6 [15]. The DNA is from a λgt-Sc hybrid with an inserted yeast DNA segment complementary to yeast 26S rRNA [13]. This hybrid was selected by the physical screening procedure given in the previous section. The RNA displaces the identical DNA strand and forms a loop which has been designated an R-loop. It is formed because at high

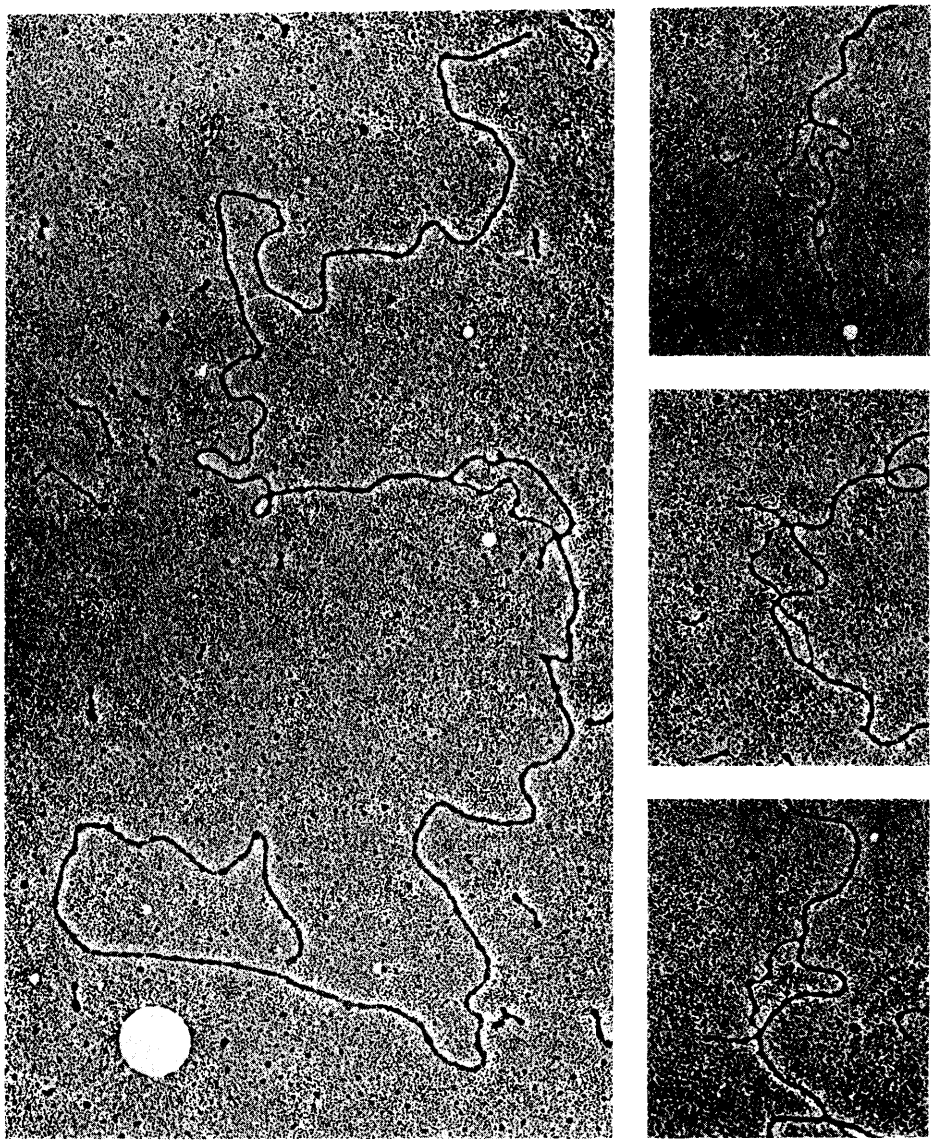

Fig. 6. The R-loops above were made by incubating 5 μg/ml λgt-Sc2056 DNA with 5 μg/ml total yeast rRNA in a solution containing 70% v/v formamide, 0.1 M PIPES, pH 7.8, and 0.01 M Na_3EDTA at 47°C for 20h. All of the 500 molecules examined by electron micrscopy contained an R-loop. Samples were mounted for electron microscopy by the basic protein film technique [2]. Grids were stained with uranyl acetate and shadowed with Pt/Pd.

formamide concentration, the RNA/DNA hybrid is more stable than the duplex DNA. The rate of R-loop formation is very dependent upon the incubation temperature (fig. 7), the rate being maximal at the denaturation temperature of the

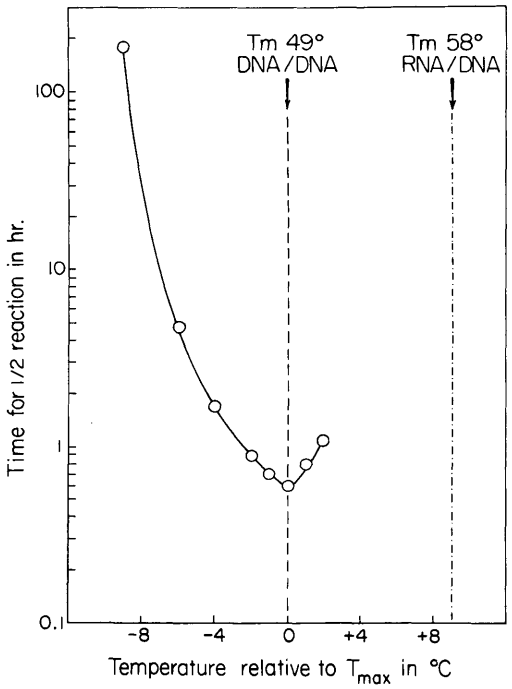

Fig. 7. Rate of R-loop formation as a function of temperature. R-loops were formed using the conditions given in fig. 6. The rate was determined by scoring the fraction of reacted DNA molecules in the electron microscope as a function of time. The melting temperatures (Tm) of the duplex DNA and of the duplex RNA/DNA hybrid were assessed by determining the temperature at which the duplex structures are converted to collapsed single strand structures in the electron microscope.

duplex DNA [15]. R-loops have considerable kinetic stability since the formamide can be removed by dialysis, and R-loop containing DNA can be cleaved with a restriction endonuclease without appreciable loss of R-loops (fig. 8). The ability to cleave R-loop containing DNA with restriction endonucleases should prove very useful in the detailed mapping of the location of the hybridized RNA.

R-loops may serve as the basis of a number of methods for physically selecting DNA sequences. One is based on the fact that the displaced DNA strand in the R-loop causes a DNA molecule containing an R-loop to be retained on a benzoylated DEAE cellulose (BD cellulose) column in the presence of high salt. This method has been used to enrich for the rDNA genes from *Dictyostelium discoideum* (fig. 9) and obtain λ clones containing these sequences [16]. Work is in progress to improve this technique so that it can be used to isolate single copy sequences from any organism.

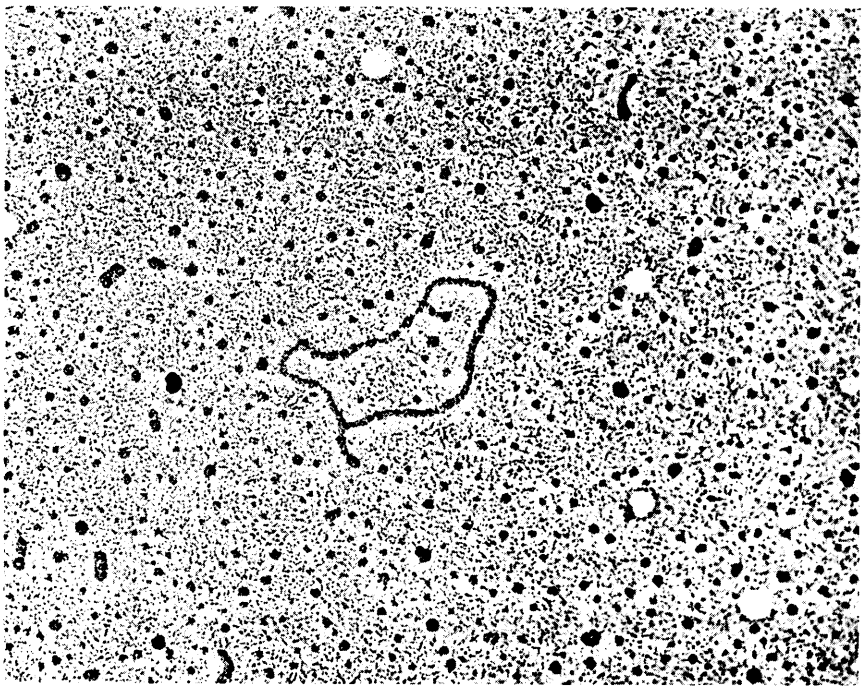

Fig. 8. DNA containing R-loops was dialyzed into 0.1 M NaCl, 0.05 M Tris, pH 7.5, 10^{-4} M EDTA, and subsequently cleaved with *Eco*Ri endonuclease. Above is one example of such a molecule prepared from a DNA sample containing 100% R-loops. The digestion was done in the presence of 10 mM $MgSO_4$ at 37 C for 10 min. The sample was mounted for electron microscopy as described in fig. 6. The R-loop shown contains a fragment of rRNA.

References

[1] N.W. Murray and K. Murray, Manipulation of restriction targets in phage λ to form receptor chromosomes for DNA fragments. Nature 251, 476 (1974).

[2] M. Thomas, J.R. Cameron and R.W. Davis, Viable molecular hybrids of bacteriophage lambda and eukaryotic DNA. Proc. Nat. Acad. Sci. USA 71, 4579 (1974).

[3] J.R. Cameron, S.M. Panasenko, I.R. Lehman and R.W. Davis, In vitro construction of bacteriophage λ carrying segments of the *Escherichia coli* chromosome: selection of hybrids containing the gene for DNA ligase. Proc. Nat. Acad. Sci. USA *72*, 3416 (1975).

[4] K. Murray and N.W. Murray, Phage lambda receptor chromosomes for DNA fragments made with restriction endonuclease III of *Haemophilus influenzae* and restriction endonuclease I of *Escherichia coli*. J. Mol. Biol. 98, 551 (1975).

[5] M. Mandel and A. Higa, Calcium-dependent bacteriophage DNA infection. J. Mol. Biol. 53, 159 (1970).

[6] C. Aaij and P. Borst, The gel electrophoresis of DNA. Biochim. Biophys. Acta 269, 192 (1972).

[7] P.A. Sharp, B. Sugden and J. Sambrook, Detection of two restriction endonuclease activities in *Haemophilus parainfluenzae* using analytical agarose-ethidium bromide electrophoresis. Biochemistry 12, 3055 (1973).

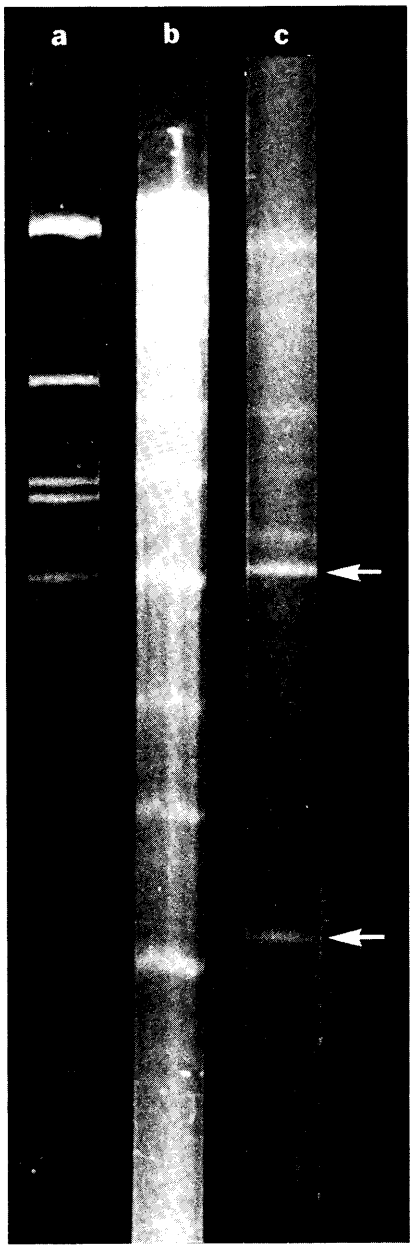

Fig. 9. Selection of rDNA cistrons from *Dictyostelium discoideum* by BD cellulose chromatograph of R-loops. R-loops were formed with 10 μg of *Eco*RI endonuclease cleaved *Dictyostelium discoideum* DNA using 5 μg 28S and 18S rRNA in 70% formamide, 0.5 M NaCl, 0.1 M PIPES, pH 7.8, 0.01 M Na_3EDTA and incubating at 52 C for 2 h. DNA molecules containing R-loops were separated from totally duplex DNA by selectively adsorbing the R-loop containing DNA to a BD cellulose [16] column in 1 M NaCl, 0.1 M Tris, pH 7.5, 0.01 M Na_3EDTA. The R-loop containing DNA was eluted by washing the column with 50% formamide, 1 M NaCl, 0.1 M Tris, pH 7.5, 0.01 M Na_3EDTA. The RNA was removed from the R-loop by RNase treatment. The resulting duplex DNA was electrophoresed on a 0.7% agarose gel as given in fig. 3. The arrows mark the bands that hybridize to 28S and 18S rRNA. a) *Eco*RI cleaved λ DNA. b) *Eco*RI cleaved *Dictyostelium discoideum* DNA. c) *Eco*RI cleaved *Dictyostelium discoideum* DNA with R-loops after adsorption and elution from BD cellulose.

[8] R.B. Helling, H.M. Goodman and H.W. Boyer, Analysis of endonuclease R. *Eco*RI fragments of DNA from lambdoid bacteriophages and other viruses by agarose gel electrophoresis. J. Virol. 14, 1235 (1974).
[9] M. Thomas and R.W. Davis, Studies on the cleavage of bacteriophage lambda DNA with *Eco*RI restriction endonuclease. J. Mol. Biol. 91, 315 (1974).
[10] J.R. Cameron and R.W. Davis, manuscript in preparation.
[11] K. Struhl, J.R. Cameron and R.W. Davis, Functional genetic expression of eukaryotic DNA in *E. coli*, Proc. Nat. Acad. Sci. USA 73, 1471 (1976).
[12] K.W. Jones and K. Murray, A procedure for detection of heterologous DNA sequences in lambdoid phage by in situ hybridization. J. Mol. Biol. 96, 455 (1975).
[13] R. Kramer, J.R. Cameron and R.W. Davis, Isolation of bacteriophage λ containing yeast ribosomal RNA genes: screening by in situ RNA hybridization to plaques. Cell 8, 227 (1976).
[14] R.L. White and D.S. Hogness, manuscript in preparation.
[15] M. Thomas R.L. White and R.W. Davis, Hybridization of RNA to double stranded DNA: formation of R-loops. Proc. Nat. Acad. Sci. USA. 73, 2294 (1976).
[16] M. Thomas, A. Jacobson and R.W. Davis, manuscript in preparation.