Gene, 11 (1980) 271–282

A new high-capacity cosmid vector and its use

(Recombinant DNA; *Drosophila melanogaster;* genomic libraries; bacteriophage λ)

Elliot M. Meyerowitz *, Gregory M. Guild *, Louise S. Prestidge and David S. Hogness **

Department of Biochemistry, Stanford University School of Medicine, Stanford, CA 94305 (U.S.A.)

(Received April 13th, 1980)
(Accepted June 25th, 1980)

SUMMARY

The construction of a cosmid, MUA-3, designed for the convenient cloning of eukaryotic DNA segments up to 48 kb in length is described. The cosmid contains all of the plasmid pBR322 with approx. 400 bases of λ DNA, including the cohesive end site, inserted at the pBR322 *Pst*I endonuclease recognition site. Methods for using this vector to construct several types of *Drosophila melanogaster* genomic DNA libraries are given, and libraries made by these methods are characterized. A sheared *Drosophila* DNA-*Eco*RI linker library is shown to stably maintain average *Drosophila* DNA inserts of over 40 kb and up to 48 kb, and the efficiency of producing clones by a partial restriction and ligation method is shown to be over 3×10^5 clones/μg of *Drosophila* DNA.

INTRODUCTION

The use of hybrid plasmid-bacteriophage vectors (cosmids) for cloning long fragments of bacterial DNA in *Escherichia coli* has been described in recent publications (Collins and Hohn, 1978; Collins and Bruning, 1978). There are several advantages that application of the cosmid technique to eukaryotic DNA can be expected to confer; among these is the ability to isolate very large contiguous regions of the DNA sequences adjacent to other cloned regions, regardless of the unique or repetitive character of the adjacent DNA. This ability would allow very rapid isolation of extremely long regions of chromosomes by successive isolation of large, overlapping clones. Another potential advantage of very large clone inserts is that they may allow mapping by in situ hybridization of unique genomic clones to specific regions of mitotic chromosomes.

This paper describes the construction of a cosmid designed for convenient cloning and growth of the largest possible foreign inserts, and illustrates the use of this vector to create a recombinant library of random *Drosophila* chromosome fragments averaging over 40 and up to 48 kb in length.

* Present addresses: (E.M.M.) Division of Biology, California Institute of Technology, Pasadena, CA 91125 (U.S.A.); (G.M.G.) Department of Biology, University of Pennsylvania, Philadelphia, PA 19104 (U.S.A.).
** To whom reprint requests should be addressed.

Abbreviations: DPN, nicotinamide-adenine dinucleotide; kb, kilobases or kilobase pairs; SDS, sodium dodecyl sulfate; TCA, trichloroacetic acid.

MATERIALS AND METHODS

(a) Materials

Restriction endonucleases were purchased from New England Biolabs, except *Pst*I, which was from Bethesda Research Labs and *Eco*RI, which was either from a preparation by David Finnegan or purchased from Bethesda Research Labs. *E. coli* DNA polymerase I and T4 DNA ligase, *E. coli* DNA ligase, *E. coli* alkaline phosphatase (prepared from Worthington BAPF by centrifugation and resuspension of the resulting pellet in 10 mM Tris · HCl pH 8, 1 mM EDTA), *Eco*RI methylase, λ gene *A* protein, λ*c*I857 DNA, *Eco*RI dodecamer linkers, calf thymus terminal transferase and pBR322 DNA tailed at the *Pst*I site with an average of 8 dG residues were generously given by S. Scherer, I.R. Lehman, T.-S. Hsieh, R.W. Davis, M. Pearson, R.T. White, C.K. Itakura, R.L. Ratliff, and G. Wahl, respectively. ^{32}P-labeled nucleoside triphosphates, ^{3}H-labeled nucleoside triphosphates, bovine serum albumin (fraction V reagent grade), DNase I and T4 polynucleotide kinase were purchased from Amersham, New England Nuclear, Miles, Worthington and P-L Biochemicals, respectively.

(b) Bacteria and plasmids

The *E. coli* strain HB101 is described in Boyer and Roulland-Dussoix (1969); the plasmids pBR322 and pSC101 are described in Bolivar et al. (1977) and Cohen et al. (1973), respectively.

(c) Cosmid construction

Four micrograms of λ*c*I857 *S*7 DNA were ^{32}P end-labeled using T4 polynucleotide kinase by the method of Maxam and Gilbert (1977), except that the phage DNA was not melted. The specific activity of the labeled DNA was 1.4×10^4 cpm/μg. Additional cold λ*c*I857 *S*7 DNA (16 μg) was added to this, and the DNA was ethanol precipitated (by adding NaCl to 0.3 M, 2.5 vol. of ethanol, freezing solid on powdered dry ice and spinning 15 min at $12\,000 \times g$), washed once with 75% ethanol, dried and resuspended in TE (10 mM Tris · HCl, pH 7.5, 1 mM EDTA). The DNA was first digested (37°C, 2 h) with an excess of *Eco*RI endonuclease in a final volume of 500 μl containing 100 mM NaCl, 10 mM $MgCl_2$, 40 mM Tris · HCl, pH 7.5, and 6 mM 2-mercaptoethanol, then incubated at 50°C for 1 h to allow annealing of the λ cohesive ends without annealing of *Eco*RI sticky ends, and finally diluted with 500 μl of H_2O and 250 μl of 5 × *E. coli* ligase buffer (100 mM Tris · HCl, pH 8.1, 5 mM $MgCl_2$, 50 mM $(NH_4)_2SO_4$, 250 μM DPN and 250 μg/ml gelatin). This was followed by 3 μl of 12.5 mg/ml *E.coli* DNA ligase and incubation for 1 h at 37°C to seal the annealed λ cohesive ends. This reaction was stopped by the addition of 100 μl of 100 mM Tris · HCl, pH 8.8, 100 mM NaCl, 20 mM EDTA and 1% Sarkosyl and heated to 65°C for 5 min. The reaction mixture was then layered over two 10–40% linear sucrose gradients (in 200 mM sodium acetate, 10 mM Tris · HCl, pH 7.6, and 10 mM EDTA) and centrifuged at 34 000 rev./min in a Beckman SW40 rotor at 2°C for 15 h. The gradients were fractionated by collecting 0.5 ml aliquots dripped from the bottom of the tubes. 10 μl of alternate fractions were electrophoresed on a horizontal 0.7% agarose gel (McDonnell et al., 1977) in Tris-borate-EDTA buffer (TBE, Peacock and Dingman, 1968). Following ethidium bromide staining (Helling et al., 1974) and illumination with short-wave (254 nm) UV light the fractions containing the ligated λ ends were pooled and the DNA was ethanol-precipitated, washed, dried and resuspended in 200 μl of TE at a concentration of 50 μg/ml. Half of this DNA (5 μg) was brought to 50 mM NaCl, 10 mM Tris · HCl, pH 7.5, 10 mM $MgCl_2$, 6 mM 2-mercaptoethanol and 100 μg/ml gelatin, then digested at 37°C for 1 h with 10 units of *Hin*cII endonuclease. This digest was electrophoresed on a 2.5% horizontal agarose gel in TBE, stained with ethidium bromide, photographed in long wave UV then exposed to Kodak XR-5 X-ray film overnight at room temperature. The autoradiograph revealed which of the numerous ethidium-stained bands contained the end-labeled and ligated λ left and right ends; as expected (Robinson and Landy, 1977), it was a fragment running at about 400 base pairs by comparison to a size standard of ColE1 DNA cut with *Hha*I. This λ *cos* site-containing band was cut from the gel with a razor blade, electroeluted (McDonnell et al., 1977), isopropanol-precipitated (as ethanol precipitation, but using isopropanol) and resuspended in 20 μl of 1 mM EDTA.

Homopolymeric dC tails were added to the isolated *cos*-containing fragment using terminal transferase in

cacodylate buffer under Co^{2+} conditions, as described in Roychoudhury et al. (1978). The average tail length was calculated to be 50 bases assuming that all of the tailed DNA was 400 base pairs in length. 0.5 ng of this dC-tailed λ *cos* fragment were annealed to an approximately equimolar amount of pBR322 that had been restricted with *Pst*I endonuclease and tailed with terminal transferase to give 8-base-long dG tails. Annealing (in 10 mM Tris · HCl, pH 7.4, 0.2 mM EDTA and 100 mM NaCl) was in a volume of 14.3 μl and at a temperature of 65°C for 2.5 min, then 42°C for 2 h; this reaction was then cooled over several hours to 4°C. HB101 cells were transformed with the annealed mix, as described in Wensink et al. (1974), except that the soft L agar was without antibiotic, and the transformed cells were plated on L agar (Lennox, 1955) plus 10 μg/ml tetracycline.

Colonies resulting from this plating were tested for tetracycline and ampicillin resistance by transferring them with sterile toothpicks to plates containing L agar and 10 μg/ml tetracycline or 250 μg/ml ampicillin, respectively, and incubating at 37°C overnight.

Cosmid DNA was extracted from the tetracycline-resistant, ampicillin-sensitive cells by either of two procedures. In the first, a cleared lysate of the appropriate cells (as in Clewell and Helinski, 1969, except that the cells were grown to stationary phase and Triton X-100 (final 0.03% concentration) was used in place of Brij/deoxycholate) was centrifuged in 1.55 g/ml CsCl and 1 mg/ml ethidium bromide at 35 000 rev./min, 20°C for 2 days in a 50 Ti rotor; or at 53 000 rev./min, 20°C overnight in a VTi 65 rotor, and the supercoil band removed, butanol-extracted four times and ethanol-precipitated.

The second procedure for cosmid isolation is a rapid preparation, and consists of making *E. coli* spheroplasts by the procedure of Cameron et al. (1977) and extracting DNA from them by SDS lysis and ethanol precipitation, as described for λ DNA isolation in the same reference.

After growth to stationary phase in the presence of tetracycline, DNA preparations from various hybrid cosmid strains have yielded between 200 ng and 2 μg of supercoiled recombinant molecules per ml of cells. This is without chloramphenicol amplification (Clewell, 1972), which has not been tried.

Gel analysis of the isolated DNA was on horizontal agarose gels as described above. Gel DNA transfers to nitrocellulose followed the method of Southern (1975); such transfers were hybridized by the procedure of Lis et al. (1978) using [^{32}P]cRNA (Wensink et al., 1974) or [^{32}P]DNA labeled by nick translation (Rigby et al., 1977) as probe.

(d) Construction of the *Eco*RI linker library

Drosophila melanogaster DNA was prepared from 0 to 19 h embryos of the Oregon-R wild-type strain by a procedure similar to that of Schachat and Hogness (1973). 2.7 mg of this DNA were sheared to an average size of 50 kb by mixing DNA (75 μg/ml in 0.5 M NaCl, 10 mM Tris · HCl, pH 8.0, 1 mM EDTA) for 50 min at 800 rev./min and 0°C in a VirTis homogenizer, as described in Hogness and Simmons (1964). This sheared DNA was sized on 10–40% sucrose gradients by centrifugation at 23 000 rev./min, 4°C for 18 h in an SW27 rotor. Each SW27 tube held a maximum of 225 μg of DNA. 1-ml fractions of the gradients were ethanol-precipitated, resuspended in TE and the DNA sized on 0.1% agarose gels (McDonnell et al., 1977), using linear monomers and dimers of phage λ DNA and λ restriction fragments as size standards. Fractions containing DNA 40 to 60 kb in length were pooled at a final concentration of 130 μg/ml.

3.4 ml of sized *Drosophila* DNA were treated with *E. coli* DNA ligase to seal any nicks introduced previously by adding 0.85 ml of 5 × *E. coli* DNA ligase buffer, bringing the reaction solution to 37°C and adding 120 units of *E. coli* DNA ligase. After 1 h the DNA was ethanol-precipitated and resuspended in TE at a concentration of 600 μg/ml. Next, the DNA was reacted with *Eco*RI methylase (Hedgepeth et al., 1972) to prevent it from being cut during the later *Eco*RI endonuclease digestion step: 30 μl of the 600 μg/ml *Drosophila* DNA was added to 20 μl of 1 M Tris · HCl, pH 8.0, 4 μl of 0.5 M EDTA, pH 8.0, 1.2 μl of 1 mM *S*-adenosyl-L-methionine and 145 μl of H_2O; this mix was brought to 37°C and 0.6 μl of 0.18 mg/ml *Eco*RI methylase added. The reaction was incubated 20 min at 37°C, then 10 min at 65°C to denature the enzyme. The reaction was then returned to 37°C, and treated with *E. coli* DNA polymerase I to produce blunt ends. This involved addition of 50 μl of 5 × polymerase I buffer (250 mM Tris · HCl, pH 7.8, 25 mM $MgCl_2$, 100 μM each of dCTP, dTTP, dATP and dGTP, 5 mM dithiothreitol and 600 μg/ml bovine serum albumin), then 5 μl 1 M

$MgCl_2$ were introduced to compensate for the EDTA present, and 6 µl of 1.5 mg/ml *E. coli* DNA polymerase I added. The reaction was incubated 30 min at 37°C, heated 10 min at 65°C and chilled on ice; then the DNA was ethanol-precipitated and resuspended in TE at a concentration of 450 µg/ml.

Following polymerase treatment, the *Drosophila* DNA was blunt-end ligated to phosphorylated dodecamer *Eco*RI linkers (prepared by heating 10 µl (2.23 µg) of linkers to 95°C for 5 min, bringing the linkers to a concentration of 50 mM glycine-NaOH, pH 9.4, 60 µM ATP, 10 mM $MgCl_2$ and 14 mM 2-mercaptoethanol in a final volume of 25 µl, and incubating for 60 min at 37°C with 5 units of T4 polynucleotide kinase). A 20 µl aliquot of the 450 µg/ml blunt-ended, methylated *Drosophila* DNA was added to a mixture containing 4 µl of 90 µg/ml phosphorylated linkers and 3 µl of 10 × T4 DNA ligase buffer (200 mM Tris · HCl, pH 7.4, 100 mM $MgCl_2$, 100 mM dithiothreitol, 50 mM ATP, 1 mg/ml bovine serum albumin), and 3 µl of 1 mg/ml T4 DNA ligase introduced. This reaction proceeded 19 h at 4°C and was stopped by a 10 min incubation at 65°C. The mixture was then brought to 37°C, 3.2 µl of 10 × RI buffer (0.9 M Tris · HCl, pH 7.4, 100 mM $MgCl_2$) added, 1.6 µl of 4500 unit/ml *Eco*RI endonuclease put in and the reaction incubated 30 min. After 10 min at 65°C and chilling on ice, the DNA was ethanol-precipitated and resuspended in 25 µl of TE, the final concentration being 360 µg/ml.

Supercoiled MUA-3 DNA was prepared from HB-101 cells bearing MUA-3 by a cleared lysate-CsCl-ethidium bromide procedure (see above). 100 µl of this DNA at 200 µg/ml were brought to 90 mM Tris · HCl, pH 7.4, 10 mM $MgCl_2$ and cut with an excess of *Eco*RI endonuclease for 30 min at 37°C. After heating 10 min at 65°C, the DNA was ethanol precipitated and resuspended in TE at a concentration of 360 µg/ml. 50 µl of this DNA were added to 40 µl H_2O and 10 µl 10 × BAP buffer (100 mM Tris · HCl, pH 7.4, 500 mM KCl, 1 mM EDTA) and 1.1 unit *E. coli* alkaline phosphatase was added. After 30 min at 37°C, the DNA was phenol-extracted 3 times, ether-extracted twice, and ethanol-precipitated. The DNA pellet was resuspended in TE at a concentration of 400 µg/ml.

The MUA-3 and *Drosophila* DNAs were then ligated: 18 µl of 360 µg/ml *Drosophila* DNA with *Eco*RI endonuclease cut linkers attached and 18 µl of 400 µg/ml *Eco*RI nuclease cut MUA-3 DNA were added to 18 µl TE plus 6 µl of 10 × T4 DNA ligase buffer, to give a molar ratio of MUA-3 to *Drosophila* DNA of 11.5 : 1. 6 µl of 1 mg/ml T4 DNA ligase were added and the reaction incubated overnight at 4°C, then stored frozen until packaged.

In vitro λ packaging was done by the method described in Blattner et al. (1978), except that the spermidine-containing buffer of Becker and Gold (1975) (spermidine buffer) was used in place of the M1 buffer (spermidine-putrescine buffer) used by Blattner et al. (1978). Use of the spermidine buffer instead of spermidine-putrescine buffer seemed to be important for the efficient production of full-sized hybrid cosmids (see DISCUSSION). Also, the centrifugation step after the mixture of buffer and freeze-thaw lysate was replaced by a 40 to 50 min incubation on ice; and after the final 60 min room temperature incubation 100 µl of 5 µg/ml DNase I were added to the packaging mix and incubated 7 min at 37°C to reduce its viscosity. The efficiency of each set of packaging extracts was assayed using Charon 9 λ DNA (Blattner et al., 1977); this efficiency was typically 1×10^8 plaques per µg Charon 9 DNA. After this step, the packaging reactions were diluted in plaque storage buffer (100 mM NaCl, 10 mM Tris · HCl, pH 7.4, 10 mM $MgCl_2$, 0.05% gelatin), brought to 1.49 g/ml in CsCl and centrifuged to equilibrium at 44000 rev./min, 4°C in a Ti 50 rotor. 3.2 µl of the ligated *Drosophila*–MUA-3 mix (0.31 µg *Drosophila* DNA) were used in each packaging reaction.

After centrifugation, the CsCl gradient was divided in 0.5 ml fractions and 10 µl of each fraction assayed for packaged cosmids by infecting HB101 cells as described below. The peak fractions were pooled and dialysed overnight against plaque storage buffer. The appropriate amount of this dialysate was added to 0.1 ml of 10 mM $MgCl_2$, 10 mM $CaCl_2$ in a sterile tube, then mixed with 0.2 ml of an overnight culture of HB101 cells (grown in L broth) and the mixture left 15 min at room temperature. Then 1 ml of L broth was added and the mix incubated at 37°C for 45 min; 3 ml of L agar (1.25% agar, liquid at 45°C) were added and after mixing the cells were poured onto plates with L agar plus 10 µg/ml tetracycline, and incubated at 37°C for 2 days.

Individual cosmid clones were isolated by picking a single colony, colony purifying by streaking on a tetracycline plate, and growing in L broth plus 10 µg/

ml tetracycline. The constant presence of tetracycline is necessary for the cell line to retain a hybrid cosmid.

DNA was made from stationary cultures of cosmid cells by the techniques already described for preparing MUA-3 DNA, except that in place of Triton X-100 in the cleared lysate procedure SDS is used. This increases the DNA yield of hybrid cosmid. Cosmid-bearing cell lines were stored by adding dimethyl sulfoxide to stationary phase liquid cultures to a final concentration of 8% and freezing at −20°C or −70°C.

(e) Production of G-C tailed libraries

Drosophila embryonic DNA was sheared to an average 45 kb length and ligated to remove nicks as previously described. Homopolymeric dC tails were added as follows: a solution containing 150 μg/ml *Drosophila* DNA, 0.57 M TES, pH 6.2, 1 mM 2-mercaptoethanol and 20 μM dCTP, including 50 μCi [^{3}H]dCTP was placed at 30 or 37°C, and $CoCl_2$ added to 1 mM. At 0 min calf thymus terminal transferase was added to 950 units/ml, and samples taken at 2 min, 5 min, and then at 5 min intervals. The samples (5 μl) were precipitated with TCA and assayed for radioactivity by scintillation counting to determine the progress of the reaction (Lobban and Kaiser, 1973). The reaction was stopped after addition of 10 to 15 bases per end (usually between 10 and 20 min) by addition of 1/5 vol. of 100 mM Tris · HCl, pH 8.8, 100 mM NaCl, 20 mM EDTA and 1% Sarkosyl. It is useful to run a small scale reaction immediately before a large one to determine the expected rate of the large reaction.

The stopped reaction was centrifuged in a 10–40% sucrose gradient for 17 h at 31 000 rev./min, 20°C in a SW27 rotor. A maximum of 150 μg of DNA was loaded per tube, 1 ml fractions were collected through the tube bottom and the peak fractions, as assayed by measuring absorbance at 260 nm, were pooled, ethanol-precipitated, resuspended in TE to 200 to 300 μg/ml and stored at −70°C.

MUA-3 DNA was prepared for the G-C pool by restricting it with an excess of *Pst*I endonuclease in 50 mM NaCl, 6 mM Tris · HCl, pH 7.4, 6 mM $MgCl_2$, 6 mM 2-mercaptoethanol and 10 μg/ml gelatin for at least 2 h at 30°C, then ethanol-precipitating the DNA and resuspending in TE at 500 μg/ml. After this, the DNA was treated with *E. coli* ligase to seal nicks and tailed to give 10 to 15 dG residues per end, as described above for dC tailing.

The dC-tailed *Drosophila* and dG-tailed MUA-3 molecules were annealed in a ratio of 1 *Drosophila* to between 7 and 30 MUA-3 molecules at a total DNA concentration of from 150 to 400 μg/ml in 0.15 M NaCl, 10 mM Tris · HCl, pH 7.4, and 2 mM EDTA by heating to 65°C 3 min, then 42°C for 2 h, then cooling over a 2 h period to room temperature (25°C) and storing at 4°C overnight or longer before packaging. Packaging and plating were as described above for the *Eco*RI linker library.

(f) Construction of an *Eco*RI partial digest library

Drosophila embryonic DNA was partially digested with *Eco*RI endonuclease (in the reaction conditions described above) to an approximate average size of 50 kb, then centrifuged on a 10–40% sucrose gradient and the 50 kb average fraction determined by 0.1% agarose gel electrophoresis. This fraction had DNA ranging from 10 to 70 kb on the gel with the bulk of the material at 50 kb, and was used to anneal to *Eco*RI endonuclease cut MUA-3 DNA. Preparation of MUA-3 DNA, annealing, ligation, packaging and plating of the library were as described above for the *Eco*RI linker library, except that the MUA-3 DNA was not treated with alkaline phosphatase.

(g) Other methods

Electron microscopy was as described in Wensink et al. (1974). The plasmid pSC101 was used as a circular size reference standard of 9.72 kb on all grids. This size was derived by comparison to λ DNA, taken as 49.0 kb (Blattner et al., 1977).

Unless otherwise mentioned, all restriction digests were in 50 mM NaCl, 10 mM Tris · HCl, pH 7.5, 10 mM $MgCl_2$, 6 mM 2-mercaptoethanol and 100 μg/ml gelatin at 37°C except for *Tha*I, which was reacted at 60°C.

All work was done under P2-EK1 conditions, as specified by the NIH Guidelines.

RESULTS

(a) Design and construction of the cosmid vectors

A useful cosmid vector requires a replication origin so that the vector and its derivatives will replicate autonomously in the host bacterium, a λ phage

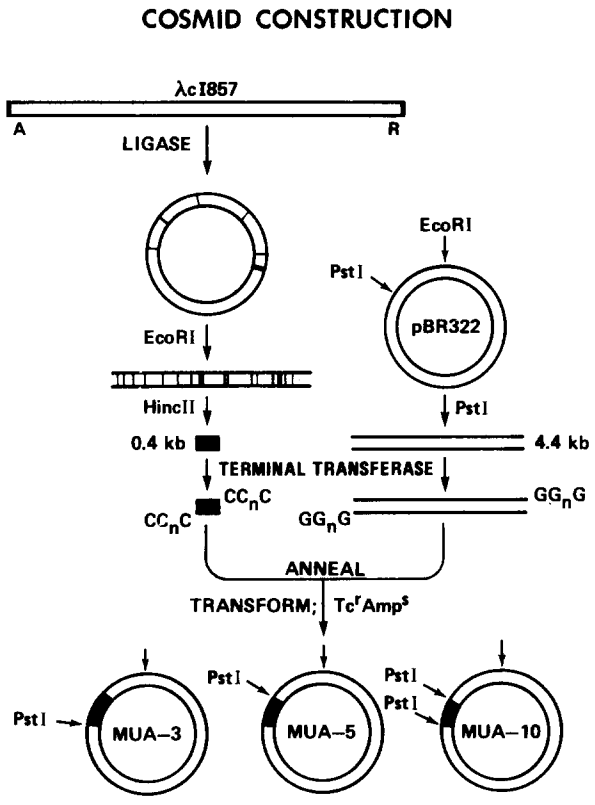

Fig. 1. Diagram illustrating steps used in construction of the MUA series of cosmids. Molecules are not drawn to scale.

cohesive end site (*cos* site, Emmons, 1974) to enable in vitro λ packaging of the vector when attached to foreign DNA, a selectable marker so that cosmid-bearing bacteria can be isolated and so that hybrid cosmids are retained in cell lines, and unique restriction endonuclease recognition sites so that the vector can be linearized and foreign DNA inserted. In addition, a cosmid designed to include the largest possible amount of foreign DNA should be as small as possible, since the total length of foreign DNA that can be cloned in a cosmid is a λ headful (37 to 53 kb, Feiss et al., 1977) minus the size of the vector. Construction of a series of cosmids with all of these properties was accomplished in several steps (Fig. 1).

First, the 400 base pair *Hin*cII restriction fragment of λ DNA that contains the λ phage cohesive end site (Robinson and Landy, 1977) was isolated by ligating the λ cohesive ends together, purifying the *Eco*RI fragment containing them on a sucrose gradient and then *Hin*cII digesting this *Eco*RI product and separating the cohesive end fragment from other *Hin*cII pieces on an agarose gel. The reason for first isolating the *cos* site on an *Eco*RI fragment and then extracting the *Hin*cII piece was that there are *Hin*cII fragments in λ *Eco*RI fragments other than the terminal ones approximately the same size as the *cos Hin*cII fragment. By first purifying the terminal *Eco*RI fragments these pieces were prevented from contaminating the final product. The pure cohesive end piece was tailed with deoxycytidine residues using terminal transferase and annealed to the complementary part of the vector, a pBR322 molecule dG tailed after being cut at its unique *Pst*I site. The pBR322 moiety provided an origin of replication and an intact tetracycline resistance element as well as several unique restriction sites that could later be used for cloning (Bolivar et al., 1977). The *Pst*I site of pBR322 interrupts an ampicillin resistance coding region. After annealing the *cos* and pBR322 molecules and transforming HB101 cells with the annealed DNA mix, 200 tetracycline-resistant colonies were isolated and colony purified. Of these 200 cell lines, 102 were ampicillin-sensitive, as expected if they contained pBR322 molecules interrupted at their *Pst*I site with λ DNA. Seven of these 102 colonies were taken, grown in liquid culture, and their plasmid DNA purified, restricted with both *Eco*RI and *Pst*I and run on an agarose gel. From the gel pattern, it was evident that three different types of plasmids had been constructed. All were pBR322 molecules with a 400 base pair insert at the *Pst*I site. One type had both possible *Pst*I sites regenerated by dG tailing (Bolivar et al., 1977), the other two had only one of the two possible *Pst*I sites regenerated (MUA-10, MUA-3 and MUA-5, respectively, see Fig. 1). MUA-10, with the λ *cos* fragment flanked by *Pst*I sites, was retained to serve as a source of *Pst*I-excisable λ *cos* sites for future cosmid constructions. MUA-3 and MUA-5 are cosmids that should be useful for cloning foreign DNA in their unique *Pst*I site as well as in the other unique pBR322 sites that do not cut in the tetracycline resistance region (such as *Eco*RI, *Ava*II, *Bal*I, *Pvu*II, *Bpa*I and *Pvu*I).

To be certain that the 400 base pair fragment inserted at the *Pst*I site of pBR322 in MUA-3, 5 and 10 was indeed the λ *cos* piece these molecules were cut with the restriction enzymes *Hin*cII, *Hha*I, *Ava*II, *Hpa*II and *Tha*I in various combinations. The gel restriction pattern obtained was in agreement with that predicted from the known sequences of PBR322

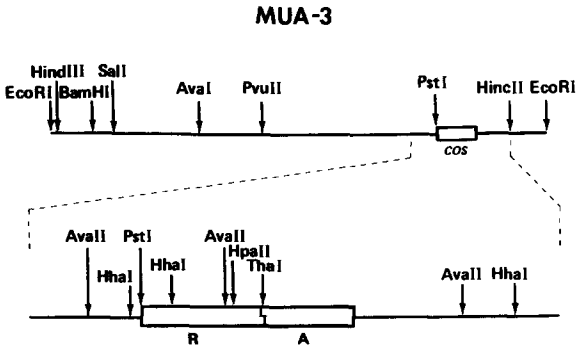

Fig. 2. Restriction map of MUA-3 after *Eco*RI digestion. The orientation of the λ *cos* fragment is the same in MUA-3, MUA-5, and MUA-10; A is the λ left arm and R the λ right arm.

(Sutcliffe, 1978) and of the λ *cos* region (Nichols and Donelson, 1978). The restriction map around the *cos* site in MUA-3 is shown in Fig. 2. As a further check, ^{32}P-labeled RNA complementary to MUA-3, 5 and 10 DNA was used as probe on gel blots (Southern, 1975) of λ*c*I857 *S*7 DNA restricted with either *Hin*dIII or *Ava*I. All three probes hybridized only to the two terminal λ restriction fragments produced by either enzyme (Szybalski and Szybalski, 1979).

(b) Use of MUA-3 in making an *Eco*RI linker recombinant library

The properties of MUA-3 as a vehicle for cloning foreign DNA were tested by producing a library of randomly sheared *Drosophila melanogaster* chromosomal DNA fragments attached to the *Eco*RI site of MUA-3 with synthetic oligonucleotide linkers (Bahl et al., 1977; Scheller et al., 1977) containing the *Eco*RI recognition sequence. The *Drosophila* DNA was from the Oregon-R wild-type strain, and was isolated from 0 to 19 h embryos. This DNA was sheared to an average size of 50 kb, and 40 to 60 kb fragments isolated by sucrose gradient sedimentation. This sized DNA was treated with *E. coli* DNA ligase (Lehman, 1974) to seal any nicks that might have been introduced in the isolation and shearing of the DNA. Nicks were sealed to prevent degradation of the DNA by *E. coli* DNA polymerase I (by translation of nicks toward each other; Kornberg, 1969) in the subsequent polymerase treatment. *E. coli* DNA ligase was used instead of the T4 enzyme because *E. coli* ligase will not join blunt-ended molecules (Lehman, 1974) which should comprise some fraction of the sheared *Drosophila* DNA (Pyeritz et al., 1972). The DNA was then treated with *Eco*RI methylase (Hedgepeth et al., 1972) to prevent its being a substrate for *Eco*RI endonuclease in the subsequent digestion step, and with *E. coli* DNA polymerase I to produce blunt-ended molecules from those with either 5′ or 3′ overlap at the ends by the 5′ to 3′ polymerase and 3′ to 5′ exonuclease activities of the enzyme (Kornberg, 1969; Backman et al., 1976). Polymerase I was used instead of S1 nuclease for this purpose to avoid the degradation of the *Drosophila* DNA that would accompany S1 treatment (Maniatis et al., 1978).

The *Drosophila* DNA was then blunt-end ligated to kinased dodecamer *Eco*RI linkers using T4 DNA ligase. The resulting molecules were reacted with *Eco*RI endonuclease to generate cohesive ends (Mertz and Davis, 1972). MUA-3 DNA was prepared for annealing and ligating to the *Drosophila* DNA by cutting to completion with *Eco*RI endonuclease, then removing the terminal 5′ phosphate groups from these linear molecules with *E. coli* alkaline phosphatase. Phosphatase treatment eliminates one source of recombinant library contamination with non-recombinant vector: *Eco*RI-cut MUA-3 molecules that have been treated with T4 DNA ligase, then in vitro λ packaged and used to infect HB101 cells can give up to 1.6×10^3 tetracycline-resistant colonies/μg of MUA-3 DNA. These colonies contain supercoiled MUA-3 molecules, as determined by DNA extraction and agarose gel analysis. If an identical population of *Eco*RI-cut MUA-3 molecules are treated with alkaline phosphatase prior to ligation, packaging and infection, no tetracycline-resistant colonies are obtained. These phosphatase-treated molecules still serve as efficient cloning vehicles for foreign DNA. The MUA-3 and *Drosophila* DNAs (11.5 : 1 molecular ratio) were mixed; several experiments showed the range of molar ratios of MUA-3 to *Drosophila* DNA between 7 and 15 to 1 to given the optimal yield of tetracycline-resistant colonies. The mixed DNAs were annealed, joined with T4 DNA ligase and in vitro λ packaged. The packaged recombinant cosmids were then centrifuged to equilibrium in an isopycnic CsCl graident to remove unknown substances that inhibit infection of *E. coli* by λ phage and that are present in the packaging extracts (Maniatis et al., 1978) and the cosmid peak from the gradient dialysed overnight against plaque storage buffer. Aliquots of the pack-

aged cosmid mix were adsorbed with HB101 cells, incubated in nutrient broth for 45 min to allow expression of tetracycline resistance, plated in soft agar onto nutrient plates with tetracycline and then incubated at 37°C for 2 days. A total of 5.5 μg of *Drosophila* DNA were packaged, and 13 201 tetracycline-resistant colonies were obtained. The efficiency of cloning was thus 2.4×10^3 colonies/μg of *Drosophila* DNA. In a number of small library constructions similar to the large scale experiment described above, the cloning efficiency varied between 10^3 and 10^4 colonies/μg of *Drosophila* DNA.

(c) Analysis of the recombinant library

The contents of the *Eco*RI linker library were tested in several ways. The first was to pick 19 random colonies, grow them in liquid culture, extract DNA by the rapid procedure and to analyze the DNA by restriction analysis and electron microscopy. *Eco*RI digestion and agarose gel electrophoresis showed that three of the 19 colonies contained nonrecombinant MUA-3 molecules, one contained a recombinant MUA-3 molecule with only 1.5 kb foreign DNA and the remaining 15 had inserts that gave rise to a large number of restriction fragments. All of the 15 clones with large inserts showed an *Eco*RI fragment that comigrated with linear MUA-3 on agarose gels, otherwise the clones were different. Nine of these 15 clones were picked at random and carefully sized by contour length measurements in the electron microscope. The clones were on average 50.3 kb with a range from 45.1 to 52.9 kb. Subtracting the MUA-3 vector contribution of 4.8 kb from each clone gives an average *Drosophila* insert of 45.5 kb, with a range from 40.3 to 48.1 kb (Table I).

A second method used to check the presence and size of inserts in the *Eco*RI linker library was to combine colonies from the library in sets of 192 and to analyze DNA from the sets. The colonies were combined by inoculating 192 different clones onto each of 60 plates (for a total of 11 520 independent clones), letting the colonies grow 2 days at 37°C and then making a slurry of the colonies on each plate in 6 ml of L broth. A 4 ml fraction of each slurry was pelleted, lysed and its supercoiled cosmid DNA population isolated. Two of the resulting DNA samples were run on a 0.5% agarose gel both uncut and cut by *Eco*RI, the remaining 58 samples were

TABLE I

Lengths of recombinant cosmids

Cosmid	Mean length ± standard deviation (kb) [a]	Length of *Drosophila* insert (kb) [b]
mDm173A1	50.0 ± 2.5	45.2
mDm173A12	52.9 ± 3.0	48.1
mDm173D6	48.3 ± 1.6	43.5
mDm173H2	49.3 ± 1.1	44.3
mDm173H11	45.1 ± 3.8	40.3
mDm174A2	51.4 ± 2.4	46.6
mDm174A11	52.9 ± 2.9	48.1
mDm174C7	50.9 ± 1.8	46.1
mDm174H1	52.0 ± 1.8	47.2

[a] The mean length and standard deviation are given for each clone, as determined by electron microscopy (N = 11 ± 1).

[b] The lengths of the *Drosophila* inserts are the mean length of the entire clone minus 4.8 kb, the length of MUA-3.

electrophoresed only after *Eco*RI digestion. Both of the uncut pools of 192 showed the same gel pattern: a very strong smeared band immediately under the sample leading well and a weaker band that comigrated with supercoiled MUA-3. Thus, most of the DNA in each pool was in very large molecules, and some was nonrecombinant MUA-3, exactly as indicated by the analysis of individual clones. The 60 *Eco*RI cut tracks showed a strong band at 4.8 kb, equal to the size of MUA-3, and a faint background of hundreds of bands ranging from over 25 to less than 1 kb. The exact pattern of bands was different in each track, as expected if each pool of 192 clones contained different fragments of *Drosophila* DNA.

In addition to the simple tests of random clones described above, individual clones representing well studied unique regions of the *Drosophila* genome have been isolated from the library (S. Artavanis-Tsakonas, M. Muskavitch, W. Bender, E. Meyerowitz, work in progress). Preliminary results indicate that 5 of 6 clones isolated by probing the library with unique restriction fragments from λ cloned *Drosophila* DNA have inserts greater than 35 kb and are identical in restriction map to the region of the *Drosophila* genome that they represent. The one anomalous clone has an insert of only 17 kb, but still contains what by restriction map is unaltered *Drosophila* genomic DNA.

The results obtained by analysis of clones isolated using a repetitive DNA probe, from the *Drosophila* histone coding region (cDm500 probe; Lifton et al., 1977), are different. The *Drosophila melanogaster* histone genes exist as an array of tandemly repeated five gene units which are interrupted at an average interval of 40 kb by nonhistone DNA array spacers (Lifton et al., 1977; Goldberg, 1979). Of 40 cosmid histone clones isolated (R.P. Lifton, unpublished), the majority have *Drosophila* DNA inserts smaller than 35 kb, the range being from 1 to 48 kb. Analysis of the 40 clones has shown that the 7 largest are the only clones containing array spacer DNA. These seven clones have inserts between 25 and 48 kb. All of the rest of the clones contain only histone gene repeat units. The number of these tandem units ranges from a fraction to four. All of these clones could have arisen by excision of entire repeat units by homologous recombination (Cohen et al., 1978). The apparently deleted clones are all stable and do not continue to excise histone repeats during growth.

(d) Other types of libraries

In addition to the *Eco*RI-linker library described above, several small *Drosophila*-recombinant libraries have been made by two other methods: homopolymer tailing (Lobban and Kaiser, 1973; Jackson et al., 1972) and restriction-ligation (Mertz and Davis, 1972).

In the first of these library types, *Drosophila* embryonic DNA was sheared to an average length of 45 kb and a 40 to 60 kb fraction was purified by sucrose gradient sedimentation. This DNA was tailed to an average of 10 dC residues per end with terminal transferase. MUA-3 DNA was cut with *Pst*I and dG tailed, also to an average of 10 bases per end, thus regenerating *Pst*I sites. *Drosophila* and MUA-3 DNAs were then annealed in concentrations of from 30 to 220 μg/ml *Drosophila* to 110 to 160 μg/ml MUA-3 DNA, and the annealed mix was packaged and plated as described for the *Eco*RI-linker library. Between 5×10^2 and 5×10^3 colonies were obtained per μg of *Drosophila* DNA in different experiments. Eight clones chosen at random from one such experiment were colony-purified, grown in liquid medium and their DNA extracted by the rapid procedure. All eight had foreign DNA inserts, by adding the gel sizes of the *Eco*RI restriction fragments the total clone sizes ranged from 39.1 to 55.0 kb, with a mean of 44.4 kb. These sizes are likely to be underestimates, since under the conditions in which the sizing gels were run, restriction fragments less than 1 kb would have been run off the end of the gel. All eight clones showed a 4.8 kb *Pst*I fragment, indicating that, as expected, the *Pst*I sites of the MUA-3 molecule were regenerated by dG tailing. That the inserts in at least two of these eight clones are identical with restriction fragments from the homologous part of the *Drosophila* genome was shown by using two of the cosmids, ^{32}P-labeled by nick translation, as probes on a Southern (1975) blot of *Eco*RI-cut *Drosophila* genomic DNA. The two clones contained only unique *Drosophila* DNA sequences, and gave a Southern hybridization pattern identical to their *Eco*RI-restriction pattern except for two bands, presumably the two derived from the termini of the sheared *Drosophila* DNA molecule.

The efficiency of homopolymer tailing and *Eco*RI-linker libraries (up to 10^4 colonies/μg foreign DNA) is adequate for making recombinant libraries of the relatively small *Drosophila* genome; under 3500 clones of average 45 kb insert size will represent each fragment of the *Drosophila* genome an average of one time. This efficiency is inadequate for convenient production of libraries of genomic DNA from organisms, such as mammals, with genomes many times the size of that of *Drosophila*. Consequently, a third type of recombinant production method was attempted. This was to partially digest *Drosophila* genomic DNA with *Eco*RI endonuclease, take an average 50 kb-size fraction from a sucrose gradient and then to anneal this DNA to a ten-fold number excess of *Eco*RI-cut MUA-3 molecules in the presence of T4 DNA ligase. This ligated DNA was packaged and plated as for the other libraries. The resulting plates showed an efficiency of 3.1×10^5 colonies/μg *Drosophila* DNA. This *Eco*RI partial-digest library was not characterized further.

DISCUSSION

These experiments demonstrate that MUA-3 can be used as a cosmid and that it is possible to use it in the construction of recombinant libraries of eukaryotic genomic DNA that have average inserts of over

40 kb, and inserts that range to almost 50 kb in length. That MUA-3 acts as a cosmid demonstrates that the part of the λ genome required for in vitro λ packaging, and thus the part of the genome recognized by the gene *A* protein (Wang and Kaiser, 1973) is less than 400 base pairs. In principle, the exact extent of the site of action of the gene *A* protein could be determined by constructing a series of plasmids containing smaller and smaller λ DNA *cos* fragments, and finding the point at which these plasmids cease to be cosmids.

Ideally, the libraries made with MUA-3 should have no uninserted vectors and no inserted cosmids less than 38 kb. The best analyzed of the recombinant libraries described above, the *Eco*RI-linker library, departs from ideal in both of these respects. The reason for the presence of uninserted MUA-3 in the library is unknown. A possible explanation is that concatemers of MUA-3 molecules were able to form despite the alkaline phosphatase treatment of the *Eco*RI-cut MUA-3 molecules due to the presence of a large number of *Eco*RI linkers, which after ligation and *Eco*RI digestion would contain some molecules with two *Eco*RI cohesive ends. It is likely that MUA-3 concatemers can be packaged and injected into *E. coli* to become viable plasmids (Umene et al., 1978, Collins and Bruning, 1978, and above). If this speculation is correct, removal of excess *Eco*RI linkers from the eukaryotic DNA before ligation to the vector would eliminate uninserted vector molecules from a recombinant library. Regardless of whether removal of excess linkers would have any effect, it is clear that contamination with nonrecombinant vector molecules occurs at a level low enough that it is of little or no consequence in screening a cosmid library for specific cloned inserts.

The reason for clones in the library smaller than the size that λ phage heads will normally package is also unknown. That some of these clones are due to intramolecular homologous recombination causing deletions is strongly suggested by the histone clone results described above, and by published examples of similar deletions occurring in plasmids growing in *recA*$^-$ cells (Cohen et al., 1978). It is also conceivable that undersized clones containing nonrepetitive DNA inserts could be formed by the same mechanism. If two small *Drosophila*-DNA fragments (for example 20 kb) were present in the population of molecules ligated to MUA-3, and a molecule was formed consisting of the two *Drosophila* fragments alternating with three MUA-3 molecules, all of the DNA between the two terminal *cos* sites, including both *Drosophila* molecules and the central MUA-3 molecule, would be likely to be packaged in one phage head (Emmons, 1974). If the large plasmid formed after infection of a bacterium with this phage had both of its MUA-3 molecules in the same orientation, homologous recombination between them would result in a cell bearing two undersized inserted cosmids. After a number of cell divisions, these plasmids would segregate and result in cells with a single half-sized inserted cosmid. If this speculation is correct, some undersized cosmids could be eliminated by very careful sizing of the DNA to be inserted. Even if such a measure were ineffective, and this model incorrect, it is clear from the results above that undersized cosmids are infrequent enough to represent only a minor inconvenience in the use of cosmid libraries.

It should be noted that insert size in recombinant cosmids may be affected by the in vitro λ packaging conditions used in producing the hybrids. All of the experiments here used spermidine buffer (Becker and Gold, 1975) for the packaging reaction. The reason for this was that several preliminary experiments showed a high frequency of cosmids with small inserts resulting from packaging in spermidine-putrescine buffer (Blattner et al., 1978), while the same DNA packaged in spermidine buffer selected almost exclusively full-sized inserts. This result is consistent with reports detailing the effects of λ genome size on in vitro λ packaging efficiency (Sternberg et al., 1977; Hohn and Murray, 1977).

The efficiency of cloning *Drosophila* DNA in MUA-3 depends on the method used in the cloning. With tailed clones efficiencies of 5×10^2 to 5×10^3 colonies/μg *Drosophila* DNA were found, with *Eco*RI linker clones efficiences from 10^3 to 10^4 clones/μg were obtained and with *Eco*RI partial clones the efficiency was well over 10^5 colonies/μg *Drosophila* DNA. It is likely that further experimentation would reveal ways to improve the efficiency of all three types of libraries. For *Drosophila* all of these efficiencies are high enough to enable libraries containing several genomes of inserts to be made without difficulty. For cloning genomes many times larger than *Drosophila*, such as those of mammals or amphibians, partial restriction libraries are clearly the method of choice.

The techniques described in this paper thus demonstrate that it is possible to create sizeable recombinant libraries containing very large fragments of eukaryotic chromosomal DNA, with little contamination by smaller fragments. By using efficient screening techniques for hybridization detection of specific colonies from pools of recombinant plasmids (Gergen et al., 1979, for example) it should be possible to easily screen such libraries for any genomic region for which there exists a probe.

ACKNOWLEDGEMENTS

This work was supported by research grants from the National Science Foundation, the National Institute of Health and the American Cancer Society. E.M.M. and G.M.G. were fellows of the Jane Coffin Childs Memorial Fund for Medical Research, G.M.G. was also supported in part by the American Cancer Society – California Division.

REFERENCES

Backman, K., Ptashne, M. and Gilbert, W.: Construction of plasmids carrying the *c*I gene of bacteriophage λ. Proc. Natl. Acad. Sci. USA 73 (1976) 4173–4178.

Bahl, C.P., Marians, K.J., Wu, R., Stawinsky, J. and Narang, S.: A general method for inserting specific DNA sequences into cloning vehicles. Gene 1 (1977) 81–92.

Becker, A. and Gold, M.: Isolation of the bacteriophage lambda *A*-gene protein. Proc. Natl. Acad. Sci. USA 72 (1975) 581–585.

Blattner, F.R., Williams, B.G., Blechl, A.E., Denniston-Thompson, K., Faber, H.E., Furlong, L.A., Grunwald, D.J. Kiefer, D.O., Moore, D.D., Schumm, J.W., Sheldon, E.L. and Smithies, O.: Charon phages: safer derivatives of bacteriophage lambda for DNA cloning. Science 196 (1977) 161–169.

Blattner, F.R., Blechl, A.E., Denniston-Thompson, K., Faber, H.E., Richards, J.E., Slightom, J.L., Tucker, P.W. and Smithies, O.: Cloning human fetal-α-globin and mouse α-type globin DNA: preparation and screening of shotgun collections. Science 202 (1978) 1279–1284.

Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W., Crosa, J.H. and Falkow, S.: Construction and characterization of new cloning vehicles, II. A multipurpose cloning system. Gene 2 (1977) 95–113.

Boyer, H.W. and Roulland-Dussoix, D.: A complementation analysis of the restriction and modification of DNA in *E. coli.* J. Mol. Biol. 41 (1969) 459–472.

Cameron, J.R., Philippsen, P. and Davis, R.W.: Analysis of chromosomal integration and deletions of yeast plasmids. Nucl. Acids Res. 4 (1977) 1429–1448.

Clewell, D.B.: Nature of ColE1 plasmid replication in *Escherichia coli* in the presence of chloramphenicol. J. Bacteriol. 110 (1972) 667–676.

Clewell, D.B. and Helinski, D.R.: Supercoiled circular DNA-protein complex in *E. coli:* purification and induced conversion to an open circular DNA form. Proc. Natl. Acad. Sci. USA, 62 (1969) 1159–1166.

Collins, J. and Bruning, H.J.: Plasmids useable as gene-cloning vectors in an in vitro packaging by coliphage λ: "cosmids". Gene 4 (1978) 85–107.

Collins, J. and Hohn, B.: Cosmids, a new type of plasmid gene-cloning vector which is packageable in vitro in lambda bacteriophage heads. Proc. Natl. Acad. Sci. USA 75 (1978) 4242–4246.

Cohen. A., Ram, D., Halvorson, H.O. and Wensink, P.C.: Deletions within *E. coli* plasmids carrying yeast rDNA. Gene 3 (1978) 135–147.

Cohen, S.N., Chang, A.C., Boyer, H.W. and Helling, R.B.: Construction of biologically functional bacterial plasmids in vitro. Proc. Natl. Acad. Sci. USA 70 (1973) 3240–3244.

Emmons, S.W.: Bacteriophage lambda derivatives carrying two copies of the cohesive end site. J. Mol. Biol. 83 (1974) 511–525.

Feiss, M., Fisher, R.A., Crayton, M.A. and Egner, C.: Packaging of the bacteriophage chromosome: Effect of chromosome length. Virology 77 (1977) 281–293.

Gergen, J.P., Stern, R.H. and Wensink, P.: Filter replicas and permanent collections of recombinant DNA plasmids. Nucl. Acids Res. 7 (1979) 2115–2136.

Goldberg, M.L.: Sequence analysis of *Drosophila* histone genes. Ph.D. dissertation, Stanford University Department of Biochemistry, 1979.

Hedgepeth, J., Goodman, H.M. and Boyer, H.W.: DNA nucleotide sequence restricted by the RI endonuclease. Proc. Natl. Acad. Sci. USA 69 (1972) 3448–3452.

Helling, R.B., Goodman, H.M. and Boyer, H.W.: Analysis of endonuclease R. *Eco*RI fragments of DNA from lambdoid bacteriophages and other viruses by agarose-gel electrophoresis. J. Virol. 14 (1974) 1235–1244.

Hogness, D.S. and Simmons, J.R.: Breakage of λdg DNA: chemical and genetic characterization of each isolated half-molecule. J. Mol. Biol. 9 (1964) 411–438.

Hohn, B. and Murray, K.: Packaging recombinant DNA molecules into bacteriophage particles in vitro. Proc. Natl. Acad. Sci. USA 74 (1977) 3259–3263.

Jackson, D.A., Symons, R.M. and Berg, P.: A biochemical method for inserting new genetic information into SV40 DNA: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of *E. coli.* Proc. Natl. Acad. Sci. USA 69 (1972) 2904–2909.

Kornberg, A.: Active center of DNA polymerase. Science 163 (1969) 1410–1418.

Lehman, I.R.: DNA ligase: structure, mechanism and function. Science 186 (1974) 790–797.

Lennox, E.S.: Transduction of linked genetic characters of the host by bacteriophage P1. Virology 1 (1955) 190–206.

Lifton, R.P., Goldberg, M.L., Karp, R.W. and Hogness, D.S.: The organization of the histone genes in *Drosophila melanogaster:* functional and evolutionary implications. Cold Spring Harbor Symp. Quant. Biol. 42 (1977) 1047–1051.

Lis, J.T., Prestidge, L. and Hogness, D.S.: A novel arrangement of tandemly repeated genes at a major heat shock site in *D. melanogaster.* Cell 14 (1978) 901–919.

Lobban, P.E. and Kaiser, A.D.: Enzymatic end-to-end joining of DNA molecules. J. Mol. Biol. 78 (1973) 453–471.

Maniatis, T., Hardison, R.C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G.K. and Efstratiadis, A.: The isolation of structural genes from libraries of eucaryotic DNA. Cell 15 (1978) 687–701.

Maxam, A.M. and Gilbert, W.: A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA 74 (1977) 560–564.

McDonnell, M.W., Simon, M.N. and Studier, F.W.: Analysis of restriction fragments of T7 DNA and determination of molecular weights by electrophoresis in neutral and alkaline gels. J. Mol. Biol. 110 (1977) 119–146.

Mertz, J.E. and Davis, R.W.: Cleavage of DNA by RI restriction endonuclease generates cohesive ends. Proc. Natl. Acad. Sci. USA 69 (1972) 3370–3374.

Nichols, B.P. and Donelson, J.E.: 178-nucleotide sequence surrounding the *cos* site of bacteriophage lambda DNA, J. Virol. 26 (1978) 429–434.

Peacock, A.C. and Dingman, C.W., Molecular weight estimation and separation of ribonucleic acid by electrophoresis in agarose-acrylamide composite gels. Biochemistry 7 (1968) 668–674.

Pyeritz, R., Schlegel, R. and Thomas, C.A., Jr.: Hydrodynamic shear breakage of DNA may produce single-chained terminals. Biochim. Biophys. Acta 272 (1972) 504–509.

Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P.: Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. 113 (1977) 237–251.

Robinson, L.H. and Landy, A.: *Hin*dII, *Hin*dIII, and *Hpa*I restriction fragment maps of bacteriophage λ DNA. Gene 2 (1977) 1–31.

Roychoudhury, R., Jay, E. and Wu, R.: Terminal labeling and addition of homopolymer tracts to duplex DNA fragments by terminal deoxynucleotidyl transferase. Nucl. Acids Res. 3 (1976) 863–877.

Schachat, F.H. and Hogness, D.S.: Repetitive sequences in isolated Thomas circles from *Drosphila melanogaster.* Cold Spring Harbor Symp. Quant. Biol. 38 (1973) 371–381.

Scheller, R.H., Dickerson, R.E., Boyer, H.W., Riggs, A.D., and Itakura, K.: Chemical synthesis of restriction enzyme recognition sites useful for cloning. Science 196 (1977) 177–180.

Southern, E.M.: Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98 (1975) 503–517.

Sternberg, N., Tiemeier, D. and Enquist, L.: In vitro packaging of a λ*D*am vector containing *Eco*RI DNA fragments of *Escherichia coli* and phage P1. Gene 1 (1977) 255–280.

Sutcliffe, J.G.: pBR322 restriction map derived from the DNA sequence; accurate DNA size markers up to 4361 nucleotide pairs long. Nucl. Acids Res. 5 (1978) 2721–2728.

Szybalski, E.H. and Szybalski, W.: A comprehensive molecular map of bacteriophage lambda. Gene 7 (1979) 217–270.

Umene, K., Shimada, K. and Takagi, Y.: Packaging of ColE1 DNA having a lambda phage cohesive end site. Mol. Gen. Genet. 159 (1978) 39–45.

Wang, J.C. and Kaiser, A.D.: Evidence that the cohesive ends of mature λ DNA are generated by the gene *A* product. Nature New Biol. 241 (1973) 16–17.

Wensink, P.C., Finnegan, D.J., Donelson, J.E. and Hogness, D.S.: A system for mapping DNA sequences in the chromosomes of *Drosophila melanogaster.* Cell 3 (1974) 315–325.

Communicated by J. Carbon.