ANALYTICAL BIOCHEMISTRY 147, 114–119 (1985)

An Improved DNA Sequencing Strategy[1]

HUN-CHI LIN, SHAU-PING LEI, AND GARY WILCOX[2]

Department of Microbiology and Molecular Biology Institute, University of California, Los Angeles, California 90024

Received September 11, 1984

A modification of Hong's systematic DNA sequencing strategy is described. The original procedure has been simplified and transfectant yield increased. After DNase I limited cleavage in the presence of Mn^{2+}, the single-cut linear DNA does not have to be separated from supercoiled or open circular DNA on an agarose gel. After ligation, the DNA is digested with a second restriction endonuclease for which a unique cleavage site resides between the insert and the first restriction endonuclease cutting site. The original intact DNA is linearized whereas the deleted subclone is not. The background is decreased to an undetectable level. This DNA sequencing strategy was tested on a 1.4-kb DNA fragment containing the *araC* regulatory gene from *Erwinia carotovora*. A set of subclones sufficient to sequence the fragment on both strands was produced in 2 days and the yield was at least 60-fold higher than in the original protocol.

KEY WORDS: DNA sequencing strategy; DNase I.

DNA sequence determination (1,2) has become a very powerful tool for the molecular biologist. Improvements have been reported (3–10) which increase the speed of generating DNA sequence data by Sanger's dideoxy chain-termination method using single-stranded bacteriophage M13 (11). The major problem encountered is the generation of subclones. The random subcloning of restriction fragments has the problem of the preference for cloning certain DNA fragments (4). The shotgun approach (4), as well as other random subcloning procedures (8), initially accumulates DNA sequence information rapidly; however, there are difficulties in finishing small gaps within otherwise continuous sequences. Several nonrandom subcloning strategies have been reported. Some require time-consuming gel fractionation (3,5,6,10) while others require special restriction sites and more manipulations (7,9).

In this paper we describe a method derived from Hong's procedure (6) that eliminates the gel purification step, increases the transfectant yield, and reduces the time required to generate subclones.

MATERIALS AND METHODS

Construction of Original Clones

A 1.4-kb DNA fragment containing the *araC* regulatory gene from *Erwinia carotovora* (Lei *et al.*, manuscript in preparation) was cloned into the *Hin*cII site of plasmid pUC12 (12) and the resulting plasmid designated pSPL8. Minilysate DNA was prepared as described (13) except Tris–EDTA buffer was used instead of Tris–EDTA–glucose buffer and lysozyme was omitted. Plasmid pSPL8 was digested with *Hin*dIII and *Bam*HI, ligated with *Hin*dIII and *Bam*HI double-digested bacteriophage M13mp10 or M13mp11 (14) RF[3] DNA, and transfected into *Escherichia*

[1] This work was supported by Grant GM30896 from the National Institutes of Health.

[2] To whom correspondence should be addressed.

[3] Abbreviations used: PEG, polyethylene glycol; RF, replicative form.

0003-2697/85 $3.00

coli strain JM103 (11). Bacteriophages containing either orientation of insert were characterized by the complementation test (5) and restriction endonuclease analysis and designated as M13mp10-SPL8 or M13mp11-SPL8. RF DNA was prepared from a 0.5 liters overnight culture (TYE broth: 15 g tryptone, 10 g yeast extract, and 5 g NaCl/liter) as described (13) and purified by cesium chloride–ethidium bromide equilibrium centrifugation.

Construction of Systematic Deletion Subclones

(a) DNase I treatment. DNase I (Worthington) was prepared as described (15) and stored at −20°C. 10× DNase I buffer [0.2 M Tris–HCl, pH 7.5, 10 mM $MnCl_2$, 1 mg/ml BSA (6)] was prepared and used within 1 h. RF DNA (12.5 μg) was suspended in 144 μl of 1× DNase I buffer and aliquoted into five 1.5-ml Eppendorf tubes. DNase I was serially diluted to concentrations of 0.5, 0.25, 0.125, 0.062, and 0.031 ng/μl. Two microliters of each diluent were added to each tube (final DNase I concentrations: 20, 10, 5, 2.5, and 1.25 pg/μl) and incubated at room temperature for 5 min. The reaction was stopped by adding an equal volume of phenol (saturated with TE buffer: 10 mM Tris–HCl, pH 7.6, 0.1 mM EDTA). The aqueous layer was removed and extracted with an equal volume of chloroform. Five microliters of each sample were mixed with 1 μl of tracking dye and run on a 1% agarose gel to determine optimal digestion conditions. In our case, the contents of the last three tubes were combined and ethanol precipitated prior to further manipulations.

(b) First restriction endonuclease digestion and PEG precipitation. The DNA pellet was resuspended in 200 μl of universal buffer (50 mM Tris–HCl, pH 7.8, 5 mM $MgCl_2$, 25 mM NaCl, 10 mM $(NH_4)_2SO_4$). M13mp10-SPL8 or M13mp11-SPL8 RF DNA was digested with *Hin*dIII or *Sma*I, respectively. After 2 h of incubation at 37°C, the reaction mixture was heated at 68°C for 10 min. Two hundred microliters of PEG solution (13% polyethylene glycol 8000, 1.6 M NaCl) were added and the mixture placed on ice for 1.5 h. After centrifugation at 4°C for 10 min, the DNA was resuspended in 100 μl of TE buffer, extracted serially with equal volumes of phenol, phenol plus chloroform (50:50) and chloroform, and ethanol precipitated.

(c) Klenow repair, self-ligation, and second restriction endonuclease digestion. The DNA pellet was resuspended in 20 μl of H buffer (1) with 5 units of Klenow fragment (DNA polymerase I large fragment) and incubated at room temperature for 15 min. The reaction was stopped by ethanol precipitation. The DNA pellet was resuspended in 200 μl of T4 DNA ligase buffer with 40 units of T4 DNA ligase and incubated overnight at 4°C. The reaction was stopped by heating at 68°C for 15 min and 100 μl of universal buffer containing 2 or 4 μl of 5 N NaCl for M13mp10-SPL8 or M13mp11-SPL8, respectively, was added. Half of each sample was saved as a control and the other half was digested. M13mp10-SPL8 or M13mp11-SPL8 was digested with *Pst*I or *Bam*HI, respectively, for 2 h at 37°C. Five microliters of each sample were transfected into 0.2 ml of competent *E. coli* strain JM103 and plated as described (11).

(d) Characterization of subclones. Approximately 100–150 plaques per plate were obtained. Isolated plaques were grown in 5 ml of TYE broth supplemented with JM103 bacteria for 8 h or overnight as described (11). Bacteriophages were harvested and minilysate RF DNA was prepared as above. The size of each deletion was determined on a 1.5% agarose gel after double digestion with *Bgl*II plus *Bam*HI or *Bgl*II plus *Hin*dIII for M13mp10-SPL8 subclones or M13mp11-SPL8 subclones, respectively.

RESULTS AND DISCUSSION

DNase I Treatment

The strategy of systematic subcloning is outlined in Fig. 1. The appropriate DNase I treatment can be easily determined by starting

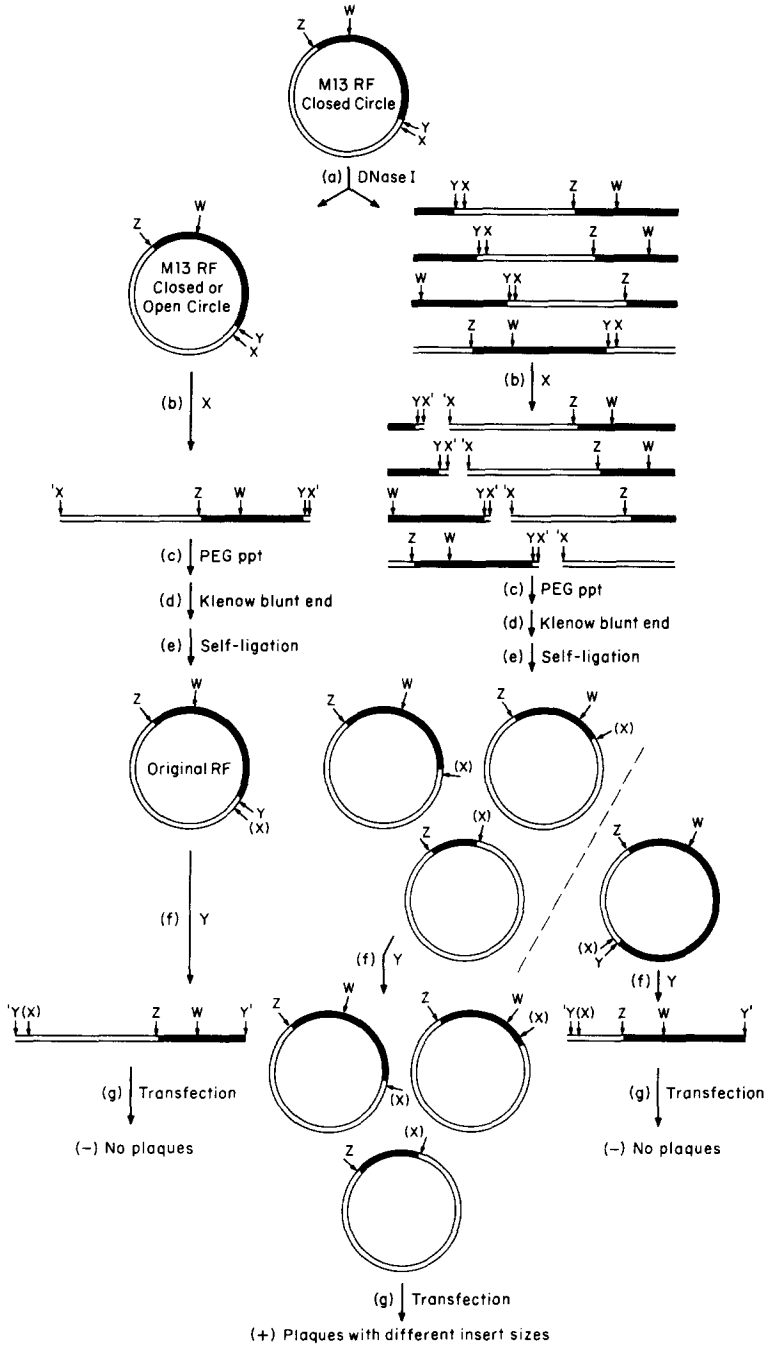

FIG. 1. Outline of the sequencing strategy. W, X, Y, and Z represent different restriction endonuclease recognition sites. Solid lines indicate the insert DNA. Open lines indicate the M13 DNA. The priming site resides next to the X site. (X) indicates the altered X site.

with the conditions described under Materials and Methods. Once the conditions are established, the same conditions can be used for different DNAs. We tested four different clones of M13 RF DNA prepared after CsCl–ethidium bromide equilibrium centrifugation with satisfactory results. In addition, we tested a DNA prepared by the alkaline method (13), which was then RNase A treated, PEG precipitated to remove RNA, phenol extracted, and ethanol precipitated. The results obtained from this DNA preparation were similar to those obtained with the DNA prepared on the CsCl–ethidium bromide gradient.

The results of DNase I treatment are shown in Fig. 2. The samples digested with 1.25, 2.5, or 5 pg/μl of DNase I (final concentrations) were pooled. The decision as to which sample to use for further manipulation is not based on how much DNA was linearized. Instead, the sample with less smear on the gel is preferred because double digestions within the insert region increase with prolonged digestion and, therefore, might confuse later the interpretation of the DNA sequence results.

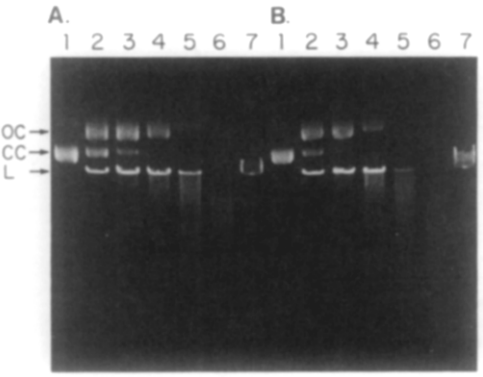

FIG. 2. DNase I treatment of bacteriophage M13 RF DNA. Panels (A) and (B) represent M13mp10-SPL8 and M13mp11-SPL8 RF DNA, respectively. Lane 1, untreated DNA. Lane 2–6, DNA treated with 1.25, 2.5, 5, 10, or 20 pg/μl of DNase I. Lane 7, single-cut linearized DNA. Approximately 0.25 μg of each DNA was electrophoresed on a 1% agarose gel. OC, CC, and L indicate the position of open circular, closed circular, and single-cut linearized DNAs, respectively.

DNase I slightly favors cleavage of the double-stranded DNA at or close to the same position on both strands in the presence of Mn^{2+} ion (4). However, we observed that most supercoiled DNA was nicked on a single-strand, converting the supercoiled structure to the open circular structure as shown in Fig. 2. In several conditions tried, no more than 15% of the total original DNA was converted to single-cut linearized DNA. Hong's protocol requires separation of linearized DNA from the supercoiled and open circular DNAs by agarose gel electrophoresis. The efficiency of isolating DNA fragments larger than 7 kb from the agarose gel is low. The M13mp series of bacteriophages (14) is larger than 7 kb. With inserts lower recovery rates from the gel will occur and therefore more DNA will be required for an experiment. In addition, the supercoiled DNA runs close to the linear DNA on the agarose gel (Fig. 2) and it is possible to have cross-contamination. The unwanted background due to contamination from the supercoiled DNA will be greatly increased if the linear DNAs are not completely separated from the supercoiled DNAs.

First Restriction Endonuclease Digestion, PEG Precipitation, Klenow Repair, Self-ligation, and Second Restriction Endonuclease Digestion

The purpose of the first restriction endonuclease digestion is to orderly release part of the insert in order to generate deletions as shown in Figure 1. PEG precipitates out the larger DNA fragments (16), removing small DNA fragments from the ligation mixture. All the DNA fragments containing the first restriction site were cleaved including those which were not linearized after DNase I treatment. After the first restriction endonuclease digestion, both ends can be blunted by Klenow enzyme in order to facilitate the

self-ligation. After ligation, most circular DNAs were the original clone with the first restriction site altered. The background reduction relies on a second restriction endonuclease treatment. The elimination of the original clones is based on the idea that the subclones should not contain the second restriction site if they have been cut within the insert by the DNase I. A single restriction site located between the first restriction site and the expected DNase I cleavage region was chosen so that digestion with the corresponding restriction endonuclease destroyed the original clones. In our case, the second restriction endonuclease was *Pst*I or *Bam*HI for M13mp10-SPL8 or M13mp11-SPL8, respectively. There was a 25-fold decrease in the number of plaques when the ligated DNA mixture was treated with the second restriction endonuclease.

Characterization of Subclones

The most efficient method to screen the subclones is restriction endonuclease digestion followed by analysis on agarose gels. We prepared RF DNAs by the alkaline method from the same tubes used for the preparation of single-stranded DNAs. Fifteen isolated medium-size plaques were picked from each orientation. The result shown in Fig. 3 demonstrated that the deletions are uniformly distributed within the insert in both orientations. Interestingly, there is no bias towards subclones with smaller deletions, which might be expected because of their larger size (9).

Starting from 12 μg of DNA, 160 plaques were obtained by the original protocol (6) and approximately 10,000 plaques were obtained by the protocol described here. The transfactant yield from the improved method is approximately 60-fold higher than that from the original method. Subcloning from a large DNA fragment is the limiting step for DNA sequencing. The procedure described here is so simple that from the DNase I digestion to the transfection and plating steps takes less than 2 days.

REFERENCES

1. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.

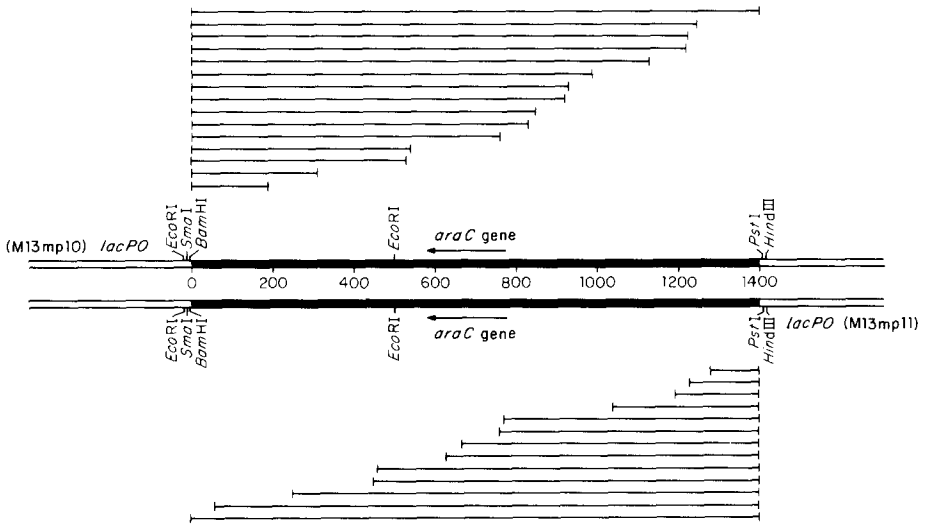

FIG. 3. Distribution of the deletion end points. The solid bar indicates insert DNA. The open bar indicates the M13 DNA. The remaining portion of each subclone is indicated by the horizontal line and the deletion end point is indicated by a vertical line.

2. Maxam, A. M., and Gilbert, W. (1980) *in* Methods in Enzymology (Grossman, L., and Moldave, K., eds.), Vol. 65, pp. 499–560, Academic Press, New York.
3. Frischauf, A. M., Garoff, H., and Lehrach, H. (1980) *Nucl. Acids Res.* **8,** 5541–5549.
4. Anderson, S. (1981) *Nucl. Acids Res.* **9,** 4087–4098.
5. Poncz, M., Solowiejczyk, D., Ballantine, M., Schwartz, E., and Surrey, S. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 4298–4302.
6. Hong, G. F. (1982) *J. Mol. Biol.* **158,** 539–549.
7. Guo, L. H., and Wu, R. (1982) *Nucl. Acids Res.* **10,** 2065–2984.
8. Deininger, P. L. (1983) *Anal. Biochem.* **129,** 216–223.
9. Henikoff, S. (1984) *Gene* **28,** 351–359.
10. Laughon, A., and Scott, M. P. (1984) *Nature (London)* **310,** 25–31.
11. Messing, J., Crea, R., and Seeburg, P. H. (1981) *Nucl. Acids Res.* **9,** 309–321.
12. Vieira, J., and Messing, J. (1982) *Gene* **19,** 259–268.
13. Ish-Horowicz, D., and Burke, J. F. (1981) *Nucl. Acids Res.* **9,** 2989–2998.
14. Messing, J., and Vieira, J. (1982) *Gene* **19,** 269–276.
15. Rigby, P. W. J., Dieckmann, M., Rhodes, C., and Berg, P. (1977) *J. Mol. Biol.* **113,** 237–251.
16. Lis, J. T. (1980) *in* Methods in Enzymology (Grossman, L., and Moldave, K., eds.), Vol. 65, pp. 347–353, Academic Press, New York.