Gene, 33 (1985) 103–119
Elsevier

GENE 1167

Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors

(Recombinant DNA; molecular cloning; polycloning sites; progressive deletions)

Celeste Yanisch-Perron, Jeffrey Vieira and Joachim Messing

Department of Biochemistry, University of Minnesota, St. Paul, MN 55108 (U.S.A.) Tel. (612) 376 1509

(Received July 27th, 1984)
(Accepted September 21st, 1984)

SUMMARY

Three kinds of improvements have been introduced into the M13-based cloning systems. (1) New *Escherichia coli* host strains have been constructed for the *E. coli* bacteriophage M13 and the high-copy-number pUC-plasmid cloning vectors. Mutations introduced into these strains improve cloning of unmodified DNA and of repetitive sequences. A new suppressorless strain facilitates the cloning of selected recombinants. (2) The complete nucleotide sequences of the M13mp and pUC vectors have been compiled from a number of sources, including the sequencing of selected segments. The M13mp18 sequence is revised to include the G-to-T substitution in its gene II at position 6 125 bp (in M13) or 6967 bp in M13mp18. (3) M13 clones suitable for sequencing have been obtained by a new method of generating unidirectional progressive deletions from the polycloning site using exonucleases HI and VII.

INTRODUCTION

Single-stranded DNA isolation has been facilitated by the properties of the single-stranded bacteriophage M13 (Messing et al., 1977). Though it is not a naturally transducing system, recombinant DNA techniques have been used to construct a general transducing system where double-stranded DNA can be introduced into the double-stranded RF of the phage. Upon transfection of appropriate host cells, the DNA strand ligated to the (+) strand of the RF is strand-separated, packaged and secreted without cell lysis as a recombinant single-stranded DNA phage.

Although inserts seven times longer than the wild-type viral genome have been cloned in M13 (Messing, 1981), the presence of large inserts can cause deletions. Accelerated growth of phage containing smaller inserts that arise from deletions makes maintaining large-fragment clones difficult (Messing, 1983). Differential growth can be observed in the plaque-size variety that results from infection of M13 clones possessing inserts of > 2000 bp.

With the introduction of a universal primer (Heidecker et al., 1980), M13 was used primarily for

---

Abbreviations: Ac, activator; Ap, ampicillin; B-broth, Bactotryptone broth; Cm, chloramphenicol; Δ, deletion; DTT, dithothreitol; EMS, ethylmethane sulfonate; Exo III and VII, exonuclease III and VII; HA, hydroxylamine hydrochloride; IPTG, isopropyl-β-D-thiogalactopyranoside; LB, Luria broth; M13UC, see RESULTS, section c2; moi, multiplicity of infection; pfu, plaque-forming units; PHS, primer hybridization site; ^{R}, resistance; RF, replicative form; RT, room temperature; Sm, streptomycin; STE, 10 mM NaCl, 10 mM Tris · HCL pH 7.5, 1 mM EDTA; Tc, tetracycline; Xgal, 5-bromo-4-chloro-indolyl-β-D-galactopyranoside; YT, yeast tryptone; [], indicates plasmid-carrier state; Δ, deletion.

0378-1119/85/$03.30 © 1985 Elsevier Science Publishers

the subcloning or shotgun cloning of small fragments (for review, see Messing, 1983). Cloning of 10 000 bp fragments now appears possible with deletions occurring less frequently than previously predicted. One of the factors interfering with the cloning of large fragments is the restriction of unmodified DNA. JM103, a restrictionless host strain, was developed to circumvent this problem (Messing et al., 1981), but later lost the mutation (Felton, 1983; Baldwin, T., personal communication, C.Y.-P. and J.M., unpubl.), and instead carried a P1 lysogen which also contains a restriction and modification system.

This paper describes the construction and characterization of a number of new strains developed for use with the M13 and pUC cloning vectors. Each strain carries a specific set of mutations that help prevent various cloning problems. A complete listing of M13mp18 and pUC19 sequences and restriction sites is also presented. The M13mp sequence was compiled from published data (Van Wezenbeek et al., 1980; Messing et al., 1977; 1981; Farabaugh, 1978; Dickson et al., 1975; Kalnins et al., 1983; Messing and Vieira, 1982; Norrander et al., 1983). Re-sequencing of pUC was necessary, as a number of mutations had been introduced into the pBR322 region of the pUC vectors. A new method of generating unidirectional deletions from a full-length M13 clone of pUC6 by Exo III and VII was used for this purpose. The results have been combined with those published earlier (Ruther, 1980; Vieira and Messing, 1982; Rubin and Spradling, 1983; Stragier, P., personal communication).

TABLE I

E. coli strains and genotypes

E. coli strain	Genotype
JM83	*ara*, Δ(*lac-proAB*), *rpsL*(= *strA*), ϕ80, *lacZ*ΔM15
JM101	*supE*, *thi*, Δ(*lac-proAB*), [F′, *traD*36, *proAB*, *lacI*^{q}*Z*ΔM15]
JM105	*thi*, *rpsL*, *endA*, *sbcB*15, *hspR*4, Δ(*lac-proAB*), [F′, *traD*36, *proAB*, *lacI*^{q}*Z*ΔM15]
JM106	*endA*1, *gyrA*96, *thi*, *hsdR*17, *supE*44, *relA*1, λ^{-}, Δ(*lac-proAB*)
JM107	*endA*1, *gyrA*96, *thi*, *hsdR*17, *supE*44, *relA*1, λ^{-}, Δ(*lac-proAB*), [F′, *traD*36, *proAB*, *lacI*^{q}*Z*ΔM15]
JM108	*recA1*, *endA1*, *gyrA*96, *thi*, *hsdR*17, *supE*44, *relA*1, Δ(*lac-proAB*)
JM109	*recA1*, *endA1*, *gyrA*96, *thi*, *hsdR*17, *supE*44, *relA*1, λ^{-}, Δ(*lac-proAB*), [F′, *traD*36, *proAB*, *lacI*^{q}*Z*ΔM15]
JM110	*rpsL*, *thr*, *leu*, *thi*, *lacY*, *galK*, *galT*, *ara*, *tonA*, *tsx*, *dam*, *dcm*, *supE*44, Δ(*lac-proAB*), [F′, *traD*36, *proAB*, *lacI*^{q}*Z*ΔM15]
DH1	F^{-}, *recA*1, *endA*1, *gyrA*96, *thi*, *hsdR*17, *supE*44, *relA*1, λ^{-}
GM48	*thr*, *leu*, *thi*, *lacY*, *galK*, *galT*, *ara*, *tonA*, *tsx*, *dam*, *dcm*, *supE*44
SL10	Hfr H, *thi*, *sup*°, Δ(*lac-proAB*), *galE*, Δ(*pgl-bio*)
TD1	MC4100, *recA*56, *srlC*300::Tn*10*
MC4100	*araD*, *rpsL*, *thi*, Δ(*lacIPOZYA*)U169
SK1592	*thi*, *supE*, *endA*, *sbcB*15, *hsdR*4
71-18	Δ(*lac-proAB*), *thi*, *supE*, [F′, *proAB*, *lacI*^{q}*Z*ΔM15]

MATERIALS AND METHODS

(a) Strains

The bacterial strains listed in Table I are *E. coli* K-12 derivatives. SK1592 was obtained from Sidney Kushner, DH1 from Jurgen Brosius, GM48 and HB101 from Raymond Rodriguez, and SL10 and TD1 from James Fuchs. Phages λ*b2c* were from Bruno Gronenborn, and f2 from David Pratt. Strains were tested for relevant markers by standard methods and as described below.

(b) Maintenance of strains

Long-term storage of desired strains was accomplished by mixing 1 ml of a stationary-phase culture with 1 ml of glycerol and freezing at −70 °C. Bacteria were revived by streaking aliquots on appropriate selective media and incubating at 37 °C. Short-term working strain stocks were maintained at −20 °C. Bacteria were revived by inoculating 10 ml of broth with 0.05 ml of the stock and shaking overnight at 37 °C. Alternatively, strains containing the F′ from JM101 were maintained on glucose minimal plates for 2–4 weeks at 4 °C.

(c) Media

Bacterial strains were grown in 2YT or LB broth (Miller, 1972) supplemented with 15 μg/ml Tc, 500 μg Sm/ml, or 50 μg/ml Ap when required. Bacteria were plated on M-9 minimal plates (M-9 minimal medium plus 1.5% agar; Miller, 1972) and supplemented with the following as required (per ml): 1 μg thiamine, 40 μg nalidixic acid, 20.1 μg 5-fluorocytosine, 40 μg amino acids, 12.5 μg Cm, and 500 μg Sm. Fusaric acid medium (Tcs broth plus 1.5% agar; Maloy and Nunn, 1981), B-broth medium (Miller, 1972) supplemented with 0.3% yeast extract and 1.4% agar, 1XA medium (Miller, 1972), or MacConkey plates were also used. B-broth plates were supplemented with 0.1 mM IPTG and 0.004% Xgal per plate. The Xgal and IPTG were stored as 2% and 0.1 M stocks, respectively.

(d) Reagents

Tc, Sm, Cm, Ap, amino acids, 5-fluorocytosine, nalidixic acid, chlorotetracycline, and IPTG were obtained from Sigma. Agar, MacConkey agar, yeast extract, and Bacto tryptone were obtained from Difco. Boehringer Mannheim supplied the Xgal. Restriction endonucleases were provided by Amersham, Bethesda Research Labs, New England Biolabs and PL Biochemicals and used as recommended by their suppliers.

(e) Transductions

Lysates of P1Cm1, clr100 and P1*vir* were prepared by heat induction of lysogenic cells (Miller, 1972). Transductions were performed as described by Miller (1972).

(f) Matings

Matings using F′ and Hfr strains were conducted as described by Miller (1972).

(g) Curing of transposon

Elimination of the Tn*10* transposon was accomplished via selection on Tcs media (Maloy and Nunn, 1981).

(h) Transformations

Preparation of competent cells for transformation by pUC plasmid and M13 RF DNA (Hanahan, 1983; Cohen et al., 1972) was modified by using 50 mM $CaCl_2$ and cell harvest at A_{550nm} 0.550–0.720 (0.800–0.890 for JM109). Transformation by M13 RF DNA was as described by Hanahan (1983) and Cohen et al. (1972) except that the heat-shocked cells and DNA were added to B-broth top agar, IPTG, Xgal, and 0.3 ml of host cells grown to an A_{550nm} of 0.720–1.200 (JM109 to 1.200–1.8). No antibiotic amplification period was required for RF DNA. Transformation of each strain were repeated five or more times.

(i) RF and plasmid DNA preparations

pUC plasmid DNA was prepared by inoculation to an A_{550nm} of 0.050 of 2 × YT medium plus Ap with an overnight culture of the plasmid-carrying host strain. The culture was shaken 8–13 h at 37°C before cell harvest. Cm amplification was not required of the multicopy pUC plasmid. RF DNA was prepared by inoculation of 2 × YT medium to an A_{550nm} of 0.05. At A_{550nm} 0.300–0.420, phage supernatant was added to an moi of 10:1, and the culture was shaken 8–13 h at 37°C. DNA was extracted by the Birnboim and Doly method (1979) and purified on CsCl gradients as described by Messing (1983).

(j) Marker tests

The restriction-minus and modification-plus phenotypes were tested by plating 0.1 ml of overnight cultures resuspended in 1XA buffer plus 0.01 M $MgSO_4$ on B-broth plates. 10 μl of the appropriate phage *λb2c* dilutions (either K-12 modified by two serial propagation cycles through JM101 or K-12 unmodified by two serial propagation cycles through HB101) were spotted on the plates. Plaques were counted after overnight incubation at 37°C. To confirm the modification-plus phenotype, an HB101 *λb2c* lysate (K-12 unmodified phage) was propagated in the questioned strain for two serial cycles with dilutions of the resultant phage lysate plated on JM101 and HB101. The HB101 lysate was used as a control. No difference in titer was correlated with a modification-plus phenotype.

Phage M13mp10a containing amber mutations in genes I and II (Messing and Vieira, 1982) and M13mp10w with no amber mutations (Norrander et al., 1983) were tested for the presence of the *su*II suppressor and the F episome that possesses mutation *lacI*^{q} and deletion *lacZ*ΔM15. The presence of the *su*II suppressor in F^{-} strains was confirmed by infection with amber phage T4amN130-N82. Tests confirming the *lacI*^{q} and *lacZ*ΔM15 deletions via blue plaque production required media supplemented with Xgal and IPTG as described (Messing et al., 1977).

Phage f2 tested for the presence of the *traD*36 mutation on the F episome. Unlike M13, the *traD* gene product is required for phage f2 propagation (Achtman et al., 1971). Phage f2's inability to infect an F′ strain confirmed presence of the NIH-recommended *traD*36 mutation on the F′.

The *recA* mutation was tested by streaking cells on M9 minimal plates and irradiating with a wavelength of 254 nm for 90 s at a distance of 22.5 cm (hand-held UV lamp model UVGL-25 from UVP, Inc. of San Gabriel, CA). As a control, half the agar plate was masked with a paper card during UV exposure. Cells were then incubated overnight at 37°C. UV-resistant growth indicated absence of the *recA* mutation.

The *dam* and *dcm* mutations in GM48 and JM110 were confirmed by propagating pUC plasmids or M13 phage in these strains for at least two overnight culture cycles, isolating the DNA as described above, and cleaving the DNA with either *Mbo*I or *Eco*RII. *Mbo*I cleavage correlated with the *dam* mutation and *Eco*RII cleavage with the *dcm* mutation. *Sau*3A cleavage served as a control.

(k) Construction of JM105

Spontaneous Sm^{R} mutants of strains SK1592 were selected for on minimal media plates containing high concentrations (500 μg/ml) of Sm. A chromosomal Δ(*lac-proAB*) deletion was introduced into SK1592*rpsL* by crossing with SL10 and selected for by growth in the presence of Sm, 5-fluorocytosine, and proline. The F′ episome, carrying mutations *traD*36, *proAB*, and *lacI*^{q}*Z*ΔM15, was transferred to SK1592-*rpsL*-Δ(*lac-proAB*) by mating with JM101. The resulting strain was called JM105 and tested for markers as described above.

(l) Construction of JM106 and JM107

RecA^{+}, Tc^{R} derivatives of DH1 were obtained by transducing DH1 with P1Cm1*clr*-100 propagated in TD1. The desired transductants, DH1-Tn*10*-*recA*$^{+}$, were selected by growth on Tc and nalidixic acid, resistance to UV irradiation, and screened for Cm sensitivity.

The Δ(*lac-proAB*) chromosomal deletion was introduced into DH1-Tn*10*-*recA*$^{+}$ by mating with SL10 and selected for by growth on minimal media plates containing nalidixic acid, 5-fluorocytosine, and proline. Positive progeny were further tested for resistance to UV irradiation and production of white colonies on MacConkey plates or B-broth plates plus Xgal and IPTG.

The *recA*$^{+}$, Δ(*lac-proAB*), *gyrA*96 progeny were tested for retention of the Hsd^{-} and Su^{+} (suppressor-plus) phenotypes. Correct progeny were then cured of the Tn*10* transposon by selection on fusaric acid medium. This strain, *recA*$^{+}$, Δ(*lac-proAB*), *endA*1, *gyrA*96, *thi*, *hsdR*17, *supE*44, *relA*1, was called JM106. JM106(F^{-}) was mated with JM101(F′), and the F′ transfer confirmed by blue plaque production when cells were infected with M13mp-10amber. The final strain was designated JM107.

(m) Construction of JM108 and JM109

The RecA^{-} phenotype was introduced into JM106 by transducing with P1*vir* propagated on JC10240 (*recA*$^{-}$, *srlC*::Tn*10*). Progeny were selected for by growth on Tc, proline, nalidixic acid, and 5-fluorocytosine, and screened for inability to grow following exposure to UV light. Further tests affirmed the Hsd^{-} and Su^{+} phenotypes and demonstrated white colonies on MacConkey plates and B-broth plates plus Xgal and IPTG.

Positive progeny were cured of the Tn*10* as before and named JM108. JM108 was mated with JM101; progeny were tested for the presence of the F′. The end product (*recA*1, *endA*1, *gyrA*96, *thi*, *hsdR*17, *supE*44, *relA*1, λ^{-}, Δ(*lac-proAB*), [F′, *traD*36, *proAB*, *lacI*^{q}*Z*ΔM15] was called JM109.

(n) Construction of JM110

Spontaneous Sm^{R} in GM48 was selected for by plating on minimal plates plus Sm. Inability to grow without leucine or threonine confirmed the desired phenotypes. The *Δ(lac-proAB)* was established in GM48-*rpsL* by mating with SL10 and selecting for on minimal plates plus proline, leucine, 5-fluorocytosine and Sm.

Correct genotype confirmation was by screening for white colony production on B-broth plates plus Xgal and IPTG, and by lysis with T4amN130-N82. JM110 was the result of mating GM48-*rpsL-Δ(lac-proAB)* with JM101. Infection by M13mp10amber affirmed the F′ presence.

(o) Generation of clones for sequencing

M13UC RF DNA (2 μg) was digested with 10 units each of *Sst*I and *Bam*HI for 1.5 h at 37°C. Next, 0.12 vol. of 80 mM EDTA, 0.4 vol. of 5 M NH_4 · acetate and 2 vols. of isopropanol were added; after vortexing and RT incubation for 15 min, the DNA was pelleted by centrifugation at 9000 × *g* for 5 min. The pellet was carefully washed with cold 70% ethanol–water and dried under vacuum. It was resuspended in 40 μl Exo III buffer (50 mM Tris · HCl pH 8, 5 mM $MgCl_2$, 1 mM DTT), 8 units of Exo III were added, and it was incubated at 37°C for 20 min, with 2 μl aliquots removed at 1-min intervals to a tube on ice containing 2 μl of 10 × Exo VII buffer (500 mM K · phosphate, pH 7, 80 mM EDTA, 10 mM DTT). Two tubes, each containing the pooled aliquots from a 10-min time period, were thus generated. Then 0.1 unit of Exo VII was added to each, and the tubes were incubated at 37°C for 45 min, and then at 70°C for 15 min. Next, 1.5 μl 0.2 M $MgCl_2$, 0,5 units large fragment DNA polymerase, and 1 μl of an 8-mM solution of dATP, dCTP, dGTP, TTP were added and incubated at RT for 30 min. To a 5-μl (200 ng of DNA) aliquot were added 5 μl 10 × ligation buffer (250 mM Tris · HCl pH 7.5, 100 mM $MgCl_2$, 25 mM hexamine cobalt chloride, and 5 mM spermidine), 5 μl 10 mM ATP, 2.5 μl 0.1 M DTT, and 2 units DNA ligase and the volume was adjusted to 50 μl with H_2O. After a 3-h incubation at RT, the DNA was precipitated as described above. The pellet was resuspended in 40 μl STE (10 mM NaCl, 10 mM Tris · HCl, pH 7.5, and 1 mM EDTA) and 10 μl were used to transform JM105.

Template preparation, sequencing reaction, gel electrophoresis, and data analysis were as described (Carlson and Messing, 1984; Messing, 1983; Larson and Messing, 1983).

RESULTS AND DISCUSSION

(a) *E. coli* hosts

(1) Conjugation mutants

The male-specific *E. coli* bacteriophage M13 requires an F episome for infection of host cells. Current NIH guidelines regarding the use of recombinant DNA discourage the use of *E. coli* strains carrying conjugation proficient plasmids like the F′ episome (Federal Register, 1980). Since the *tra* operon controls conjugation (Achtman et al., 1971), *tra* mutations have been isolated on F′*lac* DNA episomes to develop conjugation deficiencies that still allow infection by M13. M13 vector utility is also based on proper α-complementation between the phage and host β-galactosidase gene. A host strain containing the *Δ(lac-proAB)* deletion on the chromosome and an F′*lac* possessing the *tra* mutation and *lac*ΔM15 deletion was constructed and named JM101 (Messing 1979). JM101 has a suppressor that permitted growth of M13mp7, 8, 9, 10, and 11 phage that contain amber mutations in genes I and II (Messing et al., 1981; Messing and Vieira, 1982).

(2) Restriction mutants

A restrictionless host strain that facilitated cloning of unmodified DNA was constructed and called JM105. As described in MATERIALS AND METHODS, section **k**, the *Δ(lac-proAB)* chromosomal deletion was introduced into SK1592 via an Hfr cross. Although conjugation of F′ with the *traD*36 mutation was reduced by a factor of 10^{-5}, the leaky mutation allowed conjugation at a reduced rate. Therefore, JM101 was used as a donor for the F′ *traD*36*ΔproAB lacI*^{q}*Z*ΔM15 episome in mating experiments by using the complementation of proline as selection and drug resistance as counterselection. The resulting strain did not contain the *supE* mutation of JM101, so it permitted growth of wild-type M13mp10, 11, 18, and 19, but not of the amber mutants. This provides selection for transferring inserts from an M13 amber phage into a wild-type

M13 vector and for obtaining M13 recombinants with inserts in the opposite orientation (Carlson and Messing, 1984). Since JM105 is r^-m^+, any unmodified DNA cloned directly into wild-type M13 vectors and propagated in JM105 is modified but not restricted. The relevant markers have been tested as shown in Table II.

(3) Recombination mutants

E. coli K-12 restriction was not the only cause for reduced efficiency in cloning larger DNA fragments (> 2000 bp) into M13; another source of instability was recombining sequences. Since *recA* mutations reduce recombination, a host with a *recA* mutation would be useful. A new *recA*, r^-m^+, *su*II host for all

TABLE II

Marker tests

(a) Testing for the Hsd^- and Su^+ phenotypes.

λb2c propagated in modifying JM101 or unmodifying HB101 for two serial cycles was spotted on lawns of questioned strains and incubated at 37°C overnight. The JM101-modified and restricted *λb2c* was able to efficiently infect r^+m^+ strains, while HB101 unmodified and unrestricted *λb2c* phage was destroyed in r^+m^+ strains. Amber phage T4amN130-N82 was spotted upon lawns of the strains in question and incubated at 37°C overnight. Only strains possessing the *su*II suppressor gene support growth of the amber phage.

Phage	Phage dilution	Number of plaques on *E. coli* strain:			
		r^+m^+ JM101	r^-m^+ JM105	r^-m^+ JM106	r^-m^+ JM108
λb2c propagated in	10^{-2}	40	lysis	lysis	lysis
HB101 (r^-m^-)	10^{-4}	0	lysis	lysis	lysis
(1 × 10^9 pfu/ml)	10^{-6}	0	8	18	5
λb2c propagated in	10^{-2}	lysis	lysis	lysis	lysis
JM101 (r^+m^+)	10^{-4}	40	lysis	lysis	lysis
(1 × 10^8 pfu/ml)	10^{-6}	0	0	1	0
T4amN130-N82 propagated in JM101 (1 × 10^{10} pfu/ml)	10^{-2}	lysis	0	lysis	lysis

(b) Testing for F′, *traD*36, *lacI*^{q}, *lacZΔ*M15, and for Su^+, Hsd, Rec^-, and Gyr phenotypes.

Infection of plate lawns by spotted dilutions of amber phage M13mp11 + Xgal and +/− IPTG demonstrated presence of the *su*II suppressor gene and the need for IPTG induction of the *lacZ* gene for proper blue plaque production. Phage f2′s infectious incapability indicated presence of the *traD*36 mutation in the host strain. Tests for r^- and m^- were as given above in Table IIa. Plates of freshly streaked cells were subjected to UV irradiation to test for the RecA phenotype. Strain growth on plates containing nalidixic acid affirmed the presence of the *gyr* mutation.

Test	Plaques on *E. coli* strain[a]			
	JM101	JM105	JM107	JM109
Infection with M13mp11 amber				
+ Xgal + IPTG	blue	0	blue	blue
+ Xgal alone	clear	0	clear	clear
Infection with f2	0	0	0	0
Test for r^-	−	+	+	+
Test for m^-	−	−	−	−
Growth after UV exposure[a]	+	+	+	−
Growth on nalidixic acid[a]	0	0	+	+

[a] Two bottom lines refer to bacterial growth.

M13 vectors was thus constructed. Strain DH1, developed for high transformation efficiencies (Hanahan, 1983), was selected as the initial strain because it possessed a *recA*, *hsdR*17, *supE*44 genotype and high transformation efficiency. The *recA* mutation was transduced to *recA*$^+$ (Csonka and Clark, 1978) to permit deletion of the Δ(*lac-proAB*) region producing JM106. After introduction of the F′ from JM101, the resultant strain JM107 could be used in the same manner as JM101 for infection by amber or wild-type M13 or by the pUC plasmids. P1 transduction of JM106 restored the *recA*1 mutation and produced strain JM108. The introduction of JM101's episome into JM108 produced JM109. All four DH1 derivative strains, JM106, JM107, JM108, and JM109, have been screened for the correct r^-m^+ and *su*II phenotype, and JM107 and JM109 have been tested for the ΔM15 deletion, the *lacI*^{q} and *traD* mutations (Table II).

(4) Applications

These new strains broaden applications of the M13mp and pUC plasmid vector systems. JM106 and JM108 may prove useful as hosts for cosmid libraries, because deletion of the chromosomal *lac* DNA can prevent background hybridization. JM109 could be useful as the host for cDNA libraries (Helfman et al., 1983; Heidecker and Messing, 1983) and for examining the expression of mutant proteins in *E. coli*. Transformation efficiencies of all strains have been tested by the transformation protocols of Cohen et al. (1972) and Hanahan (1983). All strains approximated the efficiencies of standards DH1 and 71–18, with JM107 and JM109 proving to be slightly higher (Table III). Higher transformation efficiencies have been reported (Hanahan, 1983) and may be possible for these new strains. This work attempted only to ensure that under defined transformation conditions the new strains gave the same transformation efficiencies as the parental strains.

JM109 has proven useful by virtue of its *recA*1 mutation. Plasmids form multimers when propagated in *recA*$^+$ strains like JM83 (Bedbrook and Ausubel, 1976). JM109 maintains pUC species of a unique size whether monomer or multimer. The *recA*1 mutation destroys the mechanisms for the recombination and/or replication events that produce the multimers.

Although it is not possible to predict whether large fragments cloned into M13 and grown in JM109 will experience fewer deletions than when propagated in JM101, the following observations have been made. When a 4.5-kb fragment of the maize-controlling element activator (Ac) was cloned in M13 and propagated in JM107, deletions were found to extend from the Ac sequences into M13 sequences near the

TABLE III

Transformation efficiencies of pUC18 and M13 RF DNAs

Method	Transformants per μg pUC18 DNA[a] employing *E. coli* recipient strain:			
	DH1	JM105	JM107	JM109
$CaCl_2$	4.2×10^6	8.2×10^6	5.3×10^6	1.0×10^6
Hanahan	6.3×10^6	3.9×10^5	1.4×10^7	1.2×10^7

Method	Transformants per μg M13 RF DNA[a] employing *E. coli* recipient strain:			
	71–18	JM105	JM107	JM109
$CaCl_2$	7.3×10^5	3.9×10^5	2.6×10^5	3.4×10^5
Hanahan	3.0×10^5	1.2×10^5	2.0×10^6	1.9×10^6

[a] Transformation with the plasmid or M13 RF DNA into specified host strains, as to compare the traditional $CaCl_2$ method (Cohen et al., 1972) with the Hanahan (1983) protocol.

```
        10        20        30        40        50        60        70        80
AATGCTACTACTATTAGTAGAATTGATGCCACCTTTTCAGCTCGCGCCCCAAATGAAAATATAGCTAAACAGGTTATTGA

        90       100       110       120       130       140       150       160
CCATTTGCGAAATGTATCTAATGGTCAAACTAAATCTACTCGTTCGCAGAATTGGGAATCAACTGTTACATGGAATGAAA

       170       180       190       200       210       220       230       240
CTTCCAGACACCGTACTTTAGTTGCATATTTAAAACATGTTGAGCTACAGCACCAGATTCAGCAATTAAGCTCTAAGCCA

       250       260       270       280       290       300       310       320
TCCGCAAAAATGACCTCTTATCAAAAGGAGCAATTAAAGGTACTCTCTAATCCTGACCTGTTGGAGTTTGCTTCCGGTCT

       330       340       350       360       370       380       390       400
GGTTCGCTTTGAAGCTCGAATTAAAACGCGATATTTGAAGTCTTTCGGGCTTCCTCTTAATCTTTTTGATGCAATCCGCT

       410       420       430       440       450       460       470       480
TTGCTTCTGACTATAATAGTCAGGGTAAAGACCTGATTTTTGATTTATGGTCATTCTCGTTTTCTGAACTGTTTAAAGCA

       490       500       510       520       530       540       550       560
TTTGAGGGGGATTCAATGAATATTTATGACGATTCCGCAGTATTGGACGCTATCCAGTCTAAACATTTTACTATTACCCC

       570       580       590       600       610       620       630       640
CTCTGGCAAAACTTCTTTTGCAAAAGCCTCTCGCTATTTTGGTTTTTATCGTCGTCTGGTAAACGAGGGTTATGATAGTG

       650       660       670       680       690       700       710       720
TTGCTCTTACTATGCCTCGTAATTCCTTTTGGCGTTATGTATCTGCATTAGTTGAATGTGGTATTCCTAAATCTCAACTG

       730       740       750       760       770       780       790       800
ATGAATCTTTCTACCTGTAATAATGTTGTTCCGTTAGTTCGTTTTATTAACGTAGATTTTTCTTCCCAACGTCCTGACTG

       810       820       830       840       850       860       870       880
GTATAATGAGCCAGTTCTTAAAATCGCATAAGGTAATTCACAATGATTAAAGTTGAAATTAAACCATCTCAAGCCCAATT

       890       900       910       920       930       940       950       960
TACTACTCGTTCTGGTGTTCTCGTCAGGGCAAGCCTTATTCACTGAATGAGCAGCTTTGTTACGTTGATTTGGGTAATGA

       970       980       990      1000      1010      1020      1030      1040
ATATCCGGTTCTTGTCAAGATTACTCTTGATGAAGGTCAGCCAGCCTATGCGCCTGGTCTGTACACCGTTCATCTGTCCT

      1050      1060      1070      1080      1090      1100      1110      1120
CTTTCAAAGTTGGTCAGTTCGGTTCCCTTATGATTGACCGTCTGCGCCTCGTTCCGGCTAAGTAACATGGAGCAGGTCGC

      1130      1140      1150      1160      1170      1180      1190      1200
GGATTTCGACACAATTTATCAGGCGATGATACAAATCTCCGTTGTACTTTGTTTCGCGCTTGGTATAATCGCTGGGGGTC

      1210      1220      1230      1240      1250      1260      1270      1280
AAAGATGAGTGTTTTAGTGTATTCTTTCGCCTCTTTCGTTTTAGGTTGGTGCCTTCGTAGTGGCATTACGTATTTTACCC

      1290      1300      1310      1320      1330      1340      1350      1360
GTTTAATGGAAACTTCCTCATGAAAAAGTCTTTAGTCCTCAAAGCCTCTGTAGCCGTTGCTACCCTCGTTCCGATGCTGT

      1370      1380      1390      1400      1410      1420      1430      1440
CTTTCGCTGCTGAGGGTGACGATCCCGCAAAAGCGGCCTTTAACTCCCTGCAAGCCTCAGCGACCGAATATATCGGTTAT

      1450      1460      1470      1480      1490      1500      1510      1520
GCGTGGGCGATGGTTGTTGTCATTGTCGGCGCAACTATCGGTATCAAGCTGTTTAAGAAATTCACCTCGAAAGCAAGCTG

      1530      1540      1550      1560      1570      1580      1590      1600
ATAAACCGATACAATTAAAGGCTCCTTTTGGAGCCTTTTTTTTTGGAGATTTTCAACGTGAAAAAATTATTATTCGCAAT

      1610      1620      1630      1640      1650      1660      1670      1680
TCCTTTAGTTGTTCCTTTCTATTCTCACTCCGCTGAAACTGTTGAAAGTTGTTTAGCAAAACCCCATACAGAAAATTCAT

      1690      1700      1710      1720      1730      1740      1750      1760
TTACTAACGTCTGGAAAGACGACAAAACTTTAGATCGTTACGCTAACTATGAGGGTTGTCTGTGGAATGCTACAGGCGTT

      1770      1780      1790      1800      1810      1820      1830      1840
GTAGTTTGTACTGGTGACGAAACTCAGTGTTACGGTACATGGGTTCCTATTGGGCTTGCTATCCCTGAAAATGAGGGTGG

      1850      1860      1870      1880      1890      1900      1910      1920
TGGCTCTGAGGGTGGCGGTTCTGAGGGTGGCGGTTCTGAGGGTGGCGGTACTAAACCTCCTGAGTACGGTGATACACCTA

      1930      1940      1950      1960      1970      1980      1990      2000
TTCCGGGCTATACTTATATCAACCCTCTCGACGGCACTTATCCGCCTGGTACTGAGCAAAACCCCGCTAATCCTAATCCT

      2010      2020      2030      2040      2050      2060      2070      2080
TCTCTTGAGGAGTCTCAGCCTCTTAATACTTTCATGTTTCAGAATAATAGGTTCCGAAATAGGCAGGGGGCATTAACTGT

      2090      2100      2110      2120      2130      2140      2150      2160
TTATACGGGCACTGTTACTCAAGGCACTGACCCCGTTAAAACTTATTACCAGTACACTCCTGTATCATCAAAAGCCATGT

      2170      2180      2190      2200      2210      2220      2230      2240
ATGACGCTTACTGGAACGGTAAATTCAGAGACTGCGCTTTCCATTCTGGCTTTAATGAAGATCCATTCGTTTGTGAATAT

      2250      2260      2270      2280      2290      2300      2310      2320
CAAGGCCAATCGTCTGACCTGCCTCAACCTCCTGTCAATGCTGGCGGCGGCTCTGGTGGTGGTTCTGGTGGCGGCTCTGA

      2330      2340      2350      2360      2370      2380      2390      2400
GGGTGGTGGCTCTGAGGGTGGCGGTTCTGAGGGTGGCGGCTCTGAGGGAGGCGGTTCCGGTGGTGGCTCTGGTTCCGGTG

      2410      2420      2430      2440      2450      2460      2470      2480
ATTTTGATTATGAAAAGATGGCAAACGCTAATAAGGGGGCTATGACCGAAAATGCCGATGAAAACGCGCTACAGTCTGAC

      2490      2500      2510      2520      2530      2540      2550      2560
GCTAAAGGCAAACTTGATTCTGTCGCTACTGATTACGGTGCTGCTATCGATGGTTTCATTGGTGACGTTTCCGGCCTTGC

      2570      2580      2590      2600      2610      2620      2630      2640
TAATGGTAATGGTGCTACTGGTGATTTTGCTGGCTCTAATTCCCAAATGGCTCAAGTCGGTGACGGTGATAATTCACCTT

      2650      2660      2670      2680      2690      2700      2710      2720
TAATGAATAATTTCCGTCAATATTTACCTTCCCTCCCTCAATCGGTTGAATGTCGCCCTTTTGTCTTTAGCGCTGGTAAA

      2730      2740      2750      2760      2770      2780      2790      2800
CCATATGAATTTTCTATTGATTGTGACAAAATAAACTTATTCCGTGGTGTCTTTGCGTTTCTTTTATATGTTGCCACCTT

      2810      2820      2830      2840      2850      2860      2870      2880
TATGTATGTATTTTCTACGTTTGCTAACATACTGCGTAATAAGGAGTCTTAATCATGCCAGTTCTTTTGGGTATTCCGTT

      2890      2900      2910      2920      2930      2940      2950      2960
ATTATTGCGTTTCCTCGGTTTCCTTCTGGTAACTTTGTTCGGCTATCTGCTTACTTTTCTTAAAAAGGGCTTCGGTAAGA

      2970      2980      2990      3000      3010      3020      3030      3040
TAGCTATTGCTATTTCATTGTTTCTTGCTCTTATTATTGGGCTTAACTCAATTCTTGTGGGTTATCTCTCTGATATTAGC

      3050      3060      3070      3080      3090      3100      3110      3120
GCTCAATTACCCTCTGACTTTGTTCAGGGTGTTCAGTTAATTCTCCCGTCTAATGCGCTTCCCTGTTTTTATGTTATTCT

      3130      3140      3150      3160      3170      3180      3190      3200
CTCTGTAAAGGCTGCTATTTTCATTTTTGACGTTAAACAAAAAATCGTTTCTTATTTGGATTGGGATAAATAATATGGCT

      3210      3220      3230      3240      3250      3260      3270      3280
GTTTATTTTGTAACTGGCAAATTAGGCTCTGGAAAGACGCTCGTTAGCGTTGGTAAGATTCAGGATAAAATTGTAGCTGG

      3290      3300      3310      3320      3330      3340      3350      3360
GTGCAAAATAGCAACTAATCTTGATTTAAGGCTTCAAAACCTCCCGCAAGTCGGGAGGTTCGCTAAAACGCCTCGCGTTC

      3370      3380      3390      3400      3410      3420      3430      3440
TTAGAATACCGGATAAGCCTTCTATATCTGATTTGCTTGCTATTGGGCGCGGTAATGATTCCTACGATGAAAATAAAAAC

      3450      3460      3470      3480      3490      3500      3510      3520
GGCTTGCTTGTTCTCGATGAGTGCGGTACTTGGTTTAATACCCGTTCTTGGAATGATAAGGAAAGACAGCCGATTATTGA

      3530      3540      3550      3560      3570      3580      3590      3600
TTGGTTTCTACATGCTCGTAAATTAGGATGGGATATTATTTTTCTTGTTCAGGACTTATCTATTGTTGATAAACAGGCGC

      3610      3620      3630      3640      3650      3660      3670      3680
GTTCTGCATTAGCTGAACATGTTGTTTATTGTCGTCGTCTGGACAGAATTACTTTACCTTTTGTCGGTACTTTATATTCT

      3690      3700      3710      3720      3730      3740      3750      3760
CTTATTACTGGCTCGAAAATGCCTCTGCCTAAATTACATGTTGGCGTTGTTAAATATGGCGATTCTCAATTAAGCCCTAC

      3770      3780      3790      3800      3810      3820      3830      3840
TGTTGAGCGTTGGCTTTATACTGGTAAGAATTTGTATAACGCATATGATACTAAACAGGCTTTTTCTAGTAATTATGATT

      3850      3860      3870      3880      3890      3900      3910      3920
CCGGTGTTTATTCTTATTTAACGCCTTATTTATCACACGGTCGGTATTTCAAACCATTAAATTTAGGTCAGAAGATGAAA

      3930      3940      3950      3960      3970      3980      3990      4000
TTAACTAAAATATATTTGAAAAAGTTTTCTCGCGTTCTTTGTCTTGCGATTGGATTTGCATCAGCATTTACATATAGTTA

      4010      4020      4030      4040      4050      4060      4070      4080
TATAACCCAACCTAAGCCGGAGGTTAAAAAGGTAGTCTCTCAGACCTATGATTTTGATAAATTCACTATTGACTCTTCTC

      4090      4100      4110      4120      4130      4140      4150      4160
AGCGTCTTAATCTAAGCTATCGCTATGTTTTCAAGGATTCTAAGGGAAAATTAATTAATAGCGACGATTTACAGAAGCAA

      4170      4180      4190      4200      4210      4220      4230      4240
GGTTATTCACTCACATATATTGATTTATGTACTGTTTCCATTAAAAAAGGTAATTCAAATGAAATTGTTAAATGTAATTA

      4250      4260      4270      4280      4290      4300      4310      4320
ATTTTGTTTTCTTGATGTTTGTTTCATCATCTTCTTTTGCTCAGGTAATTGAAATGAATAATTCGCCTCTGCGCGATTTT
```

```
      4330      4340      4350      4360      4370      4380      4390      4400
GTAACTTGGTATTCAAAGCAATCAGGCGAATCCGTTATTGTTTCTCCCGATGTAAAAGGTACTGTTACTGTATATTCATC

      4410      4420      4430      4440      4450      4460      4470      4480
TGACGTTAAACCTGAAAATCTACGCAATTTCTTTATTTCTGTTTTACGTGCTAATAATTTTGATATGGTTGGTTCAATTC

      4490      4500      4510      4520      4530      4540      4550      4560
CTTCCATAATTCAGAAGTATAATCCAAACAATCAGGATTATATTGATGAATTGCCATCATCTGATAATCAGGAATATGAT

      4570      4580      4590      4600      4610      4620      4630      4640
GATAATTCCGCTCCTTCTGGTGGTTTCTTTGTTCCGCAAAATGATAATGTTACTCAAACTTTTAAAATTAATAACGTTCG

      4650      4660      4670      4680      4690      4700      4710      4720
GGCAAAGGATTTAATACGAGTTGTCGAATTGTTTGTAAAGTCTAATACTTCTAAATCCTCAAATGTATTATCTATTGACG

      4730      4740      4750      4760      4770      4780      4790      4800
GCTCTAATCTATTAGTTGTTAGTGCACCTAAAGATATTTTAGATAACCTTCCTCAATTCCTTTCTACTGTTGATTTGCCA

      4810      4820      4830      4840      4850      4860      4870      4880
ACTGACCAGATATTGATTGAGGGTTTGATATTTGAGGTTCAGCAAGGTGATGCTTTAGATTTTTCATTTGCTGCTGGCTC

      4890      4900      4910      4920      4930      4940      4950      4960
TCAGCGTGGCACTGTTGCAGGCGGTGTTAATACTGACCGCCTCACCTCTGTTTTATCTTCTGCTGGTGGTTCGTTCGGTA

      4970      4980      4990      5000      5010      5020      5030      5040
TTTTTAATGGCGATGTTTTAGGGCTATCAGTTCGCGCATTAAAGACTAATAGCCATTCAAAAATATTGTCTGTGCCACGT

      5050      5060      5070      5080      5090      5100      5110      5120
ATTCTTACGCTTTCAGGTCAGAAGGGTTCTATCTCTGTTGGCCAGAATGTCCCTTTTATTACTGGTCGTGTGACTGGTGA

      5130      5140      5150      5160      5170      5180      5190      5200
ATCTGCCAATGTAAATAATCCATTTCAGACGATTGAGCGTCAAAATGTAGGTATTTCCATGAGCGTTTTTCCTGTTGCAA

      5210      5220      5230      5240      5250      5260      5270      5280
TGGCTGGCGGTAATATTGTTCTGGATATTACCAGCAAGGCCGATAGTTTGAGTTCTTCTACTCAGGCAAGTGATGTTATT

      5290      5300      5310      5320      5330      5340      5350      5360
ACTAATCAAAGAAGTATTGCTACAACGGTTAATTTGCGTGATGGACAGACTCTTTTACTCGGTGGCCTCACTGATTATAA

      5370      5380      5390      5400      5410      5420      5430      5440
AAACACTTCTCAAGATTCTGGCGTACCGTTCCTGTCTAAAATCCCTTTAATCGGCCTCCTGTTTAGCTCCCGCTCTGATT

      5450      5460      5470      5480      5490      5500      5510      5520
CCAACGAGGAAAGCACGTTATACGTGCTCGTCAAAGCAACCATAGTACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGG

      5530      5540      5550      5560      5570      5580      5590      5600
TGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTC

      5610      5620      5630      5640      5650      5660      5670      5680
TCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCAC

      5690      5700      5710      5720      5730      5740      5750      5760
CTCGACCCCAAAAAACTTGATTTGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGAC

      5770      5780      5790      5800      5810      5820      5830      5840
GTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGGCTATTCTTTTG

      5850      5860      5870      5880      5890      5900      5910      5920
ATTTATAAGGGATTTTGCCGATTTCGGAACCACCATCAAACAGGATTTTCGCCTGCTGGGGCAAACCAGCGTGGACCGCT

      5930      5940      5950      5960      5970      5980      5990      6000
TGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCGCTGGTGAAAAGAAAAACCACCCTG

      6010      6020      6030      6040      6050      6060      6070      6080
GCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGA

      6090      6100      6110      6120      6130      6140      6150      6160
AAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCG

      6170      6180      6190      6200      6210      6220      6230      6240
GCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATTACGAATTCGAGCT

      6250      6260      6270      6280      6290      6300      6310      6320
CGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCACTGGCCGTCGTTTTACAACGTCGTGACTGG

      6330      6340      6350      6360      6370      6380      6390      6400
GAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCG

      6410      6420      6430      6440      6450      6460      6470      6480
CACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTTTGCCTGGTTTCCGGCACCAGAAGCGGTGC

      6490      6500      6510      6520      6530      6540      6550      6560
CGGAAAGCTGGCTGGAGTGCGATCTTCCTGAGGCCGATACGGTCGTCGTCCCCTCAAACTGGCAGATGCACGGTTACGAT

      6570      6580      6590      6600      6610      6620      6630      6640
GCGCCCATCTACACCAACGTAACCTATCCCATTACGGTCAATCCGCCGTTTGTTCCCACGGAGAATCCGACGGGTTGTTA

      6650      6660      6670      6680      6690      6700      6710      6720
CTCGCTCACATTTAATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACGCGAATTATTTTTGATGGCGTTCCTATTGGTT

      6730      6740      6750      6760      6770      6780      6790      6800
AAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGTTTACAATTTAAATATTTGCTTATA

      6810      6820      6830      6840      6850      6860      6870      6880
CAATCTTCCTGTTTTTGGGGCTTTTCTGATTATCAACCGGGGTACATATGATTGACATGCTAGTTTTACGATTACCGTTC

      6890      6900      6910      6920      6930      6940      6950      6960
ATCGATTCTCTTGTTTGCTCCAGACTCTCAGGCAATGACCTGATAGCCTTTGTAGATCTCTCAAAAATAGCTACCCTCTC

      6970      6980      6990      7000      7010      7020      7030      7040
CGGCATTAATTTATCAGCTAGAACGGTTGAATATCATATTGATGGTGATTTGACTGTCTCCGGCCTTTCTCACCCTTTTG

      7050      7060      7070      7080      7090      7100      7110      7120
AATCTTTACCTACACATTACTCAGGCATTGCATTTAAAATATATGAGGGTTCTAAAAATTTTTATCCTTGCGTTGAAATA

      7130      7140      7150      7160      7170      7180      7190      7200
AAGGCTTCTCCCGCAAAAGTATTACAGGGTCATAATGTTTTTGGTACAACCGATTTAGCTTTATGCTCTGAGGCTTTATT

      7210      7220      7230      7240      7250      7260      7270      7280
GCTTAATTTTGCTAATTCTTTGCCTTGCCTGTATGATTTATTGGATGTT
```

Fig. 1A. The nucleotide sequence of M13mp18. The sequence has been compiled as described in RESULTS, section **b**, and entered into an Apple II computer. Using the programs described earlier (Larson and Messing, 1983), the sequence has been printed out in the single-strand form using the original *Hin*cII site as reference point. This strand also represents the (+) or message strand. Numbers correspond to bases aligned with the last digit.

NOTE ADDED IN PROOFS:

Fig. 1 was modified in proofs because after our paper went into press, we learned from a publication by Dotto and Zinder [Nature 311 (1984) 279–280] that the M13mp phage vectors contain an altered gene II product. They used marker rescue experiments to characterize a G to T substitution at position 6125 of the M13 wild-type sequence (6967 in Fig. 1A), leading to a methionine-to-isoleucine change in the gene II protein (codon 40). The altered gene II protein is expressed at normal wild-type levels in M13mp infected cells, but compensates for the disruption of domain B of the M13 *ori* region. The presented M13 sequence has been changed accordingly.

```
 Enzyme  Site                        Pos.          Pos.          Pos.
---------------------------------------------------------------------
AccIA    (GTCGAC)            ( 6263)  6263
AhaIIC   (GGCGCC)            ( 6000)  6000
AluI     (AGCT)              (   39)    39 (   24)    63 (  140)   203
                             (   26)   229 (  104)   333 (  600)   933
                             (  554)  1487 (   29)  1516 ( 1446)  2962
                             (  313)  3275 (  336)  3611 (  484)  4095
                             ( 1330)  5425 (  204)  5629 (  331)  5960
                             (   93)  6053 (   64)  6117 (   95)  6212
                             (   25)  6237 (   45)  6282 (   93)  6375
                             (  111)  6486 (  180)  6666 (   63)  6729
                             (  220)  6949 (   27)  6976 (  201)  7177
AvaIA    (CCCGGG)            ( 6246)  6246
AvaIC    (CTCGGG)            ( 5824)  5824
AvaIIA   (GGACC)             ( 5913)  5913
BalI     (TGGCCA)            ( 5079)  5079
BamHI    (GGATCC)            ( 6251)  6251
BanIA    (GGCACC)            ( 5676)  5676 (  454)  6130 (  334)  6464
BanIB    (GGTACC)            ( 6242)  6242
BanIC    (GGCGCC)            ( 6000)  6000
BanID    (GGTGCC)            ( 1248)  1248 ( 5228)  6476
BanIIB   (GAGCTC)            ( 6236)  6236
BanIID   (GGGCTC)            ( 5642)  5642
BbVIA    (GCAGC)             (  931)   931 ( 4604)  5535 (  516)  6051
                             (  303)  6354 (   73)  6427
BbVIB    (GCTGC)             ( 1366)  1366 ( 1154)  2520 (  611)  3131
                             ( 1739)  4870 ( 1052)  5922
BglI     (GCCNNNNNGGC)       ( 6430)  6430
BglII    (AGATCT)            ( 6934)  6934
ClaI     (ATCGAT)            ( 2526)  2526 ( 4355)  6881
DdeI     (CTNAG)             (  233)   233 (  865)  1098 (  272)  1370
                             (   46)  1416 (  367)  1783 (   63)  1846
                             (   15)  1861 (   15)  1876 (   24)  1900
                             (   72)  1972 (   42)  2014 (  303)  2317
                             (   15)  2332 (   15)  2347 (   15)  2362
                             (  998)  3360 (  652)  4012 (   27)  4039
                             (   39)  4078 (   14)  4092 (   28)  4120
                             (  160)  4280 (  600)  4880 (  381)  5261
                             (  672)  5933 (  575)  6508 (  399)  6907
                             (  153)  7060 (  128)  7188
DraI     (TTTAAA)            (  189)   189 (  283)   472 ( 4149)  4621
                             ( 2162)  6783 (  290)  7073
EcoRI    (GAATTC)            ( 6230)  6230
EcoRIIA  (CCAGG)             ( 5940)  5940 (  196)  6136
EcoRIIB  (CCTGG)             ( 1013)  1013 (  952)  1965 ( 4032)  5997
                             (  330)  6327 (  127)  6454
FnuDII   (CGCG)              (   43)    43 (  304)   347 (  771)  1118
                             (   57)  1175 ( 1290)  2465 (  889)  3354
                             (   54)  3408 (  190)  3598 (  353)  3951
                             (  361)  4312 (  681)  4993 (  495)  5488
                             (   24)  5512 (   20)  5532 (  496)  6028
                             (    2)  6030 (  658)  6688 (   63)  6751
Fnu4HI   (GCNGC)             (  931)   931 (  435)  1366 (   27)  1393
```

```
 Enzyme  Site                        Pos.          Pos.          Pos.
---------------------------------------------------------------------
Fnu4HI   (GCNGC)             (  891)  2284 (    3)  2287 (   24)  2311
                             (   45)  2356 (  164)  2520 (  611)  3131
                             ( 1739)  4870 (  629)  5499 (   14)  5513
                             (   22)  5535 (  387)  5922 (  129)  6051
                             (  303)  6354 (   73)  6427
FokIA    (GGATG)             ( 3546)  3546 ( 3697)  7243
FokIB    (CATCC)             (  239)   239 ( 6121)  6360
HaeIIA   (GGCGCC)            ( 6000)  6000
HaeIIB   (AGCGCC)            ( 5558)  5558 (    8)  5566
HaeIIC   (GGCGCT)            ( 6445)  6445
HaeIID   (AGCGCT)            ( 2709)  2709 (  329)  3038
HaeIII   (GGCC)              ( 1395)  1395 (  849)  2244 (  309)  2553
                             ( 2527)  5080 (  158)  5238 (  106)  5344
                             (   69)  5413 (  311)  5724 (  214)  5938
                             (   98)  6036 (  257)  6293 (  102)  6395
                             (  117)  6512 (  169)  6681 (  341)  7022
HgaIA    (GACGC)             (  526)   526 ( 1637)  2163 (  315)  2478
                             (  758)  3236 ( 3450)  6686
HgaIB    (GCGTC)             ( 4082)  4082 ( 1075)  5157
HgiaIA   (GTGCAC)            ( 4742)  4742
HgiaIB   (GAGCTC)            ( 6236)  6236
HgiaIC   (GTGCTC)            ( 5464)  5464
HhaI     (GCGC)              (   44)    44 (  966)  1010 (   74)  1084
                             (   92)  1176 (  293)  1469 (  725)  2194
                             (  272)  2466 (  244)  2710 (  329)  3039
                             (   56)  3095 (  312)  3407 (  190)  3597
                             (  714)  4311 (  683)  4994 (  495)  5489
                             (   13)  5502 (    9)  5511 (   22)  5533
                             (   26)  5559 (    8)  5567 (  434)  6001
                             (   28)  6029 (   65)  6094 (  331)  6425
                             (   21)  6446 (  115)  6561
HincIIA  (GTCGAC)            ( 6263)  6263
HindIII  (AAGCTT)            ( 6281)  6281
HinfI    (GANTC)             (  136)   136 (   80)   216 (  274)   490
                             (   21)   511 (  212)   723 ( 1287)  2010
                             (  486)  2496 (  348)  2844 (  413)  3257
                             (  160)  3417 (  324)  3741 (   96)  3837
                             (  234)  4071 (   45)  4116 (  232)  4348
                             (  771)  5119 (  209)  5328 (   46)  5374
                             (   63)  5437 (  328)  5765 (   22)  5787
                             (  253)  6040 (  221)  6261 (  362)  6623
                             (  261)  6884 (   19)  6903 (  137)  7040
HpaII    (CCGG)              (  314)   314 (  651)   965 (  129)  1094
                             (  829)  1923 (  454)  2377 (   18)  2395
                             (  156)  2551 (  818)  3369 (  472)  3841
                             (  176)  4017 ( 1596)  5613 (  545)  6158
                             (   89)  6247 (  215)  6462 (   18)  6480
                             (  357)  6837 (  123)  6960 (   60)  7020
HphIA    (GGTGA)             ( 1375)  1375 (  398)  1773 (  135)  1908
                             (  489)  2397 (  144)  2541 (   39)  2580
                             (   39)  2619 (    6)  2625 ( 2221)  4846
```

```
Enzyme   Site                         Pos.           Pos.           Pos.
--------------------------------------------------------------------------
HphIA    (GGTGA)        (  270)  5116 (  589)  5705 (  241)  5946
                        (   33)  5979 ( 1025)  7004
HphIB    (TCACC)        ( 1502)  1502 ( 1132)  2634 ( 2288)  4922
                        ( 2108)  7030
KpnI     (GGTACC)       ( 6242)  6242
MboI     (GATC)         ( 1381)  1381 (  332)  1713 (  507)  2220
                        ( 4032)  6252 (  153)  6405 (   96)  6501
                        (  434)  6935
MboIIA   (GAAGA)        ( 2217)  2217 ( 1694)  3911 ( 2479)  6390
MboIIB   (TCTTC)        (  781)   781 ( 3293)  4074 (  196)  4270
                        (  666)  4936 (  318)  5254 (  332)  5586
                        (  917)  6503 (  301)  6804
MnlIA    (CCTC)         (  254)   254 (  119)   373 (  187)   560
                        (   27)   587 (   68)   655 (  383)  1038
                        (   49)  1087 (  143)  1230 (   66)  1296
                        (   21)  1317 (    8)  1325 (   19)  1344
                        (   71)  1415 (   90)  1505 (  391)  1896
                        (   48)  1944 (   75)  2019 (  243)  2262
                        (    6)  2268 (  404)  2672 (    4)  2676
                        (  217)  2893 (  158)  3051 (  269)  3320
                        (   31)  3351 (  351)  3702 (  604)  4306
                        (  391)  4697 (   74)  4771 (  149)  4920
                        (    5)  4925 (  421)  5346 (   69)  5415
                        (  265)  5680 (  340)  6020 (  235)  6255
                        (  277)  6532 (  423)  6955
MnlIB    (GAGG)         (  484)   484 (  141)   625 (  747)  1372
                        (  359)  1731 (  102)  1833 (   15)  1848
                        (   15)  1863 (   15)  1878 (  129)  2007
                        (  312)  2319 (   15)  2334 (   15)  2349
                        (   15)  2364 (    4)  2368 (  967)  3335
                        (  685)  4020 (  799)  4819 (   15)  4834
                        (  612)  5446 (  947)  6393 (  117)  6510
                        (  575)  7085 (  105)  7190
MstI     (TGCGCA)       ( 6424)  6424
MstII    (CCTNAGG)      ( 6507)  6507
NaeI     (GCCGGC)       ( 5612)  5612
NarI     (GGCGCC)       ( 6000)  6000
NciIA    (CCCGG)        ( 6246)  6246
NciIB    (CCGGG)        ( 1923)  1923 ( 4324)  6247 (  590)  6837
NdeI     (CATATG)       ( 2722)  2722 ( 1080)  3802 ( 3043)  6845
NlaIII   (CATG)         (  149)   149 (   47)   196 (  910)  1106
                        (  193)  1299 (  499)  1798 (  235)  2033
                        (  123)  2156 (  698)  2854 (  677)  3531
                        (   87)  3618 (   99)  3717 ( 1461)  5178
                        ( 1043)  6221 (   55)  6276 (  580)  6856
NlaIV    (GGNNCC)       ( 1061)  1061 (  187)  1248 (  292)  1540
                        (   10)  1550 (  252)  1802 (  248)  2050
                        (  323)  2373 (   18)  2391 ( 3252)  5643
                        (   12)  5655 (   21)  5676 (  190)  5866
                        (  134)  6000 (  130)  6130 (  112)  6242
                        (    9)  6251 (  213)  6464 (   12)  6476
PstI     (CTGCAG)       ( 6269)  6269
PvuI     (CGATCG)       ( 6404)  6404
PvuII    (CAGCTG)       ( 5959)  5959 (   93)  6052 (  322)  6374
RsaI     (GTAC)         (  173)   173 (  107)   280 (  741)  1021
```

```
Enzyme   Site                         Pos.           Pos.           Pos.
--------------------------------------------------------------------------
RsaI     (GTAC)         (  143)  1164 (  604)  1768 (   27)  1795
                        (   93)  1888 (   16)  1904 (   65)  1969
                        (  163)  2132 ( 1334)  3466 (  201)  3667
                        (  522)  4189 (  190)  4379 ( 1004)  5383
                        (  102)  5485 (  758)  6243 (  599)  6842
                        (  322)  7164
SacI     (GAGCTC)       ( 6236)  6236
SalI     (GTCGAC)       ( 6263)  6263
Sau96I   (GGNCC)        ( 5723)  5723 (  190)  5913 (   24)  5937
                        (  458)  6395
ScrFI    (CCNGG)        ( 1013)  1013 (  910)  1923 (   42)  1965
                        ( 3975)  5940 (   57)  5997 (  139)  6136
                        (  110)  6246 (    1)  6247 (   80)  6327
                        (  127)  6454 (  383)  6837
SfanIA   (GATGC)        (   25)    25 (  363)   388 (  965)  1353
                        ( 3496)  4849 ( 1696)  6545 (   13)  6558
SfanIB   (GCATC)        ( 3978)  3978
SmaI     (CCCGGG)       ( 6246)  6246
SnaBI    (TACGTA)       ( 1267)  1267
SphI     (GCATGC)       ( 6275)  6275
TaqI     (TCGA)         (  336)   336 (  790)  1126 (  381)  1507
                        (  441)  1948 (  579)  2527 (  927)  3454
                        (  239)  3693 (  971)  4664 ( 1018)  5682
                        (  552)  6234 (   30)  6264 (  618)  6882
XbaI     (TCTAGA)       ( 6257)  6257
XhoIIA   (GGATCC)       ( 6251)  6251
XhoIIB   (AGATCC)       ( 2219)  2219
XhoIID   (AGATCT)       ( 6934)  6934
XmnI     (GAANNNNTTC)   (  357)   357 ( 2288)  2645

 These enzymes did not appear.....AatII    AccIB    AccIC    AccID    AhaIIA
                                  AhaIIB   AhaIID   ApaI     AvaIB    AvaID
                                  AvaIIB   BanIIA   BanIIC   BclI     BsshII
                                  BstXI    BstEII   EcoRV    HgiaID   HincIIB
                                  HincIIC  HincIID  HpaI     MluI     NcoI
                                  NotI     NruI     NsiI     SacII    ScaI
                                  StuI     Tth111I  XhoI     XhoIIC   XmaIII
```

Fig. 1B. Restriction sites of M13mp18. The position (Pos.) and distances (in parentheses) between the neighbouring restriction sites are specified in nucleotides, as compiled and printed out employing the restriction mapping program. (See legend to Fig. 1A).

(A)

```
        10        20        30        40        50        60        70        80
TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGAT

        90       100       110       120       130       140       150       160
GCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGA

       170       180       190       200       210       220       230       240
GCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCC

       250       260       270       280       290       300       310       320
ATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGG

       330       340       350       360       370       380       390       400
GGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATT

       410       420       430       440       450       460       470       480
CGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCT

       490       500       510       520       530       540       550       560
GTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATG

       570       580       590       600       610       620       630       640
AGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAAT

       650       660       670       680       690       700       710       720
GAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCG

       730       740       750       760       770       780       790       800
GTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGG

       810       820       830       840       850       860       870       880
AAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCC

       890       900       910       920       930       940       950       960
GCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCG

       970       980       990      1000      1010      1020      1030      1040
TTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTC

      1050      1060      1070      1080      1090      1100      1110      1120
GGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTG

      1130      1140      1150      1160      1170      1180      1190      1200
TGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGAC

      1210      1220      1230      1240      1250      1260      1270      1280
TTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTG

      1290      1300      1310      1320      1330      1340      1350      1360
GTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAG

      1370      1380      1390      1400      1410      1420      1430      1440
TTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGA

      1450      1460      1470      1480      1490      1500      1510      1520
AAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGAT
```

```
      1530      1540      1550      1560      1570      1580      1590      1600
TTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTA

      1610      1620      1630      1640      1650      1660      1670      1680
TATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCA

      1690      1700      1710      1720      1730      1740      1750      1760
TCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGAT

      1770      1780      1790      1800      1810      1820      1830      1840
ACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTC

      1850      1860      1870      1880      1890      1900      1910      1920
CTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTG

      1930      1940      1950      1960      1970      1980      1990      2000
CGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCA

      2010      2020      2030      2040      2050      2060      2070      2080
ACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAA

      2090      2100      2110      2120      2130      2140      2150      2160
GTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGC

      2170      2180      2190      2200      2210      2220      2230      2240
TTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTC

      2250      2260      2270      2280      2290      2300      2310      2320
AATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCT

      2330      2340      2350      2360      2370      2380      2390      2400
CAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTC

      2410      2420      2430      2440      2450      2460      2470      2480
ACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAAT

      2490      2500      2510      2520      2530      2540      2550      2560
ACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTA

      2570      2580      2590      2600      2610      2620      2630      2640
TTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATC

      2650      2660      2670      2680      2690      2700      2710      2720
ATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTC
```

(B)

```
ENZYME   SITE                          POS.              POS.             POS.
------------------------------------------------------------------------------
ACCIA    (GTCGAC)            (  429)   429
ACYIA    (GGCGCC)            (  235)   235
ACYIB    (GGCGTC)            ( 2235)  2235
ACYID    (GACGTC)            ( 2617)  2617
ALUI     (AGCT)              (   43)    43 (   19)     62 (  245)    307
                             (   96)   403 (   45)    448 (   22)    470
                             (   95)   565 (   64)    629 (  118)    747
                             (  226)   973 (  136)   1109 (  257)   1366
                             (  521)  1887 (  100)   1987 (   63)   2050
ASUI     (GGGCC)             (  286)   286 ( 1534)   1820
AVAIA    (CCCGGG)            (  412)   412
AVAIIB   (GGTCC)             ( 1837)  1837 (  222)   2059
BAMHI    (GGATCC)            (  417)   417
BBVIA    (GCAGC)             (   41)    41 ( 1172)   1213 (    3)   1216
                             (  206)  1422 (  694)   2116
BBVIB    (GCTGC)             (  254)   254 (   73)    327 (  303)    630
                             (   81)   711 (   18)    729 (  419)   1148
                             (  602)  1750
BGLI     (GCCNNNNNGGC)       (  245)   245 ( 1568)   1813
DDEIA    (CTAAG)             ( 2622)  2622
DDEIB    (CTCAG)             ( 1081)  1081 (  409)   1490 (  166)   1656
DDEIC    (CTGAG)             (  171)   171 ( 2025)   2196
ECORI    (GAATTC)            (  396)   396
ECORIIA  (CCAGG)             (  354)   354 (  479)    833 (  121)    954
ECORIIB  (CCTGG)             (  545)   545 (  422)    967
FNUDII   (CGCG)              (    2)     2 (    2)      4 (  103)    107
                             (  545)   652 (    2)    654 (  198)    852
                             (  581)  1433 (  330)   1763 (  493)   2256
                             (  332)  2588
FNU4HIA  (GCCGC)             (  850)   850 (  155)   1005 ( 1084)   2089
                             (  351)  2440
FNU4HIB  (GCGGC)             (  150)   150 (  582)    732 ( 1479)   2211
HAEIB    (AGGCCA)            (  819)   819 (   11)    830
HAEIA    (TGGCCT)            ( 1282)  1282
HAEIIA   (GGCGCC)            (  235)   235
HAEIIC   (GGCGCT)            (  680)   680 (  370)   1050
HAEIII   (GGCC)              (  287)   287 (  102)    389 (  257)    646
                             (  174)   820 (   11)    831 (   18)    849
                             (  434)  1283 (  458)   1741 (   80)   1821
                             (  267)  2088
HGAIA    (GACGC)             (  908)   908 (  578)   1486
HGAIB    (GCGTC)             (  108)   108 ( 2128)   2236
HGIAIA   (GTGCAC)            (  177)   177 (  943)   1120 ( 1246)   2366
HGIAIB   (GAGCTC)            (  402)   402
HGIAIC   (GTGCTC)            ( 2281)  2281
HHAI     (GCGC)              (    3)     3 (  103)    106 (  130)    236
                             (   21)   257 (  331)    588 (   65)    653
                             (   28)   681 (   33)    714 (  270)    984
                             (   67)  1051 (  100)   1151 (  174)   1325
                             (  109)  1434 (  393)   1827 (   93)   1920
                             (  337)  2257 (  332)   2589
HINCIIA  (GTCGAC)            (  429)   429
HINDIII  (AAGCTT)            (  447)   447
HINFIA   (GAATC)             (  641)   641 (  140)    781
HINFIB   (GACTC)             (  706)   706 (  988)   1694
HINFIC   (GAGTC)             (  427)   427 (  750)   1177
HPAII    (CCGG)              (   48)    48 (   34)     82 (  331)    413
```

```
ENZYME   SITE                          POS.              POS.             POS.
------------------------------------------------------------------------------
HPAII    (CCGG)              (  111)   524 (  489)   1013 (  147)   1160
                             (   26)  1186 (  190)   1376 (  404)   1780
                             (   34)  1814 (   67)   1881 (  110)   1991
                             (  242)  2233
HPHIIA   (GGTGA)             (   12)    12 (    9)     21 ( 2152)   2173
                             (  241)  2414
HPHIB    (TCACC)             ( 1550)  1550 (  227)   1777 (  622)   2399
KPNI     (GGTACC)            (  408)   408
MBOI     (GATC)              (  277)   277 (  141)    418 (  955)   1373
                             (   75)  1448 (   11)   1459 (    8)   1467
                             (   78)  1545 (   12)   1557 (  105)   1662
                             (  341)  2003 (   18)   2021 (   46)   2067
                             (  258)  2325 (   17)   2342 (   36)   2378
MBOIIA   (GAAGA)             ( 1456)  1456
MBOIIB   (TCTTC)             (  291)   291 (  394)    685 (  862)   1547
                             (  755)  2302 (   78)   2380 (  109)   2489
MNLIA    (CCTC)              (   29)    29 (  260)    289 (  132)    421
                             (  274)   695 (  283)    978 (  878)   1856
                             (  206)  2062
MNLIB    (GAGG)              (  662)   662 (  259)    921 (  324)   1245
                             (  400)  1645 (   81)   1726 (  947)   2673
MSTI     (TGCGCA)            (  256)   256 ( 1663)   1919
PSTI     (CTGCAG)            (  435)   435
PVUI     (CGATCG)            (  276)   276 ( 1790)   2066
PVUII    (CAGCTG)            (  306)   306 (  322)    628
RSAI     (GTAC)              (  168)   168 (  241)    409 ( 1769)   2178
SACI     (GAGCTC)            (  402)   402
SALI     (GTCGAC)            (  429)   429
SFANIA   (GATGC)             (   78)    78 (  130)    208 ( 1948)   2156
SFANIB   (GCATC)             (  153)   153 (   76)    229 (  665)    894
                             ( 1052)  1946 (  440)   2386
SMAI     (CCCGGG)            (  412)   412
SPHI     (GCATGC)            (  441)   441
TAQI     (TCGA)              (  400)   400 (   30)    430 (  476)    906
                             ( 1444)  2350
XBAI     (TCTAGA)            (  423)   423
XHOIIA   (GGATCC)            (  417)   417
XHOIIB   (AGATCC)            ( 1458)  1458 (   98)   1556 (  785)   2341
XHOIIC   (GGATCT)            ( 1447)  1447 (   97)   1544 (  780)   2324
NDEI     (CATATG)            (  183)   183
NARI     (GGCGCC)            (  235)   235
MSTI     (TGCGCA)            (  256)   256 ( 1663)   1919

 THESE ENZYMES DID NOT APPEAR.....ACCIB    ACCIC    ACCID    ACYIC    ASUII
                                  AVAIB    AVAIC    AVAID    AVAIIA   AVAIII
                                  AVRII    BALI     BCLI     BGLII    BSTEIIA
                                  BSTEIIB  BSTEIIC  BSTEIID  CLAI     DDEID
                                  HAEIA    HAEIC    HAEIIB   HAEIID   HGIAID
                                  HINCIIB  HINCIIC  HINCIID  HINFID   HPAI
                                  SACII    XHOI     XHOIID   XMAIII   NCOI
```

Fig. 2. The nucleotide sequence and restriction sites of pUC19. (A) The sequence has been compiled as described and entered into an Apple II computer. The information derived from sequencing pUC6 (Fig. 4) was used to make corrections via the programs described earlier (Larson and Messing, 1983). Both the *bla* and the *lac* α-peptide gene products are read from the same strand. Since all coordinates of the pBR322 sequence refer to the opposite strand, the single-stranded DNA with the polarity published by Sutcliffe (1979) is given. The message strand is obtained with the program that produces the reverse complement of pUC19 (referred to as pUC19V). (B) Restriction site coordinates are as in Fig. 1B.

M13 origin of replication (Pohlman, R. and Messing, J., unpubl.). It should be noted that the Ac element contains numerous direct and indirect repeats in regions of low and high G + C contents (Pohlman et al., 1984). When the 4.5-kb *Pst*I-*Bam*HI Ac fragment is cloned into M13 and grown in JM109, stable recombinants have been recovered and no deletions have been detected over numerous generations (not shown).

(5) Methylation

Other *E. coli* mutations useful for recombinant DNA amplification are DNA methylation deficiencies. For example, *Mbo*I and *Bcl*I can cleave DNA propagated in a *dam*$^-$ *E. coli* strain while *Eco*RII restriction requires the absence of the *dcm* product (for review, see Roberts, 1983). If DNA is propagated in *E. coli* strains lacking the A and C methylases, it is unmodified and can be cleaved by *Mbo*I, *Bcl*I, and *Eco*RII. GM48 contains these mutations, but lacks the *Δ(lac-proAB)* deletion and the F′ *traD*36 *proAB lacI*^{q}*ZΔ*M15 episome required for M13mp and pUC vector use. Strain JM110 was developed to be *dam*$^-$ *dcm*$^-$ by introduction of the *Δ(lac-proAB)* deletion and JM101 episome into GM48. The *dam* and *dcm* mutations have been tested as described in MATERIALS AND METHODS, section **j** (not shown).

(b) M13 phage vectors

Since many specific constructions depend on the knowledge of the vector restriction map, the nucleotide sequences of M13mp18 and pUC19 have been compiled and reprinted here. The wild-type sequence of M13 has been determined by Van Wezenbeek et al. (1980). M13's unique *Hin*cII site was used as the sequence reference point. Following this reference, the C in position 3 has been converted to a T to eliminate the *Hin*cII site, the G (2220) converted to an A to eliminate the *Bam*HI site, and the C (6917) converted to a T to eliminate the *Acc*I site and introduce the *Bgl*II site (positions are for wild-type M13) (Messing et al., 1981). The *lac Hin*dII fragment has been inserted into the *Bsu*I site at position 5868 (Messing et al., 1977). The *lac* sequence has been compiled from the *lacI* gene (Farabaugh, 1978), the *lacZpo* region (Dickson et al., 1975), and the *lacZ* gene sequences (Kalnins et al., 1983). M13mp18 and M13mp19 lack the double amber mutation that M13mp10a and M13mp11a had (Messing and Vieira, 1982) but do possess two complementary polylinker regions in the *lacZ* gene (Norrander et al., 1983). The junctions of *lac* DNA and M13 DNA were as predicted from the restriction sites used for cloning the *lac Hin*dII fragment into the *Bsu*I site at position 5868. The junction sequence at the *lacZ* gene led to an early ochre termination codon that produced a *lac* α fragment of 168 amino acid residues; 18 of them represent the polylinker region. The resulting sequence of M13mp18 is presented in Fig. 1.

(c) The pUC plasmids

(1) Construction

The *lac Hae*II fragment inserted into pBR322 produced a shorter α peptide than that in M13 (Ruther, 1980), yet active. The pBR322 sequence has been modified by removing the *Eco*RI-*Pvu*II fragment containing the Tc resistance gene via a fill-in reaction and blunt-end ligation. The predicted regeneration of the *Eco*RI site failed to occur when sequencing data revealed that the deletion extends across nucleotides 4355-1-2072. Similar findings were reported by Rubin and Spradling (1983) and Stragier, P. (personal communication). Restriction sites were removed from the intermediate plasmid in the following way. EMS mutagenesis resulted in a GC to AT transition in the *Pst*I site at position 3610 (positions are for pBR322) (Vieira and Messing, 1982). Hydroxylamine treatment converted another GC to AT in the *Hin*cII site at position 3911. The *Acc*I site was eliminated by BAL31 digestion of nucleotides 2210–2250. The *lac* sequence was inserted at the *Hae*II site at position 2352. The *lac* sequence was oriented in the same direction as the *bla* gene coding for β-lactamase. The fusion of the *lacZ* sequence to the pBR322 DNA at the *Hae*II site at position 2352 resulted in an α peptide of 107 amino acid residues, 19 encoded by the polylinker region at residue 5. Translation was terminated by the UAG termination codon, which was suppressed in the *supE* strains. In *supE* strains the peptide was 15 amino acids longer and terminated by an UGA codon. These small fusion peptides were very unstable in *E. coli* and detectable only through the highly sensitive complementation test with Xgal. The nucleotide sequence and restriction map for pUC19 are given in Fig. 2.

(2) Sequencing with new method

Since the pUC plasmids have been mutagenized with EMS, HA, and treated with BAL31 before introducing the *lac* DNA into the pBR322 backbone molecule, the pBR322 portion of pUC6 (Vieira and Messing, 1982) had to be resequenced. Earlier sequencing experiments were based on M13 shotgun sequencing of pUC6 (Halling, S., Abbot, A., Kridl, J. and Messing, J., unpubl.). Because the asymmetric M13 polylinkers (Vieira and Messing, 1982) could be used to make unidirectional deletions, a different approach of generating pUC6 subclones for sequencing was tested. The polylinker permitted cleavage of the pUC plasmid or the M13 RF by two restriction endonucleases, one producing a 3′ protruding end like *Pst*I, the other a 5′ protruding end. Since Exo III was double-strand-specific and required a 3′OH end, the *Pst*I end was not accessible to this enzyme. These features simplified the nonrandom sequencing approach based on *BAL*31 treatment described below (Poncz et al., 1982) and illustrated in Fig. 3. The method included the following steps: (1a) pUC6 was linearized with *Nde*I, and the ends were made flush with the large fragment of DNA polymerase and inserted at the *Hin*cII site of M13mp19 to make M13UC. This produced, between the inserted DNA and the PHS, a unique *Sst*I site proximal to the PHS and a unique *Bam*HI site distal to it. (1b) M13UC RF was digested with *Sst*I and *Bam*HI. *Sst*I leaves a 4-bp 3′ overhang that is resistant to Exo III and protects the PHS from digestion. *Bam*HI leaves a recessed 3′ end that is sensitive to Exo III and exposes the insert to digestion. (2) The DNA was treated with Exo III for 0–20 min. Aliquots were removed at 1-min intervals. This time course generated random insert deletions while leaving the PHS intact. (3) The single-stranded region of DNA left by Exo III treatment was removed with Exo VII. Since only Exo VII is active in the presence of EDTA, addition of Exo III time-course aliquots to a tube containing Exo VII buffer plus EDTA proves a convenient way to stop the Exo III reaction. (4) To ensure formation of blunt-ended DNA, the digest was treated with DNA polymerase I. DNA ligase was added to recircularize the molecules, which were then used to transform JM105. (5) Phage isolated from transformed cells were used for direct gel electrophoresis (Messing, 1983) to determine clones of appropriate size for sequencing (Fig. 4).

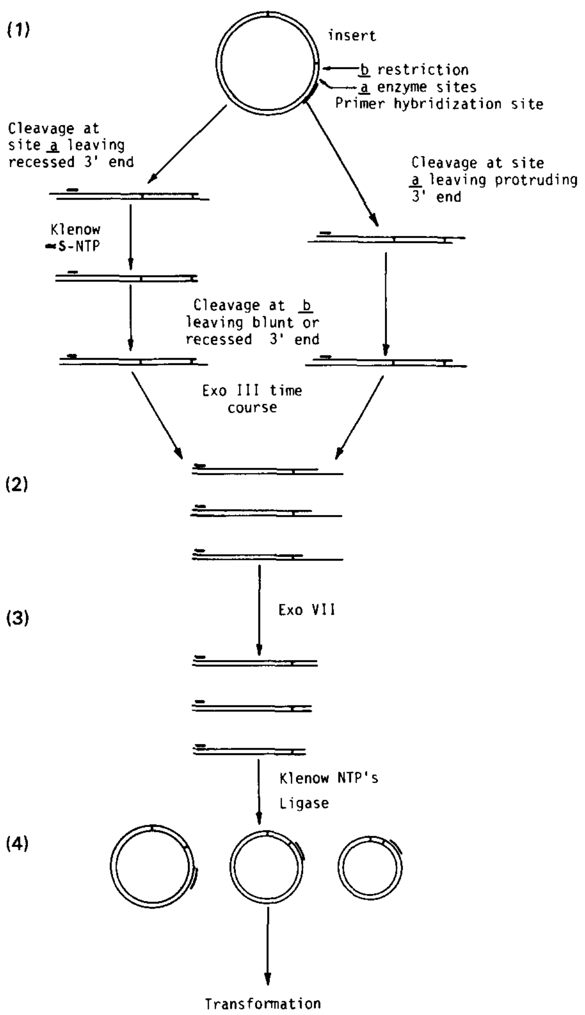
(1)
insert
b restriction
a enzyme sites
Primer hybridization site
Cleavage at site a leaving recessed 3' end
Cleavage at site a leaving protruding 3' end
Klenow
•S-NTP
Cleavage at b leaving blunt or recessed 3' end
Exo III time course
(2)
(3)
Exo VII
Klenow NTP's
Ligase
(4)
Transformation

Fig. 3. Method for generating unidirectional deletions. Details of the procedure are as follows: (1) Vector is digested with restriction enzymes a and b. Enzyme a leaves a 4-bp 3′ overhang that is resistant to Exo III and protects the PHS site. Enzyme b leaves a recessed 3′ end that is sensitive to Exo III and exposes the insert to digestion. (2) The DNA is treated with Exo III for 0–20 min and aliquots removed at 1 min intervals. This generated random insert deletions while leaving the PHS intact. (3) Exo VII removed the single-stranded DNA region left by Exo III. (4) The digest was treated with DNA polymerase I to ensure formation of blunt-ended DNA. DNA ligase was added to recircularize the various deletion products, leading to increasingly smaller circles with the PHS in the same position.

The Exo III, Exo VII, polymerase, and ligase reactions were performed sequentially by adjusting reaction buffers. Alternatively, it was possible to protect the PHS from Exo III digestion via S-NTP incorporation by DNA polymerase at a recessed 3′ end proximal to the PHS left by restriction endo-

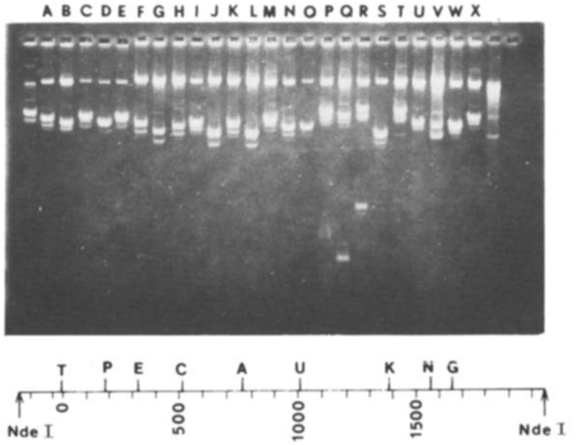

Fig. 4. Mapping of the deletion mutants. Since the position of the PHS is unaltered and all deletions occur only at the opposite end, deletion points are mapped by recombinant phage mobility changes indicated by agarose gel electrophoresis. After exonuclease and ligase treatment, the DNA is transformed into JM105. Plaques are picked from each transformation experiment, grown in small cultures, and supernatant phage used directly for agarose gel electrophoresis as described (Messing, 1983). A picture taken of the agarose gel was used to draw a physical map of the sequenced clones. The first and last lanes (unmarked) contain untreated M13UC1 and M13mp19, respectively; the other lanes are labeled alphabetically and represent individual clones. Nine clones from this gel were used to prepare a template for sequencing as described in MATERIALS AND METHODS, section **o**. The sequence has been entered into an Apple II computer and analyzed using the programs of Larson and Messing (1983). The deletion points are marked in the map by the agarose-gel-derived clone name. The map has been drawn with reference to the *Nde*I sites used to clone pUC6 into the *Hin*cII site of M13mp19. The nucleotide numbers in the map are taken from the reverse complement of pUC6, referred to as pUC6V.

nuclease digestion (Putney et al., 1981; Vieira, J., unpublished results). Cleavage of a site distal to the PHS was then needed to generate an unprotected recessed 3′ end for Exo III treatment. The consecutive steps are outlined in Fig. 3.

This approach resembles that described by Poncz et al. (1982), but hastens the construction of recombinant M13 phage needed for sequencing. As opposed to bidirectional deletions, the creation of unidirectional deletions precludes the need for recloning DNA fragments. The speed by which recombinants can be obtained resembles that of shotgun cloning. Ordering clones on a physical map is simple (Fig. 4), so the redundancies and gaps typical of shotgun sequencing are avoided. Hence, the following sequencing approach to larger DNA segments is used. Restriction sites present in the polycloning sites of M13mp18 and M13mp19 are used to clone restriction fragments in both orientations. Fragments need to be inserted such that between the PHS and the insert there exist two unique restriction enzyme sites. The restriction enzyme site proximal to the PHS must produce either a 4 bp 3′ overhang or a recessed 3′ end. The other should leave a blunt or recessed 3′ end next to the insert. Each clone pair representing both orientations is then subjected to Exo III and VII treatment. The optimum insert size for nuclease treatment is 2000–5000 bp. Also, the ExoIII unit activity of different commercial preparations can vary, necessitating the calibration of each enzyme lot. This is accomplished by the electrophoresis of DNA samples taken at two time points from an ExoIII reaction on an agarose gel for size analysis.

ACKNOWLEDGEMENTS

We thank Kris Kohn, Tammy Krogmann, Mike Zarowitz, and Stephanie Young for help in preparing the manuscript, Drs. J. Felton, T.Baldwin and P. Stragier for making results available to us prior to publication, and Betsy Kren and James Fuchs for helpful discussions. The work was supported by the Department of Energy DE-FG02-84ER13210 and NIH GM31499. J.V. is supported by the NIH training grant No. ST32 GM107094.

REFERENCES

Achtman, M., Willets, H. and Clark, A.J.: Beginning a genetic analysis of conjugational transfer determined by the F-factor in *Escherichia coli* by isolation and characterization of transfer-deficient mutants. J. Bacteriol. 106 (1971) 529–538.

Bedbrook, J.R. and Ausubel, F.M.: Recombination between bacterial plasmids leading to the formation of plasmid multimers. Cell 9 (1976) 707–716.

Birnboim, H.L. and Doly, J.: A rapid alkaline extraction procedure for screening of recombinant plasmid DNA. Nucl. Acids Res. 7 (1979) 1513–1523.

Carlson, J. and Messing, J.: Efficiency in cloning and sequencing using the single-stranded bacteriophage M13. J. Biotechnol. (1984) in press.

Cohen, S.N., Chang, A.C.Y. and Hsu, L.: Non-chromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R factor DNA. Proc. Natl. Acad. Sci. USA 69 (1972) 2110–2114.

Csonka, L. and Clark, A.: Construction of an Hfr strain useful for transferring *recA* mutations between *Escherichia coli* strains. J. Bacteriol. 143 (1980) 529–530.

Dickson, R.C., Abelson, J., Barnes, W.M. and Reznikoff, W.S.: Genetic regulation: the *lac* control region. Science 187 (1975) 27–35.

Federal Register Vol. 45, No. 147 (1980) 50525.

Felton, J.: M13 host strain JM103 contains two restriction systems. Biotechniques I (1983) 42–43.

Farabaugh, P.J.: Sequence of the *lacI* gene. Nature 274 (1978) 765–769.

Hanahan, D.: Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. 166 (1983) 557–580.

Helfman, D.M., Feramisco, J.R., Fiddes, J.C., Thomas, G.P. and Highes, S.H.: Identification of clones that encode chicken tropomyosin by direct immunological screening of a cDNA expression library. Proc. Natl. Acad. Sci. USA 80 (1983) 31–35.

Heidecker, G., Messing, J. and Gronenborn, B.: A versatile primer for DNA sequencing in the M13mp2 cloning system. Gene 10 (1980) 69–73.

Heidecker, G. and Messing, J.: Sequence analysis of zein cDNAs obtained by an efficient mRNA cloning method. Nucl. Acids Res. 11 (1983) 4891–4906.

Hu, N-T. and Messing, J.: The making of strand-specific M13 probes. Gene 17 (1982) 271–277.

Kalnins, A., Otto, K., Ruther, U. and Müller-Hill, B.: Sequence of the *lacZ* gene of *Escherichia coli*. EMBO J. 2 (1983) 593–597.

Larson, R. and Messing, J.: Apple II computer software for DNA and protein sequence data. DNA 2 (1983) 31–35.

Maloy, S. and Nunn, W.: Selection for loss of tetracycline resistance by *Escherichia coli*. J. Bacteriol. 145 (1981) 1110–1112.

Messing, J., Gronenborn, B., Müller-Hill, B. and Hofschneider, P.H.: Filamentous coliphage M13 as a cloning vehicle: insertion of a *Hin*dII fragment of the *lac* regulatory region in the M13 replicative form in vitro. Proc. Natl. Acad. Sci. USA 75 (1977) 3642–3646.

Messing, J.: A multipurpose cloning system based on the single-stranded DNA bacteriophage M13. Recombinant DNA Technical Bulletin, NIH Publication No. 79-99, Vol. 2, No. 2 (1979) 43–48.

Messing, J.: M13mp2 and derivatives: a molecular cloning system for DNA sequencing, strand specific hybridization, and in vitro mutagenesis, in Walton, A.G. (Ed.), Proceedings of the Third Cleveland Symposium on Macromolecules. Elsevier, Amsterdam, 1981, pp. 143–153.

Messing, J., Crea, R., and Seeburg, P.H.: A system for shotgun DNA sequencing. Nucl. Acids Res. 9 (1981) 309–321.

Messing, J. and Vieira, J.: A new pair of M13 vectors for selecting either strand of a double-digest restriction fragment. Gene 19 (1982) 269–276.

Messing, J.: New M13 vectors for cloning. Methods Enzymol. 101 (1983) 20–79.

Miller, J.H.: Experiments in Molecular Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1972.

Norrander, J., Kempe, T. and Messing, J.: Improved M13 vectors using oligonucleotide-directed mutagenesis. Gene 26 (1983) 101–106.

Pohlman, R., Fedoroff, N. and Messing, J.: The nucleotide sequence of the maize controlling element activator. Cell 37 (1984) 635–643.

Poncz, M., Solowiejczyk, D., Ballantine, M., Schwartz, E. and Surrey, S.: "Non-random" DNA sequence analysis in bacteriophage M13 by the dideoxy chain-termination method. Proc. Natl. Acad. Sci. USA 79 (1982) 4298–4302.

Putney, S.D., Benkovic, S.J. and Schimmel, P.R.: A DNA fragment with an alpha-phosphorothioate nucleotide at one end is asymmetrically blocked from digestion by exonuclease III and can be replicated in vivo. Proc. Natl. Acad. Sci. USA 78 (1981) 7350–7354.

Roberts, R.: Restriction and modification enzymes and their recognition sequences. Nucl. Acids Res. 12 (1984) r167–204.

Rubin, G.M. and Spradling, A.C.: Vectors for P element-mediated gene transfer in *Drosophila*. Nucl. Acids Res. 11 (1983) 6341–6348.

Ruther, U.: Construction and properties of a new cloning vehicle, allowing direct screening for recombinant plasmids. Mol. Gen. Genet. 178 (1980) 475–477.

Sanger, F., Nicklen, S., and Coulson, A.R.: DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA 74 (1977) 5463–5467.

Sutcliffe, J.G.: Complete nucleotide sequence of the *Escherichia coli* plasmid pBR322. Cold Spring Harbor Symp. Quant. Biol. 43 (1979) 77–90.

Van Wezenbeek, P.M.G.F., Hulsebos, T.J.M. and Schoenmakers, J.G.G.: Nucleotide sequence of filamentous bacteriophage M13 DNA genome: comparison with phage fd. Gene 11 (1980) 129–148.

Vieira, J. and Messing, J.: The pUC plasmids, and M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene 19 (1982) 259–268.

Communicated by A.J. Podhajska.