INFECTION AND IMMUNITY, May 1989, p. 1391–1398
0019-9567/89/051391-08$02.00/0

Vol. 57, No. 5

Cloning of Regions Required for Contact Hemolysis and Entry into LLC-MK2 Cells from *Shigella sonnei* Form I Plasmid: *virF* Is a Positive Regulator Gene for These Phenotypes

JUN-ICHI KATO, KEN-ICHIRO ITO, AKIKO NAKAMURA, AND HARUO WATANABE*

Department of Bacteriology, National Institute of Health, Kamiosaki, Shinagawa-ku, Tokyo 141, Japan

Received 21 November 1988/Accepted 17 January 1989

Two distinct regions required for both contact hemolysis and entry into LLC-MK2 cells were cloned into *Escherichia coli* from the *Shigella sonnei* form I plasmid, pSS120. The first region was cloned into an *E. coli* HB101 strain containing noninvasive Tn*1* insertion mutants of the form I plasmid, and expression of *ipa* (invasion plasmid antigen) gene products was restored. The plasmid carrying the first region was then transformed into *E. coli* lacking the form I plasmid, and additional DNA fragments from the form I plasmid were cloned into the same recipient on compatible vectors. Five of these double transformants were found to be positive for contact hemolysis activity. Restriction analysis of these five clones indicated that the previously reported *ipa* locus and the *invA* locus were present on the second plasmid region. Only the strains carrying both of these regions were active in contact hemolysis and cell invasion assays. Several proteins, including the a, b, c, and d proteins encoded by the *ipa* genes, were detected in the double transformants by Western blot (immunoblot) analysis with serum of a monkey convalescing from shigellosis. A positive regulator was suggested to exist in the first region, since the amounts of most of these proteins were simultaneously increased in the presence of this region. Subcloning and nucleotide sequencing indicated that this positive regulator gene was *virF*. Product analysis of the *virF* gene with minicells showed that two peptides (30 and 21 kilodaltons) were synthesized and that at least the 30-kilodalton protein was essential for regulation of the *ipa* genes.

Invasion of the human colonic mucosa by bacterial cells is a critical step in the pathogenesis of bacillary dysentery caused by *Shigella* spp. or enteroinvasive *Escherichia coli* (15). This invasive process is followed by intracellular multiplication, spreading to adjacent cells, and eventual cell killing; these events lead to inflammatory lesions within the lamina propria of intestinal villi (7, 15). The genus *Shigella* and enteroinvasive *E. coli* contain large plasmids that are associated with virulence and have some functions in common (26, 27, 32). These plasmids, which range in size from 120 to 140 megadaltons (MDa), carry genes for bacterial entry into epithelial cells (27), subsequent lysis of the phagocytic vacuole, multiplication within cells (28), and early killing of cells (6). In addition to these functions, the large plasmid in *Shigella sonnei* (pSS120), called the form I plasmid, encodes production of form I O antigen, which is essential for virulence (14). To understand the mechanism by which *Shigella* and enteroinvasive *E. coli* cells invade epithelial cells, it is necessary to analyze the genes on the large virulence-associated plasmids and investigate the functions of their products.

There are two approaches to analyzing these genes: analyses of transposon insertion mutants (1, 17, 29, 33) and cloning of the essential regions. Many transposon insertion mutants have been isolated, and the regions containing the genes for invasion have been identified. These analyses suggest that a number of genes in at least two distinct regions are required (29, 33). Although analyses of transposon insertion mutants identified the essential genes, including *invA* (34) and *virF* (20, 21, 29), it remained uncertain whether the regions or genes identified are sufficient for the invasive process. Molecular cloning of these genes and reconstitution of the invasive process are essential to resolve this problem. Maurelli et al. isolated the clones which restored the invasion of HeLa cells in plasmidless *Shigella* recipients from the virulence plasmid of *Shigella flexneri* serotype 5 (17). However, two points remained to be resolved. First, the *Shigella* strains with the clones did not express a full pattern of invasion: only half of the cells in the monolayer were invaded, and the cells were not invaded as extensively as were those infected by the strain with the large virulence-associated plasmid (17). Second, there was an inconsistency between the fact that these clones were not thought to contain the *virF* gene, as judged from the restriction map (17, 29), and the observation by Sasakawa et al. that the *virF* gene was essential for invasion (29).

Some invasion-negative Tn*1* insertion mutants of the form I plasmid in *S. sonnei* were isolated, and a gene, *invA*, coding for a 39-kDa protein, was identified from analyses of these mutants (33). In this work, we cloned in *E. coli* K-12 two regions required for contact hemolysis and cell invasion. One of these regions (the *invA* region) was found to contain a number of essential genes (*inv* and *ipa* genes). The other was suggested to include a positive regulator for the *inv* genes; nucleotide sequence analysis demonstrated that this positive regulator gene was identical to the *virF* gene in *S. flexneri*. Cloning of the two essential regions and reconstitution of the invasive processes of these regions has resolved the inconsistency described above.

MATERIALS AND METHODS

Bacterial strains, plasmids, and culture media. The *S. sonnei* strain used was HW383 (33). The *E. coli* strains used were HB101 (2) (F$^-$ *hsdS20 recA13 ara proA lacY galK rpsL xyl mtl supE*), DH1 (16) (F$^-$ *hsdR17 recA1 endA gyrA thi supE*), TH1219 (11) (F$^-$ *minB recA1 rpsL tsr tar*), and JM109 (35) [Δ(*lac-pro*) *hsdR17 recA1 endA gyrA thi supE relA* (F′ *traD proA lac*^{q}*Z M15*)].

* Corresponding author.

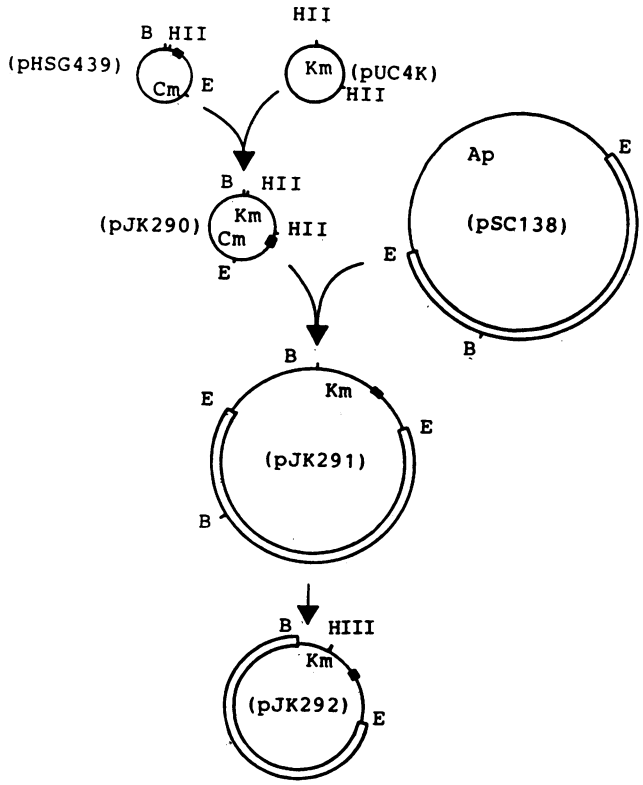

FIG. 1. Cosmid vectors derived from a mini-F plasmid. A mini-F cosmid vector, pJK292, was constructed via pJK290 and pJK291 from pSC138 (30), a derivative of the mini-F plasmid. The *Eco*RI-*Eco*RI and *Bam*HI-*Eco*RI fragments (open arcs) contain all regions essential for replication and stable partition of a mini-F replicon. Boxes represent *cos* sites. B, *Bam*HI; E, *Eco*RI; HII, *Hin*cII; HIII, *Hin*dIII.

A cosmid vector, pHSG439 (pSC101 replicon, Cm^r) (3), and a low-copy-number plasmid vector, pHSG595 (pSC101 replicon, Cm^r) (T. Hashimoto-Gotoh, personal communication), were kindly provided by T. Hashimoto-Gotoh.

Bacteria were routinely grown in broth and on 1.5% agar-solidified plates of antibiotic medium 3 (Difco Laboratories, Detroit, Mich.). Antibiotics and concentrations were as follows: kanamycin, 50 μg/ml; ampicillin, 50 μg/ml; tetracycline, 10 μg/ml; and chloramphenicol, 15 μg/ml.

Preparation and manipulation of DNA. Techniques for preparation of plasmid DNA (except for the large virulence-associated plasmid [pSS120] DNA) and manipulation of DNA were essentially those described by Maniatis et al. (16). pSS120 plasmid DNA was isolated by methods described previously (33).

Plasmid construction. The low-copy-number vector pJK292, derived from a mini-F plasmid, was constructed as shown in Fig. 1. To construct pJK290, the Km^r marker of pUC4K (31) was excised as a *Hin*cII fragment and inserted into the *Hin*cII site of pHSG439. The *Eco*RI fragment containing the essential region for replication and stable maintenance of the mini-F plasmid, pSC138 (30), was inserted into the *Eco*RI site of pJK290 to obtain pJK291; pJK292 was constructed by deletion of a *Bam*HI fragment from pJK291.

The method used for cosmid cloning was essentially that described by Maniatis et al. (16). The plasmid DNA preparation from *S. sonnei* HW383 harboring pSS120 and cryptic small plasmids was partially restricted with *Sau*3AI. The DNA fraction from a 10 to 40% sucrose gradient, which contained fragments of about 30 to 50 kilobases (kb) as well as fairly large amounts of contaminated, small-sized digests of the cryptic plasmids, was used for cosmid cloning. The *Eco*RI (dephosphorylated)-*Bam*HI and *Hin*cII (dephosphorylated)-*Bam*HI fragments of pHSG439 and the *Eco*RI (dephosphorylated)-*Bam*HI and *Hin*dIII (dephosphorylated)-*Bam*HI fragments of pJK292 were purified and used as the arms of cosmid vectors. The ligated DNA was packaged into the bacteriophage λ head by using a packaging mix (Gigapack; Stratagene, San Diego, Calif.). Recombinant clones were isolated by the selective markers of the cosmid vectors: Km^r for pJK292 and Cm^r for pHSG439.

pJK1143 was constructed by inserting the 12-kb *Sal*I fragment of pJK1115 (see Results) into the *Sal*I site of pHSG595, using JM109 as the host strain. pJK1144 and pJK1145 were constructed by inserting the 2-kb *Sma*I fragment of pJK1143 into the *Sma*I site of pUC18 (35); the plasmids carried this *Sma*I fragment in opposite orientations. The deletion derivatives of pJK1144 and pJK1145 were obtained by using a deletion kit (Takara Shuzo Corp., Kyoto, Japan) essentially as described by Henikoff (12) and Yanisch-Perron et al. (35). Each 7.5 μg of pJK1144 and pJK1145 plasmid DNA was digested with both *Pst*I and *Bam*HI, treated with 180 U of exonuclease III (37°C for 1 to 10 min), 50 U of mung bean nuclease (37°C for 60 min), and 2 U of Klenow fragment (37°C for 15 min), and ligated. The ligated DNA was transformed after digestion with *Bam*HI.

Virulence assays. The assay for contact hemolysis was performed as described elsewhere (28) except that (i) bacteria grown in antibiotic medium 3 were harvested in the early exponential phase and (ii) close contact between sheep erythrocytes and bacteria was achieved by centrifugation at 700 × *g* in a 1.5-ml sampling tube at 16°C for 15 min.

The tissue culture assay for invasion was carried out with LLC-MK2 cells as described previously (34). The assay for bacterial growth in LLC-MK2 cells was performed essentially as described elsewhere (28). Experiments were carried out in duplicate. Twelve plates inoculated with semiconfluent monolayers of LLC-MK2 cells were simultaneously infected with the strains tested and then centrifuged. After incubation at 37°C for 15 min, the plates were washed three times with Earle balanced salt solution containing 200 μg of gentamicin per ml, covered with Eagle minimal essential medium containing 200 μg of gentamicin per ml, and incubated at 37°C. After 0, 1, 3, 4, 6, and 8 h, the two plates were removed and washed once with Earle balanced salt solution containing gentamicin (200 μg/ml) and once with the same solution but without gentamicin. One plate was used for Giemsa staining to calculate the percentage of infected LLC-MK2 cells. Cells of the other plate were trypsinized, harvested, and lysed with 0.1% Triton X-100 at 37°C for 5 min. Dilutions of this suspension were plated onto antibiotic medium 3 agar, and colonies were counted after overnight growth. The average number of bacteria per infected LLC-MK2 cell was calculated as follows: (number of bacteria per plate)/(number of LLC-MK2 cells × percentage of infected LLC-MK2 cells × 0.01). Values reported are means of four experiments.

Western blot (immunoblot) analysis. Immunoblots of whole bacterial extracts run on sodium dodecyl sulfate-polyacrylamide gel were performed as previously described (4) except that the polyvinylidene difluoride filter (Immobilon transfer membrane; Millipore Corp., Bedford, Mass.) was used instead of nitrocellulose filters and horseradish peroxidase-

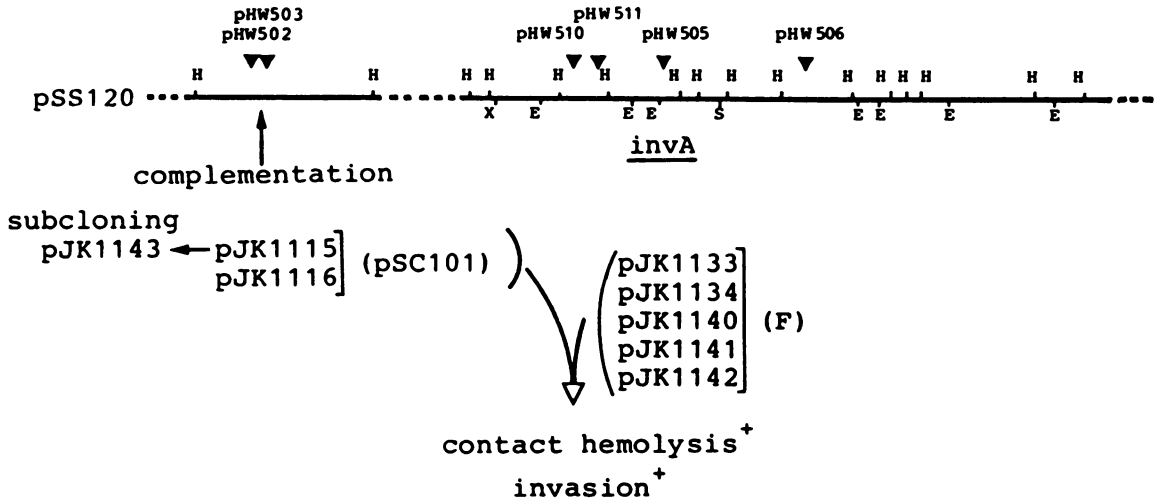

FIG. 2. Cloning strategy and restriction map of the two essential regions in pSS120. pHW502, pHW503, pHW505, pHW506, pHW510, and pHW511, described previously (33), are the noninvasive Tn*1* insertion mutants of pSS120; arrowheads indicate positions of transposon insertions. pJK1133, pJK1134, pJK1140, pJK1141, and pJK1142 carry the *invA* region, which includes all insertion sites of pHW505, pHW506, pHW510, and pHW511. In the diagram of the strategy for cloning the two regions and subcloning from pJK1115 to pJK1143, the replicons of the recombinant clones are shown in parentheses. E, *Eco*RI; HIII, *Hin*dIII; S, *Sal*I; X, *Xho*I.

conjugated anti-monkey immunoglobulin G was used instead of radioiodinated protein A. Diluted (1/1,000) serum from a monkey orally challenged with virulent *S. flexneri* 1b (about 10^{12} cells) was used to identify virulence plasmid-associated proteins.

Nucleotide sequence determination. DNA sequences of the deletion derivatives of pJK1144 and pJK1145 were determined by the dideoxy technique (23), using double-stranded DNA as the template according to the instructions of the manufacturer of the sequencing kit (Takara Shuzo).

Analysis of gene products. Minicells were prepared from TH1219 carrying appropriate plasmids and labeled with [^{35}S]methionine (1,200 Ci/mmol; Dupont, NEN Research Products, Boston, Mass.) at 37°C for 30 min unless otherwise specified. Preparation and labeling of minicells and analysis of the labeled gene products by sodium dodecyl sulfate-polyacrylamide gel electrophoresis were performed as described previously (13). Fluorography was carried out with ENLIGHTNING (Dupont, NEN).

RESULTS

Cloning of regions required for contact hemolysis. In cloning the regions essential for contact hemolysis and cell invasion, the tactic we used was to construct libraries with compatible cosmid vectors and combine cosmid clones, since analyses of noninvasive Tn*1* insertion mutants of the form I plasmid suggested that the loci for invasion genes appeared to consist of at least two distinct regions (Fig. 2; 33). Two DNA libraries of the form I plasmid were prepared by using the cosmid vectors pHSG439 (pSC101 replicon) and pJK292 (F replicon), constructed as shown in Fig. 1. We first tried to isolate the plasmids that could complement one group of noninvasive Tn*1* insertion mutants of the form I plasmid, pHW502 and pHW503 (Fig. 2). HW805 [HB101 (pHW503)] and HW811 [HB101(pHW502)] were transformed with plasmids constructed with the cosmid vector pHSG439 (pSC101 replicon). Restoration of production of the antigens reactive with convalescent monkey serum was investigated, and 2 positive strains out of 50 were obtained. The cosmid clones from the positive strains, pJK1115 and pJK1116 (Fig. 2), were transformed into HB101. Since the HB101 strains carrying pJK1115 or pJK1116 alone were inactive in antigen production and contact hemolysis, those strains were transformed with plasmids constructed with cosmid vector pJK292 (F replicon). Contact hemolysis activities of the double transformants were then assayed. Among 2,900 transformants, 1 with high activity (HB101 carrying pJK1116 and pJK292 derivative pJK1134) and 4 with low activity (HB101 carrying pJK1115 and pJK292 derivative pJK1133, pJK1140, pJK1141, or pJK1142) were isolated.

Restriction maps of plasmids pJK1133, pJK1134, pJK1140, pJK1141, and pJK1142 (Fig. 3) indicated that these plasmids shared a common DNA region of about 30 kb,

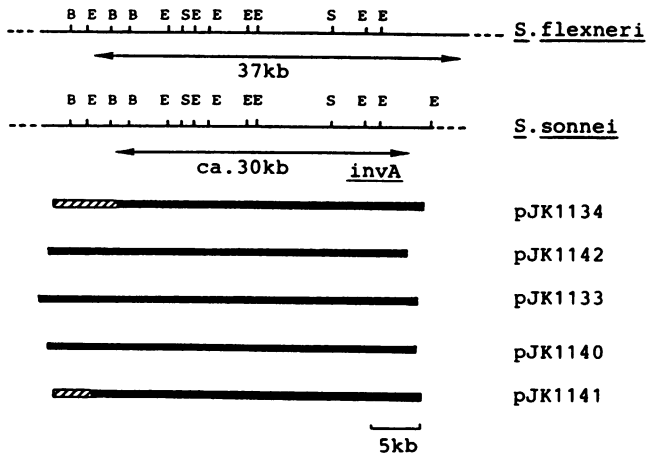

FIG. 3. Restriction maps of plasmids carrying the *invA* region. The region containing the *invA* gene of the virulence-associated plasmid (pSS120) of *S. sonnei* is shown along with the corresponding region of *S. flexneri*. Symbols: ▬, region derived from the form I plasmid; ▨, DNA sequence derived from the cryptic small plasmid of *S. sonnei* (in pJK1134) and an unknown fragment (in pJK1141). The region indicated as 37 kb is the essential region reported by Maurelli et al. (17). B, *Bam*HI; E, *Eco*RI; S, *Sal*I.

including the region of the *invA* gene identified previously (33). The presence of this gene was confirmed by Southern blot analysis, using the plasmid containing the *invA* gene (33) as a probe. Plasmids pJK1134 and pJK1141 contained extra fragments (Fig. 3), and pJK1134 hybridized with the *S. sonnei* cryptic small plasmid (data not shown). The restriction map of the *invA* region was almost the same as that of the region cloned by Maurelli et al. (Fig. 3; 17), and it was deduced that the essential region encompassed about 30 kb of the 37-kb essential region reported by Maurelli et al. (17).

*Sal*I digestion of pJK1115 and pJK1116 produced several common fragments. The plasmid (pJK1143; Fig. 2) carrying one of them, a 12-kb *Sal*I fragment, onto the vector pHSG595 (pSC101 replicon) was found to be equivalent to pJK1115 with respect to contact hemolysis and cell invasion.

Contact hemolysis and cell invasion. When the *invA* region was cloned with pJK292 (F replicon) as described above, strains with different degrees of contact hemolysis activity were obtained: those harboring pJK1134 had higher activities than did those harboring pJK1133, pJK1140, pJK1141, and pJK1142. In addition, entry into LLC-MK2 cells almost paralleled contact hemolysis activity (data not shown). In the presence of pJK1143, the contact hemolysis activity of pJK1134 at 37°C, expressed as [optical density at 550 nm (hemolysis)/optical density at 630 nm (bacterial cell mass)] × 1,000, was about 1.5-fold that of pHW1273 (935 ± 128 versus 690 ± 109), and the activity of pJK1142 was about one-fifth that of pHW1273 (138 ± 35 versus 690 ± 109). Activities of pJK1133, pJK1140, and pJK1141 were about the same as that of pJK1142. In the absence of pJK1143, however, activities of strains with each of these clones alone were almost the same as that of the host strain with no plasmid. Cells lacking one or both of these plasmids very frequently appeared even if the cells were grown under selective pressure in the presence of antibiotics. Percentages of cells carrying both plasmids in the contact hemolysis assay (i.e., Km^r Cm^r strains) were as follows: pJK1134, 72% at 30°C and 5% at 37°C; and pJK1142, 61% at 30°C and 41% at 37°C.

The fact that pJK1134 had higher activity than did the other pJK292 derivatives (pJK1133, pJK1140, pJK1141, and pJK1142) could not be explained by differences in the cloned regions as shown in Fig. 3. Possibly the copy number of pJK1134 was much higher (about 20-fold) than that of pJK1133, pJK1140, pJK1141, and pJK1142 (about one copy per cell), as judged from comparison of the densities of plasmid DNAs stained with ethidium bromide in agarose gel. This result may indicate that the activity is gene dosage dependent. The higher copy number of pJK1134 could be ascribed to the cryptic small plasmid cloned simultaneously into pJK1134 (see above; Fig. 3).

Proteins derived from the cloned regions. Immunogenic proteins synthesized from cloned regions were investigated by Western blot analysis, using serum from a monkey challenged with *S. flexneri* 1b. Products of the bacterial cells carrying no plasmids, pJK1134 alone, and both pJK1143 and 1134 were analyzed. Several bands could be detected in the bacterial cells carrying both plasmids; the amounts of these bands were simultaneously decreased in cells carrying pJK1134 alone (Fig. 4A). This effect of pJK1143 on the amounts of the products was consistent with the effect observed on contact hemolysis and invasion (see above). The fact that the amounts of several proteins were simultaneously increased in the presence of pJK1143 suggested the existence of a positive regulator on this plasmid. The proteins labeled a, b, c, and d (Fig. 4A) appeared to be the same as those identified by Hale and co-workers (10), as judged from their apparent molecular weights. In addition to these *ipa* proteins (5), 10 bands could be detected in pHW1273, pJK1134, and PJK1142. These were numbered according to molecular size (Fig. 4B).

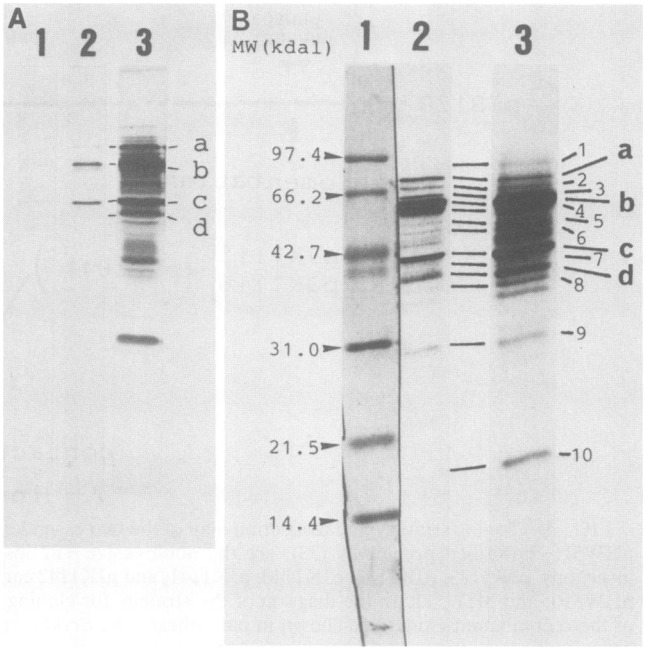

FIG. 4. Western blot hybridization analysis of peptides derived from recombinant clones. Whole bacterial cells were lysed in buffer at final concentrations of 0.062 M Tris hydrochloride (pH 6.8), 2% sodium dodecyl sulfate, 10% glycerol, and 5% 2-mercaptoethanol and boiled for 3 min. After centrifugation (10,000 × *g* for 10 min), the supernatant was used as a source of whole bacterial extracts, which were loaded for analysis of expression of peptides encoded by the recombinant clones. (A) Samples containing 10 μg of protein each. Lanes: 1, DH1 at 37°C; 2, DH1(pJK1134) at 37°C; 3, DH1(pJK1134, pJK1143) 37°C. (B) DH1 (pJK1134, pJK1143) at 37°C. Lanes: 1, molecular weight markers; 2 and 3, samples containing 8 (lane 2) and 10 (lane 3) μg of protein. Numbers on the right indicate positions of the recombinant clone-associated proteins that were detected. Proteins labeled a, b, c, and d indicate positions of the a, b, c, and d proteins reported by Hale et al. (9).

Kinetics of intracellular growth. The genes required for several virulence processes other than invasion were indicated by Sansonetti et al. (6, 28) to exist in the virulence plasmids, and the strains carrying the clones (17) were found to be able to invade but not to multiply in invaded HeLa cells. To determine whether the strains carrying the clones obtained in this study were able to proliferate in invaded LLC-MK2 cells, we investigated the kinetics of intracellular growth. The number of *E. coli* HB101 cells carrying both pJK1143 and either pJK1134 or pJK1142 in invaded LLC-MK2 cells was apparently constant, whereas the number of HB101 cells carrying pHW1273 (pSS120::Tn*1*, used as a positive control for contact hemolysis and invasion) increased after invasion (Fig. 5). This result indicates the existence of some gene region(s) essential for intracellular multiplication other than the regions cloned in this study.

Location and nucleotide sequence of the positive regulator gene. The putative positive regulator gene was subcloned from pJK1143 by transformation into DH1 carrying pJK1142 and assayed for contact hemolysis activity. A restriction map of pJK1143 was constructed, and plasmids (pJK1144

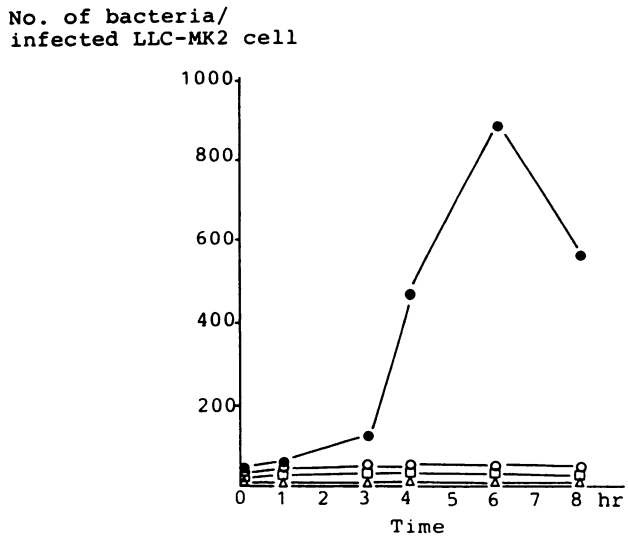

FIG. 5. Kinetics of bacterial growth in LLC-MK2 cells. Each point represents the mean of three experiments. Symbols: ●, pSS120::Tn*1*; ○, pJK1134-pJK1143; □, pJK1142-pJK1143; △, no plasmid.

and pJK1145) carrying one of the *Sma*I fragments of pJK1143 were found to be active (Fig. 6). Two series of deletion mutants were constructed from pJK1144 and pJK1145, and the contact hemolysis activity attributable to these plasmids and pJK1142 was assayed. The regulatory gene was found to be located in a 0.9-kb region (Fig. 6); the nucleotide sequence of this region is shown in Fig. 7. The nucleotide sequence was almost identical to that containing the *virF* gene, identified as the gene required for virulence and Congo red binding in *S. flexneri* (20, 21). The only difference between *virF* of *S. flexneri* and *virF* of *S. sonnei* was the existence of one extra base, A, at position 127 in the region identified (Fig. 7). One open reading frame that could encode 262 amino acids was found; this open reading frame may be significant, since the plasmids lacking parts of this putative coding region, pJK1149 and pJK1151, were inactive. The nucleotide sequence of this open reading frame was completely identical to that of *virF* of *S. flexneri*. However, the essential region of *virF* of *S. sonnei* was different from that of the *S. flexneri* gene in that both the upstream region containing the Pribnow box reported by Sakai et al. (20, 21) and the downstream region containing a typical terminator sequence were nonessential. A −35 and a −10 sequence were found in the region shown in Fig. 7. These differences might be due to differences in sensitivities of the assays used (we determined contact hemolysis activity, whereas Sakai et al. assayed for ability to bind Congo red) or to differences in the flanking regions of the *virF* genes of *S. flexneri* and *S. sonnei*.

Identification of the products of the regulatory gene. To determine whether the gene products of *virF* in *S. sonnei* and *S. flexneri* were similar, the products of pJK1148 (Fig. 6) were analyzed with minicells. To avoid interference by the product of the *bla* gene in pUC18, a plasmid vector, pACYC184, digested with *Pvu*II was inserted into the *Sca*I site within the *bla* gene to construct pUC18-pACYC184, pJK1148-pACYC184, and pJK1149-pACYC184. The products of the clones were labeled in minicells and analyzed by sodium dodecyl sulfate-gel electrophoresis. Two bands, 30 and 21 kDa, could be clearly seen in pJK1148-pACYC184 (Fig. 8, lane 2, 4, and 5). The 30-kDa protein was not produced by pJK1149-pACYC184, whereas the 21-kDa protein and the extra band (29 kDa) were detected. This result indicated that at least the 30-kDa protein was essential for the regulatory activity of pJK1148. The 21-kDa band was not produced by pJK1151 (Fig. 8, lane 6), which lacked the putative COOH-terminal region of the coding region (compare with pJK1150; Fig. 8, lane 7). These results suggested that these two proteins (30 and 21 kDa) could be derived from the same coding region by using different initiation sites.

Although a signal peptide-like amino acid sequence could be detected at the N terminus (amino acids 1 to 24 of the

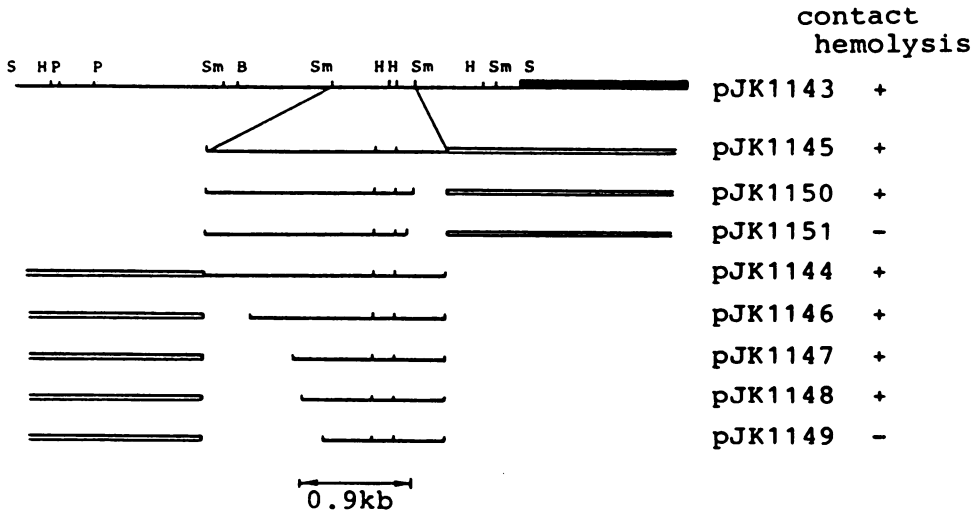

FIG. 6. Restriction maps of deletion derivatives of pJK1144 and pJK1145. pJK1150 and pJK1151 were derived from pJK1145. pJK1146, pJK1147, pJK1148, and pJK1149 were derived from pJK1144. Shown are vector pHSG595 (▬) and part of vector pUC18 (▭). Thin lines indicate cloned regions of the form I plasmid; + and − indicate positive and negative, respectively, for contact hemolysis activity in the strain carrying each of the plasmids as well as pJK1142. The 0.9-kb region containing the *virF* gene is indicated at the bottom. B, *Bam*HI; H, *Hin*dIII; P, *Pst*I; S, *Sal*I; Sm, *Sma*I.

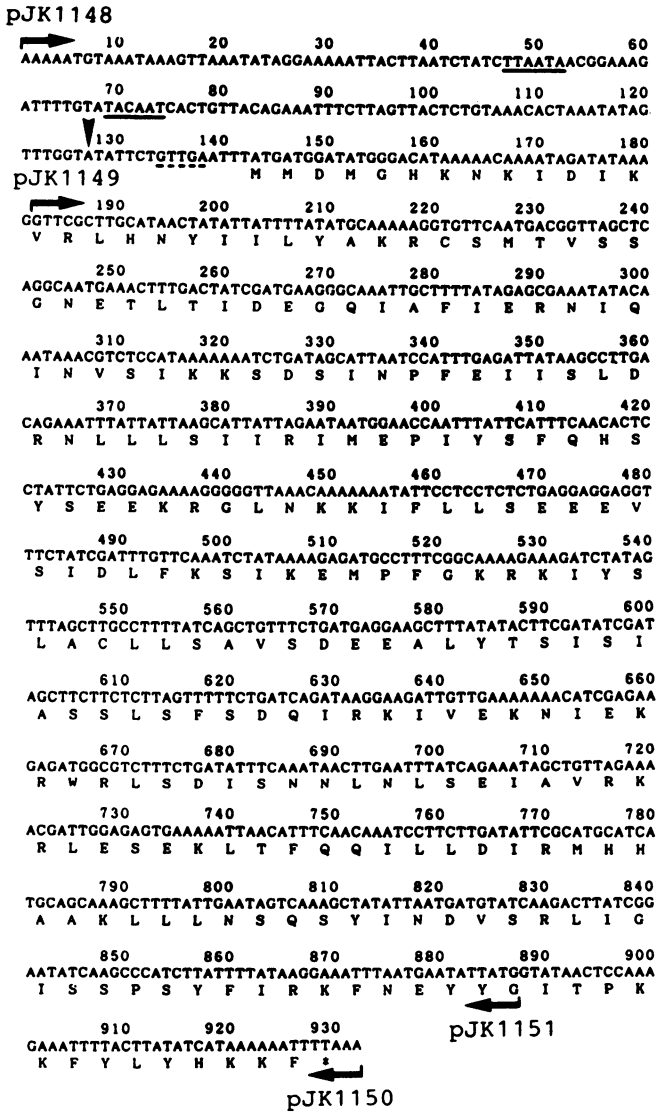

FIG. 7. Nucleotide sequence of the positive regulator gene. Shown are the DNA sequence of the cloned region from the deletion end of pJK1148 to that of pJK1150 and the deletion ends of pJK1149 and pJK1151 (◄). The deduced amino acid sequence of the open reading frame extending from bases 143 to 931 is shown beneath the corresponding DNA sequence. The most likely promoter region is underlined; a probable ribosome-binding site (Shine-Dalgarno sequence) is marked with dashed lines. One base, A, which is the only difference between the sequences of the *virF* genes of *S. flexneri* and *S. sonnei*, is indicated (▼).

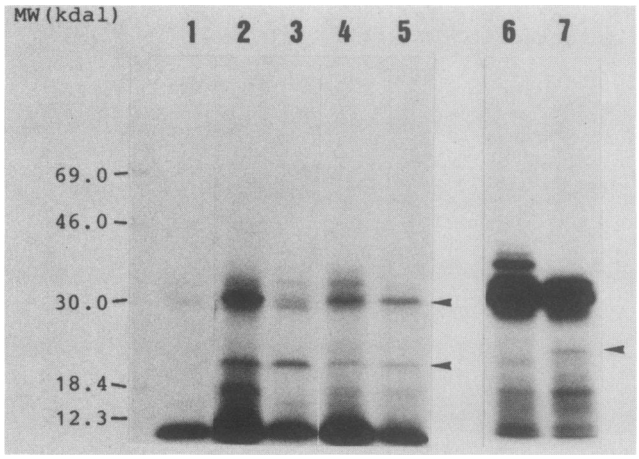

FIG. 8. Identification of the gene product. Proteins were synthesized in minicells with [^{35}S]methionine under direction of plasmids pUC18-pACYC184 (lane 1), pJK1148-pACYC184 (lanes 2, 4, and 5), pJK1149-pACYC184 (lane 3), pJK1151 (lane 6), and pJK1150 (lane 7). Lanes 4 and 5 show patterns before and after the 15-min chase that followed 5 min of labeling. Arrowheads indicate 30- and 21-kDa proteins.

product from the largest open reading frame), as reported by Sakai et al. (21), the electrophoretic pattern of the 30-kDa band was not changed by a 15-min chase after labeling for 5 min (Fig. 8, lanes 4 and 5). This observation indicates that processing of the putative signal peptide does not occur in the TH1219 background.

DISCUSSION

Two regions essential for both contact hemolysis and entry into LLC-MK2 cells were cloned in *E. coli* from the *S. sonnei* form I plasmid on compatible cosmid vectors. One region was obtained by isolating the plasmids that could complement two noninvasive Tn*1* insertion mutations, and the other was identified by examining the contact hemolysis-positive transformants carrying both the former region and cloned fragments from an *S. sonnei* cosmid library. This is the first case of cloning invasive genes in *E. coli*. The fact that the strains carrying two regions could invade LLC-MK2 cells is consistent with the observation that the *E. coli* strains carrying the form I plasmid could invade them (32). One of the invasion regions was found to contain the *invA* gene identified in our previous work (33). Restriction analysis indicated that this region is similar to the *S. flexneri* region that Maurelli et al. (17) found by analysis of Tn*5* insertion mutants and molecular cloning. Sasakawa et al. (29) also identified this invasion region by analysis of Tn*5* insertion mutants of *S. flexneri*. Our results show that the *E. coli* strains carrying the *invA* and *ipa* genes are not invasive, whereas Maurelli et al. (17) and Baudry et al. (1) reported that *S. flexneri* strains carrying the corresponding region could invade HeLa cells but were less active. This inconsistency may arise from the differences between *Shigella* strains and *E. coli* K-12; an experiment that includes construction of *S. sonnei* carrying pJK1134 is now under way. From analyses of Tn*5* insertion mutants, Sasakawa et al. (29) reported that the *virF* gene was essential for cell invasion and Congo red binding. This observation is completely consistent with our finding that both the *invA-ipa* and *virF* gene regions were necessary for invasion. The products might be synthesized without a positive regulator (*virF*) to some extent in *S. flexneri* from the clones obtained by Maurelli et al. (17), and low but definite activity may be attributable to these products.

The *virF* gene was identified as a genetic determinant for Congo red binding ability (Pcr) in *S. flexneri* 2a (20). The *E. coli* strain (MC1061) with the multicopy plasmid carrying the *virF* gene was also reported to express the Pcr$^+$ phenotype. However, we could not conclude that the *E. coli* strain (MC1061) with the multicopy plasmid carrying *virF* of *S. sonnei* expressed the Pcr$^+$ phenotype (J. Kato and H. Watanabe, unpublished data). We isolated the different regions responsible for Congo red binding from the form I

plasmid, pSS120 (H. Watanabe, unpublished data). The strain carrying this region was remarkably more Pcr^+ than was the strain carrying *virF*. Any Congo red-binding ability conferred by *virF* does not seem to be very intensive. Furthermore, Sakai and co-workers found that the *virF* products were localized in the outer membrane (T. Sakai, C. Sasakawa, and M. Yoshikawa, paper presented at the annual meeting of the Japanese Society for Bacteriology, 1987). The *virF* product may be similar to the *toxR* product (19) in acting as a positive regulator in the bacterial membrane. More detailed analysis is necessary to understand the mechanism by which the *virF* product might regulate the *inv* genes.

It was found that the large virulence-associated plasmids were interchangeable among *Shigella* species (26). Although there is somewhat diversity in the overall restriction patterns of the plasmid DNAs (10, 24), restriction maps of the essential regions for invasion showed that these regions were nearly conserved (33). Furthermore, this work shows that one of the genes required for penetration into *S. sonnei* cells was found to be identical, except for a single base pair insertion, to the counterpart in *S. flexneri*. The diversity in the patterns of digests with restriction enzymes among the large virulence-associated plasmids may be due to the recombinations caused by insertion sequences rather than to the base exchanges by point mutations.

Analyses of the kinetics of intracellular growth of the strains carrying the cloned regions suggest that one or more genes other than those cloned in this work are necessary for multiplication in the invaded cells. Our observation is consistent with the finding of Sansonetti et al. (28) that the *S. flexneri* strain carrying the *invA* region alone cannot proliferate in invaded HeLa cells. However, our results do not preclude the possibility that the apparent lack of intracellular growth of our clones in LLC-MK2 cells resulted from the genetic instability of the plasmids that we constructed. To investigate this possibility, we are now constructing genetically stable clones. If other genes are necessary for multiplication, it is possible that these genes could be isolated by transforming the plasmids from other cosmid libraries into the invasive strains constructed in this work. Although it is not clear how many steps are involved in the pathogenesis of *Shigella* spp., cloning of the determinants and reconstruction with the cloned regions as were done in this study could lead to an understanding of the mechanism.

While we were preparing this paper, Sakai et al. (22) reported that the expression of virulence antigens of an *S. flexneri* plasmid was positively regulated by the *virF* product. This result is consistent with our conclusion, although invasion and contact hemolysis activities were not directly assayed in their work. More detailed analyses of the effect of the *virF* gene product on other *inv* genes are in progress in our laboratory.

ACKNOWLEDGMENTS

We thank T. Hashimoto-Gotoh for plasmids and Y. Sasaki for technical assistance in construction of the restriction map of pJK1143.

This work was supported by grants from the Japan Health Sciences Foundation and the Ministry of Education, Science, and Culture of Japan.

LITERATURE CITED

1. **Baudry, B., A. T. Maurelli, P. Clerc, J. C. Sadoff, and P. J. Sansonetti.** 1987. Localization of plasmid loci necessary for the entry of *Shigella flexneri* into HeLa cells, and characterization of one locus encoding four immunogenic polypeptides. J. Gen. Microbiol. **133**:3403–3413.
2. **Boyer, H. W., and D. Roulland-Dussoix.** 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. J. Mol. Biol. **41**:459–472.
3. **Brady, G., H. M. Jantzen, H. U. Bernard, R. Brown, G. Schutz, and T. Hashimoto-Gotoh.** 1984. New cosmid vectors developed for eukaryotic DNA cloning. Gene **27**:223–232.
4. **Burnette, W. N.** 1981. "Western blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. Anal. Biochem. **112**:195–203.
5. **Buysse, J. M., C. K. Stover, E. V. Oaks, M. Venkatesan, and D. J. Kopecko.** 1987. Molecular cloning of invasion plasmid antigen (*ipa*) genes from *Shigella flexneri*: analysis of the *ipa* gene products and genetic mapping. J. Bacteriol. **169**:2561–2569.
6. **Clerc, P. L., A. Ryter, J. Mounier, and P. J. Sansonetti.** 1987. Plasmid-mediated early killing of eucaryotic cells by *Shigella flexneri* as studied by infection of J774 macrophages. Infect. Immun. **55**:521–527.
7. **Formal, S. B., T. L. Hale, and P. J. Sansonetti.** 1983. Invasive enteric pathogens. Rev. Infect. Dis. **5**:S702–S707.
8. **Formal, S. B., and R. B. Hornick.** 1978. Invasive *Escherichia coli*. J. Infect. Dis. **137**:641–644.
9. **Hale, T. L., E. V. Oaks, and S. B. Formal.** 1985. Identification and antigenic characterization of virulence-associated, plasmid-coded proteins of *Shigella* spp. and enteroinvasive *Escherichia coli*. Infect. Immun. **50**:620–629.
10. **Hale, T. L., P. J. Sansonetti, P. A. Schad, S. Austin, and S. B. Formal.** 1983. Characterization of virulence plasmids and plasmid-mediated outer membrane proteins in *Shigella flexneri*, *Shigella sonnei*, and *Escherichia coli*. Infect. Immun. **40**:340–350.
11. **Harayama, S., and G. L. Hazelbauer.** 1982. A gene coding for a periplasmic protein is located near the locus for termination of chromosome replication in *Escherichia coli*. J. Bacteriol. **151**: 1391–1396.
12. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28**: 351–359.
13. **Kato, J., H. Suzuki, and Y. Hirota.** 1984. Overlapping of the coding regions for α and γ components of penicillin-binding protein 1b in *Escherichia coli*. Mol. Gen. Genet. **196**:449–457.
14. **Kopecko, D. J., O. Washington, and S. B. Formal.** 1980. Genetic and physical evidence for plasmid control of *Shigella sonnei* form I cell surface antigen. Infect. Immun. **29**:207–214.
15. **LaBrec, E. H., H. Schneicher, T. J. Magnani, and S. B. Formal.** 1964. Epithelial cell penetration as an essential step in the pathogenesis of bacillary dysentery. J. Bacteriol. **88**:1503–1518.
16. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
17. **Maurelli, A. T., B. Baudry, H. d'Hauteville, T. L. Hale, and P. J. Sansonetti.** 1985. Cloning of plasmid DNA sequences involved in invasion of HeLa cells by *Shigella flexneri*. Infect. Immun. **49**:164–171.
18. **Maurelli, A. T., and P. J. Sansonetti.** 1988. Identification of a chromosomal gene controlling temperature-regulated expression of *Shigella* virulence. Proc. Natl. Acad. Sci. USA **85**: 2820–2824.
19. **Miller, V. L., R. K. Taylor, and J. J. Mekalanos.** 1987. Cholera toxin transcriptional activator ToxR is a transmembrane DNA binding protein. Cell **48**:271–279.
20. **Sakai, T., C. Sasakawa, S. Makino, K. Kamata, and M. Yoshikawa.** 1986. Molecular cloning of a genetic determinant for Congo red binding ability which is essential for the virulence of *Shigella flexneri*. Infect. Immun. **51**:476–482.
21. **Sakai, T., C. Sasakawa, S. Makino, and M. Yoshikawa.** 1986. DNA sequence and product analysis of the *virF* locus responsible for Congo red binding and cell invasion in *Shigella flexneri* 2a. Infect. Immun. **54**:395–402.
22. **Sakai, T., C. Sasakawa, and M. Yoshikawa.** 1988. Expression of

four virulence antigens of *Shigella flexneri* is positively regulated at the transcriptional level by the 30 kilodalton *virF* protein. Mol. Microbiol. **2**:589–597.

23. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
24. **Sansonetti, P. J., H. d'Hauteville, C. Ecobichon, and C. Pourcel.** 1983. Molecular comparison of virulence plasmids *Shigella* and enteroinvasive *Escherichia coli*. Ann. Microbiol. (Paris) **134A**: 295–318.
25. **Sansonetti, P. J., T. L. Hale, G. J. Dammin, C. Kapfer, H. H. Collins, Jr., and S. B. Formal.** 1983. Alterations in the pathogenicity of *Escherichia coli* K-12 after transfer of plasmid and chromosomal genes from *Shigella flexneri*. Infect. Immun. **39**:1392–1402.
26. **Sansonetti, P. J., D. J. Kopecko, and S. B. Formal.** 1981. *Shigella sonnei* plasmids: evidence that a large plasmid is necessary for virulence. Infect. Immun. **34**:75–83.
27. **Sansonetti, P. J., D. J. Kopecko, and S. B. Formal.** 1982. Involvement of a plasmid in the invasive ability of *Shigella flexneri*. Infect. Immun. **35**:852–860.
28. **Sansonetti, P. J., A. Ryter, P. Clerc, A. T. Maurelli, and J. Mounier.** 1986. Multiplication of *Shigella flexneri* within HeLa cells: lysis of the phagocytic vacuole and plasmid-mediated contact hemolysis. Infect. Immun. **51**:461–469.
29. **Sasakawa, C., S. Makino, K. Kamata, and M. Yoshikawa.** 1986. Isolation, characterization, and mapping of Tn*5* insertions into the 140-megadalton invasion plasmid defective in the mouse Sereny test in *Shigella flexneri* 2a. Infect. Immun. **54**:32–36.
30. **Timmis, K., F. Cabello, and S. Cohen.** 1975. Cloning, isolation, and characterization of replication region of complex plasmid genomes. Proc. Natl. Acad. Sci. USA **72**:2242–2246.
31. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19**:259–268.
32. **Watanabe, H., and A. Nakamura.** 1985. Large plasmids associated with virulence in *Shigella* species have a common function necessary for epithelial cell penetration. Infect. Immun. **48**: 260–262.
33. **Watanabe, H., and A. Nakamura.** 1986. Identification of *Shigella sonnei* form I plasmid genes necessary for cell invasion and their conservation among *Shigella* species and enteroinvasive *Escherichia coli*. Infect. Immun. **53**:352–358.
34. **Watanabe, H., and K. N. Timmis.** 1984. A small plasmid in *Shigella dysenteriae* 1 specifies one or more functions essential for O-antigen production and bacterial virulence. Infect. Immun. **3**:391–396.
35. **Yani**[illegible]**h-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nuclelotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.