JOURNAL OF BACTERIOLOGY, Dec. 1994, p. 7233–7243
0021-9193/94/$04.00+0

Vol. 176, No. 23

Comparative Analysis of the Replicon Regions of Eleven ColE2-Related Plasmids

SHIN-ICHIRO HIRAGA, TOMOHIKO SUGIYAMA,† AND TATEO ITOH*

Laboratory of Genetics, Department of Biology, Faculty of Science, Osaka University, Toyonaka, Osaka 560, Japan

Received 2 May 1994/Returned for modification 25 July 1994/Accepted 7 September 1994

The *incA* gene product of ColE2-P9 and ColE3-CA38 plasmids is an antisense RNA that regulates the production of the plasmid-coded Rep protein essential for replication. The Rep protein specifically binds to the origin and synthesizes a unique primer RNA at the origin. The IncB incompatibility is due to competition for the Rep protein among the origins of the same binding specificity. We localized the regions sufficient for autonomous replication of 15 ColE plasmids related to ColE2-P9 and ColE3-CA38 (ColE2-related plasmids), analyzed their incompatibility properties, and determined the nucleotide sequences of the replicon regions of 9 representative plasmids. The results suggest that all of these plasmids share common mechanisms for initiation of DNA replication and its control. Five IncA specificity types, 4 IncB specificity types, and 9 of the 20 possible combinations of the IncA and IncB types were found. The specificity of interaction of the Rep proteins and the origins might be determined by insertion or deletion of single nucleotides and substitution of several nucleotides at specific sites in the origins and by apparently corresponding insertion or deletion and substitution of amino acid sequences at specific regions in the C-terminal portions of the Rep proteins. For plasmids of four IncA specificity types, the nine-nucleotide sequences at the loop regions of the stem-loop structures of antisense RNAs are identical, suggesting an evolutionary significance of the sequence. The mosaic structures of the replicon regions with homologous and nonhomologous segments suggest that some of them were generated by exchanging functional parts through homologous recombination.

Colicin E plasmids (ColE plasmids) encode colicins which use the host *btuB* gene product on the bacterial membrane as the mediator for their bactericidal actions (12, 35). The specificity of the colicin-immunity protein interaction was used to classify colicins into nine subtypes. Restriction analyses have revealed some structural similarities among plasmids ColE2 to ColE9 (8, 32, 76, 77), whereas the ColE1 plasmid ColE1-K30 is structurally distinct (9). These ColE plasmids other than the ColE1 plasmid will be collectively called ColE2-related plasmids (Table 1). Plasmid incompatibility between some of ColE2-related plasmids has been examined by measuring mutual exclusion. Plasmids ColE3-CA38, ColE7-K317, and ColE8-J belong to one incompatibility group (11, 39), and plasmids ColE5-099, ColE6-CT14, and ColE9-J belong to another incompatibility group (11). Plasmids ColE2-P9 and ColE3-CA38 require plasmid-coded Rep proteins (19, 22) and host DNA polymerase I (25, 66) for plasmid replication. The Rep protein binds to the origin in a plasmid-specific manner (23) and synthesizes a unique primer RNA (ppApGpA) for initiation of the leading-strand DNA synthesis by DNA polymerase I at the origin (68, 69; see Fig. 1 and 5). The specificity of interaction of the origins with the Rep proteins is determined by the absence or presence of a single base pair in the origin (80). Cloned origins exclude autonomously replicating plasmids (IncB function) in a plasmid-specific manner due to competition for the Rep proteins (67). A pair of derivatives of ColE2-P9 or a pair of derivatives of ColE3-CA38 exclude each other very efficiently. The expression of the *rep* gene is regulated negatively at a posttranscriptional step by the *incA* gene product, an antisense RNA (RNA I) complementary to the 5′ untranslated region of the Rep mRNA (70, 81). RNA I rapidly forms a stable complex with the Rep mRNA in vitro, and the rate of binding is affected even by single nucleotide substitutions, which result in increase in plasmid copy numbers (62). The nucleotide sequences of the portions of ColE2-P9 and ColE3-CA38 in and around the RNA I-coding regions are identical to each other (80), and they belong to the same IncA specificity group. Between autonomously replicating ColE2-P9 and ColE3-CA38 plasmids, however, mutual exclusion is very inefficient (11, 20, 39, 67), and these plasmids have been occasionally judged to be compatible with each other. The region of about 250 bp located next to the *incA* gene (Fig. 1) is involved in stable maintenance of the plasmids by host bacteria and resolution of plasmid multimers to monomers (80). The region shows extensive structural similarity to the region of ColE1 plasmid containing the *cer* site required for similar functions (64), suggesting that ColE2-P9 and ColE3-CA38 use a host site-specific recombination system identical to that used by the ColE1-type plasmids.

Here we localized the regions of 15 ColE2-related plasmids sufficient for autonomous replication. We showed that the basic mechanisms of initiation of replication and its regulation are identical among these plasmids. We found four new IncA specificity types and two new IncB specificity types. The mosaic structures of the nucleotide sequences of the replicon regions suggest that some of these structures have arisen by exchange of functional parts of the replicon regions among several plasmids by homologous recombination. We also discuss the specificity determinants of the IncA and IncB functions.

(A preliminary account of this report was presented at the EMBO-FEMS-NATO Advanced Research Workshop on Bacteriocins, Microcins and Lantibiotics held on the Isle de Bendor, France, 22–26 September 1991 [21].)

* Corresponding author. Mailing address: Laboratory of Genetics, Department of Biology, Faculty of Science, Osaka University, Toyonaka, Osaka 560, Japan. Phone: 81-6-850-5432. Fax: 81-6-841-2449.

† Present address: Department of Molecular Protozoology, Research Institute of Microbial Diseases, Osaka University, Suita, Osaka 565, Japan.

TABLE 1. Properties of ColE2-related plasmids

Plasmid	Original host	Type of colicin	Type of immunity against colicin	Type of incompatibility		Source[a]
				IncA (RNA I)	IncB (Rep/Ori)	
ColE2-CA42	*E. coli*	E2	E2	E2-CA42	E2-CA42	1
ColE2-K321[b]	*E. coli*	E2	E2	E2-P9	E2-P9	1
ColE2-GEI288[b]	*E. coli*	E2	E2	E2-P9	E2-P9	1
ColE2-GEI554[b]	*E. coli*	E2	E2	E2-P9	E2-P9	1
ColE2-GEI602	*E. coli*	E2	E2	E5-099	E2-CA42	1
ColE2-P9[b]	*Shigella* sp.	E2	E2	E2-P9	E2-P9	2
ColE2imm-K317	*E. coli*		E2	E2imm-K317[c]	E2-CA42	1
ColE3-CA38	*E. coli*	E3	E3, E8	E2-P9	E3-CA38	2
ColE4-284[d]	*E. coli*	E4	E4	E2-CA42	E2-P9	1
ColE4-CT9[d]	*Shigella sonnei*	E4	E4	E2-CA42	E2-P9	2
ColE4-K365[d]	*E. coli*	E4	E4	E2-CA42	E2-P9	1
ColE5-099	*S. sonnei*	E5	E5	E5-099	E5-099	2
ColE6-CT14[e]	*S. sonnei*	E6	E6, E8	E5-099	E5-099	2
ColE6-Ind8[e]	*S. sonnei*	E6	E6, E8	E5-099	E5-099	2
ColE7-K317	*E. coli*	E7	E7	E2-CA42	E3-CA38	1
ColE8-J	*E. coli*	E8	E8	E2-P9[f]	E3-CA38	3
ColE9-J	*E. coli*	E9	E9, E5	E9-J	E5-099	3

[a] 1, R. Watson; 2, H. Masaki; 3, R. James.
[b] These four ColE2 plasmids gave identical restriction maps (76).
[c] Exclusion by its own *incA* gene was very slow.
[d] These three ColE4 plasmids gave identical restriction maps (76).
[e] These two ColE6 plasmids gave identical restriction maps (data not shown).
[f] ColE8 was only slowly excluded by the cloned *incA* gene of the ColE2-P9 type, but the cloned *incA* gene of ColE8 efficiently excluded other plasmids with the IncA specificity of ColE2-P9 type.

MATERIALS AND METHODS

Bacterial strains and plasmids. The *Escherichia coli* K-12 strains used are RR1*btuB* (4, 36), JM109 (79), AG1*recA1* (17), and AK101*recA*::Tn*10* (57a), a derivative of JM103 (38). Other bacterial strains have been described elsewhere (19, 70, 80). Plasmids used are pUC118 and pUC119 (74) and the 17 ColE plasmids listed in Table 1. Other plasmids have been described elsewhere (19, 72, 80, 81) except for those described below.

Construction of plasmids. An autonomously replicating derivative of each ColE2-related plasmid carrying a chloramphenicol resistance gene (pEC series) was constructed by ligating the 1.2-kb fragment containing the resistance gene (19) via *Bam*HI linkers with an appropriate fragment of each ColE2-related plasmid as follows: the 2.4-kb *Pvu*II fragments of ColE2-K321, ColE2-GEI288, ColE2-GEI554, ColE2-GEI602, ColE5-099, ColE8-J, and ColE9-J; the 2.0-kb *Pvu*II-*Pst*I fragment ColE2-CA42; the 1.5-kb *Hin*cII-*Bgl*I fragment of ColE2imm-K317; the 2.3-kb *Bgl*II-*Pvu*II fragments of ColE4-284, ColE4-CT9, and ColE4-K365; the 1.8-kb *Hin*cII-*Bgl*I fragments of ColE6-CT14 and ColE6-Ind8; and the 2.3-kb *Hin*cII-*Pst*I fragment of ColE7-K317. Plasmids carrying the two fragments in different orientations are named as described for pEC21 and pEC22 (19), and each plasmid is distinguished by a suffix representing the original plasmid; e.g., pEC22-CA42 derived from ColE2-CA42. A derivative of pNT51*inc1,2* (72) carrying the *incA* or *incB* region of each plasmid (pTI51A or pTI51B series, respectively) was constructed by replacing the *Bam*HI-*Eco*47III or *Bam*HI-*Eco*RV fragment with an appropriate fragment of each plasmid of the pEC series: the 0.6-kb *Bam*HI-*Hin*cII fragments of pEC22-CA42, pEC52, pEC72, and pEC92; the 0.7-kb *Bam*HI-*Hin*cII fragment of pEC22imm; the 0.6-kb *Bgl*II-*Hin*cII fragment of pEC42-CT9; and the 1.1-kb *Bam*HI-*Cla*I fragments of pEC62-CT14 and pEC82 for the pTI51A series; and the 1.8-kb *Bam*HI-*Hin*cII fragments of pEC22-CA42, pEC52, pEC72, and pEC92; the 0.8-kb *Bam*HI-*Hin*cII fragment of pEC22imm; the 1.7-kb *Bam*HI-*Hin*cII fragment of pEC42-CT9; and the 1.3-kb *Bam*HI-*Cla*I fragments of pEC62-CT14 and pEC82 for the pTI51B series. Each plasmid is distinguished by a suffix; e.g., pTI51A2-CA42 carries the fragment derived from ColE2-CA42.

Media, enzymes, antibiotics, and chemicals. L broth (43) and Terrific broth (55) were as described previously. Enzymes, antibiotics, and chemicals were from commercial sources.

Incompatibility test. RR1*btuB* cells harboring each of the original ColE2-related plasmids were transformed with plasmids carrying the cloned segments containing the *incA* regions (pTI51A series) or the *incB* regions (pTI51B series) of ColE2-P9 and ColE3-CA38 and the corresponding segments of representatives of other ColE2-related plasmids. The resultant transformants were selected only for the incoming plasmids and then tested for production of the colicins specified by the resident plasmids, using *E. coli* N100 as an indicator strain. N100 cells harboring each of the autonomously replicating derivatives of ColE2-related plasmids (pEC series) were transformed with plasmids of the pTI51 series and selected only for the incoming plasmids as described above. The resultant transformants were then tested for the drug resistance determined by the resident plasmids. The presence or absence of the resident plasmids was judged after incubation for 6 h at 37°C or for 12 h at 30°C. This procedure was repeated at least three times to confirm whether the resident plasmids were stably maintained.

DNA manipulations. Most DNA manipulations were performed as described previously (55). Nucleotide sequences were determined essentially as described previously (56), using a T7 sequencing kit (Pharmacia) on both strands of DNA for most of the sequenced regions described in this report; for the remaining several short regions of some plasmids, overlapping fragments of one strand were sequenced.

Computer analysis. DNA and amino acid sequence analyses were performed by using DNASIS sequence analysis software (Hitachi Software Engineering Co., Ltd., Kanagawa, Japan).

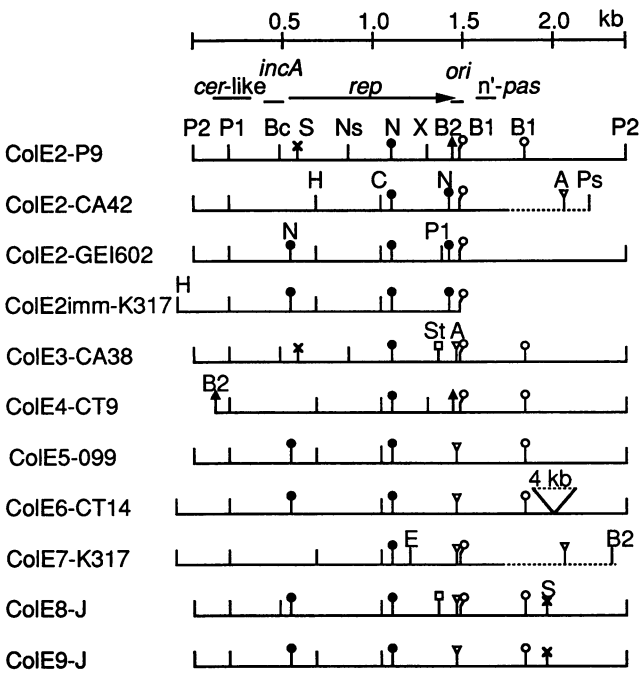


FIG. 1. Restriction maps of the segments of 11 ColE2-related plasmids containing the regions sufficient for autonomous replication (modified from those described previously [8, 19, 32, 76, 80]). Approximate positions of the known genes and sites of ColE2-P9 and ColE3-CA38 are shown above the maps: *cer*-like, the region homologous to the site of ColE1 plasmid for host site-specific recombination; *incA*, the region specifying the antisense RNA; *rep*, the region specifying the Rep protein; *ori*, replication origin; n'-*pas*, the primosome assembly site recognized by the n' protein. Dotted lines show the segments which are not homologous to the corresponding regions of other plasmids (76). Restriction sites: A (▽), *Acc*I; Bc, *Bcl*I; B1 (○), *Bgl*I; B2 (▲), *Bgl*II; C, *Cla*I; E, *Eco*RI; H, *Hin*cII; N (●), *Nsp*V; Ns, *Nsi*I; Ps, *Pst*I; P1, *Pvu*I; P2, *Pvu*II; S (×), *Spl*I; St (□), *Sty*I; X, *Xho*I. Some restriction sites are also marked with appropriate symbols as indicated in parentheses. The *Bcl*I, *Nsp*V, *Nsi*I, *Spl*I, and *Sty*I sites for plasmids other than ColE2-P9 and ColE3-CA38 were determined in this study. We were unable to find the unique *Bam*HI site between the *Cla*I and *Bgl*I sites of ColE6-CT14 reported previously (8).

DNA and RNA homology searches were done by using GenBank and EMBL sequence libraries. Protein homology searches were performed by using Protein Identification Resource and SWISS-PROT.

Nucleotide sequence accession numbers. The nucleotide sequence data reported here have been deposited in the DDBJ, EMBL, and GenBank databases under accession numbers D30054 to D30064 for ColE2-P9, ColE3-CA38, ColE2-CA42, ColE2-GEI602, ColE2imm-K317, ColE4-CT9, ColE5-099, ColE6-CT14, ColE7-K317, ColE8-J, and ColE9-J, respectively.

RESULTS AND DISCUSSION

Relationships of 15 ColE2-related plasmids to ColE2-P9 and ColE3-CA38. We examined whether the cloned *incA* and *incB* genes of ColE2-P9 and ColE3-CA38 (67) exclude the coexisting ColE2-related plasmids listed in Table 1 except for ColE2imm-K317 (data not shown). The results showed that plasmids ColE2-K321, ColE2-GEI288, ColE2-GEI554, ColE2-P9, ColE3-CA38, and ColE8-J belong to a single specificity group (ColE2-P9 group) of the IncA incompatibility mediated by an antisense RNA, RNA I (Table 1). The results also showed that plasmids ColE2-K321, ColE2-GEI288, ColE2-GEI554, ColE2-P9, ColE4-284, ColE4-CT9, and ColE4-K365 belong to a single specificity group (ColE2-P9 group) of the IncB incompatibility as a result of the specific Rep protein-origin binding and that plasmids ColE3-CA38, ColE7-K317, and ColE8-J belong to another specificity group (ColE3-CA38 group) of the IncB incompatibility (Table 1). The results together strongly suggest that these 10 ColE2-related plasmids have identical mechanisms of initiation of replication and its control. The cloned *incA* and *incB* genes of ColE2-P9 and ColE3-CA38, however, did not exclude ColE2-CA42, ColE2-GEI602, ColE5-099, ColE6-CT14, ColE6-Ind8, and ColE9-J.

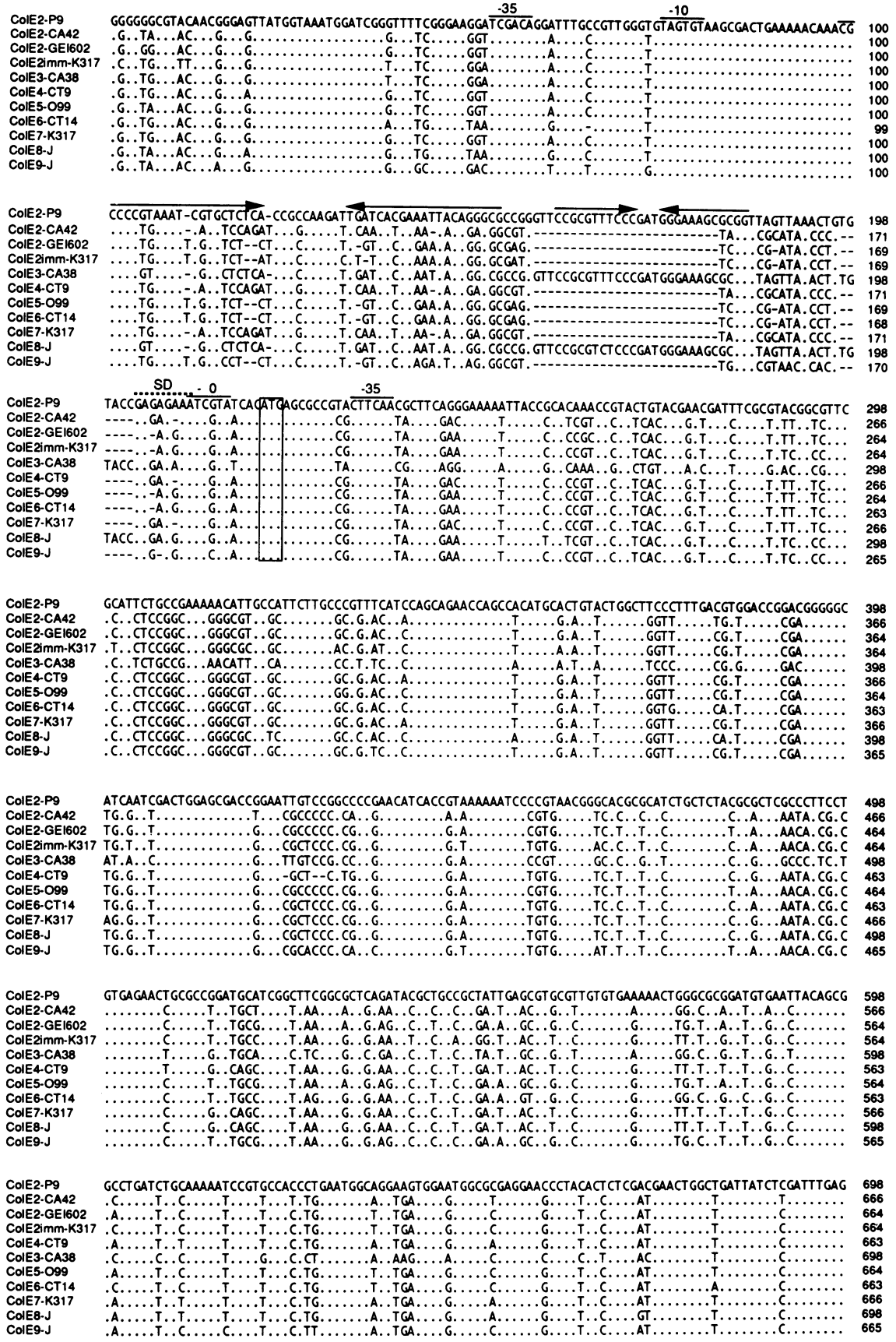
Regions of ColE2-related plasmids sufficient for autonomous replication. Restriction analyses (8, 32, 76, 77) (Fig. 1) have suggested that each of these ColE2-related plasmids contains a region which shows a certain structural similarity to the region of ColE2-P9 or ColE3-CA38 containing the 1.3-kb segments sufficient for autonomous replication (19). The fragment of each ColE2-related plasmid containing the corresponding region ligated to the fragment containing a chloramphenicol resistance gene as described in Materials and Methods results in an autonomously replicating plasmid (data not shown). Plasmids composed of these two fragments joined in either orientation were obtained. Fine restriction maps of the cloned regions of the nine representative ColE2-related plasmids (Fig. 1) further revealed structural similarities in the replicon regions of these plasmids.

Specificities of IncA and IncB functions of ColE2-related plasmids. We then cloned a segment of each of the nine representative plasmids (shown in Fig. 1) corresponding to that of ColE2-P9 or ColE3-CA38 carrying the *incA* or *incB* gene and examined whether the cloned segments exclude these ColE2-related plasmids (data not shown). We found four new IncA specificity types and two new IncB specificity types (Table 1). Consequently, there are five IncA types and four IncB types in total, and 9 combinations of them among 20 possible ones were found. All of these results demonstrated that these plasmids are interrelated by sharing either one or both of the IncA and IncB specificity types, indicating that they have identical basic mechanisms of initiation of replication and its control.

Plasmids ColE6-CT14 and ColE6-Ind8 were revealed to be indistinguishable and to carry the second replicons in the additional 4-kb regions unique for these plasmids (data not shown). The plasmid carrying only the second replicon is present at a few copies per host chromosome and seems to be different from any of these ColE2-related plasmids. The asymmetry of incompatibility against ColE5-099 and ColE9-J exhibited by ColE6-CT14 (11) and ColE6-Ind8 (data not shown) is readily explained by the fact that they are composite plasmids.

Nucleotide sequences of the replicon regions of ColE2-related plasmids. We determined the nucleotide sequences of the replicon regions of the nine representative ColE2-related plasmids (Fig. 2) and compared them with those of ColE2-P9 and ColE3-CA38 (80). These 11 plasmids are representatives for the eight different types of colicin proteins and for the nine different combinations of IncA and IncB specificity types. The sequences are highly homologous, further supporting the view that these plasmids have identical mechanisms of initiation of replication and its control.

The *rep* genes and Rep proteins. The longest open reading frame of each ColE2-related plasmid specifies a protein quite similar to the Rep protein of ColE2-P9 or ColE3-CA38 (Fig. 3), strongly suggesting that these open reading frames encode the Rep proteins of these plasmids.


-35
-10
SD
- 0
-35
ColE2-P9
ColE2-CA42
ColE2-GEI602
ColE2imm-K317
ColE3-CA38
ColE4-CT9
ColE5-O99
ColE6-CT14
ColE7-K317
ColE8-J
ColE9-J

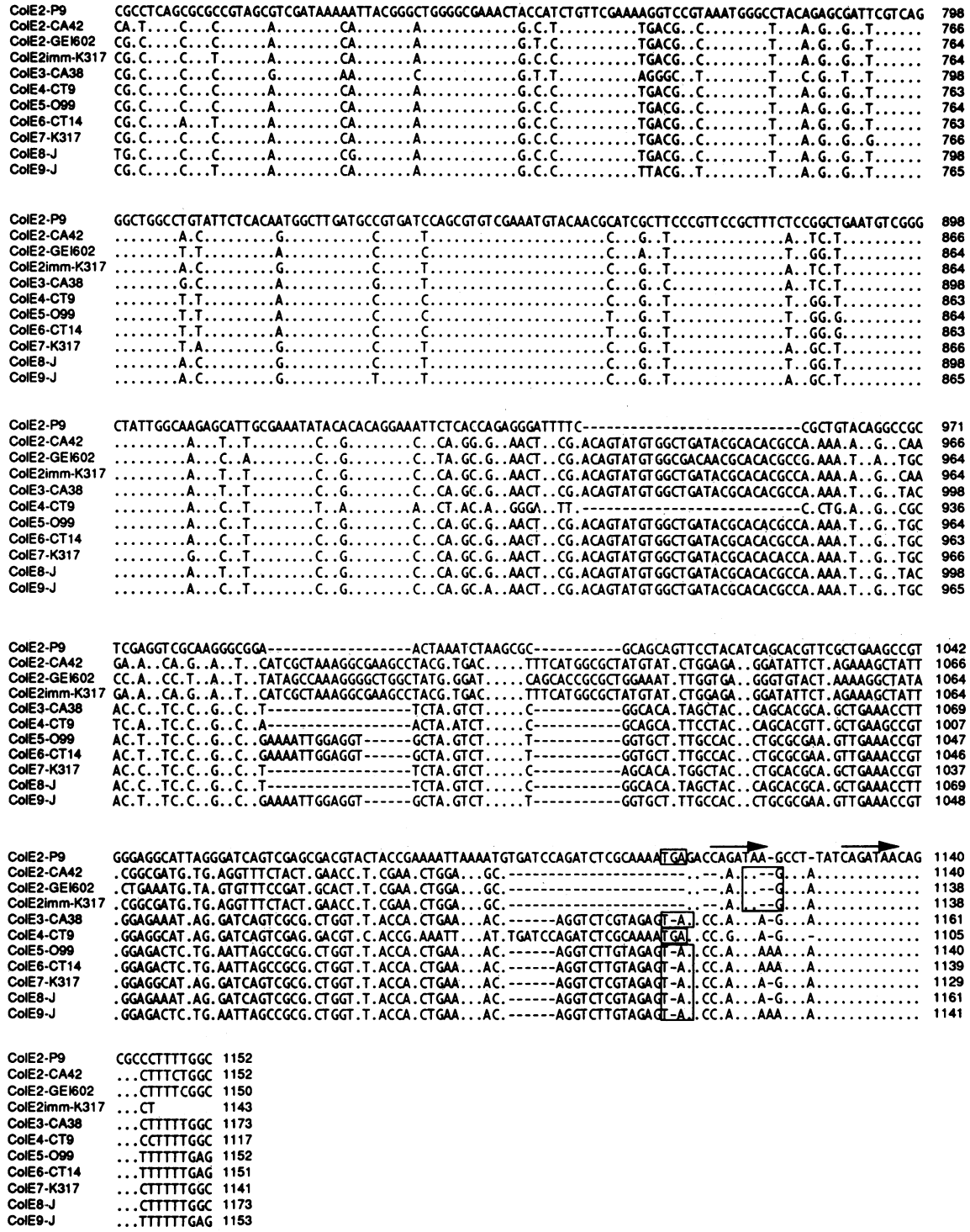


FIG. 2. Nucleotide sequences of the 1.15-kb regions sufficient for autonomous replication of the 11 ColE2-related plasmids. The complete nucleotide sequence of ColE2-P9 (80) is shown at the top. For other plasmids, only the nucleotides different from those of ColE2-P9 are shown; positions with identical nucleotides are shown by dots. Hyphens in the sequences represent deletions. Position 1 in the ColE2-P9 sequence corresponds to position 338 described previously (80). The promoters (marked as −35 and −10) for the Rep mRNA and RNA I are indicated. The putative Shine-Dalgarno (S/D) region for the Rep protein is also indicated. The initiation and termination codons for the (putative) Rep proteins are boxed. Arrows represent major direct and inverted repeat sequences.

```
ColE2-P9      MSAVLQRFREKLPHKPYCTNDFAYGVRILPKNIAILARFIQQNQPHALYWLPFDVDRTGASIDWSDRNCPAPNITVKNPRNGHAHLLYALALPVRTAPDASASALRYAAAIERALCEKLG  120
ColE4-CT9     ...A..Y.D.N...R..H.D.L.F.L..SG.GR.L..RY........QF..V..V..E..A.....RK-LV....V...V..H.......NIA......S.VK..K....I.RS.....C  119
ColE3-CA38    ...V..R.R.K...K..C.N.F.Y.V..LP.NI.I..RF........LY..P..V..T..S.....RNCPA....V...R..H.......ALP......A.AS..R....I.RA.....G  120
ColE7-K317    ...A..Y.D.N...R..H.D.L.F.L..SG.GR.L..RY........QF..V..V..E..A.....RNAPA....V...V..H.......NIA......S.VK..K....I.RS.....C  120
ColE8-J       ...A..Y.E.N...R..H.D.L.F.L..SG.GR.L..RY........QF..V..I..E..A.....RNAPA....V...V..H.......NIA......S.VK..K....I.RS.....C  120
ColE5-099     ...A..Y.E.N...R..H.D.L.F.L..SG.GR.L..GY........QF..V..V..E..A.....RNAPA....V...V..H.......NIA......A.VK..K....I.RA.....G  120
ColE6-CT14    ...A..Y.E.N...R..H.D.L.F.L..SG.GR.L..RY........QF..V..I..E..A.....RNAPA....V...V..H.......NIA......A.VR..K....I.CA.....G  120
ColE9-J       ...A..Y.E.N...R..H.D.L.F.L..SG.GR.L..RF........QF..V..V..E..A.....RNAPA....V...V..Y.......NIA......A.VK..K....I.RA.....G  120
ColE2-CA42    ...A..Y.D.N...R..H.D.L.F.L..SG.GR.L..RY........QF..V..V..E..A.....LNAPA....I...V..H.......NIA......A.VK..K....I.RA.....G  120
ColE2-GEI602  ...A..Y.E.N...R..H.D.L.F.L..SG.GR.L..RY........QF..V..V..E..A.....RNAPA....V...V..H.......NIA......A.VK..K....I.RA.....G  120
ColE2imm-K317 ...A..Y.E.N...R..H.D.L.F.L..SG.GR.L..RY........QF..V..V..E..A.....RNAPA....V...V..H.......NIA......A.VK..K....V.RS.....C  120

ColE2-P9      ADVNYSGLICKNPCHPEWQEVEWREEPYTLDELADYLDLSASARRSVDKNYGLGRNYHLFEKVRKWAYRAIRQGWPVFSQWLDAVIQRVEMYNASLPVPLSPAECRAIGKSIAKYTHRKF  240
ColE4-CT9     .............F.L..QVM.....A...D.........A.A...I..H..M...CH...MT.......I.....V..............A........A......K........K.  239
ColE3-CA38    .............C.P..QEV.....P...D.........A.A...V..N..L...CY...KG.......I.....A..............A........P......K........N.  240
ColE7-K317    .............F.L..QVM.....A...D.........A.A...I..H..M...CH...MT.......M.....E..............A........P......R........N.  240
ColE8-J       .............F.L..QVM.....A...G.........A.A...I..R..M...CH...MT.......I.....A..............A........A......K........N.  240
ColE5-099     .............F.L..LVM.....A...D.........A.A...I..H..M...CH...MT.......I.....V..............V........A......K........N.  240
ColE6-CT14    .............F.L..LVM.....A...D.........A.E...I..H..M...CH...MT.......I.....V..............V........A......K........N.  240
ColE9-J       .............F.L..QVM.....A...D.........A.A...I..H..M...CH...IT.......I.....A..............A........P......K........N.  240
ColE2-CA42    .............F.L..QVM.....A...D.........T.A...I..H..M...CH...MT.......I.....A..............A........P......K........N.  240
ColE2-GEI602  .............F.L..LVM.....A...D.........A.A...I..H..M...CY...MT.......I.....V..............A........A......K........N.  240
ColE2imm-K317 .............F.L..LVM.....A...D.........A.A...I..H..M...CH...MT.......I.....A..............A........P......K........N.  240

                        A                C                B
ColE2-P9      SPEGFS---------AVQAARGRKGG------TKSKRAAVPTSARSLKPWEALGISRATYYRKLKCDPDLAK  297
ColE4-CT9     SP.G.S---------AV..A.....------TKS..AAVPTSARSLKPWEALGISRATYYRKLKCDPDLAK  296
ColE3-CA38    TP.T.AQYVADTHTPEI..T.....------SKS..GTVATSARTLKPWEKLGISRAWYYQLKKR--GLVE  304
ColE7-K317    TP.T.AQYVADTHTPEI..A.....------SKS..STVATSARTLKPWEALGISRAWYYQLKKR--GLVE  304
ColE8-J       TP.T.AQYVADTHTPEI..T.....------SKS..GTVATSARTLKPWEKLGISRAWYYQLKKR--GLVE  304
ColE5-099     TP.T.AQYVADTHTPEI..A.....--KIGGAKS..GAVATSARTLKPWETLGISRAWYYQLKKR--GLVE  308
ColE6-CT14    TP.T.AQYVADTHTPEI..A.....--KIGGAKS..GAVATSARTLKPWETLGISRAWYYQLKKR--GLVE  308
ColE9-J       TP.T.AQYVADTHTPEI..A.....--KIGGAKS..GAVATSARTLKPWETLGISRAWYYQLKKR--GLVE  308
ColE2-CA42    TA.T.AQYVADTHTPEI..K.....IAKGEAYDD..FMALCMLENGYSQKAIAAMLEVSTRTIRNWKSGK  311
ColE2-GEI602  TP.T.AQYVATTHTPEI..A.....IAKGAGYED..STALEMLVNGCTQKAIAEMLGVSDRTIRNWKSGK  311
ColE2imm-K317 TP.T.AQYVADTHTPEI..K.....IAKGEAYDD..FMALCMLENGYSQKAIAAMLEVSTRTIRNWKSGK  311
```

FIG. 3. The amino acid sequences of the (putative) Rep proteins of the 11 ColE2-related plasmids. The complete amino acid sequence of ColE2-P9 Rep protein (23, 80) is indicated at the top. For other plasmids, only amino acids different from those of ColE2-P9 Rep protein are indicated; positions with identical amino acids are shown by dots. Hyphens in the sequences represent deletions. Plasmids are grouped according to the IncB specificity types. The regions containing group-specific insertions of amino acid sequences are indicated as A, B, and C.

The amino acid sequences of the N-terminal regions of the (putative) Rep proteins (positions 1 to around 180; Fig. 3) may be divided into two groups: one with ColE2-P9 and ColE3-CA38 and the other with the remaining plasmids. Amino acid sequences of the adjacent regions (around positions 180 to 240) are highly conserved among all of these plasmids. Many of the amino acid substitutions among them are conservative. These results suggest that the two adjacent regions may be involved in the primase activity presumably common to all of these Rep proteins.

On the other hand, the remaining C-terminal regions are almost identical in plasmids of the same IncB specificity groups but are considerably different among plasmids of different IncB specificity groups. Plasmids of ColE2-CA42, ColE3-CA38, and ColE5-099 groups possess an insertion of a stretch of nine amino acids (insertion A in Fig. 3), plasmids of the ColE2-P9 group possess an insertion of two amino acids (insertion B), and plasmids of the ColE5-099 group also possess an insertion of a stretch of four amino acids (insertion C). For plasmids of the ColE2-CA42 group, a portion of the C-terminal region (position 270 to the terminus) is quite different from those of plasmids of other IncB specificity groups. Thus, the C-terminal regions seem to play an important role in plasmid-specific interaction of the Rep protein with the origin.

There are two amino acid sequences which show some homologies to the sequence-specific DNA-binding domains (helix-turn-helix motif): one in the region around the site of insertion C (starting at position 249 of ColE2-P9 Rep protein), which is homologous to *E. coli* LexA (30) and bacteriophage immunity repressors, including P22 c protein (45) as described previously (23) (Fig. 3), and the other in the region to the left of the position of insertion B in plasmids of ColE2-P9, ColE3-CA38, and ColE5-099 types, which is homologous to the putative sequence-specific DNA-binding domains (42) of *Klebsiella pneumoniae* NtrC (16) and *Rhizobium leguminosarum* DctD (52) proteins (Fig. 4). These sequence features might suggest that the C-terminal regions of the Rep proteins contain the plasmid-specific DNA-binding domains.

The consensus amino acid sequence motif (EGYhD; h for hydrophobic aliphatic amino acid residues) for prokaryotic DNA primases (46) is not found in the (putative) Rep proteins. Computer-assisted searches revealed no significant homologies with the Rep proteins of other plasmids, including IncFII, ColIb, pSC101, F, and P1, or with the primases, including *E. coli* DnaG protein, T7 gene 4 protein, T4 gene 61 protein, and *Saccharomyces cerevisiae* PRI1 protein. We found, however,

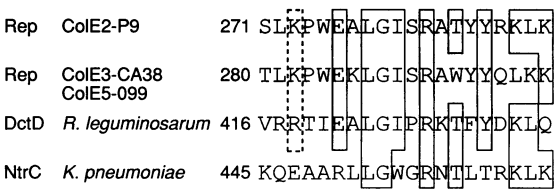


FIG. 4. Portions of the (putative) Rep proteins of ColE2-related plasmids showing homology to the DNA-binding domains (helix-turn-helix motif) of *R. leguminosarum* DctD (52) and *K. pneumoniae* NtrC proteins (16). Numbers to the left of the sequences are position numbers of amino acids at the left ends. Identical amino acids present in at least three of the four proteins are boxed, and equivalent amino acids are boxed with broken lines.

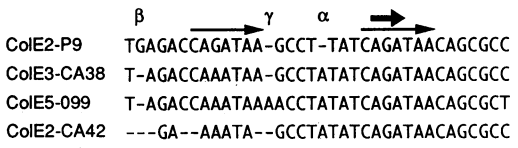


FIG. 5. Nucleotide sequences of the (putative) origins of plasmids of the four IncB specificity groups. The thick arrow indicates the primer RNA (ppApGpA) and direction of replication in ColE2-P9 (68). Thin arrows indicate major direct repeat sequences. Positions of group-specific insertions or deletions of single nucleotides are indicated as α, β, and γ.

significant homologies between the Rep proteins of the ColE2-related plasmids and the proteins specified by potential open reading frames near the putative origin regions of plasmids pAL5000 (29, 50, 51) and pJD1 (26), with about 30% identical amino acids through almost the entire regions of these proteins. We were, however, unable to find any sequence homology between the putative origin regions of these plasmids and the origins of ColE2-P9 and ColE3-CA38.

The *incB*/origin regions. The regions of these ColE2-related plasmids just next to the putative Rep coding regions are quite homologous to the origin regions of ColE2-P9 and ColE3-CA38 (Fig. 5). In particular, the nucleotide sequences of the right-half portions, where the primer RNA is synthesized, the leading-strand DNA synthesis starts, and the lagging-strand DNA synthesis terminates in ColE2-P9, are identical (68). Thus, it is strongly suggested that these regions are the origins of replication.

The nucleotide sequences of the left-half regions of the (putative) origin regions in plasmids of the same IncB specificity groups are identical, but those in plasmids of different IncB specificity groups are different. In ColE2-P9 and ColE3-CA38, a deletion or insertion of an A · T base pair in the middle (site α in Fig. 5) determines the specificity of interaction of the Rep proteins and the origins required for initiation of DNA replication (80). The origins of these two plasmids also differ by an insertion or deletion of a G · C base pair at the left end (site β). Combinations of the sites α and β from the same plasmids is important for the normal efficiency of interaction of the Rep proteins and the origins (19a). The origins of plasmids of the ColE3-CA38 and ColE5-099 groups differ only by a deletion or insertion of an A · T base pair and an adjacent G-to-A substitution (site γ). Thus, combinations of sites α, β, and γ seem to determine the specificity among plasmids of these three groups. The single base pair insertion or deletion at each of sites α, β, and γ in the origins apparently corresponds to the presence or absence of each of the amino acid sequences, insertions A, B, and C, respectively, in the Rep proteins. The left-half regions of the origins of plasmids of the ColE2-CA42 IncB specificity group are considerably different from those of other plasmids, which apparently corresponds to the differences in the amino acid sequences in the C-terminal regions of the Rep proteins. Such a correspondence might be the basis for determination of the specificity of the Rep protein-origin binding.

Two direct repeat sequences of 9 bp (5′-APyCAPuATAA-3′) in the origin regions are conserved among plasmids of the ColE2-P9, ColE3-CA38, and ColE5-099 IncB specificity groups, which are separated from each other by 5, 6, and 7 bp, respectively. The sequence might be the recognition and binding sequence for the Rep proteins. In plasmids of the ColE2-CA42 IncB specificity group, the sequence on the right side is identical to those of the other plasmids, while the sequence on the left side is considerably different.

The *incA* regions. The promoter sequences (the −35 and −10 regions) for the Rep mRNA and RNA I of ColE2-P9 and ColE3-CA38 are conserved in all the other plasmids (Fig. 2). The long inverted repeat sequences with an A stretch which specify the termination signal for transcription of RNA I are also structurally conserved. All of these features strongly suggest that small RNAs complementary to the 5′ untranslated regions of the Rep mRNAs are specified in these regions and that they are the *incA* gene products, regulating the expression of the *rep* genes, as shown for ColE2-P9 (62, 70, 81).

Stem-and-loop structures of antisense RNAs. The long stem-loop structure (structure I′) of RNA I which has been shown to be important for regulation of the *rep* gene expression in ColE2-P9 (62, 70) is structurally conserved in putative RNA I molecules of plasmids of the other four IncA specificity groups, whereas the shorter stem-loop structure (structure II′) is missing (Fig. 6). This might suggest that structures II′ are not essential for the regulation of expression of the *rep* gene.

The nine-nucleotide sequences (5′-UCUUGGCGG-3′) around the loop regions are conserved in all of the sequenced ColE2-related plasmids of the different IncA specificity groups except for plasmids of the ColE2-CA42 IncA specificity group, which contains a single base substitution (Fig. 6). For plasmids of the ColE5-099 and ColE9-J IncA specificity groups, the 22-nucleotide regions in and around the loop regions are identical, and for plasmids of the ColE2imm-K317 and ColE5-099 IncA specificity groups, the 18-nucleotide regions at the 5′ ends are identical. Nevertheless, supply of excess RNA I from the cloned *incA* regions did not raise the copy numbers of plasmids of the different IncA specificity types (data not shown), suggesting that the interaction of heterologous RNA I with the Rep mRNA, which could interfere with the homologous RNA I-Rep mRNA interaction, did not occur (71). A bulge loop and/or an interior loop is located in the stem region close to the loop portion. Such a structure might be required for a proper presentation of the loop sequences involved in the initial interaction of the two RNAs (18, 47), or instability of the stem might be required for its melting to provide additional nucleotides for pairing between the two RNAs after initial interaction at the loop tops (58). For ColE2-P9, some single base pair substitutions predicted to cause stabilization as well as those predicted to cause destabilization of the stem structure next to the loop portion of structure I′ drastically decreased the efficiency of RNA-RNA interaction (62) and the efficiency of inhibition of the *rep* gene expression by RNA I (70).

Many of the prokaryotic RNAs which are known or suggested to be antisense RNAs and/or their target RNAs have nucleotide sequences at the loop regions of the stem-loop structures similar to those of ColE2-related plasmids (Table 2). In particular, many of the antisense RNAs or their targets involved in regulation of replication of plasmids possess the sequence 5′-UUGGCG(G)-3′ or its complementary sequence at the loop regions. The initial interaction of the antisense RNAs and the target RNAs of all of these plasmids probably occurs between the identical complementary sequences, suggesting not only a common mechanism but also an evolutionary significance of the particular nucleotide sequences in RNA-RNA interaction.

A possible mechanism of inhibition of *rep* gene expression by antisense RNAs. The 5′-end portion of putative RNA I of each of these ColE2-related plasmids does not cover the putative initiation codon of the Rep mRNA and a 15-nucleotide region upstream of the initiation codon (Fig. 2). This

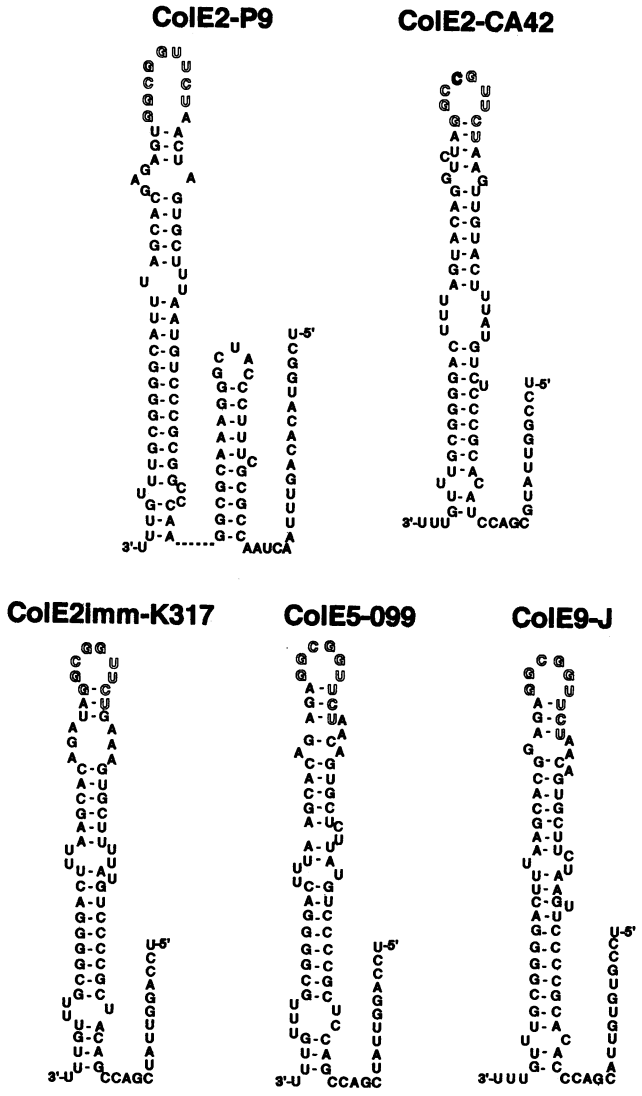


FIG. 6. Secondary structures of (putative) RNA I molecules of plasmids of the five IncA specificity types predicted by a method described elsewhere (73). The conserved sequences around the loop portions are shown by outlined letters, and the single nucleotide substitution in the loop portion of ColE2-CA42 type is shown by a boldface letter.

might suggest that inhibition of expression of the Rep protein by antisense RNA in these plasmids is not direct physical obstruction of the ribosome binding to the Shine-Dalgarno region and the initiation codon by double-stranded RNA. Thus, sequestration of additional sequence elements required for translation other than the initiation codon and Shine-Dalgarno region (37, 60) or disruption of secondary or tertiary structures required for translation (1, 2) as proposed for ColE2-P9 (81) might be also the mechanism of the inhibitory action of antisense RNAs in the other plasmids.

IncA function of ColE8-J and ColE2imm-K317. The incompatibility test with ColE8-J suggested that putative RNA I is active as an inhibitor but that the expression of the putative *rep* gene is only weakly sensitive to the action of RNA I (Table 1). Plasmid ColE2imm-K317 is only slowly excluded from the host bacteria by the cloned *incA* gene of its own (Table 1). We do not yet know whether RNA I is inactive or the expression of

TABLE 2. Comparison of nucleotide sequences around the loop portions of the stem-loop structures of various antisense RNA molecules or their target RNA molecules

RNA	Loop sequence[a]	Reference(s)
RNA I (ColE2 related[b])	UUGGCGG	80, this study
inc RNA target (IncI[c])	UUGGCGg	1, 41, 48
RNA I (ColE2-CA42 type)	UUGCCGg	This study
oop RNA (bacteriophage λ)	UUGCCGc	27
Antisense RNA (pJD1[d])	UUGGCGu	26
inc RNA target (IncF[e])	UUGGCGa	13, 34, 53, 54, 78
RNA I target (CloDF13)	UUGGGGA	61
RNA I (CloDF13)	UUCCCCA	61
RNA-OUT (Tn*10*)	UUCGCGa	6
RNA I target (CloDF13)	UUCGCAG	61
RNA I target (RSF1030)	uUGGAGG	59
Countertranscript RNA (pC221)	UUGGUCG	49
RNA I target (p15A)	UUGCAGg	57
RNA I (R1162)	UUGCAgg	24
Countertranscript RNA pC223	UUGGCUA	49
sar RNA target (P22)	UUGGAGa	33
RNA I (pT181)	UUGGUCA	28
RNA I target (ColE1)	uUCAGCA	44
RNA I (ColE1)	UUGGUAG	44
FinP RNA target (IncF)	UUGUGAG	15

[a] Bases shown in lowercase letters are revealed or predicted to be in the stem regions.
[b] ColE2-related plasmids other than those of ColE2-CA42 type.
[c] Various plasmids of IncI group and close relatives.
[d] From the nucleotide sequence (26), we propose that expression of the putative protein of pJD1 with some homology to ColE2-P9 Rep protein might be regulated by an antisense RNA, which is complementary to the 5′ portion of the mRNA containing the initiation codon and the Shine-Dalgarno region.
[e] Various plasmids of IncF group and close relatives.

the *rep* gene is insensitive. Further analyses of these two plasmids and comparison with the others might help elucidate the mechanism of inhibition of the *rep* gene expression by RNA I. Although the putative negative regulation of the *rep* gene expression by RNA I is apparently not effective, the copy numbers of these plasmids are still maintained at constant levels comparable to those of other plasmids (data not shown). Thus, there must be another mechanism for the copy number control of these plasmids.

The *cer*-like site-specific recombination sites. The *cer*-like sites of ColE2-P9 and ColE3-CA38 are well conserved among other ColE2-related plasmids (Fig. 2 and 7). The segments of these plasmids containing the putative *cer*-like sites are involved in stable maintenance of these plasmids by host cells

		Left arm	Overlap	Right arm	
ColE1	GC	GGTGCGTACAA	TTAAGGGA	TTATGGTAAAT	CC
ColE2-P9	GG	GGGGCGTACAA	--CGGGAG	TTATGGTAAAT	GG
ColE2-CA42	GG	GGTACGTACAA	--CGGGAG	TTATGGTAAAT	GG
ColE2-GEI602	GG	GGGGCGTACAA	--CGGGCG	TTATGGTAAAT	GG
ColE2imm-K317	GC	GGTGCGTTTAA	--CGGGAG	TTATGGTAAAT	GG
ColE3-CA38	GG	GGTGCGTACAA	--CGGGAG	TTATGGTAAAT	GG
ColE4-CT9	GG	GGTGCGTACAA	--CGGGAA	TTATGGTAAAT	GG
ColE5-099	GG	GGTACGTACAA	--CGGGAG	TTATGGTAAAT	GG
ColE6-CT14	GG	GGTGCGTACAA	--CGGGAG	TTATGGTAAAT	GG
ColE7-K317	GG	GGTGCGTACAA	--CGGGAG	TTATGGTAAAT	GG
ColE8-J	GG	GGTACGTACAA	--CGGGAA	TTATGGTAAAT	GG
ColE9-J	GG	GGTACGTACAA	--CAGGAG	TTATGGTAAAT	GG

FIG. 7. Nucleotide sequences of the putative *cer*-like site-specific recombination sites of the 11 ColE2-related plasmids. The core regions with the left and right arms (boxed) flanking the 6-bp overlap region as proposed previously (3) are shown. The core region of the ColE1 *cer* site is shown at the top. The left ends of the nucleotide sequences of ColE2-related plasmids correspond to positions 1 of those in Fig. 2.

and in resolution of plasmid multimers to monomers (22a, 80). For the *cer*-like sites of other systems, the core regions consisting of the right and left arms and the central overlap, where the host site-specific recombinase, XerC, acts, have been proposed (3, 10) (Fig. 7). The nucleotide sequence of the right arm in ColE1 plasmid is perfectly conserved in sequences of the ColE2-related plasmids and the putative left arms are similar to each other and to that of ColE1 plasmid, suggesting that XerC also acts on the *cer*-like sites of ColE2-related plasmids. The overlaps are almost identical among these ColE2-related plasmids except for single base pair substitutions in ColE4-CT9, ColE8-J, and ColE9-J but are different from those of other plasmids and *E. coli*. The XerC recombinase is of λ integrase type (3), and it has been shown that the recombinase of this type acts only on sites with a common overlap region (40). The 200-bp region next to the *cer*-like site with T and A clusters and the ArgR-binding site (along with the host ArgR protein as an accessory factor) are required for resolution selectivity (multimers to monomers) of the site-specific recombination at the *cer*-like sites in ColE1-like plasmids (63, 65). The 200-bp regions of ColE2-P9 and ColE3-CA38 (80) and several other ColE2-related plasmids (data not shown) to the left of the *cer*-like sites show similar characteristics. These regions of ColE2-related plasmids and the ArgR protein are also required for resolution selectivity (22a). All of these features strongly suggest that the same host site-specific recombination system is involved in stable maintenance of plasmids by host bacteria through resolution of multimers to monomers at the *cer*-like sites of these plasmids, just as in ColE1-like plasmids.

Evolution of ColE2-related replicons. All of these ColE2-related plasmids seem to have been derived from a common ancestor. We identified five IncA specificity groups and four IncB specificity groups among them (Table 1). Nine combinations (among 20 possible combinations) of the IncA and IncB specificity types were found among these plasmids, which could be the consequences of independent changes in the IncA and IncB specificity types. There could be, however, an alternative possibility.

On the basis of the pattern of distributions of the nucleotide sequence homology, the minimal replicon regions may be roughly divided into four subsegments (Fig. 8). Subsegments 1 (around positions 100 to 210 of ColE2-P9 in Fig. 2) containing the *incA* genes encoding (putative) RNA I molecules may be divided into five types, which exactly correspond to the five IncA specificity types. Subsegments 2 (around positions 210 to 770) containing the N-terminal portions of the (putative) Rep coding regions may be divided into two types. Subsegments 3 (around positions 770 to 940) containing the middle portions of the Rep coding regions are the most conserved. Subsegments 4 (around positions 940 to 1150) containing the C-terminal regions of the Rep coding regions and the (putative) origin regions may be divided into four types, which exactly correspond to the four IncB specificity types. Nine combinations of subsegments 1 and 4 are noted among these plasmids, which exactly correspond to the nine combinations of the IncA and IncB specificity types. Thus, we can see extensive mosaic structures in the replicon regions of these plasmids with abrupt transitions from homologous to nonhomologous regions or vice versa (Fig. 8). The mosaic structures might be seen in the entire genomes of these plasmids, as the patterns of combinations of the Inc specificity types in the replicon regions are apparently not linked to the species (ColE2 to E9) and types (DNase or RNase) of colicin proteins encoded by these plasmids (Table 1).

Somewhat similar structural organizations with alternating

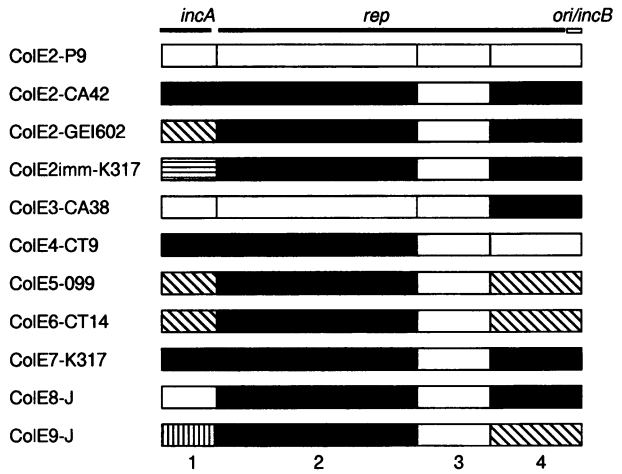


FIG. 8. Mosaic structures of the 11 ColE2-related replicons. The minimal replicon regions are divided into four segments according to the differences in distribution of sequence homology. For each segment, portions marked by the same pattern are almost identical in sequence and those marked by different patterns significantly differ from each other. Genes and sites identified in the minimal ColE2-P9 replicon are shown at the top.

homologous and nonhomologous segments have been found in the replicon regions of several ColE1-type plasmids (57, 82) and of a few plasmids of IncF and IncI groups (34, 48, 53, 54) and in the entire genomes of several lambdoid phages (see reference 5 for a review) and of several staphylococcal plasmids (49). It seems unlikely that all structures in each case have emerged only sequentially and/or divergently from a single ancestral structure as a result of accumulation of mutations. It has been postulated that some of these structures arose by exchanging functional parts of the genomes through homologous recombination and/or site-specific recombination. For ColE2-related plasmids, a similar mechanism seems to have contributed to the creation of new combinations of the IncA and IncB specificity types.

The chance of coexistence of two different ColE2-related plasmids in one bacterial cell in nature must be very low because of the rarity of close contact of two bacteria carrying different ColE2-related plasmids and because the resident plasmids exclude the entry of incoming plasmids. A homologous recombination event between two plasmids resulting in formation of a heterodimeric (composite) plasmid may also be rare in wild-type bacteria. Nevertheless, if these conditions are ever fulfilled, the second homologous recombination or the site-specific recombination at the *cer*-like sites should readily resolve the heterodimeric plasmid to produce the progeny recombinant plasmids. If these recombination events were to take place at appropriate positions, the progeny recombinant plasmids could be viable and carry new combinations of the IncA and IncB specificity types.

ColE2imm-K317 and ColE7-K317 (35, 75) and ColE8-J and ColE9-J (12), respectively, were originally carried by the same *E. coli* strains. We found that ColE6-CT14 (and also ColE6-Ind8) is a natural composite plasmid consisting of a ColE2-related replicon and another replicon. Natural composite plasmids were often found in many drug resistance plasmids, such as those of IncF groups (13). ColE3-CA38 and ColE6-CT14 carry additional immunity genes against the colicin E8 protein (7, 32). ColE9-J carries the *imm* gene and *hic* gene of ColE5-099 in addition to its own genes (8, 14), and a remnant

of a transposon-like element was found at the junction in ColE9-J (31). These might be examples of remnants or intermediary states during evolution. These ColE2-related plasmids may give us a system suitable for attempts, via in vitro evolution, to create new combinations of the IncA and IncB specificity types.

ACKNOWLEDGMENTS

We are grateful to H. Ogawa, T. Ogawa, and T. Horii for helpful discussions and continuing encouragement. We thank R. James, H. Masaki, T. Ogawa, A. Shinohara, and R. Watson for materials used in this study.

This work was supported in part by a grant-in-aid for scientific research from the Ministry of Education, Science and Culture of Japan to T.I.

REFERENCES

1. **Asano, K., A. Kato, H. Moriwaki, C. Hama, K. Shiba, and K. Mizobuchi.** 1991. Positive and negative regulation of plasmid ColIb-P9 *repZ* gene expression at the translational level. J. Biol. Chem. **266:**3774–3781.
2. **Asano, K., H. Moriwaki, and K. Mizobuchi.** 1991. An induced mRNA secondary structure enhanced *repZ* translation in plasmid ColIb-P9. J. Biol. Chem. **266:**24549–24556.
3. **Blakely, G., S. Colloms, G. May, M. Burke, and D. Sherratt.** 1991. *Escherichia coli* XerC recombinase is required for chromosomal segregation at cell division. New Biol. **3:**789–798.
4. **Bolivar, F., R. L. Rodriguez, P. P. Greene, M. C. Betlach, H. L. Heyneker, and H. W. Boyer.** 1977. Construction and characterization of new cloning vehicles. II. A multiple cloning system. Gene **2:**95–113.
5. **Campbell, A., and D. Botstein.** 1983. Evolution of the lambdoid phages, p. 365–380. *In* R. W. Hendrix, J. W. Roberts, F. W. Stahl, and R. A. Weisberg (ed.), Lambda II—1983. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
6. **Case, C. C., S. M. Roels, P. D. Jensen, J. Lee, N. Kleckner, and R. W. Simons.** 1989. The unusual stability of the IS*10* anti-sense RNA is critical for its function and is determined by the structure of its stem-domain. EMBO J. **8:**4297–4305.
7. **Chak, K.-F., and R. James.** 1984. Localization and characterization of a gene on the ColE3-CA38 plasmid that confers immunity to colicin E8. J. Gen. Microbiol. **130:**701–710.
8. **Chak, K.-F., and R. James.** 1986. Characterization of the ColE9-J plasmid and analysis of its genetic organization. J. Gen. Microbiol. **132:**61–71.
9. **Chan, P. T., H. Ohmori, J. Tomizawa, and J. Lebowitz.** 1985. Nucleotide sequence and gene organization of ColE1 DNA. J. Biol. Chem. **260:**8925–8935.
10. **Colloms, S. D., P. Sykora, G. Szatmari, and D. J. Sherratt.** 1990. Recombination at ColE1 *cer* requires the *Escherichia coli xerC* gene product, a member of the lambda integrase family of site-specific recombinases. J. Bacteriol. **172:**6973–6980.
11. **Cooper, P. C., F. K. L. Hawkins, and R. James.** 1986. Incompatibility between E colicin plasmids. J. Gen. Microbiol. **132:**1859–1862.
12. **Cooper, P. C., and R. James.** 1984. Two new E colicins, E8 and E9, produced by a strain of *Escherichia coli*. J. Gen. Microbiol. **130:**209–215.
13. **Couturier, M., F. Bex, P. L. Bergquist, and W. K. Maas.** 1988. Identification and classification of bacterial plasmids. Microbiol. Rev. **52:**375–395.
14. **Curtis, M. D., R. James, and A. Coddington.** 1989. An evolutionary relationship between the ColE5-099 and the ColE9-J plasmids revealed by nucleotide sequencing. J. Gen. Microbiol. **135:**2783–2788.
15. **Dempsey, W. B.** 1987. Transcript analysis of the plasmid R100 *traJ* and *finP* genes. Mol. Gen. Genet. **209:**533–544.
16. **Drummond, M., P. Whitty, and J. Wootton.** 1986. Sequence and domain relationships of *ntrC* and *nifA* from *Klebsiella pneumoniae*: homologies to other regulatory proteins. EMBO J. **5:**441–447.
17. **Frischauf, A., H. Lehrach, A. Poustka, and N. Murray.** 1983. Lambda replacement vectors carrying polylinker sequences. J. Mol. Biol. **170:**827–842.
18. **Hjalt, T., and E. G. H. Wagner.** 1992. The effect of loop size in antisense and target RNAs on the efficiency of antisense RNA control. Nucleic Acids Res. **20:**6723–6732.
19. **Horii, T., and T. Itoh.** 1988. Replication of ColE2 and ColE3 plasmids: the regions sufficient for autonomous replication. Mol. Gen. Genet. **212:**225–231.

19a. **Hosoi-Shinohara, M., and T. Itoh.** Unpublished data.

20. **Inselburg, J.** 1974. Incompatibility exhibited by colicin plasmids E1, E2, and E3 in *Escherichia coli*. J. Bacteriol. **119:**478–483.
21. **Itoh, T., and S. Hiraga.** 1992. Replicon evolution of ColE2-related plasmids, p. 397–411. *In* R. James, C. Lazdunski, and F. Pattus (ed.), Bacteriocins, microcins and lantibiotics. NATO ASI series, vol. H65. Springer-Verlag, Berlin.
22. **Itoh, T., and T. Horii.** 1989. Replication of ColE2 and ColE3 plasmids: in vitro replication dependent on plasmid-coded proteins. Mol. Gen. Genet. **219:**249–255.

22a. **Kantake, N., and T. Itoh.** Unpublished data.

23. **Kido, M., H. Yasueda, and T. Itoh.** 1991. Identification of a plasmid-coded protein required for initiation of ColE2 DNA replication. Nucleic Acids Res. **19:**2875–2880.
24. **Kim, K., and R. J. Meyer.** 1986. Copy-number of broad host range plasmid R1162 is regulated by a small RNA. Nucleic Acids Res. **14:**8027–8046.
25. **Kingsbury, D. T., and D. R. Helinski.** 1970. DNA polymerase as a requirement for the maintenance of the bacterial plasmid colicinogenic factor E1. Biochem. Biophys. Res. Commun. **41:**1538–1544.
26. **Korch, C., P. Hagblom, H. Öhman, M. Göransson, and S. Normark.** 1985. Cryptic plasmid of *Neisseria gonorrhoeae*: complete nucleotide sequence and genetic organization. J. Bacteriol. **163:**430–438.
27. **Krinke, L., and D. L. Wulff.** 1987. OOP RNA, produced from multicopy plasmid, inhibits λ cII gene expression through an RNaseIII-dependent mechanism. Genes Dev. **1:**1005–1013.
28. **Kumar, C. C., and R. P. Novick.** 1985. Plasmid pT181 replication is regulated by two countertranscripts. Proc. Natl. Acad. Sci. USA **82:**638–642.
29. **Labidi, A., E. Mardis, B. Roe, and R. J. Wallace, Jr.** 1992. Cloning and DNA sequence of the *Mycobacterium fortuitum* var. *fortuitum* plasmid pAL5000. Plasmid **27:**130–140.
30. **Lamerichs, R. M. J. N., A. Padilla, R. Boelens, R. Kaptein, G. Ottleben, H. Rüterjans, M. Granger-Schnarr, P. Oertel, and M. Schnarr.** 1989. The amino terminal domain of LexA repressor is α-helical but differs from canonical helix-turn-helix proteins: a two-dimensional ^{1}H NMR study. Proc. Natl. Acad. Sci. USA **86:**6863–6867.
31. **Lau, P. C. K., and J. A. Condie.** 1989. Nucleotide sequences from the colicin E5, E6 and E9 operons: presence of a degenerate transposon-like structure in the ColE9-J plasmid. Mol. Gen. Genet. **217:**269–277.
32. **Lawrence, G. M. P., and R. James.** 1984. Characterization of ColE8 plasmid, a new member of the group E colicin plasmid. Gene **29:**145–155.
33. **Liao, S.-M., and W. R. MacClure.** 1989. Regulation of gene expression by a small antisense RNA in bacteriophage P22, p. 289–297. *In* T. Cech (ed.), Molecular biology of RNA—1989. Alan R. Liss Inc., New York.
34. **Lopez, J., P. Crespo, J. Rodriguez, I. Andres, and J. M. Ortiz.** 1989. Analysis of IncF plasmids evolution: nucleotide sequence of an IncFIII replication region. Gene **78:**183–187.
35. **Males, B. M., and B. A. D. Stocker.** 1982. Colicins E4, E5, E6 and A and properties of *btuB*$^+$ colicinogenic transconjugants. J. Gen. Microbiol. **128:**95–106.
36. **Masaki, H., and T. Ohta.** 1985. Colicin E3 and its immunity genes. J. Mol. Biol. **182:**217–227.
37. **McCarthy, J. E. G., and C. Gualerzi.** 1990. Translational control of prokaryotic gene expression. Trends Genet. **6:**78–85.
38. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101:**126–133.
39. **Mock, M., and A. P. Pugsley.** 1982. The *btuB* group Col plasmids and homology between the colicins they encode. J. Bacteriol. **150:**1069–1076.

40. **Nash, H. A., and C. A. Robertson.** 1989. Heteroduplex substrates for bacteriophage lambda site-specific recombination: cleavage and strand transfer products. EMBO J. **8**:3523–3533.
41. **Nikoletti, S., P. Bird, J. Praszkier, and J. Pittard.** 1988. Analysis of the incompatibility determinants of I-complex plasmids. J. Bacteriol. **170**:1311–1318.
42. **North, A. K., K. E. Klose, K. M. Stedman, and S. Kustu.** 1993. Prokaryotic enhancer-binding proteins reflect eukaryotic-like modularity: the puzzle of nitrogen regulatory protein C. J. Bacteriol. **175**:4267–4273.
43. **Ogawa, T.** 1975. Analysis of the *dnaB* function of *E. coli* K-12 and *dnaB*-like function of P1 prophage. J. Mol. Biol. **94**:327–340.
44. **Ohmori, H., and J. Tomizawa.** 1979. Nucleotide sequence of the region required for maintenance of colicin E1 plasmid. Mol. Gen. Genet. **176**:161–170.
45. **Pabo, C. O., and R. T. Sauer.** 1984. Protein-DNA recognition. Annu. Rev. Biochem. **53**:293–321.
46. **Pansegrau, W., and E. Lanka.** 1992. A common sequence motif among prokaryotic DNA primases. Nucleic Acids Res. **20**:4931.
47. **Persson, C., E. G. H. Wagner, and K. Nordström.** 1990. Control of replication of plasmid R1: structures and sequences of the antisense RNA, CopA, required for its binding to the target, CopT. EMBO J. **9**:3767–3775.
48. **Praszkier, J., T. Wei, K. Siemering, and J. Pittard.** 1991. Comparative analysis of the replication regions of IncB, IncK, and IncZ plasmids. J. Bacteriol. **173**:2393–2397.
49. **Projan, S. J., and R. Novick.** 1988. Comparative analysis of five related staphylococcal plasmids. Plasmid **19**:203–221.
50. **Ranes, G. M., J. Rauzier, M. Lagranderie, M. Gheorghiu, and B. Gicquel.** 1990. Functional analysis of pAL5000, a plasmid from *Mycobacterium fortuitum*: construction of a "mini" *Mycobacterium-Escherichia coli* shuttle vector. J. Bacteriol. **172**:2793–2797.
51. **Rauzier, J., J. Moniz-Pereira, and B. Gicquel-Sanzey.** 1988. Complete nucleotide sequence of pAL5000, a plasmid from *Mycobacterium fortuitum*. Gene **71**:315–321.
52. **Ronson, C. W., P. M. Astwood, B. T. Nixon, and F. M. Ausubel.** 1987. Deduced products of C4-dicarboxylate transport regulatory genes of *Rhizobium leguminosarum* are homologous to nitrogen regulatory gene products. Nucleic Acids Res. **15**:7921–7934.
53. **Rosen, J., T. Ryder, H. Inokuchi, H. Ohtsubo, and E. Ohtsubo.** 1980. Genes and sites involved in replication and incompatibility of an R100 plasmid derivative based on nucleotide sequence analysis. Mol. Gen. Genet. **179**:527–537.
54. **Saadi, S., W. K. Maas, D. F. Hill, and P. L. Bergquist.** 1987. Nucleotide sequence analysis of RepFIC, a basic replicon present in IncFI plasmids P307 and F, and its relation to the RepA replicon of IncFII plasmids. J. Bacteriol. **169**:1836–1846.
55. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
56. **Sanger, F., S. Nickeln, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **75**:5463–5467.
57. **Selzer, G., T. Som, T. Itoh, and J. Tomizawa.** 1983. The origin of replication of plasmid p15A and comparative studies on the nucleotide sequences around the origin of related plasmids. Cell **32**:119–129.

57a. **Shinohara, A., and T. Ogawa.** Unpublished data.

58. **Siemering, K. R., J. Praszkier, and A. J. Pittard.** 1993. Interaction between the antisense and target RNAs involved in the regulation of IncB plasmid replication. J. Bacteriol. **175**:2895–2906.
59. **Som, T., and J. Tomizawa.** 1982. Replication origin of plasmid RSF1030. Mol. Gen. Genet. **187**:375–383.
60. **Sprengart, M. L., H. P. Fatscher, and E. Fuchs.** 1990. The initiation of translation in *E. coli*: apparent base pairing between the 16S rRNA and downstream sequences of the mRNA. Nucleic Acids Res. **18**:1719–1723.
61. **Stuitje, A. R., C. E. Spelt, E. Veltkamp, and H. J. J. Nijkamp.** 1981. Identification of mutations affecting replication control of plasmid CloDF13. Nature (London) **290**:264–267.
62. **Sugiyama, T., and T. Itoh.** 1993. Control of ColE2 DNA replication: *in vitro* binding of the antisense RNA to the Rep mRNA. Nucleic Acids Res. **21**:5972–5977.
63. **Summers, D. K.** 1989. Derivatives of ColE1 *cer* show altered topological specificity in site-specific recombination. EMBO J. **8**:309–315.
64. **Summers, D. K., and D. J. Sherratt.** 1988. Resolution of ColE1 dimers requires a DNA sequence implicated in the three dimensional organization of the cer site. EMBO J. **7**:851–858.
65. **Summers, D. K., S. Yaish, J. Archer, and D. J. Sherratt.** 1985. Multimer resolution system of ColE1 and ColK: localization of the cross over site. Mol. Gen. Genet. **201**:334–338.
66. **Tacon, W., and D. J. Sherratt.** 1976. ColE plasmid replication in DNA polymerase I-deficient strains of *Escherichia coli*. Mol. Gen. Genet. **147**:331–335.
67. **Tajima, Y., T. Horii, and T. Itoh.** 1988. Replication of ColE2 and ColE3 plasmids: two ColE2 incompatibility functions. Mol. Gen. Genet. **214**:451–455.
68. **Takechi, S., and T. Itoh.** 1994. Initiation of ColE2 plasmid DNA replication by a unique priming mechanism. Submitted for publication.
69. **Takechi, S., H. Matsui, and T. Itoh.** 1994. Primer RNA synthesis by plasmid-specified Rep protein. Submitted for publication.
70. **Takechi, S., H. Yasueda, and T. Itoh.** 1994. Control of ColE2 plasmid replication: regulation of the Rep protein expression by a plasmid-coded antisense RNA. Mol. Gen. Genet. **244**:49–56.
71. **Tomizawa, J.** 1985. Control of ColE1 plasmid replication: initial interaction of RNA I and the primer transcript is reversible. Cell **40**:527–535.
72. **Tomizawa, J., and T. Itoh.** 1981. Plasmid ColE1 incompatibility determined by interaction of RNA I with primer transcript. Proc. Natl. Acad. Sci. USA **78**:6096–6100.
73. **Turner, D. H., N. Sugimoto, and S. M. Frier.** 1988. RNA structure prediction. Annu. Rev. Biophys. Biophys. Chem. **17**:167–192.
74. **Vieira, J., and J. Messing.** 1987. Production of single-stranded plasmid DNA. Methods Enzymol. **153**:3–11.
75. **Watson, R., W. Rowsome, J. Tso, and L. P. Visentin.** 1981. Identification and characterization of Col plasmids from classical colicin E-producing strains. J. Bacteriol. **147**:569–577.
76. **Watson, R. J., T. Vernet, and L. P. Visentin.** 1985. Relationships of the Col plasmids E2, E3, E4, E5, E6 and E7: restriction mapping and colicin gene fusions. Plasmid **13**:205–210.
77. **Watson, R. J., T. Vernet, and L. P. Visentin.** 1985. Relationships of the Col plasmids E2, E3, E4, E5, E6 and E7: restriction mapping and colicin gene fusions. Plasmid **14**:97. (Erratum).
78. **Weber, P. C., and S. Palchaudhuri.** 1986. Incompatibility repressor in a RepA-like replicon of the IncFI plasmid ColV2-94. J. Bacteriol. **166**:1106–1112.
79. **Yanish-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.
80. **Yasueda, H., T. Horii, and T. Itoh.** 1989. Structural and functional organization of ColE2 and ColE3 replicons. Mol. Gen. Genet. **215**:209–216.
81. **Yasueda, H., S. Takechi, T. Sugiyama, and T. Itoh.** 1994. Control of ColE2 plasmid replication: negative regulation of the expression of the plasmid-specified initiator protein, Rep, at a posttranscriptional step. Mol. Gen. Genet. **244**:41–48.
82. **Zverev, V., and I. A. Khmel.** 1985. The nucleotide sequences of the replication origins of plasmids ColA and ColD. Plasmid **14**:192–199.