Presence of Different O Antigen Forms in Three Isolates of One Clone of *Escherichia coli*

Dan Liu and Peter R. Reeves

Department of Microbiology, The University of Sydney, Sydney, New South Wales 2006, Australia

Manuscript received January 3, 1994
Accepted for publication May 24, 1994

ABSTRACT

Escherichia coli strains ECOR2, ECOR3 and K-12 are very closely related in genotype as indicated by multilocus enzyme electrophoresis. We show that they have very different *rfb* regions indicating that recombination has occurred in this region, and we suggest that it may be associated with niche adaptation.

BACTERIAL populations have a clonal structure with long lived clones showing little variation (ACHTMAN *et al.* 1986; SELANDER *et al.* 1987). A set of 72 *Escherichia coli* isolates (OCHMAN and SELANDER 1984) is broadly representative of the genetic diversity in the species as detected by multilocus enzyme electrophoresis (MLEE) and ECOR2 and ECOR3 from this set have a close relationship with *E. coli* K-12 (R. K. SELANDER, personal communication). ECOR2 and *E. coli* K-12 have identical mobilities for the 35 enzymes used, and ECOR3 differed from *E. coli* K-12 only in the mobility of threonine dehydrogenase. In this study we present a comparison of the *rfb* regions of *E. coli* K-12, ECOR2 and ECOR3, and show that, despite their general similarity, their *rfb* regions are very different.

MATERIALS AND METHODS

Bacterial strains, plasmids and bacteriophage: The bacterial strains ECOR2 and ECOR3 are gifts of K. R. SELANDER: strain C600 (APPLEYARD 1953) was used as the K-12 strain. The allozyme profiles of the three strains were checked by electrophoretic typing of the enzymes 6-phosphogluconate dehydrogenase, malate dehydrogenase, acid phosphatase 1, acid phosphatase 2 and alcohol dehydrogenase, which allow ECOR2 and ECOR3 to be easily distinguished from most strains in the ECOR set (SELANDER *et al.* 1986). All three strains had the pattern expected for ECOR2 and ECOR3. A set of λ phage clones (KOHARA *et al.* 1987) covers the whole *E. coli* K-12 chromosome and five of the clones, K350, K351, K352 and K353, with insert from the 44-minute region, were used in this study (see Fig. 1). Phage were prepared as described by LIU and REEVES (1994). Bacterial strains and plasmids used in this study are listed in Table 1.

Enzymes and reagents: All enzymes and biochemicals were from Boehringer Mannheim or Pharmacia LKB. A λ packaging kit was from Promega, and [α-^{32}P]dCTP and [α-^{35}S]dATP were from Bresatec. Sequencing was done using a kit from U.S. Biochemical Corp., Cleveland, Ohio.

DNA methods and computer analysis: The preparations of chromosomal and plasmid DNA were as described by BASTIN *et al.* (1991), and Southern hybridization was as described by SAMBROOK *et al.* (1989). The strategies employed for cosmid cloning and restriction mapping were as described by TARTOF and HOBBS (1988). The dideoxy chain termination method (SANGER *et al.* 1977) was used for DNA sequencing. The DNA sequence was analyzed using the Australian National Genomic Information Service (ANGIS) (REISNER *et al.* 1993). The programs SEQH and SEQA (KANEHISHA 1982) were used for comparing the DNA sequences.

Methods for preparation of Lipopolysaccharide (LPS): Preparation of whole cell membranes, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and silver staining of LPS were carried out as previously described by BROWN *et al.* (1991).

RESULTS AND DISCUSSION

Relationship between *E. coli* K-12, ECOR2 and ECOR3 in the *rfb* regions: The three *E. coli* strains K-12, ECOR2 and ECOR3 are identical or near identical on MLEE (R. K. SELANDER, personal communication) and were expected to be members of a single clone and have the same O antigen, but surprisingly the O serotype of ECOR3 is O1, while that of ECOR2 is of a different unknown type (T. WHITTAM, personal communication). All extant strains of K-12 lack O antigen due to mutations in the *rfb* region, but it has recently been shown (LIU and REEVES 1994) that the parent K-12 strain had O16 antigen.

*Bam*HI, *Hin*dIII and *Eco*RI digestion of ECOR2, ECOR3 and C600 chromosomal DNA was probed with fragments from the Kohara clones covering the region from positions 2093 to 2147 (Figure 1). The regions from positions 2093 to 2103.7, and from positions 2124.2 to 2147 of *E. coli* K-12 were present in ECOR2. The regions from positions 2103.7 to 2111, and from positions 2116.2 to 2124.2 of K-12, hybridized to ECOR2 DNA, but the restriction fragment sizes were very different from those found with DNA from K-12 (data not shown). The plasmid pPR1168 (Table 1), which carries rhamnose pathway genes of K-12 (LIU and REEVES 1994), hybridized with ECOR2 and ECOR3 DNA, suggesting that the rhamnose pathway also exists in these strains. Probes A, B and C covering positions from 2111 to 2116.2, from the *rfb* region of K-12, did not hybridize with DNA from ECOR2 or ECOR3 (Figure 1). Southern hybridization of ECOR2 and ECOR3 DNA allowed us to determine the extent of similarity. The restriction maps

TABLE 1

Bacterial strains and plasmids

Strains and plasmids	Laboratory stock no.	Characteristic	Reference
C600	P334	*F⁻, thr-1, leuB6, tonA1, lacY1, tonA21, supE44, λ⁻ rfbD*	Appleyard (1953)
JM109	P3584	*supE44, nalA96, recA1, relA1, endA1, thi, hsdR17, Δ(pro-lac)*	Yanisch-Perron *et al.* (1985)
χ2819	P3851	*F⁻, lacY1, glnV44, galT22, ΔthyA57, recA56, metB1, hsdR2*, lysogenized with λc1857, b2, *red* β3,s7	Jacobs *et al.* (1986)
SØ874	P4052	*lacZ4503, try-355, upp-12, relA1, rpsL150, Δ(sbcB-rfb)86*	Neuhard and Thomassen (1976)
ECOR2	M134	Natural isolate	Ochman and Selander (1984)
ECOR3	M135	Natural isolate	Ochman and Selander (1984)
pPR1168		4.1-kb *Eco*RI-*Hin*dIII fragment from K352 in pUC18	Liu and Reeves (1994)
pPR1218		Functional *rfb* clone of ECOR2 with 40-kb insert in KT1	This study
pPR1249		9-kb *Hin*dIII fragment from pPR1218 in pUC18	This study
pPR1372		*Sal*I deletion derivative of pPR1218, with 12.7-kb insert	This study
pPR1376		*Hin*dIII deletion derivative of pPR1218, with 2.6-kb insert	This study
pPR1377		*Hin*dIII deletion derivative of pPR1218, with 5.3-kb insert	This study
KT1		High copy number cosmid	Tartof and Hobbs (1988)

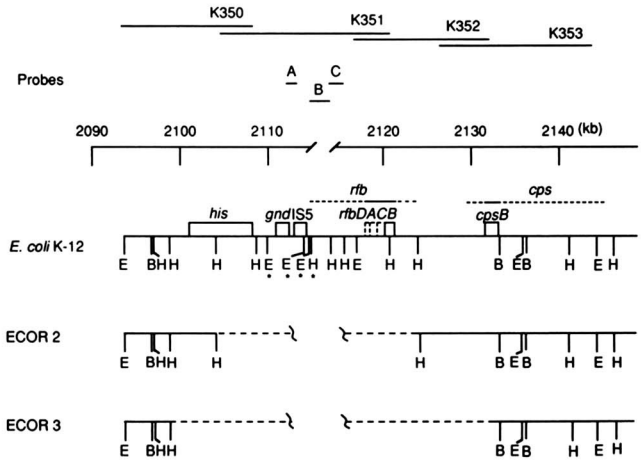

FIGURE 1.—Restriction maps of the *rfb* regions and comparison of *E. coli* K-12 with ECOR2 and ECOR3. The restriction map and scale are from the genomic physical map (Kohara *et al.* 1987) as modified by K. Rudd *et al.* (Bouffard *et al.* 1992; Rudd *et al.*, 1992). The scale is from EcoMap6 produced by the program GeneScape. The restriction sites with asterisk are not present in the physical map as published, but were found in Kohara clones in this study. The restriction maps of ECOR2 and ECOR3 were constructed by hybridization, and show only those restriction sites which are also in K-12. The regions in ECOR2 and ECOR3 which have DNA homology but differ in restriction sites compared with K-12 are indicated by dashed line, while the regions which have no homology with K-12 are shown as break. Kohara clones and probes are displayed above the restriction maps. Probe A was a polymerase chain reaction product of the ORF between the IS5 and the *gnd* gene (Liu and Reeves 1994) which is part of the last gene in *rfb* gene cluster of strain K-12, probes B and C were 2.2- and 1.6-kb *Hin*dIII fragments isolated from K351. Restriction sites: B, *Bam*HI; E, *Eco*RI and H, *Hin*dIII.

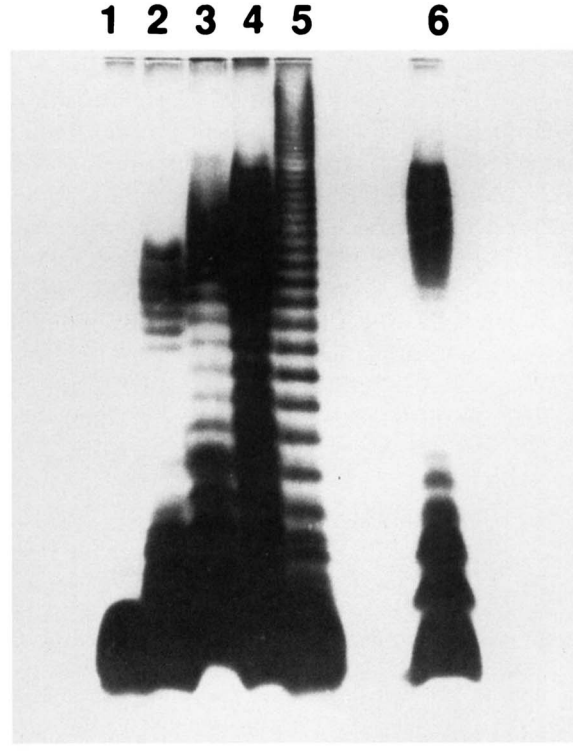

FIGURE 2.—Silver staining of LPS. The membrane extracts were electrophoresed in a 18% SDS-PAGE and silver-stained. Extracts are 1, *E. coli* K-12; 2, ECOR2; 3, pPR1218 in JM109; 4, pPR1372 in JM109; 5, pPR1372 in Sϕ874 (Δ*rfb* K-12 strain); 6, ECOR3.

of K-12 and ECOR2 are thus similar in the *his* and the *cps* regions, but different in the *gnd* and *rfb* regions (Figure 1). The restriction maps of ECOR3 and K-12 were only congruent in the *cps* region and the region beyond *his* (Figure 1). The DNA from the downstream part of the *rfb* region of *E. coli* K-12 did not hybridize to the chromosomal DNA of strains ECOR2 and ECOR3, indicating that the *rfb* regions of these three strains are substantially different.

Cloning the *rfb* region of ECOR2: The level of divergence in the *rfb* regions of the three closely related strains was surprising and to facilitate further restriction mapping and comparison we cloned the *rfb* region from strain ECOR2.

Chromosomal DNA from strain ECOR2 was partially digested with *Sau*3A and ligated into the *Bam*HI site in high copy number cosmid KT1 (Tartof and Hobbs

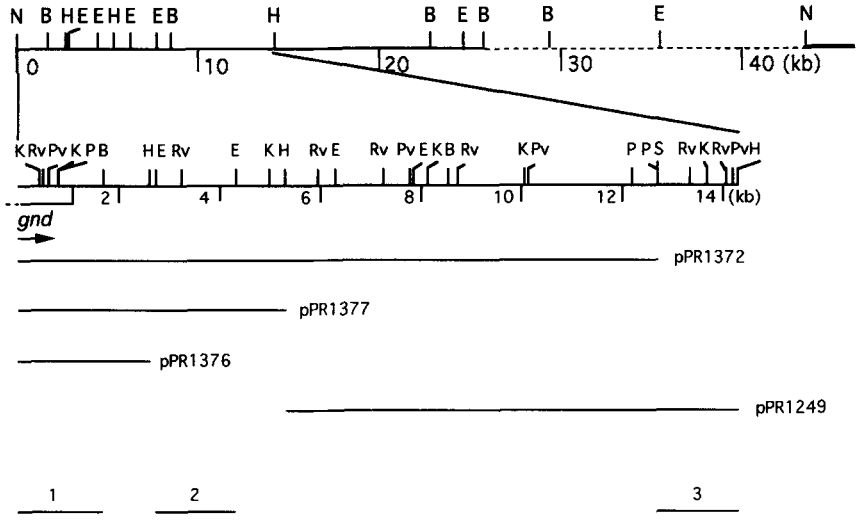

FIGURE 3.—The restriction map of ECOR2 clone pPR1218 and its derivatives. The detailed restriction map of the *rfb* region is also displayed. Part of the DNA from pPR1218 has a different restriction map compared to that constructed from Southern hybridization of chromosome of ECOR2 and is indicated by dash line. Plasmids pPR1372, pPR1376 and pPR1377 are deletion derivatives of pPR1218. Plasmid pPR1249 is in pUC18. Probes 1, 2 and 3 were isolated from pPR1376, pPR1377 and pPR1249, respectively. The position and direction of the spot sequencing of the *gnd* gene is indicated by the arrow line. Restriction sites: B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; N, *Not*I; Pv, *Pvu*II; Rv, *Eco*RV; and S, *Sal*I. The vector of pPR1218 is indicated by a heavy line.

1988). A bank of 2000 clones was made using *E. coli* strain χ2819 as host strain. Clones carrying *rfb* DNA were identified by Southern blotting using pPR1168 as the probe. The clones which hybridized to the probe were then tested for expression of LPS on SDS-PAGE. One showed smooth type LPS (Figure 2), and the plasmid was named pPR1218. The restriction map of pPR1218 (Figure 3) was determined by the method of TARTOF and HOBBS (1988) and confirmed by double digestion of derivatives of pPR1218 (Figure 3). Plasmid pPR1372, which is a *Sal*I deletion derivative of pPR1218, has a 12.7-kb insert, conferred expression of LPS on both JM109 and K-12 *rfb* delete strain SØ874 (Figure 2), showing that this DNA segment contains all the genes specific to O antigen biosynthesis.

Part of the restriction map of pPR1218, from positions 26 to 40, showed differences from the restriction map of ECOR2 chromosomal DNA (Figures 1 and 3). It is possible that two *Sau*3A fragments were ligated into the vector, but since this region is beyond the *rfb* gene cluster, this did not affect our study.

Location of the *gnd* gene of ECOR2: The *gnd* gene is located very close to the downstream end of the *rfb* gene clusters in *S. enterica* groups B (JIANG *et al.* 1991), C1 (LEE *et al.* 1992) and E1 (WANG *et al.* 1992) and in *E. coli* K-12 (LIU and REEVES 1994) and *E. coli* O7 (MAROLDA and VALVANO 1993), and marks the end of the *rfb* gene cluster. The *gnd* gene of ECOR2 was located precisely by sequencing (Figure 3). A 1.6-kb *Eco*RI (site on vector)-*Hin*dIII fragment from pPR1218 (position 0 to 1.6) was cloned into M13mp19, and the sequencing was carried out using universal primer. A 252-bp sequence was obtained (data not shown) and was homologous to the *gnd* sequence of *E. coli* K-12 from positions from 1263 to 1514 (NASOFF *et al.* 1984). There is 5.1% divergence in the *gnd* sequence between strains ECOR2 and *E. coli* K-12. This level of divergence is high for these closely related strains, indicating that the recombination event involved *gnd* and at least part of the *rfb* region in one of these two strains.

Hybridization using ECOR2 probe: Three probes derived from pPR1218 (Figure 3) were hybridized with ECOR2, ECOR3 and K-12 chromosomal DNA by Southern blotting. Probe 1 was a 1.7-kb *Eco*RI (site on vector)-*Bam*HI fragment which carries the *gnd* gene of ECOR2 (Figure 3): it hybridized to ECOR3 and *E. coli* K-12 chromosomal DNA. Probe 2 was a 1.6-kb *Eco*RI fragment carrying DNA from the *rfb* gene cluster and did not hybridize to ECOR3 or *E. coli* K-12 chromosomal DNA. Probe 3 carries DNA upstream of *rfb* and hybridized to the two strains (data not shown). These results showed again that the three strains from a clone of *E. coli* have three different *rfb* gene clusters and the regions flanking *rfb* appear to be similar within the three strains.

Conclusion: There have been previous reports of clones identified by lack of variation at many genes nonetheless having more than one O antigen type: ACHTMAN *et al.* (1986) observed O1, O18 and O2 O antigens in strains otherwise identical except at the *gnd* locus which is closely linked to *rfb*, and KAPUR *et al.* (1992) observed clones identical on similar criteria with both O2 and O78 represented. The three strains *E. coli* K-12, ECOR2 and ECOR3 have almost identical genotype as shown by MLEE of 35 enzymes (R. K. SELANDER, personal communication). However, not only do they have different O antigens, but the differences in their *rfb* gene clusters are extreme. It is clear that the differences cannot be due to mutation and that at least two of the three *rfb* regions must have been substituted by recombination during the life of the ECOR2/ECOR3/K-12 clone. Our observation of three different O antigens encoded by quite different *rfb* regions in three essentially randomly chosen isolates of a single widespread clone, together with the other observations discussed above, suggests that interclonal transfer of the *rfb* genes is much more frequent than transfer of most other genes.

The high level of polymorphism at the *rfb* locus with over 160 known forms of O antigen in *E. coli* (EWING 1986) indicates that some form of balanced selection operates. Both ØRSKOV and ØRSKOV (1976) and ACHTMAN and PLUSCHKE (1986), who observed that some O antigen forms are commonly present in *E. coli* clones with specific modes of pathogenesis, concluded that O antigen specificity is important in determining pathogenicity of clones; and one of us (REEVES 1992) has proposed a model for a form of balanced selection under which niche-specific selection for alleles of genes or gene clusters of significance in determining niche adaptation can maintain polymorphisms such as that found at the *rfb* locus.

The high frequency of transfer of *rfb* genes between clones may perhaps be due to selection for specific O antigen forms for adaptation to different niches, as O antigen variation in otherwise closely related strains can be associated with differences in disease specificity (ACHTMAN and PLUSCHKE 1986); however, in the absence of any specific details of the selection involved this remains speculative. Explanations for the high frequency of transfer based on higher intrinsic levels of recombination at the *rfb* locus do not seem plausible as the frequency appears to be much above that for other genes, but given that the G + C content of *rfb* regions is often atypical [see REEVES (1994) for review], it cannot be ruled out that this may affect recombination.

LITERATURE CITED

ACHTMAN, M., and G. PLUSCHKE, 1986 Clonal analysis of descent and virulence among selected *Escherichia coli*. Annu. Rev. Microbiol. **40:** 185–210.

ACHTMAN, M., M. HEUZENROEDER, B. KUSECEK, H. OCHMAN, D. CAUGANT *et al.*, 1986 Clonal analysis of *Escherichia coli* O2:K1 isolated from diseased humans and animals. Infect. Immun. **51:** 268–276.

APPLEYARD, R. K., 1953 Segregation of new lysogenic types during growth of a doubly lysogenic strain derived from *Escherichia coli* K-12. Genetics **39:** 440–452.

BASTIN, D. A., L. K. ROMANA and P. R. REEVES, 1991 Molecular cloning and expression in *Escherichia coli* K-12 of the *rfb* gene cluster determining the O antigen of an *E. coli* 0111 strain. Mol. Microbiol. **5:** 2223–2231.

BOUFFARD, G., J. OSTELL and K. E. RUDD, 1992 GeneScape: a relational database of *Escherichia coli* genomic map data for Macintosh computers. Comput. Appl. Biosci. **8:** 563–567.

BROWN, P. K., L. K. ROMANA and P. R. REEVES, 1991 Cloning of the *rfb* gene cluster of a group C2 *Salmonella* strain: comparison with *rfb* regions of groups B and D. Mol. Microbiol. **5:** 1873–1881.

EWING, W. H., 1986 *Edwards and Ewing's Identification of the Enterobacteriaceae.* Elsevier Science Publishers, Amsterdam.

JACOBS, W. R., J. F. BARRETT, J. E. CLARK-CURTISS and R. CURTISS, 1986 *In vivo* repackaging of recombinant cosmid molecules for analyses of *Salmonella typhimurium, Streptococcus mutans,* and mycobacterial genomic libraries. Infect. Immun. **52:** 101–109.

JIANG, X. M., B. NEAL, F. SANTIAGO, S. J. LEE, L. K. ROMANA *et al.*, 1991 Structure and sequence of the *rfb* (O antigen) gene cluster of *Salmonella* serovar typhimurium (strain LT2). Mol. Microbiol. **5:** 695–713.

KANEHISHA, M. J., 1982 Los Alamos sequence analysis package for nucleic acids and proteins. Nucleic Acids Res. **10:** 183–196.

KAPUR, V., D. G. WHITE, R. A. WILSON and T. S. WHITTAM, 1992 Outer membrane proteins mark clones of *Escherichia coli* O2 and O78 strains that cause avian septicemia. Infect. Immun. **60:** 1687–1691.

KOHARA, Y., K. AKIYAMA and K. ISONO, 1987 The physical map of the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. Cell **50:** 495–508.

LEE, S. J., L. K. ROMANA and P. R. REEVES, 1992 Sequence and structural analysis of the *rfb* (O antigen) gene cluster from a group C1 *Salmonella enterica* strain. J. Gen. Microbiol. **138:** 1843–1855.

LIU, D., and P. R. REEVES, 1994 *Escherichia coli* K-12 regains its O antigen. J. Gen. Microbiol. **140:** 49–57.

MAROLDA, C. L., and M. A. VALVANO, 1993 Identification, expression and DNA sequence of the GDP-mannose biosynthesis gene encoded by the O7 *rfb* gene cluster of strain VW187 (*Escherichia coli* O7:K1). J. Bacteriol. **175:** 148–158.

NASOFF, M. S., H. V. BAKER II and R. E. WOLF, JR., 1984 DNA sequence of the *Escherichia coli* gene, *gnd*, for 6-phosphogluconate dehydrogenase. Gene **27:** 253–264.

NEUHARD, J., and E. THOMASSEN, 1976 Altered deoxyribonucleotide pools in P2 educants of *Escherichia coli* K-12 due to deletion of the *dcd* gene. J. Bacteriol. **126:** 999–1001.

OCHMAN, H., and R. K. SELANDER, 1984 Standard reference strains of *Escherichia coli* from natural populations. J. Bacteriol. **157:** 690–693.

ØRSKOV, F., and I. ØRSKOV, 1976 Special *Escherichia coli* serotypes among enterotoxigenic strains from diarrhoea in adults and children. Med. Microbiol. Immunol. **162:** 73–80.

REEVES, P. R., 1992 Variation in O antigens, niche specific selection and bacterial populations. FEMS Microbiol. Lett. **100:** 509–516.

REEVES, P. R., 1994 Biosynthesis and assembly of lipopolysaccharide, pp. 281–314 in *Bacterial Cell Wall,* Vol. 27, edited by A. NEUBERGER and L. L. M. VAN DEENEN. Elsevier Science Publishers, Amsterdam.

REISNER, A. H., C. A. BUCHOLTZ, J. SMELT and S. MCNEIL, 1993 Australian national genetic information system, pp. 595–602 in *Proceedings of the 26th Annual Hawaii International Conference in Systems Sciences,* Vol. 1, edited by T. N. MUDGE, V. MILUTINOVIC and L. MUNTER. IEEE Computer Society Press, Los Alamos.

RUDD, K. E., G. BOUFFARD and W. MILLER, 1992 Computer analysis of *E. coli* restriction map, pp. 1–38 in *Genome Analysis,* Vol. 4, edited by K. E. DAVIES and S. M. TILGHMAN. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

SAMBROOK, J., E. F. FRITSCH and T. MANIATIS, 1989 *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

SANGER, F., S. NICKLEN and A. R. COULSON, 1977 DNA sequencing with chain termination inhibitors. Proc. Natl. Acad. Sci. USA **74:** 5463–5467.

SELANDER, R. K., D. A. CAUGANT, H. H. OCHMAN, J. M. MUSSER, M. N. GILMOUR *et al.*, 1986 Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. Appl. Environ. Microbiol. **51:** 873–884.

SELANDER, R. K., D. A. CAUGANT and T. S. WHITTAM, 1987 Genetic structure and variation in natural populations of *Escherichia coli,* pp. 1625–1648 in *Escherichia and Salmonella typhimurium: Cellular and Molecular Biology,* Vol. 2, edited by F. C. NEIDHARDT, J. L. INGRAHAM K. B. LOW, B. MAGASANIK, M. SCHAECTER *et al.* American Society for Microbiology, Washington, D.C.

TARTOF, K. D., and C. A. HOBBS, 1988 New cloning vectors and techniques for easy and rapid restriction mapping. Gene **67:** 169–182.

WANG, L., L. R. ROMANA and P. R. REEVES, 1992 Molecular analysis of a *Salmonella enterica* group E1 *rfb* gene cluster: O antigen and the genetic basis of the major polymorphism. Genetics **130:** 429–443.

YANISCH-PERRON, C., J. VIEIRA and J. MESSING, 1985 Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13 mp18 and pUC19 vectors. Gene **33:** 103–119.

Communicating editor: M. M. HOWE