

## Comparison of the Small 16S to 23S Intergenic Spacer Region (ISR) of the rRNA Operons of Some *Escherichia coli* Strains of the ECOR Collection and *E. coli* K-12

JESÚS GARCÍA-MARTÍNEZ,\* ANTONIO MARTÍNEZ-MURCIA, ANA ISABEL ANTÓN,  
AND FRANCISCO RODRÍGUEZ-VALERA

*Departamento de Genética y Microbiología, Universidad de Alicante, 03080 Alicante, Spain*

Received 21 February 1996/Accepted 27 August 1996

**Several 16S to 23S spacers of 354 bp have been sequenced from six *Escherichia coli* strains belonging to the ECOR collection. Four phylogenetically informative variable sites were identified. The results of their comparison confirm the existence of two major phylogenetic branches in this species, as previously reported. Remarkable intergenic heterogeneity was found in strain ECOR35 and its closest relatives, in which at least one of the operons has suffered a major mutagenic event or has an independent phylogenetic origin.**

rRNA operons are present in bacterial genomes in numbers varying from 1 to 11 (8, 13). In most cases they are located in the chromosome, although some examples of megaplasmid location are also known (20). In *Escherichia coli* K-12 there are seven copies per cell, located in the half of the chromosome closer to the replication origin (2, 12). They represent one of the best examples of redundant genes that are highly expressed in conditions of exponential growth. The genes for rRNA are normally located in the order 16S-23S-5S in bacteria. Between the rRNA genes there are usually short intergenic spacer regions (ISR) containing tRNA genes and some target sequences for RNase III and other recognition signals required for processing of the transcript. The seven rRNA operons of *E. coli* K-12 contain two types of 16S-23S ISR. Four of them (*rrnB*, *rrnC*, *rrnE*, and *rrnG*) contain a single tRNA<sup>Glu</sup>. The other three operons (*rrnA*, *rrnD*, and *rrnH*) have two genes for tRNA<sup>Leu</sup> and tRNA<sup>Ala</sup> (5).

Stretches of presumably nonfunctional DNA are also present and should exhibit a considerable degree of sequence variation by genetic drift (3). These sequence polymorphisms could make the spacer a fast molecular chronometer to measure short-term phylogeny, i.e., a good marker of major intraspecies lineages. Polymorphisms have been described, both between different rRNA loci in the same chromosome and between different strains (1, 6, 11, 13).

One of the most powerful tools for studying molecular evolution is the comparison of homologous sequences belonging to different representative strains. In a previous work, the spacers between genes coding for rRNA 16S and 23S genes coding for rRNA of a collection of uropathogenic *E. coli* (some of which were representatives of the ECOR collection) and other reference strains were amplified by PCR and subjected to restriction analysis. Analysis of the spacers' restriction patterns showed the presence of two markedly differentiated clusters that we named  $\alpha$  and  $\beta$  (7). In order to detect the sequence divergence between these two groups, we amplified by PCR 16S-23S ISR of six strains of the ECOR collection representing widely divergent groups of *E. coli*. Two strains (ECOR35 and

ECOR52) corresponded to the  $\alpha$  restriction fragment length polymorphism (RFLP) spacer group, and four strains (ECOR49, ECOR58, ECOR10, and ECOR44) corresponded to the  $\beta$  group. The sequences of K-12, which also belongs to the  $\beta$  group, were retrieved from the literature (10). In view of the known intergenic heterogeneity of these sequences, a minimum of five independent clones were sequenced for each strain.

Most of the strains used in this study were obtained from the American Type Culture Collection (ECOR10, ATCC 35329; ECOR35, ATCC 35354; ECOR44, ATCC 35363; ECOR49, ATCC 35368; ECOR52, ATCC 35371; and ECOR58, ATCC 35377); the K-12 strain was obtained from the Colección Española de Cultivos Tipo (CECT102). DNAs were prepared from cells grown in pure cultures with the InstaGene DNA purification matrix (Bio-Rad). PCR amplifications were carried out with *Taq* DNA polymerase (Promega) with oligonucleotides 16S14F (5'CTTGACACACCGCCCGTC3') and 23S1R (5'GGGTTTCCCCATTCGGAAATC3') as primers for the amplification of the ISR sequences; universal 16S oligonucleotide 16S0F (5'AGAGTTTGATCATGGCTCAG3') and SP $\alpha$  (5'GACTCGTAAGTCATCTTTAAC3') were used in combination to determine the presence of an unusual sequence located in the ISR of ECOR35. PCR conditions consisted of the following steps: a single cycle of denaturation at 94°C for 5 min; 35 cycles of 93°C for 1 min, 62°C for 1 min (60°C when we used the 16S0F-SP $\alpha$  combination), and 72°C for 2 min; and finally, a single step of 72°C for 10 min, which was executed to allow the completion of the amplified fragments.

Small ISR were purified from agarose gels with a GeneClean kit (Bio 101) and cloned into pUC18 with a SureClone kit (Pharmacia). Plasmid DNA was extracted with Wizard Mini-prep columns (Promega). Sequencing reactions were carried out according to the method described in reference 18 with T7 DNA polymerase (Pharmacia) and the same primers used in the amplification of the spacers. Alignment of sequences was made with the program Clustal (PC-Genie; IntelliGenetics).

Figure 1 shows the multiple alignment of the sequences found compared with those of *E. coli* K-12. Counting the 3 K-12 sequences, 14 different sequences were obtained from strains classified by RFLP as  $\beta$ , and 5 different sequences were obtained from strains classified as  $\alpha$ . The size of the small ISR was precisely 354 bp in all cases, differing from the large ISR,

\* Corresponding author. Present address: Institute of Molecular Evolutionary Genetics, Department of Biology, 518 Mueller Laboratory, Pennsylvania State University, University Park, PA 16802. Phone: (814) 863-4707. Fax: (814) 863-4706. Electronic mail address: jxg25@psuvm.psu.edu.

		<b>RsaI</b>											
<i>rrnC</i>	β	CCITTAAGAA	GCCTTCTTTG	CAGTGTCTCAC	ACAGATTGTC	TGATAGGAAG	TGAAAAGCAA	GGCGTCTTGC	GAAGCAGACT	GATACgtccc	cttcgtctag	100	
<i>rrnE</i>	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	100	
<i>rrnB2</i>	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	100	
EC10I(2)	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	100	
EC10II(3)	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	100	
EC44I(3)	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	100	
EC44II(4)	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	100	
EC49I	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	100	
EC49II(2)	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	100	
EC49III	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	100	
EC49IV	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	100	
EC58I(2)	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	100	
EC58II(2)	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	100	
EC58III	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	100	
EC35I(5)	α	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	100	
EC35II(3)	α	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	100	
EC52I(2)	α	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	100	
EC52II(2)	α	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	100	
EC52III	α	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	100	
		*****	*****	*****	*****	*****	*****	*****	*****	*****	*****		
		<b>tRNA<sup>Glu2</sup></b>											
<i>rrnC</i>	β	aggcccagga	caccgccctt	tcacggcggt	aacaggggtt	cgatccccct	aggggagccc	acTtGCTGGT	TtGTGAGTGA	AAGTCACTG	CCITAAATC	200	
<i>rrnE</i>	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	200	
<i>rrnB2</i>	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	200	
EC10I(2)	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	200	
EC10II(3)	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	200	
EC44I(3)	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	200	
EC44II(4)	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	200	
EC49I	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	200	
EC49II(2)	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	200	
EC49III	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	200	
EC49IV	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	200	
EC58I(2)	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	200	
EC58II(2)	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	200	
EC58III	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	200	
EC35I(5)	α	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	199	
EC35II(3)	α	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	199	
EC52I(2)	α	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	200	
EC52II(2)	α	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	200	
EC52III	α	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	200	
		<b>boxA</b>											
<i>rrnC</i>	β	TCAAACACTCA	TCTT	CGGGT	GATGTTTGAG	ATATTgctc	tttaaaaatc	TGGATCAAGC	TGAAAATTGA	AACACTGAAC	AACGAAAGTT	GTTCTGTGAGT	299
<i>rrnE</i>	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	299
<i>rrnB2</i>	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	299
EC10I(2)	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	299
EC10II(3)	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	299
EC44I(3)	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	299
EC44II(4)	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	299
EC49I	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	299
EC49II(2)	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	299
EC49III	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	299
EC49IV	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	299
EC58I(2)	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	299
EC58II(2)	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	299
EC58III	β	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	299
EC35I(5)	α	GTT...GATG	A...A...A...	C...A...	.....	.....	.....	.....	.....	.....	.....	.....	299
EC35II(3)	α	GTT...GATG	A...A...A...	C...A...	.....	.....	.....	.....	.....	.....	.....	.....	299
EC52I(2)	α	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	299
EC52II(2)	α	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	299
EC52III	α	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	299
		.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	
		<b>TaqI</b>											
<i>rrnC</i>	β	CTCTCAAATT	TTCGCAACAC	GATGATGAAT	<b>CGAAAGAAAC</b>	ATCTTCGGGT	TGTGA	354					
<i>rrnE</i>	β	.....	.....	.....	.....	.....	.....	354					
<i>rrnB2</i>	β	.....	.....	.....	.....	.....	.....	354					
EC10I(2)	β	.....	.....	.....	.....	.....	.....	354					
EC10II(3)	β	.....	.....	.....	.....	.....	.....	354					
EC44I(3)	β	.....	.....	.....	.....	.....	.....	354					
EC44II(4)	β	.....	.....	.....	.....	.....	.....	354					
EC49I	β	.....	.....	.....	.....	.....	.....	354					
EC49II(2)	β	.....	.....	.....	.....	.....	.....	354					
EC49III	β	.....	.....	.....	.....	.....	.....	354					
EC49IV	β	.....	.....	.....	.....	.....	.....	354					
EC58I(2)	β	.....	.....	.....	.....	.....	.....	354					
EC58II(2)	β	.....	.....	.....	.....	.....	.....	354					
EC58III	β	.....	.....	.....	.....	.....	.....	354					
EC35I(5)	α	.....	.....	.....	.....	.....	.....	354					
EC35II(3)	α	.....	.....	.....	.....	.....	.....	354					
EC52I(2)	α	.....	.....	.....	.....	.....	.....	354					
EC52II(2)	α	.....	.....	.....	.....	.....	.....	354					
EC52III	α	.....	.....	.....	.....	.....	.....	354					
		*****	*****	*****	*****	*****	*****						

FIG. 1. Comparison of 354-bp spacer sequences from six ECOR strains and K-12. The sequence of the ISR from the K-12 *rrnC* operon is used as a reference (10). Dots indicate identity of sequence with *rrnC*; lowercase letters define the conserved regions for tRNA<sup>Glu2</sup> and *boxA*, asterisks indicate conserved zones for the recognition of RNase III, and a dash marks a deletion in clones from strain ECOR35. Polymorphic sites for *RsaI* and *TaqI* are indicated in boldface characters. Numbers in parentheses indicate numbers of clones retrieved with identical sequences. α and β indicate the RFLP group to which the strain is assigned (7).

which are known to vary in length from 431 to 446 bp, with average values of 437 bp for those containing tRNA<sup>Ala</sup> and tRNA<sup>Ile</sup> (10, 15) and 440 bp for those containing tRNA<sup>Glu</sup> (10, 15, 21). The degree of sequence conservation found is remarkable, even in areas that have no known function. With the exception of clones from strain ECOR35, the numbers of nucleotide changes with regard to K-12 *rrnC* varied from one to six, with a mean nucleotide variation of approximately 1.07%. This value is smaller than the number of variations found when comparing different *E. coli* housekeeping genes (4, 16, 17, 19) and is similar to the numbers found for other spacers sequenced from different species (3, 9). Nevertheless, the level of sequence variation detected for these small ISR is lower than that inferred from the RFLP data (7), when all the ISR (large and small) are considered, with a mean value of about 2.5%, indicating that the large ISR may show a higher degree of variation. Moreover, the sequences of the small ISR did not show the *HinfI* polymorphism detected by RFLP (7), which is not necessarily surprising if we consider that it was found in very few strains from an uropathogenic *E. coli* collection, none of which belonged to the more representative ECOR strains.

There are four phylogenetically informative positions (i.e., they are present in more than two clones simultaneously) namely, 15, 21, 285, and 332. All except position 15 are located outside the stretches with known functions. At position 15 there is an A in all the  $\alpha$  strains and a T in all the  $\beta$  strains except in two clones (EC44I has an A and EC49IV has a G). This change generates a polymorphism for the enzyme *RsaI* that recognizes the sequence GTAC (the variation takes place in the A). *RsaI* is one of the informative enzymes used in the previous study in which  $\alpha$  and  $\beta$  groups were detected. At position 21 there is a consistency in sequence among clones from strains ECOR52 ( $\alpha$ ), K-12 ( $\beta$ ), and ECOR10 ( $\beta$ ) in that they share a C at that site, whereas clones from ECOR35 ( $\alpha$ ) and ECOR58 ( $\beta$ ) all have an A. In the rest of the clones from strains ECOR44 and ECOR49 (both  $\beta$ ), either an A or a C can be found. Position 285 also marks a clear difference among strains. Clones from ECOR52 ( $\alpha$ ) have a G, while those from ECOR35 ( $\alpha$ ) and most  $\beta$  clones have an A, except clones of K-12 *rrnB2* and clones EC58I and EC58III, which have a G. At position 332 strain ECOR52 has a C but ECOR35 and most  $\beta$  clones have an A, except for *rrnE* (K-12), EC10I with T, EC10II, and EC58III with C. This position generates the polymorphism *TaqI* used in our previous work (recognition sequence, TCGA; change occurs at A). These sequence distributions sustain the conjecture that the  $\alpha$  RFLP group is indeed a phylogenetic branch within the species *E. coli*, as previously suggested. The  $\beta$  group is probably much more heterogeneous and may include more than one entity of phylogenetic status  $\alpha$  (7).

Most of the variation within the  $\alpha$  group is concentrated in the two sequences found in the eight clones of the ISR amplified from strain ECOR35. These sequences have suffered a major variation between positions 200 and 220. Furthermore, generally speaking, they have more in common with the  $\beta$  than with the  $\alpha$  strains, except for the presence of the *RsaI* polymorphism used to discriminate between both groups. It is possible that the abnormal ISR or even the whole rRNA operons have been gained by horizontal genetic transfer from a phylogenetically distant strain. ECOR35 has been reported to contain important recombinatory events affecting the gene for malate dehydrogenase (19); ECOR35 may be a particularly promiscuous strain. With regard to the stretch of hypervariable sequence found between positions 200 and 220, among the larger ISR that have been sequenced in K-12, *rrnG* contains a 14-bp stretch totally different from the sequences of other

operons. This abnormal stretch is located between positions 294 and 307 that in a small ISR, having lost the *rsI* sequence, as reported for the *rrnB* (10), would correspond approximately to the positions in which most changes in ECOR35 are concentrated. Although there is no homology whatsoever between the two unusual sequences, the fact that both are located at equivalent sites in the ISR could indicate a relative lack of functional restriction, defining this zone as potentially highly variable. A PCR probe developed from this highly differentiated portion of ECOR35 clones showed that this sequence is found within the ECOR collection in a small cluster of group D strains: ECOR35, ECOR36, ECOR38, ECOR39, ECOR40, and ECOR41. None of the other strains of the ECOR collection gave a positive result with this probe, showing that the sequence of EC35I and EC35II is confined within this phylogenetic cluster. The close clustering by MLEE of the strains containing this sequence would be expected if the putative recombination event happened in the (relatively recent) common ancestor of this small cluster. In our previous work (7), when some strains ( $n = 10$ ) belonging to the RFLP cluster in which ECOR35 appeared were amplified with this primer, two gave a positive result (data not shown). The patchwork scheme found when comparing the ISR sequences would indicate that, although variations in ISR sequences are reliable enough to resolve ECOR strains into coherent groups, recombination may also play an important role in generating the diversity detected.

It is noteworthy that the vast majority of variations found for these sequences, both informative and noninformative, are concentrated in nonfunctional regions, which could be considered equivalent to the prevalence of synonymous substitutions (at the third position of the codons) detected for several genes (4, 14, 16, 17, 19). Furthermore, this fact indicates that *Taq* polymerase-induced mistakes (3) in the ISR sequences obtained were very few or nonexistent because of the nonrandomness of the unique site variation distribution along the 354 bp.

The interclonal sequence heterogeneity found when comparing different clones retrieved from the same strain is also remarkable. In fact, aside from the differences found between  $\alpha$  and  $\beta$  strains, there is almost as much variation among different operons of the same strain as among different strains, and this conclusion also applies to the three sequences of K-12 retrieved from databases. The number of different sequences found per strain is also remarkable, with a maximum of four for ECOR49. Additionally, one of the clones, EC49IV, shows six nucleotide differences with respect to EC49III, constituting a remarkable cistronic heterogeneity. In fact, from the K-12 information, the maximum number of rRNA operons containing small ISR is supposed to be 3. It is also possible that some of the larger operons (with the *rsI* sequence) lose it occasionally, changing to a small ISR, as has been detected in one of the subclones of K-12 (10). In any case, the evidence shown here indicates that rarely are identical operons found, even in the same genome. It is also clear that if ISR sequences are used for taxonomic or phylogenetic purposes, several clones should be sequenced from the same strain to include the interclonal variation in the comparison. The presence of an important variation in a sequence affecting several ISR from a well-defined cluster of strains from group D of MLEE would also need to be the subject of future research.

**Nucleotide sequence accession numbers.** Sequences were submitted to EMBL and GenBank and stored consecutively under the accession numbers U55296 to U55311.

This work was supported by grants PB 93/0930 of the DGICYT and BIO 93/0750 of the CICYT.

We thank R. K. Selander for providing the rest of ECOR strains used in this study. The secretarial assistance of K. Hernández is also gratefully acknowledged.

#### REFERENCES

1. Barry, T., G. Colleran, M. Glennon, L. K. Dunican, and F. Gannon. 1991. The 16S/23S ribosomal spacer region as a target for DNA probes to identify eubacteria. *PCR Methods Appl.* **1**:51–56.
2. Boros, I., A. Kiss, and P. Venetianer. 1979. Physical map of the seven ribosomal genes of *Escherichia coli*. *Nucleic Acids Res.* **5**:1817–1830.
3. Bourque, S. N., J. R. Valero, M. C. Lavoie, and R. C. Levesque. 1995. Comparative analysis of the 16S to 23S ribosomal intergenic spacer sequences of *Bacillus thuringiensis* strains and subspecies and of closely related species. *Appl. Environ. Microbiol.* **61**:1623–1626.
4. Boyd, E. F., K. Nelson, F.-S. Wang, T. S. Whittam, and R. K. Selander. 1994. Molecular genetic bases of allelic polymorphism in malate dehydrogenase (*mdh*) in natural populations of *Escherichia coli* and *Salmonella enterica*. *Proc. Natl. Acad. Sci. USA* **91**:1280–1284.
5. Condon, C., C. Squires, and C. L. Squires. 1995. Control of rRNA transcription in *Escherichia coli*. *Microbiol. Rev.* **59**:623–645.
6. Dolzani, L., E. Tonin, C. Lagatolla, and C. Monti-Bragadin. 1994. Typing of *Staphylococcus aureus* by amplification of the 16S-23S rRNA intergenic spacer sequences. *FEMS Microbiol. Lett.* **119**:167–174.
7. García-Martínez, J., A. J. Martínez-Murcia, F. Rodríguez-Valera, and A. Zorraquino. 1996. Molecular evidence supporting the existence of two major groups in uropathogenic *Escherichia coli*. *FEMS Immunol. Med. Microbiol.* **14**:231–244.
8. Gürtler, V. 1993. Typing of *Clostridium difficile* strains by PCR-amplification of variable length 16S-23S rDNA spacer regions. *J. Gen. Microbiol.* **139**:3089–3097.
9. Hall, L. M. 1994. Are point mutations or DNA rearrangements responsible for the restriction fragment length polymorphisms that are used to type bacteria? *Microbiology* **140**:197–204.
10. Harvey, S., C. W. Hill, C. Squires, and C. L. Squires. 1988. Loss of the spacer loop sequence from the *rnbB* operon in the *Escherichia coli* K-12 subline that bears the *relA1* mutation. *J. Bacteriol.* **170**:1235–1238.
11. Jensen, M. A., J. A. Webster, and N. Straus. 1993. Rapid identification of bacteria on the basis of polymerase chain reaction-amplified ribosomal DNA spacer polymorphisms. *Appl. Environ. Microbiol.* **59**:945–952.
12. Kiss, A., B. Sain, and P. Venetianer. 1977. The number of rRNA genes in *Escherichia coli*. *FEBS Lett.* **79**:77–79.
13. Kostman, J. R., T. D. Edlind, J. J. LiPuma, and T. L. Stull. 1992. Molecular epidemiology of *Pseudomonas cepacia* determined by polymerase chain reaction ribotyping. *J. Clin. Microbiol.* **30**:2084–2087.
14. Milkman, R., and I. R. Crawford. 1983. Clustered third-base substitutions among wild strains of *Escherichia coli*. *Science* **221**:378–380.
15. Nakayashiki, T., T. Adachi, T. Miki, and H. Inokuchi. 1992. Unpublished data.
16. Nelson, K., and R. K. Selander. 1992. Evolutionary genetics of the proline permease gene (*putP*) and the control region on the proline utilization operon in populations of *Salmonella* and *Escherichia coli*. *J. Bacteriol.* **174**:6886–6895.
17. Nelson, K., T. S. Whittam, and R. K. Selander. 1991. Nucleotide polymorphism and evolution in the glyceraldehyde-3-phosphate dehydrogenase gene (*gapA*). *Proc. Natl. Acad. Sci. USA* **88**:6667–6671.
18. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
19. Selander, R. K., J. Li, E. F. Boyd, F.-S. Wang, and K. Nelson. 1994. DNA sequence analysis of the genetic structure of populations of *Salmonella enterica* and *Escherichia coli*, p. 17–49. In F. G. Priest, A. Ramos-Cormentana, and B. J. Tindall (ed.), *Bacterial diversity and systematics*. Plenum Press, New York.
20. Suwanto, A., and S. Kaplan. 1992. Chromosome transfer in *Rhodobacter sphaeroides*: Hfr formation and genetic evidence for two unique circular chromosomes. *J. Bacteriol.* **174**:1135–1145.
21. Young, R. A., R. Macklis, and J. A. Steiz. 1979. Sequence of the 16S-23S spacer region in two ribosomal RNA operons of *Escherichia coli*. *J. Biol. Chem.* **254**:3264–3271.