Transduction, Restriction and Recombination Patterns in *Escherichia coli*

Melissa McKane and Roger Milkman

Department of Biological Sciences, The University of Iowa, Iowa City, Iowa 52242-1324

Manuscript received June 11, 1994
Accepted for publication September 24, 1994

ABSTRACT

Chromosomal DNA from several *Escherichia coli* reference (ECOR) strains was transduced by bacteriophage P1 into *E. coli* strain K12 W3110 *trpA33*. Recombination patterns of the transductants were determined by restriction fragment length polymorphism over a 40-kb region centering on a single marker (*trpA*$^+$) in the tryptophan operon. These experiments demonstrate that transduction between different strains of *E. coli* can result in recombinational replacements that are small in comparison to the entrant molecule (replacements average 8–14 kb, whereas P1 packages ~ 100 kb) often in a series of discrete segments. The transduction patterns generated resemble the natural mosaic sequence patterns of the ECOR strains described in previous work. Extensive polymorphisms in the restriction-modification systems of the ECOR strains are a possible explanation for the sequence patterns in nature. To test this possibility, two transductants were back-transduced into strain K12 W3110 *trpA33*. The resulting patterns were strikingly different from the original transductions. The size of the replacements was greater, and no multiple replacements were observed, suggesting a role for restriction-modification systems in the transduction patterns and perhaps for the mosaic sequence patterns in nature.

MUCH of our understanding of clonal structure and natural recombination in *Escherichia coli* comes from recent investigations (DUBOSE *et al.* 1988; DYKHUIZEN and GREEN 1991; MILKMAN and STOLTZFUS 1988; MILKMAN and MCKANE BRIDGES 1990, 1993; BISERCIC *et al.* 1991) using comparative sequence data of laboratory strain K12 and the 72 ECOR (*E. coli* reference) strains of recent natural origin (OCHMAN and SELANDER 1984) to help reconstruct the microevolutionary history of the *E. coli* chromosome. The compared sequences of some 37 natural strains of *E. coli* near the tryptophan operon (Figure 2 of MILKMAN and MCKANE BRIDGES 1993) show a mosaic pattern of similarities suggesting the result of recombination. These results, along with the restriction digest analysis over several regions of the chromosome (MILKMAN and MCKANE BRIDGES 1990), illustrate that groups of strains predominantly share a particular *sequence type* called a *clonal frame* (the portion of the ancestral chromosome that has not been replaced by extraclonal DNA). These groups are now referred to as *meroclones* by MILKMAN (1994), meaning that members of a group share extensive, but not genome-wide, ancestry. Sequence types differ from one another by $\geq$1.0% of their nucleotides (MILKMAN and MCKANE BRIDGES 1993). The groupings inferred by compared sequences in the *trp* region agree with the classifications derived from multilocus enzyme electrophoresis (MLEE) (SELANDER *et al.* 1987), which samples the genome. The "red," "purple" and "green" groups of MILKMAN and MCKANE BRIDGES (1993) (now the "K," "70" and "51" meroclones [MILKMAN 1994]) correspond respectively to the "A," "B1" and "B2" MLEE groups as classified by HERZER *et al.* (1990). Evidence of recombination is also seen in compared sequences of the *phoA* gene (DUBOSE *et al.* 1988), which demonstrate that a clonal frame is often interrupted by discrete discontinuous segments of DNA with different ancestral origins (MILKMAN and MCKANE BRIDGES 1993). These observed *clonal segments* average on the order of 1 kb in length; this is quite small considering that two of the main modes of genetic exchange in bacteria, conjugation and transduction, involve entrant molecules of $\geq$100 kb (WOOD 1968; STERNBERG 1990). Also of interest, MILKMAN and MCKANE BRIDGES (1993) observed that in some strains a string of discontinuous clonal segments can belong to the same sequence type. This observation is most striking in the strain ECOR 37, which is very different from K12 and yet has several small K12 type replacements in a 10-kb region (MILKMAN and MCKANE BRIDGES 1993; and unpublished data). The question arises whether the string of local replacements could have resulted from one event, because very frequent exchange between the strains seems unlikely. Specifically, the "cascade" hypothesis (MILKMAN and MCKANE BRIDGES 1993) suggests that large entrant molecules may be incorporated in several discrete segments.

This investigation was undertaken to relate the mosaic patterns observed in compared sequences with the products of recombination. The intent was to help determine the number of cellular events (conjugation or transduction) it might take to arrive at the observed replacement patterns. It was of particular interest to

Corresponding author: Melissa McKane, 138 Biology Bldg., Department of Biological Sciences, The University of Iowa, Iowa City, IA 52242-1324.

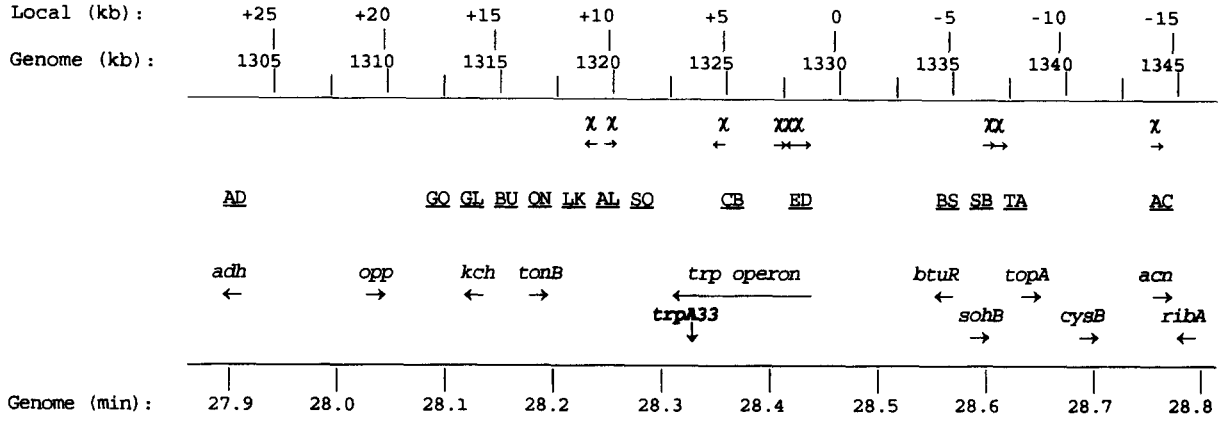

FIGURE 1.—Region map for transduction experiments. The local numbering system is with reference to the start of the *trp* operon. It is also the numbering system used in Figures 2–5. As integrated by RUDD (1992), the physical scale (KOHARA 1987) in kb pairs is also shown together with a normalized version of BACHMANN'S (1990) genetic scale in minutes. χ, known chi sites and their orientation (5′→3′) for the region 1302.5–1345.2 kb (81% of this region has been sequenced and could thus be screened for chi). Each of the two letter designations represents a PCR fragment, typically 1500 bp in length. *adh* . . . *ribA*, genes and their orientation. The site of the mutation in the recipient strain *trpA33* is indicated in bold.

test the cascade hypothesis. Clearly, analysis of single cellular events (transduction in this case) is a useful first step toward understanding the replacement patterns in nature.

The avenue of genetic transmission utilized in this study was generalized transduction by bacteriophage P1, which introduces into the cell an ~100-kb DNA molecule that is typically homologous to a chromosomal region of equal extent. Only one selectable marker (*trp*$^+$) was used to screen for transductants so that incorporation patterns would be typical; thus, there was no selection for multiple replacements or for replacements of a particular size. The region of the chromosome transduced (Figure 1) centers on the same region in which many sequences of the ECOR strains have been determined over a 12-kb continuous stretch (MILKMAN and MCKANE BRIDGES 1993, and unpublished data) using polymerase chain reaction (PCR) fragments. The sequence and restriction analyses of these strains provide a basis for analyzing the recombinants on a much finer scale than is easily possible with genetic markers alone.

The strains chosen for the experiments represent the main meroclones of the ECOR collection and one strain, ECOR 47, that is not identified with a meroclone. The recipient in all the experiments is K12 W3110 *trpA33*, which belongs to the K meroclone (MILKMAN 1995). Donor strains ECOR 27 and 65 represent the 70 and 51 meroclones, respectively. In the 10.4-kb region comparatively sequenced, the DNA of strain ECOR 47 differs an average of 2.1% in its nucleotides from K12 DNA. In this same region ECOR 27 DNA differs from K12 DNA in ~0.9% of its nucleotides, and ECOR 65 DNA differs from K12 DNA in 2.9% of its nucleotides. These averages are not uniformly distributed, because these sequences are mosaics and show local areas of more or less sequence dissimilarity. This is expected to be true for most of the genome, as evidenced by restriction analysis of several regions of the chromosome (MILKMAN and MCKANE BRIDGES 1990).

The findings of the transduction experiments indeed show that the resulting replacements, ranging from a few hundred to a few thousand base pairs, are much smaller than the entrant molecule. Also, a single transduction event often results in multiple replacements.

MATERIALS AND METHODS

Bacterial and phage strains: Strain *clr100* of bacteriophage P1 is temperate below 37° and contains a gene for chloramphenicol resistance, providing a way to select for lysogens (MILLER 1992). As a prophage, P1 exists as an extrachromosomal plasmid (IKEDA and TOMIZAWA 1969). During the lytic phase a small proportion of the phage heads will encapsulate ~100-kb pieces of the bacterial chromosome (STERNBERG 1990) in a nearly random manner (NEWMAN and MASTERS 1980).

The recipient strain used in the transductions is K12 W3110 *trpA33*, which will be usually referred to subsequently as K12 *trpA*. This strain has two consecutive nucleotide substitutions, either of which is sufficient to inactivate the *trpA* gene product. GAG→ATG replaces Glu49 with Met. The reversion of this mutation has never been known to occur (C. YANOFSKY, personal communication).

The standard ECOR collection (OCHMAN and SELANDER 1984) is a set of 72 *E. coli* strains of recent natural origin. The strains from this set that were used as donors in the transduction experiments are ECOR 27, 47 and 65.

Strain ER2437 was constructed by Dr. ELISABETH RALEIGH at New England Biolabs, starting with strain K12 W3110 *trpA33*. This derivative strain is greatly reduced in its restriction activities, as determined by a series of standard phage assays at New England Biolabs.

A K12 derivative strain R594 was used to titer P1 lysates, because P1 plaques are slightly larger and easier to see on this strain.

Lysates and transductions: Procedures were modified from

SILHAVY *et al.* (1984) and MILLER (1992). Donor strains were first made into lysogens. The strains were grown in Luria broth (LB) at 37° to an OD_{540} of 0.3 ($\sim 1 \times 10^8$ cells/ml). Cells were spun down and resuspended in one-half the volume of a 10 mM solution of $MgSO_4$. To 3 ml of molten L top agar, 0.1 ml of cell suspension, 0.1 ml of a 0.25 M $CaCl_2$ solution and 25 μl of a 34 mg/ml chloramphenicol solution were added, mixed and poured onto an L plate. Then 10 μl of P1 *clr100* lysate was spotted, and the plate was placed at 30° overnight. Because P1 *clr100* carries a gene for chloramphenicol resistance, only lysogens grew on the plates.

An overnight culture of lysogen was used to subculture (1:100) into 5 ml LB containing 10 mM $MgSO_4$ and 2 μg/ml chloramphenicol. Cultures were grown at 30° to an OD_{540} of 0.2. They were then placed at 43° with shaking for 20 min before transferring to 37° with further shaking until visible lysis occurred. Chloroform was added, the cells were pelleted and the supernatants were saved. Lysates were titered on *E. coli* strain R594.

The recipient strain (K12 *trpA*) was prepared for transduction by growth in 40 ml of LB at 37° to an OD_{540} of 0.2 ($\sim 7 \times 10^7$ cells/ml). The cells were spun down and resuspended in 0.5 ml LB containing 5 mM $CaCl_2$ and 10 mM $MgSO_4$. Typically, 10^8 cells were mixed with 10^7 phage. Control tubes with cells alone and with phage alone were also set up. The tubes were put at 37° for 20 min. Next, 5 ml of cold SSC solution (150 mM NaCl and 54 mM NaCitrate pH 7.0) was added, the tubes were spun at 2000 × *g* for 10 min and the cells were resuspended in 0.5 ml SSC solution. They were then plated on several "ZM" plates (minimal agar containing 0.5 mg/ml acid hydrolysate of casein, which contains no appreciable tryptophan) and placed at 42°. Transductant colonies were visible after ~20 h. Typical transduction frequencies averaged 7.5×10^{-7} for all crosses, and usually two transduction experiments were done to obtain a set of 15 transductants.

Plaque assay: Lysates were titered on each *E. coli* assay strain (ECOR 47, K12 *trpA* and ER2437) grown to an OD_{540} of 0.3 in LB containing 5 mM $CaCl_2$ and 10 mM $MgSO_4$. Phage suspensions were diluted, typically by 10^3 and 10^5, and 0.1 ml of the dilution was mixed with 0.2 ml of the respective assay suspension and 3 ml of molten LB top agar (containing 5 mM $CaCl_2$ and 10 mM $MgSO_4$) and plated on LB agar plates (also containing 5 mM $CaCl_2$ and 10 mM $MgSO_4$). Several platings were made and placed at 42°. Plaques were counted 12–15 h later. The ECOR 47 lysate was made as described in the previous section. The K12 *trpA* lysate was made by picking 100 P1 plaques from a K12 lawn infected by the ECOR 47 lysate. The plaques were vortexed in 500 μl LB and a few drops of chloroform, and then the agar and cells were centrifuged down. The supernatant was titered on the three strains as before.

DNA isolation: Transductants were streaked on ZM plates to isolate single colonies. Chromosomal DNA was isolated from transductants grown in LB by lysozyme and proteinase K treatment followed by phenol:chloroform:isoamyl alcohol (25:24:1) extraction and ethanol precipitation (SILHAVY *et al.* 1984).

PCR: Standard 100-μl PCR reactions were set up using ~0.5 μg of genomic DNA, 0.5 μM of each primer, 100 μM each dNTP, 2 units Taq polymerase and PCR buffer (10 mM Tris-HCl pH 8.3, 1.5 mM $MgCl_2$, 50 mM KCl). Reactions were placed in an automatic temperature cycler (Perkin-Elmer Cetus) and set for 27 cycles of 95° for 45 sec, 55° for 90 sec and 72° for 120 sec. Reactions were checked for quantity and purity by running 2 μl on a 1% agarose gel. Yields were ~1–1.5 μg per 100 μl reaction. Typically, 24-nucleotide primers were used, and the fragments generated were ~1.5 kb long. Primer sites and sequences are available from the authors.

Restriction digests: Typically, 150 ng of the unpurified 1.5-kb fragments were digested for 1 hr. The digests were run on a 5% polyacrylamide vertical gel at 37 V hr cm^{-1} to a standard dye position. Transductant digests were compared to previously interpreted digests of laboratory strain K12 and ECOR 27, 47 and 65 (MILKMAN and MCKANE BRIDGES 1990; and unpublished data).

From the restriction digest data the average donor segment length was roughly estimated in the following manner. An individual donor segment that was bounded on both sides was estimated to be the mean of the minimum and maximum lengths. For example, the donor segment for strain 47K-3 contained donor DNA from positions 7999 to 3780. It was therefore ≥4220 bp long. It was recipient-type at positions 9543 and 3703. It was therefore no longer than 5841 bp. Its length was estimated to be (4220 + 5841)/2, rounded to 5030 bp long. If an individual donor segment was bounded by recipient DNA on only one side, then its maximum length was estimated to be the minimum length plus 1/2 the average of all bounded segments for that group. This was done so to minimize the bias of the estimate for the unbounded segments. For example, strain 47K-4 was donor DNA from position 10238 to position −14338 (and maybe beyond). Its estimated length was [24,576 + 3111/2], rounded to 26,132 bp long. Then for each set of transductants, the estimated individual donor segment lengths were averaged.

Sequence determination: The PCR fragments were sequenced as previously described (MILKMAN and MCKANE BRIDGES 1993).

Notation: The phrase "transduction of donor strain X into recipient strain Y is represented as X→Y.

RESULTS

Restriction analysis of transductants: Several PCR fragments were generated for each transductant and analyzed by restriction endonucleases. Figures 2–4 summarize the restriction analyses of the transductants. In this 40-kb region all donor DNA segments are bounded by recipient DNA on at least one side, and most are bounded on both sides. Rough estimates of the size of these segments have been made (see MATERIALS AND METHODS). Donor DNA segments covering the *trp*$^-$ marker range from 1.5 to 30 kb in length. The overall average replacement size for the 47→K12 transduction is ~8 kb. Replacements in both the 27→K12 and the 65→K12 transductions average ~14 kb. Furthermore, multiple discrete donor segments were observed in several transductants: 8/18 transductants for 47→K12, 1/15 transductants for 27→K12 and 4/15 transductants for 65→K12. Although the sample size is too small for an adequate statistical test, it should still be noted that there is no apparent correlation between the number of multiple replacements observed and the percentage of sequence dissimilarity involved in each of the crosses, nor does the number of multiple replacements observed seem to be dependent on the number of sites available for analysis (36 sites in 47→K12, 28 sites in 27→K12 and 53 sites in 65→K12). It is evident from these experiments that relatively short incorporations

PCR fragment	A D	G O	G L	B U	O N	L K	A L	S Q		C B				E D				B S		S B		T A			A C	
RE[a]	H P	H P	R s	H 3	M s	S 3	S 3	M s	H 3	H 3	H 3	M s	H 3	H f	S 3	T q	•	B U	B U	B U	B U	B U	B U	M s	M s	M s	M s	H P	H P	H P	H P	M s	M s	M s	S 3	H P
Change K ↓ 47	T ↓ C	n o	C ↓ T	A ↓ C	n o	T ↓ G	n o	C ↓ T	T ↓ C	n o	n o	n o	n o	T ↓ C	n o		GA ↓ AT	T ↓ C	T ↓ G	C ↓ T	G ↓ A	C ↓ G	G ↓ A	G ↓ C	T ↓ C	A ↓ G	n o	n o	T ↓ C	n o	n o	n o	n o	A ↓ G	n o	A ↓ C
Position number	~2 6 0 5 0 0 0	2 5 5 0 0 0	1 6 0 5 3 5	1 5 4 3 7 4	1 4 3 3 2 9	1 3 3 3 0 7	1 2 3 9 0 2 4	1 2 2 4 7 5	1 2 1 3 3 7	1 1 3 3 7 8	1 1 0 8 7 7	1 1 0 8 7 8	1 0 2 5 3 8	9 5 4 3	7 9 1 9	7 1 2 0	6 3 5 0/1	5 1 6 1 8 4	4 1 7 5 8 6	4 1 7 5 8 8	3 1 7 8 0 0	3 1 7 6 0 3	3 1 7 6 9 4	1 1 9 2 5 7	1 1 2 5 5 5	1 1 5 6 7 2	1 5 7 8	-4 9 1 7	5 0 1 3 7	5 1 3 3 5	6 1 3 5 6	6 4 3 5 3	7 4 9 2 0	8 0 3 0 6	8 8 1 1	1 3 9 7 6	1 4 3 3 8
Strain																																				
47K-1	-	-	-	-	-	-	-	-	-	-	-	-	-	D	D	D	**D**	D	D	D	D	D	D	D	D	D	D	-	-	-	-	-	-	-	D	-
47K-2	-	D	-	-	-	D	D	D	-	-	-	-	D	D	D	D	**D**	D	D	D	D	D	D	-	-	-	-	-	-	-	-	-	-	-	-	-
47K-3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	D	D	**D**	D	D	D	D	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
47K-4	-	-	-	-	-	-	-	-	-	-	-	-	-	D	D	D	**D**	D	D	D	D	D	D	D	D	D	D	D	D	-	D	D	D	D	D	D
47K-5	-	-	-	-	-	-	-	-	-	D	D	-	-	-	-	-	**D**	D	D	D	D	D	D	D	D	D	D	-	-	-	-	-	-	-	-	-
47K-6	D	D	D	D	D	D	-	-	-	-	-	-	D	D	D	D	**D**	D	D	D	D	D	D	-	-	-	-	-	-	-	-	-	-	-	-	-
47K-7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**D**	D	D	D	D	D	D	D	D	D	D	D	D	Δ[b]	-	-	-	-	-	-
47K-8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**D**	D	D	D	D	D	D	-	-	-	-	D	-	D	D	D	D	D	D	D
47K-9	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	**D**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
47K-10	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	D	**D**	-	D	D	D	-	-	D	-	-	-	-	-	-	-	D	D	D	-	-
47K-11	-	-	-	-	-	-	-	-	-	-	-	-	-	-	D	D	**D**	D	D	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
47K-12	-	-	D	D	D	D	D	D	D	D	D	D	D	-	D	D	**D**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
47K-13	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	D	**D**	D	D	D	D	D	D	D	D	D	D	-	-	-	-	-	-	-	-	-
47K-14	-	-	D	-	-	-	-	-	-	-	D	D	D	D	D	D	**D**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
47K-15	-	-	D	D	D	D	D	D	D	D	D	D	D	D	D	D	**D**	D	D	D	D	D	D	D	D	D	D	-	-	-	-	-	-	-	-	-
47K-16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	D	**D**	D	D	D	D	D	D	-	-	-	-	-	-	-	-	-	-	-	-	-
47K-17	-	-	D	D	D	D	D	D	D	D	D	D	D	D	D	D	**D**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
47K-18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	D	D	**D**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

FIGURE 2.—47→K12 transductants (restriction analysis). Polymorphic restriction sites are designated as either donor DNA (D) or recipient DNA (-). The inferred nucleotide substitution is given for the gain of a restriction site (in comparison to the recipient), and a vertical "no" indicates the loss of a site. The selected marker for the transductants is indicated in bold. Position numbers are in reference to the tryptophan operon, and positions written in italics are negative. [a] Restriction enzymes: HP, *Hin*PI; Rs, *Rsa*I; H3, *Hae*III; Hf, *Hin*fI; Ms, *Ms*PI; S3, *Sau*3AI; Tq, *Taq*I; BU, *Bst*UI. [b] Deletion from 5135–5223 that is not present in either recipient or donor.

can often result from a single transduction event. Moreover, multiple incorporations are not rare.

Restriction analysis of back transductants: Two transductant strains that had extensive and evidently continuous donor DNA were chosen for the back crosses: 47K-9, which has ECOR 47 DNA extending ~20 kb from PCR fragment AD to the selected marker, and 47K-4, which has ECOR 47 DNA extending ~25 kb from PCR fragment LK to fragment AH. All 15 transductants of the 47K-9→K12 back transduction were identical to the 47K-9 donor strain (which is contained in Figure 2). The results of the 47K-4→K12 back transduction can be seen in Figure 5. No multiple incorporations were found for either of the back transduction experiments. In addition, only two of the 30 total transductants have discernible donor/recipient borders: 4B-2 has a border on one side, and 4B-3 has borders on both sides, making a donor segment that is 9–15 kb long. The average lengths of the backcross replacements are much greater than those of the original transductants. This, plus the absence of multiple discrete donor segments in the back transductants, implies that restriction played an important role in creating the observed recombination patterns in the original cross. (The question has been raised as to whether mismatch repair activity could have acted in the original transduction and not in the back transductions, thus accounting for the different patterns observed. However, we can see no clear mechanism, either with long-patch mismatch repair or with short-patch mismatch repair [see RADMAN 1988 for review], that would act with such disparity between entering molecules containing an 80–100-kb stretch of sequence dissimilarity and those containing a 20–25-kb stretch of sequence dissimilarity.)

Phage-plating efficiency: The early definition of *restriction* came from the observation that the growth of a bacteriophage on one bacterial host sometimes greatly reduced its plating efficiency on a second bacterial host (LURIA 1953). A test of the ability of P1 *clr100* lysates to form plaques on strains ECOR 47, K12 *trpA* and ER2437 was used to confirm that restriction-modification differences exist between the strains and to further indicate that restriction played a role in the recombination patterns observed for the ECOR 47→K12 *trpA* transduction. Table 1 shows that P1 phage grown on strain ECOR 47 plates over 90 times more efficiently on strain ER2437

PCR fragment	A D	G L	B U	O N	L K	A L	S Q	•	C B	E D	B S	S B	T A	A C
RE[a]	H P	B U	M s	D d	H 3	M s	H f	S 3	H 3	H 3	A u	A s	•	T q	S 3	B U	S 3	S 3	M s	A u	A u	M s	B U	B U	H P	H P	H P	B U	H P
Change K ↓ 27	n o	n o	A ↓ C	n o	n o	n o	n o	G ↓ A	n o	T ↓ G	G ↓ A	n o	GA ↓ AT	T ↓ C	A ↓ T	C ↓ T	C ↓ T	G ↓ A	G ↓ C	C ↓ G	n o	A ↓ G	N E W	n o	n o	n o	n o	n o	n o
Position number[b]	~2	1	1	1	1	1	1																						
	6	5	4	3	2	1	0	9	9	9	8	8	6	4	4	4	3	3	1	1	1	1	-4	5	6	6	6	8	4
	0	4	3	0	4	9	8	6	2	0	4	1	3	8	6	1	9	6	9	6	4	1	9	7	2	4	7	9	7
	0	8	3	1	7	5	9	0	2	4	3	2	5	0	5	5	9	9	2	7	0	6	6	7	2	3	2	2	6
	0	0	7	4	3	0	4	6	5	0	4	0	0/1	9	6	8	9	4	7	5	9	2	4	0	6	3	6	5	9
Strain																													
27K-1	-	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	-	-	-	-	-	-	-	-	-	-	-
27K-2	-	-	-	-	-	-	-	-	-	-	-	-	D	D	D	D	D	D	D	D	D	D	-	-	-	-	-	-	-
27K-3	-	-	-	-	-	-	-	-	-	-	-	-	D	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
27K-4	-	-	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	-	-	-	-	-	-	-	-	-	-	-
27K-5	-	-	-	-	-	-	-	-	-	-	D	D	D	D	D	D	D	D	D	D	D	D	-	-	-	-	-	-	-
27K-6	-	-	-	-	-	-	-	-	-	-	-	-	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	-
27K-7	-	-	-	-	-	-	-	-	-	-	-	D	D	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
27K-8	-	-	-	-	-	-	-	-	-	-	-	-	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D
27K-9	-	-	-	-	-	D	D	D	D	D	D	D	D	D	D	D	D	D	-	-	-	-	-	-	-	-	-	-	-
27K-10	-	-	-	-	-	-	-	-	-	-	-	-	D	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
27K-11	-	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D
27K-12	-	-	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	-	-
27K-13	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	-	-	-	-	-	-	-	-	-	-	-	-
27K-14	-	D	D	D	D	D	D	D	D	D	D	D	D	-	-	-	-	-	D	D	D	D	D	D	D	D	D	-	-
27K-15	-	D	D	D	D	D	D	D	D	D	D	D	D	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

FIGURE 3.—27→K12 transductants (restriction analysis). Same format as Figure 2. [a] Restriction enzymes: HP, *Hin*PI; BU, *Bst*UI; Ms, *Msp*I; Dd, *Dde*I; H3, *Hae*III; Hf, *Hin*fI; S3, *Sau*3AI; Au, *Alu*I; As, *Ase*I; Tq, *Taq*I.

than on K12 *trpA*. As mentioned in MATERIALS AND METHODS, ER2437 was derived directly from K12 *trpA* and has had its restriction activities greatly reduced. The greater plating efficiency on ER2437 suggests that restriction systems in K12 *trpA* are acting on the phage DNA produced in strain ECOR 47.

Nevertheless, there may be other factors involved in the reduction in phage infectivity besides endonucleases. The greater plating efficiency of the ECOR 47 lysate on strain ER2437 than on strain ECOR 47 suggests that something other than restriction (probably adsorption) is affecting the ability of P1 phage to infect strain ECOR 47 (because a strain would not restrict its own DNA).

As a follow up, phages from plaques on a K12 *trpA* lawn infected with ECOR 47 lysate were used to infect the three strains again. After one passage through strain K12 *trpA*, the phage now infected both K12 *trpA* and ER2437 almost equally well. These combined results (Table 1) are consistent with the view that K12 has at least one R-M system that ECOR 47 lacks. The very low plating efficiency ratio of ECOR 47 compared with ER2437 suggests that there may also be at least one R-M system in ECOR 47 that K12 lacks, though the interpretation is uncertain because of ECOR 47's adsorption properties previously mentioned. More extensive experiments using a variety of ECOR strains both in plaque assays and in transductions should help to clarify these issues.

Sequence analysis: The nucleotide sequence was determined for several of the 47→K12 transductants, concentrating mainly on regions where a border between donor-type DNA and recipient-type DNA was indicated by the restriction analyses. The total 30 kb of sequence information obtained (data not shown) is informative in two important ways. First, the sequences serve to refine the borders within the limits of sequence diversity, indicating that the transitions from donor-type to recipient-type DNA are complete (meaning that frequent alternations between donor-type DNA and recipient-type DNA at the ends of long donor segments were not observed). For the 29 borders sequenced the breakpoints were estimated to occur within 145 ± 125 bp (mean ± SD). Second, the sequence data indicate, within the limits of resolution, that recombination does not promote nucleotide substitutions, because not a single nonparental base pair was observed; one deletion was noted, however (see Figure 2). In addition, the thoroughness of the restriction analysis is demonstrated, because the sequence analysis did not reveal any donor- or recipient-type DNA segment that had not already been indicated.

DISCUSSION

Transduction patterns: The transduction experiments of this project are a step toward understanding the recombinational replacement patterns observed in

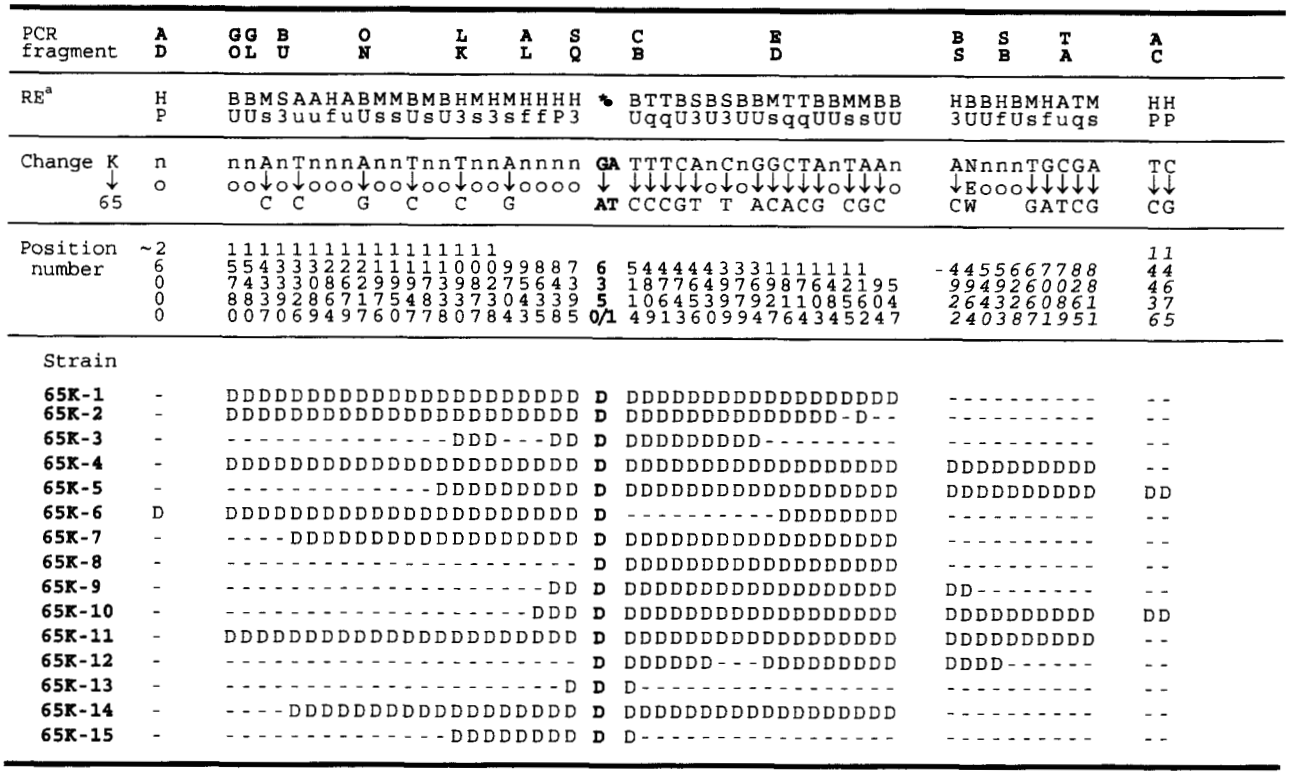

FIGURE 4.—65→K12 transductants (restriction analysis). Same format as Figure 2. [a] Restriction enzymes: HP, *Hin*PI; BU, *Bst*UI; Ms, *Msp*I; S3, *Sau*3AI; Au, *Alu*I; Hf, *Hin*fI; H3, *Hae*III; Tq, *Taq*I.

nature, because transduction is a likely mechanism of genetic exchange in nature (GOLDBERG *et al.* 1974; ROBESON *et al.* 1980). The experiments used *E. coli* strains that are representative of the natural population. Three of the strains employed, K12 W3110, ECOR 27 and ECOR 65, were from each of the three major meroclones of *E. coli* (the clonal frame groups of MILKMAN and MCKANE BRIDGES [1993]), and one strain, ECOR 47, was of a less common sequence type (MILKMAN and MCKANE BRIDGES 1993).

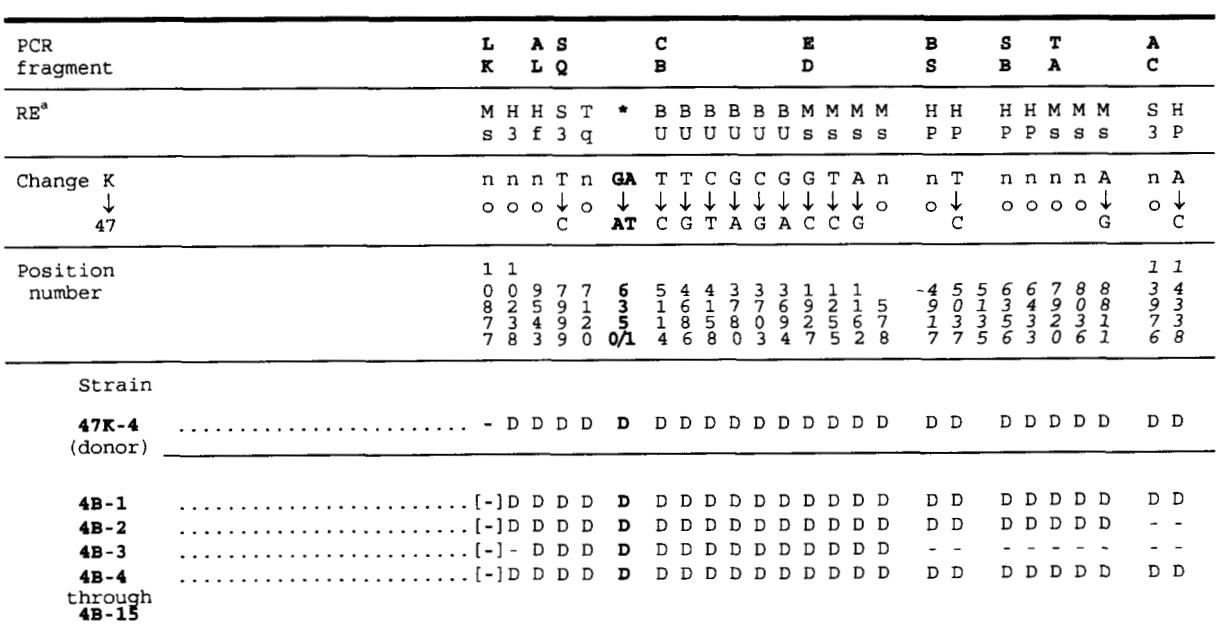

FIGURE 5.—47K back-transductants (restriction analysis). Same format as is used for Figure 2. [a] Restriction enzymes: Ms, *Msp*I; H3, *Hae*III; Hf, *Hin*fI; S3, *Sau*3AI; Tq, *Taq*I; BU, *Bst*UI. [-], this position is K12 in *both* donor (47K-4) and recipient.

TABLE 1

Titration on strains ECOR 47, K12 *trpA* and ER2437 of P1 phage lysates made on strains ECOR 47 and K12 *trpA*

	Titer (PFU/ml)		
	ECOR 47	K12 *trpA33*	ER2437
ECOR 47 lysate[a]	$(9.6 \pm 0.3_{S.E.M.}) \times 10^5$	$(1.9 \pm 0.1_{S.E.M.}) \times 10^5$	$(1.7 \pm 0.4_{S.E.M.}) \times 10^7$
K12 *trpA* lysate[b]	$(4.9 \pm 0.4_{S.E.M.}) \times 10^5$	$(4.2 \pm 0.1_{S.E.M.}) \times 10^8$	$(5.4 \pm 0.1_{S.E.M.}) \times 10^8$

[a] The 47 lysate was diluted to 10^{-3}, and 0.1 ml of the dilution was plated 15 times for each strain. For ER2437 plaques were counted on 1/11.5 of the total area of each plate.

[b] The K12 *trpA* lysate was diluted to 10^{-3} for plating on ECOR 47 and 10^{-5} for plating on K12 and ER2437. Each dilution (0.1 ml) was plated 10 times for each strain.

The replacement patterns of the original transductants appear random with respect to the distribution of borders, and they share some common properties. The bounded replacements are all small relative to the entrant molecule and range from a few hundred to several thousand base pairs (among the 65 replacements there are 10 for which only 1 border is known). Moreover, the incorporation of multiple replacements is often observed. This last observation can account for much of the mosaicism of sequence types seen in the ECOR strains, though not all, because the replacements observed in the ECOR strains are generally 1 kb or less. The transductant replacements average 8–14 kb in length, and because there is not yet any evidence of replacements on that order in the comparative sequences, it seems likely that overlapping replacements have contributed to the patterns observed thus far in nature. In addition, P1 phage transduction may not be the typical method of genetic transfer in nature; bacteriophages (*e.g.*, T1) that transfer smaller pieces of the chromosome may be more prevalent. Also, entrant molecules can be incorporated as large replacements if the strains share the same R-M systems, as implied by the back transduction experiments. Therefore, in principle, large, intraclonal replacements could wipe out or shorten interclonal replacements. (Evidence of such an event has appeared in comparative sequences from the observation of shared replacements of varied length within a meroclone [MILKMAN and MCKANE BRIDGES 1993].) Clearly, the interpretation of replacement patterns in nature remains complicated, but the information obtained from the transductants illustrates some sizes and arrangements that can be expected to occur in nature.

One pattern observed in the compared sequences that is not addressed by the transduction experiments in the multiplicity of sequence types that is often detected within a short region. This suggests frequent recombination with several donors. Of course, if the entrant molecule is itself a mosaic of sequence types (presumably as a result of one or more previous events), then it is possible to introduce more than two sequence types in a region by just one transduction event. For example, a mosaic donor fragment of sequence types "A" and "B" could recombine with a homologous recipient region of sequence type "C" to give the following: AABABAAA → CCCCCCCC ⇒ CCBABACC.

Conjugation and transformation: Different modes of transfer between the ECOR strains may result in different recombination patterns. The "long-chunk" model of SMITH (1991) suggests that 80% of exconjugants should contain long single incorporations of donor DNA. This theory is supported by the recent data of MATIC *et al.* (1994), which showed that in crosses between *E. coli* and various *mut*$^+$ and *mut*$^-$ strains of *Salmonella typhurmurium,* only 3/81 exconjugants had multiple donor replacements. Nevertheless, early observations resemble the transduction patterns in that exconjugant DNA from a restricting recipient shows far less linkage between selected and unselected markers than DNA from a nonrestricting recipient (in the classic sense of phage propagation) (PITTARD 1964; HARRIS and CHRISTENSEN 1966). So the exconjugant patterns from ECOR strains and the possible effects of differing R-M systems on these patterns still need to be explored.

Natural transformation (STEWART 1989) seems to be a less likely mode of transfer in *E. coli,* but it hasn't been ruled out altogether. Regarding transformation in *Bacillus subtilis,* ROBERTS and COHAN (1993) found that the effect of sequence polymorphisms in reducing the frequency of genetic exchange was much greater than the previously seen effect of restriction (COHAN *et al.* 1991).

Regarding both methods of transfer, it has been observed that a transformed plasmid is much more susceptible to host-mediated restriction than is the same conjugated plasmid (READ *et al.* 1992; STEIN *et al.* 1988).

Back-transductants: The results of the back transduction experiments make a striking contrast with the results of the original 47→K12 cross. The replacements in the back-transductants are all single large incorporations. This suggests that restriction endonucleases were probably responsible for the irregular patterns in the original 47→K12 cross, cutting the entrant molecule into fragments (DUBOSE *et al.* 1988; also see PITTARD 1964 and HARRIS and CHRISTENSEN 1966), of which only

one or a few would subsequently be incorporated. The entrant molecules in the back-crosses would not be subject to this type of degradation process, because the restriction-modification genes (R-M genes) of a transductant are likely to be identical to those of the original recipient because the known R-M genes for *E. coli* K12 are too far away to have been replaced in the recipient strain. The nearest R-M system, McrA (KELLEHER and RALEIGH 1991), is ~107 kb away from the selected marker (RUDD 1992). The possibility of any of the transductants possessing new R-M genes from ECOR 47 in the transduced region is unlikely, though not impossible (because of large unanalyzed segments in this region), but even if this were the case, it would be unlikely to significantly affect the outcome of a back transduction.

Restriction-modification polymorphisms: The transductions of 27→K12 and 65→K12 also seem to have been affected by restriction. The replacements in these crosses are also small (compared to the entrant molecule and also the back-transductant replacements) and sometimes multiple. This suggests the possibility that there is extensive polymorphism in the restriction-modification genes among the meroclones of *E. coli* in nature. Furthermore, strains K12, B and ECOR 1 (cited as RM74A) were all found to differ in their restriction-modification systems (DANIEL *et al.* 1988) even though they all belong to the A group (based on the multilocus enzyme electrophoresis of SELANDER *et al.* 1987 and HERZER *et al.* 1990), which also corresponds to the K meroclone (MILKMAN 1994). So although the present experiments did not involve more than one representative of any meroclone, the general possibility exists that there is R-M polymorphism even within each of the several meroclones of *E. coli*.

Further support for extensive polymorphism in the R-M systems of closely related strains comes from the comparative sequence analysis of *hsdM* genes of homologous type I R-M systems (EcoK being the archetype) in *E. coli* and *S. typhimurium* (SHARP *et al.* 1992). Interstrain divergence of the genes within a species was found to be higher than interspecific divergence, suggesting that phage-instigated selection has played a role in the evolution of these genes (SHARP *et al.* 1992).

The acquisition of extensive polymorphism in the restriction-modification systems of closely related strains may not be difficult to achieve. One pertinent example of a recent specificity change in a type I DNA restriction-modification system is a study centered on the *Eco*R124 system and the *Eco*R124/3 system, which arose from the former some 20 years ago. PRICE *et al.* (1989) found that the difference in recognition sequences, GAA(N_6)RTCG and GAA(N_7)RTCG respectively, between these two systems is due to a 12-bp sequence that appears tandemly twice in the *Eco*R124 *hsdS* gene and three times in the *Eco*R124/3 *hsdS* gene. This change in specificity arose spontaneously, presumably by unequal crossing over. Moreover, the expression of the two specificities can switch reversibly at a low level (10^{-7}) *in vivo* (GLOVER and FIRMAN 1982). Another investigation (BULLAS *et al.* 1976) demonstrated a novel restriction specificity in a P1 transductant, whose *hsdS* gene was later determined to be a hybrid of the two parental *hsdS* genes (FULLER-PACE *et al.* 1984). These observations suggest that similar specificity changes may not be rare in nature.

Restriction-modification systems may provide bacteria with more than just a defense mechanism against bacteriophage (PRICE and BICKLE 1986). They may also play a significant role in the genetic exchange process, providing recombinogenic ends (DUBOSE *et al.* 1988) for exonucleases or helicases to further act upon. Restriction endonucleases acting on entrant DNA molecules could also keep the extent of a replacement relatively small, sometimes on the order of a single gene or less. The benefit of small replacements might be the ability to receive generally favorable alleles while retaining others nearby. Small and multiple replacements may also be very advantageous in creating a favorable combination of alleles or indeed codons that would otherwise take much longer to form by large overlapping replacements or mutation. One region that might benefit from such a mechanism surrounds the *rfb* gene (SELANDER *et al.* 1987), which influences the cell surface antigen and probably favors great variation as a protective mechanism. Nearby genes, such as *gnd*, show a high degree of allelic diversity in MLEE, restriction fragment length polymorphism and sequence analysis (MILKMAN 1973; SELANDER *et al.* 1987; DYKHUIZEN and GREEN 1991; BISERCIC *et al.* 1991). It is not known if selection also favors variation in the *gnd* gene, or whether *gnd* and other neighboring loci are merely passive recipients of extensive variation because of their proximity to *rfb*.

It will be of interest to see whether conjugation between wild *E. coli* strains displays patterns resembling those seen in transduction. When this is known, the respective mating affinities of various wild *E. coli* strains will have to be investigated as we progress toward an explicit, quantitative understanding of chromosomal evolution in *E. coli*.

We thank CHARLES YANOFSKY and ERICH SIX for gifts of bacteria and phage strains, ELISABETH RALEIGH for strain construction and discussions, JOHN ROTH for advice and KERRI POHLMAN and JOANNA WERNER for valuable technical assistance. This work was supported by National Science Foundation grant BSR 90-20173.

LITERATURE CITED

BACHMANN, B., 1990 Linkage map of *Escherichia coli* K-12, ed. 8. Microbiol. Rev. **54:** 130–197.

BISERCIC, M., J. Y. FEUTRIER and P. R. REEVES, 1991 Nucleotide sequences of the *gnd* genes from nine natural isolates of *Escherichia coli*: evidence of intragenic recombination as a contributing factor in the evolution of the polymorphic *gnd* locus. J. Bacteriol. **173:** 3894–3900.

BULLAS, L. R., C. COLSON and A. VAN PEL, 1976 DNA restriction and modification systems in *Salmonella*. SQ, a new system derived by recombination between the SB system of *Salmonella typhimurium* and the SP system of *Salmonella potsdam*. J. Gen. Microbiol. **95:** 166–172.

COHAN, F. M., M. S. ROBERTS and E. C. KING, 1991 The potential for genetic exchange by transformation within a natural population of *Bacillus subtilis*. Evolution **45:** 1383–1421.

DANIEL, A. S., F. V. FULLER-PACE, D. M. LEGGE and N. E. MURRAY, 1988 Distribution and diversity of *hsd* genes in *Escherichia coli* and other enteric bacteria. J. Bacteriol. **170:** 1775–1782.

DUBOSE, R. F., D. E. DYKHUIZEN and D. L. HARTL, 1988 Genetic exchange among natural isolates of bacteria: recombination within the *phoA* gene of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **85:** 7036–7040.

DYKHUIZEN, D. E., and L. GREEN, 1991 Recombination in *Escherichia coli* and the definition of biological species. J. Bacteriol. **173:** 7257–7268.

FULLER-PACE, F. V., L. R. BULLAS, H. DELIUS and N. E. MURRAY, 1984 Genetic recombination can generate altered restriction specificity. Proc. Natl. Acad. Sci. USA **81:** 6095–6099.

GLOVER, S. W., and K. FIRMAN, 1982 Genetic control of DNA specificity, pp. 331–338 in *Genetic Exchange*, edited by U. N. STREIPS. Marcel Dekker, New York.

GOLDBERG, R. B., R. A. BENDER and S. L. STREICHER, 1974 Direct selection for P1-sensitive mutants of enteric bacteria. J. Bacteriol. **118:** 810–814.

HARRIS, D. J., and J. R. CHRISTENSEN, 1966 P1 lysogeny and bacterial conjugation. J. Bacteriol. **91:** 898.

HERZER, P. J., S. INOUYE, M. INOUYE and T. S. WHITTAM, 1990 Phylogenetic distribution of branched RNA-linked multicopy single-stranded DNA among natural isolates of *Escherichia coli*. J. Bacteriol. **172:** 6175–6181.

IKEDA, H., and J. TOMIZAWA, 1969 Prophage P1, an extrachromosomal replication unit. Cold Spring Harbor Symp. Quant. Biol. **33:** 791–798.

KELLEHER, J. E., and E. A. RALEIGH, 1991 A novel activity in *Escherichia coli* K-12 that directs restriction of DNA modified at CG dinucleotides. J. Bacteriol. **173:** 5220–5223.

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.

LURIA, S. E., 1953 Host-induced modifications of bacterial viruses. Cold Spring Harbor Symp. Quant. Biol. **18:** 237–244.

MATIC I., M. RADMAN and C. RAYSSIGUIER, 1994 Structure of recombinants from conjugational crosses between *Escherichia coli* donor and mismatch-repair deficient *Salmonella typhimurium*. Genetics **136:** 17–26.

MILKMAN, R., 1995 Recombinational exchange among clonal populations, in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, Ed. 2, edited by F. C. NEIDHARDT. American Society for Microbiology, Washington, DC. (In press).

MILKMAN, R., and M. MCKANE BRIDGES 1990 Molecular evolution of the *E. coli* chromosome. III. Clonal frames. Genetics **126:** 505–517.

MILKMAN, R., and M. MCKANE BRIDGES, 1993 Molecular evolution of the *E. coli* chromosome. IV. Sequence comparisons. Genetics **133:** 455–468.

MILKMAN, R., and A. STOLZFUS, 1988 Molecular evolution of the *E. coli* chromosome. II. Clonal segments. Genetics **120:** 359–366.

MILLER, J. H., 1992 *A Short Course in Bacterial Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

MURRAY, N. E., B. S. GOUGH and T. A. BICKLE, 1982 Structural homologies among type I restriction-modification systems. EMBO J. **1:** 535–539.

NEWMAN, B. J., and M. MASTERS, 1980 The variation with which markers are transduced by phage P1 is primarily a result of discrimination during recombination. Mol. Gen. Genet. **180:** 585–589.

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

PITTARD, J., 1964 Effect of phage-controlled restriction on genetic linkage in bacterial crosses. J. Bacteriol. **87:** 1256–1257.

PRICE, C., and T. A. BICKLE, 1986 A possible role for DNA restriction in bacterial evolution. Microbiol. Sci. **3:** 296–299.

PRICE, C., J. LINGNER, T. A. BICKLE, K. FIRMAN and S. W. GLOVER, 1989 Basis for changes in DNA recognition by the *Eco*R124 and *Eco*R124/3 type I DNA restriction and modification enzymes. J. Mol. Biol. **205:** 115–125.

RADMAN M., 1988 Mismatch repair and genetic recombination, pp. 169–192 in *Genetic Recombination*, edited by R. Kucherlapati and G. R. Smith, American Society for Microbiology, Washington, D.C.

READ, T. D., A. T. THOMAS and B. M. WILKINS, 1992 Evasion of type I and type II DNA restriction systems by IncI1 plasmid CoIIb-P9 during transfer by bacterial conjugation. Mol. Microbiol. **6:** 1933–1941.

ROBERTS, M. S., and F. M. COHAN, 1993 The effect of DNA sequence divergence on sexual isolation in Bacillus. Genetics **134:** 401–408.

ROBESON, J. P., R. M. GOLDSCHMIDT and R. CURTISS III, 1980 Potential of *Escherichia coli* to propagate cloning vectors. Nature **283:** 104–106.

RUDD, K. E., 1992 Alignment of *E. coli* DNA sequences to a revised, integrated genomic restriction map, pp. 2.3–2.43 *in A Short Course in Bacterial Genetics*, edited by J. H. MILLER. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

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 coli and Salmonella typhimurium: Cellular and Molecular Biology*, edited by F. C. NEIDHARDT. American Society for Microbiology, Washington, DC.

SHARP, P., J. A. KELLEHER, A. S. DANIEL, G. M. COWAN and N. MURRAY, 1992 Roles of selection and recombination in the evolution of type I restriction-modification systems in enterobacteria. Proc. Natl. Acad. Sci. USA **89:** 9836–9840.

SILHAVY, T. J., M. L. BERMAN and L. W. ENQUIST, 1984 *Experiments With Gene Fusions*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

SMITH, G. R., 1991 Conjugational recombination in *Escherichia coli:* myths and mechanisms. Cell **64:** 19–27.

STEIN D. C., S. GREGOIRE and A. PIEKAROWICZ, 1988 Restriction of plasmid DNA during transformation but not conjugation in *Neisseria gonorrhoeae*. Infect. Immun. **56:** 112–116.

STERNBERG, N., 1990 Bacteriophage P1 cloning system for the isolation, amplification and recovery of DNA fragments as large as 100 kilobase pairs. Proc. Natl. Acad. Sci. USA **87:** 103–107.

STEWART, G. J., 1989 The mechanism of natural transformation, pp. 139–164 in *Gene Transfer in the Environment*, edited by S. B. Levy and R. V. Miller. McGraw-Hill Publishing, New York.

WOOD, T. H., 1968 Effects of temperature, agitation, and donor strain on chromosome transfer in *Escherichia coli* K-12. J. Bacteriol. **96:** 2077–2084.

Communicating editor: W.-H. LI