JOURNAL OF BACTERIOLOGY, July 1980, p. 355-365
0021-9193/80/07-0355/11$02.00/0

Vol. 143, No. 1

Conservation and Variation of Nucleotide Sequences Within Related Bacterial Genomes: *Escherichia coli* Strains

ALGIS ANILIONIS† AND MONICA RILEY*

Department of Biochemistry, State University of New York, Stony Brook, New York 11794

Changes in the patterns produced by annealing restriction endonuclease digests of bacterial genomes with probe deoxyribonucleic acids (DNAs) containing small portions of a bacterial genome provide sensitive indicators of the degree of nucleotide sequence relatedness that exists in localized regions of the genomes of closely related bacteria. We have used five probe DNAs to explore the relatedness of parts of the genomes of six laboratory *Escherichia coli* strains. A range in the amount of variability in the positions of restriction enzyme cleavage sites in the selected portions of the genomes was found. Portions of the genome that are believed to be inactive were more variable than portions that contained functional genes: the sites in and near regions of homology to phage λ DNA in the genome showed the greatest variability. These regions probably represent remnants of cryptic prophages. Variability was assessed pairwise among four of the *E. coli* strains and ranged from 5 to >25% base pair substitutions in the λ-related regions. In contrast, the endonuclease cleavage sites in the *trp*, *tna*, *lac*, *thy* regions, and one other as-yet-unidentified segment of the genome were more highly conserved. It seems likely that these sites lie in genetic locations that are subject to functional constraints.

Information on nucleotide sequence relatedness among the genomes of closely related enterobacteria has been derived from studies on hybridization of total bacterial DNAs (5). Thermal stabilities of such hybrid DNAs have been used as measures of the extent of nucleotide sequence divergence between strains of a single species (6).

To learn more about evolution of the bacterial genome, it would be useful to have more detailed information on the ways in which closely related bacterial genomes differ from one another. The experiments reported here provide information on nucleotide sequence differences in selected small portions of the genomes of six *Escherichia coli* strains. The Southern transfer technique (26) was used in conjunction with DNA-DNA hybridization to determine the distribution of restriction enzyme target sites near and within selected portions of the genome. Phage λ DNA and four transducing phage DNAs were used as ^{32}P-labeled probes to identify those chromosomal restriction fragments that contained homologs to phage λ sequences, to the *trp*, *tna*, *thy* regions of the genomes, to a small unidentified portion of the genome, and to a portion of the *lac* operon containing the *lacZ* gene. The experimental approach allows determination of the extent of variability of restriction enzyme-sensitive sites in the regions of the bacterial genomes that are homologous to these probe DNAs.

The technique used in these experiments allows a sampling of the nucleotide sequences of a region. In comparing two genetic regions by this method, one would find no difference in the sizes of homologous fragments, provided the hexanucleotide sequences that constitute the enzyme cleavage sites appear the same number of times in the chromosomal region being examined and are identically spaced. Thus, the size of a chromosomal restriction fragment serves as a monitor of the total number of nucleotide changes within a region. Fragment size differences result from changes that cause the loss or gain of a hexanucleotide sequence and/or changes in the numbers of base pairs that lie between them.

MATERIALS AND METHODS

Phages. The constitution of the phages used in these experiments is shown in Fig. 1. λ *S*am7 phage (hereafter referred to as λ for simplicity) was used as a source of λ DNA. The λ *trp* transducing phage A70 carries the structural genes of the entire *E. coli* K-12 *trp* operon of strain W1485 (A. Anilionis, unpublished experiments). The digestion pattern of the λ *trp* DNA suggests that the bacterial DNA is between 9 and 12 kilobase (kb) pairs in extent, of which 7.4 kb pairs constitute *trp* operon DNA.

The *lac* transducing phage is a derivative of λ p*lac*5 (15) which carries the *E. coli* K-12 *lacZ* gene intact

† Present address: The Wistar Institute, Philadelphia, PA 19104.

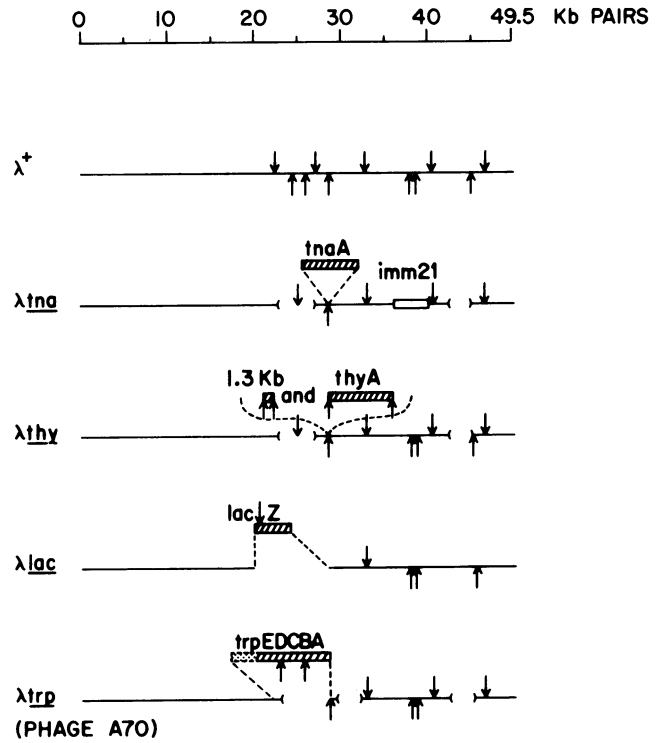

FIG. 1. *Constitution of phage DNAs. The length of the phage DNAs are drawn to scale taking λ⁺ DNA to be 49.5 kb pairs long (P. Philippsen and R. W. Davis, personal communication). Single solid lines represent lambda DNA, the open box in λ tna shows the immunity substitution from phage 21 present in this recombinant. Hatched rectangles denote bacterial DNA carried by the transducing phages, and the stippled area in λ trp shows the region of uncertainty of 3 kb pairs in extent which may be either bacterial or phage in origin. Brackets denote deletions of lambda material. Arrows show the positions of restriction endonuclease cleavage sites: downward pointing for EcoRI and upward pointing for HindIII; those lying between brackets representing hybrid cleavage sites formed by the fusion of restriction fragment ends.*

along with the *lac* control sequences and portions of the *lacI* and *lacY* genes. The positions of the two *Eco*RI sites in the λ *lac* phage were confirmed by agarose gel patterns of the *Eco*RI digest of λ *lac*. This derivative of λ p*lac* appears to have an SR2-like deletion (19) of part of the material between the λ *att* site and the *lac* substitution. The *lacZ* gene constitutes 3.1 kb pairs (29) of the 4.2 kb pairs of bacterial DNA present in this phage.

The λ *tna* phage carries the *E. coli* K-12 tryptophanase (*tnaA*) gene and is described by Borck et al. (8). This phage carries a *Hin*dIII fragment of 6.2 kb pairs in addition to the known λ vector fragments (data not shown). The tryptophanase monomer molecular weight is 55,000 (22), corresponding to a gene of approximately 1.5 kb pairs which must be present within the 6.2 kb bacterial insert.

The λ *thy* phage is a derivative of the *thyA* phage constructed by Borck et al. (8) which carries the *E. coli* K-12 thymidylate synthetase gene. *Hin*dIII digests of λ *thy* DNA revealed the presence of two non-λ fragments of 7.4 and 1.3 kb pairs; both were shown to originate from the *E. coli* chromosome (see Results). The *E. coli* thymidylate synthetase has a molecular weight of 67,000 (10), leading to an estimated gene size of 1.8 kb pairs for *thyA*. Thus, the *thyA* gene most likely resides in the 7.4 kb chromosomal fragment of the λ *thy* DNA.

Bacteria and preparation of bacterial DNA. The *E. coli* strains and their sources are as follows: K-12 (λ), *E. coli* Genetic Stock Center 5073; K-12 W1485, *E. coli* Genetic Stock Center 5024; B/1,5, Martha Baylor; W, Henry Vogel; C-1, Richard Calendar; 15T⁻, Arthur Pardee. Total bacterial DNA was prepared from each strain by a modification of the method of Marmur (20). The following steps were incorporated. (i) Lysed suspensions of minimal grown cells were treated with Pronase (self digested for 1 h at 37°C) at a final concentration of 100 μg/ml for 3 h at 37°C with gentle agitation. (ii) After phenol extraction, RNA was removed by digestion with RNase (stock solution at 10 mg/ml in 10 mM Tris-hydrochloride [pH 7.5], heat treated at 80°C for 10 min) at 50 μg/ml for 1 h at 37°C. (iii) Purified DNA samples were dialyzed against 1.0 M NaCl in TE buffer (10 mM Tris-hydrochloride [pH 7.5], 1 mM EDTA) for 12 to 18 h and then exhaustively against TE buffer. DNA concentrations were determined from optical density readings at 260 nm (OD_{260})

on samples with OD_{260}/OD_{280} ratios in the range of 1.8 to 2.0.

Preparation of ^{32}P-labeled phage DNAs. DNA was prepared from λ *Sam*7 (12) and λ transducing phages essentially as was described previously (16). Phage DNAs were labeled by nick translation as described by Rigby et al. (24) except that standard reactions containing 0.5 μg of DNA, 180 pmol each of [α-^{32}P]dATP and [α-^{32}P]dCTP (400 Ci/mmol, Amersham Corp.), and 180 pmol of unlabeled dGTP and dTTP were performed in a total volume of 30 μl at 14°C for 1 h. Specific activities obtained were routinely 1×10^8 to 2×10^8 cpm/μg of DNA. These labeled phage DNA probes were used within a month of preparation.

Restriction enzyme digests and agarose gel electrophoresis. Restriction endonuclease reaction mixtures routinely contained 1 to 2 μg of bacterial DNA in 20 μl of reaction buffer (20 mM potassium phosphate [pH 7.4], 50 mM NaCl, 10 mM $MgCl_2$, 0.2% Triton X-100, and 2 mg of gelatin per ml) and 0.3 to 2.0 U of *Eco*RI or *Hin*dIII enzymes (Boehringer Mannheim Co.). Reaction mixtures were incubated overnight at 37°C and stopped by adding one-fifth volume of 0.25 M EDTA (pH 8.0). Aliquots of 0.5 μg of digested DNA were loaded on 1% agarose gels (vertical slabs [16 by 17 by 0.25 cm] in 40 mM Tris-acetate [pH 8.2]-20 mM sodium acetate-1 mM EDTA) and subjected to electrophoresis at 40 mA (80 V) for 6 h. Gels were stained in a 0.5-μg/ml of ethidium bromide solution and photographed on a UV-light box (UV Products, San Gabriel, Calif.) using a 23A orange filter and Kodak Royal Pan 4141 film.

Transfer and hybridization of chromosomal DNAs. Denatured digests of bacterial DNAs were transferred to nitrocellulose filters from agarose gels after electrophoresis and baked essentially as described by Southern (26). Hybridization was carried out with 20 μl of probe solution per cm^2 of filter area. Probe solution contained 6 ng of ^{32}P-labeled phage DNA per ml in 4× SSC-50% formamide (SSC is 0.15 M NaCl-0.015 M sodium citrate [pH 7.0]; 4× is four times concentrated; formamide is Eastman "Spectro Grade"). Filters were sealed in polyethylene bags and hybridized at 37°C for 19 h, washed once briefly in 10 to 20 ml of 4× SSC-50% (vol/vol) formamide (Fischer; reagent grade) at room temperature and then in an excess (~200 ml) of the same solution for 3 to 4 h at 37°C, rinsed in 3 mM Tris-hydrochloride (pH 7.5), and air dried. Filters were exposed to sensitized Kodak XR-1 X-ray film and intensified with a Dupont Cronex Lightning Plus screen (YF 102909) at −70°C for 1 to 4 days.

RESULTS

Representatives of five commonly used *E. coli* strains were chosen for study: strains K-12, C, W, B, and 15. To represent the K-12 strain, the progenitor of many current laboratory strains, K-12 (λ) (1), was chosen. Since phage λ DNA was present in all probes to be used in this study, the closely related but non-lysogenic strain K-12 W1485 (1) was also included in the experiments to simplify the K-12 hybridization patterns.

Total DNA was isolated from each bacterial strain and was digested with either *Eco*RI or *Hin*dIII restriction enzymes. The electrophoretic distribution in an agarose gel of DNA fragments from both *Eco*RI and *Hin*dIII digestions of the two *E. coli* K-12 strains and *E. coli* strains C, W, B, and 15 are presented in Fig. 2. Although the patterns of the two K-12 strains were similar, they did have definite differences. Each digest pattern differed from the others, showing that cleavage sites for each of the restriction enzymes are distributed differently in each of these bacterial genomes. Prominent bands in *E. coli* W digests originate from a plasmid of about 10 kb that is present in this strain.

Hybridization experiments were carried out to locate in these chromosomal digests the restriction fragments that were homologous to each of five probe DNAs. Replicate gels were prepared of the six *E. coli* chromosomal DNAs similar to those pictured in Fig. 2. The DNA on the gels was denatured and transferred to filters. Five radioactive probe DNAs, phage λ DNA and four λ-transducing DNAs, were hybridized to replicate filters. The genetic constitution of each of the probe DNAs is shown in Fig. 1. Autoradiograms were made to locate probe DNA bound to the filters, thus permitting identification of the chromosomal restriction fragments that contained segments of DNA that were homologous to probe DNA sequences.

λ probe. The autoradiograms resulting from hybridization with phage DNA are shown in Fig. 3, and molecular weights of chromosomal fragments that contained homologs to phage λ DNA are listed in Table 1.

The *Hin*dIII and *Eco*RI digests of K-12 (λ) DNA each contained bands on agarose gels which were homologous to λ DNA. Berg and Drummond (4) showed that λ-homologs reside in *Eco*RI digests of *E. coli* K-12 DNA. The *Eco*RI fragments of strain K-12 W1485 that contained λ-homologs (Fig. 3 and Table 1) were about the same size as those reported earlier by Berg and Drummond, and similar-sized bands were present in the K-12 (λ) digest. However, differences between K-12 (λ) and K-12 W1485 were seen in the *Hin*dIII digests (Table 1). The >30-kb and 10-kb fragments in the K-12 W1485 *Hin*dIII digest were not seen in the K-12 (λ) digest. These results show that the positions of *Hin*dIII target sites are not identical in the genomes of even these two closely related K-12 sublines. Although K-12 (λ) and K-12 W1485 are separated by a few known mutational steps (1), the two strains have been maintained in parallel in laboratory culture for about 40 years, so that

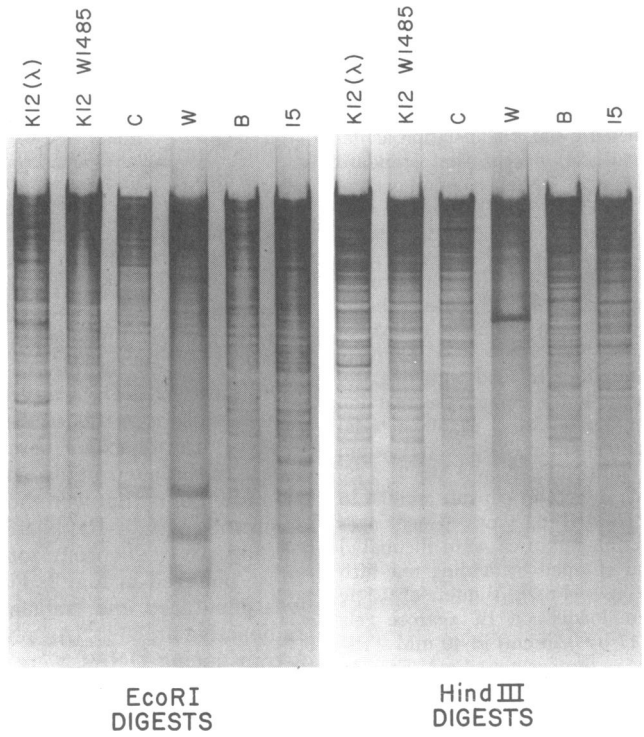

FIG. 2. *Agarose gel electrophoresis of EcoRI and HindIII digests of E. coli DNAs. Digests (0.5 μg each) were subjected to electrophoresis through 1% agarose and stained with ethidium bromide and photographed as described in the text.*

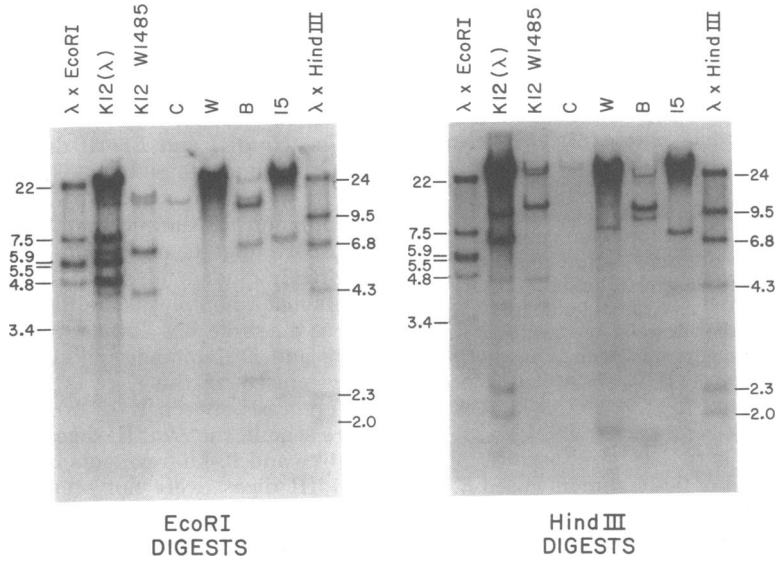

FIG. 3. *Hybridization of ^{32}P-labeled λ DNA to EcoRI and HindIII digests of E. coli DNAs. Nitrocellulose filters carrying restriction enzyme digests of E. coli DNAs (0.5 μg each) from gels identical to those shown in Fig. 2, with additional outer tracks of 0.5 ng of EcoRI and HindIII digests of λ DNA, were hybridized with ^{32}P-labeled λ DNA and autoradiographed as described in the text.*

TABLE 1. *Sizes of λ-homolog fragments in HindIII and EcoRI digests of E. coli DNAs*

Mol wt (kb) of λ-homologous fragments[a]											
*Eco*RI digests						*Hin*dIII digests					
K-12 (λ)	K-12 W1485	C	W	B	15	K-12 (λ)	K-12 W1485	C	W	B	15
			>30		>30		>30	>30	>30		>30
23				25		25	26			22	
15	15		15							(18.5)[b]	
13	13	13		13			10			10	
				11.5		9.2					
					7.5					8.7	
7.3									7.8		7.8
6.5	6.5			6.7		6.6					
5.8						4.6	4.6				
4.7											4.4
4.4	4.4		(4.3)		(4.5)						(4.2)
					2.95	3.8	3.8				
			2.5	2.5						3.5	
					2.25	(3.2)	(3.2)				
									(2.5)		
						2.25					
						2.0					
									1.85	1.80	
									~1.0		

[a] Molecular weights were determined by measurement of the autoradiogram shown in Fig. 3 and comparison with a calibration curve drawn from known molecular weights of the λ markers.

[b] Parentheses indicate faint bands.

additional changes may have accumulated in the two genomes.

Turning now to *E. coli* strains other than K-12, we found that the λ-homologs in the chromosomal digests of strains C, W, B, and 15 are highly variable in numbers and size distributions (Fig. 3 and Table 1). Although some fragment sizes were conserved in these *E. coli* strains, the extent of homology to DNA in these genomes, and the positions of some of the *Eco*RI or *Hin*dIII sites in or near the homologous regions, differed in each of the *E. coli* strains we examined.

Upholt (27) has formulated an analysis which uses the fraction of cleavage sites conserved between two DNAs to estimate the amount of sequence divergence that has occurred between the two DNAs. Assuming that changes in cleavage sites have been brought about by base substitutions alone, and not by small additions, deletions, or rearrangements, then Upholt's treatment can be applied to estimate the differences in nucleotide sequences in or near the λ-homolog portions of the genomes of some of the *E. coli* strains used in this study. The lysogenic strain K-12 (λ) was excluded from the comparison since λ-homologs could not be distinguished with certainty from the two end fragments of the prophage. Strain C was excluded because it contained only a single fragment in each digest, a number too small to be subjected to this analysis. For the four remaining strains, the fraction of fragments that were conserved were counted for each pairwise comparison of the four strains and each of the two digests, and the corresponding fraction of base substitutions was determined for each pair (Table 2).

Transducing λ DNA probes. Transducing phage DNAs were used as probes to compare the patterns of hybridization of small, specific regions of the *E. coli* K-12 genome with the equivalent regions of the genomes of closely related *E. coli* strains. The amount of bacterial DNA present in each probe ranged from 4.2 kb to about 10 kb (Fig. 1). For convenience of discussion, the bacterial DNA in the transducing phages will be referred to simply by the designation of the bacterial gene(s) known to be present: *trp*, *thy*, *tna*, and *lac*.

λ *thy* probe. The λ *thy* probe was found to contain two unrelated segments of bacterial DNA. Neither the 7.4-kb *thyA* fragment nor the "extra" 1.3-kb bacterial fragment in the λ *thy* probe DNA carries an *Eco*RI cleavage site (Fig. 1), and the *E. coli* strains tested all carry these two regions on two large *Eco*RI fragments (Table 3). If this same configuration is present in the genome from which λ *thy* phage was derived (*E. coli* K-12 W3110, see reference 8), then the 7.4-kb and 1.3-kb fragments cannot be contig-

uous (since the *Eco*RI fragments would also be contiguous, necessarily introducing an *Eco*RI cleavage site within one of the probe fragments). Thus, we believe that the two bacterial DNA segments present in the λ *thy* phage are independently derived from the same genome. The genetic location of the 1.3-kb piece is unknown.

TABLE 2. *Conservation of λ-homolog fragments in E. coli chromosomal digests*

Pairs of *E. coli* strains compared	Fraction of fragments conserved[a]			Inferred % base substitutions[b]
	*Hin*dIII digests	*Eco*RI digests	Avg	
K-12 W1485/W W/K-12 W1485	2/11	4/8	0.34	6
K-12 W1485/15 15/K-12 W1485	2/10	2/9	0.22	10
K-12 W1485/B B/K-12 W1485	4/12	4/9	0.39	5
W/15 15/W	4/9	4/9	0.44	5
W/B B/W	2/11	2/9	0.20	10
B/15 15/B	0/10	0/10	0	>25

[a] Obtained by counting the number of pairs of conserved fragments, a, in paired digests (Table 1) which contain, respectively, x and y total fragments; the fraction of fragments conserved is calculated as 2a/(x + y).

[b] The fraction of conserved fragments is related to fraction of base substitutions as $P = 1 - \left[\frac{-F + \sqrt{F^2 + 8F}}{2}\right]^{(1/n)}$, where P = fraction of base substitutions, F = fraction of conserved fragments, and n = number of bases in a restriction site (see reference 27).

The hybridization pattern for the λ *thy* probe is presented in Fig. 4. Visible in the autoradiogram are the λ-homologs that were previously located (Fig. 3) plus additional bands that can be attributed to restriction fragments that hybridized to the two bacterial fragments. These are indicated with pointers in Fig. 4, and their sizes are listed in Table 3. Some of the λ-homologs seen in Fig. 3 are not present in Fig. 4; these are presumed to be homologs for those portions of the probe DNA that were substituted or deleted in the λ vector.

The 7.4-kb *thy* homologs and the 1.3-kb bacterial DNA homologs in the *Hin*dIII digests of each of the *E. coli* strains examined were not detectably different in size within the limits of accuracy of these experiments.

In the *Eco*RI digests of the *E. coli* DNAs, a high-molecular-weight fragment of ~30 kb and a smaller (21 to 22 kb) fragment was evident in all digests, with the exception of strain W and possibly strain 15 (see Table 3 and footnote *b*).

The positions of the *Hin*dIII and *Eco*RI cleavage sites flanking the *thyA* gene and the 1.3-kb unassigned *E. coli* DNA fragment are all highly conserved, no significant size variants having been detected in any of the six genomes examined.

λ *tna* probe. Hybridization patterns produced with λ *tna* are shown in Fig. 5, and molecular weights of restriction fragments containing the *tna*-homologs are given in Table 3. In

TABLE 3. *Sizes of homologous fragments (kb) in HindIII and EcoRI digests of E. coli DNAs*[a]

Homolog	Size of homologous fragments (kb) in:											
	*Eco*RI digests						*Hin*dIII digests					
	K-12 (λ)	K-12 W1485	C	W	B	15	K-12 (λ)	K-12 W1485	C	W	B	15
thy	~30 22	~30 22	~30 22	—[b] —	~30 22	— 21	7.5 1.3	7.5 1.3	7.4 1.3	7.4 1.3	7.4 1.3	7.4 1.3
tna	8.8	8.8	8.8	16	15	14	6.2	6.2	6.2	6.2	6.2	6.2
trp	—	~30 6.3	~30	—	~30	—					11.8 6.7	
							5.9 2.95 2.7	5.9 2.95 2.6	5.9 2.95 2.6	2.95 2.6 (1.9)[c]	2.95 2.6	5.9 2.95 2.6
lac	— ~21	~21	~20	— —	~30	19	—	27	27	—	27	27

[a] Not all homologs may have been detected: hybrid fragments whose mobilities are the same as those of λ-homologs (Fig. 3) may also be present.

[b] —, λ-homologs obscure the 20- to 30-kb region.

[c] Parentheses indicate faint band on autoradiogram.

the *Hin*dIII digest, the *tna*-homologs have the same molecular weight for all six *E. coli* strains. No detectable differences exist in the position of *Hin*dIII-specific sites in the *tna* regions of the genomes of these bacteria. The *Eco*RI restriction fragments that contain *tna*-homologs are the same size, 8.8 kb, in strains K-12 (λ), K-12 W1485, and C, but are larger, around 15 kb, in strains W, B, and 15 (Table 3). A λ-homolog and a *tna*-homolog comigrate in the *Eco*RI digest of

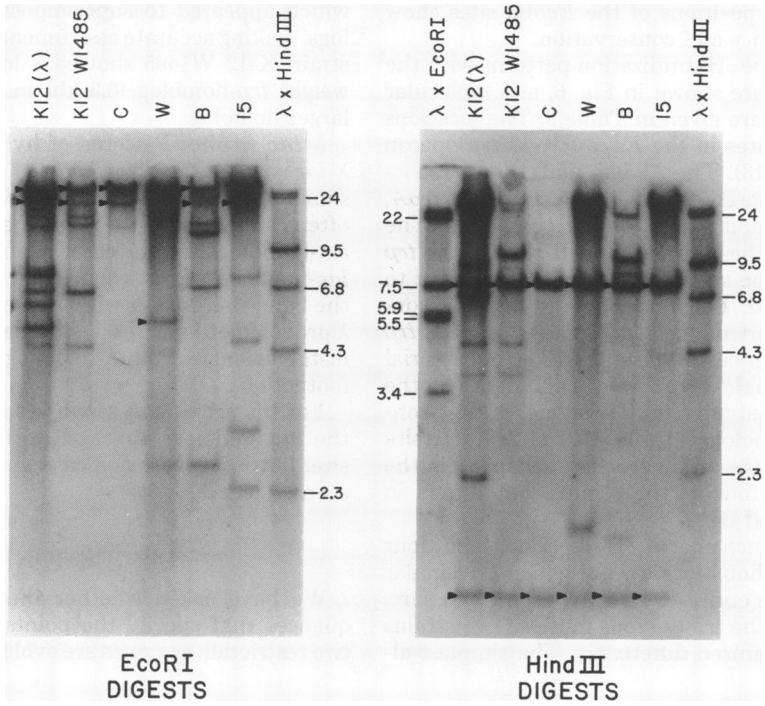

FIG. 4. *Hybridization of ^{32}P-labeled λ thy DNA to EcoRI and HindIII digests of E. coli DNAs. See legend to Fig. 3 for details. Pointers show λ thy-homologous bands not present in Fig. 3.*

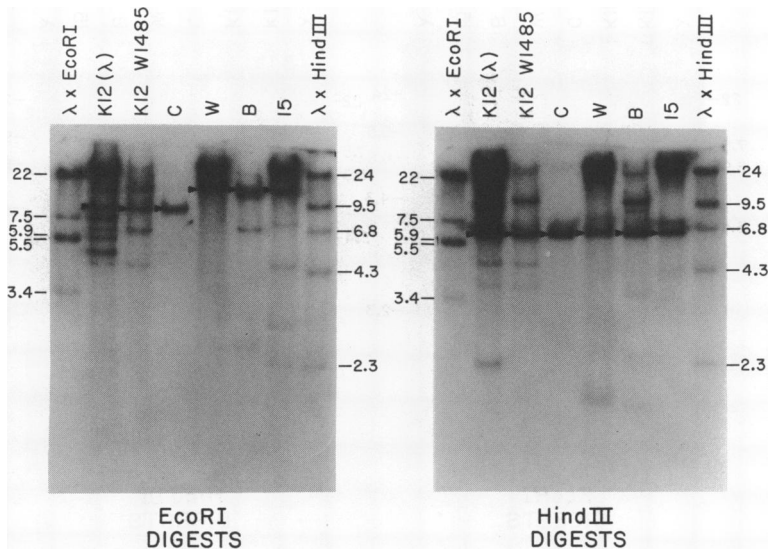

FIG. 5. *Hybridization of ^{32}P-labeled tna DNA to EcoRI and HindIII digests of E. coli DNAs. See legend to Fig. 3 for details. Pointers show λ tna-homologous bands not present in Fig. 3.*

strain W, but identification was possible because of increased relative band intensity in the λ-*tna* autoradiogram.

These results show that, as for the *thy*-homologs, the positions of the *Hin*dIII sites in the *tna* region are the same for all six *E. coli* strains, whereas the positions of the *Eco*RI sites show both variability and conservation.

λ *trp* probe. Hybridization patterns with the λ *trp* probe are shown in Fig. 6, and molecular weight data are given in Table 3. The locations of *Hin*dIII sites in the *E. coli* K-12 *trp* operon are known (13). The 2.95-kb and 2.6-kb restriction fragments contain, between them, the *trpA*, *B*, *C* genes and a part of the *trpD* gene. The positions of these three *Hin*dIII sites in the *trp* operons of the six *E. coli* strains appear not to have changed. The longer 5.9-kb fragment contains the start of the *trpD* gene, *trpE*, the *trp* control sequences, and less than 3 kb of bacterial DNA proximal to *trpE* (Fig. 1). Judging by the sizes of the larger *Hin*dIII fragments that contain *trp*-homologs (Table 3), the *Hin*dIII site upstream of the *trp* operon is unchanged in the genomes of four of the strains, but differs in strains W and B.

The chromosome of strain B contains four *Hin*dIII *trp*-homologs, two of which are similar in size to the entire K-12 *trp* operon (7 kb pairs long). Thus the *trp* operons in these two strains must be organized differently. The simplest alternative arrangement for strain B would be two separate locations of *trp* genes, each of which contributes one of the high-molecular-weight fragments.

In the *Eco*RI digests the *trp*-homologs resided in large-molecular-weight fragments, some of which appeared to superimpose with λ-homologs, making accurate assignments difficult. Only strain K-12 W1485 showed a lower-molecular-weight *trp*-homolog (6.3 kb) in addition to the larger homolog.

λ *lac* probe. Patterns of hybridization with λ *lac* probe are shown in Fig. 7. The chromosomal fragments that contained a *lac*-homolog often comigrated with fragments that contained a λ-homolog. In some cases, identification of the *lac*-homolog was possible since the intensity of the band was greater with λ *lac* probe than with λ probe. In other cases, large fragments were not distinguishable from λ-homologs (Table 3 and footnote *b*).

For the six *E. coli* strains examined, most of the *lac*-homologs appeared to be conserved in size, but in at least one case a size variant was seen.

DISCUSSION

We have asked whether the nucleotide sequences that specify the points of cleavage of two restriction enzymes are evolutionarily stable

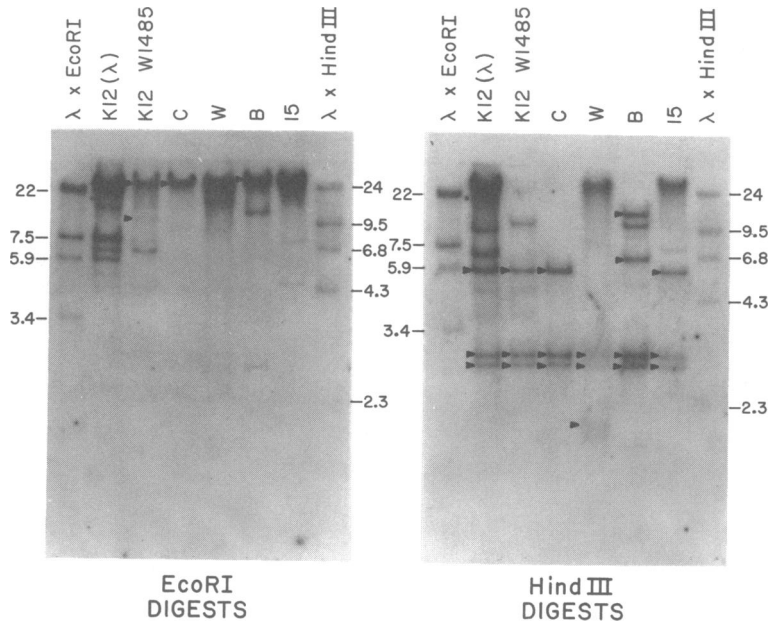

FIG. 6. *Hybridization of ^{32}P-labeled λ trp DNA to EcoRI and HindIII digests of E. coli DNAs. See legend to Fig. 3 for details. Pointers show λ trp-homologous bands not present in Fig. 3.*

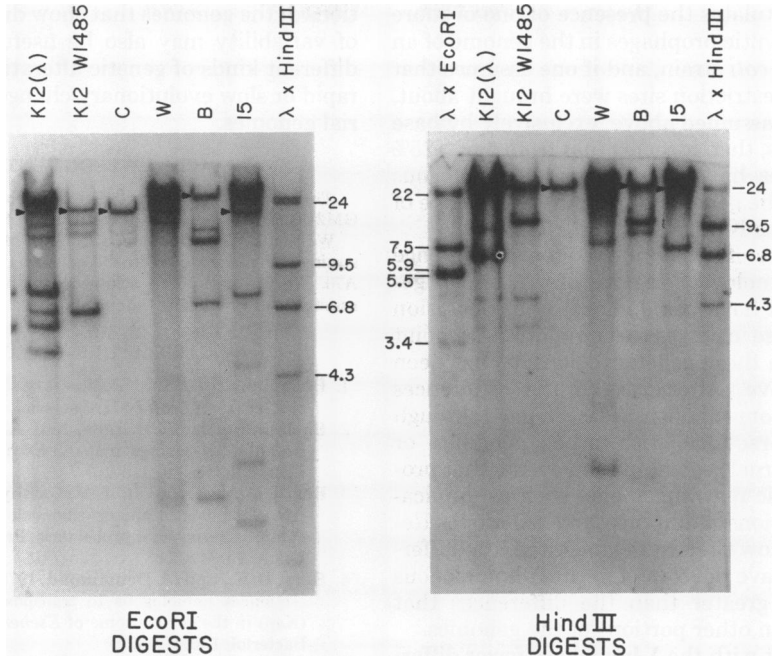

FIG. 7. *Hybridization of ^{32}P-labeled λ lac DNA to EcoRI and HindIII digests of E. coli DNAs. See legend to Fig. 3 for details. Pointers show λ lac-homologous bands not present in Fig. 3.*

and are in the same positions in the genomes of six *E. coli* strains, or whether evolutionary changes in nucleotide sequences in the genomes have altered the numbers or positions of restriction enzyme target sites. In Fig. 2 it can be seen that the restriction fragments in the digests of the DNA of *E. coli* strains K-12, C, W, B, and 15 have different size distributions in both *Eco*RI and *Hin*dIII restriction fragments. Nucleotide sequence changes have caused alterations in the positions of restriction enzyme target sites in these bacterial genomes during their divergence.

To learn more about these changes in the locations of restriction-sensitive sites, we have examined selected portions of the genome in more detail. One probe examined those portions of the genome that are homologous to phage DNA. The other four probes carried genetically defined portions of the K-12 genome that are widely scattered on the genetic map; *lac*, *trp*, *thy*, and *tna*, at map positions 8, 27, 60.5, and 82, respectively (2), as well as a small undefined portion of unknown genetic location.

The phage DNA probe identified portions of the *E. coli* genomes that are probably vestiges of lambdoid prophages, confirming and extending the work of Berg and Drummond (4). Two λ-homologs, possibly remnants of a cryptic prophage, reside at the *rac* locus (9, 11, 17), and it is known that a lambdoid *QSR* region resides in the host K-12 genome (15). We have found recently that another λ-homolog resides near the *gal* operon (unpublished observations). A fourth λ-homolog has not yet been located on the *E. coli* map.

The experiments in which λ DNA was hybridized to chromosomal digests of six *E. coli* strains showed that the positions of restriction sites in or near the λ-homologs were highly variable (Table 1), more variable than the positions in or near any of the other portions of the genome that were assessed in the same way (Table 3). Such susceptibility to change may reflect a difference in functional constraints. If the chromosomal λ-homologs are in fact vestiges of one or more lambdoid prophages, one might expect such nonfunctional cryptic phage remnants to be free to accept and retain all or most of the mutational changes that have taken place in those parts of the genome.

The genetic distance between the λ-homologous portions of the genomes of four of the *E. coli* strains was estimated in terms of the frequency of base substitutions required to produce the observed changes in restriction sites (Table 2). By this measure, *E. coli* K-12 is related to about the same extent to strains W, B, and 15; strain W is related to about the same extent to strains K-12, B, and 15; however, strains B and 15 are most widely separated from each other.

If one postulates the presence of one or more lambdoid cryptic prophages in the genome of an ancestral *E. coli* strain, and if one assumes that changes in restriction sites were brought about, as has been assumed above, exclusively by base substitutions, then it seems that from 5 to >25% base changes have occurred in λ-homologous portions of the genome during the divergence of *E. coli* strains K-12, B, W, and 15.

This degree of nucleotide sequence difference among λ-homologous regions of the *E. coli* genomes could have come about by a succession of single base changes, accumulating as point mutations in these genetic regions, as has been implied above. Alternatively, the differences could have come about, at least in part, through genetic interactions with other genomes or through internal recombination events that produced genetic rearrangements such as duplications, inversions, additions, and deletions. Regardless of how they were generated, the differences that have developed in the λ-homologous regions are greater than the differences that were found in other portions of the genomes.

In contrast with the λ-homologs, fewer differences were found in the other portions of the *E. coli* genome that were examined (Table 3). All of the *Hin*dIII sites flanking the *tnaA* and *thyA* genes and the "1.3 kb" segment were conserved in the six strains, as were a majority of the restriction sites in the other gene regions that were studied. It seems likely that the highly conserved cleavage sites lie in sequences that must remain unchanged to preserve gene function.

The different degrees of variability that exist in separate portions of the *E. coli* genome provides the opportunity to study evolutionary distances in bacteria over a wide range of phylogenetic separation by selecting the appropriate portions of the genome for study. Highly conserved RNA genes are useful for studying relationships between bacteria in different genera (7), families (3), and higher orders (14, 28). Structural genes are more useful for studying relationships between bacteria in different genera (e.g., 18). Highly variable regions such as the λ-homologs should be useful in studying relationships between very closely related bacteria, e.g., within a single species.

In summary, we have found that different portions of the genomes of *E. coli* laboratory strains have diverged from one another to different extents. It seems likely that nonfunctional regions have evolved at a faster rate than have functional regions. Fast-evolving regions of the genome should be useful for studying evolutionary changes in relatively unconstrained portions of the genomes of closely related bacteria. Portions of the genomes that show different degrees of variability may also be useful for studying different kinds of genetic alterations that affect rapid or slow evolutionary change within bacterial genomes.

ACKNOWLEDGMENTS

This work was supported by Public Health Service GM21316 from the National Institutes of Health.

We thank W. J. Brammar and N. E. Murray for bacterial strains and phages used in the construction of λ *trp* phage A70. We are pleased to acknowledge the expert technical assistance of Philomena Ostapchuk.

LITERATURE CITED

1. **Bachmann, B. J.** 1972. Pedigrees of some mutant strains of *Escherichia coli* K-12. Bacteriol. Rev. **36**:525–557.
2. **Bachmann, B., K. B. Low, and A. L. Taylor.** 1976. Recalibrated linkage map of *Escherichia coli* K-12. Bacteriol. Rev. **40**:116–167.
3. **Bendich, A. J., and B. J. McCarthy.** 1970. Ribosomal RNA homologies among ribonucleic acid cistrons in enterobacteria and myxobacteria. Proc. Natl. Acad. Sci. U.S.A. **65**:349–356.
4. **Berg, D. E., and M. Drummond.** 1978. Absence of DNA sequences homologous to transposable element Tn*5* (Kan) in the chromosome of *Escherichia coli* K-12. J. Bacteriol. **136**:419–422.
5. **Brenner, D. J., and S. Falkow.** 1971. Molecular relationships among members of the Enterobacteriaceae. Adv. Genet. **16**:81–118.
6. **Brenner, D. J., G. R. Fanning, F. J. Skerman, and S. Falkow.** 1972. Polynucleotide sequence divergence among strains of *Escherichia coli* and closely related organisms. J. Bacteriol. **109**:953–965.
7. **Brenner, D. J., G. R. Fanning, A. G. Steigerwalt, M. A. Sodd, and B. P. Doctor.** 1977. Conservation of transfer ribonucleic acid and 5S ribonucleic acid cistrons in *Enterobacteriaceae*. J. Bacteriol. **129**:1435–1439.
8. **Borck, K., J. D. Beggs, W. J. Borammar, A. S. Hopkins, and N. E. Murray.** 1976. The construction *in vitro* of transducing derivatives of phage lambda. Mol. Gen. Genet. **146**:199–207.
9. **Campbell, A.** 1977. Defective bacteriophages and incomplete prophages. Comp. Virol. **8**:259–328.
10. **Crusberg, T. C., R. Leary, and R. L. Kisliuk.** 1970. Properties of thymidylate synthetase from dichloromethotrexate-resistant *Lactobacillus casei*. J. Biol. Chem. **245**:5292–5296.
11. **Evans, R., N. R. Seeley, and P. L. Kuempel.** 1979. Loss of *rac* locus DNA in merozygotes of *Escherichia coli* K12. Mol. Gen. Genet. **175**:245–250.
12. **Goldberg, A. R., and M. Howe.** 1969. New mutations in the S cistron of bacteriophage λ affecting host cell lysis. Virology **38**:200–202.
13. **Hopkins, A. S., N. E. Murray, and W. J. Brammar.** 1976. Characterization of λ*trp*-transducing bacteriophages made *in vitro*. J. Mol. Biol. **107**:549–569.
14. **Hori, H., and S. Osawa.** 1979. Evolutionary change in 5S RNA secondary structure and a phylogenetic tree of 54 5S RNA species. Proc. Natl. Acad. Sci. U.S.A. **76**: 381–385.
15. **Ippen, K., J. A. Shapiro, and J. R. Beckwith.** 1971. Transposition of the *lac* region to the *gal* region of the *Escherichia coli* chromosome: isolation of λ*lac* transducing bacteriophages. J. Bacteriol. **108**:5–9.
16. **Kaiser, A. D.** 1962. The production of phage chromosome fragments and their capacity for genetic transfer. J. Mol. Biol. **4**:275–287.
17. **Kaiser, K., and N. E. Murray.** 1979. Physical character-

ization of the "Rac prophage" in *E. coli* K12. Mol. Gen. Genet. **175**:159–174.

18. **Lee, F., K. Bertrand, G. Bennett, and C. Yanofsky.** 1978. Comparison of the nucleotide sequences of the initial transcribed regions of the tryptophan operons of *Escherichia coli* and *Salmonella typhimurium*. J. Mol. Biol. **121**:193–217.
19. **Malamy, M. H., M. Fiandt, and W. Szybalski.** 1972. Electron microscopy of polar insertions in the *lac* operon of *Escherichia coli*. Mol. Gen. Genet. **119**:207–222.
20. **Marmur, J.** 1961. A procedure for the isolation of deoxyribonucleic acid from microorganisms. J. Mol. Biol. **3**: 208–218.
21. **McConaughy, B. L., C. D. Laird, and B. J. McCarthy.** 1969. Nucleic acid reassociation in formamide. Biochemistry **8**:3289–3295.
22. **Morino, Y., and E. E. Snell.** 1967. The subunit structure of tryptophanase. J. Biol. Chem. **242**:5591–5601.
23. **Nichols, B. P., and C. Yanofsky.** 1979. Nucleotide sequences of *trp*A of *Salmonella typhimurium* and *Escherichia coli*: an evolutionary comparison. Proc. Natl. Acad. Sci. U.S.A. **76**:5244–5248.
24. **Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237–251.
25. **Sato, K., and A. Cambell.** 1970. Specialized transduction of galactose by lambda phage from a deletion lysogen. Virology **41**:474–487.
26. **Southern, E.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
27. **Upholt, W. B.** 1977. Estimation of DNA sequence divergence from comparison of restriction endonuclease digests. Nucleic Acids Res. **4**:1257–1265.
28. **Woese, C. R., and G. E. Fox.** 1977. Phylogenetic structure of the prokaryotic domain; the primary kingdoms. Proc. Natl. Acad. Sci. U.S.A. **74**:5088–5090.
29. **Zabin, I., and A. V. Fowler.** 1978. β-Galactosidase, the lactose permease protein and thiogalactoside transacetylase, p. 89–121. *In* J. H. Miller and W. S. Reznikoff (ed.), The operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.