GENETIC DIVERSITY AND TEMPORAL VARIATION IN THE *E. COLI* POPULATION OF A HUMAN HOST

DOMINIQUE A. CAUGANT, BRUCE R. LEVIN* AND ROBERT K. SELANDER

Department of Biology, University of Rochester, Rochester, New York 14627, and
**Department of Zoology, University of Massachusetts 01003*

Manuscript received October 27, 1980
Revised copy received May 18, 1981

ABSTRACT

Electrophoretic techniques were employed to study variation in chromosomal genes encoding enzymes and in the distribution of cryptic plasmids in the *E. coli* population of a human host over an 11-month period. Thirteen of the 15 enzymes studied were polymorphic, and mean genetic diversity per locus was 0.39. Among 550 clones isolated from fecal samples, protein electrophoresis revealed 53 distinct electrophoretic types (ETs). Most ETs appeared on only one or a few days and were considered *transients*, but two (ET-12 and ET-13) were observed many times over extended periods and represented *residents*. Complete turnover in the transient ETs in the population occurred in periods of from two weeks to a month. ETs appearing in one month showed no particular genetic similarity to those of the previous month. —— All but 4 of the 53 ETs carried one or more "cryptic" plasmids with molecular weights ranging from 1 to 80 megadaltons. With few exceptions, the plasmid composition of each ET was unique. In the course of the 11-month sampling period, there were changes in the plasmid profiles of the resident strains ET-12 and ET-13, and also in the profile of a recurrent strain, ET-2, which was isolated on four days. Modification of the plasmid profile of ET-12 involved the sequential addition of relatively high molecular weight bands. For ET-2 and ET-13, the changes in the plasmid profiles were radical, suggesting invasions of new cell types rather than merely the addition and deletion of plasmids. —— The results of this study provide three lines of evidence that recombination plays a minor role in the generation of genetic diversity in the *E. coli* population of a single host. (1) Several pairs of loci were in strong linkage disequilibrium; compared to a randomly generated array of genotypes, the sample of ETs contained an excess of pairs differing at one or two loci and too many pairs with highly distinctive combinations of electromorphs. (2) In most cases where pairs of ETs differed at a single locus and, therefore, could reasonably have been generated by phage- or plasmid-mobilized gene transfer, the plasmid profiles of the pair members were radically different and/or the potentially transmitted alleles were not present in other ETs in the population. (3) Although ET-12 was abundant, being represented by 252 of the 550 clones sampled, the electrophoretic type most similar to ET-12 differed from it at six loci, and ET-12 carried two unique alleles. We conclude that most of the genetic diversity observed in this human host is a consequence of successive invasions of *E. coli* genotypes.

FUNDAMENTAL to much of population genetic theory is the concept of the Mendelian population, in which recombination occurs at a rate sufficiently

high for the population to evolve as a unit, rather than as an array of genetically independent lineages or clones. In most higher organisms, recombination is coupled with reproduction, and individuals in the same geographic area constitute a Mendelian population. Even for many predominantly "asexual" eukaryotes, rates of recombination resulting from occasional sexual reproduction are high enough to make populations rather than clones the primary units of evolution (MAYNARD SMITH 1978).

In bacteria, the transfer of genes between individuals is an infectious process that is independent of reproduction. Under laboratory conditions, recombination has been shown to occur in many species of bacteria; and in one of them, *Bacillus subtilis*, experiments performed in semi-natural conditions suggest that recombination by means of transformation is involved in short-term adaptation to the environment (GRAHAM and ISTOCK 1978, 1979). However, it is not clear how frequent recombination is in natural populations of any bacterial species. For the common intestinal bacterium *Escherichia coli*, recent studies suggest that recombination occurs at a low rate and that populations exist and evolve as groups of nearly independent clones.

There are two lines of evidence indicating that the rate of recombination is low in natural populations of *E. coli*. The first is based on *in vitro* estimates of the rate parameters of phage adsorption and conjugative plasmid transfer (SCHLESINGER 1932; LEVIN, STEWART and CHAO 1977; LEVIN, STEWART and RICE 1979). Even with substantial rates of mobilization of chromosomal genes, the rate of gene replacement by recombination that is mediated by phage and conjugative plasmid vectors would not be expected to exceed the rate of gene modification by mutation. This argument is considered in detail by LEVIN (1981b).

The second line of evidence is derived from an electrophoretic survey of genetic variation at 20 enzyme loci in 109 strains of *E. coli* isolated from humans and other mammals (SELANDER and LEVIN 1980). In spite of a mean genetic diversity per locus of nearly 50%, five pairs of clones with identical electromorphs for all 20 enzyme loci were isolated from unassociated hosts. Included in this collection of "wild" clones was one that is electrophoretically indistinguishable from the common type of the laboratory strain K12, the ancestor of which was isolated from a human host nearly 60 years ago. SELANDER and LEVIN (1980) attributed the persistence of genotypes to a low rate of recombination, coupled with a purging of genetic variation by periodic selection.

Although these *a priori* and interpretative arguments for a low rate of recombination are strong, the formal possibility remains that conditions for plasmid- and phage-mediated gene exchange are more favorable in natural populations than in laboratory cultures. It is also possible that chromosomal genes are transmitted at a high rate in the process of the spread of plasmids that results from selection for antibiotic resistance or other plasmid-determined characters. Finally, it can be argued that the stability of electrophoretic genotypes is a consequence of natural selection for specific arrays of isozymes, rather than a failure of recombination to disrupt gene complexes.

In an effort to understand the population structure and the role of recombination in the adaptation and evolution of *E. coli*, we have surveyed genetic variation in a population in a natural habitat, the intestine of a human. In this investigation, we employed electrophoresis to examine variation in genes encoding enzymes and in complements of plasmids among the bacterial clones. We present evidence for a large amount of genetic diversity in the *E. coli* population of an individual host, and we confirm earlier reports, based on serotyping, of the existence of both transitory and persistent clones and of rapid changes in the genetic composition of populations (SEARS and BROWNLEE 1952; COOK, EWINS and SHOOTER 1969). Our study provides additional evidence of limited genetic recombination in natural populations of *E. coli*, indicating that most of the genetic diversity in single hosts is a consequence of infection from external sources. The work also supports the hypothesis that "cryptic" plasmids carry genes that play a significant role in the adaptation of *E. coli*.

MATERIALS AND METHODS

Isolation and identification of E. coli: Clones were isolated from fecal samples taken from one of the authors (BRL) over a period of 11 months. In this period, the host had two head colds and a short-term case of intestinal influenza, but was otherwise healthy. He was not under antibiotic treatment during the sampling period or in the five years before the study was made.

Fecal samples on sterile cotton swabs were suspended in buffer and streaked on lactose minimal plates to isolate single colonies. After the plates had incubated at 37° for 24 hr, colonies were chosen without regard to color, size or morphology. To eliminate contaminating bacteria and fungi, the clones were streaked for two cycles on minimal lactose and one cycle on complete medium (tetrazolium lactose). Single colonies from each of the final complete-medium plates were then tested for growth on citrate minimal medium. Those that were citrate negative were stored for electrophoresis. (Less than 1% of the colonies isolated in this way were capable of growing on citrate.) For protein electrophoresis, the clones were stored in stabs and, for DNA electrophoresis, in glycerol at —20°. One clone of each electrophoretically distinct type was tested with the API 20E (TM) system for the identification of Enterobacteriaceae, and all were classified as members of the "species" *E. coli*.

Electrophoresis of enzymes: Each clone was grown overnight in 100 ml of nutrient broth (Difco), supplemented with 0.24 gm/l IPTG (Sigma) to induce the synthesis of β-galactosidase. After being washed with deionized water, the bacterial pellet was suspended in 2 ml of buffer (0.01 M Tris and 0.001 M EDTA, pH 6.8) and sonicated (in pulses, with cooling) for 1 min. The resulting suspension was centrifuged at 37,000 × g for 20 min at 4°, and the supernatant (protein extract) was stored at —70° until used for electrophoresis.

Techniques of starch-gel electrophoresis were similar to those described by SELANDER *et al.* (1971), and 15 enzymes were stained (Table 1). It is likely that all 15 of the enzymes assayed are coded by chromosomal rather than plasmid-born genes. The genes for 11 of them have been identified and mapped (BACHMANN and LOW 1980), and all 15 enzymes were present in several clones in which we were unable to detect plasmids.

Extraction and electrophoresis of plasmid DNA: Clones were grown to late log-phase in Luria broth, at which time 200 μg/ml of chloramphenicol were added for 2 to 3 hr to increase the synthesis of plasmid DNA.

Two standard procedures were used to extract low molecular weight plasmid DNA. In most cases, the technique of BIRNBOIM and DOLY (1979) was employed, with the following modifications: (1) 1.5 ml of cell suspension, rather than 0.5 ml, were used; (2) only one ethanol precipitation was used, which was kept overnight at —20°; (3) 4 μl of a heat-treated 1 μg/ml solution of Ribonuclease A was added to the final DNA suspension to eliminate the low molecular weight RNAs. For some of the samples, the technique of GUERRY, LEBLANC and

TABLE 1

Electrophoretic conditions

Enzyme	Potential (volts)	Distance of migration of amaranth dye (cm)	Buffer system
Phosphoglucomutase (PGM) Malate dehydrogenase (MDH) β-Galactosidase (βGA) 6-Phosphogluconate dehydrogenase (6PG)	130	9	Tris-citrate + NADP, pH 8.0
Phenylalanyl-leucine peptidase (PE2) Leucyl-glycyl-glycine peptidase (PE3)	130	8	Tris-citrate, pH 8.0
Isocitrate dehydrogenase (IDH) Phosphoglucose isomerase (PGI) Aconitase (ACO)	130	10	Tris-citrate, pH 6.7
Mannose-6-phosphate isomerase (MPI) Glutamic oxaloacetatic transaminase (GOT)	325	8	Lithium hydroxide, pH 8.2
Adenylate kinase (AK)	100	7	Phosphate, pH 7.0
Glyceraldehyde-3-phosphate dehydrogenase (G3P) Glucose-6-phosphate dehydrogenase (G6P) Alcohol dehydrogenase (ADH)	250	10	Poulik, pH 8.7

FALKOW (1973) was used. In both procedures, the final pellet was resuspended in TES buffer (50 mM NaCl, 5 mM EDTA, 30 mM Tris, pH 8.0). Both methods gave similar results with respect to plasmid banding on gels.

For electrophoresis of plasmid DNA, we used 0.7% agarose gels and a procedure similar to that described by MEYERS *et al.* (1976).

RESULTS

Electrophoretic variation: In total, 550 clones were isolated and examined in the 11-month sampling period. These represent 53 distinct combinations of electromorphs (designated, henceforth, as electrophoretic types or ETs), which are shown, in order of their first occurrence, in Table 2. Two of the 15 enzymes, GOT and G3P, were monomorphic, and the number of electromorphs identified for the remaining 13 enzymes ranged from 2 for PE3, MPI and PGM to 12 for βGA. The mean number of electromorphs per locus was 4.1. Mean genetic diversity per locus among the 53 ETs was 0.39, which is only 20% less than that found by SELANDER and LEVIN (1980) among 109 clones representing 98 ETs isolated from 89 individuals of a large variety of mammalian species. ET-2, which was represented by nine clones collected on four days, is electrophoretically identical to laboratory strain K12, but it is not sensitive to most K12 phages and does not carry the genetic markers of the K12 strains used in our laboratories.

TABLE 2

Electromorph genotypes in 550 clones of E. coli *from a single human host*

Electrophoretic type	N*	D†	Enzyme locus‡												
			PGI	IDH	PE2	PE3	AK	G6P	βGA	6PG	MPI	PGM	ACO	ADH	MDH
1	15	2	5	5	4	3	3	4	2	5	7	5	4	1	2
2	9	4	3	2	4	5	3	4	15	5	3	5	6	1	2
3	5	1	3	2	7	5	3	1	15	17	3	5	6	4	2
4	2	1	3	2	4	5	3	4	22	5	3	5	6	1	2
5	2	1	3	2	4	3	3	4	22	5	3	5	6	1	2
6	2	1	3	2	4	5	3	1	22	5	3	5	6	1	2
7	1	1	3	2	7	3	3	4	22	13	8	5	6	4	2
8	1	1	3	2	7	3	3	4	22	5	7	5	6	1	2
9	1	1	3	5	4	3	3	4	15	5	8	5	6	4	2
10	1	1	3	2	4	3	3	4	15	17	3	5	6	1	2
11	1	1	3	2	7	5	3	4	15	5	7	5	6	1	2
12	252	16	6	5	4	3	4	4	6	3	7	1	6	5	4
13	91	11	3	2	4	5	3	4	15	4	3	5	6	1	2
14	1	1	3	5	7	5	3	4	6	5	3	5	5	1	2
15	2	2	3	5	7	3	3	4	15	5	3	5	6	1	2
16	1	1	3	5	7	3	3	4	15	5	3	1	6	4	2
17	1	1	3	5	7	3	3	4	15	5	3	5	5	1	2
18	1	1	6	5	4	5	3	4	2	5	9	5	5	5	2
19	2	2	3	1	7	5	3	1	15	5	3	5	6	1	2

TABLE 2—Continued

Electrophoretic type	N*	D†	Enzyme locus‡												
			PGI	IDH	PE2	PE3	AK	G6P	βGA	6PG	MPI	PGM	ACO	ADH	MDH
20	30	6	3	5	7	5	3	1	15	8	3	5	6	1	2
21	3	2	6	6	4	3	1	4	12	5	3	5	5	5	2
22	1	1	6	6	4	3	1	4	15	5	3	5	5	5	2
23	1	1	6	5	4	3	3	4	12	5	8	5	4	5	4
24	1	1	3	2	7	3	3	4	22	11	8	5	5	4	2
25	1	1	3	2	4	5	3	4	15	18	3	5	6	1	2
26	1	1	3	4	4	5	3	4	15	4	3	5	6	1	2
27	2	1	3	5	4	5	3	4	15	5	3	5	6	1	2
28	1	1	3	5	4	3	3	5	15	5	9	5	4	5	2
29	1	1	6	5	4	5	3	4	12	16	9	5	5	5	2
30	1	1	6	5	4	3	3	5	15	5	9	5	5	5	2
31	9	1	3	2	4	3	3	4	19	5	7	5	6	4	2
32	5	1	3	2	7	5	3	1	15	5	3	5	6	1	2
33	18	1	3	2	7	3	3	4	20	5	8	5	6	1	2
34	1	1	3	2	7	5	3	4	15	5	7	5	6	4	2
35	2	1	3	2	4	3	3	4	23	8	3	5	6	1	2
36	3	1	3	2	4	3	3	4	10	5	8	5	6	1	2
37	1	1	3	2	4	3	3	4	22	5	8	5	6	1	2
38	1	1	3	5	7	3	3	1	15	5	8	5	6	1	2
39	2	1	6	5	4	5	3	5	2	8	8	5	5	5	2
40	1	1	3	5	4	3	5	4	22	5	8	5	8	4	2

TABLE 2—Continued

Electrophoretic type	N*	D†	Enzyme locus‡												
			PGI	IDH	PE2	PE3	AK	G6P	βGA	6PG	MPI	PGM	ACO	ADH	MDH
41	1	1	3	2	4	3	3	4	19	14	7	5	6	1	2
42	11	2	6	2	4	3	1	4	15	5	9	5	5	5	2
43	22	2	3	5	4	5	3	4	15	7	3	5	6	4	2
44	10	2	3	5	4	3	3	5	2	8	9	5	5	5	2
45	4	2	3	5	2	3	3	4	22	8	8	5	5	1	2
46	2	2	6	5	4	5	3	4	11	5	9	5	6	2	2
47	3	2	6	5	4	3	1	4	15	5	9	5	5	5	2
48	3	1	6	2	4	3	1	4	15	5	9	5	5	1	2
49	4	1	6	5	4	5	3	4	5	5	7	5	5	5	2
50	1	1	6	5	4	3	3	4	15	5	9	5	5	5	2
51	1	1	6	4	4	3	1	4	15	5	9	5	5	5	2
52	2	1	6	5	4	5	3	5	14	5	9	5	6	5	2
53	1	1	4	5	4	5	3	4	2	5	9	5	6	1	2
Number of alleles per locus Mean = *4.1*			4	5	3	2	4	3	12	11	4	2	4	4	2
Genic diversity§ Mean = *0.3867*			0.472	0.583	0.432	0.486	0.266	0.351	0.713	0.499	0.717	0.073	0.529	0.607	0.073

* N = Number of clones representing each electrophoretic type.
† D = Number of days on which an electrophoretic type was isolated.
‡ Numbers designate electromorphs of each enzyme, ranked in order of decreasing electrophoretic mobility. The numbering system is cognate with that previously employed by SELANDER and LEVIN (1980).
§ Genic diversity for a locus is calculated as $1 - \sum_{i}^{e} x_i^2$, where x_i is the frequency of the ith electromorph.

Sampling variation: We would have liked to have obtained random samples of the *E. coli* population from various parts of the intestinal tract of the host, but owing to the difficulty and inconvenience of taking repeated samples *in situ*, we sampled the clones present in the feces. In an attempt to determine the magnitude of error that this procedure engenders, samples were obtained in three ways: (1) by inserting sterile cotton-tipped swabs into the anus shortly after defecation (swab samples); (2) by suspending a day's output of feces in a sterile saline solution and rapidly agitating the suspension to release bacteria (full fecal sample); and (3) by inserting cotton-tipped swabs into different parts of a single fecal mass (fecal probes). The number of ETs obtained by each of these procedures is shown in Table 3.

Our study did not demonstrate an association between the number of clones analyzed and clonal diversity. For example, 11 ETs were observed in a swab sample of 21 clones taken on May 20, but only one ET was found in a sample of 74 clones isolated from 7 probes of a fecal mass on February 16. From the full fecal sample (taken on February 9), only 13 ETs were identified among the 50 clones isolated. For all samples, the coefficient of correlation between number of clones and number of ETs is 0.227, which is not significant ($P > 0.05$).

Although the data are not sufficient for a definitive evaluation of the relative merits of the various sampling techniques, it is apparent that the swab technique is not grossly inferior to the others. Better estimates of total diversity might be obtained by full fecal sampling, but the swab technique has the advantage of convenience.

Data from the fecal probe samples provide some evidence of spatial and/or temporal heterogeneity in the intestinal flora. As shown in Table 4, the distribution of ETs along the fecal mass sampled on March 7 was nonrandom.

Finally, it is important to note that, because selective media were not used in isolating clones, our sampling procedures would detect only those ETs represented by large populations.

Temporal variation in electrophoretic types: The dates of appearance and numbers of clones of each isolated ET are indicated in Table 3. Except for some ETs that were present in the 13-day period in January, when samples were taken daily, most ETs appeared in only one sample and were represented by only one or a few clones. Three ETs, however, were observed over extended periods. One of these persistent (or resident) types, ET-12, was first detected in August, 1979, and continued to be recovered until the end of the sampling period in March, 1980. Some 252 of the 550 clones isolated were ET-12. The second resident type, ET-13, was present in August and for a good part of the January sampling period, but it was not detected after January, despite the fact that more than 200 additional clones were examined. ET-2 was isolated on only four days, but these were distributed over a major part of the sampling period, from May, 1979, to February, 1980.

The overall pattern of temporal variation in clonal composition was as follows: (1) The resident clone ET-12 fluctuated markedly in abundance. It was common on the first 6 days of the 13-day sampling period in January, then all but dis-

TABLE 3

Number of clones of each electrophoretic type recorded on 22 sampling dates

	Date and number of clones analyzed																					
ET	5/17 11	5/20 21	8/22 20	1/12 16	1/13 20	1/14 20	1/15 20	1/16 14	1/17 20	1/18 19	1/19 20	1/20 20	1/21 20	1/22 20	1/23 18	1/24 19	2/4 20	2/9* 50	2/16† 74	3/5† 35	3/6† 36	3/7† 35
1	11	4																				
2		1								4	3							1				
3		5																				
4		2																				
5		2																				
6		2																				
7		1																				
8		1																				
9		1																				
10		1																				
11		1																				
12			17	14	16	20	7	13	20		1				2	8	20	5	74	35	7	1
13			3	1	3		12			9	2	7	16	19	14	3						
14				1																		
15					1			1														
16							1															
17										1												
18										1												
19										1	1											
20										3	11	10	3		3	1						
21											2					1						
22												1										
23												1										
24												1										
25													1									
26														1								

TABLE 3—Continued

ET	Date and number of clones analyzed																					
	5/17 11	5/20 21	8/22 20	1/12 16	1/13 20	1/14 20	1/15 20	1/16 14	1/17 20	1/18 19	1/19 20	1/20 20	1/21 20	1/22 20	1/23 18	1/24 19	2/4 20	2/9* 50	2/16† 74	3/5† 35	3/6† 36	3/7† 35
27																2						
28																1						
29																1						
30																1						
31																		9				
32																		5				
33																		18				
34																		1				
35																		2				
36																		3				
37																		1				
38																		1				
39																		2				
40																		1				
41																		1				
42																					8	3
43																					11	11
44																					2	8
45																					3	1
46																					1	1
47																					1	2
48																					3	
49																						4
50																						1
51																						1
52																						2
53																						1

* Full fecal sample (FFS); †fecal probe (FP); all other samples taken by swab (S).

TABLE 4

Number of clones of various electrophoretic types in fecal probe samples taken on three successive days

Date	Sample	Electrophoretic type												
		12	42	43	44	45	46	47	48	49	50	51	52	53
3/5/80	1	12												
	2	12												
	3	11												
3/6/80	1	3	4	4	1									
	2	1	3	4		3	1							
	3	3	1	3	1			1	3					
3/7/80	1			7						4	1			
	2	1	2	2	4	1	1	1						
	3		1	2	4			1				1	2	1

appeared between days 7 and 12, but subsequently again became predominant. It is noteworthy that in the week (January 18–23) when ET-12 was at low frequency, ET-13, the other common resident clone, increased in frequency and several new ETs appeared. (2) Except for the residents ET-12 and ET-13 and the recurrent ET-2, there was a complete replacement of ETs in a period of two weeks, from January 12–17 to January 18–24. From January to February, the turnover was also complete, and, again, there was a replacement of ETs from February to March.

Plasmid diversity: A DNA gel showing representative clones of the resident ETs isolated at different times in the sampling period is shown in Figure 1; Figure 2 is a diagram of the plasmid profiles (PP) of these clones. PPs of the transitory ETs are shown in Figure 3. *E. coli* strain χ2556, which carries eight plasmids of known molecular weight (MACRINA *et al.* 1978), was run on each gel, and the molecular weights of plasmids from "wild" *E. coli* clones were estimated from their mobilities relative to those of this strain.

To determine which bands represent covalently closed-circle plasmid DNA, we applied the criterion described by BIRNBOIM and DOLY (1979). However, it is possible that some bands represent nicked circles or other denatured forms of plasmids. In the case of very low molecular weight plasmids, the intensity of bands occasionally varied among preparations of the same clone, apparently as a result of variation in degree of preciptation. For this reason, we attached no significance to the absence of these bands.

Clones of a given ET that were isolated on the same day usually had the same PP. There were, however, a few exceptions. One of the three clones of ET-13 collected on August 22, 1979, differed from the other two in having two bands of slightly different mobility, and a similar phenomenon was noted among clones of ET-2 isolated in January. There was also one case of a transitory ET showing heterogeneity in plasmid profile among clones collected on the same day.

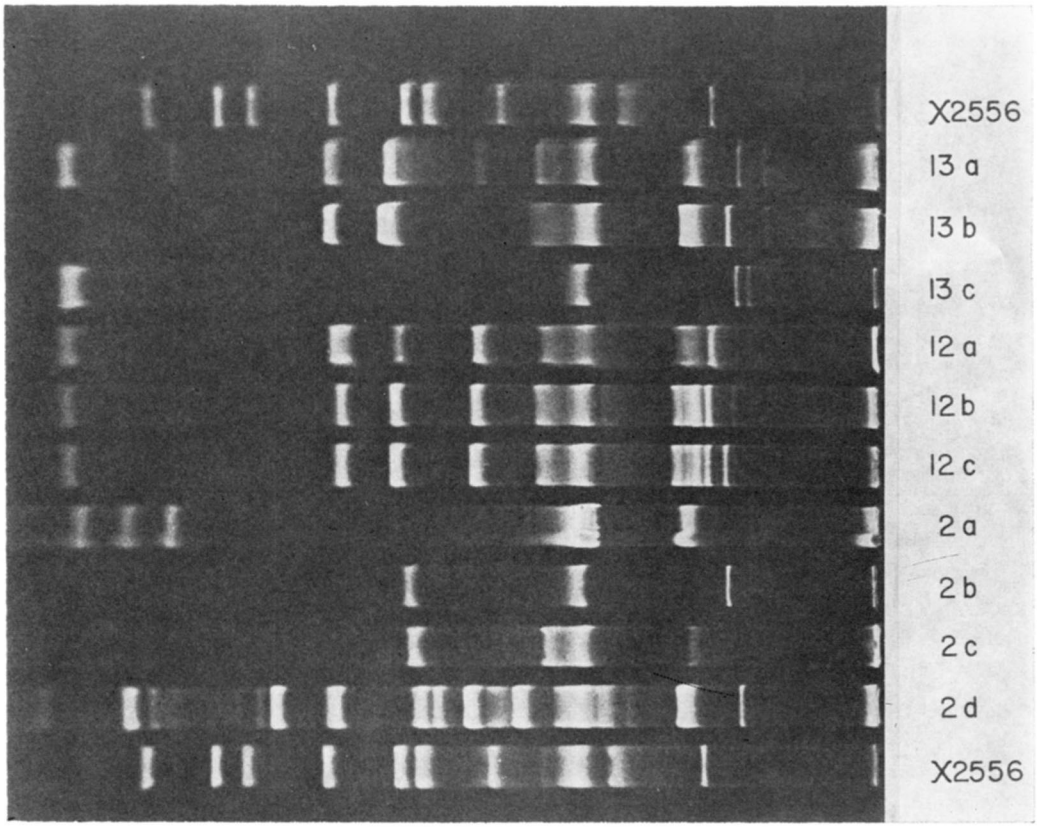

FIGURE 1.—Agarose gel of the low molecular weight (plasmid) DNA of the resident and recurrent electrophoretic types ET-12, ET-13 and ET-2, and molecular weight standard χ2556. DNA for this gel prepared by procedure of GUERRY, LEBLANC and FALKOW (1973).

As shown in Figures 1 and 2, the PPs of each of the three persistent ETs changed in the 11-month sampling period. For ET-2 and ET-13, plasmid composition was completely altered. If we were to consider only the PPs, we would regard ET-2 and ET-13 as six or seven distinct types of *E. coli*, each of which had a shorter tenure in the host than that suggested by the protein data alone. For the resident ET-12, the temporal changes in PP were considerably less dramatic. Between August, 1979, and March, 1980, high molecular weight plasmid bands were added to the PP, but the other bands remained unchanged.

Of the 50 transitory ETs, all but four had at least one plasmid band (see Figure 3). The PPs of most of these ETs were unique; and in those cases where pairs of ETs had indistinguishable PPs, the pair members were very similar electrophoretically.

DISCUSSION

Earlier work on populations of *E. coli* in individual humans, by the use of serotyping, led to the recognition of resident and transient strains (SEARS, BROWN-

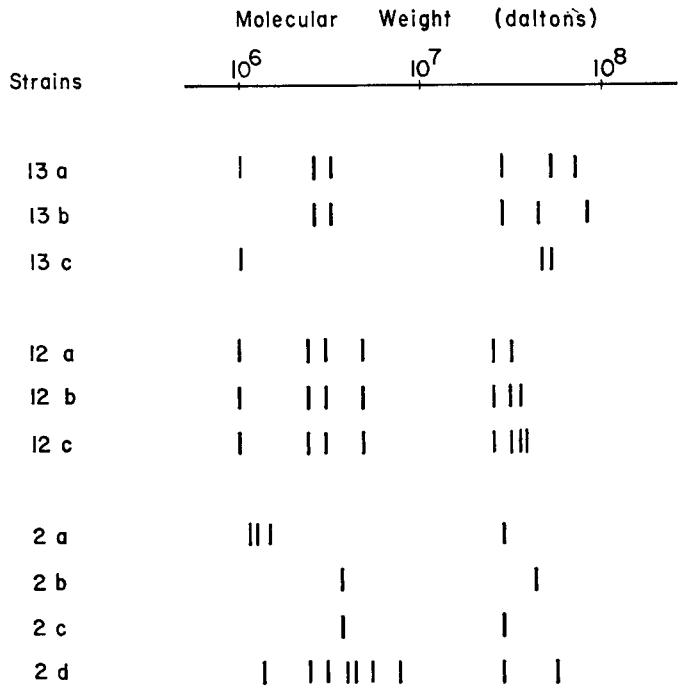

FIGURE 2.—Diagram of the plasmid profiles (PPs) of the resident and recurrent electrophoretic types ET-12, ET-13 and ET-2.

LEE and UCHIYAMA 1950; SEARS and BROWNLEE 1952). Resident strains persist for months or years, and transients are present for only a few days or weeks. SEARS *et al.* (1956) determined that individual humans (and dogs as well) harbor only one or two resident strains at any one time; the results of our electrophoretic analysis are fully in agreement. SEARS, BROWNLEE and UCHIYAMA (1950) reported that transient populations rarely consist of more than three or four strains on any one day; and SEARS and BROWNLEE (1952) usually identified no more than 10 serotypes from an individual, even in studies extending over periods of several years. In contrast, we were able to distinguish a dozen or more electrophoretic types on a single day and a total of 53 over an 11-month period. The difference presumably reflects the greater power of protein electrophoresis to distinguish clones, but it remains possible that this host carried a more diverse array of genotypes than did most hosts sampled in earlier studies. The number of distinct types of *E. coli* is further increased when the data on plasmid profiles are considered.

The serotype studies of SEARS and BROWNLEE (1950) and COOKE, EWINS and SHOOTER (1969) indicated that the genetic composition of the *E. coli* population of a host changes dramatically over short periods of time. The present electrophoretic analysis clearly confirms this finding.

Sources of clonal diversity: The primary question raised by our finding of large numbers of electrophoretic types and rapid changes in clonal structure is,

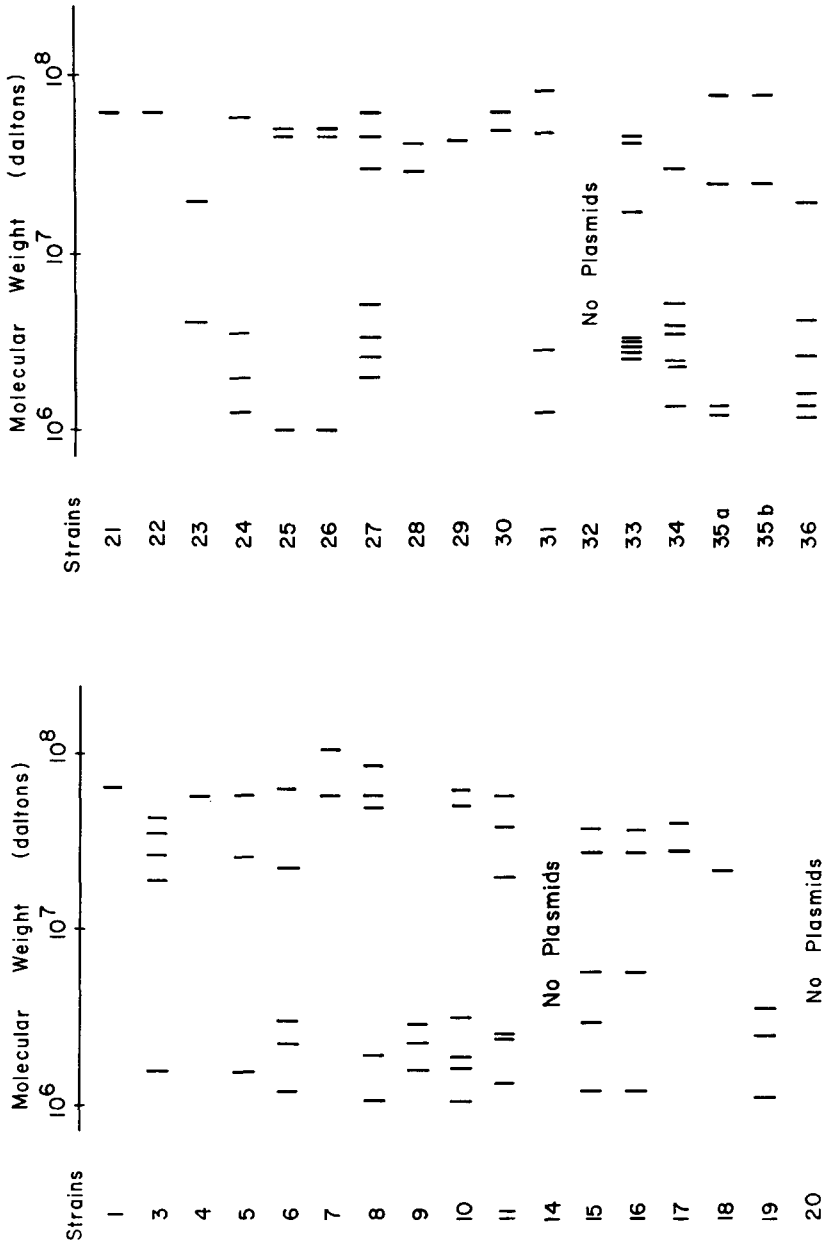

FIGURE 3.—Diagram of the plasmid profiles (PPs) of the transient electrophoretic types.

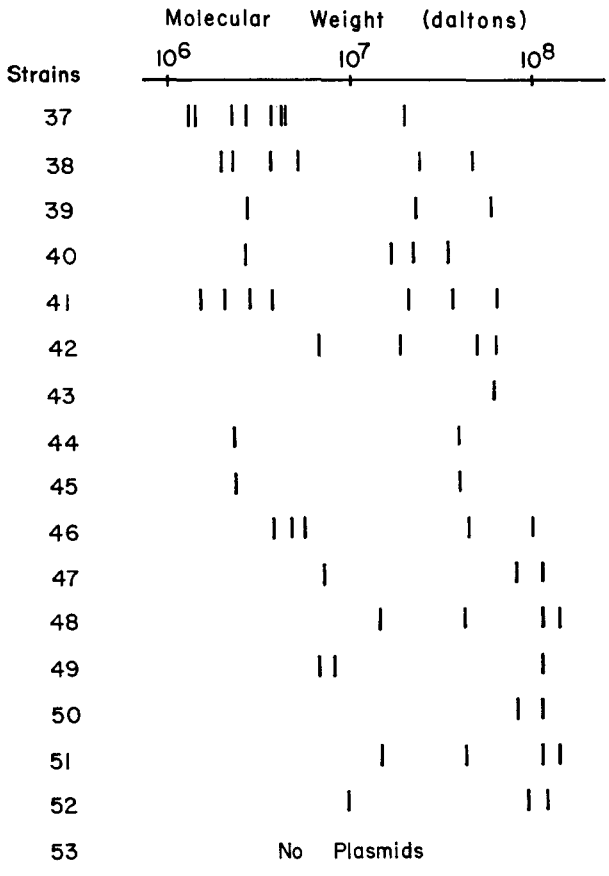

Figure 3, continued

what are the proximate sources of the genetic diversity of the *E. coli* population in a host? MILKMAN (1975), observing several combinations of electromorphs at three polymorphic loci in clones from individual mammalian hosts, concluded that "the likeliest explanation is recombination within each host." Extending this hypothesis of the *in situ* generation of diversity, we could also include mutation occurring within the host as a second source of new genotypes.

An alternative hypothesis, which we will call the immigration model, is that most of the observed clonal diversity in a host results from repeated infection from environmental sources. According to this view, relatively little of the diversity in the population of a host is generated *in situ*.

The case for low rates of recombination: If recombination in *E. coli* occurred at rates sufficiently high to overcome periodic selection (associated linkage) (ATWOOD, SCHNEIDER and RYAN 1951) and to render a population essentially "panmictic," as suggested by MILKMAN (1975), we would expect a nearly random distribution of electromorphs among the ETs. But this is not what we observed. Rather, there is strong linkage disequilibrium between several pairs of loci, as

shown in Table 5. Of the 78 tests of association performed (some of which involved the combining of electromorphs of similar mobility to satisfy cell-size requirements, which weakens the test), 16 yielded P values of less than 0.01.

The nonrandomness of the distribution of electromorphs can also be seen in an analysis of pairwise comparisons of all 53 ETs with respect to the number of loci by which they differ, together with the distribution of differences between pairs in a theoretical population consisting of genotypes formed randomly from pools of electromorphs in which the electromorph frequencies were the same as in the sample of 53 ETs examined (Figure 4). In the empirical sample, there are too many pairs of ETs that differ at only one or two loci, and too many that differ at many loci. Among the latter is the resident of ET-12, for which the average number of differences from other ETs is 9.1 loci, with a range of from 6 to 12.

Although these results clearly demonstrate that recombination does not occur frequently enough to randomize the distribution of electromorphs, it is possible that some of the observed diversity was generated in the host by recombination. In attempting to estimate how much of the within-host genetic diversity is generated by *in situ* recombination, we have to consider the fact that gene transfer in *E. coli* is a nonreciprocal process (only recipients are recombinants) and that few, very closely linked genes would be transferred in any given exchange event (see HAYES 1968). For the set of 15 enzyme loci we have assayed, co-transfer of more than one gene would be unlikely; therefore, we assume that single *in situ* recombination events would produce only single electromorph changes in genotype. Additionally, we assume that no more than one gene-transfer event would have occurred for any single cell lineage. With these considerations in mind, we have employed the following criteria in attempting to assess the magnitude of *in situ* recombination: (1) Where pairs of ETs differ at a single protein locus, but the members have very different plasmid profiles, we attribute the difference to

TABLE 5

Significance levels for tests of linkage disequilibrium between pairs of electrophoretic types of E. coli

Locus	IDH	PE2	PE3	AK	G6P	βGA	6PG	MPI	PGM	ACO	ADH	MDH
PGI	*0.004*	*0.003*	0.549	0.512	0.145	0.011	0.495	*0.001*	0.512	*0.000*	*0.000*	0.055
IDH		0.249	0.731	0.275	0.101	*0.001*	0.076	0.232	0.275	*0.003*	*0.005*	0.275
PE2			0.626	0.626	*0.007*	0.655	0.680	0.106	0.412	0.111	*0.008*	0.626
PE3				0.364	0.107	0.081	0.638	0.052	0.364	0.065	0.446	0.364
AK					0.967	0.182	0.833	0.415	*0.005*	0.430	0.285	*0.001*
G6P						0.430	0.765	0.101	0.967	0.043	0.021	0.967
βGA							0.698	0.018	0.368	0.032	0.056	0.368
6PG								0.966	0.833	0.329	0.470	0.833
MPI									0.639	*0.010*	*0.002*	0.415
PGM										0.430	0.285	*0.001*
ACO											*0.000*	0.665
ADH												0.211
MDH												

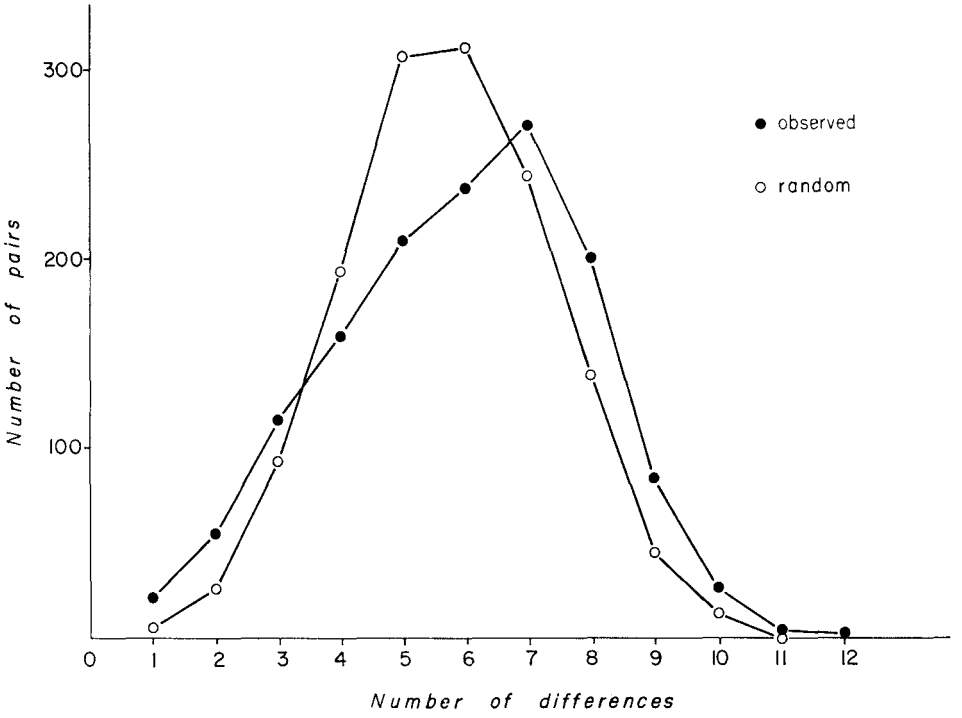

FIGURE 4.—Distribution of differences at 15 loci in 1378 pairwise comparisons of 53 ETs.

divergence occurring before the host was infected by the strains, rather than to *in situ* recombination or mutation. (2) Where two ETs differ at a single locus and have identical or nearly identical plasmid profiles (differing by, at most, one plasmid band), the difference in electromorphs is attributed to *in situ* mutation rather than to recombintion, provided that neither electromorph was observed in other clones in the sample. (3) Where two ETs differ at a single locus and have nearly identical plasmid profiles, and both electromorphs were represented in other ETs, we consider recombination to be a more likely source than mutation of the electromorph difference.

Because *E. coli* can exchange genes with other species of bacteria, it could be argued that some of the electromorphs that we have attributed to mutation in *E. coli* actually were transferred through recombination with other species. It seems unlikely, however, that interspecific recombination is an important source of the variation that we observed, since rates of recombination between species, even closely related forms such as *Shigella flexneri* and *E. coli* under laboratory conditions, are much lower than within species (FALKOW *et al.* 1963; SCHNEIDER and FALKOW 1964).

Of the 1378 pairwise combinations of the 53 ETs we identified, 21 differed at a single locus. Applying the above criteria for the *in situ* generation of the ET differences, we find that only 3 of these meet all criteria for recombination being more likely than mutation; and in the case of three other pairs, mutation is more

likely than recombination (see Table 6). For the remaining pairs, we believe that divergence from a common ancestor occurred before the infection of the host.

The array of electromorphs of the resident ET-12 provides additional support for the hypothesis that *in situ* recombination is rare and further suggests that all strains are not equally likely to serve as donors and recipients of genes (also see CURTISS 1969; JONES and CURTISS 1970). Although ET-12 was present in most samples (see Table 2) and was represented by 252 of the 550 clones examined, the most similar genotype (ET-23) differed from it by six loci and had a distinctive plasmid profile. Hence, it is unlikely that this or any of the other ETs arose by gene transfer to ET-12 or by mutation of ET-12. Furthermore, we found no evidence that ET-12 served as a donor of genes; potentially, we might have detected such a transfer because ET-12 carries unique electromorphs at two loci (6PG and AK; see Table 2).

The second resident strain, ET-13, is a member of a complex of strains that are closely related to the laboratory strain K12. This complex includes ET-2, which is indistinguishable electrophoretically from K12, and also ET-25 and ET-26. ET-26 differs from ET-13 in having a "null" (4) *versus* 2 electromorph at IDH. Since null IDH was recorded in only one other ET, which was isolated in March after ET-13 had disappeared, we presume that the change from 2 to null represents a mutation. ET-25 differs from ET-13 in having an 18 rather than a 4 electromorph at 6PG; because electromorph 18 is unique, we again attribute the change to mutation. ET-13 differs from ET-2 in having a 5 rather than a 4 electromorph at 6PG, but since their plasmid profiles are very different, we cannot infer that one of these ETs was directly derived from the other. Thus, since the resemblances between the resident ET-13 and other ETs are most readily explained by mutation, there is no need to invoke recombination.

It is of some interest that, for an unexpectedly large number of ETs that differ by single enzymes, the pair members were collected on the same day. For reasons presented above, we do not believe that many of these pairs were generated by mutation or recombination in the host. Rather, we suggest that most or all of

TABLE 6

Comparison of ETs differing at only one locus

ET pair	Locus	Electromorph	Number of plasmids*	Number of identical plasmids	Probable origin of difference
42–51	IDH	2 — 4†	4 — 4	4	Mutation
13–26	IDH	2 — 4†	3 — 3	3	Mutation
13–25	6PG	4 — 18‡	3 — 3	3	Mutation
42–48	ADH	5 — 1	4 — 4	4	Recombination
30–50	G6P	5 — 4	2 — 2	1	Recombination
47–50	AK	1 — 3	3 — 2	2	Recombination

* In cases where an ET has several plasmid profiles, the ETs used for the pairwise comparison were those with the most similar profiles.

† The number of the electromorph corresponds to a "null" allele.

‡ Unique electromorph.

them were derived from common ancestors by these processes at an earlier time, and that the co-occurence of genetically similar types is a consequence of local selection for clones of the same general genotype, or of the occurrence of related forms in the source populations.

Among the few ETs for which we cannot rule out *in situ* recombination as a source of variation, some are one- or two-enzyme variants of ET-13, which differs from the common form of the laboratory strain *E. coli* K12 at only one of the 15 enzyme loci (Table 6). (As noted earlier, ET-2 is identical to the common K12 for all fifteen enzymes.) That many "wild" *E. coli* are genetically similar to K12 is consistent with the view that the ancestor of this laboratory strain is a common *E. coli* type in the human intestinal tract (SELANDER and LEVIN 1980). However, it should be noted that the plasmid profiles of many of these K12-like ETs differ markedly from each other and from the profile of laboratory K12 (which may or may not have a single large plasmid, F); these "wild" K12s are resistant to most of the phage that attack laboratory strains of K12.

The case for immigration: We now turn to a consideration of immigration as an explanation for the extensive clonal diversity that we have observed. Several lines of evidence support the interpretation that much of the diversity in human hosts arises through repeated colonization of the intestinal tract from outside sources. First, it has been demonstrated that humans frequently acquire strains of *E. coli* from food and water. In a study in a hospital ward, COOKE *et al.* (1970) found that 63 of 873 samples of food tested contained *E. coli* in concentrations ranging from 25 to 10^4 per gram. Significantly, *E. coli* serotypes recovered from fecal samples of patients in the ward were generally similar to those present in the food, and several cases of contamination of patients by specific serotypes in food were recorded. COOKE *et al.* (1970) concluded that temporal variation in the genetic structure of *E. coli* populations in the patients was caused in large part by changes in the serotypes occurring in food.

Experimental attempts to establish new strains of *E. coli* in humans and animals have yielded varying results. COOKE *et al.* (1970) mentioned experiments in which the ingestion of 10^4 or more *E. coli* in milk led to the presence of a strain in the feces for periods of from two to eight weeks. However, two attempts by SEARS, BROWNLEE and UCHIYAMA (1950) to establish strains resulted in their presence in feces for only a few days. ANDERSON (1975) showed that, if ingested in sufficiently large numbers, even a "weakened" laboratory strain of K12 can survive for several days in the human intestine and may even multiply there for a short time (see also LEVY, MARSHALL and ROWSE-EAGLE 1980).

Another line of support for the immigration model is an absence of any particular similarity between ETs collected in consecutive months. This may be illustrated by comparing the 14 ETs isolated in the latter part of the 13-day January sampling period with the 11 ETs recorded in February (Table 3). In this analysis, we used the Euclidean distance between pairs of ETs, based on all polymorphic loci, with the potential contribution of all loci to the distance being equalized by a ranging transformation. For the pairwise comparisons of the January and February ETs, mean distance was 1.257; and for a corresponding set of 100 randomly

selected pairwise comparisons of ETs, mean distance was 1.244. Thus, the appearance of a new set of strains in February was more probably a result of colonization than of mutational and recombinational transformation of the strains present in January.

A third line of evidence involves changes in the plasmid profiles of single ETs that were isolated at different times (Figures 1 and 2). ET-2 was observed in three widely-spaced sampling periods, and the clones isolated in each period carried a unique plasmid complement. Because these three very different profiles cannot readily be constructed from one another by the simple transfer of plasmids, it seems likely that the ET-2 clones with different profiles were independently derived from different contaminated foods or other sources. Even the radical change in plasmid profile that occurred in the resident ET-13 can more readily be accounted for by independent invasions of clones carrying different arrays of plasmids than by the *in situ* transfer of plasmids.

Finally, the immigration hypothesis is also supported by the sharing of ETs among individuals of the same family or those otherwise closely associated. A study of the flora of the BRL family demonstrated extensive sharing of ETs among family members and even household pets (CAUGANT, LEVIN and SELANDER, in preparation; also see SEARS and BROWNLEE 1952, on infant twins).

In summation, there is reason to believe that immigration occurs frequently enough to produce most of the diversity and changes in the clonal composition of the *E. coli* population that we have observed in a single host. This diversity cannot be readily accounted for by *in situ* recombination and mutation.

Plasmid diversity: On the basis of earlier observations on "wild" *E. coli* (M. RICHMOND and R. CURTISS III, personal communication), we expected DNA electrophoresis to reveal the presence of plasmids in many of the clones. Nevertheless, since the host was healthy and had not been under antibiotic treatment, the high frequency of plasmid-bearing cells, the large number of plasmids carried by individual ETs and the diversity of plasmids in the population at large were surprising. Though all but one of these plasmids are "cryptic" in the sense that we do not know the phenotypes for which they code (see the APPENDIX for a description of a fortuitous exception), it is likely that most of them code for specific bacterial-host functions and are maintained in the population by selection for these characters. This interpretation is based on both *a priori* considerations of the conditions for the existence of plasmids and circumstantial evidence obtained in this survey and in other studies.

Theoretically, there are conditions under which conjugationally transmitted plasmids that are not favored by selection can become established in bacterial populations (STEWART and LEVIN 1977; LEVIN and STEWART 1980). A necessary condition for the existence of unselected and negatively selected plasmids is that their rate of infectious spread by conjugation exceeds their rate of loss by vegetative segregation and selection against the cells carrying them. For self-transmissible, conjugative plasmids, these conditions are relatively broad; for plasmids that transfer at rates similar to those that are permanently derepressed for conjugative-pili synthesis (see LEVIN, STEWART and RICE 1979), these condi-

tions have been considered biologically realistic (STEWART and LEVIN 1977). However, most naturally occurring plasmids are repressed for conjugative-pili synthesis (MEYNELL 1973), and in heterogeneous natural populations, unlike those considered in the mathematical models upon which this theory is based, only a portion of the cells not carrying a plasmid could serve as recipients for it. For these and other reasons, LEVIN (1981a, in press) suggested that it is highly unlikely that "unselected" conjugative plasmids would be maintained in bacterial populations.

For nonconjugative plasmids, in the absence of positive selection, the conditions for existence are narrow and are unlikely to obtain, even if these non-self-transmissible plasmids are readily mobilized and are transmitted by very fertile conjugative elements (LEVIN and STEWART 1980). Thus, based on existing theory and estimates of parameters, we would not expect plasmids to be maintained for long periods of time unless there were at least occasional episodes when selection favored cells carrying them over cells without plasmids. This condition would, of course, exist only when a plasmid codes for a phenotype that enhances host-cell fitness.

The plasmids detected by our DNA electrophoresis procedure ranged in size from 1 to 80 megadaltons. Even if those represented by the lower molecular weight DNA bands are frugal in the number of genes committed to insuring their replication, they could have only a few genes available for the coding of characters that enhance host-cell fitness. But many of the higher molecular weight "cryptic" plasmids, which are as large as or larger than known colicinogenic plasmids and conjugative R-plasmids, could have sufficient DNA to code for many characters that could enhance the fitness of their hosts.

One line of circumstantial evidence supporting the hypothesis that these plasmids are under selection is provided by changes in the PPs of the resident and recurrent ETs. In the cases of ET-2 and ET-13, the temporal changes in PPs were sufficiently radical to suggest that clones of ET-2 and ET-13 with different arrays of plasmids had colonized the human host at different times. These different arrays of plasmids presumably would provide phenotypes that are favored in different environmental conditions. For the resident ET-12 clones, the temporal changes in PP were more subtle, suggesting a sequential addition of higher molecular weight plasmids to an array already carried by the ET-12 cell line. Possible sources of the new plasmids are a transposon added to one of the existing ET-12 plasmids (BENNETT, RICHMOND and PETROCHEILOU 1980) and conjugative plasmids transmitted from other cells in the host. Regardless of the source of the added plasmids, we would not expect most cells of the resident strain to carry them unless they conferred a selective advantage.

A recent study by LEVY, MARSHALL and ROWSE-EAGLE (1980) of the survival of *E. coli* host-vector systems in the human intestine offers additional support for the hypothesis that cryptic plasmids enhance host-cell fitness. In this investigation, *E. coli* K12 cells carrying a nonconjugative pBR322 plasmid had a higher survival rate than did members of the same cell line not having the plasmid. There are no known genes on pBR322 that can account for this increased rate of survival.

Environment and adaptation: Our study and an earlier one based on serotyping (SEARS, BROWNLEE and UCHIYAMA 1950) indicate that the short-term adaptation and evolution of populations of *E. coli* in human hosts occur primarily through clonal replacement, a process about which little is known, either from a genetic or ecological standpoint. Among the questions raised by research on the genetic structure of populations in human hosts are the following: What are the ecological and genetic circumstances under which a particular genotype becomes established as a resident, and what environmental changes in the intestine lead to fluctuations in the population size of a resident strain and its eventual replacement by another? SEARS, BROWNLEE and UCHIYAMA (1950) found that a change in resident strains is sometimes accompanied by diarrhea, but that resident strains may persist in spite of this. Are there, in fact, two ecologically and biochemically different groups within the species *E. coli*, one consisting of strains that normally are residents in humans and other mammals and another composed largely of transients, or do all strains have the potential of becoming residents, given the appropriate conditions? Do co-resident strains occuy different regions of the intestine or different "niches" in the same regions? How does the ecology of a resident differ from that of a transient? Do resident strains show greater temporal stability of the genotype than do transient strains as a consequence of relatively strong normalizing selection in their "niches"?

Long-term studies now in progress on the *E. coli* flora of family members and hosts who have experienced extinction and re-establishment of their *E. coli* flora may provide answers to some of these questions.

We thank LOUISE BAND for technical assistance, DAVID FOLTZ for writing a computer program, ROY CURTISS III and JON GOUGEN for providing the χ2556 strain, and PAT PATTISON for typing the manuscript. The senior author was supported by a research grant from the Ministère des Affaires Etrangères, France. This research was supported by Public Health Service Grant GM-19848 and Public Health Service Research Career Development Award 1 K04 GM-00112 to BRL and Public Health Service Grant GM-22126 to RKS.

LITERATURE CITED

ANDERSON, E. S., 1975 Viability of, and transfer of a plasmid from, *E. coli* K12 in the human intestine. Nature **255**: 502–504.

ATWOOD, K. C., L. K. SCHNEIDER and F. J. RYAN, 1951 Periodic selection in *Escherichia coli*. Proc. Natl. Acad. Sci. U.S. **37**: 146–155.

BACHMANN, B. J. and B. K. LOW, 1980 Linkage map of *Escherichia coli*, K-12. Edition 6. Microbiol. Rev. **44**: 1–56.

BENNETT, P. M., M. H. RICHMOND and V. PETROCHEILOU, 1980 The inactivation of *tet* genes on a plasmid by the duplication of one inverted repeat of a transposon-like structure which itself mediates tetracycline resistance. Plasmid **3**: 135–149.

BIRNBOIM, H. C. and J. DOLLY, 1979 A rapid alkaline extraction procedure for screening plasmid DNA. Nucleic Acid Res. **7**: 1513–1523.

COOKE, E. M., S. EWINS and R. A. SHOOTER, 1969 Changing faecal populations of *Escherichia coli* in hospital medical patients. Brit. Med. J. 1969(**4**): 593–595.

COOKE, E. M., R. A. SHOOTER, P. J. KUMAR, S. A. ROUSSEAU and A. L. FOULKES, 1970 Hospital food as a possible source of *Escherichia coli* in patients. Lancet **28**: 436–437.

CURTISS, R., III, 1969 Bacterial conjugation. Ann. Rev. Microbiol. **23**: 69–136.

FALKOW, S., H. SCHNEIDER, L. S. BARON and S. B. FORMAL, 1963 Virulence of *Escherichia-Shigella* genetic hybrids for the guinea pig. J. Bacteriol. **86**: 1251–1258.

FOWLER, A. V. and I. ZABIN, 1977 The amino acid sequence of β-galactosidase of *Escherichia coli*. Proc. Natl. Acad. Sci. U.S. **74**: 1507–1510.

GRAHAM, J. B. and C. A. ISTOCK, 1978 Genetic exchange in *Bacillus subtilis* in soil. Molec. Gen. Genet. **166**: 287–290. ——, 1979 Gene exchange and natural selection cause *Bacillus subtilis* to evolve in soil culture. Science **204**: 637–639.

GUERRY, P., D. J. LEBLANC and S. FALKOW, 1973 General method for the isolation of deoxyribonucleic acid. J. Bacteriol. **116**: 1064–1066.

GUISO, N. and A. ULLMANN, 1976 Expression and regulation of lactose genes carried by plasmids. J. Bacteriol. **127**: 691–697.

HAYES, W., 1968 *The Genetics of Bacteria and Their Viruses*. J. Wiley & Sons, New York.

JONES, R. T. and R. CURTISS, 1970 Genetic exchange between *Escherchia coli* strains in the mouse intestine. J. Bacteriol. **103**: 71–80.

LEVIN, B. R., 1981a Conditions for the existence of R-plasmids in bacterial populations. pp. 197–202. In: Proc. IVth International Symp. on Antibiotic Resistance. Edited by S. MITSUHASI, L. ROSIVAL and V. KRČMERÝ. Avicenum, Prague, and Springer-Verlag, Berlin. ——, 1981b Periodic selection, infectious gene exchange and the genetic structure of *E. coli* populations. Genetics **99**: (In press.)

LEVIN, B. R. and V. A. RICE, 1980 The kinetics of transfer of nonconjugative plasmids by mobilizing conjugative factors. Genet. Research. **35**: 241–259.

LEVIN, B. R. and F. M. STEWART, 1980 The population biology of bacterial plasmids: *a priori* conditions for the existence of mobilizable nonconjugative factors. Genetics **94**: 425–443.

LEVIN, B. R., F. M. STEWART and L. CHAO, 1977 Resource limited growth, competition and predation: A model and some experimental studies with bacteria and bacteriophage. Amer. Naturalist **111**: 3–24.

LEVIN, B. R., F. M. STEWART and V. A. RICE, 1979 The kinetics of conjugative plasmid transmission: Fit of simple mass action models. Plasmid **2**: 247–260.

LEVY, S. B., B. MARSHALL and D. ROWSE-EAGLE, 1980 Survival of *Escherichia coli* host-vector systems in the mammalian intestine. Science **209**: 391–394.

MACRINA, F. L., D. J. KOPECKO, K. R. JONES, D. J. AYERS and S. M McCOWEN, 1978 A multiple plasmid-containing *Escherichia coli* strain: Convenience of size reference plasmid molecules. Plasmid **1**:417–420.

MAYNARD SMITH, J., 1978 *The Evolution of Sex*. Cambridge University Press, Cambridge.

MEYERS, J. A., D. SANCHEZ, L. ELWELL and S. FALKOW, 1976 Simple agarose gel electrophoretic method for the identification and characterization of plasmid deoxyribonucleic acid. J. Bacteriol. **127**: 1529–1537.

MEYNELL, G. G., 1973 *Bacterial Plasmids*. M.I.T. Press, Cambridge, Massachusetts.

MILKMAN, R., 1975 Allozyme variation in *E. coli* of diverse natural origins. pp. 273–285. In: *Isozymes*, Vol. IV. Edited by C. L. MARKERT. Academic Press, New York.

SCHLESINGER, M., 1932 Adsorption of bacteriophage to homologous bacteria. II. Quantitative investigation of adsorbtion velocity and saturation, estimation of particle size of the bacteriophage. Z. Hyg. Imunitatsforsch **114**: 149–160. (Trans. in *Papers on Bacterial Viruses*. Edited by G. STENT. Little Brown, Boston.)

SCHNEIDER, H. and S. FALKOW, 1964 Characterization of an HFR strain of *Shigella flexneri*. J. Bacteriol. **88**: 682–689.

SEARS, H. J. and I. BROWNLEE, 1952 Further observations on the persistence of individual strains of *Escherichia coli* in the intestinal tract of man. J. Bacteriol. **63:** 47–57.

SEARS, H. J., I. BROWNLEE and J. UCHIYAMA, 1950 Persistence of individual strains of *Escherichia coli* in the intestinal tract of man. J. Bacteriol. **59:** 293–301.

SEARS, H. J., H. JONES, R. SALOUM, I. BROWNLEE and L. F. LAMOREAUX, 1956 Persistence of individual strains of *Escherichia coli* in man and dog under varying conditions. J. Bacteriol. **71:** 370–372.

SELANDER, R. K. and B. R. LEVIN, 1980 Genetic diversity and structure in populations of *Escherichia coli*. Science **210:** 545–547.

SELANDER, R. K., M. H. SMITH, S. Y. YANG, W. E. JOHNSON and J. B. GENTRY, 1971 Biochemical polymorphism and systematics in the genus *Peromyscus*. I. Variation in the old-field mouse (*Peromyscus polionotus*). *Studies in Genetics* VI. Univ. Texas Publ. **7103:** 49–90.

STEWART, F. M. and B. R. LEVIN, 1977 The population biology of bacterial plasmids: *A priori* conditions for the existence of conjugationally transmitted factors. Genetics **87:** 209–228.

Corresponding editor: D. L. HARTL

APPENDIX

β-Galactosidase Gene on a Plasmid

One strain, ET-7, showed five bands on starch gels stained for β-galactosidase. Since β-galactosidase is a tetramer (FOWLER and ZABIN 1977), it seemed likely that this "heterozygous" phenotype was produced by the combination of two dissimilar subunits of the enzyme, one being coded by a chromosomal gene of the bacterial-host strain and another by a gene carried on a plasmid. Results of tests were consistent with this hypothesis. By culturing the ET-7 strain with an *E. coli* K12 strain that was deleted for the *lac* operon, CSH50 nal [Δ(lac pro) ara nal rspE], we were able to transmit a lactose phenotype to the latter. There were no plasmid bands on DNA gels of the CSH50 nal strain, but the *lac*+ transconjugants of the strain show two high-molecular-weight plasmid bands that are indistinguishable from the two bands shown by the ET-7 strain. The *lac*+ transconjugants display a single band of β-galactosidase activity that is the same as one of the presumed homotetamer bands seen in the ET-7 strain. Further tests with the phage lambda and the *E. coli* K12 transconjugants of this lactose plasmid, which we shall designate as pDXX1, indicate that the plasmid also carries genes for the restriction and modification of DNA, and that this restriction modification system is different from that carried by *E. coli* K12. There are many reports of lactose plasmids in species that are normally unable to use this sugar (see, for example, GUISO and ULLMANN 1976). Why *lac*+ cells such as those of ET-7 carry plasmids with extra lactose-fermenting genes, and why plasmids themselves carry genes for the restriction and modification of DNA remain unanswered questions.