INFECTION AND IMMUNITY, Sept. 1972, p. 308-315
Vol. 6, No. 3

Polynucleotide Sequence Relatedness Among Three Groups of Pathogenic *Escherichia coli* Strains

DON J. BRENNER, G. R. FANNING, A. G. STEIGERWALT, I. ØRSKOV, AND F. ØRSKOV

Division of Biochemistry, Walter Reed Army Institute of Research, Washington, D. C. 20012, and International Escherichia Centre (WHO), Statens Seruminstitut, Copenhagen, Denmark

Received for publication 18 May 1972

Escherichia coli strains that cause dysentery-like disease, parenteral infection, and infantile diarrhea form specific groups based on mobility of O and K antigens in immunoelectrophoresis. Members from each of these groups were assayed for gross nucleotide sequence relatedness. The method used was interspecific deoxyribonucleic acid (DNA) reassociation reactions carried out free in solution. Reassociated DNA was separated from unreacted DNA by passage through hydroxyapatite. DNA relatedness between these groups was approximately 80%. The groups containing those strains causing parenteral infection and those responsible for dysentery-like disease showed preferentially high intragroup DNA relatedness. The group containing strains responsible for infantile diarrhea did not show preferentially high intragroup DNA relatedness with the reference strain employed. These strains, however, did exhibit preferentially high DNA relatedness to a second reference strain.

The species *Escherichia coli* encompasses a large number of strains that are extremely similar in morphology and metabolic patterns. Three parameters of deoxyribonucleic acid (DNA) relatedness in a large number of *E. coli* strains were assayed in a recent study (3). Data obtained in this study indicated up to a 4% difference in guanine plus cytosine content, and as much as 5×10^8 dalton differences in genome size in *E. coli* strains. Furthermore, as much as one-fourth of the DNA from some typical *E. coli* strains had diverged to a point where there were no demonstrable nucleotide sequences held in common. The frequency of relatedness between *E. coli* strains forms a Gaussian distribution. The above parameters can be used in the molecular description of a species (3).

With the increasing incidence or recognition of human pathogenic strains, or both, it has become increasingly important to determine the serotype of strains isolated from clinical sources. It is known that strains with certain specific O and K antigens are closely associated with pathological conditions, whereas other serotypes are normally nonpathogenic. In the pathogenic strains, certain serotypes are predominantly associated with infantile diarrhea, other serotypes are associated with parenteral infection, and still others with dysentery-like infections. Using immunoelectrophoretic (IE) patterns of simple extracts containing the O lipopolysaccharide antigen and a possible K acidic polysaccharide antigen, it was found that the serogroups that are highly frequent in the normal intestinal canal (which are the same serogroups frequent in parenteral disease) give one IE pattern—1A. Strains associated with human infantile diarrhea had a similar cathodic O antigen mobility but no special K antigen line, and thus gave an IE pattern which was a variant of 1A, termed 1B. *E. coli* strains from dysentery-like disease gave a third IE pattern, 2, with no K antigen line and a negatively charged O antigen. Finally, it should be stressed that the remaining large number of *E. coli* strains, about which nothing is known regarding possible relatedness to specific diseases or special ecological situations, can be grouped in these three IE groups (8).

It appears, therefore, that a strong correlation exists between the electrophoretic mobility of surface antigens and the type of pathological condition produced by *E. coli* strains. In the present study, we sought to determine whether the relationship between type of pathogenicity and electrophoretic mobility is reflected in preferentially high polynucleotide sequence relatedness.

MATERIALS AND METHODS

Organisms and media. The strains used in this study are listed in Table 1. All Statens Seruminstitut strains are antigenic test strains of the respective *E. coli* O

TABLE 1. *Bacterial strains employed*

Strain	Source[a]
E. coli 01A	CDC
E. coli 02	SSI
E. coli 02A	CDC
E. coli 04	SSI
E. coli 04	CDC
E. coli 06	SSI
E. coli 06	CDC
E. coli 07	SSI
E. coli 07	CDC
E. coli 08	F. J. Skerman
E. coli 09	CDC
E. coli 025	CDC
E. coli 026	SSI
E. coli 028	SSI
E. coli 032	SSI
E. coli 068	SSI
E. coli 075	SSI
E. coli 075	CDC
E. coli 0111B4	S. Falkow
E. coli 0112	SSI
E. coli 0114	SSI
E. coli 0115	SSI
E. coli 0119	SSI
E. coli 0124	SSI
E. coli 0128	SSI
E. coli 0136	SSI
E. coli 0142	SSI
E. coli 0145	SSI
E. coli K-12	University of Washington
Alkalescens-Dispar 01	CDC
Alkalescens-Dispar 02	CDC
Alkalescens-Dispar 03	WRAIR
Alkalescens-Dispar 03 (ceylonensis)	CDC
Alkalescens-Dispar 04	CDC
Alkalescens-Dispar 05	CDC
Alkalescens-Dispar 06	CDC
Alkalescens-Dispar 07	CDC
Alkalescens-Dispar 08	CDC
Shigella boydii 1	WRAIR
S. boydii (etousa) 7	WRAIR
S. dysenteriae I	WRAIR
S. dysenteriae II	WRAIR
S. dysenteriae III	WRAIR
S. flexneri 24570	S. Falkow
S. flexneri (newcastle) 6	WRAIR
S. sonnei	WRAIR
S. sonnei virulent	WRAIR
S. sonnei avirulent	WRAIR

[a] CDC, Center for Disease Control, Atlanta, Georgia. SSI, Statens Seruminstitut, Copenhagen, Denmark. WRAIR, Walter Reed Army Institute of Research, Washington, D. C.

antigens. Bacteria were maintained on meat extract agar or nutrient slants. Unlabeled DNA was obtained from organisms grown to stationary phase at 37 C in 4.5 liters or more of Brain Heart Infusion broth. $^{32}PO_4$-labeled DNA was obtained from cells grown to stationary phase in a glucose salts broth containing minimal amounts of inorganic phosphates and 10 to 15 mCi of $^{32}PO_4$ per liter (1).

DNA preparation. The purification of both labeled and unlabeled DNA, as well as the shearing of DNA to an average single-stranded fragment size of 1.25×10^5 daltons, has been described (1).

DNA reassociation. Relatedness studies based on relative percentage of DNA reassociation are dependent on both the specificity and the completeness of DNA duplex formation. To determine nucleotide sequence relatedness, DNA from one or more reference organisms is labeled with $^{32}PO_4$ and then reacted free in solution with a series of unlabeled DNA preparations. DNA reassociation is a concentration-dependent reaction. In these experiments, the concentrations of labeled and unlabeled DNA were 0.1 μg/ml and 150 μg/ml, respectively. These concentrations were chosen so that, under the conditions of incubation, little or no reassociation occurs between labeled fragments, whereas maximal reassociation occurs between labeled DNA fragments and complementary unlabeled fragments. Specificity of reassociation is controlled by using salt concentration and incubation temperatures at which sequences containing only randomly complementary bases cannot form stable reassociation products. Specifically, the labeled and unlabeled DNA species, in 0.28 M PB (phosphate buffer, an equimolar mixture of Na_2HPO_4 and NAH_2PO_4; *p*H 6.8), were denatured in a boiling water bath for 3 to 4 min. Samples were then cooled to 60 C or 75 C and incubated at this temperature for 16 hr. In control reactions at these criteria, 0.1 μg of labeled DNA per ml reassociates at less than 1%, whereas unlabeled DNA from the reference strains employed shows 80% or greater reassociation.

After incubation the samples are diluted in deionized water to 0.14 M PB, and single-stranded DNA is separated from duplexed DNA by chromatography on hydroxyapatite (HA, reference 2). HA is equilibrated with 0.14 M PB + 0.4% sodium dodecyl sulfate (SDS) and held at 60 or 75 C. Single-stranded DNA does not bind to HA at this salt concentration, whereas reassociated DNA is bound to HA. Four 15-ml washes with 0.14 M PB + 0.4% SDS at the incubation temperature serve to remove unreacted DNA from HA. Reassociated DNA is eluted either in four 15-ml washes with 0.4 M PB or, when the thermal stability of the duplexed DNA is of interest, in a series of 15-ml washes with 0.14 M PB at increasing 5 C temperature increments up to 100 C. Radioactivity is assayed by Cerenkov counting.

RESULTS

With few exceptions, antigens from the strains tested exhibited the following immunoelectrophoretic behavior: (i) group 1A, O antigen migration towards the anode and the cathode and a special anodic K antigen (strains highly frequent in normal human colon and causing parenteral infection); (ii) group 1B, O antigen migration towards the cathode (strains associated with infantile diarrhea); (iii) group 2, O antigens migrate towards

the anode (strains causing dysentery-like disease). Representative strains from each of these groups were tested for DNA relatedness within and across group lines. Also included in the study were *E. coli* strains whose O antigen groups were known, but which had not been subjected to immunoelectrophoretic analysis. However, previous experience with many strains belonging to these O groups indicate 100% correlation between O group and IE pattern. These strains were assigned to the same group as strains with identical O antigens that had been assayed immunoelectrophoretically. In addition, several *Shigella* strains were tested.

E. coli 075 was the arbitrarily chosen DNA reference strain from group 1A. Labeled DNA from this strain was reacted with DNA from ten other strains belonging to group 1A, as well as with DNA from seven group 1B strains and six group 2 strains as shown in Table 2. Strain 075 is closely related to members of all three immunoelectrophoretic groups as shown by high relative binding at 60 C and the negligible drop in relative binding obtained at 75 C as compared to that at 60 C. Despite the high level of binding of strain 075 to all strains tested, it is evident that preferentially high reaction occurred with members of group 1A (89%) as opposed to strains from group 1B (77%), group 2 (81%), or strain K-12 (81%).

Strains 075, 01A, 02A, 04, 06, and 07 (strains in parentheses in Table 2) were arbitrarily assigned to groups 1A and 1B without being subjected to electrophoretic analysis. With the possible exception of strain 02A (group 1A) their relatedness to 075 was similar to that shown by strains placed in the group on the basis of immunoelectrophoretic analysis of O antigens. The data from group 1A suggest that some strains with common O antigens are extremely similar (04, 06, 07), whereas other strains with similar O antigens (02 and 075) appear to be significantly different in terms of their total DNA.

Strain 0114 was chosen as the DNA reference strain for group 1B. This strain exhibits approximately equal average relatedness to the three groups of *E. coli* strains (Table 3). Strain 0114 was also tested for relatedness to a group of *Shigella* strains. As expected, DNA reassociation between the shigellae and 0114 was somewhat lower than the average obtained between *E. coli* strains (3). From this data it appears that the grouping of pathogens based on electrophoretic mobility of O antigens in group 1B does not reflect preferentially high overall DNA relatedness among members of this group.

Results obtained using strain 0112 as the representative from group 2 are shown in Table 4.

TABLE 2. *DNA relatedness of E. coli 075 to E. coli strains*

Source of unlabeled DNA[a]	% Relative binding	
	At 60 C	At 75 C
Group 1A *E. coli* strains		
075	100[b]	100[b]
(075)	90	92
(01A)	87	85
02	92	88
(02A)	80	76
04	92	89
(04)	92	91
06	93	94
(06)	92	93
07	85	82
(07)	84	80
Average	89	87
Group 1B, *E. coli* strains		
(09)	77	77
026	80	82
068	72	72
0114	79	79
0119	77	78
0142	80	82
0145	76	76
Average	77	78
Group 2 *E. coli* strains		
075	75	73
032	78	76
0112	78	73
0115	88	81
0124	77	74
0136	75	71
Average	79	75
E. coli K-12	81	78

[a] Strains in parentheses were not tested immunoelectrophoretically.

[b] Not included in average.

Relatedness values obtained from optimal, 60 C reassociation reactions are highest between 0112 and other strains from group 2. Data from similar reactions carried out at 75 C indicate that while the highest average reaction still occurs between 0112 and members of group 2, it is more difficult to distinguish between group 2 and group 1B reactions.

Strain 0115, also a member of group 2, exhibits preferential overall DNA sequence relatedness to strains from group 1B; in fact, in 60 C reactions, it showed least relatedness to strains from group 2 (Table 5). A reexamination of the data reveals that, in addition to 0115, there are

TABLE 3. *DNA relatedness of E. coli 0114 to E. coli and Shigella strains*

Source of unlabeled DNA[a]	% Relative binding	
	At 60 C	At 75 C
Group 1B *E. coli* strains		
0114	100[b]	100[b]
(08)	86	81
(09)	86	85
(025)	75	73
026	81	80
068	74	75
(0111B4)	79	77
0119	88	85
0142	78	76
0145	83	83
(128)	79	77
(Alkalescens-Dispar 08)	79	79
Average	81	79
Group 1A *E. coli* strains		
(02A)	86	85
02	79	74
04	79	69
06	78	72
07	85	85
075	78	76
(Alkalescens-Dispar 01)	79	77
(Alkalescens-Dispar 03)	78	75
(Alkalescens-Dispar 04)	76	78
(Alkalescens-Dispar 05)	79	78
Average	80	77
Group 2 *E. coli* strains		
028	74	74
032	79	76
0112	81	80
0115	88	86
0124	80	79
0136	78	77
Average	80	79
E. coli K-12	78	78
Shigella strains		
S. boydii	71	66
S. boydii (etousa)	76	75
S. dysenteriae I	72	70
S. dysenteriae II	75	73
S. dysenteriae III	72	70
S. flexneri	77	73
S. flexneri (newcastle)	74	71
S. sonnei	78	74
S. sonnei virulent	77	73
S. sonnei avirulent	77	72
Average	75	72

[a] Strains in parentheses were not tested immunoelectrophoretically.

[b] Not included in average.

other strains in which the pattern of DNA relatedness does not agree with strain groupings based on pathogenicity and O antigen immunoelectrophoretic mobility. These include 02A and 025. Despite these exceptions, and the seeming inability to distinguish group 1B strains from other strains on the basis of DNA reassociation, it appears that both group 1A and group 2 strains show preferentially high intragroup DNA relatedness.

Data obtained from DNA reassociation experiments were subjected to statistical analysis in order to assess the validity of grouping based on DNA relatedness. When the data were plotted on probability paper, 075 (group 1A) reactions formed a Gaussian distribution with group 1A strains, but not with strains from groups 1B or 2. Similarly 0112 (group 2) reactions comprise a Gaussian distribution with group 2 strains more so than with strains from groups 1A or 1B. Strain 0114 (group 1B) did not form a Gaussian distribution in reactions with strains from any of the groups. Strain 0115 (group 2) reactions exhibited a Gaussian distribution only with group 2 strains, despite the fact that the reactions with group 1B strains were substantially higher than reactions with group 2 strains. The data were also analyzed using the Kruskol-Wallis rank comparison test or X^2 significance test. There is a significant difference ($P = <0.05$) in the rank correlation of the different reference strains to group 1A strains. The ranking is 075 (group 1A) > 0115 (group 2) > 0112 (group 2) > 0114 (group 1B). For group 1B strains, the rank correlation is 0115 (group 2) > 0112 (group 2) > 0114 (group 1B) > 075 (group 1A), with $P = <0.01$. For strains in group 2, the rank correlation is highly significant ($P = <0.001$) with 0112 (group 2) > 0115 (group 2) > 0114 (group 1B) > 075 (group 1A).

Genome size variations from 2.3×10^9 daltons to 3.0×10^9 daltons exist among *E. coli* strains as shown by nonreciprocal binding in reciprocal reassociation reactions and by differences in initial rates of reassociation determined spectrophotometrically (3, 4). Differences in genome size can account for what may appear to be abnormally high or low binding. For example, in reciprocal reactions, 91% of *E. coli* K-12 DNA is equivalent to 74% of DNA from *E. coli* strain 02A due to the larger genome size of 02A as compared to that of K-12. It is, therefore, quite possible that all or part of the differences in DNA binding observed between the three groups of pathogenic *E. coli* strains reflect genome size differences, rather than comparably sized genomes that contain differing amounts of related nucleotide sequences. Significant genome size differences existing in the reference strains would be

TABLE 4. *DNA relatedness of E. coli strain 0112 to E. coli strains*

Source of unlabeled DNA[a]	% Relative binding	
	At 60 C	At 75 C
Group 2 *E. coli* strains		
0112	100[b]	100[b]
028	86	79
032	84	70
0115	84	79
0124	93	83
0136	90	79
Average	87	78
Group 1A *E. coli* strains		
(01A)	81	
02	83	76
(02A)	90	85
04	81	66
(04)	75	67
06	74	71
(06)	75	74
07	79	76
(07)	83	72
075	81	69
(075)	75	65
(Alkalescens-Dispar 01)	86	76
(Alkalescens-Dispar 02)	85	77
(Alkalescens-Dispar 03)	77	72
(Alkalescens-Dispar 03 celonensis)	82	78
(Alkalescens-Dispar 04)	84	83
(Alkalescens-Dispar 05)	80	81
(Alkalescens-Dispar 06)	79	78
(Alkalescens-Dispar 07)	76	71
Average	81	74
Group 1B *E. coli* strains		
(08)	82	79
(09)	88	74
(025)	76	61
026	83	80
(0111B4)	80	76
0114	84	81
0119	85	85
0142	85	73
0145	85	75
(128)	82	76
(Alkalescens-Dispar 08)	83	77
Average	83	76
E. coli K-12	83	73

[a] Strains in parentheses were not tested immunoelectrophoretically.
[b] Not included in average.

evident as nonreciprocal relatedness values among these strains. The reciprocal binding data presented in Table 6 indicate that genome sizes of the reference strains are within 5% of one another. Therefore, genome size differences cannot account for the differences in binding observed between 075 and group 1A strains (89%) and group 1B or group 2 strains (77 and 79%, respectively). Similarly, the differences in binding of 0112 and 0115 to members of the different groups cannot be accounted for solely on the basis of possible genome size differences.

When relative relatedness data are plotted as a frequency distribution, the mode for *E. coli* K-12 reaction with *E. coli* strains is 90 to 94%, and the range of relatedness values around this mode approximates a Gaussian distribution (3). Frequency distribution graphs of relatedness for each of the four reference strains employed in this study are shown in Fig. 1. Despite the relatively small number of strains, especially

TABLE 5. *DNA relatedness of E. coli strain 0115 to E. coli strains*

Source of unlabeled DNA[a]	% Relative binding	
	At 60 C	At 75 C
Group 2 *E. coli* strains		
0115	100[b]	100[b]
028	81	79
032	83	82
0112	81	81
0124	82	78
0136	80	80
Average	81	80
Group 1A *E. coli* strains		
(01A)	81	70
02	81	78
(02A)	85	79
04	84	79
(04)	79	72
06	83	75
(06)	81	74
07	92	86
(07)	85	75
075	86	83
(075)	81	74
Average	84	77
Group 1B *E. coli* strains		
(09)	94	85
026	93	85
068	81	79
0114	95	93
0119	95	92
0142	81	75
0145	92	89
Average	90	85

[a] Strains in parentheses were not tested immunoelectrophoretically.
[b] Not included in average.

TABLE 6. *Reciprocal binding as an indication of relative genome size*

Reaction[a]	% Relative binding at 60 C	Maximum difference in genome size based on reciprocal binding
075/0112	78	075 = 4% larger
0112/075	81	
075/0114	79	0114 = 1% larger
0114/075	78	
075/0115	88	0115 = 2% larger
0115/075	86	
0112/0114	84	0114 = 4% larger
0114/0112	81	
0112/0115	84	0115 = 4% larger
0115/0112	81	
0114/0115	88	0114 = 8% larger
0115/0114	95	

[a] Number preceding slash line indicates source of labeled DNA.

in group 2, the relative relatedness data for each reference strain with each group of strains approximate a Gaussian distribution. With the exception of 0114, the group to which the reference strain belongs can be separated from the other two groups on the basis of frequency distribution. A great many more strains should be tested with one or more reference strains from each group in order to determine whether each group shows a distinct relatedness pattern within the overall frequency distribution of the species *E. coli*.

DISCUSSION

In a recent study (3), a large number of *E. coli* strains have been investigated by interstrain DNA-DNA reassociation, genome size, and guanine plus cytosine determinations in order to assess the degree of difference or divergence in genetic material that is tolerable within a species. From this sampling, it is apparent that *E. coli* strains vary by at least 4% in guanine plus cytosine content and 23% in genome size (enough DNA to specify some 600 average-sized genes). As much as 25% divergence occurred between *E. coli* strains, based on their ability to form stable DNA duplexes between strains at 75 C. It is, therefore, evident that many *E. coli* strains display substantial differences in their DNA. Within the distribution of *E. coli* strains, one can identify several subgroups. For instance, the Alkalescens-Dispar strains form a tight subgroup with 88 to 95% relatedness to *E. coli* K-12, whereas *Shigella* strains form a subgroup on the low end of the *E. coli* frequency distribution (3).

In the present study we sought to determine whether pathogenic *E. coli* strains grouped on the basis of immunoelectrophoretic mobility show preferential DNA relatedness. The data indicate preferential relatedness in group 1A strains and in group 2 strains. There are exceptions including strain 02A, and strain 0115, a member of group 2 that shows preferential relatedness to group 1B strains. The reference strain from group 1B, 0114, reacted to approximately the same extent with strains from all three groups. Statistical treatments and frequency distribution graphs of these data substantiate the conclusions drawn from mean relative relatedness values.

A variety of functions have been identified as essential to the pathogenic capacity of different bacteria. These include enzymes allowing invasiveness, toxin production, coagulase production, resistance to phagocytosis, and susceptibility to specific lysogenic converting phage. There are, however, no data on exactly how much DNA in a given organism is concerned with the ability to produce disease. Since both pathogenic strains and nonpathogenic strains comprise the species *E. coli*, data from the present study are of interest in this regard.

Average relatedness among the three groups of strains is approximately 80%. Thus, preferential relatedness of strains within a given group must either be due to sequences held in common in the remaining 20% of DNA, or due to significant differences in genome size. Although it is true that genome size differences as large as 25% do exist between certain *E. coli* strains, reciprocal binding reactions do not indicate significant genome size differences among the reference strains employed in the present study. These data are therefore consistent with an interpretation in which about one-half of the DNA not common to all three groups is conserved among group 1A strains (where mean relative relatedness is 89%). Similarly, one-third of the DNA not common to all three groups is conserved among group 2 strains. It follows that the "additional" DNA conserved in each of these groups is either a different set of sequences, or a set of sequences originally present in an ancestor common to both groups that has diverged to a point where it is disimilar in group 1A and group 2.

Data obtained from group 1B strains are more difficult to interpret. The reference strain from this group, 0114, exhibits the same average relatedness to members of all three groups. This result is consistent with an 80% core of DNA relatedness between all three groups where the remaining 20% differs among group 1B strains to the same extent that it differs between group 1B strains and strains from the other two groups. It is also possible that 0114 is not representative of DNA relationships among group 1B strains. Strain 0115, the atypical group 2 strain, shows preferential relatedness to group 1B strains. If

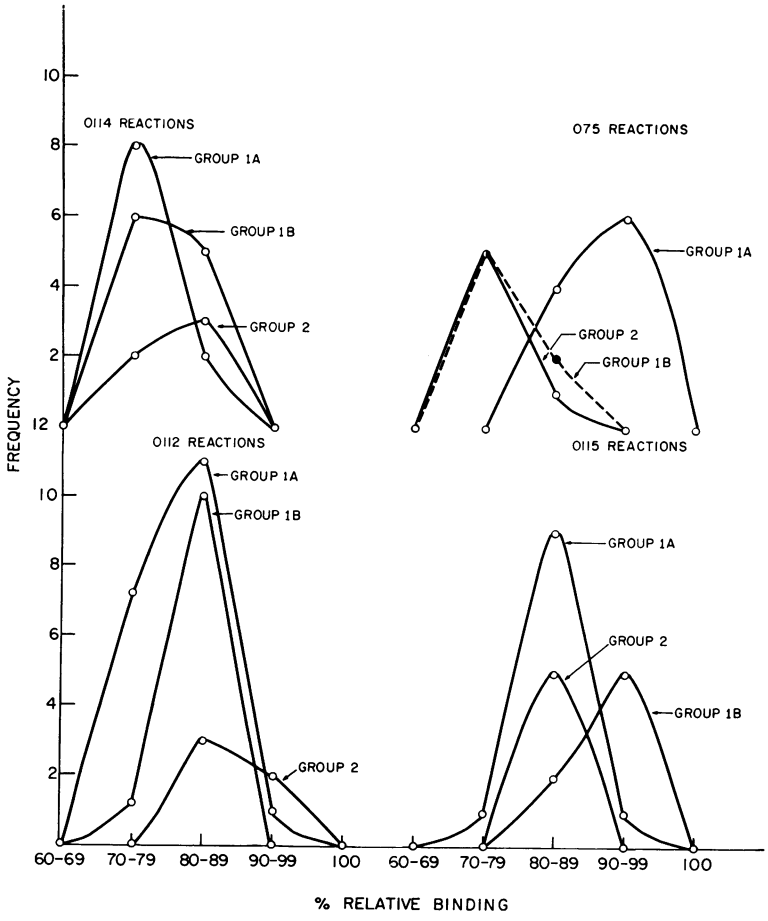

FIG. 1. *Frequency distribution of relatedness among E. coli strains. The values shown are obtained from 60 C DNA reassociation reactions in which E. coli 075, 0112, 0114, and 0115 were the reference strains.*

0115 reactions are considered typical for group 1B relatedness, we see preferentially high relatedness among these strains to at least the same extent as is present in groups 1A and 2. It will be worthwhile to isolate the 20% of DNA not common to the three groups and rereact it with each group.

There is no way to determine how much of the DNA that is preferentially conserved among pathogenic strains of *E. coli* is specifically associated with pathogenesis. The upper limit of DNA specifically concerned with the pathogenic process is approximately 10%, enough to specify some 250 to 300 genes of average size. Whereas this figure provides an upper limit, it is possible that the bulk of preferentially shared sequences are not concerned with pathogenicity. In fact, one may argue that pathogenicity is largely determined by single mutations.

Preferentially high nucleic acid relatedness among pathogens occurs in several groups of bacteria. *Brucella* species are virtually identical (5), although this result may be due to taxonomic splitting of virtually identical strains with different host specificity. Studies on the genus *Neisseria* that contains both pathogens and nonpathogens also indicate preferentially high relatedness among pathogens. Strains of *N. gonorrhoeae* and *N. meningitidis* share at least 80% of their DNA, whereas relatedness between seven nonpathogenic species varied from 15 to 75% (6, 7). Another case in point is relatedness in shigellae compared to relatedness among *Enterobacter* species. Strains of *S. flexneri, S. boydii, S. dysentereae,* and *S. sonnei* exhibit 80% or greater relatedness, whereas DNA species from the normally nonpathogenic species *E. aerogenes, E. cloacae, E. liquifaciens,* and *E. hafniae* have diverged in the range of 50 to 80% (Brenner, *unpublished observations*).

How does one account for the apparent conservation of DNA in pathogenic bacteria? It is usually argued that pathogens have evolved into an ecological niche that is specialized enough to prohibit gross divergence. In other words, the host-parasite relationship is in delicate balance so that significant divergence will be lethal or prevent a bacterium from competing successfully in its environment. Alternatively, nonpathogenic or rarely pathogenic bacteria have much more varied ecological niches and therefore divergence is successful at a greater frequency.

ACKNOWLEDGMENTS

The authors are grateful to Helen C. Sing who performed the statistical analyses. We wish to thank W. H. Ewing and G. J. Hermann for providing cultures of several of the *E. coli* strains used in this study.

LITERATURE CITED

1. Brenner, D. J., G. R. Fanning, K. E. Johnson, R. V. Citarella, and S. Falkow. 1969. Polynucleotide sequence relationships among members of the *Enterobacteriaceae*. J. Bacteriol. **98**:637–650.
2. Brenner, D. J., G. R. Fanning, A. Rake, and K. E. Johnson. 1969. A batch procedure for thermal elution of DNA from hydroxyapatite. Anal. Biochem. **28**:447–459.
3. 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.
4. Gillis, M., J. DeLey, and M. DeCleene. 1970. The determination of molecular weight of bacterial genome DNA from renaturation rates. Eur. J. Biochem. **12**:143–153.
5. Hoyer, B. H., and N. B. McCullough. 1968. Polynucleotide homologies of *Brucella* deoxyribonucleic acids. J. Bacteriol. **95**:444–448.
6. Kingsbury, D. T. 1967. Deoxyribonucleic acid homologies among species of the genus *Neisseria*. J. Bacteriol. **94**:870–874.
7. Kingsbury, D. T., G. R. Fanning, K. E. Johnson, and D. J. Brenner. 1969. Thermal stability of interspecies *Neisseria* DNA duplexes. J. Gen. Microbiol. **55**:201–208.
8. Ørskov, F., I. Ørskov, B. Jann, and K. Jann. 1971. Immunoelectrophoretic patterns of extracts from all *Escherichia coli* O and K antigen test strains: correlation with pathogenicity. Acta Pathol. Microbiol. Scand. Sect. B **79**:142–152.