Proc. Natl. Acad. Sci. USA
Vol. 77, No. 6, pp. 3529–3533, June 1980
Evolution

Evolution of a new enzymatic function by recombination within a gene

[evolved β-galactosidase (ebg)/β-galactosidase/experimental evolution/*Escherichia coli*]

BARRY G. HALL AND TIMOTHY ZUZEL

Microbiology Section, U-44, Biological Sciences, University of Connecticut, Storrs, Connecticut 06268

Communicated by Herschel L. Roman, March 10, 1980

ABSTRACT **Mutations that alter the *ebgA* gene so that the evolved β-galactosidase (ebg) enzyme of *Escherichia coli* can hydrolyze lactose fall into two classes: class I mutants use only lactose, whereas class II mutants use lactulose as well as lactose. Neither class uses galactosylarabinose effectively. In this paper we show that when both a class I and a class II mutation are present in the same *ebgA* gene, ebg enzyme acquires a specificity for galactosylarabinose. Although galactosylarabinose utilization can evolve as the consequence of sequential spontaneous mutations, it can also evolve via intragenic recombination in crosses between class I and class II *ebgA*$^+$ mutant strains. We show that the sites for class I and class II mutations lie about 1 kilobase, or about a third of the gene, apart in *ebgA*. Implications of these findings with respect to the evolution of new metabolic functions are discussed.**

The *ebg* (evolved β-galactosidase) system of *Escherichia coli* is being employed as a model system to study the evolution of new metabolic functions. This model system has demonstrated, in the laboratory, the requirement for both structural and regulatory gene mutations for the evolution of a particular new metabolic function (1–3). It has also been used to demonstrate the existence of an evolutionary pathway, in that three sequentially selected spontaneous mutations in a structural gene were required for the evolution of a particular new metabolic function (4). We now turn our attention to the role of intragenic recombination in the evolution of new metabolic functions.

Strains of *E. coli* that bear deletions within the *lacZ* (β-galactosidase) gene, but in which the *lacY* (lactose permease) gene is intact, are unable to utilize lactose or other β-galactoside sugars as sole carbon and energy sources. A second β-galactosidase, enzyme ebg^o, is the product of the wild-type allele of the *ebgA* gene located at 66 min on the *E. coli* K-12 map (5, 6). Expression of the *ebgA* gene is under control of the tightly linked *ebgR* gene, which specifies a repressor (7). The wild-type enzyme, ebg^o, has little activity toward natural β-galactoside compounds, and even *ebgR*$^-$ strains, which synthesize about 5% of their soluble protein as ebg^o enzyme, are unable to utilize lactose, lactulose, or galactosylarabinose (Gal-Ara) (2, 4). Spontaneous single-point mutations can evolve the wild-type allele, *ebgA*^{o}, to *ebgA*$^+$, resulting in enzyme with greatly increased activity toward lactose (8).

Previous studies (2, 4) have shown that a number of different mutations in the *ebgA* gene can lead to enzyme with increased β-galactosidase activity. Selection for lactose utilization results in two classes of *ebgA*$^+$ mutants: class I strains grow rapidly on lactose but are unable to utilize lactulose; class II strains grow more slowly on lactose and grow at a moderate rate on lactulose (4). Table 2 shows the first-order growth rate constants for representative members of each class. Studies of growth rates, coupled with kinetic analyses of purified enzymes, have shown that a number of nonidentical mutations can occur within each class (2, 4). Selection for lactulose, rather than lactose, utilization results in strain previously designated class III (4). Because these *ebgA*$^+$ strains are phenotypically indistinguishable from class II, they will be referred to as class II strains in this and subsequent articles.

Class I and II strains carry single mutations within the *ebgA* gene (8), and all grow extremely slowly, and sometimes not at all, on Gal-Ara. When growth occurs, doubling times of 23–36 hr are typical (4). When class I strains (lactose-positive, lactulose-negative) are subjected to selection for lactulose utilization, class IV strains arise. Class IV strains carry *two* mutations in the *ebgA* gene. They differ from the previous classes in that (*i*) they grow faster on lactose than on lactulose, and (*ii*) they grow at a significant rate (doubling times less than 7 hr) on Gal-Ara. Class IV strains are of particular interest because they seem to be obligatory intermediates on the pathway to evolving *ebg* enzyme so that it can hydrolyze lactobionic acid (4).

We previously described an "obligatory" pathway for the evolution of lactobionate utilization consisting of a class I mutation, followed by a second mutation to give a class IV strain, followed by a third mutation to give lactobionate utilization (4). Class II strains were considered evolutionary dead ends simply because we did not know what selective pressure to apply in order to evolve further substrate specificities (4). We recently decided to apply selective pressure for Gal-Ara utilization to class II strains in order to determine whether Gal-Ara$^+$ strains would resemble class IV strains.

MATERIALS AND METHODS

Bacterial and Phage Strains. All bacterial strains are *E. coli* K-12 and bear the *lacZ* deletion W4680. Unless otherwise indicated in Table 1, all strains are *ebgR*$^-$, and thus synthesize ebg enzyme constitutively. The *ebgR*$^-$ allele in strain 5A11 was selected with phenyl-β-galactoside as in ref. 7. SJ-20 is a recombinant from a mating between 1B1 and SJ-7. SJ480 is a recombinant from a mating between D2 and SJ-7. Spontaneous *thyA*$^-$ mutants were selected by growth for 20 generations in minimal medium containing thymidine and trimethoprim. Genotypes are given in Table 1. Bacteriophage P1 cam ts100 was used for transductions as described (7).

Media. Minimal medium was described previously (1). As required, methionine and arginine were employed at 100 mg/liter and streptomycin sulfate was employed at 300 mg/liter. Lactose (4-*O*-β-D-galactopyranosyl-D-glucose), lactulose (4-*O*-β-D-galactopyranosyl-D-fructose), and Gal-Ara (3-*O*-β-D-galactopyranosyl-D-arabinose) were used at 1 g/liter. All media employing a β-galactoside also contained 0.2 mM isopropyl β-D-thiogalactopyranoside for the sole purpose of inducing synthesis of the lactose permease (1). MacConkey in-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviation: ebg, evolved β-galactosidase.

Table 1. Relevant genotypes of bacterial strains

Strain	Relevant genotype	Ref.
1B1	*ebgA°*, HfrC, *spc* (= *rpsE*)	7
A2	*ebgA2*, HfrC, *spc*	1
D2	*ebgA168*, HfrC, *spc*	4
5A1	*ebgA51*, HfrC, *spc ebgR*$^+$	7
5A11	*ebgA51*, HfrC, *spc ebgR*$^-$ mutant of 5A1	This study
5A2	*ebgA52*, HfrC, *spc*	7
C1	*ebgA139*, HfrC, *spc*	4
C2	*ebgA141*, HfrC, *spc*	4
A23	*ebgA134*, HfrC, *spc*	4
5A11GA	*ebgA204*, HfrC, *spc*	This study
5A2GA	*ebgA203*, HfrC, *spc*	This study
C1GA	*ebgA201*, HfrC, *spc*	This study
C2GA	*ebgA202*, HfrC, *spc*	This study
SJ-7	*ebgA°*, F$^-$, *strA*, *metC*, *argG*, *tolC*, *ebgR*$^+$	7
SJ-8	*ebgA2*, F$^-$, *strA*, *metC*, *argG*	7
SJ-8T	*ebgA2*, F$^-$, *strA*, *metC*, *argG*, *thyA*	This study
SJ-12	*ebgA52*, F$^-$, *strA*, *metC*, *argG*	7
SJ-12T	*ebgA52*, F$^-$, *strA*, *metC*, *argG*, *thyA*	This study
SJ-20	*ebgA°*, F$^-$, *strA*, *tolC*	This study
SJ-480	*ebgA168*, F$^-$, *strA*, *metC*, *argG*	This study
SJ-480T	*ebgA168*, F$^-$, *strA*, *metC*, *argG*, *thyA*	This study
R41	*ebgA2*, F$^-$, *strA*	This study
R42	*ebgA143*, F$^-$, *strA*	This study
R61	*ebgA144*, F$^-$, *strA*	This study
R62	*ebgA52*, F$^-$, *strA*	This study
RT512	*ebgA108*, F$^-$, *strA*	This study
RT522	*ebgA107*, F$^-$, *strA*	This study
RT51168	*ebgA105*, F$^-$, *strA*	This study
RT52168	*ebgA106*, F$^-$, *strA*	This study
SJ12R/F'A2	F'122, *ebgA2*, *ebgR*$^-$/*ebgA52*, *metC*, *argG*, *recA*	This study
SJ8R/F'5A2	F'122, *ebgA52*, *ebgR*$^-$/*ebgA2*, *metC*, *argG*, *recA*	This study

dicator medium was prepared from MacConkey agar base (Difco), and contained the indicated fermentable β-galactoside at 10 g/liter. MacConkey agar kills *tolC*$^-$ cells (7), as does sodium deoxycholate, which was added to minimal medium at a concentration of 1 g/liter in some selective plates.

Growth Rates. Rates were measured as described in ref. 4.

RESULTS

Selection of Gal-Ara$^+$ mutants from class II ebgA$^+$ strains

Initial attempts to select spontaneous Gal-Ara$^+$ mutants on Gal-Ara minimal medium were hampered by significant background growth when more than 10^8 cells were spread per plate. That background growth was attributable to the marginal growth of class II strains on Gal-Ara. As an alternative, Gal-Ara$^+$ mutants were selected by a serial transfer method. The class II strains 5A2, 5A11, C1, and C2 were grown overnight in glycerol minimal medium containing 0.2 mM isopropyl thiogalactoside to fully induce the lactose permease, and inoculated into Gal-Ara minimal medium at an initial density of 10^8 cells per ml. The cultures were shaken at 37°C, and the optical density was monitored daily until the density had risen to 5×10^8 cells per ml. This required 6 days for strains C2 and 5A11, and 7 days for strains C1 and 5A2. The cultures were diluted 1:50 into the same medium and shaken for 3 days, by which time the densities of all cultures exceeded 5×10^8 cells per ml. The cultures were again diluted 1:50 and grown for 2 days to a density in excess of 5×10^8 cells per ml. Each culture was streaked onto a MacConkey Gal-Ara plate. The plates exhibited a preponderence of red (Gal-Ara$^+$) colonies, although a number of white (Gal-Ara$^-$) colonies were present on each plate. A single Gal-Ara$^+$ colony was isolated from each culture. The new strains were named 5A2GA, 5A11GA, C1GA, and C2GA. These strains, like their parents, were sensitive to the male-specific bacteriophage R17 and were resistant to the antibiotic spectinomycin.

To be sure that the Gal-Ara$^+$ phenotype was the result of mutation within the *ebgA* gene, the new mutations were mapped by transduction with bacteriophage P1. Strain SJ-20 (*ebgA°*, *ebgR*$^-$, *tolC*$^-$) was transduced with phage grown on each of the new strains, and *tolC*$^+$ transductants were selected on MacConkey lactose medium *ebgA*$^+$ cotransductants were obtained at the following frequencies: 5A2GA, 37%; 5A11GA, 43%; C1GA, 38%; and C2GA, 40%. These values are in good agreement with the previously reported value of 43.3% cotransduction between *ebgA* and *tolC* (7). All of the *ebgA*$^+$ cotransductants were replicated to MacConkey Gal-Ara plates were found to be Gal-Ara$^+$. Thus, the mutation that allows class II strains to utilize Gal-Ara effectively is in the *ebgA* gene.

Characteristics of Gal-Ara$^+$ strains

Table 2 shows the growth rates of the Gal-Ara$^+$ strains on lactose, lactulose, and Gal-Ara minimal medium. Comparison of these strains with their parental class II strains shows that in each case there was a great increase in the rate of Gal-Ara utilization such that all doubled in less than 7 hr. Likewise, in each case the growth rate on lactose increased by a factor of at least

Table 2. Growth rates on three β-galactoside sugars

Strain	Class	Growth rate constant, hr^{-1}		
		Lactose	Lactulose	Gal-Ara
1B1	WT	0	0	0
A2*	I	0.446 ± 0.033	0	0.029 ± 0.001
SJ8	I	0.365 ± 0.009	0	ND
D2*	I	0.451 ± 0.042	0	0.032 ± 0.005
5A2*	II	0.118 ± 0.016	0.255 ± 0.015	0.027 ± 0.001
SJ-12	II	0.128 ± 0.006	0.252 ± 0.003	ND
5A11	II	0.171 ± 0.004	0.233 ± 0.012	0
C1	II	0.181 ± 0.029	0.235 ± 0.010	0
C2*	II	0.227 ± 0.008	0.312 ± 0.006	0.034 ± 0.001
A23*	IV	0.360 ± 0.026	0.196 ± 0.005	0.139 ± 0.009
5A2GA	IV'	0.355 ± 0.025	0.145 ± 0.048	0.102 ± 0.022
5A11GA	IV'	0.352 ± 0.012	0.159 ± 0.018	0.136 ± 0.040
C1GA	IV'	0.376 ± 0.035	0.181 ± 0.025	0.145 ± 0.022
C2GA	IV'	0.360 ± 0.048	0.147 ± 0.024	0.102 ± 0.040
Recombinant strains				
R41	I	0.377 ± 0.009	0	0
R42	II	0.140 ± 0.003	0.167 ± 0.004	0.031 ± 0.006
R61	I	0.389 ± 0.005	0	0
R62	II	0.162 ± 0.004	0.212 ± 0.005	0
RT512	IV	0.201 ± 0.004	0.092 ± 0.002	0.165 ± 0.012
RT522	IV	0.220 ± 0.006	0.090 ± 0.004	0.112 ± 0.002
RT51168	IV	0.230 ± 0.002	0.070 ± 0.008	0.101 ± 0.006
RT52168	IV	0.222 ± 0.003	0.069 ± 0.002	0.100 ± 0.011
Diploid strains				
SJ12R/F'A2	II/F'I	0.312 ± 0.002	0.140 ± 0.004	0.041 ± 0.004
SJ8R/F'5A2	I/F'II	0.300 ± 0.001	0.182 ± 0.003	0.045 ± 0.011

Values shown are the first-order growth rate constants ± the 95% confidence interval. WT, wild type, ND, not determined.
* Data taken from ref. 4.

1.6, while the growth rate on lactulose decreased by at least a factor of 1.3. Class II strains all grow faster on lactulose than on lactose (4). Apparently mutations that permit significant growth rates on Gal-Ara reverse this relationship and result in more rapid growth on lactose than on lactulose. An examination of Table 2 shows that there is a striking similarity between the growth rates of the new Gal-Ara$^+$ strains and class IV strains on lactose, lactulose, and Gal-Ara. The Gal-Ara$^+$ strains, like class IV strains, carry two point mutations within the *ebgA* gene. The strikingly similar growth rates suggested that the *second* mutation in these strains might be equivalent to a class I mutation and that likewise the second mutation in a class IV strain might be equivalent to a class II mutation. On the basis of the similarity in growth rates the new strains are considered to be class IV; however, for the purposes of discussion they will be designated class IV′ to distinguish them from previously obtained class IV strains.

Recovery of single-point mutations from recombination within *ebgA*

If classes IV and IV′ are simply the result of a class I plus a class II mutation, it should be possible to recover *both* class I and class II type strains from crosses between a class IV or IV′ gene and a wild-type (*ebgA^o*) gene.

The class IV strain A23 and the class IV′ strain 5A2GA were each mated with the F$^-$, *tolC*$^-$, *ebgA^o*, *ebgR*$^-$, *strA*$^-$ strain SJ-20, and the mating mixture was plated onto MacConkey lactose agar containing streptomycin. The *tolC*$^+$ *strA*$^-$ recombinants that arose (about 1000 per plate) were replicated to MacConkey lactulose and MacConkey Gal-Ara plates. The replicated plates were screened for colonies that were Gal-Ara$^-$ but both lactose$^+$ and lactulose$^+$ (class II) and for colonies that were Gal-Ara$^-$ and lactulose$^-$ but lactose$^+$ (class I). Appropriate colonies were reisolated and the phenotypes were confirmed. Both of the expected recombinant types were recovered from each cross. From the A23 × SJ-20 cross the class I recombinant was designated R41, and the class II recombinant was designated R42. Likewise, from the 5A2GA × SJ-20 cross the class I recombinant was designated R61, and the class II recombinant was designated R62. Growth rates of these recombinant strains are shown in Table 2. The allele present in strain R41 should be identical with *ebgA2*, present (in the same genetic background) in strain SJ-8. The allele present in strain R62 should be identical to *ebgA52*, which is present in strain SJ-12. The alleles present in strains R42 and R61 did not previously exist, and were created by recombination.

Recovery of both class I and class II *ebgA*$^+$ alleles from these crosses is strong evidence that both class IV and class IV′ alleles are simply the sum of class I and a class II mutation in the *ebgA* gene.

Mapping the order of the class I and class II sites

To map the order of the class I and class II sites within the *ebgA* gene, strains A23 and 5A2GA were mated with strain S-J7. The mating mixture was plated onto lactose minimal medium containing sodium deoxycholate and streptomycin. The *tolC*$^+$ *strA*$^-$ recombinants that arose were replicated to MacConkey Gal-Ara plates to score for Gal-Ara$^-$ colonies, which would arise as the result of recombination between the class I and class II sites within the *ebgA* gene. In these crosses *all* of the Gal-Ara$^-$ recombinants are expected to be of that class whose site is closest to the *ebgR* gene. This is because *ebgR*$^+$ strains are lactose$^-$ regardless of the *ebgA* allele present (3) (see Fig. 1). In the A23 × SJ-7 cross an aliquot of the diluted mating mixture was also plated onto MacConkey lactose streptomycin plates to measure the recombination frequency between *ebgA* and *tolC* directly.

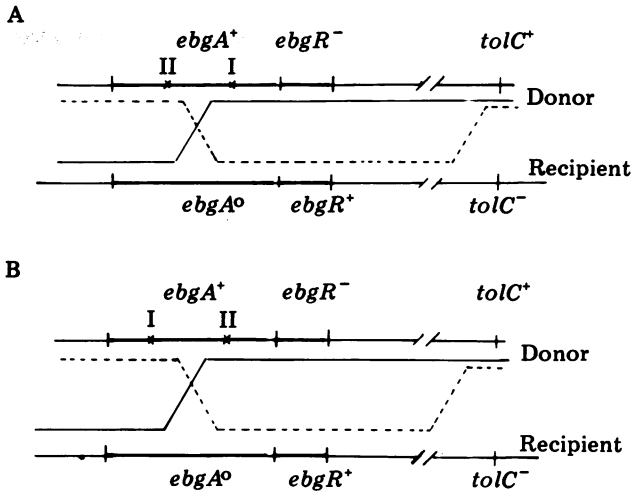

FIG. 1. Diagram of crosses between class IV *ebgA*$^+$ and *ebgA^o* alleles. The order *ebgA-ebgR-tolC* is as given in ref. 7 and confirmed by ref. 9. The distance between *ebgA* and *ebgR* is about 1.6% recombination (7). Colonies on the selective medium have received *tolC*$^+$ from the donor strain. A single crossover between *ebgR* and *tolC* results in an *ebgA^o ebgR*$^+$ strain, which is lactose-negative and thus does not form colonies on the selective medium (7). A single crossover between *ebgR* and *ebgA* results in an *ebgR*$^-$ *ebgA^o* strain, also lactose-negative (3). A single crossover to the left of *ebgA* yields a lactose-positive class IV or IV′ donor parental type, which will form colonies. A crossover between *tolC* and *ebgR* accompanied by a crossover within *ebgA* between the class I and class II sites (broken line) results in an *ebgR*$^+$ strain that cannot utilize lactose (3). A single crossover between the class I and class II sites within the *ebgA* gene (solid line) yields *ebgR*$^-$ nonparental lactose-positive recombinants. Alternative *A* shows that if the order is class II site-class I site-*ebgR*, all the nonparental colonies will be class I. Alternative *B* shows that if the order is class I site-class II site-*ebgR*, all nonparental colonies will be class II. Because of the tight linkage between *ebgA* and *ebgR* (7, 9) it was considered unlikely that double recombinants with one crossover between *ebgA* and *ebgR*, and the other crossover between the class I and the class II sites, would be detected. In fact no such double recombinants were observed.

From the cross between strain A23 and strain SJ-7, 1190 *tolC*$^+$, *ebgA*$^+$, *strA*$^-$ recombinants were recovered. Of these 18 were Gal-Ara$^-$, and all 18 were class I strains. This gives a recombination frequency of 1.5% between the class I and class II sites within the *ebgA* gene, and the order is class II site-class I site-*ebgR* gene. In that cross the recombination frequency between *ebgA* and *tolC* was 34.4%. Hartl and Dykhuizen (9) have carefully mapped the region around *ebgA*, and they estimate the distance between *ebgA* and *tolC* to be 0.6 min, or about 24.6 kilobases of DNA (6). If 34.4% recombination is equivalent to 24.6 kilobases, then 1.5% recombination is equivalent to about 1 kilobase.

That distance and order were confirmed by the cross between strain 5A2GA and SJ-7. Of 235 *tolC*$^+$, *ebgA*$^+$, *strA*$^-$ recombinants, 3, or 1.3%, were Gal-Ara$^-$. All three were class I strains.

Generation of Gal-Ara-utilizing mutants by recombination

If the class I and class II sites are about 1 kilobase apart, it should be relatively easy to obtain Gal-Ara$^+$ mutants from crosses between class I and class II *ebgA*$^+$ strains. Such recombinants would be expected to exhibit growth rates typical of class IV strains.

Two particular class I and two class II alleles were chosen because the enzymes specified by these alleles had been previously purified, characterized, and shown to have different

Table 3. Generation of Gal-Ara$^+$ strains by intragenic recombination

	Recipient					
	Strain SJ480T *ebgA168* Class I		Strain SJ8T *ebgA2* Class I		Strain SJ12T *ebgA52* Class II	
Donor						
Strain 5A11 *ebgA51* Class II	*thyA*$^+$	13,700	*thyA*$^+$	10,600	*thyA*$^+$	4800
	Gal-Ara$^+$	132	Gal-Ara$^+$	106	Gal-Ara$^+$	0
	Recombination freq. 0.96%		Recombination freq. 1.0%			
	Recombinant strain name: RT51168		Recombinant strain name: RT512			
Strain 5A2 *ebgA52* Class II	*thyA*$^+$	11,300	*thyA*$^+$	11,800		
	Gal-Ara$^+$	110	Gal-Ara$^+$	118		
	Recombination freq. 0.97%		Recombination freq. 1.0%			
	Recombinant strain name: RT52168		Recombinant strain name: RT522			
Strain A2 *ebgA2* Class I	*thyA*$^+$	12,000				
	Gal-Ara$^+$	0				

catalytic constants (K_m and V_{max}) with respect to hydrolysis of lactose and *o*-nitrophenyl β-D-galactoside (2, 4). Three of the *ebgA*$^+$ alleles were introduced into F$^-$ strains that were subsequently made *thyA*$^-$. In each cross shown in Table 3 an Hfr strain carrying one *ebgA*$^+$ allele was crossed with an F$^-$ *thyA*$^-$ strain carrying a different *ebgA*$^+$ allele (*thyA* is distal to *ebgA* in these crosses). The mating mixture was plated onto Gal-Ara minimal agar containing arginine, methionine, and streptomycin, but lacking thymidine, to select Gal-Ara$^+$ *thyA*$^+$ recombinants. Each colony must arise from a zygote that has received the *thyA*$^+$ gene from the donor and has also undergone recombination within the *ebgA* gene. To determine the number of *thyA*$^+$ zygotes present, a 1:10 dilution of the mating mixture was plated on glucose minimal medium containing arginine, methionine, and streptomycin but lacking thymidine.

Table 3 shows that all crosses of a class I with a class II allele yielded Gal-Ara$^+$ recombinants at a frequency of about 1%. On the other hand, crosses between two different class I alleles, or two different class II alleles, failed to yield any Gal-Ara+ recombinants. This may mean that double class I or double class II mutants are Gal-Ara$^-$, or it may mean that recombination is too rare to have been observed. Twenty-five Gal-Ara$^+$ recombinants were picked from each cross and tested on MacConkey Gal-Ara plates. All were confirmed to be Gal-Ara$^+$ by this test. One recombinant from each cross was saved, and, in order to facilitate comparisons, the *ebgA*$^+$ gene was transduced into strain SJ-20. The growth rates of these isogenic strains (designated RT) are shown in Table 2. The results clearly demonstrate that the class IV phenotype can arise from recombination between a class I *ebgA*$^+$ gene and a class II *ebgA*$^+$ gene.

Complementation between class I and class II

F′ episomes were constructed that carried the class I allele *ebgA2* and the class II allele *ebgA52*. Details of the construction of these episomes will be presented elsewhere. The F′ *ebgA2* was introduced into an F$^-$ *recA*$^-$ strain carrying *ebgA52*, and the F′ *ebgA52* was introduced into an F$^-$ *recA*$^-$ strain carrying *ebgA2*. In neither case did the diploid strain grow at a rate expected for a class IV strain (Table 2). These results suggest that although the class IV phenotype can arise by recombination it is not generated by complementation.

DISCUSSION

The results presented here show that significant growth on Gal-Ara arises as a consequence of two sequential mutations within the *ebgA* gene. One of the mutations must be a class I mutation, which by itself permits utilization of lactose but not lactulose. The other mutation must be a class II mutation, which by itself permits utilization of both lactose and lactulose. The two mutations may be selected in either order, depending upon the selection pressure applied.

Two lines of evidence show that class IV alleles are simply the sum of a class I and a class II mutation: (*i*) Both class I and class II alleles can be recovered as a result of recombination between a class IV allele and a wild-type (*ebgA*^{o}) allele. This is independent of the method of selection of the class IV allele. (*ii*) Class IV alleles can be generated by recombination between a class I allele and a class II allele.

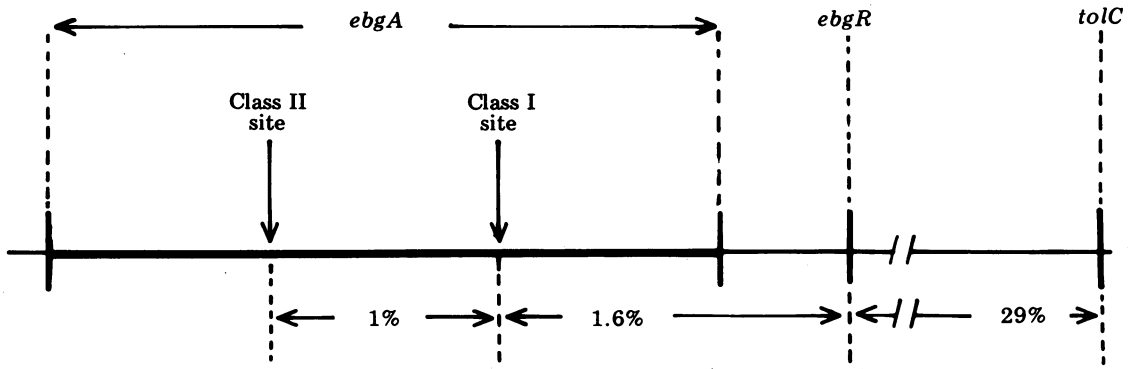

FIG. 2. Map of the *ebg* region based upon data presented here and in ref. 7. Distances are given in percent recombination. The class I site is placed 1.6% recombination from *ebgR* because the original mapping of *ebgR* with respect to *ebgA* employed the class I allele *ebgA2* (7).

The class I and class II sites within the *ebgA* gene were ordered with respect to the *ebgR* gene. The distance between the sites is about 650–1000 nucleotides, based upon recombination frequencies between 1% and 1.5%. As estimated on the basis of a subunit molecular weight of 120,000 for ebg enzyme (2), the sites are separated by 25–35% of the length of the gene. Fig. 2 shows a map of the *ebg* region.

These results show that a new enzymatic function, Gal-Ara hydrolysis, can evolve via recombination within the *ebgA* gene. Class I and class II *ebgA*$^+$ strains grow extremely slowly, and often not at all, on Gal-Ara as a sole carbon and energy source (Table 2). Because Gal-Ara-utilizing strains can evolve via recombination in the laboratory, we suggest that the same thing may happen in nature. We can envision a situation in which different alleles of the same gene have, under different selective pressures, diverged to give enzymes with somewhat different substrate specificities. Remixing of the populations would afford opportunities for intragenic recombination to generate a new allele with a substrate specificity present in neither parent. The new substrate specificity could permit the recombinant to exploit resources, or a new ecological niche, unavailable to either parent. Mechanisms of this sort at the intergenic recombination level have been discussed as mechanisms of speciation (10), but here we have directly demonstrated the evolution of a new metabolic function by intragenic recombination.

The existence of these alleles derived by recombination within the *ebgA* gene provides a unique opportunity for studying and comparing *functional* enzymes that carry different substitutions singly and together in the same molecule. We can, for instance, compare the wild-type enzyme ebg^o with the class I enzyme ebg^{+a}, with the class II enzyme ebg^{+b}, and with the recombinant class IV enzyme ebg^{+ab}. These studies should be of particular interest in that all of these mutations act to increase, rather than to impair, the activity of the enzyme.

This work was supported by Grant AI 14766 from the National Institute of Allergy and Infectious Diseases.

1. Hall, B. G. & Hartl, D. L. (1974) *Genetics* **76,** 391–400.
2. Hall, B. G. (1976) *J. Mol. Biol.* **107,** 71–84.
3. Hall, B. G. & Clarke, N. D. (1977) *Genetics* **85,** 193–201.
4. Hall, B. G. (1978) *Genetics* **89,** 453–465.
5. Hartl, D. L. & Hall, B. G. (1974) *Nature (London)* **248,** 152–153.
6. Bachman, B. J., Low, K. B. & Taylor, A. L. (1976) *Bacteriol. Rev.* **40,** 116–167.
7. Hall, B. G. & Hartl, D. L. (1975) *Genetics* **81,** 427–435.
8. Hall, B. G. (1977) *J. Bacteriol.* **129,** 540–543.
9. Hartl, D. L. & Dykhuizen, D. (1979) *Genetics* **91,** S45 (abstr.).
10. Lewontin, R. C. (1974) *The Genetic Basis of Evolutionary Change* (Columbia University Press, New York).