Mol Gen Genet (1983) 189:337–339

MGG

Short Communication

Genotypic Exclusion: A Novel Relationship Between the Ribitol-Arabitol and Galactitol Genes of *E. coli*

Christopher D. Link* and Albey M. Reiner
Department of Microbiology, University of Massachusetts, Amherst, MA 01003, USA

Summary. Genetic studies indicate that the *E. coli* C chromosomal genes which are responsible for catabolism of the pentitol sugars, ribitol and D-arabitol, are not present in the closely related *E. coli* K12 strains (Reiner 1975). Molecular studies of these tightly linked genes reveal that they are surrounded by 1.4 kilobase inverted repeats of imperfect homology (Link and Reiner 1982). Here we report that *E. coli* C lacks genes for catabolism of the hexitol sugar galactitol, genes which are present in *E. coli* K12. Furthermore, the ribitol-arabitol and galactitol genes, which show no mutual homology, are mutually exclusive when exchanged (by homologous recombination) between *E. coli* C and K12. Physical characterization of λ specialized transducing phages carrying the ribitol-arabitol or galactitol genes demonstrates that this exclusion results because these genes have identical locations in their respective chromosomes. This novel type of allelic relationship between non-homologous genes has not been previously described in prokaryotes. Analysis of the catabolic capabilities of a collection of natural *E. coli* strains suggests that this exclusion relationship extends to strains in the natural *E. coli* population. We suggest an insertion/deletion model to account for the origins of this unusual gene arrangement.

The ribitol-arabitol genes of *E. coli* C lie between *metG* and *his* in the bacterial chromosome (Scangos and Reiner 1978a). After transduction by bacteriophage P1 from *E. coli* C into K12, these genes occupy the same position in the K12 chromosome, a region in which genes for catabolism of galactitol had been previously mapped (Lengeler 1977). We found that when the ribitol-arabitol genes of *E. coli* C (which is galactitol-negative (Gat^-)) are transduced into a Gat^+ K12 strain, the recipient strain invariably loses the ability to catabolize galactitol. Similarly, transducing the galactitol genes from *E. coli* K12 into C causes this strain to lose the ability to catabolize ribitol and D-arabitol (>400 transductants screened). This observed mutual exclusion could result from either a physiological interference or a recombinational displacement of genes. When ribitol-arabitol genes are introduced into a Gat^+ K12 strain by lysogenizing this strain with λdAR, a specialized transducing phage containing these genes (Link and Reiner 1982; Scangos and Reiner 1978a), the resulting Rtl^+ Atl^+ strains remain Gat^+. Similarly, when galactitol genes are introduced into a Rtl^+ Atl^+ C strain by lysogenizing this strain with λgat (see below), the resulting Gat^+ strains remain Rtl^+ Atl^+. These results demonstrate that the catabolic pathways do coexist when their genes are located at different positions in the chromosome, thus making the physiological explanation extremely unlikely.

Two models could explain the apparent recombinational displacement of these genes. In the first model, *E. coli* C contains only the ribitol-arabitol genes and K12 contains only the galactitol genes. These genes are not homologous and have the same location in their respective chromosomes. Introduction of one set of genes by homologous recombination in surrounding regions must necessarily result in displacement of the other. In a second model, *E. coli* C contains non-functional galactitol genes closely linked to functional ribitol-arabitol genes, while K12 contains non-functional ribitol-arabitol genes closely linked to functional galactitol genes. Because of these close linkages, non-functional genes are highly cotransducible with functional genes, leading to the apparent mutual exclusion. However, previous genetic studies have shown that there are no ribitol-arabitol sequences in K12 (Reiner 1975). Furthermore, the occasional Rtl^+ Atl^+ Gat^+ transductants which would result from recombinational cross-overs within an homologous ribitol-arabitol-galactitol region have never been observed.

To distinguish conclusively between these models, we allowed the displacement event to occur on λdAR, which subsequently was physically characterized. If the first model were correct, the displacement event would lead to the replacement of a region of λdAR DNA by a new, non-homologous DNA segment. If the second model were correct, the displacement would lead to replacement by a homologous segment, causing no gross change in λdAR sequences. λdAR, (which has no galactitol-transducing ability), was lytically propagated on Gat^+K12 strain RW361, and the resulting lysate was titered for galactitol- and ribitol-transducing ability. Galactitol-transducing phages (designated λgat) were recovered at high frequency, roughly 1% the frequency of recovered ribitol-transducing phages. These λgat phages do not transduce ribitol or D-arabitol catabolism. Formation of λgat phage depends on the presence of the ribitol-arabitol genes, since none are recovered when

Offprint requests to: A.M. Reiner

* *Present Address:* Department of Molecular, Cellular and Developmental Biology, University of Colorado, Campus Box 347, Boulder, CO 80309, USA

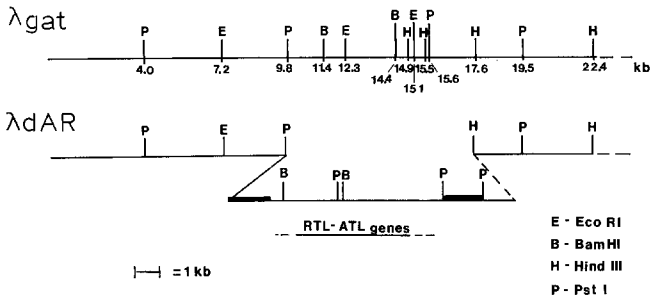

Fig. 1. Restriction analysis of the λgat and λdAR genomes. Shown are the left arms of these phages, which contain the bacterial-derived sequences. λdAR was derived from λY199 using the in vivo technique of Schrenk and Weisberg (Schrenk and Weisberg 1975). λgat was obtained by lytically propagating λdAR on strain RW361 (K12 Gat$^+$ *Δλ att*) and selecting galactitol-transducing phage from the resulting lysate. Both phages are defective and were grown by temperature induction of double lysogens containing λY199 as the helper phage. Transducing particles were purified by banding in CsCl equilibrium gradients (Miller 1972) and phage DNA was prepared by phenol extraction. Darkened bars on the λdAR map indicate the position of fragments which hybridize to the inverted-repeat-specific probe (Link and Reiner 1982). Determination of the positions of the ribitol-arabitol genes has been previously described (Link and Reiner 1982)

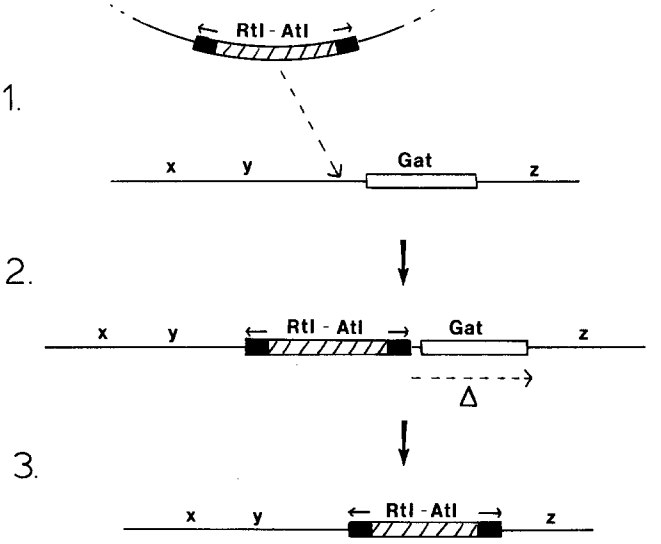

Fig. 2. Insertion/deletion model for origin of "allelic" relationship between the ribitol-arabitol and galactitol genes. Step one illustrates the insertion of a ribitol-arabitol transposon near the galactitol genes of a K12-like ancestral strain. The short arrows surrounding the ribitol-arabitol genes indicate the position of inverted repeats. Neighboring chromosomal sequences are designated by the letters x, y and z. Step two illustrates a deletion, originating from one of the inverted repeats of the ribitol-arabitol transposon, which removes the galactitol genes. This results in a strain (Step 3) which has the ribitol-arabitol genes in the same relative position that the galactitol genes occupied in the original strain

λY199, the parent phage of λdAR which lacks bacterial sequences, is propagated on a Gat$^+$K12 strain.

Restriction enzyme analyses of λdAR and λgat (Fig. 1) show that the formation of λgat results from the replacement of the ribitol-arabitol region of λdAR with a new segment of DNA. This new segment was shown to contain the galactitol genes by sub-cloning it into a recombinant plasmid (designated pGAT), which then could transform cells to Gat$^+$. The regions of homology and non-homology between λgat and λdAR, surmised by comparison of their restriction maps, were confirmed by Southern blot hybridization studies using as a probe pAR-2, a recombinant plasmid containing the *E. coli* C ribitol-arabitol region (Link and Reiner 1982). These results show that the ribitol-arabitol and galactitol genes are surrounded by the same sequences in their respective chromosomes, so that a recombination between these sequences would lead to the substitution of the galactitol genes for the ribitol-arabitol genes. These results also rule out the possibility that the ribitol-arabitol genes are located in the midst of galactitol sequences in *E. coli* C, since, were this the case, λgat phages would have an apparent deletion of λdAR ribitol-arabitol sequences, but would not acquire any new sequences.

How might this "allelic" relationship between the ribitol-arabitol and galactitol genes have arisen? The observation that the ribitol-arabitol genes of *E. coli* C are surrounded by imperfect inverted repeats has led to the suggestion that these genes may constitute a vestigial transposon (Link and Reiner 1982). The following model, outlined in Fig. 2, shows how the insertion of a ribitol-arabitol transposon could have led to the current "allelic" relationship between these genes. First, a ribitol-arabitol transposon inserted near the galactitol genes of a K12-like ancestral strain. Then, a subsequent deletion, originating from the terminus of one of the transposon inverted repeats, removed the galactitol genes. (The termini of inverted repeats of transposable elements are known to be hot spots for deletion formation (Kleckner et al. 1979; Kleckner 1981).) The resulting strain, like *E. coli* C, would contain the ribitol-arabitol genes in the same relative position that the galactitol genes occupied in the original K12-like strain.

A prediction of this insertion/deletion hypothesis is that the λdAR sequences which are displaced in λgat will correspond only to sequences which constitute the presumed ribitol-arabitol transposon. Within the limits of resolution of the restriction analysis, the inverted repeats which surround the ribitol-arabitol genes lie at the ends of the λdAR DNA segment which is lost from λgat. The absence of these sequences in λgat was confirmed by demonstrating that a probe specific for the inverted-repeat sequences (Link and Reiner 1982) does not hybridize to λgat. These results are consistent with the hypothesis that the endpoints of the λdAR segment displaced in λgat are the outside termini of the inverted repeats surrounding the ribitol-arabitol genes.

Some natural *E. coli* strains can catabolize ribitol, D-arabitol, and galactitol (see below). The genetic structure of one such strain, RM48B, was examined by cloning its entire ribitol-arabitol-galactitol chromosomal region into pBR322. The restriction map of the resulting plasmid, designated pGAR-1, is shown in Fig. 3. Comparison of this map with those of pAR-2 and pGAT indicates that the ribitol-arabitol and galactitol genes are adjacent in strain RM48B. This strain therefore has a structure similar to that of the proposed intermediate strain of the insertion/deletion model (see Fig. 2).

The ribitol-arabitol/galactitol exclusion observed between K12 and C strains is reflected also in the natural *E. coli* population, at least at the phenotypic level. Ninety-one natural *E. coli* strains (primarily from the R. Milkman collection (Milkman 1973)) were screened for their ability

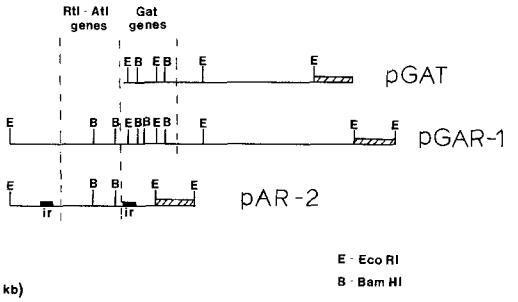

Fig. 3. Comparison of restriction maps of recombinant plasmids containing sequences cloned from *E. coli* strains K12, C and RM48B. Plasmid pGAT was constructed by subcloning two contiguous HindIII fragments of λgat into pBR322. This plasmid contains *E. coli* K12 and phage sequences and can transform *E. coli* strains to Gat^+. Plasmid pGAR-1 was constructed by ligating partial EcoRI digest fragments of RM48B chromosomal DNA into pBR322 and selecting recombinant plasmids which could transform recipients to Rtl^+. These transformants were then screened for utilization of D-arabitol and galactitol; transformants containing pGAR-1 could catabolize ribitol, D-arabitol and galactitol. Plasmid pAR-2 contains an EcoRI fragment of *E. coli* C chromosomal DNA which contains the ribitol-arabitol genes (Link and Reiner 1982). The darkened bars labeled "ir" indicate the position of the inverted repeats which surround the ribitol-arabitol genes. Note that pGAR-1 contains two adjacent regions with restriction site positions nearly identical to those of pAR-2 and pGAT, respectively. pGAR-1 contains an extra Bam HI site in its galactitol region which is not present in pGAT; this is presumably a result of divergence between these two sequences

Table 1. Ribitol and galactitol catabolism of natural *E. coli* strains

Frequency of trait in population	Phenotype	Number of strains	Probability
Gat^+ – 60.5%	Rtl^+ Gat^-	16	
	Rtl^- Gat^+	43	<0.025
Rtl^+ – 30.5%	Rtl^+ Gat^+	12	
	Rtl^- Gat^-	20	

The 91 natural *E. coli* strains used in this screen were primarily from the R. Milkman collection (Milkman 1973), and were obtained from B. Levin. The catabolic traits of these strains were determined by transferring cells of each strain with a sterile toothpick to a minimal salts plate containing either ribitol or galactitol as the sole carbon source. Strains which produced confluent growth after three days growth at 30° C were scored as positive, as were strains which threw off good-growing revertants after five days. The probability that these distributions of traits were due to chance were calculated using a chi-squared test with one degree of freedom

to utilize ribitol and galactitol. As shown in Table 1, these are variable traits among natural strains, and Rtl^+Gat^+ strains are significantly rarer ($P<0.025$) than one would predict from the frequencies of the individual genes.

We do not know what, if any, selective pressures may have led to the current "allelic" relationship between the ribitol-arabitol and galactitol genes. Strains containing both sets of genes do not appear to suffer any measurable deleterious effects (unpublished data), although under certain conditions the catabolic pathways coded by these genes are known to produce toxic intermediates (Reiner 1977; Scangos and Reiner 1978b). If the observed genotypic exclusion is a consequence of transposition/deletion (Fig. 2), non-homologous "alleles" may be expected to be found in other prokaryotic and eukaryotic genomes. Our results add to the growing evidence (Riley and Anilionis 1978) that variability among closely related strains is not restricted to the base sequence level, but include heterogeneities at higher levels of genetic organization.

Acknowledgements. We would like to thank Tom Carbone, Anne Ferguson and Jeanne Bertonis for excellent technical assistance during the course of this work, and Dr. Bruce Levin for providing the Milkman strains. This work was supported by N.S.F. grant PCM-7811161. C.D.L. was supported by N.I.H. graduate traineeship grant 074 73-04.

References

Kleckner N, Reichardt K, Botstein D (1979) Inversions and deletions of the *Salmonella* chromosome generated by the translocatable tetracycline resistance element tn10. J Mol Biol 127:89–115

Kleckner N (1981) Transposable elements in prokaryotes. Annu Rev Genet 15:341–404

Lengeler J (1977) Analysis of mutations affecting the dissimilation of galactitol (dulcitol) in *Escherichia coli* K12. Mol Gen Genet 152:83–91

Link CD, Reiner AM (1982) Inverted repeats surround the ribitol-arabitol genes of *E. coli* C. Nature 298:94–96

Milkman R (1973) Electrophoretic variation in *Escherichia coli* from natural sources. Science 182:1024–1026

Miller JH (1972) Preparation of λh80 d*lac* phage and transformation with λh80 d*lac* DNA. In: Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, pp 319–321

Reiner AM (1975) Genes for ribitol and D-arabitol catabolism in *Escherichia coli:* their loci in C strains and absence in K12 and B strains. J Bacteriol 123:530–536

Reiner AM (1977) Xylitol and D-arabitol toxicities due to derepressed fructose, galactitol, and sorbitol phosphotransferases of *Escherichia coli.* J Bacteriol 132:166–173

Riley M, Anilionis A (1978) Evolution of the bacterial genome. Ann Rev Microbiol 32:519–560

Scangos GA, Reiner AM (1978a) Ribitol and D-arabitol catabolism in *Escherichia coli.* J Bacteriol 134:492–500

Scangos GA, Reiner AM (1978b) Acquisition of ability to utilize xylitol: disadvantages of a constitutive catabolic pathway in *Escherichia coli.* J Bacteriol 134:501–505

Schrenk WJ, Weisberg RA (1975) A simple method for making new transducing lines of coliphage lambda. Mol Gen Genet 137:101–107

Communicated by G. O'Donovan

Received October 27, 1982