Published in final edited form as:

Science. 1994 July 15; 265(5170): 407–409.

Adaptive Reversion of a Frameshift Mutation in *Escherichia coli* by Simple Base Deletions in Homopolymeric Runs

Patricia L. Foster and Jeffrey M. Trimarchi

Department of Environmental Health, Boston University School of Public Health, Boston
University School of Medicine, Boston, MA 02118, USA

Abstract

Spontaneous mutations are thought to occur primarily in growing cells. However, spontaneous mutations also arise in nutritionally deprived cells, and in some cases this process appears to be adaptive. Here it is reported that when a Lac⁻ strain of *Escherichia coli* is under selection for lactose use, the spectrum of Lac⁺ mutations that arises is different, and simpler, than that arising without selection. Mutations appearing during selection were mainly one-base deletions in runs of iterated bases. Similar mutations occurring in repetitive DNA elements are associated with a variety of human hereditary diseases and are increased in cells that cannot correct heteroduplex DNA.

The mechanisms by which spontaneous mutations arise in growing cells have been the subject of much research. Two causes of spontaneous mutation that are often cited are intrinsic polymerase errors and endogenous DNA lesions. Indeed, many of the mutations that arise spontaneously in bacteria and in their viruses are the same types of errors as those made by DNA polymerases replicating damaged and undamaged templates in vitro (1,2). However, it is unclear if such mechanisms can account for the mutations that arise among cells that are apparently not growing or replicating their DNA (3). Many theories have been proposed to explain how mutations occur in nongrowing cells and why those mutations are adaptive (3–11). Comparisons between the types of mutations that arise during nonselected growth and those that arise under selective conditions may help to choose among these theories (12).

In previous publications (7,8), we described the appearance of Lac⁺ revertants among populations of a strain of *Escherichia coli*, FC40, that cannot use lactose because of a frameshift mutation affecting the *lacZ* gene. Lac⁺ revertants occurring during exponential growth, which were detected 2 days after cells were plated with lactose as the only carbon source, arose at about 10⁻⁹ per cell per generation, a rate that is well within the normal range for mutations of this type. After day 2, Lac⁺ revertants continued to arise at a nearly constant rate of about 10⁻⁹ per cell per hour. The post-plating mutants quickly became the main class, and 90 to 95% of the Lac⁺ revertants that had appeared after a few days were apparently the result of mutations that occurred after plating. This class of mutants did not appear if the cells were starved in the absence of lactose or in the presence of lactose if there was another, unfulfilled growth requirement (7). In addition, mutation to Lac⁺ during lactose selection, but not during prior growth of the cultures, required some function or functions of the major recombination pathway, RecABC (7,11). In the work reported here, we

Correspondence to: Patricia L. Foster.

Note added in proof: After submission of this manuscript, we learned that others (24) had obtained similar results.

determined the sequence changes in Lac⁺ revertants that arose each day after FC40 was plated on minimal lactose plates.

The mutant *lac* allele, *lacI33*, carried by FC40 on an F' episome, derives from an in-frame fusion of the *lacI* gene to the *lacZ* gene (13) but has a +1 frameshift mutation in *lacI* that is polar on *lacZ* (14). Because the *lacI* coding sequence is not essential, this allele can be reverted by any mutation that restores the reading frame but does not create a nonsense mutation. Such events include simple –1-bp deletions as well as more complex DNA rearrangements that restore the reading frame within the 130-bp target shown (Fig. 1).

Newly arising, independent Lac⁺ mutants of FC40 were collected on days 2 to 5 after the cells were plated on minimal lactose plates (15). The DNA from these mutants was analyzed by amplification and sequencing (16), and the results are summarized in Tables 1 and 2 and in Fig. 1. Nearly all of the Lac⁺ mutants that arose after day 2 carried –1-bp deletions, and most of these occurred in runs of three to five bases. In contrast, about 50% of the Lac⁺ revertants that were isolated on day 2 carried mutations that resulted in deletions, duplications, and rearrangements. Thus, complex mutations are frequent in the absence of selection but are rare during selection. Considering only these two classes of intragenic revertants, there were nine complex mutations and 11 –1-bp deletions among the mutants isolated on day 2 but only one complex mutation and 30 –1-bp deletions among the mutants isolated after day 2 ($\chi^2 = 10.9$, $P = 0.001$). Indeed, the one late-arising deletion mutant, isolated on day 3, was only weakly Lac⁺ (see Table 3), which suggests that this complex mutation may also have arisen during prior growth of the culture but took an extra day to appear as a Lac⁺ colony.

Because late-arising Lac⁺ colonies might have been the result of mutations that gave only a weak Lac⁺ phenotype, we measured the β -galactosidase activity of an assortment of mutants (Table 3). The mutant with the largest deletion had the greatest activity, 422 Miller units, which suggests that this deletion creates a strong promoter. Not surprisingly, the three extragenic mutants that were isolated had rather low β -galactosidase activity. Among the –1-bp deletion revertants, the levels of β -galactosidase activity were about the same—most ranged from about 200 to 300 Miller units regardless of the day on which they appeared.

These results limit the types of mechanisms that might be responsible for the Lac⁺ revertants of FC40 that arose during selection. Although we did not sequence regions outside the target region, we read approximately 200 bases of sequence for each mutant and found no sequence changes other than the ones that reverted the frameshift. In addition, the reverting mutations were found at seven different sites. The lack of silent mutations and the diversity of mutational sites argue against the theory that the late-arising mutations resulted from recombination with homeologous (similar but nonhomologous) sequences located elsewhere in the genome (10). A distinctive feature of the late-arising frameshift mutations is that 90% occurred in runs of three or more bases. This site specificity is typical of –1-bp frameshift mutations made by DNA polymerases in vitro (2) and can be explained by replication of a misaligned template (17). Thus, the late-arising revertants of FC40 might be the result of simple polymerase errors, which supports our previous conclusion, based on genetic evidence, that adaptive mutations require some form of DNA replication (8). In vivo, mutations at iterated sequences are greatly enhanced by the loss of mismatch repair functions (18), which raises the possibility, previously suggested (5), that in nutritionally deprived cells, error correction activity may be low, or that mispaired or misaligned DNA may be inaccessible to the error-correcting enzymes. Similar types of mutations, apparently due to slippage during replication, repair, or recombination of repetitive DNA elements, have recently been associated with a variety of human hereditary diseases (19).

The late-arising revertants of FC40 are a subset of the kinds of mutations that can revert *lacI33*, yet they are distinguished from the mutations that arise during nonselective growth by their requirement for RecABC (20). One straightforward explanation for this result is that – 1-bp deletions are produced by the same mechanism both before and during selection but that there is an extra requirement for some function or functions of RecABC in the latter case. For example, if the – 1-bp deletions are the result of polymerase errors, RecABC may be required to initiate DNA synthesis, to preserve the products, or both. RecABC can initiate DNA synthesis by producing D-loops (21) and (perhaps by the same mechanism) by catalyzing unequal crossing-over between regions of homology (22). Both of these mechanisms would require that more than one copy of the *lac* region be present in a cell. Thus, the mutations might arise in a subpopulation that has duplicated the *lac* region (8,23) or simply among cells (probably the majority) that have more than one copy of the episome. If DNA synthesis is initiated only infrequently or is restricted in extent, yet is error-prone, this would help to explain why the amount of DNA synthesis measured in non-dividing cells appears to be inadequate to account for the mutations that arise (3). If the products of this synthesis are transitory unless the cell achieves a useful mutation and begins to divide, this would explain why the only mutations that are recovered are adaptive (4,7–9).

Acknowledgments

We thank J. Cairns, J. Drake, E. Eisenstadt, M. Marinus, and F. Stahl for discussions and comments on this manuscript. Supported by NSF grant MCB-9213137 to P.L.F.

REFERENCES AND NOTES

1. Drake JW. Annu Rev Genet 1991;25:125. [PubMed: 1812804] Lindahl T. Nature 1993;362:709. [PubMed: 8469282] Ripley LS. Annu Rev Genet 1990;24:189. [PubMed: 2088167] Echols H, Goodman MF. Annu Rev Biochem 1991;60:477. [PubMed: 1883202]
2. Kunkel TA, Bebenek K. Biochim Biophys Acta 1988;951:1. [PubMed: 2847793]
3. Foster PL. Annu Rev Microbiol 1993;47:467. [PubMed: 8257106]
4. Cairns J, et al. Nature 1988;335:142. [PubMed: 3045565]
5. Stahl FW. :112. *ibid.* Boe L. Mol Microbiol 1990;4:597. [PubMed: 2191182]
6. Ryan FJ, et al. Z Vererbungsl 1961;92:38. [PubMed: 13744985] Hall BG. Genetics 1990;126:5. [PubMed: 2227388] Steele DF, Jinks-Robertson S. 1992;132:9. *ibid.*
7. Cairns J, Foster PL. Genetics 1991;128:695. [PubMed: 1916241]
8. Foster PL, Cairns J. 1992;131:783. *ibid.*
9. Stahl FW. 1992;132:865. *ibid.*
10. Hastings, PJ.; Rosenberg, SM. Encyclopedia of Immunology. Roitt, IM.; Delves, PJ., editors. Saunders; London: 1992. p. 602 Higgins NP. Trends Biochem Sci 1992;17:207. [PubMed: 1323887]
11. Harris RS, et al. Science 1994;264:258. [PubMed: 8146657]
12. Hall BG. Genetica 1991;84:73. [PubMed: 1756965] Prival MJ, Cebula TA. Genetics 1992;132:303. [PubMed: 1427030]
13. Müller-Hill B, Kania J. Nature 1974;249:561. [PubMed: 4599764]
14. Calos MP, Miller JH. J Mol Biol 1981;153:39. [PubMed: 7040679]
15. Lac⁺ mutants were collected from two separate experiments, each using 30 independent cultures of FC40 plated on M9 0.1% minimal lactose plates (7). In the first experiment, about 10⁸ FC40 cells were plated with 10⁹ nonreverting scavengers [FC29 (7)], and one Lac⁺ colony was isolated from each plate on days 2 to 4. In the second experiment, only 10⁷ FC40 cells were plated with 10⁹ scavengers, which allowed one well-separated Lac⁺ colony to be isolated per plate through day 5. Each Lac⁺ colony was streaked on minimal-lactose plates, and a single colony was saved for analysis.

16. Genomic DNA was prepared from purified Lac⁺ mutants, and oligonucleotide primers hybridizing to the beginning of the *lacI* sequence (base pairs 21 to 46 or 211 to 233) and to the *lacZ* coding sequence (base pairs 1680 to 1703 or 1684 to 1703) were used to amplify the *lac* region (numbers are from the Ecolac sequence in GenBank). With none of 62 mutants tested did the *lac* region fail to amplify with these primers; thus, major rearrangements that eliminate the beginning of the *lacI* gene, or separate it too far from *lacZ* for the intervening region to be amplified, are rare. Mutants whose products differed in size from that of unrevverted FC40 (included as a control for all amplifications) were not further analyzed, except that the mutant with the smallest detectable difference was sequenced and proved to be a 112-bp deletion. Forty-two amplified products that showed no size changes were sequenced by cycle sequencing (Circumvent kit, New England Biolabs) with an oligonucleotide primer to base pairs 918 to 939 (Fig. 1). Products from unrevverted FC40 were frequently sequenced in parallel to check that polymerase errors made during amplification were not appearing in the sequences. In three mutants, two arising on day 4 and one on day 5, no compensating mutation was found. These proved to carry extragenic mutations because they failed to transfer the Lac⁺ phenotype with their episomes.
17. Streisinger G, et al. Cold Spring Harbor Symp Quant Biol 1966;31:77. [PubMed: 5237214]
18. Cupples CG, et al. Genetics 1990;125:275. [PubMed: 2199309] Schaaper RM, Dunn RL. Proc Natl Acad Sci USA 1987;84:6220. [PubMed: 3306672] Strand M, et al. Nature 1993;365:274. [PubMed: 8371783] Fishel R, et al. Cell 1993;75:1027. [PubMed: 8252616] Leach FS, et al. : 1215. *ibid.* Parsons R, et al. :1227. *ibid.*
19. Kunkel TA. Nature 1993;365:207. [PubMed: 8371775]
20. We have sequenced a number of Lac⁺ mutants isolated on day 2 from *recA*-deficient derivatives of FC40 and have found examples of all the classes of mutations that occur in the *recA*⁺ strain, including -1-bp deletions in runs. Thus, none of the types of mutations that can revert FC40 are intrinsically *recA*-dependent during growth.
21. Asai T, Kogoma T. J Bacteriol 1994;176:1807. [PubMed: 8144445]
22. Anderson RP, Roth JR. Proc Natl Acad Sci USA 1981;78:3113. [PubMed: 6789329]
23. Roth, JR.; Stahl, FW. personal communication.
24. Rosenberg SM, et al. Science 1994;265:405. [PubMed: 8023163]
25. We have sequenced three additional jackpots of FC40 from other experiments. The mutations were: -AT at base pair 1042, a two-bp duplication (CpT on the sense strand) at base pair 1089, and duplication of base pairs 1067 to 1086.
26. Miller, JH. Experiments in Molecular Genetics. Cold Spring Harbor Laboratory; Cold Spring Harbor, NY: 1972.

```

921      931      941      951      961
ATG TGAATACCG GGTCAACCA CAGCAACA GATTTTTC CTCCTGGGC
Primer+
971      981      991      1001      1011
AAACCAAGCT GACCCCTTC CTCACACTC CTCGGGCA GGGCTGAAG
-1 Stop
1021      1031      1041      1051      1061
GTAATACAGC TTTTGGGCT CTCCTGGTC AAAAGAAAA CACCTCTGC
+1 Frameshift
1071      1081      1091      1101      1111
GCGATATAGC CAACCTCTC CTCCTGGTC GTTGGCTAT TCAATATGC
1121      1131      1141
ACCTGGAGAG AGAGTTTC GACCTATTC
+3 Stop

```

Fig. 1.

The target for reverting mutations in the *lacI33-lacZ* allele. Sites at which –1-bp deletion mutations were found more than once are indicated in bold. Numbering is as in the Ecolac sequence in Gen-Bank up to base pair 1144, but the extra C at base pairs 1036 to 1038, which creates the *lacI33* allele, is not numbered. The coding strand is shown. The *lacI-lacZ* gene fusion (13) eliminates the last five residues of LacI, all of the *lac* promoter and operator, and the first 23 residues (or 24 if the initiating methionine is included) of LacZ (base pairs 1145 to 1356 of the Ecolac sequence). The fusion protein is transcribed from the mutant *lacI^q* promoter. The frameshift mutation at base pairs 1036 to 1038 was induced by ICR191 (14). This mutation creates a stop codon at base pair 1145 (base pair 1358 of the Ecolac sequence), and a –1 frameshift mutation upstream of base pair 1016 would create a stop codon at base pair 1016. Thus, to revert *lacI33*, the reading frame must be restored between base pairs 1016 and 1145.

Table 1

Summary of Lac⁺ revertants.

Mutation	Day 2	Day 3	Day 4	Day 5	Days 3 to 5
Loss of bases [*]	5	1	0	0	1
Gain of bases [†]	4	0	0	0	0
−1-bp deletion	11	9	12	9	30
Extragenic	0	0	2	1	3
Total	20	10	14	10	34

^{*} Detected by a decrease in the size of the amplified product.

[†] Detected by an increase in the size of the amplified product, except for one, a duplication of base pairs 1087 to 1094 (Fig. 1), that was detected by sequencing.

Table 2

Number of occurrences of various frameshift mutations. Blank spaces indicate that no mutant was found.

Mutation*	Length of run	Day 2	Day 3	Day 4	Day 5	Days 3 to 5
–GC at 1020	3		1	1	1	2
–GC at 1036	4	4	7	8	5	20
–GCat 1039	1				1	1
–AT at 1056	5	2			2	2
–AT at 1067	1	1 [‡]				
–GCat 1068	2		1			1
–GCat 1072	3	2		3		3
–AT at 1078	1	1				
–GC at 1093	4	1				
–GC at 1106	2		1			1

* Only the first base pair in a run is given.

[‡] The culture contained a jackpot of 193 mutants per 10⁷ cells. A total of five mutants were amplified and three were sequenced, all of which were the same, which confirmed that the –AT frameshift was the mutation giving rise to the jackpot (25).

Table 3

The β -galactosidase activity of various mutants. Lac⁺ mutants were grown to saturation in M9 0.1% minimal glycerol medium (7). The β -galactosidase activity was measured as in (26), and the results are given in Miller units. The numbers are for independently isolated mutants of each class. Blank spaces indicate either that no mutant was found or that none were assayed

Mutation*	Day 2	Day 3	Day 4	Day 5
Loss of bases	76; 180; 422	80		
Gain of bases	73; 183; 272			
-GCat 1020			269	283
-GCat 1036	223	240		265
-GCat 1039				285
-AT at 1056				182
-AT at 1067	225			
-GCat 1068		165		
-GCat 1072			127; 170; 241	
-AT at 1078	216			
-GCat 1093	135			
-GCat 1106		165		
Extragenic			53; 59	76

* Only the first base pair in a run is given.