The *lacI* Gene as a Target for Mutation in Transgenic Rodents and *Escherichia coli*

Johan G. de Boer and Barry W. Glickman

Centre for Environmental Health, University of Victoria, Victoria, British Columbia, Canada V8W 3N5

ABSTRACT

The *lacI* gene has been used extensively for the recovery and analysis of mutations in bacteria with various DNA repair backgrounds and after exposure to a wide variety of mutagens. This has resulted in a large database of information on mutational mechanisms and specificity of many mutagens, as well as the effect of DNA repair background on mutagenicity. Most importantly, knowledge about the mutational sensitivity of the *lacI* gene is now available, yielding information about mutable nucleotides. This popularity and available knowledge resulted in the use of the *lacI* gene in transgenic rodents for the study of mutagenesis in mammals, where it resides in ~40 repeated copies. As the number of sequenced mutations recovered from these animals increases, we are able to analyze the sites at which mutations have been recovered in great detail and to compare the recovered sites between bacteria and transgenic animals. The nucleotides that code for the DNA-binding domain are nearly saturated with base substitutions. Even after determining the sequences of ~10,000 mutations recovered from the animals, however, new sites and new changes are still being recovered. In addition, we compare the nature of deletion mutations between bacteria and animals. Based on the nature of deletions in the animals, we conclude that each deletion occurs in a single copy of the gene.

MUTATIONAL spectra: Occasional changes in the sequence of nucleotides are referred to as mutations. Such mutations, when in moderation, are the driving force behind evolution. Mutations, however, may also result in cancer and inherited disease. Endogenous processes in the cell related to cellular metabolism and DNA replication can bring about mutations (Lutz 1990; Bridges 1996; Drake 1991). Mutations can also be the result of exogenous insults to the cell. These can include physical agents such as both ionizing and nonionizing radiation, heat, or chemical agents, which interact with DNA directly or after metabolic activation. Studying mutations for many years has led to the concept of a mutational fingerprint (Glickman 1990). This paradigm states that different treatments result in different mutagenic events that are characterized by both the nature of the mutational alteration and the sequence context of the change. In essence, a specific mutagen correlates with a specific pattern of mutagenic events. For example, ultraviolet light results in specific mutations in *Escherichia coli* that are targeted to dipyrimidine sequences (Miller 1985), while dimethylnitrosamine causes specifically G:C→A:T transitions (Glickman *et al.* 1987). Treatment with *N*-methyl-*N′*-nitro-*N*-nitrosoguanidine or *N*-ethyl-*N*-nitrosourea results in G:C→A:T transitions at G nucleotides preceded 5′ by a purine (the "PuG effect"; Burns *et al.* 1987, 1988), while the G:C→A:T transitions produced by ethyl methanesulfonate do not depend on the PuG context (Burns *et al.* 1986). Benzo[*a*]pyrene diolepoxide induces G:C→T:A transversions in mammalian cells, mainly at runs of guanines flanked by adenine residues (Mazur and Glickman 1988). Even spontaneous or background mutation, a special pastime of Drake (1996), reveals a characteristic pattern. Spontaneous mutation in mammalian cells is characteristically dominated by G:C→A:T transitions at methylated cytosines in 5′-CpG-3′ dinucleotide sequences. This specificity of mutation is what makes it possible to infer an induced response from the determination of the sequence alterations in mutants recovered after potential exposure. It is thus the changes in mutational spectra that can reveal mutagenic exposures.

The facile analysis of mutation requires a nonessential target gene in which mutations can be detected efficiently. Several such systems are currently available, including the mammalian hypoxanthine-guanine phosphoribosyl transferase (*hprt*) and adenine phosphoribosyl transferase (*aprt*), and the bacterial *gpt* and *supF* genes, which code for a bacterial homolog of *hprt* and a suppressor tRNA molecule, respectively, and the bacterial *lacI* and *lacZ* genes.

The *lacI* gene: The *lacI* gene of the *lac* operon has been used extensively for the recovery and analysis of mutations in bacteria with various DNA repair backgrounds and after exposure to a wide variety of mutagens. The LacI protein product represses the transcription of the adjacent *lacZ* gene by binding to its operator sequence (Gilbert and Muller-Hill 1970). The active LacI protein is a homotetramer of the LacI polypeptide and binds to two copies of the *lac* operator (Sauer 1996; Lewis *et al.* 1996). The lacI monomer is composed of

Corresponding author: Johan G. de Boer, Centre for Environmental Health, University of Victoria, P.O. Box 3020, Victoria, BC Canada V8W 3N5. E-mail: jdboer@uvic.ca

a DNA-binding domain (amino acids 1–59) and a core domain (amino acids 60–359) involved in oligomerization and inducer binding. The DNA-binding domain [or negative complementation (NC^+) region (Miller 1972)] is highly sensitive to base substitution mutation. Many amino acids in this domain interact with nucleotides in the operator, as is evidenced by the resolved repressor-DNA crystal structure (Lewis *et al.* 1996). When mutant LacI protein is unable to form a tetramer and bind the operator, transcription of the *lacZ* gene occurs. In most constructs, an amino-terminal fragment, or an α-lacZ fragment, is produced. This fragment may complement a carboxy-terminal or omega fragment that is provided by an appropriate host cell. This full complement has β-galactosidase activity, which can be used in a screening and in a selective assay.

The first studies using the *lacI* gene in bacteria for mutational analysis were pioneered by the group of Miller (Miller and Low 1984; Miller 1982, 1985; Coulondre *et al.* 1978; Foster *et al.* 1983) by using an analysis of amber and ochre mutants and deletion mapping (Miller *et al.* 1977). Once DNA sequencing of the *lacI* gene became feasible (Calos and Miller 1981), the facile generation of mutational spectra became a possibility. Dozens of publications, many from the laboratory headed by J. W. Drake, centered around the use of the *lacI* gene. A clever recombination assay for the recovery of the mutants (Schaaper *et al.* 1985) aided in the analysis of mutation induced by a large variety of agents (Gordon *et al.* 1991; Lambert *et al.* 1991; Schaaper *et al.* 1986, 1987, 1990). This has resulted in a large database of information on mutational mechanisms and specificity of many mutagens as well as the effect of DNA repair background on mutagenicity. Most importantly, knowledge about the mutational sensitivity of the *lacI* gene is now available, yielding information about the available sites and the effect of the "protein filter" (de Boer and Glickman 1991). This great understanding of the *lacI* gene has set the stage for use of the gene in transgenic animals.

Recently, the *lacI* gene was used as the mutation target in transgenic mice (Kohler *et al.* 1991a,b). The pLIZ plasmid containing the *lacI* gene was inserted into a bacteriophage lambda genome and injected into mouse embryonic cells. This resulted in the generation of mouse lines transgenic for the *lacI* gene (the Big Blue mouse). The construct is present in the animals in ~40 tandem copies. A similar construct has been made with Fischer F344 rats (Dycaico *et al.* 1994) and a rat embryonic fibroblast cell line (Wyborski *et al.* 1995), while an embryonic fibroblast cell line has been derived from the C57BL/6 Big Blue mouse (Wyborski 1997). The recovery of the *lacI* gene from tissues of these animals is greatly facilitated by the bacteriophage lambda construct. The addition of a lambda packaging extract to genomic DNA results in the formation of bacteriophage particles containing single bacteriophage lambda genomes and, therefore, single *lacI* genes. The presence of a mutated *lacI* gene is detected with a color assay by plating the phage particles on a lawn of bacteria in the presence of 5-bromo-4-chloro-3-indolyl-β-d-galactopyranoside (X-gal; Rogers *et al.* 1995; Young *et al.* 1995).

TABLE 1

The *lacI* gene in nucleotides and base pairs

	G	C	G:C	A	T	A:T
Whole gene	311	299	610	240	233	473
NC^+ (29–205)	52	45	97	47	33	80
Core (206–1111)	259	254	513	193	200	393

Origin of the mutants: The DNA sequences of mutants recovered from bacteria have been obtained from published literature (~14,000 mutants) and entered into a computer database (de Boer 1995). Mutations in the Big Blue mouse, rat, and cell line have been obtained from our own unpublished material (~10,000 mutants) and from published literature (~1500 mutants) and entered into the database. The literature references to these mutants can be found on the Big Blue web site (http://darwin.ceh.uvic.ca/bigblue/bigblue.htm).
The intensity of blue plaque color was estimated for a group of mutants by replating on X-gal using the CM0, CM1, CM2, and CM3 color intensity standards for comparison (Rogers *et al.* 1995).

The nucleotide targets in the *lacI* gene: The coding sequence of the *lacI* gene contains 1083 nucleotide pairs (including the termination codon) and codes for a 359-amino acid polypeptide. The numbering of the base positions used is according to Farabaugh (1978). The first transcribed nucleotide is at position 1, the first translated codon is at position 29–31, and the termination codon is at position 1109–1111. The sense strand of the gene is comprised of a total of 311 guanines and 299 cytosines, for a total of 610 G:C base pairs, and 240 adenines and 233 thymines for a total of 473 A:T base pairs (Table 1).

The G:C content of the gene is 56% compared to an average of 44% in the mouse genome (Davidson 1960) and 52% in the *E. coli* genome (Schildkraut *et al.* 1962). Because of its bacterial origin, the gene has a relatively high percentage of 5′-CpG-3′ dinucleotide sequences, compared to mammalian genes. The significance of this difference is discussed in later sections.

Mutational saturation: Knowing the complete set of sites in a gene where mutations can be recovered greatly facilitates the interpretation of mutational spectra. Figure 1 shows the number of base substitution sites that have been recovered and the total number of different changes at these sites as a function of the number of mutants from Big Blue animals that we have sequenced in our laboratory. Surprisingly, even after sequencing 10,000 mutants, the gene has obviously not yet been

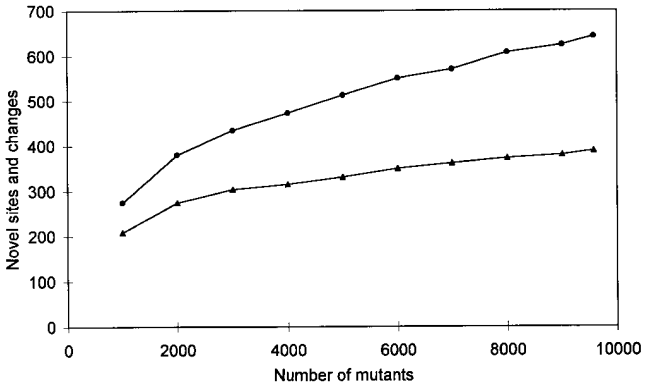


Figure 1.—Saturation of the *lacI* gene. Number of sites recovered (▲) and number of individual substitutions (●) as a function of the number of mutants sequenced in our laboratory.

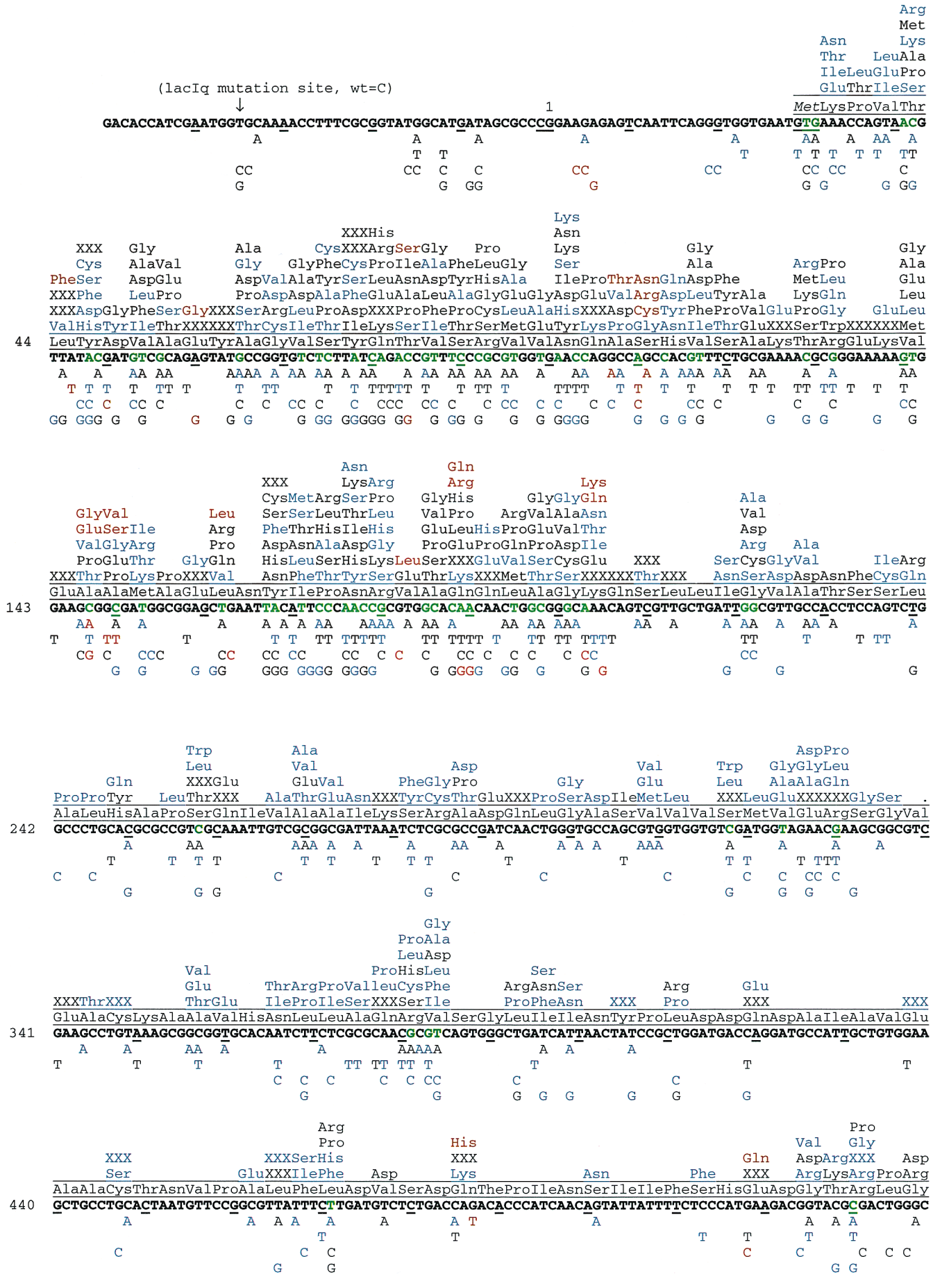
saturated with base substitution mutations. The data include mutants recovered from various tissue types from control animals and after treatment with a wide variety of mutagens. More sites will be recovered as additional mutagenic treatments with different sequence specificity are included. The number of sites and base substitution changes found in bacteria (also based on ~10,000 substitution mutations) is relatively small, compared to those found in the transgenic animals. It should be kept in mind, however, that in many published studies using bacteria, only mutations in the DNA-binding (NC^+) region were analyzed. Because this involves only approximately the first 200 base pairs, this would result in fewer sites recovered overall. Even when only the first 200 base pairs are considered in both mice and bacteria, however, the Big Blue database still has more recovered sites and changes that are unique to the transgenic system than the bacterial database, even though considerably more bacterial mutants are available for this part of the gene. At a number of sites, however, mutations were recovered in bacteria but not in the animal (Figure 2). The *lacI* gene is the only gene in which such a degree of mutational saturation has been achieved.

Figure 2 presents all nucleotide positions at which substitution mutations have been recovered. It also shows which new nucleotides were recovered at those positions, and the amino acids that are coded for by the new sequences. The data from bacterial and animal studies are summarized in Table 2, A–C. A total of 86 As, 151 Cs, 157 Gs, and 113 Ts have been found mutated in the *lacI* gene recovered from both transgenic animals and bacteria (Table 2C). The largest class in this tabulation is the C→T transition, undoubtedly because of its contribution to spontaneous or background mutation in the animal. In the NC^+ region (from base pairs 29–205), we found 75.7% (134 out of 177) of the sites mutated in the bacterial plus Big Blue data combined. A total of 305 individual changes were found at these 134 sites. At 62 sites in the NC^+ region, all three possible base changes have been recovered (green colored nucleotides in Figure 2). We note that the classic definition of the NC^+ region ends where a large clustering of mutation ends. The recovery approximates the theoretical limit, when most changes at third positions in codons are considered silent. This region can therefore be considered mutationally saturated, at least when considering nucleotide positions.

Knoll *et al.* (1994, 1996) reported mutants recovered after direct selection on lactose. A total of 115 different base changes were found at 81 nucleotide positions (only the first 195 coding positions were used in their studies). Interestingly, when compared to the BigBlue/X-gal database, fully 24.3% of these changes was not found in the large X-gal collection. Equally, when the lactose-selected mutants were compared to the smaller bacterial data set, 20.9% of the changes were novel. Furthermore, only 22–26% of the recovered blue mutant colonies was found to harbor a mutation in the NC^+ region of the gene. This suggests that direct selection with lactose has a different selection threshold than screening with X-gal or selection with phenyl-β-d-galactopyranoside (P-gal), and as such is reminiscent of the results obtained in the *lacI* system used by LeClerc *et al.* (1988). These authors used the *lacI*^{s} gene, which is alluded to in a later section, and found that mutation was not clustered to the NC^+ region.

The carboxy-terminal end of the gene (after position 1015) is particularly devoid of mutation (Figure 2). This is true in the data collection recovered from bacterial as well as from Big Blue and, therefore, less likely to be an artifact of sequencing strategy. Coding of this part of the gene includes the terminal α-helix that is involved in the tetrahelix bundle formation. This helix bundle is important for the formation of a tetramer structure. Failure of the formation of a tetramer because of mutation does not abolish repression ability (Brenowitz *et al.* 1991), which may explain the lack of recovery of mutations in this region. Mutations were also recovered in the 5′-flanking sequence, *viz.* in the −35 and −10 promoter boxes, around position 5, and at the Shine-Delgarno sequence around position 20.

Differences in mutation spectrum recovered in different parts of the gene: We have analyzed the mutational spectrum, as recovered from Big Blue animals, as a function of position in the *lacI* gene. Figure 3 shows the main components of the spontaneous spectrum (G:C→A:T transitions, G:C to A:T at CpG sites, G:C→T:A transversions, and deletions combined from various sources) in the first 400 and the next 600 base pairs. This analysis involves ~2000 spontaneous mutants recovered from various tissues. Interestingly, the contribution of G:C→A:T transitions, especially their distribution over 5′-CpG-3′ sites, is particularly concentrated in the first 400 base pairs, where 65–85% of them are located at 5′-CpG-3′ sites. In the remaining part of the gene, only


(lacIq mutation site, wt=C)
1
MetLysProValThr
GluThrIleSer
GACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAATTCAGGGTGGTGAATGTGAAACCAGTAACG
XXXHis
Lys
Asn
XXX Gly Ala CysXXXArgSerGly Pro Lys Gly Gly
Cys AlaVal Gly GlyPheCysProIleAlaPheLeuGly Ser Ala ArgPro Ala
PheSer AspGlu AspValAlaTyrSerLeuAsnAspTyrHisAla IleProThrAsnGlnAspPhe MetLeu Glu
XXXPhe LeuPro ProAspAspAlaPheGluAlaLeuAlaGlyGluGlyAspGluValArgAspLeuTyrAla LysGln Leu
XXXAspGlyPheSerGlyXXXSerArgLeuProAspXXXProPheProCysLeuAlaHisXXXAspCysTyrPheProValGluProGly GluLeu
ValHisTyrIleThrXXXXXXThrCysIleThrIleLysSerIleThrSerMetGluTyrLysProGlyAsnIleThrGluXXXSerTrpXXXXXXMet
44
LeuTyrAspValAlaGluTyrAlaGlyValSerTyrGlnThrValSerArgValValAsnGlnAlaSerHisValSerAlaLysThrArgGluLysVal
TTATACGATGTCGCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTG
Asn Gln
XXX LysArg Arg Lys
CysMetArgSerPro GlyHis GlyGlyGln Ala
GlyVal Leu SerSerLeuThrLeu ValPro ArgValAlaAsn Val
GluSerIle Arg PheThrHisIleHis GluLeuHisProGluValThr Asp
ValGlyArg Pro AspAsnAlaAspGly ProGluProGlnProAspIle Arg Ala
ProGluThr GlyGln HisLeuSerHisLysLeuSerXXXGluValSerCysGlu XXX SerCysGlyVal IleArg
XXXThrProLysProXXXVal AsnPheThrTyrSerGluThrLysXXXMetThrSerXXXXXXThrXXX AsnSerAspAspAsnPheCysGln
GluAlaAlaMetAlaGluLeuAsnTyrIleProAsnArgValAlaGlnGlnLeuAlaGlyLysGlnSerLeuLeuIleGlyValAlaThrSerSerLeu
143
GAAGCGGCGATGGCGGAGCTGAATTACATTCCCAACCGCGTGGCACAACAACTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTG
Trp Ala AspPro
Leu Val Asp Val Trp GlyGlyLeu
Gln XXXGlu GluVal PheGlyPro Gly Glu Leu AlaAlaGln
ProProTyr LeuThrXXX AlaThrGluAsnXXXTyrCysThrGluXXXProSerAspIleMetLeu XXXLeuGluXXXXXXGlySer .
AlaLeuHisAlaProSerGlnIleValAlaAlaIleLysSerArgAlaAspGlnLeuGlyAlaSerValValValSerMetValGluArgSerGlyVal
242
GCCCTGCACGCGCCGTCGCAAATTGTCGCGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGGTGTCGATGGTAGAACGAAGCGGCGTC
Gly
ProAla
LeuAsp
Val ProHisLeu Ser
Glu ThrArgProValleuCysPhe ArgAsnSer Arg Glu
XXXThrXXX ThrGlu IleProIleSerXXXSerIle ProPheAsn XXX Pro XXX XXX
GluAlaCysLysAlaAlaValHisAsnLeuLeuAlaGlnArgValSerGlyLeuIleIleAsnTyrProLeuAspAspGlnAspAlaIleAlaValGlu
341
GAAGCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACTATCCGCTGGATGACCAGGATGCCATTGCTGTGGAA
Arg Pro
Pro His Val Gly
XXX XXXSerHis XXX Gln AspArgXXX Asp
Ser GluXXXIlePhe Asp Lys Asn Phe XXX ArgLysArgProArg
AlaAlaCysThrAsnValProAlaLeuPheLeuAspValSerAspGlnThrProIleAsnSerIleIlePheSerHisGluAspGlyThrArgLeuGly
440
GCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGC

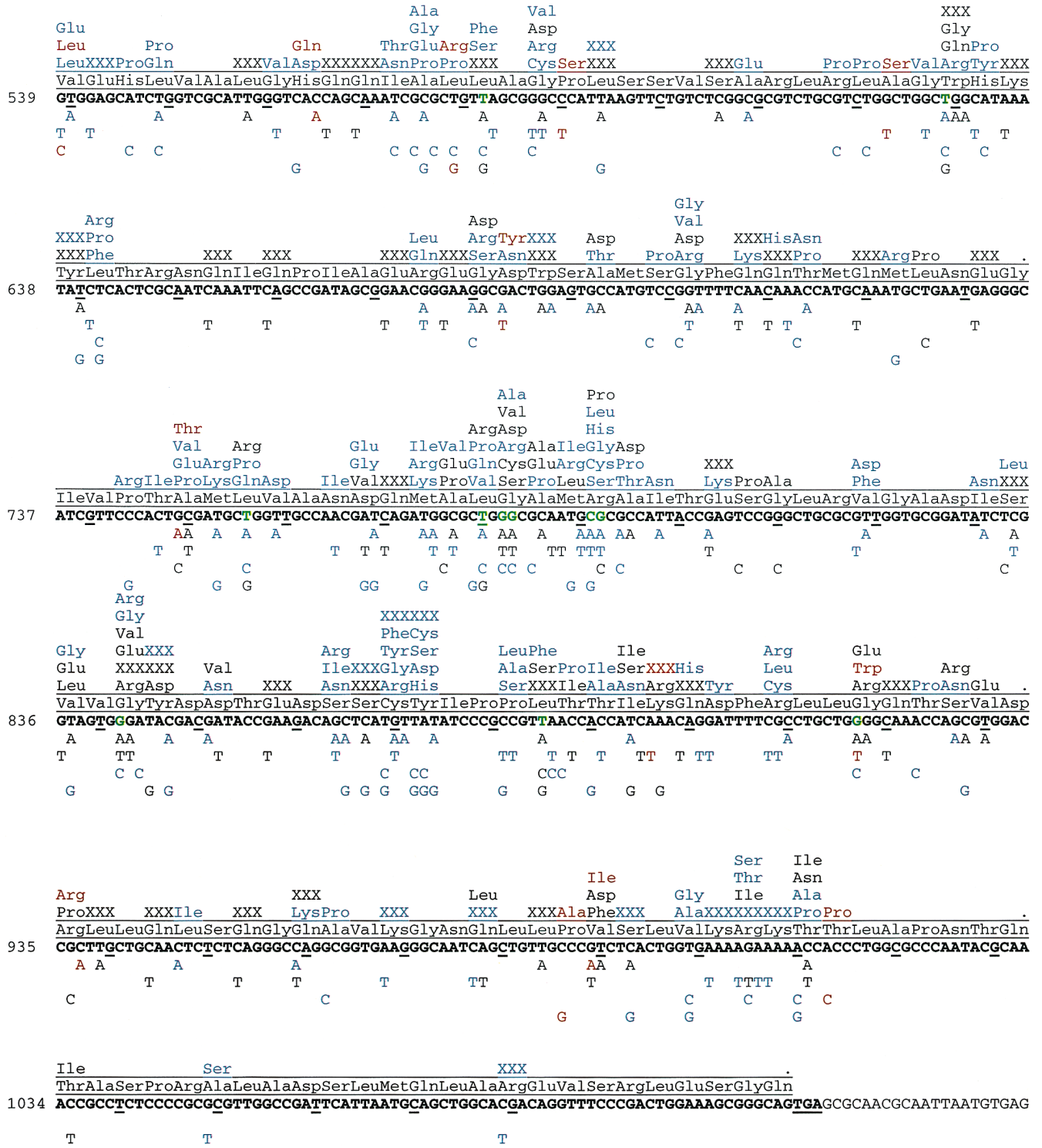


Figure 2.—Numbering starts at 1 at the mRNA transcription start. The preceding nucleotide is −1 (there is no "0"). The translation start is at position 29–31. The initiation codon is GTG, which codes for Val, but is read as fMet at the initiation codon (*e.g.*, a GTG to GTC mutation would not change the Val coding, but because it is the initiation codon, it prevents initiation). Promoter mutations are present in the −10 box and the −35 box. The Shine-Delgarno sequence is around position 20. XXX represents a termination codon. Nucleotides in green, all three changes have been recovered; black, found in both bacteria (P-gal) and Big Blue (X-gal); blue, found in Big Blue only (X-gal); red, found in bacteria only (P-gal).

20% of the recovered G:C→A:T transitions are found at 5′-CpG-3′ sites. In contrast, deletion events are threefold more common after the first 400 base pairs. It can be concluded that the characteristic large percentage of G:C→A:T transitions that occurs at 5′-CpG-3′ sites in the *lacI* gene in animals reflects the prevalence of these

TABLE 2

A. Changes in BigBlue (X-gal), based on 12,078 mutants

		Wild-type base is mutated to:			
Wild type	Total sites	A	C	G	T
A	63	—	34	24	43
C	127	74	—	47	82
G	131	69	61	—	76
T	101	70	51	59	—
	422				

B. Changes in bacteria (P-gal), based on 13,459 mutants

		Wild-type base is mutated to:			
Wild type	Total sites	A	C	G	T
A	30	—	17	12	18
C	70	38	—	14	45
G	71	28	28	—	36
T	47	31	17	27	—
	218				

C. All changes recovered in BigBlue (X-gal) and bacteria (P-gal)

		Wild-type base is mutated to:			
Wild type	Total sites	A	C	G	T
A	86	—	46	34	56
C	151	86	—	54	111
G	157	95	79	—	90
T	113	82	64	66	—
	507				

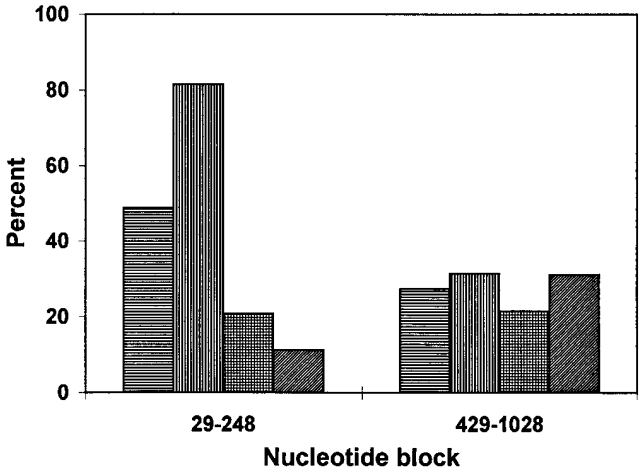


Figure 3.—Comparison between the first and second part of the gene. First column, G:C→A:T transitions; second column, G:C→A:T transitions at 5′-CpG-3′ sites; third column, G:C→T:A transversions; fourth column, nonsubstitution mutations.

mutations in the DNA-binding region. The two parts of the *lacI* gene can therefore be considered as two different sequences and we show here that these exhibit quite different mutational characteristics. Even identical sequences, which are present in different parts of the genome, can be subject to different mutational events. This was shown in the *hrpt* gene by Lichtenauer-Kaligis *et al.* (1996). Mutations recovered from five copies of the *hrpt* cDNA, which were inserted at different places in the genome, differed significantly from each other, both in mutant frequency and in spectrum.

The density of 5′-CpG-3′ sites (total nucleotides = 2 × CpGs) as a function of 200 base pair blocks along the *lacI* gene is shown in Figure 4. In addition, the number of those sites (as nucleotides) that have actually been found mutated and as a fraction are shown. The largest fraction of affected sites is found in the first part of the gene. The largest contribution of CpG sites in mutations at G:C base pairs is in the first 400 base pairs of the gene. The density of 5′-CpG-3′ dinucleotide sequences is similar for the NC^+ region and the rest of the gene (0.096 CpGs per base pair in the NC^+ region *vs.* 0.086 CpGs per base pair in the remainder of the gene.) The fraction of G:C base pairs that can be found at a CpG sequence is 0.351 for the NC^+ region and 0.256 for the remainder of the gene.

Differences in methylation patterns between mice and bacteria: Spontaneous deamination of methylated cytosines results in G:C→A:T transitions (Ehrlich *et al.* 1990). In bacteria, methylation takes place at the inner cytosine of 5′-CC(A/T)GG sequences and methylation-directed transition mutations can therefore be recovered as hotspots at these sites. These nucleotides are found at positions 104, 419, 959, and 1016, and transitions have been recovered multiple times at positions 104, 419, and 959 (Figure 2). Methylation of cytosines in mammalian cells does not occur at these sequences, but instead occurs at 5′-CpG-3′ dinucleotide sequences. The spectrum of spontaneous mutations in mammals shows a predominance of G:C→A:T transitions at 5′-CpG-3′ dinucleotide sequences, in reflection of this methylation. As a consequence, ~75% of all G:C→A:T transitions in the *lacI* gene in Big Blue are found at these sites. The gene in mice is reported to be fully methylated in all copies present in the genome (Provost and Short 1994). This high methylation pattern in the *lacI* gene results in an increased spontaneous background, which may mask differences in spontaneous mutation when different tissues are compared (de Boer *et al.* 1998). There are 95 CpG sites in the coding sequence of *lacI*, a density of 88 per 1000 base pairs. Mammalian genes generally have a lower density of 5′-CpG-3′ sequences in the coding portion of their genes to minimize the accumulation of spontaneous mutation. For example, the mouse *hrpt* gene has only 15 CpG sites per 1000 base pairs in its coding sequence, while the human p53 has 35 per 1000 base pairs. The frequency of spontaneous transitions can therefore be expected to be five- to sixfold higher in the *lacI* gene compared to the *hrpt* gene based on the number of

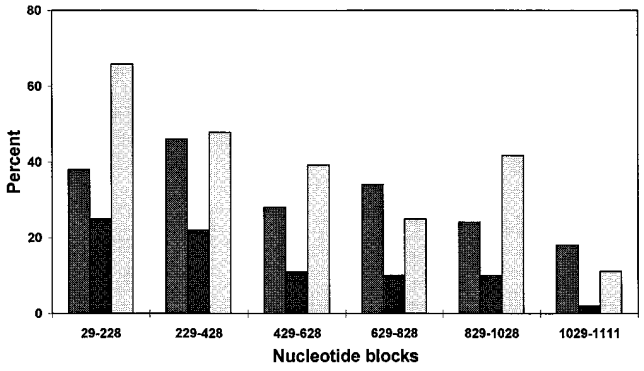


Figure 4.—Percent of G:C base pairs that are part of a 5′-CpG-3′ dinucleotide sequence (first column), recovered as mutation (second column), and as a percent of those found at 5′-CpG-3′ sites that are mutated (third column), as a function of position in the *lacI* gene.

these sites alone. This density of potential methylation sites in the *lacI* gene in the mouse compares to four per 1000 base pairs for methylation hotspots in *E. coli*. The sensitivity of the *lacI* system to mutation in Big Blue has been compared to that of the *hrpt* gene in the same animal (Skopek *et al.* 1995). The fold increase after chemical treatment of mutation in the *hrpt* gene is significantly higher than in the *lacI* gene. This may partially be explained by the higher background caused by potential methylation sites in the *lacI* gene. A *lacI* gene with only 13 CpG sequences has been constructed by Skopek *et al.* (1996) to provide a target with a potentially lower background mutant frequency. The mutant frequency in this construct, however, was only marginally lower than that of the native *lacI* sequence, and the distribution of mutations, including G:C→A:T transitions, was now shifted to the 5′ end of the gene, where mutation at the remaining CpG sites was enhanced (T. Skopek, personal communication).

Of the 610 G:C base pairs in the *lacI* gene, a total of 190 are part of the 95 5′-CpG-3′ sites. This amounts to 31% of all G:C base pairs. Randomly distributed, a maximum of 31% of mutations at G:C base pairs would therefore be expected to be found at CpG sites, when all changes would be recoverable. In reality, this percentage is lower because not all of those changes will result in detectable phenotypic changes. Mutations in Big Blue have been recovered at 258 G:C base pairs. G:C→A:T transitions at CpG sites have only been found at 47 nucleotides, while G:C→T:A transversions at CpG sites have been found at 51 nucleotides. This results in 18.2 and 19.8%, respectively, that can be expected to be found at 5′-CpG-3′ dinucleotide sequences. About 75% of all G:C→A:T transitions are typically found at CpG sites. In addition, we notice that ~45% of all G:C→C:G and G:C→A:T transversions in all spontaneous mutations are found at CpG sites. Both of these percentages are significantly higher than expected. The transitions are typically attributed to deamination events at methylated cytosines, when they are part of CpG sequences.

An additional explanation has been offered, however; 5′-CpG-3′ dinucleotide sequences, or CpG steps, are unusually malleable (Lefebvre *et al.* 1995; El Antri *et al.* 1993). The various conformations that can be adopted depend on the sequence context. This may offer an explanation for the higher fraction of mutations that is recovered at such sites, in addition to, or partly the reason for the more commonly accepted explanation of deamination of methylated cytosines. Mauffret *et al.* (1991) found that elliptinium, an intercalating drug, has a preference for binding at CpG and TpG sites. The sequence context effect may contribute to different mutability at different CpG sequences. Ketterling *et al.* (1994) and Bottema *et al.* (1991) discuss the transversions observed at 5′-CpG-3′ sites in the human factor IX gene. Their analysis indicates that the mutation rate for these events is markedly elevated compared to transversions occurring at other sites. The free radical product 8-hydroxyguanine pairs preferentially with cytosine and adenine, producing G→T transversions. If this radical reaction or the mispairing is favored by CpG methylation or by CpG structural features, we would expect to see an increase in transversions at CpG sites.

Two systems to detect mutations: Mutations in the *lacI* gene can be detected by two principal methods, both of which use the presence or absence of β-galactosidase activity. The first method relies on the color of a plaque or a colony for the screening for mutants. A mutation in the *lacI* gene results in the expression of β-galactosidase. Besides galactose, which is its normal substrate, this enzyme can also cleave X-gal. Cleavage of X-gal, a chromogenic compound, results in the production of an insoluble dye that colors the colony or plaque blue. A *lacI* gene with an altered promoter, $lacI^q$ (Muller-Hill *et al.* 1968), is usually used. This involves a C→T base substitution at position −35 in the "−35" promoter box. This mutation results in a higher level of expression of LacI protein and, therefore, in better repression of background *lacZ* expression. Reversion mutants at this site have been found among mutants recovered from the transgenic animals (Figure 2).

A second system for the detection of mutations in the *lacI* gene uses a selective assay. In addition to galactose and X-gal, β-galactosidase can also cleave P-gal. This results in the production of galactose, which can provide a sole carbon source, thus permitting the growth of bacterial colonies. A wild-type *lacI* gene will, therefore, not support colony formation. This system has been used by most researchers in bacterial studies.

A great many publications resulted from the use of the *lacI* gene in the study of mutational mechanisms and mutagenic specificity. Mutations in the recovered mutants can be detected in either the entire *lacI* gene or only in the portion coding for the DNA-binding domain, up to nucleotide position 205, which is highly sensitive to base substitutions. Mutants in this region can be selected for by a NC^+ assay (Miller 1972; Gu *et al.* 1994).

Some researchers used the *lacI*^{s} gene instead, which carries a single C→T transition at position 617, resulting in a replacement of Arg 197 by Cys (Christensen *et al.* 1985; LeClerc *et al.* 1988). This slightly "destabilizes" the LacI protein. Together with X-gal screening, this results in a system that is more sensitive to amino acid alterations, and only ~10% of the recovered mutations are now found in the NC^{+} region of the gene, rather than ~60%, resulting in an entirely different distribution of the recovered mutants.

The nature of the *lacI* gene (such as ^{q} or ^{s} mutations) and the detection system are determining factors in the nature of the mutations that are recovered, as has been already pointed out by Gordon and Halliday (1995). In the majority of the studies involving the *lacI* gene in *E. coli*, mutants were recovered after growth selection on P-gal. In contrast, the Big Blue mouse mutants were scored on the basis of their plaque color phenotype on X-gal. Knoll *et al.* (1994, 1996) used a lactose selection system for the recovery of mutants from Big Blue mice. The stringency of mutant detection may not be the same. Detection by color requires the formation of a blue dye, the amount of which depends on the level of repression of the *lacZ* gene. Growth on P-gal depends on the cleaving of enough P-gal to yield sufficient galactose to provide a carbon source for colony or phage growth. Using the P-gal selection system with the *lacI*^{q}-L8 gene (L8 is a *lacZ* promoter mutant) results in a majority of the mutants (~70%) having a deletion or an insertion of a four-base pair sequence (5′-CTGC-3′) at the spontaneous deletion/insertion hotspot at position 620–632, and only 11% is recovered as base substitutions (Halliday and Glickman 1991; Schaaper *et al.* 1986). The use of the X-gal color detection system with either the *lacI*^{q} or *lacI*^{s} gene, however, results in an increase in the fraction of base substitutions (50–87%) and a minor contribution of plus or minus four base pair mutations at the hotspot (0–14%; LeClerc *et al.* 1988; Yatagai and Glickman 1990), which is similar to what is recovered from Big Blue animals. This indicates that the use of X-gal results in a much greater sensitivity for missense mutations in the *lacI* gene. This prompted Gordon and Halliday (1995) to conclude that the bacterial spectra obtained with X-gal are a better comparison for the mouse-derived mutations than the traditionally selected spectra using P-gal. This is certainly true when substitution mutations and the plus or minus four-base pair hotspot mutations are compared. We compared the base substitution component of these spectra recovered from bacteria (*lacI*^{s}) with X-gal (Yatagai and Glickman 1990), from bacteria (*lacI*^{q}) with P-gal (Halliday and Glickman 1991), and from mouse (*lacI*^{q}) with X-gal. It is striking that the ratios of the various substitution classes are quite similar (Table 3), with the exception, perhaps, of the data from LeClerc *et al.* (1988), a compilation of data from vegetative and ex-conjugant cells.

We would expect differences in the spectra of mutants collected with X-gal and P-gal, where this would result in a LacI protein activity level that may discriminate between the two requirements. When the large collections of mutants recovered from bacteria with P-gal and from Big Blue animals with X-gal screening are compared, some differences are apparent. Several mutations were recovered in bacteria with P-gal selection that have not been seen with X-gal screening in the large Big Blue database, notably around nucleotide positions 110 and 147. Many sites where mutations were found in Big Blue were not recovered in bacteria; however, this may partially result from the lower number of substitutions recovered in bacteria.

Plaque color and threshold of detection: Bacteriophage particles with mutated *lacI* genes are detected in the Big Blue assay on the basis of a blue plaque phenotype. The intensity of the blue color, however, depends on the residual functionality of the LacI protein. Plaques that are a very light blue color may therefore be missed during the screening. The standardized assay uses four color standards (the CM series) that consists of four mutants with increasing color intensity (Rogers *et al.* 1995). We determined the plaque color of more than 120 mutants upon replating, using the CM series as color standards. Figure 5 shows the color that was observed for these mutants as a function of the position of the mutated nucleotide, regardless of the actual change. Mutations in the first 100 base pairs are generally dark blue. This probably reflects the importance of the amino acids in this region for their interactions with the DNA. Light-colored plaques are especially common in the range of positions 200 and 330. Large variations in color are seen in this region between mutants with changes in adjacent base pairs. This region coincides well with a concentration of I^{s} mutants (which are unable to respond to inducer) that was seen by Kleina and Miller (1990) after extensive *in vitro* mutagenesis of the *lacI* gene. Mutations in several other regions in the gene also result in a dark blue phenotype, including between positions 430 and 500, and other regions in the portion of the gene that codes for the core domain.

Interestingly, frameshift and stop codon mutations can also result in the very light blue phenotype. Table 4 shows selected frameshift and stop codon mutations, as well as their positions, that were recovered as very light blue (CM0 or CM1) or as dark (CM3). CM0 itself is a translation termination at base pair position 530. In general, however, a termination results in a dark color, even near the end of the gene. Frameshifts generally result in dark-colored plaques. Interestingly, frameshifts at 503 and 507 and the introduction of a stop codon at 530 result in a faint blue plaque color. The frameshift at 503 and 507 adds the same 27 and 26 nonnative amino acids, respectively, before terminating at the same position. This deletion would remove approximately half of the protein but would possibly leave

TABLE 3

Comparison of spontaneous base substitution spectra

	Bacteria / P-gal[a]	Bacteria / X-gal[b]	Big Blue / X-gal[c]
G:C→A:T	56.3	42.5	58.7
A:T→G:C	6.3	5.0	7.2
G:C→T:A	16.3	20.0	22.1
G:C→C:G	3.8	16.3	4.7
A:T→T:A	8.8	11.3	3.4
A:T→G:C	8.8	5.0	3.8
Total substitution mutants	80	80	235

[a] Halliday and Glickman (1991)
[b] LeClerc *et al.* (1988)
[c] de Boer *et al.* (1997)

dimerization potential intact. Visualization of the protein structure using RasMol molecular modeling software (Sayle and Milner-White 1995) suggests that one of two major surface areas involved in dimer formation is still intact. In contrast, several frameshift mutants near the end of the gene, *e.g.*, a +1 frameshift at 1006 that yields a slightly larger protein, result in a dark plaque color.

When the detection threshold for the blue color increases, the light-colored mutants would be missed, potentially skewing the mutational spectrum. This may happen when the plating media do not conform to standardized protocols (Rogers *et al.* 1995). We have determined the DNA alterations in a collection of 22 mutants that were scored as CM0 or CM1 from control animals in various experiments (data not shown). The collection is small, but an Adams and Skopek (1987) comparison shows a *P* value of 0.9 when compared with a collection of 282 spontaneous mutants recovered from liver (de Boer *et al.* 1997). This indicates that even though the mutant frequency may be underestimated, the mutational spectra are unlikely to be biased by omitting light colored mutant plaques from a spectrum. This is comforting because it allows some variation in the sensitivity of the assay for the analysis of the spectrum and the detection of mutant induction on the basis of the spectrum.

A comparison of deletion mutations: Deletion of DNA sequence is frequently mediated by repeated sequences (Streisinger *et al.* 1966; Krawczak and Cooper 1991; Jego *et al.* 1993). Misalignment of the template of nascent strand during extension or after DNA breaks can result in the loss of the sequence between these repeats plus one of the repeats itself. We have analyzed the deletions that have been recovered in the *lacI* gene isolated from *E. coli* and *lacI* transgenic rodents. In transgenic animals, the *lacI* gene is present as part of a 45-kb lambda genome that is repeated ~40 times in a tandem array for a total length of ~1.8 Mb (Dycaico *et al.* 1994). Deletions that have one endpoint in one copy and the other endpoint in another copy would reveal themselves upon recovery as a simple intragenic deletion or as a duplication event, depending on the location of the endpoints of the actual deletion. Approximately 200 deletion mutations in bacteria and 200 from Big Blue were used in this analysis. The deletions recovered range in length from two to >800 base pairs in length. Approximately 65% of these deletions are flanked by short, perfectly repeated sequences from one

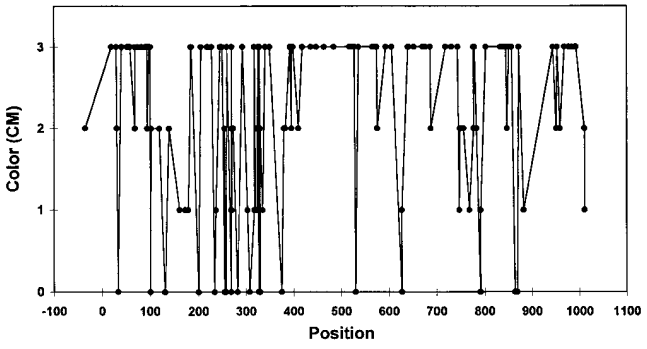


Figure 5.—The plaque color as a function of the mutated nucleotide position. Color is compared to the standard CM series.

TABLE 4

Plaque color of selected termination mutants

Position	Mutation	Color
503	−1 FS, adds 27 amino acids	CM0
507	−1 FS, adds 26 amino acids	CM0
530	Arg→Stop	This is CM0
883	−1 FS, adds one amino acid	CM3
916	−1 FS, adds 16 amino acids	CM3
1006	+1 FS, adds 19 amino acids	CM3

Plaque color is compared to the CM color standards supplied by Stratagene (La Jolla, CA).

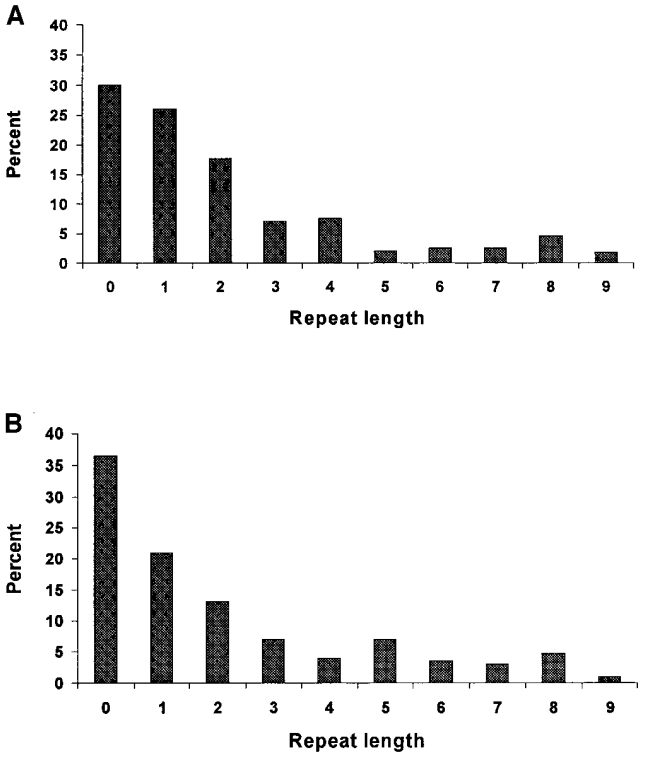


Figure 6.—The participation of repeated sequences in deletion mutations recovered from Big Blue and from bacteria. Shown is the percentage of deletions at each repeat length recovered from (A) Big Blue animals and (B) *E. coli*.

to eight base pairs long. Intergenic deletion events may equally well result in duplication of *lacI* sequence. Duplication events recovered from the Big Blue animals, however, are exceedingly rare. In addition, the participation of terminal repeat sequences in the deletion events is very comparable in bacteria and rodents (Figure 6). These observations indicate that deletion events in the Big Blue animals are mainly intragenic deletions and do not span multiple copies. This can be understood when it is realized that the target for such deletion endpoints would only be 1000 base pairs in each 50,000-base pair lambda unit. The recovery of such deletions, therefore, is unlikely to be very efficient.

Summary: No other gene has been used so frequently in mutation studies as the *lacI* gene. The number of mutations and its mutational saturation is unparalleled. With the advent of transgenic technology, *lacI* has entered a new era of mutational study in which mutation can be studied in various tissues and in multiple species of different ages, both males and females. Thus, the *lacI* gene will continue to be an excellent target for the study of mutational specificity and mechanism.

The authors thank Dave Walsh, Ken Sojonki, James Holcroft, Pam Warrington, and Naheed Hague for excellent technical assistance. This work was supported by contract NO1-ES-35365 from the National Institute for Environmental Health Sciences, USA, and the National Cancer Institute of Canada.

LITERATURE CITED

Adams, W. T., and T. R. Skopek, 1987 Statistical test for the comparison of samples from mutational spectra. J. Mol. Biol. **194:** 391–396.

Bottema, C. D., M. J. Bottema, R. P. Ketterling, H. S. Yoon, H. L. Janco *et al.*, 1991 Why does the human factor IX gene have a G+C content of 40%? Am. J. Hum. Genet. **49:** 839–850.

Brenowitz, M., N. Mandal, A. Pickar, E. Jamison and S. Adhya, 1991 DNA-binding properties of a Lac repressor mutant incapable of forming tetramers. J. Biol. Chem. **266:** 1281–1288.

Bridges, B. A., 1996 Mutation in resting cells: the role of endogenous DNA damage. Cancer Surv. **28:** 155–167.

Burns, P. A., F. L. Allen and B. W. Glickman, 1986 DNA sequence analysis of mutagenicity and site specificity of ethyl methanesulfonate in UvrB$^+$ and UvrB$^-$ strains of *Escherichia coli*. Genetics **113:** 811–819.

Burns, P. A., A. J. E. Gordon and B. W. Glickman, 1987 Influence of neighbouring base sequence on N-methyl-N′-nitro-N-nitrosoguanidine mutagenesis in the *lacI* gene of *Escherichia coli*. J. Mol. Biol. **194:** 385–390.

Burns, P. A., A. J. E. Gordon, K. Kunsmann and B. W. Glickman, 1988 Influence of neighboring base sequence on the distribution and repair of N-ethyl-N-nitrosourea-induced lesions in *Escherichia coli*. Cancer Res. **48:** 4455–4458.

Calos, M. P., and J. H. Miller, 1981 Genetic and sequence analysis of frameshift mutations induced by ICR-191. J. Mol. Biol. **153:** 39–66.

Christensen, R. B., J. R. Christensen and C. W. Lawrence, 1985 Conjugation-dependent enhancement of induced and spontaneous mutation in the *lacI* gene of *E. coli*. Mol. Gen. Genet. **201:** 35–37.

Coulondre, C., J. H. Miller, P. J. Farabaugh and W. Gilbert, 1978 Molecular basis of base substitution hotspots in *Escherichia coli*. Nature **274:** 775–780.

Davidson, M., 1960 *The Biochemistry of Nucleic Acids*, Ed. 8. Chapman & Hall, London.

de Boer, J. G., 1995 Software package for the management of sequencing projects using *lacI* transgenic animals. Environ. Mol. Mutagen. **25:** 256–262.

de Boer, J. G., and B. W. Glickman, 1991 Mutational analysis of the structure and function of the adenine phosphoribosyltransferase enzyme of Chinese hamster. J. Mol. Biol. **221:** 163–174.

de Boer, J. G., H. Erfle, D. Walsh, J. Holcroft, J. S. Provost *et al.*, 1997 The spectrum of spontaneous mutants in liver tissue of *lacI* transgenic mice. Environ. Mol. Mutagen. **30:** 273–286.

de Boer, J. G., S. Provost, N. Gorelick, K. Tindall and B. W. Glickman, 1998 Spontaneous mutation in *lacI* transgenic mice: a comparison of tissues. Mutagenesis (in press).

Drake, J. W., 1991 Spontaneous mutation. Annu. Rev. Genet. **25:** 125–146.

Drake, J. W., 1996 The antievolutionary component of antimutagenesis and anticarcinogenesis: where do mutation rates come from and where are they going? Mutation Res. **350:** 5–8.

Dycaico, M. J., G. S. Provost, P. L. Kretz, S. L. Ransom, J. C. Moores *et al.*, 1994 The use of shuttle vectors for mutation analysis in transgenic mice and rats. Mutation Res. **307:** 461–478.

Ehrlich, M., X.-Y. Zhang and N. M. Inamdar, 1990 Spontaneous deamination of cytosine and 5-methylcytosine in DNA and replacements of 5-methylcytosine residues with cytosine residues. Mutat. Res. **238:** 277–286.

El Antri, S., O. Mauffret and M. Monnot, 1993 Structural deviations of CpG provide a plausible explanation for the high frequency of mutations at this site. Phosphorus nuclear magnetic resonance and circular dichroism studies. J. Mol. Biol. **230:** 373–378.

Farabaugh, P. J., U. Schmeissner, M. Hofer and J. H. Miller, 1978 Genetic studies of the lac repressor. VII. On the molecular nature of spontaneous hotspots in the *lacI* gene of Escherichia coli. J. Mol. Biol. **126:** 847–857.

Foster, P. L., E. Eisenstadt and J. H. Miller, 1983 Base substitution mutations induced by metabolically activated aflatoxin B1. Proc. Natl. Acad. Sci. USA **80:** 2695–2698.

Gilbert, W., and B. Muller-Hill, 1970 The lactose repressor, in *The Lactose Operon*, edited by J. R. Beckwith and D. Zipser. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Glickman, B. W., 1990 Study of mutational specificity in the *lacI* gene of *Escherichia coli* as a window on the mechanisms of mutation. Environ. Mol. Mutagen **16:** 48–54.

Glickman, B. W., M. J. Horsfall, A. J. Gordon and P. A. Burns, 1987 Nearest neighbor affects G:C to A:T transitions induced by alkylating agents. Environ. Health Perspect. **76:** 29–32.

Gordon, A. J., and J. A. Halliday, 1995 Transgenic system for in vivo mutational analysis. Letter to the editor. Mutation Res. **306:** 103–105.

Gordon, A. J., J. A. Halliday, M. J. Horsfall and B. W. Glickman, 1991 Spontaneous and 9-aminoacridine-induced frameshift mutagenesis: second-site frameshift mutation within the N-terminal region of the *lacI* gene of *Escherichia coli.* Mol. Gen. Genet. **227:** 160–164.

Gu, M., A. Ahmed, C. Wei, N. Gorelick and B. W. Glickman, 1994 Development of a lambda-based complementation assay for the preliminary localization of *lacI* mutants from the Big Blue mouse: implications for a DNA-sequencing strategy. Mutat. Res. **307:** 533–540.

Halliday, J. A., and B. W. Glickman, 1991 Mechanisms of spontaneous mutation in DNA repair-proficient *Escherichia coli.* Mutat. Res. **250:** 55–71.

Jego, N., G. Thomas and R. Hamelin, 1993 Short direct repeats flanking deletions, and duplicating insertions in p53 gene in human cancers. Oncogene **8:** 209–213.

Ketterling, R. P., E. Vielhaber and S. S. Sommer, 1994 The rates of G:C to T:A and G:C to C:G transversions at CpG dinucleotides in the human factor IX gene. Am. J. Hum. Genet. **54:** 831–835.

Kleina, L. G., and J. H. Miller, 1990 Genetic studies of the Lac repressor. XIII. Extensive amino acid replacements generated by the use of natural and synthetic nonsense suppressors. J. Mol. Biol. **212:** 295–318.

Knoll, A., D. P. Jacobson, P. L. Kretz, K. Lundberg, J. M. Short *et al.*, 1994 Spontaneous mutations in *lacI*-containing lambda lysogens derived from transgenic mice: the observed patterns differ in liver and spleen. Mutat. Res. **311:** 57–67.

Knoll, A., D. P. Jacobson, H. Nishino, P. L. Kretz, J. M. Short *et al.*, 1996 A selectable system for mutation detection in the Big Blue *lacI* transgenic mouse system: what happens to the mutational spectra over time. Mutat. Res. **352:** 9–22.

Kohler, S. W., G. S. Provost, A. Fieck, P. L. Kretz, W. O. Bullock *et al.*, 1991a Analysis of spontaneous and induced mutations in transgenic mice using a lambda ZAP/*lacI* shuttle vector. Environ. Mol. Mutagen. **18:** 316–321.

Kohler, S. W., G. S. Provost, A. Fieck, P. L. Kretz, W. O. Bullock *et al.*, 1991b Spectra of spontaneous and mutagen-induced mutations in the *lacI* gene in transgenic mice. Proc. Natl. Acad. Sci. USA **88:** 7958–7962.

Krawczak, M., and D. N. Cooper, 1991 Gene deletions causing human genetic disease: mechanisms of mutagenesis and the role of the local DNA sequence environment. Hum. Genet. **86:** 425–441.

Lambert, I. B., T. A. Chin, D. W. Bryant, A. J. E. Gordon, B. W. Glickman *et al.*, 1991 The mutational specificity of 2-(2-furyl)-3-(5-nitro-2-furyl)-acrylamide (AF2) in the *lacI* gene of *Escherichia coli.* Carcinogenesis **12:** 29–34.

LeClerc, J. E., J. R. Christensen, P. V. Tata, R. B. Christensen and C. W. Lawrence, 1988 Ultraviolet light induces different spectra of *lacI* sequence changes in vegetative and conjugating cells of *Escherichia coli.* J. Mol. Biol. **203:** 619–633.

Lefebvre, A., O. Mauffret, B. Hartmann, E. Lescot and S. Fermandjian, 1995 Structural behaviour of the CpG step in two related oligonucleotides reflects its malleability in solution. Biochemistry **34:** 12019–12028.

Lewis, M., G. Chang, N. C. Horton, M. A. Kercher, H. C. Pace *et al.*, 1996 Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. Science **271:** 1247–1254.

Lichtenauer-Kaligis, E. G. R., J. Thijssen, H. Den Dulk, P. Van De Putte, J. G. Tasseron-De Jong *et al.*, 1996 Comparison of spontaneous *hrpt* mutation spectra at the nucleotide level in the endogenous *hrpt* gene and five other genomic positions. Mutat. Res. **351:** 147–155.

Lutz, W. K., 1990 Endogenous genotoxic agents and processes as a basis of spontaneous carcinogenesis. Mutat. Res. **238:** 287–295.

Mauffret, O., B. Rene, O. Convert, M. Monnot, E. Lescot *et al.*, 1991 Drug-DNA interactions: spectroscopic and footprinting studies of site and sequence specificity of elliptinium. Biopolymers **31:** 1325–1341.

Mazur, M., and B. W. Glickman, 1988 Sequence specificity of mutations induced by benzo[a]pyrene-7,8-diol-9,10-epoxide at endogenous *aprt* gene in CHO cells. Somatic Cell Mol. Genet. **14:** 393–400.

Miller, J. H., 1972 *Experiments in Molecular Genetics.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Miller, J. H., 1982 Carcinogens induce targeted mutations in *Escherichia coli.* Cell **31:** 5–7.

Miller, J. H., 1985 Mutagenic specificity of ultraviolet light. J. Mol. Biol. **182:** 45–65.

Miller, J. H., and K. B. Low, 1984 Specificity of mutagenesis resulting from the induction of the SOS system in the absence of mutagenic treatment. Cell **37:** 675–682.

Miller, J. H., D. Ganem, P. Lu and A. Schmitz, 1977 Genetic studies of the Lac repressor I. Correlation of mutational sites with specific amino acid residues: construction of a colinear gene-protein map. J. Mol. Biol. **109:** 275–301.

Muller-Hill, B., L. Crapo and W. Gilbert, 1968 Mutants that make more lac repressor. Proc. Natl. Acad. Sci. USA **59:** 1259–1264.

Provost, G. S., and J. M. Short, 1994 Characterization of mutations induced by ethylnitrosourea in seminiferous tubule germ cells of transgenic B6C3F1 mice. Proc. Natl. Acad. Sci. USA **91:** 6564–6568.

Rogers, B. J., G. S. Provost, R. R. Young, D. L. Putman and J. M. Short, 1995 Intralaboratory optimization and standardization of mutant screening conditions used for a lambda/*lacI* transgenic mouse mutagenesis assay (I). Mutat. Res. **327:** 57–66.

Sauer, R. T., 1996 Lac repressor at last. Structure **4:** 219–222.

Sayle, R. A., and E. J. Milner-White, 1995 RASMOL: biomolecular graphics for all. Trends Biochem. Sci. **20:** 374.

Schaaper, R. M., B. N. Danforth and B. W. Glickman, 1985 Rapid repeated cloning of mutant *lac* repressor genes. Gene **39:** 181–189.

Schaaper, R. M., B. N. Danforth and B. W. Glickman, 1986 Mechanisms of spontaneous mutagenesis: an analysis of the spectrum of spontaneous mutation in the *Escherichia coli lacI* gene. J. Mol. Biol. **189:** 273–284.

Schaaper, R. M., R. L. Dunn and B. W. Glickman, 1987 Mechanisms of UV-induced mutation: mutational spectra in the *Escherichia coli lacI* gene for a wild type and excision deficient strain. J. Mol. Biol. **198:** 187–202.

Schaaper, R. M., N. Koffel-Schwartz and R. P. Fuchs, 1990 N-acetoxy-N-acetyl-2-aminofluorene-induced mutagenesis in the *lacI* gene of *Escherichia coli.* Carcinogenesis **11:** 1087–1095.

Schildkraut, C. L., J. Marmur and P. Doty, 1962 J. Mol. Biol. **4:** 430–443.

Skopek, T. R., K. L. Klein and D. R. Marino, 1995 Relative sensitivity of the endogenous *hrpt* gene and *lacI* transgene in ENU-treated Big Blue B6C3F1 mice. Environ. Mol. Mutagen. **26:** 9–15.

Skopek, T. R., D. R. Marino, K. L. Kort, J. Miller and T. Pippert, 1996 Synthesis of a *lacI* gene analogue with reduced CpG content. Mutat. Res. **349:** 163–172.

Streisinger, G., Y. Okada, J. Emrich, J. Newton, A. Tsugita *et al.*, 1966 Frameshift mutations and the genetic code. *Cold Spring Harbor Symp. Quant. Biol.* **31:** 77–84.

Wyborski, D. L., 1997 Big Blue rodent cell cultures for mutagenesis studies. Strategies **10:** 7.

Wyborski, D. L., S. Malkhosyan, J. Moores, M. Perucho and J. Short, 1995 Development of a rat cell line containing stably integrated copies of a lambda/*lacI* shuttle vector. Mutat. Res. **334:** 161–166.

Yatagai, F., and B. W. Glickman, 1990 Specificity of spontaneous mutation in the *lacI* gene cloned into bacteriophage M13. Mutat. Res. **243:** 21–28.

Young, R. R., B. J. Rogers, G. S. Provost, J. M. Short and D. L. Putman, 1995 Interlaboratory comparison: liver spontaneous mutant frequency from lambda/*lacI* transgenic mice (Big Blue) (II). Mutat. Res. **327:** 67–73.