Characterization of *in vivo* Somatic Mutations at the Hypoxanthine Phosphoribosyltransferase Gene of a Human Control Population

Karolyn Burkhart-Schultz,[1] Cynthia B. Thomas,[1] Claudia L. Thompson,[2] Cheryl L. Strout,[2] Eleanor Brinson,[1] and Irene M. Jones[1]

[1]Biology and Biotechnology Research Program, Lawrence Livermore National Laboratory, Livermore, CA 94551 USA; [2]Division of Biometry and Risk Assessment, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709 USA

The ability to recognize a change in mutation spectrum after an exposure to a toxic substance and then relate that exposure to health risk depends on the knowledge of mutations that occur in the absence of exposure. Toward this end, we have been studying both the frequency and molecular nature of mutations of the hypoxanthine phosphoribosyltransferase (*hprt*) gene in peripheral blood lymphocytes as surrogate reporters of genetic damage. We have analyzed mutants, one per donor to ensure independence, from a control population in which the quantitative effects of smoking and age on mutant frequency have been well defined. Analyses of cDNA and genomic DNA by polymerase chain reaction and sequencing have identified the mutations in 63 mutants, 45 from males and 18 from females, of which 34 were smokers and 29 were nonsmokers. Slightly less than half of the mutations were base substitutions (28); they were predominantly at GC base pairs (19). Different mutations at the same site indicated that there are features of the *hprt* polypeptide that affect the mutation spectrum. Two pairs of identical mutations indicated that there may also be hot spots. Mutations not previously reported have been detected, indicating that the mutation spectrum is only partly defined. The remainder of the mutations were deletions (32) or insertions/duplications (3); deletions ranged from one base pair to complete loss of the locus. Despite a small average increase in mutant frequency for smokers, an increased proportion of base substitutions at AT base pairs in smokers ($p = 0.2$) hinted at a smoking-associated shift in the mutation spectrum. Expansion of the study to include individuals with larger, smoking-associated increases of mutant frequency will determine the significance of this observation. This background mutation study provides insight into factors that determine the mutation spectra of the *hprt* locus and provides data for comparison with mutation spectra of other populations.

Key words: *hprt*, lymphocytes, mutation spectrum, somatic mutation. *Environ Health Perspect 101(1):68–74*

There is considerable evidence that cancer is a disease produced by interactions between individual genetic susceptibility, environmental exposure, and target-tissue biology. The mutations in many tumors have been identified in a proto-oncogene or tumor-suppressor gene per se or at other sites in the genome that alter their expression, e.g., translocations affecting proto-oncogenes. In some cases, such as p53 mutations of lung tumors, the mutation spectrum is consistent with an exposure-dependent etiology. In others such a pattern is not evident. Interpreting the mutation spectrum in genes causally related to cancer requires understanding the role of mutation spectra in tumor progression as well as the basis of tissue differences in metabolism, DNA repair, and routes of exposure. In addition, individual susceptibility factors may affect all tissues of a person. Recent reviews provide excellent overviews of these highly intertwined issues in molecular cancer epidemiology (*1,2*).

In the past decade, a number of surrogate biomarkers have been developed that may be useful in epidemiologic studies to ascertain the contributions of environment and genetics to the development of cancer. These surrogates have the advantage that they are not part of the tumor progression process, and they have the potential to indicate various facets of exposure and individual susceptibility. Biomarkers such as sister chromatid exchange (SCE) and DNA or protein adducts help measure exposure. Assays for chromosome translocations or micronucleus formation each monitor specific types of genetic damage throughout the genome. Several assays for specific gene somatic mutation detect the biological consequences of exposure at the gene level: the well-established assays for mutation in lymphocytes at the hypoxanthine phosphoribosyltransferase (*hprt*; *3,4*) and HLA-A (*5,6*) genes; in erythrocytes for the glycophorin A gene (*7*); and the newer, less widely used lymphocyte mutation assay for T-cell receptor genes (*8,9*).

The lymphocyte forward-mutation studies are unique in leading to both the frequency and the molecular basis of somatic mutations. Each individual's mutant frequency integrates his or her exposure and susceptibility factors and provides a measurement based on the genetic locus and cells involved. At the molecular level, the *hprt* and HLA mutation assays have different biases: the X-linked *hprt* locus reports primarily nonrecombinigenic mechanisms of gene inactivation such as base substitutions, frameshifts, deletions, and rearrangements (*10*), whereas the HLA assay also reports recombination-based mechanisms of mutation (~30%) (*11*). Shifts in the *hprt* mutation spectrum have been reported after radiotherapy; the incidence of large deletion and rearrangement mutations increased substantially, and an altered mutation spectrum persisted in some individuals for several years after exposure (*12,13*). A similar shift in *hprt* mutation spectrum after radiation exposure has been seen *in vitro* (*14,15*).

Molecular definition of mutations responsible for loss of *hprt* function is possible due to recent technical developments. The DNA sequence for the total genetic locus is known (*16*). Polymerase chain reaction (PCR) methods based on these sequences have been developed for both genomic DNA (*17*) and mRNA (*18–20*). These methods have been used to determine mutations in the inherited Lesch Nyhan disease (*17,19,21,22*), in cells treated in vitro with mutagens (*23,24*), and in lymphocytes recovered from normal individuals (*25–28*).

The spectrum of the mutations recovered at the *hprt* gene of control populations has the potential to reveal mechanisms of background mutation that lead to gene inactivation independent of exposure to exogenous agents and independent of cancer progression per se. Such events have been implicated in the genetic divergence that occurs in precancerous proliferative cells. The background mutation spectrum will also provide a reference for studies of exposed populations.

Address correspondence to I. M. Jones, Lawrence Livermore National Laboratory, L-452, PO Box 808, Livermore, CA 94551 USA.

We thank R.J. Albertini and P. O'Neill, University of Vermont, for sharing methods and growth factor in the development phases of this work and for providing the TK-6 cells; R. Gibbs, Baylor College of Medicine, for the *hprt* multiplex primers; A. Raynor and D. Farmer, University of California at San Francisco, for providing the LAK supernatant; K. Kelsey, Harvard Medical School, for the cell line 91-XC-4; J. Fuscoe, EHRT, Inc., for analyzing selected mutants for recombinase-mediated deletion; and G. Lucier, K. Tindall, W. Piegorsch, J. Taylor, and D. Bell, NIEHS, and T. Skopek, N. Cariello, and V. Walker, University of North Carolina at Chapel Hill, for discussions and comments on this work. This work was performed under the auspices of the U.S. Department of Energy by LLNL under contract W-7405-ENG-48 with support from Interagency Agreement Y01-ES-80171 from NIEHS.

Definition of the background mutation spectrum for the human *hprt* gene of lymphocytes requires analysis of a large number of independent mutations. The *hprt* gene is a broad mutation target, one that can be inactivated by a large variety of mutations. The mutation spectrum is not yet saturated; new mutations are still being found despite the collection of more than 1000 mutations at human *hprt* genes in *in vivo* and *in vitro* studies (*29*). Population-based studies are required to define the background *in vivo* mutation spectrum and the effects of genetic susceptibility and exposure. In this report we begin to define a population-based mutation spectrum of the *hprt* gene using peripheral blood lymphocytes as an indicator for background mutations.

Materials and Methods

The recruitment of subjects and methods of obtaining samples were reviewed by Institutional Review Boards at both Lawrence Livermore National Laboratory and the National Institute of Environ-mental Health Sciences, and all subjects gave informed consent before participating in this study. The selection of subjects was based on self-reporting of no recent illness; no exposure to radiation or chemicals other than smoking or over-the-counter medications; no prior history of heart disease, diabetes, tuberculosis, high blood pressure, cancer, or hepatitis; no passive cigarette/cigar/marijuana smoke exposure at home; and age between 19 and 45 years. All nonsmokers self-reported as never having smoked regularly. This population has been described in detail (*30*).

Isolation of Thioguanine-Resistant Mutants

The frequency of thioguanine-resistant mutants was determined by limiting dilution methods (*30*). To isolate mutants, we cultured cells for up to 40 hr at 1×10^6 cells/ml with the mitogen phytohemagglutinin (PHA; 1 µg/ml; HA17; Burroughs Wellcome), then counted and plated them in round-bottomed microtiter wells with 10% (v/v) lymphokine-activated killer cell (LAK) supernatant (containing 15,000 U/ml human interleukin-2, serum-free medium, and any factors produced during the 3–4-day activation of peripheral blood monocytes), reduced PHA (0.1 µg/ml), irradiated lymphoblastoid feeder cells (5,000–20,000 viable irradiated-TK-6/well; 5 Gy Cesium-137 delivered at 4.2 Gy/min), and ß-mercaptoethanol (50 mM), with or without thioguanine (1 µg/ml) selection (*30–32*). Throughout this work the basal medium was RPMI 1640 supplemented with 5% fetal bovine serum, 20% HL-1 (Ventrex Laboratories Inc.), penicillin (100 U/ml), and streptomycin (100 mg/ml). All cultures were incubated at 37°C at 5% CO_2. Individual thioguanine-resistant clones were expanded to $5–40 \times 10^6$ cells in LAK supernatant-supplemented medium with PHA at 0.1 µg/ml, by transfer to successively larger numbers of wells, or larger wells, in the presence of lethally irradiated lymphoblastoid cells (either TK6 or a derivative thereof, 91-XC-4; that is, missing the X-linked *hprt* sequences). We refreshed thioguanine (1 µg/ml) once at the first stage of cell expansion after plating in microtiter wells. We froze expanded clones in 8% dimethyl sulfoxide, 20% fetal bovine serum, and 72% RPMI using a controlled freezing program (Cryomed) and stored the clones in liquid nitrogen. A single mutant from each of 63 human subjects was studied, hence the mutations reported are unambiguously independent events.

Preparation of RNA and DNA

We prepared cytoplasmic extracts from aliquots of frozen cells: $5–10 \times 10^6$ cells were thawed, rinsed once in RPMI 1640 medium with 15% fetal bovine serum, then washed twice with cold phosphate-buffered saline (PBS). We suspended the cells in 250 µl cold lysis buffer (10 mM Tris-HCl pH 7.8, 150 mM NaCl) containing 10 µl RNAse Block II RNAse inhibitor (Stratagene) per 250 µl buffer. We added 2% Nonidet P-40 (25 µl/250 µl lysis suspension), vortexed the sample, and held it on ice for 5 min to lyse cells. The nuclei were pelleted by centrifugation at 4°C. We transferred the supernatant containing cellular RNA to clean tubes and added 250 µl extraction buffer (40 mM Tris-HCl, pH 7.8, 40 mM EDTA, 0.7 M NaCl, and 2% sodium dodecyl sulfate; SDS. The suspension was extracted 5 times with phenol:chloroform:isoamyl alcohol (25:24:1), then once with chloroform:isoamyl alcohol (24:1). We precipitated RNA with two volumes of cold 100% ethanol. Samples were stored at -20°C. RNA was recovered by centrifugation and washed twice with a 3:1 solution of ethanol: 0.1 M sodium acetate, pH 5.2. RNA pellets were dried and resuspended in diethylpyrocarbonate-treated deionized water. We determined RNA concentrations by optical density readings. RNase Block II was added, and the RNA samples were stored at -20°C or -80°C.

We suspended the nuclear pellets in 750 µl lysis buffer (10 mM Tris pH 8, 0.4 M NaCl, and 2 mM EDTA), 125 ml proteinase K digestion solution [1% sodium dodecyl sulfate (SDS), 2 mM EDTA, and 250 mg Proteinase K], and 50 ml of 10% SDS. The pellets were incubated at 60–65°C for 24–72 hr. We removed protein and SDS using salt extraction by the addition of 750 µl deionized water and 750 ml of saturated (6 M) NaCl. After vigorous mixing for 30 sec, the precipitated SDS and proteins were removed by two successive centrifugations. The DNA was recovered by ethanol precipitation with 2 volumes of 100% ethanol at 4°C. We rinsed precipitated DNA twice with cold 70% ethanol. Pellets were air dried and resuspended in TE (10mM Tris, pH 8.0, and 0.1 mM EDTA). We determined DNA concentrations by OD readings.

Synthesis and Analysis of cDNA

We synthesized cDNA by a modification of the Superscript RNAse H-reverse transcriptase reaction (RT Rx) procedure (Gibco BRL) using 2 mg of cytoplasmic RNA/20 µl reaction. RNA, 0.2 µg oligo d(T)18, and 0.5 ml RNAse Block II were mixed in a total volume of 10.5 µl, heated for 2 min at 95°C, held at 45°C while adding 7.5 ml reaction mix [4 µl 5X Superscript reaction buffer, 2 µl of 0.1M dithiothreitol (DTT), 1.0 µl of dNTP mix containing 25 µM each of dNTP, and 0.5 µl RNAse Block II] and then adding 2.0 µl of 200 U/ml RNase H-Superscript reverse transcriptase. We incubated samples at 45°C for 1 hr, then heated them to 90–95°C for 10 min before incubating them 20 min at 37°C with 2 U RNAse H (Gibco BRL) and storing them at -20°C.

We carried out second-strand cDNA synthesis and cDNA amplification by the Taq DNA polymerase (AmpliTaq; Perkin Elmer Cetus) using *hprt*-specific primers (see Table 1) in two successive PCRs using a Perkin-Elmer DNA Thermal Cycler. In the first PCR amplification, we used three pairs of primers in separate reactions: 1) 250 nM each of IRJ040 and IRJ041 amplified a partial adenosine phosphoribosyltransferase (APRT) cDNA, as a positive control for RNA quality; 2) 62 nM each of BRIN006 and IRJ048 amplified a 620 bp *hprt* cDNA, approximately two-thirds of the *hprt* cDNA from bp 140 through the 3′ end; 3) 62 nM each of IRJ050 and IRJ048 amplified the full-length *hprt* cDNA. Reactions (50 mM KCl, 10 mM Tris-HCl pH 8.3, 1.5 mM MgCl2, 150 µM each of dNTP, 2 µl of the reverse transcrptase and one pair of primers) were covered with mineral oil, heated at 95°C for 5 min, and held at 85°C during the addition and mixing of 1.25 U of Taq DNA polymerase into the reaction (total volume of 50 µl). The 35 cycles of cDNA amplification consisted of 1 cycle of 3-min denaturation at 93°C, 2-min primer annealing at 55°C, and 3-min elongation at 72°C, followed by 24 cycles of 2-min denaturation at 93°C, 1-min primer annealing at 55°C, and 3-min elongation at 72°C, and then

10 cycles having the elongation step altered by the addition of 18 sec per round. We visualized PCR products in 10 µl of the reaction on an agarose gel (2.5% Nusieve plus 1% Seakem agarose in 1× 90 mM Tris-borate and 2.0 mM EDTA, pH 8.0), stained with ethidium bromide.

In the second PCR amplification, 1/50 of the first PCR for full-length cDNA was the template, and primers were nested with respect to the first PCR primers in a 100 µl reaction (50 mM KCl, 10 mM Tris-HCl pH 8.3, 1.5 mM MgCl2, 200 µM each of dNTP, 0.25 µM each of IRJ046 and IRJ049, and 2 U of Taq DNA polymerase). We based the number (*N*) of amplification cycles on the yield of the first PCR: $N = 35$ cycles if a full-length cDNA product was not visible; $N = 20–25$ cycles if it was visible. Cycle 1 consisted of 4-min denaturation at 93°C, 2-min primer annealing at 55°C, and 3-min elongation at 72°C, followed by (*N*-11) cycles of 2-min denaturation at 93°C, 1-min primer annealing at 55°C, and 3-min elongation at 72°C and the last 10 cycles with the elongation step altered as above. Products were evaluated as above.

We gel purified the *hprt* cDNA PCR products and recovered them from gel slices using GeneClean (BIO 101, Inc.). Comparison of the cDNA template to DNA size markers on agarose gels stained with ethidium bromide was used to quantify the cDNA.

Analysis of Genomic DNA

Individual exons and flanking intron sequences were simultaneously amplified from genomic DNA (gDNA) by multiplex PCR (mPCR) (*17*). We analyzed products on 1.4% Seakem agarose gels stained with ethidium bromide. Individual exon bands were excised from the gel, minced in sterile deionized water and refrigerated overnight. We used aliquots of the eluted cDNA as templates to amplify exon-specific fragments and introduce the M13-21 "universal" priming sequence for sequencing the product (*17*). The PCR products were ethanol precipitated and quantified on agarose gels before sequencing. In the case of one mutant with no product in the mPCR, template integrity and amplifiability were confirmed using non-*hprt* primers.

Sequencing

We used four fluorescent dye-labeled primers for sequencing *hprt* cDNA (Table 1). We used IRJ 051 and IRJ 054 first; IRJ 052 and IRJ 053 were used when necessary to clarify results. For gDNA sequencing, we used the fluorescent dye-labeled universal primer -21M13 from Applied Biosystems Inc. For dideoxy sequencing reactions, we used a modification of the ABI Taq polymerase-based cycle-sequencing protocol. Individual base reactions were performed with larger volumes: 14 µl for A and C, each with 16.8 ng template, and 28 µl for T and G, each with 33.6 ng template. Dideoxy/deoxynucleotide mixtures did not use modified bases. Cycling times and temperatures were 10 cycles with denaturation at 93°C for 2 min, annealing at 55°C for 1 min, and elongation at 72°C for 2 min, followed by 15 cycles with denaturation at 93°C for 30 sec and elongation at 70°C for 1 min. We analyzed sequencing reaction products of cDNA and gDNA templates using the ABI 373A DNA Sequencer. Sequence data were compared to wild-type sequence using the SeqEd 675: DNA Sequence Editor (Applied Biosystems Inc.).

Results

We cloned mutant lymphocytes by limiting dilution from 63 donors, 34 of whom were smokers (26 male smokers, 8 female smokers) and 29 of whom were nonsmokers (19 male nonsmokers, 10 female nonsmokers). The results of analysis of cDNA and gDNA are summarized in Table 2 and detailed in Tables 3–7. The results (Tables 2–7) are the conclusions of a series of analyses. The strategy was to first analyze the cDNA of a mutant by sequencing it. If the cDNA analysis was negative due to absence of cDNA or inconclusive results, gDNA was studied.

The amount of cDNA synthesized in the first *hprt* PCR reaction varied between mutants, probably due to both experimental sources and mutation-induced change in copy number of *hprt* RNA (mRNA). Several controls for RNA and PCR conditions were used. Synthesis of a short 3′ segment of *aprt* cDNA served as a control for the integrity of the RNA and the reverse transcriptase reaction. Due to the higher efficiency of 3′ end cDNA synthesis, a partial 3′ *hprt* cDNA (two-thirds of full-length cDNA) was synthesized to help judge the presence or absence of *hprt* cDNA. This reaction always gave a stronger signal than the full-length product. The cDNA of some mutants was not visible as a full-length product until the second, nested reaction. The use of the full-length and two-thirds 5′-end primers helped on occasions when a mutant was deleted for the site of one of the 5′ primers; for example, the full-length cDNA was not made despite the presence of the "two-thirds" product for SM16.1 4B2 (Tables 6 and 7), and the full-length cDNA was present in the absence of the two-thirds product of the exon 2-3 deletion mutations such as of NM24.1 56A6 and NM30.1 17B (Tables 6 and 7).

In some cases cDNA sequencing identified a point mutation, but in others the cDNA sequencing results were not conclusive, and analysis of gDNA was required. Simple exon skips (a clean loss of one or multiple exons from the cDNA) or absence of exon sequences adjacent to introns were taken as evidence of incorrect splicing of the gene transcript. The mutation respon-

Table 2. Summary of independent *hprt* mutations in lymphocytes of a control population

Mutation type	Number
Base substitutions	
Transitions	9
Transversions	18
Other: TGG to A	1
Total	28
Deletions/insertions	
±1 bp	5
±2 bp	3
-3 bp	2
Small deletions[a]	14
Large deletions[b]	10
+30 bp duplication	1
Total	35
Total mutations	63

[a]Deletion of 6–200 bp.
[b]Deletion spanning one or more exons.

Table 1. Primers used for cDNA synthesis and sequencing

Application	Product size	Primer ID	Primer sequence
aprt cDNA, 3' end	230 bp	IRJ 040	5'-ACT.ACA.TCG.CAG.GCC.TAG.ACT-3'
		IRJ 041	5'-TTC.ATG.GTT.CCA.CCA.GTG.GCC-3'
hprt cDNA	877 bp	IRJ 050[a]	5'-GCG.AAC.CTC.TCG.GCT.TT-3'
		IRJ 048[b]	5'-AAG.CTC.TAC.TAA.GCA.GAT.GGC.CAC.AGA.ACT-3'
2/3 *hprt* cDNA[c]	620 bp	BRIN 006	5'-ACG.TCT.TGC.TCG.AGA.TGT.GAT.G-3'
		IRJ 048	5'-ACG.TCT.TGC.TCG.AGA.TGT.GAT.G-3'
nested *hprt* cDNA	789 bp	IRJ 046[d]	5'-CTC.TGC.TCC.GCC.ACC.GGC.TTC.CT-3'
		IRJ 049[e]	5'-AAC.ATT.GAT.AAT.TTT.ACT.GGC.GAT-3'
hprt cDNA sequencing	—	IRJ 051	5'-TGT.CAA.TAG.GAC.TCC.AGA.TGT.TTC-3'
	—	IRJ 052	5'-TGT.AAG.GAT.TAT.ACT.GCC.TGA.CC-3'
	—	IRJ 053	5'-TGT.AAT.TGA.GCA.CAC.AGA.GGG.CT-3'
	—	IRJ 054	5'-TGT.AAA.GTT.GAG.AGA.TCA.TCT.CCA.C-3'

[a]IRJ050, primer HT5', shortened by 1 bp (*21*).
[b]IRJ048, primer HT3', shortened by 3 bp (*21*).
[c]Primers amplify a 3' end fragment approximately 2/3 the length of the full cDNA.
[d]IRJ046, primer 581 (*19*).
[e]IRJ049, similar to primer PCR-2 (*20*).

sible for the missplicing phenotype was sought by analysis of the indicated region of gDNA. Multiplex PCR (*17*) revealed deletions in exon-specific regions (as in NM24.1 56A6). In some cases no gross alteration of the gDNA was evident, and specific exon/intron regions of gDNA were sequenced (*17*) to identify the mutation (as in NM23.1 65B/C12 and SM6.1 24C9).

The base substitution mutations described in detail in Tables 3 and 4 indicate that *hprt* reports all possible types of base substitution mutations. The relative proportions of various base substitutions reflect many factors, as discussed below.

The deletion mutations also display a full range of mutation mechanisms as described in Tables 5–7. Deletion sizes stated in Table 6 are precise if identified by sequencing cDNA (within exons) or estimated by electrophoretic mobility when identified by multiplex PCR. The ability to detect deletions by mPCR, as was done when no cDNA was made or the cDNA had exon skips, depends on the primers currently used and the associated PCR fragment sizes (*17*). Changes in molecular weight of a few percent can be seen in any mPCR product, but the variation in size of products means that sensitivity varies by exon.

Discussion

The results presented demonstrate three major points: the diversity of mechanisms of gene inactivation mutation detectable with the *hprt* locus of human lymphocytes; the fact that the sequence of a gene target affects its mutation spectrum; and the challenge of distinguishing between the depen-

Table 3. Summary of base substitution mutations

	Smokers	Nonsmokers	Total
Transitions[a]	6	3	9
GC to AT	4	3	7
AT to GC	2	0	2
Transversions	8	10	18
GC to TA	1	2	3
GC to CG	3	6	9
AT to TA	2	2	4
AT to CG	2	0	2
Other: TGG to A		1	1
Number of donors	14	14	28
Average	10×10^{-6}	6×10^{-6}	

[a]The base listed is in the nontranscribed strand of the DNA.

Table 5. Distribution of deletion, insertion, and duplication mutations among types of donors

Type of mutation	SM	SF	NM	NF	Total
Deletion of 1–3 bp	2	3	2	1	8
Deletion of 6–200 bp[a]	9	1	3	1	14
Deletion of ? kilobase[b]	4[c]	0	5	1[c]	10
Insertion of 1–2 bp	1	0	1	0	2
Duplication	0	0	1	0	1
Total	16	4	12	3	35

S, smoker; N, nonsmoker, M, male, F, female.
[a]Includes intraexonic deletions identified by sequencing cDNA and deletions of exonic regions detected by gel analysis of multiplex poolymerase chain reaction products.
[b]Loss of one or more exonic region in multiplex polymerase chain reaction products.
[c]One mutation in which only one exon is lost.

Table 4. Base substitution mutations in human lymphocyte *hprt* mutants

Donor	Mutant	Mf, $\times 10^6$	Exon 2	Exon 3	Exon 4	Exon 5	Exon 6	Exon 7	Exon 8	Exon 9	Amino acid change/comments
SM23.1	25H11	7.1	**C>G 74**[a]								**pro > arg**
SM54.1	77B5	16.4	**C>G 74**[a]								**pro > arg**
SM25.1	49A1	6.9	**C>T 74**								**pro > leu**
NF27.1	5E2	1.9	TGG >A 117-119[c]								Term next codon
SF15.1	43A8	16.2		T>G 146							leu > arg
SF9.1	33A10	7.0		**G>T 197**							**cys > phe**
NM28.1	24E7	3.8		**G>A 197**							**cys > tyr**
SM5.1	14B6	2.1		G>A 208							gly > arg
NF13.1	33G8	6.6		C>G 222[a]							phe > leu
SF6.1	14E1	7.5		T>C 233[c]							leu > pro
SM30.1	66D7	7.1			(G>A cIn4:1)[a,c]						Donor site (ex 4 skip)
NM19.1	34F5	4.6				G>A 400					glu > lys
SM40.1	59F3	15.9					A>T 421				lys > term
NM27.1	9H9	7.0					T>A 437[a]				leu > term
NM11.1	55G2	8.8					**T>A 449**[a]				**val > asp**
SM52.1	18F6	14.5					**T>G 449**[a]				**val > gly**
NM23.1	65B/C12	4.3					(G >C E6:-1)[c]				Acceptor site (ex 6 skip)
NF9.1	34H3	5.1						C>G 486			ser > arg
NF4.1	24F5	11.8						C>T 508			arg > term (at CpG)
NM32.1	47H11	5.2						G>T 529			asp > tyr
SM26.1	59B3	10.4							**C>G 551**		**pro > arg**
NF14.1	48D6	3.2							**C>G 551**		**pro > arg**
SM19.1	17B6	12.5							T>A 563		val > asp
NF15.1	27A7	7.4							G>C 574[a]		ala > pro
NM44.1	68B3	11.3							G>C 606[a]		leu > phe
SF21.1	35D8	7.4								**G>A 617**	**cys > tyr**
NF32.1	49B11	3.1								**G>T 617**	**cys > phe**
SM6.1	24C9	11.6								(A>G E9:2)[a,c]	Acceptor site (> ex 9 alternative splicing)

S, smoker; N, nonsmokers; M, male; F, female; Mf, mutant frequency. Bold type indicates base pairs where we have detected base substitutions, either the same base change or a different base change, in more than one donor.
[a]Not previously seen to the best of our knowledge.
[b]Complex change: 1 base replaces 3.
[c]Base substitution in an intron at an mRNA splice site.

Table 6. Deletion (del), insertion (ins), and duplication (dup) mutations under 200 bp

Donor	Mutant	Mf, $\times 10^6$	Change	Position[a]
SM20.1	26 E3	5.6	del <50 bp	cDNA did not include exon 1 or 12–20 bp upstream; deletion <50 bp in exon 1 region of genomic DNA[b]
SM32.1	37E 10	6.0	del <54 bp	cDNA deleted for bp –29 to 108, including all of exons 1+2 genomic DNA has a deletion <54 bp in the exon 1 region
SM16.1	4B2	8.1	del <54 bp	cDNA is synthesized but missing 5' primer sites, giving a negative PCR result for full-length cDNA; genomic DNA has a deletion of <54 bp of the exon 1 region
SM57.1	26H10	30.1	del >50 bp	cDNA did not include exon 1; genomic DNA exon 1 region is deleted >50 but <185 bp
SM2.1	13E 2	13.6	del 10 bp	cDNA deleted for bp 41–50, exon 2; deletion in exon 2 region in genomic DNA
NF28.1	57E 1	3.9	del 53 bp	cDNA deleted for bp 42–94, exon 2; deletion in exon 2 region of one copy of genomic DNA
SF18.1	16B10	21.3	del 2 bp	cDNA deleted for bp 56–57, exon 2
SM55.1	6C10	15.2	del 6 bp	cDNA deleted for bp 103-108, exon 2
SM21.1	78E 2	4.8	ins 2bp	cDNA has +GG inserted after 108, exon 2
SM3.1	5C8	4.4	del 1 bp	cDNA deleted for bp 142, exon 3
SM29.1	46B4	6.5	del 1bp	cDNA deleted for bp 182, exon 3
NF8.1	15C8	10.1	del 2bp	cDNA deleted for bps 288–289 or 289–290, exon 3
SF4.2	5G12	5.0	del 3 bp	cDNA deleted for bps 301–303, exon 3
NM1.4	4E 8	7.1	del 1bp	cDNA deleted for bp 354, exon 4
SM58.1	35H12	19.0	del ~140 bp	cDNA skips exon 4; in genomic DNA exon 4 region is deleted ~140 bp
NM36.1	65 E8	6.8	del ~25 bp	cDNA skips exon 4; in genomic DNA exon 4 region has ~25 bp deleted
NM26.1	5E 3	3.0	del 1 bp	cDNA deleted for bp 440 or 441 or 442, exon 6
SM31.1	4F7	9.4	del 9 bp	cDNA deleted for bps 466–474, exon 6; genomic DNA has a small deletion in the exon 6 region
NM9.1	46C1	6.0	dup 30 bp	cDNA has tandom duplication of bps 408–437 or 407–436 within exon 6; in genomic DNA exon 6 is increased in size
SF29.1	49D11	7.8	del <167 bp	cDNA skips exon 8; genomic DNA has a deletion in the exon 7/8 region. Sequencing genomic DNA found a deletion in intron 7 ending at E8:-5[c]
NM40.1	98B3	9.2	del <50 bp	cDNA has deletion of 42 bp at the 5' end of exon 8; genomic DNA has a deletion in the exon 7/8 region
NM42.1	44C8	10.4	del <30 bp	genomic DNA has a deletion in the exon 7/8 region
NM3.1	34 E12	4.8	ins 1 bp	cDNA skips exon 8; Sequencing of genomic DNA for exon 9 region found +1 bp (AT) inserted at E9:-133 in intron 8
SF23.1	27A3	14.7	del 3 bp	cDNA deleted for bps 622–624, or 623–625, exon 9
SM33.1	57A9	12.4	del ~50 bp	cDNA has 7 bp inserted after bp 609 between exons 8 and 9; exon 9 region has a deletion in genomic DNA

S, smoker; N, nonsmoker; M, male; F, female; Mf, mutant frequency; PCR, polymerase chain reaction.
[a]Base 1 is the adenine of the protein synthesis initiation codon. Mutations are listed in order of exon affected.
[b]Exons and the flanking intron regions for each exon were amplified using sets of primers for multiplex PCR analysis (*17*).
[c]E8:-5 refers to the fifth base upstream of exon 8. Nomenclature from Rossi et al. (*27*).

dence of the mutation spectrum on the reporter system used, exposure parameters, and donor characteristics. Here we emphasize the principles demonstrated by our results rather than the numerical aspects of our results. Our goal is complete analysis of 200 independent mutations from this population. Due to the limits of analyses applied so far, detection of deletion mutations, especially those of males, may be preferentially ascertained relative to base substitutions. The emphasis on analysis of mutations in males represents the composition of the population we have selected for ease of molecular analysis. It has been possible to detect deletions in the *hprt* gene of females using mPCR when a shift in mPCR fragment size has occurred but not when a loss of exons on the active X chromosome has occurred. The heterozygosity of X-linked *hprt* sequences in mutants from females requires analyses other than those used here when the mutation is not defined in the cDNA. No gender dependence of mutant frequency response was detected in the larger population from which these donors were drawn (*30*); hence, there is no quantitative indication of a gender-specific mutation spectrum.

A wide variety of mutations that inactivate the *hprt* gene have been found in this population. All possible base substitutions were found. Among the base substitutions, mutations at GC base pairs predominate. This imbalance may be due to the multiplicity of events that affect GC base pairs, e.g., depurination of adducted guanines, deamination of cytosine, misincorporation at adducted guanines, and in part to the mutation-sensitive sequences in *hprt*, as discussed below. Short (one to two base pairs) insertions and deletions characteristic of polymerase errors have been noted in exons and, in one case, in an intron. Deletions have been found throughout the gene; included were events of varied size, though only one case of total gene deletion was seen. Sequence duplication was also seen only once.

The variety of mutations reflects the wealth of vulnerable features of the *hprt* gene and its protein product and provides insight into facets of the mutation spectrum that may be used in comparisons of populations. Interspecies comparisons of the amino acid sequence of *hprt* demonstrate considerable conservation, indicating areas most likely to be associated with functional domains or other critical features (*33,34*). It is striking that the TA to GC change at base 449 leading to a valine to glycine amino acid substitution resulted in an *hprt*-deficient phenotype. This conservative amino acid substitution indicates that this must be a critical residue in the protein; all missense mutations that occur at this residue should be recoverable as *hprt*-deficient mutations and should be in the *hprt* mutation spectrum. Similarly, proline residues seem to critically affect the the *hprt* polypeptide; five mutations that lead to missense replacements of proline were identified. The specific base substitutions that produced these changes included four cases of GC to CT transversion, leading to a proline to arginine mutation, and one case of a GC to AT transition, leading to a proline to leucine mutation. The proline-associated transversion mutations were responsible for four of the nine GC to CG transversions and largely defined one feature of the mutation spectrum in this population. The relative proportion of GC to AT to GC to CG mutations at these sites may discriminate between conditions that lead to transitions rather than transversions. The tetrameric state of the HPRT enzyme (*35*) also is a target of mutation. Required cross-linking by cysteines may be the reason that 4 of the 28 base substitutions (at bases 197 and 617) led to missense substitutions for cysteine. Both GC to TA and GC to AT mutations were recovered, and there is potential to look at the relative proportions of these in relation to exposure or other variables. Base pair 617 is a hot spot; mutations have been recovered in other *in vivo* human *hprt* studies, both as lymphocyte mutations (*25,26*) and as inherited *hprt* genes of Lesch Nyhan patients (*19*). In contrast,

Table 7. Mutations spanning exons and/or introns in genomic DNA

Donor	Mutant	Mf, $\times 10^6$	Position
NM24.1	56A6	5.8	cDNA skips exons 2-3; missing exons 2 + 3 in genomic DNA: is not a recombinase event[a]
NF33.1	59B4	6.9	cDNA skips exons 2-3; missing exons 2 + 3 in genomic DNA of one gene copy: is a recombinase mediated event[a]
NM30.1	17B2	4.6	cDNA skips exons 2-3; missing exons 2 + 3 in genomic DNA: is a recombinase-mediated event[a]
SM8.1	25G5	5.4	no cDNA made; missing exons 4-9 in genomic DNA
NM29.1	55A5	5.0	no cDNA made; missing exons 4-9 in genomic DNA
SM50.1	44H5	14.9	no cDNA made; missing exons 4-9 in genomic DNA.
NM37.1	77H8	3.7	cDNA skips exon 5; missing exon 5 in genomic DNA.
SM14.1	14G1	6.6	cDNA skips exons 5-6; missing exons 5-6 in genomic DNA
SM28.1	34A7	11.5	cDNA skips exon 6; missing exon 6 region in genomic DNA
NM31.1	24 E8	5.4	no cDNA made; missing all exons in the genomic DNA

S, smoker; N, nonsmoker; M, male; f, female; Mf, mutant frequency; PCR, polymerase chain reaction.
[a]Analysis of recombinase role in exon 2-3 deletions from Fuscoe et al. Fuscoe (*36*).

the absence in this study of missense mutations affecting exons 1 and 4 indicates that the HPRT enzyme is relatively tolerant of amino acid changes in these regions.

Introns are also important targets of mutation due to their role in splicing mRNA. Both deletion and base substitution mutations in the eight introns may lead to missplicing and *hprt* deficiency (*26–28*). The evolution of intron sequences is responsible for one class of deletion mutations. Introns 1 and 3 have sequences sufficiently like the consensus sequences for the recombinase involved in rearrangement of T-cell receptor genes and antibody genes that aberrant recombinase recombination produces large deletions encompassing exons 2 and 3 in a small percentage of mutants (*36,37*). The frequency of these recombinase-dependent events is not known to be affected by exposures to toxic agents and is believed to represent the persistence of mutant lymphocytes produced during early T-cell development (*38*). These exon 2-3 deletion events therefore are part of the background spectrum for human lymphocyte *hprt* mutation.

Mutations at CpG sequences do not dominate the *hprt* mutation spectrum. In this regard, *hprt* mutations are unlike mutations of p53 of some tumors (*39*), APC genes in colorectal tumors (*40*), and genes associated with human genetic diseases (*41,42*). The low incidence of CpG-associated mutations is a benefit to somatic mutation studies, as these mutations do not dilute other, potentially more informative, events. This low incidence suggests lack of methylation of control elements in *hprt* on the active X chromosome, as *hprt* is a continuously transcribed housekeeping gene (*43*). One mutation at a CpG was identified in this study, a GC to AT at base pair 508. It may be an event independent of the CpG sequence, as GC to AT was the most common class of transition. It could also be the product of mitotic recombination with a mutated *hprt* gene on the inactive X; the donor was a female.

The mutations reported here are each from a different human donor, hence they sample somatic mutations in a population. This approach, though burdensome, ensures that the mutations are biologically independent. This approach also provides the opportunity to relate mutation spectrum to mutant frequency, exposure history, and potentially, individual susceptibility genotype. Any set of mutations collected is a reflection of the population studied. Previous studies of *hprt* mutations in human lymphocytes have focused on a single or small number of donors, with the intent of determining the types of mutations present, not of relating these to population characteristics (*25–28*). Published reports have not always made clear which mutants came from the same donor. An *hprt* database (*29*) is being compiled to provide this information.

Mutant frequency provides an integrated measure of individual response to smoking and other exposure and endogenous factors. The population studied here is well defined (*30*). The age range was quite narrow, 19–45 years, averaging 32 years. The population was prescreened to ensure that smoking was the primary exposure variable. Factors that affect mutant frequency in this population included age, years of smoking, and the proportion of lymphocytes that form clones or cloning efficiency (*30*). The cloning efficiency may relate mutant frequency to cell proliferation, with more mutants forming when more cell proliferation occurs. If this hypothesis is valid, then proliferation-dependent mechanisms of mutation may be larger contributors to the mutation spectrum in individuals with higher mutant frequency in association with lower cloning efficiency.

Mutant frequency is useful in judging the likelihood that an exposure or some other factor has affected mutation. The mutant frequencies of donors from whom base substitution mutations were isolated mimick that of the overall population, being higher in smokers (10×10^{-6}) than nonsmokers (6×10^{-6}). The smokers were heterogeneous in their response to smoking as measured by the frequency of *hprt* mutant lymphocytes, however. Mutation spectra of smokers with high and low mutant frequencies will differ to the extent that smoking changes the types of mutations as well as the frequency of mutation. Though the numbers are small and the statistical significance of the difference is small ($p = 0.2$ for χ^2 analysis), it is intriguing that more base substitution events occurred at AT base pairs in smokers (6/14) than in nonsmokers (2/14). This pattern could relate to a recent observation that repair in human lymphoblastoid cells of O^4-ethylthymine and O^2-ethylthymine was less efficient than repair of O^6-ethylguanine (*43*). The average mutant frequency of these eight donors with mutations at AT base pairs was 12×10^{-6}, due largely to the six smokers whose mutant frequency averaged 13×10^{-6}. It has been noted that p53 mutations in leukemias and lymphomas show relatively more p53 mutations at AT base pairs (*39*). Differences in exposure and tissue biology probably contribute to these observations. Dose dependence of the mutation spectrum has been suggested by *in vitro hprt* mutation studies with benzo[*a*]pyrene (*23,44*).

A long-term goal of molecular epidemiology studies such as this one is to identify the impact on health risk of exposure variables, genetic susceptibility factors, and interactions between them. Given the complexities of mutation mechanisms per se and of each potential genetic target and cell type, studies of mutation spectra in any system will require population studies with analysis of large numbers of mutations. There is potential for a disease process to progressively alter the biology of the cells it affects. Surrogate cells therefore provide a valuable perspective on mutation mechanisms, one different from intermediate markers of disease. The studies of *hprt* mutation in lymphocytes reported here represent one such approach. As more is learned about the mutation spectrum of this locus, increasingly more efficient and powerful methods for comparative population studies will be possible.

Note added in proof. Mutations from 9 smokers and 12 nonsmokers have recently been reported by Vrieling et al. in *Carcinogenesis* 13:1625–1631(1992)

References

1. Harris CC. Chemical and physical carcinogenesis: advances and perspectives for the 1990s. Cancer Res 51:5023s–5044s (1991).
2. Cohen SM, Ellwein, LB. Genetic errors, cell proliferation, and carcinogenesis. Cancer Res 51:6493–6505(1991).
3. Albertini RJ, Castle KL., Borcherding WR. T-

cell cloning to detect the mutant 6-thioguanine resistant lymphocytes present in human peripheral blood. Proc Natl Acad Sci USA 79:6617–6621(1982).
4. Morley AA, Trainor KJ, Seshadri R, Ryall RB. Measurement of in vivo mutations in human lymphocytes. Nature 302:155–156(1983).
5. Janatipour M, Trainor KJ, Kutlaca R, Bennett G, Hay J, Turner DR, Morley AA. Mutations in human lymphocytes studied by an HLA selection system. Mutat Res 198:221–226(1988).
6. McCarron MA, Kutlaca A, Morley, AA. The HLA-A mutation assay: improved techniques and normal results. Mutat Res 225:89–193(1989).
7. Langlois RG, Bigbee, WL, Jensen RH. Measurement of the frequency of human erythrocytes with gene expression loss phenotypes at the glycophorin A locus. Hum Genet 74: 353–362(1986).
8. Kyoizumi SM, Akiyama M, Hirai Y, Kusunoki Y, Tanabe T, Umeki S. Spontaneous loss and alteration of antigen receptor expression in mature CD4+ T cells. J Exp Med 171:1981–1999(1990).
9. Kyoizumi S, Umeki S, Akiyama M, Hirai Y, Kusunoki Y, Nakamura N, Endoh K, Konishi J, Sasaki MS, Mori T, Fujita S, Cologne JB. Frequency of mutant T lymphocytes defective in the expression of T-cell antigen receptor gene among radiation-exposed people. Mutat Res 265: 173–180(1992).
10. Albertini RJ, Nicklas JA, O'Neill JP, Robison SH. *In vivo* somatic mutations in humans: measurement and analysis. Annu Rev Genet 24:305–326(1990).
11. Grist SA, McCarron M, Kutlaca A, Turner DR, Morley AA. In vivo somatic mutations: frequency and spectrum with age. Mutat Res 266:189–196(1992).
12. Nicklas JA, Falta MT, Hunter TC, O'Neill JP, Jacobson-Kram D, Williams JR, Albertini RJ. Molecular analysis of *in vivo* mutations in human T-lymphocytes. V. Effects of total body irradiation secondary to radioimmunoglobulin therapy (RIT). Mutagenesis 5:461–468(1990).
13. Nicklas JA, O'Neill JP, Hunter TC, Falta MT, Lippert MJ, Jacobson-Kram D, Williams JR, Albertini RJ. I*n vivo* ionizing irradiations produce deletions in the *hprt* gene of human T-lymphocytes. Mutat Res 250:383–396(1991).
14. O'Neill JP, Hunter TC, Sullivan LM, Nicklas JA, Albertini RJ. Southern-blot analyses of human T-lymphocyte mutants induced *in vitro* by g-irradiation. Mutat Res 240:143–14(1990).
15. Liber HL, Call KM, Little, JB. Molecular and biochemical analyses of spontaneous and X-ray-induced mutants in human lymphoblastoid cells. Mutat Res 178:143–153(1987).
16. Edwards A, Voss H, Rice P, Civitello A, Stegemann J, Schwager C, Zimmermann J, Erfle H, Caskey CT, Ansorge W. Automated DNA sequencing of the human HPRT locus. Genomics 6:593–608(1990).
17. Gibbs, RA, Nguyen, P-N, Edwards, A, Civitello, AB, and Caskey, CT. Multiplex DNA deletion detection and exon sequencing of the hypoxanthine phosphoribosyltransferase gene in Lesch-Nyhan families. Genomics 7: 235-244 (1990).
18. Simpson D, Crosby RM, Skopek TR. A method for specific cloning and sequencing of human *hprt* cDNA for mutation analysis. Biochem Biophys Res. Commun 151:487–492(1988).
19. Gibbs RA, Nguyen P-N, McBride LJ, Koepf SM, Caskey CT. Identification of mutations leading to Lesch-Nyhan syndrome by automated direct DNA se-quencing of in vitro amplified DNA. Proc Natl Acad Sci USA 86: 1919–1923(1989).
20. Yang J-L, Maher VM, McCormick JJ. Amplification and direct sequencing of cDNA lysate from the lysate of low numbers of diploid human cells. Gene 83: 347–354 (1989).
21. Davidson BL, Tarle SA, Palella TD, Kelley WN. Molecular basis of hypoxanthine-guanine phosphoribosyltransferase deficiency in ten subjects determined by direct sequencing of amplified transcripts. J Clin Invest 84:342–346(1989).
22. Tarle SA, Davidson BL, Wu VC, Zidar FJ, Seegmiller JE, Kelley WN, Pallela TD. Determination of the mutations responsible for the Lesch-Nyhan syndrome in 17 subjects. Genomics 10:499–501(1991).
23. Yang J-L, Chen R-H, Maher VM, McCormick JJ. Kinds and locations of mutations induced by (+/-)-7B, 8a-dihydroxy-9a,10a, epoxy-7,8,9,10-tetrahydrobenzo(*a*)pyrene in the coding region of the hypoxanthine (guanine) phosphoribosyltransferase gene in diploid human fibroblasts. Carcinogenesis 12:71–75(1991).
24. Cariello NF, Swenburg JA, Skopek TR. In vitro mutational specificity of cisplatin in the human hypoxanthine guanine phosphoribosyltransferase gene. Cancer Res 52:2866–2873(1992).
25. Rossi AM, Thijssen JCP, Tates AD, Vrieling H, Natarajan AT, Lohman PHM, van Zeeland AA. Mutations affecting RNA splicing in man are detected more frequently in somatic than in germ cells. Mutat Res 244:353–357(1990).
26. Recio L, Cochrane J, Simpson D, Skopek TR, O'Neill, JP, Nicklas, JA, Albertini, RJ DNA sequence analysis of in vivo *hprt* mutations in human T lymphocytes. Mutagenesis 5: 505–510(1990).
27. Ross AM, Tates AD, van Zeeland AA, Vrieling H. Molecular analysis of mutations affecting *hprt* mRNA splicing in human T-lymphocytes in vivo. Environ Mol Mutagen 19:7–13(1992).
28. Steingrimsdottir H, Rowley G, Dorado, G, Cole J, Lehmann, A Mutations which alter splicing in the human hypoxanthine-guanine phosphoribosyltransferase gene. Nucleic Acids Res 20:1201–1208(1992).
29. Cariello NF, Craft TR, Vrieling H, van Zeeland, AA, Adams T, Skopek, TR. Human HPRT mutant database: software for data entry and retrieval. Environ Mol Mutagen 20:81–83(1992).
30. Jones IM,. Moore DH II, Thomas CB, Thompson CL, Strout CL, Burkhart-Schultz K. Factors affecting *hprt* mutant frequency in T lymphocytes of smokers and nonsmokers. Cancer Epidemiol Bio-markers Prev (in press).
31. O'Neill JP, McGinniss MJ, Berman JK, Sullivan LM, Nicklas JA, Albertini, RJ Refinement of a T-lymphocyte cloning assay to quantify the in vivo thioguanine-resistant mutant frequency in humans. Mutagenesis 2:87–94(1987).
32. O'Neill JP, Sullivan LM, Booker JK., Pornelos BS, Falta MT, Greene CJ, Albertini, RJ. Longitudinal study of the in vivo *hprt* mutant frequency in human T-lymphocytes as determined by a cell cloning assay. Environ Mol Mutagen 13:289–293(1989).
33. Konecki DS, Brennand J, Fuscoe JC, Caskey CT. Hypoxanthine-guanine phosphoribosyltransferase genes of mouse and Chinese hamster: construction and sequence analysis of cDNA recombinants. Nucleic Acids Res 10:6763–6775(1982).
34. King A, Melton DW Characterisation of cDNA clones for hypoxanthine-guanine phosphoribosyltransferase from the human malarial parasite, *Plasmodium falciparum*: comparisons to the mammalian gene and protein. Nucleic Acids Res 15:10469–10481(1987).
35. Bochkarev MN, Kulbakina NA, Zhdanova NS, Rubtsov NB, Zakina SM, Serov OL. Evidence for tetrameric structure of mammalian hypoxanthine phosphoribosyltransferase. Biochem Genet 25:153–160 (1987).
36. Fuscoe JC, Zimmerman LJ, Lippert, MJ, Nicklas, JA, O'Neill JP, Albertini RJ. V(D)J recombinase activity mediates *hprt* gene deletion in human fetal T-lymphocytes. Cancer Res. 51:6001–6005(1991).
37. Fuscoe JC, Zimmerman LJ, Harrington-Brock K, Burnette L, Moore MM, Nicklas JA, O'Neill JP, Albertini RJ. V(D)J recombinase medicated deletion of the *hprt* gene in T-lymphocytes from human adults. Mutat Res 283:13–20(1992).
38. McGinniss MJ, Falta MT, Sullivan LM, Albertini RJ. In vivo *hprt* mutant frequencies in T-cells of normal human newborns. Mutat Res 240:117–126(1990).
39. Hollstein M, Sidransky D Vogelstein B, Harris CC. p53 mutations in human cancers. Science 253:49–53(1992).
40. Powell SM, Zilz N, Beazer-Barclay Y, Bryan TM, Hamilton SR, Thibodeau SN, Vogelstein B, Kinzler KW. APC mutations occur early during colorectal tumorigenesis. Nature 359: 235–237(1992).
41. Giannelli F, Green PM, High KA, Sommer S, Lillicrap DP, Ludwig M, Olek K, Reitsma PH, Goossens M, Yoshioka A, Brownlee, GG. Haemophilia B: database of point mutations and deletions—second edition. Nucleic Acids Res 19:2193– 2219(1991).
42. Tuddenham EGH, Cooper DN, Gitschier J, Higuchi M, Hoyer LW, Yoshioka A, Peake IR, Schwaab R, Olek K. Kazazian HH, Lavergne J-M, Giannelli F, Antonarakis, SE. Haemophilia A: database of nucleotide substitutions, deletions, insertions and rearrangements of the factor VII gene. Nucleic Acids Res 19:4821–4833(1991).
43. Wolf S, Jolly DF, Lunnen KD, Friedmann T, Migeon BR. Methylation of the hypoxanthine phosphoribosyltransferase locus on the human X chromosome: Implications of X-chromosome inactivation. Proc Natl Acad Sci USA 81: 2806-2810(1984).
43. Bronstein SM, Skopek TR Swenberg JA. Efficient repair of O^6-ethylguanine, but not O^4-ethylthymine or O^2-ethylthymine, is dependent upon O^6-alkyltransferase and nucleotide excision repair activities in human cells. Cancer Res 52:2008–2011(1992).
44. Wei S-J, Chang RL, Wong C-Q, Bhachech N, Cui XX, Hennig W, Yagi H, Sayer JM, Jerina DM, Preston DD, Conney, AH Dose-dependent differences in the profile of mutations induced by an ultimate carcinogen from benzo(*a*)pyrene. Proc Natl Acad Sci USA 88: 11227–11230(1991).