INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA

UMI®

800-521-0600

Mutation Monitoring in Human Populations

by

John Duncan Curry
B.Sc., York University, 1989
M.Sc., York University, 1993

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Biology

We accept this thesis as conforming
to the required standard

Dr. Barry W. Glickman, Supervisor (Department of Biology)

Dr. Francis Y. M. Choy, Departmental Member (Department of Biology)

Dr. Benjamin F. Koop, Departmental Member (Department of Biology)

Dr. Wolfgang Kusser, Departmental Member, (Department of Biology)

Dr. Edward E. Ishiguro, Outside Member, (Dept. of Biochemistry & Microbiology)

Dr. Barbara Shane, External Examiner, (Louisiana State University)

© JOHN DUNCAN CURRY, 1999

University of Victoria

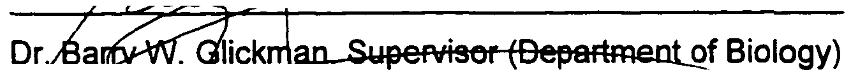
All rights reserved. Dissertation may not be reproduced in whole or in part, by photocopying or other means, without the permission of the author.

Supervisor: Dr. Barry W. Glickman

Abstract

Currently, the most widely used *in vivo* mutation monitoring system in humans is the hypoxanthine-guanine phosphoribosyltransferase (*hprt*) T-cell clonal assay. This dissertation examines the current state of the *hprt* monitoring system and the usefulness of *hprt* mutational spectra in revealing environmental exposures. The nature of spontaneous mutational spectra recovered through the implementation of this system is detailed. An examination of *hprt* mutation frequencies obtained from a set of monozygotic twins revealed a striking influence of genetic factors. As age increases, the influence of genetic factors controlling mutation frequency appears to be modified by environmental factors. Mutational spectra obtained from Russian individuals living in Moscow were distinct from the spectrum of mutation observed in age-matched Western controls. Analysis of the relationship between mutation frequency and subject age clearly demonstrated the lack of any relationship for subjects after the age of 55. This finding contradicts many previously published reports on the relationship between mutation frequency and age. Finally, the influence of tobacco smoking on mutational frequency is clear, however, no change in the mutational spectrum of smokers was revealed. Changes in mutational spectrum are analyzed in the context of the T-cell biology and reveal that the dynamics of this tissue are likely responsible for the observations made in this dissertation. Although the *hprt* gene is a highly robust and suitable target for the analysis of mutation, the target has not yet been saturated, and new single base-pair substitutions are still being characterized. The data clearly suggest that the T-cell clonai assay in its current state may not be a suitable mutational monitoring system for human populations. This dissertation concludes that new mutational assays need to be developed for monitoring mutations in human populations.

Examiners:

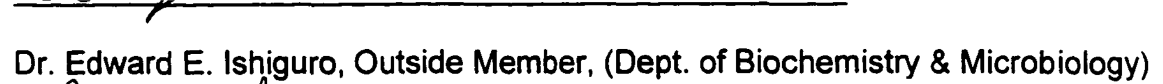


Dr. Barry W. Glickman, Supervisor (Department of Biology)

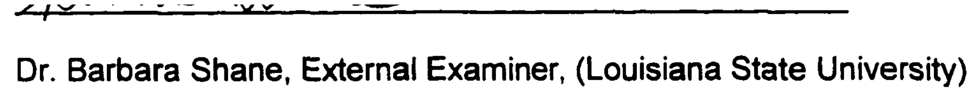
Dr. Francis Y.M. Choy, Departmental Member (Department of Biology)

Dr. Benjamin F. Koop, Departmental Member (Department of Biology)

Dr. Wolfgang Kusser, Departmental Member, (Department of Biology)



Dr. Edward E. Ishiguro, Outside Member, (Dept. of Biochemistry & Microbiology)



Dr. Barbara Shane, External Examiner, (Louisiana State University)

Table of Contents

Title Page i

Abstract ii

Table of Contents iv

List of Tables vi

List of Figures vii

Acknowledgments viii

Dedication ix

Chapter I: Introduction 1
- T-cell Clonal Assay: Overview
- T-lymphocyte *In Vivo* Clonal Expansion
- Characterization of *hprt* Mutations
- Mutational Spectra
- Outline and Rational of Dissertation
- Authors Contribution to the Multiple Authored Publications

Chapter II: Mutational Specificity and Cancer Chemoprevention 14

Chapter III: Similar Mutant Frequencies Observed Between Pairs of Monozygotic Twins. 34

Chapter IV: Mutational Spectrum of Russian Monozygotic Twins 46

Chapter V: The Influence of Sex, Smoking, and Age on Human *hprt* Mutation Frequencies and Spectra. 61

Chapter VI: Saturation of the *hprt* Mutational Target 96

Chapter VII: Discussion 113

References (Introduction and Discussion Sections) 121

Appendices

I. Spontaneous Mutations In Mammalian Cells. 124

II. Variable aberrant cDNAs in single diphtheria toxin-resistant human fibroblasts. 145

III. Moloney Murine Leukemia Reverse Transcriptase Suspect in the Production of Multiple Misincorporations during *hprt* cDNA Synthesis. 160

List of Tables

Table II.1 Mutational spectra of bacterial *lacI* mutations for the NC+ region only. .16

Table II.2 Desirable characteristics of a mutational monitoring system. 27

Table III.1 Calculation of plating and mutant frequencies. 38

Table IV.1a Mutations observed in the cDNA of 6-TG resistant T-lymphocyte clones. 51

Table IV.1b Mutants exhibiting aberrant cDNA splicing. 53

Table IV.2 A comparison of the spectrum of mutations observed in Russians 54

Table V.1 Summary of the assembled mutant frequency database by author. 65

Table V.2 Summary of regression analysis for the influence of sex and smoking on the age versus lnMF relationships. 70

Table V.3 Summary of mutant database for spectral analysis. 74

Table V.4 Subject summary data for spectra. 75

Table V.5 Mutational spectra for smoking status, sex, and age (two partitions). 77

Table V.6 Analysis of the frequency of deletion between females and males. 79

Table VI.1 *In vivo* mutational data from a *hprt* database (Curry *et al.* 1999) listed by publication year...102

Table VI.2 *In vitro* mutational data from a *hprt* database (Cariello *et al.* 1994) listed by publication year...103

List of Figures

Figure III.1 The comparison of the mean lnMF determined for each subject...............39

Figure III.2 The mutant frequency of half of the subjects versus the mutant frequency of their twin...40

Figure III.3 The mutant frequency of monozygotic twins, with age less than 40 years versus that of their twin. 41

Figure III.4 The mutant frequency of monozygotic twins, with age greater than 40 years versus that of their siblings...42

Figure V.1 Linear regression lines for the lnMF=f(age) relationships. 72

Figure VI.1 Prediction of total *in vivo* target sites for *hprt*. 104

Figure VI.2 Mutations required to saturate the *in vitro hprt target*. 104

Figure VI.3 Prediction of total *in vitro* target Sites for *hprt*.......................................105

Figure VI.4 Mutations required to saturate the *in vitro hprt target*.............................105

Figure VI.5 Prediction of target sites using both *in vitro* and *in vivo* data................106

Figure VI.6 Mutations required to saturate the *hprt* target (all data) 106

Acknowledgments

During the course of my work which is presented here, I have been supported by all the members of Dr. Barry Glickman's laboratory. In particular I wish to thank; Dr. Elliot Drobetsky, Dr. Adonis Skandalis, James Holcroft, Heather Erfle, Dr. Magomed Khaidakov, Dr. Larissa Karnaukhova, Gabriel Guenette, Dr. Barry Ford, Dr. Moyra Brackley, Dr. David Young, Jui Zhao, David Walsh , Dr. Gynn Bebb, Francis Choy, Simon Cowell, Dr. Aparecido da Cruz, Jian Chen, Dr. Johan de Boer, Henry Yu, Trent Garner, Roderick Haesevoets, Dr. Ben Koop, Paul Kotturi, Michael Parlee, Greg Stuart, Andrew Suri, Shulin Zhang, Haiyan Yang, to name just a few.

This scientific research is also supported by members of the Centre for Environmental Health at the University of Victoria. Pauline Tymchuk's support over the years has been intrinsic to the production of this dissertation, and of the published components contained within. The support of the Department of Biology at the University of Victoria is gratefully acknowledged. Throughout the writing of this thesis I have relied upon many sources to keep my sanity, and none more than the intricate sounds and rhythms of John Digweed and Paul Oakenfold.

Finally, I am very grateful for the decade of enthusiasm, encouragement, and financial support that Dr. Barry Glickman has provided during my work in his laboratory.

Dedication

This dissertation is dedicated to my loving parents Frank and Mary-Lou, who have supported me throughout my academic career. My interest in and tenacious appetite for science was fueled early by their up bringing.

Mutation Monitoring in Human Populations

I. Introduction

Mutation is extremely relevant to humans in terms of our past, current and future health. The range of mutational events is remarkable, from creating antibody diversity essential to our immune defenses, to viruses with inaccurate polymerases that accentuate mutation processes thus providing the viral gene pool with potential mechanisms to escape elimination. Mutation has dire and painful consequences to the individual in terms of genetic diseases and cancer. Aging also has been postulated to have a direct connection to the process of mutation. Several theories suggest that mutation rates increase with age and the DNA repair mechanism themselves fail, furthering mutation in the individual.

Interest in mutagenesis is further propelled by concerns that the environment is causing mutation. We must undertake an assessment of the impact of the petrochemical revolution on the human gene pool as well as those of a myriad of biological species that form the integral web of life on this planet. The nature of spontaneous mutation must also be characterized, as it too has direct consequences on human health. We are becoming increasingly aware of the great intra-individual differences, borne of mutation, in our gene pool and to the relevance of that variation on human health.

There are currently several well-established systems to monitor *in vivo* mutation in humans. These are detailed in appendix I. The most widely used system selects for mutation at the hypoxanthine-guanine phosphoribosyltransferase (*hprt*) gene as employed by the T-cell clonal assay (Albertini *et al.* 1982). The *hprt* housekeeping gene is well

suited as a surrogate target for mutational studies. Loss of HPRT function is not injurious to the cell as *de novo* pathways exist for the production of the HPRT salvaged products (Stout and Caskey 1985). As the *hprt* gene is located on the X-chromosome (Xq26.1), it is both functionally and truly hemizygous in females and males respectively (Becker *et al.* 1979). The gene target has been fully characterized and the entire genomic sequence which is spread over 44 kilobases (kb) is available (Edwards *et al.* 1990). The processed mRNA molecule consisting of 9 exons is less than 1.6 kb with only 647 bps of actual coding sequence. All classes of mutation can be recovered from this target barring large-scale deletions which may disrupt surrounding essential genes (Kronenberg *et al.* 1995).

T-cell Clonal Assay: Overview

T-lymphocytes are the basis of the clonal assay. This reflects the need to be able to grow and select *hprt* mutants. Most of the T-lymphocyte populations are in a resting state (G_0). However, they are 'primed' and ready to replicate as soon as a complementing antigen is recognized by the highly variable T-cell receptors (TCR). Once triggered by antigen, extensive cell propagation begins, and a single activated T-cell may become tens of millions in days. This proliferation potential is at the heart of the T-cell clonal assay, as the name of the assay reflects. Resting T-cells can be activated *in vitro* and their proliferation potential utilized for the selection of *in vivo* mutants. *In vitro* activation is accomplished using the mitogenic lectin phytohaemagglutinin (PHA) which specifically cross-links the CD3/TCR complex. Under assay conditions, T-cells

are activated by PHA and stimulated into continual clonal proliferation with the T-cell growth factor, interleukin-2 (IL-2). Additionally, the selective agent 6-thioguanine is added, which effectively kills those cells with functional HPRT enzymes (Waters and Swann 1997). Thus, under assay conditions only mutant T-cells are able to clonally expand. Use of multi-welled culture dishes permits a limiting dilution such that only one mutant cell will clonally expand in any given well. Counting the number of wells with proliferating mutants, as well as the overall cloning efficiency of the T-cells in a non-selective medium, permits calculation of the frequency of *in vivo* mutants for a subject. Additionally, clonally expanded mutants provide materials for determining the nature of the mutation that resulted in the loss of HPRT activity and survival under the selective conditions.

T-lymphocyte *In Vivo* Clonal Expansion

Clonal expansion of T-cell mutants poses a problem when estimating an actual mutation frequency (MuF) for subjects (Curry *et al.* 1995). Mutant frequency (MF) represents the relative number of *hprt* mutants in a subject's peripheral blood. Mutation frequency represents the number of mutational events. A single mutant T-cell has the potential to clonally expand *in vivo*, potentially producing millions of mutant T-cell clones. This expansion produces a relatively high mutant frequency, but is really a single mutational event. Thus, MF values obtained from subjects do not necessarily reflect actual MuF values.

Clones from a single clonally expanded mutant will contain a specific *hprt* mutation and a unique TCR rearrangement. Hence, they can be identified as single mutational events. Sequencing of a mutant's *hprt* gene permits characterization of the mutation that deactivated the gene. When mutants with similar mutations are recovered from a single subject, the TCR gene rearrangements can be examined to determine if the mutants are clonally related and the result of a single mutational event (Curry *et al.* 1993). Mutational frequencies are determined from corrected mutant frequency data by characterizing mutants recovered from subjects and when warranted, analyzing the TCR genes of similar mutants.

For the most part, MF values reflect good approximations of MuF, and corrections need only be applied when MF values are extraordinarily high. Still, MF has a component that is the result of the clonal expansion of single mutational events. When sufficient numbers of mutants are characterized from a single subject, some degree of *in vivo* mutant expansion will be evident (Curry *et al.* 1995). In one unusual case, an extraordinary degree of clonal expansion was detected. Despite good health, researchers noted a 100-fold increase in the subject's MF in a relatively short period of time (Nicklas *et al.* 1988). Recently these researchers have reported that within this massive clonal expansion, secondary mutations have been identified (Albertini *et al.* 1998). This and other observations indicate that it is the clonal nature of T-cells *in vivo* that ultimately may be the cause of mutation at the *hprt* gene. This hypothesis is detailed later in Chapter V.

Characterization of clonal events by means of the TCR is one of the strengths of the T-cell clonal assay. Currently there is no other *in vivo* or *in vitro* mutation monitoring system that permits the identification of clonally related mutants. For this reason, other systems, particularly *in vitro* systems, must employ strict methodologies to counter possible recovery of clonally related mutations.

Characterization of *hprt* Mutations

Several methods are currently utilized for the molecular characterization of mutant T-lymphocytes. One of the earliest methods used was a Southern blotting technique (Albertini *et al.* 1985). This method permitted deletion mutations to be characterized, though the precise sequence nature of the deletion was lacking. A benefit of this system was that the genomic blots could be washed and re-probed with TCR sequences in order to reveal the clonal relationships of mutants obtained from a single subject (Nicklas *et al.* 1986). This method has been extensively used but is limited, as *hprt* mutations involving less than 50 base pairs (bps) are not detectable. A requirement for mutant clones to expand to the size of millions, permitting enough genomic material to be extracted, is another drawback of this methodology. Routine *in vitro* clonal expansion of mutants to those high numbers is technically demanding and has not always been possible. However, researchers in the United Kingdom demonstrated that approximately 99% of mutant T-lymphocytes could be clonally expanded to greater than ten million cells (Beare *et al.* 1993).

To properly characterize mutant T-lymphocytes at the DNA sequence level several molecular techniques are required. These techniques permit a full range of mutations to be detected (base substitutions, frameshifts, insertions, duplications, and some deletions). One of the most widely used methodologies is based upon the methods of Yang *et al.* (1989), who utilized mutant *hprt* mRNA molecules recovered from the cell pellet lysates for the characterization process. Briefly, cell pellets are lysed in a soap solution containing the Moloney murine leukemia virus (MoMLV) reverse transcriptase enzyme, required DNA precursors and a *hprt* specific DNA primer. In this mixture, the first strand of cDNA is copied from the *hprt* mRNA template. Polymerase chain reactions (PCR) are used to amplify the single stranded cDNA and produce sufficient numbers of double-stranded cDNA molecules. Typically, two nested PCR amplifications are required to generate sufficient amounts of cDNA for direct sequencing (Curry *et al.* 1995), although at this point molecular cloning procedures can be used for the cDNA sequencing. Direct sequencing can be carried out on the cDNA and the molecular nature of the mutation characterized. Methods for determining the clonal relationships of T-cell mutants (de Boer *et al.* 1993) are incorporated into the *hprt* cDNA characterization methodologies (Curry *et al.* 1993). These techniques are well suited for the characterization of mutants that transcribe *hprt* mRNA. However, such techniques do occasionally generate strange artifacts (Appendix III), dictating the requirement for circumspect analysis.

The nature of the *hprt* gene presents another problem during the molecular characterization of mutants. The genomic sequence of *hprt* contains nine exon

sequences interrupted by eight introns. During normal operation the cell's splicing machinery while generating mRNA, splices these introns. Mutations that occur at sites crucial for correct exon-splicing result in mutant mRNA molecules that are characterized by missing exons. In some cases, point mutations result in the creation of novel donor or acceptor sites and cause the splicing machinery to splice only part of a given exon or to include part of the intron sequence in the transcript. Mutations that alter hairpin loop conformation, required for exon splicing, can cause aberrations in the product (Steingrimsdottir *et al.* 1992). As such, mutations that do not alter the amino acid sequence can cause *hprt* mutations in the form of defective mRNA molecules. Such mutations can occur distally from donor and acceptor sites. As such target sequences available for selectable mutations extend beyond the coding sequences. This complexity of the *hprt* gene and its effects on the potential mutational target size are discussed further in Chapter VI.

Exon splice mutations can be further characterized using a methodology termed multiplex PCR (Gibbs *et al.* 1990). A set of eight PCR primers are available, which permit the amplification of the nine *hprt* exons bounded by a portion of the surrounding intron sequences. Two of the exons that are joined by a short intron sequence are amplified together. This method can also be used rather than the cDNA method, but entails substantially more DNA sequencing efforts. The approach is particularly useful for determining the location of deletion events contained within the *hprt* gene. Specific exons with parts of their 5' and 3' introns can be sequenced to determine the location and nature of mutations affecting aberrant splicing. Splice mutations that fail to be

characterized by these methodologies may require extensive sequencing of the intron sequences. However, due to the size of some of the introns, sequencing may not be economically feasible.

Mutational Spectra

The development of mutational analysis and employment of mutational spectra is detailed extensively in Chapter II and Appendix I. Ideally, comparison of mutational spectra from an exposed group with that of a non-exposed group should reveal differences that characterize the mutagenic effects of the exposure. Mutagenic treatments performed *in vitro* are the most effective means of generating changes in the mutational spectrum observable above the spontaneous background. Changes in mutational spectrum effected through *in vivo* exposures may be much harder to detect above the background mutational spectrum. This is because *in vivo* exposures may not generate specific mutations in sufficient numbers to be visible above the background. Conversely, a mutagenic treatment or exposure may result in specific mutations clearly visible against a spontaneous spectrum devoid of such mutations.

Further complicating the generation of *in vivo* mutations are the detoxification pathways available to whole organisms. Detoxification systems have the potential to convert pre-mutagenic compounds into mutagens. Such pathways can eliminate powerful mutagens before they are able to generate mutations. However, these pathways may be overloaded with high doses of mutagen so that mutations will be observable. The use of

transgenic rodent assays is permissive to high doses, and numerous examples of mutational spectra shifts are available.

When comparing mutational spectra, the number of mutations in the spectra must be sufficiently large enough. For example, an earlier examination of mutational spectra for tobacco related changes, failed in this regard. The comparison of a non-smoking spectrum containing 55 mutations with a smoker spectrum containing 43 mutations tentatively revealed a significant difference (Vrieling *et al.* 1992). With the later comparison of larger numbers of mutations characterized from both smokers and non-smokers, no difference in spectra between the two groups could be detected (Burkhart-Schultz *et al.* 1993, 1996). Chapter VI details the largest comparison to date of mutational spectra obtained from smokers and non-smokers as well as examining sex and age effects.

A clear demonstration of a significant difference between two mutational spectra can be observed when comparing somatic mutations from normal individuals and germline mutations from individuals affected with Lesch-Nyhan's disease (Seegmiller *et al.* 1967). Although the number of germ-line mutations available is comparatively small, change in mutational spectra is startling. The normal spectrum contains approximately 30% exon splice mutations while the germ-line spectrum is essentially devoid of such mutations. This difference probably reflects the basis of the selection and the nature of exon splice mutations. Cells bearing exon splice mutations do generate some level of correctly spliced *hprt* mRNA, and thus some amount of functional HPRT, and such cells appear to be able to survive the selective agent 6-TG. Thus, individuals with germ-line

mutations that result in aberrant exon splicing potentially possess some degree of HPRT activity, permitting them to escape the debilitating effects of full HPRT loss and thus Lesch-Nyhan's disease.

Outline and Rationale of Dissertation

This dissertation primarily details the current state of the most widely used system for monitoring *in vivo* mutations in humans. The mutagenic effects of mutagens can be revealed through the employment of various mutational assays (Chapter II, Appendix I). The usefulness of mutational spectra analysis in revealing environmental exposures is currently under investigation and this dissertation discusses the shortcomings and strengths of this type of analysis.

Chapter III clearly demonstrates the potential role of genetics in determining human mutation frequencies. The environment, however, does have a considerable effect on mutation frequencies, which may become more pronounced with age (Chapter III). A potential mutational spectra deviation observed between two geographically different human populations is explored in terms of environmental exposures (Chapter IV). This possibility is examined in the context of the limitations of spectral analysis offered below.

The *hprt* T-cell clonal assay has proven to be a modestly effective system for mutation monitoring in humans. It is not without its shortcomings and the system may be a rather poor choice for monitoring mutations. Mutation frequency is closely related to age as detailed in chapter V, however, the tissue used to study mutation likely poses the biggest problem when interpreting these results. The nature and life cycle of T-cells, despite intensive study, are still poorly understood. The clonal nature of T-cells, which is

the very basis of the T-cell clonal assay, poses the most complicating problem for the study of *in vivo* mutations. In addition, the complexity of the *hprt* mutational target also limits the effectiveness of mutational spectra comparisons. Specifically, the number of target sites in *hprt* is not completely known. Saturation of the mutational target has not yet been obtained, as novel sites of mutation are still being reported (Chapter VI). The final chapter deals extensively with this issue. Finally, the need for a different mutational monitoring system is explored, and the requirements for such described. Appendix II details the details a possible system and the limitations of that system are revealed in the concluding discussion of this dissertation.

Author's Contribution to Multiple Authored Publications

The second chapter of this dissertation details the history and development of mutational spectra analysis. The chapter represents an invited paper and many of my colleagues contributed sections pertaining to their field of expertise or graduate student projects. As primary author, I was solely responsible for the organization and editing of an extensive paper.

My next two chapters stem from work that the Centre for Environmental Health was pursuing during the investigation of the potential effects of cosmic radiation on cosmonauts during their long duration flights aboard space station Mir. As principal author on these papers, the writing and editing were solely my responsibility. Dr. Khaidakov, a fellow graduate student, obtained these samples from the Russian Space Agency, but was investigating the samples obtained from the Russian space travelers (Khaidakov *et al.* 1997). A series of blood samples were obtained from seven sets of

Russian monozygotic twins and were to be used to provide control mutations for the determination of the spontaneous mutation background. However, during my extensive analysis of their mutant frequency data, a startling revelation materialized. The MF data for this group of Russian monozygotic twins appears to be tightly correlated between the twin pairs, such that the MF values fluctuate together. In addition, the mean MF determined for this group of subjects was found to be three-fold higher than age-matched Western controls. The relevance of these observations in terms of genetic control and environmental factors affecting MF are discussed. The fourth chapter is a detailed comparison of the mutational spectra obtained from these monozygotic twins with a larger mutational spectra obtained mainly from Western counterparts. Russian spontaneous spectra proved significantly different from the spontaneous spectra obtained from Westerners. This is the first report of significant regional differences in spontaneous mutational specificity.

Chapter V originally started as an investigation into the effects of age on mutant frequency and mutational spectra. As this analysis progressed, the datasets I constructed were increased to permit the analysis of sex and cigarette smoking. To date this is the largest collection of MF and mutational data that has been analyzed. While the findings confirm many previous reports on the effects of age and smoking on MF, this analysis reveals that the relationship between age and MF is more complex than has been reported previously. I also have detailed in this paper a possible difference in mutational spectra between males and females. Lack of spectral changes observed with tobacco smoke is also discussed.

My final chapter outlines a methodology to predict the effective target size of mutational targets based upon reported mutations. This concept was borne from the last chapter, and proved to be a complexity that could not be included in that chapter.

This dissertation contains three published works as appendices. The first of the appendices details the nature of spontaneous mutation specifically in mammalian cells. A number of mutational assays are detailed. I was extensively involved in the research, table construction and writing of this paper. The second appendix outlines a novel mutational assay. I was directly responsible for the mutant characterization for this project. Through extensive work that has not been reported in this dissertation, limitations of the mutational target were realized. However, the power of this novel selection system warrants its mention in this dissertation, particularly as this dissertation draws the conclusion that new mutational monitoring systems need to be developed.

The third appendix, dealing with the discovery that enzymes used for molecular characterization can actually cause strange multiple mutations, emerged from initial DNA sequencing carried out during my graduate work at York University. These unexpected results were not understood at that time. It was not until I was in Victoria, where I again found these multiple mutations, that I investigated this matter further. Others at the CEH also found similar events among their mutants. Further confirmation came during a trip to the laboratory of Dr. Samei Hou in Sweden. They had also been observing mutants characterized with multiple mutations. This publication is the result of these observations and further investigation into their potential cause.

Mutational Specificity and Cancer Chemoprevention

John Curry, Mohammed Khaidakov, Aparecido da Cruz, Larissa Karnaoukhova, Wolfgang C. Kusser, Johan de Boer, Joyce Moffat, and Barry W. Glickman

Centre for Environmental Health and the Department of Biology, University of Victoria, Victoria, British Columbia

Abstract

Mutational specificity describes the composite of all of the genetic alterations in a collection of mutations arising from a specific treatment. The information includes not only the nature of the genetic change (e.g., a base substitution or a frameshift), but also information about nucleotide position and hence the DNA context. As both the type of DNA damage and its position can be expected to reflect the nature of the chemical and physical mutagen, mutational specificity can be expected to provide insights into mechanisms of mutation. Conversely, mutational spectra should also provide insights into the identity of the mutagen. Indeed, the pioneering work on mutational specificity in *E. coli* indicates that each physical or chemical treatment produces a unique spectrum of mutations.

With the application of biotechnology to the field of genotoxicology, the database of sequenced mutations has become quite substantial. Both *in vitro* and *in vivo* data has been obtained following exposure to a variety of agents. In this communication we will critically assess whether the reality of mutational specificity has fulfilled the expectations and to examine what potential remains to be explored, especially in the area of monitoring human populations. The usefulness of both mutational spectra analysis and population monitoring with regards to chemoprevention are discussed.

©1997 Wiley-Liss, Inc.

Introduction: The Early History Of Mutational Specificity

The earliest investigations to use mutational specificity to study mutational mechanisms preceded the development of effective DNA sequencing technologies. These studies were dependent on protein sequences and were made possible by the elucidation of the genetic code. In a classic example of such work, Streisinger *et al.* (1966) using a bacteriophage lysozyme gene, first proposed the slippage mechanism for frameshift mutagenesis from the elegant use of amino acid data.

The first study of mutational specificity using direct DNA sequencing involved a bacteriophage M13 reversion assay, in which Brandenburger *et al.* (Brandenburger *et al.* 1981) sequenced 125 mutants recovered after UV and ionizing radiation treatments. Despite the limited number of sites available, they concluded that most of the mutations induced by UV light occurred at pyrimidine dinucleotides indicating that pyrimidine dimers may be involved in the induction of mutations.

While a range of studies, particularly on protein structure and function produced some insight into the nature of mutation, by far the most important system was that of J.H. Miller and his colleagues (1983) which used the suppression patterns of amber and ochre mutations to deduce single base pair substitutions in the *E. coli lacI* gene. Five of the six possible substitutions could be detected by this system. Only the A:T → G:C transition could not be detected since the TAA (ochre) to TAG (amber) or TGA (opal) changes represent the only codons available in which a single base substitution could produce a nonsense mutation. Using the *lacI* nonsense system, the mutational specificity of dozens of chemicals was determined. In addition, the impact of DNA repair has been studied by selecting the mutations in host strains of diverse genetic backgrounds. The observation that each physical or chemical agents examined demonstrate a unique spectrum of mutation was made first in this system, and remains one of the most important principles of mutational specificity (Miller *et al.* 1983).

This uniqueness of mutational specificity reflects two components: the nature and distribution of the initial DNA damage; and the sequence specificity of DNA repair. As a consequence, the nature and the distribution of mutations (i.e., mutational specificity) is unique for each mutagen.

The next development in the study of mutational specificity arose from advances in DNA sequencing technology. Using these technologies many thousands of *lacI* mutants have been sequenced over the past decade (Schaaper and Glickman 1982). One of the obvious advantages of direct DNA sequencing is that it permitted the direct detection of all six classes of base substitutions as well as frameshifts, deletions, duplications and

©1997 Wiley-Liss, Inc.

	Spontaneous	DMS	EMS
Transitions			
G:C→A:T	137 / 33.3	90 / 74.4	1241 / 98.0
A:T→G:C	38 / 9.2	3 / 2.5	8 / 0.6
Transversions			
G:C→T:A	23 / 5.6	14 / 11.6	3 / 0.2
G:C→C:G	12 / 2.9	2 / 1.6	2 / 0.2
A:T→T:A	35 / 8.5	4 / 3.3	3 / 0.2
A:T→C:G	48 / 11.7	4 / 3.3	2 / 0.2
Others			
+1 Frameshift	0 / 0.0	0 / 0.0	0 / 0.0
-1 Frameshift	18 / 4.4	4 / 3.3	4 / 0.3
Deletions	69 / 16.8	0 / 0.0	1 / 0.1
Insertions	32 / 7.8	0 / 0.0	2 / 0.2
Complex	0 / 0.0	0 / 0.0	0 / 0.0
Tandem	0 / 0.0	0 / 0.0	0 / 0.0
TOTAL	412 / 100%	121 / 100%	1266/ 100%

Table 1. Mutational spectra of bacterial *lacI* mutations for the NC+ region only. Spontaneous collection was collected by Schaaper *et al.* (1991), the dimethylsulfate (**DMS**) collection Zielenska *et al.* (1989), while the ethyl methanesulfonate (**EMS**) collection was compiled from Pienkowska *et al.* (1003), Halliday *et al.* (1990), and Burns *et al.* (1986).

insertions. The sequencing data confirmed for example, that most UV-induced mutations occurred at di-pyrimidine sites (Schaaper *et al.* 1982) as had been suggested from the analysis of *lacI* nonsense mutations recovered after UV irradiation (Miller *et al.* 1983). The sequencing studies however, also revealed something that had not been suspected: the occurrence of tandem double mutations. These generally took the form of CC→TT changes at CC sites and have become the hallmark of UV mutagenesis (Miller *et al.* 1983) from studies in *E. coli*, mammalian cells (Drobetsky *et al.* 1988), to human skin cancer (Ziegler *et al.* 1993). In addition to suggesting a mutagenic role for CC cyclobutane dimers in targeting mutations, these studies, especially when coupled with the analysis of damage distribution, led to the conclusion that mutations could be targeted by both the originally suspected lesion, the cyclobutane dimer and the (6-4) pyrimidine-pyrimidone photolesion (Drobetsky *et al.* 1987).

©1997 Wiley-Liss, Inc.

Mutational spectra differ depending upon the mutagenic treatment. Table 1 summarizes the nature of mutation in the NC+ or DNA binding region of the *lacI* gene as characterized in *E. coli* without exposure and after exposure to dimethylsulfate (DMS), and ethylmethanesulfonate (EMS). Note that the most common events in the spontaneous spectrum are single base-pair transitions (33.3%), particularly G:C→A:T. The spectra of induced mutations demonstrate a dramatic shift away from the spontaneous spectrum. Both the EMS and the DMS induced spectra show a large increase of G:C→A:T transitions (98 and 74 %, respectively).

The entire *lacI* gene includes a hotspot for spontaneous mutations consisting of 3 repeated TGGC sequences (position 620-632). When this hotspot is included in the analysis, the majority of spontaneous mutations occur at this site (duplications or deletions of one of these repeats at a ratio of 4:1, respectively). Clearly different mutagenic treatments yield strikingly different mutational spectra. It should be pointed out that mutational specificity depends not only upon the genetic target and mutagen, but also upon the biological system. For example, the spontaneous spectrum in *lacI* in bacteriophage M13 lacks the spontaneous frameshift hotspot that predominates in the bacterial spectrum (Yatagai and Glickman 1990).

©1997 Wiley-Liss, Inc.

The Study of Mutational Specificity in Mammalian Cells in Culture

The study of mutation has been facilitated by the development of selection systems that permit the assessment of mutagenic potential at a range of targets in a variety of cell types (review Glickman et al. 1993). Some of the genetic targets examined are amenable to analysis at the DNA sequence level. The most important of these are the *aprt*, *hprt*, TK, *dhfr*, and the ouabain systems. In addition, an artificial retroviral shuttle vector construct containing the *E. coli* xanthine (guanine) phosphoribosyltransferase (*gpt*) gene integrated into the chromosomal DNA of mouse cells, permitting the recovery of mutant genes has been used. An important aspect of this system is that the mutational target being characterised is integrated into the chromosomal DNA rather than existing as part of autonomously replicating plasmids. Databases of DNA sequence alterations are available for several of these gene targets each with

a number of mutagenic treatments (Provost et al. 1993).

The Development of Transgenic Animals for the Study of Mutational Specificity

The development of transgenic animals carrying target genes on recoverable shuttle vectors has provided important novel ways to assess mutagenicity and mutational specificity in animals *in vivo*. For the first time questions of strain, species, sex, age, and tissue and organ specificity of mutation can be examined *in vivo* in a short term assay. Our discussion specifically focuses on BigBlue® transgenic mice and rats which carry the *E. coli lacI* gene as the mutational target (Provost *et al.* 1993). Such targets can be recovered from mouse genomic DNA by packaging with λ packaging extracts and plating onto an *E. coli* host. A DNA fragment of 1254 base pair, containing the *lacI* gene can then be amplified by PCR and directly sequenced. The availability of a BigBlue® rat allows important inter-species comparison yielding essential information for extrapolation to humans. For *in vitro* experiments, a transgenic rat cell line, based on the same *lacI* construct, has also been established.

The nature of spontaneous mutations recovered from the animal is extremely similar for most tissues (de Boer *et al.* 1998) with the largest single class of events being G:C→A:T transitions. The majority (approximately 75%) of these transitions are recovered at 5'-CpG-3' dinucleotide sequences which suggests that they are the result of spontaneous deamination at methylated cytosines. The mutagenic response in these systems has been evaluated for a number of carcinogens including DMBA (Gorelick *et al.* 1995). The importance of mutational specificity is especially evident when the increase in mutant frequency is small. For example, we have recently determined the DNA sequence changes in mutants recovered after treatment with the flame retardant tris(2,3-dibromopropyl)phosphate (TDBP). TDBP causes tumours preferentially in the kidney of experimental animals. It is in this tissue specifically that a small increase (40%) in mutant frequency was observed; other tissues, notably stomach and liver did not demonstrate an increase in mutant frequency. A clear shift in the nature of the mutations was evident only in mutants recovered from kidney tissue (unpublished results). Both the fraction of G:C→A:T transitions and the

©1997 Wiley-Liss, Inc.

proportion recovered at 5'-CpG-3' sequences is reduced in a dose responsive manner, while the loss of G:C base pairs increased significantly.

In a similar fashion, Recio *et al.* (1995) determined the sequence of alterations in the *lacI* gene in BigBlue* mice after the inhalation of 1,3-butadiene. Mutant frequencies were increased in 2 to 3.5-fold in spleen and bone marrow. DNA sequence analysis demonstrated that substitutions at A:T base pairs were enhanced 3 to 4-fold. These substitutions are found only rarely in spontaneously arising mutants. It is clear that DNA sequence analysis increases the information provided by these studies as well as increasing the power of mutant detection when the overall increase in mutant frequency is so slight as to be statistically uncertain.

Although the transgenic animal models have only recently become available, the results from studies in these systems suggest that monitoring mutations in human populations when accompanied by DNA sequence analysis, should provide insights into the genotoxic effects of diet, lifestyles, and occupational and domestic exposures.

Monitoring Mutation in Human Populations Using the HPRT system

The T-cell hypoxanthine guanine phosphoribosyltransferase (*hprt*) clonal assay has been widely employed to monitor *in vivo* mutations in people (Albertini *et al.* 1982). The assay which permits the enumeration of 6-thioguanine resistant (TG^R) T-lymphocytes in both humans (review, Cole and Skopek 1994) and some animals , also permits mutant characterization at the DNA sequence level (Curry *et al.* 1995). The gene is relatively large, 44 kb, but the processed mRNA is relatively small, 1.6 kb, permitting the 647 bps of actual coding sequence to be easily sequenced. Mutations that occur in the non-coding regions such as the introns are detectable when the characterization of *hprt* cDNA yields specifically truncated species. These species are termed exon loss events and indicate the presence of a mutation that affects the proper splicing of *hprt* mRNA. Thus the mutational target size extends beyond the coding sequence.

T-lymphocytes undergo *in vivo* clonal expansion as part of their role in the immune system. As a consequence independent

©1997 Wiley-Liss, Inc.

mutational events can be over-represented should a mutant undergo division *in vivo*. Moreover, the actual act of cell division may be a critical step for mutation fixation. An inherent advantage of the T-cell system is that the mutant frequency can be adjusted for clonality. This is facilitated as each T-cell possesses a unique rearrangement of the T-cell receptor (TCR), which makes it possible to adjust mutant frequencies into actual mutation frequencies (Curry *et al.* 1995).

As the *hprt* clonal assay depends upon the peripheral T-lymphocyte populations, the status of that pool has direct consequences for the assay. The peripheral lymphocyte pool, about 5 x 10^9 cells, is only a small fraction of the total T-lymphocyte population (3 x 10^{12} cells) which is largely sequestered in the marginal pool (Bender *et al.* 1988). Most T-lymphocyte subsets are tissue specific and come from the gut, skin, or lymph nodes via the lymphatic system, and then preferentially return to those tissues. Migration of the 'naive' subset to the lymph nodes is observed to increase after antigen challenge (Mackay *et al.* 1992). In addition, there is a daily influx of approximately 10^9 'naive' T-lymphocytes from the thymus. Despite the complexity and heterogeneity of the T-lymphocyte peripheral pools, the ability to culture peripheral T-cells and to select mutants, makes this a convenient tissue for the study of *in vivo* mutations with the currently available protocols.

An additional complication with the use of T-cells for monitoring populations is related to their natural turnover rates in the peripheral blood. Attempts have been made to establish the half-life of T-lymphocytes by determining the rate of elimination of chromosome aberrations from these populations. However, the conclusions reported are highly variable. Estimates of half-lifes range considerably: 0 .3 years; 1.45 years; 3 years, to approximately 4 years (reviewed in da Cruz *et al.* 1996) Based upon observations of the reduction in *hprt* mutant frequency over time we have used the Buckton model and calculated a T-cell half-life of 2.1 years (Cruz *et al.* 1996). We believe that the *hprt* mutant half-life in peripheral T-cells is relatively short. Such a short half-life limits the effective use of the *hprt* assay as a method of

©1997 Wiley-Liss, Inc.

choice to monitor exposures occurring in the distant past (> 5 years).

One significant advantage of the HPRT system is the extensive database for both *in vivo* and *in vitro* mutations (Cariello *et al.* 1994). The current release of the human HPRT data base includes the sequences of more than 1300 independent mutations. Despite the size of this collection, the mutational spectrum is by no means saturated. We have recently reported four novel (not previously reported in the data base) single base pair substitutions out of twenty independent mutations from a single male subject (Curry *et al.* 1995) and have a much more striking example from a larger collection of independent mutations collected from Russian monozygotic twins. In this case, 32 novel single base pair substitutions were identified from a total of only 62 mutations.

In order to demonstrate the mutational specificity of an exposure or treatment, the spontaneous mutational spectrum must first be revealed. We have recently analyzed the *hprt* database and extracted mutants from both smokers (n=161) and non-smokers (n=290) and have failed to demonstrate that the mutational spectrum of smokers differs, even though sufficient numbers of mutants were available for a reliable comparison. Either the numbers of mutants that where available to make this comparison were insufficient or the system does not have the sensitivity to reveal a smoking effect at the *hprt* gene in T-cells.

Monitoring Mutations Induced by Chemotherapy: The case of VP-16

Etoposide (VP-16), a semisynthetic derivative of epipodophyllotoxins, is widely used as an anticancer drug and is a known mutagen (see review Anderson and Berger 1994). VP-16 has been implicated as a causative agent of secondary leukemia (Smith *et al.* 1994). Such epipodophyllotoxin-related leukemia are often (>50%) associated with abnormalities at chromosome band 11q23, particularly translocations (Felix *et al.* 1995) . It is feasible that the achievable plasma concentrations of etoposide during cancer chemotherapy can lead to the accumulation of DNA mutations and rearrangements in the blood cells which may become a cause of secondary leukemia. To investigate the possibility of etoposide induced mutation following chemotherapy in cancer

©1997 Wiley-Liss, Inc.

patients we employed the *hprt* T-cell cloning assay (Albertini *et al.* 1982) to study 12 individuals with small cell lung cancer. No increase in MF has been observed after as many as four monthly treatment cycles (Karnaoukhova *et al.* 1997).

We are currently pursuing the hypothesis that the failure to detect the induced mutation reflects the induction of apoptosis in the target cells. Etoposide has been shown to induce apoptosis in a variety of systems *in vitro* (Solary *et al.* 1994). Matsubara *et al.* (1994) showed that leukemia cells freshly obtained from patients before therapy undergo apoptosis within 6 hours following *in vitro* etoposide-prednisolone treatment. However, no apoptotic cells or fragmentation of DNA derived from peripheral blast cells were detected at any preparation following etoposide-prednisolone chemotherapy, suggesting that apoptotic cells cannot be detected due to rapid removal from the circulation (Matsubara *et al.* 1994). Cytotoxic drugs, such as etoposide primarily target rapidly dividing cells. The circulating lymphocytes with etoposide-induced DNA damage and mutation are the likely candidates to undergo apoptosis. As a result it is possible that the *hprt* assay is not able to detect etoposide-induced mutants.

©1997 Wiley-Liss, Inc.

Spectra of mutations in the *p53* gene from human cancers

As a consequence of efforts towards understanding the molecular events occurring during carcinogenesis, a wealth of mutational data related to genes involved in cancer has been generated. The prime example and largest database concerns the human tumour suppressor gene . *p53* (reviewed Zambetti and Levine 1993). The *p53* mutational database contains over 3000 entries (Hollstein et al. 1994). The *p53* gene is well suited for mutational analysis. More than 100 different codons that have been recovered as mutated from diverse human cancers (Hollstein et al. 1994) and the large number of mutable sites available in the coding sequence makes the *p53* gene very attractive for the study of mutational specificity.

Our current understanding of the etiology of cancer implies that the process leading to clinical disease is characterised by a series of genetic alterations including point mutations in *p53* as an important step. The cause of these mutations can be endogenous (e.g., deamination, replication errors, DNA repair defects) or exogenous (exposure to

environmental agents). The precise nature of these changes at the DNA sequence level offers insights into the nature of the mutagen.

The most striking example is probably found in human skin cancers. The spectrum of *p53* mutations found in invasive squamous cell carcinomas demonstrated the mutational specificity consistent with previous exposure to UV-light (Brash et al. 1991). Specifically, *p53* mutations occurred preferentially at dipyrimidines and included tandem double base substitutions, often CC→TT. These tandem double base substitutions accounted for 18% of the *p53* gene mutations characterised in skin cancer. In contrast, less then 0.1% of the *p53* mutations (in all other cancer sites combined) are tandem base substitutions (Hollstein et al. 1994). This feature is considered the hallmark of UV-induced mutations (Drobetsky et al. 1988).

The mutations in *p53* from colon cancer are characterised by G:C→A:T transitions (63%, n=960), and most of these (47% of the total *p53* mutations in colon cancer) occur at CpG sites (Hollstein et al. 1994). Transitions at CpG sites are characteristic of spontaneous mutational events (de Jong et al. 1988) and are thought to derive from deamination of 5-methylcytosine. The high incidence of G:C→A:T transitions suggests that mutational events in colon cells might be dominated by spontaneous mutations including deamination and replication errors. This is further supported by the finding that a defect in mismatch repair has been shown to be the cause of hereditary, non-polyposis colon cancer (HNPCC) (Aaltonen et al. 1993).

A correlation has also been observed between exposure to aflatoxin B1, the excretion of aflatoxin-B1-N^7-guanine in urine (Donahue et al. 1982), and the incidence of liver cancer (Hsu et al. 1991) in geographic regions with aflatoxin exposure from exogenous sources. One such region is Qidong Province in China where the mutational spectrum of *p53* mutations recovered from liver tumours has been investigated. The mutational spectrum shows G:C→T:A transversions at codon 249 of *p53* accounting for over 50% (12/21) of the mutations (Hsu et al. 1991). Aflatoxin targets G:C base pairs and the predominance of the transversion event and the appearance of a hotspot is consistent with induction by this exposure. Aflatoxin B1 induces mostly G:C→T:A transversions and the mutational spectrum from

©1997 Wiley-Liss, Inc.

hprt shows a distinct G:C→T:A transversion hotspot at position 209 (Cariello et al. 1994).

The spectrum of *p53* mutations recovered from bladder tumours also shows unique features. *p53* mutations show a high proportion of G:C→C:G transversions and double mutations both in smokers and non-smokers (Spruck et al. 1993, Kusser et al. 1994). The G:C→C:G transversion is a relatively rare mutational event. This transversion represents only 8% of all the *p53* mutations recovered (3% in colon cancer, n=960). However, in bladder cancer G:C→C:G transversions account for 21% of the sequenced mutations. Oxidative DNA damage has been suggested as an underlying cause (Spruck et al. 1993), and indeed this rare event has been induced by the singlet oxygen generating mutagen methylene blue plus light (McBride et al. 1992). Other factors need to be considered as well, 4-aminobiphenyl derived adducts at G:C base pairs have been found as a major DNA lesion in bladder biopsy samples from smokers (Talaska et al. 1991) and G:C base pairs are the major targets for 4-aminobiphenyl induced base substitution mutations (Levine et al. 1994). Metabolism may also contribute, as allelic differences in the carcinogen-metabolizing genes glutathione-S-transferase (GSTM1) and N-acetyltransferase (NAT2) have been identified as bladder cancer risk factors (Bell et al. 1993). The presence of a second *p53* mutation in a single tumour may be related to a general genetic instability of the tumour cell.

©1997 Wiley-Liss, Inc.

Future insights will depend on the accumulation of specificity data from patients having better defined occupational and lifestyle based exposures. Such conditions are difficult to achieve in real people and the human cancer database may have to be complemented by mutational analyses using transgenic animals. The *p53*-null mouse model (Donehower et al. 1992) and other transgenic systems seem ideally suited to elucidate the molecular fingerprints of carcinogens in specific genes and tissues.

The Future: Monitoring Mutation and Chemoprevention

Our ability to monitor mutational specificity in populations is severely limited by current technologies. These limitations include both biological factors such as the gene targets and tissue types to be used for selection, and the labor-intensive nature of the existing systems. We stress our belief that the problems identified with the *hprt* system largely reflect the target tissue and not the sensitivity of the target gene. Nevertheless, *hprt* is a surrogate endpoint for a biological effect, which we hope to use to assess carcinogenic potential. As a surrogate endpoint, it permits the detection of mutation much earlier than the appearance of malignancy. On the other hand, *p53* is an intermediate endpoint rather than a surrogate endpoint and thus more directly related to the development of cancer.

Both markers provide important information but it is after the fact since damage is already present. It would be of great value to monitor mutation in an earlier genetic marker for carcinogenesis. However, such a genetic marker/s remains to be identified. From a therapeutic standpoint knowing the progression of disease is of interest. These markers may offer advantages in drug trials for chemoprevention studies since they can be used as early markers for disease recurrence, particularly *p53* which has already been used for this purpose (Brennan *et al.* 1995). When a specific exposure is suspected and mutational specificity data available, it is possible to estimate *p53* mutation frequencies *in vivo* (Pourzand and Cerutti 1993). The approach involves the screening for the loss of restriction sites at mutational hotspots. The nature of the mutation is then confirmed by DNA sequencing. While exceedingly elegant, it is unlikely to provide a practical approach for monitoring populations. However, with future advances, it may ultimately be possible to use mutational specificity in humans to specifically identify exposure.

The mutational systems described in this article could be used to determine the effectiveness of chemopreventive agents, in that they would permit the changes in mutational specificity to be determined. *In vitro* studies employing human cell lines could be used to determine the spontaneous spectrum, an induced spectrum (exposed), and an induced spectrum in

©1997 Wiley-Liss, Inc.

cells treated with a chemopreventive agent. Comparisons then of the mutation frequencies, and mutational spectra would permit an assessment of the chemopreventive agent. It is, however, important to realize that extrapolation of any *in vitro* testing results to the human situation will prove considerably difficult. Should a specific chemopreventive agent work to reduce the mutation rate in a cell line, then the *in vitro* system may hold promise. However, if the chemopreventive agent operates in a different manner, such as to reduce inflammation, the efficacy of the agent may not be revealed in these simple *in vitro* systems.

To explore the effectiveness of chemopreventive agents directly in humans with existing monitoring systems will prove exceedingly difficult for several reasons. With the current *hprt* system, the wide range of mutant frequencies values observed for human populations requires that a large population be examined for any possible shift in mutant frequencies alone. In addition, our understanding of the *in vivo* spontaneous mutational spectrum is still limited such that any comparison to an exposed spectrum is difficult. Indeed any comparisons will require large numbers of mutants collected from suitable populations. Induced human spectra with numbers less than 100 are not useful, as the spontaneous spectrum can potentially mask any induced mutants.

Recently, the chemopreventive properties of oltipraz, an inducer of GST in mice and rats (Liu *et al.* 1988), and conjugated linoleic acid (CLA), have been investigated. Treatment of rats with oltipraz results in induction of GST Yc gene expression. This is of importance in the detoxification of *e.g.* aflatoxin-B_1 in the liver. Humans can hypothetically be made relatively resistant to the effects of AFB_1 through dietary supplementation with oltipraz; however, this hypothesis assumes that human liver, like rat liver, contains a GST isozyme that can be induced and which can detoxify AFB_1. Human clinical trials are currently underway to evaluate the chemoprotective effect of oltipraz in the diet. CLA, a mixture of positional and geometric isomers of linoleic acid which occurs naturally in dairy products and cooked meat demonstrates chemopreventive properties in rodents (Ip *et al.* 1991), and anticancer activity in human skin.

©1997 Wiley-Liss, Inc.

colorectal, and mammary cells *in vitro* (Shultz *et al.* 1992). Prevention is manifest at concentrations which are close to the levels consumed by humans in the diet. One percent of CLA in the diet suppresses mammary gland carcinogenesis in rats given a high dose (10 mg) of dimethylbenzanthracene (Ip *et al.* 1991). The use of transgenic animals to determine the modulation of mutation induction will prove to be an attractive method to investigate the properties of chemopreventive agents. In addition the determination of chemically induced mutation spectra in the absence and presence of chemopreventive compounds may elucidate some of the mechanisms by which these chemicals act.

Advances in DNA sequencing technology may make the direct detection of mutants possible without phenotypic selection. This would offer the advantage that cells would not need to be grown. Systems such as gradient denaturing gels (Fodde *et al.* 1994), constant denaturing gels (Borresen *et al.* 1991) or constant denaturing capillary electrophoresis (Kuypers *et al.* 1993) present some promise, but currently fall short of the sensitivity required to detect mutation at frequencies below one per 10^5 copies. It should be recognized that the accuracy required is greater than that provided by most available DNA polymerases utilized for PCR reactions.

Future systems to monitor mutations in human populations can be expected to fulfill certain practical and scientific demands. Table 2

1	Tissue does not need to be cultured.
2	No phenotypic selection.
3	Multiple tissue choices.
4	High sensitivity.
5	Universal endpoints - deletions and base substitutions.
6	Transcribed gene, expressed in all tissues in a similar manner.
7	Limited size for convenience (small cDNA) but reasonably large as a gene target (large gDNA).
8	Not sex linked.
9	Two homologous present so LOH can be studied. (another reason for non-phenotypic selection, genetics neutral. Do not need a heterozygous or hemizygous state of dominant gene mutation.)
10	Copy number stable in normal cells.
11	Well mapped chromosomal locations of polymorphism's for LOH and inheritance studies.

Table 2. Desirable characteristics of a mutational monitoring system

©1997 Wiley-Liss, Inc.

presents the wish list for an ideal system. Foremost, the system should avoid the need for *in vitro* culturing. Any tissue that can be painlessly obtained could then be used. This would include blood, buccal, hair follicles, skin as well as exfoliated bladder cells. One way to achieve freedom from the need to culture the cells, is to use a non-phenotypic method of mutant selection. This could also broaden the spectrum of recoverable mutants from a target gene. The utility of a system would also be enhanced if different classes of genetic alterations could be detected. In the near future, the human genome project will likely provide novel mutational targets, and mutational specificity studies can be expected to become more sophisticated. As each genetic target will offer both unique advantages and technical challenges, the system of the future may not rely upon a single target or single method of analysis.

©1997 Wiley-Liss, Inc.

Acknowledgements

We wish to dedicate this paper to the memory of Joyce Moffat who inspired this contribution. Joyce succumbed October 23, 1996 after a brave battle against pancreatic cancer.

©1997 Wiley-Liss, Inc.

References

Aaltonen LA, Peltomaki P, Leach FS, Sistonen P, Rylkkanen L, Mecklin JP, Jarvinen H, Powell SM, Jen J, Hamilton SR, Petersen GM, Kinzler KW, Vogelstein B, de la Chapelle A.: Clues to the pathogenesis of familial colorectal cancer. Science 260:812-816, 1993.

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.

Anderson R, Berger N: International commission for protection against environmental mutagens and carcinogens. Mutagenicity and carcinogenicity of topoisomerase-interactive agents. Mut Res 309:109-142, 1994.

Bell DA, Taylor JA, Paulson DF, Robertson CN, Mohler JL, Lucier GW: Genetic risk and carcinogen exposure: a common inherited defect of the carcinogen-metabolism gene glutathione S-transferase M1 (GSTM1) that increases susceptibility to bladder cancer. J Natl Cancer Inst 85:1159-1164, 1993.

Bender MA, Awa AA, Brooks AL, Evans HJ, Groer PG, Littlefield LG, Pereira C, Preston RJ, Wachholz BW : Current status of cytogenetic procedures to detect and quantify previous exposures to radiation. Mut Res 196:103-159, 1988.

Borresen A, Hovig E, Smith-Sorensen B, Malkin D, Lystad S, Andersen TI, Nesland JM, Isselbacher KJ, Friend SH: Constant denaturing gel electrophoresis as a rapid screening technique for p54 mutations. Proc Natl Acad Sci (USA) 88:8405-8409, 1991.

Brandenburger A, Godson GN, Radman M, Glickman BW, van Sluis CA, Doubleday OP: Radiation-induced base substitution mutagenesis in single-stranded DNA phage M13. Nature 294:180-182, 1981.

Brash DE, Rudolph AR, Simon JA, Lin A, McKenna GJ, Baden HP, Halperin AJ, Ponten J: role for sunlight in skin cancer: UV-induced *p53* mutations in squamous cell carcinoma. Proc Natl Acad Sci (USA) 88:10124-10128, 1991.

Brennan JA, Mao L, Hruban RH, Boyle JO, Eby YJ, Koch WM, Goodman SN, Sidransky D: Molecular assessment of histopathological staging in squamous-cell carcinoma of the head and neck. New England Journal of Medicine 332:429-435, 1995.

Burns PA, Allen FL, Glickman BW: DNA sequence analysis of mutagenicity and site specificity of ethyl methanesulfonate in Uvr⁻ and UvrB strains of *Escherichia coli*. Genetics 113:811-819, 1986.

Cariello NF, Beroud C, Soussi T: Database and software for the analysis of mutations at the human *p53* gene. *Nucleic Acids Res* 22:3549-3550, 1994.

Cariello NF, Cui L, Skopek TR In vitro mutational spectrum of aflatoxin B_1 in the human hypoxanthine-guanine phosphoribosyl transferase gene. Cancer Res. 54:4436-4441, 1994.

Cole J, Skopek TR: Somatic mutant frequency, mutation rates and mutational spectra in the human population in vivo. Mut Res 304:33-105, 1994.

Curry J, Rowley GT, Saddi V, Beare D, Cole J, Glickman BW : Determination of *hprt* mutant and mutation frequencies and the molecular characterization of human derived *in vivo* T-lymphocyte mutants. Env Molec Mut 25:169-179, 1995.

da Cruz AD, Curry J, Curado MP, Glickman BW: Monitoring hprt mutant frequency over time in T-lymphocytes of people accidentally exposed to high doses of ionizing radiation. Env Molec Mut 27:165-175, 1996.

de Boer JG, Provost S, Gorelick N, Tindall K, Glickman BW: Spontaneous mutation in *lacI* transgenic mice: a comparison of tissues. Mutagenesis 13(2):109-114, 1998.

©1997 Wiley-Liss, Inc.

De Jong PJ, Grosovsky AJ, Glickman BW: Spectrum of spontaneous mutations at the *APRT* locus of Chinese hamster ovary cells: an analysis at the DNA sequence level. Proc Natl Acad Sci (USA) 85:3499-3503, 1988.

Donahue PR, Essigman JM, Wogan GN: Aflatoxin DNA adducts: Detection in urine as a dosimeter of exposure. In Bridges BA, Butterworth BE, Weinstein IB (eds): "Indicators of Genotoxic Exposure" Cold Spring Harbor, NY, 1982, pp 221-229.

Donehower LA, Harvey M, Slagle BL, Arthur MJ, Mongomery Jr Ca, Butel JS, Bradley A: Mice deficient in p53 are developmentally normal but susceptible to spontaneous tumors. Nature 356:215-221, 1992.

Drobetsky EA, Glickman BW, Grosovsky AJ: Mutagenic specificity of UV light in bacterial and mammalian cells. **In** "Mechanisms and Consequences of DNA Damage Processing". UCLA Symposia on Molecular and Cellular Biology, 1988, pp 499-503.

Drobetsky EA, Grosovsky AJ, Glickman BW: The specificity of UV-induced mutations at an endogenous locus in mammalian cells. Proc Natl Acad Sci (USA) 84:9103-9107, 1987.

Felix C, Holster M, Winick N, Masterson M, Wilson A, Langer B: ALL-1 gene rearrangements in DNA topoisomerase II inhibitor-related leukemia in children. Blood 85:3250-3256, 1995.

Fodde R, Losekoot M: Mutation detection by denaturing gradient gel electrophoresis (DGGE). Human Mut 3:83-94, 1994.

Glickman BW, Saddi VA, Curry J: Spontaneous mutation in mammalian cells. Mut Res 304:19-32, 1994.

Gorelick NJ, Andrews LA, Gu M, Glickman BW: Mutational spectra in the *lacI* gene in skin from 7,12-dimethylbenz[a]anthracene-treated and untreated transgenic mice. Molec Carcinogenesis 14:53-62, 1995.

Halliday J, Zielenska M, Awadallah S, Glickman BW : Colony hybridization in *Escherichia coli*: A rapid procedure for determining the distribution of specific classes of mutations among a number of preselected sites. Env Molec Mut 16:143-148, 1990.

Hollstein M, Rice K, Greenblatt MS, Soussi T, Fuchs R, Sorlie T, Hovig E, Smith-Sorensen B, Montesano R, Harris CC: Database of *p53* somatic mutations in human tumors and cell lines. Nucleic Acids Res 22:3551-3555, 1994.

Hsu IC, Metcalf RA, Sun T, Welsh JA, Wang NJ, Harris CC: *p53* gene mutational hot spot in human hepatocellular carcinoma from Qidong, China. Nature (London) 350:427-428, 1991.

Ip, C., Chin, S.F., Scimeca, J.A., and Pariza, M.W., Mammary cancer prevention by conjugated dienoic derivative of linoleic acid. Cancer Res 51:6118-6124, 1991.

Karnaoukhova, L., Moffat J., Martins H., Glickman B.W., Mutation frequency and spectrum in lymphocytes of small cell lung cancer patients receiving etoposide chemotherapy. Cancer Res. 57(19): 4393-4407, 1997.

Kusser WC, Miao X, Glickman BW, Friedland JM, Rothman N, Hemstreet GP, Mellot J, Swan DC, Schulte PA, Hayes R: *p53* mutations in human bladder cancer. Env Molec Mut 24:156-160, 1994.

Kuypers AWHM, Willems PMW, Van der Schans MJ, Linssen PCM, Wessels HMC, de Bruijn CHMM, Everaerts FM, Mensink EJBM: Detection of point mutations in DNA using capillary electrophoresis in a polymer network. J Chromtography 621:149-156, 1993.

Levine JG, Schaaper RM, DeMarini DM: Complex framesift mutations mediated by plasmid pKM101: Mutational mechanism deduced from 4-aminobiphenyl-induced mutation spectra in Salmonella. Genetics 136:731-746, 1994.

©1997 Wiley-Liss, Inc.

Liu, Y.L., Roebuck, B.D., Yager, J.D., Groopman, J.D., and Kensler, T.W., Protection by 5-(2-pyrazinyl)-4-methyl-1,2-dithiol-3-thione (oltipraz) against the hepatotoxicity of aflatoxin B1 in the rat. Toxicology and Applied Pharmacology 93:442-451, 1988.

Mackay CR, Marston WL, Dudler L, Spertini O, Tedder TF, Hein WR : Tissue-specific migration pathways by phenotypically distinct subpopulations of memory T-cells. Euro J Immun 22:887-895, 1992.

Matsubara K, Kubota M, Adachi S, Kuwakado K, Hirota H, Wakazono Y, Akiyama Y, Mikawa H: Induction of apoptosis in childhood acute leukemia by chemotherapeutic agents: failure to detect evidence of apoptosis *in vivo*. Eur J Haematol 52:47-52, 1994.

McBride TJ, Schneider JE, Floyd RA, Loeb LA: Mutations induced by methylene blue plus light in single-stranded M13mp2. Proc Natl Acad Sci (USA) 89:6866-6870, 1992.

Miller JH: Mutational specificity in bacteria. Annual Review of Genetics 17:217-238, 1983.

Pienkowska M, Ferreira A, Anderson M, Zielenska M, Glickman BW: Large scale mutational analysis of EMS induced mutation in the *lacI* gene of *Escherichia coli*. Mut Res 288:123-131, 1993.

Pourzand C, Cerutti P: Genotypic mutation analysis by RFLP/PCR. Mut Res 288:113-121, 1993.

Provost GS, Kretz PL, Hamner RT, Matthews CD, Rogers BJ, Lundberg KS, Dycaico MJ, Short JM.: Transgenic systems for *in vivo* mutation analysis. Mut Res 288: 133-149, 1993.

Recio, L, Meyer KG: Increased frequency of mutations at A:T base pairs in the bone marrow of B6C3F1 *lacI* trangenic mice exposed to 1,3-butadiene. Env Molec Mut 26:1-8, 1995.

Schaaper RM and Dunn RL: Spontaneous Mutation in the Escherichia-Coli *lacI* Gene. *Genetics* 129:317-326, 1991.

Schaaper RM, Glickman BW: Mutability of bacteriophage M13 by ultraviolet light: Role of pyrimidine dimers. Mol Gen Genet 185:404-407, 1982.

Shultz, T.D., Chew, B.P., Seaman, W.R., and Luedecke, L.O. Inhibitory effect of conjugated dienoic derivatives of linoleic acid and beta-carotene on the in vitro growth of human cancer cells. Cancer Letters 63:125-133, 1992b.

Smith M, Rubinstein L, Ungerleider R: Therapy-related acute myeloid leukemia following treatment with epipodophyllotoxins: estimating the risks. Med Pediatr Oncol 23:86-98, 1994.

Solary E, Bertrand R, Pommier Y: Apoptosis induced by DNA Topoisomerase I and II inhibitors in human leukemic HL-60 cells. Leukemia and Lymphoma 15:21-32, 1994.

Spruck III CH, Rideout III WM, Olumi AF, Ohneseit PF, Yang AS, Tsai YC, Nichols PW, Horn T, Hermann GG, Steven K, Ross RK, Yu MC, Jones PA: Distinct pattern of *p53* mutations in bladder cancer: relationship to tobacco usage. Cancer Res 53:1162-1166, 1993.

Streisinger G, Okada Y, Emrich J, Newton J, Tsugita A, Terzaghi E, Inouye M: Frameshift mutations and the genetic code. Cold Spring Harbor Symp Quant Biol 31:77-84, 1966.

Talaska G, Al-Juburi AZ, Kadlubar FF: Smoking related carcinogen-DNA adducts in biopsy samples of human urinary bladder: identification of N-(deoxyguanosine-8-yl)-4-aminobiphenyl as a major adduct. Proc Natl Acad Sci (USA) 88:5350-5354, 1991.

Yatagai F, Glickman BW: Specificity of spontaneous mutation in the *lacI* gene cloned into bacteriophage M13. Mut Res 243:21-28, 1990.

Zambetti GP, Levine AJ: comparison of the biological activities of wild-type and mutant *p53*. FASEB Journal 7:855-865, 1993.

©1997 Wiley-Liss, Inc.

Ziegler A, Leffell DJ, Kunala S, Sharma HW, Gailani M, Simon JA, Halperin AJ, Baden HP, Shapiro PE, Bale Ae: Mutation hotspost due to sunlight in the p53 gene of nonmelanoma skin cancers. Proc Natl Acad Sci (USA) 90:4216-4220, 1993.

Zielenska M, Horsfall MJ, Glickman BW: The dissimiliar mutational consequences of S_N1 and S_N2 DNA alkylation pathways: clues from the mutational specificity of dimethylsulphate in the *lacI* gene of *Escherichia coli*. Mutagenesis 5:230-234, 1989.

©1997 Wiley-Liss, Inc.

Human Mutation 9:445-451 (1997)

Similar Mutant Frequencies Observed Between Pairs of Monozygotic Twins

John Curry, Gwyn Bebb, Joyce Moffat, Daniel Young, Magomed Khaidakov, Alan Mortimer*, and Barry W. Glickman

Centre for Environmental Health and the Department of Biology, University of Victoria, Victoria, British Columbia V8W 3N5, Canada, Fax: (250) 472-4075
*Space Life Sciences, Canadian Space Agency, Ottawa, Ontario, K1L 8E3, Canada

Abstract

The relative contribution of both genetic and environmental factors to spontaneous mutation frequency in humans is unknown. We have investigated the contribution of genetic factors to this phenomenon by determining the *in vivo* mutation frequency at the hypoxanthine-guanine phosphoribosyltransferase (*hprt*) locus in circulating T-lymphocytes obtained from pairs of monozygotic twins. *Hprt* mutant frequencies were determined three times over fourteen days in six sets of monozygotic male twins (mean age 30) taking part in a Russian Space Program inclined bed rest experiment. Blood samples were obtained prior to, during, and immediately following the experiment. Mononuclear cells were separated, frozen, and flown to Canada for analysis using the *hprt* T-lymphocyte clonal assay. There is no evidence within this data set to demonstrate that the period of inclined bed rest to simulate the effects of weightlessness had any effect upon the observed mutant frequency. However, the average mutant frequency for the six sets of Russian twins was found to be three times higher than that of Western counterparts. More surprisingly, the spontaneous mutant frequency of monozygotic twins was found to be much more similar within pairs than between pairs of twins. These data suggest that the contribution of genetics in the determination of mutation frequency is substantial. However, whether high concordance within twin pairs reflects shared environmental experience as well as common genetic factors is not entirely clear. More data will be required to distinguish genetic from environmental factors, and to determine the degree to which mutant frequency is genetically determined.

©1997 Wiley-Liss,Inc.

Introduction

There is substantial evidence that spontaneous mutation rates are genetically controlled (Drake, 1970, 1991). However, the relative contribution of both genes and the environment to the generation of background mutation is unclear. The extent to which mutations are accumulated by an individuals must in part be determined by that individual's genome. Accordingly the inherent plasticity in the genome must allow environmental

factors to influence spontaneous mutation as well. A previous study employing the classic twin study design with the sister chromatid exchance (SCE) assay demonstrated that after accounting for smoking and coffee drinking status, genetic factors accounted for ~30% of the variation in SCE rates (Hirsh *et al.*, 1992). In this study, we have determined the spontaneous *in vivo* mutant frequency for six sets of monozygotic twins in an attempt to examine the potential genetic and environmental influences on this process.

Inter-individual variation in spontaneous mutant frequency has been well documented (Cole *et al.*, 1988; Tates *et al.* 1991; Branda *et al.* 1993; Cole and Skopek, 1994). This variation at the cellular level might be expected since considerable inter-individual variation in sensitivity has been observed *in* vivo to known mutagens such as radiation and cigarette smoke (Paterson *et al.* 1985a; Caporaso and Landi 1994). Many factors including polymorphisms in metabolizing enzymes, differences in DNA repair capability, and differences in chromatin structure could contribute to this variation. (Spitz, 1995; Hittleman and Pandita, 1994). Such individual variation in sensitivity has relevance in the clinical setting. A wide variation in radiosensitivity has been observed in normal individuals (Little *et al.* 1989) as well as in individuals with known DNA repair disorders such as Ataxia Telangiectasia, which are associated with an increased susceptibility for cancer (Paterson *et al.*, 1984). For example, it is not uncommon for a patient undergoing radiotherapy to have the treatment terminated due to adverse effects stemming from their intrinsic radiation sensitivity (Paterson *et al.* 1985b; Arlett and Harcourt 1980; Chen *et al.* 1978; Little *et al.* 1988; Nagasawa and Little 1988). In an extreme example of radiation sensitivity, Badie *et al.* (1995) described a leukemia patient who died during radiation therapy. The patient was later realized to be defective in the repair of DNA double-strand breaks that was possibly a causative factor in the development of the initial disease. (Badie *et al.* 1995).

Clearly, the philosophy of "equal exposure = equal risk" is not valid. Yet, generally accepted safety standards of exposure to hazardous agents are based on this assumption, and fail to take into account the broad spectrum of potential human responses. Precisely identifying factors which influence individual variation in the spontaneous

© 1997 Wiley-Liss, Inc.

mutation frequency and susceptibility to mutagens has far-reaching implications. Appropriate protection or restriction from exposure to radiation or chemicals may be required if an individual's sensitivity to these agents can be established. Variation in individual response could be used as the basis of a screening program with the possibility that life style counseling be considered for particularly susceptible individuals. Before that is possible however, a clearer understanding of the genetic basis for mutation and the influences of the environment on mutagenesis will be required.

The ability to grow *hprt* mutant T-cells *in vitro* has allowed researchers to establish *in vivo* mutant frequencies (Albertini *et al.* 1982; Cole *et al.* 1988, Tates *et al.* 1991, Branda *et al.* 1993) and then characterize the nature of mutation at the molecular level (Rossi *et al.* 1992, Recio *et al.* 1990, Gibbs *et al.* 1990, Fuscoe *et al.* 1991; Steingrimsdottir *et al.* 1992, 1993; Cole and Slopek 1994; Curry *et al.* 1993, 1995). Thus, in populations exposed to environmental mutagens, *hprt* provides a molecular marker for mutation.

This study involves a set of six male monozygotic twins participating in a Russian Space Agency study designed to simulate the physiological effects of weightlessness in space. The simulation involves bedrest at an inclination (head down) of 7°, for a period of 2 weeks. There was no expectation that this regimen would affect mutant frequencies, but the study did provided access to healthy, well monitored, monozygotic twins. The individual sets of twins had always lived together; and the entire group was housed and fed together, two weeks prior to and throughout the bedrest protocol. Thus, each of the twin sets had experienced similar environments for most of their lives and throughout the course of the experiment. The question posed here is whether monozygotic twins have similar mutant frequencies. The finding of similar mutant frequencies would suggest that the environmental conditions and/or genetic factors affect each member of a twin pair in a similar manner.

© 1997 Wiley-Liss, Inc.

Methods

Sample Collection

The Russian Space Agency identified seven sets of monozygotic twins to participate in an inclined bed rest experiment that simulates the physiological effects of weightlessness. These sets of twins were identified from a national database established to facilitate a variety of experimental

applications. These sets of twins were demonstrated to be true monozygotic twins. The twins were not exposed to any other experimental factors except the bed rest, and only one set of twins smoked. Mononuclear cells (MNC's) were recovered from the volunteers (mean age 30 ± 2 years) using Leucoprep tubes (Becton-Dickinson, Franklin Lakes, NJ) following the manufacturer's protocols. Blood was drawn from each set of twins at the same time of day. Coded samples were collected prior to the bed rest experiment (t=0 days), during the bed rest (t=14 days), and following the conclusion (t=21 days) of the simulation. MNC's were counted and diluted to 10^{7} cells/ml in 50% calf serum (CBS), 10% DMSO, and 40% RPMI 1640. The cells were frozen in a styrofoam box placed into a -80°C freezer (Cole *et al.* 1988), transported to Canada on dry ice, and stored under liquid nitrogen until required.

T-Cell Clonal Assay

Frozen MNC's were thawed quickly in a 37°C waterbath, washed twice in RPMI 1640 supplemented with 10% CBS, and pre-incubated in growth medium (no phytohematogglutin (PHA) or interleukin 2 (IL-2)) overnight (12-16h) (Cole *et al.*, 1988). MNC's were plated in 200 µl of selective (6-thioguanine) and non-selective media (2 plates) in microtitre plates (96 well, flatbottom) at 10^4 and 3 cells per well, respectively, along with 10^4 lethally irradiated feeder cells (RJK 853 cells, a lymphoblastoid cell line, derived from a Lesch-Nyhan patient found to have a complete deletion of the *hprt* gene (Yang *et al.* 1984)). Our growth medium consisted of RPMI 1640, 20% HL-1, 5% CBS, 5% human AB serum (see Table 1), 5U/ml IL-2, 0.25µg/ml PHA, 2mM L-glutamine, 2mM pyruvic acid, 100U/ml penicillin, 100µg/ml streptomycin and 4% Fungizone. Selection plates contained 10^{-5} M 6-thioguanine. Plates were incubated on a sloped shelf (5°) in a 5% CO_2, 37°C, humidified incubator for a period of 14 days. Twenty-four hours prior to scoring, the plates were rotated 180° on the sloped shelf then scored visually using an inverted phase contrast microscope, for wells that contained expanding colonies. All plates were re-scored three days later to allow small colonies which may have been missed during the first scoring to be counted. Individual mutant clones were collected from the selective plates and further expanded when possible, and stored to permit the

© 1997 Wiley-Liss, Inc.

Subject Sample	Non-Selection Wells	6-TG Selection Wells	PE	MF (x 10^{-5})	95% Upper (x 10^{-5})	95% Lower (x 10^{-5})
R2B1*	163	946/960	0.05	2.7	5.1	1.4
R2A2*	178	473/480	0.03	5.8	14.4	2.3
R2B2*	182	666/672	0.02	5.0	13.8	1.8
R2A3*	175	761/768	0.03	3.0	7.1	1.2
R2B3*	166	760/768	0.05	2.2	4.8	1.0
R3A1*	141	404/408	0.10	1.0	2.6	0.3
R3B1*	139	397/400	0.11	0.7	2.2	0.2
R3A2*	150	119/120	0.08	1.0	7.4	0.1
R3B2*	157	451/456	0.07	1.6	4.2	0.6
R3A3*	178	288/288	0.03	0.0	-	-
R3B3*	174	96/96	0.03	0.0	-	-
R4A1*	188	1048/1056	0.01	10.8	36.0	3.3
R4B1*	186	355/360	0.01	13.2	43.3	4.0
R4A2	85	398/414	0.27	1.5	2.5	0.9
R4B2	106	366/376	0.20	1.4	2.6	0.7
R4A3	113	386/400	0.18	2.0	3.6	1.1
R4B3	124	209/216	0.15	2.3	4.9	1.0
R5A1	91	165/168	0.25	0.7	2.3	0.2
R5B1	99	104/104	0.22	0.0	-	-
R5A2	125	142/144	0.14	1.0	4.0	0.2
R5B2	123	142/144	0.15	0.9	3.8	0.2
R5A3	88	302/312	0.26	1.3	2.4	0.7
R5B3	71	348/360	0.33	1.0	1.9	0.6
R6A1	76	196/208	0.31	1.9	3.5	1.1
R6B1	53	177/192	0.43	1.9	3.2	1.1
R6A2	100	22/24	0.22	4.0	16.3	1.0
R6B2	105	107/112	0.20	2.3	5.6	0.9
R6A3	95	126/136	0.23	3.3	6.3	1.7
R6B3	96	229/240	0.23	2.0	3.8	1.1
R7A1	117	128/136	0.17	3.7	7.6	1.8
R7B1	84	209/224	0.28	2.5	4.3	1.5
R7A2	118	117/120	0.16	1.6	5.0	0.5
R7B2	143	259/264	0.10	1.9	4.9	0.8
R7A3	106	516/536	0.20	1.9	3.1	1.2
R7B3	111	380/392	0.18	1.7	3.1	0.9

Table 1. Calculation of plating and mutant frequencies. Subject samples are coded such that R# refers to a set of twins (# = 2-7) where A or B indicates the individual twin. The next number following refers to the sample collection period: (1) before (t=0), (2) during (t=14), and (3) after (t=21) the bed rest experiment. The age of the twins is as follows: R2=31, R3=35, R4=35, R5=24, R6=24, R7=31 years. The numbers given for the non-selection (192 wells plated) and 6-TG selection wells, are those wells in which no colony was observed. The 95% lower and upper confidence intervals for the mutant frequency are calculated following the methods of Branda *et al.* (1993). Subject codes identified (*) have lower plating efficiencies than those not identified. These samples were plated in our laboratories in Toronto, Canada. At that time we had been using RPMI 1640 and HL-1 prepared in our laboratory and were not using human serum. The other samples, those with higher efficiencies, were plated in Victoria, Canada with pre-made RPMI 1640, HL-1, and human serum was used. We attribute the difference observed in plating efficiencies to these factors.

characterization of the *hprt* and γ-TCR sequences at a later date.

Results

Mutant frequencies (MF) were determined for each of the three samples taken from the subjects (Table 1). From the 36 blood samples subjected to the T-cell clonal assay, 33 yielded usable MF data and those which were zero were excluded from further analysis. The mean plating efficiency (PE) and MF for all subjects was 15.3 ± 1.7 % and 27.5

© 1997 Wiley-Liss, Inc.

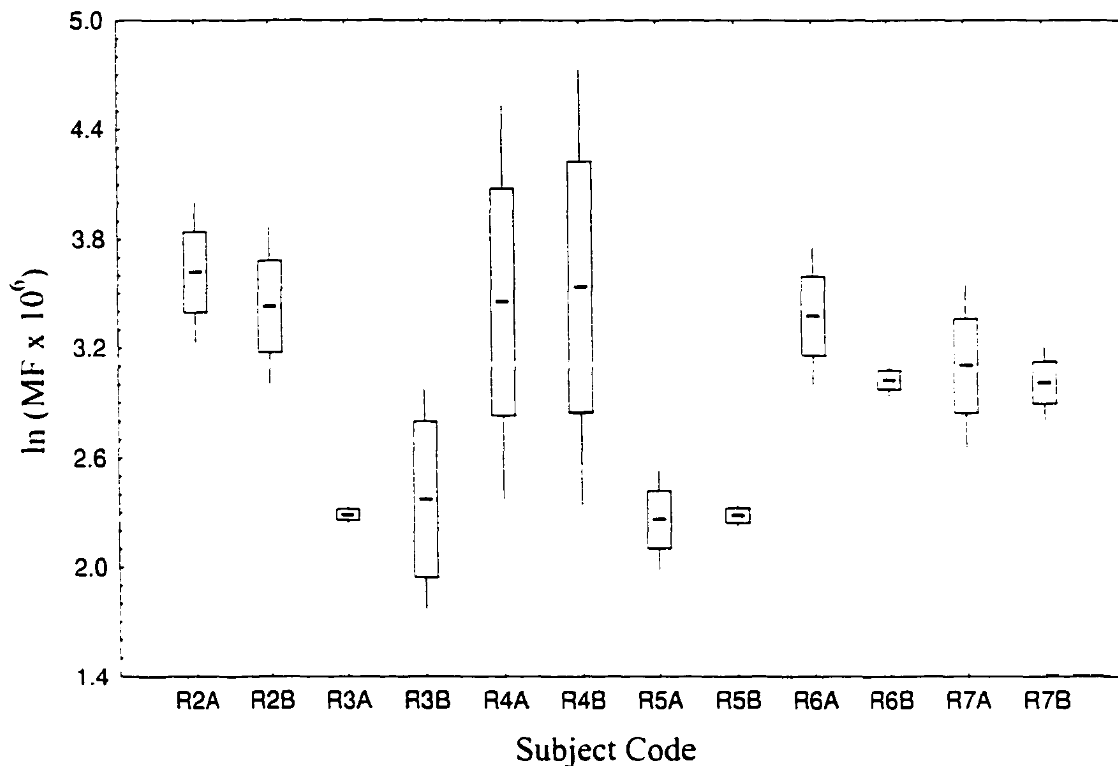


Figure 1. The comparison of the mean lnMF determined for each subject. All 3 MF determinations were obtained and used to provide the means plotted here with standard error (box) and standard deviations (line). Two subjects (R3A and R5B) yielded a zero for one of the 3 MF determinations and thus only two MF determinations were used to provide the mean lnMF.

$\pm$ 4.7 x 10^{-6}, respectively. The average MF for each subject, was determined using the three independent determinations (where possible). These averages were plotted along with standard error and deviations for each subject (Fig. 1). In addition the mutant frequencies for each of the twins were plotted against each other, and the 95% confidence bands for the calculated linear regression line ($R^2 = 0.92$, $p < 10^{-6}$) (Fig. 2). An identical plot of the PE values obtained for the twin sets reveals a similar statistically significant result ($PE^{Twin\ A} = 1.24 + 0.74\ PE^{Twin\ B}$, $R^2 = 0.933$, $p<10^{-6}$). The twin A MF determinations were then randomly matched with MF determinations obtained from the twin B group and the matches plotted as done for Figure 2. After several independent randomly assigned matches, none demonstrated any significance nor the slope of the calculated linear regression line shown (Fig. 2).

© 1997 Wiley-Liss, Inc.

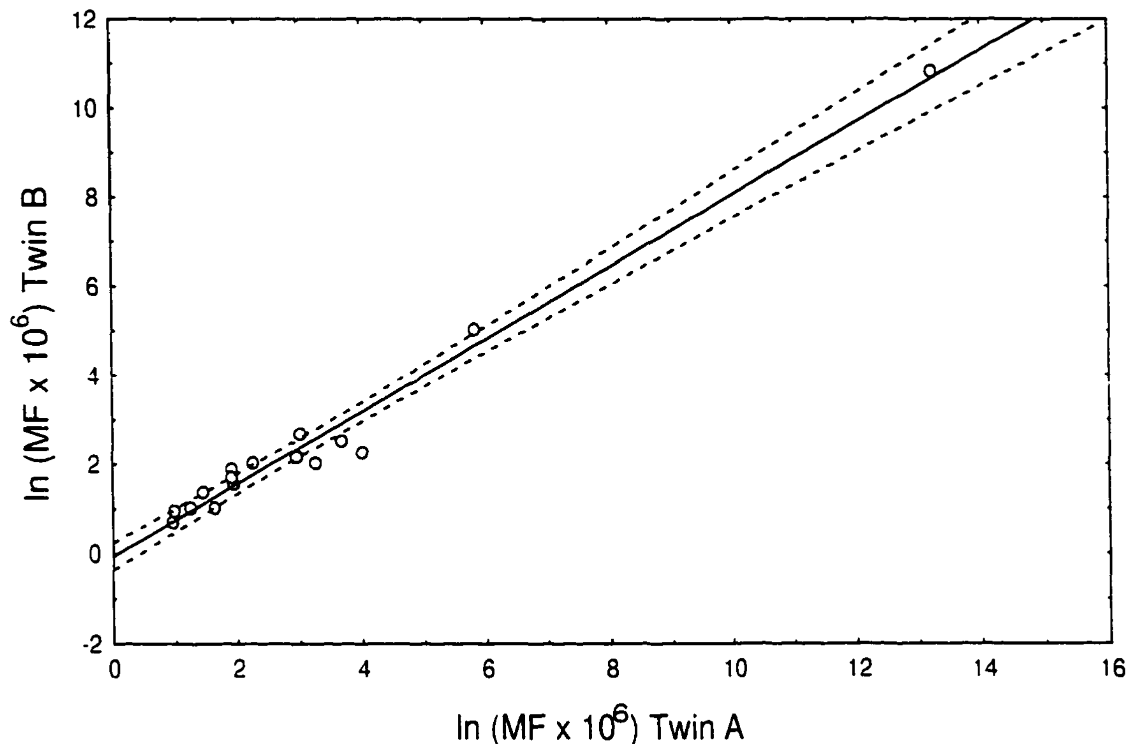


Figure 2. The mutant frequency of half of the subjects versus the mutant frequency of their twin. Twin sets were ordered into groups A or B such that the subject with the highest MF falls into group A. The calculated linear regression line (lnMF Twin B = -0.07 x 0.8 x lnMF TwinA, $p < 10^{-6}$) with 95% confidence bands is shown.

Using the calculations that relate MF and PE with age, given by both Branda *et al.* (1993) and Tates *et al.* (1991), the mean MF for the subjects presented here should be 10.0×10^{-6} and 11.4×10^{-6}, respectively. The mean MF determined for these subjects is approximately three times higher than the calculated means obtained from these two formulas indicating that the Russian subjects have a higher background MF than their Western counterparts. The relationship between PE and lnMF has been previously demonstrated by various laboratories (Cole *et al.*, 1991, Tates *et al.*, 1991; Branda *et al.*, 1993) to be negatively related, such that increased PE results in decreased lnMF values as demonstrated by Branda *et al.* (1993) (lnMF = 2.91 −1.13PE, $R^2 = 0.109$, $p < 0.001$) using a rather large collection of data. A similar relationship for the subjects presented here was observed (lnMF = 3.49 − 2.8PE, $R^2 = 0.18$, $p < 0.014$).

© 1997 Wiley-Liss, Inc.

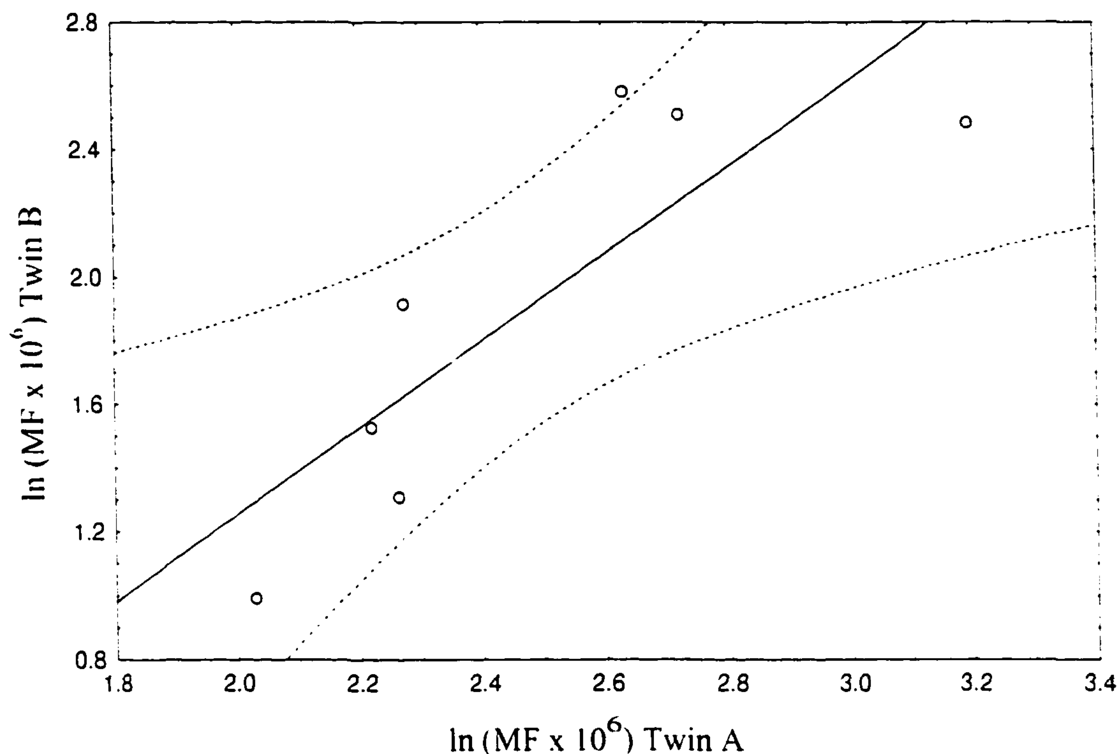


Figure 3. The mutant frequency of monozygotic twins (n=7 sets of twins), with age less than 40 years versus that of their twin (Branda *et al.* 1993). Mean age 32.9 ± 0.86; mean PE 0.46 ± 0.14; mean MF 9.8 ± 6.0 x 10^{-6} Twin sets were ordered into groups A or B such that the subject with the highest MF falls into group A. The linear regression line (p = 0.014) with 95% confidence bands is shown.

Discussion

The observation that the Russian MF were three times higher than that of Western counterparts is striking. The reasons for this increase can only be speculative at this time. One possibility could be related to an increased mutagen burden placed upon residents of this area. Another possibility may reflect differences in diet and lifestyle. Others have also observed Russian subjects demonstrating elevated MF's (Jones *et al.* 1995); however, these subjects had potential to received low doses of radiation during the clean-up of the Chernobyl nuclear accident

Throughout this study, mutant frequencies for each monozygotic twin pair are remarkably similar (Figure 2). This suggests that background mutation at the *hprt* locus could have a significant genetic basis. However, since the early life environmentals are nearly identical for each set of twins, these factors may also contribute to the within pair similarity of mutant frequencies.

© 1997 Wiley-Liss, Inc.

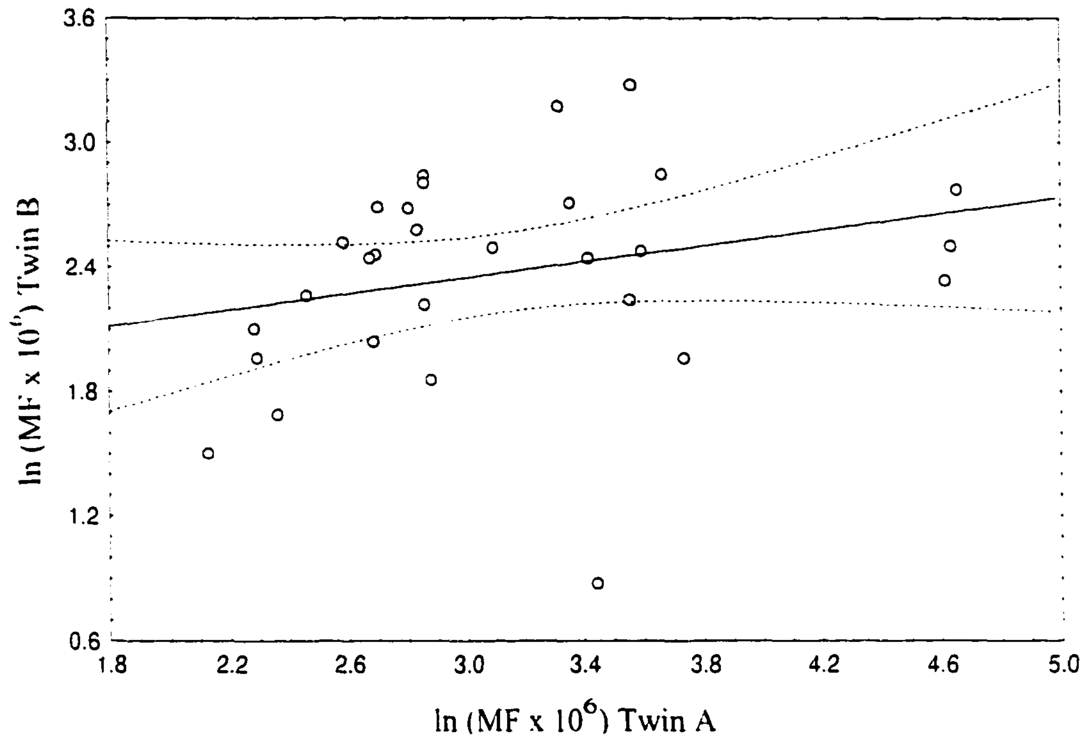


Figure 4. The mutant frequency of monozygotic twins (n=29 sets of twins), with age greater than 40 years) versus that of their siblings (Branda *et al.* 1993). Mean age 66.8 ± 3.8; mean PE 0.36 ± 0.17; mean MF 16.0 ± 9.1 x 10^{-6}. The linear regression line (R^2 = 0.08, p = 0.13) with 95% confidence bands is shown.

A contribution of genotype in background mutant frequency is not unexpected. Factors that contribute to the sensitivity of mutagens, such as the activity of enzymes involved in the metabolism of xenobiotics (Daly *et al.* 1994) and DNA repair, and determinants of chromatin structure, are in part genetically defined. Moreover, these factors are subject to considerable individual variation. The familial component of cancer is well recognized, although the relationship of these factors to carcinogenesis must be much more complex. What remains unclear from this study is the identity of these genetic components and their relative contribution to the frequency of mutation compared to environmental factors. It should be understood that neither the genetic nor the environmental factors can be partitioned from this study. As these twins were reared together and have thus for their entire lives been exposed to identical environments and share identical genetic makeup's, such questions are beyond the scope of the present study. Studies with dizygotic twins

© 1997 Wiley-Liss, Inc.

could help partition the relative contributions of environmental and genetic factors, but the most crucial study would involve monozygotic twins who have been brought up under different environmental circumstances.

One strong suggestion of the interaction between genetic and of environmental factors is revealed in other data. An environmental effect on mutation frequency would be expected to increase with age (Branda *et al.*, 1993; Cole *et al.*, 1994) so that the genetic effect would be most apparent in younger twin pairs. Analysis of additional data kindly provided by Branda *et al.* (1993) confirms this. The twin effect is apparent in twin pairs younger than 40 years of age (Fig. 3) though less so than in that presented here, but the effect has all but disappeared in sets older than 60 years (Fig. 4). It is thus clear that differences in background mutant frequencies increase with age and must result from a complex interaction between genotype and environment. While this study of monozygotic twins may be one of the first to demonstrate a genetic role in the determination of mutation frequencies in people, the interaction between genetic and environmental factors in the generation of mutation is complex, and factors such as diet and lifestyle must play a considerable role. The actual partitioning of genetic and environmental factors in determining mutation frequency will require much more sophisticated experimental designs using a large number of subjects including both monozygotic and dizygotic twins, preferably some of which have been reared separately.

Acknowledgements

This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada and the Medical Research Council of Canada. The authors thank Dr. J.A. Nicklas and the laboratory of Dr. R.J. Albertini for the additional data required to decode the twin pairs from their expansive paper (Branda *et al.* 1993).

© 1997 Wiley-Liss, Inc.

References

Albertini RJ, Castle KL, Borcherding WR (1982) Albertini RJ, Castle KL, Borcherding WR (1982) T-cell cloning to detect the mutant 6-thioguanine-resistant lymphocytes present in human peripheral blood. Proc Natl Acad Sci USA. 79:6617-6621.

Arlett CF, Harcourt SA (1980) Survey of radiosensitivity in a variety of human cell strains. Cancer Res 40:926-932.

Badie C, Iliakis G, Foray N, Alsbeih G, Pantellias GE, Okayasu R, Cheong N, Russell NS, Begg AC, Arlett CF, Malaise EP (1995) Defective repair of DNA double-strand breaks and chromosome damage in fibroblasts from a radiosensitive leukemia patient. Cancer Research 55:1232-1234.

Branda RF, Sullivan LM, O'Neill JP, Falta MT, Nicklas JA, Hirsch B, Vacek PM, Albertini RJ (1993) Measurement of HPRT mutant frequencies in T-lymphocytes from healthy human populations. Mutation Research 285: 267-279.

Caporaso NE, Landi MT (1994) Molecular epidemiology: a new perspective for the study of toxic exposures in man. A consideration of the influence of genetic susceptibility factors on risk in different lung cancer histologies. Med Lav 85:68-77.

Chen PC, Lavin MF, Kidson C, Moss D (1978) Identification of ataxia telangiectasia heterozygotes, a cancer prone population. Nature 274:484-486.

Cole J, Green MHL, James SE, Henderson L, Cole H (1988) A further assessment of factors influencing measurements of thioguanine-resistant mutant frequency in circulating T-lymphocytes. Mutation Research 204: 493-507.

Curry J, Skandalis A, Holcroft J, de Boer J, Glickman BW (1993) Coamplification of *hprt* cDNA and γ T-cell receptor sequences from 6-thioguanine resistant human T-lymphocytes. Mutation Research 288:269-275.

Curry J, Rowley GT, Saddi V, Beare D, Cole J, and Glickman BW. (1995) Determination of hprt mutant and mutation frequencies and the molecular characterization of human derived *in vivo* T-lymphocyte mutants. Environmental and Molecular Mutagenesis 25:169-179.

Daly AK, Cholerton S, Armstrong M, Idle JR (1994) Genotyping for polymorphisms in xenobiotic metabolism as a predictor of disease susceptibility. Env. Health. Perspectives 102:55-61.

Drake JW (1970) The molecular basis of mutation. Holden-Day Inc., San Francisco.

Fuscoe JC, Zimmerman LJ, Lippert MJ, Nicklas JA, O'Neill JP, Albertini RJ (1991) V(D)J recombinase-like activity mediates *hprt* gene deletion in human fetal T-lymphocytes. *Cancer Research* 51:6001-6005.

Gibbs RA, Nguyen P, Edwards A, Civitello AB, Caskey CT (1990) Multiplex DNA deletion detection and exon sequencing of the hypoxanthine phosphoribosyltransferase gene in Lesch-Nyhan families. Genomics 7:235-244.

Hittelman WN, Pandita TK (1994) Possible role of chromatin alteration in the radiosensitivity of ataxia-telangiectasia. Int J Radiat Biol 66:S109-S113.

Jones IM, Thomas CB, Tucker B, Thompson CL, Pleshanov P, Vorobtsova I, Moore II DH (1995) Impact of age and environment on somatic mutation at the hprt gene of T-lymphocytes in human. Mut Research338:129-139

Little JB, Nove J, Strong LC, and Nichols WW (1988) Survival of human diploid skin fibroblasts from normal individuals after X-irradiation. Int J Radiat Biol 54:899-910.

Little JB, Nichols WW, Troilo P, Nagasawa H, Strong LC (1989) Radiation sensitivity of cell strains from families with genetic disorders predisposed to radiation induced cancers. Cancer Res 49:4705-4714.

© 1997 Wiley-Liss, Inc.

Nagasawa H, Little JB (1988) Radiosensitivities of ten apparently normal human diploid fibroblast strains to cell killing, G2-phase chromosomal aberrations, and cell cycle delay. Cancer Research 48:4535-4538.

Paterson MC, Gentner NE, Middlestadt MV, Wienfeld M (1984) Cancer predisposition, carcinogen hypersensitivity, and aberrant DNA metabolism. J Cell Physiol Suppl 3:45-62.

Paterson MC, MacFarlane SJ, Gentner NE, Smith BP *Cellular hypersensitivity to chronic gamma radiation in cultured fibroblasts from AT heterozygotes.* **In** Ataxia telangiectasia: Genetics, Neuropathology and Immunology of a Degenerative Disease of Childhood. (Gatti and Swift Eds) Krioc Foundation Series 19: 73-87 (1985a).

Paterson MC, Gentner NE, Middlesdadt MV, Mizayans R, Weinfeld M (1985b) Hereditary and familial disorders linking cancer proneness with abnormal carcinogen response and faulty DNA metabolism. In Epidemiology and Quantitation of Environmental Risks in Humans from Radiation and Other Agents: Potential and Limitations (Castellani,A. ed) NATO Advanced Study Institute, Plenum Publishing Corporation, New York.

Recio L, Cochrane J, Simpson D, Skopek TR, O'Neill JP, Nicklas JA, Albertini RJ (1990) DNA sequence analysis of *in vivo hprt* mutation in human T-lymphocytes. Mutagenesis 5:505-510.

Rossi AM, Tates AD, van Zeeland AA, Vrieling H (1992) Molecular analysis of mutations affecting *hprt* mRNA splicing in human T-lymphocytes *in vivo*. Env Molec Muta 19:7-13.

Spitz MR (1995) Risk factors and genetic susceptibility. Cancer Treat Res 74:73-87.

Steingrimsdottir H, Rowley G, Dorado G, Cole J, Lehmann AR (1992) Mutations which alter splicing in the human hypoxanthine-guanine phosphoribosyltransferase gene. Nucl Acid Res 20:1201-1208.

© 1997 Wiley-Liss, Inc.

Steingrimsdottir H, Rowley G, Waugh A, Beare D, Cole J, Lehmann AR (1993) Molecular analysis of mutations in the *hprt* gene in circulating lymphocytes from normal and DNA-repair-deficient donors. Mutation Res 294:29-41.

Tates AD, van Dam FJ, van Mossel H, Schoemaker H, Thijssen JCP, Woldring VM, Zwinderman AH, Natarajan AT (1991) Use of the clonal assay for the measurement of frequencies of HPRT mutants in T-lymphocytes from five control populations. Mutation Res 253:199-213.

Yang TP, Patel PI, Chinault AC, Stout JT, Jackson LG, Hildebrand BM, Caskey CT (1984) Molecular evidence for new mutation at the hprt locus in Lesch-Nyhan patients. Nature 310:412-414.

A Comparison of Russian Mutational Spectra with their Western Counterparts

John Curry[1], Magomed Khaidakov[2], and Barry W. Glickman[1]

[1]Centre for Environmental Health and the Department of Biology, University of Victoria, Victoria, British Columbia, Canada

[2]Department of Genetic and Reproductive Toxicology, National Center for Toxicological Research Food and Drug Administration, Jefferson, AZ, USA

Abstract

It has been previously noted that the *hprt* mutant frequency of Russian subjects is significantly higher than age-matched Western counterparts. To further explore this difference, approximately 100 mutants collected from Russian twins reported in a previous study have been sequenced and compared to an aged-matched 'Western' mutant dataset. The mutational spectrum of the Russian subjects was significantly different (Adams and Skopek Monte Carlo test, p = 0.004). Curiously, the Russian spectrum resembles that recovered from older individuals in the West. Specifically, A:T→C:G transversions are significantly over-represented (p = 0.006) in the twin spectrum as compared to the young (age < 35) 'Western' spectrum. Even more noteworthy is the observation that 42% (23/55) of the base substitutions, almost double the expected value, have not been previously reported. These observations lead to the conclusion that this group of young Russian subjects has a mutational pattern that is distinct from the pattern of mutation observed in Western counterparts. The origin of this difference, whether related to genotoxic challenge, repair and avoidance can not be distinguished but clearly diet and other lifestyle factors are likely responsible.

Introduction

Monitoring human populations for the effects of environmental factors has been widely undertaken using the T-cell clonal assay (Albertini *et al.* 1982). In this assay T-cells deficient in hypoxanthine-guanine phosphoribosyltransferase (*hprt*) are selected *ex vivo*. This technique allows researchers to establish mutant frequencies (MF) and subsequently, by means of DNA sequencing, deduce the molecular nature of the mutational events responsible for the loss of the HPRT function. Thus, in populations exposed to

environmental mutagens, *hprt* provides an *in vivo* molecular marker for mutation.

This study involved a set of six male monozygotic twins participating in a Russian Space Agency bed-rest study designed to simulate the physiological effects of weightlessness in space. This simulation involved bed-rest at an inclination (head down) of 7°, eighteen hours a day, for a period of 2 weeks. Although there were no expectations that this regimen would effect mutant frequency or spectra, the Moscow study did provide access to healthy, well monitored, monozygotic twins. The individual sets of twins (mean age = 30) had always lived together; and the entire group, two weeks prior to and through out the bed-rest protocol, were housed and fed together.

Mutant frequency results of the T-cell clonal assays demonstrate a surprising similarity between the twin pairs (Curry *et al.* 1997). An examination of another large set of MF values obtained from monozygotic twins (Branda *et al.* 1993) again revealed a remarkable similar correlation of MF between the twin pairs (Curry *et al.* 1997). This apparent twin effect suggests a strong genetic component influences mutant frequency at the *hprt* locus. However, when older twins (greater than 40 years of age) were similarly analyzed, the twin effect on MF was no longer apparent, suggesting a dominant environmental component. As previously reported, this group of Russian twins has a three-fold increase in MF when compared to an age-corrected value based on data from Western counterparts (Curry *et al.* 1997). An unusually high MF value have been previously reported for Russian populations (Khaidakov *et al.* 1997, Jones *et al.* 1995) and potentially suggests a strong environmental component acting on these populations.

To investigate potential environmental factors affecting Russian populations, this study compares mutational spectra obtained from a group of Russian twins with a spectra obtained primarily from Western sources. A refinement of the human HPRT database (Cariello 1994), used for a previous study (Curry *et al.* 1999), serves as a comparison spectrum for the twins spectrum presented here.

Methods

Reverse Transcription and PCR of the Mutant *hprt* Gene

Mutants were collected using the T-cell clonal assay (Albertini *et al.* 1982) with modifications as outlined elsewhere (Curry *et al.* 1997). Cell pellets containing approximately 2000 mutant T-cell clones were prepared and frozen at -20°C until required. Cell pellets were thawed and 5µl of a reverse transcriptase cocktail (Yang *et al.* 1989) containing 50mM Tris-HCl, pH=8.3; 75mM KCl; 3.6mM $MgCl_2$ (Sigma); 10mM DTT; 20ng/µl BSA; 2.3U RNAGuard (Pharmacia); 2.5 % NP-40; reverse transcriptase (Gibco/BRL); 4.8 µM Primer 2 (Curry *et al.* 1993) (Dalton Chemicals, Toronto); and 2mM/each dNTP (Pharmacia) was added, and incubated. After one hour at 37°C, 40µl of the first PCR reaction cocktail (60mM KCl; 15mM Tris-HCl, pH=7.6; 2.75mM MgCl2; 0.625mM each dNTP; 6.0µM of primers P1 and P2 (Curry *et al.* 1993); 2.5U/Amplitaq (Perkin Elmer)) was added to the reverse transcriptase reaction. The reaction mixtures were placed in a thermocycler programmed for an initial denaturation step of 95°C for 5 minutes followed by 5 cycles of a slow program (94°C - 2', 60°C - 2', 72°C - 5'). The next 25 cycles utilized a fast program (94°C - 20", 60°C - 30", 72°C - 1') after which the samples were held at 4°C. The second PCR reaction was identical except that it contained 1µl of the first reaction in a total volume of 25µl along with nested primers P3 and P4 (Curry *et al.* 1993). These secondary reactions were then thermocycled for 30 rounds using the fast program. Approximately 10% of this reaction was run on a 1% agarose gel and those reactions displaying distinct PCR products were repeated using 2µl of the second PCR in a total 100µl reaction volume. Samples were pooled and the product separated on 1% low melting point gel (NuSieve), the bands cut out, and the DNA released by digestion of the agarose with beta-agarase (New England Bio). The DNA was precipitated with 95% ethanol, washed with 70% ethanol and resuspended in water.

Automated DNA Sequencing

Mutant cDNAs were sequenced on a Pharmacia A.L.F. DNA sequencer following protocols described by the manufacturer with the modifications outlined below. The concentration

of the template was measured using a Hoeffer Fluorometer. For each sequencing reaction, 2.5 pmol of fluorescently labeled primer was added to 100ng of template. Primers were prepared by Dalton Chemicals (Toronto) using Pharmacia Fluoresein Amidite. Sequencing reactions were carried out with Pharmacia Autocycle Sequencing kits except that the supplied enzyme was replaced with *Taq* polymerase. The reactions were thermocycled; 94°C for 20 seconds, 50°C for 10 seconds, one-minute extension at 72°C, preceded by an initial 2-minute denaturation step. Stop solution was added and the sample denatured at 94°C for 2 minutes prior to loading 10μl on the A.L.F. sequencer. Five sequencing primers (Karnaoukhova *et al.* 1997) were used to sequence the complete *hprt* cDNA.

To ensure the accuracy the DNA sequencing a set of 10 randomly selected mutants, which had been successfully characterized, were coded (blind) and again subjected to the characterization process. Mutants from a single subject that demonstrated the same sequence changes were further examined for clonality by determining their γ-TCR rearrangements (Curry *et al.* 1995).

Analysis of Mutant Spectra

A spectra obtained from the Russian twins was compared with the Western spectra obtained from a dataset assembled for another study (Curry *et al.* 1999). Comparison of the smoking and non-smoking spectra did not reveal any significant differences (Curry *et al.* 1999). That study did reveal a potential difference in the frequency of deletions between males and females. For this reason, the comparison Western dataset excludes mutants from females subjects.

In the mutant dataset, mutations that result in aberrant exon splicing were reported by the presumed causative mutation. Exon splice mutants recovered from the Russian subjects were not further characterized and are identified only by the exon-splicing event. The mutant dataset was re-coded such that characterized splice mutations are only reported by the actual exon-splicing event. This refinement permits a non-biased comparison of the two spectrum.

Spectra were compared using the Adams and Skopek's Monte Carlo test (MC) (Adams and

Skopek 1987). In cases where the Monte Carlo test demonstrated significance, a Fisher's exact test was used to further evaluate the differences between single mutational classes. Following the conservative Bonferroni procedure, the level of significance was lowered (p=0.05/(number of two-way comparisons) when three way or greater comparisons were performed (Wassertheil-Smoller 1995). For example, in an analysis comparing 3 spectra, only p values less than or equal to 0.017 will be considered significant.

Results

The mean plating efficiency (PE) and mutant frequency (MF) in the twins were 15.3 ±17 % and 27.5 ±4.7 x 10^{-6}, respectively (Curry *et al.* 1997). Their ages ranged from 24 to 35 with a mean value of 30.0 and a standard deviation (SD) of 4.75 (Curry *et al.* 1997). The sequences of the mutations characterized are listed in Table 1. The ten randomly picked mutants for which cDNA was isolated and sequenced twice, demonstrated the same mutations upon re-sequencing and thus reaffirm the accuracy of the sequencing methodology. A single subject (R5A1) demonstrated two identical mutations that were consequently further characterized at the γ-TCR locus. This revealed that the mutants were clonally related and thus only a single mutant was retained in the dataset.

Mutant R6A3-2 (Table 1b) was classified as a splice mutant as the 5' insertion point falls precisely along the intron/exon boundary. This mutant, however, differs from other intron inclusion splicing errors. The 5' inserted intron (Intron V, 31636 – 31702) sequence has 5 novel bases included along with the intron sequence (see Table 1b). At the 3' end, the inclusion of intron V ends at position 31703 (...ATCCTAAAGG-31703), and these last five intron bases are a direct repeat of the beginning of exon VI (TTG...AAAGG...ATA). Thus, the precise 3' insertion point of intron cannot be precisely identified. The breakpoint is also bounded by an inverted repeat. The 3' end of exon V is 5'-TGGAA-3', and 3' end of the inserted intron V (plus T base) is 5'-AAGGT-3'. There are 15 reports of this type of intron inclusion in Release 3 of the HPRT mutant database (Cariello 1994) all of which begin at precisely the end of exon V. Nine of these insertions end as shown for the mutant described

Mutant	cDNA Position	Mutation	Amino Acid Change	Sequence Context (Codon in Bold)
R2A1-3	1	A:T→C:G	MET→Leu	CCGTT...A...TGGCG
R2B1-101	2	A:T→G:C	MET→Thr	CGTTA...T...GGCGA
R2A3-30	2	A:T→C:G	MET→Arg	CGTTA...T...GGCGA
R6B1-11	3	G:C→A:T	MET→Ile	GTTAT...G...GCGAC
R2B2-21	8	G:C→A:T	Thr→Ile	GGCGA...C...CCGCA
R3B2-206	14	G:C→A:T	Ser→Asn	CCGCA...G...CCCTG
R2B1-38	47	G:C→T:A	Gly→Val	ACCAG...G...TTATG
R4B2-4	82	A:T→C:G	Tyr→Asp	ATCAT...T...ATGCT
R4A2-15	84	A:T→C:G	Tyr→STOP	CATTA...T...GCTGA
R4A2-16	84	A:T→C:G	Tyr→STOP	CATTA...T...GCTGA
R6B3-10	104	A:T→T:A	Val→Glu	AAGGG...T...GTTTA
R2A1-7	119	G:C→C:G	Gly→Ala	TCATG...G...ACTAA
R3B2-207	119	G:C→C:G	Gly→Ala	TCATG...G...ACTAA
R2B3-53	135	G:C→C:G	Arg→Ser	GACAG...G...ACTGA
R6B1-8	135	G:C→T:A	Arg→Ser	GACAG...G...ACTGA
R7B2-1	151	G:C→A:T	Arg→STOP	TTGCT...C...GAGAT
R3B2-205	190	G:C→A:T	Ala→Thr	TTGTA...G...CCCTC
R5B3-4	203	A:T→C:G	Leu→Arg	TGTGC...T...CAAGG
R4A1-402	203	A:T→C:G	Leu→Arg	TGTGC...T...CAAGG
R3A1-213	205	A:T→C:G	Lys→Gln	TGCTC...A...AGGGG
R3B2-346	211	G:C→T:A	Gly→Cys	AGGGG...G...GCTAT
R2B3-56	218	A:T→C:G	Lys→Thr	CTATA...A...ATTCT
R2A1-5	233	A:T→G:C	Leu→Pro	TGACC...T...GCTGG
R2B1-17	246	G:C→C:G	Ile→MET	TACAT...C...AAAGC
R2B2-20	254	A:T→G:C	Leu→Pro	AGCAC...T...GAATA
R3A2-210	254	A:T→G:C	Leu→Pro	AGCAC...T...GAATA
R3A1-202	260	G:C→A:T	Arg→Lys	GAATA...G...AAATA
R4A1-401	266	G:C→C:G	Ser→Thr	AAATA...G...TGATA
R2B2-19	292	G:C→A:T	Asp→Asn	CTGTA...G...ATTTT
R4A1-405	302	G:C→A:T	Arg→Lys	TATCA...G...ACTGA
R5B3-3	325	G:C→A:T	Gln→STOP	ATGAC...C...AGTCA
R6A1-2	325	G:C→A:T	Gln→STOP	ATGAC...C...AGTCA
R4B1-411	398	A:T→G:C	Val→Ala	GATTG...T...GGAAG
R4B1-413	413	A:T→G:C	Asp→Gly	AATTG...A...CACTG
R5A1-2, 3[1]	418	G:C→C:G	Gly→Arg	ACACT...G...GCAAA
R3A1-209	424	A:T→C:G	Thr→Pro	GCAAA...A...CAATG
R7B3-6	440	A:T→G:C	Leu→Pro	TTTGC...T...TTCCT
R2B3-60	449	A:T→T:A	Val→Asp	CTTGG...T...CAGGC
R2A2-35	463	G:C→C:G	Pro→Ala	ATAAT...C...CAAAG
R4A1-403	481	G:C→A:T	Ala→Thr	AGGTC...G...CAAGC

[1] Mutant R5A1-2 and R5A1-3 are clonally related.

R3B1-204	482	G:C→C:G	Ala→Gly	GGTCG...C...AAGCT
R6B1-1	488	A:T→C:G	Leu→Trp	AAGCT...T...GCTGG
R4A3-3	530	A:T→G:C	Asp→Gly	GCCAG...A...CTTTG
R7A3-8	530	A:T→G:C	Asp→Gly	GCCAG...A...CTTTG
R5B2-2	533	A:T→C:G	Phe→Cys	AGACT...T...TGTTG
R6A3-6	569	G:C→A:T	Gly→Glu	TGTAG...G...ATATG
R4B3-7	569	G:C→C:G	Gly→Ala	TGTAG...G...ATATG
R3B1-203	605	A:T→G:C	Leu→Ser	GGATT...T...GAATC
R4A2-7	610	G:C→C:G	His→Asp	TGAAT...C...ATGTT
R4B2-5	611	A:T→G:C	His→Arg	GAATC...A...TGTTT
R2A2-81	613	G:C→T:A	Val→Phe	ATCAT...G...TTTGT
R2A3-8	617	G:C→A:T	Cys→Tyr	TGTTT...G...TGTCA
R6B2-3	617	G:C→A:T	Cys→Tyr	TGTTT...G...TGTCA
R4B1-410	617	G:C→T:A	Cys→Phe	TGTTT...G...TGTCA
R7A1-6	626	G:C→T:A	Ser→Ile	CATTA...G...TGAAA
R4B3-1	459, 538	COMPLEX	Tyr&Gly→STOP	CAGTA...T...AATCC , TTGTT...G...GATTT
R7A1-1	493	COMPLEX		GTGAAAAGGA→GGGAAAGGG
R2A3-28	21-22	Deletion		CCTGG...[CG]...TCGTG
R2B1-45	533-553	Deletion		AGACT...[TTGTTG...ATTCCAG]...ACAAG
R7A2-3	162-648	Deletion		GTGAT...[GAAGGA...AAATAC]...AAAGC
R7B3-11	110-125	Deletion		TGTTTA...[TTCCTCATGGACTAAT]...TATGG
R7B2-3	443-446	Deletion		GCTTT...[CCTT]...GGTCA
R7A1-7	INSERTION	610-627		TGAAT...[AATATACTTTTGATAG]...CATGT
R4A3-8	507	_+C	+ Frameshift	ACCCC...[+C]...ACGAA
R5B2-1	629-631	_-A	- Frameshift	TAGTG...[-A]...AACTG
R6A3-10	230	_-A	- Frameshift	TGCTG...[-A]...CCTGC
R5A3-8	275 or 276	_-C	- Frameshift	TAGAT...[-C]...CATTC
R6A1-8	632	_-C	- Frameshift	TGAAA...[-C]...TGGAA
R7B3-10	275 or 276	_-C	- Frameshift	TAGAT...[-C]...CATTC
R7A2-2	447 or 448	_-G	- Frameshift	CCTTG...[-G]...TCAGG
R7B1-9	51	_-T	- Frameshift	GGTTA...[-T]...GACCT

Table 1a. Mutations observed in the cDNA of 6-TG resistant T-lymphocytes clones. The numbering of the cDNA positions is relative to the beginning (position 1) of the coding region of *hprt*.

Mutant		cDNA Change	
R2A2-33	Skipped	Exon II & III	
R4A3-1	Skipped	Exon II & III	
R4B2-1	Skipped	Exon II & III	
R5A3-1	Skipped	Exon II & III	
R5A3-4	Skipped	Exon II & III	
R6B1-9	Skipped	Exon II & III	
R6B3-5	Skipped	Exon II & III	
R7A3-14	Skipped	Exon II & III	
R7B1-1	Skipped	Exon II & III	
R7B1-3	Skipped	Exon II & III	
R4B3-5	Skipped	Exon IV	
R5B3-2	Skipped	Exon IV	
R5B3-6	Skipped	Exon IV	
R4A2-9	Skipped	Exon V	
R7A3-7	Skipped	Exon V	
R2A1-4	Skipped	Exon V & VI	
R3A1-201	Skipped	Exon VI	
R7A1-3	Skipped	Exon VI	
R2A2-36	Skipped	Exon VII	
R4A2-12	Skipped	Exon VII	
R2B3-59	Skipped	Exon VIII	
R4B2-2	Skipped	Exon VIII	
R4B2-7	Skipped	Exon VIII	
R7A3-5	Skipped	Exon VIII	
R2B2-18	28-122	Part Exon II	Partial Exon Skipping
R6A2-1	610-626	Part Exon IX	Skipped first 17 bps of Exon IX
R6A3-2	Inclusion	Part Intron V	TGGAA..(gggGaTAAgGTT.....CTAAAG)..GATAT

Table 1b. Mutants that exhibit aberrant cDNA splicing.

here, while the others end at sequences proximal to this end point. None of these reported intron inclusion splice mutants contains additional novel sequences as reported for mutant R6A3-2.

Another unusually complex mutant (R7A1-6) was observed in the twin spectra. In this mutant a sequence (GTGAAAGGA) unique to the genomic *hprt* sequence is replaced by a shorter sequence (GGGAAAGGG). The replacement sequence is found once in the genomic *hprt* sequence at position 1361-1369. A variation of the sequence (minus a 5' G nucleotide) is found at four separate sites in the genomic sequence. An examination of these four sites does not reveal any direct or inverted repeated sequences shared with the original sequence.

A comparison of the spectrum of mutation obtained from the Russian subjects with that of 'Western' males was performed. A significant difference was observed when all classes of mutation were used in the comparison (MC,

Mutation Class	Russian Twins	Males Western Dataset	Males Western Ages ≤ 35	Males Western Ages > 35
G:C->A:T	14 14.3%	111 16.9%	41 18.1%	70 16.6%
A:T->G:C	11 11.2%	34 5.2%	3 1.3%	21 5.0%
TRANSITIONS	25 25.5%	145 22.1%	44 19.5%	91 21.6%
G:C->T:A	6 6.1%	26 4.0%	12 5.3%	14 3.3%
G:C->C:G	10 10.2%	36 5.5%	11 4.9%	25 5.9%
A:T->T:A	2 2.0%	44 6.7%	17 7.5%	27 6.4%
A:T->C:G	12 12.2%	34 5.2%	7 3.1%	27 6.4%
TRANSVERSIONS	30 30.6%	140 21.3%	47 20.8%	93 22.1%
BASE SUBSTITUTIONS	**55 56.1%**	**285 43.4%**	**91 40.3%**	**184 43.7%**
Splice Mutations	27 27.6%	241 36.7%	79 35.0%	162 38.5%
Deletion	5 5.1%	46 7.0%	16 7.1%	30 7.1%
Frameshifts -1	7 7.1%	50 7.6%	25 11.1%	25 5.9%
Frameshifts +1	1 1.0%	10 1.5%	4 1.8%	6 1.4%
Complex	2 2.0%	14 2.1%	4 1.8%	10 2.4%
Tandem	0 0.0%	6 0.9%	3 1.3%	3 0.7%
Insertions	1 1.0%	5 0.8%	4 1.8%	1 0.2%
TOTAL MUTANTS	**98 100.0%**	**657 100.0%**	**226 100.0%**	**421 100.0%**

Table 2. A comparison of the spectrum of mutations observed in the Russian subjects compared with three other spectra obtained from various sources which are compiled into a refined *hprt* mutant data base (Curry *et al.* 1999).

p=0.028) (Table 2). When the comparisons were limited to just single base-pair substitutions (sbps), the Monte Carlo analysis yielded a p value of 0.015. The comparison of the remaining classes of events provided no indication of significance. The 'Western' mutant spectrum was then split into two spectra, subjects aged 35 or less and subjects aged 36 or greater (Table 2). Comparison of these two spectra with that of the twin spectrum revealed the spectrum from younger subjects to be considerably more different (MC, p = 0.01) than the spectrum from older subjects (MC, p = 0.06). Repeating the same comparisons, restricted to just sbps, reinforced the difference between the twins and the spontaneous spectrum from the younger individuals (MC, p = 0.004). No significant difference was seen during the comparison of the older spectrum (MC, p = 0.05, for a three way comparison significance is at 0.017). No significant difference was observed between the twins and the younger or older spectra when the other classes of mutation were analyzed. Comparison of the age-based spectra reveals that they are significantly different (Curry *et al.* 1999). It was previously observed that this difference reflects the low number of A:T→C:G transversions, occurring in spectrum from young individuals.

Single classes of mutation were then tested using Fisher's exact test. Comparison of the twin and spontaneous male spectra revealed that only one specific class of mutation is significantly different between the two (A:T→C:G, p = 0.012). The frequency of A:T→C:G transversions is significantly over-represented in the twin spectrum as compared to the spontaneous male spectrum. Further investigation of this over-representation was analyzed by again dividing the 'Western' dataset into younger (age < 36) and older (age ≥ 36) groups. Repeating the Fishers exact test for each class again revealed that the A:T→C:G class is significantly over represented in the twin spectrum when compared to the younger (p = 0.003) but not the older group (p = 0.06). Comparing the A:T→C:G class for all three groups (twins, young, old) clearly demonstrated a significant difference (p = 0.006). It is solely these A:T→C:G transversions that drive the Monte Carlo tests into significance. As such the twins mutational spectrum is more similar to the

spectrum of older 'Western' individuals than younger.

Exon skipping events were frequently observed in the twin mutational spectra (27.6%, 28/98). In comparison to the overall frequency of exon splicing mutations for males in the mutant database (36.8%, 242/657) the twins had a lower frequency, but there was no statistical difference between the two frequencies. There was also no statistical difference between the exon splice mutation frequencies for the males under the age of 36 or above and the twins splice frequency. Although the number of twin splice mutants was small compared to the number for the males in the mutant database, a comparison was attempted. Splice mutants were ordered into discrete categories, by those involving single, multiple and partial exon losses, as well as intron inclusions. Using the Monte Carlo test revealed no differences between the twins and the mutant data base males, or the twins and the males below or above the age of 30 years. Comparing the mean sizes of the five twin deletions (113.6 ± SD 230.1 bps) to that of the 59 male deletions (33.7 ± SD 83.5 bps) reported in the spontaneous database revealed no significant difference.

Comparing the twin sbps by position (along the coding sequence) to the positions of the 284 sbps contained in the 'Western' *hprt* dataset, revealed that 41.8% (23/55) of the twin mutations have not been previously observed. Comparatively only 24.3 % (69/284) of the sbps are novel (positions reported only once) in the Western male dataset.

Discussion

The mean MF for these Russian subjects was previously noted to be approximately three times higher than that of Western age matched counterparts (Curry *et al.* 1997). Here we report that Monte Carlo analysis of the Russian twin mutant and spontaneous male spectra reveals a significant difference. Dividing the spontaneous dataset into a younger and older group revealed the twin spectrum to be more similar to the older group (above 35 years of age). A single class of transversions (A:T→C:G) was solely responsible for this difference. Previously we had noted that the difference between mutational spectra derived from young individuals (age <= 36) and that of older individuals (age > 36) was due to the low frequency of this class of events in the

younger spectra (Curry *et al.* 1999). In the Russian spectrum, this class of mutation is clearly over-represented as compared to the younger Western spectrum. The mutational pattern observed for the Russian twins differs significantly from that of Western counterparts. Further owing to the differences observed in the Russian twin spectrum compared to their Western counterparts, is the fact that 42% of the sbps by position are novel, having not been previously reported in the literature.

To further investigate the relatively high numbers of novel sbps positions being mutated, a comparison was made with another Russian mutational spectra. A total of 86 mutations were collected from Russian cosmonauts involved in long term flight missions aboard the space station Mir (Khaidakov *et al.* 1997). Although a relatively low number of mutants to compare to those in the mutant database, an unusually high number of positions in this spectra are novel (66.7%, 12/18) as compared to the 284 sbps mutants in the 'Western' dataset (Fishers, p=0003). The significantly higher frequency of novel positions being mutated in these two Russian spectra may be indicative of mutational mechanisms or perhaps environmental exposures unique to these Russian subjects.

While the frequency of A:T→C:G transversions is significantly increased in the Russian twin spectra as compared to the age matched male spectra and is remarkably similar to the older (age >35 years) male spectra, this situation is surprisingly reversed for the Cosmonaut data. A:T→C:G transversions are considerably different ($p = 0.04$) when comparing only the sbps events between the Cosmonauts (3.8% 1/26) and the twin (21.8% 12/55) spectra. When contrasting these Cosmonaut transversions (3.8%, 1/26) to the age matched male spectra (14.8%, 27/184) this class is considerably lower, and as such is more similar to the younger male spectra (6.9%, 7/101). Fishers exact test could not be preformed as only a single A:T→C:G transversion was recovered for the Cosmonaut spectra. Significance testing of the percentages could also not be reliably performed for the same reason. As the numbers of mutants in these two distinct spectra are dissimilar, sample size may be a possible factor influencing the frequency of this

transversion class, particularly in the smaller Cosmonaut spectrum (n=26).

During the analysis of "Western' spectra for the effects of age on mutation we had previously noted that A:T→C:G transversions were significantly related to subject age (Curry *et al.* 1999). These transversions were noted to occur at low frequency in the spectrum of young individuals (age < 37) but increased in frequency in the spectrum of older individuals. Several mutagens are noted to induce A:T→C:G transversions which are normally rare in mutational spectra: Benzo(*a*)pyrene (Wei *et al.* 1993), nitrosomethylaniline (Zielenska and Guttenplan 1988), hydroxylaminopurines (Murray 1987), and N-ethyl-N-nitrosourea (Guttenplan 1990). DNA polymerase alpha operating on oxygen radical damaged DNA templates has been related to a 14-fold increase in the frequency of A→C transversions (Feig and Loeb 1994). The significant increase of this class of transversion in the twin spectra may be linked to the aforementioned or similarly operating mutagens.

Potential differences in lifestyle and diet may also contribute to the unique mutational load that challenges Russians. Higher rates of tobacco and alcohol consumption, poor nutrition, and a reduced health care system are some of the major differences between the Russian and the 'Western' subjects (Notzon *et al.* 1998). In addition, the 'Western' subjects lived in areas with more stringent environmental pollution controls than the Russians. Recently the life expectancy of Russians has declined and is now 15 years less than that of 'Western' males (Notzon *et al.* 1998). As the mutational spectra of the young Russian twins is much more similar to that of older Western counterparts it is tempting to suggest that Russians more rapidly develop the mutational patterns of older individuals. However, as the relationship between mutation accumulation and health is unknown, and the interpretation of this data must be revisited. In addition, the current Russian mutational spectrum is incomplete although the high number of novel positions found in the Russian spectrum, does suggest that the spectrum of Russian mutations is clearly a different pattern than in the Western male counterparts.

Acknowledgements

This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada and the Medical Research Council of Canada. The assistance of the Canadian Space Agency is gratefully acknowledged.

References

Adama WT, Skopek TR. 1987. Statistical test for the comparision of samples from mutational spectra. J. Mol. Biol. 194:391-396.

Albertini RJ, Castle KL, Borcherding WR. 1982. T-cell cloning to detect the mutant 6-thioguanine-resistant lymphocytes present in human peripheral blood. Proc. Natl. Acad. Sci. USA 79:6617-6621.

Branda RF, Sullivan LM, O'Neill JP, Falta MT, Nicklas JA, Hirsch B, Vacek PM, Albertini RJ. 1993. Measurement of HPRT mutant frequencies in T-lymphocytes from healthy human populations. Mutation Research 285:267-279.

Cariello NF. 1994. Database and software for the analysis of mutations at the human hprt gene. Nucleic Acids Res 22:3547-3548.

Curry J, Bebb G, Young D, Khaidakov M, Mortimer A, Glickman BW. 1997. Similar mutant frequencies observed between pairs of monozygotic twins. Human Mutation 9:445-451.

Curry J, Skandalis A, Holcroft J, de Boer J, Glickman BW. 1993. Coamplification of *hprt* cDNA and γ T-cell receptor sequences from 6-thioguanine resistant human T-lymphocytes. Mutation Res. 288:269-275.

Curry J, Rowley GT, Saddi V, Beare D, Cole J, Glickman BW. 1995. Determination of *hprt* mutant and mutation frequencies and the molecular characterization of human derived *in vivo* T-lymphocyte mutants. Environ. Mol. Mutagen. 25:169-179.

Curry J, Karnaoukhova L, Guenette GC, Glickman BW. 1999. Influence of sex, smoking, and age on human *hprt* mutation frequencies and spectra. In Press: Genetics.

Guttenplan JB. 1990. Mutagenesis by N-nitroso compounds: relationships to DNA adducts, DNA repair, and mutational efficiencies. Mutation Res. 233:177-187.

Feig, DI, Loeb LA. 1994. Oxygen radical induced mutagenesis is DNA polymerase specific. J. Mol. Biol. 235:33-41.

Jones IM, Thomas CB, Tucker B, Thompson CL, Pleshanov P, Vorobtsova I, Moore II DH. 1995. Impact of age and environment on somatic mutation at the *hprt* gene of T-lymphocytes in human. Mutation Res. 338:129-139

Karnaoukhova L, Moffat J, Martins H, Glickman BW. 1997. Mutation frequency and spectrum in lymphocytes of small cell lung cancer patiens receiving etoposide chemotherapy. Cancer Res. 57:4393-4407.

Khaidakov M, Young D, Ertle H, Mortimer A, Voronkov Y, Glickman BW. 1997. Molecular analysis of mutations in T-lymphocytes from experienced Soviet cosmonauts. Env. Molec. Mutagenesis. 30:21-30.

Murray V. 1987. Transversion-specific purine analogue mutagens and the mechanism of hydroxylaminepurine mutagenesis. Mutation Res. 177:189-199.

Notzon FC, Komarov YM, Ermakov SP, Sempos CT, Marks JS, Sempos EV. 1998. Causes of declining life expectancy in Russia. JAMA 279:793-800.

Wassertheil-Smoller S. 1995. *Biostatistics and epidemiology. A primer for health professionals.* New York: Springer-Verlag, 70 pp.

Wei SJ. Chang RL. Bhachech N. Cui XX. Merkler KA. Wong CQ. Yagi H. Jerina DM. Conney AH. 1993. Dose-dependent differences in the profile of mutations induced by (+)-7R,8S-dihydroxy-9S,10R-epoxy-7,8,9,10-tetrahydrobenzo(a)pyrene in the coding region of the hypoxanthine (guanine) phosphoribosyltransferase gene in Chinese hamster V-79 cells. Cancer Res. 53:3294-301

Yang J L. Chen RH. Maher VM. Mccormick JJ. 1991. Kinds and location of mutations induced by (+-)-7 beta,8 alpha-dihydroxy-9 alpha,10 alpha-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.

Zielenska M. Guttenplan JB. 1988. Mutagenic activity and specificity of N-nitrosomethylaniline and N-nitrosodiphenylamine in *Samonella*. Mutation Res. 202: 269-276.

Influence of Sex, Smoking, and Age on Human *hprt* Mutation Frequencies and Spectra

John Curry, Larissa Karnaoukhova, Gabriel C. Guenette and Barry W. Glickman

Centre for Environmental Health and the Department of Biology, University of Victoria, Victoria, British Columbia

Abstract

Examination of the literature for *hprt* mutant frequencies from peripheral T-cells yielded data from 1194 human subjects. Relationships between mutant frequency, age, sex and smoking were examined and the kinetics described. Mutant frequency increases rapidly with age until about age 15. Subsequently the rate of increase falls such that after age 53 the *hprt* mutant frequency is largely stabilized. Sex had no effect on mutant frequency. Cigarette smoking increased mean mutant frequency compared to non-smokers but did not alter age versus mutant frequency relationships. An *hprt in vivo* mutant database was prepared containing 795 human *hprt* mutants, from 342 individuals. No difference in mutational spectra was observed comparing smokers to non-smokers, confirming previous reports. Sex affected the frequency of deletions (>1 bp) which are recovered more than twice as frequently in females ($p = 0.008$) compared to males. There is no indication of a significant shift in mutational spectra with age for individuals greater than 19 years of age with the exception of A:T→C:G transversions. These events are more frequently recovered in older individuals.

Introduction

Aging occupies a central position in health concerns. While the causes of aging are undoubtedly complex, it has been suggested that somatic mutations play a central role in the process (review, Morley 1995). Recent models of aging suggest a role for free radicals (Harman 1992, Gutteridge 1992, Slagboom and Vijg 1989) and the diminishing efficiency of DNA repair with age (review, Holmes *et al.* 1992). The DNA damage and repair theory of aging assumes that the total number of DNA lesions increases as a function of time, and that DNA repair is unlikely to accurately correct them. In addition, the model asserts that mechanisms of DNA surveillance likely decrease with age

(Warner and Price 1989). The catastrophic model is perhaps the most extreme. This model suggests mutations accumulate with age to the point where the accuracy of molecular processes progressively deteriorates, and at the end of our lives mutation rates escalate dramatically (Holmes *et al.* 1992).

Another potential mechanism of aging involves DNA methylation (Mazin 1994). Deamination of 5-methylcytosine (m^5C) to thymine occurring at CpG sites results in the mispairing of thymine with guanine. The result is replication mediated C/G→T/A transitions. Approximately 7 x 10^6 C→T transitions per human genome per division occur (Mazin 1994). Numbers of this type of mutation accumulate with cell division, and consequently with age. Recently O'Neill and Finette (1998) reported that mutations at CpG sites in newborns (cord blood) represent 33% of the single base substitutions (6/18) while in adult populations such mutations represent only 5.1% (13/253). Their data indicate a dramatic age-related shift in mutational specificity.

More than 150 human genetic diseases have been characterized by Martin (1997) as in some way related to the normal process of human aging. However, as each of them differs significantly from "normal aging", Martin (1997) concluded "that there is no single gene that accelerates or decelerates the normal aging process". In other organisms, genes seemingly related to aging have been identified. An age-1 mutation in *C. elegans* is observed to increase life span (Johnson *et al.* 1993). This mutation increases the levels of superoxidase dismutase and catalase in an age-dependent manner (Larsen *et al.* 1993). When either enzyme is over expressed in *Drosophila*, no effect is noted on the lifespan. However, in concert they increase lifespan by one third (Orh and Sohal 1994). This is one pathway of evidence to suggest that oxygen radicals play a major role in the aging process. Specifically, oxidative damage causing 8-oxo-7-hydro-2'-deoxyguanosine (8-oxo-dG) has been shown to increase with age both *in vitro* and *in vivo* (Bohr and Anson 1995). Common sources of oxidative stress include metabolism, mitochondrial oxidative phosphorylation and radiation.

Current techniques permit the assessment of *in vivo* human somatic mutation frequencies. In addition, sequence characterization of mutations provides insight into their origin. With these advances, it has become possible to assess the influence of aging on both mutation frequency and mutational spectra. One of the targets of choice for monitoring mutations in humans is the hypoxanthine-guanine phosphoribosyltransferase (*hprt*) gene. Utilized in peripheral T-cells, the *hprt* gene is non-essential and all classes of mutation can be recovered. *In vivo* T-cell *hprt* mutants can be selected and quantified by their resistance to the base analogue 6-thioguanine (6-TG) (~10μM) as employed in the T-cell clonal assay (Albertini *et al.* 1982).

Human populations demonstrate *hprt* mutant frequency (MF) values that range over two orders of magnitude (Branda *et al.* 1993, Cole and Skopek 1994). Accordingly, studies using MF values to monitor mutation in human populations require large sample sizes to account for this high degree of variation. Sources of MF variation are reviewed elsewhere but are primarily age, cloning efficiency (CE), and clonal expansion (Cole and Skopek 1994). Recent studies suggest that dietary components may also affect individual variability observed in MFs (Branda and Albertini 1995, Duthie *et al.* 1995). No evidence for any sex influence has been reported (Cole *et al.* 1988, Tates *et al.* 1991, Branda *et al.* 1993, Cole and Skopek 1994, Huttner *et al.* 1995).

Consequences of tobacco smoking on MF have been widely studied with enigmatic results. Cole *et al.* (1988) observed that smokers had a 56% higher MF than non-smokers. They later reported that the smoking effect was age-related. The observed increase in MF for smokers was 2.9% per year compared to only 0.8% in non-smokers (Cole *et al.* 1990, Cole *et al.* 1991). This observation differs from that of Davies *et al.* (1992) and Albertini *et al.* (1988) who found no effect, and Sala-Trepat *et al.* (1990) who found smokers to have lower MF than non-smokers. A time-dependent effect of smoking was reported by Jones *et al.* (1993), where MF increased 10% per year of smoking compared to an increase of 1% per year by age. Another paper reports that a statistically significant difference between MF values of smokers and non-smokers

could be observed only for female subjects (Huttner *et al.* 1995). More recently maternal cigarette smoking failed to demonstrate any significant effect on the MF of newborns (cord blood) (Finette *et al.* 1997).

Many investigators report a linear increase of *hprt* MF with subject age (Morley *et al.* 1982, Trainor *et al.* 1984, Cole *et al.* 1988, 1991, Tates *et al.* 1991, Davies *et al.* 1992, Jones *et al.* 1993, Branda *et al.* 1993, Robinson *et al.* 1994.). Overall, these studies suggest a 1-3 % increase in MF per year at this locus. Most studies had limited numbers of adolescent subjects. Later comparison of MF values between adolescents and adults demonstrated a non-linear response in MF with age (Finette *et al.* 1994). Adolescents were observed to have a rapid increase in MF with age as compared to adult values (Finette *et al.* 1994).

Other studies employing different loci have also demonstrated linear age versus MF relationships. Frequency of mutations at the HLA-A locus were found to increase with age from a mean of 0.71 x 10^{-5} in neonates to 6.5 x 10^{-5} in the elderly (age > 60 years) (Grist *et al.* 1992). In addition to demonstrating a linear age relationship for MF at the *hprt* loci, Akiyama *et al.* (1995) resolved similar linear relationships using the glycophorin A and α/β T-cell receptor genes.

While the effect of age on mutant frequency has been the subject of several studies, the effect of aging on the mutational specificity of spontaneously arising mutants has not been well studied. The mutational spectrum of *hprt* is affected by age (O'Neill *et al.* 1998). Studies employing Southern blot and PCR techniques have demonstrated a shift in the frequency of V(D)J recombinase-mediated deletion events at the *hprt* locus which account for approximately 40% of the T-cell mutations in cord blood samples (Fuscoe *et al.* 1991) but only 2% of the events in adults (Fuscoe *et al.* 1992). It is assumed that during the ontegeny of T-cells, when V(D)J recombinase is most active, the enzyme complex occasionally works on more than just T-cell receptor (TCR) sequences. As the frequent exon II-III deletion events in cord blood are flanked by V(D)J recombinase consensus sequences, the production of this particular class of events is clearly age-related.

Both CpG transitions and V(D)J recombinase mediated deletions were found to be dramatically altered when compared between newborns (cord blood) and adults. Alterations to mutational specificity that may occur later in life have yet to be formally explored.

Methods

Mutant Frequency Analysis

A database containing published *hprt* mutant frequency (MF) data was constructed, including unpublished MF data from this laboratory. A total of 1194 individual subjects with ages ranging from 0 to 85 were identified from the available literature. The MF dataset is available upon request. An overview of the MF dataset by author is given in Table 1. Pre-treatment cancer patients were included, as those studies (Karnaoukhova *et al.* 1997, 1998) and previous studies (Caggana *et al.* 1991, Dempsey *et al.* 1985) report no difference in MF compared to

Reference	Subjects	Male/ Female	[a]Smoker/ NonSmk	Mean Age ± SD years	Mean MF ± SD x 10^{-6}
Bachand *et al.* 1991	25	17 8	6 19	51.6 ± 8.8	15.0 ± 9.9
Branda *et al.* 1993	232	77 155	106 101 25[b]	52.6 ± 16.3	14.5 ± 11.2
Cole *et al.* 1988	17	8 9	5 12	34.9 ± 11.8	5.9 ± 2.9
Curry *et al.* 1997	12	12 0	2 10	29.3 ± 4.7	19.3 ± 10.5
Curry *et al.* 1995	1	0 1	1 0	50.0	12.0
Curry *et al.* 1993	1	1 0	1 0	37.0	26.0
Davies *et al.* 1992	41	21 20	23 18	36.1 ± 12.0	2.8 ± 2.2
Finette *et al.* 1994	49	29 20	0 49	5.6 ± 4.7	2.3 ± 2.2
Finette *et al.* 1997	63	37 26	0 63	0.0 ± 0.0	0.9 ± 1.0
Hakoda *et al.* 1988	17	7 10	0 0 17	58.6 ± 7.4	3.3 ± 2.0
Henderson *et al.* 1986	23	7 7 9[c]	4 19	21.0 ± 18.3	2.6 ± 2.2
Hirai *et al.* 1995	84	25 59	14 68	64.0 ± 7.7	10.7 ± 8.6
Hou *et al.* 1995	76	76 0	0 76	43.4 ± 12.8	8.6 ± 4.6
Huttner *et al.* 1995	44	26 18	20 24	39.8 ± 8.8	7.2 ± 5.6
Jones *et al.* 1993	120	68 52	58 62	30.5 ± 6.9	8.5 ± 5.9
Karnaoukhova *et al.* 1997	12	9 3	8 2 2[b]	74.8 ± 6.0	46.4 ± 37.5
Karnaoukhova *et al.* 1998	11	7 4	4 6 1[b]	70.8 ± 9.1	45.9 ± 36.0
McGinniss *et al.* 1990	45	38 7	6 37 2[b]	0.0 ± 0.0	0.7 ± 0.4
O'Neill *et al.* 1989	6	3 3	0 6	26.3 ± 4.3	5.4 ± 1.0
Robinson *et al.* 1994	145	71 61 13[c]	52 92 1[b]	34.0 ± 16.7	9.0 ± 7.4
Saddi *et al.* 1996	8	4 4	0 8	10.8 ± 4.4	3.9 ± 4.6
Sala-Trepat *et al.* 1990	42	0 42	17 21 4[b]	50.4 ± 9.1	19.0 ± 14.0
Skandalis *et al.* 1997	9	3 6	4 5	30.8 ± 14.0	15.5 ± 7.9
Tates *et al.* 1991	111	111[c]	47 64	40.2 ± 10.9	11.2 ± 12.0
TOTALS	**1194**	**546/515/133[c]**	**378/762/54[b]**	**38.2 ± 21.6**	**10.1 ± 11.8**

Table 1. Summary of the assembled mutant frequency database by author. [a] Smokers are considered those actively smoking tobacco. [b] Smoking status unknown. [c] Sex of subjects not reported in literature.

normal populations. Where multiple MF determinations were made for a single subject sample, the average MF was determined and used in the MF database

Cloning efficiencies in the mutant frequency dataset were found to vary substantially between laboratories (n=1194, mean CE = 40.5 ± 21.6 %). Many factors likely contribute to this wide variation. Analysis of the effect of smoking status, sex, or age on cloning efficiency did not reveal any significant trends. For this reason cloning efficiencies are not considered during any of the age and MF analysis and are not reported hence forth.

Mutation frequency (MuF) is the frequency of independent mutational events occurring in a subject, rather than the frequency of mutants that may arise from one mutational event by way of *in vivo* T-cell clonal expansion. Within our own data, we corrected two subjects' mutant frequencies to estimated mutation frequencies by correcting for clonality using unique T-cell receptor (TCR) rearrangements (Curry *et al.* 1995). These two corrections were found not to have any bearing on the subsequent analysis as a whole. Such corrections were not possible in the larger dataset. However, mutant frequency values do reflect a reasonable approximation of the actual mutation frequencies. In general, correction for clonality does not exceed 20~30 % of MF with the possible exception of subjects where the MF is greater than 40×10^{-6} (O'Neill *et al.* 1994). Following the suggestions of Morley (1996) such cases have not been excluded from this analysis, despite the possibility that MF values over 40×10^{-6} may reflect clonal runs of single mutational events.

Analyses were performed using commercial statistical software packages Statistica (Statsoft, Tusla) and SAS (V6.12, SAS Institute Inc., Cary). Regression analyses of mutant frequencies are based upon their natural logarithms as these data are skewed (Henderson *et al.* 1986). Four subjects demonstrated null mutant frequencies, hence the natural logarithms could not be calculated, and they were excluded from further analysis. Relationships between lnMF and age were determined using linear regression. Standard deviations (in brackets) are given after all regression terms. In order to evaluate age trends, the dataset was partitioned

into three age groups and then re-analyzed. Differences between the Pearson's correlation coefficients were tested for significance using a Fisher's z-transformation as described by Neter *et al.* (1989). Following the conservative Bonferroni procedure, the level of significance was lowered (p=0.05/(number of two-way comparisons) when more than 2 two-way comparisons were performed (Wassertheil-Smoller 1995).

hprt Mutational Spectra Analysis

Mutational Database Construction: Release five of the *hprt* mutant database was obtained from Mutabase Software (Durham, NC). All somatic mutants from human subjects were identified, but only mutants characterized by cDNA or multiplex PCR methods were selected from the *hprt* mutant database. Unpublished mutant spectra characterized by this laboratory (n=76, Karnaoukhova *et al.* 1998) were also included.

Recent publications not yet included in the *hprt* mutant database were also added. Records from smokers, non-smokers and unknown smoking status were included and these records were then reorganized to include three new fields; age, sex, and subject ID. This additional information required examination of the original publications and was necessary in order to determine the actual number of individuals from which the data were derived. Each mutant was also checked for accuracy against the original publication. Records with no age data were excluded from the *hprt* mutant dataset.

Mutants containing more than one mutation are described solely as complex. The original *hprt* database describes such mutations with multiple records. Mutants identified with truncated cDNA's or aberrant exon splicings were identified as splice mutations. Splice mutations for which the actual mutation had been determined by genomic DNA sequencing were identified by the presumed causative mutation. Uncharacterized splice mutations, though reported in the following tables, were excluded from further analysis (n=188).

Analysis and the Elimination of Reporting Bias: Mutations collected for the analyses were obtained from several laboratories using either cDNA or multiplex PCR methods for the molecular characterization of mutants. As

differences in these two methods may result in the unequal reporting of mutation classes, or the unequal classification of mutations, a potential for bias exists. An Adams Skopek Monte Carlo test (Adams and Skopek 1987) was used to test each subset of data (by author and/or by methodologies) used against the remainder of the mutant dataset.

Mutational spectra for smokers and non-smokers were tested for differences. Similarly mutational spectra for males and females were generated and tested. Finally, mutational spectra for several age groups were identified. For example the data set was partitioned into two age groups (4-37 and 38-80) and further into three and four age groups, with roughly equal numbers of mutants where possible. As the mutant database contained only two mutant sequences obtained from a young individual (four years of age), no spectra analysis for individuals below 19 years of age was attempted.

Mutational spectra were compared using the Adams Skopek Monte Carlo test. Mutational spectra were further evaluated using Fisher's exact test to examine specific mutational classes. Chi square analysis was used in comparisons where the Fisher's exact test could not be performed.

Results

Mutant Frequency Analysis

Influence of Sex on lnMF: Differences in mean age, and lnMF between males and females were tested using a t-test. As noted in Table 1, the sex of some subjects was not reported in the literature. Distribution of age between males and females was significantly different ($p < 10^{-6}$). Cord blood samples (age = 0) are over-represented in the male population (n= 47) as compared to the female population (n=9) and result in a lower mean age for males. In order to obtain similarly distributed ages for the comparison of lnMF, the cord blood samples were excluded from the analysis. Although mean age of female subjects remained significantly higher (p = 0.001) than that of males, lnMF was not significantly different between the sexes. As sex had no influence on lnMF, both sexes were combined to assess any effects of smoking.

Influence of smoking on lnMF: Differences between the lnMF of individuals

who smoke tobacco and non-smokers were tested for significance. Mean age was significantly different between smokers and non-smokers ($p < 10^{-6}$). Subjects aged 16 or less were excluded from this analysis as they mainly comprised non-smokers and significantly lowered the mean age of the group. With this exclusion no significant difference in the mean age between the smoking and non-smoking groups was observed. However, the smoking population had a significantly higher mean lnMF compared to the non-smoking group ($p = 0.0004$).

To further explore the difference in the mean lnMF between smokers and non-smokers, these two groups were sub-divided by sex. After subjects aged 16 or less were excluded from the analysis, mean ages were not significantly different, except between male and female non-smokers ($p=0.0001$). Significant differences in the mean lnMF is observed only between male smokers and non-smokers ($p=0.005$) as well as between male non-smokers and female smokers ($p=0.001$). Male non-smokers are noted to have the lowest lnMF among the four groups. Surprisingly, no significant difference was detected between the female smokers and non-smokers.

Influence of age on lnMF: Results of several linear regression analyses are reported (Table 2). In the first regression, all data were included in the analysis of age versus lnMF. The dependent variable was lnMF and age the independent variable. Correlation coefficients of the linear age-lnMF relationships for males and females are significantly different ($p = 0.007$). As previously noted the mean age of male subjects was significantly lower than that of females. The male group contains significantly more MF determinations from cord blood samples (males n= 84, females n=35) which partially accounts for the lower male mean age as compared to females. After excluding cord blood samples from both male and female linear age lnMF regressions there was no significant difference in the correlation coefficients. The regression lines are nearly identical as indicated by the intercepts and slopes (Table 2).

Linear regressions for smokers and non-smokers revealed no significant differences in

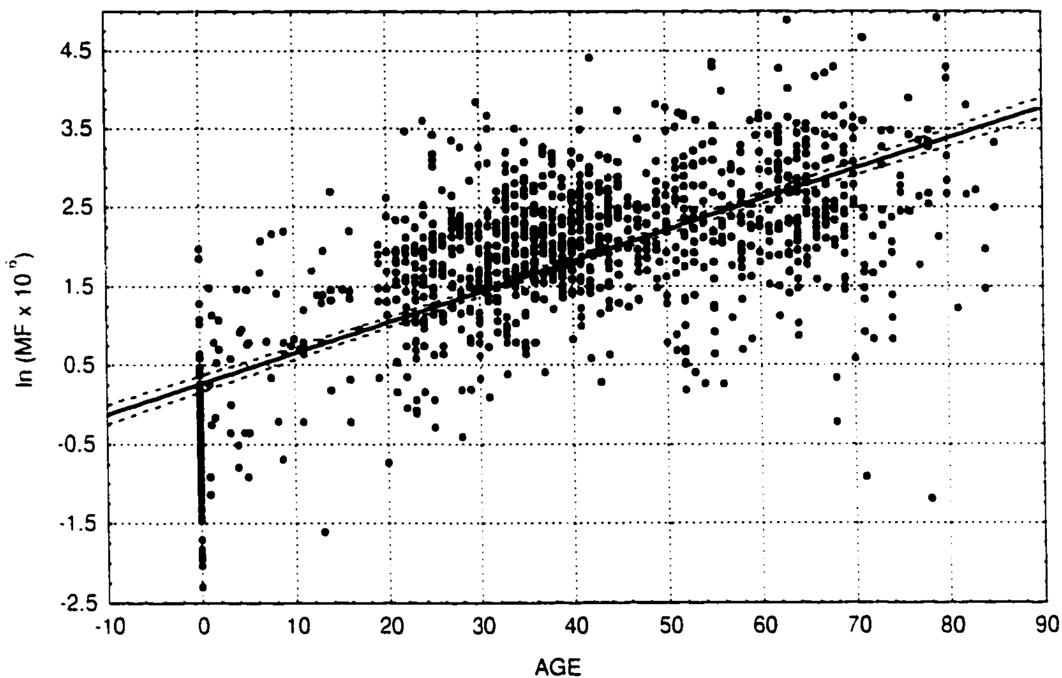
	Regression Line with (S.D.)	n [a]	r [b]	p value
ALL DATA	lnMF = 0.264 (0.052) + 0.039 (0.001) age	1190	0.690	$p < 10^{-6}$
Males	lnMF = 0.239 (0.069) + 0.039 (0.002) age	544	0.716	$p < 10^{-6}$
Females	lnMF = 0.411 (0.091) + 0.035 (0.002) age	514	0.625	$p < 10^{-6}$
Males aged > 0	lnMF = 0.927 (0.088) + 0.026 (0.002) age	462	0.527	$p < 10^{-6}$
Females aged > 0	lnMF = 1.064 (0.103) + 0.022 (0.002) age	479	0.436	$p < 10^{-6}$
Smokers aged > 16	lnMF = 1.276 (0.122) + 0.022 (0.003) age	370	0.406	$p < 10^{-6}$
Non-smokers aged>16	lnMF = 1.202 (0.091) + 0.019 (0.002) age	576	0.389	$p < 10^{-6}$
Ages 0 - 15	lnMF = -0.513 (0.072) + 0.139 (0.016) age	187	0.541	$p < 10^{-6}$
Ages 16 - 52	lnMF = 1.007 (0.111) + 0.026 (0.003) age	675	0.316	$p < 10^{-6}$
Ages 53 - 85	lnMF = 1.831 (0.442) + 0.010 (0.007) age	328	0.084	$p = 0.13$

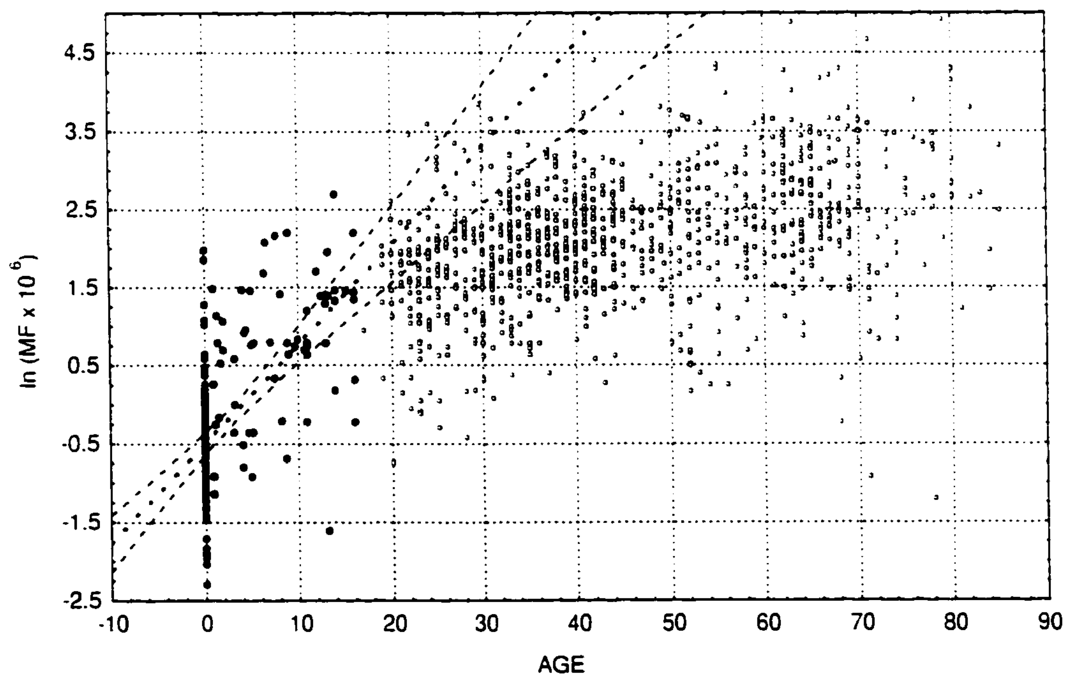
Table 2. Summary of regression analysis for the influence of sex and smoking on the age versus lnMF relationships. [a]As the sex and or smoking status of some subjects is not reported for all subjects, the n values will not sum to total sample size. [b] Pearson correlation.

the correlation coefficients, but only after subjects aged 16 years or less were excluded from the analysis. Smoking and non-smoking age versus lnMF relationships produce identical slopes, but the smoking group intercept is approximately 8% higher (Table 2).

Consistent with previous observations (Finette *et al.* 1994) data in the present study were partitioned into a child group (age 0-15) and two adult groups (age 16-52 and 53-85) (Table2). Rational for the further subdivision of the adult group is as follows. The adult population was divided at the point where the oldest partition (53-85 years of age) no longer provided a significant relationship between age and lnMF (Figure 1). A comparison of the correlation coefficients for the effect of age on lnMF in the three age partitions and for the entire data set (Table 2) was performed. All three of the partitioned relationships are significantly

different from the over all regression line. In addition, the regression lines of the three age group partitions are all significantly different from one another.


Plot A
4.5
3.5
2.5
1.5
0.5
-0.5
-1.5
-2.5
ln (MF x 10^6)
-10
0
10
20
30
40
50
60
70
80
90
AGE


Plot B
4.5
3.5
2.5
1.5
0.5
-0.5
-1.5
-2.5
ln (MF x 10^6)
-10
0
10
20
30
40
50
60
70
80
90
AGE

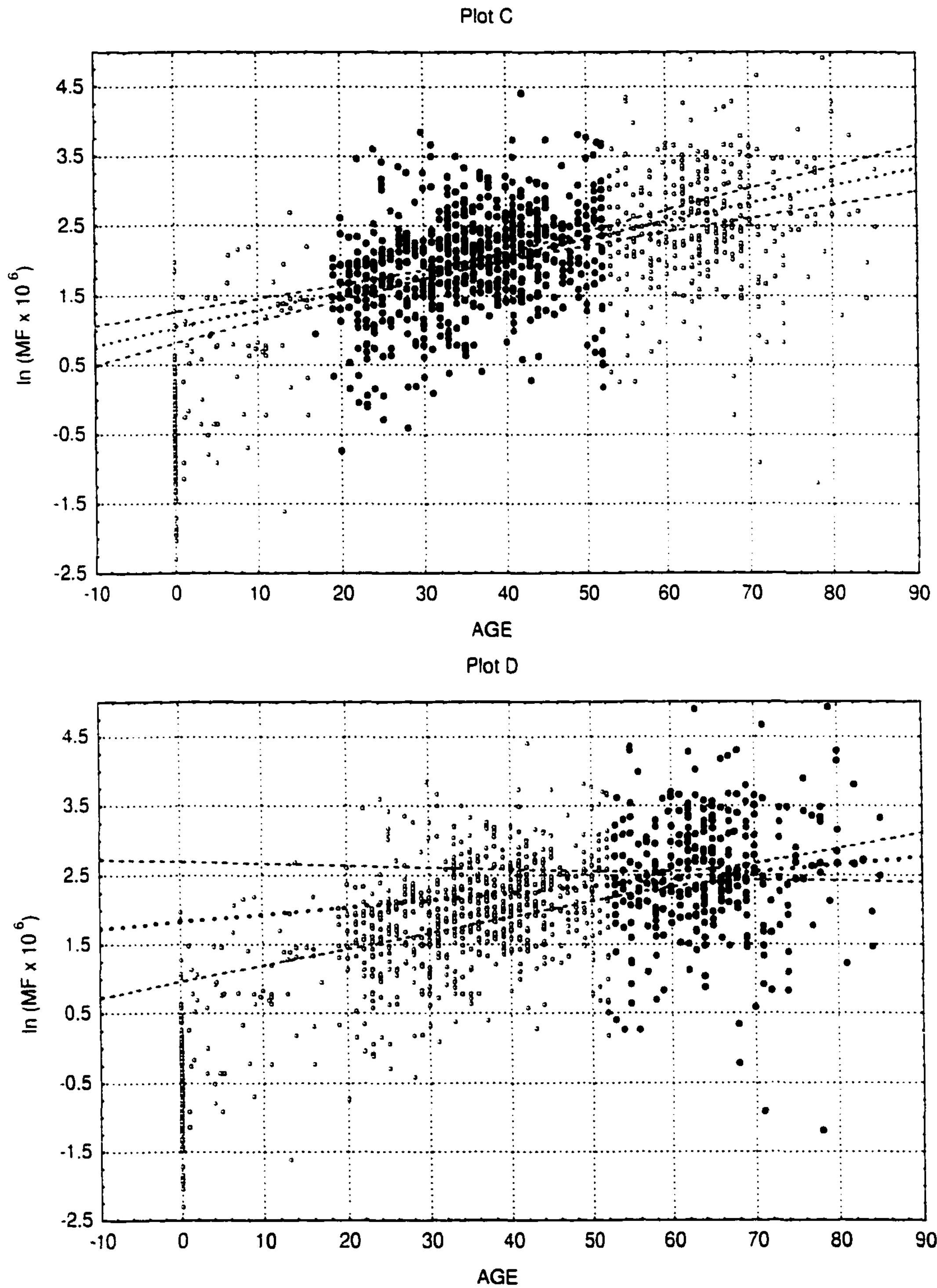


Figure 1. Linear regression lines for the lnMF= f(age) relationships for the entire dataset [Plot A] and three age groups: ages 0-15 [Plot B], ages 16-52 [Plot C] and ages 53-85 [Plot D] where no relationship with age exists. Closed circles indicate the data that is used for the respective regressions. The actual equations, R values and probability values for the equations are shown in Table 2. All four lines are significantly different from one another.

Mutational spectra analysis

A total of 795 independent mutations were collected from 342 individuals and a summary of the mutational data assembled is listed by publication (Table 3). The complete *hprt* mutant dataset is available upon request. Mean age of these 298 individuals (excludes subjects which only presented uncharacterized splice mutants) is 40.3 ± SD 15.2 years of age. Mutations are from individuals greater than 18 years of age.

Reference	Subjects[a]	Sex F/M	Nonsmoker Smoker	Mutants Sequenced	Mean Age ± SD
Anderson *et al.* 1992	4	0/4	4/0	12	25.8 ± 3.8
Burkhart-Schultz *et al.* 1993	52[b]	17/35	25/27	52	35.5 ± 8.7
Burkhart-Schultz *et al.* 1996	52	17/35	25/27	102	28.9 ± 7.2
Burkhart-Schultz *et al.* 1997	46	8/41	18/31	49	35.0 ± 9.6
Curry *et al.* 1993	1	1/0	0/1	8	37.0 ± 0.0
Curry *et al.* 1995	1	0/1	0/1	38	50.0 ± 0.0
Hou *et al.* 1993	6	0/6	6/0	14	36.3 ± 15.9
Karnaukova *et al.* 1997	11	2/9	2/9	58	74.4 ± 6.1
Karnaukova unpublished	10	3/7	4/6	76	70.0 ± 8.9
Osterholm *et al.* 1998	30	0/30	30/0	58	43.3 ± 13.0
Podlutsky *et al.* 1998	38	0/38	38/0	115	43.2 ± 11.1
Recio *et al.* 1990	1	0/1	1/0	27	23.0 ± 0.0
Rossi *et al.* 1990	9	0/9	9/0	31	37.4 ± 6.4
Shimahara *et al.* 1995	15	0/15	15/0	34	66.3 ± 8.1[c]
Skandalis *et al.* 1997	9	6/3	4/5	31	31.1 ± 14.7
Vrieling *et al.* 1992a	12	0/12	3/9	56	36.5 ± 6.8
Steingrimsdottir *et al.* 1993	4	0/4	2/2	34	41.5 ± 14.0
TOTALS	**342**	**62 / 280**	**191 /151**	**795**	**39.6 ± 14.8**

Table 3. Summary of mutant dataset for spectral analysis. [a] Subject totals corrected for subjects appearing in more than one publication. [b] Number of subjects reflects only those listed in release 5 of the *hprt* database (Mutabase Software, Durham NC). [c]Exact ages could not be determined for the thirty-four individual mutants collected from the 15 subjects but the mean age of these male subjects was 65.7 ± 7.9 years of age with a range of 58 to 81 years of age and all mutants are later placed into the oldest age partition.

with the exception of two mutations from a four-year-old. Hence, the analysis is limited to the examination of changes in mutational specificity occurring exclusively in adults. As the majority of mutations (n=513) have been obtained from the 236 male subjects, this unfortunately limits the analysis of the spectra for the effect of sex. Summary information regarding the subject samples by spectra groupings is shown (Table 4).

Test for dataset bias: When each author dataset was tested against the remainder of the *hprt* mutant dataset using the Monte Carlo test.

	Female	Male	Smokers	Non-Smokers
Total Individuals / Mutants	62 / 104	236 / 513	132 / 248	164 / 369
Smokers / Non Smokers	33 / 27	99 / 137	132 / 0	0 / 164
AGE ± SD years	36.1 ± 14.2	41.4 ± 15.3	39.5 ± 13.8	40.9 ± 16.3
	19-37 Years	*38-83 Years*		
Total Individuals / Mutants	157 / 282	139 / 335		
Smokers / Non Smokers	70 / 87	62 / 77		
AGE ± SD years	29.1 ± 5.6	52.9 ± 12.5		
	19-31 Years	*32-45 Years*	*46-83 Years*	
Total Individuals / Mutants	101 / 179	108 / 206	87 / 232	
Smokers / Non Smokers	39 / 62	62 / 46	31 / 56	
AGE ± SD years	26.0 ± 4.2	37.8 ± 3.7	60.1 ± 10.4	
	19-31 Years	*32-40 Years*	*41-54 Years*	*55-83 Years*
Total Individuals / Mutants	101 / 179	80 / 159	59 / 132	56 / 147
Smokers / Non Smokers	39 / 76	43 / 37	35 / 24	15 / 41
AGE ± SD years	26.0 ± 4.2	36.1 ± 2.4	46.3 ± 4.1	66.0 ± 8.2

Table 4. Subject summary data for spectra.

several demonstrated significant differences. In particular, the data of Burkhart-Schultz *et al.* (1996) was significantly different ($p < 10^{-6}$). In this particular dataset (n=102) the authors reported only point mutations. After mutational spectra combined from all three of the Burkhart-Schultz *et al.* (1993, 1996, 1997) papers were compared to the remaining mutational spectrum, no significant difference was detected. Their most recent paper (Burkhart-Schultz *et al.* 1997) includes mutants not reported in their earlier paper (1996). Other datasets that were found significantly different from the remainder contained fewer than 50 mutations and the comparison of such small mutational spectra likely accounts for the observed differences. When these smaller data sets were grouped together, by laboratory where possible, and the analysis repeated, no significant differences were observed.

Differential characterization of exon splice mutations represents the most consistent bias. Where causative mutations responsible for exon splice mutations were reported in the literature, the causative mutation was listed in the mutant database. As some authors did not further characterize exon splice mutants while others did, this class of mutation could possibly bias the dataset. For example, a group of mutants where the causative mutations were determined would be different from a group where no further characterization of exon splice mutants was attempted. To avoid this potential bias, splice mutants that have not been further characterized are not considered. Finally, the spectrum of mutations obtained from fully characterized splice mutants could differ from those mutations not affecting splicing. However, comparison of the mutational spectra from characterized exon splice mutations against mutations not affecting splicing revealed that these spectra are identical in terms of the other classes of mutations.

Influence of smoking on spectra: The smoking and non-smoking groups were nearly identical in terms of age distribution (Table 4). Mutants were sorted into those from smokers (248 independent mutants) and non-smokers (369 independent mutants) and the two spectra compared (Table 5). Results of the Monte Carlo

Mutation Class	Smokers	Non-smokers	Female	Male	19-37 Years	38-83 Years
GC > AT	55 *22.2%*	103 27.9%	17 16.3%	141 27.5%	70 24.8%	88 26.3%
AT > GC	22 *8.9%*	36 9.8%	11 10.6%	47 9.2%	30 10.6%	28 8.4%
Transitions	*77 31.0%*	*139* 37.7%	*28* 26.9%	*188* 36.6%	*100* 35.5%	*116* 34.6%
GC > TA	21 *8.5%*	29 7.9%	12 11.5%	38 7.4%	28 9.9%	22 6.6%
GC > CG	24 *9.7%*	34 9.2%	15 14.4%	43 8.4%	25 8.9%	33 9.9%
AT > TA	22 *8.9%*	34 9.2%	4 3.8%	52 10.1%	25 8.9%	31 9.3%
AT > CG	25 *10.1%*	24 6.5%	10 9.6%	39 7.6%	14 5.0%	35 10.4%
Transversions	*92 37.1%*	*121* 32.8%	*41* 39.4%	*172* 33.5%	*92* 32.6%	*121* 36.1%
Total Substitutions	*169 68.1%*	*260* 70.5%	*69* 66.3%	*360* 70.2%	*192* 68.1%	*237* 70.7%
[a]Deletion	32 *12.9%*	49 13.3%	22 21.2%	59 11.5%	38 13.5%	43 12.8%
Frameshift (-)	26 *10.5%*	35 9.5%	7 6.7%	54 10.5%	30 10.6%	31 9.3%
Frameshift (+)	6 *2.4%*	7 1.9%	1 1.0%	12 2.3%	7 2.5%	6 1.8%
Complex	12 *4.8%*	8 2.2%	5 4.8%	15 2.9%	6 2.1%	14 4.2%
Tandem	2 *0.8%*	6 1.6%	0 0.0%	8 1.6%	5 1.8%	3 0.9%
Insertion	1 *0.4%*	4 1.1%	0 0.0%	5 1.0%	4 1.4%	1 0.3%
Other Classes	79 *31.9%*	109 29.5%	35 33.7%	153 29.8%	90 31.9%	98 29.3%
GRAND TOTAL	**248 100%**	**369 100%**	**104 100%**	**513 100%**	**282 100%**	**335 100%**
Uncharactized Splice	88	90	34	144	54	124

Table 5. Mutational spectra for smoking status, sex, and age (two partitions). Mutational classes demonstrating significance are underlined and described in text. [a] Involving more than a single basepair.

test revealed no significant differences. A barely significant difference in the frequency of G:C→A:T between smokers and non-smokers was detected using Fisher's exact test (two-tail, p=0.05).

Eighty-seven independent mutants, (not including the uncharacterized splice mutations; 85 from smokers) were obtained from lung cancer patients (Karnaoukhova *et al.* 1997). Mutants from the treated patients did not differ from the pretreatment patients (same individuals) (Karnaoukhova *et al.* 1997) so the data were pooled. With these additional data, no significant difference in the frequency of G:C→A:T transitions between smokers and non-smokers could be detected. Consequently, the smoking and non-smoking mutational spectra datasets have been combined for further analysis.

Influence of sex on spectra: Few datasets are large enough to examine the influence of sex on mutational specificity. Subject age distribution (Table 4) was essentially equal however relatively few mutants were recovered from females subjects (n=104) as compared to males (n=513). Monte Carlo analysis did demonstrate a significant difference between the two spectra (Table 5) (p=.008; 95% confidence limits 0.006 to 0.01). Each mutation class was tested alone using Fisher's exact test (2-tail). There were significant differences in the frequency of G:C→ A:T (p = 0.02) and A:T→T:A (p = 0.04) base substitutions between males and females. However, caution is warranted considering the limited set of mutants obtained from female subjects. These differences were tested again with the addition of 87 mutants (67 from males) obtained from the etoposide study (Karnaoukhova *et al.* 1997). In both cases, significance was lost after this addition.

Deletion frequency in females is almost twice as high as in males (Table 5). Mutants from the etoposide study were added and the data re-analyzed. Addition of the etoposide mutants further strengthened the significance of the conspicuous uneven distribution of deletions between the sexes (Table 6). An analysis of the regions surrounding the endpoints of these deletions in male and female subjects did not reveal an obvious difference in regard to

	Male	Female	Fisher's
	Deletions / Total	Deletions / Total	p value
All Data	59 / 513	22 / 104	0.008
Etoposide mutants[a]	7 / 67	3 / 20	0.42
All Data + Etoposide	66 / 580	25 / 124	0.008

Table 6. Analysis of the frequency of deletion between females and males. [a] The etoposide mutant data (Karnaoukhova *et al.* 1997) includes 10 deletion events, though the publication refers to one mutant involving two separate deletions as a complex mutation.

prevalence of deletions flanked by repeated sequence. Lengths of these intragenic deletions were calculated and the means determined. Females had a mean deletion length of 24 ± 65 bps compared with males who had a mean of 34 ± 84 bps.

Characterized splice mutations were obtained disproportionately from males (n=99) as compared to females (n=16). This bias is likely due to the relative ease of characterizing the single mutant allele in males. To ensure that the bias did not affect the analysis, it was repeated excluding all splice mutations. The increased frequency of deletions from female over male subjects was confirmed.

Effect of age on spectra: In order to facilitate the analysis of age-related effects on mutational specificity, data were subdivided into three different age groups as described (Table 4) with nearly equal numbers of mutants. Group mean ages were all significantly different from one another. The Monte Carlo test was performed on mutational spectra for the two (spectra shown in Table 5), three or four partitioning groups (spectra not shown). No differences in spectra were observed for any of the age partitions. Fisher's exact test was used to compare all the classes individually (2x2 tables). The A:T→C:G class demonstrated a significant difference in the two (df=1, two tail, p=0.015), three (df=2, two tail, p=0.004), and four (df=3, two tail, p=0.012) age partitions. In each case, frequency of AT→C:G mutations was significantly lower for the youngest age partition

than the older partitions. For all three sets of comparisons, the observation remained significant after the addition of the 87 etoposide mutants.

Analysis of the distribution of mutation: Distribution of mutation within the *hprt* gene was examined for age-related differences. This analysis was based solely upon the distribution of single base-pair substitutions. Using the Monte Carlo test, the influence of smoking and age in the two-age group partition was examined. No significant differences between spectra were observed during either the smoking or age analysis (results not shown). However, several frequently mutated single base-pair substitutions (sbps) sites were identified in the mutant dataset. Sites recovered at least eight times were assembled (data not shown) and the effect of smoking, sex and age examined using the Monte Carlo test. Several of these frequently mutated sites have been previously identified as hotspot sites for mutation (positions 197, 508, 617) (Burkhart-Schutlz *et al.* 1996). No significant differences were observed in the smoking, sex, or age partitions.

Sites were then tested individually with the Fisher's exact test (cell counts for non-events were dependent upon the site nucleotide). Comparing spectra obtained from males and females, base substitutions at position 611 are more frequently recovered in female than in males (16% vs. 3% ; $p = 0.02$). The addition of the etoposide data maintained the significance ($p = 0.01$). Addition of these data was permissible because the etoposide subject group comprised both males and females.

CpG site Mutation Analysis: Frequency of G:C→A:T transition mutations occurring at CpG sites was analyzed. The *hprt* coding sequence contains eight CpG dinucleotides yielding twelve nucleotide positions where mutation will cause an amino acid substitution or stop codon. From this collection of mutants, only five CpG sites, all coding for arginine, were found mutated. No significant differences (Fisher's exact test) were observed in the frequency of mutations occurring at CpG sites for the smoking/nonsmoking, sex, or age partition comparisons. Previous investigations found a significant strand bias for mutations

arising by deamination of the methylated cytosine in the human *hprt* gene on the non-transcribed strand (Skandalis *et al.* 1994). Their analysis used the HPRT database (release 3) and included all mutations collected from both *in vitro* and *in vivo* human sources. This analysis, which limits the mutant collection to only *in vivo* spontaneous sources, confirms the non-transcribed stand bias (χ^2=10.8 df=1, $p < 0.001$).

Discussion

T-lymphocyte Biology and mutant frequency

The *hprt* T-cell clonal assay depends upon peripheral T-cells, both because of ease of acquisition, and because they can be cloned *in vitro*, which permits mutant selection. Consequently, the assay is intimately linked to the biology of T-cells (review, Cole and Skopek 1994). Composition of the available pool of T-lymphocytes changes with age. At birth, nearly all of the T-lymphocytes display CD45RA isotypes, indicating that they are "naive" or virgin T-cells. They have not met antigen, and thus not yet responded through clonal expansion. Later in adult life, a second isotype called CD45RO, is borne by nearly half of the T-lymphocyte population (Michie *et al.* 1992). These responder or "memory" T-cells have encountered antigen and responded by clonal proliferation. Clonal expansion requires cell division that provides opportunity for both DNA replication errors and their fixation. It is reasonable to suggest that the majority of spontaneous mutations arising in T-lymphocytes are the consequence of clonal expansion (Nicklas *et al.* 1988). Observations that CD45RO+ T-cells have 2 to 5 times higher mutants frequency than the naive CD45RA+ T-cells further substantiates this assertion (Baars *et al.* 1995).

Observed *hprt* mutant frequency is dependent on a number of factors. The rate at which mutation occurs is estimated at $5 \times 10^{-}$ mutations per nominal cell division (Greene *et al.* 1995). Sizes of T-lymphocyte pools fluctuate greatly over time, but are estimated at $1\text{-}5 \times 10^{11}$ cells for a healthy adult. During the course of an immunological challenge a single "responding" T-cell undergoes an estimated 20 to 50 divisions. It is likely that more than one T-cell responds to antigenic challenge. The number of challenges met both recently and over

a lifetime consequently must influence frequency of *hprt* mutants in an individual.

Clonal expansion of mutant T-cells increases the potential for recovery in the T-cell clonal assay. Single mutations, which have not clonally expanded, are nearly undetectable in the current assay, which uses 10–40 x 10^6 cells. Prevalence of mutant clonal expansion in subjects is only evident when sufficient numbers of mutants have been isolated and characterized at the *hprt* locus and subsequently, when mutant characterization indicates a possible clonal run at a TCR locus. This laboratory has previously recognized a large and two smaller clonal runs in a single subject by employing such methods (Curry *et al.* 1995) as have others (Nicklas *et al.* 1988, Albertini *et al.* 1990). O'Neill *et al.* (1994) concluded that mutant clonal expansion is a frequent occurrence after observing that of 58 individuals 35 had some degree of expansion. These observations indicate that a large component of MF is contributed by clonal expansion of single mutational events.

The kinetics at which clonal expansions are reversed remains unclear. Also unclear is the size of the original clonal pool, which remain as memory cells, and how those memory cells are maintained. An additional complication is possible selection against *hprt* mutants which may have a reduced proliferation rate (Podlutsky *et al.* 1996). Current estimates for the half-life of T-cells vary greatly but are likely in the range of two to three years (da Cruz *et al.* 1996). In summary, observed mutation frequency at the *hprt* locus in T-cells is dependent upon the nature and life-cycle of this complex tissue.

Mutant frequency analysis

Influence of sex and smoking on *hprt* MF: No effect of sex on lnMF was observed, confirming previous reports. Smoking was related to elevated mean lnMF values compared to non-smokers. Further examination of the effect of smoking revealed that the increase in lnMF was detectable only in male smokers. This contradicts the findings of Huttner *et al.* (1995) who noted a smoking-related increase in MF for females but not males.

The effect of smoking on *hprt* MF remains troublesome. Smoking has a pronounced effect on T-lymphocyte populations. Smoking increases the number of peripheral white blood

cells by approximately 30% (Howell 1970, Corre *et al.* 1971). Increased leucocyte counts have been associated with the number of cigarettes smoked. Leucocyte counts decrease after the cessation of smoking, but only after five years (Parry *et al.* 1997). Increase in size of neutrophil and CD4+ sub-populations is strongly associated with smoking (Freedman *et al.* 1996; Schaberg *et al.* 1997; Parry *et al.* 1997; Tanigawa *et al.* 1998). Mutants recovered with the T-cell clonal assay are mainly CD4+ (McGinniss *et al.* 1990).

Several suggestions have been presented concerning the smoking-related increase in CD4+ T-cells. Smoking may alter cell trafficking, decreasing the ability of the cells to adhere to cellular tissues (marginal pool) thus increasing their relative yield in the peripheral pool (Parry *et al.* 1997). Others suggest that a continual local inflammation in the respiratory tract is responsible, or that some component of tobacco smoke may act as antigen, stimulating T-cell proliferation (Tanagawa *et al.* 1998). Smoking-related increases in peripheral T-cell populations may be related to increases in *hprt* MF. Should smoking serve to increase the peripheral pool by proliferation, an increased *hprt* MF could also be expected. Smokers are observed to be more susceptible to common colds than non-smokers (Cohen *et al.* 1993) which may also explain the increased peripheral T-cell pools through proliferation against antigen.

Influence of age on *hprt* MF: Results here demonstrate that despite a highly significant overall (ages 0 to 85) lnMF versus age linear relationship, the data do provide for an alternative model. As has been demonstrated previously, lnMF increases rapidly with age in children (Finette *et al.* 1994) and this increase significantly differs from the relationship found for adult populations (Branda *et al.* 1993). Results here confirm that observation. At about the age of 15 the rate at which *hprt* mutants accumulate falls to a level maintained for another 40 years after which few *hprt* mutants appear to accumulate (Figure 1).

Rapid increases in MF occurring in children must be related to the nature of their immune systems. Specifically, a juvenile immune system is one that is growing and learning with every encounter with antigen. A newborn child enters an antigen-rich environment where it must

respond to numerous immunological challenges. During growth, the lymphoid tissues increase dramatically in size. These factors must affect *hprt* MF. In middle life, where the increase in *hprt* MF is observed to decrease as compared to juveniles, that decrease can be accounted for by both changes in the immune system and reduced antigen encounters. Children are challenged more frequently compared to adults, who encounter relatively fewer novel antigens, having already developed immunity to those previously encountered.

Changes to the immune system, most notably the involution of the thymus (review; Bodey *et al.* 1997) may have a role in the complex relationship between *hprt* MF and age. Thymus involution begins in late childhood (age 15) and by late adulthood (approximately age 60) thymic tissues have been completely replaced by fat and are no longer maturing naive T-cells (George and Ritter 1996). After this point the individual is completely dependent on the T-cell pool generated earlier in life. Of particular relevance are changes in the aging immune system specifically affecting T-lymphocytes. T-cells of aged individuals demonstrate impaired responses to antigen and decreased clonal expansion (McCarron *et al.* 1987, Globerson 1995, Grubeck-Loebenstein 1997, Pawelec *et al.* 1997). Loss of naive T-cell production and age-related T-cell impairment is conspicuously related to the last age partition where *hprt* MF no longer varies with age. A clear demonstration of thymus involution and the lack of a MF - age correlation after the age of 53 is beyond the reach of the current study, but clearly warrants further investigation.

Mutational Spectra Analysis

Influence of smoking: Using two spectra containing 43 and 55 independent mutants collected from smokers and non-smokers respectively, Vrieling *et al.* (1992a) reported finding no significant difference between the two spectra. The authors did note that both their spectra were void of any G:C→T:A transversions, which are produced efficiently by the metabolites benzo[*a*]pyrene, a tobacco carcinogen (Yang *et al.* 1991). Their relatively small sample size limited the chance for recovering such events. In the larger sample collection analyzed here, there was no difference

in the distribution of G:C→T:A transversions between the smoking and non-smoking spectra.

The liability of comparing mutational spectra of limited size is that apparent differences between the spectra may disappear as the spectral size increases. For instance, the apparent mutational 'hotspot' (exclusively G:C→A:T transitions at position 617) reported by Vrieling and his colleagues (1992b) in humans occupationally exposed to ethylene oxide (mutants=18), is frequently mutated in unexposed populations and some transversions have been recovered at this particular site (Burkhart-Schultz *et al.* 1993, 1996, Andersson *et al.* 1992).

To date the present study is the largest comparison of smoking and non-smoking mutational spectra. Analysis of smoking and non-smoking spectra using the Adams Skopek Monte Carlo test did not reveal any difference, confirming reports of Vrieling *et al.* (1992a) and Burkhart-Schultz *et al.* (1993, 1996). This conclusion confounds the observation that numerous smoking-related DNA adducts are formed in various tissues, including T-lymphocytes (Van Schooten *et al.* 1997).

Burkhart-Schultz *et al.* (1996) concluded that smoking yielded no observable change in spectrum despite an increase in MF. A plausible explanation for an enhanced MF but no change in spectrum between smokers and non-smokers relates to the fact that smoking increases the numbers of T-cells *in vivo*. As smokers are more prone to, and suffer more, respiratory illnesses they may have relatively more active immune systems with a higher T-cell turnover than non-smokers. It follows that the smoking-related increase in MF might be a reflection of more rounds of replication experienced in the T-cells of smokers. This would also explain why no change in spectrum is observed. Finally, the possibility exists that the *hprt* locus or T-cells are simply a rather poor bio-marker for the study of the effects of smoking.

Influence of sex: The mutant dataset of males is eight-fold larger than females. This reflects the relative ease of molecular analysis, due to the male hemizygous state. Despite the disparity in sample size, comparison of base substitution mutational spectra from males and females did not reveal any significant difference. Among 104 mutants recovered from females, a statistically significant (Fisher's test, $p=0.008$) increase in the frequency of deletions over that of males was observed. Analysis of the deletion endpoints for both males ($n=59$) and females ($n=22$) reveals a similar distribution of the breakpoints occurring at repeat and inverted repeat sequences.

An increase in the recovery of small deletions from female subjects is intriguing. It is possible that the difference reflects recombination-mediated repair, as a second copy of the gene is available in females. However, a larger number of mutants from females is needed to better understand the mechanisms involved.

Influence of age: The question of how age affects mutational specificity was examined by coupling the *hprt* mutant dataset with subject age. Analysis of several different modes of age partitioning using the Monte Carlo test revealed no significant differences in the complete spectra, however an increase in the frequency of A:T→C:G transversions with age was observed.

Generally, A:T→C:G transversions are relatively rare in mutational spectra. However, several chemical agents have been reported to increase the frequency of this class of transversion. Benzo(α)pyrene's ultimate carcinogenic metabolite [(+)-BPDE] was found to increase the frequency of this type of transversion in the *hprt* gene of Chinese hamster cells, but only at low doses (Wei *et al.* 1993). In the Salmonella microsome assay, *N*-nitrosomethylaniline induced predominantly this type of transversion (Zielenska and Guttenplan 1988). Eight purine base analogues are noted to induce A:T→C:G transversions in *E. coli* (five to 124 fold). The proposed mechanism is the reaction of these hydroxylaminopurines at the O^4-position of thymine (Murray 1987). *N*-ethyl-*N*-nitrosourea, a potent monofunctional ethylating agent capable of alkylating the O^4-position of thymine, has also been reported to

induce these transversions (Eckert *et al.* 1988, Guttenplan 1990). Involvement of these agents in the age-related increase of A:T→C:G transversions can not be determined.

Another potential mechanism for the origin of A:T→C:G transversions may be linked to oxygen radicals and the polymerase most active in dividing cells. *In vitro* oxygen radical induced mutagenesis was eloquently demonstrated to be DNA polymerase specific (Feig and Loeb 1994). Using oxygen radical damaged M13 viral DNA in a forward mutation assay, DNA polymerase alpha demonstrated a 14-fold increase in the frequency of A → C transversion as compared to undamaged templates. By comparison, DNA polymerase beta yielded only a 1.5 fold increase. DNA polymerase alpha accounts for more than 85% of total DNA polymerase activity in dividing cells but only 5% in quiescent cells (Kornberg and Baker 1991). The involvement of this type of mechanism to account for the age-dependent increase in A:T→C:G transversions can be only speculative at this time.

Conclusions

The major contribution to *hprt* mutation occurs early in life though mutations continue to accumulate at a diminishing rate until the early fifties. After this age there is no further significant accumulation of mutation. Despite the overall increase in mutation in adults over time, the spectrum of mutation was relatively constant with the single exception of the frequency of A:T→C:G transversions which increase with age. No evidence was recovered to support any model of aging that predicts mutation rate to accelerate with age. Either most mutations do occur at an early stage in life, or these observations are peculiar to T-cells. As the stem cells for the production of T-cells are less prolific in old age, and T-cells themselves apparently divide less vigorously, there are fewer opportunities for mutation. Thus, while the HPRT clonal assay may be the most prominent technique in use for the monitoring of mutation in humans *in vivo*, it is not necessarily well suited for the purpose. Indeed, neither long-term exposure such as in the case of tobacco smoking (this study) nor treatment with a powerful

chemotherapeutic agent such as etoposide (Karnaukova *et al.* 1997) produced any observable impact on mutational spectra. We are thus forced to conclude that further insights into the life cycle and fundamental biology of T-cells are required to properly appreciate the advantages and limitations of this system for environmental monitoring purposes.

Acknowlegdements

This analysis could not have been accomplished without the international support of research into human mutant frequencies and mutational spectra. We are extremely appreciative of the collaborative efforts of all of the laboratories working in this field for access to their databases. In addition, we also thank the anonymous donors who provided blood for the assays used in these studies. Statistical support was provided by Dr. Moyra Brackley. Support of the Natural Sciences and Engineering Research Council of Canada, the Medical Research Council and the Canadian Space Agency is gratefully acknowledged.

Literature Cited

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

Akiyama, M., S. Kyoizumi, Y. Hirai, Y. Kusunoki, K. S. Iwamoto, and N. Nakamura. 1995 Mutation frequency in human blood cells increases with age. Mutation Res. **338:** 141-149.

Albertini, R. J., K. L. Castle and W. R. Borcherding, 1982 T-cell cloning to detect the mutant 6-thioguanine-resistant lymphocytes present in human peripheral blood. Proc. Natl. Acad. Sci. USA **79:** 6617-6621.

Albertini, R. J., L. M. Sullivan, J. K. Berman, C. J. Greene, J. A. Stewart, J. M. Silveira, and J. P. O'Neill, 1988 Mutagenicity monitoring in humans by autoradiographic assay for mutant T-lymphocytes. Mut. Research **204:** 481-492.

Albertini, R. J., J. P. O'Neill, J. A. Nicklas, L. Recio and T.R. Skopek, 1990 *hprt* mutations *in vivo* in human T-lymphocytes: Frequencies, Spectra and Clonality, pp. 15-24 in *Mutation and The Environment. Part C: Somatic and Heritable Mutation, Adduction, and Epidemiology*, edited by M. L. MENDELSOHN and R. J. ALBERTINI. Wiley-Liss, Toronto.

Andersson, B., S. Falt and B. Lambert, 1992 Strand specificity for mutations induced by (+)-anti BPDE in the *hprt* gene in human T-lymphocytes. Mutation Res. **269:** 129-140.

Bachand, M., A. M. Seifert and K. Messing, 1991 Nuclear medicine patients do not have higher mutant frequencies after exposure to thallium-201. Mutation Res. **262:** 1-6.

Baars P. A., M. M. Maurice, M. Rep, B. Hooibrink, and R. A. van Lier, 1995 Heterogeneity of the circulating human CD4+ T cell population. Further evidence that the CD4+CD45RA-CD27- T cell subset contains specialized primed T cells. J. Immunol **154:** 17-25

Bodey, B., B. Bodey Jr, S. E. Siegel, and H. E. Kaiser, 1997 Involution of the mammalian thymus, one of the leading regulators of aging. In Vivo **11:** 421-440.

Bohr V. A., and R. M. Anson, 1995 DNA damage, mutation and fine structure DNA repair in aging. Mut. Research **338:** 25-34.

Branda, R. F., L. M. Sullivan, J. P. O'Neill, M. T. Falta, J. A. Nicklas, B. Hirsch, P. M. Vacek, and R. J. Albertini, 1993 Measurement of HPRT mutant frequencies in T-lymphocytes from healthy human populations. Mutation Res. **285:** 267-279.

Branda, R. F., and R. J. Albertini, 1995 Effect of dietary components on hprt mutant frequencies in human T-lymphocytes. Mutation Res. **346:** 121-127.

Burkhart-Schultz, K., C. B. Thomas, C. L. Thompson, C. L. Strout, E. Brinson, and I. M. Jones, 1993 Characterization of *in vivo* somatic mutations at the hypoxanthine phosphoribosyltransferase gene of a human control population. Environ. Health Perspect. **101:** 68-74.

Burkhart-Schultz, K. J., C. L. Thompson, and I. M. Jones, 1996 Spectrum of somatic mutation at the hypoxanthine phosphoribosyltransferase (hprt) gene of healthy people. Carcinogenesis **17:** 1871-1883.

Burkhart-Schultz, K. J. and I. M. Jones, 1997 Deletion and insertion in vivo somatic mutations in the hypoxanthine phosphoribosyltransferase (*hprt*) gene of human T-lymphocytes. Environ. Mol. Mutagen. **30:** 371-384.

Caggana, M., H. L. Liber, P. M. Mauch, C. N. Coleman, and K. T. Kesley, 1991 In vivo somatic mutation in lymphocytes of Hodgkin's disease patients. Environ. Mol. Mutagen. **18:** 6-13.

Cole, J., M. H. L. Green, S. E. James, L. Henderson and H. Cole, 1988 A further assessment of factors influencing measurements of thioguanine-resistant mutant frequency in circulating T-lymphocytes. Mutation Res. **204:** 493-507.

Cole, J., M. H. L. Green, G. Stephens, A. P. Waugh, D. Beare, H. Steingrimsdotti, and B. A., Bridge, 1990 HPRT somatic mutation data. pp. 25-35 in *Mutation and The Environment. Part C: Somatic and heritable mutation, adduction, and epidemiology*, edited by M. L. MENDELSOHN and R. J. ALBERTINI, Wiley-Liss, Toronto.

Cole, J., A. P. W. Waugh, D. M. Beare, M. Sala-Trepat, G. Stephens, and M. H. Green, 1991 HPRT Mutant Frequencies in Circulating Lymphocytes: Population Studies using Normal Donors, Exposed Groups and Cancer-Prone Syndromes. Prog. Clin. Biol. Res. **372:** 319-328.

Cole, J. and T. R. Skopek, 1994 International Commission for Protection Against Environmental Mutagens and Carcinogens. Working paper no. 3. Somatic mutant frequency, mutation rates and mutational spectra in the human population in vivo. Mut. Research **304:** 33-105.

Cohen, S., D. A. J. Tyrrell, M. A. H. Russell, M. J. Jarvis, and A.P. Smith, 1993 Smoking, alcohol consumption, and susceptibility to the common cold. Amer. J. Pub. Health **83:** 1277-1283.

Corre, J., J. Lellouch, and D. Schwartz, 1971 Smoking and leucocyte counts: results of an epidemiological study. Lancet 2 **(7725):** 632-633.

Curry, J., A. Skandalis, J. Holcroft, J. De Boer and B. W. Glickman, 1993 Coamplification of *hprt* cDNA and γ T-cell receptor sequences from 6-thioguanine resistant human T-lymphocytes. Mutation Res. **288:** 269-275.

Curry, J., G. T. Rowley, V. Saddi, D. Beare, J. Cole, and B. W. Glickman, 1995 Determination of *hprt* mutant and mutation frequencies and the molecular characterization of human derived in vivo T-lymphocyte mutants. Environ. Mol. Mutagen. **25:** 169-179.

Curry J., G. Bebb, D. Young, M. Khaidakov, A. Mortimer, and B. W. Glickman, 1997 Similar mutant frequencies observed between pairs of monozygotic twins. Human Mutation **9:** 445-451

Da Cruz, A. D, J. Curry, M. P. Curado, and B. W. Glickman, 1996 Monitoring *hprt* mutant frequency over time in T-lymphocytes of people accidentally exposed to high doses of ionizing radiation. Env. Molec. Mutagen. **27:** 165-175.

Davies, M. J., D. P. Lovell and D. Anderson, 1992 Thioguanine-resistant mutant frequency in T-lymphocytes from a healthy human population. Mutation Res. **265:** 165-171.

Dempsey, J. L., R. S. Seshadri, and A. A. Morley, 1985 Increased mutation frequency following treatment with cancer chemotherapy. Cancer Res. **45:** 2873-2877.

Duthie, S. J, M. Ross, and A. R. Collins, 1995 The influence of smoking and diet on the hypoxanthine phosphoribosyltransferase (*hprt*) mutant frequency in circulating T-lymphocytes from a normal human population. Mutation Res. **331:** 55-64.

Eckert, K. A., C. A. Ingle, D. K. Klinedinst, and N. R. Drinkwater, 1988 Molecular analysis of mutations induced in human cells by N-ethyl-N-nitrosourea. Mol. Carcinog. **1:** 50:56.

Feig, D. I., and L. A. Loeb, 1994 Oxygen radical induced mutagenesis is DNA polymerase specific. J. Mol. Biol. **235:** 33-41.

Finette, B. A., L. M. Sullivan, J. P. O'Neill, J. A. Nicklas, P. M. Vacek and R. J. Albertini, 1994 Determination of *hprt* mutant frequencies in T-lymphocytes from a healthy pediatric population: statistical comparison between newborn, children and adult mutant frequencies, cloning efficiency and age. Mutation Res. **308:** 223-231.

Finette, B. A., T. Poseno, P. M. Vacek, and R. J. Albertini, 1997 The effects of maternal cigarette smoke exposure on somatic mutant frequencies at the hprt locus of healthy newborns. Mutation Res. **377:** 115-123.

Freedman, D. S., W. D. Flanders, J. J. Barboriak, A. M. Malarcher, and L. Gates, 1996 Cigarette smoking and leucocyte subpopulations in men. Ann. Epidemiol. **6(4):** 299-306.

Fuscoe, J. C., L. J. Zimmerman, M. J. Lippert, J. A. Nicklas, J. P. O'Neill and R. J. Albertini, 1991 V(D)J recombinase-like activity mediates *hprt* gene deletion in human fetal T-lymphocytes. Cancer Res. **51:** 6001-6005.

Fuscoe J. C., L. J. Zimmerman, K. Harrington-Brock, L. Burnette, M. M. Moore, J. A. Nicklas, J. P. O'Neill, and R. J. Albertini, 1992 V(D)J recombinase-mediated deletion of the *hprt* gene in T-lymphocytes from adult humans. Mutation Res. **283:** 13-20.

George, A. J. T. and M. A. Ritter, 1996 Thymic involution with ageing: obsolscence of good housekeeping? Immunol Today **17:** 267-272.

Globerson, A., 1995, T lymphocytes and aging. Int. Arch. Allergy Immunol. **107:** 491-497.

Greene, M. H, J. P. O'Neill, and J. Cole, 1995 Suggestions concerning the relationship between mutant frequency and mutation rate at the *hprt* locus in human peripheral T-lymphocytes. Mut. Research **334:** 323-339.

Grist, S. A., M. Mccarron, A. Kutlaca, D. R. Turner and A. A. Morley, 1992 *In vivo* human somatic mutation: frequency and spectrum with age. Mutation Res. **266:** 189-196.

Grubeck-Loebenstein, B., 1997 Changes in the aging immune system. Biologicals **25:** 205-208.

Gutteridge J. M. C., 1992 Ageing and free radicals. Med. Lab. Sci. **49:** 313-318.

Guttenplan, J. B., 1990 Mutagenesis by N-nitroso compounds: relationships to DNA adducts, DNA repair, and mutational efficiencies. Mutation Res. **233:** 177-187.

Hakoda, M., M. Akiyama, S. Kyoizumi, A. A. Awa, M. Yamakido, and M. Otake, 1988 Increased somatic cell mutant frequency in atomic bomb survivors. Mutation Res. **201:** 39-48.

Harman, D., 1992 Free radical theory of aging. Mutation Res. **275:** 257-266.

Henderson L., H. Cole, J. Cole, S. E. James and M. Green. 1986 Detection of somatic mutations in man: evaluation of the microtitre cloning assay for T-lymphocytes. Mutagenesis 1: 195-200.

Hirai, Y., Y. Kusunoki, S. Kyoizumi, A. A. Awa, D. J. Pawel, N. Nakamura, and M. Akiyama . 1995 Mutant frequency at the *HPRT* locus in peripheral blood T-lymphocytes of atomic bomb survivors. Mutation Res. **329:** 183-196.

Holmes, G. E., C. Bernstein, and H. Bernstein, 1992 Oxidative and other DNA damages at the basis of aging: a review. Mutation Res. **275:** 305-315

Hou, S. -M., A. -M. Steen, S. Falt and B. Andersson, 1993 Molecular spectrum of background mutation at the *hprt* locus in human T-lymphocytes. Mutagenesis **8:** 43-49.

Hou, S. M., S. Falt, and A. M Steen, 1995 Hprt mutant frequency and GSTM1 genotype in non-smoking healthy individuals. Env. Molec. Mutagen **25(2):** 97-105

Howell, R. W., 1970 Smoking habits and laboratory tests. Lancet **2(7664):** 152.

Huttner E., B. Holzapfel, and S. Kropf, 1995 Frequency of HPRT mutant lymphocytes in a human control population as determined by the T-cell cloning procedure. Mutation Res. **348:** 83-91.

Johnson, T. E., 1993 Increased life-span of age-1 mutants in *Caenorhabditis elegans* and lower Gompertz rate of aging. Science **249:** 908-912.

Jones, I. M., D. H. Moore, C. B. Thomas, C. L. Thompson, C. L. Strout, and K. Burkhart-Schultz . 1993 Factors affecting *HPRT* mutant frequency in T-lymphocytes of smokers and nonsmokers. Cancer Epid. Biom. Prev. **2:** 249-260.

Karnaoukhova, L., J. Moffat, H. Martins, and B. W. Glickman, 1997 Mutation frequency and spectrum in lymphocytes of small cell lung cancer patients receiving etoposide chemotherapy. Cancer Res. **57(19):** 4393-4407.

Karnaoukhova, L. G., K. S. Wilson, and B. W. Glickman, 1998 Mutant frequency and mutational spectra in T-lymphocytes of multiple myeloma patients receiving melphalan/prednisone chemotherapy. Submitted to Cancer Research.

Kornberg, A., and T. Baker, 1991 *DNA replication*. W.H. Freeman & Co., New York.

Larsen, P. L., 1993 Aging and resistance to oxidative damage in *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA **90:** 8905-8909.

Martin, G. M., 1997 The genetics of aging. Hosp. Pract. **32:** 47-75.

Mazin, A. L., 1994 Enzymatic DNA methylation as an aging mechanism. Mol. Biol. (Mosk) **28:** 21-51.

McCarron, M., Y. Osborne, C. J. Story, I. L. Dempsey, D. R. Turner, and M. M. Morely, 1987 Effect of age on lymphocyte proliferation. Mech. Ageing Dev. **41:** 211-218.

Mcginniss, M. J., M. T. Falta, L. M. Sullivan and R. J. Albertini, 1990 In vivo *hprt* mutant frequencies in T-cells of normal newborns. Mutation Res. **240:** 117-126.

Michie,C. A., A. Mclean, C. Alcock, and P. C. Beverley, 1992 Memory maybe based upon a long lived clonal run rather than long lived lymphocytes. Such that memory is a pool of short lived lymphocytes that persists for live. Nature **360:** 264-265.

Morley, A. A., S. Cox, and R. Holliday, 1982 Human lymphocytes resistant to 6-thioguanine increase with age. Mech. Age. Develop. **19:** 21-26.

Morley, A. A., 1995 The somatic mutation theory of aging. Mutation Res. **338:** 19-23.

Morley, A. A., 1996 The estimation of in vivo mutation rates and frequency from samples of human lymphocytes. Mutation Res. **357:** 167-176.

Murray, V., 1987 Transversion-specific purine analogue mutagens and the mechanism of hydroxylaminepurine mutagenesis. Mutation Res. **177(2):** 189-199.

Neter, J., W. Wasserman, and M. H. Kutner, 1989 *Applied linear regression models (2nd ed.)* Homewood, IL: Irwin.

Nicklas, J. A., J. P. O'Neill, L. M. Sullivan, T. C. Hunter, M. Allegretta, B. F. Chastenay, B. L. Libbus, and R. J. Albertini, 1988 Molecular analysis of *in vivo hprt* mutations in human T-lymphocytes. II. Demonstration of a clonal amplification of hprt mutant T-lymphocytes *in vivo*. Environ. Mol. Mutagen. **12:** 271-284.

O'Neill, J. P., L. M. Sullivan, J. K. Booker, B. S. Pornelos, M. T. Falta, C. J. Greene, and R. J. Albertini, 1989 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.

O'Neill J. P., J. A. Nicklas, T. C. Hunter, O. B. Batson, M. Allegretta, M. T. Falta, R. F. Branda, and R. J. Albertini, 1994 The effect of T-lymphocyte 'clonality' on the calculated *hprt* mutation frequency occurring in vivo in humans. Mutation Res. **313:** 215-225.

O'Neill, J. P., and B. A. Finette, 1998 Transition mutations at CpG dinucleotides are the most frequent in vivo spontaneous single-base substitution mutation in the human *HPRT* gene. Env. Molec. Mutagen. **32:** 188-191.

Orr, W. C., and R. S. Sohal, 1994 Extension of life-span by overexpression of superoxidase dismutase and catalase in Drosophila melanogaster. Science **263:** 1128-1130.

Osterholm, A., and S. Hou, 1998 Splicing mutations at the HPRT locus in human T-lymphocytes in vivo. Environ. Mol. Mutagen. **32:** 25-32.

Parry, H., S. Cohen, J. E. Schlarb, D. A. Tyrrell, A. Fisher, M. A. Russell, and M. J. Jarvis, 1997 Smoking, alcohol consumption, and leukocyte counts. Am. J. Clin. Pathol. **107:** 64-67.

Pawelec, G., M. Adibzadeh, R. Solana, and I. Beckman, 1997 The T-cell in the ageing individual. Mech. Ageing Development **93:** 35-45.

Pudlutsky, A., T. Bastlova, and B. Lambert, 1996 Reduced proliferation rate of hypoxanthine-phosphoribosyltransferase mutant human T-lymphocytes in vitro. Environ. Mol. Mutagen. **28:** 13-18.

Podlutsky, A., A. Osterholm, S. Hou, A. Hofmaier, and B. Lambert, 1998 Spectrum of point mutations in the coding region of the hypoxanthine-guanine phosphoribosyltransferase (*hprt*) gene in human T-lymphocytes *in vivo*. Carcinogenesis **19:** 557-566.

Recio, L., J. Cochrane, D. Simpson, T. R. Skopek, J. P. O'Neill, J. A. Nicklas, and R. J. Albertini, 1990 DNA sequence analysis of *in vivo hprt* mutation in human T lymphocytes. Mutagenesis **5:** 505-510.

Robinson, D. R., K. Goodall, R. J. Albertini, J. P. O'Neill, B. Finette, M. Sala-Trepat, E. Moustacchi, A. D. Tates, D. M. Beare, M .H. L. Green, and J. Cole, 1994 An analysis of in vivo *hprt* mutant frequency in circulating T-lymphocytes in the normal human populations: a comparison of four datasets. Mutation Res. **313:** 227-247.

Rossi, A. M., J. C. P. Thijssen, A. D. Tates, H. Vrieling, A. T. Natarajan, P. H. Lohman, and A. A. van Zeeland AA, 1990 Mutations affecting RNA splicing in man are detected more frequently in somatic than germ cells. Mutation Res. **244:** 353-357.

Saddi, V., J. Curry, A. Nohturfft, W. Kusser, and B. W. Glickman, 1996 Increased *hprt* mutant frequencies in Brazilian children accidentally exposed to ionizing radiation. Environ. Molec. Mutagen. **28:** 267-275.

Sala-Trepat, M., J. Cole, M. H. L. Green, O. Rigaud, J. R. Vilcoq, and E. Moustacchi, 1990 Genotoxic effects of radiotherapy and chemotherapy on the circulating lymphocytes of breast cancer patients. III: Measurement of mutant frequency to 6-thioguanine resistance. Mutagenesis **5:** 593-598.

Schaberg, T., C. Theilacker, O. T. Nitschke, and H. Lode, 1997 Lymphocyte subsets in peripheral blood and smoking habits. Lung **175(6):** 387-394.

Shimahara H., T. Kato, Y. Hirai and M. Akiyama, 1995 Spectrum of *in vivo* mutations in T lymphocytes from atomic bomb survivors. I. Sequence alterations in cDNA. Carcinogenesis **16:** 583-591.

Skandalis, A., B. N. Ford, and B. W. Glickman, 1994 Strand bias in mutation involving 5-methylcytosine deamination in the human *hprt* gene. Mut. Research **314:** 21-26.

Skandalis A., A. D. Da Cruz, J. Curry, A. Nohturfft, M. Curado, and B.W. Glickman, 1997 Molecular analysis of T-lymphocyte *hprt* mutations in individuals exposed to ionizing radiation in Goiania, Brasil. Environ Molec Mutagen **29:** 107-116.

Slagboom, P. E. and J. Vijg, 1989 Genetic instability and aging: theories, facts and future perspectives. Genome **31:** 373-385.

Steingrimsdottir, H., G. Rowley, A. Waugh, D. Beare, I. Ceccherini, J. Cole, and A. R. Lehmann, 1993 Molecular analysis of mutations in the *hprt* gene in circulating lymphocytes from normal and DNA-repair-deficient donors. Mutation Res. **294:** 29-41.

Tanigawa, T. S. Araki, A. Nakata, and S. Sakurai, 1998 Increase in the helper inducer (CD4+CD29+) T lymphocytes in smokers. Ind. Health **36:** 78-81.

Tates, A. D., F. J. Van Dam, H. Van Mossel, H. Schoemaker, J. C. P. Thijssen, V. M. Woldring, A. H. Zwinderman, and A. T. Natarajan, 1991 Use of clonal assay for the measurement of frequencies of HPRT mutants in T-lymphocytes from five control populations. Mutation Res. **253:** 199-213.

Trainor, K. J., D. J. Wigmore, A. Chrysostomou, J. Dempsey, R. Seshadri, and A. A. Morely, 1984 Mutation frequency in human lymphocytes increases with age. Mech. Age. Develop. **27:** 83-86.

Van Schooten, F. J, R. W. Godschalk, A. Breedijk, L. M. Maas, E. Kriek, H. Sakai, G. Wigbout, P. Baas, N. van Veer, and N. van Zandwijk, 1997. 32P-postlabelling of aromatic DNA adducts in white blood cells and alveolar macrophages of smokers: saturation at high exposures. Mut. Research **378:** 65-75.

Vrieling, H., J. C. P. Thijssen, A. M. Rossi, F. J. Van Dam, A. T. Natarajan, A. D. Tates, and A. A. van Zeeland, 1992a Enhanced *hprt* mutant frequency but no significant difference in mutation spectrum between a smoking and a non-smoking human population. Carcinogenesis **13:** 1625-1631.

Vrieling, H., A. D. Tates, A.T. Natarajan, and A. A. Van Zeeland, 1992b Age-related accumulation of mutations in human T-lymphocytes. Ann. N.Y. Acad. Sci. **663:** 36-42.

Warner, H. R., and A. R. Price, 1989 Involvement of DNA repair in cancer and aging. J. Gerontol. **44:** 45-54.

Wassertheil-Smoller, S., 1995 *Biostatistics and epidemiology. A primer for health professionals.* Springer-Verlag, New York.

Wei S. J., R. L. Chang, N. Bhachech, X. X. Cui, K. A. Merkler, C. Q. Wong, H. Yagi, D. M. Jerina, and A. H. Conney, 1993 Dose-dependent differences in the profile of mutations induced by (+)-7R,8S-dihydroxy-9S,10R-epoxy-7,8,9,10-tetrahydrobenzo(a)pyrene in the coding region of the hypoxanthine (guanine) phosphoribosyltransferase gene in Chinese hamster V-79 cells. Cancer Res. **53(14):** 3294-301.

Yang, J. L., R. H. Chen, V. M. Maher, and J. J. Mccormick, 1991 Kinds and location of mutations induced by (+/-)-7 beta,8 alpha-dihydroxy-9 alpha,10 alpha-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.

Zielenska, M., and J. B. Guttenplan, 1988 Mutagenic activity and specificity of N-nitrosomethylaniline and N-nitrosodiphenylamine in *Samonella*. Mutation Res. **202:** 269-276.

Saturation of the *hprt* Mutational Target

John Curry, Moyra Brackley, Barry Ford, and Barry W. Glickman

Centre for Environmental Health and the Department of Biology, University of Victoria, Victoria, British Columbia

Abstract

We compiled a dataset from the literature of 361 human *hprt* single base pair substitutions (sbps) from 268 subjects. Using this dataset, we predict the number of selectable mutable sites or target size of the *in vivo hprt* mutational target, and estimate the number of mutations required to saturate that target. In order to simplify the model, we include only mutations directed at base pairs (bp) in the coding sequence. Number of unique sites recovered each author/year, and total mutations accumulated were used to determine the fraction of unique sites. The linear regression of those data (Y =6.83 – 5.26 x X, R^2 = 0.99; $p < 10^{-6}$) where the Y intercept value predicts the number of single base-pair substitution mutations required to saturate the number of available mutable sites (925 mutations). The number of available sites was predicted by plotting the number of sites mutated against the fraction of unique sites (Y = 5.78 – 3.97 x X, R^2 = 0.99, $p < 10^{-6}$) and extrapolating the Y intercept (324 ± 1 mutable sites). The human *hprt in vitro* data base (Release 3) (n=687, sbps only) was subjected to the same analysis. The number of mutations required for saturation was 1553 ± 1 sbps and the estimated number of sites available for mutation was 501 ± 1. The differences between the *in vivo* and *in vitro* predictions are discussed. Due to the non-random nature of mutation and the frequency of rare sites, full saturation of any mutational target is extremely unlikely.

Introduction

This report describes an approach for predicting the mutational target size of a gene and to estimate the number of mutations required to saturate that mutational target. Target saturation is defined as the point when the dataset of mutants includes mutations at all possible selectable target sites. The number of mutations required to saturate a mutational target is dependent upon both the actual size of the target and the functional plasticity of the protein product.

Mutational targets currently in use provide a range of target sizes. The bacterial mutational target *supF* is relatively small (< 200 bps). The majority of mutations recovered in this gene are

base substitutions (> 80%) and a few frameshifts (Kraemer and Seidman 1989, Kawanishi *et al.* 1998, Cabral *et al.* 1998). Most targets are large; ?cII with 270 bps, *gpt* with 555 bps, *aprt* with 540 bps, and *hprt* with 657 bps. For this study we have selected the moderately sized human hypoxanthine-guanine phosphoribosyltransferase (*hprt*) gene. Both *in vivo* and *in vitro* mutation data are available. The use of this gene is further facilitated by the *hprt* database available to researchers (Cariello 1994).

Hprt is the most widely used target for the study of mutations in humans. It is a non-essential housekeeping gene, not known to produce any deleterious effect when inactive, except as germline mutation when it leads to the Lesch-Nyhan syndrome (Seegmiller *et al.* 1967). The *hprt* gene is located on the long arm of the X-chromosome and is either truly or functionally hemizygous, in males and females, respectively (Becker *et al.* 1979). The gene extends over 44 kilobases (kb) of genomic sequence (Edwards *et al.* 1990) while the processed mRNA is less than 1.6 kb, containing only 647 base-pairs (bp) of actual coding sequence.

Hprt mutations are readily selected using the toxic base analogue 6-thioguanine (6-TG). Mutants can be selected both *in vivo* and *in vitro*. *In vivo* selection depends upon the use of the T-cell clonal assay (Albertini *et al.* 1982) in which selection is carried out *ex vivo*. Activated T-cells are plated at limiting dilutions in the presence of the selective agent 6-thioguanine and resistant colonies are recovered. Similar selection can be carried out *in vitro* using most human and cultured mammalian cell lines. Mutants subsequently can be characterized at the molecular level (Yang *et al.* 1989).

The *hprt* target size, that is, the number of bp in the sequence available for mutation by single base-pair substitution (sbps) is unknown. The target gene consists of 657 bps of coding sequence which, due to the degeneracy of the genetic code, there are only 411 sites at which a base substitution produces an amino acid change. There are also 25 sites where a sbps will produce termination codons. We assume that the majority of these termination events will produce a selectable phenotype. However, as not all amino acid substitutions result in a selectable phenotype, the actual size of this mutational

target remains unknown. A germline missense mutation causing a reduction in the amount of enzyme per cell was reported to be asymptomatic (Fujimori *et al.* 1998). In reality the potential target size is larger than expected as mutations that affect splicing can also be selected (review, O'Neill *et al.* 1998). The existence of splicing mutations thus further complicates the calculation of the *hprt* target size. Despite the large number of mutants being characterized., saturation of the *hprt* mutational target has not yet been obtained. This is apparent from frequent reports of novel mutations. Determination of *hprt*'s actual mutational target size and consequently of its saturation status may be instructive.

Methods

A model was developed to estimate the number of available selectable sites within a genetic target. The hypothetical target was designed with 100 mutable positions. Mutations in that target were collected using a random number generator. Several sets of random numbers ranging from 1 to 100 were each matched with random numbers drawn from 1 to 20. The first set of numbers (1-100) represent recovered mutations along a mutational target that contains 100 mutable sites, while the second set of numbers (t=1-20) details a period for which the first set were randomly drawn. For each time period (1-20), the number of unique mutations (X_t) observed (each number being counted only once in each set), the total number of individual sites recovered (S_t) and the cumulative number of mutations (M_t) were determined. The fraction of unique sites (X_t/M_t) is calculated along with the natural log of the total mutations ($ln\ M_t$). The next set of data for the next time period was added to the first and the values recalculated. After examining all 20 time periods in this manner, the fraction of unique sites was plotted against the number of mutable sites and linear regression analysis performed (Statistica, StatSoft Inc., Tulsa). The number of mutable sites was then determined by extrapolating the Y-intercept at 0 percentage of unique sites. The number of mutations required to saturate this hypothetical mutational target was determined from the Y-intercept on a graph plotting the percentage of unique sites against the natural *ln* of the total mutations.

Three simulations using randomly generated numbers were examined. The first contained just 250 random numbers representing mutations. For the second simulation, a much larger set of numbers was drawn (n=1000). For the third simulation, 1000 random numbers were drawn, but the positions (sites) were not all equally mutable/obtainable. To simulate unequal site probabilities of mutation, the probabilities of drawing mutations at 10 of the 100 sites were reduced to $p = 0.0001$ and at an additional 10 sites the probabilities of mutation were set at $p = 0.001$. The remaining 80 sites had drawing probability of $p = .0123625$. The use of differential drawing probabilities simulates the fact that not all sites of a gene are equally to mutable.

The actual mutational data from the *hprt* dataset were handled in a similar fashion. The sbps mutations (excludes sbps occurring in intron sequences) were collected for a series of published human spontaneous *in vivo hprt* data sets (Curry *et al.* 1999). The dataset was partitioned by author and ordered by publication year. The novel fraction (unique sites / total mutations), number of mutated sites, and total number of mutations for the time periods was determined cumulatively as described above. Total mutable sites and number to saturate were calculated as described above. Analysis was repeated in an identical fashion using the sbps (no intron sbps) from the *hprt* mutant database (MutaSoft, Carielo) for *in vitro* mutations. This set of data contained mutants that were both spontaneous and induced by a wide variety of mutagenic treatments.

Results

Simulation 1: Under-saturation.

Using the set of 250 randomly generated numbers (representing mutations), the number of mutable sites and the number of mutations required to saturate those sites was determined from linear regressions. All 100 sites had equal drawing probabilities, representing equally mutable sites along the simulated mutational target. The linear regression for the mutable sites was calculated as $Y = 123 - 100. x X$, ($R^2 = 0.89$, $p < 10^{-5}$), with the y-intercept (123 ± 2 indicating the number of mutable sites. Saturation of these mutable sites occurred when 1018 mutations were collected as determined from the linear regression line ($Y = 6.9 - 2.9 x X$,

$R^2 = 0.97$, $p < 10^{-6}$). Repeating these analysis for other randomly generated data sets yielded nearly identical results (Data not shown).

Simulation 2: Over-saturation.

To over-saturate the simulated mutational target, 1000 random numbers were drawn and all 100 sites had equal probabilities of being drawn. The number of mutable sites predicted in this simulation was 101 as represented by the linear regression ($Y = 101 - 64 x X$, $R^2 = 0.93$, $p < 10^{-6}$). This simulation was clearly over-saturated, and saturation was reached at 640 ($e^{6.46}$) numbers or mutations as predicted by the linear regression ($Y = 6.46 - 3.1 x X$, $R^2 = 0.87$, $p < 10^{-6}$). After 640 simulated mutations were plotted, no new sites were found, and every site (n=100) had been "mutated".

Simulation 3: A biased mutational target.

This model is a more realistic representation of the situation for a mutational target. It is well established that mutation is a non-random event. Thus for a variety of reasons, some sites are recovered more frequently mutated than other sites. To model this situation, the probabilities of drawing some of the mutable sites (from 1 to 100) were altered such that 10 sites had a moderate probability ($p = 0.001$), another 10 had a low probability ($p = 0.0001$). The remaining 80 sites had a high probability relative to other sites ($p = 0.0124$). This simulation significantly underestimates the actual number of mutable sites (100). There were a total of only 86 mutable sites ($Y = 86 - 51 x X$, $R^2 = 0.95$, $p < 10^{-6}$) and saturation at 672 mutations ($e^{6.51}$) ($Y = 6.51 - 3.7 x X$, $R^2 = 0.81$, $p < 10^{-6}$). This simulation of unequal recovery can never be expected to reach full saturation. To draw just one of the low probability sites ($p = 0.0001$) would require 10000 mutants/draws. Consequently, millions of mutations (drawings) would be required to fully saturate this site-biased target.

***In vivo hprt* data:**

The human *hprt* database (MutaSoft, NC), containing 361 sbps, was ordered by publication (Table 1). Linear regression was used to estimate the number of mutable sites and mutations required to saturate, shown in Figures 1 and 2, respectively. Regression of the natural log of the total sites observed versus the novel fraction by time period, produced a significantly

better fit ($p = 10^{-4}$) for the data ($R^2 = 0.99$) than total sites versus the novel fraction ($R^2 = 0.87$). These data predict that 324 mutable sites exist for *hprt in vivo*, and that approximately 925 random mutations would be required to saturate the target.

***In vitro hprt* data:**

The *in vitro* sbps mutational data available from published collections (human *hprt* database (MutaSoft, NC)) listed by publication in Table 2. From this table the number of mutable sites and mutations required to saturate those sites were calculated. Regressions are shown in Figures 3 and 4. The number of mutable sites given by the *in vitro* data indicate that the target is 501 bp and that 1553 random mutations will saturate the target.

***In vitro* and *in vivo* data combined:**

As *in vitro* and *in vivo* regression analysis revealed different target sizes, these two datasets were combined to estimate total *hprt* target size. The combined datasets were order by publication year (Table 3). Linear regressions to determine the total target sites and mutations required to saturate those sites are shown in Figures 5 and 6, respectively. With this combined dataset, the model predicts that there are 425 mutable sites, which will be saturated with 1619 random mutations.

Authored Data Sets (t)	Unique (X)	Total Sites (S)	Cumulative Total (M)	Novel Fraction (X/M)
Rossi *et al.* 1990	10	11	12	.833
Recio *et al.* 1990a	17	20	24	.708
Andersson *et al.* 1992	20	25	31	.645
Vrieling *et al.* 1992	30	40	52	.577
Burkhart-Schultz *et al.* 1993	35	51	75	.467
Curry *et al.* 1993	34	51	76	.447
Hou *et al.* 1993	37	54	85	.435
Steingrimsdottir *et al.* 1993	40	60	95	.421
Curry *et al.* 1995.	41	67	112	.366
Shimahara *et al.* 1995	48	78	128	.375
Burkhart-Schultz *et al.* 1996	55	100	199	.276
Skandalis *et al.* 1997	59	106	211	.280
Karnauokova *et al.* 1997	61	113	236	.258
Podlutsky *et al.* 1998	63	115	239	.264
Osterholm *et al.* 1998	77	140	326	.236
Karnauokova *et al.* 1999	76	146	361	.211

Table 1. *In vivo* mutational data from a *hprt* database (Curry *et al.* 1999) listed by publication year.

Author	Unique Sites	Total Sites	Cumulative Total	Novel Fraction
Liber *et al.* 1989	4	4	4	1.000
Mattano *et al.* 1990	12	13	14	.857
Skopek *et al.* 1990	13	14	15	.867
Cariello *et al.* 1990	14	15	16	.875
Recio *et al.* 1990b	23	27	32	.719
Chen *et al.* 1990	39	49	62	.629
Chen *et al.* 1991	47	67	98	.480
Bronstein *et al.* 1991	65	101	161	.404
McGregor *et al.* 1991a	72	112	180	.400
Yang *et al.* 1991a	73	120	206	.354
Dorado *et al.* 1991	75	123	212	.354
Yang *et al.* 1991b	90	142	245	.367
McGregor *et al.* 1991b	98	160	288	.340
Keohavong *et al.* 1991	98	160	289	.339
Lukash *et al.* 1991	96	168	331	.290
Cariello *et al.* 1992	95	168	335	.284
Andersson *et al.* 1992	91	170	355	.256
Yang *et al.* 1993	88	177	407	.216
Giver *et al.* 1993	88	177	415	.212
Wang *et al.* 1993	99	194	457	.217
Guillouf *et al.* 1993	102	199	472	.216
Bastlova *et al.* 1993	101	200	478	.211
Papadopoulo *et al.* 1993	100	203	495	.202
Denault *et al.* 1993	103	207	506	.204
McGregor *et al.* 1994	108	213	525	.206
Yang *et al.* 1994a	110	218	541	.203
Nelson *et al.* 1994	110	219	552	.199
Yang *et al.* 1994b	100	226	612	.163
Taft *et al.* 1994	103	229	617	.167
Lightfoot *et al.* 1994	104	230	618	.168
Morgenthaler *et al.* 1995	100	230	643	.156
Lichtenauer-Kaligis *et al.* 1995a	98	231	660	.148
Lichtenauer-Kaligis *et al.* 1995b	99	234	681	.145
Bao *et al.* 1995	99	235	687	.144

Table 2. *In vitro* mutational data from a *hprt* database (Cariello *et al.* 1994) listed by publication year.

Figure 1. Prediction of Total *in vivo* Target Sites for *hprt*

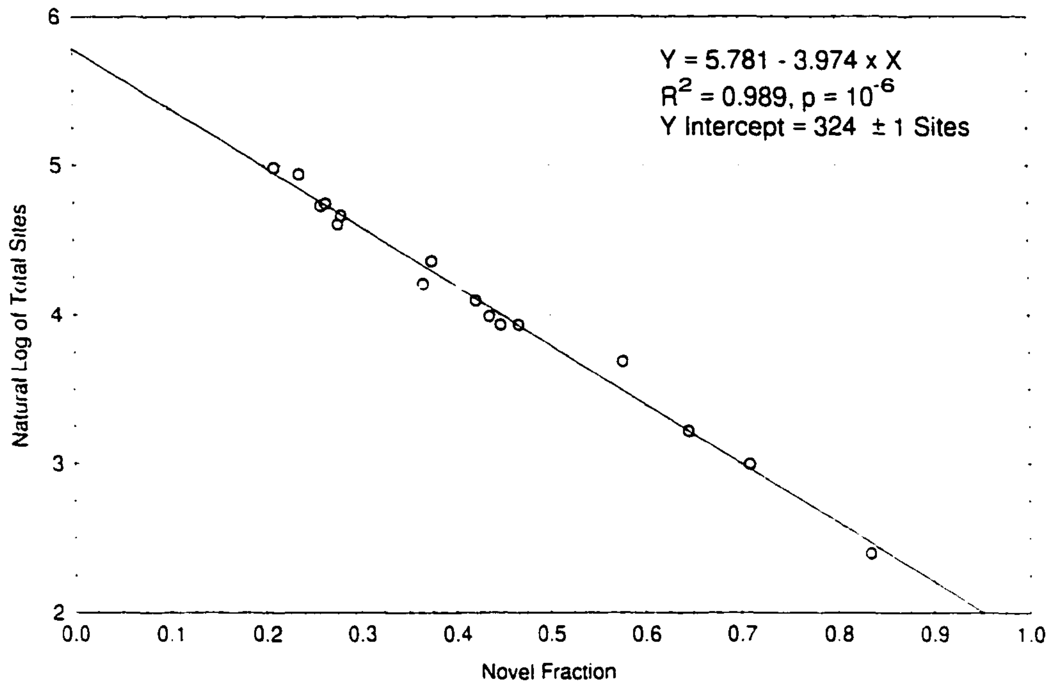


Figure 2. Mutations Required to Saturate the *in vivo hprt Target*

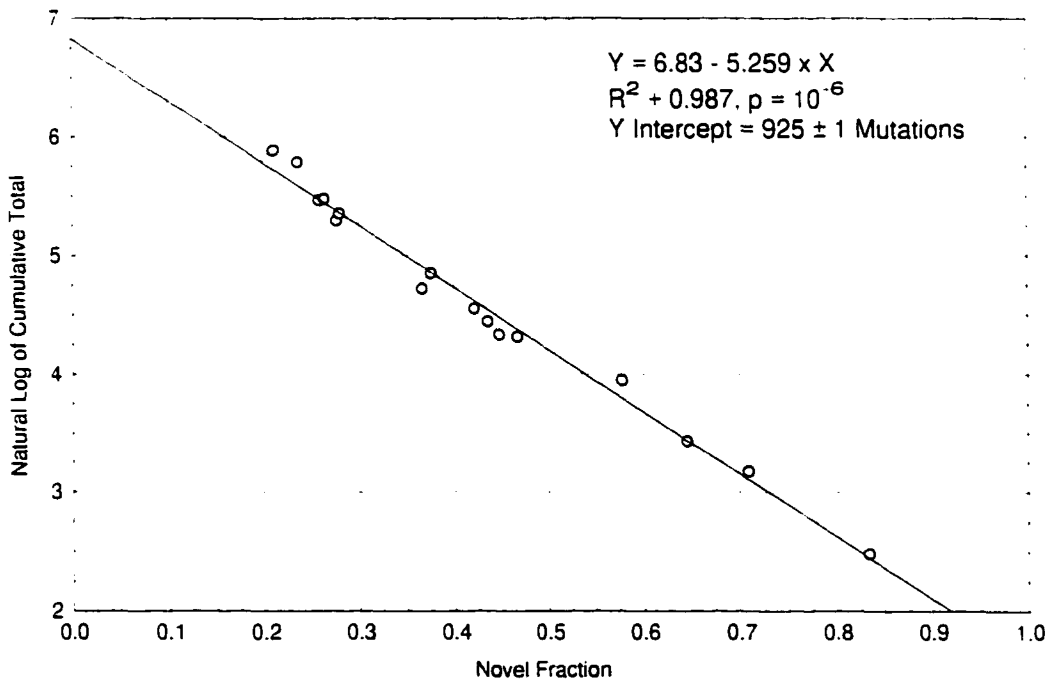


Figure 3. Prediction of Total *in vitro* Target Sites for *hprt*

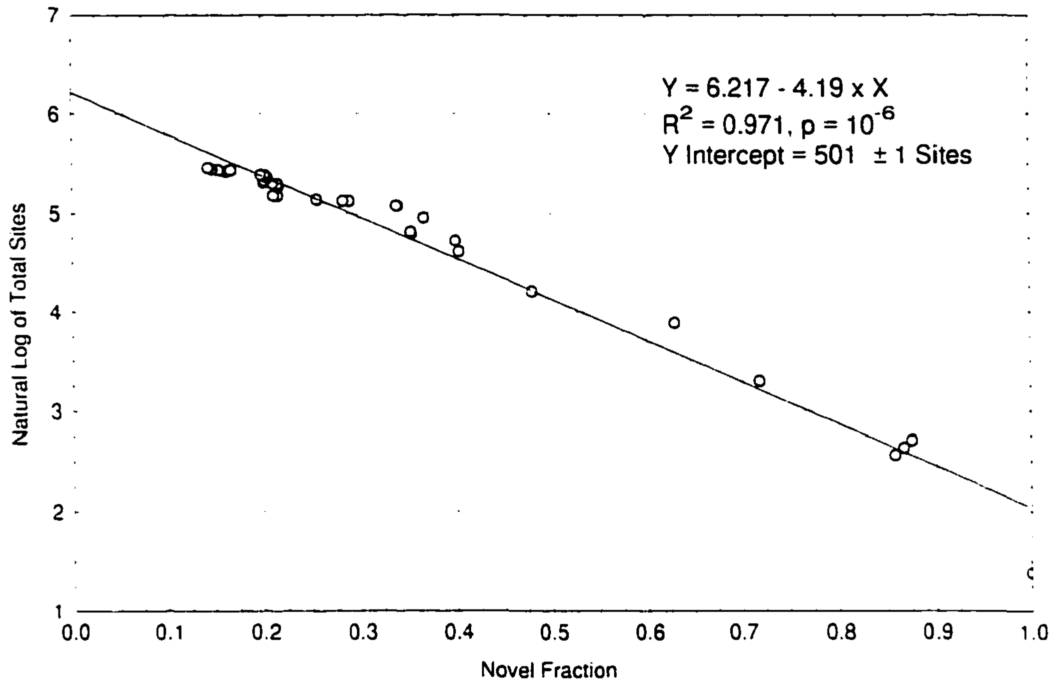


Figure 4. Mutations Required to Saturate the *in vitro hprt Target*

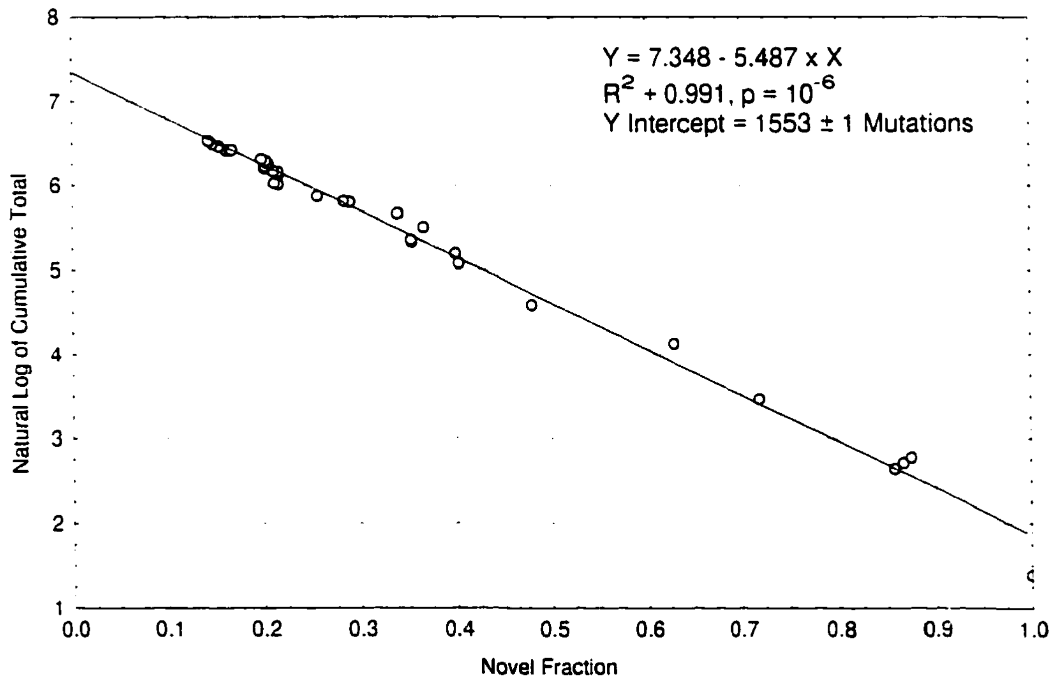


Figure 5. Prediction of target sites using both *in vitro* and *in vivo* data.

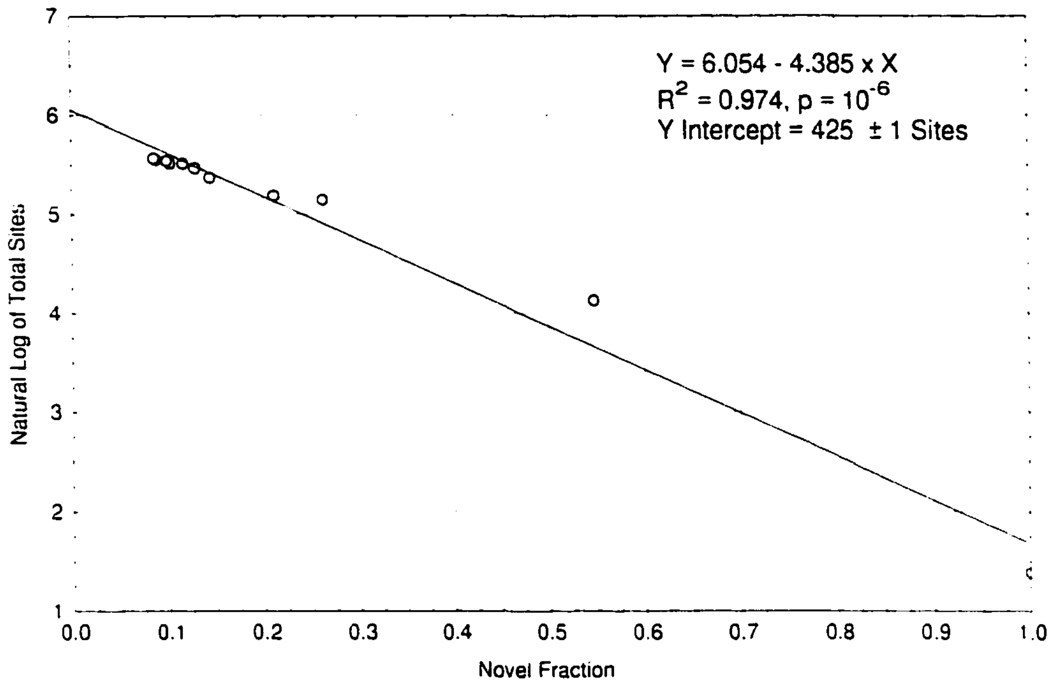
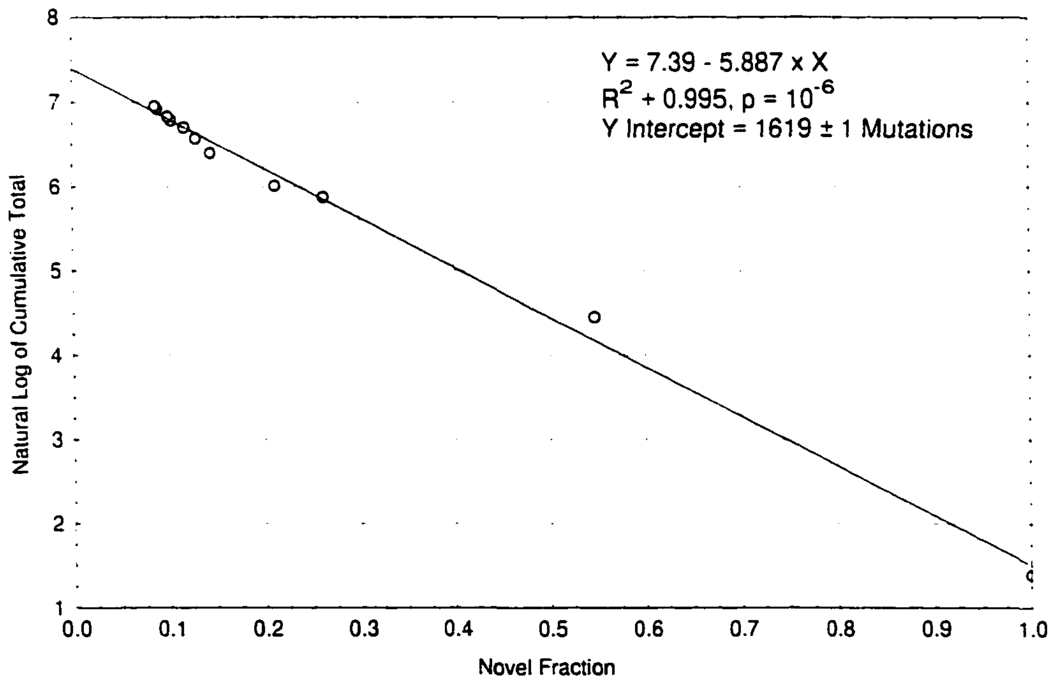


Figure 6. Mutations Required to Saturate the *hprt Target (All data)*



Discussion

Target saturation occurs when all selectable sites for mutation have been observed and no novel sites will be observed in any future collection using the same selection strategy. Testing for saturation can thus be as simple as assessing the contribution of novel sites in the literature. Our approach here involves comparing each reported mutational data set (by author) against the accumulated dataset for a single gene. To simplify this assessment only sbps in the coding sequence were considered, although in principal every base pair position is potentially selectable when frameshift events are considered.

As the *hprt* mutational target extends beyond the coding sequence, we do not yet have a complete understanding of this mutational target. This is partly due to our limited understanding of mRNA splicing, and in particular, how coding (which may not change the amino acid sequence) or non-coding sequences affect mRNA conformation during exon splicing. For this reason, we have excluded intron sbps from this assessment.

Surprisingly several collections demonstrate notably elevated percentages of novel point mutations: the Brazilian (50%, Skandalis *et al.* 1997), the Japanese collections (68.8%, Shimahara *et al.* 1995), and two Russian collections (43%, Chapter IV; 68% Khaidakov *et al.* 1997). Elevated levels of novel mutations within these datasets may be indicative of unique mutagen exposures or lifestyle differences. Equally plausible is that the total *hprt* mutant dataset is still too small so that new datasets will continue to present unique mutations. With a novel fraction of 0.086 point mutations being reported for the last point of the combined data subset, the *hprt* target is approaching saturation. However, the less frequently recovered sites are still missing, and saturation is still not imminent.

Linear regression based on the *in vivo* dataset predicted 324 mutable sites while the larger *in vitro* dataset suggested 501 sites. As noted in the simulation results, the model over-predicts the number of mutable sites when data are limited or under-saturated. However, as the actual *hprt* target is a biased target, the model likely under-predicts the actual number of target

sites available for mutation. Combining the *in vivo* and *in vitro* data increases the available data for the predictions (n=1048 mutations) and may better predict the target size. These combined data suggest that 425 sites exist for mutation and that 1600 random mutations will saturate the *hprt* target.

The use of unique mutations as an endpoint is useful as it indicates the rarity of mutation at some of the mutable sites. For a given selection procedure, the number of mutable sites is fixed but not necessarily known. However, the contribution of each mutable site is not equal, as mutation is a non-random event. The *in vivo* site utilization is lower than the *in vitro* utilization. This likely reflects the fact that the *in vitro* collection represents the sum of numerous different mutagenic treatments each with its own mutational spectrum. This diversity has the effect of increasing the rate of recovery of rare sites and hence apparently increasing the target size.

For a combination of reasons, some sites will be recovered only rarely and thus may be virtually unavailable without a very large number of mutants being characterized, preferably following exposure to diverse mutagens. Examination of the biased simulation reveals that the approach to saturation point is not linear nor is it likely obtainable. This does not defeat the potential value of target estimates for mutagenesis studies, but clearly warns of the complexity of genetic targets. Estimation of true target size will require an alternative approach such as site specific mutation such that each target position can be evaluated.

References

Albertini RJ, KL Castle and WR Borcherding, 1982 T-cell cloning to detect the mutant 6-thioguanine-resistant lymphocytes present in human peripheral blood. Proc. Natl. Acad. Sci. USA **79:** 6617-6621.

Andersson B, Falt S, Lambert B, 1992 Strand specificity for mutations induced by (+)-anti BPDE in the hprt gene in human T-lymphocytes. Mutat. Res. 269:129-140.

Bao CY, Ma AH, Evans HH, Horng MF, Mencl J, Hui TE, Sedwick WD, 1995 Molecular analysis of hypoxanthine phosphoribosyltransferase gene deletions induced by alpha- and X-radiation in human lymphoblastoid cells. Mutat. Res. 326:1-15.

Bastlova T, Andersson B, Lambert B, Kolman A, 1993 Molecular analysis of ethylene oxide-induced mutations at the HPRT locus in human diploid fibroblasts. Mutat. Res. 287:283-292.

Becker MA, RC Yen, P Itkin, SJ Goss, JE Seegmiller, B Bakay, 1979 Regional localization of the gene for human phosphoribosylpyrophosphate synthetase on the X chromosome. Science **203:**1016-1019.

Bronstein SM, Cochrane JE, Craft TR, Swenberg JA, Skopek TR, 1991 Toxicity, mutagenicity, and mutational spectra of N-ethyl-N-nitrosourea in human cell lines with different DNA repair phenotypes. Cancer Res. 51:5188-5197.

Burkhart-Shultz K, CB Thomas, CL Thompson, CL Strout, E Brinson, and IM Jones, 1993 Characterization of *in vivo* somatic mutations at the hypoxanthine phosphoribosyltransferase gene of a human control population. Environ. Health Perspect. **101:** 68-74.

Burkhart-Shultz KJ, CL Thompson, and IM Jones, 1996 Spectrum of somatic mutation at the hypoxanthine phosphoribosyltransferase (hprt) gene of healthy people. Carcinogenesis **17:** 1871-1883.

Cabral RE, AC Leitao, C Lage, A Caldeira-de-Araujo, M Bernardo-Filho, FJ Dantas, JB Cabral-Neto, 1998 Mutational potentiality of stannous chloride: an important reducing agent in the Tc-99m-radiopharmaceuticals. Mutation Res. **408:**129-135.

Cariello NF, Keohavong P, Kat AG, Thilly WG, 1990 Formaldehyde-induced and spontaneous alterations in human hprt DNA sequence and mRNA expression. Mutat. Res. 231:165-176.

Cariello NF, Swenberg JA, Skopek TR, 1992 In vitro mutational specificity of cisplatin in the human hypoxanthine guanine phosphoribosyltransferase gene. Cancer Res. 52:2866-2873.

Cariello NF, 1994 Database and software for the analysis of mutations at the human hprt gene. Nucleic Acids Res. **22**:3547-3548.

Chen R, Maher VM, McCormick JJ, 1990 Effect of excision repair by diploid human fibroblasts on the kinds and locations of mutations induced by (+-)-7 beta,8 alpha-dihydroxy-9 alpha,10 alpha-epoxy-7,8,9,10-tetrahydrobenzo[a]pyrene in the coding region of the HPRT gene. Proc. Natl. Acad. Sci. (USA) 87:8680-8684.

Chen RH, Maher VM, McCormick JJ, 1991 Lack of a cell cycle-dependent strand bias for mutations induced in the HPRT gene by (+-)-7 beta,8 alpha-dihydroxy-9 alpha,10 alpha-epoxy-7,8,9,10-tetrahydrobenzo(a)pyrene in excision repair-deficient human cells. Cancer Res. 51:2587-2592.

Curry J, A Skandalis, J Holcroft, J de Boer, and BW Glickman, 1993 Co-amplification of *hprt* cDNA and T-lymphocyte receptor sequences from 6-thioguanine resistant human T-lymphocytes. Mutation Res. **288:**269-275.

Curry J, GT Rowley, V Saddi, D Beare, J Cole, and BW Glickman, 1995 Determination of *hprt* mutant and mutation frequencies and the molecular characterization of human derived *in vivo* T-lymphocyte mutants. Environ. Mol. Mutagen. **25:**169-179.

Curry J, L Karnauokhova, GC Guenette, BW Glickman, 1999 Influence of sex, smoking,

and age on human *hprt* mutation frequencies and spectra. Genetics: In press.

Denault CM, Skopek TR, Liber HL, 1993 The effects of hypoxia and cysteamine on X-ray mutagenesis in human cells. II. hprt mRNA expression and cDNA sequence analysis of induced mutants. Radiat. Res. 136: 271-279.

Dorado G, Steingrimsdottir H, Arlett CF, Lehmann AR, 1991 Molecular analysis of ultraviolet-induced mutations in a xeroderma pigmentosum cell line. J. Mol. Biol. 217:217-222.

Edwards A, H Voss, P Rice, A Civitello, J Stegemann, C Schwager, J Zimmermann, H Erfle, CT Caskey, and W Ansorge, 1990 Automated DNA sequencing of the human HPRT locus. *Genomics* **6**:593-608.

Fujimori S, N Yamaoka, R Sakuma, M Hakoda, H Yamanaka, I Akaoka, N Kamatani, 1998 A first evidence of an asymptomatic germline missense base substitution in the hypoxanthine phosphoribosyltransferase (*hprt*) gene in humans. Adv. Exp. Med Biol. 431:319-322.

Giver CR, Nelson SL, Grosovsky AJ, 1993 Spectrum of spontaneous HPRT- mutations in TK6 human lymphoblasts. Environ. Mol. Mutagen. 22:138-146.

Guillouf C, A Laquerbe, E Moustacchi, D Papadopoulo, 1993 Mutagenic processing of psoralen monoadducts differ in normal and Fanconi anemia cells. Mutagenesis 8:355-361.

Hou SM, AM Steen, S Falt and B Andersson, 1993 Molecular spectrum of background mutation at the *hprt* locus in human T-lymphocytes. Mutagenesis **8**:43-49.

Karnaoukhova L, J Moffat, H Martins, BW Glickman, 1997 Mutation frequency and spectrum in lymphocytes of small cell lung cancer patients receiving etoposide chemotherapy. Cancer Res. **57**:4393-4407.

Karnaoukhova LG, KS Wilson, and BW Glickman, 1999 Mutant frequency and mutational spectra in T-lymphocytes of multiple myeloma patients receiving melphalan/prednisone chemotherapy. Submitted to Mutation Research.

Kawanishi M, T Matsuda, G Sasaki, T Yagi, S Matsui, H Takebe, 1998 A spectrum of mutations induced by crotonaldehyde in shuttle vector plasmids propagated in human cells. Carcinogenesis **19**:69-72.

Keohavong PK, Liu VF, Thilly WG, 1991 Analysis of point mutations induced by ultraviolet light in human cells. Mutat. Res. 249:147-159.

Khaidakov M, D Young, H Erfle, A Mortimer, Y Voronkov, BW Glickman, 1997 Molecular analysis of mutations in T-lymphocytes from experienced Soviet cosmonauts. Env. Molec. Mutagenesis. **30**:21-30.

Kraemer KH, MM Seidman, 1989 Use of supF, an Escherichia coli tyrosine suppressor tRNA gene, as a mutagenic target in shuttle-vector plasmids. Mutat Res **220**:61-72.

Liber HL, Benforado K, Crosby RM, Simpson D, Skopek TR, 1989 Formaldehyde-induced and spontaneous alterations in human hprt DNA sequence and mRNA expression. Mutat. Res. 226:31-37.

Lichtenauer-Kaligis EG, Thijssen J, den Dulk H, van de Putte P, Giphart-Gassler M, Tasseron-de Jong JG, 1995a UV-induced mutagenesis in the endogenous hprt gene and in hprt cDNA genes integrated at different positions of the human genome. Mutat. Res. 326:131-146.

Lichtenauer-Kaligis EG, Thijssen JC, den Dulk H, van de Putte P, Giphart-Gassler M, Tasseron-de Jong JG, 1995b Spontaneous mutation spectrum in the hprt gene in human lymphoblastoid TK6 cells. Mutagenesis 10:137-143.

Lightfoot T, Lewkonia RM, Snyder FF, 1994 Sequence, expression and characterization of HPRTMoose Jaw: a point mutation resulting in cooperativity and decreased substrate affinities. Hum. Mol. Genet. 3:1377-1381.

Lukash LV, Boldt J, Pegg AE, Dolan ME, 1991 Effect of O6-alkylguanine-DNA alkyltransferase on the frequency and spectrum of mutations induced by N-methyl-N'-nitro-N-nitrosoguanidine in the HPRT gene of diploid human fibroblasts. Mutat. Res. 250:397-409.

Mattano SS, Palella TD, Mitchell BS, 1990 Mutations induced at the hypoxanthine-guanine phosphoribosyltransferase locus of human T-lymphoblasts by perturbations of purine deoxyribonucleoside triphosphate pools. Cancer Res. 50:4566-4571.

McGregor WG, Maher VM, McCormick JJ, 1991a Kinds and locations of mutations arising spontaneously in the coding region of the HPRT gene of finite-life-span diploid human fibroblasts. Somatic Cell Molec. Genet. 17:463-469.

McGregor WG, Chen R, Lukash L, Maher VM, McCormick JJ, 1991b Cell cycle-dependent strand bias for UV-induced mutations in the transcribed strand of excision repair-proficient human fibroblasts but not in repair-deficient cells. Mol. Cell Biol. 11:1927-1934.

McGregor WG, Maher VM, McCormick JJ,1994 Kinds and locations of mutations induced in thehypoxanthine-guanine phosphoribosyltransferase gene of human T-lymphocytes by 1-nitrosopyrene, including those caused by V(D)J recombinase. Cancer Res. 54:4207-4213.

Morgenthaler PML, Holzhauser D, 1995 Analysis of mutations induced by 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine (PhIP) in human lymphoblastoid cells. Carcinogenesis 16:713-718.

Nelson SL, Giver CR, Grosovsky AJ, 1994 Spectrum of X-ray-induced mutations in the human hprt gene. Carcinogenesis 15:495-502.

O'Neill JP, PK Rogan, N Cariello, JA Nicklas, 1998 Mutations that alter RNA splicing of the human HPRT gene: a review of the spectrum. Mutations Res. 411:179-214.

Osterholm, A., and S. Hou, 1998 Splicing mutations at the HPRT locus in human T-lymphocytes in vivo. Environ. Mol. Mutagen. 32:25-32.

Papadopoulo D, Laquerbe A, Guillouf C, Moustacchi E, 1993 Molecular spectrum of mutations induced at the HPRT locus by a cross-linking agent in human cell lines with different repair capacities. Mutat Res. 294:167-177.

Pudlutsky, A., T. Bastlova, and B. Lambert, 1996 Reduced proliferation rate of hypoxanthine-phosphoribosyltransferase mutant human T-lymphocytes in vitro. Environ. Mol. Mutagen. **28:**13-18.

Recio L, J Cochrane, D Simpson, TR Skopek, JP O'Neill, JA Nicklas, and, RJ Albertini., 1990a DNA sequence analysis of *in vivo hprt* mutation in human T lymphocytes. Mutagenesis 5:505-510.

Recio L, Simpson D, Cochrane J, Liber H, Skopek TR, 1990b Molecular analysis of hprt mutants induced by 2-cyanoethylene oxide in human lymphoblastoid cells. Mutat. Res. 242:195-208.

Rossi AM, JCP Thijssen, AD Tates, H Vrieling, AT Natarajan, PH Lohman, and AA van Zeeland, 1990 Mutations affecting RNA splicing in man are detected more frequently in somatic than germ cells. Mutation Res. **244:**353-357.

Seegmiller JE, FM Rosenbloom, and WN Kelley, 1967 Enzyme defect associated with a sex-linked human neurological disorder and excessive purine synthesis. Science **155:** 1682-1683.

Shimahara H, T Kato, Y Hirai, and M Akiyama, 1995 Spectrum of *in vivo* mutations in T lymphocytes from atomic bomb survivors. I. Sequence alterations in cDNA. Carcinogenesis **16:**583-591.

Skandalis A, AD da Cruz, J Curry, A Nohturfft, M Curado, and BW Glickman, 1997 Molecular analysis of T-lymphocytes hprt mutations in individuals exposed to ionizing radiation in Goiania, Brasil. Environ Molec Mutagen **29**:107-116.

Skopek TR, Recio L, Simpson D, Dallaire L, Melancon SB, Ogier H, O'Neill JP, Falta MT, Nicklas JA, Albertini RJ, 1990 Molecular analyses of a Lesch-Nyhan syndrome mutation (hprtMontreal) by use of T-lymphocyte cultures. Hum. Genet. 85:111-116.

Steingrimsdottir H, G Rowley, A Waugh, D Beare, I Ceccherini, J Cole, and AR Lehmann, 1993 Molecular analysis of mutations in the *hprt* gene in circulating lymphocytes from normal and DNA-repair-deficient donors. Mutation Res. 294:29-41.

Taft SA, Liber HL, Skopek TR, 1994 Mutational spectrum of ICR-191 at the hprt locus in human lymphoblastoid cells. Environ. Mol. Mutagen. 23:96-100.

Vrieling H, JCP Thijssen, AM Rossi, FJ van Dam, AT Natarajan, AD Tates, and AA van Zeeland, 1992 Enhanced *hprt* mutant frequency but n significant difference in mutation spectrum between a smoking and a non-smoking human population. Carcinogenesis **13**:1625-1631.

Wang YC, Maher VM, Mitchell DL, McCormick JJ, 1993 Evidence from mutation spectra that the UV hypermutability of xeroderma pigmentosum variant cells reflects abnormal, error-prone replication on a template containing photoproducts. Mol. Cell. Biol. 13:4276-4283.

Yang JL, VM Maher, and JT McCormick, 1989 Amplification and direct sequencing of cDNA from the lysate of low numbers of diploid human cells. Gene **83**:347-354.

Yang JL, Chen RH, Maher VM, McCormick JJ, 1991a Kinds and location of mutations induced by (+/-)-7 beta,8 alpha-dihydroxy-9 alpha,10 alpha-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.

Yang JL, Hu MC, Wu CW, 1991b Novel mutational spectrum induced by N-methyl-N'-nitro-N-nitrosoguanidine in the coding region of the hypoxanthine (guanine) phosphoribosyltransferase gene in diploid human fibroblasts. J. Mol. Biol. 221:421-430.

Yang JL, Lin JG, Hu MC, Wu CW, 1993 Mutagenicity and mutational spectrum of N-methyl-N'-nitro-N-nitrosoguanidine in the hprt gene in G1-S and late S phase of diploid human fibroblasts. Cancer Res. 53: 2865-2873.

Yang SC, Lin JG, Chiou CC, Chen LY, Yang JL, 1994a Mutation specificity of 8-methoxypsoralen plus two doses of UVA irradiation in the hprt gene in diploid human fibroblasts. Carcinogenesis 15:201-207.

Yang JL, Lee PC, Lin SR, Lin JG, 1994b Comparison of mutation spectra induced by N-ethyl-N-nitrosourea in the hprt gene of Mer+ and Mer- diploid human fibroblasts. Carcinogenesis 15:939-945.

VII. Discussion

For the monitoring of mutation in human populations the *hprt* T-cell clonal assay is currently the most effective system available. However, it does suffer from significant drawbacks that limit its resolution for the task at hand. These limitations are of two distinct origins. The first is related to the biological complexity of the T-cell system. The second is the need for a large sample of characterized mutations, in order to assess events distinguishable above the spontaneous background.

In several independent surveys, the *hprt* T-cell clonal assay failed to detect any mutational signature that distinguishes itself from the spontaneous background. The failure of the system to detect clear mutational signatures from even tobacco smoking is perplexing, particularly since smoking-related DNA adducts can be detected. This observation highlights the limitations of the system in terms of target, and more importantly, of the tissue utilized (Chapter VI).

Several studies have examined the effects of ionizing radiation on exposed human populations using the T-cell clonal assay. Each has failed to demonstrate significant change in mutational spectrum of exposed populations compared to control populations. Potential reasons for these failures are detailed below.

Our ionizing radiation studies are centered on the Goiânia radiological accident. During the course of this accident, a large population was exposed to considerable ionizing radiation from a breached ^{137}Cesium source. The accident occurred in September of 1987, but the first sample collection took place much later in March of

1990. Results from this first sample collection clearly demonstrated that the exposed group (average external exposure 1.7 Gy) had considerably increased MF relative to the unexposed control group (Skandalis *et al.*, 1997).

However, analyses of mutational spectra obtained from the exposed and control populations failed to identify any significant differences, although the small sample size may have been a limiting factor. Any spectral analysis performed on mutational spectrum is obviously statistically limited by sample size. Similarly, a second radiation study, which provided a limited number of mutants, was of limited value for assessing mutational spectra differences (Skandalis *et al.* 1995). High costs and limited size of study groups are critical limiting factors.

A subsequent study of the same exposed population employing nearly double the original mutants from the exposed group also failed to demonstrate any significant difference compared to the control population (da Cruz and Glickman 1997). The failure of this study to reveal changes in the exposed mutational spectra was reported earlier. Monitoring the exposed population over the following three years (three different sampling times) revealed that T-cells have a half-life of only 2.1 years (da Cruz et al. 1996). Thus, mutants induced by the radiation exposure in 1987 may not have been present in the subjects for sampling times approximately three, four and five years following exposure. This observation clearly limits the use of the T-cell clonal assay for monitoring populations more than two years after the initial exposure. Recent results further erode the usefulness of the assay by suggesting that the half-life of CD4+

and CD8+ T-cells may be even shorter at 87 and 77 days, respectively (Hellerstein *et al.* 1999).

Studies of the mutagenic potential of two powerful chemotherapeutic agents (etoposide - Karnauokhova *et al.* 1997, Melphalan – Karnauokhova *et al.* 1999) also failed to demonstrate any mutational spectra differences between exposed and pre-exposed patients. These studies suggest that these chemotherapeutic agents elicit apoptosis in T-cells so that any potentially induced mutations would not be recoverable.

Another radiation study, which analyzed T-cell mutants, recovered from experienced Russian cosmonauts again suffered from a restricted number of mutations (Khaidakov *et al.* 1997). Comparison of a base substitution spectrum from cosmonauts (mutants=26) with a similar spectrum obtained from the HPRT database (mutants=215), revealed a significant excess of A:T→GC in the Russians who all lived in Moscow. Such a conclusion is highly speculative and thus somewhat risky considering the limited number of mutants obtained from the cosmonauts.

Failure of the T-cell clonal assay to detect mutagen exposures is also due to our limited understanding of spontaneous mutations in humans. Complexity of spontaneous mutational spectra in humans necessitates a vast amount of characterized mutations. Mutational targets with few mutable sites may require fewer mutations to present a complete mutational spectrum. However, as target size increases, so does the number of mutations required for an accurate representation of any mutational spectrum whether induced or spontaneous. Complex mutational targets such as the *hprt* gene, though rich in terms of the wide variety of mutational classes that can be recovered

require a considerable number of characterized mutations to realize the nature of the spontaneous spectrum.

In addition, the *hprt* target is not fully saturated with mutations (Chapter VI). Even when considering mutations collected from a broad range of both *in vivo* and *in vitro* studies, new site mutations are still regularly reported in the literature. The ability of the *hprt* gene to produce aberrant splicing products clearly demonstrates that the effective target size of this and any other spliced gene extends into the non-coding sequence.

Ideally any given mutagen would leave a unique "fingerprint" of mutation that could be clearly observed above background. It will only be in cases where the mutagen exposure has been great enough or perhaps chronic, that mutagen specific changes will be visible above the spontaneous background. It is clear that some mutagens are context specific, causing mutations at unique positions along the mutation target.

To complicate matters further, the genome's inherent ability for DNA repair has the effect of masking mutagen exposures. Repair itself can mutate the target, and leave mutations that appear spontaneous in nature, and thus may not be visible above the spontaneous background.

Genetic control of mutation frequency is clearly suggested by analysis of the Russian monozygotic twins. As detailed in chapter III, this genetic control likely weakens with age. Older twin sets did not demonstrate this remarkable tight relationship and we speculated that with time, environmental contributions mask the genetic component of mutation frequencies.

Although the mutant spectrum obtained from the Russian twins was significantly different when compared to the Western spectrum, the relatively small numbers of mutants provides the ever present caveat. When the mutational spectrum from Russian subjects is increased by adding in the mutations collected from Russian cosmonauts (Khaidakov *et al.* 1997) to that of the twins, the difference in A:T→C:G transversions disappears. Cosmonaut data (five subjects) contain only a single such event, while the twin data (14 subjects) contains twelve. It is only during the comparison of the larger age-match Western data that this mutational class is revealed as significantly different. The cosmonaut population is significantly older than the twin populations (30.0 ± S.D. 1.4 versus 47.4 ± S.D. 7.5). When this older Russian spectrum is compared to an aged matched Western spectrum, a significant difference is again realized for the A:T→C:G transversions. Comparison of the younger Western with cosmonaut spectra reveals that frequency of these transversions is identical. A significant difference in this transversion class was noted in the age partition spectrum (Chapter V) where it increases in frequency with age. For the Russian subjects it appears to decrease with age. With the Western spectral data, there are signicantly large numbers of mutations, making the observation reasonably sound, while limited numbers of Russian mutations may hinder the observation in that group. Clearly, further numbers of such mutants need to be added to Russian and even to the Western spectra to further substantiate these observations.

Mutational spectra analysis has extreme power for mutagen detection in certain systems. In its current state the *hprt* T-cell clonal assay's usefulness in monitoring environmental exposures in humans is questionable. While the mutagenic effects of high mutagen doses are detectable in animal models, the effects of environmental mutagens occurring at relatively low doses may not be detectable in humans. Although this may not be the case for populations challenged with significant environmental exposures. While some mutagens, notably UV light, leave hallmark mutational patterns, their prevalence above the spontaneous mutational background is difficult to detect. Mutational spectra containing large numbers of mutants provide a way of overcoming this difficulty.

As detailed in chapter II, a better mutational assay needs to be developed. Ideally, a system employing a readily accessible tissue where mutants can be selected without *ex vivo* growth would be an improvement. In addition, current systems require the viability of the mutants and as such, lethal mutations cannot be assessed. A system that would permit the recovery of lethal mutations at genetic targets should be considered.

As detailed in Appendix II, the elongation factor 2 (*ef*-2) gene demonstrated some prospect. Mutants with diphtheria toxin resistance mutations were recoverable, although most appeared not able to propagate. Such mutations were restricted to those that altered the diphtheria toxin binding site of *ef*-2. From the analysis of single mutants, the potential role of exon skipping in leading to diphtheria toxin resistance was realized. Exon skipping can be the result of mutations occurring in or around the exon

containing the diphthamide codon required for diphtheria poisoning. Thus, diphtheria resistance mutations clearly are not restricted to just the single diphthamide codon, and the potential mutational target size for the *ef*-2gene is much broader.

Unfortunately, analysis of the *ef*-2 exon skipping mutations did not reveal any point mutations that may have been the cause of the aberrant splicing events. Such mutations may have been contained within the skipped exon, and are thus not recoverable. As only the coding sequence is known for the human *ef*-2 gene, intron primers could not be constructed to characterize the skipped exons, and potentially reveal the causative mutations. However, use of primers within the exon sequences did permit the characterization of two relatively short *ef*-2 intron sequences. An attempt was made to obtain the *ef*-2 genomic sequence. A shotgun sequencing strategy was proposed to fully sequence the genomic sequences. However, much to our surprise, the genomic library clone that was obtained and characterized with Southern blot analysis did not contain the *ef*-2 genomic sequence. Having the full genomic sequence of *ef*-2 was critical to investigating the potential cause of the aberrant exon splicing mutations. Had *ef*-2 intron primers been available we would have further pursued the ef-2 target as a potentially new mutational target.

Monitoring human populations for mutations caused by environmental exposures is limited by the currently used mutational assay. Due to the complex nature of T-cells, the *hprt* clonal assay is not ideal for this task. Mutational spectra from exposed populations containing limited numbers of mutants are unlikely to reveal

mutagen exposures. The significant increase in A:T→C:G transversions occurring with age may be an important observation of spectral change occurring in adult human populations. The relevance of the frequency shift of this transversion remains to be explored.

Despite the criticisms detailed in this dissertation, there is a need for *in vivo* mutation analysis in human populations. The *hprt* mutational target, although highly complex, has proven useful although the limitations of the employed tissue are just now being realized. T-cells are particularly practical for mutational studies as they contain unique T-cell receptor sequences that permit the discovery of clonal events. In addition, the large database of *hprt* mutants, both *in vivo* and *in vitro*, will be invaluable for future spectral comparisons. The development of a cell sorting process for the selection of *hprt*⁻ mutants would negate the requirement for *ex vivo* growth. In combination with large scale automated DNA sequencing technologies, cell sorting would thus permit the rapid characterization of mutants and generation of sufficiently large mutational spectra. However, this dissertation concludes that both the *hprt* assay and new mutational assays need to be developed and implemented to effectively monitor mutations in human populations. Fortunately, many of the tools required are becoming increasingly available.

References

Albertini RJ, KL Castle, WR and Borcherding, 1982 T-cell cloning to detect the mutant 6-thioguanine-resistant lymphocytes present in human peripheral blood. Proceedings of the National Academy of Science: USA. **79**:6617-6621.

Albertini RJ, JP O'Neill, JA Nicklas, NH Heintz, and PC Kelleher. 1985 Alterations of *the hprt* gene in human *in vivo*-derived 6-thioguanine-resistant T lymphocytes. Nature **316**:369-371.

Albertini RJ, JA Nicklas, TR Skopeck, L Recio, and JP O'Neill. 1998 Genetic instability in human T-lyphocytes. Mutation Res. **400**:381-389.

Beare DM, KE Aldridge, MR O'Donovan, and J Cole. 1993 An improved procedure for the in vitro expansion of human T-lymphocyte clones for mutant analysis. Mutation Res. **291**: 207-212.

Becker MA, RC Yen, P Itkin, SJ Goss, JE Seegmiller, and B Bakay, 1979 Regional localization of the gene for human phosphoribosylpyrophosphate synthetase on the X chromosome. Science **203**: 1016-1019.

Burkhart-Schultz KJ, CB Thomas, CL Thompson, CL Strout, E Brinson, and IM Jones. 1993 Characterization of *in vivo* somatic mutations at the hypoxanthine phosphoribosyltransferase gene of a human control population. Environ. Health Perspect. **101**:68-74.

Burkhart-Schultz KJ, CL Thompson, and IM Jones. 1996 Spectrum of somatic mutation at the hypoxanthine phosphoribosyltransferase (*hprt*) gene of healthy people. Carcinogenesis **17**: 1871-1883.

Curry J, A Skandalis, J Holcroft, J de Boer, and BW Glickman. 1993. Co-amplification of *hprt* cDNA and T-lymphocyte receptor sequences from 6-thioguanine resistant human T-lymphocytes. Mutation Research **288**:269-275.

Curry J, GT Rowley, V Saddi, D Beare, J Cole, and BW Glickman. 1995 Determination of *hprt* mutant and mutation frequencies and the molecular characterization of human derived in vivo T-lymphocyte mutants. Environmental and Molecular Mutagenesis **25**:169-179.

da Cruz AD, J Curry, MP Curado, and BW Glickman. 1996 Monitoring *hprt* mutation frequency over time in people accidentally exposed to high doses of ionizing radiation. Environmental and Molecular Mutagenesis **27**:165-175.

da Cruz AD, BW Glickman. 1997 Nature of mutation in the human *hprt* gene following *in vivo* exposure to ionizing radiation of cesium-137. Environmental and Molecular Mutagenesis **30**: 385-395.

de Boer JG, J Curry, and BW Glickman, 1993. A fast and simple method to determine the clonal relationship between human T-lymphocytes. Mutation Research **288**:173-180.

Edwards A, H Voss, P Rice, A Civitello, J Stegemann, C Schwager, J Zimmermann, H Erfle, CT Caskey, and W Ansorge, 1990 Automated DNA sequencing of the human HPRT locus. Genomics **6**: 593-608.

Gibbs RA, PN Nguyen, A Edwards, AB Civitello, and CT Caskey, 1990 Multiplex DNA deletion detection and exon sequencing of the hypoxanthine phosphoribosyltransferase gene in Lesch-Nyhan families. Genomics **7**:235-244.

Hellerstein M, MB Hanley, D Cesar, S Siler, C Papageorgopoulos, E Wieder, D Schmidt, R Hoh, R Neese, D Macallan, S Deeks, JM McCune, 1999 Directly measured kinetics of circulating T lymphocytes in normal and HIV-1-infected humans. Nature Med. **5**:83-89.

Karnaoukhova, L, J Moffat, H Martins, and BW Glickman, 1997 Mutation frequency and spectrum in lymphocytes of small cell lung cancer patients receiving etoposide chemotherapy. Cancer Res. **57**:4393-4407.

Karnaoukhova, LG, KS Wilson, and BW Glickman, 1999 Mutant frequency and mutational spectra in T-lymphocytes of multiple myeloma patients receiving melphalan/prednisone chemotherapy. Submitted to Mutation Research.

Khaidakov M, DYoung, H Ertle, A Mortimer, Y Voronkov, BW Glickman, 1997 Molecular analysis of mutations in T-lymphocytes from experienced Soviet cosmonauts. Environmental and Molecular Mutagenesis **30**:21-30.

Kronenberg A, S Gauny, K Criddle, D Vannais, A Ueno, S Kraemer, and CA Waldren, 1995 Heavy ion mutagenesis: linear energy transfer effects and genetic linkage. Radiat Environ Biophys **34**:73-8.

Nicklas JA, JP O'Neill, RJ Albertini, 1886 Use of T-cell receptor gene probes to quantify the *in vivo* *hprt* mutations in human T-lymphocytes. Mutation Res. **173**:67-72.

Nicklas JA, JP O'Neill, LM Sullivan, TC Hunter, M Allegretta, BF Chastenay, BL Libbus, and RJ Albertini, 1988 Molecular analyses of *in vivo* hypoxanthine-guanine phosphoribosyltransferase mutations in human T-lymphocytes: II. Demonstration of a clonal amplification of *hprt* mutant T-lymphocytes *in vivo*. Environ Mol Mutagen **12**:271-284.

Seegmiller JE, FM Rosenbloom, and WN Kelley, 1967. Enzyme defect associated with a sex-linked human neurological disorder and excessive purine synthesis. *Science* **155**:1682-1683.

Skandalis A, AD da Cruz, J Curry, M Curado, and BW Glickman, 1997 Molecular analysis of T-lymphocyte hprt⁻ mutations from individuals exposed to ionizing radiation in Goiania, Brasil. Environmental and Molecular Mutagenesis **29**:107-116.

Skandalis A, J Curry, JP O'Neill, JA Nicklas, RJ Albertini, BW and Glickman, 1995 Analysis of point mutations in the *hprt* gene of cancer patients treated with radioimmunoglobulin therapy. Environmental and Molecular Mutagenesis **26**:213-217.

Steingrimsdottir H, C Rowley, G Dorado, J Cole, and AR Lehmann, 1992 Mutations which alter splicing in the human hypoxanthine-guanine phosphoribosyltransferase gene. Nucleic Acids Res. **20**:1201-1208.

Stout JT, and CT Caskey, 1985 HPRT: Gene structure, expression and mutation. Annual Review of Genetics **19**:127-148.

Vrieling H, JC Thijssen, AM Rossi, FJ van Dam, AT Natarajan, AD Tates, and AA van Zeeland, 1992 Enhanced *hprt* mutant frequency but no significant difference in mutation spectrum between a smoking and a non-smoking human population. Carcinogenesis **13**:1625-1631.

Waters TR, and PF Swann, 1997 Cytotoxic mechanism of 6-thioguanine: hMutSalpha, the human mismatch binding heterodimer, binds to DNA containing S6-methylthioguanine. Biochemistry **36**:2501-2506.

Yang JL, VM Maher, and JT McCormick, 1989 Amplification and direct sequencing of cDNA from the lysate of low numbers of diploid human cells. Gene **83**:347-354.

Mutation Research 304: 19-32 (1994)

International Commission for Protection Against Environmental Mutagens and Carcinogens Working Paper No. 2

Spontaneous Mutations In Mammalian Cells

Barry W. Glickman, Verra A. Saddi and John Curry

Centre for Environmental Health and the Department of Biology, University of Victoria, Victoria, British Columbia, Canada

Summary

Spontaneous or background mutation in mammals plays an important role in both medical and evolutionary contexts. However, establishing mutation frequencies or rates has not always been easy. When the field of mammalian mutagenesis was in its infancy, the word "variant" rather than "mutant" was often used because the genetic nature of the observed phenotypic alteration could not be adequately proven. Nowadays numerous target genes have been identified in which mutant frequencies can be measured, and occasionally even rates can be estimated. Indeed, the genetic basis for 'variants' now often comes from direct DNA sequencing. This review describes the most often used and best understood genetic markers for mutation research and examines their usefulness. In addition, mutational specificity is compared for several loci and the use of DNA-sequence data in determining the origins of spontaneous mutation is also discussed. An important observation is that spontaneous mutation frequencies of similar sized genes can vary by more than an order of magnitude. Chromosomal location, the nature of the gene product and mutational specificity may offer partial explanation.

Introduction

Spontaneous mutation has long been of interest for its impact on human health in terms of both birth defects and such diseases as cancer and arteriosclerosis as well as aging. The recent molecular analysis of mutations known to play a major role in the development of cancer has greatly stimulated our interest (Bishop, 1991). As the fields of molecular epidemiology and mutational analysis develop, the frequency and nature of mutation will provide insights into the molecular origin and biological consequences of mutation. Indeed, there has been a significant expansion of research dealing

©1994 with permission from Elsevier Science Publishing B.V.

with the molecular nature of mutation. These studies deal with increasingly complex genetic systems. The earliest papers centred on the bacteriophage T4 *rII* system (e.g., Benzer, 196*) and the *lacI* gene of *E. coli* (Miller *et al.*, 1978; Schaaper *et al.*, 1986), but soon included eukaryotic systems such as the *sup4*-o and URA3 genes of yeast (Giroux *et al.*, 1988; Lee *et al.*, 1988). Drosophila also soon became the subject of molecular analysis. In mammalian cells, a number of genetic loci including Na/K ATPase (ouabain resistance), *aprt*, *dhfr*, *hprt* are being investigated (Thacker *et al.*, 1978; Mitchell *et al.*, 1986; de Jong *et al.*, 1988; Phear *et al.*, 1989; Recio *et al.*, 1990a; Rossi *et al.*, 1990). A number of genes have also been introduced into mammalian cells on retroviral vectors, e.g., *gpt*, *hprt* and *aprt* genes (e.g., Ashman and Davidson, 1987; Ikehata *et al.*, 1989; Skandalis *et al.*, 1992). The stage has now been achieved that mutation is being analyzed in human populations *in vivo*. Mutation frequencies and the molecular characterization of mutations is being carried out at the *hprt*, HLA and glycophorin A (GPA) loci. Furthermore, the detailed analysis of mutations in people demonstrating genetic disorders is also providing insights into the nature of mutation.

©1997 Elsevier Science Publishing B.V.

The resultant data indicates that the spectra of spontaneous mutation not only differ between prokaryotes and eukaryotes, but among genes in mammalian cells. Such variation may reflect differences in cell type or origin, culture conditions, the size and location of genes, and chromosomal organization. Repair (maintenance) may differ with location, chromosomal structure and timing, as well as level of expression. This may be of particular relevance since the efficacy of DNA repair may differ throughout the genome (Smith and Mellon, 1990).

Spontaneous Mutation - Towards A Definition

The concept of spontaneous mutation requires some consideration. Historically, it has described mutations arising in the absence of any specified mutational treatment. This definition implies that spontaneous mutation encompasses those mutations whose origins can not be accounted for by exogenous factors. This view presents little difficulty when considering mutational data generated *in vitro*. While growth condition, media composition and pH, the presence of antioxidants, rates of cell division and similar factors may serve to modulate mutation, their origin can not be

associated with specific exogenous genetic challenges and the term "spontaneous" fits well. Similarly, mutations, which arise over the course of evolutionary time can be considered to be spontaneous. Their origin can not be specifically accounted for. It would appear therefore that mutations, which occur "at arms length", are readily viewed as spontaneous.

In contrast, mutations occurring *in vivo* in the absence of specific inducing treatments may be more complex than properly accounted for by the term "spontaneous". This is most readily seen in examining data from human populations where in the absence of specific treatments, life style, geographical location and genetic variability may affect both the frequency and nature of the mutations recovered. Here, the term "background" mutation may be more appropriate. Indeed, in *in vivo* experiments, such as those involving transgenic mice, numerous exogenous factors can be seen to modulate mutation in the whole animal and in a tissue specific fashion. It is for example, unlikely that a liver where mixed function oxidases have been induced will produce the same spontaneous mutation rate and spectrum as a liver where these enzymes are less active. We must therefore be cognisant of the fact that the terms spontaneous and background mutations may be used interchangeably, but often carry different implications.

Classes of Mutation

In considering both rates and nature of mutation, information must be provided concerning the nature of the genetic target. Information is required concerning the nature of the assay. Is it a forward or reverse assay? Reverse assays are often restricted to a particular class of events. Such events may be a specific base change or a frameshift or may include a larger target assessing "second site" revertants. Even forward mutational systems must be well described. For example, ouabain resistance involving an essential gene, provides a window on mutation restricted to base substitutions (Thacker *et al.*, 1978). In contrast, the *tk* or *hprt* loci, being non-essential, permit the recovery of practically all classes of mutational events.

A listing of the potential classes of mutational events is provided (see Table 1). The information in this table has been adapted from a similar table by Claes Ramel (1989). Epimutations have been discussed in detail by Holliday (1991). One such

©1997 Elsevier Science Publishing B.V.

Classes
Point mutations[a]
-base substitutions
missense
nonsense
splice junction
-frameshifts
-small deletions or insertions
Deletions
-intragenic
-multilocus
Translocations
Inversions
Duplications
Insertions[b]
-known, nearby sequences
-mobile genetic elements
Non-disjunction
Recombination
-somatic crossing-over
-gene conversion
Reduction to homozygosity
-non-disjunction, gene conversion
Amplification
-Sister-chromatid unequal exchange
-unequal somatic crossing-over
-disproportionate or specific DNA - replication
Methylation moderated gene expression (an epigenetic event)

Table 1: Classes of mutational and non-mutational events which may be recovered in mammalian systems (adapted from Ramel, 1991). [a]Point mutations have been defined as mutations which can not be detected by Southern blotting. The limits of detection thus varies with the size of the gene target being examined. In the case of *aprt* this can be as few as 25 bp (Grosovsky et al., 1988) compared to *hprt* mutations where events involving less than 500 bp are difficult to detect in genomic DNA (Ramel, 1989). [b] Insertion events have been divided into two classes because these mutations arise via very different mechanisms

class of events involves the suppression of expression by the methylation of cytosine at CpG sites. There are several examples of genes being reactivated by 5-azacytidine (5-AC) which causes demethylation (Jones and Taylor, 1981). These include the *tk* gene (Clough *et al.*, 1982; Tasseron-de Jong *et al.*, 1989) the pro^- and metallothionine deficient phenotypes of CHO cells (Holliday and Ho, 1990). None of the 70 spontaneous mutants recovered in the *aprt* gene in CHO cells (de Jong *et al.*, 1986) were reverted by 5-AC treatment (Drobetsky and Glickman, unpublished). In contrast, 2 of 14 spontaneous mutants recovered by Zhang and Jennsen (1991a) at the *hprt* locus in V79 cells were reverted by 5-AC. Their data indicate extensive system differences between the role of methylation in gene inactivation and reinforces the need for caution in interpreting mutational results. Epimutation related to DNA methylation and gene expression has been recently reviewed by Holliday (1991).

Mutation Rates

Drake (1991) has recently reviewed the question of mutation rates in six genetic systems including bacteriophages M13, lambda, T2.T4, *E. coli*, *S. cerevisiae* and *N. crassa*. The method used to calculate mutation frequencies is derived from the fluctuation test of Luria and Delbruck (1943) in which many replica cultures are grown from one mutant-free inoculum. Genome size varied by 6500.

©1997 Elsevier Science Publishing B.V.

while mutation rates per base varied by 16,000-fold. The mutation rate per genome remained remarkably constant at 0.0033 mutations per genome per replication throughout the organisms tested. These results confirm Drake's much earlier observation (Drake, 1969).

The question of rates in mammalian cells is much less clear. Nevertheless, it is unlikely that Drake's rule concerning the constancy of genome mutation rates can operate in mammalian cells. Replication accuracy would have to be so high as to be uneconomical. The human haploid complement is 3.1 x 10^9 bp. In order to maintain the same genomic mutation rate as found by Drake (1991), an error rate of $<10^{-13}$ per base pair replicated would be required. Thus, in the case of a 1000 bp gene, a mutation rate of 7.5 x 10^{-10} per replication would be required in order to be comparable. This is many times larger than actually observed (see below). Even when a correction is made for the full sized *hprt* gene including its introns, a mutation rate of 3 X 10^{-8}/replication would be expected. This is at least two orders of magnitude higher than the observed frequency.

Genetic Targets

Non-Selected Loci

Several new techniques have been designed to reduce our dependence upon phenotypic selection. These include Density Gradient Gel Electrophoresis (DGGE) (Cariello *et al.*, 1988, 1990) and Single-Strand Conformational Polymorphisms (SSCP) (Murakami *et al.*, 1991; Orita *et al.*, 1989; Suzuki *et al.*, 1991). While each of these techniques bring the promise of more effective and precise mutational determinations, our current data is almost exclusively derived from phenotypically selected mutations.

Selectable Loci

Several loci have been used in the development of mutational testing. These include markers of drug resistance, toxin resistance and metabolic deficiencies.

Aprt (adenine phosphoribosyl transferase)

The *aprt* gene is smallest of the salvage pathway genes being only 2.5 kb (Lowy et al, 1980), and thus most readily analyzed. Its DNA sequence has been established (Nalbantaglu *et al.*,

©1997 Elsevier Science Publishing B.V.

1986; de Boer *et al.*, 1989b). Despite its small size, it is a 'normal' gene complete with 5 exons. Its chromosomal location is 16q24 in humans; Z4,Z7 in CHO cells. Both hemizygous and heterozygous cell lines are available which facilitates both mutation selection and analysis. In addition, its size permits the analysis of the complete gene or just the coding region. However, despite its numerous advantages, the *aprt* locus has several shortcomings which are a reflection of its small size. It has been found to be rather insensitive to some agents because of the absence of target sequences. For example, the *aprt* gene is practically refractory to mutation by CC-1065 and several of its derivatives because the run of adenines to which the drug is targeted is absent from the target sequence. In contrast, the *hprt* gene has several such sites where mutations are indeed recovered (Mazur, Aaron and Glickman, personal communication). *Aprt* is similarly a less effective target for ionizing radiation because it presents a smaller target for deletion and rearrangement events. Its autosomal recessive nature has as an additional consequence that it can only be used to monitor mutation in people who are heterozygous at the *aprt* locus. This will restrict its application in the area of human monitoring.

©1997 Elsevier Science Publishing B.V.

Hprt

Hypoxanthine-guanine phosphoribosyltransferase

The *hprt* gene also codes for an enzyme in the purine nucleotide salvage pathway and is thus non-essential. It is a largish gene (44 kb in humans, 34 kb in mice; Konecki *et al.*, 1982; Melton *et al.*, 1986; Patel *et al.*, 1986) with nine exons. Mutations in *hprt* have been studied following the enzymatic amplification of c-DNA (749 bp) although this often leaves splice junction events unresolved without looking at the sequences of the intron exon borders in the genomic DNA. The *hprt* gene is not likely to be too close to an essential gene as large deletions and rearrangements have been detected (Turner *et al.*, 1985). This is in marked contrast to the aprt hetrozygotes system where only polar mutations are recovered as multilocus deletions, indicating that an essential gene is located to one side of aprt.

A marked advantage of the *hprt* locus is its location on the X-chromosome. This makes the *hprt* gene hemizygous in males and functionally hemizygous in females. Hprt mutations can thus essentially be analyzed in any cell line without regard to gender. Peripheral blood T-lymphocytes

have been most frequently examined in humans (Albertini *et al.*, 1982; Morley *et al.*, 1983). This makes *hprt* an attractive locus to study *in vitro* as well.

A unique advantage is that this gene has been studied in a number of mammalian systems allowing inter-species comparisons (Stout and Caskey, 1985). The fact that a defect in this gene in people *in vivo* is responsible for two human diseases, Lesch-Nyhan and gouty arthritis means that human population data is also available (Davidson *et al.*, 1988a,b, 1989a,b; Nicklas *et al.*, 1989; Fujimora *et al.*, 1988, 1990). Using the *hprt* system it is therefore possible to both *in vivo* and *in vitro* results, monitor human populations and to compare germ line and somatic mutation as well.

HLA Complex

The HLA complex is a polymorphic set of genes occupying about 6,000,000 bp on the short arm of chromosome 6. Three classes of these genes can be distinguished. These are (1) HLA-A, HLA-B and HLA-C genes coding for surface antigens that are expressed on the surface of most nucleated cells; (2) genes coding for cell surface antigens expressed primarily by B-cells, macrophages and dendritic cells; and (3) genes coding for portions of a complement pathway (for a review see Bodmer, 1984). The HLA system permits the detection of 'variants' or mutants using specific antibodies plus complement to kill cells expressing a specific surface antigen. The survivors of such a treatment are thus defective in a given antigen. HLA1, encoded for a gene of 5 kb and containing 7 exons has been extensively studied (Janatipour *et al.*, 1988; Turner and Morely, 1990; McCarron *et al.*, 1989).

Dhfr dihydrofolate reductase

DHFR is a housekeeping enzyme involved in the conversion of folate and dihydrofolate into its active reduced form, tetrahydrofolate. This compound is an essential component of the biosynthetic pathways for glycine, purine and thymidine nucleotides. In humans the gene is located in chromosome 5 and is about 30 kb with 6 exons while in CHO cells it is located on chromosome 2 and only 25 kb in length (Mitchell *et al.*, 1986). It is a chromosomally interesting site because in CHO cells its 5' end has several DNAase I hypersensitive sites and its CpG sites are unmethylated. The remaining 22 kb has not DNAase I sensitive sites and is well methylated.

Very little spontaneous data exists for

©1997 Elsevier Science Publishing B.V.

DHFR and only 7 spontaneous mutations have been isolated (Carothers, personal communication). The reason for the lack of data stems from the high cost of the selection which depends upon [^{3}H]deoxyuridine. This has restricted the assay to induce mutation where the frequencies are elevated. A noteworthy point about the DHFR system however, is the strength of the reverse selection since less than 1% of enzyme activity is required for selection.

Mutations In Mammalian Cells

In Table 2a & b we have compiled mutant frequencies obtained both *in vitro* and *in vivo* for diverse loci in a variety of mammalian cells. The

Cells	M.F. *in vivo* $\times 10^{-6}$	SYSTEM	M.F. *in vitro* $\times 10^{-6}$	Cells
[a]T-cells [b]T-cells [c]T-cells newborns [d]T-cells	2.50-20 14.80 0.64 10.40	HPRT	1.10 20.00 1.00	[g]CHEF 18-1D-3 [h]V79 [i]diploid human fibroblasts
		APRT	1.80	[m]CHO
[e]lymphocytes neonates elderly [f]lymphocytes [p]lymphocytes A2 A3	7.10 65.30 29.90 30.8 4.7	HLA-A		
		gpt	13.20	[n]AS52 CHO
		DHFR	17.00	[k]UA21 CHO
		TK	Small Colony 54.00 Large Colony 53.00 1.5-5	[l]L5178Y mouse lymphoma [q]L5178Y mouse lymphoma [j]TK6
[d]erythrocytes N0 [s]erythrocytes M0 N0 NM NN	10.9 10.0 11.0 10.0 17.0	GPA		
[t]Erythrocytes Beta-Globulin Gene Cluster	0.038	HBB		

Table 2a: Notes to table. [a]Turner et al. 1985. Mean M.F. range as determined from 7 healthy adult human males. [b]Vijayalaxmi & Evans 1984. Mean M.F. as determined from 12 male and 12 female humans aged 19 to 79 years of age. [c]McGinniss et al. 1990. Mean M.F. as determined from 37 male and 8 female human newborns. [d]Tates et al. 1991. Mean M.F. as determined from 111 humans. [e]Grist et al. 1991. Mean M.F. for both HLA-A2 and A3 as determined from 167 assays in 73 humans. [f]McCarron et al. 1989. Mean M.F.as determined from 83 assays in 21 humans. [g]Kaden et al. 1989. [h]Vrieling et al. 1985. [i]Yang et al. 1991. [j]Tindall et al.1989. [k]Carothers et al. 1986. [l]Cifonne et al. 1987. [m]Grosovsky et al. 1986. [n]Yandell et al. 1990. [o]Tates et al. 1989. [p]Janitipour et al. 1988. [q]Clive, personal communication ($54 \pm 19 \times 10^{-6} \pm 1$ std. Dev.). [r]Jensen et al. 1978. [s]Langlois et al. 1989. [t]Seshadri et al. 1987. Calculated as the mean 0 117 individuals. [u]Gupta and Siminovitch 1980. [v]Gupta and Goldstein 1981. [w]Gupta 1984. [x]Gupta and Siminovitch 1978a,b. [y]Jensen et al. 1987. [z]Gupta and Siminovitch 1976. [aa]Baker et al. 1974. [bb]Gupta and Siminovitch 1980. [bb]Gupta and Singh 1982. [cc]Gupta and Siminovitch 1978a. [dd]Gupta and Goldstein 1981. [ee]Stanely and Siminovitch 1976. [ff]Gupta and Siminovitch 1978b. [gg]Elmore et al. 1983. [hh]Seshadri et al. 1987.

©1997 Elsevier Science Publishing B.V.

System	M.F. x 10^{-6} in vitro	Cell type	Genetic Properties
DRB	0.44	CHO	Target?
Dichlororibofuranosyl benzimidazol	14.0	Human fibroblasts	essential
	0.5	CHO	codominant
	0.1 – 3.6	[bb]CHO	
DT^R or Dip^R	1 – 10	Human fibroblasts	ef-2
Diptheria Toxin	1 – 8	Human fibroblasts	essential
	0.05	CHOP	codominant
Emetine^R	0.2 – 0.5	CHO	40S Ribosome
	0.57	CHO	essential
	0.1 – 2.30	CHO	recessive
Ouabain^R	10 – 12	Mouse L cells	Na^+ K^+ ATPase
	0.75	CHO	essential
	0.2 – 6.0	[bb]CHO	recessive
MBG	0.65	CHO	Target?
	0.1 – 2.6	[bb]CHO	essential recessive
Toyocamycin	0.15 – 1000	CHO	Adenosine
	670	CHO	kinase, non-essential Recessive
Podophyllotoxin	1.74	CHO	Microtubules Essential codominant

Table 2b: Selected *in vitro* mutant frequencies.

range in forward mutant frequencies where the entire gene is the target is 0.64 x 10^{-6} obtained by McGinnis *et al.* (1990) at the *hprt* locus in new born infants. While adult mutation frequencies obtained for *hprt* in vivo range from 2.5 to 20 x 10^{-6} consistent with the observation that mutations accumulate with age (Vijayalaxmi and Evans, 1984). Mutation frequencies for this genetic locus obtained *in vitro* are comparable, ranging from 1 to 20 x 10^{-6}, regardless whether from human T-cells, fibroblasts, CHEF/18-1D-3 cell line derivatives or Chinese hamster V79 cells.

Whereas mutation frequencies for the *hprt* locus seem to fall within a narrow range, the frequencies of mutants recovered at the *tk* locus covers a significant range. Data from Cifonne *et al.*, (1987) and D. Clive (personal communication) indicate that the mutation combined mutation frequency for both large and small colonies falls is 55 x 10^{-6}. There is no obvious reason for the much lower mutation frequency found with the human TK6 cell line (Yandell *et al.*, 1990). Overall,

©1997 Elsevier Science Publishing B.V.

mutation frequencies fall into a remarkably narrow range (Table 3).

Mutational Spectra In Mammalian Cells

Table 4 summarizes mutational spectra obtained in a number of systems at the DNA sequence level. Among base substitutions, G:C → A:T transitions are the most predominant change seen in *hprt in vivo* and *aprt in vitro* although this is not the case for *hprt* in Chinese hamster cells (Zhang *et al.*, 1992). Overall however, transversions are more common than transitions, although there are four independent pathways by which transversions can occur compared with only two for transitions. In most spectra some transversions appear to be poorly recovered, but it is not always the same transversion. Among the human medical genetics cases, the A:T → C:G event was very rare while in the V79 data set, the A:T → T:A transversion is rarely recovered. This differs from the *E. coli lacI* experience (Schaaper *et al.*, 1986; Schaaper and Dunn, 1991; Holliday and Glickman, 1991) in which the G:C → A:T transition also predominates but the G:C → C:G transversion event is exceedingly rare. The *aprt* data set of de Jong *et al.*, (1988) have not been included in Table 4. A G:C → A:T hotspot at position 241 predominated the mutational spectrum. This transition hotspot was absent from the *aprt* mutant collection of Phear *et al.* (1987). The subsequent analysis of several mutant *aprt* collections by Glickman and his colleagues failed to yield mutations at this hotspot. Since this site was also a hotspot for UV mutagenesis, the possibility that fluorescent light exposures was responsible for the hotspot was considered, and following short exposures to TL-lighting, hotspot mutations were recovered (Drobetsky and Glickman, personal communication). This makes the initial findings of the 241 hotspot a likely artefact reflecting the specificity of UVA light as a mutagen.

Cell Type	Target Selection	Mutations cell Generation x 10[illegible]
CHO	Emetine	0.49
Mouse L cells	Ouabain	0.05 – 0.06
Human fibroblasts:		
normal (KD)		0.38 –0.85
transformed (KD-11A)		0.6 – 1.3
CHO	DRB	0.44
Human fibroblasts	Diptheria	0.5 – 0.6
Human fibroblasts		0.1 – 0.3
CHO	PHA	1.5
CHO	Toy	38
Human fibroblasts:	HPRT	
normal (KD)		1.6
transformed (KD-11A)		1.1 – 1.8
Human lymphocytes	HPRT	
normal		0.15
neoplastic		0.53 – 6.7

Table 3: Mutation rates in mammalian cells. References in Table 2a.

Deletions appear to be relatively rare in the

©1997 Elsevier Science Publishing B.V.

Mutation Type	Lesch-Nyhan And Gout [a]	Human [b] T-cell *hprt*	CHO Aprt [c]	CHO Gpt [d]	CHO cDNA Hprt [e]	V79 Hprt [f]
Total number	113	99	103	100	99	119
Base changes	72	38	61	26	24	64
Transitions						
G:C → A:T	26	13	25	7	0	9
A:T → G:C	7	3	10	5	2	12
Transversions						
G:C → T:A	12	10	12	0	8	13
G:C → C:G	10	6	10	7	0	13
A:T → T:A	12	3	2	5	10	1
A:T → C:G	5	3	2	2	4	16
Frameshifts	10	16	7	2	38	5
Deletions	17	6	20	67	19	13
Duplications	0	0	6	0	8	3
Compounds	2	0	1	0	2	0
Other	0	0	4	5	8	1
Splice	12	39	4	-	-	33

Table 4: Mutational Spectra at a several endogenous and shuttle vector loci. [a]Pooled data from Wilson *et al.*, 1983a,b; Wilson and Kelly, 1983; 1984; Davidson *et al.*, 1988a,b,c; 1989a,b; Fujimori *et al.*, 1988; 1989; 1990; Cariello *et al.*, 1988, Gibbs *et al.*, 1989, 1990. [b]Rossi *et al.*, 1990. [c]Phear *et al.*, 1989. [d]Ashman and Davidson, 1987. [e]Ikehata *et al.*, 1989. [f]Zhang *et al.*, 1992.

hprt data of Zhang *et al.* (1992). This is consistent with *hprt* results reported by Nicklas et al (1989) in which events detectable by Southern blot analysis accounted for 14% (48/326) mutants analyzed. These results are in contrast to the results of Turner et al. (1985) who reported that more than half (12/21) of spontaneous *hprt* mutations recovered *in vivo* in human T-lymphocytes involved genetic events detectable by Southern blotting. Similarly, an analysis of *hprt* gene mutations in cultured B-lymphoblastoid cell lines indicated that more than a third of them (39/85) involved large structural alterations (Gennett and Thilly, 1988). A very high frequency of deletion events (151/171) involving the entire locus was noted by Yandell *et al.* (1990) for the *tk* locus while Dewyse and Bradley (1989) reported a similar observation for the *aprt* gene in CHO cells where deletions were involved in 97% of the 198 mutations examined.

A total of 76 independent spontaneous mutations in the *hprt* gene of V79 cells were analyzed in 2 different laboratories (Zhang et al 1992). The two sets of 17 and 59 mutants, respectively, were isolated under different growth conditions and the cDNA was analyzed by PCR and direct DNA sequencing. The two sets produced similar spectra (see Table 5). The majority of mutations were base substitutions (>60%), with

©1997 Elsevier Science Publishing B.V.

MUTATION TYPE	Set 1	%	Set 2	%
Base Changes	7	41	31	53
Transitions	1	6	11	19
Transversions	6	35	20	34
Frameshifts	0	0	2	3
Deletions	2	12	8	14
Duplications	0	0	2	3
Others	0	0	0	0
Splice Mutations	8[a]	47	16[a]	27
TOTALS	17	100	59	100

Table 5: Comparison of spontaneous spectra for *hprt* in V79 cells from two different laboratories. [a] Of the 8 splice mutants in set 1, 5 involve the loss of exon 4; 5 of the 17 splice mutants in set 2 involved the same loss. The Chi^2 value is 1.99 ($p>0.15$).

transversions being twice as prevalent as transitions. A:T → T:A transversions were rarely recovered. Deletions comprised only 13% of the mutants. Strikingly, about one-third of all the mutants resulted in aberrant splicing. Of these, most involved exon 4. Among the spontaneous events, splice junction alterations leading to the aberrant processing of the primary transcript is the largest contributor to mutation.

The Nature of Mutation, the DNA Target and Splice Site Events

There have been many discussions of the nature of mutation at the protein level in prokaryotic cells. One of the most detailed examined the nature of more than 15,000 spontaneous mutations in the *lacI* gene (Gordon *et al.*, 1988; Holliday and Glickman, 1991). This study demonstrates that the nature of mutations selected depends upon the function of the protein target. In general, when an enzymatic function is involved, mutations which result in the production of a nonsense mutation or events at glycine codons (GGX) predominate. The data recovered is thus quite skewed. In the case of the *de novo* formation of nonsense mutations, all events with the exception of the A:T → G:C transition can be detected and considering the nature of the glycine codon, events at G:C sites (notably transitions) tend to predominate. In contrast, in structural regions or sites such as the protein DNA binding domain of the *lacI* gene, almost any change which affects the amino acid composition of the binding domain can be detected. Hence, the fraction of events at potential nonsense sites and GGX sequences is reduced. Spectra in such domains may be more reflective of events at the DNA sequence level.

Mutations in eukaryotes have an extra dimension to their genetic target when compared to prokaryotes. This is related to the recovery of mutations of RNA processing rather than

©1997 Elsevier Science Publishing B.V.

functionality. A high fraction of mutational events at the *hprt* locus affects RNA transcript processing. This has been found by Zhang *et al.* (1992) in V79 cells, Recio *et al.* (1990a) and Rossi *et al.* (1990, 1992) in human T-lymphocytes. In contrast, only a small fraction of the mutants recovered in the *aprt* gene (de Jong *et al.*, 1988; Phear *et al.*, 1989) involve splice junctions. These results indicate the importance of splice junctions as mutational targets. This is especially the case for induced mutations when the target for DNA damage is consistent with the consensus splice sites such as seen for MNU (Zhang and Jennsen, 1991b) or Cis-Platina (de Jong and Glickman, 1989a). Some data indicate that these sequences may occasionally be hotspots for mutational events. Cases in point are the high recovery of splice mutations affecting exon 4 in the *hprt* gene (Zhang *et al.*, 1992) and exon 5 of the *dhfr* gene (Mitchell *et al.*, 1986). These results have a major consequence. They indicate that individual gene sensitivity will be affected by the structure of the target gene.

Mutation Frequencies in Malignant and Transformed Cells

As described above, mutation frequencies can be affected by a number of factors including ploidy, cell type, age, nutritional state, oxygen stress, and stage of replication. It has also been suggested that error rates increase during neoplastic transformation. Thus mutation frequencies in in vitro systems which may rely on transformed cells may produce different mutation frequencies than are seen in vivo. An indication of the extent of this effect can be seen in Table 2 which compares in vivo and in vitro data for selected genetic targets.

Among this line Elmore *et al.* (1983) compared spontaneous mutation rates in normal and chemically transformed human skin fibroblasts. Two genetic loci were studied. These were thioguanine resistance involving the *hprt* locus and Ouabain-resistance involving the Na^+ K^+ - ATPase locus. The mutation rates determined using the Luria-Delbruck fluctuation assay were between 1.6 - 2.1×10^{-6} in the *hprt* assay and $3.8 - 8.5 \times 10^{-8}$ in the Oub^r assay. No differences in mutation rates between the normal and the transformed cell lines was observed, indicating that transformation itself is

©1997 Elsevier Science Publishing B.V.

not necessarily related to the alteration in mutation rates observed in neoplastic cells. In contrast, a study by Seshadri *et al.* (1987) comparing TG^r – resistant rates from 3 normal and 3 malignant lymphocyte cell lines indicated a very significant increase in mutation rate. The normal cell lines had mutation rates of 24.6, 15, and 5.5 x 10^{-8}, while the malignant lines had rates of 666, 53, and 131 x 10^{-8}.

Conclusions

This survey of spontaneous mutation in mammalian cells demonstrates that the frequency of mutation varies considerably from system to system. This variation may reflect differences in the nature and location of the genetic target, cell type, growth and mutant selection conditions as well as other experimentally defined, and not so well defined factors. Additional insight into the nature of spontaneous mutation can often be gained from studies of mutation specificity. Spectra provide information that may pinpoint the sources of mutation, and often account for major discrepancies between systems. However viewed, the variation in mutant frequencies and rates means that information on spontaneous mutation in mammalian cells will have to be used intelligently.

©1997 Elsevier Science Publishing B.V.

References

Albertini, R.J., K.L. Castle and W.R. Borcherding, 1982 T-cell cloning to detect the mutant 6-thioguanine-resistant lymphocytes present in human peripheral blood. Proc. Natl. Acad. Sciene (U.S.A.) 79: 6617-6621.

Ashman, C.R., and R.L. Davidson, 1987 Sequence analysis of spontaneous mutations in a shuttle vector gene integrated into mammalian chromosomal DNA. Proc. Natl. Acad. Sci. USA 84: 3354-3358.

Baker, R.M., D.M. Brunette, R. Mankovitz, L.H. Thompson, G.F. Whitmore, L. Siminovitch and E. Till, 1974 Ouabain-resistant mutants of mouse and hamster cells in culture. Cell 1: 9-21.

Benzer, S., 1961 On the topography of genetic fine structure. Proc. Natl. Acad. Sci. (U.S.A.) 47: 403-415.

Bishop, J. M. 1991. Molecular themes in oncogenesis. Cell 64: 235-248.

Bodmer, W., 1988 Somatic cell genetics and cancer. Cancer Surv. 7: 239-250.

Cariello, N.F., J.K. Scott, A.G. Kat, W.G. Thilly, and P. Keohavong, 1988. Resolution of a missense mutant in human genomic DNA by denaturing gradient gel electrophoresis and direct sequencing using in vitro DNA amplification: $HPRT_{Munich}$. Am. J. Hum. Genet. 42: 726-734.

Cariello, N.F., P. Keovang, A.G. Kat and W.G. Thilly, 1990 Molecular analysis of complex human cell populations. Mutational spectra of MNNG and ICR-191. Mutation Res. 231: 165-176.

Carothers, A.M., G. Urlaub, N. Ellis, and D. Grunberger. 1986. Proc. Natl. Acad. Sci. USA 83:6519-6523.

Cifone, M.A., B. Myhr, A. Eiche, and G. Bolcsfoldi, 1987 Effect of pH shifts on the mutant frequency at the thymidine kinase locus in mouse lymphoma L5178Y TK^{+-} cells. Mutation Res. 189:39-46.

Clough, D.W., L.M. Kunkel, and R.L. Davidson 1982 5-azacytidine-induced reactivation of a Herpes simplex thymidine kinase gene. Science 216, 70-73.

Davidson, B.L., M. Pashmforoush, W.N. Kelley, and T.D. Palella. 1988a. Genetic basis of hypoxanthine-guanine phosphoribosyltransferase deficiency in a patient with the Lesch-Nyhan syndrome ($HPRT_{Flint}$) Gene 63:331-336.

Davidson, B.L., S.-C. Chen, J.M. Wilson, W.N. Kelley, and T.D. Palella. 1988b. Hypoxanthine-guanine phosphoribosyltransferase: Genetic evidence for identical mutations in two deficient subjects. J. Clin. Invest. 82:2164-2167.

Davidson, B.L., T.D. Palella, and W.N. Kelley. 1988c. Human hypoxanthine-guanine phosphoribosyltransferase: a single nucleotide substitution in cDNA clones from a patient with Lesch-Nyhan syndrome ($HPRT_{Midland}$). Gene 68:85-91.

Davidson, B.L., M. Pashmforoush, W.N. Kelly, and T.D. Palella. 1989a. Human hypoxanthine-guanine phosphoribosyltransferase deficiency: The molecular defect in a patient with gout ($HPRT_{Ashville}$). J. Biol. Chem. 264: 520-525.

Davidson, B.L., S.A. Tarie, T.D. Palella, and W.N. Kelly. 1989b. Molecular basis of hypoxanthine-guanine phosphoribosyltransferase deficiency in ten subjects determined by direct sequencing of amplified transcripts. J. Clin. Invest. 84:342-346.

©1997 Elsevier Science Publishing B.V.

de Boer, J.G., and B.W. Glickman, 1989a Sequence specificity of mutation induced by the anti-tumor drug cisplatin in CHO *aprt* gene. Carcinogenesis 10: 1363-1367.

de Boer, J.G., E.A. Drobetsky, A.J. Grosovsky, M. Mazur, and B.W. Glickman, 1989b The chinese hamster *aprt* gene as a mutational target. Its sequence and an analysis of direct and inverted repeats. Mutation Res. 226: 239-244.

de Jong, P.J., A.J. Grosovsky, and B.W. Glickman. 1988. Spectrum of spontaneous mutation at the APRT locus of Chinese hamster ovary cells: An analysis at the DNA sequence level. Proc. Natl. Acad. Sci. USA 85: 3499-3503.

Dewyse, P., and W.E.C. Bradley. 1989. High-frequency deletion event at aprt locus of CHO cells: Detection and characterization of endpoints. Somat. Cell Mol. Genet. 15:19-28.

Drake, J.W. 1969 Comparative rates of spontaneous mutation. Nature 221, 1132.

Drake, J.W. (1991) A constant rate of spontaneous mutation in DNA-based microbes. Proc. Natl. Acad. Sci. (USA) 88, 7160-7164.

Elmore, E., T. Kakunaga, and J.C. Barret, 1983 Comparison of spontaneous mutation rates of normal and chemically transformed human skin fibroblasts. Cancer Res. 43: 1650-1655.

Fujimori, S., Y. Hidaka, B.L. Davidson, T.D. Palella, and W.N. Kelley. 1988. Identification of a single nucleotide change in a mutant gene for hypoxanthine-guanine phosphoribosyltransferase ($HPRT_{Ann\ Arbor}$) Human Genet. 79: 39-43.

Fujimori, S., B.L. Davidson, W.N. Kelley, and T.D. Palella. 1989 Identification of a single nucleotide change in the hypoxanthine-guanine phosphoribosyltransferase gene ($HPRT_{Yale}$) responsible for Lesch-Nyhan syndrome. J. Clin. Invest. 83: 11-13.

Fujimori, S., N. Kamantani, Y. Nishida, N. Ogasawara, and I. Akaoka, 1990 Hypoxanthine-guanine phosphoribosyltransferase deficiency: nucleotide substitution causing Lesch-Nyhan syndrome identified for the first time among Japanese, Hum. Genet. 84: 483-486.

Gennett, I.N., and W.G. Thilly, 1988 Mapping large spontaneous deletion endpoints in the human HPRT gene. Mutation Res. 201: 149-160.

Gibbs, R.A., P.-N. Nguyen, L.J. Mcbride, S.M. Koepf, and C.T. Caskey, 1989 Identification of mutations leading to the Lesch-Nyhan syndrome by automated direct DNA sequencing of *in vitro* amplified cDNA. Proc. Natl. Acad. Sci. USA 86: 1919-1923.

Gibbs, R.A., P.-N. Nguyen, A. Edwards, A.B. Civitello, and C.T. Caskey, 1990 Multiplex DNA deletion and exon sequencing of the hypoxanthine phosphoribosyltranferase gene in Lesch-Nyhan families. Genomics 7: 235-244.

Giroux, C.N., J.R.A. Mis, M.K. Pierce, S.E. Kohalmi, and B.A. Kunz, 1988 DNA sequence analysis of spontaneous mutations in the SUP4-o gene of Saccharomyces cerevisiae. Mol. Cell. Biol. 8: 978-981.

Glickman, B.W., E.A. Drobetsky, and A.J. Grosovsky, 1988 A study of the specificity of spontaneous and UV-induced mutation at the endogenous aprt gene of chinese hamster ovary cells. Banbury Rep. 28: 167-182.

Gordon, A.J.E., P.A. Burns, D.F. Fix, F. Yatagai, F.L. Allen, M.J. Horsfall, J.A. Halliday, J. Grey, C. Bernelot-Moens, and B.W. Glickman, 1988 Missense mutation in the *lacI* gene of *Escherichia coli*. Inferences on the structure of repressor protein. J. Mol. Biol. 200: 239-251.

©1997 Elsevier Science Publishing B.V.

Grist, S.A., M. McCarron, A. Kutlaca, D.R. Turner, and A.A. Morely, 1992 In vivo human somatic mutation frequency and spectrum with age. Mutation Res. 266: 189-196.

Grosovsky, A.J., E.A. Drobetsky, P.J. De Jong, and B.W. Glickman, 1986 Southern analysis of genomic alterations in gamma-ray-induced aprt-hamster cell mutants. Genetics 113: 405-415.

Grosovsky, A.J., J.G. DeBoer, P.J. De Jong, E.A. Drobetsky, and B.W. Glickman, 1988 Base substitutions, frameshifts and small deletions constitute ionizing radiation induced point mutations in mammalian cells. Proc. Natl. Acad. Sci. USA 85: 185-188.

Gupta, R.S., 1984 Genetic markers for quantitative mutagenesis studies in Chinese hamster ovary cell, Applications to mutagen screening studies, in: B.J. Kilbey, M. Legator, W. Nichols and C. Ramel (Eds), Handbook of mutagenicity Test Procedures, 2nd edn., Elsevier, Amsterdam.

Gupta, R.S. and S. Goldstein, 1981 Mutagen testing in the human fibroblast diptheria toxin resistannce (HF Dip^R) system, in: F.J. de Serres and J. Ashby (Eds.), Progress in Mutation Research, Vol.1. Evaluation of Short-Term Tests for Carcinogens, Elsevier, Amsterdam.

Gupta, R.S. and L. Siminovitch, 1976 The isolation and preliminary characterization of somatic cell mutants resistant to the protein synthesis inhibitor – emetine, Cell 9: 213-219.

Gupta, R.S., and L. Siminovitch, 1978a Isolation and characterization of mutants of human fibroblasts to diptheria toxin. Proc. Natl. Acad. Sci. (U.S.A.) 75: 3337-33430.

Gupta, R.S., and L. Siminovitch, 1978b Genetic and biochemical studies with the adenosine analogs toyocamycin and tubercidine. Mutations at the adenosine kinase locus in Chinese hamster cells. Somat. Cell Genet. 4: 715-735.

Gupta, R.S., and L. Siminovitch, 1980 DRB resistance in Chinese hamster and human cells. Genetic and biochemical characteristics of the selection system. Somat. Cell Genet. 6(2): 151-169.

Gupta, R.S., and B. Singh, 1982 Mutagenic response to five independent genetic loci in CHO cells to a variety of mutagens. Mutation Res. 94: 55-71.

Halliday, J.A. and B.W. Glickman, 1991 Mechanisms of spontaneous mutation in DNA repair-proficient *Escherichia coli*. Mutation Res. 250: 351-363.

Holliday, R., 1991 Mutations and epimutations in mammalian cells. Mutation Research 250, 351-363.

Holliday, R. and T. Ho, 1991 Gene silencing in mammalian cells by uptake of 5-methyl deoxycytidine-5'-triphosphate. Somat. Cell Mol. Genet. 17:537-542.

Ikehata, H., T. Akagi, H. Kimura, S. Akasaka, and T. Kato, 1989 Spectrum of spontaneous mutations in a cDNA of the human *hprt* gene integrated in chromosomal DNA. Mol. Gen. Genet. 219: 349-358.

Janatipour, M., K.J. Trainor, R. Kilmura, S. Akasaka and T. Kato, 1989 Spectrum of spontaneous mutations in a cDNA of human *hprt* gene integrated in chromosomal DNA. Mol. Gen. Genet., 219: 349-358.

Jenssen, R.H., W.L. Bigbee and R.G. Langlois, 1987 In vivo somatic mutations in the glycophorin A locus of human erythroid cells, in: Mammalian Cell Mutagenesis, Cold Spring Harbor Press, New York, pp. 149-159.

©1997 Elsevier Science Publishing B.V.

Jones, P.A. and S.M. Taylor 1981 Hemimethylated duplex DNAs prepared from 5-azacytidine treated cells. Nucleic Acid Res. 9: 2933-2947.

Kaden, D.A., L. Bardwell, P. Newmark, A. Anisowicz, T.R. Skopek, and R. Sager. 1989 High frequency of large spontaneous deletions of DNA in tumor derived CHEF cells. Proc. Natl. Acad. Sci. (U.S.A.) 86: 2306-2310.

Konecki, D.S., J. Brennand, J.C. Fuscoe, C.T. Caskey, and A.C. Chinault. 1982 Hypoxanthine-guanine phosphoribosyltransferase genes of mouse and Chinese hamster: construction and sequence analysis of cDNA recombinants. Nucleic Acids Res. 10: 6753-6775.

Langlois, R.G., W.L. Bigbee, R.H. Jessen and J. German. 1989 Evidence for increased *in vivo* mutation and somatic recombination in Bloom's syndrome. Proc. Natl. Acad. Sci. (U.S.A.) 86: 670-674.

Lee, G.S.F., E.A. Savage, R.G. Ritzel, and R.C. von Borstel. 1988 The base-alteration spectrum of spontaneous and ultraviolet radiation-induced forward mutations in the URA3 locus of *Saccharomyces cerevisiae*. Mol. Gen. Genet. 214: 396-404.

Lowy, I. A. Pellicer, J.F. Jackson, G.K. Sim, and S. Silverstein. 1980 Isolation of transforming DNA. Cloning the hamster aprt gene. Cell 22: 817-823.

Luria, S.E. and M. Delbruck, 1943 Mutations of bacteria from virus sensitivity to virus resistance. Genetics 28, 491-511.

McCarron, M.A., A. Kutlaca and A.A. Morely, 1989 The HLA-A mutation assay. Improved technique and normal results. Mutation Res. 240: 117-126.

McGinniss, M.J., M.T. Falta, L.M. Sullivan and R.J. Albertini, 1990 In vivo *hprt* mutant frequencies in T-cells of normal human newborns. Mutation Res. 240: 117-126.

Melton, D.W., 1987 HPRT gene organization and expression, in: N. Maclean (Ed.), Oxford Surveys on Eukaryotic Genes, Vol. 4. Oxford University Press, Oxford, pp. 34-76.

Miller, J.A., 1978 Some current prespectives on chemical carcinogenesis in humans and experimental animals. Cancer Res. 38: 1479-1496.

Mitchell, P.J., G. Urlaub, and L. Chasin. 1986. Spontaneous splicing mutations at the dihydrofolate reductase locus in Chinese hamster ovary cells. Mol. Cell. Biol. 6: 1926-1935.

Morely, A.A., K.J. Trainor, R. Seshadri, and R.G. Ryall, 1983 Measurements of in vivo mutations in human lymphocytes. Nature 302: 155-156.

Murakami, Y., K. Hayashi and T. Sekiya, 1991 Detection of aberrations of the p53 alleles and the gene transcript in human tumor cell lines by single-strand confirmation polymorphism analysis. Cancer Res. 51: 3356-3361.

Nalbantoglu, J., D. Hartley, G. Phear, G. Tear, and M. Meuth, 1986 Spontaneous deletion formation at the *aprt* locus in hamster cells, the presence of short sequence homologies and dyad symmetries at deletion termini. EMBO J. 5:1199-1204.

Nicklas, J.A. T.C. Hunter, J.P. O'Neill, and R.J. Albertini. 1989 Molecular analyses of in vivo *hprt* mutations in human T-lymphocytes: III. Longitudinal study of *hprt* gene structural alterations and T-cell clonal origins. Mutation Res. 215: 147-160.

Orita, M., H. Iwahana, H. Kanazawa, H. Hayashi and T. Sekiya, 1989 Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. Proc. Natl. Acad. Sci. (U.S.A.), 86: 2677-2770.

©1997 Elsevier Science Publishing B.V.

Patel, P.I., P.E. Framson, C.T. Caskey, and A.C. Chinault. 1986 Fine structure of the human hypoxanthine phosphoribosyltransferase gene. Mol. Cell. Biol. 6: 393-403.

Phear, G., J. Nalbantoglu, and M. Meuth. 1987 Next nucleotide effects in mutations driven by DNA precursor pool imbalances at the *aprt* locus of CHO cells. Proc. Natl. Acad. Sci. USA 84: 4450-4454.

Phear, G., W. Armstrong, and M. Meuth. 1989. Molecular basis of spontaneous mutation at the APRT locus of hamster cells. J. Mol. Biol. 209: 577-582.

Ramel, C., 1989 The nature of spontaneous mutations. Mutation Research 212: 33-42.

Recio, L., J. Cochrane, D. Simpson, T.R. Skopek, J.P. O'Neil, J.A. Nicklas, and R.J. Albertini. 1990a DNA sequence analysis of in vivo *hprt* mutation in human T-lymphocytes. Mutagenesis 5: 505-510.

Recio, L., D. Simpson, J. Cochrane, H. Liber, and T.R. Skopek. 1990a Molecular analysis of *hprt* mutants induced by 2-cyanoethylene oxide in human lympoblastoid cells. Mutation Res. 242: 195-208.

Recio, L. D. Simpson, J. Cochrane, H. Liber and T.R. Skopek. 1990b Molecular analysis of *hprt* mutants induced by 2-cyanoethylene oxide in human lymphoblastoid cells. Mutation Res. 242: 195-208.

Rossi, A.M., J.C.P. Thijssen, A.D. Tates, H. Vrieling, A.T. Natarajan, P.H.M. Lohman, and A.A. van Zeeland. 1990 Mutations affecting RNA splicing in man are detected more frequently in somatic than in germ cells. Mutation Res. 244: 353-357.

Rossi, A.M., A.D. Tates, A.A. van Zeeland, and H. Vrieling. 1992 Molecular analysis of mutations affecting HPRT mRNA splicing in human T-lymphocytes *in vivo*. Environ. Mol. Mutagen. 19: 7-13.

Schaaper, R.M. and R.L. Dunn. 1991 Spontaneous mutation in the *Escherichia coli lac*I gene. Genetics 129: 317-326.

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 lacfl gene. J. Mol. Biol. 189: 273-284.

Seshadri, R.R., J. Kutlaca, K. Trainor, C. Matthews, and A.A. Morely, 1987 Mutation rate of normal and malignant human lymphocytes. Cancer Res. 47: 407-409.

Skandalis, A., A.J. Grosovsky, E.A. Drobetsky, and B.W. Glickman, 1992 Investigation of the mutagenic specificity of X-rays using a retroviral shuttle vector in CHO cells. Environ. Mol. Mutagen. 20: 271-276.

Smith, C.A., and I. Mellon. 1990 Clues to the organization of DNA repair systems gained from studies of intragenomic repair heterogeneity. In: (Obe, E., ed) Advances in mutagenesis, Vol. 1, Springer Verlag, Berlin, pp. 153-194.

Stanley, P., and L. Siminovitch, 1976 Clues to the organization of DNA repair systems gained from studies of intragenomic repair heterogeneity. in: G. Obe (Ed.), Advances in Mutagenesis, Vol. 1, Springer, Berlin, pp. 153-194.

Stout, J.T., and C.T. Caskey, 1985 HPRT: gene structure, expression, and mutation, Ann. Rev. Genet. 19: 127-148.

Suzuki, Y., T. Sekiya and K. Hayashi, 1991 Allele-specific polymerase chain reaction, a method for amplification and sequence determination of a single component among a mixture of sequence variants. Anal. Biochem., 192: 82-84.

©1997 Elsevier Science Publishing B.V.

Tasseron-de Jong, J.G., H. der Dulk, P.van de Putte and M. Giphart Gassler, 1989 *De novo* methylation as a major event in the inactivation of transfected Herpes virus thymidine kinase genes in human cells. Biochim. Biophys. Acta 1007: 215-223.

Tates, A.D., L.F. Bernini, A.T. Natarajan, J.S. Ploem, N.P. Verwoerd, J. Cole, M.H.L. Green, C.F. Arlett, and P.N. Norris, 1989 Detection of somatic mutations in man, HPRT mutations in lymphocytes and hemoglobin mutations in erythrocytes. Mutation Res., 213: 73-82.

Tates, A.D., F.J. van Dam, H. van Mossel, H. Schoemaker, J.C.P. Thijssen, V.M. Woldring, A.H. Zwinderman and A.T. Natarajan, 1991 Use of the clonal assay for the measurement of frequencies of HPRT mutants in the T-lymphocytes from five control populations. Mutation Res. 253: 199-213.

Thacker, J., M.A. Stephens, and A. Stretch, 1978 Mutation to ouabain resistance in Chinese hamster cells. Induction by ethyl methanesulphonate and lack of induction by ionizing radiation. Mutation Res. 51: 255-270.

Tiah, M., and A. Ronen, 1991 Dominant lethal cell mutants detected by the autoradiographic assay for exotoxin A resistence. Mutation Res. 213: 205-215.

Tiah, M., and A. Ronen, 1989 Autoradiographic detection of mutation to exotoxin-A resistance in mouse fibroblasts treated with ethyl methanesulfonate, X-rays and ultraviolet light. Mutation Res. 213: 205-215.

Tindall, K.R. and L.F. Stankowski. 1989. Molecular analysis of spontaneous mutations at the gpt locus in Chinese hamster ovary (AS52) cells. Mutation Res. 220: 241-253.

Turner, D.R., A.A. Morley, M. Haliandros, R. Kutlaca, and B.J. Sanderson, 1985 In vivo somatic mutations in human lymphocytes frequently result from major gene alterations. Nature 315: 343-345.

Turner, D.R., A.A. Morely, 1990 Human somatic mutation at the autosomal HLA-A locus, in: M.L. Mendelsohn and R.J. Albertini (Eds.), Mutation and the Environment, Part C, Wiley-Liss, New York, pp. 337-345.

Vijayalaxmi, and H.J. Evans, 1984 Measurement of spontaneous and X-irradiation-induced 6-thioguanine-resistant human blood lymphocytes using a T-cell cloning technique. Mutation Res. 125: 87-94.

Vrieling, H., J.W.I.M. Simons, F. Arwert, A.T. Matarajan, and A.A. Van Zeeland, 1985 Mutation induced by X-rays at the HPRT locus in cultured Chinese hamster cells are mostly large deletions. Mutation Res. 144: 281-286.

Wilson, J.M., and W.N. Kelley, 1983 Molecular basis of hypoxanthine-guanine phosphoribosyltransferase deficiency in a patient with the Lesch-Nyhan syndrome. J. Clin. Invest. 71: 1331-1335.

Wilson, J.M., G.E. Tarr, and W.N. Kelley, 1983a Human hypoxanthine (guanine) phosphoribosyltransferase: An amino acid substitution in a mutant form of the enzyme isolated from a patient with gout. Proc. Natl. Acad. Sci. USA 80: 870-873.

Wilson, J.M., R. Kobayashi, I.H. Fox, and W.N. Kelley, 1983b Human hypoxanthine-guanine phosphoribosyltransferase: Molecular abnormality in a mutant form of the enzyme ($HPRT_{Toronto}$). J. Biol. Chem. 258: 6458-6460.

Yandell, D.W., T.P. Dryja, and J.B. Little, 1990 Molecular genetic analysis of recessive mutations at a heterozygous autosomal locus in human cells. Mutation Res. 229:89-102.

Yang, J.L., M.C. Hu, and C.W. Wu, 1991 Novel mutational spectrum induced by *N*-methyl-*N*-nitro-*N*-nitrosoquanidine in the coding region of the hypoxanthine (guanine) phosphoribosyl transferase gene in diploid human fibroblast. J. Mol. Biol. 221: 421-430.

©1997 Elsevier Science Publishing B.V.

Zhang, L.-H., and D. Jenssen, 1991a. Characterization of HAT- and HasT- resistant HPRT mutant clones of V79 Chinese hamster cells. Mutation Res. 263:151-158.

Zhang, L.-H., and D. Jenssen. 1991b. Site specificity of N-methyl-N-nitrosourea-induced transition mutations in the *hprt* gene. Carcinogenesis 12: 1903-1909.

Zhang L.H., H. Vrieling, A.A. van Zeeland, D. Jenssen, 1992 Spectrum of spontaneously occurring mutations in *the hprt* gene of V79 Chinese hamster cells. J. Mol. Biol. 223: 627-35.

©1997 Elsevier Science Publishing B.V.

Variable aberrant cDNAs in single diphtheria toxin-resistant human fibroblasts.

Amiram Ronen[¶], Miriam Broit[¶], Axel Nohturfft[†‡], John Curry[†] and Barry W Glickman[†]

[¶] Department of Genetics, The Hebrew University, Jerusalem 91904, Israel
[†] Centre for Environmental Health and the Department of Biology, University of Victoria, Victoria, British Columbia, Canada
[‡] Present address: Department of Molecular Genetics, UT Southwestern, Dallas, TX 75235-9046

Abstract

We treated transformed human fibroblasts with diphtheria toxin (DT) and isolated 40 single cells that were toxin resistant but unable to propagate. In 13 of them toxin resistance was associated with the presence of one or more aberrant transcripts of the structural gene for elongation factor 2 (EF-2). cDNA obtained from these transcripts had 164-447 bp-long deletions. Each of these deletions was associated with 2-8 base pairs-long repeats at its breakpoints. Only 10 out of 16 cDNA deletions were associated with presumed exon junctions. A role is suggested for errors in transcription in producing the aberrant transcripts, which gave rise to the deletion-bearing cDNA species.

Introduction

Diphtheria toxin (DT) and *Pseudomonas aeruginosa* exotoxin A (PEA) both kill mammalian cells by ADP ribosylation of elongation factor 2 (EF-2), thereby shutting off protein synthesis (Collier 1967, Pappenheimer *et al.* 1973, Collier 1975, Iglewski *et al.* 1977). A pre-requisite for this process is the post-transcriptional modification of a histidine residue at position 715 of the EF-2 molecule to diphthamide (Van-Ness *et al.* 1980a, 1980b). Diphthamide 715 in the mammalian EF-2 protein and its counterparts in EF-2 from other organisms are the only known targets for ADP-ribosylation by either DT or PEA (Pappenheimer and Gill 1972, Iglewski *et al.* 1977, Dunlop and Bodley 1983). Cross-resistance to DT and PEA may result from changes in the primary structure of the *ef2* gene, or from defective post-translational modification of the EF-2 protein. Toxin resistance apparently due to mutant EF-2 has

©1997 with permission from Elsevier Science Publishing B.V.

been described in rodent cell lines (Moehring and Moehring 1977, Ronen *et al.* 1984), established as well as primary human cell lines (Moehring and Moehring 1977, Gupta and Siminovitch 1978, Drinkwater 1982, Ronen *et al.* 1984, Tiah and Ronen 1989, Tiah and Ronen 1991), and yeast (Chen *et al.* 1985). Mutations in the mammalian *ef2* gene conferring toxin resistance have been described, at the DNA-sequence level, in Chinese-hamster ovary (CHO) and transformed human fibroblasts (Kohno and Uchida 1987, Kohno *et al.* 1986, Foley *et al.* 1995, Nakanishi *et al.* 1988, Curry *et al.* 1995). They often are base-pair substitutions, notably G → A transitions, in codon 717 (Kohno and Uchida 1987, Foley *et al.* 1995, Curry *et al.* 1995). It has been suggested that the ensuing amino-acid replacements protect histidine 715 from undergoing post-translational modification to diphthamide (Foley *et al.* 1995).

Petri dish-grown cultures of DT-sensitive fibroblasts cease to synthesize proteins shortly after exposure to the toxin. Pre-existing, DT-resistant cells can be detected in such cultures by virtue of their ability, in presence of the toxin, to maintain protein synthesis and incorporate radio-

©1997 Elsevier Science Publishing B.V.

labeled amino acids into cellular proteins. Ronen *et al.* (1984) and others (Gupta and Singh *et al.* 1985, Sakamoto *et al.* 1986) used autoradiography in order to detect DT resistance in cultured fibroblasts of Chinese-hamster or human origin, and Tiah and Ronen (Tiah and Ronen 1989, Tiah and Ronen 1991) used similar methods to detect PEA resistance in murine cells, as early as 18 hours after exposure.

Toxin-resistant cells can also be detected by standard microscopy if, following exposure, the sensitive cells are allowed several days to disintegrate and detach from the culture-dish surface. In both techniques, the frequency of the resistant <u>cells</u> is 1-2 orders of magnitude higher than that of the resistant *colonies* found in the conventional colony assay (Tiah and Ronen 1991). With few exceptions, these resistant cells are incapable of proliferation (we will refer to such cells as "non propagating"). Nevertheless, they may remain attached to the petri dish for up to 2 weeks following exposure. During much of that time their appearance is similar to that of DT-resistant cells, or of sensitive cells growing in toxin-free medium. It has been suggested (assay (Tiah and Ronen 1991) that these cells harbor

lethal DT-resistant mutations which, by having been selected for continued protein synthesis rather than for normal proliferative ability, represent part of the mutational spectrum in the *ef2* gene that is missed by the conventional colony assay.

In an unpublished work we (Ronen, personal communication) analyzed the products of RT PCR from a single CHO cell that had survived treatment with DT. The only mutation found in the cDNA from that cell was a 0.4 kb-long deletion. That finding prompted us to look for a similar phenomenon in human fibroblasts. Here we examine the region between codons 415 and 710 in *ef2* cDNA from non-propagating, DT-resistant human fibroblasts. We show that DT resistance in these cells is associated with the presence of an *ef2* message which is different from that found in DT-sensitive or in propagating DT-resistant cells.

Materials and Methods

Cell lines: GM637 and GM847 are SV40-transformed, DT-sensitive fibroblast lines obtained from a normal female and an HPRT-deficient male,

©1997 Elsevier Science Publishing B.V.

respectively; DTR1 is a spontaneous DT-resistant mutant, derived from GM637 by selection with 0.1 lethal factors (LF) DT/ml; DTRE4 is an ethylmethane sulfonate-induced, DT-resistant mutant similarly derived from strain GM847 Both DTR1 and DTRE4 owe their resistance to a G→A substitution in codon 717.

Cells were grown in Dulbecco's Modified medium (DMEM) containing 5% calf serum and 5% newborn-calf serum, in a humidified 37°C incubator, in 5% CO_2.

Selection for DT-resistant cells: GM637 cells growing in Dulbecco's modified Eagle medium in 50-mm petri plates were treated with 1 LF DT/ml for 3 days. After 4-5 more days they were examined with an inverted phase-contrast microscope, and a circle was drawn on the bottom surface of the dish around each live cell. Cells were judged to be alive by the appearance of the cytoplasm and the nucleus, as well as their general shape (Collier 1975, Tiah and

Ronen 1991). Exclusion of Trypan Blue or erythrosin B was not found to change judgments made without the dyes, and the use of fluorescent dyes made the examination too cumbersome. After 24 more hours the marked cells were re-examined, and those still deemed alive were marked again. The plates were then washed with phosphate-buffered saline (PBS) and air-dried. The marked cells were isolated by cutting the bottom of the petri dish around them with a hot scalpel.

RT-PCR and sequencing of DNA from resistant cells. Lysis and reverse transcription: Plates with toxin-resistant cells were washed with phosphate-buffered saline and air-dried. Small pieces of the bottom of the petri dish, carrying single DT-resistant cells, were excised with a hot scalpel and transferred to 500-µl PCR tubes with either 5 or 10 µl of lysis-RT mix. The mix contained 10 mM Tris pH 8.3, 50 mM KCl, 5 mM $MgCl_2$, 500 µM dNTPs, 0.25% NP40, 0.5 µM primer 2665R (5'GCGTCGCTGT-GTCGGGACA 3') and, per 1 µl, 2.5 units of M-MLV reverse transcriptase (BRL) and 2 U of RNAsin (Promega). The tubes were incubated 40 min at 41°C and 5 min at 99°C, and transferred to ice.

PCR1: To each tube were added 3 volumes (15-30 µl) of first PCR mix, such that the final composition of the reaction mixture was: 10 mM Tris pH9.2, 50 mM KCl, 1.5 mM $MgCl_2$, 0.5 µM of each primer, and 0.05 U *Taq* polymerase µl. The primers used were 996L (5' CAAACCCCTGCTGAAGGCTG 3') and 2665R. The thermo-cycling profile (using 500-µl, thick-walled tubes in a Perkin-Elmer Cetus Thermocycler) was: 6 min at 94°C; 30 cycles of 1 min at 94°C, 1 min at 60°C and 1.5 min at 72°C; 6 min at 72°C.

PCR2: The product of the first PCR was diluted at least 1:50 into a PCR2 mix identical with that used in PCR1, except for the primers: 1245L (5' AGTCTTCTCGGGGCTGGT 3') and 2129R (5' GGTGGATGGCGTCGGCGT). The protocol used in PCR2 was the same as in PCR1. The final PCR2 product spanned the interval between codons 415 and 709, and was expected

©1997 Elsevier Science Publishing B.V.

to be 902 bp long. _Primers used for amplifying smaller fragments and for sequencing were: 1496L (5' TCAGCGTCAGCCCTGTTG 3'), 1535R (5' GCAGGTCAGCCGGGTTCT 3'), 1623L (5' GGGAGAGCATATCATCGC 3'), 1743R (5' GACTCTTCACTGACCGTC 3'), 1851L (5' CGATAAAGGCGAGGTGTCCG 3') and 1917R (5' CCTCAGCCACGTCCCACTCG 3'). The reaction was done as described in Figure 1, except that the pH was 8.5 and KCl was at 60 mM, and thin-walled PCR tubes were used in a Perkin-Elmer Cetus 9600 Thermocycler. The thermocycling profile was: 1 min at 95°C; 30 cycles of 5 sec at 94°C, 5 sec at 60°C and 40 sec at 72°C; and 7 min at 72°C.

Separation and Sequencing of short cDNA species. To separate and purify the short DNA species, 15 μl of the PCR product were subjected to electrophoresis on a 3% low melting point agarose gel (NuSeive), and the bands were excised, melted in 250 μl Tris-EDTA, and used as template for further amplification or for sequencing. The DNA product was precipitated with iso-propanol. Cycle-sequencing was done using (γ32P) ATP end-labeled primers. The primers were those described above.

Results

40 non-propagating, DT-resistant cells were isolated and subjected to RT-PCR. 13 cells produced, in addition to the expected, 885 bp-long cDNA species, one or more aberrant cDNAs that were either shorter or longer than expected. While some of them could not be amplified for further examination, we were able to analyze 16 of the short cDNAs (Figure 1). Each of them had a deletion of 164 - 447 bp (Figure 2). In control experiments, RT-PCR of the *ef2* mRNA, performed with either single or multiple (100-300) GM637 cells which had not been treated with DT, produced only DNA of the expected size. Similarly, no aberrant cDNAs were observed among the RT-PCR products from single DTR1 or DTRE4 cells grown either with or without DT.

No point mutations were found either in the deletion-carrying or the corresponding full-sized cDNAs. The sequence found was that published by Rapp *et al.* (1989, GeneBank accession

©1997 Elsevier Science Publishing B.V.

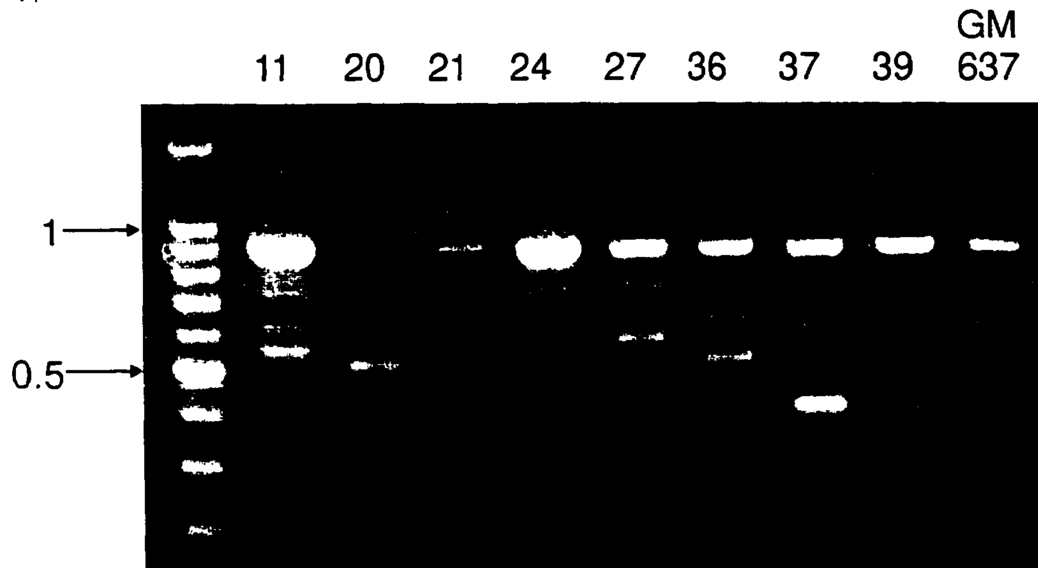


Figure 1. Aberrant cDNA species in single non-propagating DT-resistant cells. Agarose-gel electrophoresis of PCR2 products made with primers 1245L and 1917R from non-propagating DT-resistant mutants. Mutants' numbers are shown. The expected size for these products is 691bp.

number Z11692), with one exception: the nucleotide at position 2190 (3rd position in codon 730) is C rather than T. Our data also confirm C rather than T at position 1632. In a separate group of 14 non-propagating cells that had been selected for DT resistance we amplified the *hprt* as well as the *ef2* transcript, using both *hprt*-specific primers and the *ef2*-specific primers. 4 cells produced multiple aberrant *ef2* cDNAs, and they all had only the (single) expected species of *hprt* cDNA.

In 9 of the aberrant cDNAs, the 5' end of the deletions is at, or within 12 bp from a site which, in the hamster cDNA, is the junction of exons 7 and 8. In 6 of the 9, the 3' end of the deletion is at, or within 5 bp from the exon 9 exon 10 junction. In one more aberrant cDNA, the 5' end of the deletion is at the exon 9 exon 10 junction. The other 16 ends are not associated with exon junctions.

©1997 Elsevier Science Publishing B.V.

A

```
                 7/8                                   9/10
20.1  TGAAGCCAATCCAGAGAACAATCTT  1342 (376)  CCCATCAAGAAATCTGAC
                 7/8                                   9/10
40.1  TGAAGCCAATCCAGAGAACAATCTT  1344 (367)  CCCATCAAGAAATCTGAC
                 7/8                                   9/10
36.2  TGAAGCCAATCCAGAGAACAATCTT  1346 (368)  CCCATCAAGAAATCTGAC
                 7/8                                   9/10
39.1  TGAAGCCAATCCAGAGAACAATCTT  1346 (368)  CCCATCAAGAAATCTGAC
                 7/8                                   9/10
21.1  TGAAGCCAATCCAGAGAACAATCTT  1349 (366)  CCCATCAAGAAATCTGAC
                 78                                    9/10
27.1  TGAAGCCAATCCAGAGAACAATCTT  1349 (366)  CCCATCAAGAAATCTGAC
```

B

```
                 7/8
37.1  TGAAGCCAATCCAGAGAACAATCTT  1340 (246)  AATGGCCAAGTCC

                 7/8
38.1  TGAAGCCAATCCAGAGAACAATCTT  1352 (441)  CCCCAACAAGCAC
                 7/8
18.1  TGAAGCCAATCCAGAGAACAATCTT  1352 (447)  CCCCAACAAGCACAACCGG
```

C

```
22.1        TTGTGGGCCT            1412 (278)     AGGAGGACCA

11.1        TGAAGACGGGCAC         1448 (387)     CCCCGACGGCCTG

32.1        TTCAGCGTCAG           1502 (394)     GCGGGCGCGCT

36.1        GAGGCCAAGAAC          1536 (254)     TCCCCCAACAAG

11.2        CTGACCTGCCC           1551 (274)     CGGCCCCTTCCC

39.2        CTGGCCAAGTCCGACC      1591 (194)     CTCTCCAAGTCCCCCA
               9/10
24.1        ATCAAGAAAT            1714 (165)     CGTCAGGAGC
```

Figure 2. cDNA deletions in single non-propogating DT-resistant cells. Numbers on left are cell numbers followed by agarose gel-band numbers. Human ef2 cDNA sequences are from Rapp *et al.* (1989). Exon junctions (denoted by the appropriate numbers above junction sites) are from the hamster genomic ef2 map (Nakanishi *et al.* 1988). Deletions are underlined. Repeats are in bold type. The actual boundaries of each deletion may be at any homologous positions within the repeat, but are arbitrarily shown as if they are at least 20 bp outside the given sequence. The numbers inside each deletion are the map position of the 5' repeat and (in brackets) the number of missing base pairs. Deletions in groups A and B all have their 5' ends near, or at the junctions of exons 7 and 8. Those in group A also have their 3' ends near, or at the junction of exons 9 and 10.

©1997 Elsevier Science Publishing B.V.

40.1
5'-TACCTGAAGCCAATCCAGAGAACAATCTTGAT-3'
5'-CGCCTGCATCCCCATCAAGAAATCTGACCCGG-3'

39.1, 36.2
5'-ACCTGAAGCCAATCCAGAGAACAATCTTGATGA-3'
5'-CCTGCATCCCCATCAAGAAATCTGACCCGGTCG-3'

20.1
5'-CCTCTACCTGAAGCCAATCCAGAGAACAATCTTG-3'
5'-GCATCCCCATCAAGAAATCTGACCCGGTCGTCTC-3'

27.1, 21.1
5'-CTGAAGCCAATCCAGAGAACAATCTTGATGATGG-3'
5'-CCTGCATCCCCATCAAGAAATCTGACCCGGTCGT-3'

18.1
5'-AAGCCAATCCAGAGAACAATCTTGATGATGGGCC-3'
5'-AGTCCCCCAACAAGCACAACCGGCTGTACATGAA-3'

37.1
5'-GGACCTCTACCTGAAGCCAATCCAGAGAACAATCT-3'
5'-GGGCTGAAGCGGCTGGCCAAGTCCGACCCCATGGT-3'

38.1
5'-GAAGCCAATCCAGAGAACAATCTTGATGATGGGCC-3'
5'-CTCTCCAAGTCCCCCAACAAGCACAACCGGCTGTA-3'

22.1
5'-TTGTGGGAACATTGTGGGCCTCGTGGGCGTGG-3'
5'-GAAGGACCTGGAGGAGGACCACGCCTGCATCC-3'

24.1
5'-CCTGCATCCCCATCAAGAAATCTGACCCGGTC-3'
5'-AGGTGTCCGCCCGTCAGGAGCTCAAGCAGCGG-3'

11.2
5'-CAAGAACCCGGCTGACCTGCCCAAGCTGGTGGA-3'
5'-ACATGAAGGCGCGGCCCTTCCCCGACGGCCTGG-3'

32.1
5'-GGGTGATGAAGTTCAGCGTCAGCCCTGTTGTCA-3'
5'-GCTCAAGCAGCGGGCGCGCTACCTGGCCGAGAA-3'

36.1
5'-GAGTGGCCGTGGAGGCCAAGAACCCGGCTGACCT-3'
5'-GCCTCTCCAAGTCCCCCAACAAGCACAACCGGCT-3'

11.1
5'-CCAGTTCCTGGTGAAGACGGGCACCATCACCACCTTC-3'
5'-GCGCGGCCCTTCCCCGACGGCCTGGCCGAGGACATCG-3'

39.2
5'-GGCTGAAGCGGCTGGCCAAGTCCGACCCCATGGTGCAG-3'
5'-ACGTGCTCTGCCTCTCCAAGTCCCCCCAACAAGCACAAC-3'

Figure 3. Homology around deletion junctions. For each of the cDNA deletions, the sequence around the 5' and 3' ends are aligned to show possible homology. Only the sequence in the non-transcribed strand is shown. Direct repeats are underlined. Deletions 39.1 and 36.2 are identical, and so are deletions 21.1 and 27.1.

©1997 Elsevier Science Publishing B.V.

Two features shared by all 16 cDNA deletions are striking: (a) 2-8 base pairs at the 3' ends are directly repeated at the 5' end; these repeats do not allow explicit location of the events which have created the deletions. (b) Each is unique (with the possible exception of the deletions in cDNAs nos. 21.1 and 36.2, which may be identical to those in cDNAs nos. 27.1 and 39.1, respectively). It is also noteworthy that the aberrant cDNAs do not contain any sequences that are not part of the *ef2* coding sequence.

Discussion

The cDNA deletions we observe in non-propagating, DT-resistant cells are not found in sensitive cells that had not been subjected to the toxin, or in cells from DT-resistant lines grown either in the presence or absence of toxin. It therefore seems unlikely that they are artifacts resulting from e.g., template switching by the polymerases used in the RT-PCR (Paabo *et al.* 1990), or amplification of DNA from processed *ef2*-like pseudogenes (Koide *et al.* 1990). The possibility that they result from perturbation of the splicing machinery by exposure to DT is ruled out by the absence of corresponding deletions in *hprt* cDNA from the same cells where aberrant *ef2* cDNAs were found. It therefore seems likely that the aberrant cDNAs observed in the non-propagating DT-resistant cells represent aberrant *ef2* transcripts. Conceivably, such transcripts could have resulted from (a) aberrant splicing, (b) transcription of chromosomal deletions, or (c) errors made during transcription.

A prominent feature of the cDNA deletions is their association with sites which, in the hamster cDNA, are known to be exon junctions (Nakanishi *et al.* 1988).

The genomic sequence of human *ef2* has not been published. However, there is 87% homology between the human and hamster *ef2* cDNAs, and better than 99% homology between the corresponding proteins (Nakanishi *et al.* 1988, Rapp *et al.* 1989). The *locations* of introns are therefore likely to be identical in the two species. This is supported by data (J. Curry, personal communication) showing the existence of introns in the human *ef2* gene precisely where

©1997 Elsevier Science Publishing B.V.

introns 8, 9 and 10 of the hamster gene are located.

In 6/16 of the aberrant cDNAs, of both ends of the deletion are near, or at an exon junction. In another 4 aberrant cDNAs, one end of the deletion is similarly associated with an exon junction. This would suggest that the deletions resulted from exon skipping, possibly caused by chromosomal mutations (variable exon skipping caused by a single mutation could account for the presence in a single cell of more than one aberrant cDNA). Such explanation is, however, not applicable to those cases where neither end of the deletion is associated with an exon junction. Nor is it compatible with the absence, in all 16 cases, of (*1*) the GT...AG or AT...AC dinucleotide-pairs, which almost invariably constitute the ends of spliced introns (see Mount 1996); (*2*) point mutations in either the aberrant or normal cDNAs that might be responsible for exon skipping, and (*3*) intron sequences.

The remarkable association between the cDNA deletions and direct repeats suggests that a mechanism other than exon skipping may be involved in their formation. Repeats are not known to play a role in the removal of internal segments from either processed or un-processed mRNA. On the other hand, they are known to be involved in various recombinational events occurring at the DNA level, such as inter- or intra-molecular recombination which may, in turn, result in the formation of deletions (or duplications). it is, however, extremely unlikely that the cDNA deletions we observe here represent genomic-DNA deletions. Such an origin would require the presence in each cell (notably, nos. 11, 36 and 39) of more than one deletion-bearing chromosome (see Figure 1).

The conflict could be resolved if the EF-2 mRNA mutations were formed *during transcription*. Conceivably, erroneous transcription could be an indirect result of an increased need for EF-2 mRNA synthesis following the inactivation of EF-2 by DT. However, in unpublished experiments we found the cellular level of EF-2 mRNA in non-propagating cells to be no higher than that in normal, DT-sensitive cells. On the other hand, the EF-2 mRNA mutations could have resulted from interactions between *normal* RNA polymerase and direct repeats. In this scenario, the presence

©1997 Elsevier Science Publishing B.V.

of direct repeats in the template DNA may promote interactions (e.g., loop formation) that will create opportunities for RNA polymerase to hop from one position to another and, by doing so, skip a segment of the template. RNA polymerase has recently been reported capable of switching DNA templates in vitro, without losing its transcript in the process (Foley *et al.* 1995). In this case, the proximity to the site of event of exon junctions may be coincidental, merely reflecting their being accidentally located in regions where the putative template switching is likely to occur.

It is tempting to speculate that the partial homology between the corresponding regions surrounding the direct repeats (Figure 3) might facilitate template switching during transcription. In some cases, however, this homology is not greater than that found between DNA sequences picked at random from regions of the *ef2* cDNA map where deletions have not been found (data not shown). Additional factors which influence elongation, such as mRNA secondary structure (Sastry and Hoffman 1995), endogenous signals in the template DNA (Aso *et al.* 1995), and mutagen-induced DNA adducts (Sauerbier and Hercules 1978) may affect the putative template switching. This is in line with the reported increase in the frequency of non-propagating DT-resistant cells after mutagenic treatment (Ronen *et al.* 1984, Tiah and Ronen 1991, Sakamoto *et al.* 1986).

Finally, all the cDNA deletions in Figure 2 were not necessarily generated by the same mechanism. It is possible that, in some cells, the aberrant *ef2* transcripts resulted from chromosomal mutations (via exon skipping), while in others they were caused by template switching during transcription. The presence in a single cell of different cDNA deletions indicates, however, that regardless of their origin, they may represent a larger spectrum of similar events. Some aberrant transcripts may have fortuitously been missed by the RT-PCR process, or given rise to cDNA molecules that we were unable to amplify for sequencing.

The relationship between the cDNA deletions and the two phenotypes, DT resistance and the loss of ability to propagate, is enigmatic.

The part of the ef2 gene we analyzed (codons 415-710) is in the region coding for the

©1997 Elsevier Science Publishing B.V.

ribosome-binding domain of the protein (Kohno *et al.* 1986, Rapp *et al.* 1989). In CHO cells, an amino-acid replacement at position 584 confers DT resistance by protecting histidine 715 from being modified to diphthamide. This presumably is made possible by the proximity of the two amino acids on the surface of the folded EF-2 protein (Foley *et al.* 1995). A similar protective effect is unlikely to be the explanation in the case of the cDNA deletions observed here, because of their severity. For example, 10 of the 16 deletions result in reading-frame shifts with new termination codons located as close as 1 codon downstream. Conceivably, the abnormal proteins could still protect the EF-2 molecules made by the normal allele, should EF-2 act as a dimer or a multimer. However, rabbit EF-2 is reported to be active as a single polypeptide (Merrick *et al.* 1975), and we are not aware of evidence for multimer formation. Alternatively, in spite of lacking the domain containing the target for DT, the mutant EF-2 could still be able interact with the toxin and thus deplete its intracellular pool. In any case, the presence of the mutant EF-2 in the cell is not compatible with normal growth (Tiah and Ronen 1989, Foley *et al.* 1995).

Selection for DT-resistant colonies reveals a rather limited spectrum of *ef2* mutations to DT resistance, presumably because most of the mutations in that gene are lethal (Tiah and Ronen 1991, Kohno *et al.* 1986, Aust *et al.* 1984). This is in line with the highly conserved nature of the gene. Thus, in 2 separate studies of DT-resistant CHO cell lines, a total of 11 out of 14 ef2 mutations were G→A transitions in codon 717, and only 3 were base-pair substitutions elsewhere (Kohno and Uchida, Foley *et al.* 1995). It has been suggested that other amino-acid replacements result in protein synthesis whose level is insufficient for normal cell growth (Foley *et al.* 1995). The single cell assay, being free of the constraint of growth, can detects DT resistance even when it is incompatible with cell proliferation and cannot be transmitted to progeny cells. In a broader sense, it can reveal events that result in a *complex phenotype*, and are therefore masked at the multi-cell level.

©1997 Elsevier Science Publishing B.V.

Acknowledgments

Work in Israel was supported by a grant to A. Ronen from The Israel Science Foundation, administered by The Israel Academy For Sciences And Humanities. Work in Canada was supported by a grant from the National Cancer Institute of Canada and a National Science And Engineering Research Council (NSERC) University-Industrial Chair Award to BW Glickman.

References

Aso T., J.W. Conaway and R.C. Conaway (1995) The RNA polymerase II elongation complex. The FASEB Journal. 9. 1419-1428

Aust A.E., N.R. Drinkwater, K. Debien, V.M. Maher and J.J. McCormick (1984) Comparison of the frequency of diphtheria toxin and thioguanine resistance induced by a series of carcinogens to analyze their mutational specificities in diploid human fibroblasts. Mutation Res., 125, 95-104

Chen J.Y., J.W. Bodley and D.M. Livingston (1985) Diphtheria toxin-resistant mutants of Saccharomyces cerevisiae. Mol. Cell. Biol., 5, 3357-3360

Collier R.J. (1975) Diphtheria toxin: Mode of action and structure. Bacteriological Reviews, 39, 54-85

Collier R.J., (1967) Effect of diphtheria toxin on protein synthesis: inactivation of one of the transfer factors. J. Mol. Biol., 25, 83-98

Curry J., M. Broit, A. Nohturfft, A. Ronen and B.W. Glickman (1995) Sequence analysis of elongation factor-2 cDNA from single diphtheria toxin-resistant human fibroblasts. Environ. Mol. Mutagen., 25 (S 25), 10

Drinkwater N.R., R. Corner, J.J. McCormick and V.M. Maher (1982) An in situ assay for induced diphtheria-toxin-resistant mutants of diploid human fibroblasts. Mutation Res., 106, 277-289

©1997 Elsevier Science Publishing B.V.

Dunlop P.C. and J.W. Bodley (1983) Biosynthetic labeling of Diphthamide in Saccharomyces cerevisiae, J. Biol. Chem., 258, 4754-4758

Foley B.T., J.M. Moehring and T.J. Moehring (1995) Mutations in the elongation factor 2 gene which confer resistance to diphtheria toxin and Pseudomonas exotoxin A - Genetic and biochemical analyses. J. Biol. Chem., 270, 23218-23225

Gupta R.S. and B. Singh (1985) Autoradiographic Detection of Diphtheria Toxin Resistant Mutants in Human Diploid Fibroblasts. Environ. Mutagen., 7, 611-624

Gupta R.S. and L. Siminovitch (1978) Isolation and characterization of mutants of human diploid fibroblasts resistant to diphtheria toxin. Proc. Natl. Acad. Sci. USA, 75, 3337-3340

Iglewski B.H., P.V. Liu and D. Kabat (1977) Mechanism of Action of Pseudomonas aeruginosa Exotoxin A: Adenosine Diphosphate-Ribosylation of Mammalian Elongation Factor 2 In Vitro and In Vivo. Infection and Immunity, 15, 138-144

Kohno K. and T. Uchida (1987) Highly frequent single amino acid substitution in mammalian elongation factor 2 (EF-2) results in expression of resistance to EF-2-ADP-ribosylating Toxins. J. Biol. Chem., 262, 12298-12305

Kohno K., T. Uchida, H. Ohkubo, S. Nakanishi, T. Nakanishi, T. Fukui, E. Ohtsuka and M. Ikehara (1986) Amino acid sequence of mammalian elongation factor 2 deduced from the cDNA sequence: Homology with GTP-binding proteins. Proc. Natl. Acad. Sci. USA, 83, 4978-4982

Koide T., M. Ishiura, N. Hazumi, T. Shroishi, Y. Okada and T. Uchida (1990) Amplification of a long sequence that includes a processed pseudogene for elongation factor 2 in the mouse. Genomics 6, 80-88

Merrick W.C., W.M. Kemper, J.A. Kantor and W.F. Anderson (1975) Purification and properties of rabbit reticulocyte protein synthesis elongation factor 2. J. Biol., Chem. 250, 2620-2625

Moehring T.J. and J.M. Moehring (1977) Selection and Characterization of Cells Resistant to Diphtheria Toxin and Pseudomonas Exotoxin A: Presumptive Translational Mutants. Cell, 11, 447-454

Mount S.M. (1996) AT-AC introns: an ATtACk on Dogma. Science, 271, 16901692

Nakanishi T., K. Kohno, M. Ishiura, H. Ohashi and T. Uchida (1988) Complete Nucleotide Sequence and Characterization of the 5'-Flanking Region of Mammalian Elongation Factor 2 Gene. J. Biol. Chem., 263, 6384-6391

©1997 Elsevier Science Publishing B.V.

Nudler E., E. Avetissova, V. Markovtsov and A. Goldfarb (1996) Transcription processivity: protein-DNA interactions holding together the elongation complex. Science, 273, 211-217

Paabo S., D.M. Irwin and A.C. Wilson (1990) DNA damage promotes jumping between templates during enzymatic amplification. J. Biol. Chem., 265, 4718-4721

Pappenheimer A.M. Jr. and D.M. Gill (1973) Diphtheria. Science, 182, 353-357

Rapp G., J. Klaudiny, G. Hagendorff, M.R. Luck and K.H. Scheit (1989) Complete sequence of the coding region of human elongation factor 2 (EF-2) by enzymatic amplification of cDNA from human ovarian granulosa cells. Biol. Chem. Hoppe-Seyler, 370, 1071-1075

Ronen A., J.D. Gingerich, A.M.V. Duncan and J.A. Heddle (1984) Autoradiographic assay of mutants resistant to diphtheria toxin in mammalian cells in vitro. Proc. Natl. Acad. Sci. USA, 81, 6124-6128

Sakamoto H., M. Terada and T. Sugimura (1986) Detection of Mutant Diphtheria Toxin-resistant Chinese Hamster Lung Cells in situ by Autoradiography. Cancer Res., 46, 2041-2044

Sastry S.S. and P.L. Hoffman (1995) The inluence of RNA and DNA template structures during transcript elongation by RNA polymerases. Biochem. Biophys. Res. Commun., 211, 106-114

Sauerbier W. and K. Hercules (1978) Gene and transcription unit mapping by radiation effects. Annu. Rev. Genet., 12, 329-363

Tiah M. and A. Ronen (1989) Autoradiographic detection of mutation to exotoxin-A resistance in mouse fibroblasts treated with ethylmethanesulfonate, X rays and ultraviolet light. Mutation Res., 213, 205-215

Tiah M. and A. Ronen (1991) Dominant lethal cell mutants detected by the autoradiographic assay for exotoxin-A resistance. Mutation Res., 249, 211-222

Van-Ness B.G., J.B. Howard and J.W. Bodley (1980a) ADP-Ribosylation of Elongation Factor 2 by Diphtheria Toxin. J Biol Chem., 255, 10710-10716

Van-Ness B.G., J.B. Howard and J.W. Bodley (1980b) ADP-ribosylation of Elongation Factor 2 by Diphtheria Toxin. J. Biol. Chem., 255, 10717-10720

©1997 Elsevier Science Publishing B.V.

Moloney Murine Leukemia Reverse Transcriptase Suspect in the Production of Multiple Misincorporations During *hprt* cDNA Synthesis

John Curry and Barry W. Glickman

Centre for Environmental Health and the Department of Biology, University of Victoria, Victoria, British Columbia

Abstract

Our laboratory has characterized several hundred mutant *hprt* cDNA's produced using Moloney Murine leukemia reverse transcriptase to convert mRNA to cDNA. During the characterization of these mutants we have detected six T-lymphocyte mutants that demonstrate multiple G:C→A:T transitions along the *hprt* cDNA coding sequence. Attempts to repeat the mRNA to cDNA conversion and subsequent characterization have demonstrated that the multiple transitions are likely artifacts. We suggest that reverse transcriptase is directly responsible for these multiple base substitutions and as such, that multiple mutations be viewed as suspect requiring confirmation at the genomic level.

Communication

The monitoring of *in vivo* human mutations is facilitated by the use of the hypoxanthine-guanine phosphoribosyltransferase (*hprt*) T-cell clonal assay (Albertini *et al.* 1982). Not only can *in vivo* mutant frequencies be determined, the mutations responsible for the observed 6-thioguanine resistant (TG^R) phenotype can also be characterized at the DNA sequence level. This approach can provide useful data on the nature and origin of mutations occurring *in vivo*. However, the properties of the enzymes used in the molecular characterization of mutant *hprt* mRNA's require some circumspection. Specifically, the cDNA's are produced by the reverse transcription of mRNA's contained in cell lysates (Yang *et al.* 1989) or mRNA extracts (Steingrimsdottir *et al.* 1993). Reverse transcriptases commonly used to transcribe cDNA's, such as the Moloney Murine Leukemia Virus (Mo-MLV) and avian myeloblastosis virus (AMV) reverse transcriptase (RT), are characterized by high error-rates (Roberts *et al.* 1989, Bakhanashvili *et al.* 1992, Bakhanashvili *et al.* 1993). Using the M13mp2 *lacZα* gene as

©1997 with permission from Elsevier Science B.V.

template, both RT's were estimated to commit an error at least once every 30,000 nucleotides polymerized (Roberts *et al.* 1989). Indeed, the relatively high *in vivo* mutation rates characteristic of these retroviruses are thought to be a direct consequence of this error proneness.

In principal, errors that occur during the reverse transcription of *hprt* mRNA's should go undetected because cell lysates or extracts are derived from a large number of cells ($2x10^{3-5}$), each containing approximately 8-10 *hprt* mRNA molecules (Steen *et al.* 1990). Two subsequent polymerase chain reactions (PCR) directed with two sets of nested primers (Yang *et al.* 1989) are used to amplify the cDNA providing sufficient amounts of product for direct double stranded sequencing (Winship *et al.* 1989). It would thus be expected that errors produced by Mo-MLV RT would be diluted out by other copies not having a similar error.

Nevertheless, over the past several years we have upon occasion recovered human *hprt* mutant sequences that we suspect were the consequence of Mo-MLV RT mis-incorporation errors. Using the method briefly described above, as modified by Curry *et al.* (1993, 1995), multiple G:C→A:T transitions occurring in a single cDNA product have been characterized from several TG^R T-lymphocyte clones. To date

17-10		**B9-6**		**17-15**		**17-7**	
89	ΓXTΓAΓΓ	96	TTTΓΓAA	306	AXTΓAAΓ	601	AΓΓΓATT
97	TTΓΓAAA	101	AAAΓΓΓT	384	AAAΓAAT	606	TTTΓAAT
361	ΓATΓATX	118	XATΓΓAX	432	ΓXAΓAXT		
452	TXAΓΓXA	197	TXTΓTΓT			**WX1-9**	
469	AAAΓATΓ	403	ΓAAΓATA	**FB16B**		272	ATAΓATX
472	ATΓΓTXA	438	TTTΓXTT	361	ΓATΓATX	285	TATΓAXT
518	TTΓΓATA	492	ΓXTΓΓTΓ	393	XTTΓATT	334[A]	AXAΓΓΓΓ
553	XXAΓAX	569	TAΓΓATA	412	ATTΓAXA	352	ATTΓΓTΓ
	A						
580	XTTΓAXT	599	TXAΓΓΓA	432	ΓXAΓAXT	400	ΓTΓΓAAΓ
628	AΓTΓAAA	601	TTTΓAAT	606[C]	TTTΓAAT	581[B]	TTΓAXTA
664	TΓAΓAΓT						
674	ΓTTΓAΓT						

Table 1. Position and the sequence context of multiple base substitutions characterized. All of these mutations are G:C→A:T transitions except for three transversions; ([A]) is a G:C→C:G; ([B]) is a A:T→T:A; ([C]) is a G:C→T:A.

©1997 Elsevier Science B.V.

we have observed six separate cDNA characterizations where such multiple base substitutions have occurred (Table 1). Attempts to repeat these observations by generating new cDNA from fresh extracts of the same mutant clones proved unsuccessful. In contrast, when the original RT-PCR products (primary PCR) were re-amplified (secondary PCR) and characterized, the same multiple mutations were again encountered. From these and other observations, we believe that the multiple transitions are an artifact occurring in the RT phase and hence generated by the MoMLV reverse transcriptase.

The majority of the multiple base substitutions, in the cases that we have characterized, involve GA dinucleotides (26 of 37 mutants recovered), while other dinucleotides are involved less frequently: GG (8/37), GT (2/37), GC (1/37). Using a Chi square test, with expected values for all four dinucleotides being equal (9.25), the observed values are highly significant (χ^2=43.5, df=3, $p < 10^{-7}$) and demonstrate a predilection for the mechanism responsible here to act at GA dinucleotides. For those events occurring at GA dinucleotides, the preceding nucleotide appears not to have any significant consequence upon the misincorporations: TGA 13/26, AGA 8/26, GGA 5/26 (χ^2=3.8, df=2, $p < 0.14$).

We can not explain how in an excess of template the MoMLV RT reaction can produce products such as the six we describe here. Pathak and Temin (1990) report a similar phenomenon which they have termed hypermutation, while determining the *in vivo* forward mutation rate during a single round of replication of the spleen necrosis virus. Characterization of proviruses revealed 15 G→A transitions which are proposed to be the direct results of mis-incorporation during viral reverse transcription. Two of the substitutions appear to involve tandem events where GG dinucleotides in the sequence G<u>GG</u>AA are replaced with AA dinucleotides. Nine of the fifteen single-basepair substitutions occur at GA dinucleotide sequences, compared with 2/15 at GT and 2/15 at GG. None of the substitutions are preceded by A nucleotides. A simple Chi square analysis demonstrates a significant preference for GA dinucleotides (χ^2=11.9, df=3, $p < 0.008$). This

©1997 Elsevier Science B.V.

GA dinucleotide preference is the same as we observed in our *hprt* study.

Ji and Loeb (1994) employing an *in vitro* system, report several multiple mutations produced by the HIV-1 RT using the *env* gene as template. The authors observed nine multiple mutations (with either 2 or 3 mutations each) along of the *env* gene templates. In addition, the HIV-1 RT enzyme was estimated to have an error rate of approximately one error per 5000 RNA nucleotides copied and of the 64 mutations, 76% were transitions with the most frequent type being G:C→A:T.

We have not been able to develop the conditions required for the RT-PCR step to reproducibly generate multiply misincorporated cDNA's. Such cDNA's have been observed seemingly randomly during the course of many different experiments, over several years. Since they are rare events, we have not sought to develop a more error-free system. It is possible that an occasional clonal lysate contains much fewer accessible cells and thus very few mRNA molecules for reverse transcription. With just a small number of mRNA molecules, RT errors may not be 'washed' out by correct copies. This does not, however, explain the multiple nature of these transitions. Perhaps a defective or inefficient RT enzyme inefficiently and quite inaccurately copies only a few mRNA molecules, thereby producing not only a cDNA with multiple transitions, but one that predominates sufficiently to be scored at the sequencing level.

We can however, not rule out the possibility that in the initial PCR step, it is the synthesis of the 2nd cDNA strand that is the source of the multiple substitutions. Using primers specific for exon 3 of *hprt*, *Taq* polymerase was found to predominantly produce A:T→G:C transitions (Keohavong *et al.* 1989) rather than the G:C→A:T transitions. Moreover, there have been no reports of multiple misincorporations with this enzyme. As the majority of the transitions reported here involve G:C basepairs, the possibility of DNA damage due to incubation of the template at high temperatures also exists. Deamination of cytosine to uracil is the most frequent DNA damage caused in this manner, producing G:C→A:T transitions (Lindahl 1979, Eckert and Kunkel 1991a, 1991b) Nevertheless the probability that multiple DNA damages occur on a single template does seem very unlikely.

©1997 Elsevier Science B.V.

An additional factor may be that *Taq* polymerase exhibits a template preference, and for some reason preferentially amplifies the altered templates during the course of the amplification. This would require that a template altered by MoMLV RT would be preferentially amplified by *Taq* during PCR amplification. Under such circumstances the mutated template would be over represented and preferentially visible in the subsequent sequencing. While our inability to consistently reproduce these observations prevents us from exploring their cause in detail, we felt it important to report their recovery. Our first case of these multiple events arose while following the *hprt* mutant frequency in someone exposed to a potentially genotoxic work environment. These observations thus led to unnecessary fears. Also, in addition to acknowledging these rare events, some speculation on their origin seemed worthwhile.

Acknowledgements

This work was supported by a grant to Dr. Barry Glickman from the Medical Research Council of Canada. We thank Aparecido da Cruz and Magomed Khaidakov for isolating and characterizing the WX1-9 and FB16B mutants. The criticisms of Dr. Elliot Drobetsky and the encouragement of Dr. Adonis Skandalis were greatly appreciated.

©1997 Elsevier Science B.V.

References

Albertini, R.J., K.L. Castle and W.R. Borcherding (1982) T-cell cloning to detect the mutant 6-thioguanine-resistant lymphocytes present in human peripheral blood. Proc Natl Acad Sci USA., 79, 6617-6621.

Bakhanashvili, M. and A. Hizi (1992) Fidelity of the RNA-dependent DNA synthesis exhibited by the reverse transcriptases of human immunodeficiency virus types 1 and 2 and of murine leukemia virus: Mispair extension frequencies. Biochemistry, 31, 9393-9398.

Bakhanashvili, M. and A. Hizi (1993) The fidelity of the reverse transcriptases of human immunodeficiency viruses and murine leukemia virus, exhibited by the mispair extension frequencies, is sequence dependent and enzyme related. FEBS, 319, 201-205.

Curry J., A. Skandalis, J. Holcroft, J. de Boer and B.W. Glickman (1993) Coamplification of *hprt* cDNA and gamma T-cell receptor sequences from 6-thioguanine resistant human T-lymphocytes. Mutation Res., 288, 269-275.

Curry J., G.T. Rowley, V. Saddi, D. Beare, J. Cole, and B.W. Glickman (1995) Determination of *hprt* mutant and mutation frequencies and the molecular characterization of human derived *in vivo* T-lymphocyte mutants. Environ and Molec Mutagen., 25, 169-179.

Eckert, K.A. and T.A. Kunkel (1991a) DNA polymerase fidelity and the polymerase chain reaction. PCR Methods and Applications, 1, 17-24.

Eckert, K.A. and T.A. Kunkel (1991b) The fidelity of DNA polymerases used in the PCR, in: M.J. McPhearson, P. Quirke, G.R. Taylor (Eds.), Polymerase chain reaction I: A practical approach, IRL Press at Oxford University Press, Oxford, pp. 227-246.

Ji, J. and L.A. Loeb (1994) Fidelity of HIV-1 reverse transcriptase copying a hypervariable region of the HIV-1 *env* gene. Virology, 199, 323-330.

Keohavong, P. and W.G. Thilly (1989) Fidelity of DNA polymerases in DNA amplification. Proc Natl Acad Sci USA., 86, 9253-9257.

Lindahl, T. (1979) DNA glycosylases, endonucleases for apurinic/apyrimidinic sites, and base-excision repair. Prog. Nucleic Acid Res. Mol. Biol., 22, 135-189.

Pathak, V.K. and H.M. Temin (1990) Broad spectrum of in vivo forward mutations, hypermutations, and mutational hotspots in a retroviral shuttle vector after a single replication cycle: Substitutions, frameshifts, and hypermutations. Proc. Natl. Acad. Sci. U.S.A., 87, 6019-6023.

Roberts, J.D., B.D. Preston, L.A. Johnston, A. Soni, L.A. Loeb and T.A. Kunkel (1989) Fidelity of two reverse transcriptases during DNA-dependent DNA synthesis in vitro. Molec Cellular Biol., 9, 469-476.

Steen, A., H. Luthman, D. Hellgren and B. Lambert (1990) Levels of hypoxanthine phosphoribosyltransferase RNA in human cells, Experimental Cell Res., 186, 236-244.

©1997 Elsevier Science B.V.

Steingrimsdottir, H., G. Rowley, A. Waugh, D. Beare, J. Cole and A.R. Lehmann (1993) Molecular analysis of mutations in the *hprt* gene in circulating lymphocytes from normal and DNA-repair-deficient donors. Mutation Res., 294, 29–41.

Winship, P.R. (1989) An improved method for directly sequencing PCR amplified material using dimethyl sulphoxide. NAR, 17, 1266-1266.

Yang, J.L., V.M. Maher and J.T. McCormick (1989) Amplification and direct sequencing of cDNA from the lysate of low numbers of diploid human cells. Gene 83, 347-354.

©1997 Elsevier Science B.V.