

Mitochondrial Mutational Spectra in Human Bronchial Epithelial Cells of Smokers and Nonsmokers

by

Hilary A. Collier

A. B. Biochemistry and Molecular Biology
Harvard University, 1989

S. M. Technology and Policy
MIT, 1991

S. M. Toxicology
MIT, 1993

SUBMITTED TO THE DIVISION OF TOXICOLOGY IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

Ph.D. IN TOXICOLOGY
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

NOVEMBER 1997

[February 1998]

©1997 Massachusetts Institute of Technology. All rights reserved.

Signature of Author: _____

Division of Toxicology
November, 1997

Certified by: _____

William G. Thilly
Thesis Supervisor

Accepted by: _____

Peter S. Dedon
Chairman, Committee on Graduate Students

NOV 25 1997

SCIENCE

LIBRARIES

This doctoral thesis has been examined by a Committee of the Division of Toxicology as follows:

Professor David Schauer



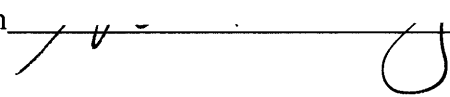
Chairman

Professor William Thilly



Supervisor

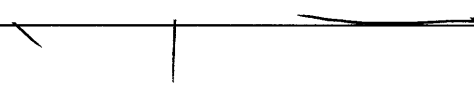
Professor Gerald Wogan



Dr. Thomas Kunkel



Professor Richard Albertini



Mitochondrial Mutational Spectra in Human Bronchial Epithelial Cells of Smokers and Nonsmokers

by
Hilary A. Collier

Submitted to the Division of Toxicology November 1997 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

ABSTRACT

By first separating mutant from nonmutant DNA sequences on the basis of their melting temperatures and then increasing the number of copies by high-fidelity PCR, observation of point mutations in biological samples at fractions at or above 10^{-6} has been achieved. Using this method, hotspot point mutations that lie in 100 bp of the mitochondrial genome (bp 10,030-10,130) in samples of cultured cells and human tissues have been detected. Seventeen hotspot mutants were isolated, their fractions ranging from 2×10^{-3} down to the limit of detection of 10^{-6} . These hotspots are primarily transitions.

Several human tissues such as colon, lung and muscle, and cultured human lymphoblastoid cells, contained the same set of mutational hotspots. Mutational spectra were also determined in bronchial epithelial cells of smokers and nonsmokers in order to specifically address whether smoking induces hotspot mitochondrial mutations. Among the individuals sampled were three pairs of monozygotic twins in which one twin never smoked and the other had a significant smoking history. The mutational spectra in smokers' samples did not differ significantly from the mutational spectra in nonsmokers. No smoking-specific hotspots were detected and the overall mitochondrial mutant fractions in this sequence in smokers' samples were not elevated compared to nonsmokers. In fact, as much variability was observed among sections of the same individual's lung as among samples from smokers and nonsmokers. The similarity of the hotspot sets in different organs and *in vitro*, and the lack of smoking-induced mutations in bronchial epithelial cells, support the conclusion that human mitochondrial point mutations in the sequence studied are primarily spontaneous in origin and arise either from DNA replication errors or reactions of DNA with endogenous metabolites.

While the same mitochondrial mutants were observed in multiple samples, the frequencies varied considerably. In approximately 2% of observations, a particular mutation was observed at greater than five-fold times the mean for the particular mutant. These cases are interpreted to represent the accumulation of multiple mutant mitochondrial DNA copies in a bronchial epithelial stem cell and its descendants.

Thesis Supervisor: William G. Thilly
Title: Professor of Toxicology

ACKNOWLEDGMENTS

The work described in this thesis reflects the efforts of many individuals, scientists and non-scientists, dedicated to the cause of understanding and preventing disease. To them, I am deeply indebted.

I would like to thank my advisor Bill Thilly for his support and enthusiasm, for helping me to think clearly and for focusing on questions that are important rather than just technically feasible. The opportunity to train with a scientist of his caliber has been invaluable.

The members of my committee have also been extremely helpful and attentive. Their advice and guidance has been extremely beneficial as these experiments progressed.

I have been the recipient of three fellowships during my tenure at MIT: the Ida Green Fellowship, the Claire Booth Luce Fellowship and the Martin Fellowship in Environmental Science. These sources of funding have permitted me to focus on my research and on building my own career as a female scientist, and the generous benefactors are gratefully acknowledged.

Further, the National Institutes of Environmental Health Sciences has supported the work specifically reported herein. Without a government supportive of basic research, studies like these would not be possible.

The many collaborators with whom I have worked have contributed substantially to the strength of these studies. I am especially indebted to the physicians and staff of the University of Rochester, including Prof. Mark Utell, Prof. Mark Frampton, and Dr. Alfonso Torres, who performed the brush biopsies and provided much guidance from the clinical perspective. Other collaborators whose contributions I would like to recognize include Dr. James Willey of the Medical College of Ohio, Dr. Sam Singer from Brigham and Women's Hospital, Prof. Jan Vijg and Prof. Jeanne Wei of the Beth Israel Division of Aging, and Prof. Barry Karger, Dr. Frantisek Foret and Dr. Alexey Belenky of the Barnett Institute of Northeastern University.

To those who volunteered for the procedure, especially the identical twins who were willing to travel from Minnesota to Rochester, NY, I am especially grateful for their altruistic contribution. It is the selfless efforts of such individuals that make these studies possible.

I have had the pleasure of mentoring three UROP students, all of whom will be outstanding and compassionate physicians. I wish to express my gratitude to them—Yvonne Chan, Suma Dutta and Wendla Silverberg—for their hard work and enthusiasm.

There are many members of the Thilly laboratory who have directly contributed to the work described here. I would like to specifically acknowledge several. Dr. Paulo André aided with technology development, characterized the *Pfu* noise on the target sequence and isolated many mutants for sequencing. Xiao-Cheng Li also contributed significantly to early feasibility experiments, and isolated and sequenced many mutants. Luisa Marcelino performed the experiments that demonstrated that it is possible to induce mutations in mitochondrial DNA. Pablo Herrero has provided substantial modeling expertise. Prof. Gengxi Hu characterized the nuclear pseudogenes of this mitochondrial sequence; in addition, his guidance as a patient teacher is gratefully acknowledged. Klaudyne Hong assisted with the development of PAPA technology and, along with Andrea Kim and Brindha Muniappan, repeated these experiments and provided helpful cell counts. Dr. John Hanekamp paved the way technologically for these studies and identified this target sequence of mitochondrial DNA. I greatly value the significant contributions to this work made by these colleagues.

I would especially like to thank Dr. Konstantin Khrapko who developed the CDCE technology, led the development effort for the mutational spectra methodology described, and actually performed some of the experiments described in this thesis. Collaborating with such a brilliant scientist has been extremely rewarding; I am privileged to have had this opportunity.

Many, many members of the Thilly laboratory have provided an outlet for scientific discussion and a refuge for emotional support. Further, the excellent administrative staff has made being a student a pleasure. To name just a few of many, I would like to thank Dr. Claudia Buser, Dr. Rita Cha, Dr. Jia Chen, Rita De Meo, Jacklene Griffith, David Katz, Paula Lee, Dr. Lucy Ling, Wen Luo, Dr. Lata Shimamé-Moré, Dr. Joyce Morrill, Dr. Riitta Mustonen, Cindy Silck-Flannery, Tracey Slayton, Peter Southam, Aoy Tomita, Beth Ann Turnquist, and Dr. Reinhold Wasserkort.

Special recognition is owed to my dear friends who have provided so much friendship, good advice and encouragement during these years. As I get older and hopefully wiser, I have learned to cherish these friendships as I value little else.

Finally and most importantly, I wish to thank my extremely supportive extended and nuclear family, including my grandmother, my sister and my parents. My father has always been inspirational as a careful, dedicated scientist, committed to understanding biology and harnessing it to cure disease. I thank both of my parents for encouraging me to pursue my ideas relentlessly, while providing the emotional support I needed when such paths proved difficult. Thank you.

TABLE OF CONTENTS

	<u>Page</u>
Title Page	1
Committee Page	2
Abstract	3
Acknowledgments	4
Table of Contents	6
List of Figures	9
List of Abbreviations	11
1. Introduction	12
2. Literature Review	13
2.1 Mutational Spectra	13
2.1.1 Introduction	13
2.1.2 Phenotypic selection	13
2.1.3 Mutational spectra to define mutational pathways	16
2.1.4 <i>En masse</i> mutational spectra	19
2.1.5 Examples of mutational spectra	22
2.1.6 Unselected mutational spectra	25
2.2 Mitochondrial DNA as a target for mutational spectrometry	26
2.2.1 Introduction	26
2.2.2 Advantages	26
2.2.3 Differences from nuclear DNA	27
2.2.4 Mitochondrial DNA and human disease	30
2.2.5 Low-frequency somatic mitochondrial DNA alterations	34
2.2.6 Distribution of mitochondrial mutants	40
2.2.7 Mitochondrial DNA and aging	41
2.2.8 Target sequence and melting map	43
2.3 Bronchial epithelial cells and tissue kinetics	48
2.3.1 Cell types	48
2.3.2 Stem cells	48
2.3.3 Cell kinetics	49
2.3.4 Cells and lung cancer	49
2.4 Smoking	52
2.4.1 Age-specific lung cancer rates	52
2.4.2 Smoking and lung cancer	52
2.4.3 Mutagenicity of selected components of cigarette smoke	57
2.4.4 Smoking and adducts	61
2.4.5 Smoking and mutations	63

	<u>Page</u>
3. Materials and Methods	68
3.1 Human cells and tissues	68
3.1.1 TK6 cells	68
3.1.2 Human samples	68
3.1.3 Biopsies obtained via brush bronchoscopy	68
3.1.4 Lung dissection	69
3.1.5 Other organs	69
3.2 Mutational spectra	69
3.2.1 DNA isolation and restriction digestion	69
3.2.2 Constant denaturant gel electrophoresis	69
3.2.3 High-fidelity PCR	70
3.2.4 Constant denaturant capillary electrophoresis	73
3.2.5 High-resolution CDCE	74
3.2.6 Capillary hybridization	74
3.2.7 Test for asymmetric species	74
4. Results	76
4.1 Mutational spectra methodology	76
4.1.1 Overview	76
4.1.2 Copy number measurement and internal standard	76
4.1.3 Pre-PCR enrichment	79
4.1.4 High-fidelity PCR	79
4.1.5 Post-PCR enrichment by CDCE	80
4.1.6 High-resolution CDCE	83
4.1.7 Identification of mutants	83
4.1.8 Quantification of mutants	86
4.2 Tests for validity	86
4.2.1 Reconstruction experiments	86
4.2.2 Spectra differ from each other	89
4.2.3 Reproducibility	94
4.2.4 <i>Pfu</i> noise characterization	94
4.2.5 Test for asymmetric species	94
4.3 Mitochondrial mutational spectra	102
4.3.1 Several organs	102
4.3.2 Hotspots	102
4.3.3 Tumor spectra	102
4.3.4 TK6 cells	109
4.3.5 Treated cells	109
4.3.6 Mutant fractions and mutation rates	116
4.4 Comparison of smokers and nonsmokers	119
4.4.1 Twins discordant for smoking	119
4.4.2 Dissected lungs	119
4.4.3 Average mutant fractions	126
4.5 Variability in mutant fractions	126

	<u>Page</u>
4.6 Mutant fractions and age	136
4.6.1 Infant	136
4.6.2 Bronchial epithelial cells	136
4.6.3 Other tissues	141
5. Discussion	143
5.1 Summary	144
5.2 Sources of observed mutants	145
5.2.1 Spontaneous	145
5.2.2 Endogenous mutagens or replicative errors?	146
5.2.3 Comparison with MNNG-induced spectrum	146
5.3 Potential sources of error	147
5.3.1 Internal standard	147
5.3.2 Peak measurement and identification	147
5.3.3 Missed mutants	148
5.4 Lack of smoking-induced spectra	148
5.4.1 Sequence is vacant	148
5.4.2 Subset of population is at risk	149
5.4.3 High spontaneous mutant fraction	149
5.4.4 Alternatives to mutagenesis	150
5.5 Distribution of mutants	152
5.5.1 Mutant-rich clusters	152
5.5.2 Modeling mutant accumulation	153
5.5.3 Mitochondrial proliferation?	158
5.6 Aging and mitochondrial DNA	159
5.6.1 Deletions	159
5.6.2 Point mutations	160
6. Conclusions	161
7. Suggestions for future research	162
7.1 Sources of human mutations	162
7.2 Mitochondrial mutants with age in different tissues	162
7.3 Mechanism of smoking-induced lung cancer	163
Literature cited	165
Appendix: Table of mutant fractions	191

LIST OF FIGURES

<u>Figure</u>	<u>Page</u>
1. Demonstration that mutational spectra are characteristic of the mutagen	14
2. <i>Hprt</i> mutant fraction as a function of age	17
3. <i>En masse</i> approach to mutational spectra	20
4. Mutational spectra in <i>hprt</i> exon 3 induced by different chemicals and detected by DGGE	23
5. Pathogenic point mutations in mtDNA	32
6. Increase in the ratio of deleted to total mtDNA with age	35
7. Tissue specific selection for mitochondrial haplotypes	38
8. Melting map of the mitochondrial target sequence used in these studies	44
9. Clover leaf of tRNA ^{Gly} in mtDNA	46
10. Proliferative indices for bronchial epithelial tissues	50
11. Age-specific lung cancer rates	53
12. Correlation between smoking and lung cancer risk	55
13. NNK metabolism	59
14. Proportions of the common deletion with age, smoking index, and lung cancer status	65
15. Velocity of DNA migration through the heated zone of the capillary vs. separation temperature	71
16. Overview of the procedure	77
17. Post-PCR enrichment of mutants by CDCE	81
18. Mutant identification by comigration with authentic standards	84
19. Capillary hybridization to confirm peak identification	87
20. Limit of detection: Reconstruction experiment	90
21. Examples of mitochondrial mutational spectra	92
22. Reproducibility of the mutational spectra procedure	95
23. Characterization of <i>Pfu</i> -induced noise on mitochondrial target sequence	97

<u>Figure</u>	<u>Page</u>
24. Test for asymmetric species: low T_m mutants	100
25. Test for asymmetric species: high T_m mutants	102
26. Mutational spectra of lung, colon and muscle samples	105
27. The mitochondrial DNA sequence (bp 10011-10215) used as a target sequence for these studies	107
28. Mutational spectra of colon and muscle tumor samples	109
29. Mutational spectra of TK6 cells	112
30. MNNG mutational spectrum in mitochondrial DNA of MT1 cells	114
31. Position, type of mutation and frequency of hotspot mutations	117
32. Sample of low T_m mutational spectra from twins bronchial epithelial cells	120
33. Mitochondrial mutational spectra of a muscle tumor sample sliced with a microtome	122
34. Mutational spectra of bronchial epithelial cell samples from three pairs of twins discordant for smoking	124
35. Mean mutant fractions in smokers and nonsmokers	127
36. Examples of mutant fraction histograms	129
37. Histogram of the number of samples versus fraction of the mean for all peaks	132
38. Ratio of mutant fractions for p1.5 and p1 observed in limiting dilution experiments with TK6 cells grown 400 generations	134
39. Mitochondrial mutational spectra of bronchial epithelial cells from an infant	137
40. Mitochondrial mutant fractions with age	139
41. High T_m mutational spectra of a heart sample from a 77-year old man	142
42. Distribution of mtDNA genotypes in colonic crypts	154
43. Modeling the accumulation of mitochondrial mutants	156

LIST OF ABBREVIATIONS

6TG	6-thioguanine
6TGr	6-thioguanine resistant
8-OH-dG	8-hydroxydeoxyguanine
ADP	adenosine diphosphate
APC	adenomatous polyposis coli
ATP	adenosine triphosphate
BaP	benzo[a]pyrene
bp	base pairs
BPDE	benzo[a]pyrene-7,8-diol-9,10-epoxide
BSA	bovine serum albumin
CDCE	constant denaturant capillary electrophoresis
CDGE	constant denaturant gel electrophoresis
CSC	cigarette smoke condensate
DGGE	denaturing gradient gel electrophoresis
EDTA	[ethylenedinitrilo]-tetraacetic acid
ELISA	enzyme-linked immunosorbent assay
FADH ₂	flavin adenine dinucleotide
GPA	glycophorin A
H ₂ O ₂	hydrogen peroxide
HNPCC	hereditary nonpolyposis colorectal cancer
ID	inner diameter
LHON	Leber's hereditary optic neuropathy
MAMA	mismatch amplification mutation assay
MELAS	mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes
MERRF	myoclonus epilepsy with ragged-red fibers
MNNG	N-methyl-N'-nitro-N-nitrosoguanidine
MnSOD	manganese superoxide dismutase
MNU	methyl-N-nitrosourea
MPTP	1-methyl-4-phenyl-1,2,5,6-tetrahydropyridine
mt	mitochondrial
NAB	N'-nitrosoanabasine
NADH	nicotinamide adenine dinucleotide
NARP	neuropathy, ataxia, and retinitis pigmentosa
NNAL	4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol
NNK	N-(methylnitrosamino)-1-(3-pyridyl)-1-butanone
NNN	N-nitrosornicotine
PAH	polycyclic aromatic hydrocarbon
PCR	polymerase chain reaction
RFLP	restriction fragment length polymorphism
rRNA	ribosomal RNA
SIDS	sudden infant death syndrome
SDS	sodium dodecyl sulfate
TAE	tris-acetic acid-EDTA
TBE	tris-boric acid-EDTA
TEMED	N, N, N', N'-tetra-methyl-ethylenediamine
T _m	melting temperature
TMR	tetramethylrhodamine
tRNA	transfer RNA

1. INTRODUCTION

While the hypothesis that chemicals in the environment induce mutations that transform normal human cells into a tumorous phenotype is plausible, except for a few clear cases, a causal link has not been demonstrated. Interest in testing this central premise has guided the experimental goal of measuring the point mutational spectra in human tissues. Studies in phage, bacteria, and human cell culture have demonstrated that treating cells with different chemicals results in characteristic mutational spectra (Benzer et al., 1958; Cariello et al., 1990; Coulondre et al., 1977; Keohavong et al., 1992). We hypothesized that the most important mutagenic agents in human tissues could be identified based on the mutational spectra observed.

Toward this end, a method was developed that permits measurement of hotspot point mutants that lie in a defined 100 bp target sequence in the mitochondrial DNA of human cell and tissue samples (Khrapko et al., 1997). Single base pair point mutants present at fractions as low as 10^{-6} are detectable with this technique. The methodology employs an initial separation of DNA copies containing mutations in the target sequence from wild-type DNA with constant denaturant gel electrophoresis (CDGE) (Hovig et al., 1991) followed by high-fidelity PCR with *Pfu* DNA polymerase. For further enrichment of mutants and visualization of mutational spectra, constant denaturant capillary electrophoresis (CDCE), a variant of capillary electrophoresis that includes a high temperature zone for separation of mutants (Khrapko et al., 1994b), is employed due to its speed, reproducibility, and excellent resolution.

The presence of several hundred to a thousand copies of mitochondrial DNA per cell (Khrapko et al., 1997; Robin et al., 1988) permits greater statistical power for a given sample size. Use of a mitochondrial DNA target has permitted reproducible detection of rare mutants from the small number of cells obtainable from bronchoscopy samples. While the accumulation of somatic point mutations in mitochondrial DNA has been suggested as an important contributor to human diseases, including cancer (Shay et al., 1987; Warburg, 1956; Yamamoto et al., 1992) and aging (Aspnes et al., 1997; Bandy et al., 1990; Cortopassi et al., 1990; de Grey, 1997), the most important contributors to mutagenesis in mitochondrial DNA are not necessarily also the most significant mutagenic agents in nuclear DNA.

Experiments were performed to determine the frequencies of mutations present in human tissue samples from multiple organs, two tumor samples and cells in culture. In addition, whether smoking induces a significant change in the kind and number of mutations in the mitochondrial DNA of smokers' bronchial epithelial cells was specifically addressed. Smoking was selected as a model for testing whether a particular environmental agent is inducing a characteristic mutational spectra. Smoking is an appealing candidate as a model because it represents a clear case of high and continuous exposure to a mixture of chemicals containing known, powerful chemical mutagens. Also, a clear dose response relationship between the number of cigarettes smoked per day and lung cancer risk has been demonstrated by several large epidemiological studies (IARC, 1986). Further, dose information is easily obtainable for smoking as opposed to many other environmental agents for which exposure is often difficult to assess.

Smokers are good candidates for such experiments because the target cell for smoking-induced lung cancer, the bronchial epithelial cell, can be harvested for analysis. In these experiments, bronchial epithelial cells were obtained by brush biopsy from several individuals including three pairs of monozygotic twins discordant for smoking. Bronchial epithelial cells were also harvested from whole, dissected lungs. The mitochondrial mutational spectra of these samples was determined in order to test for smoking-induced mutations.

2. LITERATURE REVIEW

2.1 MUTATIONAL SPECTRA

2.1.1 Introduction

Mutational spectra were discovered by Seymour Benzer and Ernest Freese (Benzer, 1961; Benzer et al., 1958) who showed that the position and frequency of mutations in the genome of the T4 bacterial virus were different for untreated and chemically treated virus populations (see Figure 1). Each mutagen was observed to induce a set of characteristic mutagenic "hotspots:" specific mutations, base substitutions or small deletions or insertions, that occur within a treated population more frequently than would be expected by chance. Since then, the specificity of mutational spectra for many mutagens has been extended to bacteria, yeast, and rodent and human cells [(Thilly, 1990) and references therein (Armstrong et al., 1990; Roy et al., 1994)]. In particular it has been recognized that spontaneous mutations are also characterized by a reproducible set of predominant point mutations (Benzer et al., 1958; Oller et al., 1992).

The experimental goal addressed in the studies presented herein has been to determine the pattern of mutations that have accumulated in somatic cells in human tissues over the course of an individual's life. Ultimately, such observations may provide information about which of the many mutagenic agents to which humans are continuously exposed are most important in terms of producing mutations.

2.1.2 Phenotypic selection

Most investigators of mutational spectra use an approach in which cells containing mutations in a target gene are selected based on a change in phenotype. DNA isolated from individual colonies, each containing a single mutant, is sequenced. Mutational spectra data are therefore generated one mutant at a time.

The most frequently employed target for these studies is the *hypoxanthine-guanine phosphoribosyl transferase* (*hprt*) gene. The HPRT enzyme permits cells to salvage purines from the cell's microenvironment and incorporate them into synthesized DNA. Cells with mutations in the *hprt* gene that render their HPRT enzymes defective, are resistant to the effects of toxic purine analogs including 6-thioguanine (6TG) (Szybalski et al., 1962). By plating cells in the presence and absence of 6TG, it is possible to determine the *hprt* mutant fraction for a given cell population (Furth et al., 1981).

Albertini and colleagues pioneered a method for plating T-lymphocytes derived from human peripheral blood in the presence and absence of 6TG to determine the *hprt* mutant fraction of human samples (Albertini, 1985; Albertini et al., 1982; Albertini et al., 1985). This approach is now being used by investigators at several laboratories. The methodology has permitted determination of the average mutant fraction and the variation in this mutant fraction within and among individuals. 6-thioguanine resistant (6TG^r) mutant fractions in lymphocytes of adults generally range from 1×10^{-6} to 4×10^{-5} (Tates et al., 1991). Multiple studies have reported linear increases in mutant frequencies with age (Robinson et al., 1994; Tates et al., 1991; Vijayalaxmi et al., 1984). In these experiments, it is important to be careful not to confuse a decrease in cloning efficiency with age with a rise in mutant fraction with age. Decreased cloning efficiency can lead to artifactual increases in mutant fraction because the presence of wild-type cells in the selection plates can suppress the recovery of HPRT mutants in the assay, and the suppression is stronger in samples with high cloning efficiencies (Khaidakov et al., 1996). A meta-analysis of the effects of age on mutant fraction from four different laboratories revealed that a doubling of age was associated with a 52% increase in mutant frequency in one laboratory, a 66% increase in mutant frequency in another, an 87% increase in mutant frequency in a third and

Fig. 1: Demonstration that mutational spectra are characteristic of the mutagen.

Distribution of mutations induced by a variety of compounds on the rII gene in bacteriophage.

Source: Benzer, 1961.

no increase in the fourth (Robinson et al., 1994) (see Figure 2). As is clear from Figure 2, infants have significantly lower *hprt* mutant fractions than adults (note that the scale is log). A mean *hprt* mutant frequency of 0.64×10^{-6} was found in 45 placental blood samples, which is approximately 10-fold lower than mean values in normal adults (McGinniss et al., 1990). Increased *hprt* mutant fractions have been observed among populations exposed to mutagens in an occupational setting [for instance, in cyclophosphamide production (Hüttner et al., 1990)] or as a treatment for cancer (Nicklas et al., 1991; Branda, 1991; Caggana, 1991).

When *hprt* mutant fractions are determined by cloning individual T-cells, it is possible to assess the contribution that expansion of a single clone makes to the *hprt* mutant fraction observed (Curry et al., 1995; Nicklas et al., 1986). Each T-cell clone expresses a T-cell receptor with a characteristic restriction pattern due to gene rearrangement. The T-cell receptor restriction pattern of each 6TG^r clone can be characterized in order to differentiate whether two clones received the same mutation independently or whether the two clones were derived from clonal expansion of a single mutant T-cell. Descendants of a clonal expansion of a single T cell carrying an *hprt* mutation have been observed to represent a substantial fraction of the mutants isolated from a single individual (Curry et al., 1995; Nicklas et al., 1988).

There are two major limitations of determining mutational spectra by plating for 6-thioguanine resistance. First, generating the necessary number of mutants to draw statistically significant results is labor intensive and expensive when mutants are isolated one at a time. Second, it is generally necessary to use a cell type with a high plating efficiency *in vitro* and that can be grown in culture for sufficient time to allow phenotypic selection. This requirement has generally [although not completely, see (Martin et al., 1996)] limited use of the *hprt* mutation assay to T-lymphocytes.

Other approaches have also been used to measure mutant fractions in human tissues for different genes. An assay for mutations in the HLA locus has been developed that permits measurement and molecular analysis of both gene and chromosomal mutations (Janatipour et al., 1988; McCarron et al., 1989). Individuals heterozygous for HLA-A2 or HLA-A3 can be studied. In such individuals, there is a single copy of the target gene (either HLA-A2 or HLA-A3) and thus mutants within it are selectable. Mutant lymphocytes are immunoselected by exposing them to monoclonal antibodies against the appropriate HLA followed by complement. Wild-type cells are killed by the treatment; HLA mutants survive. Mutant frequency can be calculated as the ratio of the cloning efficiency for selected cells to the cloning efficiency for non-selected cells. DNA isolated from expanded mutant clones can be examined for the presence or absence of polymorphic loci on chromosome 6 where the targeted HLA genes reside. Point mutations, gene deletions, mitotic recombination, gene conversion or whole chromosome hemizygosity can consequently be detected. The results of 167 assays in 73 individuals showed that mutant frequencies increased significantly with age from a mean of 0.71×10^{-5} in neonates to 6.53×10^{-5} in elderly (Grist et al., 1992). Molecular analysis of 434 mutants from 31 individuals revealed that gene deletion events (defined as loss of the selected allele and no loss of heterozygosity elsewhere) had occurred in 2.8% of cases, mitotic recombination (defined as loss of the selected allele and loss of heterozygosity at the nearby factor XIII gene) in 32.5% of cases and no change on Southern blotting (suggesting that a point mutation is responsible for the phenotypic alterations) in 64.7%.

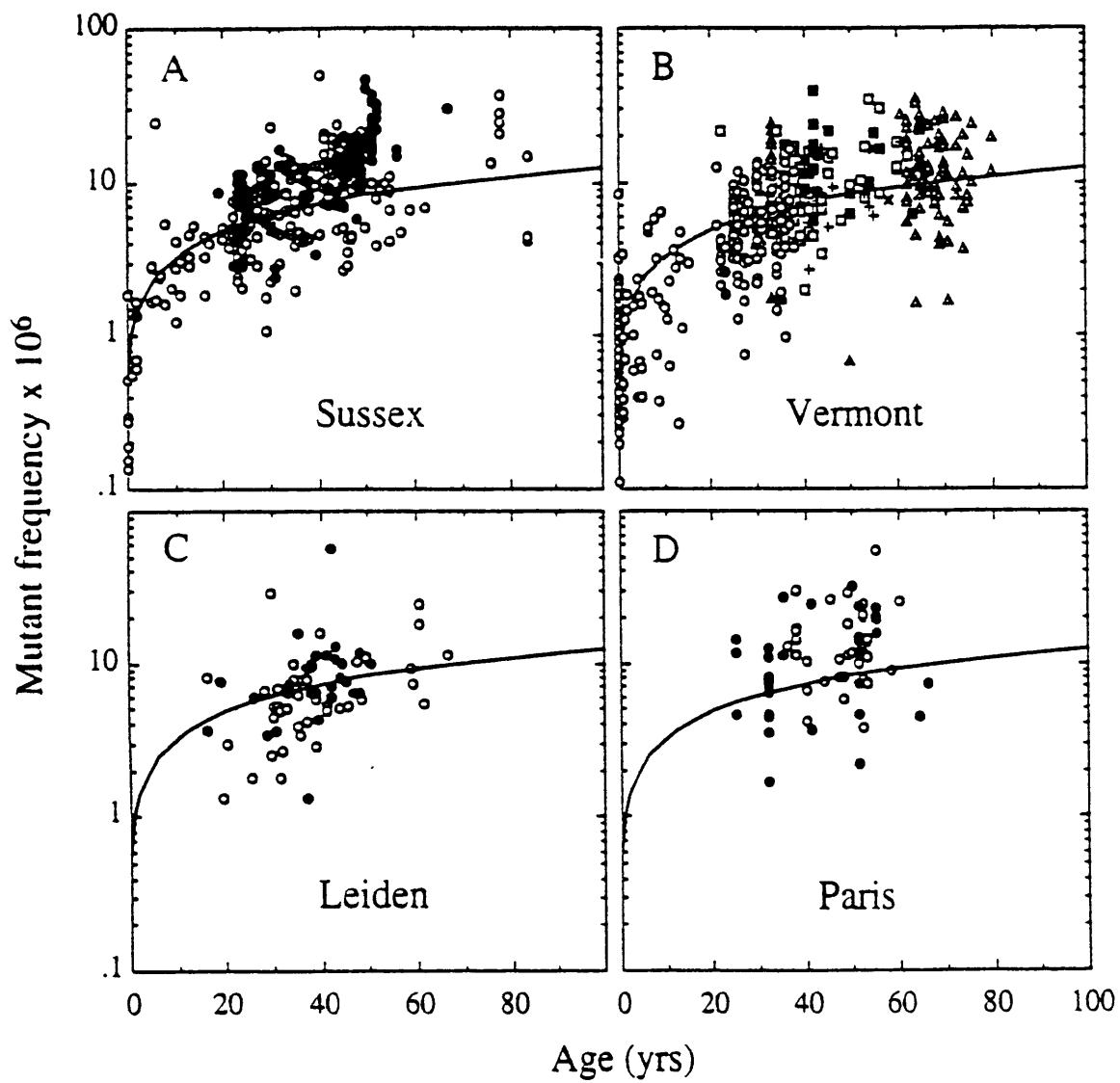
2.1.3 Mutational spectra to define mutational pathways

Mutational spectra data have been pivotal in the assignment of the most likely mutagenic pathway for several model systems. For instance, determination of the mutational spectra of the *lacI* gene in *E. coli* revealed that the two most highly mutable hotspots, which incurred mutations at 10-fold higher than the average frequency, were

Fig. 2: *Hprt* mutant fraction as a function of age.

Mutant frequency adjusted for cloning efficiency plotted against subject age. The line indicates the best fit for all the combined data assuming non-smoking status, and 50% cloning efficiency. (A) Sussex. newborns, children and adults: non-smokers ○; smokers, ●. (B) Vermont. newborns, children, and normal laboratory subjects: non-smokers, ○; smokers ●. Nurses: non-smokers, +; smokers, X. Breast cancer controls: nonsmokers, □; smokers, ■. Twins: nonsmokers, Δ; smokers, filled Δ. (C) Leiden. Normal subjects: non-smokers, ○; smokers, ●. (D) Paris. Laboratory controls and breast cancer patients before treatment: nonsmokers, ○; smokers, ●.

Source: Robinson et al., 1994



cytosines that were discovered to be methylated (Coulondre et al., 1978). When a novel methylated cytosine was engineered into the sequence, this site was also a hotspot for GC->AT transitions (Coulondre et al., 1978). Deamination of 5-methylated cytosine or misreplication past a 5-methylcytosine, are consequently candidates for the most important mutational pathway in this system.

Mutational spectra data also proved valuable in identifying the mutagenic pathways contributing to mutations observed in tumors of individuals with a disease that causes a predisposition to colon cancer termed hereditary nonpolyposis colorectal cancer (HNPCC). These patients inherited a defect in one of the proteins involved in mismatch repair (Papadopoulos et al., 1994; Hemminki, 1994; Leach, 1993; Nicolaides, 1994). Colon tumors in individuals with inherited defects in mismatch repair proteins show an increase in frameshift mutations (Aaltonen et al., 1993; Fishel, 1993; Ionov, 1993; Boyer, 1995). An increased frequency of frameshifts has also been observed in sporadic patients in addition to those with a family history (Chen et al., 1997), and this particular pathway may be important in as much as 20% of colon cancers (Heinen et al., 1995). These frameshifts may reflect polymerase slippage due to a transient misalignment of the primer and template (Streisinger et al., 1966; Kunkel, 1990). Increased frequency of frameshifts had been previously shown to be a phenotype of mismatch repair deficient cells (Levinson et al., 1987), and this correlation between the *in vivo* spectra and *in vitro* spectra

2.1.4 *En masse* mutational spectra

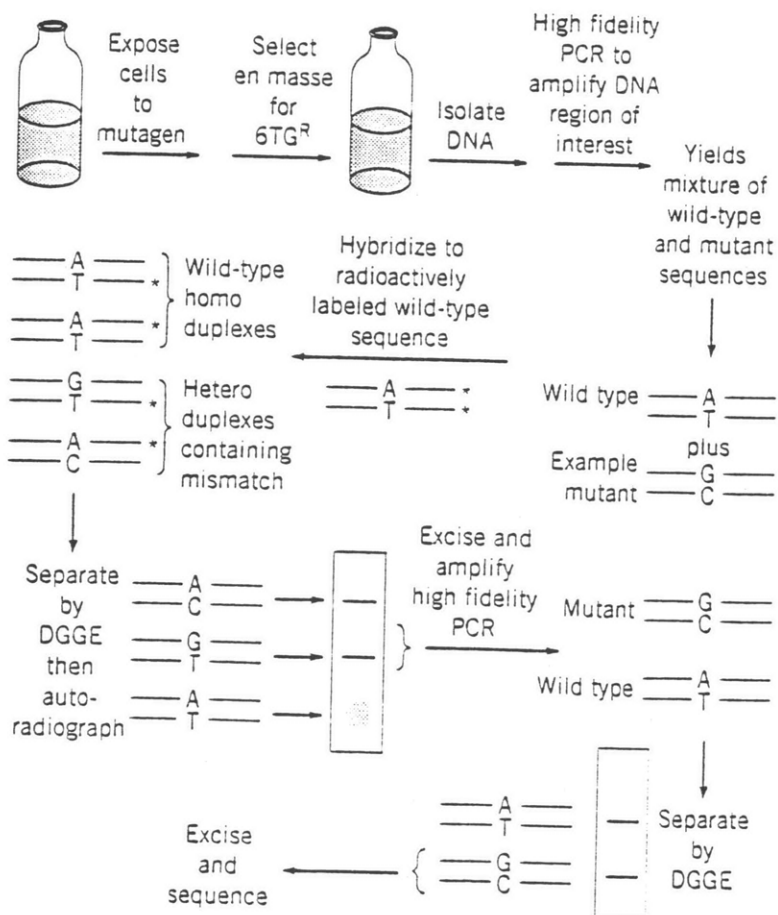
The goal of the W. Thilly laboratory has been to develop an unselected approach to mutational spectra that would permit analysis of statistically significant hotspot mutations in a variety of human tissues. One important milestone along this path was development of a method that allows processing of many mutants at once, thus permitting identification of statistically significant hotspots without sequencing many individual colonies (Cariello et al., 1990; Keohavong et al., 1992; Thilly, 1985) (see Figure 3). Human lymphoblastoid cells from the cell line TK6 (Skopek et al., 1978) were treated with mutagenic agents, and selected *en masse* for mutants in the *hprt* gene with 6TG. This selection increased the mutant fraction for an individual mutant from approximately 10^{-7} to 10^{-3} . DNA was extracted from 6TG^r cells and *hprt* exon 3 was amplified with a high-fidelity DNA polymerase. Hotspot mutants were detected with denaturing gradient gel electrophoresis (DGGE).

DGGE was developed by Fischer and Lerman (Fischer et al., 1983) as a practical application of statistical mechanics concerning cooperative equilibria in macromolecules that predict DNA would melt in cooperative domains (Gotoh et al., 1981; Poland, 1974; Poland et al., 1966). Mutations in the low melting domain of a DNA fragment with a biphasic melting profile were discovered to alter the melting temperature (Fischer et al., 1983). As DNA is electrophoresed through a gel containing an increasing gradient of denaturing chemicals, specifically urea and formamide, the low melting domain of fragments with a biphasic melting temperature "opens" at a given denaturant concentration. In this "open" conformation, the high melting domain portion of the sequence remains double stranded while the low melting domain is denatured for a significant portion of the time. DNA in the open conformation moves more slowly through the gel and therefore appears to "focus" at a denaturant concentration that reflects the melting temperature of the low melting domain. Mutations in the low melting domain consequently affect the position in the gel at which DNA focuses. Mutants can be converted into heteroduplexes, i.e., double stranded molecules that have a single base mismatch, for instance A<>C or G<>T, by denaturing and reannealing the products to force mutants to anneal with the wild-type in excess. Heteroduplexes have an even greater difference in melting temperatures from wild-type, thus creating heteroduplexes makes the hotspots more easily resolved from wild-type. This

Fig. 3: *En masse* approach to mutational spectra.

The protocol for an *en masse* approach to detecting mutational spectra is shown. Cells are exposed to a mutagen, and 6-thioguanine resistant cells are selected. DNA is isolated and subjected to high-fidelity PCR to amplify the DNA region of interest. DNA is melted and reannealed so that mutants are present as heteroduplexes containing a mismatched base. DNA is separated by DGGE to determine hotspots, and individual mutants can be excised and sequenced.

Source: E. C. Friedberg et al., (eds.) *DNA Repair and Mutagenesis*, 1995.



approach permitted detection of hotspot point mutations in DNA isomelting domains of approximately 100 bp in cultured human cells treated with mutagenic compounds (Cariello et al., 1990; Chen et al., 1994; Kat, 1992; Keohavong et al., 1992; Oller et al., 1992).

2.1.5 Examples of mutational spectra

The *en masse* selection approach was applied to several different compounds and to untreated cells. Mutational spectra were discovered to be characteristic of the mutagen used. Such experiments were performed with cells treated, for example, with MNNG (Cariello et al., 1990), benzo[a]pyrene-diol-epoxide (BPDE) (Keohavong et al., 1992), and hydrogen peroxide (Oller et al., 1992). A summary of these results is provided in Figure 4 and the spectra for spontaneous cultures and three compounds are described in more detail below.

2.1.5.1 N-methyl-N'-nitro-N-nitrosoguanidine (MNNG)

The mutational spectra induced by N-methyl-N'-nitro-N-nitrosoguanidine (MNNG) at the third exon of *hprt* was determined in two cultures of TK6 cells. The cells were treated with MNNG, one culture at 10 ng/ml and one culture at 15 ng/ml (Cariello et al., 1990). The population treated with 10 ng/ml showed hotspot GC->AT transitions at the second and third guanine in a run of six guanines with the third guanine mutated more frequently than the second. The population treated with 15 ng/ml MNNG showed a single GC->AT transition hotspot at the second guanine in the six guanine sequence.

2.1.5.2 Benzo[a]pyrene-diol-epoxide

The mutational spectra induced by the polycyclic aromatic hydrocarbon (BPDE) in human TK6 cells was determined by *en masse* treatment and selection in quadruplicate cultures (Keohavong et al., 1992). An average of 1.6×10^4 6TG^r mutants were created per culture. High-fidelity PCR and DGGE separation revealed 16 mutations in exon 3 that were extremely reproducible among the cultures. The mutations were predominantly GC->TA transversions, although GC->AT, GC->CG, and AT->TA substitutions, single base pair deletions and other more complex substitutions were also observed. The mutant frequencies of the specific mutants observed ranged from 0.7 to 2.0% of the 6-thioguanine resistant cells. Of the mutations observed, 6 occurred within a run of 6 guanines and 5 occurred in the sequence 5'-GAAGAG-3'.

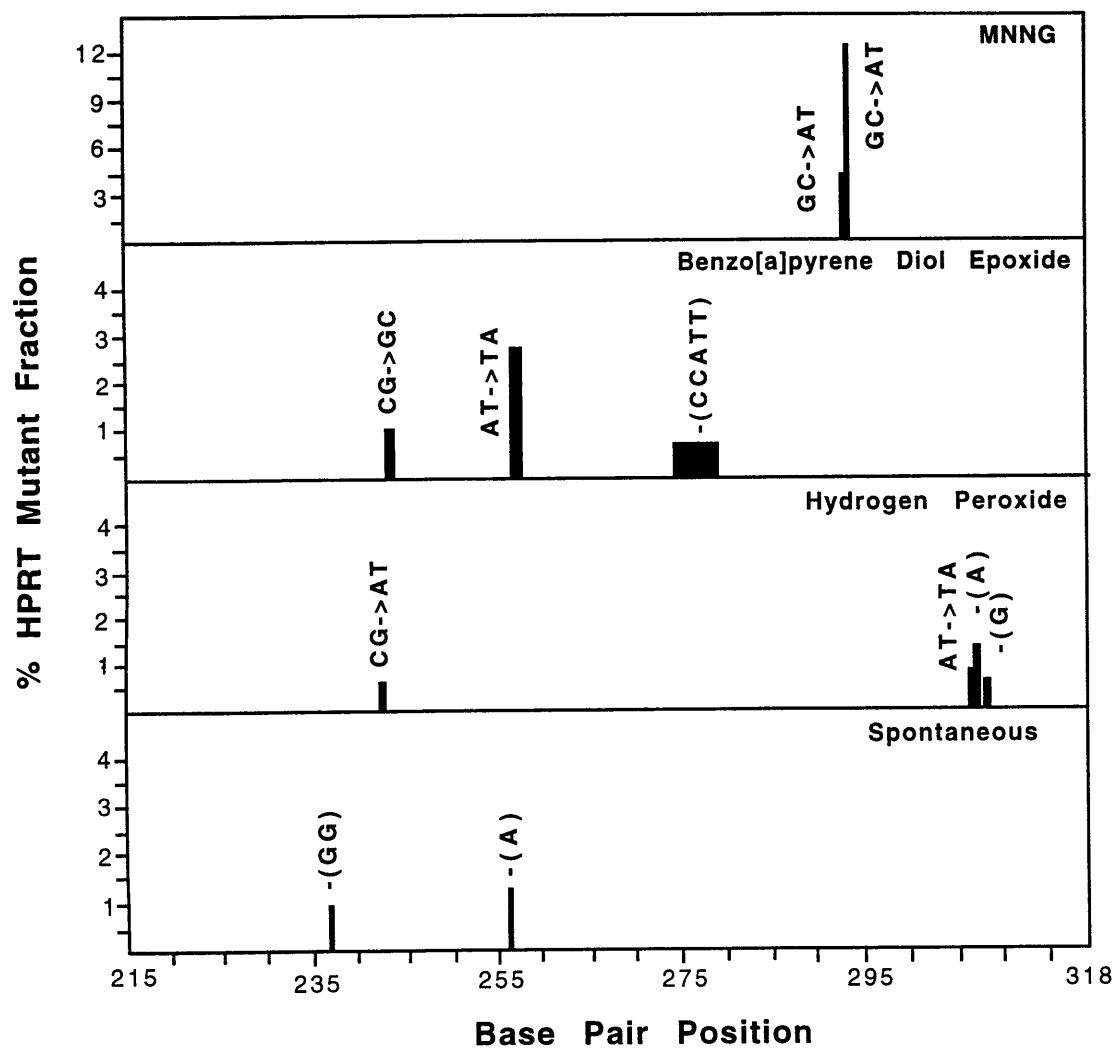
2.1.5.3 Hydrogen peroxide

Triplicate cultures of TK6 cells were treated with hydrogen peroxide to induce *hprt* mutant fractions ranging from 2.1×10^{-6} to 3.8×10^{-6} above the mutant fraction in cells before treatment (Oller et al., 1992). Hydrogen peroxide- (H₂O₂)-induced mutations (1 hr, 20 mM) differed from the spontaneous spectrum. Four H₂O₂-induced hotspots were observed in exon 3 of *hprt*. An AT->TA transversion was present at an average frequency of 3% of all 6TG^r mutants; two GC->CG transversions were detected at an average frequency of 0.6% and 0.4%, respectively; and a five bp deletion (bp 274-278) was present at an average frequency of 0.3%. The AT->TA transversion was detected in all three replicate cultures; the other three mutations were detected in two out of three replicate cultures.

Fig. 4: Mutational spectra in *hprt* exon 3 induced by different chemicals and detected by DGGE.

Comparison of the mutational spectra induced by a variety of mutagens on the same portion of the *hprt* gene using *en masse* treatment and selection followed by PCR and DGGE. Mutational spectra induced by MNNG, benzo[a]pyrene diol epoxide, and hydrogen peroxide, and the spontaneous spectra are shown.

Sources: Cariello et al., 1990; Keohavong and Thilly, 1992; and Oller and Thilly, 1992.



2.1.5.4 Spontaneous

Large cultures of 7×10^8 TK6 cells were grown for 19 days without treatment to generate spontaneous mutations, and then selected for *hprt* mutants with 6TG (Oller et al., 1992). The 6TG^r mutant fractions in the triplicate cultures after growth were 3.2×10^{-6} to 5.5×10^{-6} . DNA isolated from the cultures was analyzed for point mutational hotspots in the third exon of the *hprt* gene using DNA amplification and denaturing gradient gel electrophoresis. Two point mutational hotspots were observed. A deletion of the two adenines in the sequence GAAT and deletion of both guanines in the sequence CTGGAT were detected in triplicate cultures. Each of the hotspots comprised about 1% of the total 6TG^r mutants.

2.1.6 Unselected mutational spectra

Determining the most important mutagenic pathways in human tissues requires a methodology for determining mutational spectra that is independent of phenotypic selection for mutants. Recently, approaches to detecting rare ($<10^{-4}$) mutations in DNA based on genotype rather than phenotype have been developed. For example, allele specific PCR (Cha et al., 1992; Wu et al., 1989a), ligase chain reaction (Landegren et al., 1988; Wu et al., 1989b; Barany, 1991; Abravaya, 1995), or high efficiency restriction digestion (Felley-Bosco et al., 1991; Kan et al., 1978; Khrapko et al., 1994a; Kumar et al., 1990) permit detection of point mutations in target sequences from one to six base pairs.

In order to detect mutational spectra for 100 base pair target sequences, separation of mutant and wild-type DNA was achieved by physical-chemical means. Enriching for mutants prior to PCR is critical so that PCR noise does not obscure visualization of mutants. The errors introduced even by an extremely faithful DNA polymerase that generates 1 erroneous base pair in 10^6 base pair duplications (Keohavong et al., 1989), after 10^6 -fold amplification (20 doublings) of a 200 base pair sequence, would result in product DNA in which a 4×10^{-3} fraction of the single stranded copies contained a mutation, an average of 2×10^{-5} per base pair. This level of background errors could obscure mutations present in human tissues at the level of 10^{-6} .

K. Khrapko, through a collaboration between the W. Thilly laboratory at MIT and B. Karger's laboratory at the Barnett Institute of Northeastern University, adapted the principles of constant denaturant gel electrophoresis (Hovig et al., 1991) to a capillary gel electrophoresis (Cohen et al., 1988) format to create a separation method called "constant denaturant capillary electrophoresis" (CDCE) (Khrapko et al., 1994b). Constant denaturant electrophoresis can be used to separate DNA sequences with a biphasic melting profile that contain mutations in the low melting domain. Separation of mutants is based on the principle that the DNA fragment travels with a slower velocity when the low melting domain is partially melted than when it is closed, presumably due to differences in the cross sectional area of the fragment. At a carefully selected temperature, the molecules of a fragment with a biphasic melting temperature will be in an equilibrium between the partially melted and double stranded conformations. Even single base pair mutations in the low melting domain will affect the equilibrium. Point mutations will cause the DNA fragment to spend more or less time in the double stranded conformation, depending upon whether the mutation makes the low melting domain more or less thermodynamically stable, respectively. Differences in mobility therefore result. High T_m mutants, for instance, AT->GC mutations, exhibit increased mobility while low T_m mutants, GC->AT mutations, have decreased mobility. Constant denaturant gel electrophoresis takes advantage of these principles (Hovig et al., 1991). DNA is electrophoresed through an acrylamide slab gel

bathed in a temperature bath heated to a specific temperature designed to optimize mutant separation.

K. Khrapko and colleagues converted a capillary electrophoresis instrument into a constant denaturant capillary electrophoresis instrument by adding a water jacket to bathe a portion of the capillary at a specific temperature, thus subjecting DNA to denaturing conditions during a portion of the capillary electrophoresis run. Individual point mutants can thereby be separated as homoduplexes and heteroduplexes using this technique.

Separation by CDCE enabled K. Khrapko and colleagues to measure mutant fractions as low as 10^{-6} in model reconstruction experiments with mixtures of PCR fragments (Khrapko et al., 1994b). Experiments by K. Khrapko, X.-C. Li and P. André demonstrated the feasibility of applying the approach to detecting mutations in a mitochondrial target sequence from genomic DNA isolated from cells and tissues. After DNA isolation and restriction digestion to liberate the target fragment, the sample was boiled and reannealed with a fluorescein-labeled probe, followed by successive rounds of CDCE for mutant enrichment and high-fidelity PCR for target sequence amplification. The technique was applied to several tissue samples and TK6 cells. The capillary outputs suggested that a similar pattern of mutations was present in the tissue samples and, at a lower frequency, in TK6 cells. Many of these mutants were then purified and sequenced by K. Khrapko, X.-C. Li and P. André. CDCE enrichment, PCR conditions, and the target sequence used for these experiments, in addition to the sequences of isolated mutants, are described in more detail in the Materials and Methods and Results sections.

2.2 MITOCHONDRIAL DNA AS A TARGET FOR MUTATIONAL SPECTROMETRY

2.2.1 Introduction

Mitochondrial DNA in humans is a 16 kb circular genome that encodes 13 of the approximately 70 protein subunits in the energy production system, 22 tRNAs and 2 rRNAs. The other subunits of the respiratory chain are encoded by the nuclear genome. The respiratory chain catalyzes the phosphorylation of ADP to ATP to store as an energy source. Five protein complexes located in the mitochondrial inner membrane are required. Complexes I-IV take electrons from reduced electron carriers generated by the citric acid cycle, including NADH and FADH₂, and transfers them in a stepwise process to molecular oxygen. These complexes generate an electrochemical gradient by pumping protons across the inner mitochondrial membrane. Energy released by permitting protons to return to the mitochondrial matrix is coupled to the phosphorylation of ADP to ATP by ATP synthase (complex V).

Complex I (NADH dehydrogenase) and complex IV (cytochrome-c-oxidase) contain the largest number of mitochondrial subunits. Seven of approximately 40 subunits of complex I are mtDNA encoded and three of 13 subunits of complex IV are encoded by mtDNA. Complex II (succinate dehydrogenase) is completely nuclearly encoded, while complex III (ubiquinol cytochrome-c-reductase) has only one subunit encoded by mtDNA among its 10 proteins. Two subunits of complex V are also mitochondrially encoded.

2.2.2 Advantages

2.2.2.1 Technical development: high copy number

The goal of developing an unselected approach to mutational spectrometry led to a copy number problem. Given that 100 copies of a particular mutant is desirable for statistical significance, in order to reproducibly detect point mutants present at 10^{-7} , which would be the expected value for a mutant that represented 1% of all 6-thioguanine resistant mutants (Grist et al., 1992; Robinson et al., 1994), 10^9 copies would be required for

analysis. For a single copy gene, this implies that 10^9 cells are required, which closely approximates the total number of bronchial epithelial cells present in the upper airways of the lung. Clearly, an alternative approach was warranted.

The alternative selected was to analyze a mitochondrial DNA target sequence. Mitochondrial DNA was previously shown to be present at a few hundred to thousands of copies per cell (Robin et al., 1988). We have estimated about 1000 copies per cell for TK6 cells (Khrapko et al., 1997), about 400 copies per cell for bronchial epithelial cells, and more than a thousand copies per cell for brain and heart tissue. Use of a mitochondrial DNA target consequently reduces the size of the tissue sample and the amount of total DNA that must be handled to obtain statistically reproducible observations. In the experiments described herein, reproducible detection of rare mutants in mitochondrial DNA was possible from the small number of cells (approximately 10^5 - 10^6) obtainable from brush biopsies.

2.2.2.2 Expected high mutation rate

Prior to these experiments, accurate assessments of the point mutant fractions of mitochondrial DNA in human tissue were not available. There was, however, reason for cautious optimism that the mutant fractions would be higher than observed for nuclear sequences. Mitochondrial sequences have been reported to evolve approximately 10-times faster than nuclear sequences (Brown et al., 1979; Wallace et al., 1987). Further, mitochondrial DNA is selectively adducted by exogenous compounds (Wilkinson et al., 1975; Wunderlich et al., 1970). As will be shown in the Results section, the mutant fractions observed in mitochondrial DNA were indeed higher than the corresponding mutant fractions in nuclear DNA, which was fortuitous in that signals were easily detectable above the background noise generated by PCR, adducts, or other artifacts of the procedure.

2.2.3 Differences from nuclear DNA

While a mitochondrial target sequence is advantageous in that fewer cells are required and, as we later discovered, high mutant fractions are observed, which further decreases the number of cells needed, the choice of a mitochondrial target sequence also poses significant disadvantages. Most importantly, all known oncogenes and tumor suppressors are encoded in nuclear DNA. Since the ultimate goal of these experiments is to determine the most likely mutagenic agents that contribute to cancer, determining the mutational spectra of a tumor suppressor or oncogene, mutation of which represents the initiating event in cancer, would be ideal. Recognizing that the most important source of mutations in mitochondrial DNA is not necessarily the same as the most important mutagen for nuclear sequences, the use of a mitochondrial DNA target makes it inappropriate to draw conclusions about tumor suppressors or oncogenes. Experiments are currently being performed to adapt this procedure to nuclear sequences. Several of the most important differences between mitochondrial DNA and nuclear DNA that could contribute to differences in the most important mutational pathways are discussed below:

2.2.3.1 Intracellular state

Mitochondrial DNA is supercoiled and is not protected by histones as nuclear DNA.

2.2.3.2 Intracellular microenvironment

It has been estimated that 2% of the oxygen consumed by mitochondria is converted to superoxide free radicals ($O_2^{\bullet-}$) as a result of electrons in the oxidative phosphorylation

chain reacting with oxygen (Chance et al., 1979). Superoxide can react directly with DNA, or may be converted into hydrogen peroxide (H_2O_2), which can form highly reactive hydroxyl free radicals ($\text{OH}^\bullet$). These reactive species are generated within the mitochondrial membrane, and thus are physically close to mitochondrial DNA, which is attached to the mitochondrial inner membrane. This could theoretically make mitochondrial DNA more susceptible to adduction by reactive oxygen species, such as superoxide, H_2O_2 and hydroxyl radicals.

In 1988, Richter and colleagues reported levels of 8-hydroxydeoxyguanine (8-OH-dG), an adduct formed by reaction of hydroxyl with deoxyguanine, in rat liver (Richter et al., 1988). 8-OH-dG was reported at a level of 1 per 130,000 bases in nuclear DNA and 1 per 8000 bases in mtDNA based on high performance liquid chromatography elution profiles. These levels are not incompatible with those reported in humans. Estimates of the abundance of 8-OH-dG include an average of 0.03% of total deoxyguanosine content in the temporal cortex of individuals from 42 to 97 years of age (Mecocci et al., 1993), 0.5% in the diaphragm of an 85 year old man (Hayakawa et al., 1991) and 1.5% in a 97 year old heart (Hayakawa et al., 1991).

Experiments by other groups, however, have reported a lower amount of oxidative damage in mitochondrial DNA than reported in these studies. One of these studies, in which the authors were careful to minimize adduction during the isolation procedure, suggested a significantly lower level of 8-OH-dG in mitochondrial DNA in HeLa cells. The ratio of 8-OH-dG to total guanines was approximately 3×10^{-7} (Higuchi et al., 1995). In another study, specific repair endonucleases were used to measure oxidative modifications in mtDNA from rat liver (Hegler et al., 1993). The number of modifications sensitive to formamidopyrimidine-DNA glycosylase, which include 8-OH-dG adducts, was only 8×10^{-6} per bp in mitochondrial DNA.

2.2.3.3 Adduction by exogenous agents

Higher levels of adduction to mitochondrial than to nuclear DNA have been reported for several compounds. Preferential alkylation of mitochondrial DNA after incubation of isolated mitochondria or cell nuclei with N-methyl-N-nitrosourea (NMU) has been reported (Wunderlich et al., 1970). The specific radioactivity of the mitochondrial DNA was 3-7 times that of the nuclear DNA. *In vivo*, treating rats with 50 mg [^{14}C]NMU per kg led to greater incorporation of radioactivity into mitochondrial versus nuclear DNA with ratios ranging from 3 to 17 in different experiments (Murphy et al., 1990). An approximately two-fold increase in alkylation of mitochondrial as opposed to nuclear DNA was discovered for N⁷-methylguanine adducts in rat liver DNA when dimethylnitrosamine was used to treat rats (Wilkinson et al., 1975). In this experiment, increased alkylation levels in mitochondrial versus nuclear DNA were observed at three different dose regimens. For the weak alkylating agent, methyl methane sulphonate, however, mitochondrial and nuclear DNA were alkylated to a similar extent.

Preferential binding to mitochondrial DNA has been observed for bulky adducts as well. Measurement of the amount of [^3H]aflatoxin B₁ bound to nuclear and mitochondrial DNA in treated rats revealed that the levels were three to four times higher in mitochondrial than nuclear DNA (Niranjan et al., 1981). Further, many lipophilic mutagens including polycyclic aromatic hydrocarbons tend to accumulate in mitochondria (Schapira, 1992), possibly as a result of their high lipid content.

2.2.3.4 Repair: Nucleotide excision, base excision, mismatch

The set of enzymes available for repair of mitochondrial DNA is apparently different from that available to nuclear DNA. The lack of nucleotide excision repair for

mitochondrial DNA was first documented in 1974, when Clayton and colleagues reported the absence of pyrimidine dimer repair mechanisms in mammalian mitochondrial DNA based on the ability of the phage T4 UV endonuclease to nick covalently closed mitochondrial DNA containing pyrimidine dimers (Clayton et al., 1974).

By generating single strand breaks at the site of specific DNA lesions followed by Southern blot, the formation and removal of adducts has been monitored (Pettepher et al., 1991). Pyrimidine dimers and complex alkylation damage induced by bifunctional agents were not repaired in mitochondrial DNA, and there was minimal repair of cisplatin intrastrand crosslinks (LeDoux et al., 1992). These findings are in concordance with the absence of nucleotide excision repair. In contrast, there is efficient repair of cisplatin interstrand crosslinks and of N-methylpurines. For both of these types of lesions, about 70% of the adducts were removed by 24 hours. Several other authors have also demonstrated repair in mitochondrial DNA of alkylated purines (Meyers et al., 1988; Satoh et al., 1988) and oxidative damage (Driggers et al., 1993). A process similar to the base excision mechanism for nuclear DNA therefore apparently exists for mitochondrial DNA. O⁶-alkyl guanines, which are repaired by O⁶-methylguanine methyl transferase in nuclear DNA in *E. coli* (Samson et al., 1977) have also been observed to be repaired in the mitochondrial DNA of rat liver and kidney, although nearly no removal was observed from the DNA of brain mitochondria, suggesting tissue-specific repair (Satoh et al., 1988).

Several enzymes involved in repair of mitochondrial DNA damage in mammalian cells have been identified. Among the enzymes that have been reported include uracil DNA glycosylases from rat liver (Domena et al., 1985; Gupta et al., 1981), a mouse plasmacytoma line (Tomkinson et al., 1990) and humans (Nilsen et al., 1997); an endonuclease specific for apurinic and apyrimidinic sites from mouse cells (Tomkinson et al., 1988); a UV endonuclease from a mouse plasmacytoma cell line (Tomkinson et al., 1990); and a repair enzyme for alkali-labile sites from rat insulinoma cells (Pettepher et al., 1991). A 23 kd methyltransferase protein, similar to the one found in the nucleus, has been isolated from rat liver mitochondria (Myers et al., 1988).

Whether human mitochondrial DNA is repaired postreplication by mismatch repair enzymes is unknown as no specific enzymes have been identified. There is evidence for mismatch repair of mitochondrial DNA in yeast. The yeast MSH1 gene is a homologue of mutS, the protein that binds to DNA containing a mismatch during mismatch repair (Su et al., 1986), and is targeted to mitochondria (Reenan et al., 1992b). Mutations in *msh1* result in large-scale rearrangement of mitochondrial DNA (Reenan et al., 1992a). Since the fidelity of DNA synthesis in the nucleus has been reported to decrease 60-fold (Kat et al., 1993), 85-fold (Reenan et al., 1992b), and 340-fold (Fishel et al., 1993) when a mismatch repair protein is mutated, depending upon the species and the specific protein affected, a higher rate of spontaneous mutagenesis could be observed in mitochondrial DNA than in nuclear DNA if human mtDNA lacks this error correction mechanism.

2.2.3.5 Replication

Mitochondrial DNA is replicated by DNA polymerase- γ , a distinct polymerase from the ones that replicate the nuclear genome. The gamma polymerase was recently cloned (Ropp et al., 1996). The gene coding for the enzyme was discovered to contain a CAG trinucleotide repeat encoding polyglutamine in the first exon. This sequence shows limited instability in some mismatch-repair deficient cell lines, varying in size from 12 to 16 repeat units.

This DNA polymerase contains an associated 3'→5' exonuclease activity. In a forward mutation assay with porcine liver DNA polymerase- γ , the enzyme was discovered to have a high fidelity (misincorporation of less than one error for each million detectable nucleotides polymerized) for certain mispairs, similar to that of other exonuclease-containing DNA polymerases (Kunkel et al., 1989). In another study, the average

mutation frequency of chick embryo polymerase- γ was estimated as 46×10^{-4} , while for calf liver polymerase- γ the estimated mutation frequency was 27×10^{-4} (Kunkel, 1985). These frequencies were lower than observed for purified pol- α and pol- β (nuclear polymerases), and pol- γ was notable in the low level of frameshifts produced.

2.2.3.6 Multiple copies per cell

Each cell contains several hundred to thousands of copies of mitochondrial DNA. If one of the mitochondrial DNA copies in a cell were mutated, this would not be expected to produce a phenotypic change in the cell because wild-type mtDNA would compensate for the mutated mtDNA. Thus, for most cases of low-frequency mitochondrial mutants phenotypic changes are not predicted to occur. In one model system, phenotypic change in terms of decreased oxidative phosphorylation capacity was not observed for cells in culture until as much as 90% of the mitochondrial DNA copies contained a mutation in tRNA^{Leu}(UUR) at bp 3243 (Attardi et al., 1995).

The presence of multiple copies raises questions about the possibility that not all mitochondrial DNA copies are replicated. In one study in which this question was specifically addressed, mitochondrial DNA copies were labeled with [³H]thymidine, chased, and then labeled again with 5-bromodeoxyuridine (Bogenhagen et al., 1977). The results were consistent with mitochondrial DNA replication occurring throughout the cell cycle. Further, the copies selected for replication in the second round appeared to be random with respect to those selected in the first round. This argues against both extreme models, on the one hand, that a "queen bee molecule" is replicated repeatedly and on the other hand, that each mitochondrial DNA copy is replicated once and only once for each cell division. Not inconsistent with this result are the findings of Davis and Clayton suggesting that mtDNA replication first occurs in the vicinity of nuclear-provided materials and then distributes throughout a cellular mitochondrial network (Davis et al., 1996). Pertinent to the studies described herein is the possibility that occasional mitochondrial copies, possibly preferentially those containing adducts or premutagenic lesions, could be left unreplicated without serious consequences for the cell.

2.2.4 Mitochondrial DNA and human disease

The mitochondrial encephalomyopathies are a heterogeneous group of disorders in which a defect in mitochondrial metabolism is thought to be the primary cause of disease. These diseases are generally characterized by tissue-specific manifestation of clinical symptoms, including blindness, deafness, dementia, movement disorders, weakness, cardiac failure, diabetes, renal dysfunction and liver disease (Lombes et al., 1989; Peterson et al., 1991; Zeviani et al., 1989a). Mitochondrial disorders can be divided into two broad classes based on the type of mutation: large-scale rearrangements (deletions) and point mutations.

2.2.4.1 Deletions

Deletions of mitochondrial DNA have been frequently detected in patients with mitochondrial myopathies (Holt et al., 1989a; Holt et al., 1988; Lestienne et al., 1988; Moraes et al., 1989; Zeviani et al., 1988), especially in cases where ocular myopathy is a key feature, for example, Kearns-Sayre syndrome and progressive external ophthalmoplegia. In patients, a large fraction of the total population of skeletal muscle mtDNA molecules (up to 90%) contains large-scale deletions (Holt et al., 1989a; Holt et al., 1988; Lestienne et al., 1988; Moraes et al., 1989; Zeviani et al., 1988). A characteristic feature of these diseases is the presence of large numbers of structurally

abnormal mitochondria in muscle fiber segments referred to as "ragged red" based on their appearance with Gomori trichrome stain (Byrne et al., 1985; Johnson et al., 1983). Mitochondrial diseases involving large scale deletions are nearly always sporadic, suggesting that the mtDNA deletions arise as a clonal event in the egg or early in development.

In one report of a patient with Kearns-Sayre syndrome, for instance, the patient's symptoms included bilateral drooping eyelids, weakness of the muscles that move the eyes, decreased hearing progressing to deafness, generalized muscle weakness, and defects in atrioventricular signaling in the heart (Bresolin et al., 1987). Although cytochrome-c-oxidase was present, enzyme activity declined as the disease progressed. In a first muscle biopsy at age 31, 3% of muscle fibers appeared as ragged red. In a second muscle biopsy seven years later, 20% of muscles stained as ragged red and 60% were completely absent for cytochrome-c-oxidase activity.

Reported deletions in patients with mitochondrial myopathies range in size from 1.3-7.6 kb (Holt et al., 1989a; Moraes et al., 1989). Although the particular species of deleted mtDNA varies among patients, approximately one-third of all patients harbor the same deletion, called the "common deletion" (Mita et al., 1990; Schon et al., 1989; Shoffner et al., 1989). This deletion eliminates 4,977 bases of mtDNA flanked by a 13 bp direct repeat.

The frequency of the most prominent deletion varies among the tissues of mitochondrial myopathy patients. In one Kearns-Sayre patient, the "common deletion" represented 60% of the copies in skeletal muscle, 15% in heart muscle and kidney, and less than 5% in the liver (Obermaier-Kusser et al., 1990).

Autosomal dominant inheritance of a disease characterized by multiple deletions of mtDNA has also been reported (Suomalainen et al., 1992; Zeviani et al., 1989a). In this case, the cause must be a defective nuclear gene coding, for example, for an enzyme involved in maintaining the mitochondrial genome.

2.2.4.2 Point mutations

More than 40 pathogenetic point mutations in mitochondrial DNA have been identified to date, including mutations associated with MELAS (mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes), MERRF (myoclonus epilepsy with ragged-red fibers), NARP (neuropathy, ataxia, and retinitis pigmentosa), and LHON (Leber's hereditary optic neuropathy). In all of these disorders, the quantity of mutated mtDNA molecules is high in selected tissues (usually between 5 and 100% of total mtDNA) and can be detected easily by restriction fragment length polymorphism analysis. Mitochondrial mutations that have been identified in mitochondrial diseases are summarized in Figure 5.

MERRF was among the first diseases clearly demonstrated to originate with a mitochondrial DNA mutation. Wallace and colleagues demonstrated that a mtDNA mutation must be pathogenetic by proving maternal inheritance in a large family, and by identifying specific deficiencies in mitochondrial respiratory complexes I and IV (Wallace et al., 1988). Respiratory capacity of different patients was correlated with the nature and severity of the patients' clinical manifestations. Tissues were affected in the order of the extent to which they rely on mitochondrial energy production (first central nervous system, then type I oxidative skeletal muscle fibers, then heart). The causative mutation was subsequently identified as an AT->GC transition at nucleotide pair 8344 which alters the T Ψ C loop of the tRNA^{Lys} gene (Shoffner et al., 1990). All MERRF patients and their less-affected maternal relatives had between 83% and 98% mutated mtDNA and showed an age-related association between genotype and phenotype.

Leber's hereditary optic neuropathy (LHON) is another disease caused by mtDNA mutations. It is characterized by bilateral loss of central vision. At least five mtDNA point

Fig. 5: Pathogenic point mutations in mtDNA.

Point mutations of mitochondrial DNA which have been related to human mitochondrial diseases. *This mutation might be a Japanese polymorphism (Goto et al., 1994). ADPD: Alzheimer's disease and Parkinson's disease; CIPO: chronic intestinal pseudo-obstruction with myopathy and ophthalmoplegia; CPEO: chronic progressive external ophthalmoplegia; DEAF: maternally inherited sensorineural deafness; DMDF: diabetes mellitus and deafness; FICM: fatal infantile cardiomyopathy minus; FICP: fatal infantile cardiomyopathy plus; HCM: hypertrophic cardiomyopathy and myopathy; Leigh: Leigh's subacute necrotizing encephalomyelopathy; L IMM: lethal infantile mitochondrial myopathy; LHON: Leber's hereditary optic neuropathy; MELAS: mitochondrial encephalomyopathy, lactic acidosis and stroke-like symptoms; MERRF: myoclonic epilepsy and ragged-red fiber disease; MILS: maternally inherited Leigh's syndrome; MM, mitochondrial myopathy; NARP: neuropathy, ataxia and retinitis pigmentosa; PEO: progressive external ophthalmoplegia.

Source: Kadenbach et al., 1995 (see Kadenbach et al., 1995 for references)

Nucleotide	Mutation	Gene location	Disease	Reference
721	TA->CG	12S rRNA	ADPD	Brown et al., 1994
1555	AT->GC	12S rRNA	DEAF	Prezant et al., 1993
3196	GC->AT	16S rRNA	ADPD	Shoffner et al., 1993
3243	AT->GC	tRNA ^{Leu} UUR	MELAS/PEO	Goto et al., 1990
3243	AT->GC	tRNA ^{Leu} UUR	DMDF	van den Ouweland et al., 1992
3250	TA->CG	tRNA ^{Leu} UUR	MM	Goto et al, 1992
3251	AT->GC	tRNA ^{Leu} UUR	MELAS	Sweeney et al., 1993
3252	AT->GC	tRNA ^{Leu} UUR	MELAS	Morten et al., 1993
3253	CG->TA	tRNA ^{Leu} UUR	FCM	Ozawa et al., 1991
3256	CG->TA	tRNA ^{Leu} UUR	Multisystem/PEO	Moraes et al, 1993
3260	AT->GC	tRNA ^{Leu} UUR	HCM	Zeviani et al, 1991
3271	TA->CG	tRNA ^{Leu} UUR	MELAS	Goto et al., 1991
3291	TA->CG	tRNA ^{Leu} UUR	MELAS	Goto et al., 1991
3302	AT->GC	tRNA ^{Leu} UUR	Myopathy	Bindoff et al., 1993
3303	CG->GC	tRNA ^{Leu} UUR	Cardiomyopathy	Silvestri et al., 1993
3304	TA->CG	ND1	LHON	Brown et al., 1992
3394	TA->CG	ND1	LHON	Huoponen et al., 1991
3397	AT->GC	ND1	ADPD	Schoffner et al., 1993
3460	GC->AT	ND1	LHON	Huoponen et al., 1991
4136	AT->GC	ND1	LHON	Howell et al., 1991
4160	TA->CG	ND1	LHON	Howell et al., 1991
4216	TA->CG	ND1	LHON	Mackey and Howell, 1992
4269	AT->GC	tRNA-Ile	FCP	Taniike et al., 1992
4317	AT->GC	tRNA-Ile	FCP	Tanaka et al., 1990
4336	AT->GC	tRNA-Gln	ADPD	Shoffner et al., 1993
4917	AT->GC	ND2	LHON	Johns and Berman, 1991
5244	GC->AT	ND2	LHON	Brown et al., 1992
5554	CG->TA	tRNA-Trp	FCM	Ozawa et al, 1991
5601	GC->AT	tRNA-Ala	MELAS	Tanaka et al, 1991
5692	AT->GC	tRNA-Asn	CPEO	Seibel et al., 1994
5703	GC->AT	tRNA-Asn	Myopathy/PEO	Moraes et al., 1993
7444	GC->AT	COX I	LHON	Brown et al., 1992
8344	GC->AT	tRNA-Lys	MERRF	Shoffner et al, 1990
8356	TA->CG	tRNA-Lys	MERRF	Silvestri et al., 1992
8993	TA->GC	ATP 6	Leigh	Shoffner et al., 1992
8993	TA->GC	ATP 6	NARP	Holt et al., 1990
8993	TA->CG	ATP 6	Leigh	de Vries et al., 1993
9438	GC->AT	COI	LHON	Johns and Neufeld, 1993
9804	GC->AT	COI	LHON	Johns and Neufeld, 1993
9997	AT->GC	tRNA-Gly	HCM	Merante et al., 1993
10006	AT->GC	tRNA-Gly	CPO	Lauber et al., 1991
11084*	AT->GC	ND4	MELAS	Lertrit et al., 1992
11778	GC->AT	ND4	LHON	Wallace et al., 1988
12246	AT->GC	tRNA-Ser(GCU)	CPO	Lauber et al., 1991
13708	GC->AT	ND5	LHON	Johns and Berman, 1991
13730	GC->AT	ND5	LHON	Howell et al., 1993
14459	GC->AT	ND6	LHON	Wallace et al., 1993
14484	TA->CG	ND6	LHON	Johns et al., 1992
15257	GC->AT	Cyt b	LHON	Johns and Neufeld, 1991
15812	GC->AT	Cyt b	LHON	Johns and Neufeld, 1991
15923	AT->GC	tRNA-Thr	LIMM	Yoon et al., 1991
15924	AT->GC	tRNA-Thr	LIMM	Yoon et al., 1991
15990	CG->TA	tRNA-Pro	MM	Moraes et al., 1993

mutations, all of which alter evolutionarily conserved amino acids in the subunits of the electron-transport chain, appear sufficient in themselves to cause the disease. The most prevalent mutation, accounting for over 50% of cases, is a GC->AT transition at mtDNA bp 11778 located in subunit IV of the NADH dehydrogenase complex (I). The mutation converts a conserved arginine to a histidine (Poulton et al., 1991; Vilkki et al., 1989; Wallace et al., 1988). LHON can also result from another class of base substitutions that are not sufficient to induce the disease in themselves, but act synergistically to increase the risk for optic atrophy. These secondary polymorphisms are also found in control individuals, though at a much lower frequency, and they alter less conserved amino acids (Savontaus, 1995).

As opposed to almost all other mitochondrial diseases which are characterized by heteroplasmy, i.e., mutant molecules coexisting with wild-type mtDNA, LHON is an example of a mitochondrial disorder which usually harbors primary mutations usually in a homoplasmic state (Stone et al., 1992; Wallace et al., 1988), although persistent low-level heteroplasmy accounts for approximately 14% of cases (Holt et al., 1989b; Sweeney et al., 1992; Newman et al., 1991).

Another example of a disease that can result from a point mutation in mitochondrial DNA is diabetes. Diabetes-like symptoms have been associated with mitochondrial DNA deletions, duplications, and mutations in the tRNA^{Leu}(UUR) gene and missense mutations in the NADH dehydrogenase subunit genes (Gerbitz et al., 1995). The most common is an AT->GC transition mutation in the tRNA^{Leu}(UUR) gene at bp 3243 which, in homoplasmic form, is associated with diabetes in about 1.5% of the diabetic population in different countries and races (Kadowaki et al., 1994; Reardon et al., 1992).

Tissue specificity in the proportion of mtDNA containing a mutant has been observed for point mutations as well as deletions. In a case report of a 34 year old man with recurrent strokes, Chinnery and colleagues discovered that the proportion of an AT->GC transition at bp 3243, which has been associated with MELAS, was 85% in the individual's skeletal muscle but only 4% in the blood (Chinnery et al., 1997).

2.2.5 Low-frequency somatic mitochondrial DNA alterations

The accumulation of point mutations and deletions in mitochondrial DNA has been suggested as an important contributor to human diseases, including cancer (Shay et al., 1987; Warburg, 1956; Yamamoto et al., 1992) and aging (Bandy et al., 1990; Corral-Debrinski et al., 1992; Cortopassi et al., 1990; Münscher et al., 1993). Central to the finding of deletions in mitochondrial DNA of normal tissue has been the observation that deletion frequency increases with age in a tissue-specific manner.

2.2.5.1 Deletions

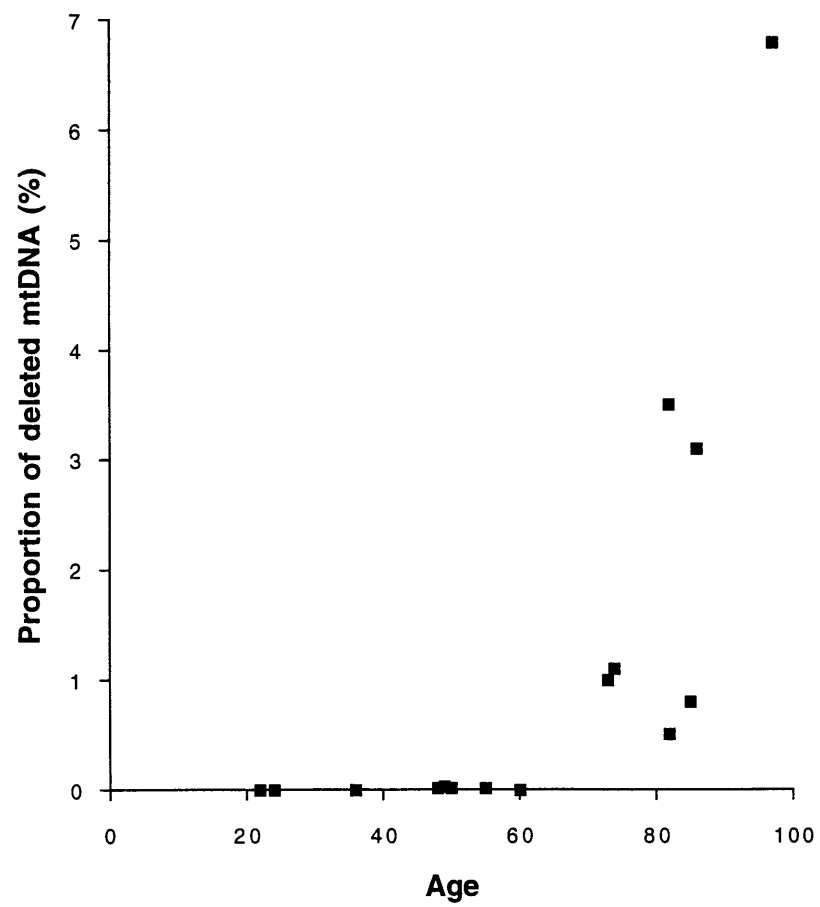
Large deletions of mitochondrial DNA have been found to increase with donor age in multiple human tissues (Cortopassi et al., 1990; Linnane et al., 1990; Simonetti et al., 1992; Yen et al., 1991). While most studies have focused on the "common deletion," the 5 kb deletion discovered frequently in patients with mitochondrial diseases, several other deletions have also been shown to increase with age (Katayama et al., 1991; Yen et al., 1992). In one study, a 7.4 kb deletion was observed in myocardial mitochondrial DNA from all subjects over 70 years of age (Sugiyama et al., 1991). The proportion of mtDNA copies containing this deletion increased supralinearly with age. The mtDNA deletion was present at 0.025% in a 36 year old; at 1.4% in a 85 year old; and at 6.7% in a 97 year old. An example of the increase in the frequency of deleted mitochondrial DNA copies with age observed in one study is shown in Figure 6.

Several studies have identified multiple mtDNA deletions with advancing age. Zhang and colleagues detected 10 different deletions in the heart, brain and skeletal muscle

Fig. 6: Increase in the ratio of deleted to total mtDNA with age.

The ratio of deleted mtDNA to the total mtDNA plotted against the age of subjects. The amounts of the total and deleted mtDNA in autopsied myocardium of various ages were determined by kinetic PCR. The ratio increases exponentially with age.

Source: Ozawa, 1995



of a 69 year old female (Zhang et al., 1992). Ozawa and colleagues developed a method based on PCR with 180 different primer pairs that permits detection of all deletions in mtDNA (Katsumata et al., 1994). In the heart mtDNA of a 28 year old normal female, 67 different types of mitochondrial DNA minicircles were observed and they are reported to represent 23% of mitochondrial DNA copies (Ozawa, 1995). Heart mtDNA of a patient with a mutation in the tRNA^{Asp} gene who died of heart failure at age 19 was also investigated. In this patient, 235 different types of deletions were detected and mtDNA minicircles represented 84% of the mtDNA copies.

Multiple deletions that increase with age have been detected in other species, including rhesus monkeys (Lee et al., 1993), mice (Wang et al., 1997) and rats (Van Tuyle et al., 1996).

2.2.5.2 Tissue specificity of mtDNA alterations

Large deletions of mitochondrial DNA have been discovered to show tissue-specific differences in frequency (Simonetti et al., 1992; Yen et al., 1991). In one report, the fraction of mitochondrial DNA that was deleted was 5.5×10^{-4} in muscle, 1×10^{-4} in heart, 0.5×10^{-4} in spleen, and much lower levels in liver, kidney and lung in an 84 year old man (Simonetti et al., 1992). Indeed, deletions are generally rare in blood cells (Yen et al., 1991) and splenic lymphocytes (Cortopassi et al., 1992). Even within an organ, tissue-specificity has been observed in the levels of the "common deletion:" among the different regions of brain, striatum has been found to contain an approximately 200-fold higher level of deleted mitochondrial DNA than cerebellum (Corral-Debrinski et al., 1992; Soong et al., 1992). This has been suggested to be related to the higher density of dopaminergic neurons because dopaminergic neurotransmitters are degraded by monoamine oxidase within the outer mitochondrial membrane, a process that is accompanied by hydrogen peroxide production (Corral-Debrinski et al., 1992; Cortopassi et al., 1992).

Tissue-specific selection for different mtDNA haplotypes has been observed in heteroplasmic mice containing two different mtDNA genotypes (BALB and NZB). The haplotypes contain 106 point mutational differences and are assumed to be neutral to selection (Jenuth et al., 1997). In these mice, the mean frequency of the BALB haplotype in a female was the same as the mean frequency in all her offspring, and further the relative proportion of the BALB haplotype was very similar among tissues in the neonates. With time, however, the frequencies of the BALB haplotype in different tissues changed significantly. Without exception (n=37 mice), the proportion of the BALB mtDNA genotype increased in blood and spleen while that of the NZB genotype increased in liver and kidney. In the liver, the NZB genotype was estimated to have a 50% selective advantage over the BALB genotype in each cell generation (see Figure 7).

2.2.5.3 Point mutations

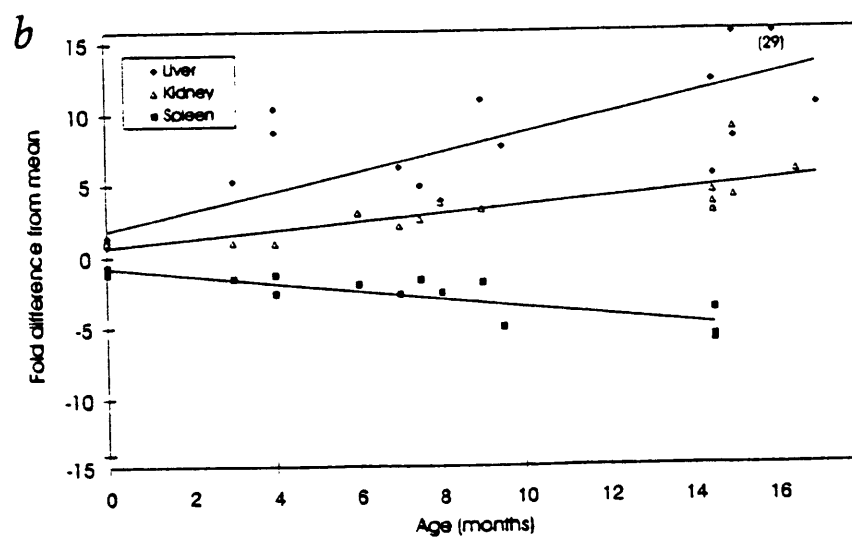
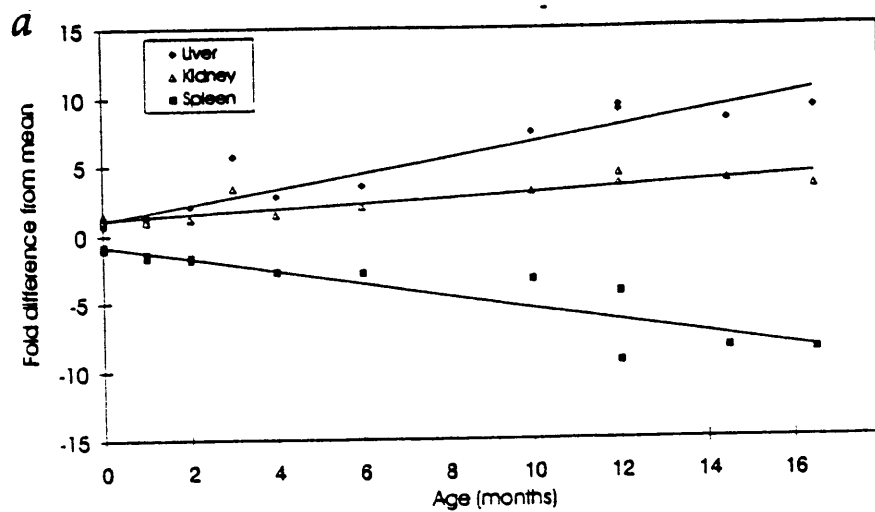
Low-frequency mitochondrial DNA point mutations associated with pathologic conditions have also been found in several undiseased organs. Using PCR/RFLP analysis with mispairing primers, Münscher and colleagues found an AT->GC transition at mtDNA bp 8344 [a mutation associated with MERRF (Shoffner et al., 1990)] in extraocular muscle in 11 of 14 normal older adults, and calculated that a maximum of 2.4% of this specific mutation was present in the oldest samples (Münscher et al., 1993).

A mutation at bp 10,006 in tRNA^{Gly}(TCC) was observed in all patients over 67 years old but not in patients 30 and under (Münscher et al., 1993). This age-related difference was not discovered for mutations at either bp 5692 (in tRNA^{Asn}) or bp 12246 [in tRNA^{Ser}(AGY)]. Mutations at bp 10,006, 8344 and 3243 were observed in most elderly individuals while mutations at 5692 and 12,246 occurred only occasionally.

Fig. 7: Tissue specific selection for mitochondrial haplotypes.

Selection for alternative mtDNA genotypes in spleen vs. liver and kidney of heteroplasmic animals with increasing age. Each point corresponds to the value obtained in an individual animal. Panels A and B refer to mice with BALB (A) and mixed BALB/NZB (B) nuclear backgrounds

Source: Jenuth et al., 1997.



Münscher and colleagues concluded that certain nucleotides of mtDNA represent “hot spots” for somatic point mutations and suggested that these may contribute to aging.

Kadenbach and colleagues performed nonquantitative allele-specific PCR for five point mutations in tRNA sequences on samples from individuals of various ages (Kadenbach et al., 1995). Mutations were below detectable levels in all individuals for four of the five point mutations. A mutation at bp 12,220 was detected in two of five elderly (>80 years old) individuals and not in two younger individuals. The same authors also performed quantitative PCR for the mutations associated with MERRF (bp 8344) and MELAS (bp 3243). Levels were not detectable for 20 year olds for either mutation. Among subjects >60 years old, the MERRF mutation was observed at frequencies up to 2% of all wild-type copies and the MELAS mutation was observed at frequencies up to 1%. The authors' claims of an "exponential" increase with time for these two mutations are not supported by the data.

In two other studies of an AT->GC transition at bp 3243 associated with MELAS, the reported frequencies were 0.1% using allele-specific PCR (Zhang et al., 1993), and from an undetectable level to 0.7% using an approach based on restriction digestion of the wild-type (Pallotti et al., 1996). Zhang and colleagues reported that this mutation was present at higher levels in adults than infants. While seven out of 38 adult heart tissues tested were found to contain this mutation at an estimated 0.1%, the mutation was not detected in any of the 16 infant tissues analyzed. Pallotti and colleagues used restriction digestion to enrich for mutant copies, followed by PCR to detect mitochondrial mutants, in the 5 bp region between bp 8991 and 8995, and in the 4 bp region between bps 3242 and 3246 (Pallotti et al., 1996). The restriction digestion/PCR approach was used to detect spontaneous point mutations in DNA isolated from skeletal muscle from normal individuals of various ages. Low levels of somatic mutations were observed above background in both regions, but there was no correlation between the amount of mutation detected and the age of the subject. These results contrasted with the age-related increase in the amount of the mtDNA "common deletion" in these samples. They therefore concluded that there is no preferential amplification or accumulation of specific point mutations in skeletal muscle over the course of a human lifespan.

Using an approach based on high efficiency restriction digestion followed by denaturing gradient gel electrophoresis (Fischer et al., 1983) to separate mutants from wild-type DNA and from each other, hotspot mutations were detected in the 16S rRNA region of mitochondrial DNA. In this assay, the DNA is restriction digested to eliminate approximately 99% of wild-type sequences. The sample is amplified, restriction digested again, reamplified and analyzed using denaturing electrophoresis. Application of this method to two lung samples and a colon sample revealed a common GC->AT hotspot in the restriction enzyme recognition sites of the 16S rRNA region (G. Hu, H. C., X.-C. Li and O. Kostas, unpublished results). This hotspot was estimated to have a frequency of 10^{-6} to 10^{-4} .

Attempts have also been made to characterize mitochondrial point mutations by direct sequencing. Gadaleta and colleagues sequenced 16,000 bases from two different regions of the genome and did not find any sequence differences between adult and senescent rat brain mtDNA (Gadaleta et al., 1992). Similarly Bodenteich and colleagues detected only one heteroplasmic mutation out of 32,000 bases sequenced from two different regions of mtDNA from the retina of a 71 year old individual (Bodenteich et al., 1991).

2.2.6 Distribution of mitochondrial mutants

The distribution of mitochondrial mutants among cells has been investigated both by histochemical staining and by *in situ* hybridization. Histochemistry has generally involved staining for lack of cytochrome-c-oxidase activity. Cytochrome-c-oxidase is the

terminal enzyme in the respiratory chain that irreversibly transfers electrons of the respiratory chain to molecular oxygen. Lack of cytochrome-c-oxidase activity implies that no oxidative phosphorylation-based ATP synthesis is occurring in mitochondria.

Human aging has been associated with focal accumulation of respiratory deficient cells in various tissues like heart (Müller-Höcker, 1989), skeletal muscle (Müller-Höcker, 1990), extraocular muscle (Müller-Höcker et al., 1992) and diaphragm (Müller-Höcker, 1990). In one study, cytochrome-c-oxidase staining was studied in 140 hearts from men obtained at autopsy (Müller-Höcker, 1989). Randomly distributed individual cardiomyocytes lacking enzyme activity were observed independent of heart disease. The proportion of hearts with cytochrome-c-oxidase deficient cardiomyocytes increased with age from 0/7 hearts aged 0-7 years to 21/21 hearts aged 81-97. The density of the defects also increased with age from 2 to 3 defects per cm^2 in the second and third decade to about 50 defects per cm^2 in advanced age. In most cases there was a complete loss of activity and the defect always involved all the mitochondria of a cell.

In situ hybridization with oligonucleotides has also been used to determine the distribution of mitochondrial mutants among cells. Mita and colleagues performed a semiquantitative *in situ* hybridization analysis and found that mutant and wild-type mtDNAs were inversely distributed in the muscle fibers of a Kearns-Sayre patient (Mita et al., 1989). High levels of deleted mtDNAs were found in oxidatively defective fiber segments, and fiber segments with high levels of wild-type mtDNA had normal cytochrome oxidase activity, suggesting that deleted and wild-type mtDNAs are functionally independent.

Shoubridge and colleagues also used *in situ* hybridization to determine the distribution of deleted mitochondrial DNA copies in muscle fiber segments (Shoubridge et al., 1990). In four patients with mitochondrial myopathies, they discovered that individual ragged red fiber segments contained an unusually large number of total mitochondrial DNA copies. Most of these were deleted mitochondrial copies (6- to 7- times normal); in addition, the normal number of wild-type copies were also present. These fibers were negative for cytochrome-c-oxidase activity. Other unaffected muscle fibers contained only the normal amount of wild-type copies. This suggested that loss of cytochrome-c-oxidase function was a result not of lack of wild-type mtDNA but of interference from deleted mtDNAs.

Fiber bundle analysis has also been employed to investigate the distribution of mitochondrial mutants. Defined numbers of fibers were dissected from a large muscle of the rhesus monkey (Schwarze et al., 1995). The number of different deletions observed declined when the number of fibers examined decreased from 75 to 10. The frequency of specific deletions from the same samples, however, increased in abundance from approximately 1% to between 5 and 13% of total mitochondrial genomes when the cell number decreased. These findings are in concordance with evidence above that mitochondrial deletions accumulate within cells.

2.2.7 Mitochondrial DNA and aging

As described previously, the levels of deleted mtDNA increase significantly with age in humans and other species. This observation has suggested that mitochondrial DNA mutations may be involved in aging. Evidence for this assertion is described below. Muscle, heart and brain function, as well as electron transport activity decline with age, while the number of cytochrome-c-oxidase deficient fibers and oxidative free radicals increases. Further, caloric restriction, the only mechanism that has been proven to increase longevity in many species, also results in lower levels of mtDNA alterations. Finally, it is interesting to note that metabolic rate and longevity are inversely correlated among species.

2.2.7.1 Organ dysfunction

Aging is associated with loss of muscle mass and strength, and loss of neurons. Such changes could be associated with mitochondrial DNA alterations either because cells with mitochondrial alterations can no longer function due to a decline in oxidative phosphorylation ability, or because mitochondrial defects promote a sequence of events resulting in neuronal apoptosis.

A decline in mitochondrial activity with aging has been documented. In addition to histological studies described previously that demonstrate an increase in cytochrome-c-oxidase deficient cells with age, biochemical analyses have also been performed to determine the function of the electron transport complexes of humans, rhesus monkeys, rats, and mice (Boffoli et al., 1994; Bowling et al., 1993; Desai et al., 1996; Sugiyama et al., 1993). Age-associated decreases in the activities of complexes I and IV have been observed in nonproliferating tissues such as skeletal muscle, heart, and brain, but usually not in proliferating organs like liver or kidney. In one study, the activities of complexes I and IV of heart mitochondria of rats aged 100 weeks decreased significantly, by 31% and 22%, respectively, compared with those of rats aged 7 weeks (Takasawa et al., 1993). No significant decline with age was observed in complexes II and III (Sugiyama et al., 1993). Since complexes I and IV contain the most mitochondrial DNA encoded subunits, this finding supports the involvement of mitochondrial DNA alterations in the decline of oxidative phosphorylation.

Mitochondrial dysfunction has also been associated with apoptosis (Cortopassi et al., 1995; Polyak, 1997; Sheehan, 1997). Death of dopaminergic neurons (Oertel et al., 1993) and sensorineural hair cells of the inner ear are characteristic features of aging. Further, the dementia associated with Alzheimer's disease results from apoptosis of cholinergic neurons, and late-onset diabetes may reflect apoptosis of pancreatic islet cells. These correlations raise the possibility that mitochondrial DNA alterations may play a role in the aging process through this mechanism as well. Indeed, deafness, dementia, ataxia and diabetes are symptoms of mitochondrially encoded diseases.

2.2.7.2 Oxidative damage

Levels of the adducted base 8-hydroxydeoxyguanine (8-OH-dG), which results from interaction of free radicals with deoxyguanine, have been extensively examined because the compound can be detected with high performance liquid chromatography. Several laboratories have reported that this particular adduct increases with age in several species (Agarwal et al., 1994; Hayakawa et al., 1991; Sohal et al., 1994). Mecocci and colleagues, for instance, determined the levels of 8-OH-dG in DNA isolated from 3 brain regions from 10 normal humans aged 42 to 97 years, and discovered that the levels of adducted nucleotide increased with age (Mecocci et al., 1993). Such findings have also been extended to rodents. The levels of 8-OH-dG increased by 130% in rats aged 100 weeks compared with levels in rats aged 7 weeks in one study (Takasawa et al., 1993).

In addition, overexpression of two major antioxidant enzymes, dismutase and catalase, has been shown to bestow longevity upon *Drosophila melanogaster* (Orr et al., 1994).

2.2.7.3 Caloric restriction

Another pertinent observation is that caloric-restricted animals age less. Caloric restriction extends the lifespan of rats by 33%, from less than three to four years of age. The long-lived rats stay youthful longer and suffer fewer late-life diseases than their normally-fed counterparts (Weindruch, 1996). Caloric restriction increases the maximum (as opposed to the average) lifespan, thereby demonstrating that the effect is not simply due

to prevention of premature death but actually a slowing of the rate of aging. Increased lifespans from dietary restriction have been observed in protozoa (Rudzinska, 1952), guppies (Comfort, 1963), and rodents (Sohal et al., 1996). Further, dietary-restricted mice show less decline in learning and motor performance with age (Ingram et al., 1987).

The explanation for the health benefits resulting from caloric-restricted diets that has the most experimental support is that caloric restriction limits damage to mitochondria by free radicals (Weindruch, 1996). Mitochondria harvested from the brain, heart and kidney of dietary-restricted mice reveal that levels of the superoxide radical and of hydrogen peroxide were markedly lower in animals subjected to long-term caloric restriction than in normally-fed controls (Sohal et al., 1996). In a recent study, muscle fibers from aged caloric-restricted rats were shown to contain significantly fewer cells deficient in cytochrome-c-oxidase activity, fewer cells that stained as ragged red, and a significantly lower number of mtDNA deletion products in two (adductor longus and soleus) of the four muscles examined (Aspnos et al., 1997).

2.2.7.4 Metabolic rate and lifespan

One possible explanation for the beneficial effects of caloric restriction is that such treatment reduces resting energy expenditure (Ramsey et al., 1997). Supporting this hypothesis is the finding that metabolic rates and lifespans correlate. When the metabolic rate, defined by ml of O₂ consumed per kg per hour, is compared with the generation time (in days) for 8 species of primates, the correlation coefficient of -0.75 supports a purported association between the factors (Martin et al., 1993). Further, rates of mitochondrial oxygen consumption and O₂^{•-} and H₂O₂ production in the kidney and heart, were shown to be inversely correlated with maximum lifespan in seven mammalian species: mouse, hamster, rat, guinea pig, rabbit, pig, and cow (Ku et al., 1993). This correlation suggests that accumulated damage to mitochondrial DNA may represent the limiting factor in generation time. Also supporting this hypothesis is the observation that mice accumulate significantly higher levels of mitochondrial deletions than humans. Using nested PCR, the absolute frequency and rate of increase of mitochondrial deletion mutations was determined (Wang et al., 1997). Deletion frequencies rose by 2.5 x 10⁵, 6300- and 4000-fold in heart, brain, and kidney, respectively, from young to old mice. The deletion frequencies rose 40-fold faster in mouse than human mtDNA per unit time, a difference that is similar to the difference in lifespan, supporting the hypothesis that death from old age occurs as a result of an accumulation of deleterious mitochondrial mutations.

2.2.8 Target sequence and melting map

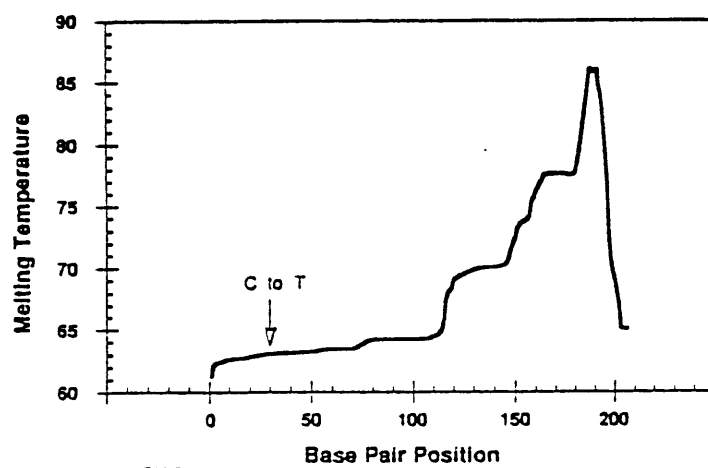
The previous sections have described the recent explosion of information about the role of mutations in mitochondrial DNA in acute diseases and possibly also in aging. For these studies, a section of the mitochondrial genome was selected for analysis based on its adaptability to separation by constant denaturing electrophoresis. The target sequence of mitochondrial DNA used for this study was a 100 bp low melting domain of a 200 base pair region of mitochondrial DNA bps 10011 to 10217. This sequence was selected because of its suitability for constant denaturant electrophoresis, which can separate DNA containing mutations in the low melting domain of a sequence with a biphasic melting profile. As can be observed from Figure 8, the selected target sequence contains a 100 bp low melting domain that will be in a rapid equilibrium between open and closed conformations while the high melting domain remains double stranded and serves as a "clamp."

The target sequence codes for two genes: tRNA^{Gly}(TCC) and NADH dehydrogenase subunit III. The proposed clover leaf structure of mitochondrial tRNA^{Gly}(TCC) is shown in Figure 9. Mutations present within the tRNA^{Gly}(TCC) but

Fig. 8: Melting map of the mitochondrial target sequence used in these studies.

Melting map of the wild-type DNA fragment calculated according to the algorithm of Lerman and Silverstein, 1987. CW7, J3 and CW7(mut) indicate the positions of the primers used for PCR and the primer used to generate the artificial mutant used as a low T_m internal standard (CW7(mut)). The vertical arrow indicates the position of the artificial mutant mutation. Base pair one on the map corresponds to bp 10,011 of the human mitochondrial genome.

Source: Khrapko et al., 1994.



CW7 →
 CW7(mut) →
 T

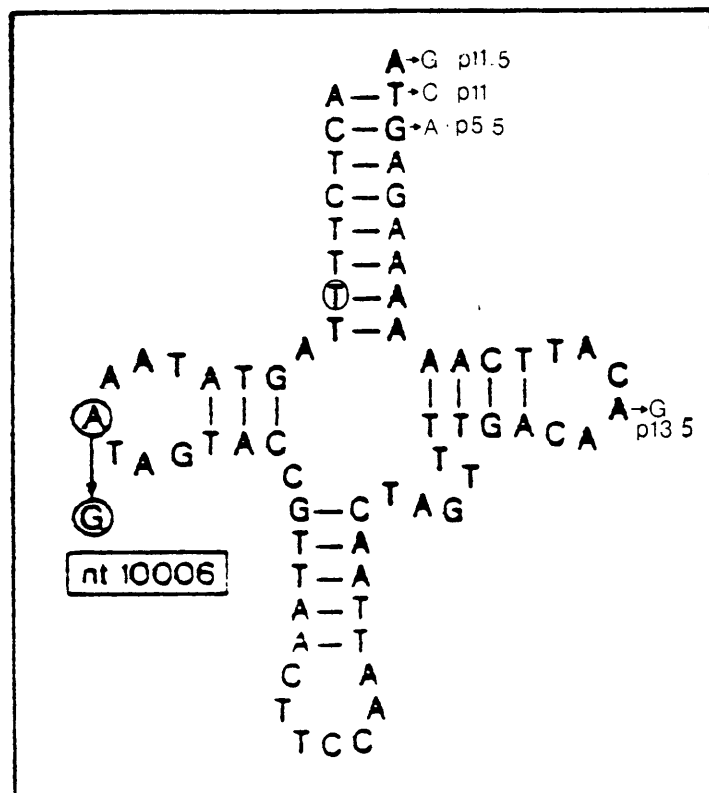
← J3°

• 5' - label (^{32}P or fluorescein)

Fig. 9: Clover leaf of tRNA^{gly} in mtDNA.

Proposed clover leaf structures of the mitochondrial tRNA gene for glycine (TCC). The substitution at bp 10006, together with two other mutations was mutated in a patient with chronic intestinal pseudo-obstruction with myopathy and ophthalmoplegia. The circled nucleotide at bp 9997 was found mutated in a patient with cardiomyopathy. Also indicated are the positions of mutations observed as hotspot point mutations in human tissue samples. The number designation for each are as follows: p13.5, AT->GC at bp 10042; p5.5, GC->AT at bp 10056; p11, TA->CG at bp 10057; p11.5 AT->GC at bp 10058.

Source: Münscher et al., 1993.



tRNA^{Gly(TCC)}

outside of our target sequence have been associated with chronic intestinal pseudo-obstruction with myopathy and ophthalmoplegia [bp 10006, (Lauber et al., 1991)] and cardiomyopathy [bp 9997, (Merante et al., 1994)].

The target sequence also includes a portion of the NADH dehydrogenase subunit III, which encodes one of 25 polypeptides that compose complex I of the respiratory chain. This enzyme complex contains a flavin mononucleotide as a tightly bound prosthetic group. The flavin oxidizes reduced NADH, and then the complex transfers electrons from reduced flavin to iron-sulfur centers and ultimately to coenzyme Q. Complex III and IV then carry electrons from coenzyme Q to cytochrome c and ultimately to oxygen to generate water.

The drug 1-methyl-4-phenyl-1,2,5,6-tetrahydropyridine (MPTP) is a specific inhibitor of complex I (Nicklas et al., 1985) and provides some information about the possible effects of mutations in genes encoding complex I subunits. MPTP administration to animals can simulate the symptoms of Parkinson's disease (Oertel et al., 1993; Schapira, 1992), a common, late-onset neurodegenerative disorder that results in part from the gradual loss of dopaminergic neurons in the substantia nigra region of the brain. MPTP may induce Parkinson's-like symptoms by causing apoptosis of substantia nigra neurons (Sheehan et al., 1997). That MPTP-induced apoptosis is mediated through superoxide is suggested by findings that mice with deficiencies in the mitochondrial superoxide dismutase, MnSOD, are hypersensitive to MPTP (Cortopassi et al., 1995). MnSOD-deficient mice exhibit massive brain toxicity when treated with MPTP.

2.3 BRONCHIAL EPITHELIAL CELLS AND TISSUE KINETICS

The bronchial epithelial cells of the airways are organized as a pseudostratified layer of cells in which the majority are ciliated (Burkitt et al., 1993). Pseudostratification results even though all cells touch the basement membrane because differences in cell size make the nuclei appear to be present at different levels, thus creating the illusion of stratification.

2.3.1 Cell types

The epithelium is composed of five to six major cell types lying on a common basement membrane. The most important cells in the upper airways are basal cells, mucous secreting cells, and ciliated cells. The mucous secreting cells include goblet cells, serous cells, small mucous granule cells and the Clara cells, the last of which is present only in the bronchioles, the smaller, distal airways of the lung. Less frequently observed are neuroendocrine cells (also called Kulchitsky cells).

2.3.2 Stem cells

There is clear evidence that there are cells in the lung capable of acting as stem cells, i.e., cells having extensive self-renewal capacity and the ability to generate differentiated progeny. Clusters of cells arising from the same stem cell have been visualized in a model system in which human bronchial epithelial cells are seeded into a denuded rat trachea and implanted into athymic mice (Zepeda et al., 1995). In this experiment, the human bronchial epithelial cells were transfected with recombinant retroviruses expressing reporter genes prior to seeding. Clonal expansion of individual retrovirus-marked cells was observed, with the largest clones comprising 10^3 - 10^4 cells.

Which of the cell types in the lung acts as stem cells is still a source of controversy. Ciliated cells are generally considered to be terminally differentiated and therefore are not candidates as stem cells. The other cell types are not terminally differentiated, do not contain cilia, and are capable of cell division to replace lost or damaged cells. There are two views reported in the literature about which cells of the tracheobronchial epithelium constitute stem cells. One model postulates that the basal cells act as the principle stem cell

in the conducting airways, giving rise to the various secretory cells and ciliated cells (Blenkinsopp, 1967; Reid et al., 1979). This conclusion is largely based on cell turnover studies of normal respiratory epithelium using [^3H]-thymidine as a marker of cell proliferation (Donnelly et al., 1982). Support for basal cells as stem cells is also derived from an *in vivo* culture system for testing proliferation. Tracheal grafts were stripped of their own epithelium and then inoculated with basal cells and transplanted subcutaneously into nude mice (Terzaghi et al., 1978). Basal cells were discovered to be capable of reconstituting an epithelium containing basal cells, secretory cells and ciliated cells, as determined by light and electron microscopy (Nettesheim et al., 1990).

An alternative model postulates that the secretory cells are the pivotal cells from which other cells develop. This model is based on studies of regeneration of chemically or physically injured epithelium which suggest that basal cells only generate basal cells and that the secretory cells of the trachea and bronchus are responsible for regeneration of the rest of the epithelium (Evans et al., 1986; Keenan et al., 1982a; Keenan et al., 1982b; McDowell et al., 1982).

Clara cells which occur exclusively in the bronchioles in mammalian species (Jeffery et al., 1977; Kuhn et al., 1974) are rich in enzymes involved in drug metabolism (Boyd, 1977). These cells were observed to be capable of reconstituting an epithelium in the tracheal graft model described above (Nettesheim et al., 1990). Clara cells have also been shown to serve as the progenitors both for themselves and for ciliated cells during the bronchiolar epithelial differentiation in the fetus in rabbits (Plopper et al., 1991).

2.3.3 Cell kinetics

The mitotic activity of bronchial epithelium in steady-state conditions is extremely low (Kauffman, 1980; Plopper et al., 1987). This has made routine use of the mitotic index to estimate cell turnover difficult because there are too few mitoses for this method to yield statistically reliable results. An estimate of human bronchial epithelial cell turnover is possible, however. In a recent study of human lung epithelium in which over 250 lungs were analyzed at autopsy, while no mitoses were observed, labeling with an antibody against Ki-67 antigen permitted determination of the fraction of cells actively synthesizing DNA (Boers et al., 1996). The proliferating fraction of all epithelial cells was $0.76 \pm 0.53\%$.

This number has led to some speculation about the organization of cells in the epithelial layer. Since 0.0076 is approximately 1/128, the proliferation rate is consistent with a model in which one cell of a "turnover unit" of 128 cells divides once per day, and one cell in the turnover unit dies each day. This also suggests that each cell divides approximately 3 times per year, which correlates well with the rate observed by Dr. E. Furth (University of Pennsylvania) for colon epithelium, as described further below in the Discussion (Alternatives to mutagenesis: Cell kinetics).

In rodents, the labeling index for bronchial epithelial cells has been determined by injecting radioactive tracer such as [^3H]-thymidine to label cells in S phase. The results of a number of studies are summarized in Figure 10. In general, about 1 in 1000 cells is undergoing mitosis at any time.

2.3.4 Cells and lung cancer

Mucous secreting cells and basal cells, because they are actively dividing and incorporate tritiated thymidine (Blenkinsopp, 1967; Evans et al., 1986), represent the most important cells of origin for lung cancer. In one study in which one hundred human primary lung carcinomas were classified based on the cell of origin (McDowell et al., 1978), basal and/or mucous cells were the cells of origin for 88% of the tumors, 7% were

Fig. 10: Proliferative indices for bronchial epithelial tissues.

Results from multiple experiments to determine proliferative indices of various cell types in the bronchial epithelium of a number of species are summarized.

Source: Shami and Evans, 1992 (for references, see Shami and Evans, 1992)

Intrapulmonary Basal Cells Proliferative Indices (%) (Trachea and Bronchi)

Animal	Age	Cell		Tissue		Authors
		Mitotic	Labeling	Mitotic	Labeling	
Hamster	1 day	0.2				Blenkinsopp, 1967 Boldue and Reed, 1976 Boron and Paradise, 1978 Donnelly et al., 1982
	3 months	0.04-0.05	0.3-1.1			
Monkey	4-9 year				0.7	Keenon et al., 1982 Lane and Gordon, 1974 McDowell et al., 1984, 1985 Wells, 1970 Wilson, 1984
Rat	1 month	2.3	7.0	2.2	1.4	
	3 months			0.18		

Intrapulmonary Basal Cells Proliferative Indices (%) (Bronchi)

Animal	Age	Cell		Tissue		Authors
		Mitotic	Labeling	Mitotic	Labeling	
Hamster	3 month		1.5			Boldue and Reed, 1976 Breuer et al., 1990
Rat	1 month			0.003-0.15	0.44	Evans et al., 1986
	3 months			0.25		

Extrapulmonary Nonciliated Cells Proliferative Indices (%) (Trachea and Bronchi)

Animal	Age	Cell		Tissue		Authors
		Mitotic	Labeling	Mitotic	Labeling	
Hamster	1 day	0.2				Blenkensopp, 1967 Boldue and Reid, 1967 Boron and Paradise, 1978 Dennelly et al., 1982
	3 month	0.02-0.1				
Rat	1 month		0.9	0.9		
	3 months			0.1		

Intrapulmonary Nonciliated Cells Proliferative Indices (%) (Bronchi)

Animal	Age	Cell		Tissue		Authors
		Mitotic	Labeling	Mitotic	Labeling	
Hamster	3 months		1.0			Boldue and Reid, 1967 Breuer et al., 1990
Rat	1 month			0.05-0.07		Evans et al., 1986
	3 months			0.05		

Intrapulmonary Bronchiolar Cells Proliferative Indices (%) (Clara Cells)

Animal	Age	Cell		Tissue		Authors
		Mitotic	Labeling	Mitotic	Labeling	
Hamster	12 months				0.16	Castleman et al., 1980 Eustis et al., 1981
Monkey	2-7 years				0.1	Evans et al., (1976, 1977, 1978, 1982) Hackett (1979, 1983)
Mouse	3 months				0.3	Lum et al., 1978 Shami et al., 1982, 1986
Rat	1 month				0.2-1.1	
	3 months				0.1-1.1	
	11 months				0.1	
	19 months				0.2	
	25 months				0.8	

derived from neurosecretory cells, 1 tumor (1%) was attributed to Clara cell origin, and 4% were unclassified.

2.4 SMOKING

2.4.1 Age-specific lung cancer rates

A close association between cigarette smoking and lung cancer has been thoroughly documented (Doll et al., 1976; General, 1989; Hammond, 1966; IARC, 1986; Wynder et al., 1950). In order to better understand the relationship between smoking and lung cancer, P. Herrero of the W. Thilly laboratory has compiled the probability of lung cancer death for individuals residing in the United States since 1900 as published in Mortality Statistics (1900-1936) and Vital Statistics of the United States (1937-1992). Lung cancer incidence was plotted by year of birth. Shown in Figure 11A and 11B are the number of cases of lung cancer death per 100,000 persons per year for both European American males and females by birth year. A significant increase in the incidence of lung cancer is apparent among those born in more recent cohorts for both males and females, a shift that is hypothesized to reflect the widespread influence of smoking. Further, the rise in male deaths occurred in earlier birth cohorts than in females, which may mirror the earlier rise in smoking prevalence among men than women. These hypotheses are substantiated by Figures 11C and 11D. Figure 11C provides the probability of death by lung cancer in men and women between 60 and 64 years of age as a function of birth year. Figure 11D represents the fraction of each birth year cohort which reported substantial cigarette smoking. The relationship between cigarette use and lung cancer death is supported by the similarity of the curves. For both lung cancer mortality and cigarette smoking prevalence, there is an approximately 30 year difference between the increase in men and in women.

2.4.2 Smoking and lung cancer

The public health significance of lung cancer is clear: lung cancer is currently the leading cause of cancer death in the United States. Lung cancer is also the most common cause of death from cancer worldwide. In 1994, there were an estimated 400,000 deaths from lung cancer in the U.S. alone, which represented 30% of all cancer deaths (Boring et al., 1994). The total national cost of smoking in 1994 was \$20.8 billion in direct health care costs, \$6.9 billion in premature disability and \$40.3 billion in premature death (McFarling, 1994). In 1912, before widespread use of cigarettes, only 374 cases of lung cancer had ever been reported in the medical literature. Overall survival statistics for lung cancer are exceedingly poor: fewer than 8% of patients are alive 5 years after diagnosis.

In the 1950s the British Doctors' Study demonstrated that an increased risk of death from lung cancer is associated with smoking, a finding fully substantiated in the 1979 Report of the Surgeon General on Smoking and Health. The study's conclusions included the following observations: Cigarette smokers taken as a whole are at an approximately 10-fold greater risk of lung cancer than nonsmokers. A close correlation exists between the risk of lung cancer and the number of cigarettes smoked per day. The increase in relative risk is 20-fold for those who smoke two packs or more per day. In addition to the Surgeon General's report, several large epidemiological studies demonstrated a clear dose-response relationship between smoking and lung cancer; these results are summarized in Figure 12.

Lung cancer is comprised of three major histological types: squamous cell (epidermoid) carcinoma, small cell carcinoma, and adenocarcinoma. Kreyberg and Doll hypothesized in the 1950's that squamous and small cell tumors were related to smoking while adenocarcinomas were not as closely associated (Doll et al., 1957; Kreyberg, 1955). More recent studies have reported that adenocarcinomas are also associated with smoking (Weiss et al., 1972).

Fig. 11: Age-specific lung cancer rates.

Number of cases of lung cancer death per 100,000 people for European American males (A) and European American females (B) by birth year. (C) Number of lung cancer deaths per 100,000 individuals for 60-64 year old male and females by birth year. (D) Fraction of US cigarette smokers by birth year.

Source: Pablo Herrero compiled from Mortality Statistics (1900-1936) and Vital Statistics of the United States (1937-1992) and Harris, 1983.

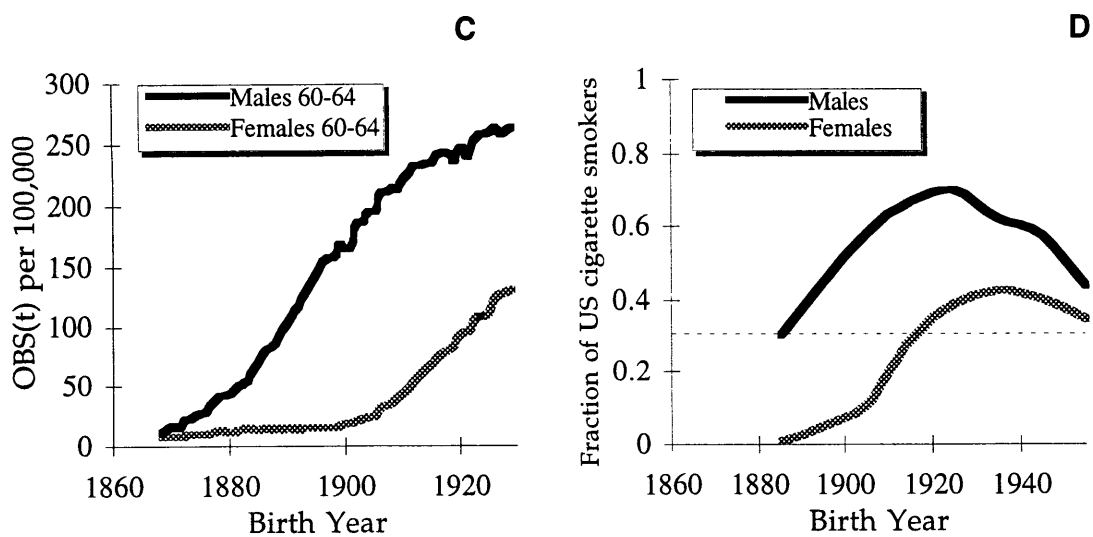
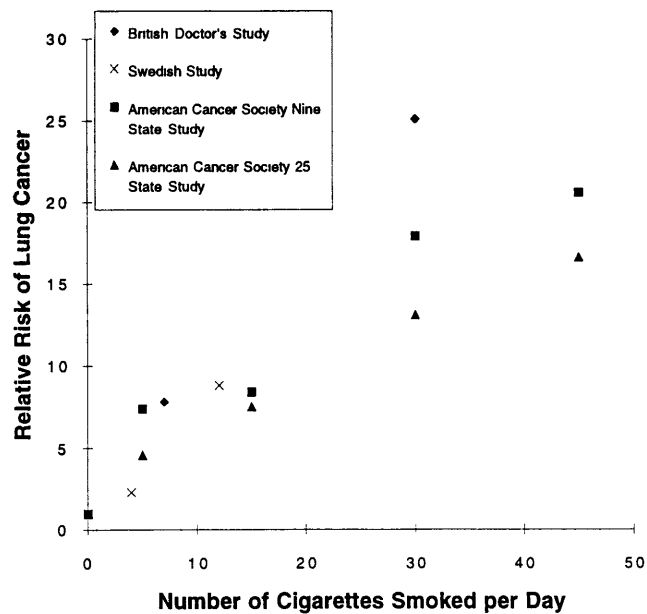


Fig. 12: Correlation between smoking and lung cancer risk.

Results of four large epidemiological studies are summarized with respect to lung cancer risk as a function of cigarette dose. In the accompanying table, the methods used in each of the studies is summarized.



Study	Year of Enrollment	Sample size: initial samples; in brackets, population for follow-up	Source of information of smoking (proportion of respondents)	Duration of follow-up and no. of deaths	Completeness of follow-up for mortality
American Cancer Society Nine State Study	1952	204,547 men [187,783]	Self-administered questionnaire	44 months, 11,870 deaths	0.989
British Doctors Study	1951	34,440 men (aged 20+)	Self-administered questionnaire (69% respondents)	20 years, 10,072 deaths	not available
American Cancer Society 25-State Study	1959-1960	1,078,894 subjects First follow-up: 440,558 men, 562,671 women (aged 35-84); second follow-up: 358,422 men, 483,519 women	Self-administered questionnaire	4.5 + 5 years, 26,448 deaths in men; 16,773 deaths in women	97.4% in women, 97.9% in men in first follow-up
Swedish Study	1963	27,342 men, 27,732 women (aged 18-69)	Self-administered questionnaire (89% respondents)	10 years, 5655 deaths (2968 autopsies)	not available

Lung tumors originate from bronchial epithelial cells of the upper airways, generally within the first few generations of airways, however the location depends upon the tumor type. Adenocarcinomas tend to occur at approximately the third generation of the airway while squamous and small cell carcinomas tend to occur within the first and second (Kvale et al., 1976). Bronchoalveolar carcinomas, a distinct tumor type, occur very peripherally, and are thought not to be related to smoking. Ashley and Davies (1967) found that smoking was as closely associated with adenocarcinomas as with other types of lung cancer and concluded that a tumor's histological type depended entirely upon which part of the respiratory tract was affected (Ashley et al., 1967).

2.4.3 Mutagenicity of selected components of cigarette smoke

2.4.3.1 Benzo[a]pyrene (BaP)

Benzo[a]pyrene is a ubiquitous polycyclic aromatic hydrocarbon mutagen and carcinogen found in cigarette smoke, atmospheric pollution and a variety of foods (IARC, 1986). The estimated daily exposure to BaP from different sources is 400-800 ng/day for 20 cigarettes; 10,000 ng per charcoal-broiled steak; 9-40 ng/day from ambient air; and 1 ng/day from drinking water (El-Bayoumy, 1992; Agency for Toxic Substances and Disease Registry, 1993)

BaP is metabolized *in vivo* into chemically reactive intermediates which can then react with cellular macromolecules including DNA (Brookes et al., 1964; Gelboin, 1969). The 7,8-diol-9,10-epoxide (BPDE) is the most active metabolite of benzo[a]pyrene in terms of binding to DNA (Ivanovic et al., 1976; Meehan et al., 1976; Sims et al., 1974), mutagenicity in Chinese hamster ovary cells (Wood et al., 1977), and tumorigenic activity in mice (Buening et al., 1978; Kapitulnik et al., 1977). Newborn mice treated with the most active diastereomer of the diol epoxide developed an average of 4.42 pulmonary adenomas per mouse, as compared with an average of 0.13 lung tumors per mouse in vehicle-treated controls (Kapitulnik et al., 1978).

The major adduct formed when BPDE reacts with DNA *in vitro* is the adduct at the N² position of guanine (Straub et al., 1977; Weinstein et al., 1976), although adduction at the N⁷ position of guanine and the N² position of adenine also occur at a considerable frequency.

When 185 nonsense mutations in the *lacI* gene of *E. coli* induced by BPDE were investigated, GC->TA transversions were the most frequently observed type of mutation (Eisenstadt et al., 1982). This led to the suggestion that an abasic intermediate may be important in BaP-induced mutagenesis, or that rotation of a modified N²-guanine around its glycosylic bond could result in mispairing. In addition to studies of the *lacI* gene, several subsequent studies have concluded that GC->TA mutations are induced by benzo[a]pyrene diol epoxide. These studies used as targets a shuttle vector replicated in either human cells (Yang et al., 1987) or in monkey cells (Roilides et al., 1988), and BPDE-treated plasmids replicated in bacteria (Chakrabarti et al., 1984). BPDE-induced mutations have also been analyzed in the endogenous *aprt* gene in Chinese hamster ovary cells (Mazur et al., 1988) and the *hprt* gene in human cells (Chen et al., 1990; Yang et al., 1991). The mutants analyzed in all of these studies were predominantly GC->TA transversions.

Benzo[a]pyrene has been used as a model compound in an experiment designed to determine the effect of dose rate on mutational spectra (Chen et al., 1996). Using a combination of *en masse* treatment with benzo[a]pyrene, selection for *hprt* mutants with 6-thioguanine, PCR and denaturing gradient gel electrophoresis (DGGE), the mutational spectrum induced by benzo[a]pyrene treatment of a human cell line containing the metabolizing enzyme aryl hydrocarbon hydroxylase (although not the epoxide hydrolase) was determined. Duplicate cultures were treated with a single toxic exposure of 30 μ M for

28 hours, a nontoxic exposure of 0.5 μM for 6 days, and an exposure approximating estimates of BaP concentration in the human lung of 20 nM for 20 days. The mutational spectra were critically dependent upon the conditions of exposure. While the long-term, low-dose and toxic dose treatments induced mainly GC->TA mutations, the intermediate dose treatment induced two AT->TA mutations and a 7 bp deletion.

2.4.3.2 Tobacco-specific nitrosamines

Another class of potent mutagenic compounds in tobacco smoke are the tobacco-specific nitrosamines which are formed by nitrosation of nicotine during tobacco processing (Burton et al., 1989). The amount of several N-nitroso compounds in cigarettes has been measured. In main stream cigarette smoked, there is 0.08-0.77 ng per cigarette of N-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK), 0.2-3.7 ng/cigarette of N-nitrosonornicotine (NNN), and 0-0.15 ng/cig of NAB (N'-nitrosoanabasine) (IARC, 1986).

N-nitroso compounds produce tumors in up to 40 animal species (Bogovski et al., 1981). They display unusual species, organ and cell specificity and a wide variety of genetic effects. NNK is a highly potent carcinogen in mice, rats and hamsters, and is specific for lung tissues (Hecht et al., 1988). The compound requires activation by cytochrome p-450 to exhibit its mutagenic and carcinogenic properties (Crespi et al., 1991; Hecht et al., 1988). The initial and crucial step of this activation is α -hydroxylation of the two carbons bound to the N-nitroso group to form reactive electrophilic species (see Figure 13). Several cytochrome P-450 isozymes have been associated with the metabolic activation of NNK including P-450 2B1, 1A1, 1A2, 2D6 and 2E1 (Crespi et al., 1991; Devereux et al., 1988).

The compound 4-(methylnitrosamino)-1-(3-pyridyl-N-oxide)-1-butanol- (NNAL-) glucuronide, a metabolic byproduct, has been detected in the urine of all of the over 100 smokers analyzed to date, but not in nonsmokers who have not been exposed to passive environmental tobacco smoke. In one study, amounts of NNAL-glucuronide ranged from 0.16 to 19.0 pmol/mg creatinine, levels consistent with the amounts of NNK reported in mainstream cigarette smoke (Carmella et al., 1995).

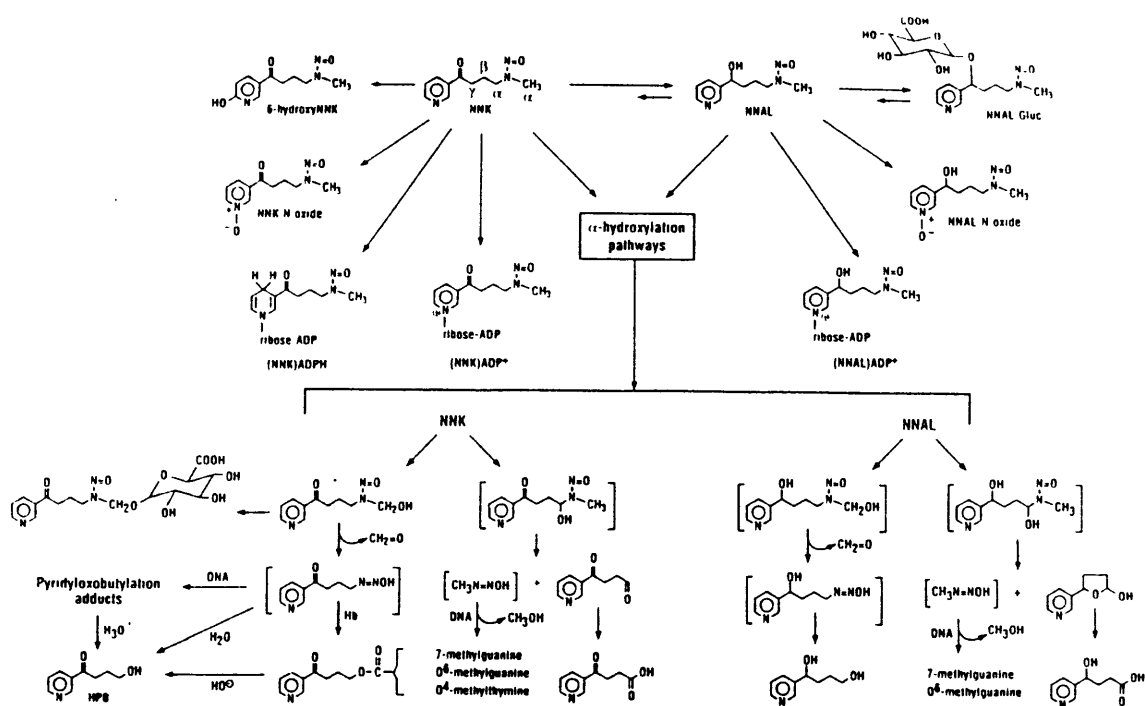
Activated NNK can methylate and pyridyloxobutylate purine bases of DNA (Hecht et al., 1980; Murphy et al., 1990). Methyl and pyridyloxobutyl adducts have been detected in DNA of tissues from animals treated with NNK (Belinsky et al., 1986; Hechts et al., 1988). Analysis of mutations in the K-ras gene isolated from tumors in A/J mice treated with NNK showed a high prevalence of GC->AT transitions in the second base of codon 12 (GGT->GAT), which is consistent with the expected mutation induced by O⁶-methylguanine (Belinsky et al., 1992; Belinsky et al., 1989; Ronai et al., 1993). Activated K-ras genes with mutations in codon 12 have been detected in 16% of human lung tumors and in 24% of adenocarcinomas of the lung (Rodenhuis et al., 1992). However, the types of K-ras mutation observed most frequently in human lungs is GGT->TGT (58%), which is not the specific mutation observed in NNK-treated mice. The NNK-induced GGT->GAT mutation represents is observed in 19% of the tumors with K-ras codon 12 mutations.

The mutational specificity of NNK has also been investigated in a mammalian cell system expressing a human cytochrome p450 2A6 and containing the *guanine phosphoribosyl transferase* gene as a target for mutagenesis (Tiano et al., 1994). A dose-dependent increase in cytotoxicity and mutant frequency was observed upon treatment with NNK, but only in a cell line containing the p450 2A6 enzyme. At the highest NNK dose (1200 $\mu\text{g/ml}$), the mutant fraction (339×10^{-6}) was 14-fold greater than the spontaneous mutant fraction (24×10^{-6}). Among 98 mutant clones analyzed by PCR, 21 contained deletions/rearrangements and 77 were putative point mutants. Sequencing showed that 81% of the point mutants were GC->AT transitions and four of the six most frequently

Fig. 13: NNK metabolism

NNK metabolism pathways, as determined in rodents and primates. Compounds in brackets are postulated intermediates.

Source: Hecht, 1996.



mutated positions were at the second G of a GGT motif, which is the type of mutation and the sequence context of the mutation discovered in codon 12 of K-ras in treated mice.

2.4.3.3 Cigarette smoke condensate

The first report of *in vitro* genotoxic activity caused by an aqueous fraction of cigarette smoke particulate matter, cigarette smoke condensate (CSC), was in onion root-tips (Venema, 1959). A 0.04 ppm fraction induced chromosome lagging, "sticky" chromosomes, and acentric fragments after only two hours, the severity of which increased after four hours. Leuchtenberger and Leuchtenberger later extended these observations to the gaseous phase of cigarette smoke (Leuchtenberger et al., 1970). Kier and colleagues were the first to demonstrate the ability of CSC to induce point mutations. In the *Salmonella typhimurium* microsomal test, the condensates from several types of cigarettes induced frameshifts in the presence of liver or lung microsomal enzymes (Kier et al., 1974). The CSC from just one-hundredth of one cigarette per plate produced detectable mutagenicity.

There are two reports in the literature of increased mutant fractions in mammalian cells after treatment with cigarette smoke condensate. Both Clive and colleagues and Mitchell and colleagues reported dose-response relationships between mutagenicity and exposure to cigarette smoke condensate in L5178/TK⁺/⁻ mouse lymphoma cells. However, the cloning efficiency of the cells declined as the mutant fraction increased, so that the product of survival and mutant fraction was similar at different doses. Such findings suggest that the induced mutant fractions may be an artifactual result of decreased cloning efficiency because, as discussed previously, the presence of wild-type cells may suppress the recovery of mutants in the HPRT assay and the effect is heightened in samples with high cloning efficiency (Khaidakov et al., 1996). The short period of time elapsed between treatment with Aroclor stimulated rat liver homogenate and plating (just 2 hours) suggests that the cells may not have retained the cloning efficiency of control cells as would be necessary in such an experiment.

In a more recent experiment Charles Crespi and Bruce Penman of Gentest Corporation (Waltham, MA) treated a number of cell lines expressing different cytochrome p450s with cigarette smoke condensate and did not observe a statistically significant increase in mutant fraction in any of the lines (B. Penman, personal communication).

2.4.4 Smoking and DNA adducts

2.4.4.1 Smoking-induced adducts

Several studies have reported an increase in the level of aromatic DNA adducts in lung tissue from human surgical specimens or autopsy material based on the ³²P postlabeling assay (Cuzick et al., 1990; Phillips et al., 1990; Randerath et al., 1989; Wiencke et al., 1995). In one study by Philips and colleagues, the significance level reached $p < 0.01$ and the values were 5.53 ± 2.13 adducts per 10^8 nucleotides for smokers versus 3.45 ± 1.62 adducts per 10^8 nucleotides for nonsmokers (Phillips et al., 1990). In a separate study, purified DNA from human lung and several other organs was analyzed for DNA adducts using the nuclease P1 modification of the ³²P postlabeling technique (Cuzick et al., 1990). Lung was the only organ for which a statistically significant increase in adduct levels in smokers was observed. Adduct levels in lungs of current smokers were 9.3 ± 6.5 adducts per 10^8 nucleotides; in ex-smokers, 6.0 ± 4.2 adducts per 10^8 nucleotides and in non-smokers, 2.5 ± 2.3 adducts per 10^8 nucleotides.

DNA adduct profiles have also been examined from bronchial biopsies obtained from 78 individuals undergoing routine diagnostic bronchoscopy (Dunn et al., 1991).

Lung cancer was present in 37 (47%) of the patients. ^{32}P postlabeling chromatograms from all of the smokers ($n=28$) showed a distinctive diagonal adduct zone which was present in 24 of 40 ex-smokers and 4 of 10 lifetime non-smokers. Adduct levels and chromatographic patterns were similar in bronchial tissue from different lobes of the lung, and in tumor and nontumor tissue taken from the same patient. Adduct levels were not correlated with cancer independent of smoking.

A different conclusion about regional variability within the lung was reached by Weston and Bowman (Weston et al., 1991). BPDE-adducted nucleotides were purified by immunoaffinity chromatography and the major product was detected by fluorescence spectroscopy. Levels of BPDE-DNA adducts in the range of 1-40 in 10^8 nucleotides were detected in 6 out of 25 samples. Adduct levels varied significantly among different parts of the same lung. In one particularly extreme case, BPDE adduct levels were 35 attamole BPDE/ μg DNA in one section of the lung and 1215 attamoles BPDE/ μg DNA in another section.

Studies of adduct levels in smokers and nonsmokers have also been performed in peripheral blood. Several studies of leukocyte DNA have shown an association between polycyclic aromatic hydrocarbon (PAH)-DNA adduct levels and smoking measured by ELISA (Kriek et al., 1993; Liou et al., 1989). Other studies have not demonstrated a statistically significant relationship between adducts and smoking by either ELISA or ^{32}P postlabeling (Jahnke et al., 1990; Perera et al., 1989; Van Schooten et al., 1992). An association between smoking and PAH adducts was observed with both ELISA and ^{32}P postlabeling using the longer-lived lymphocyte/monocyte fraction of white blood cells (Santella et al., 1992; Savelle et al., 1991). In another study, a significant reduction in mean PAH-DNA adduct levels in leukocytes was observed after 8 months of smoking cessation ($p<.05$) (Mooney et al., 1995). The mean PAH-DNA adducts for all 40 subjects was 2-fold higher while smoking [8.3 ± 8.8 adducts/ 10^8 nucleotides ($n=40$)] than 8 months after cessation [4.0 ± 4.1 adducts/ 10^8 nucleotides ($n=35$)].

Increased levels of adducts other than polycyclics have also been reported in smokers. Two studies have reported significant differences between smokers and non-smokers in levels of the alkylated base N^7 -methylguanine. In one report, mean levels of N^7 -methylguanine were 6.9 per 10^7 nucleotides in total white blood cells, 4.7 per 10^7 nucleotides in granulocytes and 23.6 per 10^7 nucleotides in lymphocytes of smokers, compared with values of 3.4, 2.8 and 13.5 per 10^7 nucleotides in nonsmokers (Mustonen et al., 1992). A second study analyzed N^7 -methylguanine in bronchial DNA samples and found mean levels of 17.3 adducts per 10^7 nucleotides in smokers and 4.7 adducts per 10^7 nucleotides in nonsmokers (Mustonen et al., 1993). It has also been reported that smoking induces oxidative adducts based on an observed 50% increase in 8-OH-dG lesions in sperm DNA samples from smokers (Fraga et al., 1996).

As discussed in the section on DNA repair in mitochondrial DNA, repair of these adducts is expected to differ between mitochondrial and nuclear DNA. While adducts recognized by base excision repair machinery including oxidative damage (Driggers et al., 1993) and alkylation are expected to be repaired (LeDoux et al., 1992), large, bulky adducts including aromatic amines and polycyclic aromatic hydrocarbons would not be expected to be repaired (Clayton et al., 1974).

2.4.4.2 Smoking and adducts in mitochondrial DNA

The extent of DNA damage in mitochondrial DNA has also been reported to be significantly higher in bronchoalveolar lavage tissues from smokers as compared with nonsmokers (Ballinger et al., 1996). In this study, DNA damage was assessed by measuring the amount of PCR product created when defined amounts of template were used for long template PCR (16,262 bp for mitochondrial DNA and 13,458 bp for nuclear

DNA). The absence of PCR product was interpreted to reflect damage to the template DNA that inhibited synthesis by the polymerase. A significant increase was observed in the level of DNA damage in smokers as compared with nonsmokers for a nuclear gene ($p=0.0056$), and an even larger difference between smokers and nonsmokers was observed for long template PCR on the mitochondrial genome ($p=0.00072$). In another study employing synchronous fluorescence spectrophotometry, Balansky and colleagues detected increased levels of DNA adducts in mitochondrial DNA of lung tissue in rats exposed to mainstream cigarette smoke (Balansky et al., 1996).

2.4.5 Smoking and mutations

2.4.5.1 Smoking-induced mutational spectra

A smoking-related increase in the frequency of peripheral T lymphocytes containing mutations in the *hypoxanthine-guanine phosphoribosyl transferase* (*hprt*) gene has been demonstrated by several laboratories (Albertini et al., 1988; Ammenheuser et al., 1994; Cole et al., 1988; Jones et al., 1993). Cole and colleagues, for example, have carefully analyzed experimental factors and donor attributes that affect *hprt* mutant fractions in blood (Cole et al., 1988). In this study, smoking (range: 5 to >30 cigarettes per day) caused an estimated 56% increase in mutant frequency ($p<0.001$). Jones and colleagues also found that years smoked was an important determinant of *hprt* mutant fraction in T-lymphocytes from 62 smokers and 58 nonsmokers (Jones et al., 1993). In a study of 11 smokers and 12 non-smokers, Hüttner and colleagues observed that the mean variant frequency in smokers was significantly higher (70%) than that in non-smokers (Hüttner et al., 1990). Tate and colleagues reported that the mean mutant fraction for smokers was 36% higher than for nonsmokers (Tate et al., 1991). This difference was statistically significant only if mutant fractions were adjusted for cloning efficiency. Other studies have not reported an increased mutant fraction in smokers [for instance, (Davies et al., 1992)].

The influence of smoking on mutant fraction has also been studied for the glycophorin A (GPA) gene, which codes for the M and N blood group erythrocyte cell surface antigens (Langlois et al., 1993). The 50% of the population with a MN blood phenotype are suitable for analysis. This phenotype results from codominant allelic expression at the GPA locus, thus permitting independent detection of nonlethal null mutations in either the M or the N allele. Using this approach, a small, not statistically significant smoking effect of a 20% increase in the fraction of singly labeled cells was reported in one study (Bigbee et al., 1989), while no correlation of mutant fraction with smoking was observed in another study (Jensen et al., 1995).

In addition to testing for effects on mutant fraction, it is also possible to sequence individual *hprt* clones to determine whether smoking affects the specific mutations present. In one study, T-lymphocytes from nine adult smokers (≥ 15 cigarettes per day) and 12 age-matched controls were plated. 6TG^r clones were isolated, and *hprt* mutations were identified by reverse transcriptase-PCR and sequencing. Vrieling and colleagues discovered an elevated mutant fraction in the smokers but did not discern a difference in the relative frequency of mutation types in the smokers and nonsmokers (Vrieling et al., 1992). In both smokers and nonsmokers, transitions accounted for approximately 15% of the total mutations; transversions for about 20%; exon skipping for approximately 45%; and deletions and frameshifts for almost 20%. GC→TA transversions, the mutations frequently observed among benzo[a]pyrene diol epoxide-induced mutations in human cell culture (Keohavong et al., 1992), was not detected in smokers.

Burkhart-Schultz and colleagues reported the sequences of 28 mutations in the *hprt* gene from smokers and nonsmokers. They concluded that there was a suggestion ($p=0.2$ by chi-squared analysis) of more mutations at AT base pairs in smokers (6/14) than nonsmokers (2/14) (Burkhart-Schultz et al., 1993).

A. Tomita of the W. Thilly laboratory has analyzed a database of all reported *hprt* mutations (Cariello, 1996) in order to discover whether smokers' *hprt* mutations differ significantly from those of nonsmokers. She compared the set of smokers' somatic mutations to the set of nonsmokers' somatic mutations divided into three categories: deletions, other alterations and insertions. The smokers were significantly different from the nonsmokers based on a chi-squared analysis at the level of $p < 0.01$, due largely to an increase in the relative number of point mutations (126/205 for smokers versus 107/212 for nonsmokers). When the types of point mutations were considered separately, the frequency of GC→TA mutations were elevated significantly in smokers compared with nonsmokers at the level of $p < 0.05$ but not $p < 0.01$ (15/97 mutations in smokers and 8/88 mutations in nonsmokers). In conclusion, there is some evidence that smoking leads to an increase in mutant fraction of approximately two-fold in nuclear sequences but a smoking-induced mutational spectra has not been characterized.

2.4.5.2 Smoking-induced deletions of mitochondrial and nuclear DNA

The ability of cigarette smoking to cause deletions has also been investigated. One study utilized mouse embryos homozygous for the pink-eyed unstable (p^{un}) mutation which is the result of duplication of a 70 kb internal fragment of the mouse *p* gene. Spontaneous reversion events occur through recombination-mediated deletion of one copy of a duplicated sequence. Reversion events in premelanocytes in the mouse embryo give rise to black spots on the gray fur of the offspring. Both whole body exposures of pregnant p^{un} mice to cigarette smoke for 4 hours and treatment with a 15 mg per kg dose of cigarette smoke condensate during their 10th day of gestation led to a significant increase in the number of DNA deletions in the embryos as evidenced by spotted offspring (Jalili et al., submitted). A concentration of 31 ± 2 mg per m^3 of total particulate matter raised the spotting frequency from the control value of 12.6% to 23.3% ($p \leq 0.05$), while exposures at 96 ± 15 mg/ m^3 resulted in 28% of offspring showing spots ($p \leq 0.05$).

Whether smoking induces mitochondrial deletions has also been directly addressed. Ballinger and colleagues (1996) examined the levels of the 4.9 kb common deletion in alveolar macrophages. Smokers contained almost 7 times the level of the common deletion as compared with nonsmokers (0.016 versus 0.0023%), although the differences were not statistically significant because interindividual values were large.

In another study, the levels of the 4.9 kb common deletion, a 7.4 kb deletion and tandem duplication in the D-loop region were investigated in hair follicles from 213 male nonsmokers and 74 male smokers (Liu et al., 1997). Current smokers had a 3.1-fold higher average incidence of the 4.9 kb deletion as compared with nonsmokers ($p < 0.001$). No dose-response relationship was observed with the amount of smoking; the heaviest smokers had lower levels of deletions than light smokers (see Figure 14). No statistically significant relationship was observed between the incidence of the 7.4 kb deletion or duplicated mtDNA and the smoking index. Intriguingly, a significant difference between lung cancer patients and controls was observed for the proportion of the 4.9 kb deletion. In lung cancer patients with high smoking indices, the deleted mtDNA represented 1.2% of hair follicle mtDNA while among healthy subjects with the same smoking index, the proportion was 0.15%.

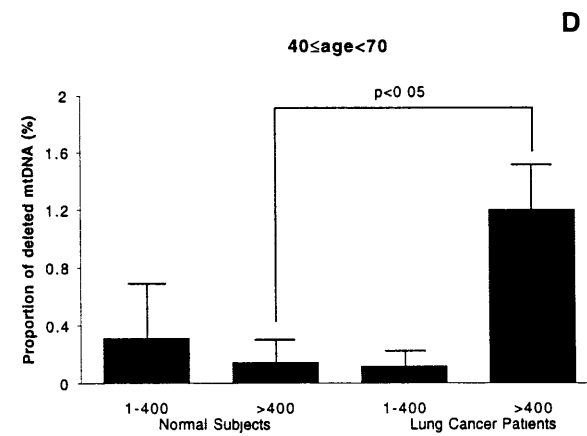
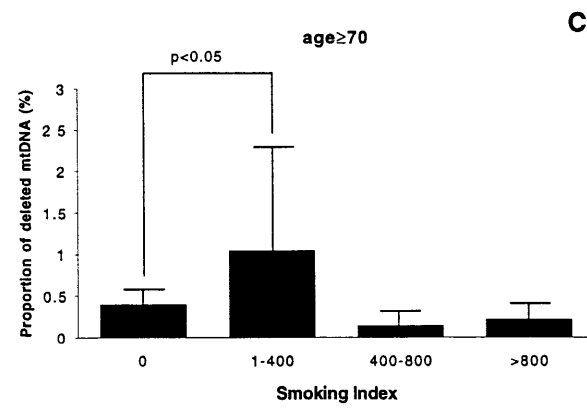
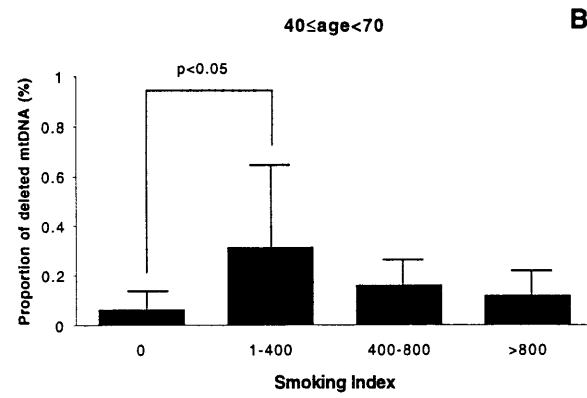
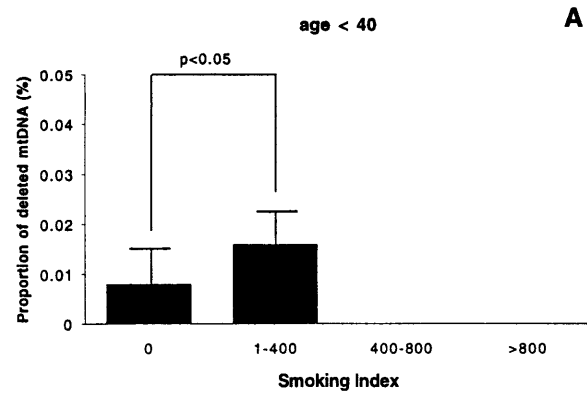
2.4.5.3 p53 mutations in lung tumors

Sequencing the p53 gene in human lung tumors has provided some evidence that smoking may induce mutations in human lungs. As opposed to most tumors, for which GC→AT transitions, mainly at CpG sites, predominate, GC→TA transversions are the mutation type most frequently observed in lung tumors of smoking subjects (Husgafvel-Pursiainen et al., 1996; Jones et al., 1992). The difference between the types of mutations

Fig. 14: Proportions of the common deletion with age, smoking index, and lung cancer status

Comparison of the proportion of the 4,977 bp deleted mtDNA in the hair follicles of subjects in three age groups (A, B, and C) with different grades of smoking index. Bars reflect the proportion (mean $\pm$ S.E.M.) of the 4,977 bp deleted mtDNA. Smoking index = (cigarettes smoked per day) $\times$ (years of smoking). (D) Comparison of the proportion of the 4,977 bp deleted mtDNA in the hair follicles of mild (smoking index within 400) and heavy (smoking index greater than 400) smokers in normal subjects and lung cancer patients, aged between 40 and 70 years.

Source: Liu et al., 1997



in smokers' lung tumors and those observed in other organs suggests that tobacco smoke may be specifically inducing the GC->TA mutations observed. While there are many mutagenic compounds in tobacco smoke, it is worth noting that, as described previously, GC->TA mutations are the predominant type induced by certain polycyclic aromatic hydrocarbons, including metabolically activated benzo[a]pyrene (Chen et al., 1990; Eisenstadt et al., 1982; Keohavong et al., 1992; Mazur et al., 1988).

A recent report on the formation of benzo[a]pyrene adducts in the p53 gene of HeLa and bronchial epithelial cells also provides evidence that supports the hypothesis that benzo[a]pyrene induces the mutations observed in the p53 gene in smokers' lung tumors (Denissenko et al., 1996). Denissenko and colleagues treated bronchial epithelial cells with 4 mM benzo[a]pyrene diol epoxide (BPDE) for 30 minutes. The distribution of adducts in p53 was determined by cleavage with an enzyme complex that digests specifically at adducted sites, the UvrABC nuclease, followed by ligase mediated PCR. The hotspots for BPDE-induced adducts were at the same positions that represent hotspots for p53 mutations in human lung tumors. While adduct spectra are not necessarily synonymous with mutational spectra [see for instance (Fuchs, 1984)], the dose of BPDE was significantly higher than the biologically relevant dose (Chen et al., 1996), and mutations in p53 have not been demonstrated to be the initiating mutations in lung tumors, nevertheless, the study does put forth a plausible mechanism by which smoking could induce lung tumors via mutations.

3. MATERIALS AND METHODS

3.1 HUMAN CELLS AND TISSUES

3.1.1 TK6 cells

TK6 cells (Skopek et al., 1978) and MT-1 cells (Goldmacher et al., 1986), a derivative of TK6 deficient in mismatch repair (Kat et al., 1993), were maintained in exponential growth as spinner cultures by daily dilution with RPMI-1640 media (Gibco BRL, Grand Island, NY) supplemented with 10% horse serum (Gibco BRL) at densities not exceeding 10^6 cells/ml. TK6 cells were grown for more than 400 generations.

3.1.2 Human Samples

Brush biopsies obtained through bronchoscopy were collected from 8 healthy volunteers who gave informed consent. The protocol was approved by the Research Subjects Review Board at the University of Rochester and the Committee on Use of Humans as Experimental Subjects at MIT. Procedures were performed at the University of Rochester. Subjects were free of significant respiratory or cardiovascular disease determined by medical history and physical examination.

Among the volunteers were 3 pairs of monozygotic twins recruited from the Minnesota Twins Registry (Lykken et al., 1990) who were discordant for smoking. For each pair of twins, one was a lifelong nonsmoker while the other had a >10 pack-years smoking history. Twins were all females, 46-54 years old. The other bronchoscopy volunteers were a 28 year old male and a 24 year old female, both of whom were smokers.

Whole lungs were harvested from organ donors at the University of Rochester Medical Center or Brigham and Women's Hospital. Procedures were approved by the Research Subjects Review Board at the University of Rochester, the Committee on Use of Humans as Experimental Subjects at MIT and the Institutional Review Board of Brigham and Women's Hospital. Informed consent was obtained from the next of kin. Lungs were obtained immediately post-mortem and were free of diseases affecting the upper respiratory tree. Two smokers' lungs (one 47 year old female and one 58 year old male), each of whom smoked two packs per day at the time of death, and one nonsmoker's lung (an 84 year old male) were available.

The lungs of an infant who died of sudden infant death syndrome (SIDS) were harvested and dissected by Dr. James Willey and his colleagues at the Medical College of Ohio (Toledo, OH).

3.1.3 Biopsies obtained via brush bronchoscopy

Subjects were premedicated with 0.75 mg of atropine intravenously prior to the procedure. The oropharynx was topically anesthetized with atomized lidocaine. A fiberoptic bronchoscope (PVA-1000 Pentax, Tokyo) with video capability was inserted orally. The vocal cords were topically anesthetized with lidocaine. The bronchoscope was then advanced to the airways. After gently wedging the bronchoscope in a subsegment of the lingula, four 50 ml aliquots of sterile normal saline were instilled and withdrawn under gentle suction. Brush biopsies were then obtained from eight positions within the first three bifurcations of airway: the right main carina, the first bifurcation of the right upper lobe, the first bifurcation of the right lower lobe, the second bifurcation of the right lower lobe, and the corresponding positions on the left side. After each biopsy, the brush was gently shaken in 3 ml of sterile phosphate buffered saline (PBS) on ice. Each position was brushed twice. Cells were pelleted and resuspended in 10 μ l PBS. Each brush yielded 3 x

10^5 to 2×10^6 cells, of which about 95% were epithelial cells based on differential counts performed on cytopsin slides.

3.1.4 Lung dissection

For each lung, the airways were dissected intact from the surrounding parenchyma. The bronchial tree was preserved from the trachea to approximately the sixth bifurcation of airway. The airways of the bronchial tree were divided into sections at each bifurcation. Each sample was rinsed with media [(LHC-8 (Biofluids, Rockville, MD) or RPMI 1640 (Gibco BRL, Grand Island, NY)] in a petri dish. Epithelial cells were obtained from each section by slicing the segment to open the lumen, then gently scraping the luminal surface with a scalpel. Cells were pelleted by gentle centrifugation at $4000 \times g$ for 3 minutes, and then resuspended in a small volume suitable for DNA isolation. Microscopic verification confirmed that the cells obtained in this manner were indeed epithelial in origin. The number of cells obtained from each section ranged from 10^5 to 10^7 .

3.1.5 Other organs

Surgical discard samples of colon and muscle were generously provided by Dr. S. Singer of the Brigham and Women's Hospital. Brain tissue genomic DNA from a 100 year old man and heart muscle genomic DNA from a 77 year old individual were generously provided through a collaboration with Drs. J. Wei and J. Vijg (Harvard Medical School and Beth Israel Hospital).

3.2 MUTATIONAL SPECTRA

3.2.1 DNA isolation and restriction digestion

DNA was isolated from suspended cells or macerated tissue by digestion with proteinase K (1 mg/ml) and RNase A (0.1 mg/ml) in 0.5% SDS followed by ethanol precipitation of DNA. DNA prepared from bronchial epithelial cells in this way was easily restricted and amplified. The method provided high DNA yields from 10^5 to 10^9 cultured cells or from milligrams to grams of tissue. DNA was digested with *RsaI*, which recognizes base pairs 10,009-10,012 of the mitochondrial genome, and *DdeI*, which recognizes base pairs 10,227-10,231 (New England Biolabs, Beverly, MA) at 2 units/ μ g DNA, overnight.

3.2.2 Constant denaturant gel electrophoresis

Constant denaturant gel electrophoresis (CDGE) can be used to separate DNA sequences containing a high melting domain adjacent to a low melting domain based on the sequence of the low melting domain (Hovig et al., 1991). Due to the cooperative melting of DNA as melting domains (Fischer et al., 1983; Poland, 1974), sequences with even a single base pair substitution in the low melting domain will have a higher or lower melting temperature as compared with wild-type, and therefore will be separable on a CDGE. DNA sequences containing mutations that convert AT base pairs present in the low melting domain into GC base pairs will have a higher melting temperature (T_m) than wild-type, and will consequently migrate more rapidly through a gel under partially denaturing conditions. Sequences with mutations that convert GC base pairs to AT base pairs will have a lower melting temperature than wild-type and will migrate more slowly under partially denaturing conditions.

The electrophoretic mobility of a DNA fragment with a biphasic melting profile decreases with increased temperature following a sigmoid curve (see Figure 15). The temperature near the midpoint of the sigmoid curve of the wild-type fragment is desirable for CDGE separations as it will ensure separation of high and low T_m mutants.

CDGE gels were cast as 8% acrylamide, 1/40 bis-acrylamide, TAE (40 mM Tris Acetate, 1 mM EDTA), 0.4 mm thick. During separation, the gels were submerged in a tank containing TAE heated to a temperature around 70°C that was precisely selected based on test gels to optimize separation. Samples of 4-6 μ l of a restriction digestion were loaded and separated for 2 hours at 10 V/cm.

Fluorescein-labeled markers and samples were separated in alternating parallel lanes. The markers were a mixture of PCR amplified *hprt* exon 3 wild-type and mutant homoduplexes and heteroduplexes with melting temperatures similar to those of the mitochondrial DNA mutants to be detected. One lane of each gel contained a second marker which was the wild-type target mitochondrial sequence with an altered high melting domain primer (J50). The melting behavior of this sequence was indistinguishable from that of the normal wild-type sequence, but it could not be amplified with the primers used for subsequent PCR. This primer was designed as such to avoid contamination of the tissue samples (the number of target sequence copies in a marker lane is three orders of magnitude higher than in the sample lane). The marker bands were visualized by illuminating the gel with a 488 nm argon laser (Idaho Technology, Salt Lake City, UT) and marked while viewing the gel through a 520 nm low pass glass filter (Oriel, Stratford, CT). The relative positions of the markers provide information on the exact position to which the desired mitochondrial mutants migrated in the gel.

Homoduplexes with higher and lower melting temperatures than wild-type were excised from the gel from positions below and above the wild-type, respectively. High T_m mutants were enriched by this procedure approximately 100 to 200-fold and low T_m mutants about 5 to 10-fold.

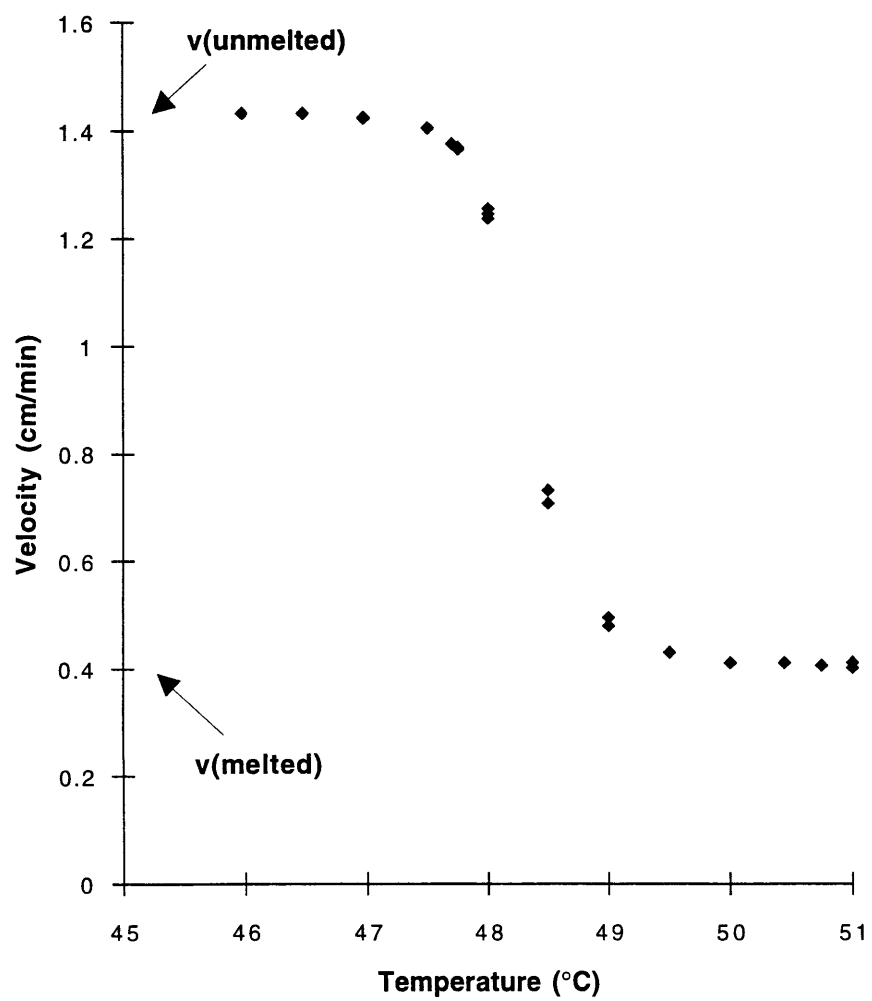
The gel slices were finely ground between two glass slides and transferred into 1.5 ml test tubes containing 200 μ l of 200 mM NaCl. DNA was eluted by shaking for 20 minutes at 50°C, 1300 min⁻¹. Gel particles were spun down, and the supernatant was collected. The DNA was precipitated with two volumes of ethanol and dissolved in about 15 μ l of water.

3.2.3 High-fidelity PCR

PCR was performed in 10 μ l capillaries in an air thermocycler (Idaho Technology, Idaho Falls, ID) with native *Pfu* thermostable DNA polymerase (Stratagene, La Jolla, CA). The PCR mix included 10 mM KCl, 6 mM (NH₄)₂SO₄, 20 mM Tris-HCl (pH 8.0), 2 mM MgCl₂, 0.1% Triton X-100, 100 μ g/ml BSA, 0.2 μ M each primer, 0.1 mM dNTPs, and 0.1 units/ μ l of *Pfu*. Denaturation at 95°C, annealing at 57°C and extension at 72°C, each for approximately 10 second intervals, constituted the amplification cycle. After the desired number of cycles, samples were incubated at 72°C for 2 minutes and then at 45°C for 30 minutes. The primers used were CW7 [(ACC GTT AAC TTC CAA TTA AC) identical to base pairs 10011-10031 of human mitochondrial genome] and J3 [(GCG GGC GCA GGG AAA GAG GT) complementary to base pairs 10196-10215]. The altered high melting domain primer used to generate CDGE markers (J50) was [(GAA GAA TTT TAT GGA GAA AGG GTG CGC CCG GGG GGA TAT AGG GTC GAA GC)]. For fluorescently-labeled primers, a fluorescein moiety was linked to a cyanoethyl phosphoramidite by a nine atom spacer arm. This species was coupled to the 5' end of the oligonucleotide during commercial synthesis (Ransom Hill Bioscience, Inc., Ramona, CA). The tetramethylrhodamine (TMR) dye was coupled to the oligonucleotide via an amino C6 linker also during commercial synthesis (Biosynthesis, Louisville, TX).

Fig. 15: Velocity of DNA migration through the heated zone of the capillary vs. separation temperature.

The target DNA fragment used in these experiments was electrophoresed through a capillary with a 10 cm heated jacket at temperatures ranging from 45°C to 51°C. The electric field was approximately 300 V/cm. Buffer conditions in the gel and the running buffer were 0.1X TBE. The full peak width at half-height is plotted. The mobility is highest when the DNA is in a double-stranded form at lower temperatures up to 47°C, and declines along a sigmoid curve to a lower mobility in the partially melted form at 51°C.



3.2.4 Constant denaturant capillary electrophoresis

Constant denaturant capillary electrophoresis (CDCE) is an adaptation of the separation principles of CDGE to a capillary electrophoresis format. PCR products containing mutations in their low melting domains are separated based on differences in their velocity as they migrate through a heated section of the capillary. An important advance required for reproducible separation by CDCE was a replaceable linear polyacrylamide matrix. The procedure for synthesizing the matrix was adapted from Ruiz-Martinez and colleagues (1993). A 5% acrylamide solution in TBE was deoxygenated by argon bubbling in an ice bath for 10 minutes and any contact of the solution with air was avoided until polymerization was complete. TEMED and ammonium persulfate were added to 0.03% and 0.003%, respectively. The polymerizing solution was immediately dispensed into 10 ml glass syringes and left at 2°C for several days. The matrix was dispensed from the 10 ml syringes into 100 µl high pressure U6K syringes (Rainin, Emeryville, CA), which were used to replace the matrix in capillaries through home-made teflon tube fittings.

In order to reduce adsorption of the matrix to the capillary and electroendosmosis, the capillary was coated with linear polyacrylamide chains in an adaptation of the method of Hjertén (Hjertén, 1985). Capillaries were treated with 1M NaOH for 2 hours, washed sequentially with 1M HCl and methanol, treated overnight with γ -methacryloxypropyltrimethoxysilane (Sigma, St Louis, MO) and washed with methanol. The capillaries were then filled with a polymerizing solution of 6% acrylamide in TBE (89mM Tris, 89 mM borate, 1mM EDTA, pH 8.4), 0.1% TEMED and 0.025% ammonium persulfate, and left for several hours to polymerize.

For CDCE separations, electrophoresis was generally performed in 75 µm I.D. capillaries approximately 30 cm in length. About 10^8 total copies were electroinjected into a capillary and run at about 200 V per cm. A portion of the capillary was heated by a water jacket connected to a constant temperature circulator. For high-resolution runs, a temperature circulator with precision to 0.01°C and an external probe was used (Neslab, Portsmouth, NH). The jacket was positioned 5 cm from the injection end of the capillary. The length of the jacket was 5 cm for the enrichment of heteroduplexes and 15 cm for high-resolution CDCE. The temperature of the water jacket was about 65°C for 1X TBE conditions and approximately 70°C for high-resolution conditions described in the next section. The precise temperature to be used in a particular experiment and the time of fraction collection was determined in test runs with appropriate standards.

The CDCE instrument used for identification of mutants included a dual dye detection system. To detect DNA, the capillary was illuminated by a 515 nm argon laser and emitted light was collected and split. Light was directed into two detectors through appropriate sets of filters. For fluorescein, a combination of a 540 nm bandpass and a 530 nm long pass filter was used; for tetramethylrhodamine (TMR), a single 580 nm bandpass filter was employed. Instruments were also used in a single wavelength detection mode, in which a 488 nm laser was used and filters for fluorescein detection were two 520 nm filters. The signal from the photomultipliers was amplified (10^8 V/A) by a current preamplifier (Oriol, Stanford, CT) and recorded by a MP100 data acquisition system (Biopac Systems, Goleta, CA). DNA fragments are observed in the output as peaks, each of which represents a different sequence, and the area under the peak is proportional to the number of copies of the specific molecule.

At a specific temperature, all heteroduplexes migrate through the heated zone within a definable range of mobilities which is differentiable from that of the wild-type homoduplex. Heteroduplexes in the amplified PCR product could therefore be collected separate from the wild-type homoduplex, thus allowing enrichment of mutants. For heteroduplex collections, samples were electroeluted from the anode end of the capillary into 0.5 ml Eppendorf tubes with 5µl of 0.1X TBE and 0.1 mg/ml bovine serum albumin.

3.2.5 High-resolution CDCE

High-resolution CDCE runs were performed at approximately 100 V/cm using gel and running buffers containing 30 mM Na⁺ (30 mM sodium borate, pH 8), in addition to TBE. Gels for high-resolution CDCE were prepared using a lower amount of ammonium persulfate (0.0015%). The higher salt concentrations in these gels necessitated higher water bath temperatures of approximately 70°C.

In the initial experiments performed to detect mitochondrial mutants in human tissue samples, the mutants observed were isolated and sequenced. As further samples were examined, the same mutants were observed to be present repeatedly in multiple samples. In order to determine whether a peak in a given sample was the same as a previously sequenced peak or if it represented a new mutant, the dual dye detection system described in the previous section was used to identify sample mutants based on comigration during a single capillary run with a panel of previously sequenced mutants. Fluorescein-labeled samples were coinjected into the capillary with authentic standard mutant peaks amplified with PCR primers labeled with tetramethylrhodamine (TMR). Using a dual-channel instrument it was possible to preliminarily identify and measure sample peaks that comigrated with known standards.

3.2.6 Capillary hybridization

In order to identify the sequence change associated with individual peaks more rigorously, capillary hybridization was developed as a test for peak identity. The CDCE instrument used for identification by hybridization included consecutive heating jackets each bathed by a separate water bath. The fluorescein-labeled sample and a panel of TMR-labeled standards were co-separated in the first jacket at the appropriate temperature (~71.5°C) under high-resolution conditions. Approximately 3×10^8 copies of each standard mutant and about 10^7 copies of fluorescein-labeled sample, in which each mutant represented approximately 10% of the copies, was injected. After the mutants were separated from each other as homoduplexes, the run was stopped. The temperature in the second jacket was increased to denature the DNA (80°C), then decreased to 55°C and maintained for 10 minutes to permit reannealing of single strands. The sample was then electrophoresed to the detector at a temperature that separates homoduplexes from heteroduplexes (~70°C). Peak identity was verified when individual fluorescein-labeled peaks were observed as homoduplexes reflecting that they had reannealed. Sample peaks that formed heteroduplexes migrated differently from the standards with which they coseparated and consequently "disappeared".

3.2.7 Test for asymmetric species

A test was designed to distinguish between signals that resulted from true, double stranded mutants present in cells versus those that originated from cellular pre-mutagenic species, like adducts and heteroduplexes, that were converted to mutants during *in vitro* PCR based on linear preamplification with one or the other primer. For these experiments, DNA eluted from a CDGE was linearly amplified with either the upstream or downstream primer for 20 cycles. Eluted DNA was also amplified with both primers for 5 cycles. Exonuclease-deficient *Pfu* DNA polymerase was used for this first amplification for all three samples in order to avoid exonuclease digestion of the single strand products, and conditions were otherwise the same as for other PCR reactions. In some cases, a heteroduplex was added to the sample prior to linear amplification to serve as an asymmetrical control expected to yield different mutant fractions depending on the primer used for preamplification. Samples were then subjected to exponential amplification with

exonuclease-proficient *Pfu* DNA polymerase. Heteroduplexes were collected on CDCE and reamplified to create homoduplexes. Homoduplexes were visualized on high-resolution CDCE as described above. Mutant fractions in the resulting homoduplexes were compared to an internal standard to determine whether each signal was significantly different depending upon the primer used for preamplification.

4. RESULTS

4.1 MUTATIONAL SPECTRA METHODOLOGY

4.1.1 Overview

Our approach to measuring mutational spectra is based on sequential enrichment of mutants that lie within a 100 bp target sequence (see Fig. 8 in Literature Review for melting map of target sequence). The target sequence comprises the low melting domain of a 200 bp DNA fragment with a biphasic melting profile suitable for separation of point mutants by partially denaturing gel electrophoresis, i.e., a low melting domain adjacent to a melting domain requiring a higher temperature for denaturation. Mutations in the low melting domain of such a sequence affect the rapid equilibrium between the partially-denatured and double stranded forms of the molecule that coexist in a given temperature range, and therefore alter the mobility of the DNA through a gel matrix. The enrichment of mutants is based on separation of all of the mutant sequences present in a sample from the large excess of wild-type sequences.

The mutational spectra methodology involves two separation methods, both of which are based on the principles of cooperative melting equilibria (Lerman et al., 1987): slab gel (CDGE) (Hovig et al., 1991) and capillary (CDCE) (Khrapko et al., 1994b) (see Figure 16 for an overview of the procedure). Enrichment of mutants prior to subjecting the DNA to PCR is essential to avoid PCR noise generated by amplifying the wild-type from obscuring visualization of low-frequency mutants. CDGE was used as the pre-PCR enrichment step to permit separations of samples containing μg amounts of total DNA. Such samples cannot be conveniently separated on capillaries with a 75 μm inner diameter. CDCE, which permits rapid and precise collection of mutants, was used for post-PCR mutant enrichment, at which point the total DNA loaded was negligible.

Prior to CDGE separation, DNA was digested with a pair of restriction endonucleases that excise the 200 bp target fragment. The samples were doped with small known fractions of particular mutants which served as internal standards. The internal standards permit monitoring of the enrichment efficiency at each step, and determination of the absolute mutant fractions of mutants in a sample. The samples were run on CDGE, and mutant-enriched DNA fractions located above and below the wild-type were eluted. Enriched fractions were subjected to high-fidelity PCR which eliminated interference from non-target DNA and also added a fluorescent label to the target sequence.

After PCR, the mutants were further enriched by CDCE separation. The pooled fraction containing all of the mutants, but not the wild-type DNA peak, was collected from the capillary and further amplified by PCR. Enriched samples in which mutants were clearly visible were run on high-resolution CDCE capable of separating the mutants from each other and displaying them as individual peaks suitable for identification and isolation for sequencing.

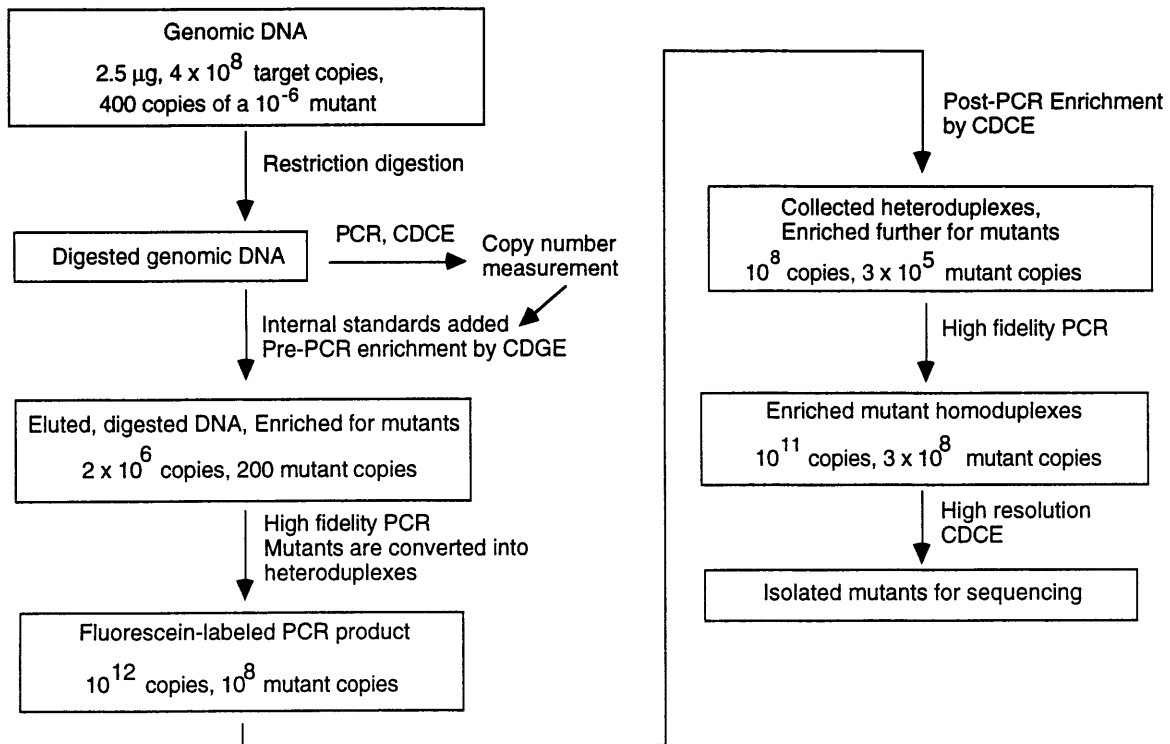
4.1.2 Copy number measurement and internal standard

It is critical for measurement of mutant fractions that the number of copies of target sequence in each DNA sample is measured accurately. Since measurements of mutant fractions are based on comparisons with internal standards, knowledge of the total number of initial copies of the target sequence is essential for accurately introducing internal standards at the desired initial fraction. Further, the number of mutant copies in the sample must be known to ensure that it is sufficient for the desired statistical power.

To measure the initial number of copies available for PCR in a DNA sample, the sample was doped with a PCR amplified target sequence containing an AT->GC mutation at bp 10072 (p13, see Figure 27 in Results) to serve as an internal standard. The number of mutant copies in the stock dilutions of internal standard was inferred from the amount of

Fig. 16: Overview of the procedure.

Flow diagram of the sample handling steps showing the total number of target copies and the number of copies of a mutant initially present in the sample at a fraction of 10^{-6} at each step.



primer used in PCR and observation on CDCE that the primer was completely exhausted. The sample DNA and a known amount of internal standard were subjected to PCR followed by CDCE. The areas of the wild-type and mutant peaks were determined. The ratio of these areas was used to calculate the initial number of copies.

Critical to this calculation is the demonstration that wild-type and internal standard amplify with the same efficiency. Comparison of the p13 to wild-type ratio before and after 20 doublings revealed that the amplification efficiency of the internal standard was equal to that of the wild-type (0.6 per cycle) within $\pm 1\%$.

This method for measuring copy number is robust; the sample can be doped at any ratio of mutant to wild-type from 0.01 to 100 with good results. It has also shown good reproducibility. Estimated copy numbers from independent copy number experiments were generally within a factor of two of each other.

4.1.3 Pre-PCR enrichment by CDGE

Prior to enrichment for mutant sequences, samples are represented by about 2.5 μg restriction digested cellular DNA containing approximately 4×10^8 copies of the unlabeled target sequence. The small fraction of target sequences that contain mutations must be separated from the vast excess of wild-type sequences and non-target DNA. This task could be simplified by PCR amplification of the sample followed by CDCE separation. However, to achieve low backgrounds and, hence, low detection limits, it is important to significantly enrich for mutants before PCR is applied to the sample (see Figure 20 in Reconstruction experiments for an illustration of the improvement gained by CDGE).

Slab gel CDGE separation (Hovig et al., 1991) was used for the necessary initial enrichment. Under the partially denaturing conditions used for separation, high T_m mutants and low T_m mutants were excised from the portions of the gel below and above the wild-type band, respectively. CDGE rather than CDCE was used at this stage because the large amount of non-target DNA and impurities in the samples create a high-resistance zone in the capillary which prevents separation. An alternative approach has since been developed, in which CDGE is replaced by CDCE in wide-bore capillaries (Li et al., 1996).

The efficiency of CDGE enrichment of mutants was estimated using a sample doped with a high T_m and a low T_m mutant as internal standards at fractions of 10^{-2} each. Mutants at this fraction can be easily and reliably observed on CDCE after PCR. (The samples used for mutational analysis are usually doped at 10^{-5} - 10^{-4} , fractions that are too low to be observed without further enrichment). The ratio of mutant fractions after CDGE to the initial fraction represents the enrichment efficiency. The enrichment was typically 100-200 for high T_m mutants and 5-10 for low T_m mutants.

Elution from the CDGE gel slices and the first PCR represents the statistical "bottleneck" in the procedure since the number of mutant copies is at a minimum. Any mutant to be measured with adequate precision (that is, not worse than $\pm 20\%$, 95% confidence limits) should be represented by at least 100 copies at this point in the procedure. Thus, if one began with 4×10^8 mitochondrial DNA copies, a mutant with a mutant fraction of 10^{-6} would be present as 400 copies. Losses in DNA isolation and enrichment steps would reduce this number to about 200, and an estimated efficiency of the first cycle of PCR of 0.5 would mean the bottleneck number would be 100 mutant copies, the minimum necessary to give the desired lower limit of statistical uncertainty.

4.1.4 High-fidelity PCR

After enrichment by CDGE, the samples were subjected to PCR, which generated 10^{12} copies of the target sequence, reduced the relative amount of non-target DNA to an

insignificant level and provided for fluorescent labeling of the target. In essence, PCR made the samples suitable for CDCE separations.

PCR introduces "noise" into the analysis. It is important to amplify with a DNA polymerase of sufficiently high fidelity so that errors created by copying excess wild-type strands do not interfere with observation of mutants in the samples. *Pfu* DNA polymerase was used in this work because it is thermostable and was reported to have a high fidelity (Lundberg et al., 1991). P. André and colleagues determined the fidelity of *Pfu* on this target sequence, and discovered an unusually low error rate, 6.5×10^{-7} errors per base pair doubling (André et al., 1997). This is significantly lower than other DNA polymerases screened for fidelity (Keohavong et al., 1989).

Pfu does, however, share with other DNA polymerases the ability to convert wild-type sequences into byproducts that have lower melting temperatures than the wild-type homoduplex. These byproducts can interfere with the mutant enrichment step on CDCE. Among these byproducts are incomplete or exonucleolytically processed products missing from one to several nucleotides. The fractions of these byproducts can be as high as 50% of the total in some PCR preparations. The problem has been reduced but not eliminated by use of an increased *Pfu* concentration (four-fold higher than the manufacturer's suggestion), and a 30 minute post-PCR incubation at 45°C. Under these conditions, the CDCE enrichment is approximately 30-fold, limited by collection of 3% of the wild-type as tail on the main peak and low melting temperature wild-type byproducts.

During PCR, the mutants were converted into heteroduplexes with the excess of wild-type strands by continuing to subject the sample to temperature cycles even after the point where the primers were exhausted. All such heteroduplexes have significantly lower melting temperatures than the parent wild-type homoduplex, a fact which allows collection of all of the heteroduplexes as a single fraction greatly reduced in the number of wild-type strands by DGGE, CDGE, or CDCE.

4.1.5 Post-PCR enrichment by CDCE

A post-PCR CDCE separation is shown in Fig. 17. A sample of TK6 cellular DNA from cells grown 400 generations is shown. The sample contains two prominent low T_m mutants with initial fractions of about 4×10^{-4} and 3×10^{-4} , in addition to the internal standard at 10^{-4} , and was selected for demonstration purposes. In an initial CDCE (Figure 17A), the four heteroduplexes of the two mutants (marked by X's) are about 0.2% of the wild-type peak (which implies 5-fold CDGE enrichment) and are hardly distinguishable among the wild-type variants (noise) which are present at a similar fraction (marked by asterisks).

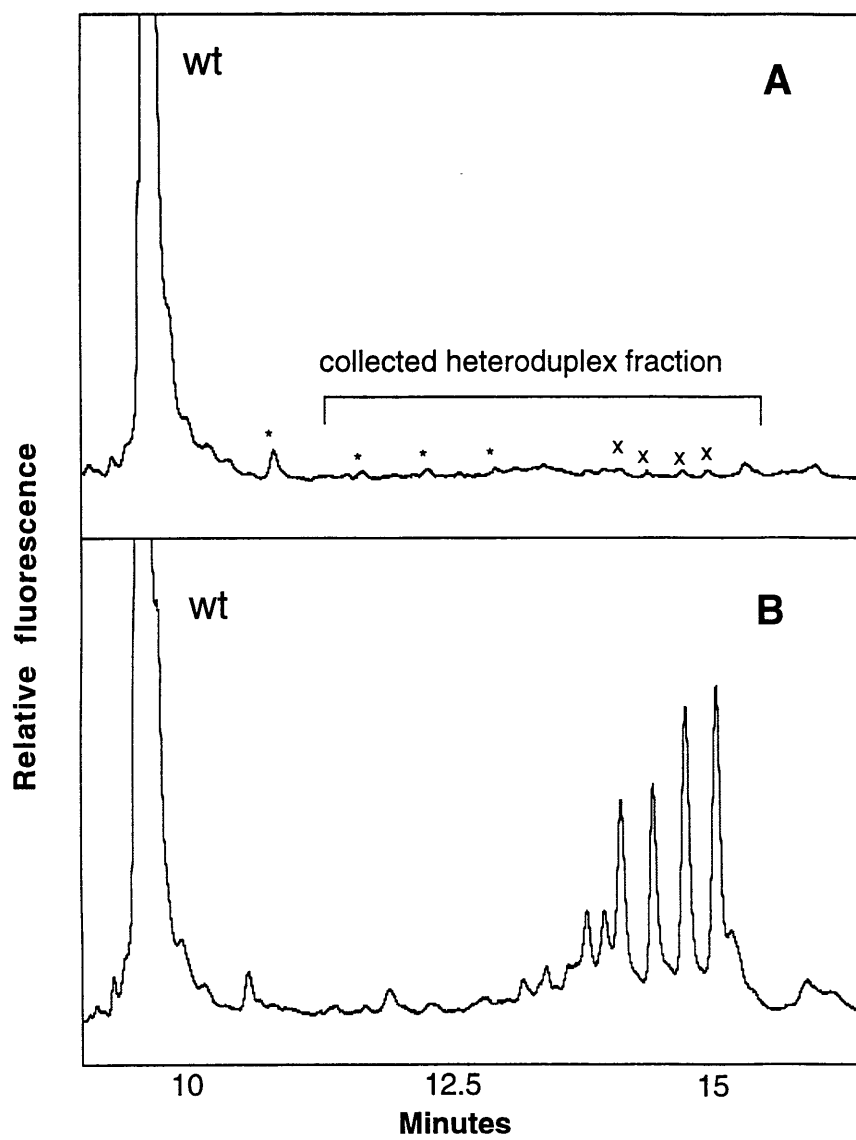
The heteroduplex fraction of the first CDCE was collected, amplified and run on a second CDCE (Fig. 17B). On the second CDCE, the four heteroduplex peaks of the two major mutants are clearly visible above the background and constitute more than 5% of the wild-type peak each, which implies a more than 25-fold enrichment by CDCE. In addition to these prominent peaks, smaller mutant peaks that were not visible on the first CDCE are visible on the second CDCE.

The combined size of all mutant heteroduplex peaks in the separation shown in Fig. 17B is comparable to the size of the wild-type peak. In this situation, another collection of the heteroduplex fraction will not significantly enrich against the wild-type because each heteroduplex contains one wild-type strand. If the initial frequency of mutants in the sample were lower, as is usually the case, the combined size of heteroduplexes at this stage would be smaller and additional CDCE separations would provide further enrichment.

Fig. 17: Post-PCR enrichment of mutants by CDCE.

A. TK6 cellular DNA eluted from the low T_m region after CDGE was subjected to 35 cycles of PCR and separated by CDCE. Mutant heteroduplex peaks marked by "x" are not clearly visible above background peaks marked by "*". The x axis reflects the time since the beginning of the run at which the peak reaches the detector and the y axis marks the relative intensity of fluorescence.

B. The heteroduplex fraction of the first CDCE shown above was collected, subjected to 35 additional cycles of PCR and run on a second CDCE. A set of mutant heteroduplex peaks is clearly distinguishable from the background.



4.1.6 High-resolution CDCE

Once the mutants have been enriched, individual mutants are then separated, measured, isolated and sequenced. These procedures are most easily performed with the mutants in homoduplex form. Mutant heteroduplexes were converted into homoduplexes by stopping PCR when the molar amount of unused primers still exceeded that of the products. Homoduplexes were then separated via CDCE under "high-resolution separation conditions" that include increasing the time the sample spends in the heated zone. This was achieved by increasing the length of the zone and decreasing the running voltage. High-resolution conditions also involved increasing the salt concentration in the electrophoretic buffer. The improvement in resolution with these modifications results from increasing the average number of partial meltings and reannealings a molecule undergoes while in the heated zone of the capillary (Khrapko et al., 1996).

Fig. 18 (upper curve) shows a high resolution separation of the same sample as the one shown in Fig. 17 after it was converted to homoduplex form. The reason that the two separations look dissimilar is that each pair of heteroduplexes which in Fig. 17 were represented as two peaks yielded a single homoduplex peak in Fig. 18.

4.1.7 Identification of mutants

In our first studies of mitochondrial hotspot mutants, each mutant peak was isolated, amplified and sequenced. It became clear soon thereafter that the same mutants appeared in multiple samples and that a more efficient mode of mutant identification was possible. All sequenced mutants were labeled with tetramethylrhodamine (TMR) and combined together to form a "standard set" of markers. This TMR-labeled "standard set" could then be coinjected with an aliquot of fluorescein-labeled sample under study and both signals could be observed on a single CDCE separation by means of a two-wavelength detector. In this way, mutants in the fluorescein-labeled sample could be identified based on comigration with the TMR-labeled previously isolated hotspot mutants. Further, any novel mutants could be identified for subsequent isolation and sequencing. The fluorescein and TMR labeled homoduplexes are shown for TK6 cells grown 400 generations in Fig. 18.

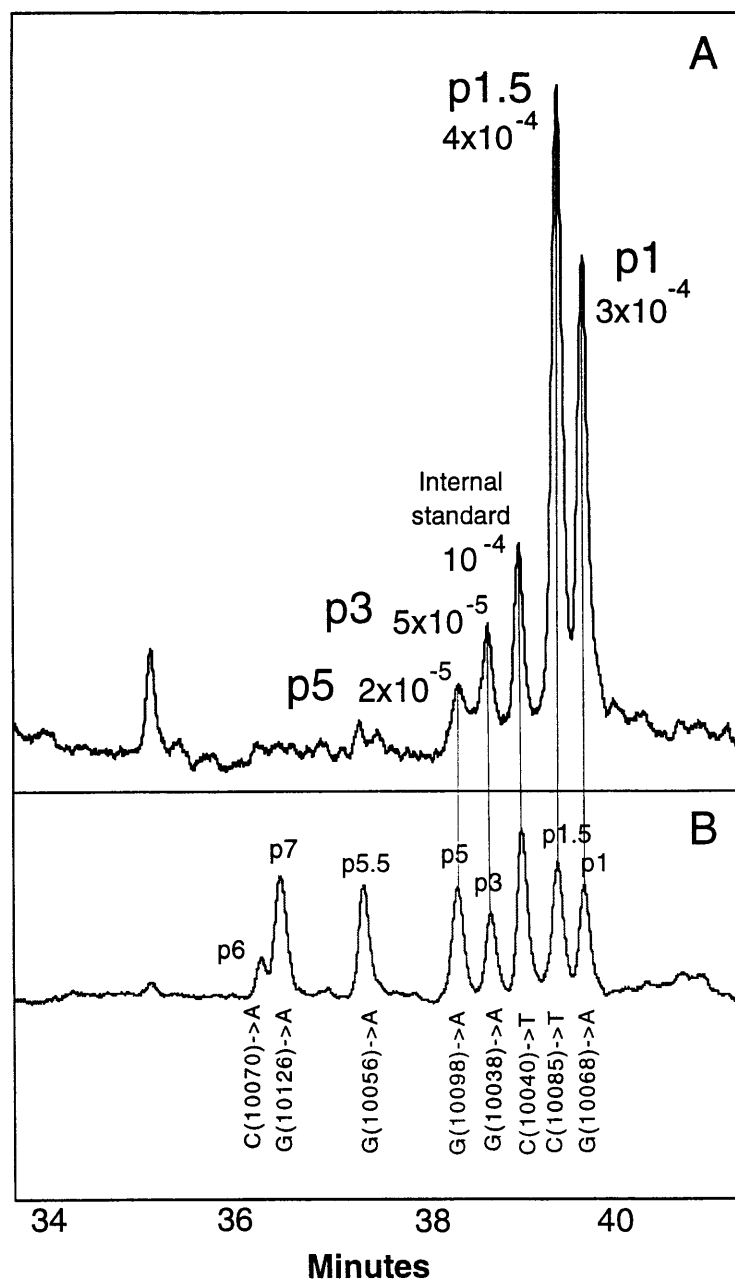
Development of a dual-dye detection method significantly facilitated mutant identification. The results supported the preliminary conclusion that the same specific mutants appear repeatedly in all samples tested. However, comigration is not proof of identity. Two distinct mutants could theoretically comigrate under a specific set of conditions. In order to test more rigorously the identity of specific peaks, an identification procedure based on capillary hybridization of sample peaks with the panel of previously isolated mutants was developed. The CDCE instrument used for identification by hybridization included two consecutive heating jackets. The sample and the standards were co-separated in the first jacket as performed for identification by co-migration, and the resulting set of peaks was stopped inside the second jacket. The temperature in the second jacket was increased to denature the DNA, then decreased to permit reannealing of single strands, and finally the sample was electrophoresed through the rest of the second jacket to the detector.

If a sample peak were identical to the standard with which it comigrated, then its reannealing to homoduplexes was forced by the high concentration of the "driver" in the reaction, that is, the standard mutants labeled with TMR. In this case, the position of the fluorescein-labeled peak with respect to the TMR-labeled homoduplexes, which reanneal efficiently due to their high concentration, is retained. In contrast, if the peak is different from the standard mutant with which it co-migrates during the pre-hybridization part of the CDCE separation, the hybridization procedure will convert the sample mutant into two heteroduplexes with the standard mutant present in excess. Such heteroduplexes are always less stable than the homoduplexes from which they are derived, and move to a

Fig. 18: Mutant identification by comigration with authentic standards.

A sample of fluorescein-labeled low T_m mutants isolated from TK6 cells in the form of homoduplexes (A) was coinjected with a tetramethylrhodamine (TMR)-labeled set of known low T_m mutants (B). Standard mutants identified as p1, p1.5, A (internal standard), p3, p5, p5.5 and p7 represent GC→AT transitions at base pairs 10068, 10085, 10040, 10038, 10098, 10056 and 10126, respectively. p6 represents a GC→TA transversion at base pair 10070. Based on a comparison of the mobilities, mutants p1, p1.5 and p3 were determined to be present in the sample. Comparison with the internal standard added to the sample at a mutant fraction of 10^{-4} permits assignment of mutant fractions of other mutants.

Relative Fluorescence



different position during the post-hybridization portion of CDCE separation. This results in the "disappearance" of the sample peak from its original position in the spectrum.

Capillary hybridization was performed for high T_m mutants from 16 twins' bronchoscopy samples, 3 dissected lung samples, TK6, muscle and colon samples. The specific mutants tested by capillary hybridization were 1, 3, 5, 5.5, 6, 7, 11, 11.3, 11.5, 13, 13.5, 14, 15.5, 16, 18, 18.5, and 19. Approximately 90% of peaks identified by comigration were confirmed as present by capillary hybridization. In most cases, the mutant fractions estimated by measurement of homoduplex runs and capillary hybridization were very similar (within 2- to 3-fold). In cases where a significant difference was observed, the mutant was often difficult to measure in the homoduplex run due to the presence of other closely-spaced peaks.

An example of the mutational spectra observed before and after capillary hybridization is shown in Fig. 19. In the top panel, the high T_m mutational spectra of a sample biopsied from a nonsmoking twin's left lower lobe is shown. The lower curve of the top panel represents several TMR-labeled mutants that were electrophoresed simultaneously with the sample. In the lower panel of Fig. 19, the same sample is shown after capillary hybridization. Peaks that remain are identical to the standard with which they comigrate, and are numbered. All sample peaks that were not identical to one of the TMR-labeled mutants injected in this run were converted to single stranded DNA or heteroduplexes. Such species were retarded during migration through the second water jacket and hence passed the detector later, after the hybridized homoduplexes. These species consequently "disappeared" from the spectrum. As an example, the peak between mutant 11 and mutant 11.5 labeled "x" in Fig. 19 was not identical to any of the TMR-labeled standards and was consequently not observed in the spectrum after capillary hybridization. This particular peak represents mutant 11.3 which was not included in the standards for this particular hybridization run.

4.1.8 Quantification of mutants.

Peak measurement is illustrated in Fig. 18. The area under each peak is proportional to the initial mutant fraction in this sample. The initial mutant fraction of mutant peaks p3, p1.5 and p1 can be determined by comparison of the areas under these peaks to the peak representing the 10^{-4} internal standard. In this sample, the original fractions of mutants represented by p3, p1.5 and p1 are estimated to be 5×10^{-5} , 4×10^{-4} and 3×10^{-4} , respectively.

4.2 TESTS FOR VALIDITY

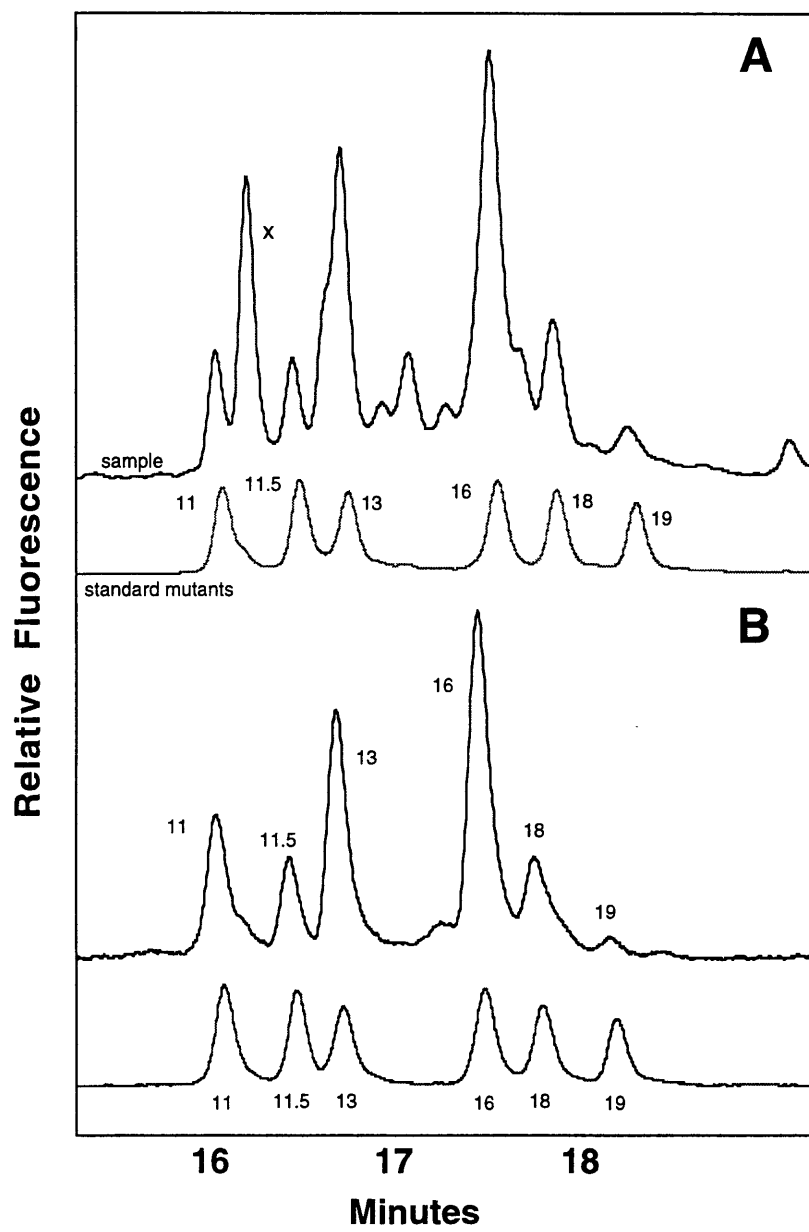
4.2.1 Reconstruction experiments

The mutants that we detect with this method are represented in our initial samples by as few as four hundred molecules. Consequently, an obvious problem is the potential for contamination from purified PCR preparations of these mutant sequences. In order to minimize contamination, DNA isolation, CDGE, and preparing aliquots for PCR were performed in a clean lab equipped with high-throughput air filters. No PCR products of purified mutant sequences entered this laboratory except for very dilute stocks used to introduce internal standards.

In order to determine whether any of the mutant peaks observed were caused by contamination, *Pfu* polymerization errors or by replicative bypass of DNA adducts created during our procedures, CDGE-purified wild-type homoduplex was doped with known amounts of internal standards containing a single point mutation and subjected to all of the steps of our procedure. Shown in Fig. 20 are the results when wild-type PCR fragments

Fig. 19: Capillary hybridization to confirm peak identification.

The high T_m homoduplex mutational spectra was determined for the left lower lobe bronchoscopy sample of a nonsmoking twin. (A) The set of homoduplexes observed as the mutational spectra, and several TMR-labeled mutants from the standard set, specifically 11, 11.5, 13, 16, 18 and 19, that were used for capillary hybridization. (B) The same sample after capillary hybridization. Individual peaks that are still present in the sample are considered to be confirmed as representing specific mutants and are numbered. The peak designated with an "x" is peak 11.3. This mutant disappeared after hybridization because there was no excess 11.3 TMR-labeled mutant with which it could reanneal.



purified from possible mutants by CDGE were doped with a low T_m mutant at 10^{-4} and with a high T_m mutant at 10^{-5} , mixed with non-amplifiable herring carrier DNA, and subjected to all the procedures described above except for the initial DNA isolation. Importantly, no prominent mutants were observed in this control sample besides the internal standards. In this experiment, the peaks representing both of the reference mutants are about ten times larger than the largest peaks of the background. Thus, the limit of detection, defined as the frequency of a mutant that yields a peak comparable to the largest background peaks, at least for purified wild-type DNA, should be considered 10^{-5} for low T_m mutants and 10^{-6} for high T_m mutants. In a variation on this control, to test whether detectable levels of DNA adducts leading to noise "mutants" were generated during cell lysis, the target human mitochondrial fragment was mixed with rat cells and subjected to the entire procedure. No interfering signals were observed in this control experiment either.

Since the limit of detection depends on the size of background peaks it is important to know the sources of background mutants. Most of the background peaks do not co-migrate with mutants originating from *Pfu* replication errors (André et al., 1997) as described further subsequently. The most likely source of the background may be adducts present in mitochondrial DNA that were converted to mutations during *in vitro* PCR. As can be seen from a comparison of Fig. 20A and 20B, CDGE substantially eliminates some of this background. Panel B displays the results from the same experiments as A without CDGE enrichment. The difference in detection limits between high and low T_m mutants could reflect that more of these lesions co-migrate with low T_m mutants than with high T_m mutants.

In order to further investigate the possibility that adducts of mitochondrial DNA contribute to background mutations in the mutational spectra procedure, the amount of enrichment of a low T_m artificial mutant PCR product spiked into genomic DNA during CDGE was determined for DNA isolated from various sources. Mutant PCR product added to TK6 genomic DNA was enriched more efficiently in two independent experiments than PCR product added to genomic DNA from tissue samples. Further, when artificial mutant was added to genomic DNA isolated from lung tissue of an 84 year old man, the enrichment was particularly poor. In one experiment, the enrichment for TK6 cells was 14-fold; for lung parenchyma, 6.4-fold and for lung epithelial cells from the 84 year old man, 1.8-fold. No difference was observed, however, in the extent of enrichment between smoking and nonsmoking twins. These findings suggest that modifications of mitochondrial DNA that lead to altered mobility on a CDGE may represent important contributors to the background in our mutational spectra determination of low T_m mutants. Further, these results suggest that there is modification of mitochondrial DNA at a level such that several percent of the copies contain a modification somewhere in the 100 bp low melting target. Finally, these experiments suggest that the extent of modification may differ for cultured cells, human tissues and tissues of the elderly. These observations are supported by measurements of 8-OH-dG in mitochondrial DNA from cell and tissue samples described in the Literature Review (Intracellular milieu).

4.2.2 Spectra differ from each other

The limit of detection available with the mitochondrial mutational spectra methodology was shown to be sufficient to observe mitochondrial hotspot mutations in human tissue samples. Fig. 21 provides high resolution CDCE runs of the following samples: (A) CDGE-purified wild-type DNA with internal standards added; (B) a sample derived from a culture of TK6 cells grown 400 generations; and (C) a sample of human bronchial epithelial cells. The cell and tissue samples in panels B and C contain a series of

Fig. 20: Limit of detection: Reconstruction experiment.

A: A sample of wild-type PCR fragment that was depleted of mutants by CDGE was mixed with nonamplifiable carrier DNA. The sample was spiked with a high T_m mutant (10^{-5}) and a low T_m mutant (10^{-4}) as internal standards, and subjected to the mutant isolation procedure (panel A, "CDGE+"). The left panel includes mutants with melting temperatures higher than wild type and the right panel depicts mutants with melting temperatures lower than wild type.

B: To demonstrate the role of pre-PCR CDGE, the same experiment was carried out with the CDGE enrichment omitted (panel B, "CDGE-").

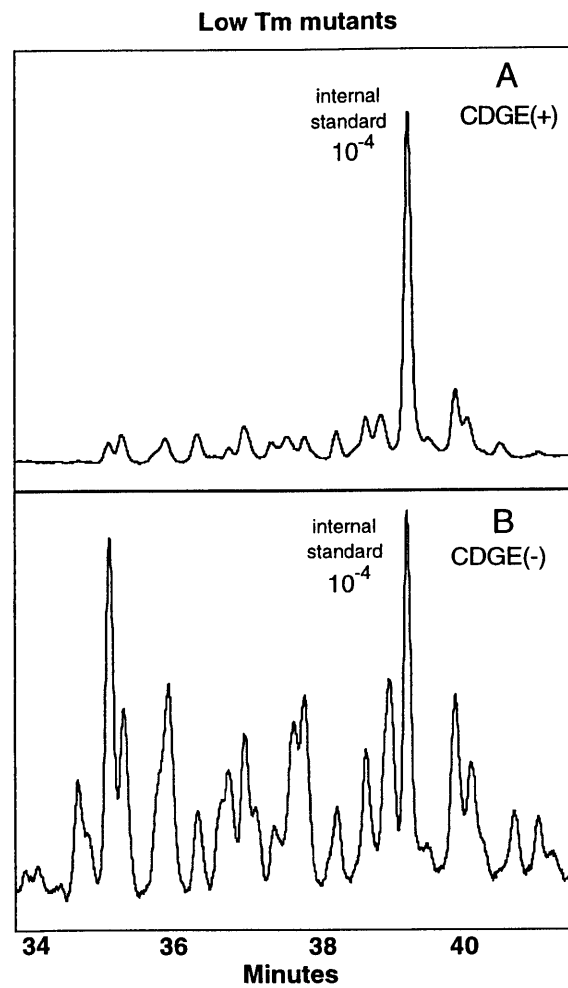
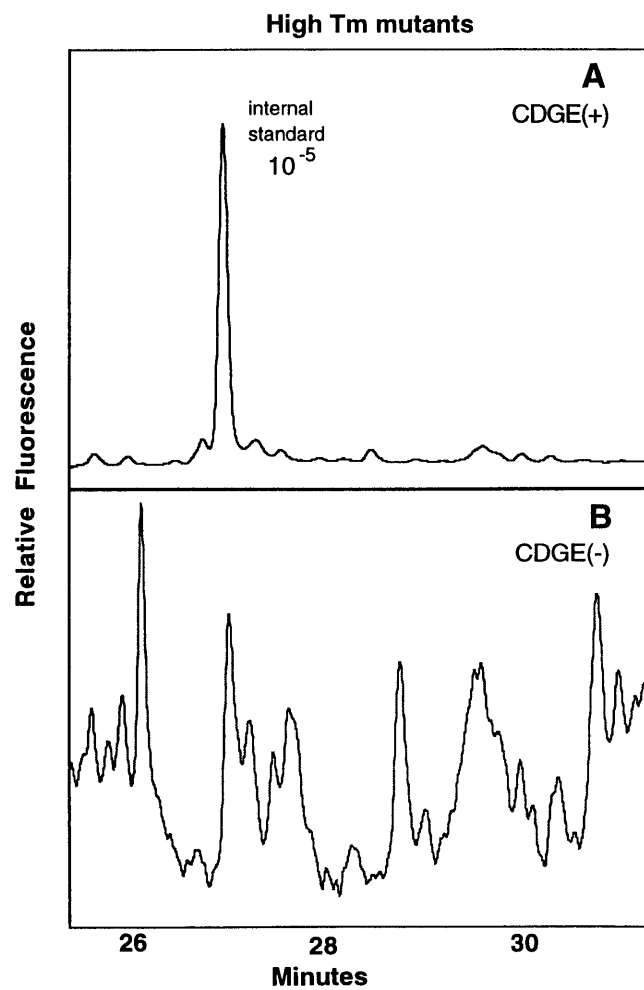
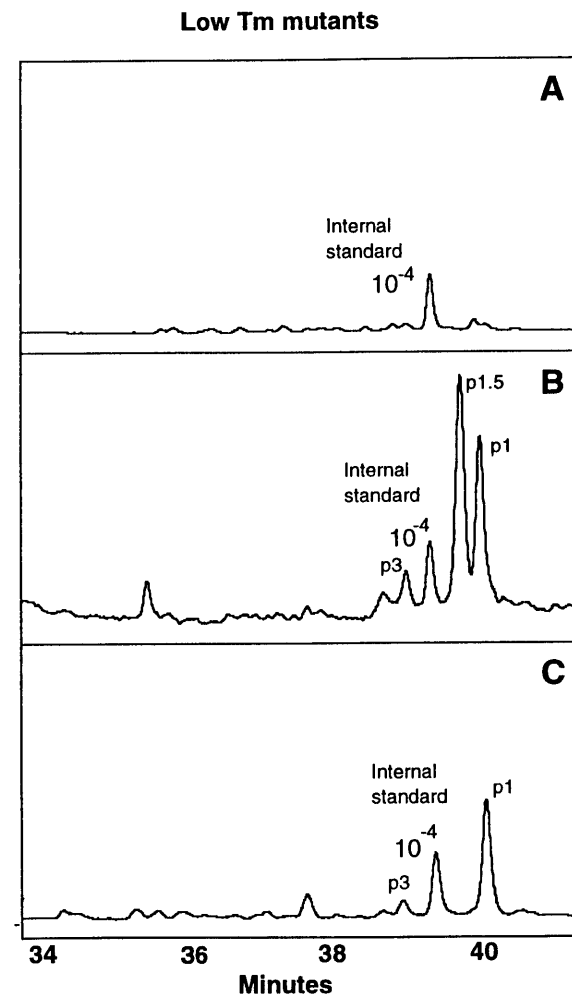
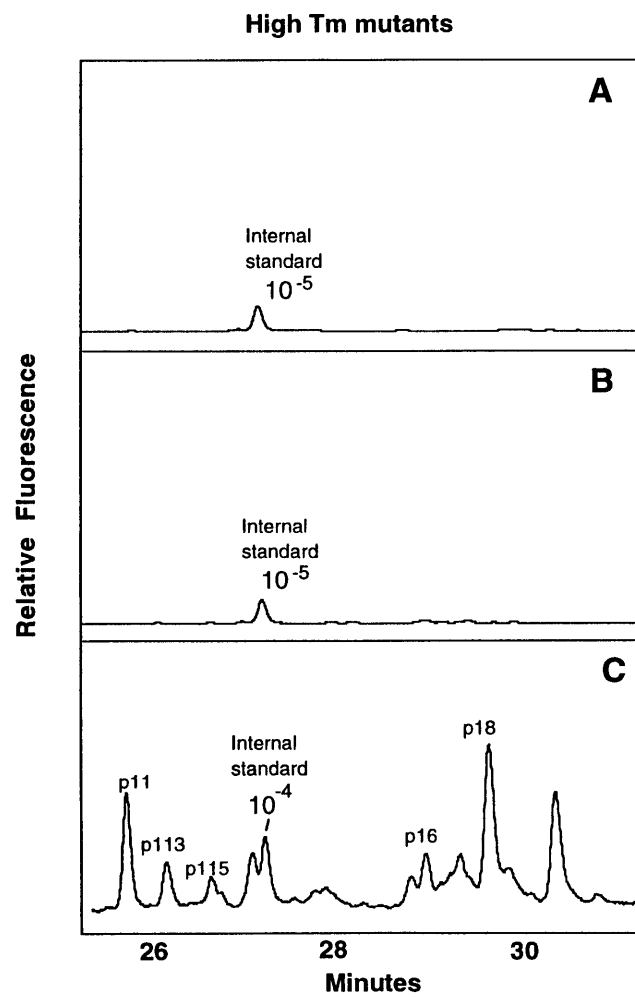


Fig. 21: Examples of mitochondrial mutational spectra.

A: Negative control. Sonicated herring DNA was mixed with CDGE-purified wild-type and internal standards and subjected to mutational spectra analysis.

B: A spectrum derived from a culture of TK6 human lymphoblastoid cells.

C: A spectrum of a human lung epithelium sample. For sequences of mutants, see Figure 27.



peaks that are not present in the negative control in panel A. The fractions of high T_m mutants in the TK6 sample (Fig. 21B) are on the order of 10^{-6} . The low T_m mutants are as high as 4×10^{-4} . The bronchial epithelium sample (Fig. 21C) displays significant peaks in both the high and low T_m sets. In the low T_m set, the bronchial tissue displays some of the same hotspot mutants as are found in the TK6 cells (p1 and p3). In the high T_m region of the bronchial epithelium sample, several peaks are seen with mutant fractions on the order of 10^{-4} . Five peaks are identified by number (see Figure 27 in Results: Hotspots, for sequence changes).

Both the cell and tissue samples contain hotspot mutants detectable significantly above the background generated by the procedure itself. Further the pattern observed in DNA isolated from TK6 cells was substantially different from the pattern observed from this bronchial epithelial cell sample, a finding that supports the conclusion that the mutations observed do not result from any of the procedures used, as this process would be expected to induce the same type and frequency of mutants in all samples.

4.2.3 Reproducibility

The reproducibility of the procedure is demonstrated in Figure 22. A single DNA sample taken from TK6 cells was split in two, and internal standards were added separately to each aliquot. Duplicate samples were analyzed independently on a constant denaturing gradient gel, followed by further enrichment via CDCE and visualization of mutants. As can be observed, the reproducibility is excellent. Samples were routinely analyzed in duplicate and the reproducibility was generally very good. When mutant fractions are measured, the ratio of mutant fractions between duplicates is generally less than two-fold. This is in sharp contrast with the much more large difference in the mutant fraction of particular peaks between samples, where the mutant fractions can vary by 30-fold or more, as discussed in Results: Variability of Mutant Fractions.

4.2.4 *Pfu* noise characterization

An important potential source of error in these procedures is to misinterpret mutants created by *Pfu* DNA polymerase *in vitro* for mutants actually present in tissue samples. In order to avoid this possibility, P. André and colleagues characterized the PCR noise induced by *Pfu* DNA polymerase in this DNA sequence (André et al., 1997). Using successive rounds of PCR to generate *Pfu*-induced mutants and CDCE to enrich for the mutants, the error rate of *Pfu* was estimated as 6.5×10^{-7} errors/bp-doubling. Five of the most frequent hot spot mutations were three transversions (GC→TA at bp 10070, mutant E; AT→TA at bp 10071, mutant D; and AT→CG at bp 10071, mutant A), one transition (AT→GC at bp 10108, mutant B) and a one base pair deletion in a run of six consecutive adenines (bps 10048-10053, mutant C) (see Figure 23). Of the peaks in the "standard set" that were consistently observed in human tissues, only one (p6, mutant E) was found to arise from the PCR process itself. The others are unlikely to originate from PCR noise.

4.2.5 Test for asymmetric species

The tests for the validity of mutational spectra results discussed to this point are concerned with systemic bias in which mutants may have been created as a result of artifacts generated by our procedures themselves. Such errors include contamination of samples and PCR-created mutants. PCR errors could arise as misincorporation events during primer extension over a wild-type substrate or as a result of misinsertion during replicative bypass of DNA adducts formed in our procedures. The errors of this class, as

Fig. 22: Reproducibility of the mutational spectra procedure.

Restriction digested DNA from TK6 cells containing the equivalent of approximately 3×10^8 copies of the mitochondrial target sequence was divided into two samples and each was doped at 10^{-4} with the artificial mutant to serve as an internal standard. Panels A and B are the results of CDCE separations of the duplicate samples.

Low Tm mutants

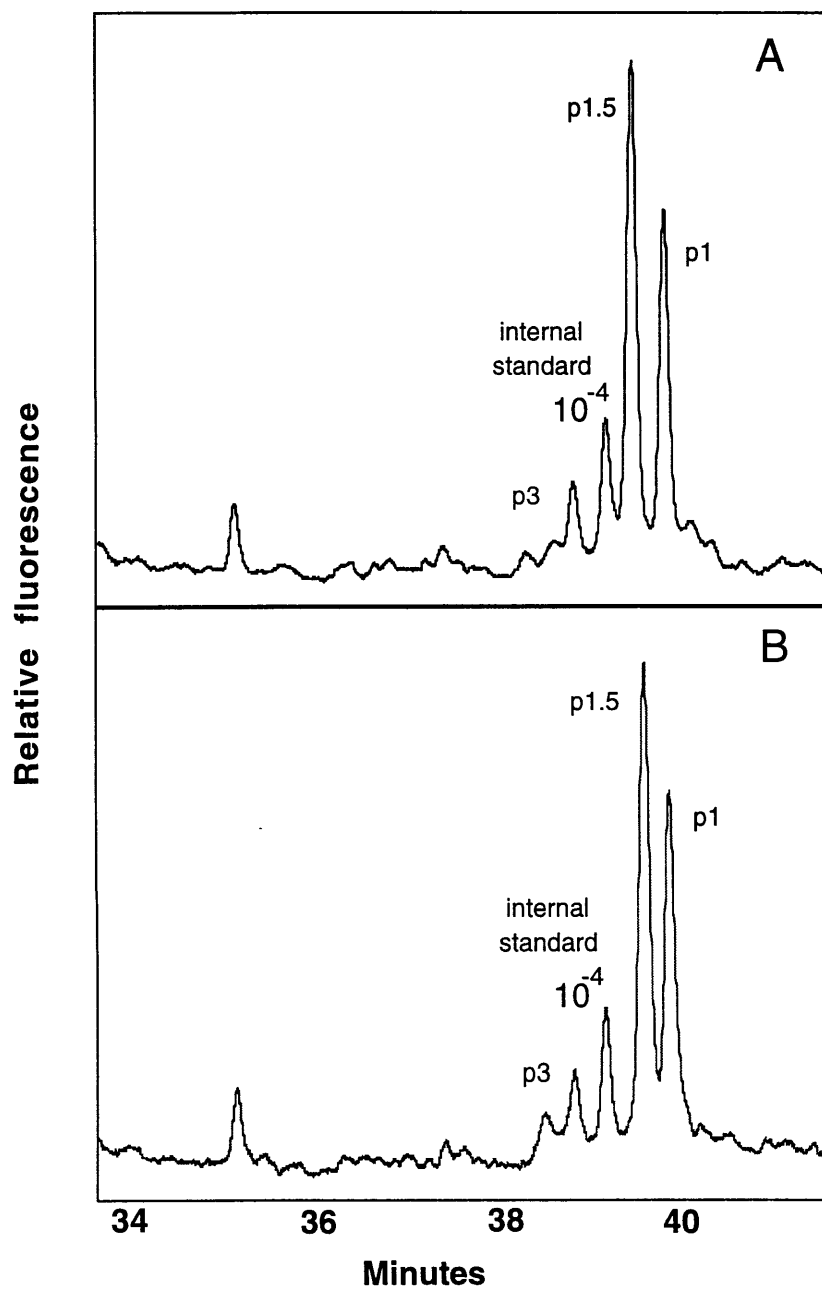
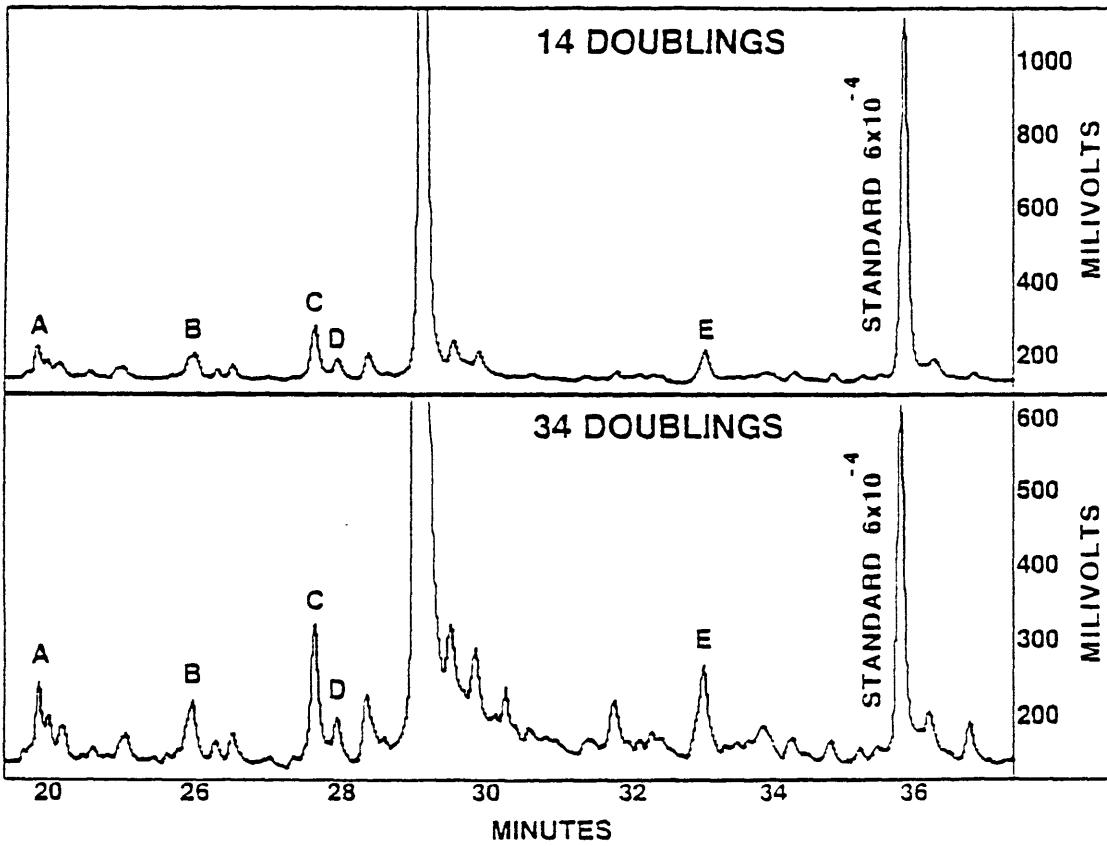


Fig. 23: Characterization of *Pfu*-induced noise on mitochondrial target sequence.

Analysis by CDCE of the purified homoduplex mutant fractions generated by amplifying wild-type DNA for 14 and 34 doublings using *Pfu* DNA polymerase. Many small peaks visible after 14 doublings were observed as significantly larger peaks after 34 doublings. Of the major peaks for which the mutant sequences were determined, peak A is an AT->CG transversion at bp 10071; peak B is an AT->GC transition at bp 10108; peak C is a deletion occurring in a run of six consecutive adenines at bp 10048-10053; peak D is an AT->TA transversion at bp 10071; and peak E is a GC->TA transversion at bp 10070.

Source: André et al., 1997



discussed above, can be addressed by comparison of a spectrum to the negative control, such as the one shown in Fig. 21A.

A second type of error could also bias the mutational spectra we observe. The errors of this class arise when cellular pre-mutagenic sites are converted into mutants by our procedures. These would include DNA adducts in the original mitochondrial sample that were misreplicated by the polymerase during PCR to create mutants. Similarly, mismatch intermediates that were formed during DNA replication but were not yet repaired at the time of DNA isolation, could be mistaken for true heritable mutations. Note that such mismatches could exist whether or not there is mismatch repair machinery in human mitochondrial DNA. In order to discriminate cellular lesions and mismatches from true mutants, a method was developed that exploits a distinguishing property of these species: adducts or sequence changes are present in only one of the two DNA strands, while the other strand is an unadducted wild-type sequence. True mutants, in contrast, contain a changed sequence at the same base pair in both DNA strands of the original sample.

To preserve this distinguishing information, a modified PCR procedure was employed which includes asymmetric pre-amplification. Two aliquots of a CDGE-enriched sample were subjected to repeated primer elongation cycles with only one or the other of the two PCR primers. During this "linear amplification" stage, only one strand of the sample DNA was used as a template. Thus after preamplification, one aliquot contained a vast excess of "Watson" copies while the other contained excess "Crick" copies as single strands. Then the other primer was added and exponential amplification followed.

Asymmetric pre-amplification of a real double stranded mutant results in equal fractions of a particular mutant when the samples preamplified with opposite DNA strands are compared. However, in the case of a mismatch or a pre-mutagenic lesion, the mutant will arise only from copying the strand which contains the mismatch or premutagenic lesion. This results in different mutant fractions for a particular mutant between the spectra obtained from the two aliquots which begin by copying only the Watson or the Crick strand.

Fig. 24 provides the results of a test for asymmetric species as applied to low T_m mutants isolated from a lung epithelium sample. The sample was preamplified to achieve a 5.7-fold for the J3 primer (Fig. 24A), 10-fold for both primers (Fig. 24B), and 8.8-fold enrichment for the CW7 primer (Fig. 24C). The sample contains the peaks for p1 and p3. The results show that p1 yields the same mutant fraction regardless of the strand used in linear amplification. We regard this finding as strong evidence that the p1 signal arises from a true mutant in this sample. Other experiments in which TK6 cells were examined, however, have suggested that p1 may have slight ($<1/3$) adduct characteristics. The signal designated p3 shows a significantly larger mutant fraction when linear amplification was performed to copy the Watson as opposed to the Crick strand. These results suggest that a significant portion of the signal designated p3 arose either as a mismatch intermediate or as an adduct located in the Watson strand that was converted into a mutant by *Pfu* polymerase. Another mutant, that was not sequenced, designated X in Fig. 24, demonstrates the opposite distribution, which suggests that it originates in significant part from a mismatch or a lesion located on the Crick strand.

Figure 25 reveals the results when high T_m mutants from a bronchial epithelial cell sample were tested for symmetry. In this experiment, the optimal enrichment (13-fold) was obtained for CW7 (Fig. 25A) and an enrichment of 7-fold was achieved for J3 (Fig. 25C). A separate aliquot was amplified with both primers CW7 and J16 for 5 cycles, which generated an enrichment of 10-fold (Fig. 25B). The pattern was essentially constant whether the sample was preamplified with one primer, the other primer, or both primers. For each of the peaks in the high T_m standard set, the ratio of the mutant fractions in a sample preamplified with one primer versus the other was less than two. In contrast, the ratio of mutant fractions for a heteroduplex introduced prior to preamplification as an asymmetrical control was six when samples preamplified with one or the other primer were

Fig. 24: Test for asymmetric species: low T_m mutants.

Low T_m mutants from lung epithelium were eluted from a CDGE and linearly preamplified for 20 cycles either with the J3 primer (A) or the CW7 primer (C). In (B), the sample was preamplified with both primers for 5 cycles. The amplification was 8.8-fold for the CW7 primer, 5.7-fold for the J3 primer and 10-fold for both primers. Preamplified samples were then subjected to the remaining protocol for determining mutational spectra and the results are shown.

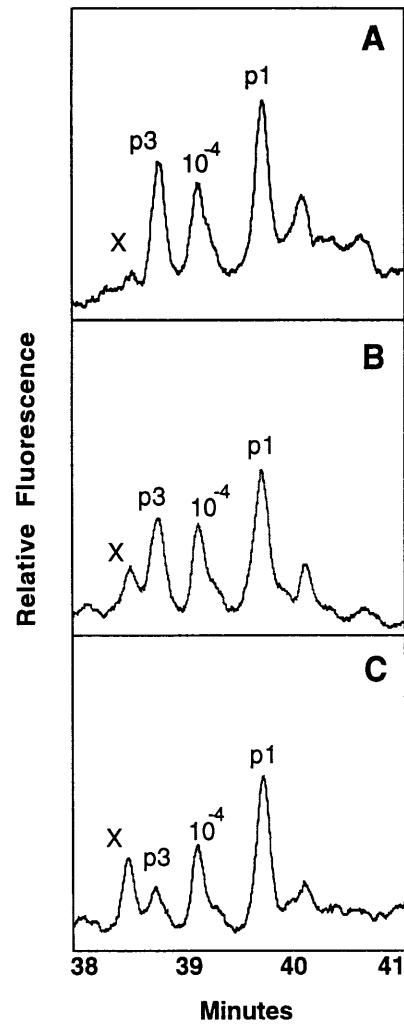
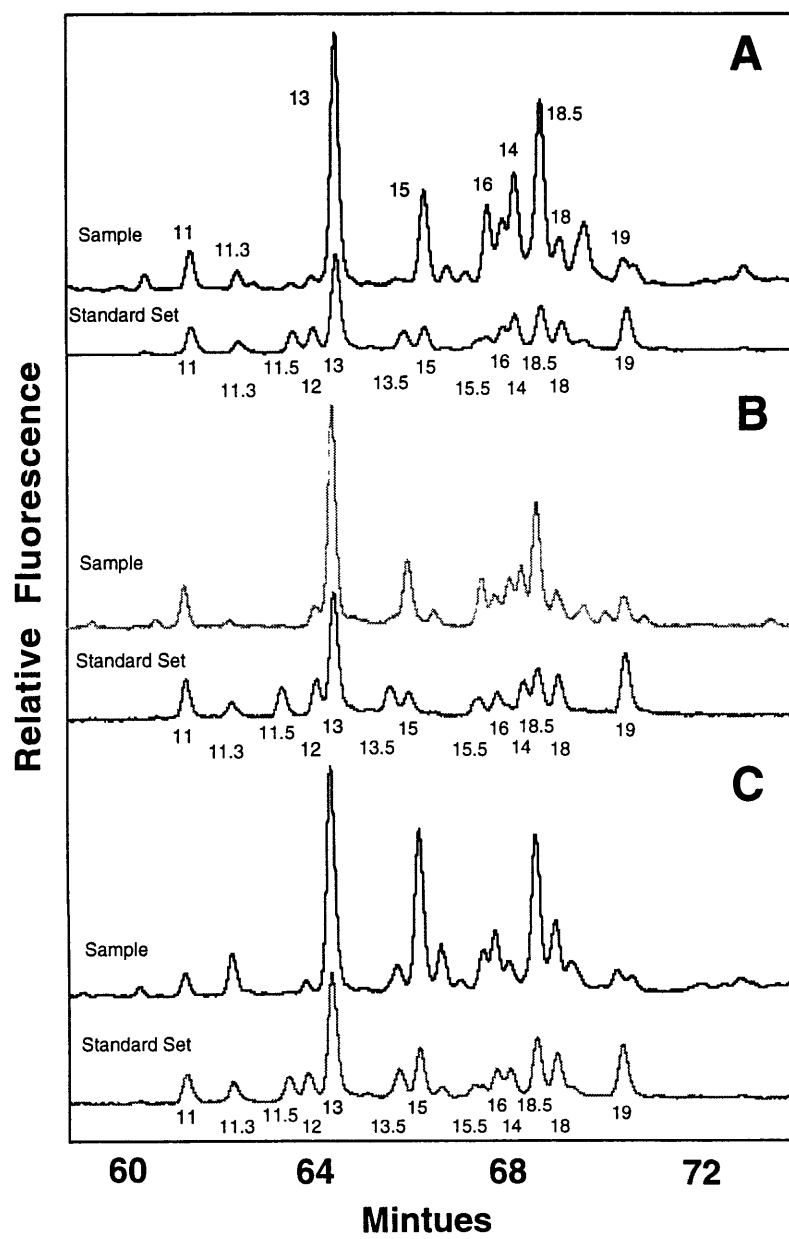


Fig. 25: Test for asymmetric species: high T_m mutants.

DNA was isolated from cells obtained via brush bronchoscopy from the right upper lobe of a twin with a smoking history. The high T_m homoduplex portion of a CDGE was eluted and subjected to linear preamplification. An enrichment of 7-fold was achieved for J3 and the optimal enrichment (13-fold) was obtained for CW7. A separate aliquot was amplified with both CW7 and J16 for 5 cycles, which generated an enrichment of 10-fold. After exponential PCR, heteroduplexes were enriched by CDCE, reamplified and converted to homoduplexes. (A) The sample after preamplification with CW7 along with coseparated standards, (B) homoduplexes for samples preamplified with both primers and coseparated standards, and (C) sample after for preamplification with primer J3 and coseparated standards.



compared. These results support the conclusion that high T_m signals represent true, double stranded mutants. Similar results concluding that high T_m signals reflect true mutants were also observed for samples from other organs besides lung.

4.3 MITOCHONDRIAL MUTATIONAL SPECTRA

The previous section describes some of the experiments performed to determine the validity of the mitochondrial mutational spectra determined by this method. In this section, the mutational spectra of a variety of human samples are provided.

4.3.1 Several organs

Figure 26 shows the homoduplex mutational spectra for four samples of normal tissues: colon, muscle and lung epithelium (2 samples). Each of the tissue samples was found to contain prominent mutational hotspots that were detectable with our procedure. Further, the same set of approximately 17 hotspots were observed repeatedly among tissue samples. These mutants were confirmed by capillary hybridization and are numbered in Figure 26. The mutant fraction of each particular mutant, however, was discovered to vary significantly among samples. Among these organs (see heart sample in Mutant Fraction and Age: Other Tissues for an exception), no specific mutant was discovered to be particularly prominent in a specific tissue. The example of the two different bronchial epithelial samples illustrates this point, as some mutants (for instance, p18), vary as much between the two bronchial epithelial samples (Fig. 26A and B) as they do between different tissues (Fig. 26A, B, C and D). The similarity in the overall level of mutations observed for all detectable point mutations in this 100 bp low melting domain region of mitochondrial DNA contrasts with the tissue specificity observed in the levels of deletions in aged individuals as discussed in the Literature Review.

4.3.2 Hotspots

Most of the mutants that are observed repeatedly in different tissue samples have been purified and sequenced. For the mutants in the standard set and the artificial mutant used as an internal standard, Figure 27 provides the specific type of mutation and the sequence context. All but two mutations are transitions, with both GC->AT and AT->GC mutations represented. The predominance of transition mutations is in concordance with studies of the types of mitochondrial point mutations observed as polymorphisms. Both GC->AT and AT->GC transitions are observed as polymorphisms of mitochondrial DNA more frequently than transversions (Aquadro et al., 1983; Horai et al., 1990). Most of the mutations in the NADH dehydrogenase protein create amino acid substitutions, but p13 and p19 are silent and p7 creates a stop codon. No differences in the number or kind of mutations present in the tRNA^{Gly} versus the NADH dehydrogenase subunit were observed. Triplet analysis of nearest neighbors did not reveal any significant deviation from chance.

4.3.3 Tumor spectra

Figure 28 provides the set of mutant peaks observed for samples derived from tumors of the colon and muscle. The same set of hotspot mutations as in normal tissues is observed in these tumors, and the mutant fractions are not significantly different from those determined for healthy tissue. Such findings suggest that the processes that generate mutations in healthy tissue also generate mutations in tumors as well.

Fig 26: Mutational spectra of lung, colon and muscle samples.

DNA was isolated from two samples of 10^6 human bronchial epithelial cells each (A and B), approximately 1 g epithelial cells from a human colon (C), and 1 g of human muscle tissue (D). DNA was restriction digested, mixed with internal standards and the mutational spectra were determined. Depicted are separations of fluorescein-labeled samples along with the co-separated TMR-labeled authentic standards (lower of each pair of curves). (X) marks a peak that has not been sequenced.

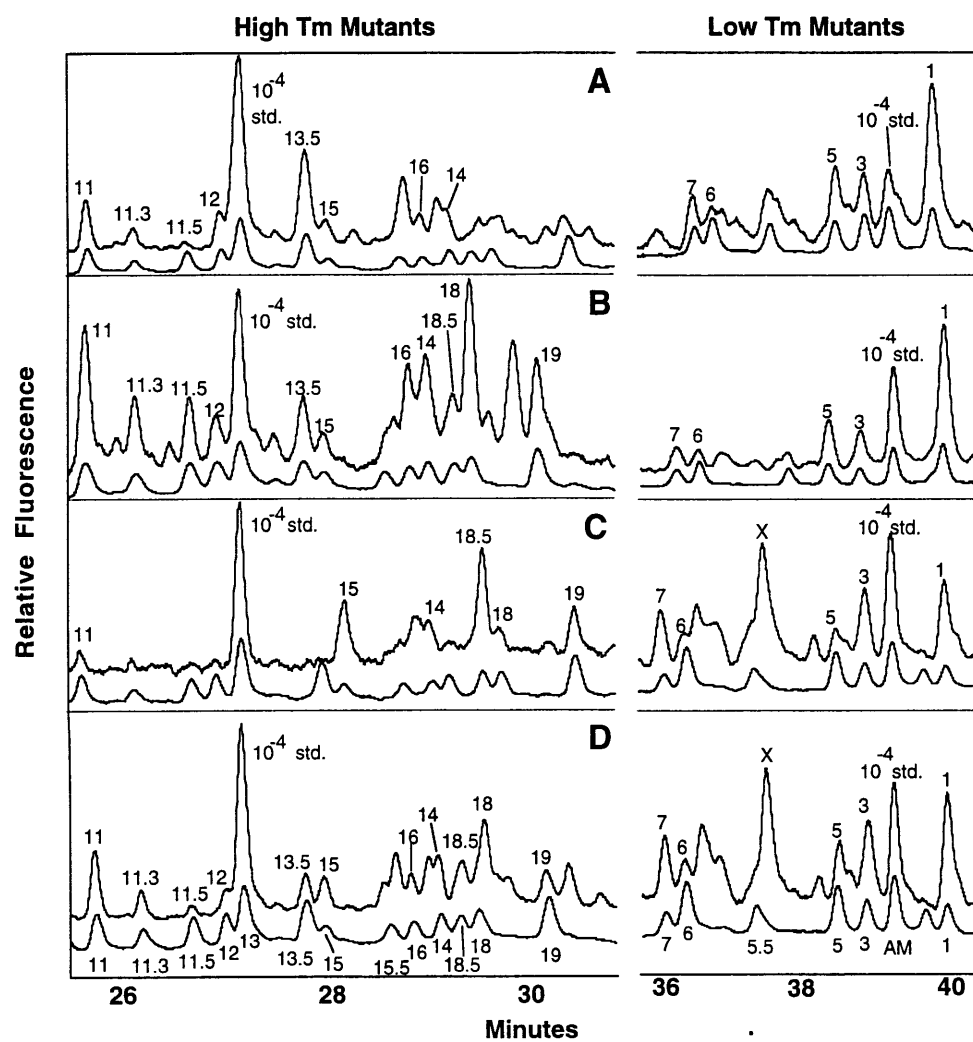


Fig. 27: The mitochondrial DNA sequence (bp 10011-10215) used as a target sequence for these studies.

The high melting and low melting domains, as well as the junction between the tRNA^{Gly} sequence and the beginning of the NADH dehydrogenase subunit III sequence, are indicated. Also shown are the specific base pair substitutions associated with each mutant in the standard set, along with the number designation. The base substitution of the artificial mutant used as an internal standard for low T_m mutants is given and designated "am". Sequences recognized by DNA primers used for PCR are underlined.

Target Sequence and Isolated Mutants

10011

ACCGTTAACT TCCAATTAAC

primer CW7

TAGTTTTGAC



 A T
 3 am

10041

AACATTCAA

$\downarrow$
 G
 13.5

◀—tRNA^{Gly} | NADH dehydrogenase subIII —▶

AAAGAGTAAT

↓ ↓ ↓ ↓
ACG C
5.5 11 11.5 11.3

AAACTTCGCC

			
G	C	T	A
18.5	18	2	6
		A	
		1	

10071

TTAATTTTAA



TAATCAACAC

CCTCCTAGCC

↓
A
5

10101

TTACTACTAA

TAATTATTAC

A DNA sequence is shown: ATTTTGACTA. Above the sequence, a horizontal arrow points left towards the 'T' at position 6, labeled "low melting domain". Below the sequence, a vertical arrow points down from the 'A' at position 7, which is labeled "7" below it.

high melting domain 
CCACA ACTCA

ACGGCTACAT

AGAAAAATCC

10161

ACCCCTTACG

AGTGC GGCTT

CGACCCTATA

10191

TCCCCCGCCC

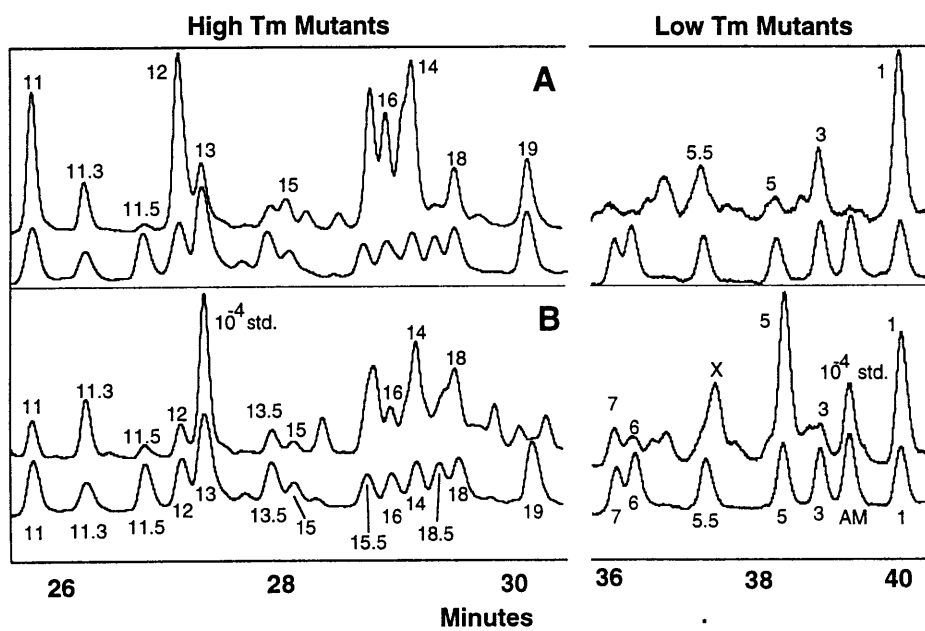
GCGTCCCTTT

CTCCA

complementary to primer J3

Fig. 28: Mutational spectra of colon and muscle tumor samples.

The mutational spectra was determined for colon (A) and muscle (B) tumor samples of 1 g and 20 mg, respectively. Shown are the CDCE separations of the mutational spectra for colon and muscle tumor samples as well as the co-separations of the TMR-labeled set (lower curve of each pair of curves).



4.3.4 TK6 cells

The mutational spectra present in TK6 cells, as an example of human cells in culture, was also determined (Fig. 29). The most prominent mutation in these cells was mutant p1. Other mutants are present at a significantly lower mutant fraction. Examination of the high T_m mutants, combined with capillary hybridization, revealed that the mutations present in TK6 cells are in fact the mutants from the standard set observed in human tissues as well. The ratio of p1 to other peaks is apparently about 10-fold higher in TK6 cells as compared with human tissue samples.

An important question to address was whether the point mutant fractions increase with generations of growth in TK6 cells. While "older" cells might contain more adducts, as a result of accumulated mutations to oxidative phosphorylation proteins or DNA repair enzymes, for instance, it would be consistent with the hypothesis that the signals observed are true mutants if they were found to increase with generations of growth in culture. K. Khrapko determined the mutational spectra for TK6 cells that were grown for up to 400 generations in culture. The relatively larger size of peak p1 made it a good candidate to study the increase in mutant fraction with cell division. Shown in Figure 29A is the increase in mutant fraction with cell divisions for p1. Separate data points represent independent mutant fraction measurements from a single culture. P1 increased almost linearly relative to the internal standard up to 200 generations. The mutant fraction did not increase linearly from 200 to 400 generations. After 400 generations of continuous culture, it is possible that a faster growing subpopulation of cells displaced the other cells in the culture, thus making mutant fractions no longer reliable. Displacement by a faster growing subpopulation has been previously documented (Kubitschek, 1970).

Figure 29B shows the set of high T_m mutant peaks for the same TK6 culture at 52 and 400 successive generations. The mutants observed in tissue samples are present in the cultured cells, as confirmed by capillary hybridization. However, the mutant fractions are significantly lower in cultured human cells than in tissues (note the change in the size of the internal standard peak area between Figures 28 and 29). The high T_m mutants, in general, increased in mutant frequency with cell division, but the increase was not linear with generations of culture. The nonlinearity of the increase may reflect the low mutant fractions of these peaks. In a single, relatively small culture from which 2/3 were discarded each day for dilution, rare mutants, especially if they were not distributed evenly among cells, would not be expected to show a linear increase in mutant fraction (Oller et al., 1989).

4.3.5 Treated cells

In order to determine whether exogenous compounds are capable of inducing mutations in mitochondrial DNA, L. Marcelino and colleagues treated a mismatch repair deficient cell line, MT-1 (Goldmacher et al., 1986), with a 4 μ M dose of the alkylating agent N-methyl-N'-nitro-N-nitrosoguanidine (MNNG)(Marcelino et al., in prep). The induced mutant fraction was 5×10^{-3} in the *hprt* gene as determined by plating in the presence and absence of 6-thioguanine, a 200-fold induction over controls. Using the methodology for determining mutational spectra in mitochondrial DNA as described in Materials and Methods, multiple hotspot mutations were observed in duplicate in MNNG-treated cultures and not in controls (see Figure 30). Sequencing of individual MNNG-induced mutants revealed that they were all GC->AT mutations, a finding consistent with MNNG-induced mutations in other systems (Burns et al., 1987; Coulondre et al., 1977; Kat, 1992). L. Marcelino also treated cells with benzo[a]pyrene diol epoxide. In order to induce a significant surviving mutant fraction in *hprt*, she subjected the cells to a series of eight cycles of treatment and recovery, and a 40-fold induction over the controls resulted. When these cells were analyzed with the procedures described in Materials and Methods, no prominent mutants induced by the treatment were observed in the target sequence.

Fig. 29: Mutational spectra of TK6 cells

A. Mutant fraction of p1 (GC->AT, bp 10,068) with generations of human cells grown in culture. TK6 cells were sampled at the number of generations designated on the horizontal axis and the mutational spectra determined. The mutant fraction for p1 is plotted on the y axis.

B. TK6 cells grown for 52 and 400 generations were sampled. DNA was isolated from an aliquot of cells and subjected to mutational spectrometry. CDCE runs of the mutational spectra for high T_m mutants are shown.

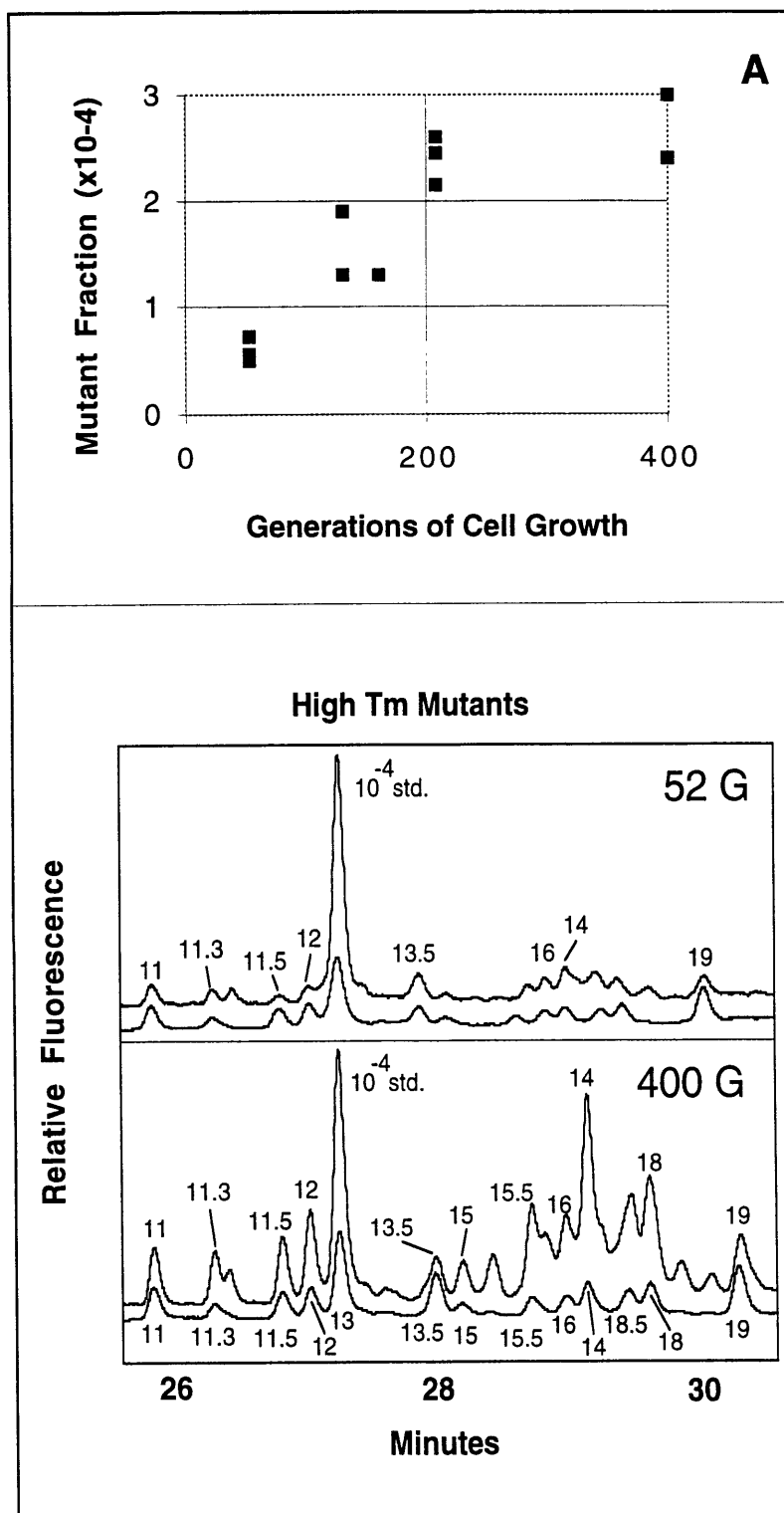
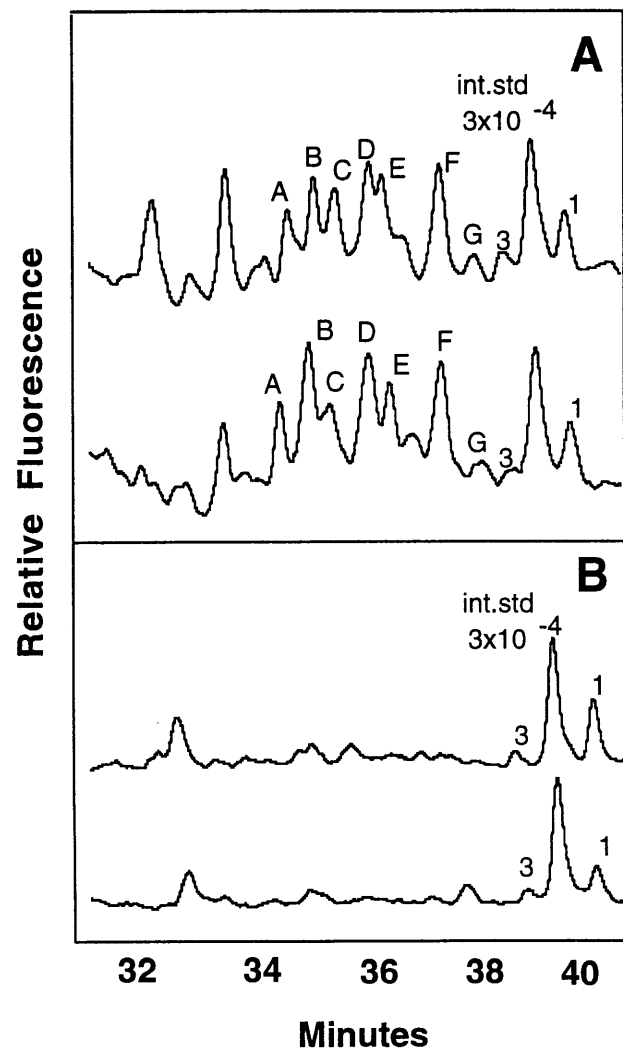


Fig. 30: MNNG mutational spectrum in mitochondrial DNA of MT1 cells.

MT-1 cells that lack mismatch repair were treated with 4 μ M MNNG or left untreated as controls in duplicate. The mutational spectra of low T_m mutants are shown for (A) 4 μ M MNNG, and (B) 0 μ M MNNG. The two curves represent duplicate samples. The artificial mutant was included as an internal standard at a mutant fraction of 3×10^{-4} . Induced mutations, labeled with letters, are clearly observed in treated and not untreated samples. Sequencing the induced mutations revealed that they were GC→AT transitions.

Source: L. Marcelino



4.3.6 Mutant fractions and mutation rate

The average mutant fraction for each of the 17 mutants of the standard set was calculated for all bronchial epithelial cell samples and plotted in a histogram along the target sequence shown in Figure 31. Also provided are the kinds of mutation and the position of the mutational hotspots in the context of the local DNA sequence. Note that the average mutant frequencies for these prominent hotspots range from 5.3×10^{-6} for p11.5 to 4.3×10^{-4} for p1.

Estimating a mutation rate from the mutant fraction determined by this procedure requires a number of assumptions for several reasons. First, it is not possible to distinguish between mutants that were generated independently and those that resulted from expansion of a cell containing the particular mutant. In the case of *hprt* mutants in T lymphocytes, the clonality of individual mutant colonies can be determined by characterizing the T-cell receptor as described in the Literature Review: Phenotypic Selection. The extent of overestimation for solid tissues has not been determined, but as demonstrated below, clonal expansion of mutants is likely to play a role in at least a small fraction of observations.

Another significant drawback in estimating a mutation rate from mutant fractions are uncertainties about the number of divisions. For this discussion, bronchial epithelial cells are assumed to divide three times per year by analogy with colon epithelium and the observations of Dr. E. Furth (Univ. of Pennsylvania), that colon epithelium divides at approximately this rate. Further, the number of mitochondrial DNA divisions per cell division is assumed to be one.

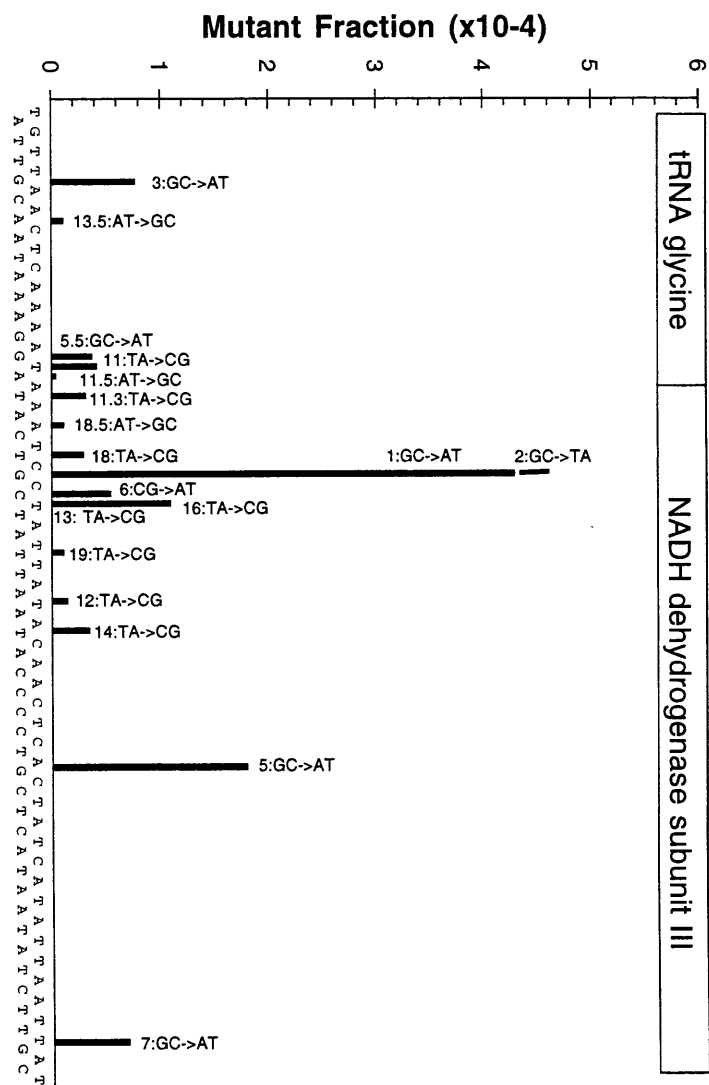
With these caveats in mind, a typical, prominent peak in the bronchoscopy samples based on comparison with an internal standard has a mutant fraction of approximately 10^{-4} . Assume that this mutation occurred after development of the lung when the lung epithelium was maintaining constant cell number. In order to model an organ with constant cell number, assume that half the cells die after DNA replication and cell division, along with half of the mutants. In each generation, Nr mutants are created where N is the number of cells and r is the mutation rate. Assuming half the cells and mutants die after each generation, $\frac{Nr}{2}$ mutants remain at the end of the generation. The number of mutants

present at any generation is $\frac{gNr}{2}$, where g is the number of generations, while the number of cells is N . The mutant fraction can therefore be approximated by $\frac{gr}{2}$. The number of generations is estimated by dividing the number of years of life by the turnover time for cells in the lung. For a 30 year old individual, assuming three divisions per year, the number of generations is 90. A mutation rate of 2×10^{-6} errors/(bp x duplication) for a prominent hotspot would consequently be predicted.

The calculated mutant frequency for a hotspot mtDNA mutation can be compared with that estimated for nuclear mutations. Experiments on T-lymphocytes from human donors suggest a mutation rate of 5×10^{-7} mutations per generation for the entire *hprt* gene (Green et al., 1995). Assuming that there are approximately 1000 base pairs at risk in *hprt*, an average nuclear mutation rate of approximately 10^{-10} mutations/(bp x generation) results. A nuclear hotspot with a mutation rate 100-fold the average would have a mutation rate of 1×10^{-8} in the nucleus as compared with 2×10^{-6} for a mitochondrial hotspot. Consequently, the best available estimate is that the mutation rate in mitochondrial DNA is apparently several hundred-fold higher than in nuclear DNA.

Fig. 31: Position, type of mutation and frequency of hotspot mutations.

The type of mutation and its position for each of the sequenced mutants is shown. The target sequence printed along the x axis is staggered for clarity. The positions of the two genes encoded by this sequence, tRNA^{Gly} and NADH dehydrogenase subunit III, are designated. The height of each peak reflects the average mutant fraction of this peak in all lung samples analyzed to date (n=39). Note that mutants 1 and 2 are different substitutions at the same base pair. Also note that mutant 13 (TA->CG, bp 10,072) was used as an internal standard to measure the size of other peaks.



4.4 COMPARISON OF SMOKERS AND NONSMOKERS

4.4.1 Twins discordant for smoking

Shown in Fig. 32 are mutational spectra for mutants with a lower melting temperature than wild-type in two sections of the lungs of a pair of twins. In the top panel (Fig. 32A) are the results for cells sampled from the right side of the main bifurcation separating the right and left lungs of the smoking twin. Individual peaks are numbered based on capillary hybridization with TMR-labeled standards. The artificial mutant included as an internal standard is identified in each sample and was introduced at 10^{-4} . Panels 32B and 32C provide the results for the same individual when DNA isolated from cells sampled from the first bifurcation in the right upper lobe was analyzed. The sample was divided into two at the DNA isolation stage and each portion was independently subjected to the procedure. The figure demonstrates that the results are almost identical spectra. Figs. 32D and 32E provide the results when samples from the corresponding smoking twin were analyzed, including samples from the right side of the first bifurcation (Fig. 32D) and the right upper lobe (Fig. 32E).

The same low T_m mutants were observed repeatedly in different bronchial epithelial cell samples. No specific low T_m mutants were present only in smoking samples, and the overall mutant fractions were similar regardless of smoking history. Further, the amount of variability was as great between two samples taken from different positions within a single individual's lung as between samples taken from different individuals. For instance, in the right main carina of the smoking twin (Fig. 32D), mutant 5.5 is the most prominent mutant, present at a mutant fraction of 2×10^{-4} . In the right upper lobe sample from the same individual (Fig. 32E), which was biopsied from just a few centimeters of airway away, this particular peak did not rise significantly above background levels. This difference is not attributable to artifacts of the procedure, as it is clearly observed in each sample in duplicate. Further, intralung variation is as great as any observed between smokers and nonsmokers.

These findings were in concordance with earlier experiments performed by K. Khrapko in which a muscle tumor was sliced into several sections with a microtome (see Figure 33). Mutational spectra were determined for two adjacent slices and a remote slice. The adjacent slices showed similar mutational spectra while the remote slice had a distinct spectra. These findings suggest a "clustering" of mitochondrial mutants within solid tissues.

Fig. 34 provides the mutational spectra, including both low and high T_m mutants, for a representative sample from each of the six twins. This figure includes the TMR-labeled "standard set" of previously purified mutants that was coseparated with each sample and recorded in a separate channel. Similar observations as reported for Fig. 32 are here extended to the high T_m mutants and other twin pairs. The same mutants were repeatedly observed as the most prominent mutations in all of the samples, although a particular mutant may not be detectable in a given sample. Further, an extra mutant not part of the standard set was occasionally observed at a significant frequency in a particular sample (for instance, peak x in Fig. 34A). No specific peaks were observed consistently in smokers' samples and not in nonsmokers. Further, the overall mutant fraction, based on comparison with internal standards, was not significantly increased in smokers.

4.4.2 Dissected lungs

The brush biopsy samples that were obtainable from twins' lungs were analyzed through the third generation of airway. In order to test whether smoking-induced

Fig. 32: Sample of low T_m mutational spectra from twins bronchial epithelial cells.

Brush bronchoscopy samples were biopsied from the right main carina and the first bifurcation of the right upper lobe from smoking and nonsmoking identical twins. Samples were subjected to the procedure for determining mitochondrial mutational spectra as described in Materials and Methods, and the low T_m mutants are depicted. Individual peaks are identified as specific mutants by number based on comigration with TMR-labeled standards recorded in a separate channel (not shown) and were confirmed by capillary hybridization. The artificial mutant was added as an internal standard at 10^{-4} prior to CDGE. Low T_m mutational spectra were determined from the nonsmoking twin right main carina (A) and right upper lobe in duplicate (B) and (C). The DNA was split after restriction digestion and internal standards were added separately. Duplicate samples were loaded on a CDGE and analyzed independently. The low T_m mutational spectra were also determined from the smoking twin's right main carina (D) and right upper lobe (E).

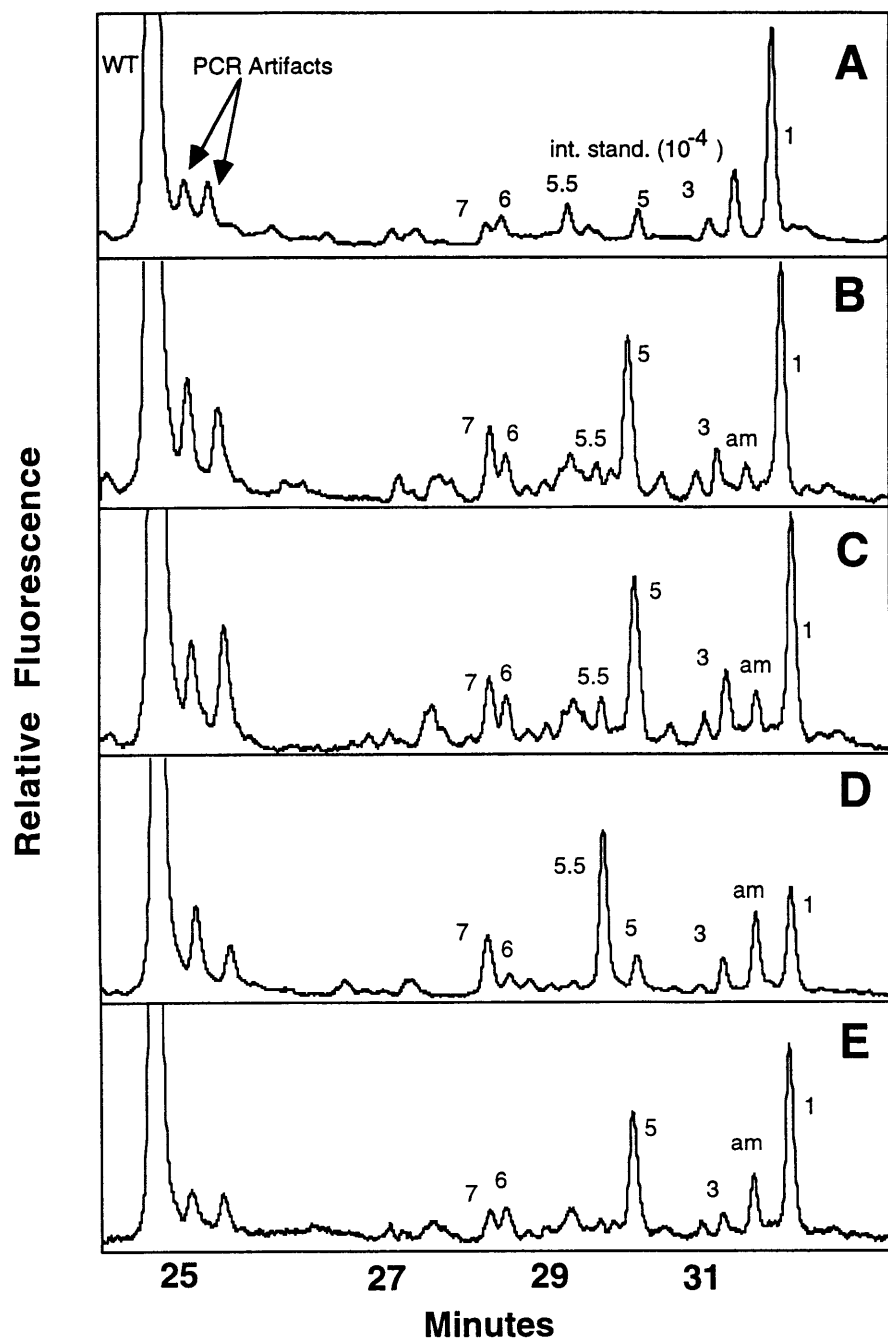


Fig. 33: Mitochondrial mutational spectra of a muscle tumor sample sliced with a microtome.

Three slices of a muscle tumor, two adjacent to each other (A and B) and one remote (C) were analyzed to determine the mitochondrial mutational spectra. Samples were run prior to high resolution gel conditions and were not subjected to capillary hybridization. Mutants are therefore lettered instead of numbered. Mutants with the same letter in different samples are likely to represent the same mutant, but capillary hybridization would be required for confirmation.

Source: K. Khrapko

High Tm Mutants

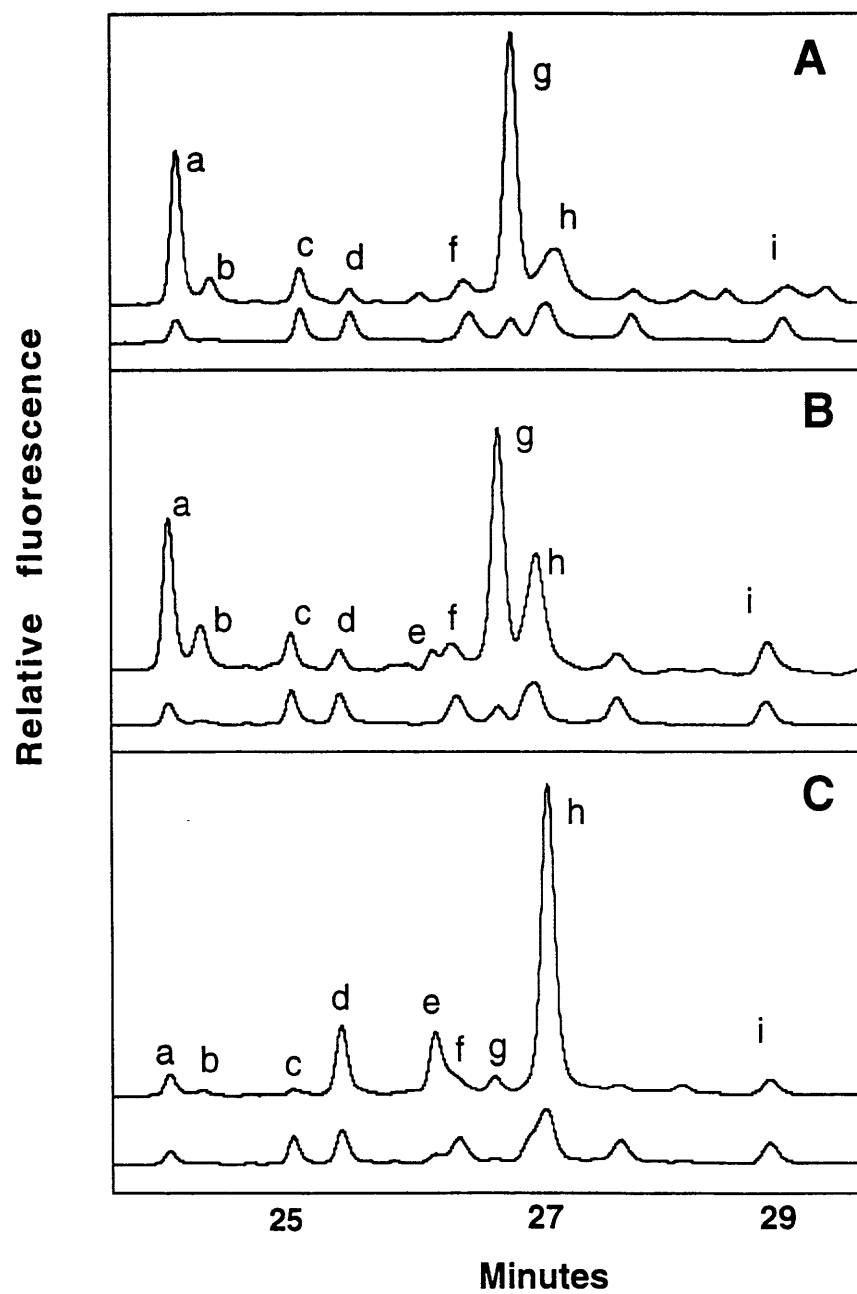
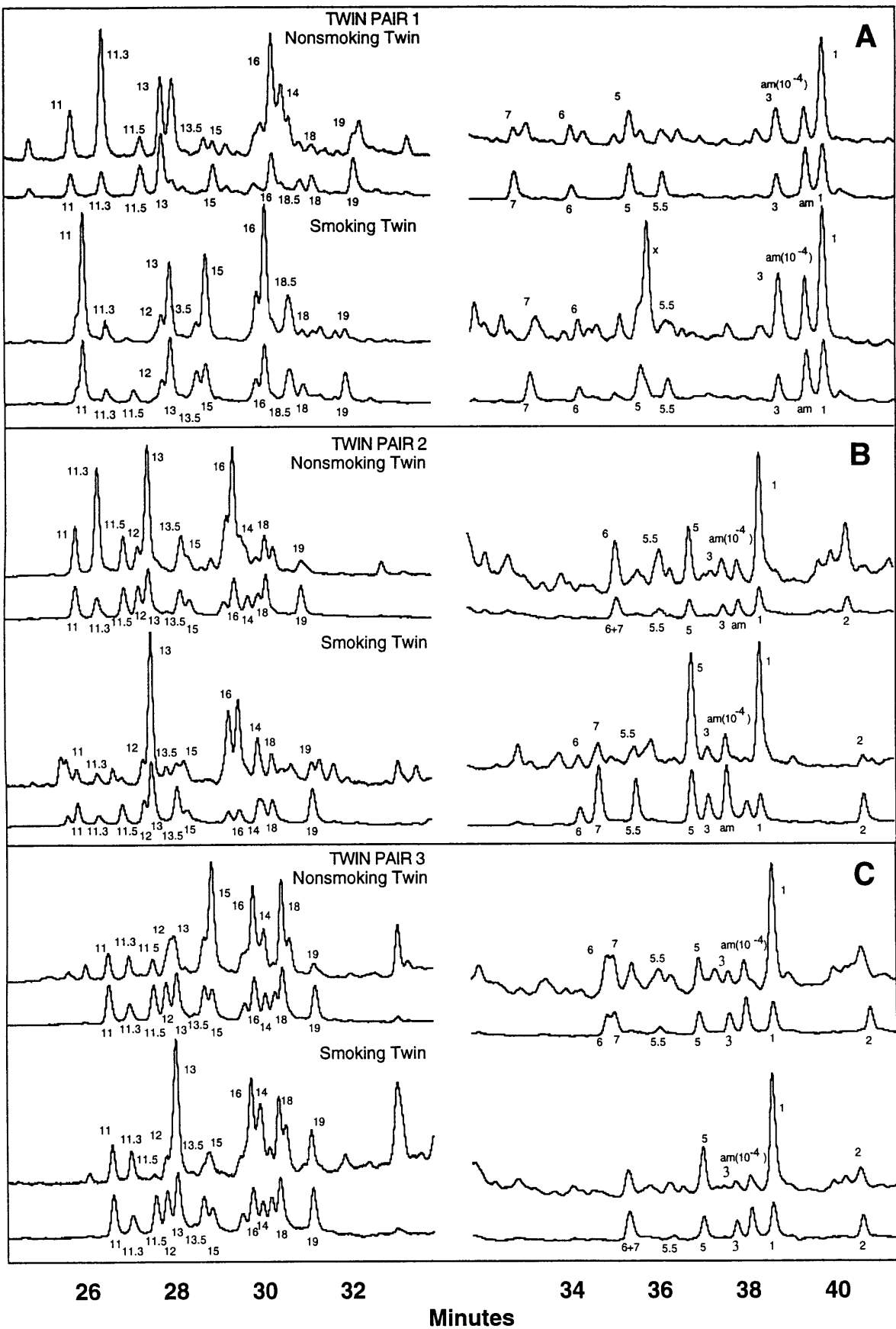


Fig. 34: Mutational spectra of bronchial epithelial cell samples from three pairs of twins discordant for smoking.

Mutation spectra of a representative sample from each of the six twins included in the study are shown. Both high T_m homoduplexes (left panels) and low T_m homoduplexes (right panel) are shown. (A) The results from twin pair 1; both samples were biopsied from the right upper lobe. (B) Samples taken from the second pair of twins. For the smoking twin, the sample was taken from the right upper lobe and for the nonsmoking twin from the left lower lobe. (C) Mutational spectra of samples biopsied from the third pair of twins, in both cases from the right lower lobe. Cells were resuspended in 100 ml PBS and the mutational spectra were determined for each sample as described in Materials and Methods. The artificial mutant and p13 were added as internal standards at 10^{-4} except in the case of twin pair 3, nonsmoking sample, in which p13 was added at 10^{-5} . Shown are fluorescein-labeled homoduplexes and the TMR-labeled standard set with which they were coseparated. Peak identity was confirmed by capillary hybridization.



mutations were present in airways distal to the third generation, bronchial epithelial cells were isolated from dissected whole lungs. Mutational spectra derived from dissected lungs, including samples from both smokers and nonsmokers, were similar to spectra from bronchoscopy samples. The same mutants observed in bronchoscopy samples were also represented in the dissected lungs, and the mutant frequencies were similar as well. No significant differences were observed in smokers compared to nonsmokers in any of the airways tested, through the eighth bifurcation.

4.4.3 Average mutant fractions

The previous figures have demonstrated that there are no new, prominent mutants observed in smokers bronchial epithelial cells and not in nonsmokers' cells. Further, they also reveal that the overall mutant fraction is not larger in smokers samples. Another important question was whether any particular one of the mutants commonly observed in the standard set was present at a significantly higher frequency in the smokers as compared with the nonsmokers. Shown in Fig. 35 is a summary of the average mutant fraction for each of the mutants commonly observed among bronchial epithelial cell samples in smokers and nonsmokers. Smokers' mean mutant fractions represent averages from 13 bronchoscopy samples taken from 4 individuals and 14 samples from dissected lungs of 2 individuals. The nonsmokers' mutant fractions represent the mean from 11 bronchoscopy samples taken from 4 individuals and 5 samples taken from 1 dissected lung. These data are not normally distributed and therefore the use of a standard deviation is not strictly appropriate. However, the standard deviations are included in order to clarify the conclusion that no specific peak was statistically significantly different between the smokers and nonsmokers.

The standard deviations on the mean mutant fraction are large as compared with the mean for many of the specific mutants. The large standard deviations reflect cases in which individual mutants were substantially overrepresented in particular samples. An example of such an instance is the large size of mutant 5.5 in the smoking twin, right main carina sample pictured in Fig. 32D. The same mutant was barely detectable in a sample taken a few centimeters of airway away as shown in Fig. 32E.

4.5 Variability in mutant fractions

While the same hotspots were observed repeatedly in multiple bronchial epithelial cell samples from many individuals, the frequency of individual mutants varied significantly in different samples. This variability could not be attributed to experimental errors, smoking status, position within the lung, age, gender, or any other variable measured.

In order to better understand the distribution of mutant fractions among samples, the mutant fractions were plotted as histograms. Mutant fractions were measured in each bronchial epithelial sample for 16 mutants that were considered unlikely to be PCR noise (p6 was eliminated) or adducts (p3 was eliminated). For the samples that were analyzed in duplicate (about 2/3 of the database), mutant fractions were averaged. In approximately 1/3 to 1/2 of the cases, the mutant fractions were confirmed by capillary hybridization. Samples likely to represent errors in the introduction of the internal standard, for instance, in which high T_m or low T_m mutants were consistently over or under represented (4 samples), were discarded from analysis. The final database contained mutant fraction data for 570 mutants from 39 samples. A table containing the 570 mutant fractions is provided in the appendix.

Histograms of the number of samples containing a particular range of mutant fractions were generated for each of the mutant peaks. Examples of three such histograms are given in Figure 36. The mutant fractions are not normally distributed. There are often

Fig. 35: Mean mutant fractions in smokers and nonsmokers.

Histogram of mean mutant fractions in smokers and nonsmokers. Peak areas were measured in high resolution homoduplex runs for each of 17 frequently observed mutants. Smokers were represented by 13 bronchoscopy samples taken from 4 individuals and 14 samples from dissected lungs of 2 individuals. The mean of the nonsmokers represents 11 bronchoscopy samples taken from 4 individuals and 5 samples taken from 1 dissected lung. Peak areas were converted to mutant fractions based on comparison with the introduced internal standard. Mutant fractions were measured and the mean and standard deviations determined. Mean mutant fractions for smokers (speckled bars) and nonsmokers (white bars) are plotted. The standard deviation is designated by a line extending from the mean to the position of the mean plus one standard deviation.

Mutant Fractions for Each Mutant in Smokers and Nonsmokers

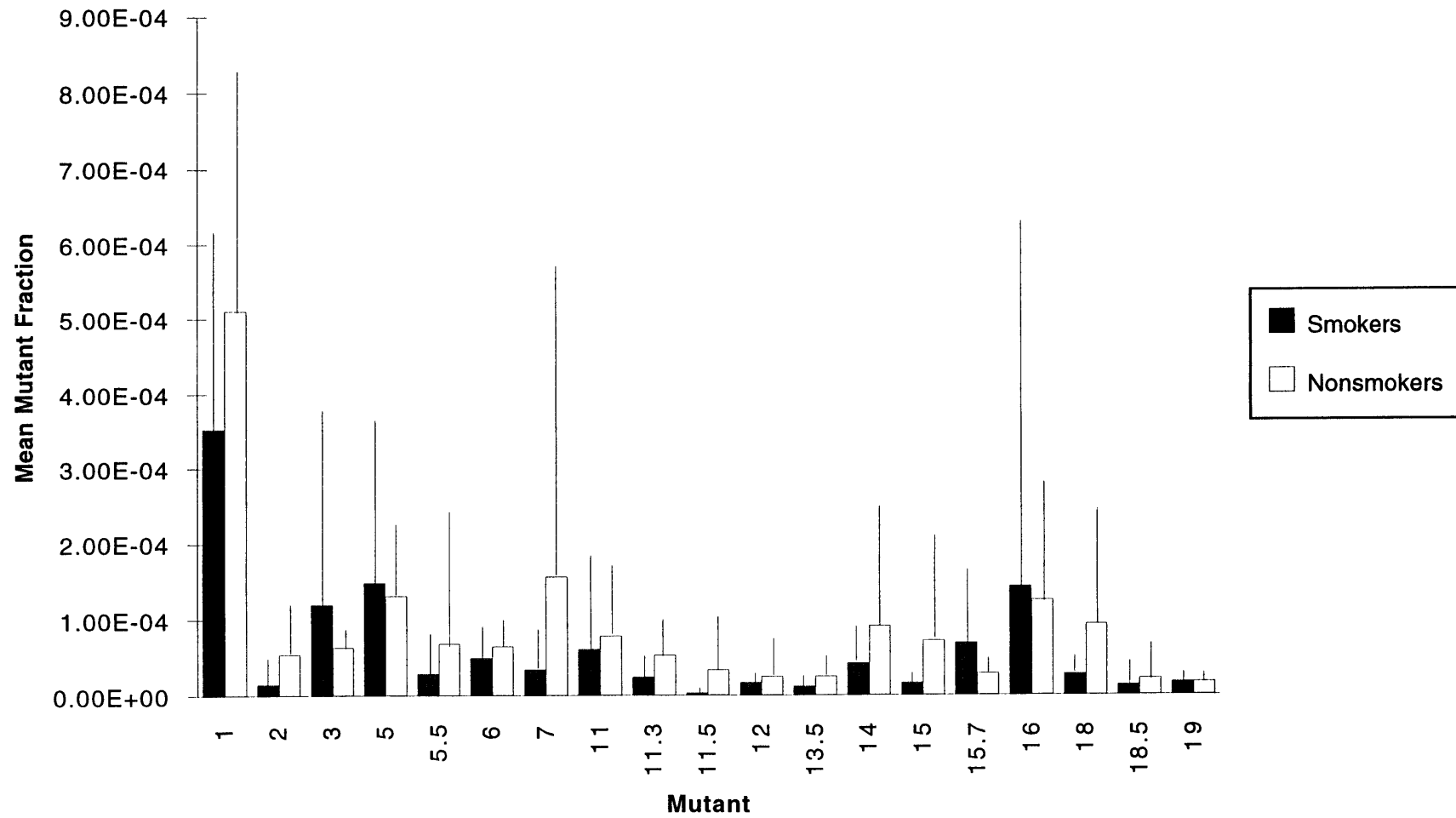
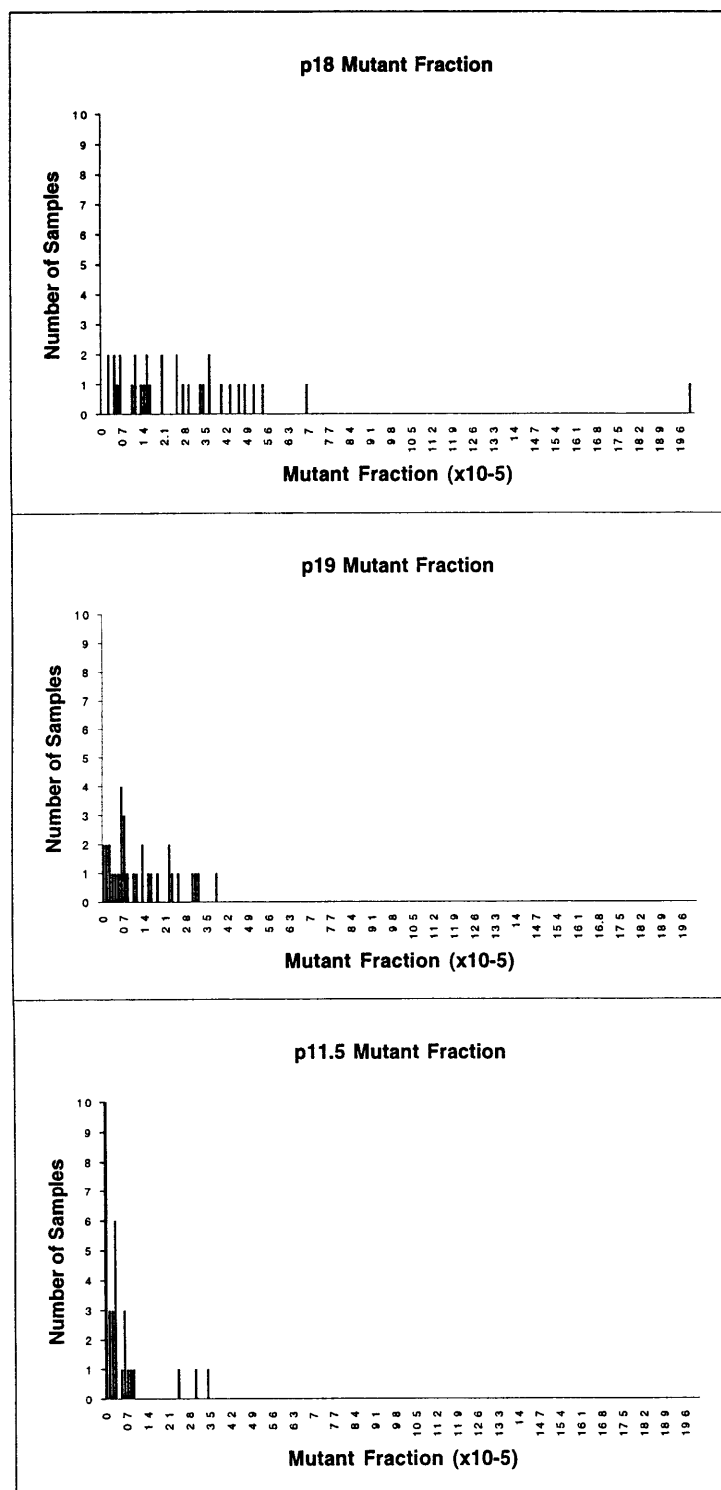


Fig. 36: Examples of mutant fraction histograms.

Mutant fractions were measured in 39 bronchial epithelial cell samples. In some cases, a particular mutant was difficult to measure and was therefore eliminated from analysis. The number of samples with a given mutant fraction was plotted and the results are shown for three mutants, p18, p19 and p11.5.



one or two samples that contain a significantly larger mutant fraction than other samples (for instance, the unusually large p18 shown in Fig. 36). These observations are unlikely to reflect errors in the procedure as they have been observed repeatedly in duplicate. Further, such mutants can vary as compared with the mean by as much as 5- to 11-fold, which is considerably larger than the variability observed when a sample is split. Such mutants are believed to represent the manifestations of clonal expansions of mutant-rich cells, as described further below.

Another observation is that, with the exception of outliers, the mutant fractions for each mutant do cluster around a consistent value characteristic of the mutant. Some mutants are clearly more frequent than others. The largest mean of 4×10^{-4} was observed for p1, while the smallest mean was for p11.5 with a mean of 5×10^{-6} . The mutant fractions for particular peaks therefore can be considered to reflect average mutant fractions characteristic of bronchial epithelial cells. Because many samples of bronchial epithelial cells have been analyzed to date, it is possible to compile a quantitative representation of a "typical" bronchial epithelial mutational spectrum. These mean mutant fractions are shown in the histogram in Figure 31.

In order to better represent the variability observed for specific mutants among the samples analyzed, the histograms for all mutants were superimposed on a single plot. First, the mean was determined for each mutant. The "q-test" was applied to this calculation; if a particular mutant increased the range of mutant fractions for the all samples by more than two-fold, it was not included in the estimation of the mean. The x-axis was then taken as the fraction of the mean for that mutant and the y axis as the number of samples with a mutant fraction in that range. This permitted superposition of all mutants to generate a single histogram (see Figure 37). The vast majority of mutants (551/570) are present in a particular sample at less than four times the mean for the particular mutant. In ten cases, a mutant was observed at 5 to 11-fold the mean for that mutant. In two cases, mutants were observed at a significantly higher frequency than in other samples (>40-fold).

Dilution experiments were also performed in order to investigate the distribution of mitochondrial mutants among cells. TK6 cells grown for 400 generations contained two prominent mutants with melting temperatures lower than wild-type, p1 (GC->AT mutation at bp 10,068) and p1.5 (GC->AT mutation at bp 10,085). p1.5 is not generally observed in tissue samples. The homoduplex spectra for a large sample of these cells is shown in Fig. 18. Aliquots of these TK6 cells were diluted and their concentration confirmed by cell counting with a hemacytometer. Defined numbers of cells were subjected to PCR: 5×10^3 cells per PCR reaction, 10^3 cells per PCR reaction and 10^2 cells per PCR reaction. The fraction of cells that were actually amplified was monitored by introduction of an internal standard. Approximately 1/3 to 1/2 of the cells present were amplified; samples containing more cells amplified more completely. The PCR products were enriched for mutants through two iterations of CDCE collection of mutants and PCR with *Pfu*. Mutant fractions were measured for p1.5 and p1 and their ratios, plotted against the number of cells per sample, are shown in Figure 38. Note that the ratio of the p1.5 to p1 mutant fraction should not be affected by the PCR efficiency. The variance in the ratios is much higher for samples with 5000 and fewer cells per sample than for the three samples with 100,000 cells, which showed very similar results. This suggests that the mitochondrial mutants are clustered among cells, and the number of mutants per cell is likely to vary over a wide range.

Fig. 37: Histogram of the number of samples versus the fraction of the mean for all peaks.

For each mutant, the mean mutant fraction was calculated and the number of samples with a given fraction of the mean was plotted. The same data are shown on two plots: (A) all of the data are included and the x axis values extend to 56, (B) two of the data points are not included and the fraction of the mean is truncated at 12.

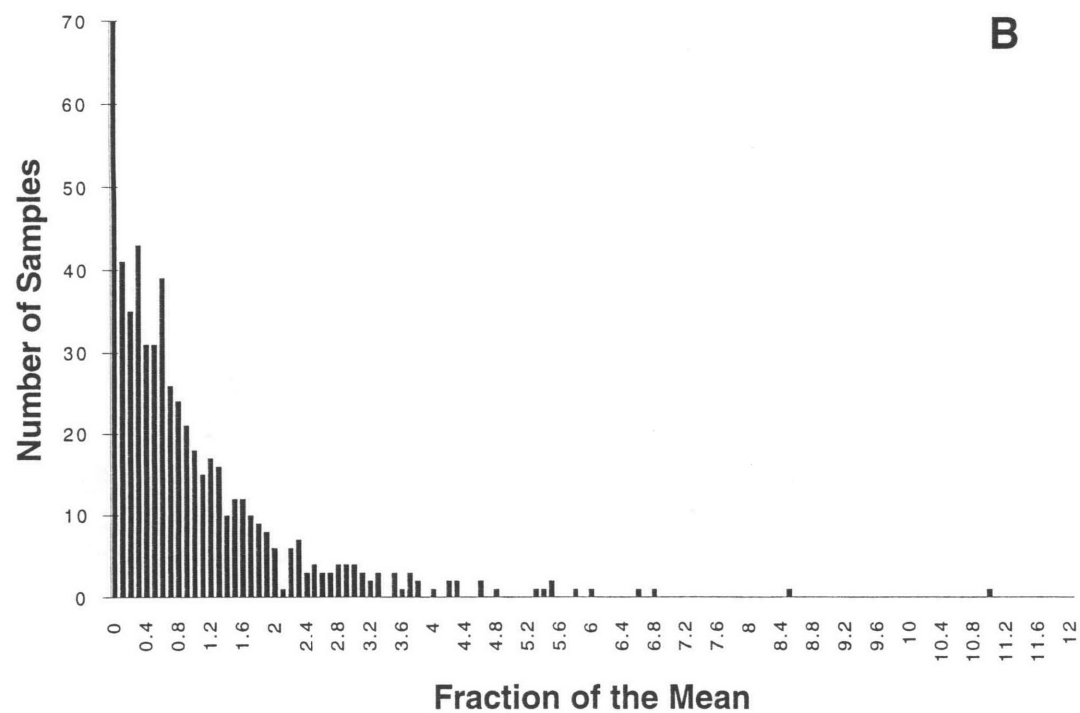
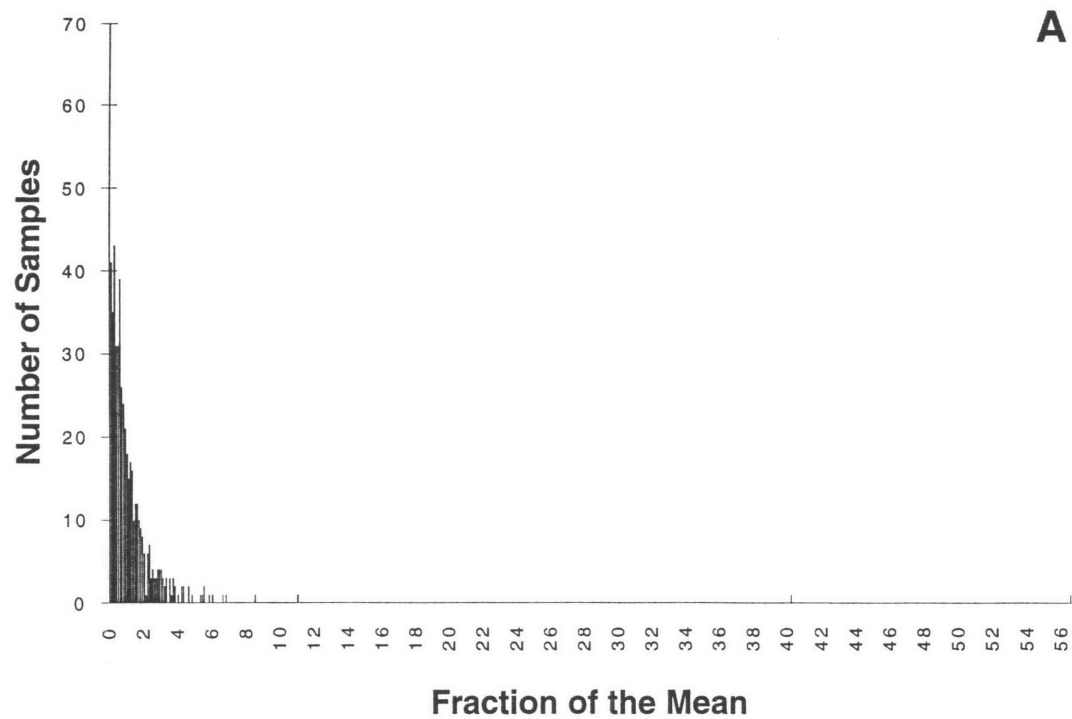
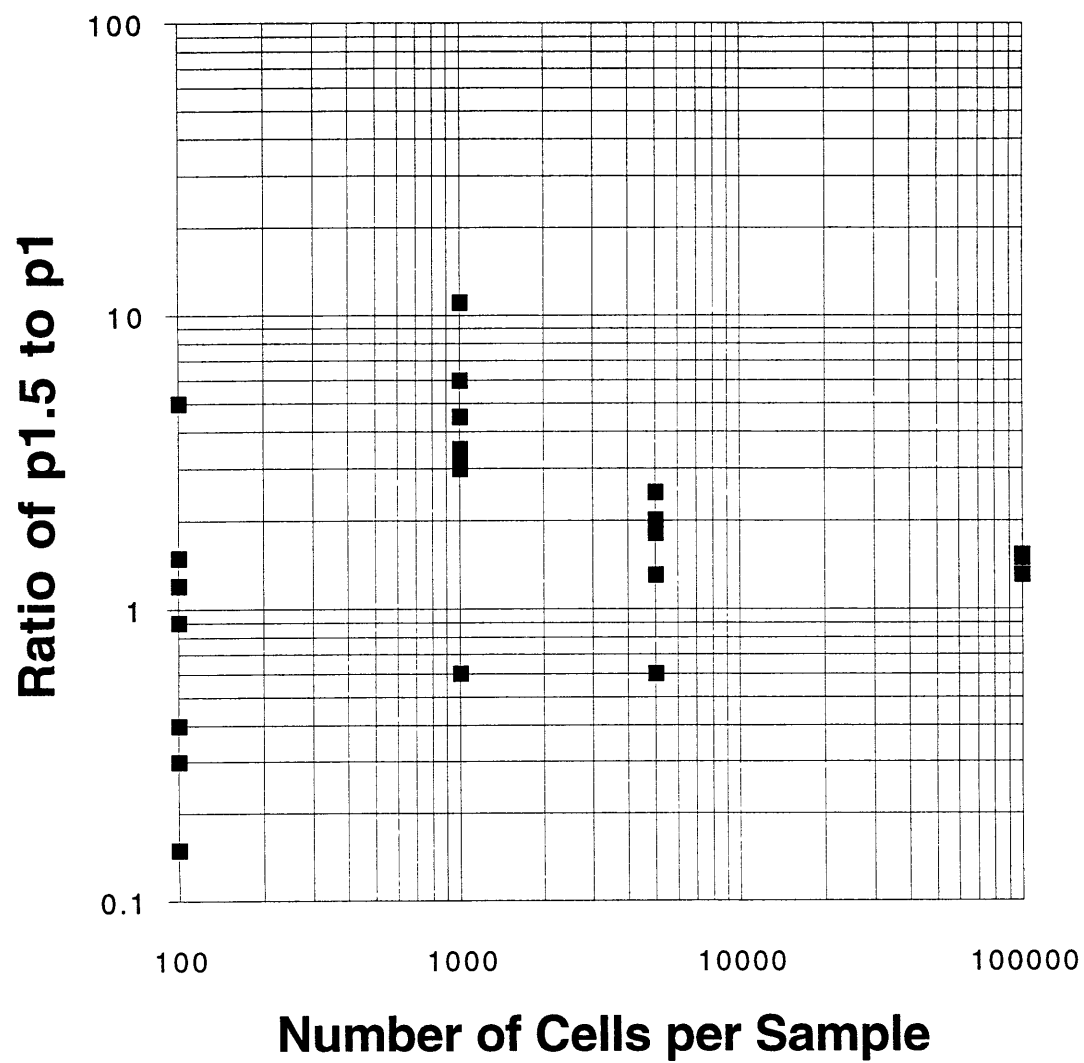


Fig. 38: Ratio of mutant fractions for p1.5 and p1 observed in limiting dilution experiments with TK6 cells grown 400 generations.

For the largest sample (100,000 cells), the mutational spectra was determined as in Materials and Methods. For smaller samples, known amounts of TK6 cells were diluted and added directly to a PCR reaction. After amplification, mutants were enriched by two rounds of CDCE enrichment and *Pfu* PCR. The ratio of the amount of p1.5 to the amount of p1 was plotted. Each point represents an independent PCR reaction.



4.6 Mutant fractions and age

4.6.1 Infant

In order to better understand the behavior of mitochondrial point mutational hotspots as a function of age, bronchial epithelial cell samples from an infant who died of sudden infant death syndrome were analyzed to determine the mutational spectra. The mitochondrial point mutation fractions present in these samples were significantly lower than for adults. An example of a mutational spectra observed for high T_m mutants for epithelial cells isolated from the left lobe of the infant is shown in Figure 39. Note that the internal standard was added at 10^{-5} , which is 10-fold lower than the mutant fraction used for adult tissue samples. The large peak on the left is assumed to be a nuclear pseudogene of this sequence, as it is consistently observed in human samples and contains multiple mutations compared with the wild-type mitochondrial sequence (Hadler et al., 1983; Hu et al., 1995; Lonsdale et al., 1983; Richter, 1983). Three different sections of the infant's lung were investigated, and all three contained mutant fractions lower than any observed in adult lung samples. These findings are consistent with the hypothesis that the mutations observed in different human adult tissues are somatic mitochondrial mutations that accumulate over the subject's lifetime.

These observations also make it extremely unlikely that the mutations detected are polymorphisms. One might argue that the 17 mutational hotspots that were repeatedly observed were actually inherited as polymorphisms, and that the differences within a single organ reflect variations due to segregation of the mutants. This would require that all of the individuals tested to date inherited the same 17 mutants as polymorphisms in this 100 base pair target. The same level of polymorphisms would be expected for the rest of the mitochondrial genome of 16.5 kb, implying that each individual should have inherited, on average, 2805 different polymorphisms according to this logic. Such an argument is extremely unlikely because, first, it is believed that there is a bottleneck in mitochondrial DNA copy number in oogenesis (Bolhuis et al., 1990; Hauswirth et al., 1985; Laipis et al., 1988). One study of mitochondrial DNA copy number during mammalian oogenesis concluded that only 5 copies are likely to be present as founders in the fertilized zygote (Hauswirth et al., 1985), while an experiment with heteroplasmic mice carrying two mtDNA genotypes concluded that the effective number of segregating units for mtDNA is approximately 200 in mice (Jenuth et al., 1996). Second, because these mutations are so low as to be essentially undetectable in the infant sample, the argument would require that coincidentally, the infant is the only one of the individuals tested so far who did not inherit these mutations.

4.6.2 Bronchial epithelial cells

Analysis of samples from four dissected lungs permitted consideration of the trends in mitochondrial mutant fractions with age. Shown in Figure 40 are the results when the sum of all of the mutants in the standard set for each of several different samples taken from four individuals was plotted against age. The infant samples contained significantly fewer mutants than samples from adults. The point mutant fractions in the 84 year old and two middle aged individuals are similar. It is not clear from these data whether the most appropriate interpretation is that the mutant fraction is increasing modestly but consistently with age, or whether the mutant fraction increases between infancy and adulthood and then plateaus. The results suggest that point mutations in human bronchial epithelial cells do not increase supralinearly with age as has been observed for deletions in post-mitotic tissues (see Figure 6 in Literature Review: mtDNA and Aging), although analysis of more individuals is required to draw firm conclusions.

Fig. 39: Mitochondrial mutational spectra of bronchial epithelial cells from an infant.

Lungs of an infant who died of sudden infant death syndrome were harvested and bronchial epithelial cells were collected from the left lobe. DNA was isolated and the mitochondrial mutational spectra was determined. Shown are the high T_m mutants that resulted. Note that the internal standard is present at a mutant fraction of 10^{-5} , ten-fold lower than the mutant fraction used for adult tissues.

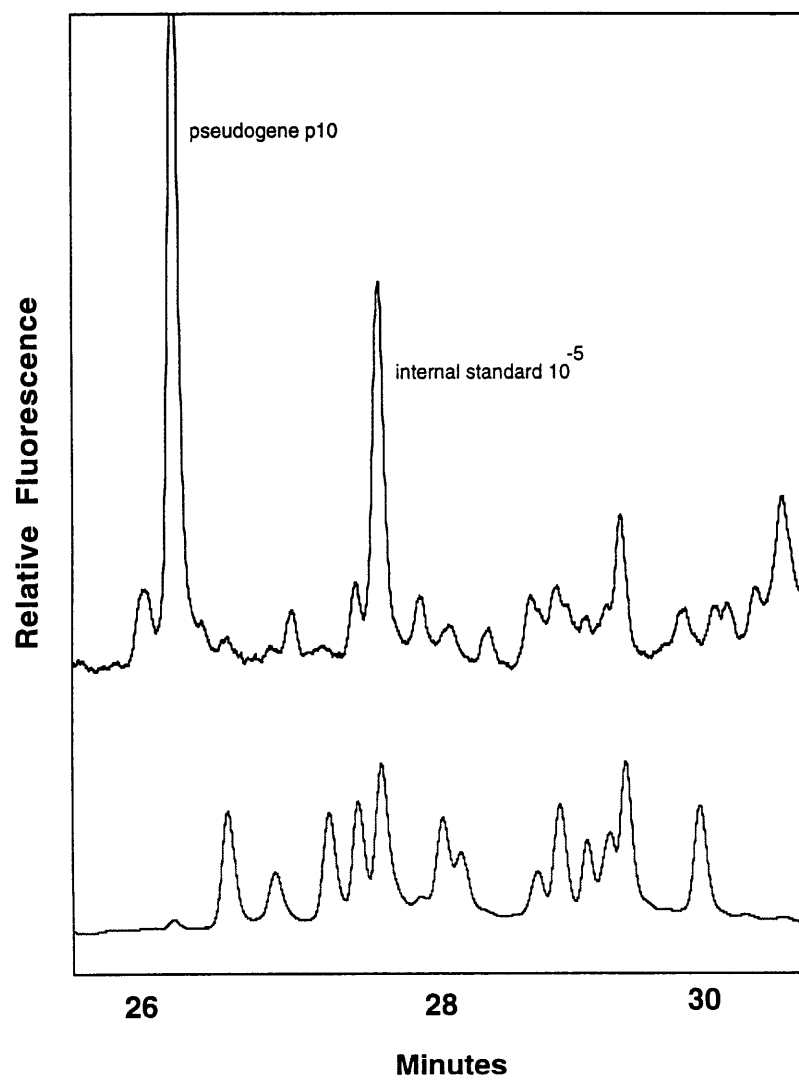
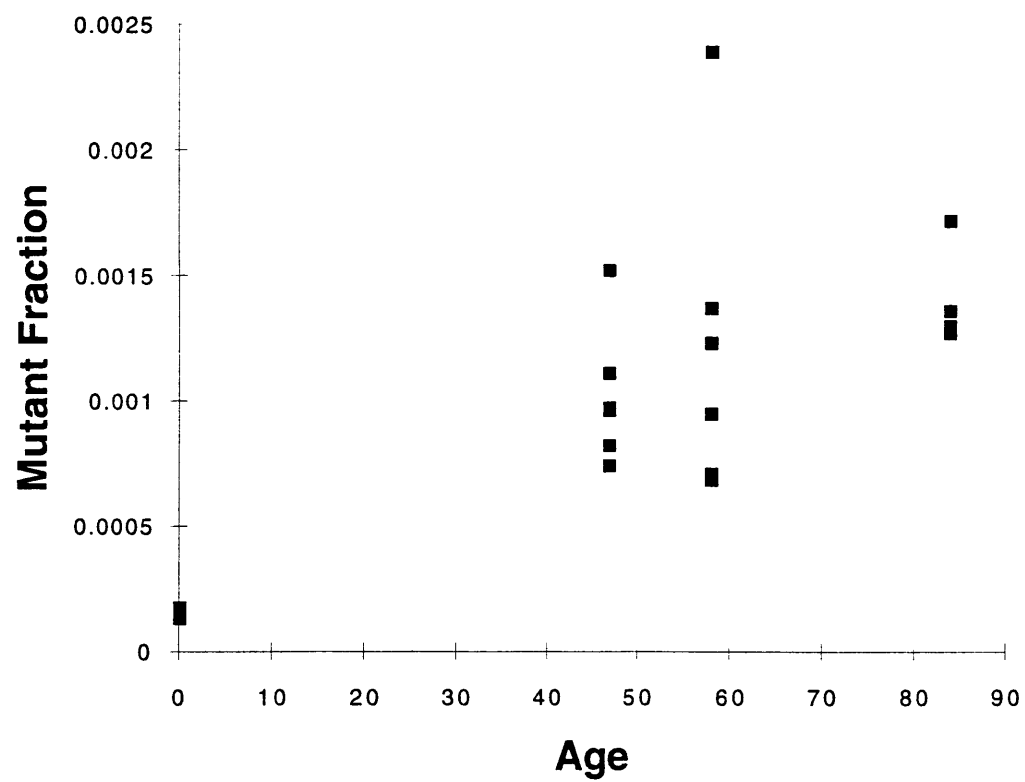


Fig. 40: Mitochondrial mutant fractions with age.

Mutant fractions for each of the mutants in the standard set were summed for each bronchial epithelial sample analyzed from four dissected lungs. The sum of the mutant fractions in different lung sectors is plotted against the donor's age.

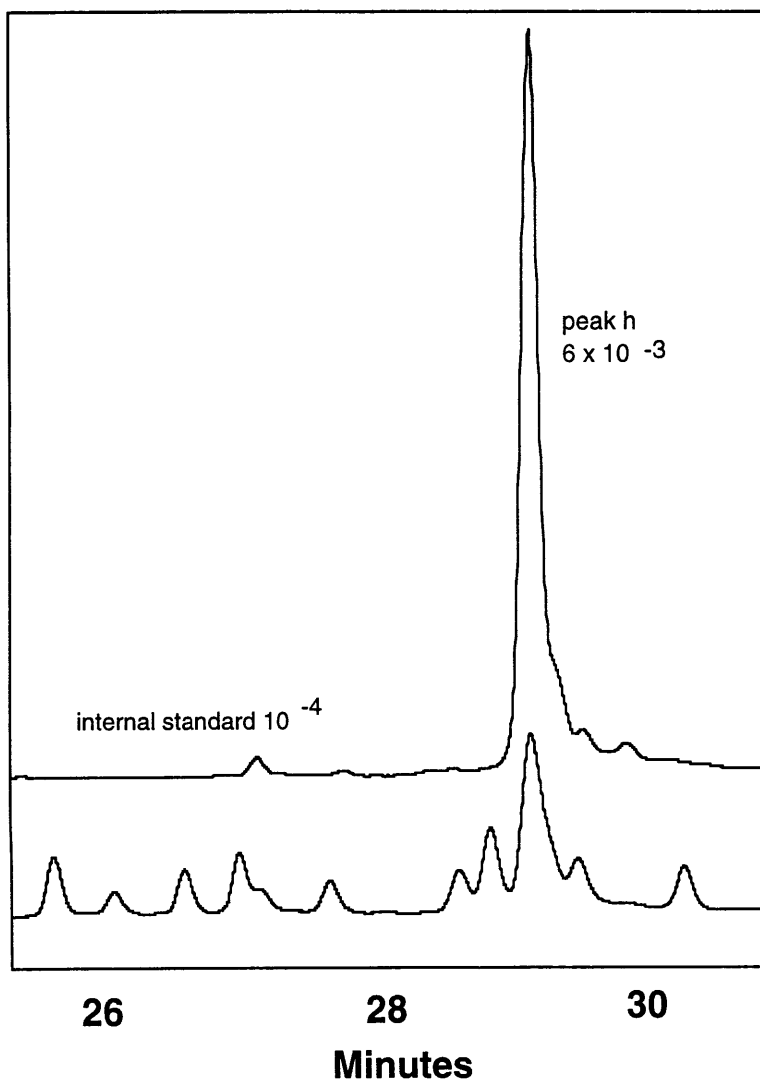


4.6.3 Other tissues

While studies on bronchial epithelial cells suggest that point mutations do not increase supralinearly with age, it is important to note that lung tissue is not among the tissues for which deletions are observed to increase substantially with age. Studies on the behavior of point mutations with age in different tissues will be performed by K. Khrapko as faculty of the Division of Aging at Harvard Medical School. As a preliminary experiment toward this goal, a single experiment was performed in which the mutational spectra present in the hippocampal region of the brain from a 100 year old man and a heart muscle sample from a 77 year old man generously provided by J. Wei and J. Vijg (Harvard Medical School) were determined. While the hippocampal section of the brain did not show prominent mutants (no mutations larger than the 10^{-4} internal standards were observed), the heart sample contained two unusually large mutations. One mutant, which was more stable than wild-type, was estimated to be present at a mutant fraction of 6×10^{-3} (shown in Figure 41), while a second mutant that was less stable than wild-type was estimated to be present at a mutant fraction of 1.5×10^{-2} . These two mutants are the largest mutations yet observed with this procedure. Neither of the mutants are members of the "standard set."

Fig. 41: High T_m mutational spectra of a heart sample from a 77-year old man.

Heart muscle DNA generously provided by Dr. J. Wei (Harvard Medical School) was analyzed to determine the mitochondrial mutational spectra. Shown are the high T_m mutants. A single high T_m mutation was present at a mutant fraction 6×10^{-3} . Not shown is a low T_m mutant also discovered in this sample present at an estimated 1.5×10^{-2} .



5. DISCUSSION

5.1 SUMMARY

A methodology has been developed that permits detection of point mutations present in a 100 bp mitochondrial DNA target sequence in human tissue samples. The procedure couples partially denaturing electrophoresis to enrich for mutant over wild-type sequences with high-fidelity PCR to amplify mutant-enriched samples with minimal introduction of PCR errors. While other methods are available for detecting and measuring low frequency point mutations, they are limited to a single known mutation using allele specific PCR (Cha et al., 1992; Wu et al., 1989a) or the ligase chain reaction (Abravaya et al., 1995; Barany, 1991; Landegren et al., 1988; Wu et al., 1989b), or mutations present in ≤ 6 base pairs in a naturally occurring or artificially created restriction enzyme recognition site (Felley-Bosco et al., 1991; Kan et al., 1978; Khrapko et al., 1994a; Kumar et al., 1990; Pallotti et al., 1996). The ability to determine a mutational spectrum in a 100 bp target sequence in any human tissue by avoiding a requirement for phenotypic selection is, to the best of our knowledge, unique.

The technical advances that we have achieved in the course of this work include the development of CDGE and CDCE conditions that permit enrichment of mutants, conditions for high-fidelity PCR, accurate use of internal standards, a double dye CDCE detection system, high-resolution CDCE conditions, capillary hybridization and tests for asymmetrical species. These methods should prove valuable in future experiments designed to determine the mutational spectra in nuclear sequences.

The methodology is sufficient to observe the prominent hotspot mutations in mitochondrial DNA isolated from human tissues. Hotspot mutations were discovered and were found to be present at a higher mutant fraction than has been detected for nuclear sequences, with the ratio estimated as several hundred-fold. The same set of approximately 17 hotspot mutants were observed repeatedly in human bronchial epithelial cells from multiple donors, TK6 cells, colon and colon tumor, muscle and muscle tumor. Specific mutants were not present as prominent peaks in some samples, while prominent peaks representing other mutants not part of the "standard set" were observed on occasion as well. Sequencing revealed that most of these mutations were predominantly transitions, both GC→AT and AT→GC.

Smokers and nonsmokers contained the same set of mutational hotspots as observed in other organs. There were no prominent smoking-specific hotspots and the overall mutant fractions were not larger in the smokers than the nonsmokers. Comparison of the average mutant fractions for each of the mutants observed repeatedly did not reveal significant differences between smokers and nonsmokers. There was as much variability in the fractions observed for individual mutants in a single lung as between smokers and nonsmokers.

When all of the bronchial epithelial cell samples are considered, most of the mutant peaks observed were within 4-fold of the mean for the specific peak (551/570). Significant outliers were also observed. These are interpreted as reflecting the accumulation of multiple mitochondrial copies containing a specific mutant within a bronchial epithelial stem cell. In two instances, the mutant fractions were significantly larger than the mean (>40-fold). Such cases may reflect proliferation of mitochondria in response to lack of oxidative phosphorylation function.

The results suggest that the most likely sources of mutation in human mitochondrial DNA are endogenous sources, although the results are not sufficient to distinguish between replication errors, adduction by endogenous compounds or another pathway.

Left unresolved is the mechanism by which smoking causes cancer. Two of the important hypotheses are first, that smoking induces prominent mutations in nuclear but not mitochondrial DNA due to the high spontaneous background in mitochondrial DNA, and

second, that smoking causes cancer by a mechanism distinct from mutagenesis, for instance, by altering the kinetics of cell proliferation in bronchial epithelial cells.

5.2 SOURCES OF OBSERVED MUTANTS

5.2.1 Spontaneous

Among the most important questions raised by these results is the source of mitochondrial mutations in human lung tissue. For several reasons, the most important source(s) of mutations appears to be endogenous. One of the major findings that supports this conclusion was the observation that most of the mutational hotspots observed in human tissue are also observed in human cells grown in culture. Although the mutant fractions are generally significantly lower in the cultured cells, the same mutants are apparently present as confirmed by capillary hybridization.

In addition to the observation that the same mutational hotspots are present both *in vivo* and *in vitro*, further evidence is derived from the observation that the same kinds and positions of hotspot mutations are observed in colon, muscle and lung. Note that the one heart sample investigated is an exception. Our finding that the same mutations were present in multiple tissues, which would be expected to be exposed to a very different array of mutagenic agents but share the possible endogenous sources of mutants, thus supports this conclusion.

Also critical is the conclusion that the mitochondrial mutational spectra in smokers' bronchial epithelial cells are indistinguishable from those in nonsmokers. The link between smoking and lung cancer is among the most clearly demonstrated for exogenous agents to which humans are routinely exposed. Further, multiple known and potent mutagenic agents are present in cigarette smoke. Consequently, if smoking is not the most important contributor to mutagenesis in the bronchial epithelial cells of smokers, then endogenous mutagenic sources are likely to be the most important contributor to mitochondrial mutagenesis in general.

The several hundred-fold higher mutant fractions for these mitochondrial hotspots relative to nuclear hotspots, and to the level of mutagenesis that might be expected to be induced by environmental factors, provides a reasonable explanation for our finding. As discussed in Results, L. Marcelino and colleagues (in prep.) demonstrated that a toxic dose of MNNG would be required to induce mutants of approximately the same frequency as observed in human tissue samples. Consequently, environmental agents may induce mitochondrial mutations, but they apparently do so at levels significantly lower than those that have been induced by endogenous sources.

A critical question in attempting to understand the relationship between exposure to mutagenic agents, adduction, and mutations in mitochondrial and nuclear DNA is the response of polymerase- γ when it encounters adducted DNA. Since there are many copies of mitochondrial DNA per cell, it is not necessary for the cell's survival to replicate each one at every division. In fact, Clayton and colleagues observed that in cells treated with ultraviolet light, mitochondrial DNA copies with pyrimidine dimers were preferentially not replicated (Clayton et al., 1974). Further, in an experiment performed *in vitro* with synthetic oligonucleotides, tetrahydrofuran, an analog of an abasic site in DNA, was found to inhibit elongation by DNA polymerase- γ (Pinz et al., 1995). The findings of L. Marcelino and colleagues described in the Results section showing that a large dose of MNNG induced detectable mitochondrial mutations, while a substantial dose of benzo[a]pyrene diol epoxide did not (Marcelino et al., in prep), might reflect the gamma polymerase falling off the molecule when it confronts bulky adducts, including benzo[a]pyrene diol epoxide. It is also noteworthy that bulky adducts are not repaired in mitochondrial DNA while alkylating damage is repaired (LeDoux et al., 1992). The

distributive nature of polymerase- γ is unlikely to be sufficient to explain completely the lack of smoking-induced mutations, however, because there are high levels of alkylating agents in addition to polycyclic aromatic hydrocarbons in cigarette smoke, and mutations induced by these agents were also not observed.

5.2.2 Endogenous mutagens or replicative errors?

The two most likely sources of endogenous mutations are errors introduced by the mitochondrial gamma polymerase during replication, and mutations that result from *in vivo* replication past adducts generated by reaction with endogenous compounds. Consideration of the types of mutations observed, which are mostly transitions, is not sufficient to support a conclusion. A spectrum dominated by AT→GC and GC→AT transitions is reminiscent of the spectrum produced by the Klenow fragment of DNA polymerase I from *E. coli* (Keohavong et al., 1989). It is also suggestive of the type of spectrum observed in the *lacI* gene in *E. coli* that are defective in mismatch repair (Leong et al., 1986; Schaaper et al., 1987). While mismatch repair has been reported for yeast mitochondrial DNA (Chi et al., 1994), it has not been reported for human mitochondrial DNA. The mutational spectra observed could therefore be plausibly hypothesized to result from errors in DNA replication in the absence of mismatch repair, which could explain not only the predominance of transitions but also the higher mutant fractions in mitochondrial than nuclear DNA.

The mutations observed may also result from reactions with endogenous agents, for instance, reactive oxygen species. Such species would be expected to be more mutagenic to mitochondrial than nuclear DNA based on their high concentrations in the inner mitochondrial membrane. The types of mutations induced by oxidative radicals have been reported in several systems. The *hprt* exon 3 mutational spectrum of hydrogen peroxide-treated TK6 cells included a prominent AT→TA transversion and two less prominent GC→CG transversions (Oller et al., 1992). As will be described further below, AT→TA and GC→CG mutations have been missed with the mutational spectra methodology employed. Grollman and colleagues investigated the types of mutations induced during polymerization past 8-OH-dG constructed into a synthetic template, and discovered that GC→TA mutations were predominantly induced (Shibutani et al., 1991), while Kuchino and colleagues reached a different conclusion, finding that all possible misinsertions across from the adducted nucleotide were observed and further that pyrimidines adjacent to the adducted residue were also misread (Kuchino et al., 1987). Finally, in a study that specifically addressed the behavior of polymerase- γ at the site of 8-OH-dG, C was incorporated into 73% of full length products and A was incorporated into 27%. Insertion of an adenine would lead to GC→TA transversions (Pinz et al., 1995). It is not well-established, however, exactly which adducts are expected to result from oxidative damage. Therefore, although the mutations observed, mainly transitions, are not precisely what would be expected from reaction with reactive 8-OH-dG, it is not yet possible to resolve this question.

5.2.3 Comparison with MNNG-induced mutational spectrum

Another interesting finding is that two of the mutations that are frequently observed in human tissues, p5 and p5.5, were also observed by L. Marcelino and colleagues as mutations induced by the model alkylating agent N-methyl-N'-nitro-N-nitrosoguanidine (MNNG) (Marcelino et al., in prep). Both of these mutations are GC→AT transitions at guanines preceded by adenines, the most frequently mutated site for alkylating agents (Burns et al., 1987). Such findings raise the interesting possibility that endogenous alkylating agents may induce these mutations in human tissues.

5.3 POTENTIAL SOURCES OF ERROR

While the mutational spectra that are produced by the mitochondrial mutational spectra methodology described are, for reasons discussed above, likely to accurately reflect the mutations present in tissue samples, there are several important sources of error in these experiments. Mutant fractions for individual peaks are estimated to be very likely to be correct within a factor of approximately three. None of the major findings and conclusions of this thesis would be altered by minor differences in mutant fraction that would result from the sources of error described below.

5.3.1 Internal standard

Accurate addition of the internal standard requires that the copy number of a given sample is measured precisely before the CDGE step. The difficulty of measuring small amounts of viscous material, like concentrated DNA samples, may lead to some inaccuracies in this measurement. Further, internal standards must be diluted significantly for use as a reference, a process that can lead to imprecision. It is also significant that accurate introduction of the internal standard requires complete restriction digestion. If the restriction digestion is incomplete, then fewer than the expected numbers of copies will migrate in the high T_m mutant range, and the high T_m internal standard will be overintroduced. Fortunately, underrestriction digestion (which was observed in one sample), can be detected when the CDGE gel is stained with ethidium bromide. The sample is observed as a bright band of high molecular weight instead of a uniform smear, and such samples can be discarded from analysis. Generally, when the copy number of a given sample was measured independently, the results were within two-fold, although on occasion variations of three-fold were observed.

5.3.2 Peak measurement and identification

Peak measurement is another factor that can lead to inaccuracies in estimates of mutant fractions. Peaks that are well-separated from other peaks during high-resolution homoduplex runs can be measured accurately, but for peaks that are not well-resolved, estimates of the area under each of the peaks can be difficult. Capillary hybridization can be extremely valuable in such cases in determining accurate mutant fractions, although quantification is better for high T_m than low T_m mutants. Peaks may also be mistaken for an incorrect mutant sequence if they are not well-resolved. Capillary hybridization can clarify essentially all such cases.

Another potential source of error involves differences among the mutants in amplification efficiency. Given that our procedure involves about 10^9 -fold amplification, the relative sizes of the peaks observed depends not only on the original mutant fractions, but also on the relative amplification advantage or disadvantage of a particular mutant (allelic preference). To discover if the mutant fractions observed deviate significantly from the original fractions due to allelic preference, the relative amplification efficiencies of most of the mutants in the standard set were tested. Mutants were mixed together and amplified 1000-fold in the same PCR reaction and the relative sizes of the peaks before and after the amplification were compared, as suggested by Keohavong and Thilly (Keohavong et al., 1989). The amplification efficiency of each of the mutants in the standard set was within $\pm 6\%$ of the efficiency of amplification of the wild-type (0.6 per cycle). Given that the overall amplification of mutants during our procedure is about 10^9 -fold, these differences in amplification efficiencies will lead to errors in the measurement of mutants within a factor of two. Since it is not possible to check the amplification efficiencies of all theoretically possible mutants, some mutants with a severe amplification disadvantage may have originally been present in the samples but were not observed. It is important to note,

though, that the artificial mutant which was created for use as a low T_m internal standard and was never observed in human tissue samples, amplified with an efficiency similar to wild-type.

5.3.3 Missed mutants

Theoretically, constant denaturant electrophoresis should be able to separate all mutants from each other, especially in heteroduplex form (Børresen et al., 1991; Ridanpää et al., 1995). While all heteroduplexes are easily separable from wild-type, some mutants may be lost during the CDGE homoduplex separation that occurs immediately after restriction digestion. Two portions of the constant denaturant gel are excised, one above and one below the wild-type. The area left for the wild-type, approximately 1.2 cm of gel, is expected to contain some homoduplex mutants that have similar melting behavior to the wild-type. Reconstruction experiments have been performed to determine which mutations are likely to be lost because they migrate too close to the wild-type. While all mutations that convert GC->AT base pairs or vice versa have always been detectable with such experiments, some mutations were not. An A deletion in a run of six adenines is represented at about 50% of its original frequency post CDGE under conditions used in these experiments, because it is present at the border of the high T_m region excised. When an AT->TA mutation was tested, this particular mutation was discovered to migrate close to the wild-type and therefore to have been excluded. Mutations that convert a GC base pairs into AT base pairs or vice versa, for both transition or transversion mutations, therefore, would be expected to be included in the spectra we observe. Single base pair deletions might or might not be included. GC->CG and AT->TA transversions have a high probability of migrating too close to the wild-type for inclusion.

5.4 LACK OF SMOKING-INDUCED SPECTRA

No significant increase was observed in the overall mutant fraction or individual mutants in bronchial epithelial cells sampled from smokers compared with nonsmokers for the mitochondrial sequence investigated. Any mutations induced by smoking in lung epithelial mitochondrial DNA must be small ($<10^{-6}$) compared to those observed at frequencies as high as several times 10^{-4} . Yet, as discussed previously, smoking is clearly associated with lung cancer (Doll et al., 1976; General, 1989; Hammond, 1966; IARC, 1986; Wynder et al., 1950). Several possible explanations for these findings are offered and discussed: (1) the sequence is vacant, (2) only a subset of the population is at risk and no members of this subset were sampled, (3) mitochondrial mutations induced by smoking are masked by spontaneous mutations in mitochondrial but not nuclear DNA, and (4) smoking induces lung cancer by a mechanism distinct from mutagenesis.

5.4.1 Sequence is vacant

One possibility is that this particular DNA sequence is "vacant" for smoking-induced mutations. Existing data suggest two ways in which the probability that a given 100 bp is vacant for a given compound can be estimated for nuclear sequences. Extensive experience with the *hprt* gene (Cariello et al., 1990; Chen et al., 1994; Kat, 1992; Keohavong et al., 1992; Oller et al., 1992), has shown that in approximately 15% of cases, a particular compound will not induce a prominent hotspot in a given 100 bp region. The low melting domain of *hprt* exon 3, for instance, is vacant for MNU (Kinkaid, 1993), MNNG (Cariello et al., 1990) and ICR-191 (Cariello et al., 1990). An alternative approach is to determine the frequency of mutations when one compound is used to induce hotspots across an entire gene. Investigations of the mutational spectra of MNNG in MT-1

cells revealed 11 point mutational hotspots in 1025 base pairs scanned, or 1 for every 93 base pairs (Kat, 1992).

The low fraction of GC base pairs in our low melting domain (27%) makes it a relatively poor target for compounds that specifically adduct at guanines. However, MNNG, a model alkylating agent with a preference for adduction at guanines, induced eight hotspot mutations in this same sequence.

It remains theoretically possible that smoking-induced hotspots exist in other regions of mitochondrial DNA, and would be detected by performing similar experiments on a different mitochondrial sequence.

5.4.2 Subset of population is at risk

A second possibility is that only a fraction of individuals are susceptible to smoking-induced mutagenesis. Indeed, while a very high fraction of those who develop lung cancer in industrialized nations are smokers, only a small fraction of all smokers develop lung cancer. Genetic factors may play a role in determining which smokers actually develop tumors (Sellers et al., 1990). An inherited predisposition could exert an influence in any of several aspects of lung cancer susceptibility. Large interindividual differences in the rate of particle clearance from the lungs have been observed (Albert et al., 1969; Wilkey et al., 1980), and studies of monozygotic and dizygotic twins indicate that genetic factors have a major influence on this host defense (Camner et al., 1972). A fraction of the population may also have inherited increased activity of drug-metabolizing enzymes that convert tobacco-derived promutagens to their DNA-binding form (Ayesh et al., 1984; Bartsch et al., 1991; Caporaso et al., 1990), which could affect lung cancer risk. Indeed, the inducibility of aryl hydrocarbon hydroxylase activity in mitogen-stimulated lymphocytes has been associated with a higher risk of lung cancer in smokers (Kärki et al., 1987; Kellerman et al., 1973; Kouri et al., 1982). Further, an increased level of adducts in human lungs has been correlated with a null genotype for the detoxifying enzyme glutathione S-transferase (Shields et al., 1993).

The participants in this study were recruited because they were monozygotic twins discordant for smoking. None of the participants had lung cancer. It could be argued that smoking-induced mutations would have been observed if the study had included members of a smaller, at-risk population. In order to test this hypothesis, healthy lung tissue from smokers who have actually developed lung cancer would need to be tested for smoking-specific mutational patterns.

5.4.3 High spontaneous mutant fraction

A third possibility is that smoking does induce mutations in mitochondrial and nuclear DNA but that in mitochondrial DNA, in contrast to nuclear DNA, those mutations are masked by the generally higher level of spontaneous mutations. Indeed, the mutant fractions detected in these mitochondrial sequences are several hundred-fold higher than expected for nuclear sequences, as discussed previously. Possible explanations include differences in the level of oxidative damage to mitochondrial versus nuclear DNA, or the absence of nucleotide excision repair or potentially mismatch repair in mitochondrial DNA.

The expected mutant fraction of smoking-induced mutations can be estimated based on long-term, low-dose treatment of human cells with benzo[a]pyrene (Chen et al., 1996). Treatment for 20 days with 0.02 μM BaP, a concentration similar to that achieved by smoking 10 cigarettes a day, induced a mutation rate of 8.4×10^{-7} *hprt* deficient mutants per generation. Assuming an individual's lung cells divide three times a year, 20 years of smoking 10 cigarettes per day would imply an induced mutant fraction of 5×10^{-5} for the entire *hprt* gene for benzo[a]pyrene. A corresponding "1% hotspot" would be present at 5×10^{-7} . If the induced level per base pair in mitochondrial DNA were similar to that for

nuclear DNA, then induced mutants at such a level might not have been observed given the much more prominent mutations that result from endogenous mutagens.

5.4.4 Alternatives to mutagenesis

The final possible explanation for the observations reported herein that will be discussed is that smoking induces cancer by a mechanism distinct from induction of point mutations. Two possible pathways are by increasing recombination and by altering cell kinetics.

5.4.4.1 Recombination

Smoking could theoretically cause cancer not via point mutations but by inducing larger scale chromosomal damage, for instance recombination events, that would not have been detected in this assay. As described in the Literature Review: Smoking-induced deletions, the ability of cigarette smoke to induce homologous recombination in the nuclear genome of mice has been reported (Jalili et al., submitted).

Until recently, it was assumed that there is no homologous recombination in mitochondrial DNA. Studies suggest that it is at least an infrequent event. Ohno and colleagues described a patient with both the MELAS point mutation at bp 3243 and a mtDNA deletion (Ohno et al., 1996). The deletion was found only in mtDNA molecules that carried the 3242 mutation and not in those that were wild-type, suggesting that crossover recombination between the deletion and point mutation is rare. However, Thyagarajan and colleagues have recently reported that extracts of mitochondria from human cells contain enzymes that can catalyze the homologous recombination of DNA plasmids (Thyagarajan et al., 1996). These findings raise the possibility that recombination may occur in mitochondrial DNA.

It is possible that the deletions of mitochondrial DNA observed in patients and, at a lower frequency, the elderly (see Literature Review), result from intramitochondrial homologous recombination events. Indeed, the deletions are often flanked by direct repeats and have been shown to increase with age (Ashkenas, 1997; Ozawa, 1995). As discussed previously, these deletions have been reported to be elevated in smokers compared with nonsmokers (Ballinger et al., 1996; Liu et al., 1997).

5.4.4.2 Cell kinetics

Another possibility is that smoking leads to cancer by increasing cell proliferation (Ackerknecht, 1953). Histological analysis has documented a significant increase in basal cell hyperplasia among smokers (Auerbach et al., 1961), and that the extent of hyperplasia is associated with the number of cigarettes smoked per day. In a study of over 22,000 slides of lungs from autopsy samples, the percentage of slides with 3 or more cell rows (hyperplasia) was 9.4% for never smokers, 35.5% for those who smoked <0.5 packs per day, 40.3% for 0.5-1 pack per day smokers, 63.8% for 1-2 packs per day smokers, and 76.2% for 2 or more packs per day smokers. In addition to observed hyperplasia, increased cell turnover has been documented in rodent cells in response to smoking. [³H]-thymidine labeling of hamster epithelial cells exposed to cigarette smoke revealed increased mitotic activity after exposure (Boren, 1970). These findings support the conclusion that smoking causes cell proliferation.

Compounds that affect cell turnover can directly affect the probability of developing cancer. P. Herrero and W. Thilly have been developing mathematical models to describe cancer development that include explicit terms for the rates of cell division, cell death and mutation. The model assumes that two mutations are sufficient to create an "initiated" cell that gives rise to a dysplastic colony, and a third mutation is sufficient for the

transformation to a carcinoma. Cell turnover parameters are critical in determining the probability that an initiated cell survives stochastic extinction. Cell division and death rates are also used to model the amount of time necessary for a colony of initiated cells to obtain a third mutation that would convert it into a carcinoma. The higher the division rate, the shorter the period of time required for progression.

In conjunction with the mathematical models, W. Thilly and colleagues have proposed to measure the cell division and death rates in healthy bronchial epithelium of smokers and nonsmokers. Such experiments have already been performed in normal colon epithelium, colon adenomas and colon carcinomas by Dr. E. Furth (University of Pennsylvania). Expressed in rates of division or death per cell year (α and β , respectively), normal colonic epithelium has equal death and division rates of 2.92 per stem cell-year. Adenomas have higher division (α) and death rates (β) of approximately 8.8, with the division rate being slightly greater than the rate of cell death. Carcinomas do not have an increased rate of cell death as compared with adenomas, but the rate of cell division is increased to about 30 per year, which implies substantial growth for a colony. Once these values were incorporated into the equation describing the probability of cancer at a given age, the mutation rates for the first two initiating mutations could be estimated and these were discovered to correlate well with expectations (3×10^{-7}) (Robinson et al., 1994; Grist et al., 1992).

P. Herrero and W. Thilly have applied similar reasoning to the data for lung cancer mortality in smokers and nonsmokers. By assuming that the rate of cell death is the for lung as colon, and applying the same mathematical model to lung cancer data for cohorts of individuals born before and after cigarette smoking became widespread (see Figure 11), they estimated the parameter ($\alpha-\beta$), which represents net growth of an "initiated" colony. For females, the value ($\alpha-\beta$) increased from 0.12 before widespread introduction of cigarettes to 0.19 after cigarettes were generally available; for males, the increase was from 0.15 to 0.20. This suggests that smoking may play a role in inducing lung cancer via an effect on cell kinetics.

If smoking were affecting cell turnover, then two changes actually observed in the age-specific lung cancer death rates would be predicted. First, by increasing the probability that an initiated cell survives to become a dysplastic colony, the probability of having cancer would be greater at any given age. This change is clearly observed as cohorts are compared progressing to those born most recently. Secondly, an increase in the number of cell divisions per doubling of the dysplastic colony would be manifested by a shift in the curve towards having cancer at earlier ages, a change that is also observed. It should be pointed out, however, that modeling the lung cancer data also led to the conclusion that smoking increases the mutation rate for the initiating mutation as well (from 1.2×10^{-6} mutations/cell-year in the target gene before cigarettes to 1.8×10^{-6} mutations/cell-year in the target gene after cigarettes for females, and from 1.1×10^{-6} mutations/cell-year in the target gene before cigarettes to 1.6×10^{-6} mutations/cell-year in the target gene after cigarettes for males). Yet, an increase in the mutation rate for initiating mutations could explain the higher probability of cancer at a given age, but would not be sufficient to explain the shift in the curve. Experiments to determine empirically the rate of cell division and death in human bronchial epithelial cells from smokers and nonsmokers will aid in assessing the contribution, if any, that changes in cell turnover make to lung cancer incidence.

5.5 DISTRIBUTION OF MUTANTS

An important observation that resulted from these studies was that a particular mutant in a given sample could, on occasion, be observed at a mutant fraction significantly above the mean for that mutant. One possible explanation for these findings is discussed.

5.5.1 Mutant-rich clusters

After determining the mean mutant fraction for each of 16 mutants, the number of samples containing a particular range of mutant fractions was plotted against the fraction of the mean mutant fraction for that mutant (see Figure 37 in Results). In the vast majority of cases (551/570), the mutant fraction for a given sample was within 4-fold of the mean for that mutant. There were 10 cases (out of 570 observations) in which the mutant fraction for a given sample was between 5 and 11-fold above the mean. Such cases are unlikely to be attributable to errors in the introduction of the internal standard or statistical fluctuation during analysis due to their large size and because they were consistently observed in duplicate.

Since this variability was observed even within a single organ and could not be attributed to smoking status, gender, position within the lung, or any other trait, small samples are likely to be the source of the variability. If mitochondrial mutants were distributed randomly among cells, it would be extremely unlikely that mutant fractions at 5-11-fold the mean would be observed simply as a result of sampling. For a mutant with a mean mutant fraction of 5×10^{-5} , a sample of 10^6 cells containing 400 mitochondrial DNA copies per cell with this mutant represented at 8-fold the mean would contain $8 \times (5 \times 10^{-5}) \times 400 \times 10^6 = 1.6 \times 10^5$ copies. The average number of mutants per sample would be 2×10^4 copies. If this were a normal distribution (which it clearly is not), then the standard deviation would be the square root of 2×10^4 or 141 copies. In such a case, the outlier sample would be 993 standard deviations above the mean. Clearly, the default hypothesis of random mutant distribution must be rejected.

Assuming that mitochondrial mutants may be clustered within cells makes such an outcome more likely. Since fewer mutant-rich cells would be required to have been sampled in order to generate the elevated mutant fractions observed, the variance would be expected to be higher.

The accumulation of mitochondrial mutants in a particular cell can be pictured as follows: Consider a cell containing a mitochondrial DNA mutation in the target sequence in one of several hundred copies of mitochondrial DNA. When the mitochondrial DNA in this cell is replicated, on average, a cell containing two mutant mitochondrial DNA copies will result. When the cell divides, the two copies may separate such that each of the two daughter cells receives a single copy. Alternatively, one daughter cell may receive both copies. If this process continues over many generations, by chance, a small number of cells containing many mitochondrial DNA copies would be expected to accumulate over time.

It is not possible to estimate the size of such a colony without making assumptions about the number of mutant copies per mutant-rich cell. Assuming 400 mutant copies per cell, the size of the colonies for mutants 5- to 11-fold higher than the mean are generally in the range of several hundred cells, depending upon the mean mutant fraction and the number of cells used for the DNA isolation. Returning to the example above, if an excess of cells containing 400 mutant copies were included in one sample of 10^6 cells causing it to have a mutant fraction 8-fold times the mean, then this would require an extra 350 completely mutant cells.

We suggest that this could occur if a bronchial epithelial stem cell accumulated many copies of a particular mitochondrial mutant. Cells derived from such a mutant-rich cell would be expected also to contain many mitochondrial mutants. Thus, specific DNA

preparations may have been observed to include an unusually large mutant because a colony of mutant-rich cells was sampled.

Clustering of nuclear mutants in solid tissues has been observed previously and presumed to result from mutant stem cells. Dr. R. Cha divided mammary gland epithelium into sections and used allele-specific PCR [the mismatch amplification mutation assay (MAMA)] to measure the levels of H-ras mutants in DNA from each section (Cha et al., 1994). Some sectors contained significantly higher levels of the mutant, suggesting the possibility that a cluster of mutant cells derived from a stem cell containing an H-ras mutant was sampled. Clustering of mutant cells likely to have derived from a mutant stem cell has also been observed by immunohistochemical staining for p53 in epidermis (Jonason et al., 1996). Further, Z.-Y. Chen, H. Zarbl and K. Hong have detected sections of bronchial epithelial cells from autopsy samples containing unusually high levels of a K-ras mutation also by allele-specific PCR (MAMA).

Finally, there is evidence for a similar phenomenon from heteroplasmic mice carrying two mtDNA genotypes that differ from each other at 106 bases (Jenuth et al., 1997). The mtDNA genotypes in colonic crypts, which are each derived from the clonal expansion of a single founder cell, were examined. In crypt samples from two heteroplasmic animals aged 4 and 15 months, the mean proportion of the BALB haplotype was approximately 4% in the population of crypts from both animals. There were, however, significant differences in the distribution of frequencies of the two mtDNA genotypes among individual crypts in the young and old animal (see Figure 42). In individual crypts from the 4-month-old animal, the proportion of the BALB haplotype varied from 0 to 25%. In contrast, two distinct populations were observed in the 15-month-old animal; most crypts had lost the BALB genotype completely (0% BALB), but a small number of crypts contained predominantly (>50%) BALB mtDNA. These findings support the following model. A stem cell containing a small number of BALB mtDNA copies has some chance of losing each BALB copy to the daughter cell that becomes a transition cell and ultimately a terminally differentiated cell, with each division. While most stem cells lost all of the lower frequency BALB mtDNA copies by 15 months of life, a small fraction of stem cells by chance retained the less frequent BALB genotype and accumulated many mtDNA mutants.

5.5.2 Modeling mutant accumulation

Modeling the expected number of mutant-rich clusters in bronchial epithelial cells was undertaken by P. Herrero. He assumed that there are 256 cells in a "turnover unit", all of which derive from a single stem cell. The cells are arranged in a pyramidal structure such that there is 1 stem cell, and 1 transition cell in the first layer, then 2 cells in the second layer, 4 cells in the third layer and so on, up to 128 cells in the final layer. One cell from the 128 cell layer dies each day. When 2 cells from the outermost, 128 cell layer die, then a cell from the 64 cell layer must divide to replace them. When 2 cells from the 64 cell layer have divided, a cell from the 32 cell layer must divide to replace them. And so on. When the stem cell divides, it creates another stem cell and a transition cell. The model developed by P. Herrero also assumes a mutation rate of 2×10^{-6} , that there are 400 mtDNA copies per cell, and that mitochondrial mutants are distributed randomly to cells, that is, once a cell contains two mitochondrial mutants, there is a 50% chance that one mutant copy will be transmitted to each daughter cell, and a 50% chance that both mutants will go to a single daughter cell. Using these assumptions, P. Herrero calculated the probability that a stem cell contained different amounts of mutants after various generation times. The results are given in Fig. 43. After 150 generations (approximately 50 years), the probability that a cell contained greater than 300 mutants (integration of the area under 300 to 400 mutants) under these assumptions is approximately 10^{-9} .

Fig. 42: Distribution of mtDNA genotypes in colonic crypts.

Frequency histogram of the distribution of mtDNA genotypes in individual colonic crypts from animals 4 and 15 months of age.

Source: Jenuth et al., 1997.

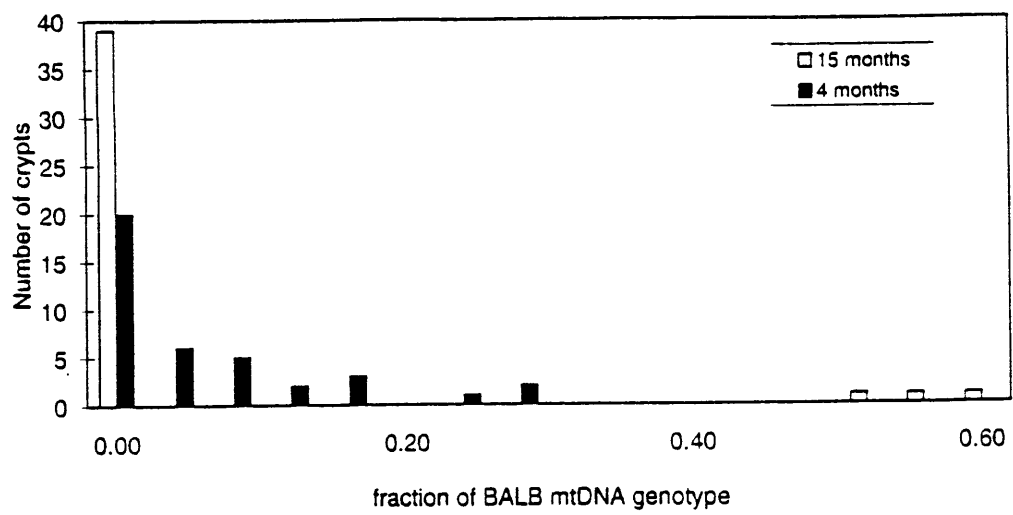
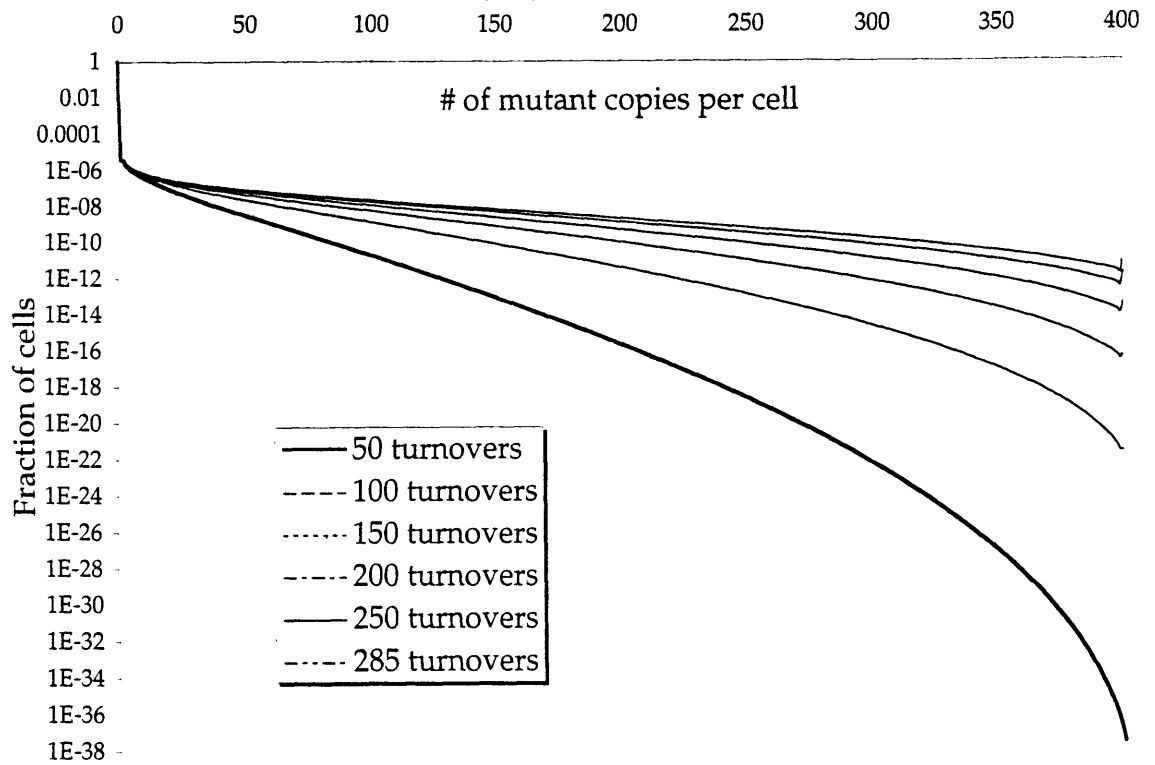


Fig. 43. Modeling the accumulation of mitochondrial mutants.

The probability of a cell containing different numbers of mitochondrial mutants was determined for different numbers of generations. Assumptions included 256 cells per turnover units, 400 mtDNA copies per cell and random segregation of mutant copies with cell division.

Source: P. Herrero

Distribution of number of mutants in stem cell population



Mutational rate = 2×10^{-6}

This result can be compared to that observed empirically. Assuming 10^6 cells per sample and that there were approximately $10^6/128 = 8000$ stem cells per sample, largely mutant stem cells were observed in the following fraction of instances: $12 \text{ (high mutant observations)} / (570 \text{ total observations} \times 8000 \text{ cells} \times 16 \text{ hotspots}) = 2 \times 10^{-7}$. This number is 200-fold higher than the expected value, although it should be recognized that a large number of assumptions were made.

Two possible explanations for the greater than expected likelihood of a sample containing a mutant with an unusually large mutant fraction are offered. One possibility is that the model does not accurately reflect reality because, when a mitochondrial mutant divides to become two mitochondrial mutants, the probability that both mutants are transmitted to the same daughter cell upon cell division is greater than 50%. For instance, if mitochondria as opposed to mitochondrial DNA were the segregating unit, as has been substantiated experimentally (Attardi et al., 1995), segregation would be significantly more rapid. Another possibility is that there is some selection for specific mitochondrial mutants as opposed to wild-type. The data available are not sufficient to distinguish between these two possibilities. Information about the phenotypic effects of cells containing many copies of a particular mitochondrial mutant could shed light on which of these mechanisms is responsible. Selection would only be expected to have an effect on mutants that confer a phenotypic effect, while all mutants, including silent mutants, would be expected to be affected by nonrandom segregation during cell division. Mutant p19 is therefore of particular interest since it is a silent substitution and leads to no change in the amino acid coding sequence of the NADH dehydrogenase subunit III protein. The histogram for p19 is provided in Figure 36 in Results: Variability in Mutant Fractions. Note that no samples with unusually large p19 mutant fractions have been observed, thus making either of the offered explanations possible.

5.5.3 Mitochondrial proliferation?

Among the mutants in the histogram shown in Figure 37 are two samples in which a particular mutant is present at a value very significantly above the mean for that mutant in all other samples. Mutant p7 and mutant p16 were each observed in one case at >40-fold the mean. In both instances, the unusually large frequencies were observed in duplicate and the identity of the peaks in both samples was confirmed. These mutants are significantly larger than would be expected from a completely mutant turnover unit and are therefore hypothesized to result from proliferation of the mitochondria (and consequently the mutant mitochondrial DNA therein) in a number of cells.

Cells with large numbers of mitochondria have been observed in muscle as ragged red fibers and often lack cytochrome-c-oxidase activity. These may result when cells attempt to compensate for reduced oxidative potential by stimulating mitochondrial proliferation (Tremblay, 1969). In animal models, skeletal muscle mitochondria proliferate in response to increased oxidative demand (Pette et al., 1985; Salmons et al., 1981; Williams et al., 1986). Further, Shoubridge and colleagues demonstrated by *in situ* hybridization that some of the muscle fibers of two mitochondrial myopathy patients contained 6- to 7-fold higher levels of deleted mtDNA than observed in adjacent, normal fibers (Shoubridge et al., 1990). Finally, oncocytes, epithelial cells with a high content of often abnormal mitochondria, are found in nearly all epithelial organs, especially with advancing age (Müller-Höcker, 1992).

A cluster of cells with proliferation of mitochondria in response to lack of oxidative phosphorylation is one possible explanation for the two cases of outstanding mutants. This proliferative response would require that the mutations cause a phenotypic change. Mutant p7 represents the only example of a nonsense mutation, and would lead to a significantly truncated and presumably nonfunctional NADH dehydrogenase subunit III protein. The effect on the enzyme complex is unknown. For p16, the mutation is a missense mutation

and its effects on the protein are unknown. It is interesting to note that peak h from Fig. 33, which depicts the mutational spectra in muscle tumor slices, is likely to represent p16 and is very large in the sample in panel C.

5.6 AGING AND MITOCHONDRIAL DNA

5.6.1 Deletions

As described previously, mitochondrial deletions increase in variety and frequency with age in humans and other species, especially in postmitotic tissues. There are several, not incompatible, possible mechanisms for this increase with age. One option is that the deletions are induced independently, and that they are observed to increase supralinearly with age because the causative factor increases substantially with age. Alternatively, selection for deleted mitochondrial DNA copies may exist.

Several researchers have suggested that mitochondrial deletions represent a manifestation of a vicious cycle and that aging is a result of accumulated damage to mitochondria from free radicals [for example, (Ames et al., 1995)]. The model usually includes the following argument. Oxidative free radicals, including superoxide and hydroxyl radicals, react with mitochondrial DNA and induce mutations. The mutations induced would be expected to adversely affect mitochondrially encoded proteins, either by causing mutations in the proteins themselves or by affecting their translation via mutations in mitochondrially encoded tRNA. If a protein in the electron transport pathway is damaged, then the enzyme complexes upstream of the damage will be in a reduced state, and more likely to contribute electrons to oxygen to form superoxide (Bandy and Davidson, 1990). Increased levels of superoxide would produce even more free radicals and more damage to mitochondrial DNA. Hence, a vicious cycle of increased free radicals and damage, each accelerating the other, would result. Supporting this argument is the observation that the levels of free radicals in postmitotic cells increase with age (Mecocci et al., 1993). Also, caloric restriction decreases the levels of free radicals in muscle tissue (Dubey et al., 1996; Sohal et al., 1996) and decreases the levels of deleted mitochondrial DNA molecules (Aspnes et al., 1997), in addition to delaying aging.

A second possibility is that mitochondrial deletions increase with age because there is a mechanism that selects for these deletions. Critical findings supporting selection for preexisting mutants rather than the generation of new mutants are that the mitochondria of a single respiration-deficient muscle fiber carry the same mutation, and different fibers carry different mutations (Müller-Höcker et al., 1993). This observation supports a model in which a single mutant molecule is amplified via selection. It has been suggested that the selection may occur because deleted copies of mitochondrial DNA have an advantage during replication due to their shorter length (Wallace, 1989), or that *cis*-acting elements within mitochondrial DNA responsible for regulating replication are deleted (Shoubridge et al., 1990). Both of these hypotheses require deletions for a selective advantage. Therefore, the observation that 5 of 13 human myoblast cell lines carrying the mtDNA point mutation responsible for MELAS were discovered to undergo a rapid shift of their genotype toward the pure mutant type (Yoneda, 1992), while the other 8 cell lines, 6 of which were almost completely homoplasmic, maintained a stable genotype, argues that the hypotheses that require the mutation to be a deletion may not be correct.

Another possible explanation for selection of mitochondrial mutants is that loss of mitochondrial function may confer a selective advantage. Selection could take the form of a positive selection, for instance, individual organelles that do not have functioning mitochondria could be replicated more frequently, possibly as a proliferative response as described in the section in the Discussion on Mitochondrial Proliferation. Alternatively, there may be selection against wild-type organelles. For instance, de Grey has argued that mitochondria with reduced respiratory function due to a mutation or deletion affecting the

respiratory chain may suffer less frequent lysosomal degradation because they inflict free radical damage more slowly on their own membranes (de Grey, 1997).

5.6.2 Point mutations

As described previously, several authors have attempted to measure the relationship between the frequency of point mutations in mitochondrial DNA and age (see Literature Review) (Münscher et al., 1993; Pallotti et al., 1996), and the results are unclear. The methodology described herein should be valuable for determining the relationship between mitochondrial point mutations and aging, given the possibility of measuring mutations, including silent mutations, at levels as low as 10^{-6} from any tissue sample.

While the results are very preliminary, multiple samples of the lungs of an 84 year old showed point mutational spectra in this sequence very similar to those observed in 50-year olds. While there may be a modest increase with age in point mutational spectra in bronchial epithelial cells, there does not appear to be a dramatic increase in frequency with age as observed for deletions in post-mitotic cells. These findings make unlikely a model in which aging in all tissues is characterized mainly by a vicious cycle of increased free radicals contributing to increased point mutations in mitochondrial DNA.

The analysis of a single heart sample from a 77 year old provided an unexpected result: two extremely frequent mutations were seen, neither of which is among those routinely observed. These mutations may be the result of a selection process in heart tissue for mitochondrial mutants.

6. CONCLUSIONS

Similar mitochondrial mutations were observed repeatedly in different samples but at different frequencies.

Mutant frequencies varied considerably even from closely spaced positions within a single organ.

Smoking did not induce new, prominent mitochondrial mutants.

Smoking did not increase overall mitochondrial mutant fractions in this sequence.

None of the mutants of the that were consistently observed in human tissue samples were elevated in smokers.

The large variability in the size of individual mutants may reflect the accumulation of mitochondrial mutants in bronchial epithelial stem cells, and in two extreme cases, proliferation of mutant mitochondria.

Mitochondrial point mutants do not appear to increase supralinearly with age in human bronchial epithelial cells, but selection for point mutants may occur in post-mitotic tissues.

7. SUGGESTIONS FOR FUTURE RESEARCH

7.1 SOURCES OF HUMAN MUTATIONS

The discovery that mitochondrial mutations are likely to result largely from endogenous agents suggests that one or both of two major pathways may be critical for mitochondrial mutagenesis: replication errors and endogenous damage. Given that a set of specific mutants arise as hotspots in human tissues, it is possible to test whether the mitochondrial DNA polymerase- γ induces these same mutations *in vitro*. W.-M. Zheng in the W. Thilly laboratory has amplified purified wild-type copies of the mitochondrial target sequence used in these experiments with a preparation of polymerase- γ generously provided by Dr. William Copeland (National Institute of Environmental Health Science)[see (Ropp et al., 1996) for cloning of the gene]. Capillary electrophoresis has been used to enrich for mitochondrial mutants generated during amplification. Characterization of the mutational spectra induced by polymerase- γ on this sequence will be performed as for *Pfu* DNA polymerase (André et al., 1997). These experiments may clarify whether the specific mutations observed as mitochondrial hotspots in human tissue samples are likely to result from DNA polymerase errors.

Further, the suggestion that endogenous factors may be the most important mutagenic pathway in mitochondrial DNA raises (but certainly does not prove) the possibility that the most important pathways in nuclear mutagenesis are spontaneous as well. This assertion is supported by analysis of databases for *hprt* and *adenomatous polyposis coli* (APC), which suggest that polymerase-mediated slippage events (Kunkel, 1990) and mutations at CpG sites (Coulondre et al., 1978) are primary mutagenic pathways. One approach to assessing this possibility is to test human DNA polymerase- δ , the putative replication polymerase for human nuclear DNA, to determine whether its replicative hotspots correlate with the mutagenic hotspots observed *in vivo*. B. Muniappan of the Thilly laboratory is planning to use this enzyme to amplify the APC gene, which is a particularly good sequence for such studies, as several prominent hotspot mutations have been observed in colon cancer. A correlation between the *in vivo* mutations observed in APC and the replicative hotspots induced by pol- δ would suggest that pol- δ errors are the source of such mutations.

Another approach to studying whether spontaneous pathways are responsible for the mutations that occur in humans is to determine the *in vitro* mutational spectrum of a single gene in several different organisms. A group led by A. Tomita of the W. Thilly laboratory in collaboration with the Demple and Walker laboratories is studying the spontaneous hotspot mutations observed in *hprt* exon 3 in cultured human cells, and bacteria and yeast with *hprt* exon 3 introduced into their genomes, in order to compare the spontaneous hotspots observed in the three different model systems. A. Tomita's determination of the *in vitro* *hprt* mutational spectrum can be correlated with the *in vivo* data compiled in available databases.

7.2 MITOCHONDRIAL MUTANTS AND AGE IN DIFFERENT TISSUES

The ability to measure point mutations in mitochondrial DNA provides an opportunity to monitor the frequency of these mutations with age. Also, eliminating phenotypic selection implies that the mutational spectra can be measured in a variety of tissues, including post-mitotic tissues. The ability to measure mutants in any tissue is critical given the many previous observations that different tissues accumulate mitochondrial mutations at different rates. K. Khrapko has been appointed to the faculty at the Harvard Medical School Division of Aging and will continue these investigations. He

will test whether point mutations increase with age in humans, and whether there are significant differences among tissues. Which specific mutations, if any, increase with age will be determined to define the implications for the mechanism of their generation. He will also investigate the mitochondrial mutational spectra of caloric-restricted rats in order to better understand the effect of the age-lengthening procedures on mitochondrial point mutations. The ultimate goal will be to differentiate whether an increase in mitochondrial mutations actually causes aging, or is merely correlated with aging.

7.3 MECHANISM OF SMOKING-INDUCED LUNG CANCER

7.3.1 Nuclear mutational spectra

One of the most important questions left unresolved by this work is how smoking causes cancer. Direct experimentation to determine whether smoking causes cancer by inducing mutations in critical nuclear sequences, but does not induce mutations above the spontaneous level in mitochondrial DNA, would be very valuable. Denissenko and colleagues showed that the hotspots of the p53 adduct spectra formed when BPDE is permitted to react with bronchial epithelial cells in culture are the same as the hotspots observed in the p53 gene of smokers' tumors (Denissenko et al., 1996). Such findings suggest but do not prove that smoking induces mutations that lead to cancer in critical tumor suppressors. Further studies determining the mutational spectra in nuclear sequences of healthy bronchial epithelial cells in smokers and nonsmokers will be undertaken by K. Hong of the W. Thilly laboratory.

While the methodology described herein is sufficient for identifying and measuring hotspot point mutations in mitochondrial DNA, further technical improvements are needed to measure mutational spectra in nuclear DNA, where mutant fractions and the number of target sequence copies per cell are both lower. X.-C. Li of the W. Thilly laboratory has developed a method using sequence-specific hybridization coupled with a biotin-streptavidin capture system on magnetic beads that permits greater than 10,000-fold enrichment of a target sequence from total genomic DNA with a 70% yield (Li, 1997). The first enrichment for mutants can then be performed using a wide-bore capillary separation (Li et al., 1996). A sensitivity of 10^{-6} has already been demonstrated.

7.3.2 Recombination

The recent discovery that mitochondrial DNA is capable of performing homologous recombination in cell extracts (Thyagarajan et al., 1996), combined with the observation that a common mitochondrial deletion may be more prominent in smokers as compared with nonsmokers (Ballinger et al., 1996; Liu et al., 1997), has raised the possibility that smoking induces deletions in mitochondrial DNA by stimulating homologous recombination. A. Kim of the W. Thilly laboratory is planning experiments to investigate the role of homologous recombination in the development of human tumors. She will specifically address whether there are hotspots for chromosomal rearrangement throughout the genome, as there are mutational hotspots.

7.3.3 Cell kinetics

Another possibility is that smoking does not induce cancer via a mutagenic mechanism but by affecting cell kinetics, that is, the rate of cell division and cell death in bronchial epithelial cells. As discussed previously, P. Herrero and W. Thilly have developed a mathematical model that separates the variables of mutation rate, mitotic rate and apoptotic rate. Computer simulations with the data on age-specific lung cancer rates in the U.S. prior to and after the widespread introduction of cigarettes suggests that smoking

has an effect on the factor $(\alpha - \beta)$, that is, the rate of cell division minus the rate of cell death. Through a collaboration with Dr. E. Furth (Univ. of Penn.) who has already measured mitotic and apoptotic rates in colon, the rates of cell division and death will be measured *in vivo* in normal human lung epithelium and preneoplastic lesions in smokers and nonsmokers. This approach may clarify whether smoking affects the probability of lung cancer by increasing the mutation rate in genes that would lead to cancer by influencing cell kinetics or both.

LITERATURE CITED

- Aaltonen, L. A., P. Peltomäki, F. S. Leach, P. Sistonen, L. Pylkkänen, J.-P. Mecklin, H. Järvinen, S. M. Powell, J. Jen, S. R. Hamilton, G. M. Petersen, K. W. Kinzler, B. Vogelstein, and A. de la Chapelle. (1993). Clues to the pathogenesis of familial colorectal cancer. *Science* **260**: 812-816.
- Abravaya, K., J. J. Carrino, S. Muldoon, and H. H. Lee. (1995). Detection of point mutations with a modified ligase chain reaction (Gap-LCR). *Nucl. Acids Res.* **23**: 675-682.
- Ackerknecht, E. H. Rudolph Virchow. Madison, WI: University of Wisconsin Press, 1953.
- Agarwal, S., and R. S. Sohal. (1994). DNA oxidative damage and life expectancy in houseflies. *Proc. Natl. Acad. Sci. USA* **91**: 12332-12335.
- Agency for Toxic Substances and Disease Registry. (1993). Polycyclic aromatic hydrocarbon toxicity. *Am. Fam. Phys.* **47**: 623-624.
- Albert, R. E., M. Lippmann, and W. Briscoe. (1969). The characteristics of bronchial clearance in humans and the effects of cigarette smoking. *Arch. Environ. Health* **18**: 738-755.
- Albertini, R. J. (1985). Somatic gene mutations *in vivo* as indicated by the 6-thioguanine-resistant T-lymphocytes in human blood. *Mutat. Res.* **150**: 411.
- Albertini, R. J., K. L. Castle, and W. R. Borchering. (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., J. P. O'Neill, J. A. Nicklas, N. H. Heinz, and P. C. Kelleher. (1985). Alterations of the *hprt* gene in human *in vivo*-derived 6-thioguanine-resistant T lymphocytes. *Nature* **316**: 369-371.
- Albertini, R. J., L. M. Sullivan, J. K. Berman, C. J. Green, J. A. Stewart, J. M. Silveira, and J. P. O'Neill. (1988). Mutagenicity monitoring in humans by autoradiographic assay for human T lymphocytes. *Mutat. Res.* **204**: 481-92.
- Ames, B., M. Shigenaga, and T. Hagen. (1995). Mitochondrial decay in aging. *Biochim. Biophys. Acta* **1271**: 165-170.
- Ammenheuser, M., A. B. Berenson, N. J. Stiglich, E. B. J. Whorton, and J. B. J. Ward. (1994). Elevated frequencies of *hprt* mutant lymphocytes in cigarette-smoking mothers and their newborns. *Mutat. Res.* **304**: 285-94.
- André, P., A. Kim, K. Khrapko, and W. G. Thilly. (1997). Fidelity and mutational spectra of *Pfu* DNA polymerase. *Genome Meth.* **7**: 843-852.
- Aquadro, C. F., and B. D. Greenberg. (1983). Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences from seven individuals. *Genetics* **103**: 287-312.

Armstrong, J. D., and B. A. Kunz. (1990). Site and strand specificity of UVB mutagenesis in the SUP4-o gene of yeast. *Proc. Natl. Acad. Sci. USA* **87**: 9005-9009.

Ashkenas, J. (1997). Homologous recombination in human mitochondria? *Am. J. Hum. Genet.* **61**: 18-22.

Ashley, D. J., and H. D. Davies. (1967). Cancer of the lung histology and biology behaviours. *Cancer* **20**: 165-174.

Aspnes, L. E., C. M. Lee, R. Weindruch, S. S. Chung, E. B. Roecker, and J. M. Aiken. (1997). Caloric restriction reduces fiber loss and mitochondrial abnormalities in aged rat muscle. *Faseb J.* **11**: 573-81.

Attardi, G., M. Yoneda, and A. Chomyn. (1995). Complementation and segregation behavior of disease-causing mitochondrial DNA mutations in cellular model systems. *Biochim. Biophys. Acta* **1271**: 241-248.

Auerbach, O., A. P. Stout, C. Hammond, and L. Garfinkel. (1961). Changes in bronchial epithelium in relation to cigarette smoking and in relation to lung cancer. *New Eng. J. Med* **265**: 253-267.

Ayesh, R., J. R. Idle, J. C. Ritchie, M. J. Crothers, and M. R. Hetzel. (1984). Metabolic oxidation phenotypes as markers for susceptibility to lung cancer. *Nature* **312**: 169-170.

Balansky, R., A. Izzotti, L. Scatolini, F. D'Agostini, and S. De Flora. (1996). Induction by carcinogens and chemoprevention by N-acetylcysteine of adducts to mitochondrial DNA in rat organs. *Cancer Res.* **56**: 1642-1647.

Ballinger, S. W., T. F. Boudier, G. S. Davis, W. A. Judice, J. A. Nicklas, and R. J. Albertini. (1996). Mitochondrial genome damage associated with cigarette smoking. *Cancer Res.* **56**: 5692-5697.

Bandy, B., and A. J. Davison. (1990). Mitochondrial mutations may increase oxidative stress: implications for carcinogenesis and aging? *Free Radical Biol. Med.* **8**: 523-539.

Barany, F. (1991). Genetic disease detection and DNA amplification using cloned thermostable ligase. *Proc. Natl. Acad. Sci. USA* **88**: 189-193.

Bartsch, H., S. Petruzzelli, S. De Flora, E. Hietanen, A. M. Camus, M. Castegnaro, O. Geneste, A. Camoirano, R. Saracci, and C. Giuntini. (1991). Carcinogen metabolism and DNA adducts in human lung tissues as affected by tobacco smoking or metabolic phenotype: A case-control study on lung cancer patients. *Mutat. Res.* **150**: 103-113.

Belinsky, S. A., T. R. Devereux, J. F. Foley, R. R. Maronpot, and M. W. Anderson. (1992). Role of the alveolar type II cell in the development and progression of pulmonary tumors induced by 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone in the A/J mouse. *Cancer Res.* **52**: 3164.

Belinsky, S. A., T. R. Devereux, R. R. Maronpot, G. C. Stoner, and M. W. Anderson. (1989). The relationship between the formation of promutagenic adducts and the activation of the K-ras proto-oncogene in lung tumors from A/J mice treated with nitrosamines. *Cancer Res.* **49**: 5305.

- Belinsky, S. A., C. M. White, J. A. Boucheron, F. C. Richardson, J. A. Swenberg, and M. Anderson. (1986). Accumulation and persistence of DNA adducts in respiratory tissue of rats following multiple administrations of the tobacco specific carcinogen 4-(N-methyl-N-nitrosamino)-1-(3-pyridyl)-1-butanone. *Cancer Res* **46**: 1280.
- Benzer, S. (1961). On the topography of the genetic fine structure. *Proc. Natl. Acad. Sci. USA* **47**: 403-415.
- Benzer, S., and E. Freese. (1958). Induction of specific mutations with 5-bromouracil. *Proc. Natl. Acad. Sci. USA* **44**: 112.
- Bigbee, W. L., R. G. Langlois, M. Swift, and R. H. Jensen. (1989). Evidence for an elevated frequency of *in vivo* somatic cell mutations in ataxia telangiectasia. *Am. J. Hum. Genet.* **44**: 402-408.
- Blenkinsopp, W. (1967). Proliferation of respiratory tract epithelium in the rat. *Exp. Cell Res.* **46**: 144-154.
- Bodenteich, A., L. G. Mitchell, and C. R. Merrill. (1991). A lifetime of retinal light exposure does not appear to increase mitochondrial mutations. *Gene* **108**: 305-310.
- Boers, J., J. den Brok, J. Koudstaal, J. Arends, and F. Thunnissen. (1996). Number and proliferation of neuroendocrine cells in normal human airway epithelium. *Am. J. Respir. Crit. Care Med.* **154**: 758-763.
- Boffoli, D., S. C. Scacco, R. Vergari, G. Solarino, G. Santacrose, and S. Papa. (1994). Decline with age of the respiratory chain activity in human skeletal muscle. *Biochim. Biophys. Acta* **1226**: 73-82.
- Bogenhagen, D., and D. A. Clayton. (1977). Mouse L cell mitochondrial DNA molecules are selected randomly for replication throughout the cell cycles. *Cell* **11**: 719-727.
- Bogovski, P., and S. Bogovski. (1981). Animal species in which N-nitroso compounds induce cancer. *Int. J. Cancer* **27**: 471-474.
- Bolhuis, P. A., E. M. Bleeker-Wagemakers, N. J. Ponne, M. J. Van Schoneveld, A. Westerveld, C. Van den Bogert, and H. F. Tabak. (1990). Rapid shift in genotype of human mitochondrial DNA in a family with Leber's Hereditary Optic Neuropathy. *Biochem. Biophys. Res. Commun.* **170**: 994-997.
- Boren, H. G. "Pulmonary cell kinetics after exposure to cigarette smoke." *Inhalation Carcinogenesis*. Oak Ridge, TN: U.S. Atomic Energy Commission, 1970. 229-241.
- Boring, C. C., T. S. Squires, T. Tong, and S. Montgomery. (1994). Cancer statistics. *CA Cancer J. Clin.* **44**: 7-26.
- Børresen, A.-L., E. Hovig, B. Smith-Sorensen, D. Malkin, S. Lystad, T. I. Andersen, J. M. Nesland, K. J. Isselbacher, and S. H. Friend. (1991). Constant denaturant gel electrophoresis as a rapid screening technique for p53 mutations. *Proc. Natl. Acad. Sci., USA* **88**: 8405-8409.
- Bowling, A. C., E. M. Mutisya, L. C. Walker, D. L. Price, L. C. Cork, and M. F. Beal. (1993). Age-dependent impairment of mitochondrial function in primate brain. *J. Neurochem.* **60**: 1964-1967.

- Boyd, M. R. (1977). Evidence for the Clara cell as the site of cytochrome P-450 dependent mixed functions oxidase activity in lung. *Nature* **269**: 713-715.
- Boyer, J. C., A. Umar, J. I. Risinger, J. R. Lipford, M. Kane, S. Yin, J. C. Barrett, R. D. Kolodner, and T. A. Kunkel. (1995) Microsatellite instability, mismatch repair deficiency, and genetic defects in human cancer cell lines. *Cancer Res.* **55**: 6063-6070.
- Branda, R. F., J. P. O'Neill, L. M. Sullivan, and R. J. Albertini. (1991) Factors influencing mutation at the *hprt* locus in T-lymphocytes: Women treated for breast cancer. *Cancer Res.* **51**: 6603-6607.
- Bresolin, N., M. Moggio, L. Bet, A. Gallanti, A. Prella, E. Nobile-Orazio, L. Adobbati, C. Ferrante, G. Pellegrino, and G. Scarlato. (1987). Progressive cytochrome c oxidase deficiency in a case of Kearns-Sayre Syndrome: Morphological, immunological, and biochemical studies in muscle biopsies and autopsy tissues. *Ann. Neurol.* **21**: 564-572.
- Brookes, P., and P. D. Lawley. (1964). Evidence for the binding of polynuclear aromatic hydrocarbons to the nucleic acids of mouse skin: Relation between carcinogenic power of hydrocarbons and their binding to deoxyribonucleic acid. *Nature* **202**: 781-784.
- Brown, W. M., M. J. George, and A. C. Wilson. (1979). Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* **76**: 1967-71.
- Buening, M. K., P. G. Wislocki, W. Levin, H. Yagi, D. R. Thakker, H. Akagi, M. Koreeda, D. M. Jerina, and A. H. Conney. (1978). Tumorigenicity of the optical enantiomers of the diastereomeric benzo[a]pyrene 7,8-diol-9,10-epoxides in newborn mice: Exception activity of (+)-7b,8a-dihydroxy-9a,10a-epoxy-7,8,9,10-tetrahydrobenzo[a]pyrene. *Proc. Natl. Acad. Sci. USA* **75**: 5358-5361.
- 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.
- Burkitt, H. G., B. Young, and J. W. Heath. Wheater's Functional Histology. Edinburgh: Churchill Livingstone, 1993.
- Burns, P. A., A. J. E. Gordon, and B. W. Glickman. (1987). Influence of neighboring base sequence on N-methyl-N'-nitro-N-nitrosoguanine mutagenesis in the *lacI* gene of *Escherichia coli*. *J. Mol. Biol.* **194**: 385-390.
- Burton, H. R., G. H. Childs, R. A. Anderson, and P. D. Fleming. (1989). Changes in chemical composition of Burley tobacco during senescence and curing. 3. Tobacco-specific nitrosamines. *J. Agric. Food Chem.* **37**: 426-430.
- Byrne, E., X. Dennett, I. Trounce, and R. Henderson. (1985). Partial cytochrome oxidase (aa3) deficiency in chronic progressive external ophthalmoplegia. Histochemical and biochemical studies. *J. Neurol. Sci.* **71**: 257-271.
- Caggana, M., H. L. Liber, P. M. Mauch, C. N. Coleman, and K. T. Kelsey. (1991). *In vivo* somatic mutations in the lymphocytes of Hodgkin's disease patients. *Environ. Mol. Mutagen.* **18**: 6-13.

Camner, P., K. Philipson, and I. Friberg. (1972). Tracheobronchial clearance in twins. *Arch. Environ. Health* **24**: 82-87.

Caporaso, N. E., M. A. Tucker, R. N. Hoover, R. B. Hayes, L. W. Pickle, H. J. Issaq, G. M. Muschik, L. Green-Gallo, D. Buivys, S. Aisner, J. H. Resau, B. F. Trump, D. Tollerud, A. Weston, and C. C. Harris. (1990). Lung cancer and the debrisoquine metabolic phenotype. *J. Natl. Cancer Inst.* **82**: 1264-1272.

Cariello, N. (1996). Relational database model for DNA mutations and a software program for implementation of the model. *Mutat. Res.* **359**: 103-117.

Cariello, N. F., P. Keohavong, A. G. Kat, and W. G. Thilly. (1990). Molecular analysis of complex human cell populations: mutational spectra of MNNG and ICR-191. *Mutat. Res.* **231**: 165-176.

Carmella, S. G., S. Akerkar, J. P. Richie, Jr., and S. S. Hecht. (1995). Intraindividual and interindividual differences in metabolites of the tobacco-specific lung carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) in smokers' urine. *Cancer Epidemiol. Biomarkers Prev.* **4**: 635.

Cha, R., W. Thilly, and H. Zarbl. (1994). N-Nitroso-N-methylurea-induced rat mammary tumors arise from cells with preexisting oncogenic Hras-1 gene mutations. *Proc. Natl. Acad. Sci. USA* **91**: 3749-3753.

Cha, R. S., H. Zarbl, P. Keohavong, and W. G. Thilly. (1992). Mismatch Amplification Mutation Assay (MAMA): Application to the c-H-ras gene. *PCR Meths. Apps.* **2**: 14-20.

Chakrabarti, S., H. Mizusawa, and M. Seidman. (1984). Survival and mutagenesis of bacterial plasmids with localized carcinogen adducts. *Mutat. Res.* **126**: 127-37.

Chance, B., H. Sies, and A. Boveris. (1979). Hydroperoxide metabolism in mammalian organs. *Physiol. Rev.* **59**: 527-605.

Chen, J., and W. G. Thilly. (1994). Mutational spectrum of chromium (VI) in human cells. *Mutat. Res.* **323**: 21-27.

Chen, J., and W. G. Thilly. (1996). Mutational spectra vary with exposure conditions: Benzo[a]pyrene in human cells. *Mutat. Res.* **357**: 209-217.

Chen, R.-H., V.-M. Maher, and M. J. J. (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, W., J. Chen, J. Liu, W. Lin, K. King, J. Whang-Peng, and W. Yang. (1997). Microsatellite instability in sporadic-colon-cancer patients with and without liver metastases. *Int. J. Cancer* **74**: 470-474.

Chi, N. W., and R. D. Kolodner. (1994). The effect of DNA mismatches on the ATPase activity of MSH1, a protein in yeast mitochondria that recognizes DNA mismatches. *J. Biol. Chem.* **269**: 29993-7.

Chinnery, P. F., D. M. Turnbull, T. J. Walls, and P. J. Reading. (1997). Recurrent strokes in a 34-year-old man. *The Lancet* **350**: 560.

- Clayton, D. A., J. N. Doda, and E. C. Friedberg. (1974). The absence of a pyrimidine dimer repair mechanism in mammalian mitochondria. *Proc. Natl. Acad. Sci. USA* **71**: 2777-2781.
- Cleeter, M. W., J. M. Cooper, and A. H. Schapira. (1992). Irreversible inhibition of mitochondrial complex I by 1-methyl-4-phenylpyridinium: evidence for free radical involvement. *J. Neurochem.* **58**: 786-789.
- Cohen, A. S., D. R. Najarian, A. Paulus, A. Guttman, J. A. Smith, and B. L. Karger. (1988). Rapid separation and purification of oligonucleotides by high-performance capillary gel electrophoresis. *Proc. Natl. Acad. Sci. USA* **85**: 9660-9663.
- 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. *Mutat. Res.* **204**: 493-507.
- Comfort, A. (1963). Effect of delayed and resumed growth on the longevity of a fish (*Lebistes reticulatus*, Peters) in captivity. *Gerontologia* **8**: 150.
- Corral-Debrinski, M., T. Horton, M. T. Lott, J. M. Shoffner, M. F. Beal, and D. C. Wallace. (1992). Mitochondrial DNA deletions in human brain: regional variability and increase with advanced age. *Nat. Genet.* **2**: 324-329.
- Cortopassi, G., and E. Wang. (1995). Modelling the effects of age-related mtDNA mutation accumulation: complex I deficiency, superoxide and cell death. *Biochim. Biophys. Acta* **1271**: 171-176.
- Cortopassi, G. A., and N. Anheim. (1990). Detection of a specific mitochondrial DNA deletion in tissues of older humans. *Nucl. Acids Res.* **18**: 6927-6933.
- Cortopassi, G. A., D. Shibata, N.-W. Soong, and N. Arnheim. (1992). A pattern of accumulation of a somatic deletion of mitochondrial DNA in aging human tissues. *Proc. Natl. Acad. Sci. USA* **89**: 7370-7374.
- Coulondre, C., and J. H. Miller. (1977). Genetic studies of the lac repressor. IV. Mutagenic specificity in the *lacI* gene of *Escherichia coli*. *J. Mol. Biol.* **117**: 577-606.
- Coulondre, C., J. H. Miller, P. J. Farabnaugh, and W. Gilbert. (1978). Molecular basis of base substitution hotspots in *Escherichia coli*. *Nature* **274**: 775-780.
- Crespi, C. L., B. W. Penman, H. V. Gelboin, and F. J. Gonzalez. (1991). A tobacco smoke derived nitrosamine, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone, is activated by multiple human cytochrome P-450s including polymorphic human cytochrome P-450 2D6. *Carcinogenesis* **12**: 1197-1201.
- Curry, J., G. T. Rowley, V. Saddi, D. Beare, J. Cole, and B. 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.
- Cuzick, J., M. N. Routledge, D. Jenkins, and R. C. Garner. (1990). DNA adducts in different tissues of smokers and non-smokers. *Int. J. Cancer* **45**: 673-678.

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

Davis, A., and D. Clayton. (1996). *In situ* localization of mitochondrial DNA replication in intact mammalian cells. *J. Cell Biol.* **135**: 883-893.

de Grey, A. (1997). A proposed refinement of the mitochondrial free radical theory of aging. *Bioessays* **19**: 161-166.

Denissenko, M. F., A. Pao, M.-s. Tang, and G. P. Pfeifer. (1996). Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in p53. *Science* **274**: 430-432.

Desai, V. G., R. Weindruch, R. W. Hart, and R. J. Feuers. (1996). Age-associated alterations in mouse gastrocnemius electron transport system are attenuated by dietary restriction. *Arch. Biochem. Biophys.* **333**: 145-151.

Devereux, T. R., M. W. Anderson, and S. A. Belinsky. (1988). Factors regulating activation and DNA alkylation by 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone and nitrosodimethylamine in rat lung and isolated lung cells and the relationship to carcinogenicity. *Cancer Res.* **48**: 4215-4221.

Doll, R., A. B. Hill, and L. Kreyberg. (1957). The significance of cell type in relation to the aetiology of lung cancer. *Br. J. Cancer* **11**: 43-8?

Doll, R., and R. Peto. (1976). Mortality in relation to smoking: Ten years' observations of British doctors. *Br. med. J.* **ii**: 1525-1536.

Domena, J. D., and D. W. Mosbaugh. (1985). Purification of nuclear and mitochondrial uracil-DNA glycosylase from rat liver. Identification of two distinct subcellular forms. *Biochemistry* **24**: 7320-7334.

Donnelly, G. M., D. G. Haack, and C. S. Heird. (1982). Tracheal epithelium: cell kinetics and differentiation in normal rat tissue. *Cell Tissue Kinet.* **15**: 119-130.

Driggers, W. J., S. P. LeDoux, and G. L. Wilson. (1993). Repair of oxidative damage within the mitochondrial DNA of RINr 38 cells. *J. Biol. Chem.* **268**: 22042-22045.

Dubey, A., M. J. Forster, H. Lal, and R. S. Sohal. (1996). Effect of age and caloric intake on protein oxidation in different brain regions and on behavioral functions of the mouse. *Arch. Biochem. Biophys.* **333**: 189-97.

Dunn, B. P., S. Vedal, R. H. C. San, W.-F. Kwan, B. Nelems, D. A. Enarson, and H. F. Stich. (1991). DNA adducts in bronchial biopsies. *Int. J. Cancer* **48**: 485-492.

Eisenstadt, E., A. J. Warren, J. Porter, D. Atkins, and J. H. Miller. (1982). Carcinogenic epoxides of benzo[a]pyrene and cyclopenta[cd]pyrene induce base substitutions via specific transversions. *Proc. Natl. Acad. Sci. USA* **79**: 1945-1949.

El-Bayoumy, K. (1992). Environmental carcinogens that may be involved in human breast cancer etiology. *Chem. Res. Toxicol.* **5**: 585-590.

- Evans, M. J., S. G. Shami, L. J. Cabral-Anderson, and N. P. Dekker. (1986). Role of nonciliated cells in renewal of the bronchial epithelium of rats exposed to NO₂. *Am. J. Pathol.* **123**: 126-133.
- Felley-Bosco, E., C. Pourzand, J. Zijlstra, P. Amstad, and P. Cerutti. (1991). A genotypic mutation system measuring mutations in restriction recognition sequences. *Nucl. Acids Res.* **19**: 2913-2919.
- Fischer, S. G., and L. S. Lerman. (1983). DNA fragments different by single base-pair substitutions are separated in denaturing gradient gel: correspondence with melting theory. *Proc. Natl. Acad. Sci., USA* **80**: 1579-83.
- Fishel, R., M. K. Lescoe, M. R. S. Rao, N. G. Copeland, N. A. Jenkins, J. Garber, M. Kane, and R. Kolodner. (1993). The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell* **75**: 1027-1038.
- Fraga, C., P. Motchnik, A. Wyrobek, D. Rempel, and B. Ames. (1996). Smoking and low antioxidant levels increase oxidative damage to sperm DNA. *Mutat. Res.* **351**: 199-203.
- Friedberg, E. C., G. C. Walker, and W. Siede. DNA Repair and Mutagenesis. Washington, D.C.: ASM Press, 1995. 698 pp.
- Fuchs, R. P. P. (1984). DNA binding spectrum of the carcinogen N-acetoxy-N-2-acetylaminofluorene significantly differs from the mutation spectrum. *J. Mol. Biol.* **177**: 173-180.
- Furth, E. E., W. G. Thilly, B. W. Penman, H. L. Liber, and W. M. Rand. (1981). Quantitative assay for mutation in diploid human lymphoblasts using microtiter plates. *Analytical Biochem.* **110**: 1-8.
- Gadaleta, M. N., G. Rainaldi, A. M. S. Lezza, F. Milella, F. Fracasso, and P. Cantatore. (1992). Mitochondrial DNA copy number and mitochondrial DNA deletion in adult and senescent rats. *Mutat. Res.* **275**: 181-193.
- Gelboin, H. W. (1969). A microsome-dependent binding of benzo[a]pyrene to DNA. *Cancer Res.* **29**: 1272-1276.
- General, S. Reducing the Health Consequences of Smoking. Rockville, MD: US Department of Health and Human Services, 1989.
- Gerbitz, K.-D., J. M. W. van den Ouweland, J. A. Maassen, and M. Jaksch. (1995). Mitochondrial diabetes mellitus: a review. *Biochim. Biophys. Acta* **1271**: 253-260.
- Goldmacher, V. S., R. A. Cuzick, Jr., and W. G. Thilly. (1986). Isolation and partial characterization of human cell mutants differing in sensitivity to killing and mutation by methylnitrosourea and N-methyl-N'-nitro-N-nitrosoguanidine. *J. Biol. Chem.* **261**: 12462-12471.
- Gotoh, O., and Y. Tagashira. (1981). Stabilities of nearest-neighbor doublets in double-helical DNA determined by fitting calculated melting profiles to observed profiles. *Biopolymers* **20**: 1033-1042.

Green, M. H. L., 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. *Mutat. Res.* **334**: 323-339.

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

Gupta, P. K., and M. A. Sirover. (1981). Stimulation of the nuclear uracil DNA glycosylase in proliferating human fibroblasts. *Cancer Res.* **41**: 3133-3136.

Hadler, H. I., B. Dimitrijevic, and R. Mahalingam. (1983). Mitochondrial DNA and nuclear DNA from normal rat liver have a common sequence. *Proc. Natl. Acad. Sci., USA* **80**: 6495-99.

Hammond, E. C. (1966). Smoking in relation to the death rates of one million men and women. *Natl. Cancer Inst. Monogr.* **19**: 127-204.

Harris, J. E. (1983). Cigarette smoking among successive birth cohorts of men and women in the United States during 1900-1980. *J. Natl. Cancer. Inst.* **71**: 474-478.

Hauswirth, W. W., and P. J. Laipis. "Transmission genetics of mammalian mitochondria: A molecular model and experimental evidence." Achievements and Perspectives of Mitochondrial Research Vol. II: Biogenesis. Ed. E. Quagliariello. Elsevier Science Publishers, 1985. 49-59.

Hayakawa, M., K. Torii, S. Sugiyama, M. Tanaka, and T. Ozawa. (1991). Age-associated accumulation of 8-hydroxydeoxyguanosine in mitochondrial DNA of human diaphragm. *Biochem. Biophys. Res. Commun.* **179**: 1023-1029.

Hecht, S. S., and D. Hoffmann. (1988). Tobacco-specific nitrosamines, an important group of carcinogens in tobacco and tobacco smoke. *Carcinogenesis* **9**: 875-884.

Hecht, S. S., R. Young, and C. B. Chen. (1980). Metabolism in the F344 rat of 4-(N-methyl-N-nitrosamino)-1-(3-pyridyl)-1-butanone, a tobacco specific carcinogen. *Cancer Res.* **40**: 4144.

Hechts, S. S., T. E. Spratt, and N. Trushin. (1988). Evidence for 4-(3-pyridyl)-4-oxobutylolation of DNA in F344 rats treated with the tobacco specific nitrosamines 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone and N'-nitrososornicotine. *Carcinogenesis* **9**: 161.

Hegler, J., D. Bittner, S. Boiteux, and B. Epe. (1993). Quantification of oxidative DNA modifications in mitochondria. *Carcinogenesis* **14**: 2309-2312.

Heinen, C., D. Richardson, R. White, and J. Groden. (1995). Microsatellite instability in colorectal adenocarcinoma cell lines that have full-length adenomatous polyposis coli protein. *Cancer Res.* **55**: 4797-4799.

Hemminki, A., P. Peltomäki, J.-P. Mecklin, H. Järvinen, R. Salovaara, M. Nyström-Lahti, A. de la Chapelle, and L. A. Aaltonen. (1994). Loss of wild type MLH1 gene is a feature of hereditary nonpolyposis colorectal cancer. *Nat. Genet.* **8**: 405-410.

Higuchi, Y., and S. Linn. (1995). Purification of all forms of HeLa cell mitochondrial DNA and assessment of damage to it caused by hydrogen peroxide treatment of mitochondria or cells. *J. Biol. Chem.* **270**: 7950-6.

Hjertén, S. (1985). High-performance electrophoresis: Elimination of electroendosmosis and solute adsorption. *J. Chromatogr.* **347**: 191-198.

Holt, I. J., A. E. Harding, J. M. Cooper, A. H. V. Schapira, A. Toscano, J. B. Clark, and J. A. Morgan-Hughes. (1989a). Mitochondrial myopathies: clinical and biochemical features of 30 patients with major deletions of muscle mitochondrial DNA. *Ann. Neurol.* **26**: 699-708.

Holt, I. J., A. E. Harding, and J. A. Morgan-Hughes. (1988). Deletions of muscle mitochondrial DNA in patients with mitochondrial myopathies. *Nature* **331**: 717-719.

Holt, I. J., D. H. Miller, and A. E. Harding. (1989b). Genetic heterogeneity and mitochondrial DNA heteroplasmy in Leber's hereditary optic neuropathy. *J. Med. Genet.* **26**: 739-743.

Horai, S., and K. Hayasaka. (1990). Intraspecific nucleotide sequence differences in the major noncoding region of human mitochondrial DNA. *Am. J. Hum. Genet.* **46**: 828-842.

Hovig, E., B. Smith-Sørensen, A. Brøgger, and A.-L. Børresen. (1991). Constant denaturant gel electrophoresis, a modification of denaturing gradient gel electrophoresis, in mutation detection. *Mutat. Res.* **262**: 63-71.

Hu, G., and W. Thilly. (1995). Pseudogenes reveal recent acute changes in the human genome. *Curr. Genet.* **28**: 410-414.

Husgafvel-Pursiainen, K., and A. Kannio. (1996). Cigarette smoking and p53 mutations in lung cancer and bladder cancer. *Environ. Health Perspect.* **104 S3**: 553-556.

Hüttner, E., U. Mergner, R. Braun, and J. Schöneich. (1990). Increased frequency of 6-thioguanine-resistant lymphocytes in peripheral blood of workers employed in cyclophosphamide production. *Mutat. Res.* **243**: 101-107.

IARC. Monograph on the Evaluation of the Carcinogenic Risk of Chemicals to Humans. Vol. Vol 38: Tobacco Smoking of Lyon: International Agency for Research on Cancer, 1986.

Ingram, D. K., R. Weindruch, E. L. Spangler, J. R. Freeman, and R. L. Walford. (1987). Dietary restriction benefits learning and motor performance of aged mice. *J. Gerontol.* **42**: 78-81.

Ionov, Y., M. A. Peinado, S. Malkhosyan, D. Shibata, and M. Perucho. (1993). Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature* **363**: 558-561.

Ivanovic, V., N. E. Geacintov, and I. B. Weinstein. (1976). Cellular binding of benzo[a]pyrene to DNA characterized by low temperature fluorescence. *Biochem. Biophys. Res. Commun.* **70**: 1172-1179.

- Jahnke, G. D., G. L. Thompson, C. L. Walker, G. E. Gallagher, G. W. Lucier, and R. P. Di Quagustine. (1990). Multiple DNA adducts in lymphocytes of smokers and nonsmokers determined by ^{32}P -postlabeling analysis. *Carcinogenesis* **11**: 205-211.
- Jalili, T., G. G. K. Murthy, and R. H. Schiestl. Cigarette smoke induces DNA deletions in the mouse embryo. (submitted).
- Janatipour, M., K. J. Trainor, R. Kutlaca, G. Bennett, J. Hay, D. R. Turner, and A. A. Morely. (1988). Mutations in human lymphocytes studied by an HLA selection system. *Mutat. Res.* **198**: 221-226.
- Jeffery, P. K., and L. M. Reid. "Ultrastructure of airway epithelium and submucosal gland during development." Lung Biology in Health and Disease, Vol. 6, Development of the Lung. Ed. W. A. Hodson. Marcel Dekker, 1977. 87-134.
- Jensen, R., R. Langlois, W. Bigbee, S. Grant, D. N. Moore, M. Pilinskaya, I. Vorobtsova, and P. Pleshanov. (1995). Elevated frequency of glycophorin A mutations in erythrocytes from Chernobyl accident victims. *Radiat. Res.* **141**: 129-135.
- Jenuth, J. P., A. C. Peterson, K. Fu, and E. A. Shoubridge. (1996). Random genetic drift in the female germline explains the rapid segregation of mammalian mitochondrial DNA. *Nat. Genet.* **14**: 146-151.
- Jenuth, J. P., A. C. Peterson, and E. A. Shoubridge. (1997). Tissue-specific selection for different mtDNA genotypes in heteroplasmic mice. *Nat. Genet.* **16**: 93-5.
- Johnson, M. A., D. M. Turnbull, D. J. Dick, and H. S. A. Sherratt. (1983). A partial deficiency of cytochrome c oxidase in chronic progressive external ophthalmoplegia. *J. Neurol. Sci.* **60**: 31-53.
- Jonason, A. S., S. Kunala, G. J. Price, R. J. Restifo, H. M. Spinelli, J. A. Persing, D. J. Leffell, R. E. Tarone, and D. E. Brash. (1996). Frequent clones of p53-mutated keratinocytes in normal human skin. *Proc. Natl. Acad. Sci. USA* **93**: 14025-14029.
- Jones, I., 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 Epidemiol. Biomarkers Prev.* **2**: 249-60.
- Jones, P. A., W. Rideout, J.-C. Shen, C. H. Spruck, and Y. C. Tsai. (1992). Methylation, mutation and cancer. *BioEssays* **14**: 33-36.
- Kadenbach, B., C. Münscher, V. Frank, J. Müller-Höcker, and J. Napiwotzki. (1995). Human aging is associated with stochastic somatic mutations of mitochondrial DNA. *Mutat. Res.* **338**: 161-172.
- Kadowaki, T., H. Kadowaki, Y. Mori, K. Tobe, R. Sakuta, Y. Suzuki, Y. Tanabe, H. Sakura, T. Awata, Y. Goto, T. Hayakawa, K. Matsuoka, R. Kawamori, T. Kamada, S. Horai, I. Nonaka, R. Hagura, Y. Akanuma, and Y. Yazaki. (1994). A subtype of diabetes mellitus associated with a mutation of mitochondrial DNA. *N. Eng. J. Med.* **330**: 962-968.
- Kan, Y. W., and A. M. Dozy. (1978). Antenatal diagnosis of sickle-cell anaemia by DNA analysis of amniotic-fluid cells. *Lancet* **ii**: 910-912.

Kapitulnik, J., W. Levin, A. H. Conney, H. Yagi, and D. M. Jerina. (1977). Benzo[a]pyrene 7,8-dihydrodiol is more carcinogenic than benzo[a]pyrene in newborn mice. *Nature* **266**: 378-380.

Kapitulnik, J., P. G. Wislocki, W. Levin, H. Yagi, D. M. Jerina, and A. H. Conney. (1978). Tumorigenicity studies with diol-epoxides of benzo[a]pyrene which indicate that (±)-trans-7b,8a,dihydroxy-9a,10a-epoxy-7,8,9,10-tetrahydrobenzo[a]pyrene is an ultimate carcinogen in newborn mice. *Cancer Res.* **38**: 354-358.

Kärki, N., R. Pokela, L. Nuutinen, and O. Pelkonen. (1987). Aryl hydrocarbon hydroxylase in lymphocytes and lung tissues from lung cancer patients and controls. *Int. J. Cancer* **39**: 565-570.

Kat, A. "MNNG mutational spectra in the human hprt gene." Ph.D. thesis. MIT, 1992.

Kat, A., W. G. Thilly, W. Fang, M. Longley, G. Li, and P. Modrich. (1993). An alkylation-tolerant, mutator human cell line is deficient in strand-specific mismatch repair. *Proc. Natl. Acad. Sci., USA* **90**: 6424-6428.

Katayama, M., M. Tanaka, H. Yamamota, T. Ohbayshi, Y. Nimur, and T. Ozawa. (1991). Deleted mitochondrial DNA in the skeletal muscle of aged individuals. *Biochem. Int.* **1991**: 894-899.

Katsumata, K., M. Hayakawa, M. Tanaka, S. Sugiyama, and T. Ozawa. (1994). Fragmentation of human heart mitochondrial DNA associated with premature aging. *Biochem. Biophys. Res. Commun.* **202**: 102-110.

Kauffman, S. L. (1980). Cell proliferation in the mammalian lung. *Int. Rev. Exp. Pathol.* **22**: 131-191.

Keenan, K., J. W. Combs, and E. McDowell. (1982a). Regeneration of hamster tracheal epithelium after mechanical injury: I. Focal lesions: Quantitative morphologic study of cell proliferation. *Virchows Arch. [Cell. Pathol.]* **41**: 193-214.

Keenan, K. P., J. W. Combs, and E. M. McDowell. (1982b). Regeneration of hamster tracheal epithelium after mechanical injury. II. Multifocal lesions: stathmokinetic and autoradiographic studies of cell proliferation. *Virchows Arch. Pathol.* **41**: 215-229.

Kellerman, G., G. R. Shaw, and M. Luyten-Kellermann. (1973). Aryl hydrocarbon hydroxylase inducibility and bronchiogenic carcinoma. *New Engl. J. Med.* **289**: 934-937.

Keohavong, P., and W. G. Thilly. (1989). Fidelity of DNA polymerases in DNA amplification. *Proc. Natl. Acad. Sci. USA* **86**: 9253-9257.

Keohavong, P., and W. G. Thilly. (1992). Mutational spectrometry: A general approach for hot-spot point mutations in selectable genes. *Proc. Natl. Acad. Sci. USA* **89**: 4623-4627.

Khaidakov, M., and Glickman, B. W. (1996) Possible factors leading to a misjudgement of mutant frequencies in HPRT assay. *Mutat. Res.* **354**: 9-14.

Khrapko, K., P. André, R. Cha, B. Hu, and W. G. Thilly. (1994a). Mutational spectrometry: Means and ends. *Prog. Nucl. Acid Res. Mol. Biol.* **49**: 285-311.

Khrapko, K., H. Collier, P. André, X.-C. Li, F. Foret, A. Belenky, B. L. Karger, and W. G. Thilly. (1997). Mutational spectrometry without phenotypic selection: Human mitochondrial DNA. *Nucl. Acids Res.* **25**: 685-693.

Khrapko, K., H. Collier, and W. G. Thilly. (1996). Efficiency of separation of DNA mutations by constant denaturant capillary electrophoresis (CDCE) is controlled by the kinetics of DNA melting equilibrium. *Electrophoresis* **17**: 1867-1874.

Khrapko, K., J. S. Hanekamp, W. G. Thilly, A. Belenkii, F. Foret, and B. L. Karger. (1994b). Constant denaturant capillary electrophoresis (CDCE): a high resolution approach to mutational analysis. *Nucl. Acids Res.* **22**: 362-369.

Kier, L. D., E. Yamasaki, and B. N. Ames. (1974). Detection of mutagenic activity in cigarette smoke condensates. *Proc. Natl. Acad. Sci., USA* **71**: 4159-4163.

Kinkaid, T. "Mutational spectrum of MNU in TK-6 cells." S.M. Thesis. MIT, 1993.

Kouri, R. E., C. E. McKinney, D. J. Slomiany, D. R. Snodgrass, N. P. Wray, and T. L. McLemore. (1982). Positive correlation between high aryl hydrocarbon hydroxylase activity and primary lung cancer as analyzed in cryopreserved lymphocytes. *Cancer Res.* **42**: 5030-5037.

Kreyberg, L. (1955). Lung cancer and tobacco smoking in Norway. *Br. J. Cancer* **9**: 495-500.

Kriek, E., F. J. Van Schooten, M. J. X. Hillebrand, F. E. Van Leeuwen, L. DenEngelse, A. J. A. De Looff, and A. P. G. Dijkmans. (1993). DNA adducts as a measure of lung cancer risk in humans exposed to polycyclic aromatic hydrocarbons. *Environ. Health Perspect.* **99**: 71-75.

Ku, H. H., U. T. Brunk, and R. S. Sohal. (1993). Relationship between mitochondrial superoxide and hydrogen peroxide production and longevity of mammalian species. *Free Radic. Biol. Med.* **15**: 621-627.

Kubitschek, H. E. Introduction to Research with Continuous Cultures. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1970.

Kuchino, Y., F. Mori, H. Kasai, H. Inoue, S. Iwai, K. Miura, E. Ohtsuka, and S. Nishimura. (1987). Misreading of 8-OH-dG at modified base and at adjacent bases: Misreading of DNA templates containing 8-hydroxydeoxyguanosine at the modified base and at adjacent residues. *Nature* **327**: 77-9.

Kuhn III, C., L. A. Callaway, and F. B. Askin. (1974). The formation of granules in the bronchiolar Clara cells of the rat. *Electron Microscopy J. Ultrastruct. Res.* **49**: 387-391.

Kumar, R., S. Sukumar, and M. Barbacid. (1990). Activation of ras oncogenes preceding the onset of neoplasia. *Science* **248**: 1101-1104.

Kunkel, T. A. (1985). The mutational specificity of DNA polymerases- α and - γ during *in vitro* DNA synthesis. *J. Biol. Chem.* **260**: 12866-12874.

Kunkel, T. A. (1990). Misalignment-mediated DNA synthesis errors. *Biochemistry* **29**: 8003-8011.

- Kunkel, T. A., and D. W. Mosbaugh. (1989). Exonucleolytic proofreading by a mammalian DNA polymerase γ . *Biochemistry* **28**: 988-995.
- Kvale, P. A., F. R. Bode, and S. Kini. (1976). Diagnostic accuracy in lung cancer: Comparison of techniques used in association with flexible fiberoptic bronchoscopy. *Chest* **69**: 752-757.
- Laipis, P. J., M. J. Van de Walle, and W. W. Hauswirth. (1988). Unequal partitioning of bovine mitochondrial genotypes among siblings. *Proc. Natl. Acad. Sci. USA* **85**: 8107-8110.
- Landegren, U., R. Kaiser, J. Sanders, and L. Hood. (1988). A ligase-mediated gene detection technique. *Science* **241**: 1077-1080.
- Langlois, R. G., M. Akiyama, Y. Kusunoki, B. R. Dupont, D. H. Moore, II, W. L. Bigbee, S. G. Grant, and R. H. Jensen. (1993). Analysis of somatic cell mutations at the glycophorin A locus in atomic bomb survivors: a comparative study of assay methods. *Radiat. Res.* **136**: 111-117.
- Lauber, J., C. Marsac, B. Kadenbach, and P. Seibel. (1991). Mutations in mitochondrial tRNA genes: a frequent cause of neuromuscular disease. *Nucl. Acids Res.* **19**: 1393-1397.
- Leach, F. S., N. C. Nicolaides, N. Papadopoulos, B. Liu, J. Jen, R. Parsons, P. Peltomäki, P. Sistonen, L. A. Aaltonen, M. Nyström-Lahti, X.-Y. Guan, J. Zhang, P. S. Meltzer, J.-W. Yu, F.-T. Kao, D. J. Chen, K. M. Cerosaletti, R. E. K. Fournier, S. Todd, T. Lewis, R. J. Leach, S. L. Naylor, J. Weissenbach, J.-P. Mecklin, H. Järvinen, G. M. Petersen, S. R. Hamilton, J. Green, J. Jass, P. Watson, H. T. Lynch, J. M. Trent, A. de la Chapelle, K. W. Kinzler, B. Vogelstein. (1993) Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. *Cell* **75**:1215-1225.
- LeDoux, S. P., G. L. Wilson, E. J. Beecham, T. Stevnsner, K. Wassermann, and V. A. Bohr. (1992). Repair of mitochondrial DNA after various types of DNA damage in Chinese hamster ovary cells. *Carcinogenesis* **13**: 1967-1973.
- Lee, C. M., S. S. Chung, J. M. Kaczowski, R. Weindruch, and J. M. Aiken. (1993). Multiple mitochondrial DNA deletions associated with age in skeletal muscle of rhesus monkey. *J. Gerontol.* **48**: B201-B205.
- Leong, P. M., H. C. Hsia, and J. H. Miller. (1986). Analysis of spontaneous base substitutions generated in mismatch-repair-deficient strains of *Escherichia coli*. *J. Bacteriol.* **168**: 412-416.
- Lerman, L. S., and K. Silverstein. (1987). Computational simulation of DNA melting and its application to denaturing gradient gel electrophoresis. *Meths. Enzymol.* **155**: 482-501.
- Lestienne, P., and G. Ponsot. (1988). Kearns-Sayre syndrome with muscle mitochondrial deletion. *Lancet* **i**: 885.
- Leuchtenberger, C., and R. Leuchtenberger. (1970). Differential cytological and cytochemical responses of various cultures from mouse tissues to repeated exposures to puffs from the gas phase of charcoal-filtered fresh cigarette smoke. *Exp. Cell Res.* **62**: 161-72.

- Levinson, G., and G. A. Gutman. (1987). High frequencies of short frameshifts in poly-CA/TG tandem repeats borne by bacteriophage M13 in *Escherichia coli* K-12. *Nucl. Acids Res.* **15**: 5323-5338.
- Li, X.-C. (1997). Sequence-specific enrichment of a nuclear single-copy gene fragment from human genomic DNA. (submitted).
- Li, X.-C., and W. G. Thilly. (1996). Use of wide-bore capillaries in constant denaturant capillary electrophoresis. *Electrophoresis* **17**: 1884-1889.
- Linnane, A. W., A. Baumer, R. J. Maxwell, H. Preston, C. F. Zhang, and S. Marzuki. (1990). Mitochondrial gene mutation: the ageing process and degenerative disease. *Biochem. Int.* **22**: 1067-1076.
- Liou, S., D. Jacobson-Kram, M. C. Poirier, D. Nhuyen, P. T. Strickland, and M. S. Tockman. (1989). Biological monitoring of fire fighters: sister chromatid exchange and polycyclic aromatic hydrocarbon-DNA adducts in peripheral blood cells. *Cancer Res.* **49**: 4929-4935.
- Liu, C.-S., S.-H. Kao, and Y.-H. Wei. (1997). Smoking-associated mitochondrial DNA mutations in human hair follicles. *Environ. Mol. Mutagen.* **30**: 47-55.
- Lombes, A., E. Bonilla, and S. DiMauro. (1989). Mitochondrial encephalomyopathies. *Rev. Neurol. (Paris)* **145**: 671-689.
- Lonsdale, D. M., T. P. Hodge, C. J. Howe, and D. B. Stern. (1983). Maize mitochondrial DNA contains a sequence homologous to the ribulose-1,5-bisphosphate carboxylase large subunit gene of chloroplast DNA. *Cell* **34**: 1007-1014.
- Lundberg, K. S., D. D. Shoemaker, M. W. Adams, J. M. Short, J. A. Sorge, and E. J. Mathur. (1991). High-fidelity amplification using a thermostable DNA polymerase isolated from *Pyrococcus furiosus*. *Gene* **108**: 1-6.
- Lykken, D. T., T. J. J. Bouchard, M. McGue, and A. Tellegen. (1990). The Minnesota Twin Family Registry: some initial findings. *Acta Genet. Med. Gemellol. (Roma)* **39**: 35-70.
- Marcelino, L., P. C. André, K. Khrapko, H. A. Collier, J. Griffith, and W. G. Thilly. Chemically-induced mutation in mitochondrial DNA of human cells. (in prep).
- Martin, A. P., and S. R. Palumbi. (1993). Body size, metabolic rate, generation time, and the molecular clock. *Proc. Natl. Acad. Sci. USA* **90**: 4087-4091.
- Martin, G. M., C. E. Ogburn, L. M. Colgin, A. M. Gown, S. D. Edland, and R. J. Monnat, Jr. (1996). Somatic mutations are frequent and increase with age in human kidney epithelial cells. *Hum. Mol. Gen.* **5**: 215-221.
- Mazur, M., and B. W. Glickman. (1988). Sequence specificity of mutations induced by benzo[a]pyrene-7,8-diol-9,10 epoxide at endogenous *aprt* gene in CHO cells. *Somat. Cell Mol. Gen.* **14**: 393-400.
- McCarron, M. A., A. Kutlaca, and A. A. Morley. (1989). The HLA-A mutation assay: improved technique and normal results. *Mutat. Res.* **225**: 189-193.

McDowell, E. M., L. A. Barrett, F. Glavin, C. C. Harris, and B. F. Trump. (1978). The respiratory epithelium. I. Human bronchus. *J. Natl. Cancer Inst.* **61**: 539-549.

McDowell, E. M., K. P. Keenan, and M. Huang. (1982). Restoration of mucociliary tracheal epithelium following deprivation of vitamin A. A quantitative morphologic study. *Virchows Arch. Pathol.* **45**: 221-240.

McFarling, U. L. "What if we really made smoking history?" *The Boston Globe* April 11, 1994: 25.

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. *Mutat. Res.* **240**: 117-126.

Mecocci, P., U. MacGarvey, A. E. Kaufman, D. Koontz, J. M. Shoffner, D. C. Wallace, and M. F. Beal. (1993). Oxidative damage to mitochondrial DNA shows marked age-dependent increases in human brain. *Ann. Neurol.* **34**: 609-16.

Meehan, T., K. Straub, and M. Calvin. (1976). Elucidation of hydrocarbon structure in an enzyme-catalyzed benzo[a]pyrene-poly(G) covalent complex. *Proc. Natl. Acad. Sci. USA* **73**: 1437-1441.

Merante, F., I. Tein, L. Benson, and B. H. Robinson. (1994). Maternally inherited hypertrophic cardiomyopathy due to a novel T-to-C transition at nucleotide 9997 in the mitochondrial tRNA(glycine) gene. *Am. J. Hum. Genet.* **55**: 437-446.

Meyers, K. A., R. Saffhill, and P. J. O'Conner. (1988). Repair of alkylated purines in the hepatic DNA in mitochondria and nuclei in the rat. *Carcinogenesis* **9**: 523-539.

Mita, S., R. Rizzuto, C. T. Moraes, S. Shanske, E. Arnaudo, G. Fabrizi, Y. Koga, S. DiMauro, and E. A. Schon. (1990). Recombination via flanking direct repeats is a major cause of large-scale deletions of human mitochondrial DNA. *Nucl. Acids Res.* **18**: 561-567.

Mita, S., B. Schmidt, E. A. Schon, S. DiMauro, and E. Bonilla. (1989). Detection of "deleted" mitochondrial genomes in cytochrome-c-oxidase-deficient muscle fibers of a patient with Kearns-Sayre syndrome. *Proc. Natl. Acad. Sci. USA* **86**: 9505-9513.

Mooney, L. A., R. M. Santella, L. Covey, A. M. Jeffrey, W. Bigbee, M. C. Randall, T. B. Cooper, R. Ottman, W. Y. Tsai, L. Wazneh, A. H. Glassman, T.-L. Young, and F. P. Perera. (1995). Decline of DNA damage and other biomarkers in peripheral blood following smoking cessation. *Cancer Epidemiol. Biomarkers Prev.* **4**: 627-634.

Moraes, C. T., S. DiMauro, M. Zeviani, A. Lombes, S. Shanske, A. F. Miranda, H. Nakase, E. Bonilla, L. C. Werneck, S. Servidei, I. Nonaka, Y. Koga, A. J. Spiro, A. K. W. Brownell, B. Schmidt, D. L. Schotland, M. Zupanc, D. C. DeVivo, E. A. Schon, and L. P. Rowland. (1989). Mitochondrial DNA deletions in progressive external ophthalmoplegia and Kearns-Sayre syndrome. *N. Engl. J. Med.* **320**: 1293-1299.

Müller-Höcker, J. (1989). Cytochrome-c-oxidase deficient cardiomyocytes in the human heart—an age-related phenomenon. *Am. J. Pathol.* **134**: 1167-1171.

Müller-Höcker, J. (1990). Cytochrome-c-oxidase deficient fibres in the limb muscle and diaphragm of men without muscular disease: an age related alteration. *J. Neurol. Sci.* **100**: 14-21.

Müller-Höcker, J. (1992). Random cytochrome-c-oxidase deficiency of oxyphil cell nodules in the parathyroid gland. A mitochondrial cytopathy related to cell ageing? *Pathol. Res. Pract.* **188**: 701-6.

Müller-Höcker, J., K. Schneiderbanger, F. H. Stefani, and B. Kadenbach. (1992). Progressive loss of cytochrome-c-oxidase in the extraocular muscles during ageing. A cytochemical-immunocytochemical study. *Mutat. Res.* **275**: 115-124.

Müller-Höcker, J., P. Seibel, K. Schneiderbanger, and B. Kadenbach. (1993). Different *in situ* hybridization patterns of mitochondrial DNA in cytochrome c oxidase-deficient extraocular muscle fibres in the elderly. *Virchows Arch. [Pathol. Anat. Histopathol.]* **422**: 7-15.

Münscher, C., J. Müller-Höcker, and B. Kadenbach. (1993). Human aging is associated with various point mutations in tRNA genes of mitochondrial DNA. *Biol. Chem. Hoppe Seyler* **374**: 1099-1104.

Murphy, S. E., A. Palomino, S. S. Hecht, and D. Hoffmann. (1990). Dose-response study of DNA and hemoglobin adduct formation by 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone in F344 rats. *Cancer Res.* **50**: 5446-5452.

Mustonen, R., and K. Hemminki. (1992). 7-Methylguanine levels in DNA of smokers' and non-smokers' total white blood cells, granulocytes and lymphocytes. *Carcinogenesis* **13**: 1951-1955.

Mustonen, R., B. Schoket, and K. Hemminki. (1993). Smoking-related DNA adducts: ³²P-postlabeling analysis of 7-methylguanine in human bronchial and lymphocyte DNA. *Carcinogenesis* **14**: 151.

Myers, K. A., R. Saffhill, and P. J. O'Connor. (1988). Repair of alkylated purines in the hepatic DNA of mitochondria and nuclei in the rat. *Carcinogenesis* **9**: 285-292.

Nettesheim, P., A. M. Jetten, Y. Inayama, A. R. Brody, M. A. George, L. B. Gilmore, T. Gray, and G. E. R. Hook. (1990). Pathways of differentiation of airway epithelial cells. *Environ. Health Perspect.* **85**: 317-320.

Newman, N. J., M. A. Lott, and D. C. Wallace. (1991). The clinical characteristics of pedigrees of Leber's hereditary optic neuropathy with the 11778 mutation. *Am. J. Ophthalmol.* **111**: 750-762.

Nicklas, J. A., J. P. O'Neill, and R. J. Albertini. (1986). Use of T-cell receptor gene probes to quantify the *in vivo* *hprt* mutations in human T-lymphocytes. *Mutat. Res.* **173**: 65-72.

Nicklas, J. A., J. P. O'Neill, T. C. Hunter, M. T. Falta, M. J. Lippert, D. Jacobson-Kram, J. R. Williams, and R. J. Albertini. (1991). *In vivo* ionizing irradiations produce deletions in the *hprt* gene of human T-lymphocytes. *Mutat. Res.* **250**: 383-396.

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* 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.

Nicklas, W. J., I. Vyas, and R. E. Heikkila. (1985). Inhibition of NADH-linked oxidation in brain mitochondria by 1-methyl-4-phenyl-pyridine, a metabolite of the neurotoxin, 1-methyl-4-phenyl-1,2,5,6-tetrahydropyridine. *Life Sci.* **36**: 2503-2508.

Nicolaides, N C, N. Papadopoulos, B. Liu, Y.-F. Wei, K. C. Carter, S. M. Ruben, C. A. Rosen, W. A. Haseltine, R. D. Fleischmann, C. M. Fraser, M. D. Adams, J. C. Venter, M. G. Dunlop, S. R. Hamilton, G. M. Petersen, A. de la Chapelle, B. Vogelstein, and K. W. Kinzler. (1994) Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. *Nature* **371**: 75-80.

Nilsen, H., M. Otterlei, T. Haug, K. Solum, T. A. Nagelhus, F. Skorpen, and H. E. Krokan. (1997). Nuclear and mitochondrial uracil-DNA glycosylases are generated by alternative splicing and transcription from different positions in the UNG gene. *Nucl. Acids Res.* **15**: 750-755.

Niranjan, B. G., N. K. Bhat, and N. G. Avadhani. (1981). Preferential attack of mitochondrial DNA by aflatoxin B1 during hepatocarcinogenesis. *Science* **215**: 73-75.

Obermaier-Kusser, B., J. Müller-Höcker, I. Nelson, P. Lestienne, C. Enter, T. Riedele, and K.-D. Gerbitz. (1990). Different copy numbers of apparently identically deleted mitochondrial DNA in tissues from a patient with Kearns-Sayre syndrome detected by PCR. *Biochem. Biophys. Res. Comm.* **169**: 1007-1015.

Oertel, W. H., and A. Kupsch. (1993). Pathogenesis and animal studies of Parkinson's disease. *Curr. Op. Neurol. Neurosurg.* **6**: 323-332.

Ohno, K., M. Yamamoto, A. G. Engel, C. M. Harper, L. R. Roberts, G. H. Tan, and V. Fatourech. (1996). MELAS- and Kearns-Sayre-type commutation with myopathy and autoimmune polyendocrinopathy. *Ann. Neurol* **39**: 761-766.

Oller, A. R., P. Rastogi, S. Morgenthaler, and W. G. Thilly. (1989). A statistical model to estimate variance in long term-low dose mutation assays: Testing of the model in a human lymphoblastoid mutation assay. *Mutat. Res* **216**: 83-97.

Oller, A. R., and W. G. Thilly. (1992). Mutational spectra in human B-cells. Spontaneous, oxygen and hydrogen peroxide-induced mutations at the *hprt* gene. *J. Mol. Biol.* **228**: 813-826.

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

Ozawa, T. (1995). Mechanism of somatic mitochondrial DNA mutations associated with age and diseases. *Biochim. Biophys. Acta* **1271**: 177-189.

Pallotti, F., X. Chen, E. Bonilla, and E. A. Schon. (1996). Evidence that specific mtDNA point mutations may not accumulate in skeletal muscle during normal human aging. *Am. J. Hum. Genet.* **59**: 591-602.

Papadopoulos, N., N. C. Nicolaides, Y.-F. Wei, S. M. Ruben, K. C. Carter, C. A. Rosen, W. A. Haseltine, R. D. Fleischmann, C. M. Fraser, M. D. Adams, J. C. Venter, S. R. Hamilton, G. M. Petersen, P. Watson, H. T. Lynch, P. Peltomäki, J.-P. Mecklin, A. de la Chapelle, K. W. Kinzler, and B. Vogelstein. (1994). Mutation of a mutL homolog in hereditary colon cancer. *Science* **263**: 1625-1629.

Perera, F. P., J. Mayer, A. Jaretzki, S. Hearne, D. Brenner, T. L. Young, H. K. Fischman, M. Grimes, S. Grantham, M. X. Tang, W. Y. Tsai, and R. M. Santella. (1989). Comparison of DNA adducts and sister chromatid exchange in lung cancer cases and control. *Cancer Res.* **49**: 4446-4451.

Peterson, L. A., and S. S. Hecht. (1991). O6-methylguanine is a critical determinant of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone tumorigenesis in A/J mouse lung. *Cancer Res.* **51**: 5557.

Pette, D., and G. Vrbova. (1985). Neural control of phenotypic expression in mammalian muscle fibers. *Muscle Nerve* **8**: 676-689.

Pettepher, C. C., S. P. LeDoux, V. A. Bohr, and G. L. Wilson. (1991). Repair of alkali-labile sites within the mitochondrial DNA of RINr 38 cells after exposure to the nitrosourea streptozotocin. *J. Biol. Chem.* **266**: 3113-3117.

Phillips, D. H., B. Schoket, A. L. Hewer, E. Bailey, S. Kostic, and I. Vicze. (1990). Influence of cigarette smoking on the levels of DNA adducts in human bronchial epithelium and white blood cells. *Int. J. Cancer* **46**: 569-575.

Pinz, K. G., S. Shibutani, and D. F. Bogenhagen. (1995). Action of mitochondrial DNA polymerase γ at sites of base loss or oxidative damage. *J. Biol. Chem.* **270**: 9202-9206.

Plopper, C. G., and D. L. Dungworth. "Structure, function, cell injury and cell renewal of bronchiolar and alveolar epithelium." *Lung Carcinomas*. Ed. E. M. McDowell. London: Churchill Livingstone, 1987. 94-128.

Plopper, C. G., S. J. Nishio, J. L. Alley, P. Kass, and D. M. Hyde. (1991). The role of the nonciliated bronchiolar epithelial (Clara) cell as the progenitor cell during bronchiolar epithelial differentiation in the perinatal rabbit lung. *Am. J. Respir. Cell. Mol. Biol.* **7**: 606-613.

Poland, D. (1974). Recursion relation generation of probability profiles for specific-sequence macromolecules with long-range correlations. *Biopolymers* **13**: 1859-1871.

Poland, D., and H. A. Scheraga. (1966). Occurrence of a phase transition in nucleic acid models. *J. Chem. Phys.* **45**: 1464-1469.

Polyak, K., Y. Xia, J. L. Zweier, K. W. Kinzler, B. Vogelstein. (1997) A model for p53-induced apoptosis. *Nature* **389**: 300-305.

Poulton, J., M. E. Deadman, J. Bronte-Stewart, W. S. Foulds, and R. M. Gardiner. (1991). Analysis of mitochondrial DNA in Leber's hereditary optic neuropathy. *J. Med. Genet.* **28**: 765-770.

Pryor, W. "Free Radicals in Biology." *Superoxide Radical and Hydrogen Peroxide in Mitochondrial*. New York: Academic Press, 1982. V: 65-90.

Ramsey, J. J., E. B. Roecker, R. Weindruch, and J. W. Kemnitz. (1997). Energy expenditure of adult male rhesus monkeys during the first 30 mo. of dietary restriction. *Am. J. Physiol.* **272**: E901-E907.

Randerath, E., R. H. Miller, D. Mittal, T. A. Avitts, H. A. Dunsford, and K. Randerath. (1989). Covalent DNA damage in tissues of cigarette smokers as determined by ^{32}P -postlabeling assay. *J. Nat. Cancer Inst.* **81**: 341-347.

Reardon, W., R. J. M. Ross, M. G. Sweeney, L. M. Luxon, M. E. Pembrey, A. E. Harding, and R. C. Trembath. (1992). Diabetes mellitus associated with a pathogenic point mutation in mitochondrial DNA. *Lancet* **340**: 1376-1379.

Reenan, R. A., and R. D. Kolodner. (1992a). Characterization of insertion mutations in the *Saccharomyces cerevisiae* MSH1 and MSH2 genes: Evidence for separate mitochondrial and nuclear functions. *Genetics* **132**: 975-985.

Reenan, R. A. G., and R. D. Kolodner. (1992b). Isolation and characterization of two *Saccharomyces cerevisiae* genes encoding homologs of the bacterial HexA and MutS mismatch repair proteins. *Genetics* **132**: 963-973.

Reid, L., and R. Jones. (1979). Bronchial mucosal cells. *Fed. Proc.* **38**: 191.

Richter, C. (1983). Do mitochondrial DNA fragments promote cancer and aging? *FEBS Lett.* **241**: 1-5.

Richter, C., J. Park, and B. Ames. (1988). Normal oxidative damage to mitochondrial and nuclear-DNA is extensive. *Proc. Natl. Acad. Sci. USA* **85**: 6465-6467.

Ridanpää, M., K. Vurvall, L. H. Zhang, K. Husgafvel-Pursiainen, and A. Önfelt. (1995). Comparison of DGGE and CDGE in detection of single base changes in the hamster *hprt* and human *N-ras* genes. *Mutat. Res.* **334**: 357-364.

Robin, E. D., and R. Wong. (1988). Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells. *J. Cell. Physiol.* **136**: 507-513.

Robinson, D. R., K. Goodall, R. J. Albertini, J. P. O'Neill, B. Finette, M. Sala-Trepat, E. Moustacchi, A. D. Bates, 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 population: a comparison of four datasets. *Mutat. Res.* **313**: 227-247.

Rodenhuis, S., and R. J. C. Slebos. (1992). Clinical significance of *ras* oncogene activation in human lung cancer. *Cancer Res.[Suppl.]* **52**: 2665s.

Roilides, E., J. E. Gielen, N. Tuteja, A. S. Levine, and K. Dixon. (1988). Mutational specificity of benzo[a]pyrene diol epoxide in monkey cells. *Mutat. Res.* **198**: 199-206.

Ronai, Z., S. Gradia, L. a. Peterson, and S. S. Hecht. (1993). G to A transitions and G to T transversion in codon 12 of the *Ki-ras* oncogene isolated from mouse lung tumors induced by 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) and related DNA methylating and pyridyloxobutylating agents. *Carcinogenesis* **14**: 2419.

Ropp, P. A., and W. C. Copeland. (1996). Cloning and characterization of the human mitochondrial DNA polymerase, DNA polymerase- γ . *Genomics* **36**: 449-458.

Roy, A., and R. P. Fuchs. (1994). Mutational spectrum induced in *Saccharomyces cerevisiae* by the carcinogen N-2-acetylaminofluorene. *Mol. Gen. Genet.* **245**: 69-77.

Rudzinska, M. A. (1952). Overfeeding and lifespan in *Tokophya infusionum*. *J. Gerontol.* **7**: 544-552.

Ruiz-Martinez, M. C., J. Berka, A. Belenkii, F. Foret, A. W. Miller, and B. L. Karger. (1993). DNA sequencing by capillary electrophoresis with replaceable linear polyacrylamide and laser-induced fluorescence detection. *Anal. Chem.* **65**: 2851-8.

Salmons, S., and J. Henriksson. (1981). The adaptive response of skeletal muscle to increased use. *Muscle Nerve* **4**: 94-105.

Samson, L., and J. Carins. (1977). A new pathway for DNA repair in *Escherichia coli*. *Nature* **267**: 281-283.

Santella, R. M., R. A. Grinberg-Funes, T. L. Young, C. Dickey, V. N. Singh, L. W. Wang, and F. P. Perera. (1992). Cigarette smoking related polycyclic aromatic hydrocarbon-DNA adducts in peripheral mononuclear cells. *Carcinogenesis* **13**: 2041-2045.

Satoh, M. S., N.-H. Huh, M. F. Rajewsky, and T. Kuroki. (1988). Enzymatic removal of O6-ethylguanine from mitochondrial DNA in rat tissues exposed to N-ethyl-N-nitrosourea *in vivo*. *J. Biol. Chem.* **263**: 6854-6856.

Savelle, K., and K. Hemminki. (1991). DNA adducts in lymphocytes and granulocytes of smokers and nonsmokers detected by the ³²P-postlabeling assay. *Carcinogenesis* **13**: 2041-2045.

Savontaus, M.-L. (1995). mtDNA mutations in Leber's hereditary optic neuropathy. *Biochim. Biophys. Acta* **1271**: 261-263.

Schaaper, R. M., and R. L. Dunn. (1987). Mechanisms of mutagenesis in the *Escherichia coli* mutator mutD5: role of DNA mismatch repair. *Proc. Natl. Acad. Sci. USA* **84**: 6220-6224.

Schapira, A. H. (1992). MPTP and other Parkinson-inducing agents. *Curr. Opin. Neurol. Neurosurg.* **5**: 396-400.

Schon, E. A., R. Rizzuto, C. T. Moraes, H. Nakase, M. Zeviani, and S. DiMauro. (1989). A direct repeat is a hotspot for large-scale deletion of human mitochondrial DNA. *Science* **244**: 346-349.

Schwarze, S. R., C. M. Lee, S. S. Chung, E. B. Roecker, R. Weindruch, and J. M. Aiken. (1995). High levels of mitochondrial DNA deletions in skeletal muscle of old rhesus monkeys. *Mech. Ageing Dev.* **83**: 91-101.

Sellers, T. A., J. E. Bailey-Wilson, R. C. Elston, A. F. Wilson, G. Z. Elston, W. L. Ooi, and H. Rothschild. (1990). Evidence for Mendelian inheritance in the pathogenesis of lung cancer. *J. Natl. Cancer Inst.* **82**: 1272-1279.

Shami, S. G., and M. J. Evans. "Kinetics of pulmonary cells." Treatise on Pulmonary Toxicology. Vol I: Comparative Biology of the Normal Lung. Ed. R. A. Parent. Boca Raton, FL: CRC Press, 1992. 145-155.

Shay, J. W., and H. Werbin. (1987). Are mitochondrial DNA mutations involved in the carcinogenic process? *Mutat. Res.* **186**: 149-160.

Sheehan, J. P., P. E. Palmer, G. A. Helm, and J. B. Tuttle. (1997). MPP+ induced apoptotic cell death in SH-SY5Y neuroblastoma cells: An electron microscope study. *J. Neurosci. Res.* **48**: 226-237.

Shibutani, S., M. Takeshita, and A. P. Grollman. (1991). Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxoG. *Nature* **349**: 431-434.

Shields, P. G., E. D. Bowman, A. M. Harrington, V. T. Doan, and A. Weston. (1993). Polycyclic aromatic hydrocarbon-DNA adducts in human lung and cancer susceptibility genes. *Cancer Res.* **53**: 3486-3492.

Shoffner, J. M., M. R. Lott, A. S. Voljavec, S. A. Soueidan, D. A. Costigan, and D. C. Wallace. (1989). Spontaneous Kearns-Sayre/chronic external ophthalmoplegia plus syndrome associated with a mitochondrial DNA deletion: A slip-replication model and metabolic therapy. *Proc. Natl. Acad. Sci. USA* **86**: 7952-7956.

Shoffner, J. M., M. T. Lott, A. M. S. Lezza, P. Seibel, S. W. Ballinger, and D. C. Wallace. (1990). Myoclonic epilepsy and ragged-red fiber disease (MERRF) is associated with a mitochondrial DNA tRNA^{lys} mutation. *Cell* **61**: 931-937.

Shoubridge, E. A., G. Karpati, and K. E. M. Hastings. (1990). Deletion mutants are functionally dominant over wild-type mitochondrial genomes in skeletal muscle fiber segments in mitochondrial disease. *Cell* **62**: 43-49.

Simonetti, S., X. Chen, S. DiMauro, and E. A. Schon. (1992). Accumulation of deletions in human mitochondrial DNA during normal aging: analysis by quantitative PCR. *Biochim. Biophys. Acta* **1180**: 113-122.

Sims, P., P. L. Grover, A. Swaisland, K. Pal, and A. Hewer. (1974). Metabolic activation of benzo[a]pyrene proceeds by a diol-epoxide. *Nature* **252**: 326-328.

Skopek, T. R., H. L. Liber, B. W. Penman, and W. G. Thilly. (1978). Isolation of a human lymphoblastoid line heterozygous at the thymidine kinase locus: Possibility for a rapid human cell mutation assay. *Biochem. Biophys. Res. Commun.* **84**: 411-416.

Sohal, R., and R. Weindruch. (1996). Oxidative stress, caloric restriction, and aging. *Science* **273**: 59-63.

Sohal, R. S., S. Agarwal, M. Candas, M. J. Forster, and H. Lal. (1994). Effect of age and caloric restriction on DNA oxidative damage in different tissues of C57BL/6 mice. *Mech. Ageing Dev.* **76**: 215-224.

Soong, N. W., D. R. Hinton, G. Cortopassi, and N. Arnheim. (1992). Mosaicism for a specific somatic mitochondrial DNA mutation in adult human brain. *Nat. Genet.* **2**: 318-23.

Stone, E. M., N. J. Newman, N. R. Miller, D. R. Johns, M. T. Lott, and D. C. Wallace. (1992). Visual recovery in patients with Leber's hereditary optic neuropathy and the 11778 mutation. *J. Clin. Neuroophthalmol.* **12**: 10-14.

Straub, K., T. Meehan, A. Burlingame, and M. Calvin. (1977). Identification of the major adducts formed by reaction of benzo[a]pyrene diol epoxide with DNA *in vitro*. *Proc. Natl. Acad. Sci. USA* **74**: 5285-89.

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

Su, S., and P. Modrich. (1986). *Escherichia coli* mutS-encoded protein binds to mismatched DNA base pairs. *Proc. Natl. Acad. Sci. USA* **83**: 5057-5061.

Sugiyama, S., K. Hattori, M. Hayakawa, and T. Ozawa. (1991). Quantitative analysis of age-associated accumulation of mitochondrial DNA with deletion in human hearts. *Biochem. Biophys. Res. Commun.* **180**: 894-899.

Sugiyama, S., M. Takasawa, M. Hayakawa, and T. Ozawa. (1993). Changes in skeletal muscle, heart and liver mitochondrial electron transport activities in rats and dogs of various ages. *Biochem. Mol. Biol. Int.* **30**: 937-944.

Suomalainen, A., A. Majander, M. Haltia, H. Somer, J. Lönnquist, M.-L. Savontaus, and L. Peltonen. (1992). Multiple deletions of mitochondrial DNA in several tissues of a patient with severe retarded depression and familial progressive external ophthalmoplegia. *J. Clin. Invest.* **90**: 61-66.

Sweeney, M. G., M. B. Davis, A. Lashwood, M. Bronckington, A. Toscano, and A. E. Harding. (1992). Evidence against an X-linked locus close to DXS7 determining visual loss susceptibility in British and Italian families with Leber hereditary optic neuropathy. *Am. J. Hum. Genet.* **51**: 741-748.

Szybalski, W., and E. H. Szybalski. (1962). Drug sensitivity as a genetic marker for a human cell line. *U. Mich. Med. Bull.* **28**: 277-293.

Takasawa, M., M. Hayakawa, S. Sugiyama, K. Hattori, T. Ito, and T. Ozawa. (1993). Age-associated damage in mitochondrial function in rat hearts. *Exp. Gerontol.* **28**: 269-280.

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 T-lymphocytes from five control populations. *Mutat. Res.* **253**: 199-213.

Terzaghi, M., P. Nettesheim, and M. L. William. (1978). Repopulation of denuded tracheal grafts with normal preneoplastic and neoplastic epithelial cell populations. *Cancer Res.* **38**: 4546-4553.

Thilly, W. "Potential use of gradient denaturing gel electrophoresis in obtaining mutational spectra from human cells." *Carcinogenesis*. Eds. E. Huberman and S. Barr. New York: Raven Press, 1985. 10: 511-528.

Thilly, W. G. (1990). Mutational spectrometry in animal toxicity testing. *Annu. Rev. Pharmacol. Toxicol.* **30**: 369-385.

Thyagarajan, B., R. A. Padua, and C. Campbell. (1996). Mammalian mitochondria possess homologous DNA recombination activity. *J. Biol. Chem.* **271**: 27536-27543.

Tiano, H. F., R.-L. Wang, M. Hosokawa, C. Crespi, K. R. Tindall, and R. Langenbach. (1994). Human CYP2A6 activation of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK): mutational specificity in the *gpt* gene of AS52 cells. *Carcinogenesis* **15**: 2859-2866.

Tomkinson, A. E., R. T. Bonk, J. Kim, Bartfeld, and S. Linn. (1990). Mammalian mitochondrial endonuclease activities specific for ultraviolet-irradiated DNA. *Nucl. Acids Res.* **18**: 929-935.

Tomkinson, A. E., R. T. Bonk, and S. Linn. (1988). Mitochondrial endonuclease activities specific for apurinic/aprimidinic sites in DNA from mouse cells. *J. Biol. Chem.* **263**: 12532-12527.

Tremblay, G. (1969). Oncocytes, a review. *Meth. Ach. Exp. Path.* **4**: 121-140.

Van Schooten, F. J., M. J. X. Hillebrand, F. E. Leeuwen, N. V. Zandwijk, H. M. Jansen, L. D. Engelse, and E. Kriek. (1992). Polycyclic aromatic hydrocarbon-DNA in white blood cells from lung cancer patients: no correlation with adducts in lung. *Carcinogenesis* **13**: 987-993.

Van Tuyle, G. C., J. P. Gudikote, V. R. Hurt, B. B. Miller, and C. A. Moore. (1996). Multiple, large deletions in rat mitochondrial DNA: Evidence for a major hot spot. *Mutat. Res.* **349**: 95-107.

Venema, G. (1959). The influence of substances extracted from cigarette-smoke on mitotic processes in *Allum cepa*. *Chromosoma* **10**: 679-685.

Vijayalaxmi, and H. Evans. (1984). Measurement of spontaneous and X-irradiation-induced 6-thioguanine-resistant human blood lymphocytes using a T-cell cloning technique. *Mutat. Res.* **125**: 87-94.

Vilkki, J., M.-L. Savontaus, and E. K. Nikoskelainen. (1989). Genetic heterogeneity in Leber hereditary optic neuroretinopathy revealed by mitochondrial DNA polymorphism. *Am. J. Hum. Genet.* **45**: 206-211.

Vrieling, H., J. C. P. Thijssen, A. M. Rossi, F. J. van Dam, A. T. Natarajan, A. D. Bates, and A. A. van Zeeland. (1992). Enhanced *hprt* mutant frequency but no significant difference in mutation spectrum between a smoking and nonsmoking human population. *Carcinogenesis* **13**: 1625-31.

Wallace, D., X. Zheng, M. Lott, J. Shoffner, J. Hodge, R. Kelley, C. Epstein, and L. Hopkins. (1988). Familial mitochondrial encephalomyopathy (MERRF): Genetic, pathophysiological, and biochemical characterization of a mitochondrial DNA disease. *Cell* **55**: 601-610.

Wallace, D. C. (1989). Mitochondrial DNA mutations and neuromuscular disease. *Trends Genet.* **5**: 9-13.

Wallace, D. C., G. Singh, M. T. Lott, J. A. Hodge, T. G. Schurr, A. M. S. Lessa, L. J. Elsas, and E. K. Nikoskelainen. (1988). Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy. *Science* **242**: 1427-1430.

Wallace, D. C., J. Ye, S. N. Neckelmann, G. Singh, K. A. Webster, and B. G. Greenberg. (1987). Sequence analysis of cDNAs for the human and bovine ATP synthase

β subunit: mitochondrial DNA genes sustain seventeen times more mutations. *Curr. Genet.* **12**: 81-90.

Wang, E., A. Wong, and G. Cortopassi. (1997). The rate of mitochondrial mutagenesis is faster in mice than humans. *Mutat. Res.* **377**: 157-166.

Warburg, O. (1956). On the origin of cancer cells. *Science* **123**: 309-314.

Weindruch, R. (1996). Retardation of aging by caloric restriction in rodents and primates. *Toxicol. Pathol.* **24**: 742-5.

Weinstein, I. B., A. M. Jeffrey, K. W. Jenette, S. H. Blobstein, R. G. Harvey, C. Harris, H. Autrup, H. Kasai, and K. Nakanishi. (1976). Benzo[a]pyrene diol epoxides as intermediates in nucleic acid binding *in vitro*. *Science* **193**: 592-595.

Weiss, W., K. R. Boucot, H. Seidman, and W. J. Carnahan. (1972). Risk of lung cancer according to histologic type and cigarette dosage. *JAMA* **222**: 799-801.

Weston, A., and E. D. Bowman. (1991). Fluorescence detection of benzo[a]pyrene-DNA adducts in human lung. *Carcinogenesis* **12**: 1445-1449.

Wiencke, J. K., K. T. Kelsey, A. Varkonyi, K. Semey, J. C. Wain, E. Mark, and D. C. Christiani. (1995). Correlation of DNA adducts in blood mononuclear cells with tobacco carcinogen-induced damage in human lung. *Cancer Res.* **55**: 4910-4914.

Wilkey, D. D., P. S. Lee, F. J. Hass, T. R. Gerrity, D. B. Yeates, and R. V. Lourenco. (1980). Mucociliary clearance of deposited particles from the human lungs: Intra- and inter-subject reproductivity, total and regional lung clearance, and model comparisons. *Arch. Environ. Health* **35**: 294-302.

Wilkinson, R., A. Hawks, and A. E. Pegg. (1975). Methylation of rat liver mitochondrial deoxyribonucleic acid by chemical carcinogens and associated alterations in physical properties. *Chem.-Biol. Interactions* **9**: 157-167.

Williams, R. S., S. Salmons, E. A. Newsholme, R. E. Kaufman, and J. Mellor. (1986). Regulation of nuclear and mitochondrial gene expression by contractile activity in skeletal muscle. *J. Biol. Chem.* **261**: 376-380.

Wood, A. W., R. L. Chang, W. Levin, H. Yagi, D. R. Thakker, D. M. Jerina, and A. H. Conney. (1977). Differences in mutagenicity of the optical enantiomers of the diastereomeric benzo[a]pyrene 7,8-diol-9,10-epoxides. *Biochem. Biophys. Res. Commun.* **77**: 1389-1396.

Wu, D. Y., L. Ugozzoli, B. K. Pal, and R. B. Wallace. (1989a). Allele-specific enzymatic amplification of beta-globin genomic DNA for diagnosis of sickle cell anemia. *Proc. Natl. Acad. Sci. USA* **86**: 2757-2760.

Wu, D. Y., and R. B. Wallace. (1989b). The ligation amplification reaction (LAR)--amplification of specific DNA sequences using sequential rounds of template-dependent ligation. *Genomics* **4**: 560-569.

Wunderlich, V., M. Schütt, M. Böttger, and A. Graffi. (1970). Preferential alkylation of mitochondrial deoxyribonucleic acid by N-methyl-N-nitrosourea. *Biochem. J.* **118**: 99-109.

Wynder, E. L., and E. A. Graham. (1950). Tobacco smoking as a possible etiologic factor in bronchiogenic carcinoma. *J. Am. Med. Assoc.* **143**: 329-336.

Yamamoto, H., M. Tanaka, M. Katayama, T. Obayashi, Y. Nimura, and T. Ozawa. (1992). Significant existence of deleted mitochondrial DNA in cirrhotic liver surrounding hepatic tumor. *Biochem. Biophys. Res. Commun.* **182**: 913-920.

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.

Yang, J.-L., V. M. Maher, and J. J. McCormick. (1987). Kinds of mutations formed when a shuttle vector containing adducts of ($\pm$)-7b,8a-dihydroxy-9a,10a-epoxy-7,8,9,10-tetrahydrobenzo[a]pyrene replicates in human cells. *Proc. Natl. Acad. Sci. USA* **84**: 3787-3791.

Yen, T.-C., C.-Y. Pang, R.-H. Hsieh, C.-H. Su, K.-L. King, and Y.-H. Wei. (1992). Age-dependent 6 kb deletion in human liver mitochondrial DNA. *Biochem. Int.* **26**: 457-468.

Yen, T.-C., J.-H. Su, K.-L. King, and Y.-H. Wei. (1991). Ageing-associated 5 kb deletion in human liver mitochondrial DNA. *Biochem. Biophys. Res. Commun.* **178**: 124-131.

Yoneda M., A. Chomyn, A. Martinuzzi, O. Hurko, and G. Attardi. (1992). Marked replicative advantage of human mtDNA carrying a point mutation that causes the MELAS encephalomyopathy. *Proc. Natl. Acad. Sci., USA* **89**: 1164-11648.

Zepeda, M. L., M. R. Chinoy, and J. M. Wilson. (1995). Characterization of stem cells in human airway capable of reconstituting a fully differentiated bronchial epithelium. *Somatic Cell Mol. Genet.* **21**: 61-73.

Zeviani, M., E. Bonilla, D. C. De Vivo, and S. DiMauro. (1989a). Mitochondrial diseases. *Neurol. Clin.* **7**: 123-156.

Zeviani, M., C. T. Moraes, S. di Mauro, H. Nakase, E. Bonilla, E. A. Schon, and L. P. Rowland. (1988). Deletions of mitochondrial DNA in Kearns-Sayre syndrome. *Neurology* **38**: 1339-1346.

Zeviani, M., S. Servidei, C. Gellera, E. Bertini, S. DiMauro, and S. DiDonato. (1989b). An autosomal dominant disorder with multiple deletions of mitochondrial DNA starting at the D-loop region. *Nature* **339**: 309-311.

Zhang, C., A. Baumer, R. J. Maxwell, A. W. Linnane, and P. Nagley. (1992). Multiple mitochondrial DNA deletions in an elderly human individual. *FEBS Lett.* **297**: 34-38.

Zhang, C., A. W. Linnane, and N. Phillip. (1993). Occurrence of a particular base substitution (3243 A to G) in mitochondrial DNA of tissues of ageing humans. *Biochem. Biophys. Res. Commun.* **195**: 1104-1110.

APPENDIX: TABLE OF MUTANT FRACTIONS

Bronch or Dissected Lung?	Subject	Sex	Age	Twin?	Smoker?	Gel	Lobe	Sample	Bifurcation #	Peak Number	Mutant Fraction
bronchoscopy	BG	F	51	Yes	Smoker	J17	right main carina	1A	1	1	1.20E-4
bronchoscopy	BG	F	51	Yes	Smoker	J17	right main carina	1A	1	2	0
bronchoscopy	BG	F	51	Yes	Smoker	J17	right main carina	1A	1	3	3.82E-5
bronchoscopy	BG	F	51	Yes	Smoker	J17	right main carina	1A	1	5	2.52E-5
bronchoscopy	BG	F	51	Yes	Smoker	J17	right main carina	1A	1	5.5	1.18E-4
bronchoscopy	BG	F	51	Yes	Smoker	J17	right main carina	1A	1	6	1.63E-5
bronchoscopy	BG	F	51	Yes	Smoker	J17	right main carina	1A	1	7	7.80E-5
bronchoscopy	BG	F	51	Yes	Smoker	J17	right main carina	1A	1	11	7.89E-5
bronchoscopy	BG	F	51	Yes	Smoker	J17	right main carina	1A	1	11.3	8.42E-5
bronchoscopy	BG	F	51	Yes	Smoker	J17	right main carina	1A	1	12	6.32E-5
bronchoscopy	BG	F	51	Yes	Smoker	J17	right main carina	1A	1	13	1.00E-4
bronchoscopy	BG	F	51	Yes	Smoker	J17	right main carina	1A	1	13.5	2.21E-5
bronchoscopy	BG	F	51	Yes	Smoker	J17	right main carina	1A	1	14	1.84E-4
bronchoscopy	BG	F	51	Yes	Smoker	J17	right main carina	1A	1	15	3.47E-6
bronchoscopy	BG	F	51	Yes	Smoker	J17	right main carina	1A	1	16	2.26E-3
bronchoscopy	BG	F	51	Yes	Smoker	J17	right main carina	1A	1	19	6.84E-6
bronchoscopy	BG	F	51	Yes	Smoker	J19	right upper lobe	2A	2	1	3.37E-04
bronchoscopy	BG	F	51	Yes	Smoker	J19	right upper lobe	2A	2	2	3.89E-05
bronchoscopy	BG	F	51	Yes	Smoker	J19	right upper lobe	2A	2	3	5.33E-5
bronchoscopy	BG	F	51	Yes	Smoker	J19	right upper lobe	2A	2	5	4.33E-5
bronchoscopy	BG	F	51	Yes	Smoker	J19	right upper lobe	2A	2	5.5	0.00E+00
bronchoscopy	BG	F	51	Yes	Smoker	J19	right upper lobe	2A	2	6	3.73E-05
bronchoscopy	BG	F	51	Yes	Smoker	J19	right upper lobe	2A	2	7	4.73E-05
bronchoscopy	BG	F	51	Yes	Smoker	J19	right upper lobe	2A	2	11	7.90E-06
bronchoscopy	BG	F	51	Yes	Smoker	J19	right upper lobe	2A	2	11.3	7.27E-06
bronchoscopy	BG	F	51	Yes	Smoker	J19	right upper lobe	2A	2	11.5	0
bronchoscopy	BG	F	51	Yes	Smoker	J19	right upper lobe	2A	2	12	1.21E-05
bronchoscopy	BG	F	51	Yes	Smoker	J19	right upper lobe	2A	2	13	1.00E-04
bronchoscopy	BG	F	51	Yes	Smoker	J19	right upper lobe	2A	2	13	1.00E-04
bronchoscopy	BG	F	51	Yes	Smoker	J19	right upper lobe	2A	2	13.5	1.00E-05
bronchoscopy	BG	F	51	Yes	Smoker	J19	right upper lobe	2A	2	14	2.02E-05
bronchoscopy	BG	F	51	Yes	Smoker	J19	right upper lobe	2A	2	15	1.21E-05
bronchoscopy	BG	F	51	Yes	Smoker	J19	right upper lobe	2A	2	15.5	0.00E+00
bronchoscopy	BG	F	51	Yes	Smoker	J19	right upper lobe	2A	2	15.7	3.31E-05
bronchoscopy	BG	F	51	Yes	Smoker	J19	right upper lobe	2A	2	16	6.00E-05
bronchoscopy	BG	F	51	Yes	Smoker	J19	right upper lobe	2A	2	18	1.40E-05
bronchoscopy	BG	F	51	Yes	Smoker	J19	right upper lobe	2A	2	18.5	3.00E-06
bronchoscopy	BG	F	51	Yes	Smoker	J19	right upper lobe	2A	2	19	1.10E-05
bronchoscopy	BG	F	51	Yes	Smoker	J26	right lower lobe	4A	3	1	9.49E-04
bronchoscopy	BG	F	51	Yes	Smoker	J24	right lower lobe	4A	3	2	0.00E+00
bronchoscopy	BG	F	51	Yes	Smoker	J26	right lower lobe	4A	3	3	5.24E-05
bronchoscopy	BG	F	51	Yes	Smoker	J26	right lower lobe	4A	3	5	2.93E-04
bronchoscopy	BG	F	51	Yes	Smoker	J26	right lower lobe	4A	3	5.5	0
bronchoscopy	BG	F	51	Yes	Smoker	J26	right lower lobe	4A	3	6	1.16E-04
bronchoscopy	BG	F	51	Yes	Smoker	J24	right lower lobe	4A	3	7	0.00E+00
bronchoscopy	BG	F	51	Yes	Smoker	J24	right lower lobe	4A	3	11	6.84E-05
bronchoscopy	BG	F	51	Yes	Smoker	J24	right lower lobe	4A	3	11.3	1.91E-05
bronchoscopy	BG	F	51	Yes	Smoker	J24	right lower lobe	4A	3	11.5	3.21E-06
bronchoscopy	BG	F	51	Yes	Smoker	J24	right lower lobe	4A	3	12	0
bronchoscopy	BG	F	51	Yes	Smoker	J24	right lower lobe	4A	3	13	1.00E-04
bronchoscopy	BG	F	51	Yes	Smoker	J26	right lower lobe	4A	3	13	1.00E-04
bronchoscopy	BG	F	51	Yes	Smoker	J24	right lower lobe	4A	3	13.5	1.82E-05
bronchoscopy	BG	F	51	Yes	Smoker	J24	right lower lobe	4A	3	14	2.54E-05
bronchoscopy	BG	F	51	Yes	Smoker	J24	right lower lobe	4A	3	15	1.39E-05
bronchoscopy	BG	F	51	Yes	Smoker	J24	right lower lobe	4A	3	15.7	2.18E-04
bronchoscopy	BG	F	51	Yes	Smoker	J24	right lower lobe	4A	3	16	1.93E-05
bronchoscopy	BG	F	51	Yes	Smoker	J24	right lower lobe	4A	3	18	1.68E-05
bronchoscopy	BG	F	51	Yes	Smoker	J26	right lower lobe	4A	3	18.5	0

Branch or Dissected Lung?	Subject	Sex	Age	Twin?	Smoker?	Gel	Lobe	Sample	Bifurcation #	Peak Number	Mutant Fraction
bronchoscopy	BG	F	51	Yes	Smoker	J24	right lower lobe	4A	3	19	3.81E-05
bronchoscopy	BG	F	51	Yes	Smoker	J24	left lower lobe	8A	3	1	4.07E-5
bronchoscopy	BG	F	51	Yes	Smoker	J24	left lower lobe	8A	3	2	2.03E-6
bronchoscopy	BG	F	51	Yes	Smoker	J24	left lower lobe	8A	3	3	8.94E-6
bronchoscopy	BG	F	51	Yes	Smoker	J24	left lower lobe	8A	3	5	3.25E-5
bronchoscopy	BG	F	51	Yes	Smoker	J24	left lower lobe	8A	3	5.5	4.47E-6
bronchoscopy	BG	F	51	Yes	Smoker	J24	left lower lobe	8A	3	6	1.14E-5
bronchoscopy	BG	F	51	Yes	Smoker	J24	left lower lobe	8A	3	7	0.00E+0
bronchoscopy	BG	F	51	Yes	Smoker	J24	left lower lobe	8A	3	11	9.62E-05
bronchoscopy	BG	F	51	Yes	Smoker	J24	left lower lobe	8A	3	11.3	7.88E-05
bronchoscopy	BG	F	51	Yes	Smoker	J24	left lower lobe	8A	3	11.5	8.46E-06
bronchoscopy	BG	F	51	Yes	Smoker	J24	left lower lobe	8A	3	12	2.31E-05
bronchoscopy	BG	F	51	Yes	Smoker	J24	left lower lobe	8A	3	13	1.00E-04
bronchoscopy	BG	F	51	Yes	Smoker	J24	left lower lobe	8A	3	13.5	5.77E-06
bronchoscopy	BG	F	51	Yes	Smoker	J24	left lower lobe	8A	3	14	7.69E-06
bronchoscopy	BG	F	51	Yes	Smoker	J24	left lower lobe	8A	3	15	6.73E-06
bronchoscopy	BG	F	51	Yes	Smoker	J24	left lower lobe	8A	3	15.7	2.31E-05
bronchoscopy	BG	F	51	Yes	Smoker	J24	left lower lobe	8A	3	16	2.40E-05
bronchoscopy	BG	F	51	Yes	Smoker	J24	left lower lobe	8A	3	18	1.10E-05
bronchoscopy	BG	F	51	Yes	Smoker	J24	left lower lobe	8A	3	18.5	1.35E-05
bronchoscopy	BG	F	51	Yes	Smoker	J24	left lower lobe	8A	3	19	1.50E-05
bronchoscopy	BO	F	51	Yes	Nonsmoker	J17	right main carina	1A	1	1	2.82E-4
bronchoscopy	BO	F	51	Yes	Nonsmoker	J17	right main carina	1A	1	2	5.33E-6
bronchoscopy	BO	F	51	Yes	Nonsmoker	J17	right main carina	1A	1	3	2.23E-5
bronchoscopy	BO	F	51	Yes	Nonsmoker	J17	right main carina	1A	1	5	4.64E-5
bronchoscopy	BO	F	51	Yes	Nonsmoker	J17	right main carina	1A	1	5.5	4.47E-5
bronchoscopy	BO	F	51	Yes	Nonsmoker	J17	right main carina	1A	1	6	3.36E-5
bronchoscopy	BO	F	51	Yes	Nonsmoker	J17	right main carina	1A	1	7	1.00E-04
bronchoscopy	BO	F	51	Yes	Nonsmoker	J19	right upper lobe	2A	2	1	5.54E-04
bronchoscopy	BO	F	51	Yes	Nonsmoker	J19	right upper lobe	2A	2	2	5.88E-05
bronchoscopy	BO	F	51	Yes	Nonsmoker	J19	right upper lobe	2A	2	3	1.64E-04
bronchoscopy	BO	F	51	Yes	Nonsmoker	J19	right upper lobe	2A	2	5	3.60E-04
bronchoscopy	BO	F	51	Yes	Nonsmoker	J19	right upper lobe	2A	2	5.5	1.03E-04
bronchoscopy	BO	F	51	Yes	Nonsmoker	J19	right upper lobe	2A	2	6	1.19E-04
bronchoscopy	BO	F	51	Yes	Nonsmoker	J19	right upper lobe	2A	2	7	1.19E-04
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	right lower lobe	4A	3	1	1.34E-3
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	right lower lobe	4A	3	2	1.71E-4
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	right lower lobe	4A	3	3	1.04E-4
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	right lower lobe	4A	3	5	1.83E-4
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	right lower lobe	4A	3	5.5	5.00E-05
bronchoscopy	BO	F	51	Yes	Nonsmoker	J26	right lower lobe	4A	3	6	6.00E-05
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	right lower lobe	4A	3	7	1.07E-03
bronchoscopy	BO	F	51	Yes	Nonsmoker	J26	right lower lobe	4A	3	11	7.50E-06
bronchoscopy	BO	F	51	Yes	Nonsmoker	J26	right lower lobe	4A	3	11.3	4.50E-06
bronchoscopy	BO	F	51	Yes	Nonsmoker	J26	right lower lobe	4A	3	11.5	5.30E-06
bronchoscopy	BO	F	51	Yes	Nonsmoker	J26	right lower lobe	4A	3	12	0
bronchoscopy	BO	F	51	Yes	Nonsmoker	J26	right lower lobe	4A	3	13	1.00E-05
bronchoscopy	BO	F	51	Yes	Nonsmoker	J26	right lower lobe	4A	3	13.5	6.60E-06
bronchoscopy	BO	F	51	Yes	Nonsmoker	J26	right lower lobe	4A	3	14	7.90E-06
bronchoscopy	BO	F	51	Yes	Nonsmoker	J26	right lower lobe	4A	3	15	3.90E-06
bronchoscopy	BO	F	51	Yes	Nonsmoker	J26	right lower lobe	4A	3	15.7	5.30E-06
bronchoscopy	BO	F	51	Yes	Nonsmoker	J26	right lower lobe	4A	3	16	8.80E-06
bronchoscopy	BO	F	51	Yes	Nonsmoker	J26	right lower lobe	4A	3	18	2.80E-06
bronchoscopy	BO	F	51	Yes	Nonsmoker	J26	right lower lobe	4A	3	18.5	0.00E+00
bronchoscopy	BO	F	51	Yes	Nonsmoker	J26	right lower lobe	4A	3	19	2.00E-06
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	left lower lobe	8A	3	1	6.30E-04
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	left lower lobe	8A	3	2	0
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	left lower lobe	8A	3	3	1.00E-4

Bronch or Dissected Lung?	Subject	Sex	Age	Twin?	Smoker?	Gel	Lobe	Sample	Bifurcation #	Peak Number	Mutant Fraction
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	left lower lobe	8A	3	5	3.33E-4
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	left lower lobe	8A	3	5.5	1.00E-4
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	left lower lobe	8A	3	6	1.20E-04
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	left lower lobe	8A	3	7	1.00E-05
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	left lower lobe	8A	3	11	5.10E-05
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	left lower lobe	8A	3	11.3	9.38E-05
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	left lower lobe	8A	3	11.5	3.40E-05
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	left lower lobe	8A	3	12	2.00E-05
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	left lower lobe	8A	3	13	1.00E-04
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	left lower lobe	8A	3	13.5	9.50E-06
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	left lower lobe	8A	3	14	0
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	left lower lobe	8A	3	15	3.85E-06
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	left lower lobe	8A	3	15.7	2.46E-05
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	left lower lobe	8A	3	16	1.40E-04
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	left lower lobe	8A	3	18	2.90E-05
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	left lower lobe	8A	3	18.5	1.20E-05
bronchoscopy	BO	F	51	Yes	Nonsmoker	J24	left lower lobe	8A	3	19	4.30E-06
bronchoscopy	JC	F	54	Yes	Nonsmoker	G26	right main carina	1A	1	1	7.90E-5
bronchoscopy	JC	F	54	Yes	Nonsmoker	G26	right main carina	1A	1	2	5.97E-6
bronchoscopy	JC	F	54	Yes	Nonsmoker	G26	right main carina	1A	1	3	3.23E-5
bronchoscopy	JC	F	54	Yes	Nonsmoker	G26	right main carina	1A	1	5	5.81E-6
bronchoscopy	JC	F	54	Yes	Nonsmoker	G26	right main carina	1A	1	5.5	2.74E-5
bronchoscopy	JC	F	54	Yes	Nonsmoker	G26	right main carina	1A	1	6	1.13E-5
bronchoscopy	JC	F	54	Yes	Nonsmoker	G26	right main carina	1A	1	7	6.61E-6
bronchoscopy	JC	F	54	Yes	Nonsmoker	G26	right main carina	1A	1	11	1.20E-05
bronchoscopy	JC	F	54	Yes	Nonsmoker	G26	right main carina	1A	1	11.3	1.93E-5
bronchoscopy	JC	F	54	Yes	Nonsmoker	G26	right main carina	1A	1	11.5	0.00E+0
bronchoscopy	JC	F	54	Yes	Nonsmoker	G26	right main carina	1A	1	11.7	1.30E-6
bronchoscopy	JC	F	54	Yes	Nonsmoker	G26	right main carina	1A	1	12	1.10E-05
bronchoscopy	JC	F	54	Yes	Nonsmoker	G26	right main carina	1A	1	13	1.00E-4
bronchoscopy	JC	F	54	Yes	Nonsmoker	G26	right main carina	1A	1	13.5	3.00E-6
bronchoscopy	JC	F	54	Yes	Nonsmoker	G26	right main carina	1A	1	14	1.00E-6
bronchoscopy	JC	F	54	Yes	Nonsmoker	G26	right main carina	1A	1	15	8.86E-6
bronchoscopy	JC	F	54	Yes	Nonsmoker	G26	right main carina	1A	1	15.5	8.50E-7
bronchoscopy	JC	F	54	Yes	Nonsmoker	G26	right main carina	1A	1	16	3.64E-5
bronchoscopy	JC	F	54	Yes	Nonsmoker	G26	right main carina	1A	1	18	1.34E-5
bronchoscopy	JC	F	54	Yes	Nonsmoker	G26	right main carina	1A	1	18.5	4.60E-5
bronchoscopy	JC	F	54	Yes	Nonsmoker	G26	right main carina	1A	1	19	5.20E-6
bronchoscopy	JC	F	54	Yes	Nonsmoker	J24	right lower lobe	4A	3	1	3.03E-04
bronchoscopy	JC	F	54	Yes	Nonsmoker	J24	right lower lobe	4A	3	2	2.64E-05
bronchoscopy	JC	F	54	Yes	Nonsmoker	J24	right lower lobe	4A	3	3	3.65E-05
bronchoscopy	JC	F	54	Yes	Nonsmoker	J24	right lower lobe	4A	3	5	8.42E-05
bronchoscopy	JC	F	54	Yes	Nonsmoker	J24	right lower lobe	4A	3	5.5	9.00E-05
bronchoscopy	JC	F	54	Yes	Nonsmoker	J24	right lower lobe	4A	3	6	8.22E-05
bronchoscopy	JC	F	54	Yes	Nonsmoker	J24	right lower lobe	4A	3	7	7.30E-05
bronchoscopy	JC	F	54	Yes	Nonsmoker	J26	right lower lobe	4A	3	11	1.50E-05
bronchoscopy	JC	F	54	Yes	Nonsmoker	J26	right lower lobe	4A	3	11.3	6.40E-06
bronchoscopy	JC	F	54	Yes	Nonsmoker	J26	right lower lobe	4A	3	11.5	9.60E-06
bronchoscopy	JC	F	54	Yes	Nonsmoker	J26	right lower lobe	4A	3	11.7	1.20E-05
bronchoscopy	JC	F	54	Yes	Nonsmoker	J26	right lower lobe	4A	3	12	5.40E-06
bronchoscopy	JC	F	54	Yes	Nonsmoker	J26	right lower lobe	4A	3	13	1.00E-05
bronchoscopy	JC	F	54	Yes	Nonsmoker	J26	right lower lobe	4A	3	13.5	1.20E-05
bronchoscopy	JC	F	54	Yes	Nonsmoker	J26	right lower lobe	4A	3	14	1.90E-05
bronchoscopy	JC	F	54	Yes	Nonsmoker	J26	right lower lobe	4A	3	15	2.60E-05
bronchoscopy	JC	F	54	Yes	Nonsmoker	J26	right lower lobe	4A	3	15.5	0.00E+00
bronchoscopy	JC	F	54	Yes	Nonsmoker	J26	right lower lobe	4A	3	16	4.30E-05
bronchoscopy	JC	F	54	Yes	Nonsmoker	J26	right lower lobe	4A	3	18	4.60E-05
bronchoscopy	JC	F	54	Yes	Nonsmoker	J26	right lower lobe	4A	3	18.5	5.00E-06

Bronch or Dissected Lung?	Subject	Sex	Age	Twin?	Smoker?	Gel	Lobe	Sample	Bifurcation #	Peak Number	Mutant Fraction
bronchoscopy	JC	F	54	Yes	Nonsmoker	J26	right lower lobe	4A	3	19	6.30E-06
bronchoscopy	KH	F	24	No	Smoker	J15	right main carina	1A	1	1	7.62E-5
bronchoscopy	KH	F	24	No	Smoker	J15	right main carina	1A	1	2	0.00E+0
bronchoscopy	KH	F	24	No	Smoker	J15	right main carina	1A	1	3	3.62E-5
bronchoscopy	KH	F	24	No	Smoker	J15	right main carina	1A	1	5	8.67E-5
bronchoscopy	KH	F	24	No	Smoker	J15	right main carina	1A	1	7	0.00E+0
bronchoscopy	KW	F	46	Yes	Smoker	J20	right main carina	1A	1	1	9.25E-05
bronchoscopy	KW	F	46	Yes	Smoker	J20	right main carina	1A	1	2	2.42E-06
bronchoscopy	KW	F	46	Yes	Smoker	J20	right main carina	1A	1	3	3.40E-5
bronchoscopy	KW	F	46	Yes	Smoker	J20	right main carina	1A	1	5	2.00E-5
bronchoscopy	KW	F	46	Yes	Smoker	J20	right main carina	1A	1	5.5	9.74E-05
bronchoscopy	KW	F	46	Yes	Smoker	J20	right main carina	1A	1	6	5.36E-6
bronchoscopy	KW	F	46	Yes	Smoker	J20	right main carina	1A	1	7	6.09E-6
bronchoscopy	KW	F	46	Yes	Smoker	J21	right upper lobe	2A	2	1	5.68E-4
bronchoscopy	KW	F	46	Yes	Smoker	J21	right upper lobe	2A	2	2	1.84E-5
bronchoscopy	KW	F	46	Yes	Smoker	J21	right upper lobe	2A	2	3	2.31E-4
bronchoscopy	KW	F	46	Yes	Smoker	J21	right upper lobe	2A	2	5	7.50E-5
bronchoscopy	KW	F	46	Yes	Smoker	J21	right upper lobe	2A	2	5.5	2.34E-4
bronchoscopy	KW	F	46	Yes	Smoker	J21	right upper lobe	2A	2	6	2.00E-5
bronchoscopy	KW	F	46	Yes	Smoker	J21	right upper lobe	2A	2	7	1.30E-04
bronchoscopy	KW	F	46	Yes	Smoker	J21	right upper lobe	2A	2	11	1.52E-04
bronchoscopy	KW	F	46	Yes	Smoker	J21	right upper lobe	2A	2	11.3	3.20E-05
bronchoscopy	KW	F	46	Yes	Smoker	J21	right upper lobe	2A	2	11.5	0
bronchoscopy	KW	F	46	Yes	Smoker	J21	right upper lobe	2A	2	11.7	1.20E-06
bronchoscopy	KW	F	46	Yes	Smoker	J21	right upper lobe	2A	2	12	3.00E-05
bronchoscopy	KW	F	46	Yes	Smoker	J21	right upper lobe	2A	2	13	1.00E-04
bronchoscopy	KW	F	46	Yes	Smoker	J21	right upper lobe	2A	2	13	1.00E-04
bronchoscopy	KW	F	46	Yes	Smoker	J21	right upper lobe	2A	2	13.5	1.93E-05
bronchoscopy	KW	F	46	Yes	Smoker	J21	right upper lobe	2A	2	14	3.50E-05
bronchoscopy	KW	F	46	Yes	Smoker	J21	right upper lobe	2A	2	15.7	6.45E-05
bronchoscopy	KW	F	46	Yes	Smoker	J21	right upper lobe	2A	2	16	1.68E-04
bronchoscopy	KW	F	46	Yes	Smoker	J21	right upper lobe	2A	2	18	5.90E-06
bronchoscopy	KW	F	46	Yes	Smoker	J21	right upper lobe	2A	2	18.5	2.80E-05
bronchoscopy	KW	F	46	Yes	Smoker	J21	right upper lobe	2A	2	19	1.30E-05
bronchoscopy	KW	F	46	Yes	Smoker	J24	right lower lobe	4A	3	1	4.67E-04
bronchoscopy	KW	F	46	Yes	Smoker	J24	right lower lobe	4A	3	2	
bronchoscopy	KW	F	46	Yes	Smoker	J24	right lower lobe	4A	3	3	5.33E-04
bronchoscopy	KW	F	46	Yes	Smoker	J24	right lower lobe	4A	3	5	1.50E-04
bronchoscopy	KW	F	46	Yes	Smoker	J24	right lower lobe	4A	3	6	1.50E-04
bronchoscopy	KW	F	46	Yes	Smoker	J24	right lower lobe	4A	3	7	1.50E-04
bronchoscopy	KW	F	46	Yes	Smoker	J24	right lower lobe	4A	3	11	1.88E-05
bronchoscopy	KW	F	46	Yes	Smoker	J24	right lower lobe	4A	3	11.3	1.00E-05
bronchoscopy	KW	F	46	Yes	Smoker	J24	right lower lobe	4A	3	11.5	6.25E-06
bronchoscopy	KW	F	46	Yes	Smoker	J24	right lower lobe	4A	3	12	1.56E-06
bronchoscopy	KW	F	46	Yes	Smoker	J24	right lower lobe	4A	3	13	1.00E-04
bronchoscopy	KW	F	46	Yes	Smoker	J24	right lower lobe	4A	3	13.5	4.38E-06
bronchoscopy	KW	F	46	Yes	Smoker	J24	right lower lobe	4A	3	14	1.25E-05
bronchoscopy	KW	F	46	Yes	Smoker	J24	right lower lobe	4A	3	15	1.06E-05
bronchoscopy	KW	F	46	Yes	Smoker	J24	right lower lobe	4A	3	15.7	9.38E-06
bronchoscopy	KW	F	46	Yes	Smoker	J24	right lower lobe	4A	3	16	9.38E-06
bronchoscopy	KW	F	46	Yes	Smoker	J24	right lower lobe	4A	3	18	5.13E-05
bronchoscopy	KW	F	46	Yes	Smoker	J24	right lower lobe	4A	3	18.5	6.88E-07
bronchoscopy	KW	F	46	Yes	Smoker	J24	right lower lobe	4A	3	19	1.63E-06
bronchoscopy	MD	F	46	Yes	Nonsmoker	J20	right main carina	1A	1	1	6.72E-4
bronchoscopy	MD	F	46	Yes	Nonsmoker	J20	right main carina	1A	1	2	2.64E-5
bronchoscopy	MD	F	46	Yes	Nonsmoker	J20	right main carina	1A	1	3	9.24E-5
bronchoscopy	MD	F	46	Yes	Nonsmoker	J20	right main carina	1A	1	5	1.06E-4
bronchoscopy	MD	F	46	Yes	Nonsmoker	J20	right main carina	1A	1	5.5	2.65E-5

Branch or Dissected Lung?	Subject	Sex	Age	Twin?	Smoker?	Gel	Lobe	Sample	Bifurcation #	Peak Number	Mutant Fraction
bronchoscopy	MD	F	46	Yes	Nonsmoker	J20	right main carina	1A	1	6	4.69E-5
bronchoscopy	MD	F	46	Yes	Nonsmoker	J20	right main carina	1A	1	7	7.04E-5
bronchoscopy	MD	F	46	Yes	Nonsmoker	J20	right main carina	1A	1	11	1.46E-04
bronchoscopy	MD	F	46	Yes	Nonsmoker	J20	right main carina	1A	1	11.3	6.27E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J20	right main carina	1A	1	11.5	3.00E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J20	right main carina	1A	1	11.7	5.64E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J20	right main carina	1A	1	12	1.02E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J20	right main carina	1A	1	13	1.00E-04
bronchoscopy	MD	F	46	Yes	Nonsmoker	J20	right main carina	1A	1	13.5	7.57E-06
bronchoscopy	MD	F	46	Yes	Nonsmoker	J20	right main carina	1A	1	14	3.40E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J20	right main carina	1A	1	15	6.27E-06
bronchoscopy	MD	F	46	Yes	Nonsmoker	J20	right main carina	1A	1	15.7	2.32E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J20	right main carina	1A	1	16	9.75E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J20	right main carina	1A	1	18	1.99E-04
bronchoscopy	MD	F	46	Yes	Nonsmoker	J20	right main carina	1A	1	18.5	1.02E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J20	right main carina	1A	1	19	3.18E-05
bronchoscopy	MD	F	51	Yes	Nonsmoker	J21	right upper lobe	2A	2	1	2.32E-4
bronchoscopy	MD	F	51	Yes	Nonsmoker	J21	right upper lobe	2A	2	2	2.36E-5
bronchoscopy	MD	F	51	Yes	Nonsmoker	J21	right upper lobe	2A	2	3	7.62E-5
bronchoscopy	MD	F	51	Yes	Nonsmoker	J21	right upper lobe	2A	2	5	6.72E-5
bronchoscopy	MD	F	51	Yes	Nonsmoker	J21	right upper lobe	2A	2	5.5	2.58E-5
bronchoscopy	MD	F	51	Yes	Nonsmoker	J21	right upper lobe	2A	2	6	4.29E-5
bronchoscopy	MD	F	51	Yes	Nonsmoker	J21	right upper lobe	2A	2	7	1.69E-5
bronchoscopy	MD	F	46	Yes	Nonsmoker	J21	right upper lobe	2A	2	11	6.01E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J21	right upper lobe	2A	2	11.3	1.59E-04
bronchoscopy	MD	F	46	Yes	Nonsmoker	J21	right upper lobe	2A	2	11.5	6.23E-06
bronchoscopy	MD	F	46	Yes	Nonsmoker	J21	right upper lobe	2A	2	11.7	2.86E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J21	right upper lobe	2A	2	12	4.90E-06
bronchoscopy	MD	F	46	Yes	Nonsmoker	J21	right upper lobe	2A	2	13	1.00E-04
bronchoscopy	MD	F	46	Yes	Nonsmoker	J21	right upper lobe	2A	2	13	1.00E-04
bronchoscopy	MD	F	46	Yes	Nonsmoker	J21	right upper lobe	2A	2	13.5	1.74E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J21	right upper lobe	2A	2	14	7.49E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J21	right upper lobe	2A	2	15	6.73E-06
bronchoscopy	MD	F	46	Yes	Nonsmoker	J21	right upper lobe	2A	2	15.5	0
bronchoscopy	MD	F	46	Yes	Nonsmoker	J21	right upper lobe	2A	2	15.7	2.00E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J21	right upper lobe	2A	2	16	4.21E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J21	right upper lobe	2A	2	18	1.57E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J21	right upper lobe	2A	2	18.5	5.90E-06
bronchoscopy	MD	F	46	Yes	Nonsmoker	J21	right upper lobe	2A	2	19	2.50E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J26	right lower lobe	4A	3	1	2.81E-04
bronchoscopy	MD	F	46	Yes	Nonsmoker	J26	right lower lobe	4A	3	2	6.19E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J26	right lower lobe	4A	3	3	2.29E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J26	right lower lobe	4A	3	5	7.62E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J26	right lower lobe	4A	3	5.5	0
bronchoscopy	MD	F	46	Yes	Nonsmoker	J26	right lower lobe	4A	3	6	5.71E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J26	right lower lobe	4A	3	7	0
bronchoscopy	MD	F	46	Yes	Nonsmoker	J26	right lower lobe	4A	3	11	5.31E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J26	right lower lobe	4A	3	11.3	0
bronchoscopy	MD	F	46	Yes	Nonsmoker	J26	right lower lobe	4A	3	11.5	0
bronchoscopy	MD	F	46	Yes	Nonsmoker	J26	right lower lobe	4A	3	12	0
bronchoscopy	MD	F	46	Yes	Nonsmoker	J26	right lower lobe	4A	3	13	1.00E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J26	right lower lobe	4A	3	13.5	2.31E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J26	right lower lobe	4A	3	14	1.38E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J26	right lower lobe	4A	3	15	2.08E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J26	right lower lobe	4A	3	15.5	2.31E-06
bronchoscopy	MD	F	46	Yes	Nonsmoker	J26	right lower lobe	4A	3	16	7.15E-05
bronchoscopy	MD	F	46	Yes	Nonsmoker	J26	right lower lobe	4A	3	18	7.69E-06
bronchoscopy	MD	F	46	Yes	Nonsmoker	J26	right lower lobe	4A	3	18.5	3.85E-06

Bronch or Dissected Lung?	Subject	Sex	Age	Twin?	Smoker?	Gel	Lobe	Sample	Bifurcation #	Peak Number	Mutant Fraction
bronchoscopy	MD	F	46	Yes	Nonsmoker	J26	right lower lobe	4A	3	19	7.38E-06
bronchoscopy	ND	F	54	Yes	Smoker	G26	right main carina	1A	1	1	5.00E-5
bronchoscopy	ND	F	54	Yes	Smoker	G26	right main carina	1A	1	2	1.82E-6
bronchoscopy	ND	F	54	Yes	Smoker	G26	right main carina	1A	1	3	7.91E-6
bronchoscopy	ND	F	54	Yes	Smoker	G26	right main carina	1A	1	5	2.09E-5
bronchoscopy	ND	F	54	Yes	Smoker	G26	right main carina	1A	1	5.5	4.55E-6
bronchoscopy	ND	F	54	Yes	Smoker	G26	right main carina	1A	1	6	9.09E-6
bronchoscopy	ND	F	54	Yes	Smoker	G26	right main carina	1A	1	7	5.64E-6
bronchoscopy	ND	F	54	Yes	Smoker	G26	right main carina	1A	1	11	9.20E-06
bronchoscopy	ND	F	54	Yes	Smoker	G26	right main carina	1A	1	11.3	6.21E-6
bronchoscopy	ND	F	54	Yes	Smoker	G26	right main carina	1A	1	12	5.50E-05
bronchoscopy	ND	F	54	Yes	Smoker	G26	right main carina	1A	1	13	1.00E-4
bronchoscopy	ND	F	54	Yes	Smoker	G26	right main carina	1A	1	13.5	2.80E-05
bronchoscopy	ND	F	54	Yes	Smoker	G26	right main carina	1A	1	14	1.52E-5
bronchoscopy	ND	F	54	Yes	Smoker	G26	right main carina	1A	1	15	7.93E-6
bronchoscopy	ND	F	54	Yes	Smoker	G26	right main carina	1A	1	16	1.40E-05
bronchoscopy	ND	F	54	Yes	Smoker	G26	right main carina	1A	1	18	2.50E-05
bronchoscopy	ND	F	54	Yes	Smoker	G26	right main carina	1A	1	18.5	6.40E-6
bronchoscopy	ND	F	54	Yes	Smoker	G26	right main carina	1A	1	19	6.30E-6
bronchoscopy	ND	F	54	Yes	Smoker	J16	right upper lobe	2A	2	1	9.02E-5
bronchoscopy	ND	F	54	Yes	Smoker	J16	right upper lobe	2A	2	2	5.12E-6
bronchoscopy	ND	F	54	Yes	Smoker	J16	right upper lobe	2A	2	3	1.39E-5
bronchoscopy	ND	F	54	Yes	Smoker	J16	right upper lobe	2A	2	5	1.15E-4
bronchoscopy	ND	F	54	Yes	Smoker	J16	right upper lobe	2A	2	5.5	1.32E-5
bronchoscopy	ND	F	54	Yes	Smoker	J16	right upper lobe	2A	2	6	1.15E-5
bronchoscopy	ND	F	54	Yes	Smoker	J16	right upper lobe	2A	2	7	3.66E-6
bronchoscopy	ND	F	54	Yes	Smoker	J16	right upper lobe	2A	2	11	1.10E-05
bronchoscopy	ND	F	54	Yes	Smoker	J16	right upper lobe	2A	2	11.3	6.00E-6
bronchoscopy	ND	F	54	Yes	Smoker	J16	right upper lobe	2A	2	11.5	0.00E+0
bronchoscopy	ND	F	54	Yes	Smoker	J16	right upper lobe	2A	2	12	3.00E-6
bronchoscopy	ND	F	54	Yes	Smoker	J16	right upper lobe	2A	2	13	1.00E-4
bronchoscopy	ND	F	54	Yes	Smoker	J16	right upper lobe	2A	2	13.5	9.70E-07
bronchoscopy	ND	F	54	Yes	Smoker	J16	right upper lobe	2A	2	14	1.50E-5
bronchoscopy	ND	F	54	Yes	Smoker	J16	right upper lobe	2A	2	15	8.00E-5
bronchoscopy	ND	F	54	Yes	Smoker	J16	right upper lobe	2A	2	15.5	0.00E+0
bronchoscopy	ND	F	54	Yes	Smoker	J16	right upper lobe	2A	2	16	1.10E-5
bronchoscopy	ND	F	54	Yes	Smoker	J16	right upper lobe	2A	2	18	1.10E-05
bronchoscopy	ND	F	54	Yes	Smoker	J16	right upper lobe	2A	2	18.5	1.10E-04
bronchoscopy	ND	F	54	Yes	Smoker	J16	right upper lobe	2A	2	19	3.70E-06
bronchoscopy	ND	F	54	Yes	Smoker	J24	right lower lobe	4A	3	1	8.22E-4
bronchoscopy	ND	F	54	Yes	Smoker	J24	right lower lobe	4A	3	2	1.10E-4
bronchoscopy	ND	F	54	Yes	Smoker	J24	right lower lobe	4A	3	3	6.11E-5
bronchoscopy	ND	F	54	Yes	Smoker	J24	right lower lobe	4A	3	5	3.33E-4
bronchoscopy	ND	F	54	Yes	Smoker	J24	right lower lobe	4A	3	5.5	6.67E-5
bronchoscopy	ND	F	54	Yes	Smoker	J24	right lower lobe	4A	3	6	8.89E-5
bronchoscopy	ND	F	54	Yes	Smoker	J24	right lower lobe	4A	3	7	8.89E-5
bronchoscopy	ND	F	54	Yes	Smoker	J24	right lower lobe	4A	3	11	2.90E-05
bronchoscopy	ND	F	54	Yes	Smoker	J24	right lower lobe	4A	3	11.3	2.80E-05
bronchoscopy	ND	F	54	Yes	Smoker	J24	right lower lobe	4A	3	11.5	3.60E-06
bronchoscopy	ND	F	54	Yes	Smoker	J24	right lower lobe	4A	3	12	3.20E-05
bronchoscopy	ND	F	54	Yes	Smoker	J24	right lower lobe	4A	3	13	1.00E-04
bronchoscopy	ND	F	54	Yes	Smoker	J24	right lower lobe	4A	3	13.5	7.30E-06
bronchoscopy	ND	F	54	Yes	Smoker	J24	right lower lobe	4A	3	14	3.50E-05
bronchoscopy	ND	F	54	Yes	Smoker	J24	right lower lobe	4A	3	15	1.39E-05
bronchoscopy	ND	F	54	Yes	Smoker	J24	right lower lobe	4A	3	15.7	5.56E-06
bronchoscopy	ND	F	54	Yes	Smoker	J24	right lower lobe	4A	3	16	7.10E-05
bronchoscopy	ND	F	54	Yes	Smoker	J24	right lower lobe	4A	3	18	4.80E-05
bronchoscopy	ND	F	54	Yes	Smoker	J24	right lower lobe	4A	3	18.5	2.20E-05

Bronch or Dissected Lung?	Subject	Sex	Age	Twin?	Smoker?	Gel	Lobe	Sample	Bifurcation #	Peak Number	Mutant Fraction
bronchoscopy	ND	F	54	Yes	Smoker	J24	right lower lobe	4A	3	19	1.80E-05
bronchoscopy	WS	M	28	No	Nonsmoker	J15	right main carina	1A	1	1	6.00E-5
bronchoscopy	WS	M	28	No	Nonsmoker	J15	right main carina	1A	1	2	0.00E+0
bronchoscopy	WS	M	28	No	Nonsmoker	J15	right main carina	1A	1	3	3.60E-5
bronchoscopy	WS	M	28	No	Nonsmoker	J15	right main carina	1A	1	5	6.90E-5
bronchoscopy	WS	M	28	No	Nonsmoker	J15	right main carina	1A	1	7	2.40E-5
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	1	3	1	6.88E-04
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	1	3	2	?
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	1	3	3	7.50E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	1	3	5	2.56E-04
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	1	3	5.5	3.13E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	1	3	6	4.81E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	1	3	7	4.38E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	1	3	11	1.07E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	1	3	11.3	1.55E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	1	3	11.5	0
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	1	3	12	0
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	1	3	13	1.00E-04
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	1	3	13.5	0
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	1	3	14	2.10E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	1	3	15	3.83E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	1	3	15.7	1.72E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	1	3	16	1.72E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	1	3	18	7.93E-06
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	1	3	18.5	0
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	1	3	19	0
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	8	4	1	5.00E-04
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	8	4	2	2.28E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	8	4	3	7.22E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	8	4	5	1.22E-04
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	8	4	5.5	5.00E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	8	4	6	5.56E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	8	4	7	7.22E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	8	4	11	3.70E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	8	4	11.3	8.10E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	8	4	11.5	0
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	8	4	11.7	9.40E-06
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	8	4	12	4.90E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	8	4	13	1.00E-04
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	8	4	13.5	1.10E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	8	4	14	4.80E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	8	4	15	1.14E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	8	4	15.7	6.70E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	8	4	16	3.90E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	8	4	18	2.00E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	8	4	18.5	1.10E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	8	4	19	2.30E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	10	5	1	1.11E-03
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	10	5	2	5.26E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	10	5	3	8.42E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	10	5	6	8.95E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	10	5	7	3.74E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	10	5	11	4.23E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	10	5	11.3	5.64E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	10	5	11.5	7.82E-06
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	10	5	12	1.03E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	10	5	13	1.00E-04
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	10	5	13.5	2.95E-05

Branch or Dissected Lung?	Subject	Sex	Age	Twin?	Smoker?	Gel	Lobe	Sample	Bifurcation #	Peak Number	Mutant Fraction
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	10	5	14	5.13E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	10	5	15	1.67E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	10	5	15.7	4.23E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	10	5	16	1.92E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	10	5	18	6.92E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	10	5	18.5	7.31E-06
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	10	5	19	?
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	11	6	11	4.30E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	11	6	11.3	3.97E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	11	6	11.5	3.90E-06
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	11	6	12	2.00E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	11	6	13	1.00E-04
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	11	6	13.5	4.30E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	11	6	14	1.28E-04
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	11	6	15	3.72E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	11	6	15.5	2.56E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	11	6	15.7	7.69E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	11	6	16	1.40E-04
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	11	6	18	4.30E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	11	6	18.5	3.00E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	11	6	19	2.20E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	12	7	1	7.63E-04
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	12	7	2	?
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	12	7	3	8.13E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	12	7	5	2.19E-04
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	12	7	5.5	4.38E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	12	7	6	6.06E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	12	7	7	3.81E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	12	7	11	3.41E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	12	7	11.3	6.35E-06
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	12	7	11.5	2.12E-06
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	12	7	12	2.12E-06
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	12	7	13	1.00E-04
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	12	7	13.5	9.41E-06
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	12	7	14	3.41E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	12	7	15	1.76E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	12	7	15.7	1.41E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	12	7	16	4.82E-06
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	12	7	18	2.59E-05
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	12	7	18.5	6.12E-06
dissected lung	28-Mar	M	84	No	Nonsmoker	J28	right upper lobe	12	7	19	?
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	1	2	1	2.19E-4
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	1	2	2	7.10E-5
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	1	2	3	9.00E-5
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	1	2	5	1.57E-4
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	1	2	5.5	1.00E-6
dissected lung	14-Nov	F	47	No	Smoker	J28	right lower lobe	1	2	5.5	1.13E-05
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	1	2	6	7.12E-5
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	1	2	7	1.75E-4
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	1	2	11	2.65E-05
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	1	2	11.3	1.90E-05
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	1	2	11.5	2.40E-05
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	1	2	11.7	3.90E-6
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	1	2	12	3.70E-06
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	1	2	13	1.00E-4
dissected lung	14-Nov	F	47	No	Smoker	J28	right lower lobe	1	2	13	1.00E-04
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	1	2	13.5	9.60E-06
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	1	2	14	3.50E-05

Branch or Dissected Lung?	Subject	Sex	Age	Twin?	Smoker?	Gel	Lobe	Sample	Bifurcation #	Peak Number	Mutant Fraction
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	1	2	15	1.47E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	1	2	15.7	5.60E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	1	2	16	6.70E-05
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	1	2	18	3.30E-05
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	1	2	18.5	2.30E-05
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	1	2	19	1.30E-05
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	2	3	1	4.80E-4
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	2	3	2	6.67E-5
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	2	3	3	4.53E-5
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	2	3	5	1.87E-4
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	2	3	5.5	4.67E-5
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	2	3	6	1.13E-4
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	2	3	7	1.13E-4
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	2	3	11	5.34E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	2	3	11.3	7.81E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	2	3	11.5	2.74E-6
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	2	3	12	2.60E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	2	3	13	1.00E-4
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	2	3	13.5	1.51E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	2	3	14	8.22E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	2	3	15	9.73E-6
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	2	3	15.7	8.22E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	2	3	16	8.22E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	2	3	18	2.05E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	2	3	18.5	5.48E-6
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	2	3	19	7.53E-6
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	3	3	1	4.41E-4
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	3	3	2	?
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	3	3	3	2.59E-5
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	3	3	5	7.06E-5
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	3	3	5.5	1.41E-5
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	3	3	6	5.41E-5
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	3	3	7	3.00E-5
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	5	4	1	3.00E-4
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	5	4	2	2.73E-6
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	5	4	3	2.09E-5
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	5	4	5	7.45E-5
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	5	4	5.5	1.00E-5
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	5	4	6	7.00E-5
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	5	4	7	1.64E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	5	4	11	2.30E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	5	4	11.3	1.03E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	5	4	11.5	6.08E-6
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	5	4	12	1.22E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	5	4	13	1.00E-4
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	5	4	13.5	5.27E-6
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	5	4	14	3.92E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	5	4	15	1.35E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	5	4	15.7	1.22E-4
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	5	4	16	1.35E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	5	4	18	5.41E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	5	4	18.5	2.70E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	5	4	19	6.76E-6
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	13	5	1	3.60E-4
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	13	5	2	?
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	13	5	3	4.20E-5
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	13	5	5	8.00E-5
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	13	5	5.5	3.00E-6

Branch or Dissected Lung?	Subject	Sex	Age	Twin?	Smoker?	Gel	Lobe	Sample	Bifurcation #	Peak Number	Mutant Fraction
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	13	5	6	3.25E-5
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	13	5	7	3.25E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	13	6	11	3.92E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	13	6	11.3	1.31E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	13	6	11.5	1.00E-6
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	13	6	12	6.15E-6
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	13	6	13	1.00E-4
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	13	6	13.5	1.23E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	13	6	14	2.08E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	13	6	15	7.69E-6
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	13	6	15.7	5.31E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	13	6	16	3.08E-7
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	13	6	18	2.77E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	13	6	18.5	4.62E-6
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	13	6	19	2.31E-6
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	22	8	1	3.58E-4
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	22	8	2	1.43E-6
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	22	8	3	6.35E-5
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	22	8	5	2.31E-4
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	22	8	5.5	5.20E-6
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	22	8	6	4.93E-5
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	22	8	7	4.39E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	22	8	11	1.76E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	22	8	11.3	1.18E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	22	8	11.5	3.25E-6
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	22	8	12	7.53E-6
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	22	8	13	1.00E-4
dissected lung	14-Nov	F	47	No	Smoker	J28	right lower lobe	22	8	13	1.00E-04
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	22	8	13.5	6.18E-6
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	22	8	14	1.73E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	22	8	15	2.20E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	22	8	15.7	5.80E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	22	8	16	2.90E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	22	8	18	3.48E-5
dissected lung	14-Nov	F	47	No	Smoker	J28	right lower lobe	22	8	18.5	0
dissected lung	14-Nov	F	47	No	Smoker	J28	right lower lobe	22	8	19	0
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	24	9	1	2.94E-4
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	24	9	2	9.80E-6
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	24	9	3	7.06E-5
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	24	9	5	1.18E-4
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	24	9	5.5	2.94E-5
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	24	9	6	8.82E-5
dissected lung	14-Nov	F	47	No	Smoker	J22	right lower lobe	24	9	7	3.92E-6
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	24	9	11	1.60E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	24	9	11.3	1.80E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	24	9	11.5	1.10E-6
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	24	9	12	7.00E-6
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	24	9	13	1.00E-4
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	24	9	13.5	3.50E-6
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	24	9	14	2.50E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	24	9	15	2.70E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	24	9	15.5	3.00E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	24	9	16	2.50E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	24	9	18	3.60E-5
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	24	9	18.5	6.00E-7
dissected lung	14-Nov	F	47	No	Smoker	J23	right lower lobe	24	9	19	1.60E-5
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	1	2	1	6.08E-04
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	1	2	2	2.00E-05

Branch or Dissected Lung?	Subject	Sex	Age	Twin?	Smoker?	Gel	Lobe	Sample	Bifurcation #	Peak Number	Mutant Fraction
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	1	2	3	7.92E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	1	2	5	3.00E-04
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	1	2	5.5	?
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	1	2	6	6.67E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	1	2	7	0
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	1	2	11	6.67E-5
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	1	2	11.3	2.44E-5
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	1	2	11.5	3.56E-6
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	1	2	12	2.22E-5
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	1	2	13	1.00E-4
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	1	2	13.5	1.78E-5
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	1	2	14	3.33E-5
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	1	2	15	2.00E-5
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	1	2	15.7	4.44E-5
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	1	2	16	3.40E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	1	2	18	1.00E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	1	2	18.5	0.00E+0
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	1	2	19	2.20E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	2	2	1	3.50E-04
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	2	2	2	3.75E-06
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	2	2	3	6.88E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	2	2	5	1.81E-04
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	2	2	5.5	0
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	2	2	6	6.25E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	2	2	7	6.25E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	2	2	11	6.93E-5
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	2	2	11.3	1.07E-5
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	2	2	11.5	2.86E-6
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	2	2	12	2.86E-6
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	2	2	13	1.00E-4
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	2	2	13.5	1.43E-5
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	2	2	15	5.71E-6
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	2	2	15.7	1.79E-5
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	2	2	16	1.79E-5
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	2	2	18	4.07E-5
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	2	2	18.5	3.57E-6
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	2	2	19	3.21E-5
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	3	3	1	3.83E-04
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	3	3	2	5.56E-06
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	3	3	3	1.11E-04
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	3	3	5	7.22E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	3	3	5.5	0
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	3	3	6	0
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	3	3	7	0
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	3	3	11	3.21E-5
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	3	3	11.3	1.07E-5
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	3	3	11.5	3.14E-6
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	3	3	12	2.86E-6
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	3	3	13	1.00E-4
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	3	3	13.5	1.43E-5
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	3	3	14	2.85E-5
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	3	3	15	4.64E-5
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	3	3	15.7	0.00E+0
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	3	3	16	2.86E-5
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	3	3	18	4.29E-6
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	3	3	18.5	0.00E+0
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	3	3	19	7.14E-6
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	5	4	1	9.30E-04

Bronch or Dissected Lung?	Subject	Sex	Age	Twin?	Smoker?	Gel	Lobe	Sample	Bifurcation #	Peak Number	Mutant Fraction
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	5	4	2	0
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	5	4	3	1.50E-04
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	5	4	5	1.10E-03
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	5	4	6	2.00E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	5	4	7	2.00E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	5	4	11	1.57E-5
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	5	4	11.3	1.71E-5
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	5	4	11.5	1.57E-6
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	5	4	12	1.21E-5
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	5	4	13	1.00E-04
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	5	4	13.5	0
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	5	4	14	4.29E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	5	4	15	1.43E-06
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	5	4	15.7	2.50E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	5	4	16	2.50E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	5	4	18	1.57E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	5	4	18.5	2.86E-06
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	5	4	19	1.07E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	8	6	1	8.65E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	8	6	2	2.70E-06
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	8	6	3	1.11E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	8	6	5	8.65E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	8	6	5.5	0
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	8	6	6	1.89E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	8	6	7	1.89E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	8	6	11	2.57E-06
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	8	6	11.3	6.29E-06
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	8	6	11.5	4.29E-07
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	8	6	12	0
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	8	6	13	1.00E-04
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	8	6	13.5	0
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	8	6	14	8.57E-07
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	8	6	15	7.14E-06
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	8	6	15.7	0
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	8	6	16	1.63E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	8	6	18	2.69E-06
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	8	6	18.5	5.71E-07
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	8	6	19	1.37E-06
dissected lung	20-Dec	M	58	No	Smoker	J26	right upper lobe	9	6	1	6.15E-04
dissected lung	20-Dec	M	58	No	Smoker	J26	right upper lobe	9	6	2	3.85E-06
dissected lung	20-Dec	M	58	No	Smoker	J26	right upper lobe	9	6	3	5.77E-05
dissected lung	20-Dec	M	58	No	Smoker	J26	right upper lobe	9	6	5	1.23E-04
dissected lung	20-Dec	M	58	No	Smoker	J26	right upper lobe	9	6	5.5	0
dissected lung	20-Dec	M	58	No	Smoker	J26	right upper lobe	9	6	6	5.38E-05
dissected lung	20-Dec	M	58	No	Smoker	J26	right upper lobe	9	6	7	0
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	9	6	11	5.97E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	9	6	11.3	2.70E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	9	6	11.5	0
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	9	6	12	5.19E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	9	6	13	1.00E-04
dissected lung	20-Dec	M	58	No	Smoker	J26	right upper lobe	9	6	13	1.00E-04
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	9	6	13.5	0
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	9	6	14	5.70E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	9	6	15	2.56E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	9	6	15.7	2.33E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	9	6	16	6.75E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	9	6	18	3.63E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	9	6	18.5	0

Bronch or Dissected Lung?	Subject	Sex	Age	Twin?	Smoker?	Gel	Lobe	Sample	Bifurcation #	Peak Number	Mutant Fraction
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	9	6	19	3.06E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	10	3	1	3.19E-04
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	10	3	2	2.13E-06
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	10	3	3	1.28E-04
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	10	3	5	6.17E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	10	3	5.5	?
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	10	3	6	2.77E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	10	3	7	2.77E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	10	3	11	2.62E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	10	3	11.3	3.93E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	10	3	11.5	0
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	10	3	12	8.20E-06
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	10	3	13	1.00E-04
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	10	3	13.5	8.20E-06
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	10	3	14	0
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	10	3	15	1.97E-05
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	10	3	15.5	1.64E-06
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	10	3	18	4.92E-06
dissected lung	20-Dec	M	58	No	Smoker	J25	right upper lobe	10	3	19	8.20E-06