Do Women Who Die of Early Breast Cancer Experience a Burst of Mutations During the Growth of the Mammary Gland in Puberty?

by

Paula Maria Lee

S.B. Civil and Environmental Engineering , Massachusetts Institute of Technology

(June, 1993)

Submitted to the Division of Toxicology

in Partial Fulfillment of the Requirements for the Degree of

Master of Science

in Toxicology

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1996

Author_____
Division of Toxicology
May 30, 1996

Certifiedby_____
Professor William G. Thilly
Thesis Supervisor

Acceptedby_____
Professor Peter Dedon
Chairman, Department Committee on Graduate Students

ACKNOWLEDGMENTS

I would first like to thank my advisor, Professor William Thilly for giving me the opportunity to work on this project and for his guidance throughout. His scientific rigor and enthusiasm for numerical analysis motivated me to always strive beyond the surface analysis.

I would also like to thank Gengxi Hu, who graciously taught me the molecular biology skills necessary for surviving in the lab and for instilling in me a healthy dose of scientific cynicism. I would especially like to thank Paulo André, aka (Paulinho) for all his support and patient late night and early morning discussions of the meaning of my "peaks". Special thanks also to Konstantin Khrapko for his discussions of my PCRs and John Hanekamp for his kind assistance in obtaining tissue samples from Dr. Sam Singer. I would also like to especially acknowledge Rita DeMeo for her endless help and Beth-Ann Turnquist for always bailing me out of Bill Gates' computer traps. I am also indebted to Pablo Herrero for his expertise and extensive help in converting the population data to actual graphs.

To my friends in the Thilly lab, both past and present: Andrea, Aoy, Beth-Ann, Brindha, Cindy, Enda, Gengxi, Henny, Hilary, Jackie, Jia, Klaudyne, Luisa, Paul, Pablo, Paulo, Rita (Cha), Rita (DeMeo), Wei-Ming, Wendy, Wen and Xiao-Cheng, thank you for sharing the ups and downs of daily life and for providing unending patience and support. I would specially like to acknowledge my cell culture friends, Jackie and Aoy. Jackie, thanks for keeping me out of trouble, feeding me well and always keeping me on my toes. Aoy, thanks for all the encouragement, coo-fee breaks and always keeping me focused on the big picture. To the lunch time crew, Amba, Suki and Al, thanks for always providing me with stimulating discussions and lots of good food.

This thesis is dedicated to my parents and to Ted, who have always provided stability and a sense of perspective. My parents have always supported my pursuits of knowledge and encouraged me throughout. To Ted - thank you for standing by me and always believing in me.

This work was supported by a training grant from the National Institute of Environmental Health Sciences (NIEHS).

Do Women Who Die of Early Breast Cancer Experience a Burst of Mutations During the Growth of the Mammary Gland in Puberty?

by

Paula Maria Lee

Submitted to the Division of Toxicology
on May 30, 1996, in partial fulfillment of the
requirements for the degree of
Master of Science

Abstract

This thesis outlines the theoretical basis of the hypothesis that women who die of early breast cancer experience a burst of mutations in any number of oncogenes or tumor suppressor genes during the development of the mammary gland in puberty. The basis of this hypothesis is derived from previous experiments in which the mammary epithelium of untreated sexually immature 50 day old female Fischer 344 rats contained clusters of c-H-ras codon 12 G->A mutants. The average size of these mutant clusters and their distribution within the glands indicated that they arose earlier in the pubertal development of the gland. We have examined the birth year cohort age-specific breast cancer mortality curves for both European American and African American females born in the years 1867, 1877, 1887, 1897, 1907, 1917, 1927 and 1937. We have postulated that the overall shape of the curve which may be separated into two distinct populations indicates a possible epidemiological basis of our hypothesis.

We have laid out the development of the mammary gland during each of its hormonal stages, i.e. puberty, pregnancy and menopause and have estimated within an order of magnitude the number of epithelial cells at risk during each of these stages. We have also modeled the epithelial cell kinetics during the development of the gland during

puberty and have outlined the possible scenarios for the appearance of spontaneous mutant cell clusters during this period.

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

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ACF	Aberrant crypt foci
APC	Adenomatous polyposis coli
BCNU	1,3-bis(2-chloroethyl)-1-nitrosourea
BSA	Bovine serum albumin
CDCE	Constant denaturing capillary electrophoresis
CNU	Chloroethylnitrosourea
DMBA	Dimethylbenz(a)anthracene
dNTP	deoxynucleotide triphosphate
EBV	Epstein-barr virus
ENU	N-ethyl-N-nitrosourea
FACS	Fluorescence activated cell sorting
FBS	Fetal bovine serum
FFTB	First full term birth
FFTP	First full term pregnancy
FTB	Full term birth
FTP	Full term pregnancy
gDNA	Genomic DNA
H&E	Hematoxylin and eosin
MAMA	Mismatch amplification mutation assay
MEGM	Mammary epithelium growth medium
MGMT	O6-methylguanine methyltransferase

MNNG	N-methyl-N'-nitro-N-nitrosoguanidine
mRNA	messenger RNA
NMU	N-nitroso-methylurea
PAGE	Polyacrylamide gel
SDS	Sodium dodecyl sulfate
TID	Terminal interlobular ducts

1. INTRODUCTION

1.1 The Mutational Basis of Cancer

The field of genetic toxicology has had at its foundation the premise that humans are exposed to exogenous chemicals which induce mutations in cellular genes, which subsequently lead to the development of cancer in a susceptible organ. Somatic mutations have been found in at least two classes of genes (proto-oncogenes and tumor suppressor genes) in both human and experimental animal tumors. The spectrum of mutations found in these tumors range from chromosomal rearrangements such as translocations or gene amplification to large scales insertions or deletions to point mutations. For example, the activation of the K-ras oncogene and the inactivation of at least 3 tumor suppressor genes, including the adenomatous polyposis coli (APC) gene have been found to occur in human colon tumors (Fearon et al., 1990). The normal cellular counterparts of oncogenes, termed "proto-oncogenes" function in controlling cell growth and differentiation, while normal tumor suppressor genes act as negative regulators of cell division.

Molecular characterization of spontaneous and chemical induced tumors generated in several rodent model systems revealed the presence of activated ras oncogenes, which were first discovered in avian and murine oncoviruses. Activated ras oncogenes have also been found in quite high percentages in colon carcinomas, pancreatic carcinomas, lung adenocarcinomas, cholangiocarcinomas, certain types of thyroid tumors, endometrial

adenocarcinomas, mucinous carcinomas of the ovary, squamous cell carcinomas and to a much lesser extent in several other types of human tumors (Anderson et al., 1992). It was difficult to define the precise role of oncogene activation in the initiation and development of human cancers. The availability of animal model systems, in particular, rodent model systems, in which oncogenes could be reproducibly activated by chemical treatment made the goal of defining the role of the oncogenes in the multistep process of carcinogenesis more realistic. The reproducible activation of the ras oncogene in chemical induced rodent tumors made it possible to correlate the activating mutations with the pre-mutagenic adducts which were supposedly formed directly by the carcinogen. In particular, rodent mammary models proved useful for studying the mechanisms of human mammary carcinogenesis.

1.2 Mammary Carcinogenesis

1.2.1Breast CancerMortality Data

The current estimate of every 1 in 8 American women developing breast cancer during the course of a lifetime has caused an increase in research activity in this field. We have obtained the breast cancer mortality data (Vital Statistics of the United States) and have separated it according to birth years and plotted it accordingly. The composite birth year cohort age-specific mortality breast cancer curves for European American females and non-European American females born in the years 1867, 1877, 1887, 1897, 1907, 1917, 1927 and 1937 are shown in Figures 1 and 2 respectively.

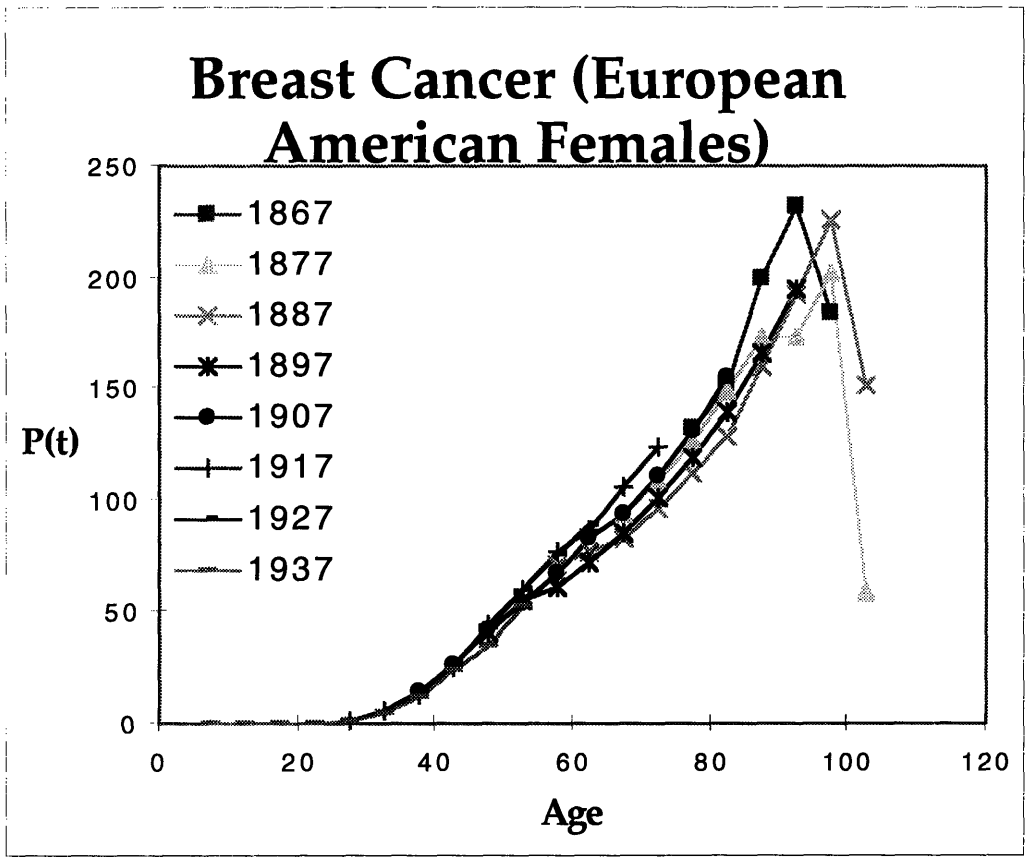


Figure 1 - Birth year cohort age-specific breast cancer mortality curve for European American Females

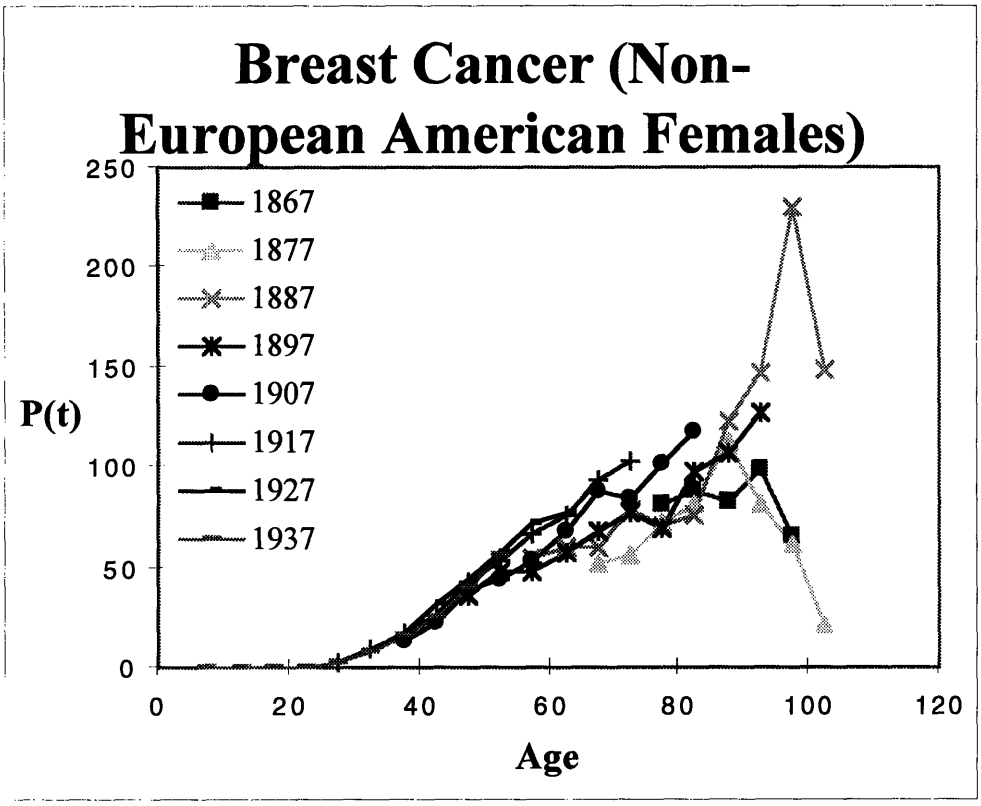


Figure 2 - Birth year cohort age-specific breast cancer mortality curve for Non-European American Females

Examination and analysis of the mortality curves reveals a reproducibly distinct inflection in the curves around the ages of 55-65. There is a peak in the curve somewhere around age 55, followed by a relatively flat 5 year period. After this point, the curves begin to rise again in an approximately linear fashion. This inflection suggests to us that there might be two separate populations of women who die of breast cancer. We may separate the overall curve into two separate populations by more or less symmetrically extrapolating the first part of the curve to yield "Population 1". The second population curve may be obtained by subtracting the extrapolated values from "Population 1" from the overall values of the combined curve. We may then postulate that the breast cancers in both populations occur by distinct mechanisms, the kinetics of which may be considered separately. It must however be noted that if the rates of dying of breast cancer for the women in "Population 2" begin to increase, it is possible that we may no longer observe the inflection in the overall mortality curve. It is interesting to note that at values of "t" corresponding to "Population 1", the slope of the curve of "P(t)" vs "t" does not increase with birth year. In contrast, the slope of the curve of "P(t)" vs ""t" increases with advancing birth year for "Population 2".

Two clearly distinct populations are also observed in the birth year cohort age-specific mortality curves for testicular cancer and Hodgkin's lymphoma. Figures 3, 4 and 5 depict the birth year cohort age-specific mortality curves for European American Female Hodgkin's lymphoma, European American Male Hodgkin's lymphoma and European American Male Testicular cancer respectively. Examination of Figures 3 and 4 show that the first population begins dying off around age 15, peaks at age 35 and then falls. The second population then begins to rise around age 50. In the case of testicular cancer, an examination of Figure 5 reveals a first population reminiscent of that in Hodgkin's lymphoma. Similarly, the curve of the first population begins to rise around age 15, peaks

between ages 30 to 35 and then begins to fall. The second population in testicular cancer however begins to rise around ages 60 to 65, in contrast to age 50 in Hodgkin's lymphoma.

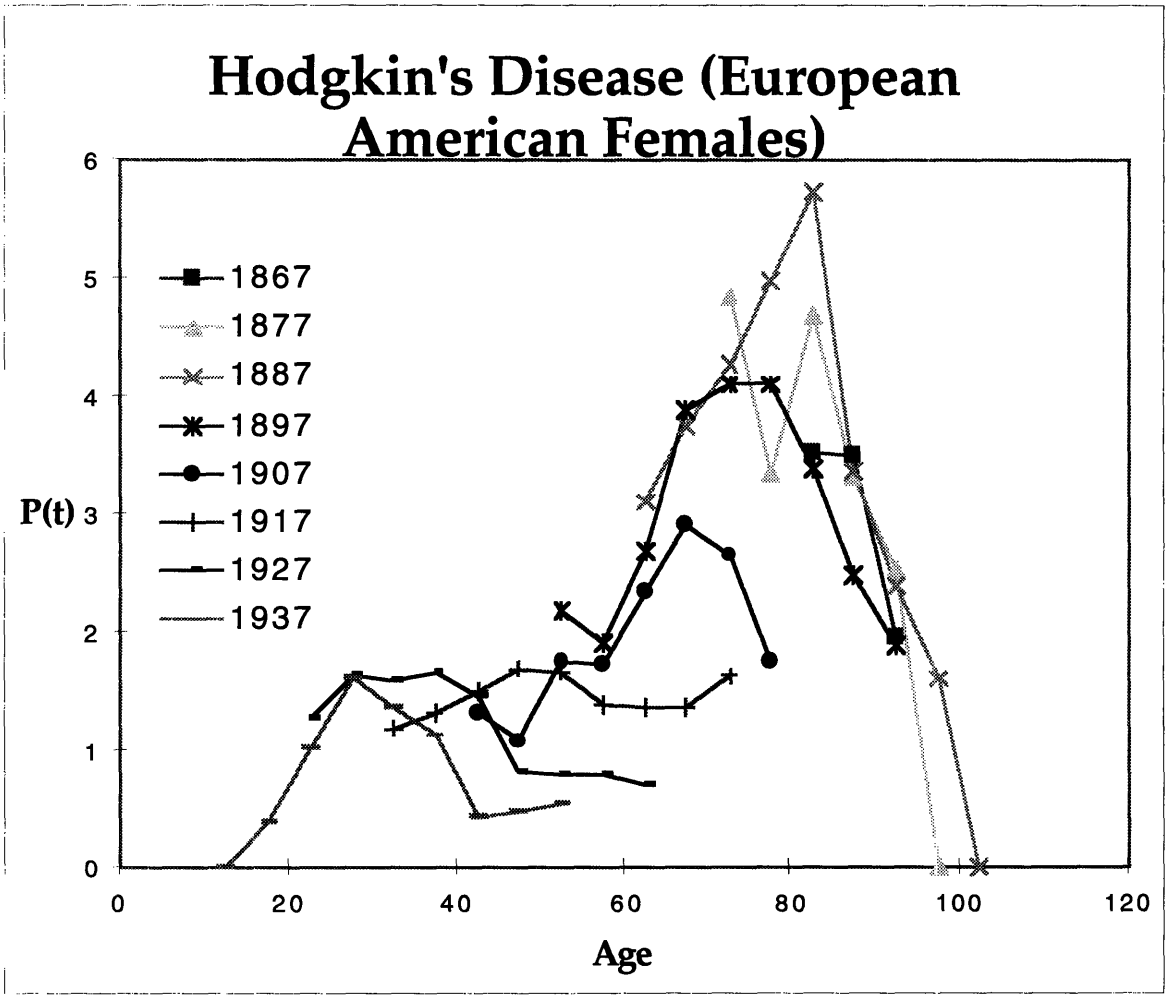


Figure 3 - Birth year cohort age-specific Hodgkin's lymphoma mortality curve for European American females

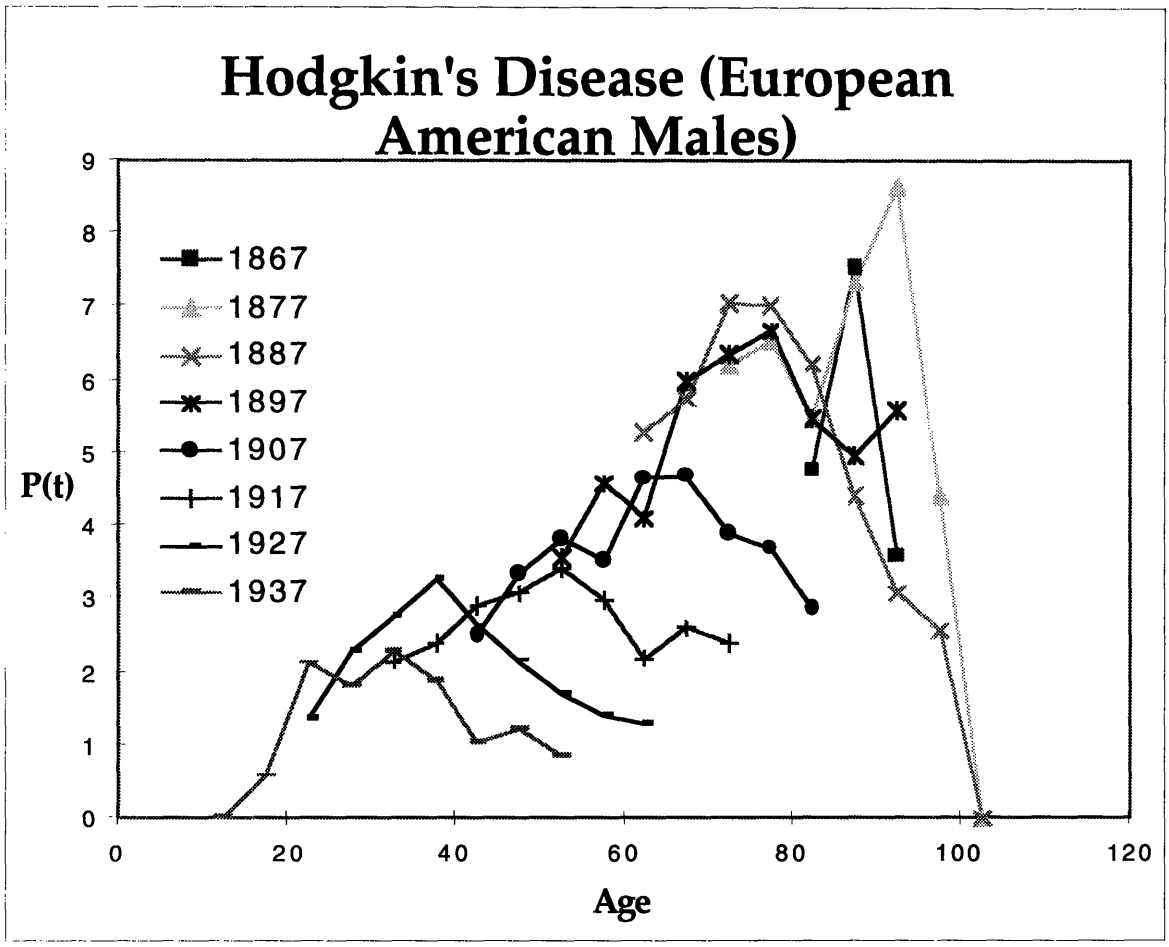


Figure 4 - Birth year cohort age-specific Hodgkin's lymphoma mortality curve for European American males

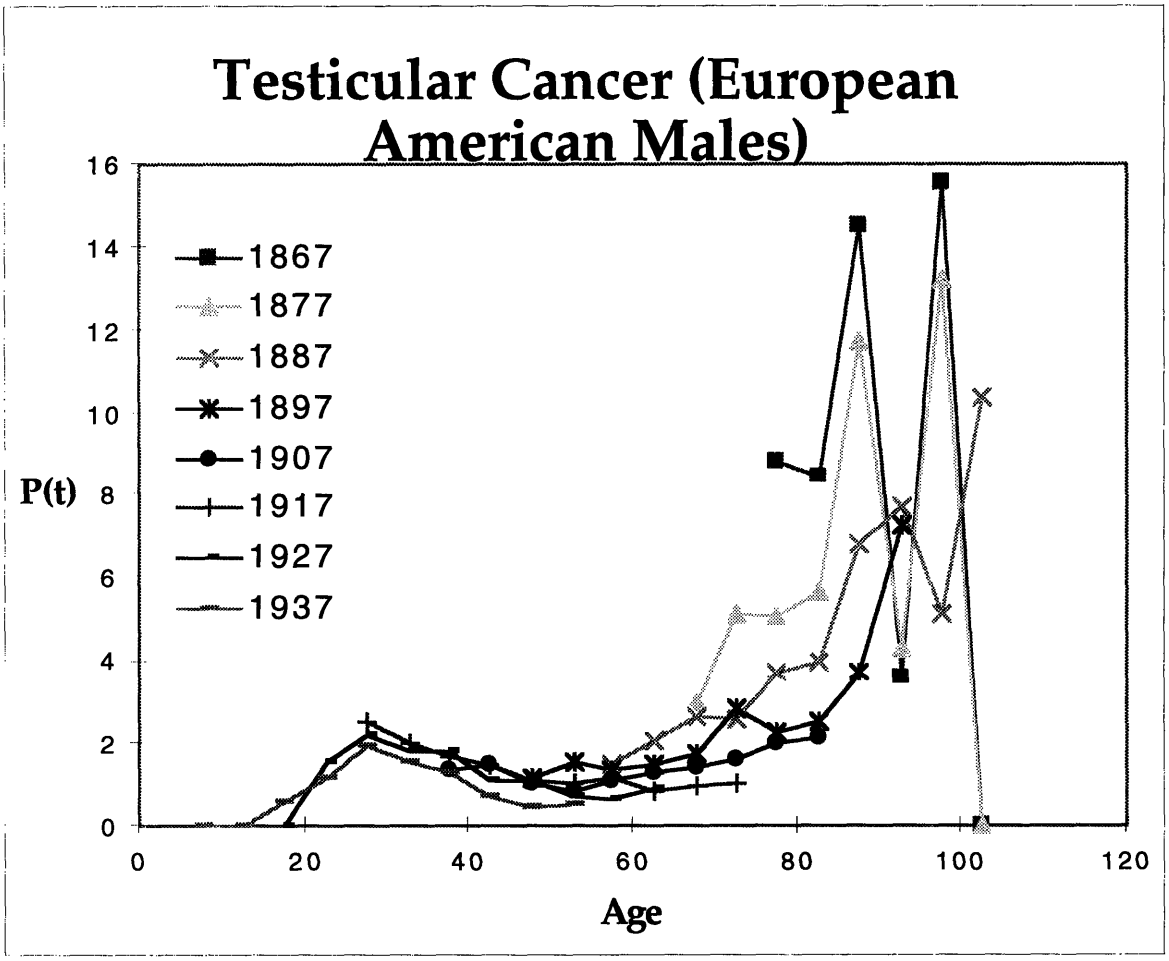


Figure 5 - Birth year cohort age-specific testicular cancer mortality curve for European American males

1.3 The Alternative Hypothesis of Rodent Mammary Carcinogenesis

The reproducible induction of mammary carcinomas after treating sexually immature female rats with the methylating agent, N-nitroso-methylurea (NMU) provided a useful model for the study of human mammary carcinogenesis. Specific G->A mutations were consistently found in codon 12 of the c-H-ras gene in these tumors. The collective hypothesis that arose from these experiments (see Appendix for detailed descriptions) is that chemical carcinogens act as initiators in the multi-stage theory of carcinogenesis by

directly inducing the oncogenic mutations which are later found in tumors. It was presumed that as tumors are monoclonal, the nature of the predominant genetic changes in the tumors is indicative of both the timing of the event and the chemical responsible for the changes.

For many years, the idea that NMU induced the specific mutation in the mammary epithelial cells which later underwent clonal expansion and developed into mammary tumors dominated the field of oncogenesis. However, there is an alternative hypothesis which was never seriously investigated. It is quite possible that the chemicals are not directly inducing the mutations which are observed in the resulting tumors, but are simply selecting the pre-existing mutant cells. It may be that the mutations arose prior to treatment with the chemical, perhaps during the window of time between the onset of puberty and treatment with the chemical. In the case of rats, treatment with the chemicals usually takes place at about 50-55 days of age and the onset of puberty occurs around 35 days of age (Dr. Rita Cha, Ph.D. Thesis, MIT 1992). Therefore, there is a window of approximately 15-20 days in which the mammary gland is undergoing development via branching of the ducts and proliferation of the mammary epithelial cells lining the ducts. In this stage of rapid cell proliferation, it is possible that spontaneous mutations, in this case in the c-H-ras gene may occur. Treatment with NMU may have some independent effect on the c-H-ras mutant cells, possible selecting them and promoting their proliferation into mammary tumors.

Despite the fact that the resulting mammary tumors may contain predominantly codon 12 mutations in the c-H-ras gene, this may not mean that these are the predominant mutations induced by NMU, (assuming that NMU induced the mutations). A tumor is a clonal population which was selected in vivo and for this reason, we may not assume that the predominant mutation is the one induced by the chemical. The only information that

may be garnered is that the cells with the predominant mutation have a selective growth advantage and that these mutant cells were present earlier in the development of the mammary tumor. It is possible that the chemical induces other mutations which do not have a selective growth advantage and may not be detected because the mutant frequency is below the level of detection of the specific assay.

1.4 Testing the Alternative Hypothesis

Cha et al., (1994) were the first to directly test this alternative hypothesis, by trying to prove that the oncogenic mutations which are detected in chemical-induced rodent tumors are directly induced by the chemical carcinogens themselves. They set out to determine if a carcinogenic dose (30 mg/kg) of NMU administered to female Fischer 344 rats induced mutations in the the c-H-ras oncogene of mammary epithelial cells or if the mutations were pre-existing.

Using the mismatch amplification mutation assay (MAMA) which was key to the testing of the hypothesis and had a demonstrated sensitivity of 10^{-5} (Cha et al., 1992), they determined that the G->A transition mutation at codon 12 of the c-H-ras gene existed in the mammary epithelium of these rats prior to treatment with NMU (Cha et al., 1994). They hypothesized that NMU acted in a manner independent of inducing the codon 12 mutations and was perhaps selecting for the mammary epithelial cells bearing the mutations, inducing them to expand and develop into a mammary tumor. The evidence for this conclusion is based on the fact that (i) 70% of untreated animals contained detectable levels of c-H-ras mutants (an average of about 128 mutants per animal), (ii) the mutants were clustered as sectors within the mammary gland and (iii) that treatment with the carcinogenic dose of NMU did not result in a significant increase in the number of mutants, the fraction of organ sectors with mutant cells or the fraction of animals containing

detectable levels of c-H-ras mutants. Table 1 shows the distribution of c-H-ras mutant cells in sectored mammary tissue of 50 day old untreated Fischer 344 rats (taken from Cha et al., 1994). By determining the average number of c-H-ras mutants per sector and their distribution in the rat mammary gland, Dr. Cha and her colleagues were able to conclude that there was a "burst" of mutations and to pinpoint the "timing" of the burst close to the first estrus[1]. The clustering of the c-H-ras mutants as sectors in the gland is consistent with their origin as descendants of single cells which mutated early in the development of the gland. Otherwise, if the mutants arose independently or at a time closer to the attainment of maximum cell number of the organ, they would be randomly and evenly distributed throughout the organ. This is based on the logic that the induced mutants would not have had sufficient time/doublings to clonally expand and develop into colonies. It must be remembered however, that the limit of sensitivity of the MAMA assay does not allow us to distinguish between truly independent mutants and colonies which have not yet attained sufficient mutant cells (mutant fraction required >= 10^{-5}), which would allow them to be scored as positive sectors.

---

[1] Dr. Cha determined the timing of the burst using the following method (Dr. R. Cha, doctoral thesis, 1992).

Sectors that she analyzed from 50 day old rats contained on average 128 (2^7) mutant colonies. Assuming that each sector arose from a single cell, she determined that the single mutation in these sectors arose 8 cell divisions prior to the time of observation. Using the observation that the cell division cycle in young virgin rats varies between 10 and 28 hours (Russo and Russo, 1988), she was able to calculate that the early event occurred around 40 days of age.

Table1. Distribution of H-ras mutant cells in sectored mammary tissue of 50 day old untreated Fischer 344 rats (from Cha et al., 1994)

ANIMAL	SECTOR	MUTANT/SECTOR	MUTANT FRACTION
50-1	**D**	150	2.1×10^{-5}
50-2	**E**	100	1.4×10^{-5}
50-4	**A**	110	1.6×10^{-5}
	D	50	0.7×10^{-5}
	E	80	1.2×10^{-5}
50-5	**B**	200	2.8×10^{-5}
	C	150	2.1×10^{-5}
50-6	**C**	120	1.7×10^{-5}
Average # of mutants		32[2]	

[2] This includes both positive and negative mammary tissue sectors

2. 1 Hypothesis

The results of Cha et al. (1994) have altered our perspective on the mechanism of rat mammary carcinogenesis. The mutations which are found in the rat mammary tumors, if indeed spontaneous in origin, will force us to consider the contributions to the overall role of carcinogenesis of endogenous agents in human mammary carcinogenesis. Their observations prompted us to think that perhaps a similar phenomenon (a burst of spontaneous mutations during puberty) might be occurring in humans. We then decided to determine if there was any theoretical evidence that would support our notion of a pubertal burst occurring during the growth of the mammary gland in puberty. Using the shape of mortality curves for a number of different cancers as a reference, we may predict the shapes of the breast cancer mortality curves, if indeed a pubertal burst did occur. As shown in Figure 6, there would be an early population which would begin dying off beginning somewhere around age 25. The curve would then peak and begin to fall until age 40-45. The predominant pattern which is found in tumors in which there is no expected pubertal burst would be an approximately linear increase in the curve beginning around age 55-60, which is also shown in Figure 6. The composite of both of these curves is shown in Figure 7, and is remarkably reminiscent of the breast cancer mortality curves (birth year cohort specific data for the years 1867-1937) shown in Figures 1 and 2.

HYPOTHESIS

The basis of this thesis was to determine if women who die of early breast cancer experience a burst of mutations during the growth of the mammary gland during puberty in any number of oncogenes or tumor suppressor genes.

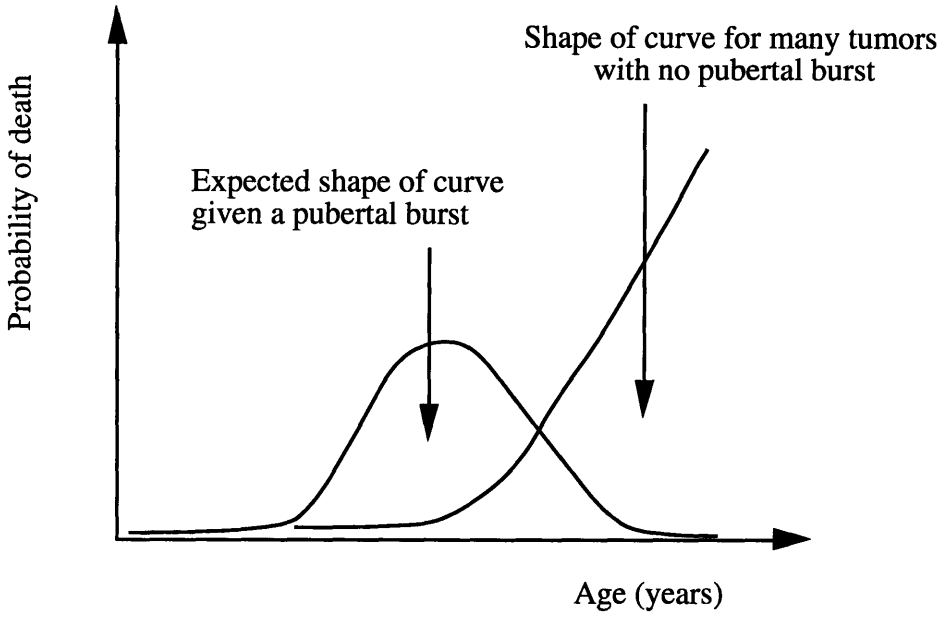


Figure 6 - Theoretical shape of mortality curves based upon pubertal burst hypothesis

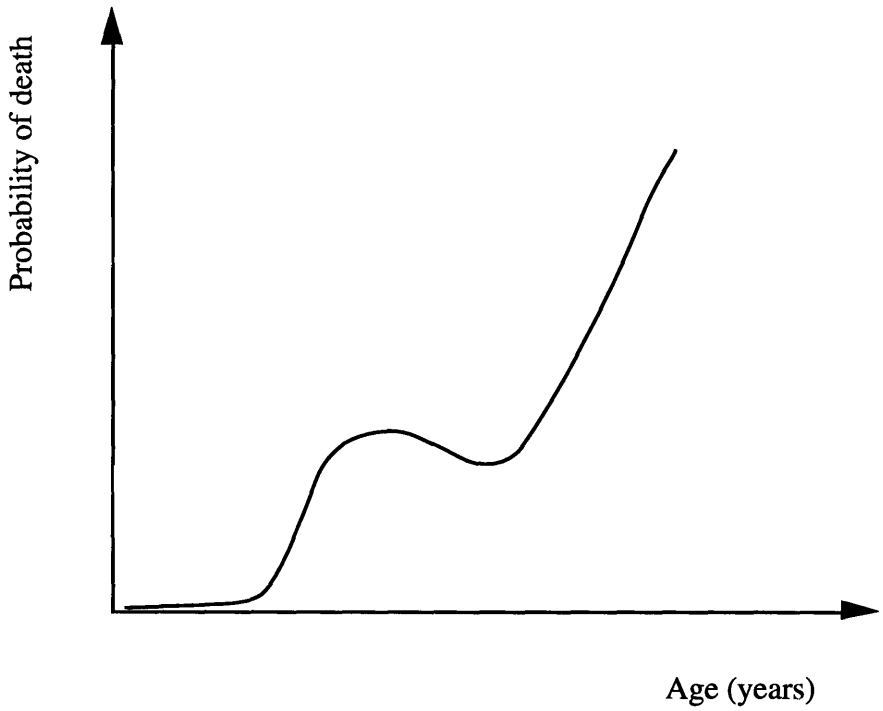


Figure7 - Composite theoretical curve of both pubertal burst and non-pubertal burst

2.2 The Human Mammary Gland

2.2.1 General Structure of Mammary Gland

The structure of the mammary gland varies with the hormonal status of the individual. The human mammary gland may be divided into three structural compartments:

- the glandular tissue, i.e. epithelial cells of the compound tubulo-alveolar type, which is further divided into lobes
- connective tissue which surrounds the glandular tissue and separates lobes from each other
- interlobular adipose tissue.

The glandular component is composed of a tree-like network of branching ducts and terminal secretory lobules, which becomes increasingly ramified as it extends away from the nipple in both the radial and longitudinal directions. The gland consists of 15-25 lobes, each of which converges to a lactiferous duct which then opens up into the lactiferous sinus at the apex of the nipple. The lactiferous ducts lead into the terminal interlobular ducts (TID), each of which further divides into several intralobular ducts. Gray's Anatomy (38th edition) cautions that even though the lobes of the mammary gland are depicted as discrete anatomical regions in the gland, they typically grow into each other at the edges and do not look like discrete structural entities during surgery. Lobules are the terminal secretory portion of the branching structure and their structures are dependent on the hormonal status of the gland. In the mature breast (specialized for milk production and secretion), each lobule consists of several blind ending extensions, termed as alveolar buds, ductules or alveoli (acini) depending on their stage of development. Each of these branches converge and form an alveolar duct through which the milk is secreted.

The connective tissue stroma of the mammary gland loosely encapsulates and surrounds the lobules, thus accomodating easy, rapid expansion of the glandular portion of the organ both during puberty and pregnancy. Adipose tissue occurs between the lobules in a lobe and the quantity is also dependent both on the individual and the hormonal status of the individual.

The ducts are lined by columnar epithelium for most of their length. Larger ducts are typically lined by two layers of cells, while the smaller ones are lined by a single layer of epithelium (luminal epithelium). The bases of the luminal epithelial cells are in close contact with numerous myoepithelial (basal) cells which invaginate their bases and form a separate layer surrounding the ducts and alveoli. The lactiferous ducts are lined with stratified cuboidal epithelium and close to their openings at the apex of the nipple, the stratified cuboidal epithelium is replaced by keratinized stratified squamous epithelium. Figure 8 depicts the ultrastructure of the resting mammary gland (Krstic, 1991).

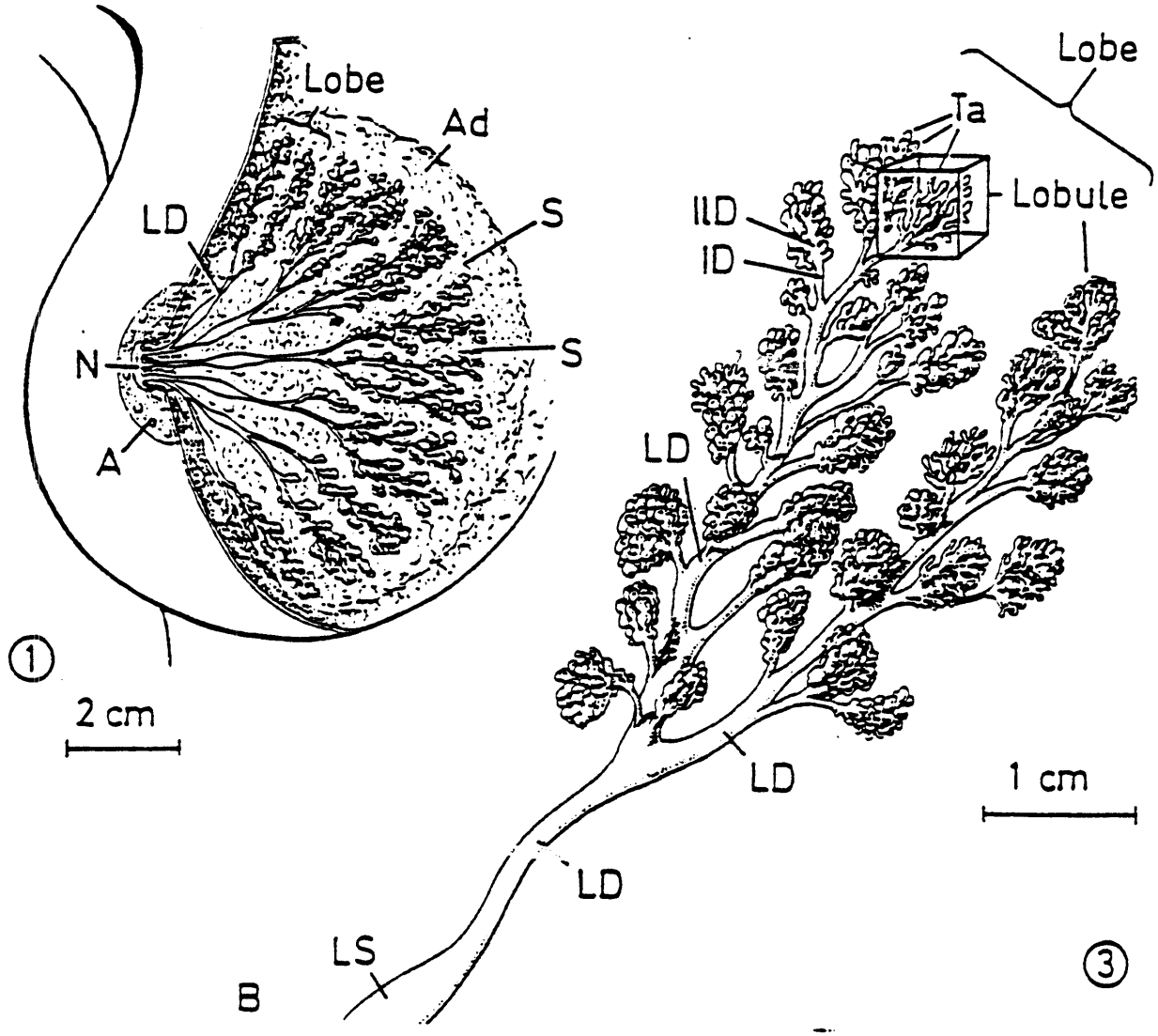


Figure 8 - Ultrastructure of the resting human mammary gland (Krstic, 1991)

2.2.2Pre-Natal Development of the Human Mammary Gland

The first evidence of the mammary gland appears as the milk streak during the 4th week of gestation, when the fetus is approximately 2.5 mm long. The milk streak then becomes the mammary ridge or the milk line during the 5th week. According to the qualitative ultrastructural observations of Dabelow (1957) described by Russo and Russo, the process of gland development in the fetus may be characterized into 9 different stages. The first stage is the thickening of the mammary ridge stage which is termed the "milk hill" stage in the 6th week, followed by the "mammary disc" stage when parenchymal cells start to invade the underlying stroma. The "mammary disc" stage then progresses to the globular stage between the 7th and 8th week. Subsequent to this is the cone stage which occurs at the 9th week and is characterized by further inward growth of the mammary parenchyma. Epithelial buds sprout from the invading parenchyma and these buds become lobular in shape between the 12th and 13th week, with notching at the epithelial-stromal border, otherwise known as "indentation". The next stage is called the "branching" stage and is marked by the appearance of 15-25 epithelial strips or solid cords in the 15 week old fetus. The canalization stage occurs next between 20 and 32 weeks of gestation and the solid cords become "canalized" by desquamation and lysis of the central epithelial cells. The final stage is the end-vesicle stage in which the end vesicles contain colostrum and have a monolayer of epithelial cells. Lobules have thus far never been observed in the fetal mammary gland. Their formation has been associated only with the onset of puberty. The male and female mammary glands are structurally similar at birth, with the exception that the male mammary gland never undergoes hormone-mediated development.

2.2.3 Development of the Human Mammary Gland During Puberty

Development of the mammary gland is initiated during puberty and varies from woman to woman. Mammary gland development may be characterized by determining the mammary gland area, volume, degree of branching of the ductal network, the degree of structures, such as lobules whose appearance shows the level of differentiation of the gland, or a combination of each of these variables. At the onset of puberty, somewhere between ages 11 and 13, an increase in the secretion of estrogen and progesterone (see Figure 9 for details on the age correlated systemic levels of the hormones) drives the proliferation and branching of the ductal network and an accompanying increase in the density of the surrounding stroma, characterized by an increase in deposition of collagenous connective tissue and adipose tissue. This accounts primarily for the increase in size of the mammary gland during puberty. The ducts grow and divide in both a dichotomous (repeated bifurcation) manner and a sympodial manner (formation of an apparent main axis form successive secondary axes) on a dichotomous basis. The ducts lengthen, divide and form club-shaped terminal end buds, which later become small ductules or alveolar buds. The alveolar buds then become acini or alveoli. The ductules/alveolar buds are morphologically more developed than the terminal end buds but less developed than the acini (alveoli). According to Russo and Russo, the alveolar buds cluster around a terminal duct forming a type 1 lobule. These terminal ducts or alveolar buds are lined by two layers of epithelial cells, whereas the terminal end buds are lined by up to four layers of epithelium. The acini are the secretory unit of the mammary gland. Lobule formation occurs within 1-2 years after the first menarche.

2.2.4 The Mammary Gland During Pregnancy

The breast achieves its maximal state of development during pregnancy. There are two distinct stages; the first is characterized by the proliferation of the distal portions of the ductal network (otherwise referred to as lobule formation), resulting in the transformation of ductules or alveolar buds to acini (alveoli). Rapid proliferation and morphological development of the functional secretory units occurs during the first three months of pregnancy. The epithelial cells in the acini of the newly formed lobules increase greatly in number as they are actively undergoing cell division and they also increase in size due to cytoplasmic enlargement. It was not possible to garner any information in the literature about whether there is a distinct "stem cell" population which is responsible for most of the cell division and increase in epithelial cell population or if all the cells in the ductal network are capable of undergoing cell division. I would postulate that it would be more efficient for the gland to have a reserved "stem cell" population which would remain at the distal portions of the ductal network. As the network grows radially and orthogonally away from the apex of the nipple, the cells in the distal portion of the ducts would automatically give rise to the components of the new lobules. We must also note that as the glandular portion of the breast increases, the stromal component actually decreases to accomodate the expanding glandular compartment.

During the middle of pregnancy, the lobules are further enlarged and increased in number. The lobules at this stage still contain a combination of terminal ducts and acini. Certainly, by the end of the first half of pregnancy, the groundwork for the structure of the expanded ductal network is already established. The second half of pregnancy is characterized mostly by differentiation into the milk secreting acini, which are the fully differentiated structures. These acini are not susceptible to neoplastic transformation and hence would theoretically reduce the number of cells at risk. This structure is maintained

throughout lactation. However, the post-lactational stages of the mammary gland are marked by a regression of the alveolar/ductal system, an increase in apoptosis of the epithelial cell population and a decrease in their cell size.

2.2.5 The Mammary Gland During Menopause

The decrease in the secretion of ovarian hormones, beginning around age 45 causes a regression in the glandular structure of the mammary gland. The main ducts and their branches remain, while there is shrinkage and collapse of the lobules. There is an increase in the fat and connective tissue content of the gland.

2.3 The Hormonal Status of the Pre-Pubertal and Pubertal Mammary Gland

Female puberty is characterized by a number of endocrine changes. Some of these changes are actually initiated several years before the first menarche (average age of 12 [Gray's Anatomy, 38th Edition]). The hormonal status of the pre-pubertal and the pubertal mammary gland is an important factor in our hypothesis. The link between the physiology and the endocrinology of the gland during the different stages of its development and the molecular changes underlying the hypothesized mutational burst is unknown. Such vital information is unfortunately unavailable in the existing literature. Vihko and Apter gathered data on some 200 health girls aged 7 -18 years and were able to come up with average systemic concentrations of several hormones, including estradiol and progesterone. We focused on the concentrations of estraogen/estradiol and progesterone as these are the strategic hormones driving the maturation and development of the mammary gland during puberty. Figure 9 represent the graphs of the concentrations of estradiol and

progesterone respectively which have been adapted from the data presented by Vihko and Apter.

Examination of Figure 9 shows that the concentration of estradiol increases two fold between the ages of 7 and 9 (0.025 nM - 0.05 nM). Over the course of the following four years (ages 9-13), the concentration of estradiol increases five fold, leveling off at 0.2375 nM for about two years until age 15. An additional slight increase occurs again from age 15 until age 17 and finally flattens at 0.275 nM. The overall concentration of estradiol increases 11 fold in the period of approximately 11 years. The concentration of progesterone also changes significantly throughout the pre-pubertal and pubertal years. The increase in concentration is however more gradual than that of estradiol and increases overall by approximately 2.3 fold (0.3 nM - 0.7 nM) over the span of 11 years. The progesterone concentration increases 1.3 fold during the first 4 years (ages 7-11), 1.25 fold between 11 and 12, holding steady until age 14 and finally increasing 1.4 fold between ages 14 and 18. Data on the post-pubertal concentrations of these key enzymes have not been found in the literature.

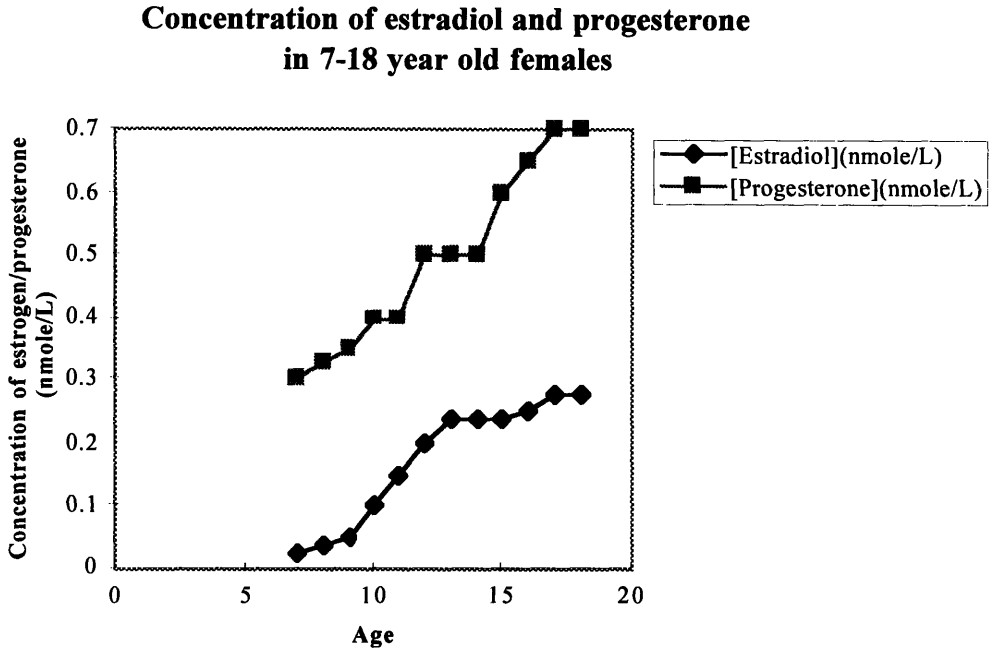


Figure 9 - Concentration of Estradiol and progesterone in 7-18 year old females

2.4 Estimated Number of cells as a Function of the Hormonal Status

In an effort to obtain an **order of magnitude estimate** of the number of epithelial cells at risk of transformation in the human mammary gland as a function of the hormonal status of the gland, we have gathered data from several sources and made several assumptions. Russo and Russo determined the level of differentiation of lobules in the breast tissue of 22 adult women, ranging in age from puberty to 63 years of age who were undergoing reduction mammoplasties. Morphologically, lobules can be characterized as type 1, type 2, type 3 and type 4. They found that lobules of type 1 were found mainly in pubertal and young nulliparous pubertal women, while type 2 and type 3 lobules were found in parous, non-pregnant women and type 4 lobules were found primarily in

pregnant women. They have also measured parameters such as the number of cells per alveolar bud/ductule/acinus depending on the lobule type and the number of alveolar buds/acini/ductules per lobule. A summary of this data is shown in Table 2. The hierarchy of structures in the mammary gland is as follows: epithelial cell -> ductule/alveolar bud/acinus ->lobule -> lobe. Since the hormonal status of the gland has been determined as a function of the type of lobules found in breast tissue, we are assuming that the parameters for the different lobule types may be used to represent the different hormonal stages in the mammary gland. Therefore, parameters for lobules type 1 may be used to estimate the number of cells at risk for young post-pubertal women, lobules type 2 and 3 for parous post-pubertal non-pregnant and nulliparous young women and lobules type 4 for pregnant women.

	Lobule Type 1	Lobule Type 2	Lobule Type 3	Lobule Type 4
Number /lobule[3]	11.20	47.0	81.0	180.0
# cells/cross-section[4]	32.43	13.14	11.0	10.0
Cross-sectional area/lobule (mm2)	0.048	0.060	0.129	0.250

Table 2 Number of epithelial cells per cross section of alveolar bud/ductule/acinus, number of alveolar buds/ductules/acini per lobule and cross-sectional area per lobule (mm^2)

[3] # of alveolar buds/ductules or acini per lobule
[4] # of epithelial cells per cross-section of alveolar bud/ductule or acini

The number of epithelial cells at risk in each lobe may be determined by multiplying the number of cells per cross-section of either alveolar bud/ductule/acinus by the number of lobules in each lobe. We may then obtain the number of epithelial cells per breast as we know that there are between 15-25 lobes per breast depending on the individual. For the purposes of our calculations, we have chosen to use an average of 20 lobes per breast. Since a measurement of the number of lobules per lobe is not available, we have estimated the value of this parameter for lobules types 1 through 4 using the following assumptions. We know that the mass of each breast in post-pubertal women is between 150-200 g and increases to about 500 g during pregnancy (Gray's Anatomy, 38th edition). Based on this, we have estimated the mass of each breast in late puberty females to be about 100 g. Therefore, using the density of body fat (0.78 mg/mm^3) and the relative parenchyma-stroma ratio measurements made using whole mount studies of 14 human breasts of pubertal, post-pubertal and pregnant women by Russo and Russo, we are able to determine the total volume of the mammary gland occupied by the lobules. Russo and Russo have found that 10% of the mammary gland in late pubertal females is composed of parenchyma, and this value increases to 30% in post-pubertal and young nulliparous women. During pregnancy, this ratio changes significantly and during the first half of pregnancy, 55% of the gland is composed of parenchyma while it increases to 73% during the second half of pregnancy. The volume of each lobule may be estimated as we have the cross-sectional area of lobules types 1-4, and we are assuming that the lobule may be modeled as a sphere. Therefore, the number of lobules per lobe may be obtained by dividing the total volume of each breast occupied by lobules by the volume of each lobule. The number of cells per lobe may then be obtained by multiplying the number of lobules/lobe by the number of cells/lobule (Table 2). Multiplying this parameter by 20 (average number of lobes/breast) gives an estimate of the number of epithelial cells at risk in each breast. Table 3 shows the values of all the estimates of the parameters which have been obtained using the above assumptions. It must also be noted that the number of

epithelial cells at risk during the second half of pregnancy is overestimated as a large number of the epithelial cells in the first half of pregnancy have become fully differentiated milk producing acini and are no longer susceptible to neoplastic transformation.

	Type 1	Type 2	Type 3	Type 4 - 1st half pregnancy	Type 4 - 2nd half of pregnancy
Total volume of breast occupied by lobules	7.8×10^3 mm^3	$3.51\text{-}4.68 \times 10^4$ mm^3	$3.51\text{-}4.68 \times 10^4$ mm^3	2.145×10^5 mm^3	2.85×10^5 mm^3
radius of lobule	0.124 mm	0.138 mm	0.203 mm	0.282 mm	0.282 mm
Volume/lobule	7.9×10^{-3} mm^3	1.1×10^{-2} mm^3	3.5×10^{-2} mm^3	9.39×10^{-2} mm^3	9.39×10^{-2} mm^3
#lobules/lobe	9.8×10^5	$3.19\text{-}4.25 \times 10^6$	$1.0\text{-}1.33 \times 10^6$	2.28×10^6	3.03×10^6
#cells/lobe	1.1×10^7	$1.49\text{-}1.99 \times 10^8$	$8.1 \times 10^7\text{-}1.08 \times 10^8$	4.1×10^8	5.45×10^8
#cells/breast	2.2×10^8	$2.98\text{-}3.98 \times 10^9$	$1.62\text{-}2.16 \times 10^9$	8.2×10^9	1.09×10^{10}

Table 3 - Values of parameters obtained necessary for the order of magnitude estimate of the number of epithelial cells at risk in the breast

2.5 Potential Relationship Between Breast Cancer Risk Factors and Tissue Kinetics

Early age at menarche (age 11 and before), late menopause (> 55 years of age), and age at first full term pregnancy (FFTP) have been identified as "risk factors" for the development of breast cancer [Moolgavkar et al., (1980), Pike et al., (1981), MacMahon et al., (1970), Trichopoulos et al., (1972)]. We can try to relate these risk factors with the events which are taking place at the cellular level as the organ proceeds through its

hormone-mediated stages of development. Pike et al., (1981) found that the relative risk for developing breast cancer in women aged 32 or younger is twice as high for those who experienced menarche at 11 versus those who experienced menarche at 13 or older. MacMahon et al., (1970) showed that nulliparous women in an age-controlled study had a relative risk that was twice the value for women who had a FFTP before the age of 20. If the FFTP was between the ages of 20 to 24, 25 to 29 or 30-35, the relative risk increased by 20%, 56% and 88% respectively. It appears that the key factor is the age at FFTP as the relative risk does not decrease significantly with subsequent full term births (FTB), controlling for age at FFTB. Trichopoulos et al., (1972) have also estimated that women whose natural menopause occurred before age 45 had about half the breast cancer risk as the women whose natural menopause occurred after age 55. Based on our hypothesis, these kinetics probably affect "Population 2".

For the sake of speculating on the possible tissue kinetics which might be underlying the epidemiological observations, we shall think of the mammary gland as a homogeneous population of cells which are susceptible to transformation. As mentioned previously, the cells in the mammary gland which are susceptible to the development of carcinomas are the ductal epithelial cells. We are making the simple assumption that breast cancer risk is determined solely by the number of susceptible epithelial cells and topographic factors such as the positions of these epithelial cells in the architecture of the gland do not have any significant influence on the overall risk. We are also assuming that the two main determinants for decreasing the size of this population are death and differentiation.

The status of the population of epithelial cells can be characterized at each of the hormonally distinct stages in a woman's life - puberty, pregnancy and menopause. Epithelial cell proliferation is stimulated at the onset of puberty by increased levels of

estrogen. The mammary gland is continuously in a state of development and the epithelial cells undergo a limited amount of hormonally-induced proliferation during each menstrual cycle. During the first half of pregnancy, the maturation of the gland continues and there is much ductal branching and epithelial cell division. However, during the second half of pregnancy, much of the epithelial cell population undergoes differentiation and becomes specialized for the task of milk production. Regression in the levels of the hormones which maintain the breast epithelium at the onset of menopause causes the involution of the breast. The size of the population of susceptible cells is significantly reduced as a result of either an increase in the death rate of the cell population, an increase in the rate of differentiation or a decrease in the growth rate or a combination of all three factors (Moolgavkar et al., 1980).

An early age at menarche induces early proliferation of the breast epithelium and consequently leads to an increase in the number of susceptible epithelial cells earlier in life. This increases the length of time for which they can acquire the rate limiting mutations in order to acquire the selective growth phenotype. Similarly, a late menopause also increases the length of time the epithelium is susceptible to incurring the rate limiting mutations. At the onset of menopause (45-55), one would expect a population of cells which have already acquired one of the two rate limiting mutations would exist. This would imply that the longer this population of cells stays around, the more likely it would be that they would acquire the second rate limiting mutation, thus setting the stages for the development of a clinically detectable tumor. The protective effect of a full term pregnancy probably arises from the fact that in the second half of pregnancy, the significant fraction of the cells which arose from a rapid rate of cell turnover during the first half of the pregnancy are removed by differentiation. This translates into a decrease in the value of N and hence, a decrease in the value of the slope. Moolgavkar et al., (1980) also caution us that a FTP not only decreases the population of cells which have not

acquired any of the rate limiting mutations, but it also decreases the population of cells which have acquired one of the two mutations. Phenotypically, these two populations should be identical and there should be no selection for one population over the other in the process of differentiation. Therefore, the earlier the FFTP occurs, the faster the number of susceptible cells is decreased and the smaller the population of cells available for acquiring the rate limiting mutations. This smaller number of cells effectively decreases the probability of mutation induction (assuming that N is decreased by at least one order of magnitude).

2.6 Modeling the Cell Kinetics During Development of the Mammary Gland

The kinetics of epithelial cell growth in the developing human breast is crucial to the prediction of the expected number of mutant clusters and the average size of the clusters. Let us consider a developing mammary gland and assume continuous growth in time. We assume that the human mammary gland begins with 10^5 epithelial cells at the onset of puberty at age 11, and most of the cell growth is complete by age 16, at which time the mature breast contains approximately 7.5×10^8 cells[5]. Therefore, the mammary gland undergoes a 10^4 fold expansion in the space of about 5 years.

During the critical growth phase, the mammary gland will exhibit growth such that the rate of growth of cells is proportional to the number of cells in the organ. In other words, the growth of the gland is exponential and may be described by equation 2 and Figure 11. In equation 2, k is the proportionality constant and describes the net growth of

[5] 4 g of reduction mammoplasty tissue yields 1.5×10^7 epithelial cells on collagenase digestion (PML). The average mass of the mammary gland in a non-pregnant post-pubertal female is about 150-200g (Gray's Anatomy, 38th edition). Therefore, the estimated number of epithelial cells in an adult non-pregnant gland is about 7.5×10^8, which is about 40% lower than the theoretical estimate derived in Table 3.

epithelial cells during a fixed time interval and has units of inverse time ($time^{-1}$). Separation of the variables in the linear ordinary differential equation expressed in equation 1 yields equation 2, which is the explicit solution.

$$\frac{\partial N}{\partial t} = kN \qquad (1)$$

$$N = N_0 e^{kt} \qquad (2)$$

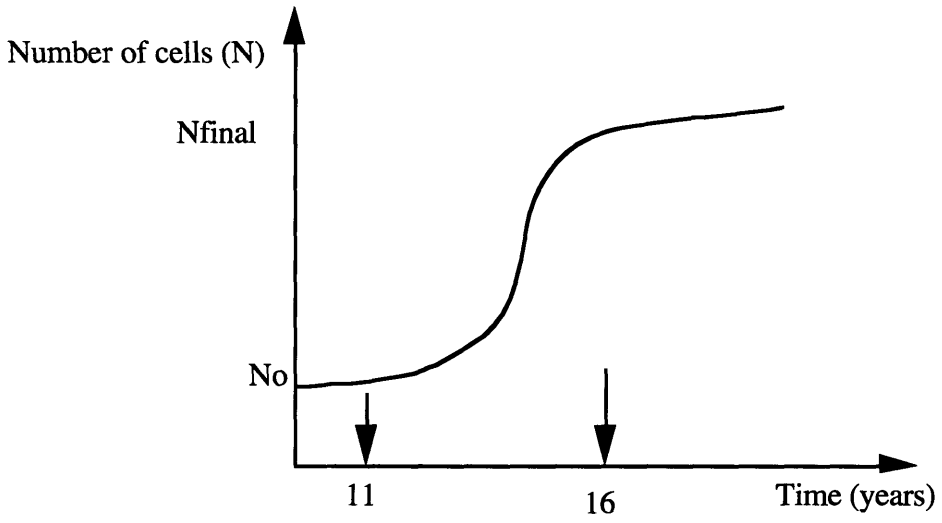


Figure 11 - Graph depicting exponential growth of epithelial cells in human mammary gland during puberty

Based on Equation 1, we may obtain values for k by making an exponential fit to the two data points. We may vary

(1) N_0 in the range of 10^3-10^6 holding N_{final} constant at $2x10^9$ or

(2) N_{final} in the range of 10^8-10^{11} holding N_0 constant at 10^5 or

(3) vary both N_0 and N_{final}.

After obtaining a reasonable value for k, we may then use equation 2 to obtain an estimate of the doubling time of the epithelial cells. Tables 3 and 4 show the various values of k and doubling times (t_2) for different combinations of N_0 and N_{final}.

Table 4 - Values of k and doubling times (t_2) for various values of N_o holding N_{final} constant

N_o	N_{final}	k/cell.yr	Doubling time (t_2)
10^3	$2x10^9$	2.9	0.24 yr = 88 days
10^4	$2x10^9$	2.44	0.28 yr = 102 days
10^5	$2x10^9$	1.98	0.35 yr = 128 days
10^6	$2x10^9$	1.52	0.46 yr = 168 days

Table 5 - values of k and doubling times (t_2) for various values of N_{final} holding N_o constant

N_o	**N_{final}**	**k/cell.yr**	**Doubling time (t_2)**
10^5	10^8	1.38	0.50 yr = 183 days
10^5	10^9	1.84	0.38 yr = 139 days
10^5	10^{10}	2.30	0.30 yr = 110 days
10^5	10^{11}	2.76	0.25 yr = 91 days

Returning to the central hypothesis, we are attempting to determine if there is a burst of mutations during the growth of a human breast during puberty. It has already been established by the work of Cha et al., 1994 that such a burst occurs in the rat mammary gland in the c-H-ras oncogene, beginning around the first estrus. It is therefore necessary for us to determine if:

(i) a similar burst occurs in other gene sequences in the pubertal rat mammary gland

(ii)a similar burst occurs in humans in the c-H-ras oncogene

(iii) a similar burst occurs in other gene sequences in the pubertal human mammary gland

In the event that the burst occurs, we must determine

(i) is the burst limited to a specific base change, for example G->A

(ii) determine when it occurs and

(iii) determine the distribution of mutants and the average size of the clusters.

Let's consider the case in which the total number of epithelial cells is less than 10^7 cells, as is the case for the first net 7 doublings of the pubertal breast (10^5-10^7). In the case in which N = 10^5 cells, the expected number of mutant cells at the time of doubling of N would be 10^{-2}, using a mutation rate of 10^{-7}/cell.generation. Since we cannot speak of 0.01 or 0.1 mutants arising from a population of 10^5 or 10^6 cells at the time of doubling, this is not a good physical representation of the process as the generation of spontaneous mutants can be considered to be a binomial event, that is, a mutant is generated or it isn't. Therefore, when N<10^7 cells, we must think of the probability of finding a cluster of a specific size from a sample of women rather than thinking of the expected number of individual mutants which are generated at the time of doubling. Table 6 summarizes the expected size of a single mutant cluster and the probability of finding a cluster of a specific size in a sample of women, as a function of the number of epithelial cells in a growing mammary gland. We must also consider the fact that even at N>10^7 cells, there is still a possibility of not observing any mutants, the probability of which can be determined using the Poisson distribution.[6] Table 5 show the probabilities of not observing a single mutant

[6] The Poisson distribution can be described by the equation
$P(n)= \lambda n e^{-1}/n!$ where P(n) is the probability of observing n events and λ is the expected number of events.

cell and the expected number of independent mutants generated at each doubling as a function of the changing values of N in the developing gland. Therefore, it will be necessary to model the total expected number of mutants and to the determine their distribution and average colony size.

Table 6 - Probability of not observing a single mutant as a function of the number of epithelial cells in the developing gland (N)

Number of cells at risk (N)	**Expected number of independent mutants**	**P(observing no mutants)**
1.625×10^5	(0.01625)	0.98
3.25×10^5	(0.0325)	0.97
6.25×10^5	(0.0625)	0.94
1.25×10^6	(0.125)	0.88
2.5×10^6	(0.25)	0.78
5×10^6	(0.5)	0.61
10^7	1	0.37
2×10^7	2	0.14
4×10^7	4	1.8×10^{-2}
8×10^7	8	3.35×10^{-4}
1.6×10^8	16	1.1×10^{-7}
3.2×10^8	32	1.2×10^{-14}
6.4×10^8	64	1.6×10^{-28}
1.28×10^9	128	2.6×10^{-56}

$P(o)=e^{-1} = 1/e = 0.367$ as $\lambda = 1$

Table 7 - Expected size of mutant cluster and probability of observing a cluster of such a size in a cross-section of women as a function of the number of cells present in the breast at the time at which the mutation is obtained

Number of cells at risk (N)	**[7]Expected size of single mutant cluster in mature breast**	**[8]Mutant fraction per sector if single cluster**	**Probability of observing cluster of appropriate size**
1.625×10^5	$2^{14} = 16{,}384$	4×10^{-5}	2/100 women
3.25×10^5	$2^{13} = 8{,}192$	2×10^{-5}	3/100 women
6.25×10^5	$2^{12} = 4{,}096$	1×10^{-5}	6/100 women
1.25×10^6	$2^{11} = 2{,}048$	5.1×10^{-6}	1/10 women
2.5×10^6	$2^{10} = 1{,}024$	2.55×10^{-6}	2-3/10 women
5×10^6	$2^9 = 512$	1.3×10^{-6}	5/10 women
10^7	$2^8 = 256$	6.5×10^{-7} (ND)	every woman
2×10^7	$2^7 = 128$	3.3×10^{-7} (ND)	every woman
4×10^7	$2^6 = 64$	1.65×10^{-7} (ND)	every woman
8×10^7	$2^5 = 32$	8.3×10^{-8} (ND)	every woman
1.6×10^8	$2^4 = 16$	4.2×10^{-8} (ND)	every woman
3.2×10^8	$2^3 = 8$	2.1×10^{-8} (ND)	every woman

[7]Assume that each single mutant generated at the time of doubling of the breast may give rise to a single mutant cluster or colony

[8] This assumes that the mutant cells do not possess any proliferative advantage over normal cells and both grow at the same rate.

Each sector will contain 4×10^8 epithelial cells ($2 \times 10^9/5$ sectors)

After the gland reaches 10^7 cells, it is expected that there will be more than 1 positive sector per breast ($>10^7$ cells) and more than one mutant cluster per sector ($>5 \times 10^7$ cells). However, the mutant fraction will not be expected to increase significantly to the point at which the expected mutant fraction increases over ~6×10^{-7}.

ND means not detectable using the current limits of MAMA.

$6.4x10^8$	$2^2 = 4$	$1.05x10^{-8}$ (ND)	every woman
$1.28x10^9$	2	$5.3x10^{-9}$ (ND)	every woman
$2.56x10^9$	1	$2.6x10^{-9}$ (ND)	every woman

This work is based on the detection of mutants in the human pubertal breast. It is therefore imperative that we consider the different mechanisms forcing the generation of mutants in the whole mammary gland. In the growing breast, there are two sources of mutants. The first source is the spontaneous generation of mutants from normal cells which is driven by the parameter r, which is the spontaneous mutation rate in equation 3. The second source of mutants is the growth of mutant cells which were generated spontaneously during previous cell divisions, which will be referred to as preexisting mutants hereafter. The parameter p in equation 3 is similar to the parameter k for normal cells and describes the net growth rate for the mutant cells. Again, the effect of this component is exponential as in the case of normal cells, the rate of growth of the mutant cells is dependent on the number of existing mutant cells. For predictive purposes, we will assume that p = k. However, in reality, p may be different from k because the net turnover rates of mutant epithelial cells may be significantly different from that of normal cells. In addition, perhaps only a select fraction of the mutant cell population may actually be undergoing division.

If we attempt to model the total number of mutants as a function of the number of epithelial cells N in the developing mammary gland, we will arrive at an equation such that

$$\Delta M = rN\Delta t + pM\Delta t \qquad (3)$$

where ΔM is the change in the number of mutants and Δt is the change in time. For an infinitely small Δt, equation 3 approaches the partial differential equation and becomes equation 4, where rN is the contribution from the spontaneous generation of mutants and pM is the contribution from the growth of pre-existing mutants. The premise is that every mutant cell which arises in the growth of an organ is capable of giving rise to an independent colony of mutant cells, driven by the component pM.

$$\frac{\partial M}{\partial t} = rN + pM \qquad (4)$$

When the linear differential equation expressed in equation 8 is solved using integrating factors, the following solutions are obtained. Equation 5 is the solution when p is different from k. When $k > p$, $e^{kt} - e^{pt} > 0$, $k-p > 0$, and $M > 0$ growing as e^{kt}. When $p < k$, $e^{kt} - e^{pt} < 0$, $k-p < 0$, and $M > 0$ growing as e^{pt}. When $p = k$, the solution becomes equation 6 because in this case, equation 5 becomes an indeterminant (0/0). The ratio of p to k determines the total number of mutants such that when $p > k$, M/N grows exponentially, when $p = k$, M/N grows proportionally to time and when $p < k$, M/N decays exponentially. The ratio of p to r determines the size distribution of the clusters, such that when $p >> r$ and $p << r$, the situations depicted in Figures 12 and 13 result. The solutions to Equation 4 will allow us to determine the actual value after the number of mutants is determined from the tissue samples.

$$M = \frac{PN_o}{k-p}[e^{kt}-e^{pt}] \qquad (5)$$

$$M = pN_o te^{kt} \qquad (6)$$

The first specific condition which we will consider is the case where $p << r$, that is the effect of the spontaneous generation of mutants far outweighs the effect of the net

growth of pre-existing mutants. The effect of p would be to give rise to smaller colonies, while the effect of r would be the initiation of a large number of independent mutants. The net effect of this condition would be predicted to be a large number of colonies of smaller average size. The expected distribution in the mature breast is depicted in Figures 12 and 13.

The second condition which we will consider is the case where $p >> r$, that is the effect of the net growth of pre-existing mutants far outweighs the effect of the spontaneous generation of mutants. The lower value of r would implicate the initiation of a smaller number of mutant cells capable of giving rise to independent colonies. However, the value of p would mean that each of these single independent colonies would on average contain a greater number of mutant cells. The expected distribution in the mature breast is depicted in Figures 12 and 13. In the case in which $p = r$, an intermediate situation would be expected. Figure 11 is a plot of the expected number of colonies versus the expected number of mutants per colony for the cases $p << r$ and $p >> r$. In any of the specific situations described, it is clear that mutants which arise soon before the mammary gland attains its final N will not have the chance to show themselves as a detectable mutant colony (see Table 6).

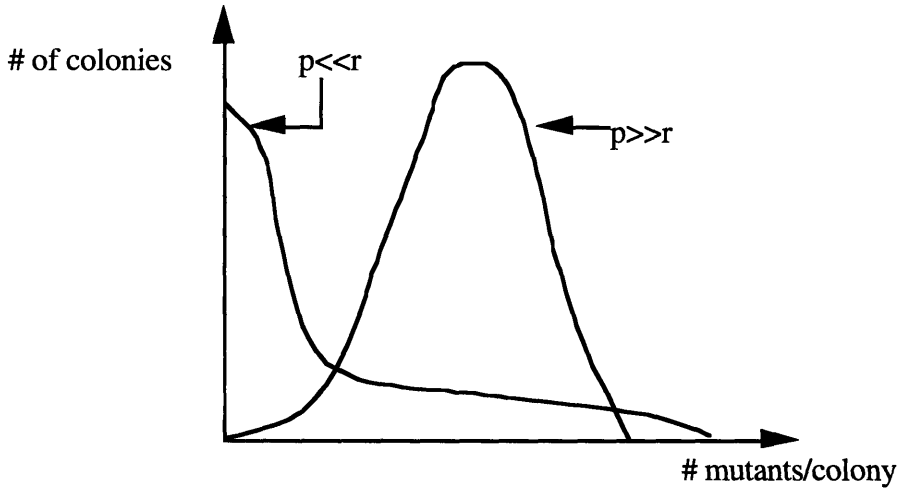


Figure 11- Expected distribution of number of mutant colonies versus the number of mutants per colony

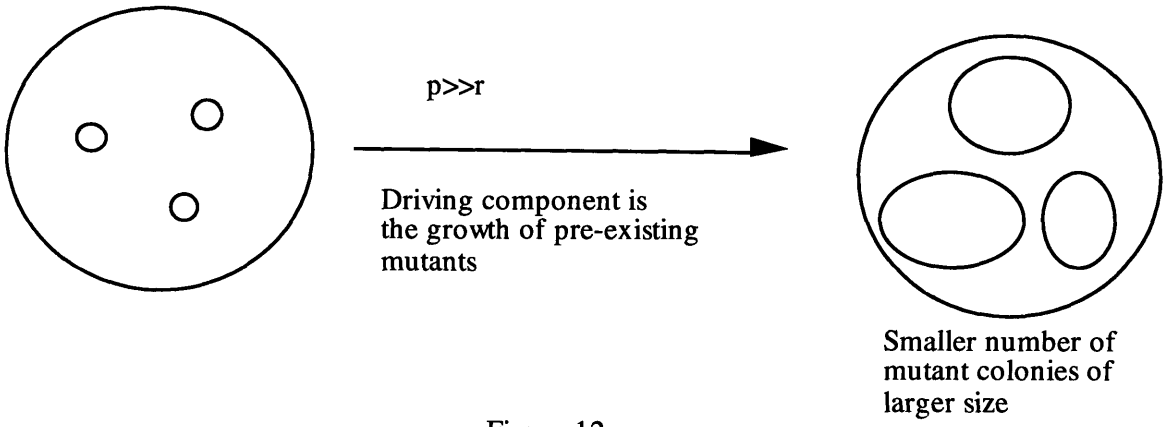


Figure 12

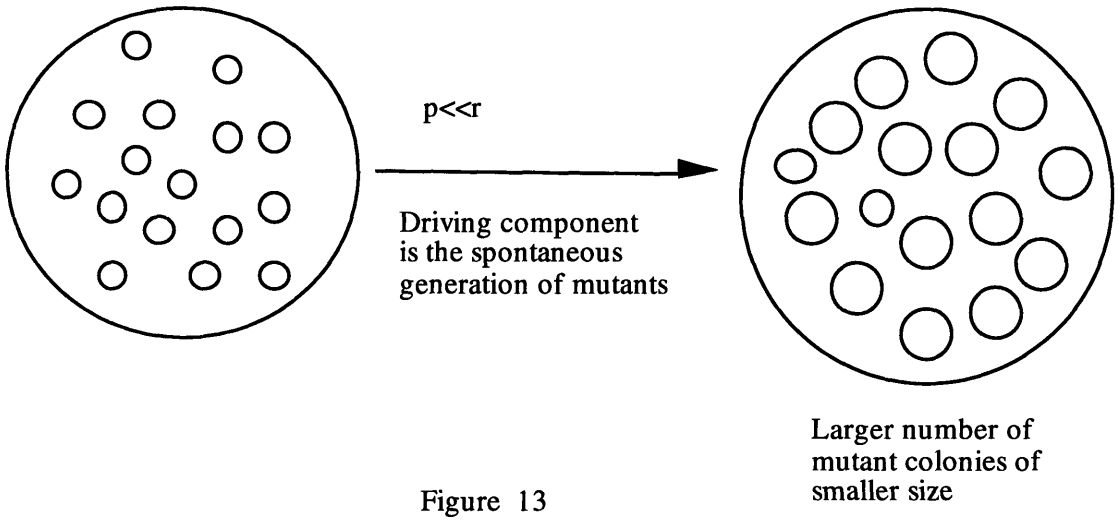


Figure 13

In order to detect these mutant colonies, a minimum mutant fraction of 10^{-6}-10^{-5} is required. This means that in the mature breast of approximately $2x10^9$ epithelial cells which is divided into 5 equally sized sectors containing approximately $4x10^8$ cells, there must be at least 1 colony containing a minimum of 400 to 4000 mutant cells (see Table 6).

3. MATERIALS AND METHODS

3.1 Isolating Mammary Epithelial Cells from a Reduction Mammoplasty Sample

One of the major challenges of this work is to obtain a purified fraction of the epithelial cells lining the ducts. These are the cells which are at risk of getting mutations and giving rise to carcinomas. The approach which has been used is an adaptation of the method outlined by Freshney, (1992). The sterile [9]surgical discard tissue which was obtained from reduction mammoplasties was first cut into sections of predetermined masses so that we may obtain sufficient numbers of epithelial cells for molecular analysis. The "whitish" areas of the tissue were excised from the remainder of the tissue, which according to Freshney, (1992) are the regions which contain the ducts. Each tissue section had an average mass of [10]6 g. The tissue was then minced using a razor blade and then placed into a 50 ml centrifuge tube containing approximately 30 ml of tissue processing media, 10 ml of 5X enzyme solution and 5 ml of fetal bovine serum (FBS). The tissue processing media contains Ham's F12 nutrient mixture supplemented with insulin (10 μg/ml), penicillin (100 U/ml), streptomycin (100 μg/ml), polymixin B (50 U/ml) and Fungizone (5 μg/ml). The enzyme solution consists of collagenase type IV (1,500 U/ml) dissolved in an appropriate amount of tissue processing media, filtered sequentially through 0.8 μm, 0.45 μm, and 0.20 μm Nalgene filters and hyaluronidase (100 U/ml) dissolved in an appropriate amount of tissue processing media and filtered through a 0.20

---

[9] The tissue used in this work has been kindly provided by Dr. Samuel Singer of the Brigham and Women's Hospital in Boston.

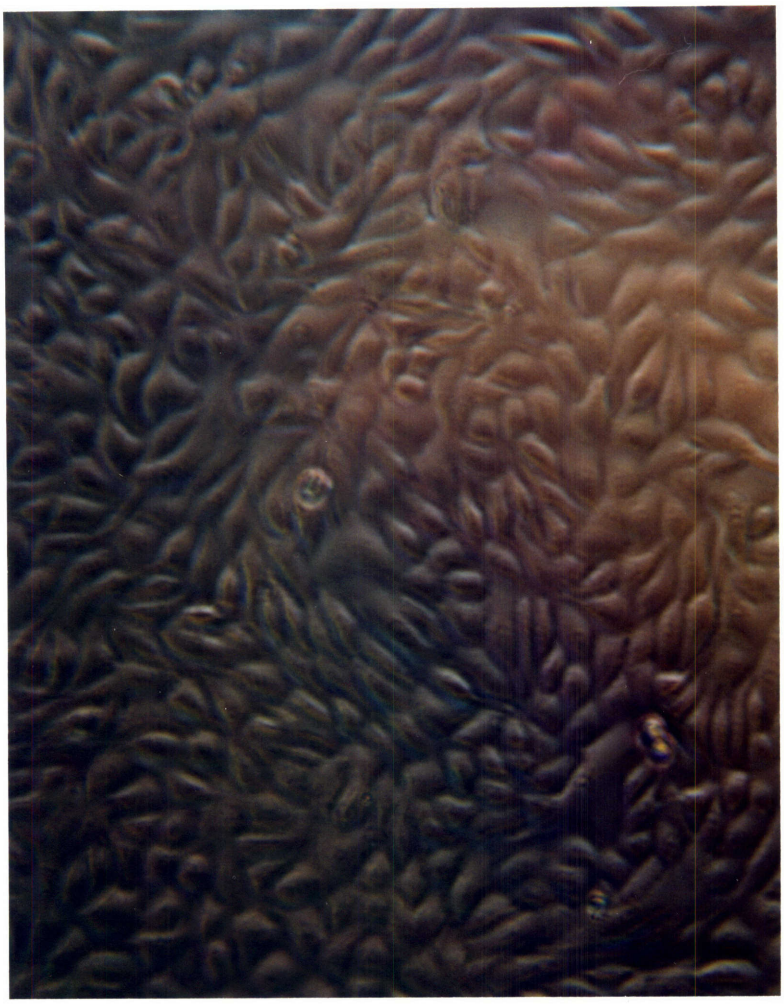
[10] A 4g mass of breast tissue obtained from a reduction mammoplasty (Dr. Sam Singer) yielded an average of $3x10^7$ cells (samples HBDRA1 (PML) and HBDR1E (PML)). Assuming that this cell fraction contains about equal proportions of epithelial cells and fibroblasts, then there are $1.5x10^7$ epithelial cells in this sample. Assuming that on average 1 g of tissue contains about 10^8 cells, then 4 g of tissue should contain $4x10^8$ cells. Therefore in the 4 g sample, the epithelial cells constituted approximately 3.75% of the whole tissue sample.

μm Nalgene filter. Equal volumes of the sterile collagenase solution and the hyaluronidase solution were mixed together to give the 5X solution.

The second step in the overall process is the overnight digestion of the tissue sections with the 5X enzyme solution, the tissue processing media and the fetal bovine serum at 37°C in a shaker water bath. Collagenase digestion yields primarily fibroblasts and epithelial cells. The centrifuge tubes were then spun at 800g for 10 minutes and the overlying fat layer and the media layer were both removed. The remaining pellet was then resuspended in 5 ml of mammary epithelial growth medium (MEGM - Clonetics, CA) using a 5 ml pipetteman. One drop was placed on a microscope slide to determine the degree of completion of digestion. If microscope examination revealed that there were still epithelial cells attached to the surrounding stroma, then digestion with the mix of 5X enzyme solution, tissue processing media and FBS was carried out again for another 3 hours. This process is repeated until microscope examination revealed that digestion was complete. After digestion was complete, the cell pellet was resuspended in 4 ml of MEGM and 1 ml was counted using a Coulter counter and the remaining 3 ml were split between three 25 ml T flasks. 5 ml of MEGM was added to each flask.

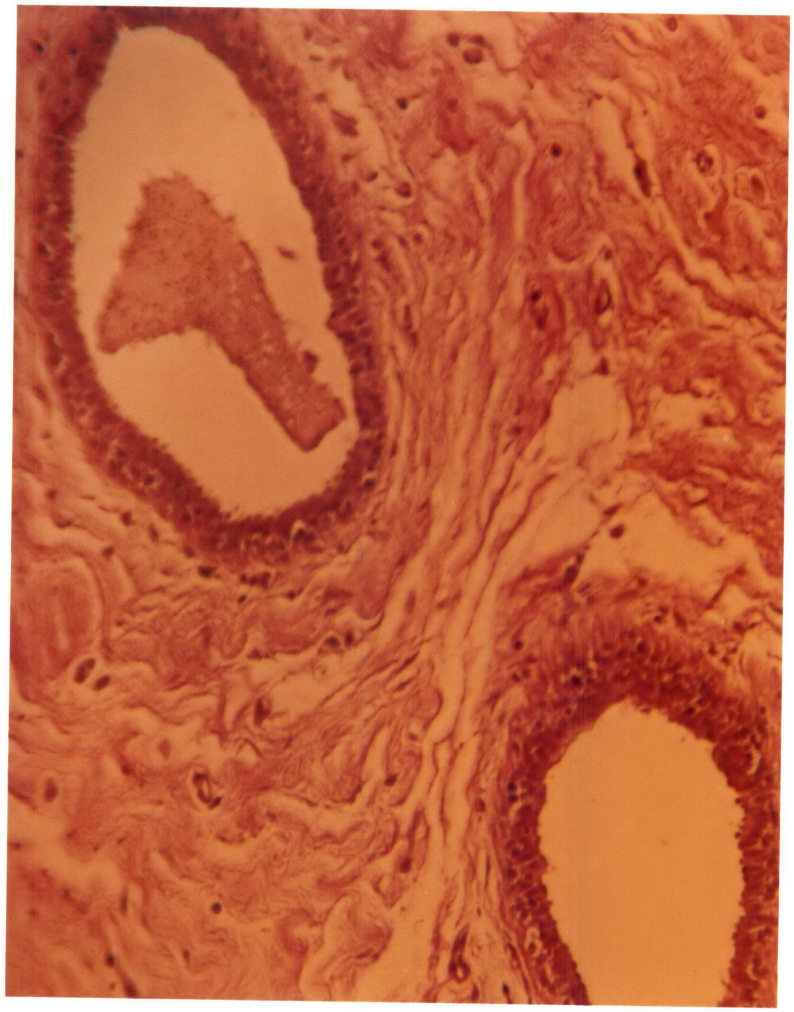
The next step is primarily for separating the fibroblasts from the epithelial cells and takes advantage of the fact that fibroblasts attach more readily to the surface of a plastic tissue culture dish. The fibroblasts attach within 1-2 hours of seeding the culture, whereas epithelial cells will attach several hours later. Two hours after the first set of cultures were seeded, the media from each of the flasks was transferred to a new flask. Again, two hours after the second set of cultures were seeded, the media from each of the flasks was transferred to another set of new flasks. New media was added to both the first and second set of flasks. The three sets of cultures were grown for several days in an effort to determine the types of cells which attached and survived.

Indeed, after 5 days of growth in an incubator kept at 37°C, two of the three flasks from the third set contained only epithelial cells and the third flask contained mostly epithelial cells with about 20% clones of fibroblasts. The first and second sets of flasks all contained fibroblasts. These results could mean that the differential attachment procedure was successful in separating the epithelial cells and the fibroblasts at the end. On the other hand, it could be that there were actually epithelial cells which had attached to the surface of the flasks in the first and second sets of flasks. However, the rapidly growing sturdy fibroblasts in culture overgrew the epithelial cells. In any case, the predominance of epithelial cells in the third set of flasks indicate that the separation was successful, and if indeed there were epithelial cells which were lost in the first and second set of flasks, the efficiency of the procedure would have to be improved. Figure 14 is a picture of the attached epithelial cells in culture. The cells were typed based on their morphologies. However, definitive typing of the cells should be done immunohistochemically using antibodies against cytokeratins, which appear on the surface of epithelial cells only.



3.2 Histology of the Reduction Mammoplasty Sample

A single 4 g piece of tissue from the same reduction mammoplasty sample was fixed in 40 ml of neutral buffered formalin (NBF) for 10 days at 4 °C. The sample was then paraffin embedded and sections were made which were stained with the dyes, hematoxylin and eosin (H&E). Figure 15 is a photo of the slide of a cross section of the ducts surrounded primarily by adipocytes and some loose connective tissue.



3.3 Alternative Approaches for Separating Fibroblasts and Mammary Epithelial Cells

Other possibilities for the separation of the population of fibroblasts and epithelial cells which are derived from the collagenase digestion of the surgical samples include

- separation of the cells using a medium (Percoll) composed of colloidal silica coated with polyvinylpyrrolidone (PVP) for density gradient centrifugation of the cells (Pharmacia Biotech)
- use of magnetic beads (Dynal) conjugated to antibodies to cytokeratin 19 on the surface of epithelial cells or
- use of fluorescent antibodies to cytokeratin 19 on the surface of the epithelial cells and separation of the cell types using a fluorescence activated cell sorter (FACS).

3.4 Cell Culture for Obtaining Positive and Negative Controls

The positive controls which were to be used for MAMA in reconstruction experiments and the actual samples were obtained from cell lines derived from tumors. The GGC->GTC (glycine->valine) mutant which was used to assay the codon 12 c-H-ras mutation was derived from a human bladder carcinoma cell line, T24 (Capon et al., 1983). The GTC->TTC (valine->phenylalanine) mutant which was used to assay the exon 5 - codon 157 mutation was derived from the H2087 cell line, derived from a small-cell lung carcinoma (SCLC) and was a generous gift from Dr. Herb Oie (National Cancer Institute). TK6 (human B lymphoblastoid cell line) DNA was used as the negative control and was a gift from Ms. Jackie Goodluck-Griffith (MIT Center for Environmental Health Sciences).

H2087 cells were grown in RPMI 1640 media (Gibco-BRL) supplemented with 10% horse serum (Gibco-BRL). They were grown in 75 ml plastic T-flasks. TK6 cells were grown in suspension using RPMI 1640 media (Gibco-BRL) supplemented with 10% bovine horse serum (Gibco-BRL). T24 cells were grown in 75 ml plastic T-flasks using RPMI 1640 media supplemented with 10% bovine calf serum (BCS, Gibco-BRL). Cell stocks were counted using a Coulter counter. Cell stocks which needed to be frozen were first centrifuged at 1000g for about 15 minutes in either a 10 ml or 50 ml centrifuge tube and were resuspended in fresh media supplemented with 10% DMSO.

3.5 Genomic DNA Extraction

3.5.1 Extraction Protocol

The extraction of genomic DNA (gDNA) from cell lines which are used as both positive and negative controls is done using the protocol described by Khrapko et al. (Manuscript in preparation). The protocol may also be used for the extraction of DNA from a purified fraction of mammary epithelial cells or from a tissue section, in the case of an organ which is composed mainly of a single cell type.

The volumes quoted in the following description have been optimized for the extraction of genomic DNA from approximately $5x10^6$ cells of cells in suspension or from a 20 mg tissue section. The cells are resuspended as evenly and as gently as possible in 450 µL of standard TE buffer, pH 8.0 (50 mM Tris/HCl, 10 mM EDTA, pH 8.0). 25 µL of proteinase K (20 mg/ml) is added to the suspension, followed by 25 µL of 10% sodium dodecyl sulfate (SDS) and it is digested for 3 hours at 55°C. This is then followed by the addition of 1µL of 10 mg/ml RNase A and is digested for 45 minutes at 37°C in a shaker

water bath. The digest is then centrifuged for 15 minutes at 12,000 rpm and the supernatant (approximately 80-90% of the original volume) is removed. Centrifugation allows the high molecular weight DNA to stick to the sides at the base of the tube. 20 µL of 5M NaCl is then added to the remaining suspension. The DNA in the remaining suspension is then precipitated by the addition of 900 µL of 100% ethanol and increasing the speed of vortexing. The 100% ethanol is then removed and the DNA spool is washed by adding 800 µL of 70% ethanol. A small volume of the 70% ethanol is left in the tube so that the DNA spool never becomes completely dry. Otherwise, it would be very difficult to dissolve the DNA. The resulting DNA spool is then transferred to approximately 500 µL of a 20X dilution of standard TE buffer , pH 8.0 (2.5 mM Tris/HCl, 0.5 mM EDTA, pH 8.0).

3.5.2 Determining the yield of DNA after extraction

A rough determination of the yield of DNA is determined by running the final DNA solution dissolved in a 20X dilution of TE buffer, pH 8.0 in parallel with standards of 3,10, 30, 100 and 300 ng of high molecular weight herring DNA on an 8% PAGE gel. The gel is run for 2 hours at 250 V and a comparison of the intensities of the smeared bands representing the actual samples with the standards allows an estimation of the final concentration of DNA.

Two more accurate methods of determining the yield of the DNA extraction procedure include (i) measuring the absorbance of the DNA sample at 260 nm using a uv spectrophotometer and (ii) coamplifying a specific sequence of a single copy gene such as the APC gene with a known number of copies of an internal standard.

Expecting approximately 6 μg of DNA from 10^6 cells, I achieved a yield of 120% for gDNA extraction from T24 cells and 150% from TK6 cells as determined by the absorbance at 260 nm on the uv spectrophotometer. This indicates that either (i) the estimate of 6 μg of DNA/10^6 cells is an underestimate, as not all the cells at the time of freezing may have been strictly diploid (they are diploid but are doubling their DNA), depending on which phase of the cell cycle the cell is in or (ii) the cell count given by the Coulter counter is an underestimate as cells may be clumping.

3.5.3 Determining copy number using an internal standard

The number of copies of a specific sequence may be determined if the assumption is made that there has been no gene amplification. For example, the number of copies of the p53 tumor suppressor gene and the c-H-ras oncogene from a specific DNA preparation may be determined using the number of copies of the wildtype APC tumor suppressor gene as a benchmark. An APC point mutant which runs as separate peak from the wild type using constant denaturing capillary electrophoresis (CDCE) at a temperature of 64°C was used. A population of these mutants at a concentration of 10^5copies/μl was obtained from Dr. Gengxi Hu. An estimated 1:1 mixture of APC mutant:(TK6/T24/H2087 DNA) was amplified using primers which did not differentially amplify either mutant or wildtype. The ratio of the wildtype peak area to the mutant peak area multiplied by the concentration of mutant copies yields the concentration.

3.6 Mismatch Amplification Mutation Assay (MAMA)

3.6.1 The General Method

The mismatch amplification mutation assay (MAMA) was developed by Dr. Rita Cha and is a variation of PCR which allows selective amplification of a mutant allele versus a wildtype allele. This is achieved by the use of a mismatch primer which has a double mismatch with the wildtype allele and a single mismatch with the mutant allele at the 3' ultimate or penultimate end (Cha et al., 1992). MAMA has been shown to have a sensitivity of 10^{-5} for a specific G->A mutation at the second nucleotide of codon 12 in the c-H-ras gene. Thus it can detect 30 mutant copies among $3x10^6$ copies of the wildtype.

For each specific mutation which needs to be assayed, a series of mismatch primers each with a different mismatch needs to be screened using pure mutant DNA as a template, in order to determine the primers with the best chances for achieving a MAMA assay with the highest sensitivity. Primers which contain mismatches with T have the lowest probability of extending the mutant allele in a MAMA assay (Personal communication, Dr. Rita Cha). Therefore, primers have been designed to incorporate G-G, A-G, A-A, A-C and C-C mismatches , in order of decreasing extension efficiencies.

After the initial screening of the set of primers, the conditions for each potential primer must be optimized before the limit of sensitivity of the MAMA assay may be determined via a reconstruction experiment. A typical 50 µL MAMA reaction is composed of: approximately 1-5 µg of genomic DNA ($3x10^5$-$1.5x10^6$ copies), 200 nM of each primer or $6x10^{12}$ copies of each primer (one of which is fluorescein labeled), 4 mM $MgCl_2$, 0.15 mM dNTPs or $4.5x10^{15}$ dNTP molecules, 5 Units of Stoffel DNA

polymerase (Perkin-Elmer Cetus), 0.5 μL of 100X BSA and 1X Stoffel buffer (10 mM Tris-HCL, 10 mM KCl, pH 8.3). The $MgCl_2$ is provided separately as a 25 mM stock solution. The Stoffel fragment of Amplitaq DNA polymerase is being used as it lacks an additional level of proofreading activity over Taq DNA polymerase. It lacks both 5' to 3' and 3' to 5' exonuclease activities, while Taq polymerase only lacks the 3' to 5' exonuclease activity. It is a 61 kD, highly thermostable enzyme which lacks the N-terminal 289 amino acid portion of AmpliTaq. For this reason, the Stoffel fragment of Amplitaq can be used to discriminate and extend more efficiently, primers which have a single mismatch with the wildtype and a perfect match with the mutant.

The variables which need to be optimized prior to the reconstruction experiments are the annealing temperature, the concentration of NaCl, and the concentration of glycerol[11]. The annealing temperature in general is increased as close as possible to the lowest melting temperature of the two primers, in order to decrease the amount of non specific amplification arising from hybridization with random sites in the genome. For a 20 bp primer, the probability of annealing to any site with perfect complementarity in the genome is $(1/4)^{20}$. The addition of 25 mM NaCl (final concentration) increases the specificity of annealing of the primers (unpublished observation, PML). The positively charged sodium ions tend to act as a sort of "glue" to bind the negatively charged primers and DNA template.

[11] It has been shown that 5-10 % vol/vol of glycerol increases the specificity of MAMA (Cha et al., 1992). However, levels greater than 15% actually inhibit the reaction.

3.6.2 Sequence of Mismatch Primers and Design

The sequences of the primers used in the MAMA assays for the detection of both G->T mutations in codon 12 of the c-h-ras gene and in codon 157 of exon 5 are listed below. There is an A/G mismatch between the mismatch primers and the wildtype sequence for both mutations.

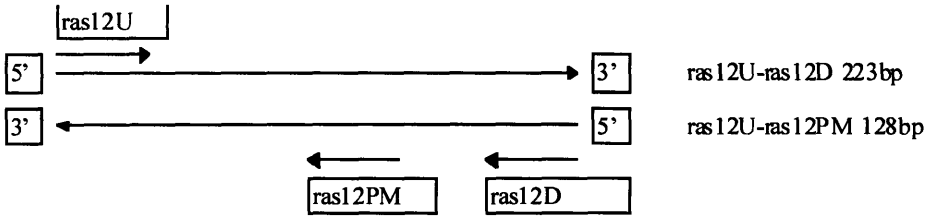


Figure 16 Schematic of codon 12 H-ras G->A mutant MAMA primers and product lengths

ras12U - 5' GCT GAG CAG GGC CCT CCT TG 3'

ras12D - 5' AGC TGC TGG CAC CTG GAC GG 3'

ras12PM - 5' CGC ACT CTT GCC CAC ACC GA 3'

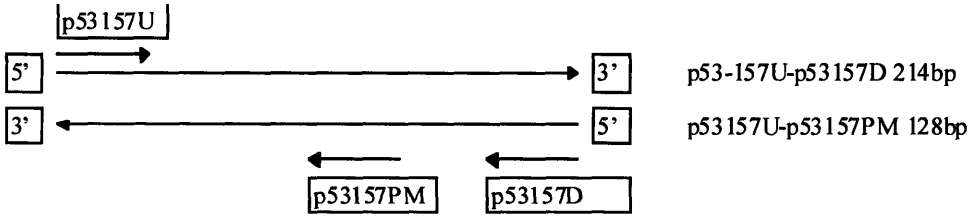


Figure 17 Schematic of p53 codon 157 G->T mutant MAMA primers and product lengths

p53157U - 5' CTT CCT GCA GTA CTC CCC TG 3'

p53157D - 5' GCC CCA GCT GCT CAC CAT CG 3'

p53157PM - 5' TAG ATG GCC ATG GCG CGG AA 3'

3.7 Optimizing the System - The Story of the Positive Controls

Before the reconstruction experiments can be carried out for determining the sensitivity of the assay, the system must first be calibrated using a dilution series for the positive control. Such a system has been tested only for the c-H-ras codon 12 mutation using T24 DNA as the positive control. High molecular weight herring DNA was used as the carrier DNA i.e. the DNA without any copies of the mutant h-ras sequence. This is important so that there will be no background signal generated. Both the herring DNA and the T24 DNA were subjected to a simultaneous 24 hour restriction digestion using the enzyme PstI. The completeness of digestion was checked on an 8% PAGE gel.

The carrier DNA is important for keeping the total amount of DNA constant from dilution to dilution. A series dilution was made, adding the appropriate amount of herring DNA to each eppendorf tube to pre-coat the walls. At very low dilutions of the positive control (10-1000 copies/ul), it is possible that the restriction digested DNA sticks to the side of the walls and are unavailable for amplification, thus artificially rendering the system non-linear. Restriction digestion of the DNA increases the amplification efficiency by approximately 20% (unpublished data, PML). The absolute positive control was 10^6 copies of T24 DNA or approximately 3.3 ug. The absolute negative control was 3.3 ug of herring DNA. For the series dilutions, the range of initial copy numbers of T24 postive control was from 10^5, 10^4, 10^3, 100 and 10 copies per 25 ul PCR reaction. 4 units of

Stoffel polymerase were used per reaction with 4 mM MgCl2. The cycling profile was 95°C for 15 seconds, 62°C for 25 seconds and 72°C for 25 seconds. The reactions containing 10^6 and 10^5 copies of T24 DNA were run for 30 cycles, 10^4 copies for 35 cycles, 10^3 copies for 40 cycles, 100 copies for 45 cycles and 10 copies for 50 cycles. This protocol of cycle numbers was used to avoid overamplification of the product and possible nonlinearity in the amplification results.

APPENDIX

A1.1 History of Rodent Mammary Neoplasia

The use of rodent models has advanced our understanding of the mechanisms of mammary neoplasia. Guillino and his colleagues began laying the foundation for the field as far back as 1975 when they treated 50 day old female Buf/N rats with the potent carcinogen and mutagen, N-nitroso-N-methylurea (NMU) at a dose of 30 mg/kg (Guillino et al., 1975). They discovered that 90% of the treated rats developed mammary carcinomas about 60 days after exposure to the carcinogen. About 8 years later, Sukumar et al. (1983) were able to demonstrate that each of the 9 mammary carcinomas which arose subsequent to the treatment of 50 day old Buf/N rats with NMU contained a transforming H-ras allele. The field moved closer to determining whether chemicals cause mammary tumors by mutating specific cellular targets, in particular, oncogenes with the experimental results of Zarbl et al., (1985). They demonstrated that when 50 day old female Sprague-Dawley, Buffalo and inbred Fischer 344 (F344) rats were treated with a carcinogenic dose of either NMU (30 mg/kg) or dimethylbenz(a)anthracene (DMBA), approximately 75% of the mammary tumors developed in the rats treated with NMU carried a G->A transition mutation at the second nucleotide of codon 12 in the c-H-ras-1 gene. The DMBA induced mammary tumors on the other hand carried specific A->T transversion mutations in codon 61 of the c-H-ras-1 gene. They reasoned that since the particular G->A mutations were present in the NMU-induced tumors only and not in the DMBA-induced tumors, that NMU induced the mutations in the c-H-ras gene and perhaps contributed to the initiation of carcinogenesis. This development set the stage for the ideas that (1) each chemical had a specific cellular target and perhaps a hallmark or preferred base change (GC->AT for NMU and AT->TA for DMBA) and that (2) the genetic change was responsible for the

development of the mammary tumor (first evidence of a link between mutagenesis and carcinogenesis).

A1.2 Timing of Initiation

After it was established that NMU induced tumors reproducibly carried a G->A mutation at codon 12 of the c-H-ras gene, may researcher began to focus on determining the exact timing of this genetic insult. Lu and Archer, (1992) were trying to understand the molecular basis of the differential susceptibilities of the Buf/N and the Copenhagen rat to mammary carcinogens. Earlier experiments which they had performed showed that kinetics of formation and repair of O6-methylguanine in the mammary tissue of both strains of rats were identical. In addition, the extent of methylation of a restriction fragment containing exons 1-4 of the c-H-ras gene by NMU and the level of H-ras expression in mammary tissue was not different for the two strains. They therefore decided to measure the frequency of activated codon 12 H-ras alleles in the mammary epithelium of both strains of rats following exposure to 30 mg/kg or 50 mg/kg or NMU. They thought that perhaps the differential susceptibility between the two rats lay in either the timing of initiation or a lack of initiation in the Copenhagen rat.

Interestingly, 30 days after treatment with NMU, a few of the glands tested in the Buf/N rats had mutant fractions of about 10^{-5}. At 60 days post-treatment, several of the glands had mutant fractions between 10^{-4} and 10^{-3}, an increase of about 10-100 fold. This increase in the number of mutant cells is in agreement with the normal proliferation of the mammary epithelium during puberty. In contrast, none of the 40 glands tested from the control rats contained the mutations (mutant fraction > 10^{-5}). In the case of the Copenhagen rat, 30 days after treatment, there were four positive glands containing mutant fractions of 10^{-4} to 10^{-5} and 15 glands with mutant fractions of 10^{-5}. Sixty days after

treatment, 19 of the glands tested were positive with mutant fractions of 10-5. None of the 40 glands tested in the control Copenhagen rats were positive (had mutant fractions > 10-5). The authors concluded that since both the mammary epithelium from both the Buf/N and Copenhagen rats contained ras positive mutants, then the difference in susceptibility to mammary carcinoma induction was not a result of lack of initiation of the ras alleles. They hypothesized that perhaps it is due to the expression of a dominant autosomal suppressor gene in the mammary parenchyma of the Copenhagen rats, which prevents the proliferation of the activated ras epithelial cells into tumors.

There are two main problems with their interpretation which does not allow me to fully accept their conclusions. In the first place, the claimed sensitivity of their assay is 10^{-5}. However, in the molecular analyses, the authors used 1 ug of DNA which is equivalent to $3x10^5$ copies of the ras allele. This means that at a mutant fraction of 10^{-5}, there should theoretically be 3 copies of the ras allele. However, according to the statistical variation as predicted by the Poisson distribution, the actual number of copies can range from 0 to 7 copies of the allele per sample. Therefore, it is possible that there are positive glands, which may contain activated ras alleles at mutant fractions of 10^{-5} or lower and which cannot be detected by the assay. Such glands have effectively been scored as zero by the authors. Secondly, molecular analysis was not performed on all of the 12 glands in the rats. Therefore, it is possible that there are positive glands which have not been detected both in the treated animals and in the control animals. Such uncertainties motivate me to entertain the possibilities that the mammary epithelium of both strains of rats (both treated and control) might contain activated ras alleles.

A1.3 NMU Background

NMU is most commonly used for the laboratory synthesis of diazomethane. It has been used extensively in cancer research as a model compound to determine the molecular mechanisms underlying exogenous and endogenous nitrosamine carcinogenesis. It is a cytotoxic, mutagenic, teratogenic and carcinogenic compound. It has been tested in numerous animal species thus far and all were susceptible to NMU induced carcinogenesis. It is a potential human carcinogen based on its mutagenicity and animal carcinogenicity.

It is highly reactive and spontaneously decomposes quite rapidly at neutral or basic pH. It has an estimated half life of 30 minutes in vivo. NMU is readily sequestered within cells and for this reason, rapidly disappears from the blood stream of a rat and is uniformly distributed among the different organs of an animal within hours of i.v. administration. NMU decomposes to form the highly reactive metabolites, methyldiazohydroxide and isocyanate. The first metabolite is a methylating agent while the second is a carbamoylating agent. Methyl diazohydroxide reacts primarily with proteins, DNA and RNA inside a cell, while the cyanate ion reacts primarily with proteins. NMU reacts with both eukaryotic and prokaryotic DNA to form many promutagenic alkylated adducts. Figure 18 describes the spontaneous decomposition of NMU at neutral or basic pH

CH_3-N_2O-CO-NH_2	---------->	CH_3N=NOH +	HNCO
NMU		methyl diazohydroxide	isocyanate

Figure 18 - Spontaneous decomposition products of NMU

A1.4 Mutational Specificity of NMU in bacterial and mammalian cells

Further evidence for the hypothesis linking mutagenesis and carcinogenesis came from experiments detailing the mutational specificity of NMU in bacterial and mammalian cells and other rodent carcinogenicity models. Other chemicals which were used in the rodent assays included the polycyclic hydrocarbon DMBA and N-methyl-N'nitro-N-nitrosoguanidine (MNNG).

Richardson and colleagues (Richardson et al., 1987) characterized the DNA base changes and measured the levels of DNA alkylation following in vivo exposure of *E.coli* to NMU and N-ethyl-N-nitrosourea (ENU). They subjected *E.coli* carrying the xanthine phosphoribosyl transferase gene (gpt) as the genetic target on the plasmid pSV2gpt with 7 mM MNU or 35 mM ENU. They then compared the resulting mutational spectra (frequency of base changes, types of base changes and the sequence specificity of base changes) with the concentration of alkylation adducts (O6-alkyl guanine and O4 alkyl thymidine) in the genomic DNA after exposure. Numerous in vitro misincorporation studies, in vivo site directed mutagenesis studies and in vivo studies comparing the specific adducts formed after treatment with alkylating agents and the resulting mutation frequencies have suggested that the principal promutagenic lesions are O4 alkylthmidine and O6-alkylguanine. For NMU induced mutagenesis, it was hypothesized that the O6-methyl guanine adducts were the most important lesions. In order to determine if O6-methylguanine lesions were responsible for the mutations induced in intact cells, O6-methylguanine adducts were incorporated into a specific site of a viral genome and allowed to replicate in vivo. In vivo replication in bacterial systems (Loechler et al., 1984) and in mammalian systems (Mitra et al., 1989) resulted in GC->AT transition mutations. Singer et al., (1986) and Saffhill et al., (1985) have shown that an alkylated guanine causes

mispairing with a thymine during DNA replication in vitro, thereby fixing a G->A mutation.

Richardson et al., (1987) found that 100% of the NMU mutations were GC->AT transitions, in contrast to 73% of the ENU mutations. 21% of the ENU mutations were AT->GC transitions, 3% were GC->CG transversions and 3% AT->CG transversions. The profile was very different for the spontaneous mutants which comprised 41% insertions, 47% deletions, 6% GC->AT transitions and 6% GC->TA transversions. NMU demonstrated a positional specificity with 82% of the GC->AT transitions occurring at a middle guanine flanked by another guanine at the 5' position and either an A or a T at the 3' position. 71% of the ENU GC->AT transitions showed this specificity.

The ratio of O4-alkylthmidine to O6-alkylguanine was 0.014 for NMU and 0.28 for ENU, suggesting that perhaps the specific ratio of promutagenic adducts may be responsible somewhat for the different mutational spectra observed. The mutation frequencies for both NMU and ENU were however remarkably similar, with each having an average of $4x10^{-5}$. The spontaneous mutation frequency was 13% that of NMU and ENU and was $0.55x10^{-5}$. The similarity in mutation frequencies despite beginning with 5 fold less NMU and a combined alkylation due to NMU being 2.8 fold higher than that due to ENU suggests (1) that NMU is about 14 fold more reactive with the DNA at the specific site and that (2) the ethylated adducts are more efficient at fixing the promutagenic lesions into mutations, maybe due to enhanced mispairing or reduced or less efficient repair.

In the realm of NMU mutagenesis in human cells, DuBridge et al., (1987) used stable autonomously replicating Epstein-Barr virus (EBV) vectors carrying the bacterial lacI gene as the mutational target and established them in late passage 293 human

embryonic kidney cells. The vectors were later returned to E.coli for colony scoring for mutation detection and analysis. An NMU stock solution of concentration 10 mg/ml was added to the growth medium of the kidney cells at 30-40% confluence.

97% of all of the nonsense mutations and 97% of all the missense mutations were GC->AT transitions, implying that NMU has quite an overwhelming preference for GC->AT base changes. Inspection of the local sequence surrounding the mutated guanines in the missense mutations reveal that 74% of them have a 5' guanine, indicating a sequence or positional specificity. In fact, 48% of the mutated guanines occur within a run of 3 guanines. The 3' base specificity of an A or T observed by Richarson et al., (1987) was not observed by Dubridge et al., (1987). Theses studies provided support for the notion that NMU has a preferred base change and positional specificity and may indeed by responsible for the G->A mutations observed in the mammary carcinomas of rats treated with NMU.

A1.5 Further evidence for the mutational specificity of NMU from rodent models

The conclusions of Zarbl et al. were further corroborated by a number of rodent experiments in which a carcinogenic chemical other than NMU was used to induce carcinogenesis. Quintanilla et al. (1986) also found AT->TA transversions at codon 61 in over 90% of skin tumors induced in mice by treatment with DMBA. The molecular analysis was performed using an XbaI restriction enzyme polymorphism (the mutant was XbaI digestible). They also found that in a total of 12 tumors obtained from NIH mice treated with MNNG (6 papillomas and 6 carcinomas), none had the XbaI polymorphism, indicating that the two chemicals induced different mutations. Further support for the positional specificity of NMU was also obtained from the work of You et al. (1989) who compared the nature of proto-oncogene activation between spontaneously and NMU

induced lung tumors in strain A mice. They found that in the spontaneously derived tumors, 60% of the point mutations occurred at codon 12 of the Ki-ras gene while 30% occurred at codon 61 of same gene. This contrasts with the NMU induced tumors in which 100% of the mutations occurred at the middle guanine of codon 12 in the Ki-ras gene and were GC->AT (GGT->GAT) transitions. This further shows that while each chemical has its preferred base change, it has a different genetic target and target organ, which is dependent on the species treated.

A1.6 Formation of O6-methylguanine in DNA and its repair by O6-methylguanine methyltransferase (MGMT)

In addressing the mechanism of induction of mammary neoplasia in rats after treatment with NMU, it is imperative that a discussion of kinetics of formation and removal of the DNA lesion, O6-methylguanine and its removal by the repair enzyme O6-methylguanine methyltransferase (MGMT) be included. O6-methylguanine has been shown to lead to G->A mutations and to be the primary promutagenic lesion incurred in DNA after treatment with any member of the class of alkylating agents. There have been attempts to correlate the level of this specific adduct after treatment with an alkylating agent with the frequency of appearance of phenotypic expressions of transformation, such as tumors or preneoplastic lesions, using animal models. Qin et al. (1991) quantified the levels of DNA adducts formed in vivo in rat colon mucosa after a single treatment with varying concentrations of NMU. The phenotypic end point for their work was the appearance of aberrant crypt foci (ACF), some 4-16 weeks after treatment. They determined that the formation of adducts was dose-dependent, ranging from an average of 5-308 fmol/ug of DNA for treatment by NMU concentrations ranging from 1.5mM to 75 mM. The formation of ACF was also dose-dependent, increasing up until the 8th week. In contrast, there were no observable lesions in the control rats. Interestingly, the authors

in a previous experiment found specific G->A transition mutations at the 2nd nucleotide of codon 12 of the Ki-ras gene in the cells of the ACF.

Evidence from a number of animal and cell culture experiments show that repair of the O6-alkyl guanine adducts (independent of their origin) protects cells against malignant conversion. The O6-alkylguanine adduct has been shown to be repaired by the enzyme, O6-methyl(alkyl)guanine-DNA-methyltransferase (O6-MGMT). It is known that NMU induces tumors preferentially in tissues with low O6-MGMT activities, either because of low rates of synthesis of the protein or production of less efficient molecules. O6-MGMT repairs alkylated bases by removing the alkyl group and adding it to a cysteine residue within the protein. This reaction inactivates the protein and is referred to as "suicide inactivation". For this reason, the reaction for the resolution of alkylated bases is stoichiometric. Qin and his colleagues had previously proven that increasing the levels of MGMT activity decreased the susceptibility of animals to tumor induction by N-nitroso compounds. They determined this by the use of transgenic mice in which the E.coli MGMT gene (ada) attached to the Chinese hamster metallothionein I gene promoter had been introduced.

There has also been some synergy between two separate aspects of research on alkylating agents. On the one hand, there are the experiments which have quantified the levels of DNA adducts and MGMT in normal cells, in order to determine the levels of alkylating agents and the threshold values of alkylated adducts necessary for the malignant transformation of normal cells. On the other hand, there is another set of experiments which attempt to quantify the levels of MGMT in tumor tissue and non-target tissue, in order to determine their susceptibilities to a class of alkylating agents which are used as chemotherapeutic agents (methylating and chloroethylating nitrosoureas). Chemicals such

as chloroethylnitrosourea (CNU) and 1,3-bis(2-chloroethyl)-1-nitrosourea (BCNU) are included in this class.

Maze and Sampson, (1996) recently published some work in which they found that increasing the levels of MGMT in bone marrow stem cells rescues mice from the cytotoxic effects of BCNU. BCNU is a common chemotherapeutic agent but its success is compromised by the accompanying severe bone marrow toxicity which results in the cumulative depletion of all hematopoietic lineages (pancytopenia). It is used in high doses for the elimination of lymphomas, breast tumors, lung tumors, gastrointestinal tumors and childhood and adult glial tumors. The overall conceptual approach of eliminating the "bad guys" and leaving the innocent "bystanders" untouched is of practical application in the field of human chemotherapy. It is hypothesized that the cytotoxic effects of BCNU on bone marrow cells is mediated by the inefficient repair of the O6-chloroethyl-guanine DNA adduct which subsequently rearranges to form interstrand crosslinks, severely inhibiting DNA replication.

The levels of MGMT expressed are tissue and cell-dependent. It has been found that bone marrow cells, epithelial cells lining the colonic mucosa and brain cells all express relatively low levels of MGMT, leading to inefficient and reduced repair of alkylating lesions in these tissues. Maze and Sampson introduced a retroviral vector expressing the human MGMT gene under the control of the phosphoglycerate kinase promoter (PGK-MGMT) into the accessible mouse bone marrow stem cells. Bone marrow cells harvested from the transplanted animals displayed extensive resistance to BCNU cytotoxicity in vitro. In addition, the mice also showed a decreased mortality due to pancytopenia subsequent to BCNU treatment.

Other groups have also measured the level of MGMT in various human normal and tumor tissues. In particular, Musarrat et al. (1995) measured the levels of O6-MGMT in 6 normal breast samples (reduction mammoplasty), 17 benign breast samples and 18 breast tumor samples. The average levels in the three groups of samples were 0.221, 0.224, and 0.502 pmol/mg protein respectively. The levels obtained by another group of researchers (Grafstrom et al., 1984) for other tissues were 0.261 pmol/mg for the human normal colon, 0.217 pmol/mg for the normal human esophagus, 0.122 pmol/mg for the normal human lung, 0.076 pmol/mg for the normal human brain and 0.873 pmol/mg for the normal human liver.

Preuss et al. (1995) have also measured the levels of O6-MGMT in breast and brain tumors, primarily in an effort to determine the overall resistance of the tumor cells to the cytotoxic activities of chemotherapeutic alkylating agents. They assayed 68 breast carcinomas (ductal invasive carcinoma, lobular invasive carcinoma, medullary invasive carcinoma and recurrencies) and 38 brain tumors (pikocytic astrocytoma, low grade astrocytoma, anaplastic astrocytoma and glioblastoma multiformae) and found that the mean level in the breast tumors was 321 fmol/mg protein, which is about 40% lower than the mean level determined by Musarrat et al., (1995) (ranging from below the level of detection to 863 fmol/mg protein) In brain tumors, it was 6-fold lower at 55 fmol/mg protein (ranging from below the level of detection to 238 fmol/mg protein).

There appears to be much interindividual variation in the levels of the MGMT enzyme, perhaps suggesting a subgroup of people who are more susceptible to alkylation induced carcinogenesis. It is unknown whether the interindividual variation reflects the natural genetic diversity of the population, inducibility of the repair enzymes perhaps in individuals who are exposed to higher levels of environmental carcinogens or just the natural mosaic cellular composition of the tissues. Fritz et al. (1991), Laval et al. (1991)

and Chan et al. (1992) have all shown that MGMT gene expression is inducible since they observed that the MGMT mRNA and protein levels increased upon exposure to genotoxic stress in rat liver cells. The mechanism of MGMT inducible gene expression is unclear. It could be a lack of transcription factors activating the promoter, changes in the methylation pattern of the gene or chromosome loss or simple dilution of the protein due to extensive cell division. It would be interesting to determine if this interindividual variation is reflected in any disease pattern. For example, maybe the subset of the women who get early breast cancer have naturally lower levels of MGMT, making them more susceptible to both endogenous and exogenous DNA alkylation damage.

MGMT definitely displays tissue, cell and species specific patterns of expression. It is expressed in nearly all non-established human fibroblast cell lines but is lacking in about 20% of human tumor and various immortalized cell lines (Mex-). Day et al., (1980) have also shown that it can be lost in the course of cell transformation by tumor virus infection. It would definitely be interesting to determine if a lack of expression of MGMT or a lower than normal level of expression plays a role in human mammary carcinogenesis. Such issues as (1) do mex- cells have a selective growth advantage either by some direct or indirect mechanism and (2) are the mex- cells more vulnerable to transformation need to be adressed.

BIBLIOGRAPHY

Anderson, M.W., Reynolds, S.H., You, M. and Maronpot, R.M., Role of proto-oncogene activation in carcinogenesis, Environmental Health Perspectives (1992), 98: 13-24.

Bradshaw, R.A. and Prentis, S., (Eds) Oncogenes and growth factors, Elsevier Science Publishers (1987).

Burns, P.A., Gordon, A.J.E. and Glickman, B.W., Mutational specificity of N-methyl-N-nitrosourea in the *lacI* gene of *E.coli*, Carcinogenesis (1988), 9: 1607-1610.

Capon, D.J., Chen, E.Y., Levinson, A.D., Seeburg, P.H. and Goeddel, D.V., Complete nucleotide sequence of the T24 human bladder carcinoma oncogene and its normal homologue, Nature (1983), 302: 33-37.

Cha, R.S., Zarbl, H., Keohavong, P. and Thilly W.G., Mismatch amplification mutation assay (MAMA): Application to the c-H-ras gene, PCR Methods Appl. (1992), 2: 14-20.

Chen, J-M., Zhang, Y-P., Wang, C., Sun, Y., Fujimoto, J. and Ikenage, M., O6-methylguanine-DNA methyltransferase activity in human tumors, Carcinogenesis (1992), 13: 1503-1507.

Dr. Rita S. Cha, Ph.D. Thesis, MIT November 1992, N-methyl-N-Nitrosourea induced rat mammary tumors arise from cells harboring spontaneous oncogenic Ha-ras-1 gene mutations.

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

DuBridge, R.B., Tang, P., Hsia, H.C., Leong, P-M., Miller, J. and Calos, M.P., Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system, Molecular and Cellular Biology (1987), 7: 379-387.

Fearon, E.R., and Vogelstein, B., A genetic model for colorectal tumorigenesis, Cell (1990) 61:759-767.

Fong, L.Y.Y., Jensen, D.E. and Magee, P.N., DNA methyl-adduct dosimetry and O^6-alkylguanine-DNA alkyl transferse activity determinations in rat mammary carcinogenesis by procarbazine and N-methylnitrosourea, Carcinogenesis (1990), 11: 411-417.

Freshney, I.R. (Ed), Culture of epithelial cells, (1992) Wiley-Liss.

Gerson, S.L., Trey, J.E., Miller, K. and Berger, N.A., Comparison of O^6-alkylguanine-DNA alkyltransferase activity based on cellular DNA content in human, rat and mouse tissues, Carcinogenesis (1986), 7: 745-749.

Gray's Anatomy, 38th edition (1995), Churchill Livingstone Publishers.

Guillino, M., Pettigrew, H.M. and Grantham, F.H., Nitrosomethylurea as mammary gland carcinogen in rats, JNCI (1975), 2: 401-409.

Kirstic, R.V., Human Microscopic Anatomy, Springer-Verlag (1991).

Lu, S-J. and Archer, M., Ha-ras oncogene activation in mammary glands of N-methyl-N-nitrosourea-treated rats genetically resistant to mammary adenocarcinogenesis, Proc. Natl. Acad. Sci. (1992), 89: 1001-1005.

Maze, R., Carney, J.P., Kelley, M.R., Glassner, B.J., Williams, D.A. and Samson, L., Increasing DNA repair methyltransferase levels via boen marrow stem cell transduction rescues mice from the toxic effects of 1,3-bis(2-chloroethyl)-1-nitrosourea, a chemotherapeutic alkylating agent, Proc. Natl. Acad. Sci (1996), 93: 206-210.

Moolgavkar, S.H., Day, N.E. and Stevens, R.G., Two-stage model for carcinogenesis: Epidemiology of breast cancer in females, JNCI (1980), 65: 559-569.

Musarrat, J., Wilson, J.A., Abou-Issa, H. and Wani, A.A., O6-alkylguanine DNA alkyltransferase activity levels in normal, benign and malignant human female

breast, Biochemical and Biophysical Research Communications (1995), 208: 688-696.

Neville, M.C. and Daniel, C.W.(Eds.), The Mammary Gland: Development, Regulation, and Function. (1987) Plenum Publishing.

Preuss, I., Eberhagen, I., Haas, S., Eibl, R.H., Kaufmann, M., von Minckwitz, G. and Kaina, B., O6-methylguanine-DNA methyltransferase activity in breast and brain tumors, Int. J. Cancer (1995), 61: 321-326.

Qin, X., Zarkovic, M., Nakatsuru, Y., Arai, M., Oda, H. and Ishikawa, T., DNA adduct formation and assessment of aberrant crypt foci in vivo in the rat colon mucosa after treatment with N-methyl-N-nitrosourea, Carcinogenesis (1994), 15: 851-855.

Quintanilla, M., Brown, K., Ramsden, M. and Balmain, A., Carcinogen-specific induced mutation and amplification of Ha-ras during mouse skin carcinogenesis, Nature (1986), 322: 78-80.

Richardson, K.K., Richardson, F.C., Crosby, R.M., Swenberg, J.A. and Skopek, T.R., DNA base changes and alkylation following in vivo exposure of E. coli to N-methyl-N-nitrosourea or N-ethyl-N-nitrosourea, Proc. Natl. Acad. Sci. (1987), 84: 344-348.

Russo, J., and Russo, I.H., Biology of Disease: Biological and Molecular bases of mammary carcinogenesis, Laboratory Investigation (1987), 57: 112-137.

Sukumar, S., Notaria, V., Martin-Zanca, D. and Barbacid, M., Induction of mammary carcinomas in rats by nitroso-methylurea involves malignant activation of H-ras-1 locus by single point mutations, Nature (1983), 306: 658-661.

Vihko, R. and Apter, D., Endocrine maturation in the course of female puberty, Hormones and Breast Cancer.

You, M., Candrian, U., Maronpot, R., Stoner, G.D. and Anderson, M.W., Activation of the Ki-ras protooncogene in spontaneously occurring and chemically induced lung tumors of the strain A mouse, Proc, Natl. Acad. Sci. (1989), 86: 3070-3074.

Zarbl, H., Sukumar, S., Arthur, A., Martin-Zanca, D. and Barbacid, M., Direct mutagenesis of Ha-ras-1 oncogenes by N-nitroso-N-methylurea during initiation of mammary carcinogenesi in rats, Nature (1985), 315: 382-385.

Zhang, R., Haag, J.D. and Gould, M.N., Reduction in the frequency of activated ras oncogenes in rat mammary carcinomas with increasing N-methyl-N-nitrosourea doses or increasing prolactin levels, Cancer Research (1990), 50: 4286-4290.