2017

# NUCLEOTIDE EXCISION REPAIR: IMPACTS OF ENVIRONMENTAL CARCINOGENS AND ITS ROLE IN CANCER SUSCEPTIBILITY IN APPALACHIAN KENTUCKY

Nathaniel C. Holcomb
*University of Kentucky*, ncholc2@uky.edu
Digital Object Identifier: https://doi.org/10.13023/ETD.2017.030

NUCLEOTIDE EXCISION REPAIR: IMPACTS OF
ENVIRONMENTAL CARCINOGENS AND ITS ROLE IN
CANCER SUSCEPTIBILITY IN APPALACHIAN KENTUCKY

_____

DISSERTATION
_____

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Medicine at the University of Kentucky

By
Nathaniel Charles Holcomb

Lexington, Kentucky

Director: Dr. Isabel Mellon, Associate Professor of Toxicology and Cancer Biology

Lexington, Kentucky

2017

ABSTRACT OF DISSERTATION

NUCLEOTIDE EXCISION REPAIR: IMPACTS OF
ENVIRONMENTAL CARCINOGENS AND ITS ROLE IN
CANCER SUSCEPTIBILITY IN APPALACHAIN KENTUCKY

Lung cancer is a particularly devastating disease, accounting for the most deaths among all cancer types in the United States. Despite a reduction in the country's smoking rates, cigarette smoking remains the number one risk factor for lung cancer. Additionally arsenic exposure, which occurs primarily through contaminated drinking water in the U.S., is associated with increased lung cancer incidence. The nucleotide excision repair (NER) pathway is critical for maintenance of genomic fidelity, removing DNA lesions that could otherwise promote DNA mutations and drive carcinogenesis. Tobacco smoking introduces significant amounts of DNA damage and produces characteristic DNA mutations found in lung cancers of smokers, and arsenic increases lung cancer risk in smokers beyond the risk of smoking along. The contributions of these chemicals to DNA damage and cancer have been well documented, but few studies have examined their effects on DNA repair pathways, particularly the nucleotide excision repair (NER) pathway. Arsenic, while not directly mutagenic, promotes the carcinogenicity of other compounds including agents that produce DNA damage that is repaired by the NER pathway. In this dissertation I investigated the effects of cigarette smoke condensate (CSC, a whole-smoke tobacco surrogate) and arsenic on NER. I observed that CSC or arsenic treatment inhibited NER as measured by a slot-blot assay using UV-induced photolesions as model substrates to measure NER. The abundance of Xeroderma Pigmentosum complementation group C (XPC), a critical NER protein, was significantly reduced in all lines treated with either chemical, while XPA protein was unaffected. CSC and arsenic also affected RNA levels of certain NER genes. Finally, proteasome-regulated XPC turnover was affected by CSC and arsenic treatment, suggesting a potential mechanism for XPC protein inhibition. The observed impairment of NER by CSC is critically important in tobacco cancer etiology – CSC introduces DNA damage, some of which is repaired exclusively by NER, and CSC inhibits the NER pathway as well, providing a two-sided assault on cellular genetic fidelity. I then adapted the NER assay to measure repair in lymphocytes isolated from human subjects

of a study investigating the high incidence of lung cancer in Appalachian Kentucky.  I observed an age-dependent decline in NER efficiency that was modulated by subject smoking status and a reduced NER efficiency among current smokers in the lung cancer patient population compared to control subjects in the youngest age group, suggesting individual DNA repair capacity measured with this repair assay may be a biomarker for lung cancer susceptibility.

_____Nathaniel Holcomb_____

_____2-22-17_____
(Date)

NUCLEOTIDE EXCISION REPAIR: IMPACTS OF
ENVIRONMENTAL CARCINOGENS AND ITS ROLE IN
CANCER SUSCEPTIBILITY IN APPALACHAIN KENTUCKY


By

Nathaniel Charles Holcomb


<div align="right">

_____Isabel Mellon_____
Director of Dissertation

_____Isabel Mellon_____
Director of Graduate Studies

_____2-22-2017_____
(Date)

</div>

**Acknowledgements**

Kentucky has been my home my entire life; I have never lived outside its borders. I have grown up and lived in predominantly urban settings in an otherwise rural state, but the health issues that plague our state are still visible, particularly the state's smoking and obesity problems. My fascination with science and my love of my home state have guided my career ambitions and have lead me to the place I currently reside, in a position to help illuminate a crucial problem we as a Commonwealth face. Kentuckians live at a convergence of high smoking, poor health, low income, and potentially dangerous environmental pollutants, which have together contributed to the nation's worst rate of lung cancer incidence and mortality. Something must be done to address this problem, as lung cancer is easily the most deadly cancer that we face which comes as a result of very preventable carcinogenic exposures.

The large portion of cancer research surrounds detection and treatment, trying to reduce the impact the disease has on the population. My research has focused on cancer prevention through increased understanding of causes of cancer risk that can be used as educational tools for healthcare providers and the general population. We have looked at cancer prevention through the lens of DNA repair; factors that diminish DNA repair stochastically increase cancer risk, and therefore understanding what can impact DNA repair can lead to a greater breadth of knowledge of how to mitigate the risk of cancer development.

I work in Dr. Isabel Mellon's lab in the Department of Toxicology and Cancer Biology at the University of Kentucky's College of Medicine. Dr. Mellon is a renowned researcher in the field of nucleotide excision repair (NER), and she was part of Dr. Hanawalt's research team at Stanford credited with the discovery of a sub-pathway in NER, transcription-coupled NER (TC-NER). It has been under her remarkable guidance and understanding that I have been able to pursue my research interests. I would like to thank her for wonderful mentorship and for challenging me to critically evaluate the results of my experiments and teaching me how to truly understand the scientific method. I would not be the scientist I am today without her.

I would like to thank my amazing coworker Dr. Mamta Goswami who has been indispensable in my graduate research. Much of what I've learned in the lab has been from her instruction, and most of the research in this paper was done under her tutelage in some form or fashion. She truly has been an essential component in my growth as a scientist. Dr. Sung Gu Han also provided much of the framework upon which my cell culture experiments were based, and he has my thanks for that.

My committee members, Drs. Orren, Rangnekar, Huang, Shelton and former members Drs. Gairola and Kaetzel have been instrumental in my progress as a scientist, and I have nothing but the greatest respect for them.

A wonderful collaboration was established with several researchers and Markey Cancer staff in order to perform the human population study I am

TABLE OF CONTENTS

**Chapter One:** Introduction

**Chapter Two:** Experimental Methods

**Chapter Three:** *Exposure of human lung cells to tobacco smoke condensate inhibits the nucleotide excision repair pathway*

**Chapter Four:** *Trivalent inorganic arsenic inhibits the nucleotide excision repair pathway and impairs expression of XPC protein.*

**Chapter Five:** *Establishment of a reproducible measure of individual NER efficiency in peripheral blood mononuclear cells*

**Chapter Six:** *Variations in nucleotide excision repair efficiency in a human population*

**Chapter 7:** *Unanswered questions, a case-control study, and a future direction for the NER functional assay*

List of Figures

List of Tables

# Chapter 1 - Introduction

## 1.1 Cancer in the United States

Cancer is defined by the National Cancer Institute as a general term for diseases in which abnormal cells divide uncontrollably and possess the potential to invade nearby tissues or remote locations in the body. While several genes such as *K-Ras* and p53 are frequently mutated in cancer [1, 2], cancer is not the result of a singular genetic defect, but rather cancers arise through a multitude of DNA mutations that may be innate or acquired over a lifetime of exposure to DNA damaging agents. Several cellular processes must be altered for a normal cell to undergo cellular transformation into a cancer cell, and the types of changes required for transformation have been well documented as hallmarks of cancer development. These hallmarks initially included unlimited replicative potential, evasion of apoptosis, self-sufficiency with growth factors, resistance to growth inhibition, evasion of growth suppression, and angiogenesis [3]. More recently, additional hallmarks - avoiding host immune system and deregulating cellular energetics - as well as the enabling characteristics of inflammation and genomic instability or DNA mutation were added as contributors to the pathogenesis of many, if not all, cancers [4]. One of the enabling hallmarks of cancer, genomic instability, is a condition of high mutation frequency within the genome, and can be the result of DNA repair pathways not functioning properly. DNA is under significant pressure to repair the damage it is constantly subject to, as failure to do so may lead to an increase in DNA mutation frequency which can

1

promote (i.e. enable) all the cellular changes listed above and thus drive carcinogenesis.

In the United States, cancer is responsible for the second most number of deaths among all causes, producing more annual deaths than the next five categories combined (http://www.cdc.gov/nchs/fastats/leading-causes-of-death.htm 2014 data). Only cardiovascular disease ranks higher than cancer.  Cancer is a disease of aging - the risk of developing cancer increases in both genders across all sites as individuals grow older [5]. The death rate from cancer has increased dramatically since the beginning of the last century [6], a result of increased longevity and dramatic declines in the conditions that previously produced high levels of mortality. From 1900 to 2010, the average U.S. life expectancy increased from 47.3 to 78.7 years [7] , and during the same period of time cancer went from the eighth leading cause of death to the second. In 1900, conditions like pneumonia, tuberculosis, and GI infections were the greatest causes of death [6], but these conditions have all but been eliminated as causes of death in the United States today, contributing to the increase in life expectancy.  Indeed, the death rate produced by the top ten leading causes of death has dropped by almost 50% between 1900 and 2010. However, cancer and heart disease account for more deaths now than over a century ago.  In fact, cancer mortality rates have tripled in that time frame [6], and cancer has consistently been the second leading cause of death behind heart disease for the last 75 years [8]. Cancer incidence and death rates have dropped slightly in recent years from a peak in the early 1990's [5], but cancer still remains a significant health risk.

Another contribution to the rise of cancer incidence and mortality in the U.S. over this time period is that the average age at which individuals are diagnosed with cancer across all types is 65 (http://seer.cancer.gov/statfacts/html/all.html). This is almost twenty years greater than U.S. life expectancy was at the beginning of last century [7].  Since our population is living longer, and since cancer rates increase with increasing age, more people now are at risk of developing cancer than ever before.  In recent years there has been a stalling of life expectancy growth  [9, 10] and a reduction in new cancer diagnoses, but there will still be an estimated 1.7 million new cancer diagnoses and about 600,000 cancer-related deaths in 2016.  Over 14 million people are living with cancer in the United States as of 2013 (http://seer.cancer.gov/statfacts/html/all.html).

## 1.1.1 Lung Cancer in the U.S. and Kentucky

Among the different types of cancer lung cancer produces the highest number of deaths in the U.S., three times more than the second leading cause of cancer mortality, colon cancer [5].  The five year survival curve for individuals diagnosed with lung cancer is 17.7% (2006-2012 data, http://seer.cancer.gov/statfacts/html/lungb.html).  This low overall survival is due in part to the average stage at diagnosis for lung cancer.  Over half of all lung cancers are stage three or worse upon first diagnosis [5].  Five year survival for patients with stage three lung cancer is under 5%.  In contrast, stage one lung cancers, while only 16% of diagnosed lung cancers, have a five year survival of over 50%.  This stresses the importance of early detection in lung cancers.  The fact that lung cancers are diagnosed at late stages more frequently than early stages

3

contributes to a chilling statistic.  In the United States, the median age for lung

cancer diagnoses is 70, and the average age at which those with lung cancer die

is only 72 (http://seer.cancer.gov/statfacts/html/lungb.html).

Lung cancer is responsible for 27.7% of all estimated cancer deaths in the

United States in 2015 and 34.8% of the estimated cancer-related deaths in

Kentucky [5].  While the national age-adjusted rate of lung cancer incidence is

62.74 per 100,000 individuals (based on 2009-2013 data; http://www.cancer-

rates.info/naaccr/), Kentucky's rate is 50% higher than the national average, with 96.4

lung and bronchus cancer diagnoses per 100,000 individuals (based on 2009-

2013 data) (Figure 1.1).  Lung cancer is not the only cancer that is causing health

issues in Kentucky; the state has the highest age adjusted overall cancer

incidence from 2009-2013 (http://www.cancer-rates.info/naaccr/) (Figure 1.2) and the

highest over-all cancer death rate among all states [5]

## 1.1.2 Smoking and lung cancer in Kentucky

The primary risk factor in the United States for developing lung cancer is

cigarette smoking [11].  According to the Centers for Disease Control and

Prevention, 25.9% of Kentucky adults are current smokers (every day or some

days) (https://www.cdc.gov/tobacco/data_statistics/fact_sheets/adult_data/cig_smoking/).  Kentucky's

percentage of current smokers was 10% higher than the national average of

15.1% in 2015 and ranks first among all states for tobacco smoking incidence.

Tobacco smoking is undoubtedly the principle risk factor for the development of

lung cancer in Kentucky.  However, within the state there exists a geographical

disparity of lung cancer incidences. The eastern portion of Kentucky lies within

the multi-state region of the United States called Appalachia.  In Kentucky, the

Appalachian region has a higher lung cancer incidence than in the rest of the

state (http://cancer-rates.info/ky/index.php), likely due to the high rate of tobacco use,

which is greater than the rate in non-Appalachian Kentucky [12] and 60-80%

higher than the national average.  When viewed on a county-by-county basis, the

high lung cancer incidence counties cluster in the eastern portion of the state, the

Appalachian region (Figure 1.2a).  By dividing the state into "Non-Appalachia"

and "Appalachia" counties, it becomes clear that average lung cancer incidence

of counties in the Appalachian region is over 20% higher than of those in the

non-Appalachian region (Figure 1.2b).  Within the greater Appalachian region of

the U.S., a higher cancer burden is observed across all sites compared to non-

Appalachian areas in the United States, particularly with regard to tobacco-

related cancers [13].

While lung cancer rates in Appalachian Kentucky are primarily a result of

tobacco smoking rates that are over 50% higher than the national average and

greater than the rates in the rest of the state [12], there are clusters of counties in

Kentucky with lung cancer rates that are higher than the rest of the state after

controlling for gender, age, and smoking status [14], suggesting that additional

factors may be contributing to the high lung cancer incidence (Figure 1.3).  Two

of the three high lung cancer incidence clusters reside within Appalachian

Kentucky, while a third cluster contains many counties that reside within a large

metropolitan area and lung cancer rates may be high as a result of increased

pollution associated with such a region. Additionally, a multi-state study of

Appalachian United States found that the increased lung cancer rates in coal mining areas within the region could not be explained by tobacco use alone [15], and the high lung cancer incidence clusters in Appalachian Kentucky are also areas of high coal production [14]. These studies open the possibility that some as yet undetermined factors are contributing to increased lung cancer incidence in Appalachian Kentucky and perhaps the greater Appalachian region.

## 1.2 DNA Repair

DNA is a highly reactive, massively large molecule (a single human diploid cell contains almost seven billion nucleotides) that cannot be replaced if damaged. Instead, DNA must be repaired in response to the persistent damage it is subject to, or otherwise the damage may become permanent, which is the definition of a DNA mutation. There are several types of DNA damaging agents including exogenous agents (such as UV light and tobacco smoke) and endogenous byproducts of normal cellular metabolism. DNA can also undergo spontaneous nucleotide hydrolysis (producing abasic sites) and spontaneous deamination (which can produce base changes) [16, 17]. The reactivity of DNA is due in large part to the unsaturated ring structure of the DNA bases, producing a nucleophilic environment subject to reactions with electrophiles [18]. Additionally, the negatively charged phosphodiester backbone attracts positively charged ions capable of reacting with DNA, including metal ions that can undergo proximal redox cycles (Fenton chemistry) to generate reactive oxygen species. The typical result of DNA damage is a local disruption of standard

duplex structure, which forms the basis for recognition by DNA repair proteins across multiple DNA repair pathways.

Humans possess three highly efficient DNA excision-repair pathways to preserve genomic integrity.  These pathways are nucleotide excision repair (NER), base excision repair (BER), and mismatch repair (MMR), and they are genetically conserved processes in eukaryotes whose disruptions often lead to increased genomic instability and ultimately increased cancer incidence.  In addition to these, there are pathways involved in repairing double-strand breaks, a type of damage that excision repair cannot address as a suitable template does not exist.  All of these processes are necessary to ensure proper genomic maintenance and cellular survival.

The excision repair pathways are discussed in more detail below. Combined they respond to the majority of DNA damage cells face from exogenous and endogenous sources.  It should be noted that while all three pathways are critical and none can be fully complemented by the other two, NER is the sole focus of my dissertation.  The experiments laid out herein involve environmental factors that may promote carcinogenesis by producing lesions recognized by NER, interfering with normal NER function, or both.  While all three excision repair pathways are important for preventing DNA mutations that can promote carcinogenesis as I will describe below, the MMR and BER pathways were not the subjects of our research, and are included here to highlight the fact that mutations and polymorphisms in genes involved in all three excision repair pathways can increase cancer risk.

## 1.2.1 NER

NER is a versatile pathway that removes a wide variety of structurally diverse DNA lesions including those generated by metabolites of chemical carcinogens such as polycyclic aromatic hydrocarbons (PAHs) in tobacco smoke as well as cyclobutane pyrimidine dimers (CPDs) and 6-4 photoproducts (6-4PPs) induced by exposure to UV light [19]. In mammals, at least 20 different protein factors participate in NER, including the XPA-G factors that are singly defective in the 7 corresponding complementation groups of the human disease Xeroderma Pigmentosum (XP). NER efficiency is also regulated by the tumor suppressor factor p53 through transcriptional regulation of the *XPC* and *DDB2* gene products in response to DNA damage [20-23]. The NER pathway, also referred to as global genome nucleotide excision repair (GG-NER), is responsible for removing DNA damage from nontranscribed DNA. Damaged present in the transcribed strands of active genes are removed by transcription-coupled nucleotide excision repair (TC-NER) [24].

The following is an abridged explanation of the steps in the NER pathway (reviewed quite well in [19, 25]). The NER pathway is initiated when a helically distorting DNA adduct forms (Figure 1.4A). The manner of recognition is different between two sub-pathways. In GG-NER the DNA recognition protein is XPC [26]. In order to be able to respond to the multitude of DNA adducts that are substrates for NER, XPC, stabilized by its obligate binding partner HR23B, recognizes the helical distortion in DNA produced by an adduct [27] and binds to the undamaged strand opposite the lesion (Figure 1.4B). Recent evidence

suggests that the rate limiting step in NER is the efficiency with which DNA damage can be recognized, implicating XPC as the rate limiting enzyme in GG-NER [28-30]. As TC-NER functions exclusively on the transcribed strand of actively transcribed DNA, damage recognition during TC-NER is initiated when the RNA polymerase II complex is stalled by the DNA adduct, which stabilizes the CBS protein that is present in the complex and promotes recruitment of the subsequent NER machinery by CSB [24, 31, 32].

After damage recognition, the two pathways converge for the remainder of the repair process. Damage recognition proteins (XPC in GG-NER and CSB in TC-NER) recruit the TFIIH transcriptional complex containing the helicases XPB and XPD to the site of the damage. XPB pries open the DNA and permits XPD to travel across the damaged strand, unwinding DNA until it stalls at the damaged nucleotide (Figure 1.4C) [33, 34]. The unwinding permits subsequent NER factors access to the DNA damage site, including the endonuclease XPG (which is recruited through interactions with TFIIH), the single strand binding protein RPA, and XPA (Figure 1.4D). XPA is an indispensable component of the NER pathway, but the role it plays is still somewhat vague. XPA can interact with a multitude of NER factors including XPC, THIIH, RPA, XPF/ERCC1, and more. It may serve to verify that all necessary factors are in place ahead of the DNA incision step. The other directional endonuclease complex, XPF/ERCC1, is recruited to the damage site through interactions with XPA [35]. After recruitment of the XPF/ERCC1 complex, it cuts 5' of the damaged nucleotide (Figure 1.4E) This 5' incision produces a 3' OH group that serves as a template on which a

polymerase can bind and begin synthesis of DNA to replace the damaged strand. After synthesis has begun, XPG performs the second single-stranded cut 3' of the damaged site, releasing the damage-containing oligonucleotide (Figure 1.4F). The original integrity of the DNA is restored when the repair polymerase synthesizes an approximately 30 nucleotide segment of DNA and a DNA ligase seals the strand (Figure 1.4G) [36]. A recent investigation into the post-incision steps in NER reveals that the process is more complicated than previously thought. While DNA polymerases δ and ε were initially thought to be the principle polymerases in the gap filling step, the error prone polymerase κ also plays a role in repair synthesis [37]. Further complicating the post-incision process, the ligase used to seal the nicked DNA after repair synthesis can be either DNA ligase I or IIIα, and which one is used may depend on the replicative status of the cell [38].

There are three primary inherited disorders resulting from mutations in NER genes in humans; xeroderma pigmentosum (XP), Cockayne syndrome (CS), and trichothiodystrophy (TTD) [39], although there are additional syndromes stemming from NER deficiencies, depending on the functional consequences of mutation in the repair protein [40, 41]. Of the primary disorders, XP is characterized by extreme photosensitivity and a dramatically elevated incidence of skin cancer, as well as an increased internal cancer mortality [40]. The average age at diagnosis of nonmelanoma skin cancer for individuals with XP is 50 years sooner than the general population [40, 42].

While the disorder is rare (as infrequent as an estimated 1 in 1,000,000 in the U.S.), there are regions in the world where XP frequency is considerably higher. In Japan, for example, the frequency of XP-A in the general population is about 1 in 22,000 [43]. This high incidence also produces a carrier frequency of XP-A in Japan that is thought to be as high as 1 in 100. The carriers of XP-A do not exhibit repair deficiencies, as measured by an unscheduled DNA synthesis (UDS) assay [44], nor do XP-A heterozygote mice show any increased incidence of tumor development compared to wildtype mice following UV exposure [45]. However, there is some evidence to suggest that other XP heterozygotes may have an intermediate repair phenotype. XP-C heterozygotes have reduced XPC RNA levels compared to genotypically normal individuals [46], and reduced expression of XPC has been associated with increased risk for developing certain head and neck cancers [47], suggesting that certain XP carriers may be at an increased risk for cancer development.

Polymorphisms in NER genes have also been explored for their potential to contribute to cancer risk, which will be discussed in more detail in Chapter 5. Exploring the health consequences of NER gene heterozygosity or NER gene polymorphisms is part of a process of determining whether these variations exist in a population and can be used as biomarkers to predict cancer risk for individuals who may have suboptimal function of their nucleotide excision repair pathway.

**1.2.2 MMR**

The MMR pathway in eukaryotes is responsible for a 50-100 fold increase in DNA fidelity [48] by identifying single base mismatches or small insertions/deletions that result from improper DNA replication.  The pathway is remarkably similar across eukaryotes and prokaryotes, although it is modified in E. coli.  In bacteria, a homodimer of the MutS protein recognizes mismatched bases or small insertions/deletions [48].  This process is similar in eukaryotes, with three MutS homologs (MSH2, MSH3, and MSH6) forming the heterodimeric complexes MutSα (MSH2/MSH6) and MutSβ (MSH2/MSH3). MutSα recognizes single base mismatches and small insertion/deletion loops (IDLs) while MutSβ recognizes larger IDLs [49].

After recognition of the mismatch/IDL, a homodimer of the MutL protein is recruited in bacteria to nick the nascent strand near the site of the mismatch.  In eukaryotes, three heterodimer MutL homologs , MutLα, MutLβ, and MutLγ are formed by complexing MLH1 with PMS2, PMS1, and MLH3 respectively [50].  Of these, MutLα is believed to be responsible for the majority of MMR in eukaryotes [51].  Upon strand incision, an exonuclease (EXO1/UvrD in eukaryotes/prokaryotes) is recruited to remove the strand with the replicative error, a polymerase (Polδ/Pol III) fills in the gap, and a ligase seals the nick, completing the repair process.

Human homologs of the bacterial MutL/S proteins are mutated in a number of cancers, including cancers arising from Lynch syndrome, an inherited cancer syndrome resulting from a germline mutation in one of the MMR genes

[52], primarily the MLH1 or MSH2 genes.  This is not surprising given MSH2 is

present in both MutS complexes, and MLH1 is present in all three MutL

complexes, and thus losing either of these proteins abolishes that component of

the MMR pathway entirely.  The inherited mutation occurs in only one allele of

the MMR gene, so a second somatic mutation is required, which is why the

disorder manifests later in life, as opposed to the inherited NER disease XP.

Only about 5% of colorectal cancers are a result of Lynch syndrome, but the

disease also contributes to other cancers including endometrial, ovarian, and

stomach cancers.  Of all hereditary nonpolyposis colon cancers, upwards of 90%

are a result of a mutation or epigenetic silencing of a MMR gene [53].  There is

also mounting evidence that, similar to what has been reported with NER genes,

polymorphisms in MMR genes may also contribute to increased cancer risk [54-

56]

### 1.2.3 BER

The BER pathway is responsible for detecting and eliminating a variety of

DNA lesions targeted specifically to the purine and pyrimidine bases that

compose the inner portion of the DNA complex.  Lesion recognition in the BER

pathway is achieved through a number of glycosylases which recognize specific

base modifications or abnormal base pairings.  After recognition, the glycosylase

removes the modified base by cleavage of the N-glyosidic bond, but leaves the

sugar moiety intact, producing an abasic site (often referred to as an AP site for

apurinic/apyrimidinic).  The presence of an AP site recruits APE-1 (AP

endonuclease 1) to cut the sugar-phosphate backbone at the abasic site's 5' end,

producing a 3' hydroxyl site. Once the DNA backbone is cleaved, DNA Pol β can use the 3' hydroxyl to resynthesize the missing nucleotide. Cleavage of the extra phosphate group (also by Pol β) and ligation of the DNA backbone (XRCC-1/DNA ligase IV) finish the repair process. When only one nucleotide is removed in this fashion, as is generally the case, the repair pathway is referred to as "short patch BER." Occasionally more nucleotides than the one recognized by the glycosylase are removed in a process known as "long patch BER."

Germline mutations in various glycosylases involved in BER have been associated with increased cancer risk. A homozygous mutation in the NTHL1 glycosylase is associated with increased colorectal cancer [57]. A compound heterozygous mutation of the glycosylase MYH, the human homolog of bacterial mutY, has been attributed to an inherited cancer disorder MAP (MUTYH associated polyposis) [58, 59].

BER proteins downstream of the glycosylase are not traditionally inactivated in human cancers. The remaining components of the core BER pathway (APE-1, Pol-β, FEN1, and DNA-ligase) are critical for life – knocking them out causes embryonic lethality [60], and as such a cell acquiring a mutation that inhibits any of these proteins is much more likely to die than to undergo carcinogenesis. However, there is considerable evidence that polymorphisms in certain BER genes modulate cancer predisposition. Abnormal estrogen metabolism can lead to an accumulation of depurinating adducts, leading to formation of AP sites [61]. Variation in BER efficacy due to genetic polymorphisms may affect the repair of these AP sites and perhaps contribute to

14

an increase in mutagenesis and ultimately breast cancer risk [62]. Multiple APE-1 polymorphisms have been associated with increased breast cancer development [63].

## 1.2.4 Concluding remarks on DNA repair

As I have described above, all three excision repair pathways - NER, MMR, and BER - are vital for maintenance of genomic fidelity. A failure in any of these pathways can lead to corruption of the genetic code through induction of DNA mutations, producing an increase in cancer risk. There are several ways in which a repair pathway's efficiency can be reduced. Polymorphisms producing semi-functional proteins can have negative effects on repair efficiency, as can expression level variations like those seen in some heterozygotes, or those seen when transcript levels of a repair gene are reduced. Variations expression of repair proteins and polymorphisms in repair genes can and have been measured in general populations as indicators of individual relative risk of developing various types of cancers, which will be discussed in more detail in chapters 5 and 6. By their nature, expression studies and polymorphism studies are limited to the genes/proteins of interest. Repair pathways, NER in particular, employ dozens of proteins to respond to DNA damage, and even more proteins are involved in regulating repair protein stability through post-translational modifications and expression levels through transcription regulation. As a result, these experiments may not measure all the factors that contribute to an effective repair pathway's performance. Instead, an assay that measures the functional efficiency of a repair pathway would be a more effective indicator of individual

variations in repair efficiency and may be the most useful measure of an individual's cancer development risk.

## 1.3 Tobacco smoking and DNA damage

It has been over 50 years since the initial publication of the Surgeon General's first report on the health consequences of cigarette smoking with an emphasis on lung cancer, and since that reporting tobacco smoking rates have been on the decline in the United States [11].  While adult smoking rates are down from over 40% when the report was published to about 15% today, tobacco use continues to be a considerable health hazard in the United States, with an estimated 36.5 million people still smoking in 2015 [64].  Tobacco smoking remains the number one risk factor for lung cancer development despite the drop in rates, with smoking contributing to approximately 80-90% of lung cancers [11]. Tobacco smoke contains over 70 known human carcinogens [65, 66] including a number of polycyclic aromatic hydrocarbons (PAH) [67]. Perhaps the most well researched of the PAH chemicals in tobacco smoke is benzo[a]pyrene (B[a]P). Through a three stage process involving cytochrome p450s (usually CYP1A1 or CYP1B1) and microsomal epoxide hydrolase, B[a]P is metabolized by into one of many isoforms of benzo[a]pyrene-7-8,dihydrodiol-9,10-epoxide (BPDE) [68, 69]. BPDE is a DNA reactive compound that covalently binds to guanines at the N2 position [70-72] and causes predominantly G to T transversion mutations if left unrepaired [73].  While the BPDE adduct is a substrate for the nucleotide excision repair pathway [74, 75], failure to remove the adduct leads to increased mutation formation which drives carcinogenesis.

The relationship between NER and B[a]P mutagenesis is dependent on a cellular response that occurs when NER fails to remove the adduct. When NER is unable to process the lesion before DNA replication occurs, a low fidelity lesion-bypass polymerase is recruited to synthesize DNA across from the adducted guanine [76]. In yeast, the bypass polymerase is able to correctly insert cytosine about 75% of the time, but 14% of all insertions across from (+)-*trans*-*anti*-BPDE-N$^2$-dG, the predominant stereoisomer of BPDE [71], were adenine [76]. This mismatch, upon another round of replication, would produce a G-T transversion mutation in one of the daughter cells. An increase in DNA damage burden or a decrease in NER function could result in an increased persistence of the DNA adduct, increasing the need for the lower fidelity replicative polymerase and thereby increasing the rate of mutation formations.

The significance of this one carcinogen out of the scores present in tobacco smoke can be seen in the mutational signature of the *p53* gene in lung cancers from tobacco smokers. The tumor suppressor gene *p53* is mutated in around 60% of human lung cancers [77]. BPDE adducts produce mutational hotspots in P53 *in vitro* that are at the same major codons mutated in human lung cancers (including codons 157, 158, 245, 248, and 273) [78, 79], suggesting that P53 mutations observed in lung cancers are produced at least in part by unrepaired BPDE adducts. Our current understanding as to why BPDE lesions are found at these hotspots relates to the presence of CpG islands in the P53 gene. It has been observed that *p53* is methylated at all CpG dinucleotides within the exon 5-8 coding region, the same location of the mutational hotspots

mentioned above [80]. These methylated CpGs are preferential targets for BPDE adduct formation [81], providing further evidence that connects tobacco smoking, via BPDE adduct generation, to the P53 mutational spectra observed in lung cancers of smokers. Therefore, the relationship between tobacco smoking and lung cancer involves DNA damage produced by tobacco smoke that the NER pathway has not repaired, resulting in damage tolerance mechanisms of certain lesion bypass polymerases that increase the potential for mutagenesis.

## 1.4 Another factor for lung cancer in Kentucky - Arsenic

There are of course many more risk factors for the development of lung cancer. One such risk factor is arsenic exposure, which is associated with increases in cancer risk at several organ sites, including the lung [82, 83]. The relationship between arsenic exposure and cancer risk will be discussed in greater detail in chapter 4. In the United States there are several routes of arsenic exposure, including contaminated drinking water, occupational exposure, and even as a component in tobacco smoke. Arsenic exposure increases the risk for developing lung cancer, among many others. Interestingly, arsenic is not a direct mutagen [84], and the mechanism most commonly proposed for its carcinogenicity is a co-carcinogenic function, increasing the carcinogenic potential of other agents.

As mentioned above, there is an increased incidence of lung cancer in Appalachian Kentucky relative to the rest of the state that cannot be fully explained by a higher incidence smoking. There are many possible explanations for this disparity. The high incidence could be a consequence of higher smoking

quantity (the amount the smokers in the region smoke), exposure to other cancer causing agents such as radon or heavy metals, a genetic predisposition to cancer development that is over-represented in the region, or some combination of these and other factors. One possible explanation for the increased incidence could be related to arsenic exposure in the region. High levels of arsenic have been observed in well waters central Appalachia, including Eastern Kentucky [85], as well as in shale outcroppings in the region [86]. Appalachian Kentucky contains high levels of arsenic found in coal deposits [87-89]. A process of coal mining in the region known as mountaintop removal may introduce arsenic into local water systems [90]. This arsenic could find its way into unregulated private water wells, which has the potential to produce locally high arsenic exposure. Indeed, areas in which coal mining has been associated with high incidence of lung cancer include the greater Appalachian region [15] as well as a southeastern region of Kentucky within Appalachia [14], suggesting that arsenic, or some other component of Appalachian coal, may contribute to lung cancer rates in the region.

Mechanistically, several hypotheses have been put forth to explain arsenic carcinogenesis including inhibition of DNA repair pathways, promoting the carcinogenesis of other compounds, and this will be discussed in greater extent in chapter 4. As tobacco smoke is the greatest risk factor in lung cancer development, and since it is well established that tobacco smoke generates DNA damage repaired by NER that contributes to lung cancer risk in smokers, if arsenic was able to inhibit of the repair of tobacco smoke-induced DNA damage

(via an inhibition of the NER pathway), that would exacerbate the damage, increase its mutagenic potential, and ultimately increase lung cancer risk.

## 1.5 Research Hypothesis

My research centers on the impact of environmental factors on the nucleotide excision repair pathway and the consequences of variations in NER efficiency with regards to cancer risk, particularly lung cancer risk in Appalachian Kentucky.  Tobacco smoking is undoubtedly the primary factor in the development of lung cancer.  Tobacco smoke produces DNA damage that is eliminated by the nucleotide excision repair pathway.  A reduction in the efficiency of the NER pathway could result in a reduction in the repair of DNA damage caused by tobacco smoke and an increase in the risk of developing lung cancer.  Based on those observations, I propose that tobacco smoke condensate, used as a surrogate for whole tobacco smoke, inhibits NER function, which may be an additional mechanism through which it promotes carcinogenesis.  Chapter 3 will address this hypothesis using lung cells in a cell culture model.  As mentioned above, another lung carcinogen, arsenic, increases the carcinogenicity of other DNA damaging agents. I propose that arsenic inhibits NER function, which may be one mechanism of the carcinogenesis of arsenic. Chapter 4 will address this hypothesis using lung and skin cells in a cell culture model

The results of the cell culture studies coupled with the high lung cancer incidence in eastern Kentucky prompted an epidemiological study to address the following hypothesis: the increased lung cancer incidence in Appalachian

Kentucky is due in part to an impaired NER efficiency of individuals in the region, which may be a consequence of environmental exposure from agents such as tobacco smoke and arsenic.  In other words, we hypothesize that individuals with lung cancer in Appalachian Kentucky have reduced NER efficiency relative to individuals without lung cancer in the region.  To address this hypothesis, a DNA repair assay was adapted to measure the efficiency of NER in a human population, and the process of confirming the reliability of this assay is presented in chapter 5, along with the effect of experimental variables on the measurement of NER efficiency in the study population.  Chapter 6 focuses on the population demographic factors that impacted the measurement of the efficiency of NER in a study population.  Lastly, in chapter 7 I address unanswered questions from the previous chapters, and I present the findings of a comparison of the efficiency of NER between lung cancer cases and healthy controls. Additionally, I discuss how this study will help establish guidelines for proper case-control comparisons using the NER repair assay in the future.

**Figure 1.1 Lung cancer and total cancer incidence in the United States.**

Age-Adjusted Invasive Cancer Incidence Rates in the United States
Lung and Bronchus, 2009-2013
By State

**A**

Age-Adjusted to the 2000 U.S. Standard Million Population
United States Rate 62.74 / per 100,000

- 28.80 - 58.71
- 58.97 - 63.09
- 64.59 - 70.72
- 71.24 - 96.41
- Data unavailable

Kentucky: 96.41

Age-Adjusted Invasive Cancer Incidence Rates in the United States
All Sites, 2009-2013
By State

**B**

Age-Adjusted to the 2000 U.S. Standard Million Population
United States Rate 457.00 / per 100,000

- 392.49 - 446.35
- 447.53 - 459.91
- 464.22 - 478.59
- 480.71 - 520.88
- Data unavailable

Kentucky: 520.88

http://www.cancer-rates.info/naaccr/

**Figure 1.1. Lung cancer and total cancer incidence in the United States.** (A)

Map of lung cancer incidence in the United States. Lung cancer rates for each of

the 50 states are shown. Dark red states indicate a high incidence of lung

cancer. Kentucky has the highest age-adjusted incidence of lung cancer in the

country at 96.41 cases per 100,000 people, over 50% higher than the national

average of 62.74. (B) Map of cancer incidence in the United States. Aggregate

cancer rates at all sites for each of the 50 states are shown. Dark red states

indicate a high incidence of lung cancer. Kentucky has the highest age-adjusted

incidence of cancer at all sites in the country at 520.88 cases per 100,000

people, compared to the national average of 457. Data is from 2009-2013, and

is from the NAACCR (http://www.cancer-rates.info/naaccr/)

**Figure 1.2 Variation in lung cancer frequency within Kentucky.**



Age-Adjusted Invasive Cancer Incidence Rates in Kentucky
Lung and Bronchus, 2009-2013
**A**   By County
Age-Adjusted to the 2000 U.S. Standard Million Population
Kentucky Rate 96.4 / per 100,000

68.2 - 90.3
90.4 - 100.2
100.8 - 113.0
113.1 - 157.5
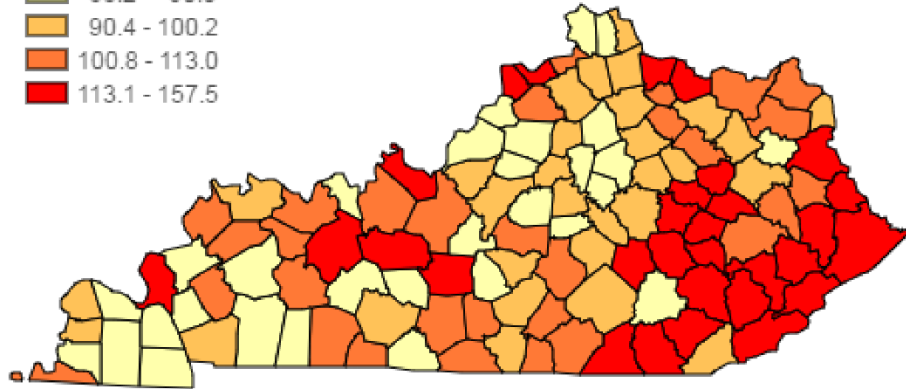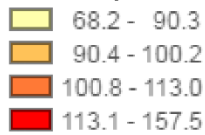


Age-Adjusted Invasive Cancer Incidence Rates in Kentucky
Lung and Bronchus, 2009-2013
**B**   By Appalachian Region
Age-Adjusted to the 2000 U.S. Standard Million Population
Kentucky Rate 96.4 / per 100,000

91.2
109.2
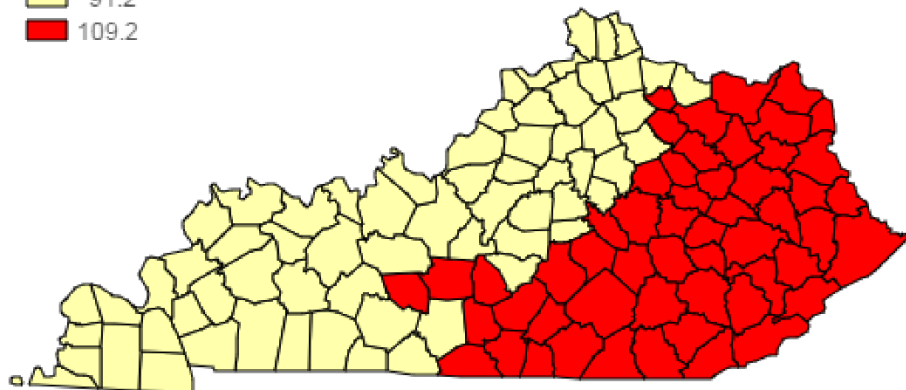
All rates per 100,000.
Data accessed September 29, 2016. Based on data released Jan 2016.
© 2016 Kentucky Cancer Registry.

http://cancer-rates.info/ky/index.php

**Figure 1.2. Variation in lung cancer frequency within Kentucky.** (A) Lung

cancer rates for each of the 120 counties within Kentucky are shown. Dark red

counties indicate a high incidence of lung cancer. The high incidence counties

are present throughout the state but are at the greatest density in the far-eastern

region, the area known as Appalachia. (B) A breakdown of cancer incidence in

Kentucky between counties designated as Appalachian or non-Appalachian

shows the burden of lung cancer is considerably higher (by about 20%) in

Appalachian Kentucky as compared with the rest of the state.

# Figure 1.3. Clusters of high lung cancer incidence in Kentucky.



Crude lung cancer incidence per 100,000, 1995-2007
- 66.6 - 89.0
- 91.7 - 107.9
- 109.6 - 126.3
- 127.8 - 160.1

Cluster 3
RR=1.04
p=0.01

Cluster 1
RR=1.21
p<0.01

Cluster 2
RR=1.17
p<0.01
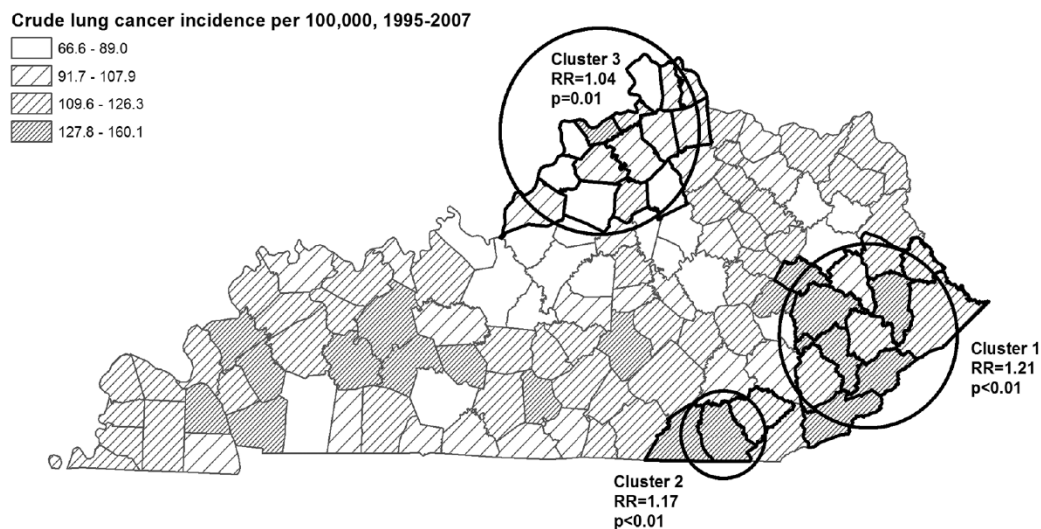
*(Christian et. al 2011)* Reprinted with permission from SAGE Publications © 2011

**Figure 1.3. Clusters of high lung cancer incidence in Kentucky.** Lung cancer rates cannot be explained by smoking incidence in certain regions of Kentucky. Using a spatial scanning statistic that groups together counties in an expanding radius from a central point, *Christian et al. 2011* discovered three clusters of counties in Kentucky whose average lung cancer rate was higher than the state average after correcting for smoking, age and gender. Two of these clusters (clusters 1 and 2) lie within Appalachian Kentucky.

**Figure 1.4. Steps in the Global Genomic Nucleotide Excision Repair (GG-NER) pathway.**

**Figure 1.4. Steps in the Global Genomic Nucleotide Excision Repair (GG-NER) pathway.** (A) A bulky, helically distorting DNA adduct forms. (B) DNA damage recognition protein XPC, complexed with its binding partner HR-23B recognizes the distortion in DNA caused by the adduct and binds on the undamaged strand opposite the adduct. (C) Binding of XPC promotes recruitment of the TFIIH complex, which uses dual helicases (XPB and XPD) to unwind the DNA, forming a single stranded bubble.  (D) Opening of the DNA promotes recruitment of several proteins the complex, including the endonuclease XPG, the single-strand DNA binding protein RPA, and XPA, which functions to confirm the presence of all the components necessary to perform subsequent steps.  (E)  ERCC1-XPF endonuclease is recruited to the damage site and it cuts 5' to the DNA adduct on the damaged strand.  (F) After the 5' incision a DNA polymerase (normally delta or epsilon) is recruited to the site along with RFC and PCNA to resynthesize the gap produced by the endonucleases.  XPG then cuts 3' of the damage site, releasing the single stranded oligonucleotide containing the damage site. (G)  The sugar-phosphate backbone is sealed by DNA ligase I or III to complete the repair process.

# Chapter 2 - Experimental Methods

## 2.1 Cell Culture

The human lung fibroblast cell line IMR-90 (obtained from the American Type Culture Collection) was grown in minimal essential medium (Eagle) containing Earle's salts (Mediatech) supplemented with 0.1 mM non-essential amino acids (Lonza), 2 mM glutamine (Mediatech), 100 units/ml penicillin, 100 µg/ml streptomycin and 10% fetal bovine serum (Sigma).

The human bronchial epithelial cell line, BEAS-2B (obtained from the American Type Culture Collection), was grown in Dulbecco's Modified Eagle's Medium supplemented with 2 mM glutamine, 100 units/mL penicillin, 100 µg/mL streptomycin and 10% heat inactivated fetal bovine serum.

Cultures of primary mouse keratinocyte cells were isolated as described [91]. Briefly, cells were taken from the epidermis of newborn C57BL/6J mice 1-3 days after birth and grown in keratinocyte growth media composed of an equal volume of low calcium (LoCal) media containing (MEM Eagle with Earle's salt, without calcium, 8% chelexed FBS, 2 mM glutamine, 100 units/mL penicillin, 100 µg/mL streptomycin, 0.25 µg/mL fungizone and 0.06 mM $CaCl2$ and fibroblast conditioned LoCal medium. In order to condition the LoCal medium, primary mouse fibroblast cells were grown in it until confluent, and the medium was removed, filter sterilized with a 0.22 micron filter, and stored in minus 80°C until use. All cell lines were maintained in a 5% $CO_2$ incubator at 37ºC.

**2.2 Preparation of CSC**

Cigarette smoke condensate (CSC) was prepared using the University of Kentucky Reference 3R4F Cigarettes as previously described (Hsu et al 1991). Briefly, the particulate phase of smoke was collected on Cambridge filters using a 30-port smoking machine (Borgwaldt) under standard Federal Trade Commission conditions, [92] dissolved in DMSO at a concentration of 40 µg/µl, and stored in small aliquots at -80ºC. At the time of use, aliquots were thawed in a 37°C water bath and discarded after use.

**2.3 Treatment of cells with CSC**

IMR-90 cells were grown to confluency prior to treatment and then treated with CSC for 24 h unless otherwise stated. BEAS-2B cells were seeded at a density of 1.2 million cells per 10 cm plate, allowed to grow for 24 h and log phase cells were then treated with CSC for 16 h. The growth of BEAS-2B cells does not exhibit contact inhibition so studies could not be performed using confluent cultures. For both cell lines, control (mock treated) cells were treated with DMSO, the solvent used for the preparation of CSC, at a volume equal to that used for the highest dose of CSC in each experiment.

**2.4 Treatment of cells with sodium arsenite**

IMR-90 and mouse keratinocyte cells were grown to confluency and then treated with arsenite for 24 h unless otherwise stated. Control (untreated) cells were treated by adding comparable volumes of water, the solvent used to dissolve sodium arsenite, to the medium. In the repair assays, sodium arsenite was added to the medium for the duration of the repair interval.

**2.5 Treatment of cells with MG-132**

MG-132, a potent inhibitor of proteasome mediated proteolysis, was prepared from a 2mM stock solution from EMD Bioscience (catalog # 475791) stored at -20⁰C.  The solution was thawed immediately prior to treatment and diluted for use.  To investigate how changes in the abundance of XPC protein are influenced by UV-C irradiation, confluent IMR-90 cells were treated with MG-132 for 4 h prior to irradiation.  The cells were irradiated with 20 J/m$^2$ UV-C and either lysed immediately or returned to medium containing MG-132 for incubation after UV.  XPC levels were examined using the Western method described above. To investigate a potential involvement of the proteasome in the turnover of XPC after cells are treated with CSC or arsenite, confluent IMR-90 cells were treated with the indicated concentrations of CSC or arsenite for 16 h in the presence or absence of MG-132.  After treatment, cells were counted and lysed, and XPC levels were measured by Western Blotting.

**2.6 Analysis of cell viability for CSC and arsenite treatments**

Cultured cells were grown and treated with CSC or arsenic as described above. After the appropriate treatment duration, cells were trypsinized, washed with chilled PBS, centrifuged, and resuspended in 10 mL of chilled PBS. Cells were counted in triplicate (at minimum) for each dose using trypan blue dye exclusion.

**2.7 Isolation and cryopreservation of peripheral blood mononuclear cells**

PBMCs were isolated from blood samples using standard Ficoll density gradient centrifugation methods [93].  Upon receipt, the blood was warmed to

room temperature, diluted 1:1 in serum-containing medium, (90% RPMI, 10% FBS) and layered onto the density gradient. The samples were centrifuged in an Allegra X-22R centrifuge (400g, 25ºC, 30 m), and the PBMC cell layer (buffy coat) was extracted and washed in serum containing medium. The cells were counted, pelleted, and resuspended in 4ºC freezing medium (50% FBS, 40% RPMI, 10% DMSO) at a density of 5 x $10^6$ cells/ml [94]. The cells were placed in a Mr. Frosty, a storage device that uses isopropanol to slowly reduce the temperature of the freezing medium, and then stored overnight in a -80°C freezer before placing into a liquid nitrogen jacketed tank for long-term storage in the Biospecimen and Tissue Procurement Shared Research Facility (BSTP SRF).

## 2.8 Thawing and stimulation of peripheral blood mononuclear cells

Samples of PBMCs were removed from liquid nitrogen storage and incubated at 37°C. When the solution thawed, the cells were immediately transferred to tubes containing 37ºC thawing medium (50% FBS, 50% RPMI) at a ratio of 1 ml thawed cells to 10 ml of thawing medium. An aliquot of cells was incubated with Trypan blue dye to evaluate cellular viability (which was greater than 90%). The remainder of the cells were then centrifuged (1200 RPM, 10 minutes) to pellet the cells. The supernatant was then removed and the cells were resuspended at a density of 1 million cells/ml of growth medium (20% FBS, 80% RPMI, pen/strep and glutamine) containing the mitogen phytohaemagglutinin (PHA-P, from Sigma Aldrich) at a concentration of 10 µg/ml. PBMCs were then incubated at 37° C in 5% $CO_2$ for 72 h and the repair assay was performed as described below. In the absence of PHA, which stimulates

cellular growth and division, we found that lymphocytes are not proficient in NER, similar to a previous observation that CPD repair was significantly impaired in unstimulated cells [95].

**2.9 Measurement of NER – the immunoblot assay**

The removal of 6-4PPs and CPDs from total genomic DNA was measured using an immunoblot assay as previously described [96, 97] with some modifications. Cells were grown and treated with CSC or arsenic (when testing the impact of these agents on NER) as described above.  They were then washed twice with PBS and irradiated with UV-C light to introduce photolesions; 20 J/m$^2$ UV-C to measure the removal of 6-4 PPs and 2 J/m$^2$ UV-C to measure the removal of CPDs. After irradiation, cells were either lysed immediately or after incubation for increasing periods of time to permit repair.  When testing the impact of CSC or arsenic on repair, the respective chemical was added back into the medium for the repair period after irradiation.  DNA was isolated using a Promega wizard genomic DNA isolation kit.  The concentration of the DNA samples from each time-point was measured using a fluorometer and Hoechst dye, which emits blue fluorescence when bound to dsDNA. The DNA was then denatured and equal amounts of DNA from each time-point were loaded onto a Hybond nitrocellulose membrane (Biorad) using a slot blot apparatus (100-200 ng of DNA per slot for the detection of 6-4 PPs and 20 ng of DNA per slot for detection of CPDs). The membrane was baked in a vacuum oven at 80$^o$C for 1 h, treated with 5 % nonfat dry milk (Blotto, Santa Cruz Biotechnology) in TBST (10 mM Tris pH 8.0, 150 mM NaCl and 0.1% Tween-20) for 1 h, and incubated

overnight at 4° with mouse monoclonal antibodies (1:10,000 dilution) specific for

either 6-4 PPs or CPDs (CAC-NM-DND-002, CAC-NM-DND-001, Cosmo Bio

USA) in 1% dry milk and TBST.  The following day, the membrane was washed

extensively with TBST and then incubated with goat anti-mouse horseradish

peroxide-conjugated secondary antibodies (1:10,000 dilution, Thermo Scientific)

in 1% dry milk and TBST for 2 h at room temperature. After washing,

chemiluminescence (ECL – Plus, GE Healthcare Bio-Sciences Corp) and

fluorimaging were used to detect the photolesions.  Image Quant computer

software was employed to quantify the signal detected by the phosphorimager.

The percent repair at each time point was calculated by dividing the signal

strength of the slot-blot band obtained at each repair time point by the signal

strength of the band obtained at the zero time point. For most membranes, a

DNA ladder of irradiated DNA was used as an internal control for uniformity of

slotting, antibody incubation and development with chemiluminescence.

Unirradiated samples of DNA were also loaded to detect any nonspecific binding

of the antibodies to DNA which in general was found to be insignificant.

The immunoblot assay used to measure NER functions by inducing DNA

damage and measuring clearance over time of the induced lesions.  The

damaging agent used in the experiment was UV-C light, which introduces DNA

photolesions.  Biologically, UV-C is absorbed by the ozone layer and as a result

humans are almost never exposed to this wavelength of light.  Our rationale

behind using this wavelength of light instead of UV-B/A was the quantity of

photolesions introduced by UV-C from a given voltage is much higher than higher

wavelength UV bulbs.  This allows us to rapidly introduce DNA damage in a short period of time, minimizing the time cells are out of medium, and essentially eliminates the potential for repair to begin during the course of treating the cells with the damaging agents.  This is also why a chemical agent was not used to introduce DNA damage in the NER assay.  A variety of chemicals including BaP, acrolein, and AAF produce DNA damage that the NER complex would recognize and repair.  However, these chemicals require time to be absorbed into the cell, to reach the DNA, and to adduct to the DNA, and some even require cellular metabolism to become reactive to DNA, capable of binding and generating a DNA adduct.  Therefore, these chemicals require a significantly longer incubation time in the cellular medium than it takes to irradiate cells (30 seconds).  During a longer incubation period, the NER pathway is actively NER attempting to remove the DNA damage as it is being introduced.  A baseline damage point for all treatments is thus hard to achieve as any negative effect on repair would change the amount of DNA lesions present at "time zero" between treatments.

**2.9.1 Immunoblot assay for lymphocytes**

Lymphocytes stimulated with PHA were placed into 10 cm tissue culture plates at a density of 1 million cells per 1 ml stimulation medium (80% RPMI, 20% FBS, 10 µg/mL PHA) and irradiated with 20 J/m$^2$ UV-C to introduce 6-4 PP lesions. Cells were either lysed immediately after irradiation or after a period of time to permit repair of the damaged DNA.  At three points after irradiation (1, 2, 4 hours) a portion of the irradiated cells were harvested and lysed.  The remainder of the repair assay was then carried out in the same fashion as described in section 2.9.

## 2.10 Western blot analysis

The effect of CSC treatment on the abundance of XPC, XPA, and β-actin proteins in IMR-90 cells and BEAS-2B cells and the effect of arsenic treatment on the abundance of XPC, XPA, and β-actin proteins in IMR-90 cells and mouse keratinocyte cells were determined by Western blotting using mouse monoclonal antibodies (XPA for mouse cells: ab65963, Abcam;  XPA for human cells: sc-56813; XPC: sc-74411, Santa Cruz Biotechnology; β-actin: A3854, Sigma). Cells were treated with different concentrations of CSC, sodium arsenite, or mock treated as described above for 24 h unless stated otherwise.  After treatment, cells were washed with PBS and trypsinized. Approximately five million cells were collected by centrifugation for each treatment, washed with PBS and stored at -20$^o$ or -80$^o$C. Each sample of frozen cells was thawed on ice, resuspended in 200 µl RIPA buffer (50 mM Tris (7.4), 150 mM NaCl, 1% NP-40, 1 mM EDTA, 0.25% Na-Deoxycholate) containing 0.6 mM PMSF, 1% Protease Cocktail Inhibitor (Sigma) and 2 units of DNase 1 (New England BioLabs), lysed by sonication, centrifuged at 13,500 G to remove debris, and the amount of protein in each of the supernatants was quantified using the Bradford method. Alternatively, cells were counted, pelleted, directly lysed in loading dye and samples were loaded based on cell number.  Samples containing 100 µg of protein were mixed with loading dye, boiled at 100$^o$C for 5 min, resolved in 8% (XPC/Actin) or 12% (XPA/Actin) SDS-PAGE gels and transferred to a PVDF membrane (Bio-Rad). The membranes were incubated for one h with 5% nonfat dry milk in TBST for blocking.  For XPC and ß Actin, the upper half of the

membrane was probed with a 1:1,000 dilution of XPC antibody and the lower half

with 1:100,000 dilution of ß Actin antibody and incubated overnight at 4°C. The

membrane was then washed extensively with TBST and incubated for 2 h with

1:5,000 diluted HRP-conjugated goat anti-mouse antibodies (GE Healthcare Bio-

Sciences Corp) at room temperature. After extensive washing with TBST, the

binding of antibodies was detected using enhanced chemiluminescence (ECL –

Plus, GE Healthcare Bio-Sciences Corp) and fluorimaging. For XPA and ß Actin,

the full membrane was probed with 1:1,000 dilution of XPA antibody overnight at

4°C. Washing and development for XPA was performed in the same manner as

XPC. The membrane was then stripped using a stripping buffer (Pierce) at 37°C

for 15 min, washed with TBST, blocked and probed with ß Actin antibody. The

amount of XPC or XPA in each lane was normalized to the amount of ß Actin in

the same lane.

## 2.11 Real-time PCR

The effect of CSC or arsenite on the expression of *XPC, XPA DDB2*

(arsenite only), and *GAPDH* RNA in IMR-90 cells was measured using Real-

Time PCR. Confluent cells were treated as described with CSC, arsenite, or their

respective controls for 24 hours, washed with PBS, trypsinized, centrifuged,

washed with PBS and pelleted. Total RNA was isolated from one million cells for

each treatment using a GenElute mammalian total RNA Miniprep kit (Sigma).

The RNA 6000 Nano Chip kit from Agilent was used to determine the quality and

concentration of RNA isolated from the cells. For each treatment, 1.5 µg of RNA

was reverse transcribed to cDNA using random primers and a high capacity

cDNA RT kit (Applied Biosystems). Primer sets specific to each gene were identified using the Universal Probe Library (Roche), and synthesized by Integrated DNA Technologies. Reactions were carried out on a Roche LightCycler 480 system using a 96-well plate format. Each reaction contained 1X Master Mix, 100 nM fluorescent reporter probe, 200 nM of each forward and reverse primer and 5 µl of cDNA (diluted 1:10). Samples were first incubated for 10 min at 95ºC followed by 40 cycles of amplification (95º C for 15 s denaturation and 60º C for 1 min annealing).  An equal volume of 1:10 diluted cDNA from each experimental treatment was mixed together to create the DNA that was used as a standard. Serial dilutions of mixed cDNA (1:1, 1:10, 1:100, 1:1,000, and 1:10,000) were used to create a standard curve. The expression levels of RNA in the treated cells were determined by comparing the crossing point (CP) values of the treated cells with the CP values of a standard curves for the corresponding gene or RNA. The crossing point is cycle of amplification during which the fluorescence of a sample achieves a threshold fluorescence above background. The amounts of *XPC*, *XPA*, or *DDB2* RNA were normalized to the amounts of *GAPDH* RNA for each treatment.

# Chapter 3

# Exposure of Human Lung Cells to Tobacco Smoke Condensate Inhibits the Nucleotide Excision Repair Pathway

## 3.1 Introduction

Lung cancer is a deadly disease and a leading cause of cancer-related mortality in the US and in the world [98-100]. In 2012, the most recent year data is available, lung cancer accounted for 1.8 million cases of cancer and 1.6 million deaths worldwide [101, 102]. Exposure to tobacco smoke is the predominant risk factor for the development of lung cancer and it is estimated to account for 85-90% of all lung cancer cases [103, 104]. It is also associated with the formation of tumors at additional sites in the body that are not directly exposed to smoke including the bladder, pancreas, liver, stomach and bone marrow [11, 105]. Tobacco use remains prevalent in certain regions of the world [106] and while its use has declined in the US, approximately 50% of newly diagnosed lung cancers occur in former smokers [103]. Hence, lung cancer and other forms of cancer associated with tobacco smoke exposure remain a tremendous health burden in the US and world-wide. Continued elucidation of the molecular mechanisms that lead to the formation of cancers associated with tobacco smoke is essential for prevention, treatment and identification of individuals who are at greatest risk for the development of cancer.

Thousands of compounds have been identified in the vapor and particulate phases of cigarette smoke and they include carcinogens, co-carcinogens, mutagens and tumor promoters. Approximately 70 of these

compounds have been classified as carcinogens [66, 104]. Different classes of chemical carcinogens are present in tobacco smoke including the polycyclic aromatic hydrocarbons (PAHs) such as benzo[a]pyrene (B[a]P), dibenz[a,h]anthracene and dibenzo[a,l]pyrene. The DNA-reactive metabolites of PAHs are considered to be among the primary tobacco smoke carcinogens [67, 104].  Metabolic activation of these and other chemical compounds found in tobacco smoke can generate intermediates that react with DNA bases and produce DNA adducts. Hence, DNA adducts are likely continually formed in the lung tissues of people who smoke, and if they are not removed by DNA repair processes, their persistence could lead to the formation of mutations. Many different types of genetic alterations are found in lung cancer and they include point mutations, genomic rearrangements, amplifications and large scale insertions and deletions. Mutations in KRAS and TP53 are frequently found in lung tumors and lung tissues of smokers [103, 107, 108], and the accumulation of mutations in these and other important oncogenes and tumor suppressor genes are driving forces in the development of lung cancer.

PAH-induced DNA damage is removed by the nucleotide excision repair (NER) pathway [74, 75, 109-113] and hence, NER activity is likely critical to the prevention of carcinogen-induced mutations that contribute to neoplasia associated with smoke exposure. NER is a versatile pathway that removes a wide variety of structurally diverse DNA lesions including those generated by metabolites of chemical carcinogens as well as those generated by exposure to ultraviolet (UV) light. The cyclobutane pyrimidine dimer (CPD) and 6-4

photoproduct (6-4 PP), produced by UV light, are model substrates commonly studied when measuring NER activity as they are rapidly generated by a brief exposure to UV light [25]. In mammals, at least 20 different protein factors participate in NER, including the XPA-G factors that are singly defective in the 7 corresponding complementation groups of the human disease, xeroderma pigmentosum (XP).  The tumor suppressor factor p53 also impacts NER efficiency probably by transcriptional regulation of the *XPC* and *DDB2* gene products [20, 21, 114, 115]. The NER pathway is comprised of two sub-pathways that differ in their mechanism of damage recognition: global genomic NER (GG-NER) which can remove damage from anywhere in the genome and transcription-coupled NER (TC-NER) which selectively removes damage from the transcribed strands of expressed genes.  In GG-NER, DNA damage recognition is accomplished by XPC, which is stabilized by its binding partners RAD23B, and CENTRIN2 [26].  In TC-NER, damage is recognized by the stalling of the RNA polymerase complex at the site of damage (reviewed in [24]).  After DNA damage recognition, many of the subsequent steps are the same for GG-NER and TC-NER. The helicase activities of TFIIH produce additional unwinding of DNA where upon the endonuclease activities of the XPF/ERCC1 complex and XPG produce single-strand incisions flanking the damaged site.  The original integrity of the DNA is restored after an approximately 30 nucleotide region of DNA containing the lesion is removed, and the gap is filled by pol δ or pol ε, using the undamaged strand as a template (reviewed in [25]).

Several different types of animal models have been used to investigate the molecular mechanisms of lung cancer development caused by exposure to tobacco smoke. Unfortunately, many smoke inhalation studies have had limited success [116-118]. The A/J mouse model has been used extensively but it is confounded by the spontaneous lung tumors in control animals, need for long smoke exposure and recovery regimens, low tumor induction by smoke inhalation and responses that do not adequately mimic those found in humans exposed to tobacco smoke [116, 119]. Greater tumor incidence has been achieved by exposing female B6C3F1 mice to lifetime, whole-body mainstream tobacco smoke [118, 120]. The utilization of mice with targeted disruptions in tumor suppressor genes or oncogenes associated with lung cancer development in humans may yield improved animal models and additional mechanistic insights [121].

As animal models for tobacco smoke exposure have had limited success, cigarette smoke condensate (CSC) has been used as a surrogate for cigarette smoke exposure to study its effects in model cell culture systems [105]. It is the particulate phase of cigarette smoke collected on Cambridge filters and resuspended in DMSO. It is mutagenic and genotoxic and produces several different types of mutations including point mutations, deletions, loss of heterozygosity (LOH),  microsatellite instability, sister chromatid exchanges and micronuclei [122-127]. It introduces DNA damage [127-130]. It also induces human and mammalian cell transformation [131-133] and tumor formation when applied to mouse skin [134-136]. While it is well established that cigarette smoke

42

introduces a variety of different types of DNA damage, it is less clear how smoke exposure influences DNA repair efficiency and DNA damage response pathways.

Loss of DNA repair capability results in increased mutagenesis and carcinogenesis. In mouse models, deficiencies in NER have been associated with tumorigenesis at many organ sites including the lung [137]. Compared to normal mice, NER-deficient mice have a higher incidence of lung tumors when exposed to B[a]P [138, 139] and XPC-deficient mice have elevated levels of spontaneous lung tumors [140]. By extension, even a partial loss of NER efficiency in people is likely to increase mutagenesis and cancer incidence, particularly in cases of chronic DNA damage induction, as occurs in the lung tissue of smokers.

We have studied the effects of CSC on the NER pathway using an immuno-slot blot assay to measure NER in two human lung cell lines; IMR-90 and BEAS-2B. We find a dose-dependent inhibition of the efficiency of NER when both cell lines are treated with increasing concentrations of CSC. Additionally, the impact of CSC on the abundance of various NER proteins and their respective RNAs was investigated. We find that the abundance of XPC protein is significantly reduced when cells are treated with increasing concentrations of CSC, while the abundance of XPA protein is unaffected by treatment. Both XPC and XPA RNA levels are modestly affected by CSC treatment. Finally, a possible mechanism in which CSC inhibits XPC expression by altering protein turnover was investigated. Our results provide evidence for the first time that CSC can directly interfere with the normal NER process, both in

terms of overall efficiency as well as at the protein and RNA level of NER factors, suggesting a possible new manner by which tobacco smoke may promote carcinogenesis.

**3.2 Statistical Considerations**

All statistical evaluations were done using Graph Pad Prism 6. For the experiments comparing multiple treatments to a control, a 1-way ANOVA with a Holm-Sidak test for multiple-comparison was employed. For the experiments comparing one treatment to a control, a Student's T-test was employed. For all analyses, $p < 0.05$ was used as the threshold for significance. Statistical significance is described in the results section and not presented on the images as some graphs contained too many significant treatments to indicate on the graphs.

**3.3 Results**

**3.3.1 CSC inhibits NER in IMR-90 cells**

We first investigated the effects of CSC on the NER pathway by studying a human lung fibroblast cell line, IMR-90. The effect of CSC on cell viability was evaluated by treating cells with a range of different doses, 0 to 200 µg/ml, for 24 h. No significant effect on viability was observed even after treatment of cells with 200 µg/ml, the highest concentration of CSC used (Fig. 3.1A). NER was studied using an immuno-slot blot assay that measures the removal of 6-4PPs or CPDs introduced by UV irradiation. These lesions are repaired exclusively by NER in human cells and 6-4PPs are model substrates to measure NER efficiency in total genomic DNA. In untreated cells, 6-4PPs are rapidly removed from DNA with the

majority of lesions removed within 3 h after UV irradiation. Treatment with CSC inhibits the removal of 6-4PP lesions in a dose-dependent fashion and the inhibition was statistically significant for all doses of CSC used (except 60 μg/ml) at one or more time points (Fig. 3.1B-C). In general, NER in CSC treated samples was slowed but reached completion within 24 h after UV irradiation except when cells were treated with the highest dose of CSC, where both the kinetics and extent of repair were reduced. In contrast to the rapid removal of 6-4PPs from genomic DNA found in untreated IMR-90 cells, the removal of CPDs in untreated cells was much slower. This observation of slow and inefficient removal of CPDs is similar to what other investigators have previously found in human cell lines using comparable doses of UV light [141, 142]. Treatment with CSC had only a minor inhibitory effect on the removal of CPDs (Fig. 3.1D-E), with a statistically significant difference only observed at the 48 h time point (p=.0435). The significance of this inhibition is difficult to evaluate since it was only seen at 48 h after UV irradiation. While the IMR-90 cells were studied in a confluent state, small differences in the amount of DNA replication at the 48 h time point could contribute to the differences in repair measured in the CSC-untreated and treated samples.

**3.3.2 CSC reduces the abundance of XPC but not XPA protein in IMR-90 cells**

The effect of CSC on the abundance of XPC and XPA protein was determined by western blotting using whole cell lysates of IMR-90 cells treated with 120, 160 and 200 μg/ml of CSC for 24 h. These are the same doses and

treatment time used to study the effects of CSC on NER. Whole cell lysates from mock treated cells (DMSO) were used as controls. The abundance of XPC protein was reduced in IMR-90 cells treated with increasing concentration of CSC for 24 h (Fig. 3.2A-B); XPC protein levels were reduced by 56% in cells treated with 200 µg/ml of CSC. The reduction in XPC protein compared to the mock treated cells was statistically significant in all treatments.  In contrast, there was not any statistically significant change in the abundance of XPA protein in cells treated with CSC (Fig. 3.2C-D).  The amounts of XPC and XPA proteins were normalized to the amounts of β actin present in each lane.

### 3.3.3 A timecourse of CSC treatment shows a correlation between the reduction of XPC protein and the inhibition of NER

Treatment of IMR-90 cells with CSC for 24 h inhibits NER and results in reduced expression of XPC protein. To measure the kinetics of these inhibitions, cells were treated with 200 µg/ml of CSC and the abundance of XPC protein was measured at 4 h intervals over a 24 h period and NER was measured after treatment with CSC for 8 h intervals over a 24 h period. A reduction in XPC protein compared to untreated cells was observed as early as 8 h and it was maximally inhibited by 16 h. The inhibition was statistically significant for all timepoints except 4 h (Fig. 3.3A-B).  Significant inhibition of NER was observed at all treatment times beginning with the 8 h treatment (Fig. 3.3C-D). NER was maximally inhibited with the longest length of treatment, 24 h, and a statistically significant difference from cells not treated with CSC was observed for all repair timepoints.

### 3.3.4 CSC modestly reduces XPC and XPA RNA levels in IMR-90

The effect of CSC treatment on the abundance of XPC or XPA RNA in

IMR-90 cells was measured using the same doses of CSC that were used to

evaluate its effect on protein levels. No statistically significant changes in the

abundance of XPC RNA were observed when cells were treated with 120 µg/ml

or 160 µg/ml CSC, whereas a modest but statistically significant decrease was

observed when cells were treated with 200 µg/ml of CSC (Fig. 3.4).  XPA showed

modest, but significant, reduction in RNA expression across all CSC treatments.

### 3.3.5 CSC inhibits NER and reduces the abundance of XPC protein in

### BEAS-2B cells

We also investigated the impact of treatment with CSC on NER in a

bronchial epithelial cell line, BEAS-2B.  We chose BEAS-2B cells due to the

relevance of epithelial cells as sites of lung cancer formation, as they are the

cells that line the respiratory tract and directly interact with inhaled carcinogens.

The effect of CSC on cell viability was evaluated by treating cells with the same

dose range chosen for IMR-90 viability (Fig. 3.5A).  Some toxicity was observed

for BEAS-2B cells treated with CSC compared to little or no toxicity observed in

IMR-90 cells (Fig. 1A). BEAS-2B were actively replicating at the time of

treatment, and as a result they were likely more sensitive than IMR-90 cells to

the toxicity of CSC. The effect of CSC on NER was then studied by measuring

the removal of 6-4PPs after treatment with 175 µg/ml CSC for 16 h (Fig. 3.5B).

In cells not treated with CSC, NER was rapid and efficient. Treatment with CSC

resulted in significant inhibition of the removal of 6-4 PPs at all three time points

(Fig. 3.5C).  BEAS-2B cells were also treated with different doses of CSC for 16

h and probed for XPC and XPA protein expression (Fig 3.5D).  A reduction in the

abundance of XPC protein, but not XPA, was observed across the range of

treatments of BEAS-2B cells with CSC (Fig 3.5E); similar to our findings in IMR-

90 cells.

**3.3.6 The reduction of XPC protein in IMR-90 cells by CSC treatment is**

**mediated through the proteasome**

After observing that treatment of IMR-90 cells with CSC results in a

significant reduction in the abundance of XPC protein, we investigated the

potential influence of CSC treatment on ubiquitin mediated turnover of XPC.  UV

irradiation induces ubiquitination of XPC protein [143] and XPC protein

abundance has been previously linked to ubiquitin modification. After UV

irradiation, XPC protein levels quickly drop in a proteasome dependent manner

[144-146], although this appears to be both temporary and UV-dose dependent

[143].  We also addressed this by treating IMR-90 cells with MG-132, which

permits ubiquitin-linkages but prevents ubiquitin-mediated proteasomal

degradation, and then exposing those cells to UV irradiation.  XPC levels were

reduced after UV exposure, but in the presence of MG-132, this reduction was

significantly diminished (Fig, 3.6A).  This confirms that UV-mediated XPC protein

levels are altered in a proteasome-dependent manner.  We then asked whether

the observed inhibition of XPC by CSC was also mediated by the proteasome.

We treated IMR-90 cells with CSC in the presence of MG-132 for 16 h and

observed that the reduction in XPC protein level produced by treatment with CSC

alone was almost completely abrogated by the addition of MG-132 (Fig. 3.6B).

120 µg/ml CSC reduced XPC expression in the absence of MG-132, but in the

presence of 1 and 5 µM MG-132, CSC did not inhibit XPC protein expression,

suggesting that the inhibition of XPC protein expression by CSC requires

functional proteasomal activity.

## 3.4 Discussion

In this study, we find that cigarette smoke condensate, a surrogate for

tobacco smoke exposure, can inhibit NER function and reduce the expression of

XPC, a key protein required for DNA damage recognition in the NER pathway.

Consequently tobacco smoke exposure can affect the integrity of DNA in two

fundamentally different ways. It is well established that it can introduce DNA

damage, an important contributor to lung carcinogenesis. However, our findings

indicate that, in addition, it can also inhibit the DNA repair pathway that is

essential for the removal of some of the carcinogenic DNA damage introduced by

smoke carcinogens. Hence, cells of the lung exposed to smoke would likely

suffer an even greater DNA damage burden than previously held. Certain

individual constituents of tobacco smoke have been implicated in inhibiting DNA

repair.  Acrolein, a combustion product of cigarette smoke, has been shown to

inhibit the nucleotide excision repair pathway and it inhibits XPC expression in a

proteasome-dependent fashion [147-149].  In addition, arsenic, a metal

constituent of tobacco smoke, has been found to reduce XPC expression [150].

Together, these studies suggest that components of tobacco smoke can impact

lung carcinogenesis by inhibiting NER. CSC can also inhibit the base excision

repair (BER) pathway, which repairs oxidative DNA damage, another type of tobacco smoke damage [130, 131]. Hence, it will be important to investigate how tobacco smoke-induced alterations in different DNA repair pathways contribute to lung carcinogenesis.

DNA damage recognition is an early, key step in NER and several studies suggest that DNA damage recognition by XPC is or can be the rate-limiting step in the pathway. Reduced expression of XPC has been associated with reduced repair of UV-induced photoproducts [151] and increased cancer incidence [47, 152]. Biochemical and cellular studies indicate that the binding affinity of XPC for the DNA lesion or the time it takes XPC to find the DNA lesion may be the specific rate limiting step [28-30]. Conversely, complementation of an XPA deficient cell line with very low levels of XPA protein fully restores DNA repair activity [153] and XPA becomes rate limiting only when levels are reduced by over 90% [154, 155]; these studies suggest that the participation of XPA protein is usually not a rate limiting step in NER. We find that treatment of cells with CSC inhibits NER and there is a concomitant reduction in XPC protein while XPA protein levels remain unchanged. If XPC protein is rate limiting in NER, then the inhibition of NER by CSC may be a direct consequence of the reduction in XPC protein. Additional genetic studies that manipulate the abundance of XPC during or after CSC treatment are needed to more directly establish the relationship between CSC induced alterations in XPC protein levels and its inhibition of NER. We find that treatment with CSC results in only a modest reduction in the levels of XPC RNA and it is only detected well after the amounts of XPC protein are

reduced. This suggests that most of the reduction in XPC protein produced by treatment with CSC is not caused by a reduction in XPC RNA. Lastly, an additional protein, UV DNA Damage Binding Protein 2 (UV-DDB2), is specifically required for the recognition and removal of CPDs in cells and functions in the turnover of XPC that can impact the removal of both CPDs and 6-4PPs. We attempted to examine the effect of CSC treatment on the abundance of DDB2 but were unsuccessful in identifying an antibody that yielded results specific to the protein.

Previous studies have demonstrated that XPC protein is polyubiquitinated after exposure to UV light and its polyubiquitination is mediated by the UV-DDB-Ubiquitin ligase complex [143, 145, 156]. There are at least two distinct types of ubiquitin modifications to XPC following UV-induced DNA damage. A lysine-48-linked polyubiquitin linkage appears to promote degradation of XPC by the proteasome [157]. In contrast, a lysine-63-linked ubiquitination of XPC can be critical for the removal of XPC from the lesion site and allow downstream NER factors access to the DNA damage [158]. In addition, although the type of ubiquitin linkage was not determined, ubiquitin modification of XPC can actually increase its binding affinity for undamaged DNA [143] XPC is also modified by sumoylation in response to UV damage [145] and inhibition of this modification reduces XPC stability after UV irradiation. XPC SUMO modification has been implicated in promoting lysine-63 mediated XPC polyubiquitination [158, 159]. A recent report revealed seven unique UV-induced ubiquitination sites on XPC [160]. All of these studies demonstrate the complexity of the post-translational

modifications to XPC in response to DNA damage and they likely represent [81] the intricacies of regulating such an important and complex pathway.

We observe that treatment with CSC results in a reduction in the abundance of XPC protein in the absence of UV damage, and that this reduction is reversed when cells are treated with MG-132.  This suggests that the reduction in the abundance of XPC protein produced by exposure to CSC is a consequence of enhanced proteasome-dependent turnover of the protein that is mediated by ubiquitination. XPC is intrinsically unstable as a monomer [161] and knocking down one of its binding partners, HR23B, promotes XPC degradation in a proteasome dependent manner [162].  Studies have also found that ubiquitination and proteasome-dependent turnover of XPC is important in maintaining steady-state levels of the protein in the absence of DNA damage [163]. In addition, the deubiquitinating enzyme USP-7 removes a UV-induced polyubiquitin chain from XPC that would otherwise target it for proteasome-mediated degradation [146].  Hence, HR23B and USP-7 both function to stabilize XPC by inhibiting ubiquitin-mediated degradation. Since treatment with CSC can result in the introduction of different forms of DNA damage (discussed below), it is possible that the induced turnover of XPC by the proteasome is mediated through the introduction of certain types of DNA damage. Additional studies are needed to determine if CSC-mediated turnover of XPC functions by promoting ubiquitination or inhibiting deubiquitination and to characterize the sites of ubiquitination. CSC has been previously shown to enhance the proteasome-mediated turnover of Akt, a protein kinase [164], which together with our studies

may indicate that smoke exposure can impact multiple pathways by targeting specific proteins to the proteasome for degradation.

The assay used to measure NER in this study measures the introduction and removal of UV light-induced photoproducts. We chose to use this approach because it provides a method to rapidly introduce DNA damage and it precisely quantitates the kinetics of removal of a model substrate for the NER pathway. Moreover, the use of UV light as a DNA damaging agent is not confounded by variations in cellular uptake or metabolic activation as can be the case when using chemical agents to introduce DNA damage. It is likely that CSC would inhibit the removal of other substrates for NER but additional studies are required to investigate its effect on the removal of chemical modifications to DNA introduced by tobacco smoke. Tobacco smoke contains a number of compounds capable of producing DNA lesions, several of which are repaired by the NER pathway, including BPDE. The BPDE lesion can be particularly important in the etiology of smoking-related lung cancer, as the G-T transversion signature mutation of BPDE [70, 76, 81, 105, 165] is found more frequently in at CpG mutational hotspots in the *P53* gene of lung tumors from smokers compared with lung tumors from non-smokers [78, 166, 167]. Mice exposed topically to B[a]P develop skin tumors with the same signature mutations in p53 [168]. Hence, measuring the effect of CSC on the removal of BPDE adducts is important to directly demonstrate that tobacco smoke exposure inhibits the repair of DNA damage introduced by the tobacco smoke itself. In our study, it is likely that any adducted base damage directly introduced by the CSC treatment was very small

compared to the levels of photoproducts introduced by UV irradiation. Thus, it is unlikely that the inhibition of NER by CSC treatment was due to a competition between different types of DNA damage introduced during the experiments or titration of the NER pathway by damage directly introduced by CSC.

## 3.5 Conclusions

Cigarette smoking remains the highest risk factor for lung cancer development in the United States and the world.  However, not all smokers develop lung cancer.  Our results suggest that CSC, a commonly accepted surrogate for tobacco smoke exposure, inhibits the NER pathway, increasing the persistence of lesions in DNA.  In addition, CSC reduces the abundance of XPC protein, a key protein required for DNA damage recognition in NER, and this reduction is likely produced by targeting XPC to the proteasome for degradation. It is well established that variations in DNA repair capacity contribute to cancer risk, and our findings indicate that inhibition of NER is another mechanism by which smoking may contribute to development of lung cancer. Therefore, it may be prudent to consider measuring individual DNA repair capacity as a means of evaluating lung cancer risk, particularly among people who have other risk factors.

**Figure 3.1. CSC inhibits NER in IMR-90 human lung fibroblasts.**
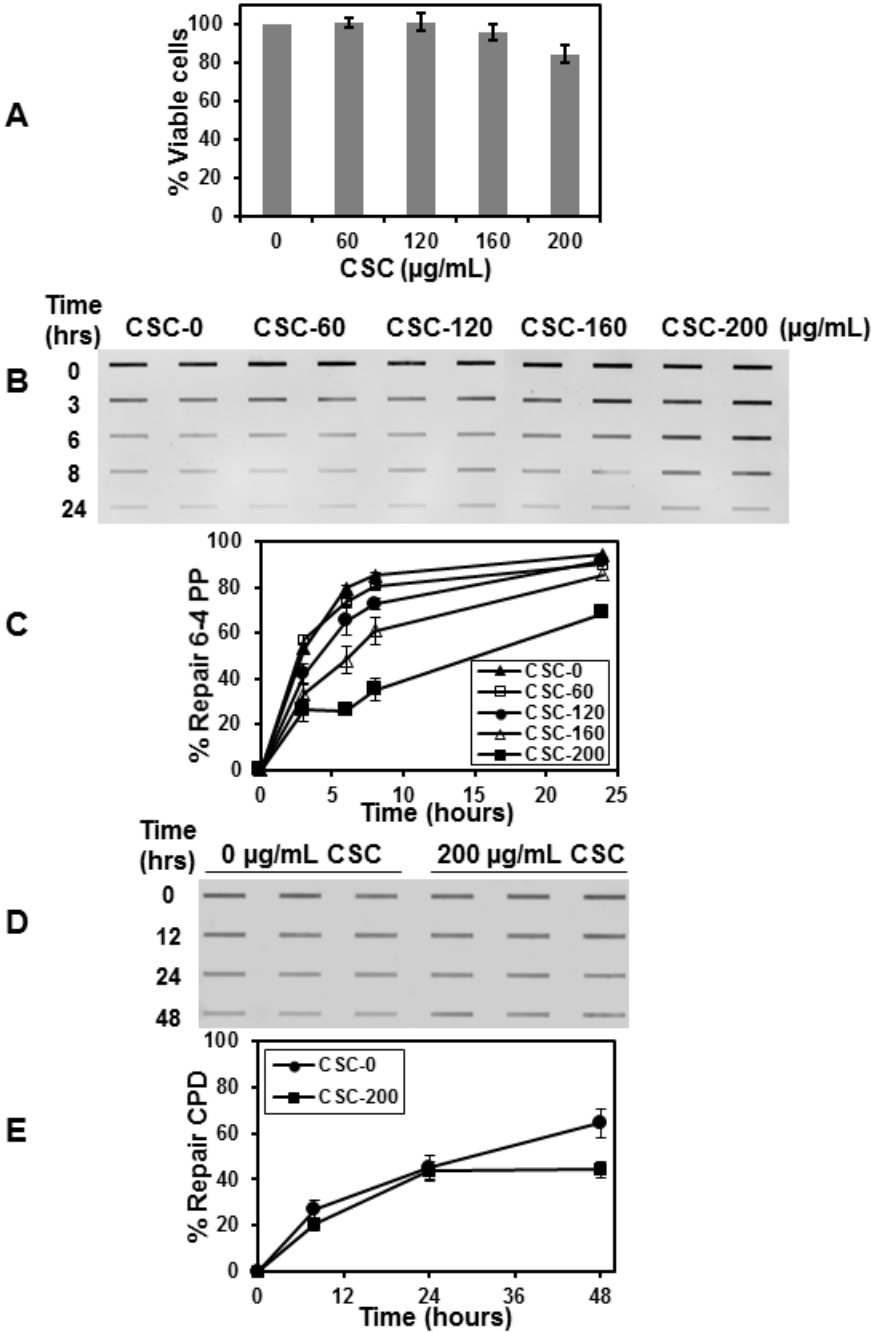
**Figure 3.1. CSC inhibits NER in IMR-90 human lung fibroblasts.** (A) Cell

viability after CSC treatment. Confluent IMR-90 cells were treated with CSC (or

mock treated with DMSO) with the concentrations shown for 24 h and the

percentage of viable cells was measured using Trypan blue dye exclusion. The

data represent the mean ± SE (Standard Error) from four independent

experiments (except for 60 µg/ml CSC which included two independent

experiments). (B) Removal of 6-4 PPs. Cells were treated with the concentrations

of CSC shown or with DMSO for 24 h and irradiated with 20 J/m$^2$ UVC to

introduce photolesions. After irradiation, cells were either lysed immediately or

after incubation in medium containing CSC or DMSO for the times (h) shown to

permit repair.  The immunoblot assay to detect 6-4 PPs was performed and

samples were loaded in duplicate for each repair time point. (C) Removal of 6-4

PPs. A graphical representation of results obtained from multiple immunoblots

measuring the removal of 6-4 PPs is shown. Each data point represents the

mean ± SE of three repeats from two independent experiments.  (D)  Removal of

CPDs. IMR-90 cells were treated with 200 µg/mL CSC or with DMSO for 24 h

and irradiated with 2 J/m$^2$ UVC to introduce photolesions.  After irradiation, cells

were either lysed immediately or after incubation in medium containing CSC or

DMSO for the times shown to permit repair.  An immunoblot assay to detect CPD

lesions was performed and samples were loaded in triplicate for each time point.

(E) Removal of CPDs. A graphical representation of results obtained from

multiple immunoblots measuring the removal of CPDs is shown.  Each data point

represents the mean ± SE of three repeats from two independent experiments.

**Figure 3.2. CSC reduces the abundance of XPC, but not XPA, protein in IMR-90 cells.**
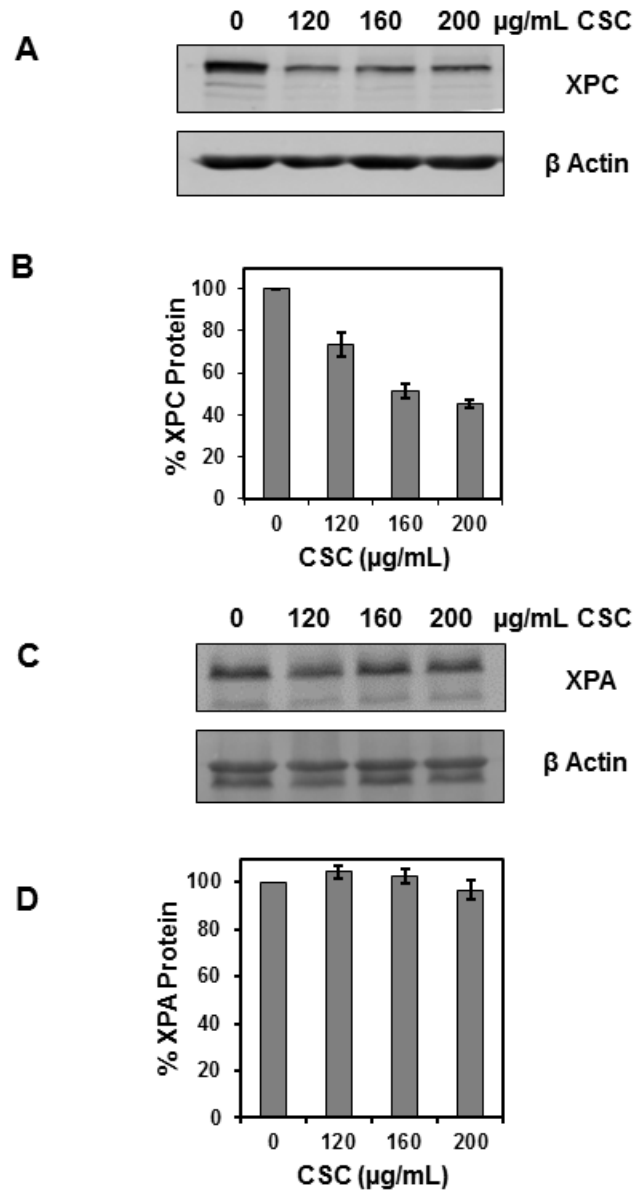
**Figure 3.2. CSC reduces the abundance of XPC, but not XPA, protein in IMR-90 cells.** (A) Cells were treated with different concentrations of CSC as shown for 24 h, and the abundance of XPC was examined by western blot analysis. (B) A graphical representation of multiple western blots for XPC expression is shown. The data presented are the mean ± SE from two repeats each of three independent experiments, and XPC expression was normalized to β-actin. (C) Cells were treated as shown for 24 h, and the abundance of XPA was measured by western blot analysis. (D) A graphical representation of multiple western blots for XPA expression is shown. The data presented are the mean ± SE of three repeats from three independent experiments, and XPA expression was normalized to β-actin.

**Figure 3.3. The impact of CSC on XPC protein and NER efficiency depends on treatment duration.**
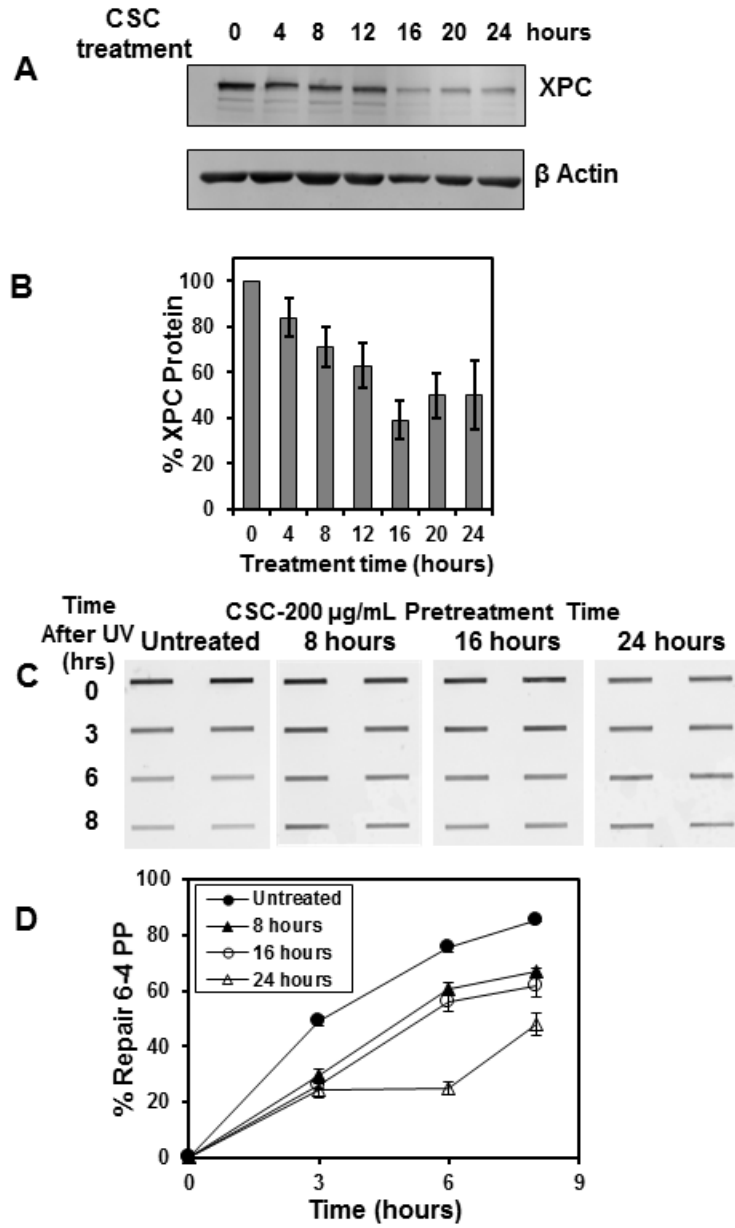
**Figure 3.3. The impact of CSC on XPC protein and NER efficiency depends on treatment duration.** (A) IMR-90 Cells were treated with 200 µg/ml of CSC for the times indicated and XPC expression was examined by western blot analysis. (B) A graphical representation of multiple western blots examining the time course of inhibition for XPC expression after CSC treatment is shown. The data presented are the mean ± SE of three repeats from one experiment, and XPC expression was normalized to β-actin. (C) Results of an immunoblot showing the time course of the effect of CSC on repair of 6-4 PPs in IMR-90 cells. Cells were treated with 200 µg/mL CSC for 8, 16, or 24 h (or DMSO for 24 h) and irradiated with 20 J/m$^2$ UVC to introduce photolesions. After irradiation, cells were either lysed immediately or after incubation in medium containing CSC or DMSO for the times shown (3, 6 and 8 h) to permit repair. An immunoblot assay was performed and samples were loaded in duplicate to measure the removal of 6-4 PPs. (D) A graphical representation of multiple immunoblots for 6-4 repair is shown. Each data point represents the mean ± SE of three repeats from one experiment.

**Figure 3.4. The effect of CSC on the abundance of XPC and XPA RNA in IMR-90 cells.**
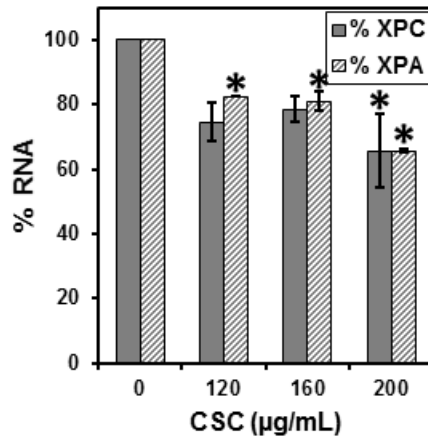


**Figure 3.4. The effect of CSC on the abundance of XPC and XPA RNA in IMR-90 cells.** Cells were treated with the indicated concentrations of CSC (or DMSO) for 24 hours. RNA was isolated and Real Time PCR was performed as described in the methods section. The expression of XPC or XPA RNA was normalized to GAPDH RNA for each treatment. The data presented are the mean ± SE of one analysis from three independent experiments.

**Figure 3.5. CSC inhibits NER and the abundance of XPC protein in BEAS-2B cells.**
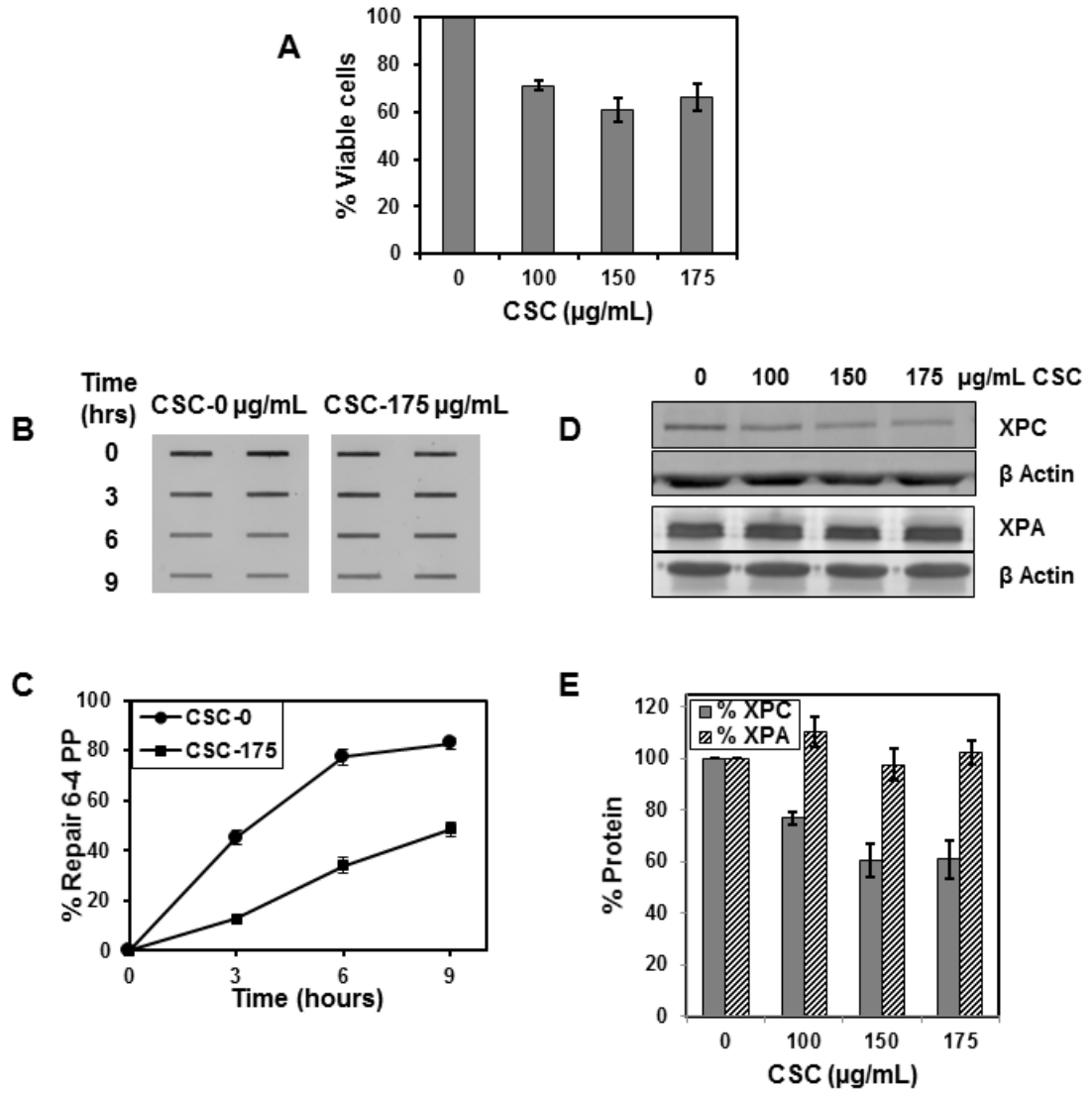
**Figure 3.5. CSC inhibits NER and the abundance of XPC protein in BEAS-2B cells.** (A) Cells were treated with the concentrations of CSC shown for 16 h and the percentage of viable cells was measured using Trypan blue dye exclusion. The data presented are the mean ± SE of four biological experiments.  (B) Cells were treated with 175 µg/mL CSC (or DMSO) for 16 h, and irradiated with 20 J/m$^2$ UVC to introduce photolesions.  After irradiation, cells were either lysed immediately or after incubation in medium containing CSC or DMSO for increasing periods of time to permit repair.  An immunoblot assay was performed and samples were loaded in duplicate to measure the removal of 6-4 PPs. (C) A graphical representation of multiple immunoblots for 6-4 repair is shown.  Each data point represents the mean ± SE of four repeats from one experiment.  (D) Cells were treated with the different concentrations of CSC shown for 24 h, lysed, and the abundance of XPC and XPA were examined by western analysis. (E) A graphical representation of multiple western blots for XPC and XPA protein is shown.  The data presented are the mean ± SE of two repeats from two independent experiments for XPC and three repeats of one experiment for XPA.  XPC and XPA values were normalized to β-actin.

**Figure 3.6. Involvement of the proteasome in the reduced expression of XPC protein in IMR-90 cells after treatment with UV or CSC.**
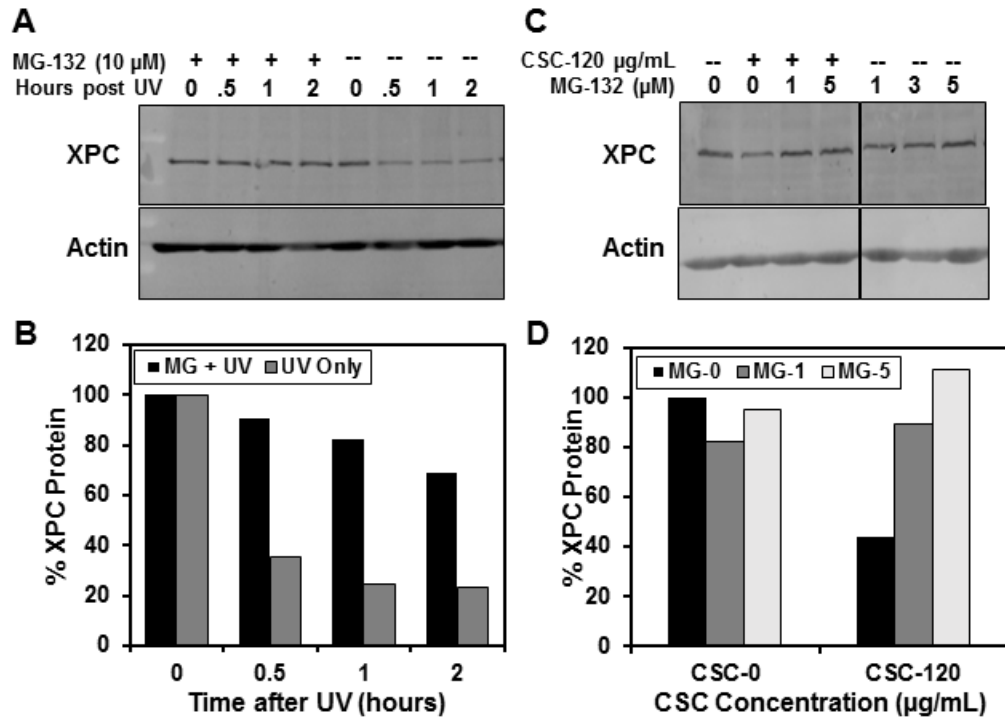
**Figure 3.6. Involvement of the proteasome in the reduced expression of XPC protein in IMR-90 cells after treatment with UV or CSC.** (A) Cells were treated with 10 µM MG-132 or untreated for 4 hours, and then irradiated with 20 J/m$^2$ UV-C.  After irradiation, cells were either lysed immediately or after incubation for increasing periods of time in the same type of medium as was used for the pretreatment; medium containing MG-132 or not containing MG-132. XPC expression was measured using Western blot analysis.  (B) A graphical representation of the experiments from (A) is shown.  XPC expression was normalized to β-actin.  The data presented are the mean of two repeats from one experiment.  The percent XPC protein was calculated by comparing post-UV time points to the appropriate 0 h (MG-treated or not MG-treated) time point. Treatment with MG-132 for 4 h had a negligible impact on XPC expression before irradiation, so both 0 h values were set to 100%. (C) Cells were treated with a combination of CSC and/or MG-132 at the indicated concentrations for 16 h.  XPC levels were measured by Western blot analysis and normalized to Actin. (D) A graphical representation of the average obtained from two different blots for the experiments in (C) is shown; treatment with 3 µM MG-132 was not included. XPC expression was normalized to β-actin.

# Chapter 4

## Inorganic arsenic inhibits the nucleotide excision repair pathway and reduces the expression of XPC protein

### 4.1 Introduction

Inorganic arsenic is a metalloid found in moderate levels in the Earth's crust. It is one of the world's oldest known carcinogens [169] and it is classified as a class I carcinogen [170]. Humans can be exposed to arsenic by drinking contaminated water [171], inhaling certain substances in occupational settings [170], ingesting contaminated food (summarized in [172]) and using tobacco products [173]. It is estimated that approximately 200 million people in the world are exposed to drinking water contaminated with potentially harmful levels of arsenic [174]. In the US, exposure to contaminated drinking water occurs in many regions throughout the country [175]. According to the Agency for Toxic Substances and Disease Registry (ATSDR), arsenic ranks number one on the priority list of hazardous substances found at National Priority List sites (http://www.atsdr.cdc.gov/SPL/index.html).

Exposure to arsenic increases the risk for cancer development at multiple organ sites including the skin [176-178], lung [179, 180], liver (summarized in [181]), and bladder [182].  Arsenic has been used for centuries for a range of medicinal purposes, and topical treatment with an arsenic compound called Fowler's solution (discovered in 1786) was used to treat a host of diseases including malaria, syphilis, asthma, chorea, eczema, and psoriasis [183]. It was over one hundred years later that Hutchison proposed that Fowler's solution was

a human skin carcinogen [184]. Correlations between arsenic exposure and internal cancer development were first observed in regions of the world where populations were exposed to levels of arsenic between 10-100 fold greater than the 2002 EPA standard and current World Health Organization Guideline level of 10 µg/L in drinking water [185]. In these regions of the world, such as parts of Bangladesh, India, China, Argentina, Chile, and other countries (Reviewed in [175]), high arsenic contamination in the water is a consequence of local geographical deposits of the metal and poor procedures for decontaminating water. In the US, exposure to arsenic is generally more moderate. However, it has been estimated that arsenic concentrations exceed 20 µg/L in 5% of regulated water systems [175] and arsenic exposure in the US has been tentatively linked to skin [186], bladder [187] and lung cancers [83]. Interestingly, arsenic has also been adopted as a chemotherapeutic agent, with arsenic trioxide (ATO) used with success to treat acute promyelocytic leukemia [188-190].

Arsenic exposure likely contributes to cancer development by multiple mechanisms. Evidence indicates that arsenic impacts genotoxicity, although it does not directly interact with DNA to produce DNA damage (reviewed in [84]). Rather than directly producing DNA damage, arsenic is often considered a co-carcinogen that influences the mutagenicity and carcinogenicity of other agents. Co-treatment of cells with arsenic and UV light results in a reduction in the removal of photoproducts produced by UV light [191-193] and an increase in mutation rates produced by UV light [193-196]. Co-treatment of cells with

arsenic increases the levels of DNA adducts produced by benzo[a]pyrene diol-epoxide (BPDE) which is a metabolite of benzo[a]pyrene (B[a]P), a compound present in tobacco smoke [197, 198]. Consistent with this observation, arsenic reduces the removal of BPDE adducts [199, 200] and increases the frequency of mutations formed by BPDE [201]. These observations are directly relevant to the types of cancer associated with exposure to arsenic; specifically, arsenic increases non-melanoma skin cancer and lung cancer incidences in humans for which UV light and tobacco smoke exposure, respectively, are risk factors [177, 202].

Several studies have investigated how arsenic may act as a co-carcinogen using mouse models.  When added to drinking water, arsenic enhances the production of DNA damage and the mutagenicity of topically applied B[a]P, but it does not produce mutations or DNA damage in the absence of B[a]P [203, 204].  Arsenic increases the multiplicity and size of skin tumors produced by dimethylbenz[a]anthracene (DMBA), a polycyclic aromatic hydrocarbon (PAH), but it does not alter skin carcinogenesis in the absence of DMBA [205, 206].  It also greatly increases the formation of UV-induced skin tumors in hairless mice [207], but it does not induce tumor formation in the absence of UV damage [208].

Many of the agents that arsenic interacts with as a co-mutagen or co-carcinogen, including UV light and tobacco smoke, produce DNA lesions that are removed by the nucleotide excision repair (NER) pathway. UV light produces cyclobutane pyrimidine dimers (CPDs) and 6-4 photoproducts (6-4 PPs), both

substrates for NER [209]. Tobacco smoke contains several classes of chemicals which can react with DNA, forming adducts that are removed by NER. Multiple polycyclic aromatic hydrocarbons (PAHs) from combustion of tobacco (as well as fossil fuels and other organic matter) produce bulky, DNA distorting adducts that are generally repaired by NER. PAH-induced adducts include (+)-*trans*-BPDE-$N^2$-dG, the primary stereoisomeric adduct formed by B[a]P, which is clearly removed by NER (reviewed in [75]). Tobacco smoke also contains a class of aromatic amines called 4-aminobiphenyls (4-ABP) which can produce DNA adducts recognized by NER, including *N*-(deoxyguanosin-8-yl)-4-aminobiphenyl (dG-C8-ABP) [210].

The NER pathway removes helix-distorting DNA lesions which can cause mutations and drive carcinogenesis. This is clearly illustrated by decades of investigation of the disease Xeroderma Pigmentosum (XP). Patients with XP are deficient in NER and have greatly elevated levels of skin cancer and other forms of cancer [40, 42, 211]. In mammals, at least 20 different proteins participate in NER, including the XPA-G factors that are singly defective in the 7 corresponding complementation groups of XP [25, 212, 213]. The tumor suppressor factor p53 also impacts NER efficiency probably by transcriptional regulation of the *XPC* and *DDB2* genes [20, 21, 114, 115]. NER, sometimes referred to as global genomic NER (GG-NER), can remove damage from anywhere in the genome. A subpathway of NER called transcription-coupled NER (TC-NER) selectively removes damage from the transcribed strands of expressed genes. These two pathways differ in their mechanism of DNA damage recognition. In NER, DNA

damage recognition is accomplished by XPC, which is stabilized by its binding

partners RAD23B and CENTRIN2 [26] and is assisted by the UV-damaged DNA

binding protein DDB2 (the product of the *XPE* gene). In TC-NER, damage is

recognized by the stalling of the RNA polymerase complex at the site of damage

(reviewed in [24]). After DNA damage recognition, the subsequent steps are the

same for NER and TC-NER. The multi-subunit complex TFIIH contains helicase

activities that produce additional unwinding of DNA, which produces double-

strand/single-strand DNA junctions. After DNA unwinding several NER

components are recruited to the site of the lesion, including XPA, which is likely

used to verify the presence of the DNA lesion and that the required NER factors

are present for the subsequent steps of the pathway. Next, the endonuclease

activities of the XPF/ERCC1 complex and XPG produce single-strand incisions

flanking the damaged site. The original integrity of the DNA is restored after an

approximately 30 nucleotide region of DNA containing the lesion is excised, and

the gap is filled by pol δ or pol ε, using the undamaged strand as a template, with

DNA ligase IIIα or DNA ligase I filling in the final nick in the DNA (reviewed in

[25]).

The molecular mechanisms by which arsenic may act as a co-carcinogen

are under debate. One possibility is that arsenic inhibits the removal of DNA

damage produced by carcinogens such as UV light and certain compounds

present in tobacco smoke, thus exacerbating their mutagenic effects. NER has

been suggested as a candidate pathway since it removes the DNA damage

introduced by these agents [150, 200, 214-217]. In the present study, we have

examined the impact of arsenic (as the trivalent ion arsenite) on NER using an immuno-blot assay to directly measure DNA lesions specifically removed by NER. NER function was inhibited with increasing concentrations of arsenite in both human fibroblasts and mouse keratinocytes. Additionally, NER protein and RNA levels were measured in both cell types in response to arsenite treatment, and a concentration-dependent decrease in XPC protein and XPC, XPA, and DDB2 RNA levels was observed. Finally, a possible mechanism in which arsenite inhibits XPC protein expression by altering protein turnover was investigated. Our findings support the hypothesis that arsenic can promote carcinogenesis by interfering with the NER-specific repair of DNA damage introduced by other carcinogens, and provides insights into how arsenic may additionally function as an anti-cancer drug, especially in tandem with DNA damage-inducing chemotherapeutics.

## 4.2 Results

### 4.2.1 Statistical considerations

All statistical evaluations were done using Graph Pad Prism 6. All experiments involved comparing multiple treatments to a control, and therefore a 1-way ANOVA with a Holm-Sidak test for multiple-comparison was employed. For all analyses, $p < 0.05$ was used as the threshold for significance.

### 4.2.2 Arsenite inhibits NER in primary mouse keratinocyte cells

We investigated the impact of treatment with arsenite on NER in primary mouse keratinocytes. Arsenic can be a co-carcinogen in skin cancer development and squamous cell and basal cell carcinoma originate from

keratinocytes located in the skin.  First, the effect of arsenite on mouse

keratinocyte cell viability was evaluated across a range of arsenite exposures

(Fig. 1A).  Toxicity was moderate across the treatment range, with the highest

exposure producing ~50% toxicity.  After examining the effects of arsenite on cell

viability, the effect of arsenite on NER was studied using an immuno-slot blot

assay that uses lesion-specific antibodies to measure the removal of 6-4 PPs or

CPDs introduced by UV irradiation.  These lesions are removed exclusively by

NER in rodent and human cells, although CPDs are a poorer substrate for NER

and require a longer period of time to be removed from DNA [218, 219].  In cells

exposed to arsenic, they were pretreated with arsenic for 24 h, irradiated with UV

and arsenic was added to the medium for the duration of the repair interval. In

keratinocytes not treated with arsenite, 6-4 PPs were rapidly removed from DNA,

with the majority of lesions removed within 4 h after UV irradiation. In contrast,

treatment with arsenite resulted in a concentration-dependent inhibition of the

removal of 6-4 PPs compared to untreated cells (Fig. 1B).  20 µM arsenite

significantly inhibited 6-4 PP removal at the 8 and 24 h timepoints, while 40 µM

arsenite significantly inhibited 6-4 PP removal at the 4, 8, and 24 h timepoints.

Arsenite treatment also inhibited the removal of CPDs in mouse keratinocytes

(Fig. 1C).  The UV-C dose used to measure the removal of CPDs was

considerably lower than that used to measure removal of 6-4 PPs (2 J/m$^2$

compared to 20 J/m$^2$) since CPDs are introduced at a much higher frequency

than 6-4 PPs. 20 µM arsenite significantly inhibited the removal of CPDs at the

48 h timepoint, and the 40 µM treatment significantly inhibited the removal of CPDS across all of the timepoints measured.

**4.2.3 Arsenite reduces the abundance of XPC, but not XPA, protein in primary mouse keratinocyte cells**

The effect of arsenite on the abundance of XPC and XPA was determined by western blotting using whole cell lysates isolated from mouse keratinocytes treated for 24 h with the indicated concentrations of arsenite (Fig 2A). The amounts of XPC and XPA at each concentration are presented in reference to their amounts in untreated cells, with normalization for the amount of β-actin present in the same lane (Fig. 2B-C). The abundance of XPC was reduced in cells treated with increasing concentration of arsenite (Fig. 2B); XPC levels were reduced by 40% in cells treated with 10 µM arsenite, the lowest concentration tested.  The reduction increased to 59% and 81% for the 20 and 40 µM arsenite treatments.  The amount of reduction compared to untreated cells was statistically significant for each concentration of arsenite administered. In contrast, the abundance of XPA was not significantly affected by arsenite treatment, except at the highest exposure (Fig. 2C).

**4.2.4 Arsenite inhibits NER in IMR-90 cells**

We also investigated the effects of arsenite on the NER pathway by studying the human lung fibroblast cell line, IMR-90.  First, the effect of arsenite on cell viability was evaluated by treating cells with the indicated concentrations for 24 h (Fig. 3A). Treatment with arsenite produced some moderate toxicity at exposures ≥15 µM, with 40 µM producing an approximate 50% reduction in

viability.   NER was studied as described.  Treatment with ≥20 μM arsenite

resulted in a concentration-dependent inhibition of the removal of 6-4 PPs

compared to untreated cells.  The removal of 6-4 PP lesions was significantly

slowed in all the arsenite treatments, across all timepoints measured (except As-

20 at 2, 4 and 6 h) (Fig. 3B). No significant effect on the removal of 6-4 PPs was

observed when cells were treated with 10 μM arsenite for 24 h (data not shown).

Second, IMR-90 cells were treated for a longer period of time, 48 h,  with a

reduced range of arsenic concentrations (0-15 μM arsenite) and cell viability and

the removal of 6-4 PPs was measured (Supplemental Fig. 1). When cells were

treated with arsenic for 48h, the medium was removed after 24 h of treatment

and replaced with fresh medium containing the same concentration of arsenic.

Treatment of cells with 10 or 15 μM arsenite for 48 h produced little toxicity

(Supplemental Fig. 1A) but resulted in a significant inhibition of the removal of 6-

4 PPs compared to untreated cells at all timepoints except the 3 h timepoint of

the 10 μM arsenite treatment (Supplemental Fig 1B).  At these lower doses, the

amount of inhibition observed was similar, likely due to technical limitations of the

assay to differentiate small differences in NER efficiency. Thus, we observe that

arsenic inhibits the removal of 6-4 PPs after treatment with higher concentrations

for 24 h and after treatment with lower concentrations for 48 h.  The impact of

arsenite on the removal of CPDs was also investigated.  When cells were treated

with 20 or 40 μM arsenite for 24 h and irradiated with 2 $J/m^2$ UV light, the

removal of CPDs was inhibited in a statistically significant fashion at multiple

timepoints (12 and 48 h timepoints for 20 µM arsenite, all timepoints for 40 µM arsenite) (Fig. 3C).

## 4.2.5 Arsenite reduces the abundance of XPC, but not XPA, protein in IMR-90 cells

The effect of arsenite on the abundance of XPC and XPA was determined by western blotting using whole cell lysates isolated from IMR-90 cells that were treated for 24 h with different concentrations of arsenite. Whole cell lysates isolated from untreated cells were used as controls. The amounts of XPC and XPA were normalized to the amount of β-actin in the same lane. The abundance of XPC was reduced in IMR-90 cells treated with increasing concentrations of arsenite (Fig. 4A); XPC levels were reduced by 40% in cells treated with 10 µM arsenite, the lowest exposure tested. Inhibition increased to 70 and 85% for the 20 and 40 µM arsenite treatments. The reduction in the abundance of XPC in all three treatments compared to untreated cells was statistically significant (Fig 4B). XPA levels were also examined after arsenite treatment (Fig 4C). In contrast to XPC, only the highest concentration of arsenite produced a modest but significant 25% reduction in the abundance of XPA (Fig 4D).

## 4.2.6 Arsenite reduces XPC, XPA, and DDB2 RNA levels in IMR-90 cells

The effect of arsenite treatment on the abundance of *XPC*, *XPA*, and *DDB2* RNA in IMR-90 cells was measured by real-time PCR, using the same treatment conditions used to evaluate its effect on protein levels. Arsenite significantly reduced RNA expression of all three genes in a concentration-dependent fashion (Fig. 5). Changes in the abundance of *XPC* and *XPA* RNA

were observed at 20 and 40 µM arsenite, whereas *DDB2* expression was severely reduced even at the lowest exposure, 10 µM arsenite.

## 4.2.7 The reduction of XPC protein in IMR-90 cells by arsenite treatment is mediated through the proteasome

After observing that treatment of IMR-90 cells with arsenite results in a significant reduction in the abundance of XPC protein, we investigated the potential influence of arsenite treatment on ubiquitin-mediated degradation of XPC. As XPC is modified by ubiquitin [143, 145, 156], we hypothesized that arsenite may inhibit XPC abundance in a manner that is regulated by the proteasome. To test this hypothesis, we treated IMR-90 cells with arsenite in the presence or absence of the proteasomal inhibitor MG-132 for 16 h and measured XPC by Western blotting (Fig. 6A). In cells treated with arsenic alone, we observed a reduction in XPC expression similar to the levels seen in Fig. 4A. When cells were co-exposed to arsenic and MG-132, the reduction in XPC expression was eliminated, and expression levels were maintained at a level similar to untreated cells (Fig. 6B). These findings suggest that proteasomal degradation of XPC is at least partially responsible for the observed reduction in XPC protein levels upon treatment with arsenite.

## 4.3 Discussion

In this study, we use an immuno slot-blot method to directly measure the removal of UV photoproducts and we find that arsenic, in the trivalent form of sodium arsenite, inhibits the efficiency of NER. Treatment with arsenite significantly inhibits the removal of 6-4 PPs and CPDs in two different cell types;

human IMR-90 lung fibroblasts and primary mouse keratinocytes.  In addition, we find that treatment with arsenite reduces the abundance of XPC protein and *XPC*, *XPA* and *DDB2* RNA. Others have investigated the effects of arsenic on the efficiency of NER using a variety of different assays. Arsenite has been found to interfere with the production of repair-related DNA strand breaks in general and alters DNA incision and ligation frequencies in human fibroblasts exposed to UV light, more indirect measurements of NER [214-217]. An early observation of an inhibitory effect of arsenite on the efficiency of NER was made studying the removal of thymine dimers using chromatography in a Chinese hamster ovary (CHO) cell line measured 24 h after UV irradiation [192].  More recently, arsenite treatment of human lymphoblastoid TK6 cells was found to slow the removal of UV-induced DNA damage as measured by the Comet assay 2 h after UV irradiation [193] and arsenite treatment of a mouse keratinocyte cell line was found to delay the removal of 6-4 PPs introduced by solar simulated UV radiation [191]. Therefore, our findings extend and add to the understanding of how arsenic alters NER and possibly contributes to arsenic-mediated carcinogenesis.

We observed some increase in cellular toxicity when cells were treated with different concentrations of arsenic.  While the inhibition of NER can sensitize cells to cell death and apoptosis [220-222], it is unclear how apoptosis or cell death may impact the efficiency of NER.  There was a 10 fold difference in the doses of UV light used in our experiments to measure the removal of 6-4 PP and CPD lesions (20 J/m$^2$ for 6-4 PPs and 2 J/m$^2$ for CPDs).  However, we found that the inhibition of NER found after treatment with arsenic was significant regardless

of whether the cells were treated with a lower or higher dose of UV light; this suggests that the inhibition of NER in response to treatment with arsenic was not solely brought about changes in toxicity. In addition, when we treated cells with lower concentrations of arsenic for a longer period of time, 48 h instead of 24 h, toxicity was markedly reduced, but we still observed a significant inhibition in removal of 6-4 PPs.

Our observations of the impact of arsenic on cellular expression of NER proteins and RNA levels also extend previous studies. Treatment of human skin fibroblasts with arsenic reduced the expression of XPC and DDB2 at both the protein and RNA levels, but the efficiency of NER was not measured [150]. We observe that treatment with arsenic significantly reduces XPC protein in human IMR-90 lung fibroblasts and primary mouse keratinocytes but we were unable to measure DDB2 protein since we were unsuccessful in identifying an antibody that yielded results specific to the protein. We also observe a reduction in *XPC*, *XPA*, and *DDB2* RNA levels in IMR-90 cells treated with arsenic. Our observation of a reduction in *XPA* RNA but not XPA protein after treatment with some concentrations of arsenic might be due to a relatively high abundance of XPA protein or low turnover of the protein. Interestingly, arsenic exposure has also been linked to a reduction in RNA levels of several NER genes in a human population, including *ERCC1*, *XPF*, and *XPB*, but not *XPA* (*XPC* and *DDB2* RNA levels were not measured) [223].

Several studies have proposed mechanisms by which arsenic can inhibit NER function. One study observed that arsenic reduces the expression of XPC

(and DDB2) protein and RNA in human skin fibroblasts, and these reductions also correlated with a reduction in the recruitment of XPC to sites of DNA damage [150]. The mechanism by which arsenic reduced XPC expression or whether treatment with arsenic impacted the efficiency of the removal of DNA damage was not explored, in contrast to our study. Additionally, arsenic inhibits the removal of BPDE adducts in A549 human lung cells and it can displace zinc from a peptide representing the zinc finger domain of XPA at high exposures (as judged by 19% zinc release when the peptide was treated with 100 µM arsenite) [199]. Since the zinc finger domain has been shown to be important for the function of XPA in NER [224], arsenic may disrupt NER by altering the structure of XPA, but additional studies are needed to explore these findings. In addition, methylated metabolites of arsenite were found to be considerably more efficient than the parental arsenical in displacing zinc from XPA, and inhibited the removal of BPDE lesions at lower concentrations [199].

We observe that treatment with arsenic inhibits NER function and reduces XPC expression. XPC is ubiquitylated in response to UV-induced DNA damage [143, 145, 156]. There are multiple ubiquitin modifications to XPC following the introduction of UV-induced DNA damage. These include a lysine-48-linked polyubiquitin linkage that appears to promote degradation or recycling of XPC [157] and a lysine-63-linked ubiquitination of XPC that can be critical for the removal of XPC from the damaged site to allow downstream NER factors access to the DNA damage [158]. The deubiquitinating enzyme USP-7 is responsible for removing a UV-dependent ubiquitin linkage to XPC, and this promotes XPC

stability [146].  We tested whether the reduction of XPC protein expression that we observe after treatment of IMR-90 cells with arsenic is related to proteasomal turnover of XPC.  We observe that the proteasomal inhibitor MG-132 abrogated the effects of arsenite in reducing XPC protein levels. This indicates that the reduction in the abundance of XPC produced by exposure to arsenic is, at least in part, a consequence of enhanced proteasome-dependent turnover of the protein that is mediated by ubiquitination.  Additional studies are needed to determine if arsenic-mediated turnover of XPC functions by promoting ubiquitination or inhibiting deubiquitination and to characterize the sites of ubiquitination. Furthermore, arsenic's effects on the proteasome may not be limited to altering the expression of XPC and may have an influence on the stability of other proteins and pathways.

Given our observation that XPC RNA is also reduced by arsenite treatment, arsenic-mediated reduction in XPC expression may not be solely a result of increased proteasomal turnover of XPC.  However, inhibiting proteasomal activity with MG-132 fully restored XPC expression in the presence of arsenic, indicating that proteasomal turnover of XPC is responsible for the majority of the observed reduction of XPC protein abundance.  Had the reduction in XPC levels by arsenic treatment been primarily a consequence of reduced transcript levels, MG-132 would not have been able to restore XPC expression to levels seen in untreated cells.  Additionally, treatment with MG-132 alone did not result in an increase in XPC abundance. This suggests that the steady state turnover of XPC is fairly low in undamaged cells, and this observation is

consistent with results of a previous study where XPC levels were unchanged for up to nine hours when XPC turnover was inhibited [225]. Consequently, the increase in XPC protein in cells treated with MG-132 and arsenic relative to cells treated with arsenic alone is not a result of MG-132 increasing XPC above basal levels, but rather a function of inhibiting the turnover of XPC mediated by arsenic. Nevertheless, the relative contributions of the reduction in XPC RNA and the increase in XPC turnover to the overall arsenite-dependent reduction in XPC needs further study.

We studied the effects of arsenic on the NER pathway in light of previous associations made between exposure to arsenic and elevated frequencies of lung and skin cancer in humans [176, 179] and elevated frequencies of skin cancer in UV-exposed mice [207, 208]. The primary risk factor for lung cancer is exposure to tobacco smoke and the primary risk factor for skin cancer is exposure to solar UV irradiation. Tobacco smoke introduces many types of DNA damage, including the highly studied PAHs such as B[a]P. BPDE, the ultimate metabolite of B[a]P, produces large bulky adducts that are recognized and removed by the NER pathway [75, 109, 111]. BPDE adducts produce a characteristic mutational signature, and this signature is found more frequently in lung cancers of smokers than non-smokers [78, 166, 167]. Sun exposure introduces DNA lesions which are also recognized by NER [25, 209]. In non-melanoma skin cancer, mutational signatures characteristic of exposure to UV light are observed in the p53 gene and are present at mutational hotspots [226], indicating that these NER substrates can be causative in the production of non-

melanoma skin cancer. Melanocytes are more resistant to UV induced cytotoxicity than are keratinocytes, the epidermal skin cells responsible for most non-melanoma skin cancers. This difference may explain the observation that arsenic increases the cancer incidence of non-melanoma skin cancers [227] while the relationship between arsenic and melanoma development is less clearly established [228, 229]. Since deficiencies in the removal of BPDE adducts or UV-induced photolesions by NER result in increased mutation levels, the inhibition of NER by arsenite could explain how arsenic can act as a co-carcinogen in smoking-related lung and UV-related skin cancers.

In general, humans are exposed to arsenic for a substantial period of time, such as when they are exposed to contaminated drinking water (which can occur for years), or when they are treated with arsenic-based chemotherapies (which can occur for several weeks [188]). In contrast, treatment of cells in culture usually involves a single, much shorter exposure to arsenic which often necessitates the use of higher than physiological concentrations to observe measurable effects. Chronic treatment of cells with low concentrations of arsenite can induce cellular transformation [230-233], global DNA hypomethylation [231-234], local DNA hypermethylation [234, 235], alterations in chromatin structure and splicing patterns [230], and increases in oncogenic gene expression [232, 236]. Recent research has indicated that chronic arsenic exposure alters methylation patterns of DNA repair factors involved in NER, leading to reduced RNA expression of ERCC1 and ERCC2 [237]. These epigenetic changes likely contribute to arsenic-mediated carcinogenesis, but chronic treatment of cells in

culture with lower, physiological concentrations would likely complicate investigations and interpretations of direct effects on DNA repair. Additional studies in human populations are important to determine if chronic exposure to lower concentrations of arsenic through contaminated drinking water or arsenic-based chemotherapy inhibits the efficiency of NER in a human populations.

Several previous studies, some described above, and our study have investigated the effects of arsenic on NER using single treatments of at least 10 μM arsenic [150, 192, 199, 200]. These concentrations are higher than what humans would usually encounter during long term chronic exposure to contaminated drinking water. However, they are within the range of concentrations used to sensitize chemo-resistant cells to killing by chemotherapeutic agents such as doxorubicin [238], paclitaxel [239], erlotinib [240], and cisplatin [241]. Of particular interest is the finding that arsenic treatment sensitizes chemo-resistant cancer cells to killing by cisplatin. Cisplatin generates DNA adducts that are substrates for NER. Hence, co-treatment with arsenic may sensitize cells to cisplatin by inhibiting the NER pathway [241]. The same study also found that treatment with arsenic reduced expression of XPC, and that disruption of the *XPC* gene in the absence of arsenic treatment sensitized the chemo-resistant cancer cells to killing by cisplatin. These studies and our study suggest that disruption of the NER pathway by treatment with arsenic should be explored to enhance the efficacy of chemotherapeutic agents used to treat cancer.

While several studies have indicated that arsenic increases the mutation frequency of other carcinogens such as UV-light [193-196] and components of tobacco smoke [201, 203], relatively few studies have investigated the impact arsenic plays on the types of mutations found in co-exposed systems.  While not a measurement of mutagenesis, co-exposure of arsenic and benzo[a]pyrene increased the frequency of adducts produced by benzo[a]pyrene in mice [204]. In cultured CHO cells, co-treatment with arsenic and UV-light increases mutation frequency by a factor of two relative to cells exposed to UV-light alone, and the mutational spectrum in the co-treated cells retained a UV-signature [242]. Very few studies have examined the spectra of mutations produced in human skin cancers related to human exposure to arsenic.  In one single study of an arsenic-exposed population of skin cancer patients, the mutations found in p53 were at different codons than the normal hotspots seen in the p53 gene of UV-induced skin cancers [243]. However, all the tumor samples in this single study were obtained from areas of the body not generally exposed to sunlight, and hence it is unlikely that UV light played any major role in the etiology of these tumors. Consequently, it is clear that additional studies are important to investigate the mutational spectra found in skin tumors derived from sun exposed areas of the body in individuals exposed to arsenic. These studies should aid in understanding whether the alterations to the efficiency of NER that we have observed in our study also occur in humans exposed to arsenic, and if so, whether they play a role in increasing mutational burden and the risk of developing cancer.

**Fig 4.1. Arsenic inhibits the removal of 6-4 PPs and CPDs in primary mouse keratinocytes.**
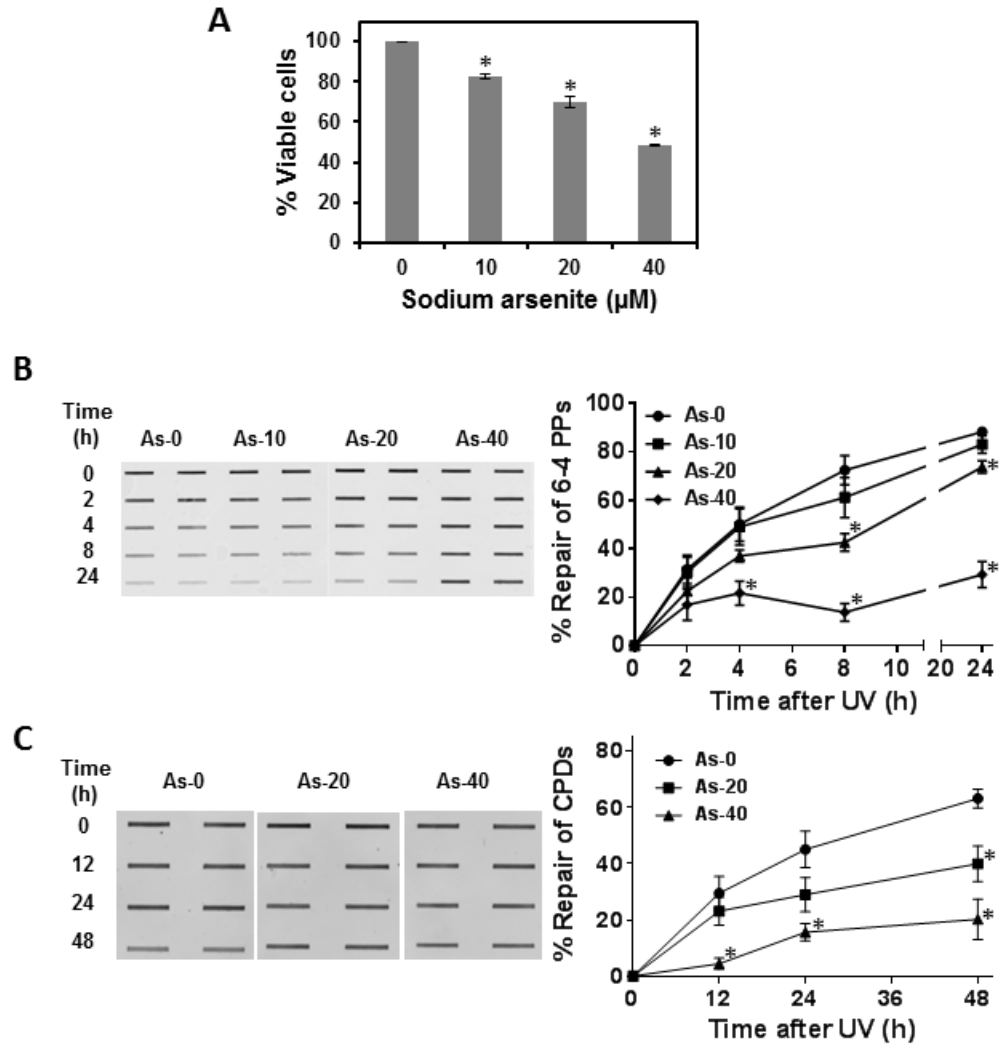
**Figure 4.1.** (A) Primary mouse keratinocytes were treated with the indicated concentrations of arsenite for 24 h and the percentage of viable cells was measured using Trypan blue dye exclusion. The data represent the mean ± SE (Standard Error) from two independent experiments. (B) Primary mouse keratinocytes were treated with the indicated concentrations of arsenite for 24 h, irradiated with 20 J/m$^2$ UV-C, and incubated in the same concentration of As-containing medium for the times shown. An immuno slot-blot assay was performed and removal of the 6-4 PP lesion was measured. A graphical representation of the results obtained from multiple immunoblots measuring the removal of 6-4 PPs is also shown. Each data point represents the mean ± SE of 4-5 repeats of one biological experiment (5 repeats of 0 and 10 doses, 4 repeats for 20 and 40 doses). (C) Primary mouse keratinocytes were treated with the indicated concentrations of arsenite for 24 h, irradiated with 2 J/m$^2$ UV-C, and incubated in the same concentration of As-containing medium for the times shown. An immuno slot-blot assay was performed and the removal of CPDs was measured. A graphical representation of the results obtained from multiple immunoblots measuring the removal of CPDs is also shown. Each data point in the graph represents the mean +/- SE of three repeats from one independent experiment. Treatments with arsenic that produced a significant difference (P <0.05) compared to the control treatment are designated with an asterisk. For the repair assay, each treatment was compared to the untreated cells at the same timepoint after irradiation.

**Figure 4.2. Arsenic reduces XPC, but not XPA, protein in primary mouse keratinocytes.**
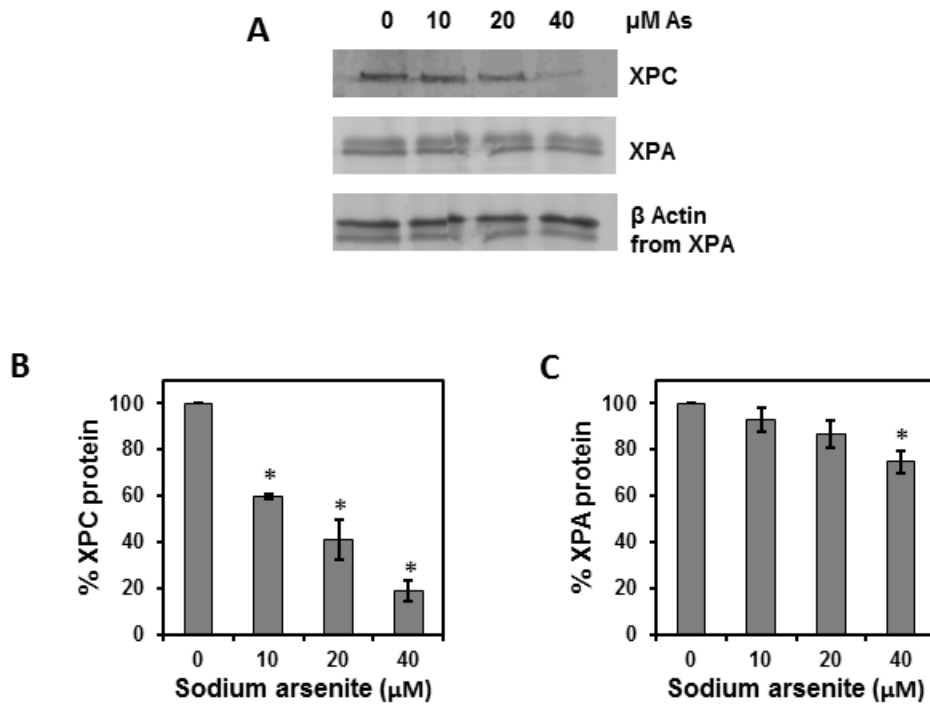


**Figure 4.2.** (A) Primary mouse keratinocytes were treated with the indicated concentrations of arsenite for 24 h, and the abundance of XPC and XPA protein was examined by western blot analysis. Graphical representations of multiple western blots for XPC (B) and XPA (C) protein are shown. The data presented are the mean ± SE from two repeats each of one independent experiment, and the amount of XPC or XPA protein were normalized to β-actin. Treatments with arsenic that produced a significant difference (P <0.05) compared to the control treatment are designated with an asterisk.

**Figure 4.3. Arsenic inhibits the removal of 6-4 PPs and CPDs in IMR-90 human lung fibroblasts**
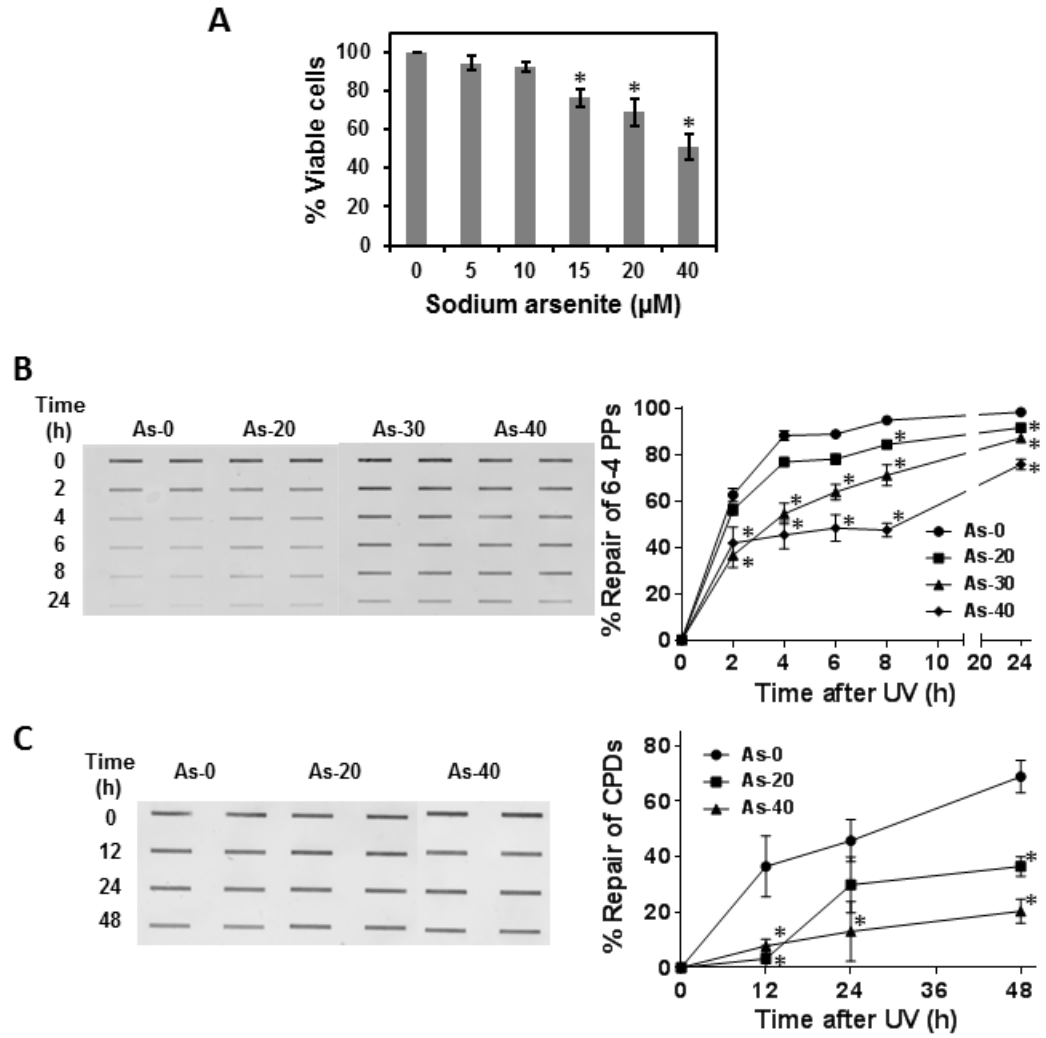
**Figure 4.3.** (A) IMR-90 cells were treated with the indicated concentrations of arsenite for 24 h and the percentage of viable cells was measured using Trypan blue dye exclusion. The data represent the mean $\pm$ SE from seven independent experiments. (B) IMR-90 cells were treated with the indicated concentrations of arsenite for 24 h, irradiated with 20 J/m$^2$ UV-C, and incubated in the same concentration of As-containing medium for the times shown. An immuno slot-blot assay was performed and removal of the 6-4 PP lesion was measured. A graphical representation of the results obtained from multiple immunoblots measuring the removal of 6-4 PPs is also shown. Each data point in the graph represents the mean +/- SE of four repeats from one experiment. (C) IMR-90 cells were treated with the indicated concentrations of arsenite for 24 h, irradiated with 2 J/m$^2$ UV-C, and incubated in the same concentration of As-containing medium for the times shown. An immuno slot-blot assay was performed and the removal of CPDs was measured. A graphical representation of the results obtained from multiple immunoblots measuring the removal of CPDs is also shown. Each data point in the graph represents the mean +/- SE of at least three repeats from one experiment (two repeats of one experiment for the 40 µM dose). Treatments with arsenic that produced a significant difference (P <0.05) compared to the control treatment are designated with an asterisk. For the repair assay, each treatment was compared to the untreated cells at the same timepoint after irradiation.

**Figure 4.4. Arsenic reduces the abundance of XPC, but not XPA, protein in IMR-90 cells.**
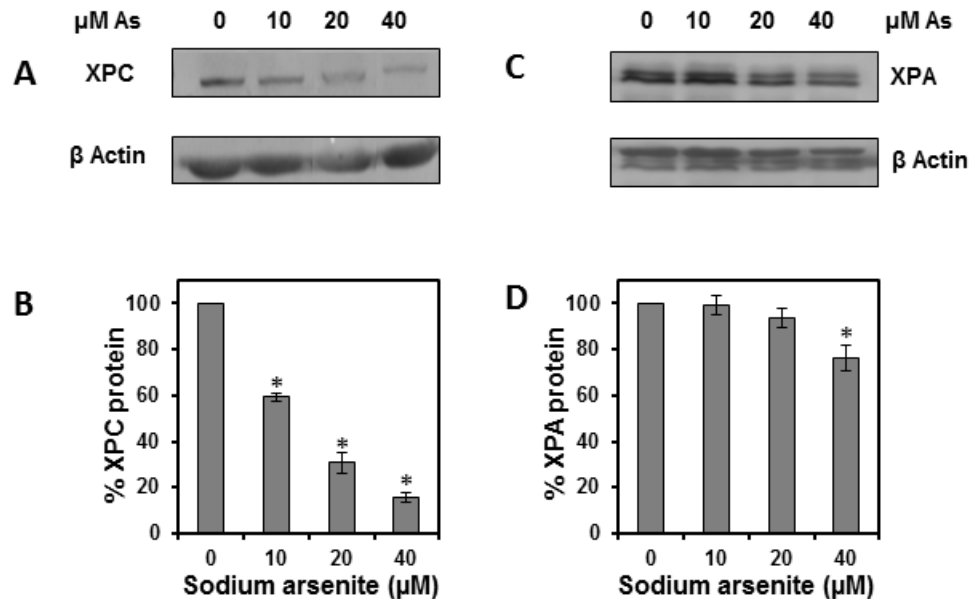


**Figure 4.4.** (A) IMR-90 cells were treated were treated with the indicated

concentrations of arsenite for 24 h, and the abundance of XPC was examined by

western blot analysis.  (B) A graphical representation of multiple western blots for

XPC expression is shown.  The data presented are the mean ± SE from two

repeats each of three independent experiments, and XPC expression was

normalized to β-actin. (C) IMR-90 cells were treated with the indicated

concentrations of arsenite for 24 h, and the abundance of XPA was measured by

western blot analysis.  (D) A graphical representation of multiple western blots for

XPA expression is shown. The data presented are the mean ± SE of three

repeats from three independent experiments, and XPA expression was

normalized to β-actin.  Treatments with arsenic that produced a significant

difference (P <0.05) compared to the untreated cells are designated with an

asterisk.

**Figure 4.5. The effect of sodium arsenite on the abundance of XPC and XPA RNA in IMR-90 cells.**
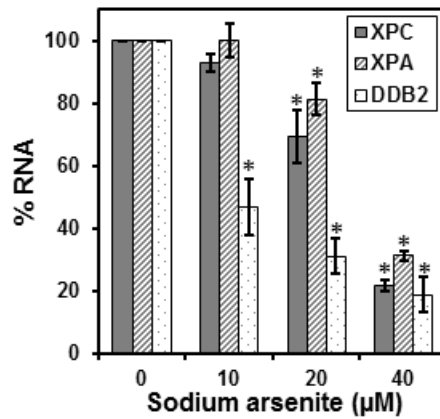


**Figure 4.5.** IMR-90 cells were treated with the indicated concentrations of arsenite for 24 h. RNA was isolated and Real Time PCR was performed as described in the methods section. The expressions of XPC, XPA, and DDB2 RNA were normalized to GAPDH RNA for each treatment. The data presented are the mean ± SE of two repeats of three independent experiments for XPC, and XPA, and one repeat of three independent experiments for DDB2. Treatments with arsenic that produced a significant difference (P <0.05) compared to the untreated cells are designated with an asterisk.

**Figure 4.6. Involvement of the proteasome in As-mediated XPC inhibition**
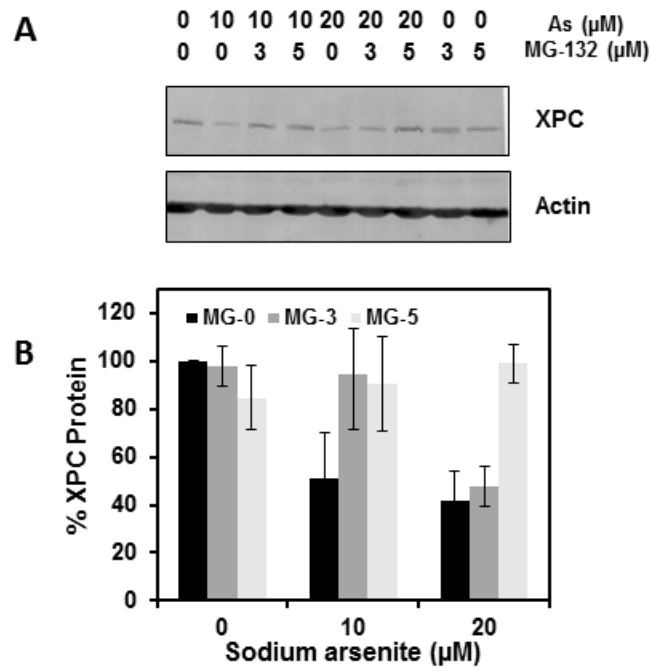


**Figure. 4.6.** (A) IMR-90 cells were treated with a combination of arsenite and/or the proteasome inhibitor MG-132 at the indicated concentrations for 16 h.  XPC levels were measured by Western blot analysis and normalized to Actin. (B) A graphical representation of multiple Western blots for XPC is shown.  The data presented are the mean of two repeats from two experiments (except MG-only treatments, which are the mean of two treatments of one experiment), and XPC expression was normalized to β-actin.

**Supplemental Table 4.1. Primers used in quantitative real-time PCR analysis**

| Primer name | Sequence |
|---|---|
| XPC-41 forward | 5'GCCAAAGGTCTGCTCATCA3' |
| XPC-41 reverse | 5'GGAGCTGCTGCCTCACTCT3' |
| XPC-35 forward | 5'GGGGCTACCATGAATGAAGA3' |
| XPC-35 reverse | 5'CCCAGCACAGGCTCACTAAG3' |
| XPA-71 forward | 5'CGAGTATCGAGCGGAAGC3' |
| XPA-71 reverse | 5'TTTGGGGCTGCTTTTACATT3' |
| XPA-63 forward | 5'GCAGCCCCAAAGATAATTGAC3' |
| XPA-63 reverse | 5'TCGCATATTACATAATCAAATTCCA3' |
| GAPDH-10 forward | 5'GCTGCATTCGCCCTCTTA3' |
| GAPDH-10 reverse | 5'GAGGCTCCTCCAGAATATGTGA3' |

**S. Table 4.1.** Multiple primer pairs were used to measure XPC and XPA RNA expression, and the resulting analysis did not produce any measurable difference between primer pairs within a single gene, so both primer pairs were included in the analysis of these genes.

**Supplemental Figure 4.1. Extending arsenite treatment time enhanced arsenite impairment of 6-4 PP removal**
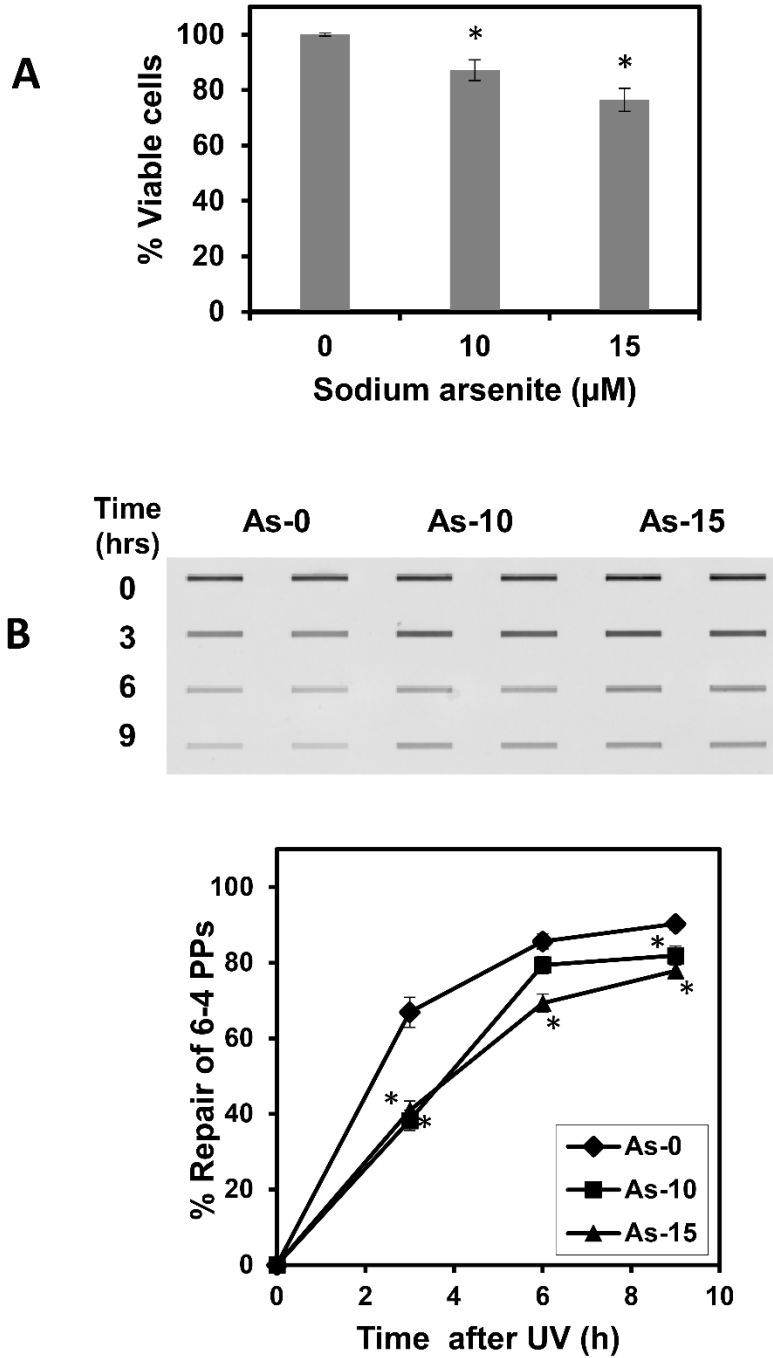
**S. Figure 4.1** A) IMR-90 cells were treated with the indicated concentrations of arsenite for 48 hours with the medium changed every 24 hours and the percentage of viable cells was measured using Trypan blue dye exclusion. The data represent the mean ± SE (Standard Error) from three measurements of one independent experiment.  (B) IMR-90 cells were treated with the indicated concentrations of arsenite for 48 hours, irradiated with 20 $J/m^2$ UV-C, and incubated in the same concentration of As-containing medium for the times shown. A representative slot blot and a graph of multiple immunoblots measuring removal of 6-4 PP are shown.  Each data point in the graph represents the mean +/- SE of three repeats from one independent experiment. Asterisks denote a significant difference between the treated sample and the control at a given timepoint, as measured by a 1-way ANOVA with a multiple comparison correction.

# Chapter 5

# Establishment of a reproducible measure of individual NER efficiency in peripheral blood mononuclear cells

## 5.1 Introduction

Nucleotide Excision Repair (NER) is a multi-protein complex designed to repair a variety of exogenous and endogenous DNA lesions and prevent mutation formation [244]. The NER pathway is divided into two sub-pathways that recognize different components of the genome. Global genomic NER (GG-NER) utilizes the XPC protein in complex with HR23B and Centrin-2 to recognize a remarkable array of DNA damage found on non-transcribed DNA, which is the majority of DNA [26]. Damage present on the transcribed strand of DNA is recognized by an RNA polymerase stalling at a transcription blocking lesion, leading to recruitment of downstream factors in the transcription-coupled NER (TC-NER) sub-pathway (reviewed in [24]). After DNA damage recognition, both pathways use common factors for the remaining steps. The TFIIH transcriptional complex is recruited to the site of damage to unwind the DNA around the lesion, with two helicases creating further opening the DNA, which permits subsequent NER factors to assemble at the site of the lesion. Two endonucleases, XPF/ERCC1 and XPG, then nick the damaged strand on the 5' and 3' sides of the lesion, respectively. The oligonucleotide containing the DNA adduct is removed and a DNA polymerase resynthesizes the gap using the undamaged strand as a template [25]. The NER pathway is described in more detail in chapter 1.

Individuals suffering from xeroderma pigmentosum (XP), one of three rare genetic disorders associated with deficiencies in NER, exhibit a pronounced increase in cancer incidence, both external and internal [25], with certain types of skin cancer showing incidence rates over a thousand fold greater than the general population [245]. XP is a hereditary autosomal recessive disorder, with XP A-G subtypes corresponding to the XPA-XPG proteins involved in NER [246]. The frequency of XP varies greatly between geographical regions, with a frequency of approximately 1 in 1,000,000 in Western Europe and the United States [247] but 1 in 22,000 in Japan [43]. Carriers of XP have been studied to determine the health consequences of this heterozygosity. XP-A heterozygotes have proficient DNA repair faculties [44], while XP-C heterozygotes have an intermediate phenotype regarding XPC mRNA and protein expression [46]. This may explain why low levels of XPA can correct the UV sensitivity of XP-A deficient cell lines [153] while reduced XPC expression correlates with reductions in repair of UV-induced photo-damage [248] and increased cancer risk for certain head and neck cancers [47] . Another study of the impact of NER gene expression levels on head and neck cancer risk indicated that XPC expression did not correlate to increased cancer risk, while ERCC1, XPB, XPG, and CS-B levels did [249], indicating that expression levels of individual NER factors may not be the most appropriate marker for cancer risk.

In addition to expression levels of NER factors, polymorphisms in genes involved in NER have also been examined for their potential to contribute to cancer susceptibility. The polymorphism frequency of XP genes varies

97

considerably (reviewed in [250]), so it is hard to give a single estimate of XP

polymorphism frequency. Despite that, XP polymorphism frequency is

sufficiently large to contribute to the population variation in DNA repair capacity

[251, 252]. Recent meta-analyses have reached different conclusions regarding

XP gene polymorphisms and cancer risks, with some showing certain XP

polymorphisms contribute to carcinogenesis [253-255], while others show no

relationship [256] or conflicting results from the same polymorphism between

studies [250, 257]. While certainly not conclusive, these results suggest that

certain XP gene polymorphisms may contribute to increased cancer incidence.

These studies also indicate there is a potential for many different genetic

polymorphisms to contribute to cancer risk. Given this, polymorphisms in one or

even a handful of genes may not fully reveal and individual's cancer

susceptibility. Instead, measuring DNA repair is an aggregate measure of

steady-state NER function and as a single value can be used as indicator of

individual cancer risk.

Previous studies have investigated DNA repair capacity (DRC) in isolated

lymphocytes using methods including the host-cell reactivation assay [246, 258-

260] and the comet assay [94, 261-263]. Several studies have used the

measurement of DRC in circulating lymphocytes as an indicator of individual

cancer risk [259, 264-266]. These lymphocytes are used as surrogate cells when

investigating DRC in a target tissue because of the non-invasive manner by

which they can be obtained and because target tissue may have undergone

exposure-mediated somatic changes [19]. Additionally, a positive correlation

between DRC in lymphocytes and target tissues in XP patients has been observed [267], reinforcing the utility of lymphocytes as an appropriate measure of patient DRC.

Current research on DNA repair of isolated human lymphocytes often employs the host cell reactivation (HCR) assay, which functions by measuring repair of a damaged plasmids transfected in to the cell.  The HCR assay also requires transcription of the reactivated gene on the plasmid that is being measured as an indicator of effective repair, involving other pathways beyond DNA repair in order to measure repair.  The immuno-blot assay we have adapted to work with isolated lymphocytes accurately and reproducibly measures NER efficiency directly, and the lesions introduced are done so in a stochastic manner that does not involve cellular uptake or metabolism.  Also, the output value from a host-cell reactivation assay is a percentage of activity of the reporter gene. It does not provide any information about NER kinetics

An additional method for measuring DNA repair in isolated lymphocytes, the Comet assay, can be adapted to specifically measure repair of UV-induced photolesions.  The comet assay uses an electronic charge to separate double-stranded DNA from single-stranded DNA.  Greater quantities of DNA damage produce larger amounts of single-stranded DNA.  In order to make the assay specific to one type of DNA damage, like UV-induced photolesions, damage-specific endonucleases can be added.  T4 endonuclease V cuts at sites of UV-induced CPDs, and as such provides a manner with which one can adapt the Comet assay to specifically measure the amount of CPDs present in a cell.

However, the 6-4 PP lesion is repaired much more proficiently than the CPD lesion in human cells, and as such greater variations in removal of the 6-4 PP lesion can be seen in a shorter period of time than compared to removal of the CPD lesion.

Recent work in our lab has provided insights into how the NER pathway can be modified by exposure of cells in culture to certain carcinogens such as tobacco smoke [268] and arsenic (Chapter 4). These studies suggest that human exposure to these agents may contribute to changes in NER efficiency, but a human study has not been performed using the same repair assay, and it is unknown what other factors may affect NER efficiency in a human population. Herein we report our findings pertaining to the establishment a novel method for measuring DRC in isolated lymphocytes using an assay to measure NER efficiency, specifically the sub-pathway known as global genomic NER, using a modified immuno-slot blot repair assay. We found reproducible measurements of NER efficiency across multiple repeated measures from the same sample, as well as between experiments performed on multiple batches of lymphocytes isolated from the same individual and even between different blood draws from the same individual. Additionally, we observed a distribution of repair values in our study population that was similar to DRC distributions found using the HCR assay. Our findings suggest this method is a reliable measurement of individual DNA repair capacity and that it is capable of capturing population variation in DNA repair, which indicate that the measurement of NER efficiency may be a

suitable biomarker of cancer risk and may be used to study the effect of environmental exposures on NER efficiency.

## 5.2 Results

### 5.2.1 Statistical analysis

The immunoblot repair assay provided a repair curve that showed the percentage of 6-4 PPs removed at different timepoints after irradiation. We determined that the most effective way to present this data was to create a single value that could explain the repair efficiency of each individual. We chose to calculate a half-life statistic, a value that represented the time, in hours, that was required by each individual to remove 50% of the introduced photolesions. Briefly, an immunoblot (Fig. 5.1A) is used to calculate the percentage of photolesions that have been removed at each timepoint, and these values are plotted on a linear axis (Fig. 5.1B). This is then converted from percentage lesions removed to percentage lesions remaining, and those values are plotted on a log-scale y-axis (Fig. 5.1C). When a log-scale axis plot such as this produces a linear graph, the data is best fit to an exponential decay curve. So, an exponential decay regression analysis was performed and used to calculate "half-life." The process of calculating this value from the regression analysis is explained in detail in chapter 6.

Normality tests for population distribution of half-life were performed using SAS 9.3 software, and comparisons of half-lives between populations were done using Graph Pad Prism 6, using a 2-sided Student's T-test with a threshold of significance of $P < 0.05$.

### 5.2.2 Study population

The subjects for this study were part of a case-control study examining lung cancer rates in Appalachian Kentucky. Lung cancer patients of the case-control study were identified through the Kentucky Cancer Registry, and healthy control subjects were identified from the Kentucky voter registration records. Control subjects were age and gender matched to the lung cancer cases using a frequency matching method. All participants were required to be over 17, and had to have no prior history of cancer diagnosis, excepting stage I or II non-melanoma skin cancers. They were required to have a working phone, to be willing to consent to an in-person interview, and to be able to speak English without the use of an interpreter. The individuals whose lymphocytes were analyzed for repair represent a random sampling of the subjects in the case-control study at a 1:3 ratio of lung cancer cases to control subjects. In all, 42 cancer patients and 156 control subjects were analyzed for DRC. A greater number of subjects provided blood samples, but low lymphocyte yields from the blood or high non-lymphocyte contamination of the buffy coat prevented us from performing the repair analysis on those samples. Initially, additional 6 lung cancer subjects were part of the study population, but they received radiation therapy. After it was determined that those who received radiation therapy had a significantly different (higher) average half-life than those who did not, they were removed from the study population. In fact, only 25% of samples from individuals who received radiation therapy were stimulated and provided a measurable repair value – the remainder had insufficient lymphocyte yields to perform the

stimulation or did not respond properly to the stimulation process. We also evaluated half-life values from individuals who received chemotherapy and those who had undergone surgery related to their lung cancer, and neither of those populations had a different half-life than the remainder of the cancer subjects, so they remained in the study (Supplemental Figure 5.1). Additionally, a smaller second study of 33 lung cancer patients from Kentucky was analyzed using the same repair assay, and the data from that assay was used to confirm the levels of experimental variation in individual DRC measurements. The repair assay was performed blindly to reduce any potential biases in conducting the repair experiments. A statistician with information about the demographics of the study participants provided us with a list of samples to stimulate that was age and gender matched between cases and controls (although some exclusions after the analysis shifted the gender distribution slightly). The identification of the subjects as lung cancer cases or controls was withheld until the analysis was completed, and were revealed after completion of the repair analysis in order to perform comparisons between cases and controls, which will be discussed in Chapter 7.

### 5.2.3 Experimental variation of repair capacity

Of the 198 samples analyzed, the average coefficient of variation (CV, standard deviation/average) of the half-life was 11%, slightly higher than a reported CV using HCR as a measure of repair in isolated lymphocytes [258], but considerably lower than more recent reporting of 23% [266] and 26.2% [269]. The population showed a wide range of NER efficiency, with half-lives ranging from 1.5 to10.7 hours. The average and standard deviation are 4.17 and 1.77

hours respectively.  A distribution of the data indicates that the data seems fairly normally distributed with a slight right tail (Fig. 5.2A).  However, a Kolmogorov-Smirnov (K-S) statistical test for normality indicates that it is not normally distributed (p < 0.01).  Certain statistical analyses such as ANOVA assume a normal distribution (although they can tolerate minor deviations from normality) and therefore skewed data are often transformed with a mathematical operation (such as a natural-log transformation) to normalize the data set.  The distribution of the natural-log transformed half-life values (Fig. 5.2B) is normally distributed using the same K-S test for normality (p > 0.150).  Thus, when needed, a normal data set can be generated from the half-life values produced by the repair assay, strengthening value of the repair assay as it relates to the types of analyses that can be performed on the half-life values generated.

### 5.2.4 Replication of the repair assay

Experimental reproducibility is critical.  We performed two independent stimulations of lymphocytes isolated from twelve study participants (7 from the larger study and 5 from a smaller, second study).  These stimulations were tested independently for their half-life, and Table 5.1 shows the half-life and standard error of these repeat stimulations, along with a p-value from a Student's T-test comparing the two half-life values.  The first and second stimulations of these lymphocytes produced half-life values that were not statistically different.  Of particular interest, one study participant (individual 12 in Table 1) had blood drawn on two separate occasions, and the repeat stimulations came from these

separate blood draws, suggesting the half-life value from the repair assay is reproducible across different blood draws of the same individual.

## 5.2.5 Impact of different experimental parameters on NER efficiency

We examined the effect of three recovery parameters - blood volume, lymphocyte yield, and DNA recovery – on the half-life value produced by the repair assay. The lymphocyte yield is the number of cells in a milliliter of blood counted before the cells are frozen for storage, and the DNA recovery is the concentration of DNA recovered from the first timepoint in the repair assay divided by the number of cells stimulated for that timepoint. These three factors had quite large ranges: the volume of blood collected for the study ranged from 5.5 ml to 18.1 ml, the amount of lymphocytes recovered from this blood ranged from 0.76 to 5.84 million cells/ml blood, and the DNA recovery ranged from 0.41 to 4.07 micrograms of DNA/million cells. With these large ranges, particularly the lymphocyte and DNA recoveries, it was unknown what effect, if any, they might have on the reliability of the assay. These recovery parameters were plotted against the half-life of each individual in the study. A graph showing the relationship between blood volume and half-life (Fig. 5.3A) indicates there is no significant impact on half-life due to the blood volume. A linear regression analysis of the scatter plot provided an $R^2$ value of 0.0148 (p = 0.09). Likewise, a comparison of lymphocyte recovery to half-life (Fig. 5.3B) also indicates that cell recovery did not have a significant impact on half-life. A linear regression analysis provided an $R^2$ value of 0.018 (p = 0.06). Interestingly, a plot of DNA recovery verses half-life showed a minor reduction in half-life (increase in repair)

with increasing DNA recovery (Fig. 5.3C).  While technically significant (p = 0.037), the linear regression equation indicates that an increase in DNA recovery of 2 µg DNA per 1 million cells stimulated (over half of the entire range of DNA recoveries in the study) only produces a 1-hour increase in half-life, and it is unclear whether the DNA recovery was a cause or a result of changes in NER efficiency.

## 5.2.6 Impact of the time delay between blood draw and lymphocyte isolation on NER efficiency

Previous reports have suggested that there is an optimal maximum time delay between blood draw and lymphocyte isolation, in order to maximize post-thaw cell viability and response to mitogen stimulation.  This delay is the most important factor affecting cellular recovery and viability [270], with lymphocyte recovery dropping substantially when samples sit for 24 hours rather than 8 hours before isolation.  It was unknown, however, how this delay would affect cellular DNA repair using our repair assay, so we compared the time delay between blood draw and lymphocyte isolation to several experimental parameters.  Increasing the time delay increased the number of cells recovered per milliliter of blood provided (Fig. 5.4A).  This may be misleading, however, as non-monocyte contamination may have increased in the buffy coats of samples that were delayed in their shipment [271].  Prolonged storage of blood alters granulocyte buoyancy and impairs proper separation from mononuclear cells by use of a density gradient, increasing granulocyte contamination in the monocyte buffy coat layer [272].  Increasing processing time also reduced DNA recovery at

the first timepoint in the repair assay, perhaps a consequence of a reduced

response to mitogen stimulation in samples with longer processing time (Fig.

5.4B).  Despite the change in cell count at isolation and DNA recovery during the

repair assay, there was no change in NER efficiency based on processing time

(Fig. 5.4C).

The majority of samples (n = 179) had time delays that clustered around

24 hours. A small number of samples (n = 19) had time delays that clustered

around 48 hours, a result of a shipping issue that was corrected after the sample

collections were underway.  As most studies do not exceed a 24 hours

processing delay, this provided a unique opportunity to observe the

consequences of such a delay on the outcome of the DNA repair assay. There

was a substantial increase in cellular recovery in the samples that sat a day

longer compared to those that did not (3.14 vs 2.14 million cells/ml blood,

$p<0.0001$) (Fig. 5.4A). Additionally, there was a substantial decrease in the

average DNA recovery in the samples that sat longer compared to those that did

not (1.04 vs 1.51 µg/million cells, $p=0.002$) (Fig. 5.4B).  These changes in

recovery did not have an effect on DNA repair, however. The average half-life of

the samples that took an extra day to process was not significantly different than

those isolated after being kept on ice one night.  The samples that did not have

an extra day delay had an average half-life of 4.16 hours with a standard

deviation of 1.80 hours, compared to those that did have an extra day delay,

which had an average half-life of 4.27 hours with a standard deviation of 1.44

hours ($p=0.80$) (Fig. 5.4C).

## 5.2.7 Impact of storage time and storage and isolation dates on NER efficiency

After isolating the lymphocytes from the blood samples, the cells were stored in liquid nitrogen. The duration of the storage varied from a few days to almost three years, as experimental parameters were still being established when the lymphocyte collection began. Samples therefore remained in liquid nitrogen storage until the stimulation protocol was fully completed, and were removed in batches for stimulations. Previous DRC studies using the host-cell reactivation assay have suggested that the length of time the lymphocytes are cryopreserved could potentially modulate DNA repair [260, 273], so the time between lymphocyte isolation and stimulation was compared to half-life of each individual in the study population to look for a possible relationship. Half-life was not impacted by the length of lymphocyte storage in liquid nitrogen (Fig. 5.5A). A linear regression analysis of the scatter plot produced an $R^2$ value of 0.128 (p =0.113). Additionally, we examined the chronology of our isolations and stimulations to determine if some factor, such as a change in the batch of some processing material, may have caused a temporal change in our repair assay, as temporal variation in DRC had been previously suggested [273]. Across three years of isolations and stimulations, there was little impact in half-life based on when the sample was isolated or when the sample was stimulated for repair (Fig. 5.5B-C). Linear regression analyses of the graphs of isolation date and stimulation date plotted against half-life produced $R^2$ values of 0.0016 (p = 0.576) and 0.0147 (p = 0.089) respectively.

## 5.3 Discussion

Measuring individual DNA repair capacity (DRC) from isolated lymphocytes in a human population as a means of estimating cancer risk was first proposed 25 years ago [258]. The classical methods employed to measure individual DRC from isolated lymphocytes have been the host-cell reactivation assay and variations on the comet assay [274-277]. Our findings indicate that DNA repair capacity can be measured using an alternative assay to those presently available, employing cellular machinery to repair DNA damage. The DNA repair assay utilized here is designed to monitor the removal of DNA lesions introduced by UV-C exposure. These lesions are removed exclusively by the nucleotide excision repair pathway, making the assay a specific measure of NER efficiency. We have reported here the successful implementation of this repair assay on a human population for the first time.

We examined the impact of several factors that had the potential to confound our measurement of NER efficiency. As this assay had not been implemented in a human population before, it was necessary to measure the reliability of the repair assay. To that end, we tested recovery parameters, temporal fluctuation on the repair process, and multiple stimulations of the same samples to look at how well the half-life measure of NER efficiency performed. We observed that the processing time, stimulation and isolation dates, and time between stimulation and isolation did not contribute to a significant variation in NER efficiency, as measured by the repair assay endpoint, half-life. Additionally, blood volume and lymphocyte recovery did not have an impact on half-life. Blood

volumes as low as 5 ml provided NER efficiency measures that were no different than 15 ml samples. This is important as it provides us with an idea of how much biological material must be consumed in order to get a reliable measure of NER efficiency from an individual.

The time required to process the samples after the blood was drawn did not significantly impact NER efficiency. It did, however, have a significant impact on two other experimental parameters, cell recovery per volume of blood and DNA recovery per number of cells stimulated. The cell recovery was increased and the DNA recovery was decreased as processing delay increased. We observed that samples whose processing time was one day longer than the normal time had a higher incidence of red blood cell contamination in the buffy coat during lymphocyte isolation. Based on the observation that samples with potential contaminating cells were difficult to stimulate, we concluded that DNA from nucleated cells present in the sample was causing the difficulty with our stimulation. Not all samples that were delayed an extra day during the shipping process were affected in this manner, and those unaffected seemed to have no discernable difference in repair compared to samples shipped properly, but the increased incidence of a failure to stimulate among samples that did have extended shipping time highlights the importance of proper shipping time in order to maximize the success of the DNA repair assay.

Our results here indicate that NER efficiency can be directly measured from isolated lymphocytes from blood samples, preferably when the isolation occurs within 24 hours of the blood draw. This is a longer period of time than

most studies, as previous measures of DNA repair from isolated lymphocytes have typically drawn blood at the same location where lymphocytes are isolated and cryopreserved, or the shipping time has been kept to a minimum. Our findings indicate that a study in which blood samples are provided remotely and shipped to a central location for processing can achieve a reproducible measure of repair.

## 5.4 Conclusions

The ability of our cells to remove DNA damage in a timely manner is paramount to survival. The NER pathway responds to an array of environmental and endogenous mutagens, repairing damaged DNA and preventing cancer-causing mutations from forming. We observed a wide range in individual NER efficiencies within the study population. Fluctuations in repair efficiency may provide an early indicator for cancer risk. Those individuals with a low NER efficiency may be at a higher risk of failing to repair DNA damage, which could result in an increased risk of developing cancer.

**Table 5.1 Repeated measurement of NER efficiency in isolated PBMCs**

| Individual | Half-life 1 | SE | Half-life 2 | SE | p-value |
|---|---|---|---|---|---|
| 1 | 3.32 | 0.27 | 3.18 | 0.5 | 0.88 |
| 2 | 5.14 | 1.34 | 4.37 | 0.18 | 0.39 |
| 3 | 3.9 | 0.33 | 3.12 | 0.38 | 0.18 |
| 4 | 1.57 | 0.14 | 1.82 | 0.13 | 0.2 |
| 5 | 6.3 | 0.55 | 6.34 | 1.06 | 0.85 |
| 6 | 4.56 | 0.38 | 4.85 | 0.95 | 0.61 |
| 7 | 7.42 | 2.09 | 5.07 | 1.07 | 0.28 |
| 8 | 3.11 | 0.4 | 3.24 | 0.52 | 0.54 |
| 9 | 3.64 | 0.26 | 3.17 | 0.36 | 0.43 |
| 10 | 3.34 | 0.46 | 2.69 | 0.24 | 0.25 |
| 11 | 2.62 | 0.05 | 2.71 | 0.2 | 0.64 |
| 12 | 2.33 | 0.15 | 2.47 | 0.29 | 0.62 |

**Table 5.1** Lymphocytes from twelve individuals were stimulated two different times and DRC was measured for each stimulation.  Half-life values are the average of three repeated slot blots of the same DNA.  Half-life and standard error (SE) are both in units of hours.  P-value taken from a two-sided Student's T-test.

**Figure 5.1 DNA repair assay performed on lymphocytes isolated from whole blood**
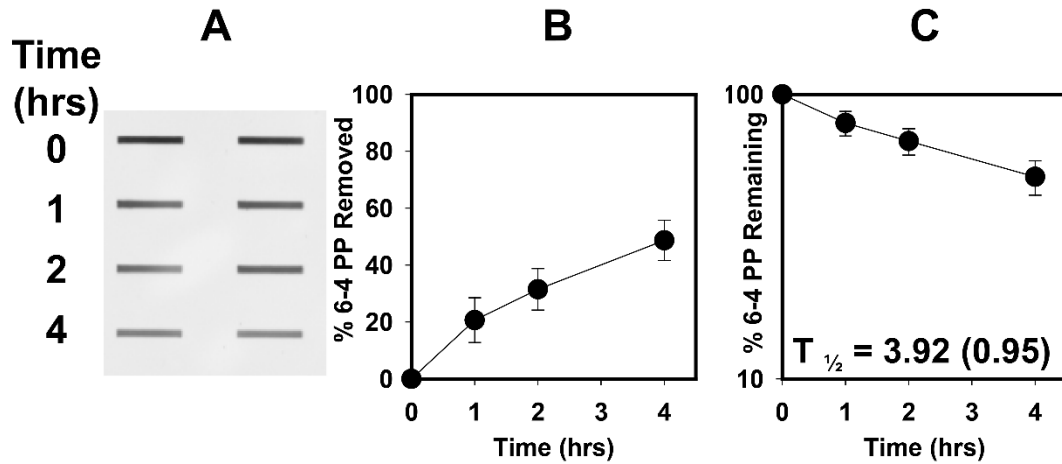


**Fig 5.1** (A) The results of one slot blot from the DNA repair assay. Equal

amounts of DNA isolated from each indicated timepoint after irradiation are

slotted in duplicate, and an antibody specific to 6-4 PPs is used to visualize

removal of DNA damage over time. (B) The percent 6-4 photoproducts removed

are calculated from three repeats of the slot-blot per individual, with the average

repair and standard error at each timepoint plotted on the graph. (C) The percent

repair values are converted into % 6-4 photoproducts remaining at each

timepoint, and the average and standard error values are plotted as shown.

From this data, an exponential decay regression analysis was performed to

generate the half-life (and standard error) values shown.

**Figure 5.2 The distribution of NER efficiency across the study population.**
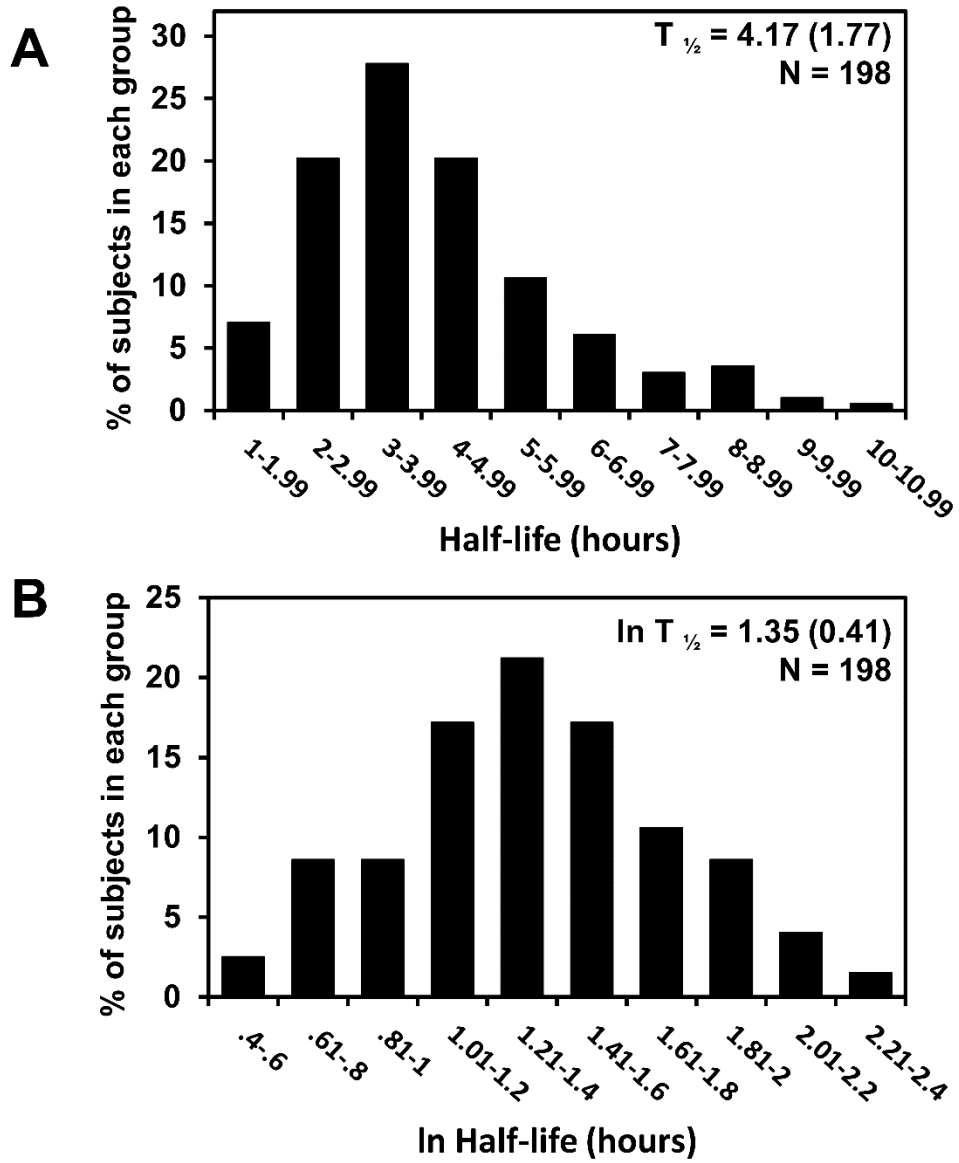
**Fig 5.2** (A) A Histogram of half-life distribution across the entire population is shown, as well as the average and standard deviation of the population half-life. The NER efficiency of the study population has a predominantly Gaussian distribution but is skewed towards slower repair values (right tailed).  The Kolmogorov-Smirnov (KS) test for data normality indicates that the half-life values are not normally distributed ($p < 0.0001$).  (B) In order to produce a normally distributed data set, a natural log transformation of the data was performed, and a histogram of the natural log transformed data is shown, along with the average and standard deviation of the transformed half-life of the study population.  The KS normality test confirms that the natural log-transformed half-life values are normally distributed ($p = 0.3807$)

# Fig. 5.3 The impact of collection parameters on half-life.
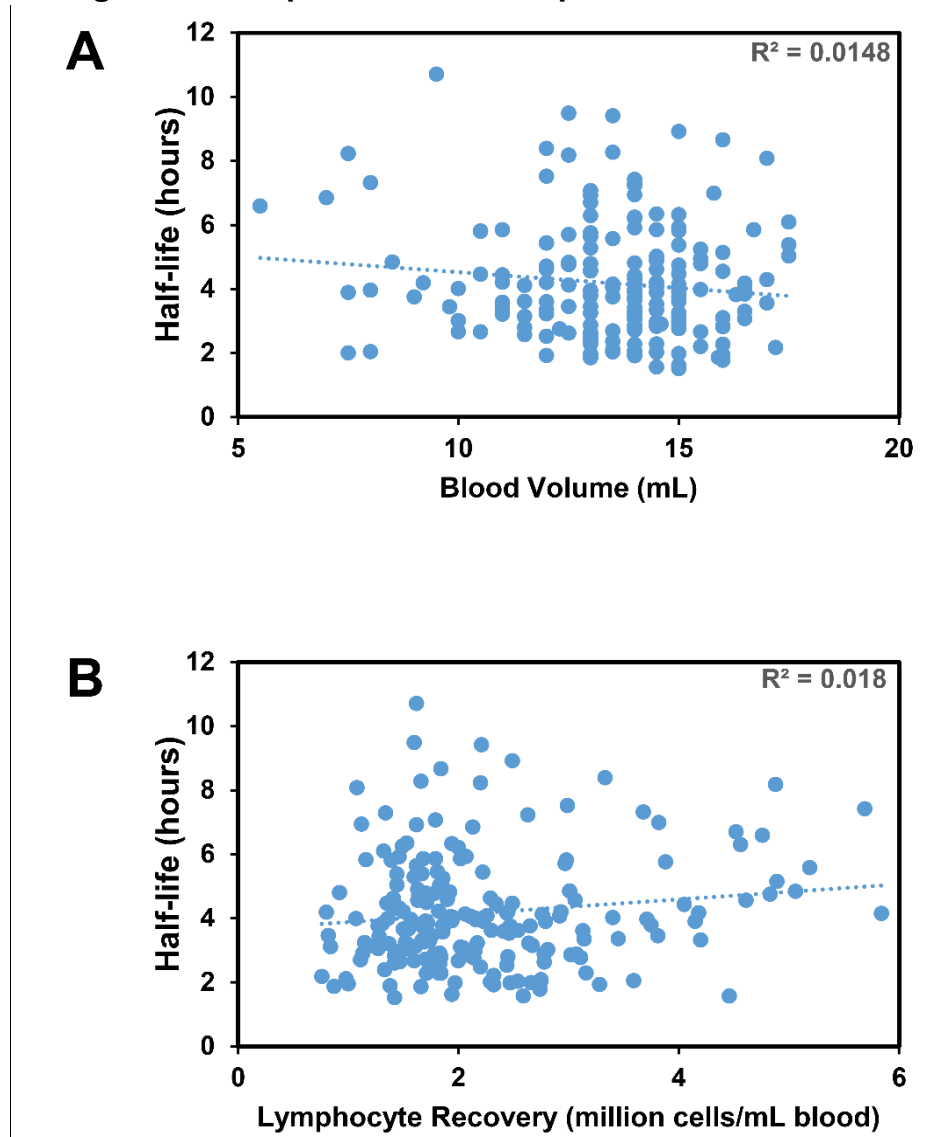


**Fig 5.3** (A) A scatter plot of blood volume verses half-life shows a slight reduction in half-life with increasing sample volume, but the trend is not significant. A linear regression analysis of the data produced an $R^2$ value of 0.0148 (p =0.09). (B) A scatter plot of lymphocyte recovery verses half-life shows a slight increase in half-life with increasing lymphocyte recovery, but the trend is not significant. A linear regression analysis of the data produced an $R^2$ value of 0.018 (p =0.06).

**Figure 5.4 The impact of processing delay on experimental parameters.**
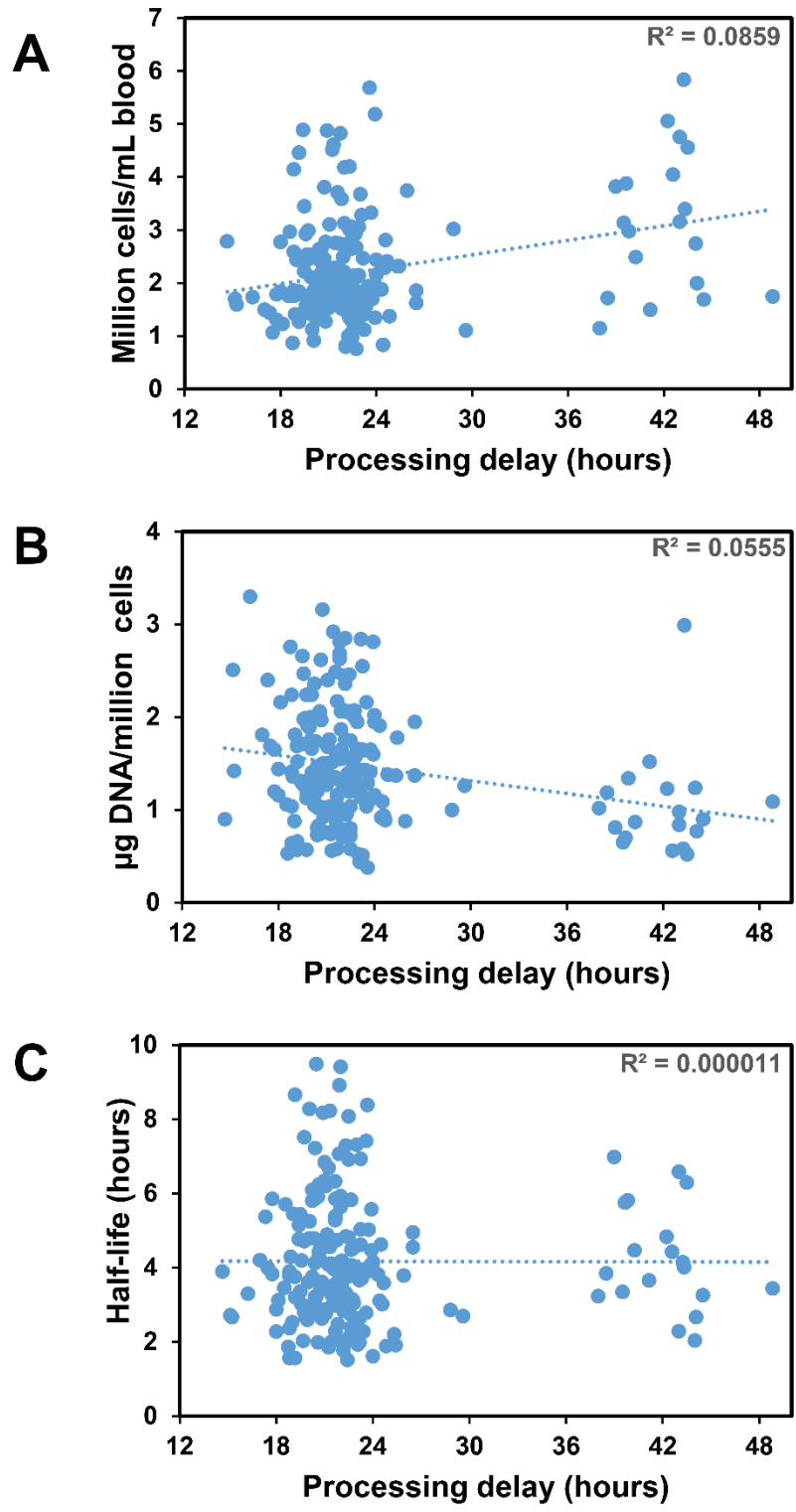
**Fig 5.4** (A) A scatter plot of processing delay verses lymphocyte recovery shows

a significant increase in cells recovered with increasing processing delay. A

linear regression analysis of the data produced an $R^2$ value of 0.0859

(p =0.00003). (B) A scatter plot of processing delay verses DNA recovery

indicates that higher delay times significantly reduce DNA recovery, perhaps as a

result of an increase in non-lymphocyte cell contamination with increasing

processing delay. A linear regression analysis of the data produced an $R^2$ value

of 0.0555 (p = 0.0008). (C) A scatter plot of processing delay verses half-life

shows that despite changing recovery parameters, the time delay between blood

draw and lymphocyte isolation does not have any impact on half-life. A linear

regression analysis of the data produced an $R^2$ value of 0.000011 (p = 0.963).

**Figure 5.5 The impact of storage time and processing dates on half-life.**
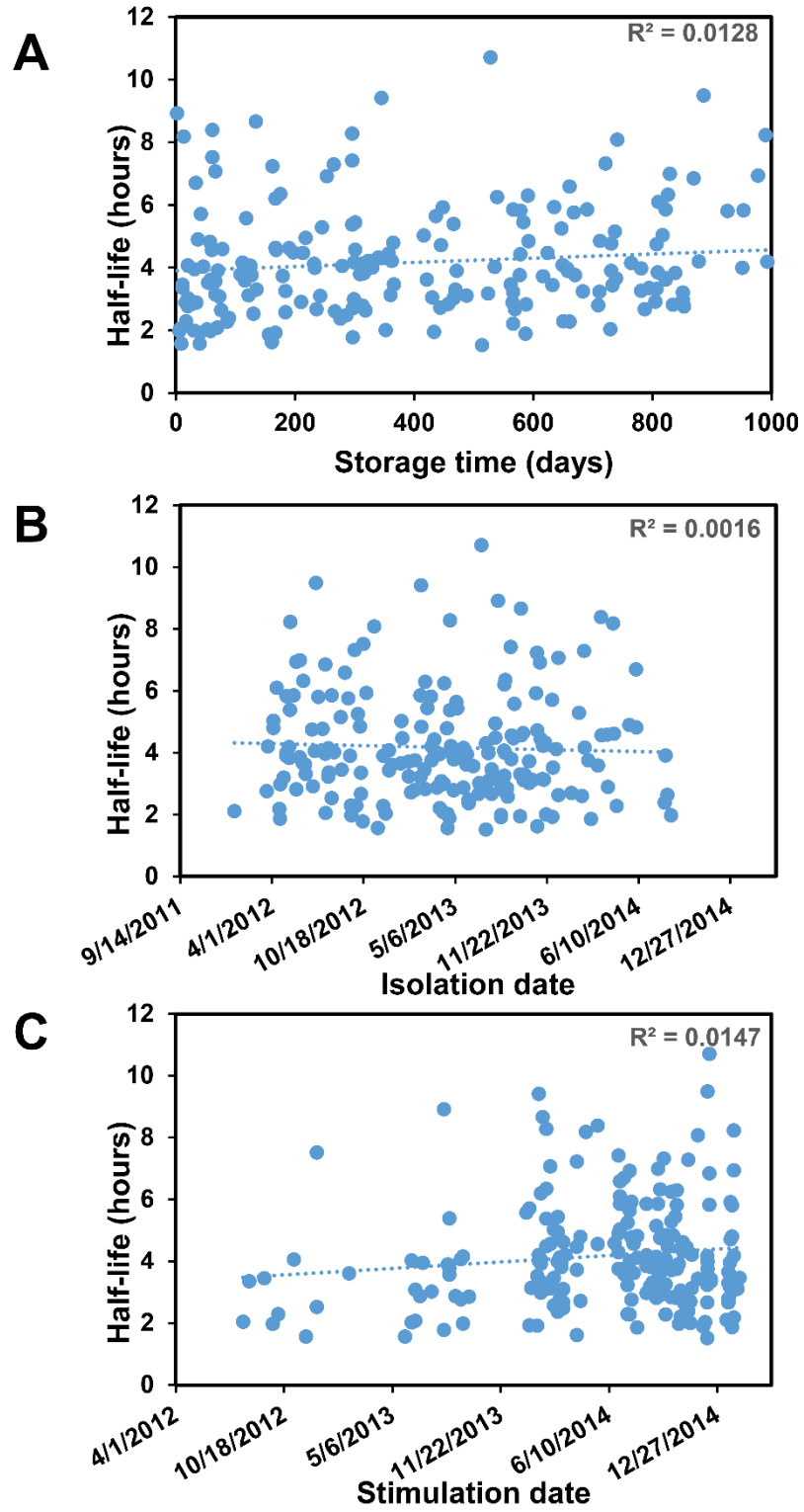
**Fig 5.5** (A) A scatter plot of processing delay verses half-life shows that despite changing recovery parameters, the time delay between blood draw and lymphocyte isolation does not have any impact on half-life. A linear regression analysis of the data produced an R2 value of 0.0128 (p = 0.113). (B) A scatter plot of isolation date verses half-life shows no temporal effect on the repair endpoint. A linear regression analysis of the data produced an R2 value of 0.0016 (p =0.576). (C) A scatter plot of stimulation date verses half-life shows no significant change in half-life across the two year period in which stimulations were performed. A linear regression analysis of the data produced an R2 value of 0.0147 (p = 0.089).

**Supplemental Fig. 5.1 The impact of cancer treatment therapies on half-life.**
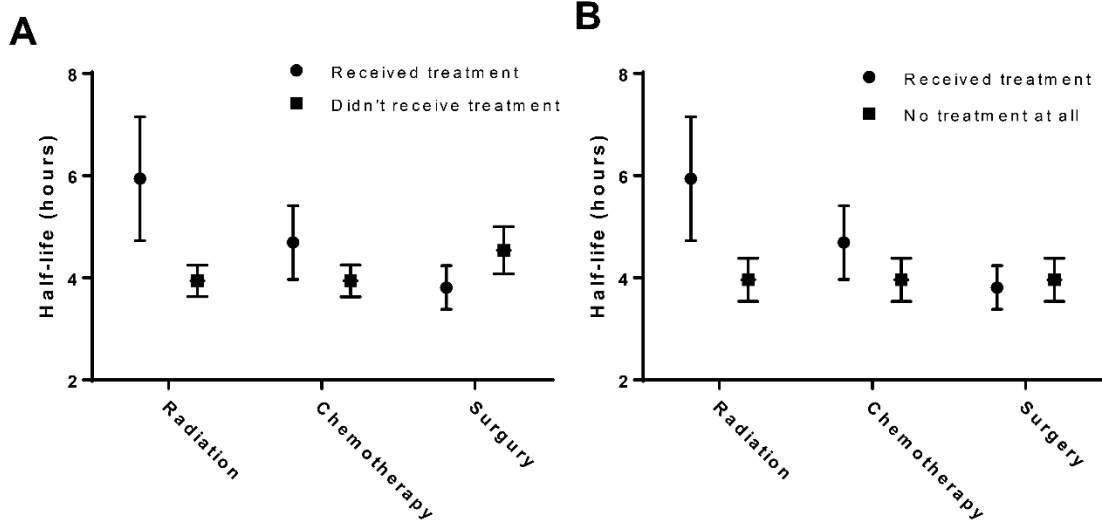


**C**

| Factor | Lung Cancer Subject Type | | | | P-value |
|---|---|---|---|---|---|
| Therapy | Received therapy | | Did not receive therapy | | |
| | N | Mean ± SD | N | Mean ± SD | |
| Radiation | 6 | 5.94 ± 2.97 | 42 | 3.94 ± 1.99 | 0.036 |
| Chemotherapy | 16 | 4.69 ± 2.89 | 32 | 3.94 ± 1.76 | 0.270 |
| Surgery | 23 | 3.81± 2.05 | 25 | 4.54 ± 2.32 | 0.256 |

**D**

| Factor | Lung Cancer Subject Type | | | | P-value |
|---|---|---|---|---|---|
| Therapy | Received therapy | | Did not receive any therapy | | |
| | N | Mean ± SD | N | Mean ± SD | |
| Radiation | 6 | 5.94 ± 2.97 | 12 | 3.96 ± 1.47 | 0.0725 |
| Chemotherapy | 16 | 4.69 ± 2.89 | | | 0.4319 |
| Surgery | 23 | 3.81± 2.05 | | | 0.8238 |

**Supplemental Fig. 5.1** (A) Lung cancer subjects were placed into groups based on whether or not they received a particular treatment – radiation therapy, chemotherapy, or surgery. Half-life values and standard errors were determined for each population and are shown.  (B) A subset lung cancer subjects were separated into groups of individuals who received a particular treatment and those who received no treatment.  Half-life values and standard errors are shown. (C) P-values were obtained by comparing individuals who did or did not receive a particular type of treatment (as shown in (A). (D) P-values were obtained by comparing individuals who received a particular treatment and those who received no treatment at all (as shown in (B)).

# Chapter 6

# Variations in nucleotide excision repair efficiency in a human population

## 6.1 Introduction

Maintenance of genomic fidelity is critically important to the survival of living organisms.  DNA damage, from endogenous and exogenous sources, is a constant, huge burden on the cell.  If left unrepaired, this damage can result in DNA mutations, leading to an increase in cancer risk, or the damage can trigger apoptotic cell death, an event associated with organismal aging.  In humans, several multi-protein pathways contribute to the maintenance of genomic fidelity, including the nucleotide excision repair (NER) pathway [213].  The NER pathway is responsible for recognizing and removing a wide variety of structurally diverse DNA lesions.  Substrates for NER include adducts generated by metabolites of chemical carcinogens such as polycyclic aromatic hydrocarbons in tobacco smoke, UV-induced photoproducts, and platinum adducts formed by the chemotherapeutic agent cisplatin [19].

In humans, more than twenty different proteins are involved in the NER pathway.   These proteins include the XPA-G factors that are singly defective in seven corresponding complementation groups of the human disorder Xeroderma Pigmentosum (XP) [25].  The tumor suppressor factor p53 also impacts NER efficiency, likely through transcriptional regulation of the *XPC* and *DDB2* gene products, both of which are DNA damage recognition proteins [20, 21, 114, 115]. The NER pathway is comprised of two sub-pathways that differ in their

mechanism of damage recognition: global genomic NER (GG-NER) which recognizes and removes DNA damage from anywhere in the genome and transcription-coupled NER (TC-NER) which selectively removes damage from the transcribed strands of actively transcribed genes.  In GG-NER, DNA damage recognition is accomplished by XPC, which recognizes helically distorting lesions and is stabilized by its binding partners RAD23B, and CENTRIN2 [26, 27]. Additionally DDB2 assists XPC in recognition of certain DNA photolesions such as the cyclobutane pyrimidine dimer [142], which is not as well recognized as the 6-4 photoproduct (6-4 PP) by XPC [34].  In TC-NER, damage is recognized by the stalling of the RNA polymerase complex at the site of damage during transcription (reviewed in [24]).  After the DNA damage recognition step, many of the subsequent steps are the same for GG-NER and TC-NER. The helicase activities of TFIIH produce additional unwinding of DNA where upon the endonuclease activities of the XPF/ERCC1 complex and XPG produce single-strand incisions flanking the damaged site.  The original integrity of the DNA is restored after an approximately 30 nucleotide region of DNA containing the lesion is removed, and the gap is filled by pol δ or pol ε, using the undamaged strand as a template (reviewed in [25]).

Genetic deficiencies in NER are the origins of rare human disorders including Xeroderma Pigmentosum (XP), Cockayne Syndrome, and Trichothiodystrophy [39].  Of these, XP is characterized by dramatically increased skin cancer rates as well as increased internal neoplasms [40, 211, 278].  While XP is rather uncommon, polymorphisms in NER genes can have

much greater incidences [257, 279, 280]. Several of these polymorphisms have been associated with increased risk of developing cancer, including colorectal [281], prostate [282], liver [283], esophagus [284], lung [285], and skin cancers [279, 286].  In addition, the expression levels of several NER proteins have been associated with increased cancer risk [249, 287].

As these studies show, there are many factors controlling the proteins involved in NER that have the potential to contribute to cancer risk.  As such, a functional analysis of population variation in DNA repair capacity may be of greater value in predicting cancer risk. To this end, several studies have employed functional assays to investigate the variation in DNA repair within a human population.   The two most common methods of measuring population distribution of DNA repair have been using the Comet assay [274, 277, 288] and the host cell reactivation (HCR) assay [258, 259, 269, 289].  The Comet assay relies on measuring the presence of DNA strand breaks as an indicator of DNA repair.  Combining a damage source with an endonuclease specific to an adduct generated by that, such as the use of bacterial T4 endonuclease V to cleave at sites of CPDs, allows the Comet assay to measure the repair of a specific GG-NER substrate [290]. The HCR assay, on the other hand, measures repair of a transcriptionally active gene, typically on a transfected plasmid, via detection of the gene product, and as a result the HCR assay is actually measuring the TC-NER subpathway of NER.

We have employed a modified immunoblot assay to measure NER efficiency in isolated peripheral blood mononuclear cells (PBMCs) in a human

population.  The repair assay measures the removal of a UV-induced photolesion, the 6-4 photoproduct.  The 6-4 photoproduct (6-4 PP) and cyclobutane pyrimidine dimer (CPD) DNA adducts, produced by UV light, are model substrates for measuring NER activity as they are rapidly generated by a brief exposure to UV light and are repaired exclusively by NER in humans [25]. 6-4 PPs are repaired much more quickly than CPDs, likely due to the degree with which the two lesions are recognized by NER [291], and as such we chose to measure 6-4 PP removal in individuals to maximize detection of variations in NER efficiency.

We explored the efficiency of NER in individuals residing in the Appalachian region of Kentucky in preparation for a larger case-control study investigating lung cancer incidence in the region.  Appalachian Kentucky has some of the highest cancer rates in the nation, including a lung cancer rate (109.2 per 100,000, http://cancer-rates.info/ky/index.php) that is almost double the national average (59.4 in 100,000, https://nccd.cdc.gov/uscs/toptencancers.aspx).  We observed an age-dependent reduction in NER efficiency in the study population, and a slight tendency for females to have higher NER efficiencies than males. We also observed that the age-dependent reduction of NER efficiency can be modulated by smoking status.  These findings will aid in implementing this repair assay in future studies investigating the relationship between DNA repair and cancer susceptibility.

## 6.2 Results

### 6.2.1 Statistical Analysis

The results of three slot blots obtained from a single biological repair experiment were averaged and used to calculate the NER half-life value for each individual. When the percentage of adducts remaining at each time point was plotted using a log scale, the individual plots were largely linear indicating that the data exhibit an exponential decay pattern. Therefore, we performed an exponential decay regression analysis of the averaged values for the percentage of adducts remaining that we obtained in each experiment using Sigma Plot and the equation: $f(x) = Ae^{-bx}$. In this equation, $f(x)$ is the percentage of 6-4 PPs remaining at time $x$. Setting $f(x)$ equal to 50% and solving for $x$, which is the time taken to remove 50% of the 6-4 PPs, produces the equation: $x$ = -ln(50/A)/$b$. The regression analysis provided the exponential decay variables ($A$ and $b$) necessary to calculate the half-life value.

The difference in average NER efficiency between age groups was evaluated by a 1-way ANOVA trend analysis and the difference in average NER efficiency between groups separated by smoking status or gender were evaluated by a 1-way student's T-test. When comparing NER efficiency and subject age as a continuous variable, a linear regression analysis was performed on the data and the coefficient of determination ($R^2$) obtained from that regression analysis was used to obtain the Pearson correlation coefficient (R). A multiple comparison analysis was performed between age, gender, and smoking and their combined effects on NER efficiency. Lastly, an interaction effect was examined evaluating

the impact of smoking and gender on the effect of subject age on NER efficiency. In all cases, the threshold for significance was $P < 0.05$. All statistical analyses were performed using the Graph Pad Prism 6 and SAS 9.3 software.

## 6.2.2 Study Population

Individuals were identified from Kentucky voter registration records of the 5th Congressional District of Kentucky located within the region defined as Appalachian Kentucky. All participants were required to be over the age of 17 years and have no previous diagnosis of cancer, with the exception of stage I or stage II non-melanoma skin cancer. Participants were required to have a working phone and be willing to consent to an in-person interview held in at their residence without the use of an interpreter. A detailed questionnaire which included questions about smoking history, work history, and medical history was administered by an employee of Kentucky Homeplace, who also drew blood samples. The blood samples were transferred into sodium heparin-coated tubes, placed on ice and shipped overnight by courier to the Biospecimen and Tissue Procurement Shared Research Facility (BSTP SRF) at the University of Kentucky and then distributed to Dr. Mellon's laboratory for the isolation of PBMCs.

## 6.2.3 Study population demographics

Subject recruitment numbers for this study as well as several population demographics are shown in Table 6.1. The average age of the study population, 61.7 years, is high due to the frequency matching of these individuals to a group of lung cancer patients for a larger case-control study. There are an equal number of males and females in the study population. Study participants were assigned

to one of three smoking groups based on a questionnaire that asked whether they were current smokers, former smokers (a history of smoking, but no smoking in the last 6 months), or never smokers (less than 100 cigarettes in their lifetime). Half of the participants reported being never smokers (n = 79, 50.6%), while the other half were either current (n= 26, 16.7%) or former (n= 51, 32.7 %) smokers. This distribution of cigarette usage in our study population is consistent with recently reported data for smoking status and smoking history among individuals residing in Appalachian Kentucky, the geographical area in which all of our study participants lived [12].

## 6.2.4 Measurement and distribution of NER efficiency in the study population

We measured the efficiency of NER in each person within our study population using PBMCs isolated from individual blood samples and an immuno-slot blot method to quantify the introduction and removal of 6-4 PPs produced by irradiation with UV light. We chose to utilize the removal of 6-4 PPs from live cells as a representation of NER efficiency in people because a number of biochemical, cellular and animal studies have shown that 6-4 PPs are model substrates for NER [292-296]. We chose PBMCs to study because acquiring blood from donors is relatively less invasive than acquiring most solid tissue samples. In addition, most assays that measure NER function/activity require viable cells but they have not been successful when applied to solid tissue samples. The time course of 0 – 4 h was established in preliminary studies using blood samples obtained from healthy volunteers at the University of Kentucky

who removed most 6-4 PPs during this time period after treatment of PBMCs with 20 J/m$^2$ UV-C. The dose of 20 J/m$^2$ was chosen to achieve a strong antibody-dependent chemiluminescence signal well above background levels using 100 ng of DNA per slot and to also achieve a strong signal since PBMCs are grown in suspension rather than as a monolayer which necessitated irradiating them in the presence of growth medium.

Representative immuno-slot blots illustrating the kinetics and extent of repair in three different subjects are shown (Fig. 6.1A). For each subject, three independent slot-blots were performed and the results were averaged to calculate the percentage of 6-4 PPs removed at each time point (Fig. 6.1B). Since our goal was to compare the efficiency of NER among all individuals in our study, we chose to use the results obtained from all 3 repair time points to calculate a single value for lesion half-life; the time it took for each individual to remove 50% of the 6-4 PPs introduced into the DNA of their PBMCs. To achieve this, the percentage of 6-4 PPs remaining at each time point was calculated and plotted using a logarithmic y-axis. This produced a linear response (Fig. 6.1C) and confirmed that the data fit an exponential decay model.  An exponential decay regression analysis was performed on the data as described in the methods section.  The half-lives for the individuals shown ranged from 1.57 h to 3.45 h (Fig. 6.1C). This analysis was performed on 156 subjects and a histogram of these results illustrates the range and the distribution of NER efficiency across the population (Fig. 6.2A).  The population had an average half-life of 4.24 h, with a standard deviation of 1.70 h. Small half-live values indicate faster NER while

larger half-life values indicate slower NER. The half-life values ranged from the fastest, 1.52 h, to the slowest, 10.71 h. These results clearly show that the efficiency of NER can significantly differ among different individuals. It is unlikely that differences in NER efficiency can be attributed to differences in the introduction of UV damage. Identical conditions were followed for the irradiation of all samples and the intensity of the slot-blot band for the 0 h time point provided an indication of the uniformity of the introduction of damage across samples.

The distribution of half-lives across this population shows a largely Gaussian distribution with some skewing towards the slower values, as expected given the nature of the repair assay. Proficient repair values would be clustered around a peak value, and deficiencies in repair would produce larger half-life vaues and generate a right-skewed tail. According to the Kolmogorov-Smirnov (KS) test for a normal (Gaussian) distribution, the half-life values are not completely normally distributed ($p < 0.01$). The null hypothesis of the KS test states that the data are normally distributed, and thus a "significant" p-value means the data are not normally distributed. Certain statistical analyses such as ANOVA and T-tests require (assume) a Gaussian distribution so we performed a natural-log transformation of the data and tested for normality. This is a common approach to normalize a slightly skewed distribution. The resulting distribution is normal ($p > 0.15$) as evaluated by the KS test (Fig. 6.2B). We performed several comparisons of NER efficiency between subpopulations using either the untransformed or transformed data sets and the results were similar regarding

trends and significance.  Hence, we chose to present the analysis of the non-transformed data set in subsequent figures and all tables since a half-life value represented in hours rather than the natural-log of the hours is biologically more relevant.

## 6.2.5 The efficiency of NER is reduced with increasing subject age

We investigated the relationship between NER efficiency and the age of each individual.  Initially, the study population was separated into three different age groups and the average half-life obtained from each age group was compared (Fig 6.3a).  NER efficiency slowed with increasing subject age when subjects were stratified into the three age groups shown, as indicated by an increase in half-life across these age groups.  A one-way ANOVA analysis examining the trend of increasing half-life with age showed that this trend was significant ($p = 0.0023$). Additionally, we examined the relationship between subject age and NER efficiency when age was evaluated continuously instead of in discrete groups.  A graph of the half-life value plotted against age for each individual in the study is shown (Fig 6.3b).  A linear regression analysis of this data produced a coefficient of determination ($R^2$) value of 0.655.  The Pearson score ($R = .2559$) indicates a significant relationship between subject age and half-life ($p = 0.0013$) using the continuous measurement of subject age.

Factors such as lymphocyte and DNA recovery were also evaluated for their potential to modulate the observed effect of subject age on NER efficiency and were determined to have minimal impact (data not shown, but similar work was done in chapter 5 and a similar conclusion was drawn for that population).

## 6.2.6 The relationship between NER efficiency and other population demographics

We investigated whether the smoking status of each individual could influence their efficiency of NER. There is some limited evidence that current smokers may have faster repair than former or never smokers [289, 297, 298], although these experiments utilized the HCR assay and therefore measured TC-NER. A graph of the average half-lives and their standard deviations found for current, former and never smokers is shown (Fig 6.4a). The current smoking population has a faster half-life (3.83 h) than the former (4.37 h) and never smoking (4.28 h) populations, although the difference is not statistically significant (p = 0.0914). However, the group of current smokers also has a lower average age than the other two groups, and since we have found that subject age influences that efficiency of NER, the faster NER found in current smokers is less pronounced after accounting for age (See below and Table 6.1). As the difference in smoking groups was most pronounced between current smokers and "non-current" that was the comparison that was presented. Comparisons between individual groups, including current vs former and current vs never have even higher p-values than those between current and non-current smokers as a result of diminished sample sizes.

The impact of gender on NER efficiency was also investigated. Previous studies have suggested slight increases in DNA repair capacity in males compared with females [273, 299]. In our study, however, females had a somewhat faster NER efficiency than males, as indicated by an average half-life

133

of 4.12 ± 1.66 h for females compared to 4.35 ± 1.74 h for males (Fig 6.4b). This difference was not statistically significant (p = 0.2029) using a 1-way student's T-test. However, this trend was present in every subpopulation examined; females had a faster NER efficiency than males in all three age groups and all three smoking groups, although these differences were not statistically significant, perhaps due to small sample sizes (Table 2).

## 6.2.7 The reduction in NER efficiency with increasing subject age is not seen in all subpopulations

As with the total population, the male and female populations both show a trend of decreasing NER efficiency (increasing half-life) with increasing subject age (Fig 6.5a and 6.5b), and the trend is significant in both males and females. The strength of the relationship was stronger in the male population than the female population, as evident in the p-values for age associated decline in NER efficiency in males (p = 0.013) and females (p = 0.038). However, separating the study participants by smoking status reveals that only current and former smokers show an age-dependent decline in NER efficiency (Fig 6.6a and 6.6b respectively) while non-smokers show no such trend (Fig 6.6c). Supplemental table 6.1 contains the $R^2$ values and p-values derived from linear regression analysis of graphs of the relationship between age and NER efficiency in each population. Since non-smokers do not show an age-dependent decline in NER efficiency, the smoking status of individuals (in addition to their age) must be accounted for when drawing meaningful comparisons of NER efficiencies between groups of individuals, such as in case-control studies.

**6.2.8 Multivariable analysis confirms efficiency of NER is effected by subject age**

We performed a multivariable regression analysis to determine which factors were significant modulators of NER efficiency when examined together. As indicated in Table 1 in the "adjusted p-value column", subject age remains the only significant factor effecting NER efficiency when age, gender, and smoking status are all considered ($p = 0.0138$). The analysis indicates that neither gender ($p = 0.2005$) nor smoking ($p = 0.3987$) are significant contributors to variation in NER efficiency. The multivariable regression analysis did not alter the significance of age or gender on NER efficiency, but the p-value for smoking (which compared current smokers to former and nonsmokers) changed from 0.0914 to 0.3987, likely due to the large difference in ages between the smoking groups (current smokers were considerably younger than former/never smokers). Since subject age had a large effect on NER efficiency, the difference in ages between the smoking groups was ultimately responsible for most of the difference in NER efficiency between smoking groups. As the male and female populations had very similar ages, the multivariable regression analysis did not change the p-value for the effect of gender on NER efficiency. Finally, an interaction effect was incorporated into the multivariable regression analysis to evaluate the impact of smoking status and gender on the observed age effect on NER efficiency. Smoking status ($p = 0.0226$) had a significant impact on the observed age-dependent effect on NER efficiency, while gender ($p > 0.80$) did

not. This finding is consistent with our observation that both the male and female population showed a strong reduction in NER efficiency with increasing age.

We also examined the relationship between lymphocyte and DNA recovery and the effect of subject age on NER efficiency. Similar to the results presented in chapter 5, these two collection parameters did not have any significant impact on the subject age-dependent modulation of NER efficiency.

## 6.3 Discussion

Nucleotide Excision Repair is a critical pathway involved in removal of DNA damage that would otherwise promote mutations and increase cancer risk. We have established that this assay can reproducibly measure individual NER efficiency, as well as measure across a large range of values in a population. NER efficiency was measured in a rural population in Appalachian Kentucky, an area that suffers high cancer rates, poor overall health, and limited access to healthcare. The objectives of the study were to determine how NER efficiency was distributed across a population when measured using this assay and establish whether this repair assay could measure differences in DNA repair between individuals or groups of individuals in a population. The study population exhibited a wide range of NER efficiencies, with almost an order of magnitude separating the half-lives of the fastest and slowest repairing individuals. The range of repair values is consistent with HCR ranges, which also can exhibit an order of magnitude difference between the smallest and largest values [300]. We observed a significant decline in DNA repair specific to the NER pathway with increasing age of study subjects, both male and female. We also found that

smoking status impacts the age-dependent decline in NER efficiency. This information will guide further implementation of the repair assay for use in measuring NER efficiency as a potential biomarker of cancer susceptibility.

In addition to contributing to cancer risk, defective DNA repair may have a role in the process of aging. Aging can be viewed as a reduction in biological functions that occurs with time as a result of an accumulation of somatic DNA damage [301, 302]. This is due in part to an increase in genomic instability (reviewed in [303]). There are two primary factors that regulate genomic stability; the rate of DNA damage induction and the efficiency of DNA repair. A reduction in DNA repair would promote an increase in genomic instability, promoting cellular and potentially organismal aging.

Genetic deficiencies in NER lead to increases in an aging phenotype, including premature cell death and reduced lifespans, in both humans [42, 213, 304] and mice [305, 306]. These studies use repair deficient systems and thus show the consequences of impaired NER, but they do not provide information on the relationship between normal DNA repair and aging. Several studies have investigated the relationship between proficient DNA repair and aging (reviewed in [307]). Early studies using the unscheduled DNA synthesis (UDS) assay found that UV-induced DNA damage was repaired more slowly with increasing age in rats [308] and more slowly in later passages of WI-38 human lung cells compared to earlier passages [309]. A decline in NER efficiency, as measured by reduced removal of UV-induced photolesions, with increasing donor age was observed in cultured human dermal fibroblasts [310].

A number of studies have explored the relationship between DNA repair and aging in isolated lymphocytes in a human population using the host-cell reactivation assay (HCR). However, as stated above, this assay measures TC-NER, and the results of these studies have been quite varied. DNA repair capacity (DRC) was reduced with increasing subject age in a study of lung cancer patients [260]. In contrast, other studies have shown that DRC decreased with age in healthy individuals but increased with age in cancer patients [311, 312]. DRC has also decreased with increasing subject age regardless of cancer status in both groups significantly [259], or only significantly in the control population [313-315]. Some studies have reported that age did not contribute significantly to DRC in either the case or control populations [273, 316], and DRC has even remained unchanged in controls while increasing with age in cancer patients [299].

The previous findings from human population studies suggest that the current understanding of the relationship between DNA repair capacity and subject age from isolated lymphocytes in a human population is not complete. We saw a very strong influence of age on NER efficiency, and we also observed that this effect was modulated by smoking status (as indicated by the significance of the interaction variable of smoking and age on NER). It is possible that this interaction may explain some of the disparities in the age effect on DNA repair seen in these studies – if smoking status is not uniform across age groups it may obscure the effect of subject age on DNA repair.

Our findings indicate that subject age must be properly matched when comparing NER efficiency between populations, as when exploring the effect of DNA repair on cancer susceptibility.  Additionally, while smoking status did not significantly alter NER efficiency, it had a significant impact on the age-dependent decline in NER efficiency. As a result, we propose that smoking status should also be consistent both between study populations and across age groups within the same population in order to most accurately measure differences in NER efficiency between populations.

## Table 6.1 Factors affecting NER efficiency

| | Subject Demographics | | | | P-value for Half-life | Adjusted P-value | Interaction effect P-value (with age) |
|---|---|---|---|---|---|---|---|
| Variable | No. | % | Age ± SD | Half-life (hr) ± SD | | | |
| All subjects | 156 | 100 | 61.7 ± 8.8 | 4.24 ± 1.70 | | | |
| Age, years | | | | | | | |
| Under 50 | 19 | 12.2 | 45.4 ± 5.6 | 3.22 ± 1.25 | 0.0023 | 0.0138 | |
| 50-65 | 76 | 48.7 | 58.3 ± 4.0 | 4.23 ± 1.63 | | | |
| Over 65 | 61 | 39.1 | 71.1 ± 5.0 | 4.57 ± 1.80 | | | |
| Smoking status | | | | | | | |
| Current | 26 | 16.7 | 57.7 ± 8.8 | 3.83 ± 1.75 | 0.0914 | 0.3987 | 0.0226 |
| Former | 51 | 32.7 | 63.4 ± 9.5 | 4.37 ± 1.77 | | | |
| Never | 79 | 50.6 | 62.0 ± 10.0 | 4.28 ± 1.64 | | | |
| Sex | | | | | | | |
| Male | 78 | 50 | 61.3 ± 9.3 | 4.35 ± 1.74 | 0.2029 | 0.2005 | > 0.80 |
| Female | 78 | 50 | 62.2 ± 9.3 | 4.12 ± 1.66 | | | |

**Table 6.1** Demographic information and NER efficiency are provided for the indicated groups in the study population.  P-values from comparisons of half-lives between different subpopulations are shown. For comparing half-life values across the three age groups, the p-value was obtained from a 1-way ANOVA trend analysis.  For comparing half-life values in groups based on smoking status or gender, the p-value was obtained from a 1-way Student's T-test. The adjusted p-value is the result of a multivariable regression analysis that factored age, smoking status, and gender.  The interaction effect p-value was derived from the multivariable regression analysis and indicates whether the indicated variable had a significant influence on the observed effect of subject age on NER efficiency.

**Table 6.2 Influence of gender on NER efficiency in various study populations.**

| Group | NER efficiency in Males versus Females | | | | p-value |
|---|---|---|---|---|---|
| | **Male** | | **Female** | | |
| | N | T ½ Mean ± SD | N | T ½ Mean ± SD | |
| All | 78 | 4.35 ± 1.74 | 78 | 4.12 ± 1.66 | 0.2029 |
| **Age Group** | **Male** | | **Female** | | |
| Under 50 | 13 | 3.31 ± 1.46 | 6 | 3.02 ± 0.69 | 0.3261 |
| 50-65 | 34 | 4.36 ± 1.48 | 42 | 4.11 ± 1.75 | 0.2560 |
| Over 65 | 31 | 4.77 ± 1.97 | 30 | 4.36 ± 1.60 | 0.1848 |
| **Smoking status** | **Male** | | **Female** | | |
| Current | 17 | 3.76 ± 1.62 | 9 | 3.69 ± 1.98 | 0.4522 |
| Former | 32 | 4.61 ± 1.77 | 19 | 3.97 ± 1.74 | 0.1087 |
| Never | 29 | 4.38 ± 1.76 | 50 | 4.26 ± 1.58 | 0.3715 |

**Table 6.2** A comparison of NER efficiency between males and females in different groups of the study population. Females tend to repair faster than males across all populations measured, but the difference is not significant.

# Figure 6.1 DNA repair assay performed on lymphocytes isolated from three different subjects
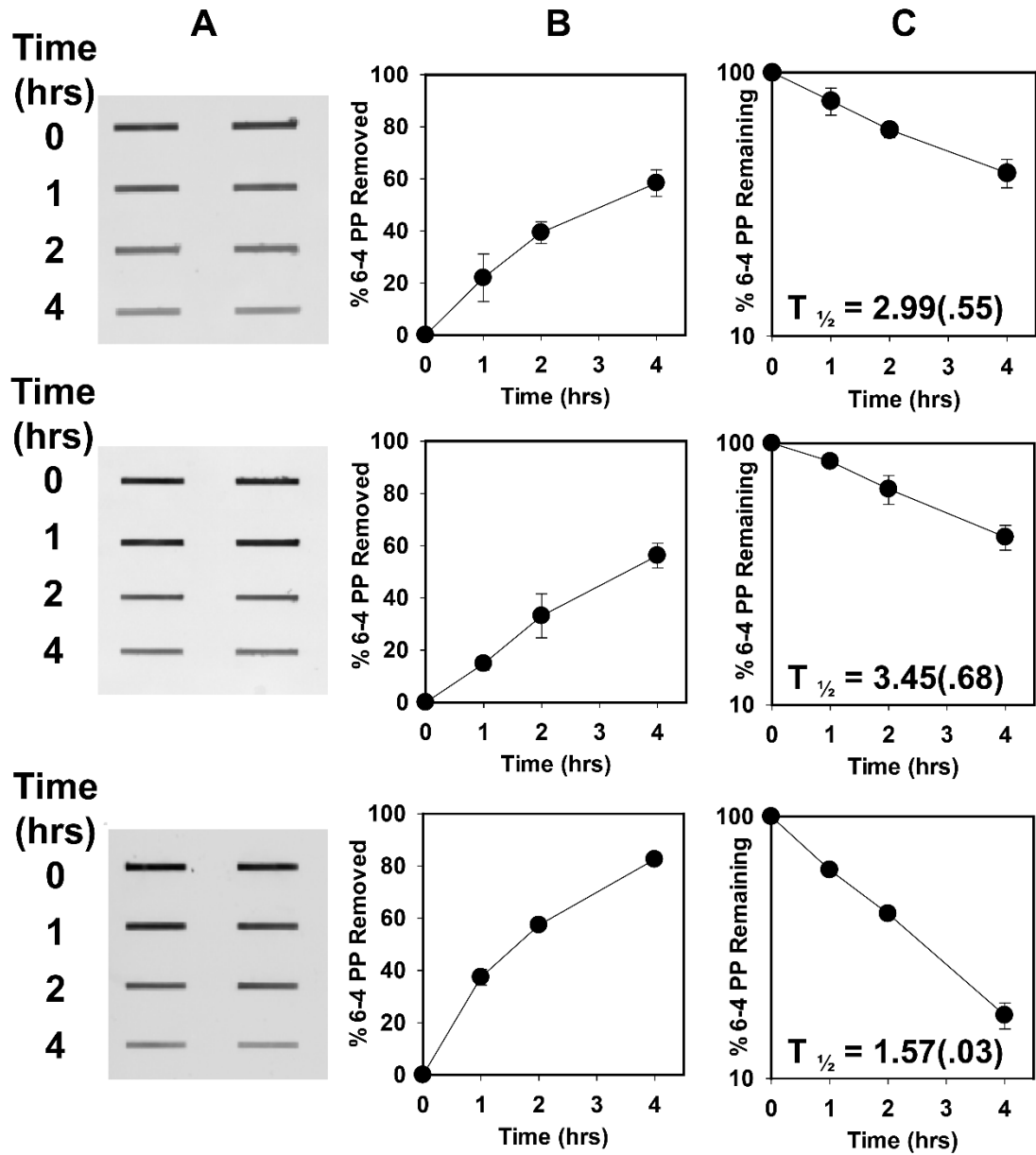
**Figure 6.1** (A) A slot-blot from each of three individuals in the study is shown.

(B) The percent 6-4 photoproducts removed are calculated from three repeats of

the slot-blot per individual, with the average repair and standard error at each

timepoint plotted on the graph. (C) The percent repair values are converted into

% 6-4 photoproducts remaining at each timepoint, and the average and standard

error values are plotted as shown. From this data, an exponential decay

regression analysis was performed to generate the half-life (and standard error)

values shown for each individual.

**Figure 6.2 The distribution of NER efficiency across the study population**



A

% of subjects in each group

T ½ = 4.24 (1.70)
N = 156

Half-life (hours)

B

% of subjects in each group

ln T ½ = 1.37 (0.39)
N = 156

ln Half-life (hours)
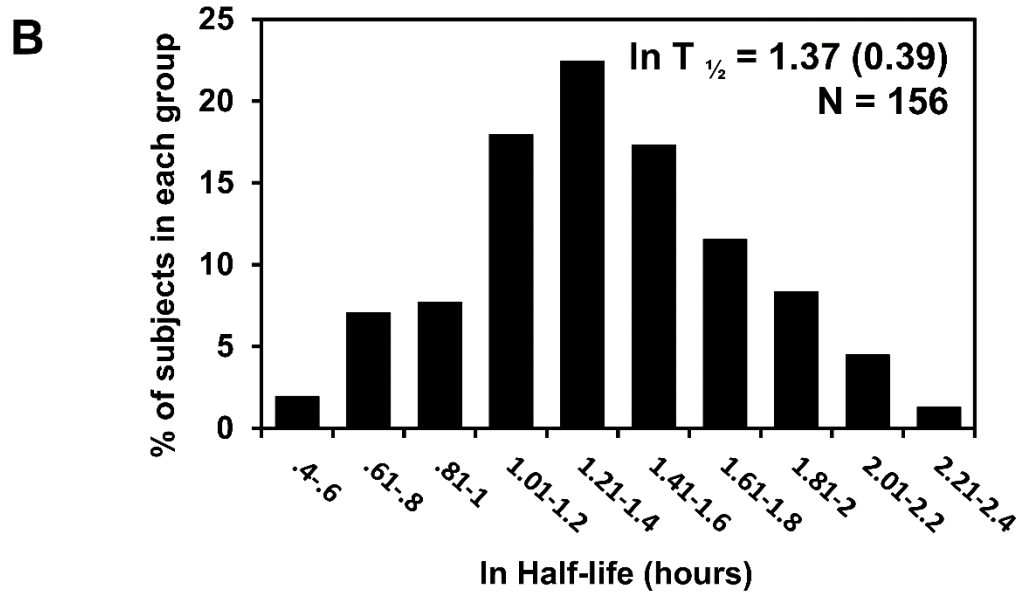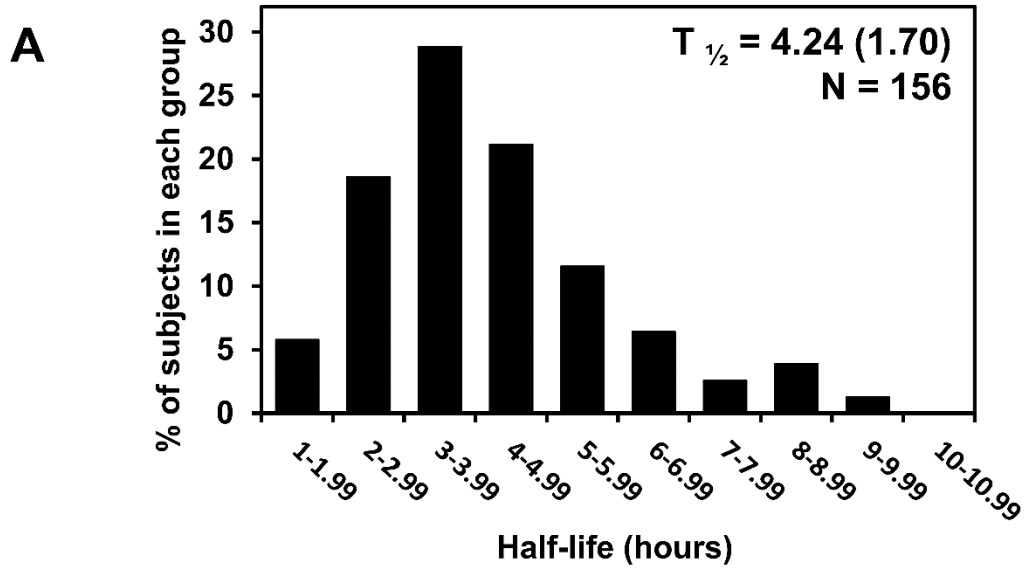
**Figure 6.2** (A) A Histogram of half-life distribution across the entire population is shown, as well as the average and standard deviation of the population half-life. The study population's NER efficiency has a mostly Gaussian distribution but is slightly skewed towards slower repair values. The Kolmogorov-Smirnov (KS) normality test indicates that the half-life values are not normally distributed (p < 0.01). (B) In order to produce a normally distributed data set, a natural log transformation of the data was performed, and a histogram of the natural log transformed data is shown, along with the average and standard deviation of the transformed half-life of the study population. The KS normality test confirms that the natural log-transformed half-life values are normally distributed (p > 0.15)

**Figure 6.3 NER efficiency is reduced with increasing subject age**
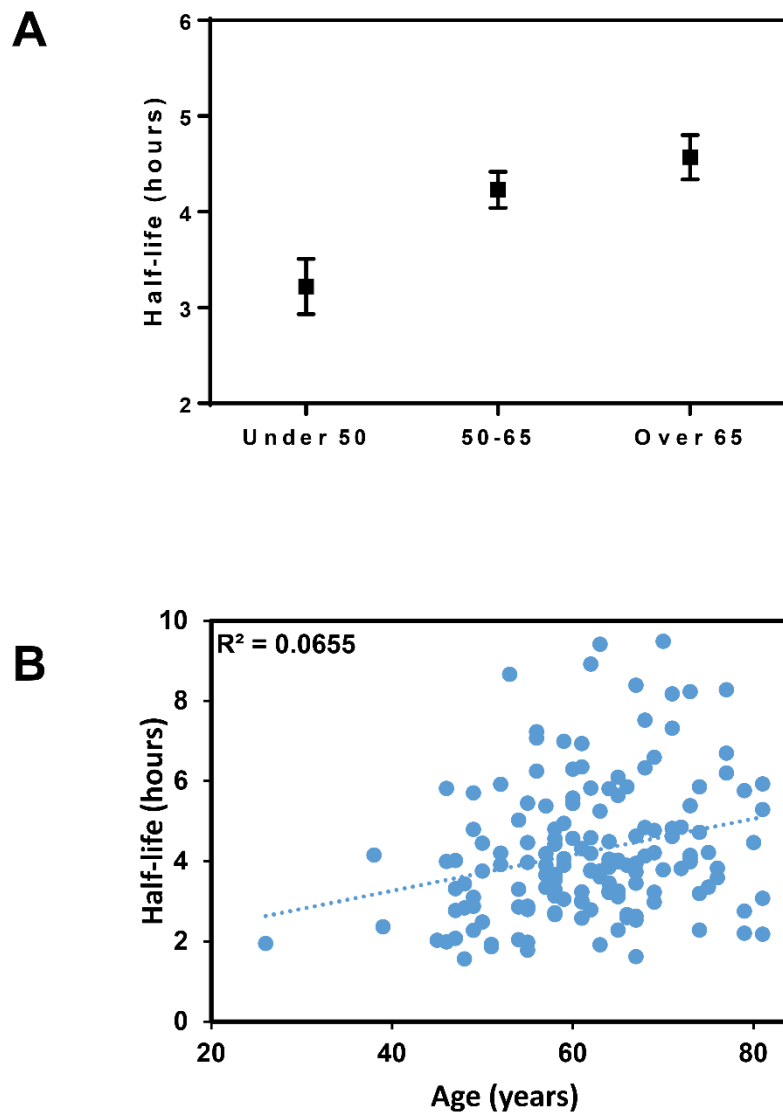
**Figure 6.3** (A) A graph of the average half-life and standard error for the population separated into three age groups shows an increase in half-life with increasing subject age. A 1-way ANOVA trend analysis for increasing half-life produced a p-value of 0.0023. (B) A scatter plot of the half-life vs subject age of each individual in the study also shows an increase in half-life with increasing subject age. A linear regression analysis of the data produced an $R^2$ value of 0.0655 (p = 0.0013)

**Figure 6.4 The effect of smoking status and gender on NER efficiency**
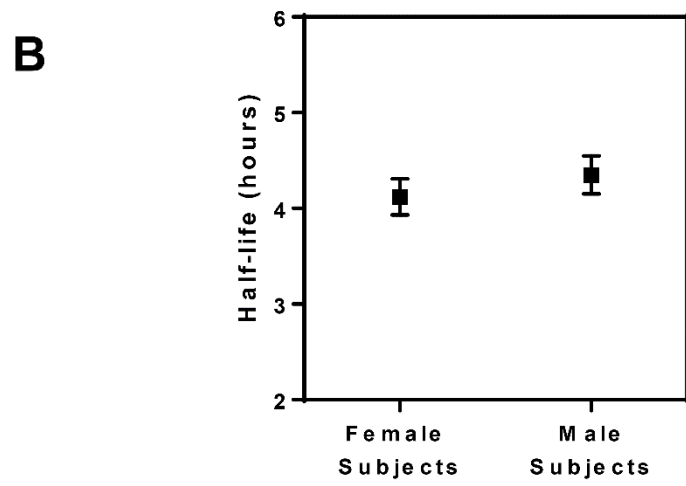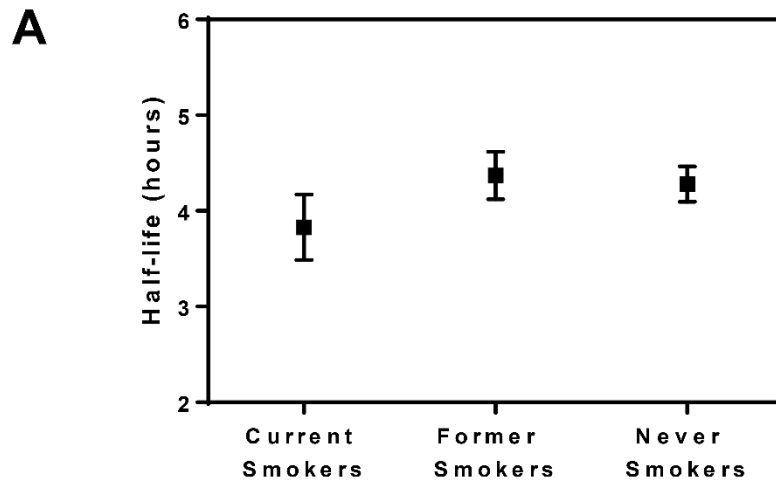
**Figure. 6.4** (A) The effect of smoking status on NER efficiency.  The average half-life and standard error from each of the three smoking groups in the study are plotted on the graph.  The current smokers have a faster NER efficiency compared to the other smoking groups, but the difference is not significant ($p = 0.0914$). (B) The effect of gender on NER efficiency. The average half-life and standard error of females and males in the study are plotted on the graph. Female subjects have a moderately faster NER efficiency than males, but this difference is not statistically significant ($p = 0.2029$).

**Figure 6.5 NER efficiency decreases with subject age in both genders in the study population**
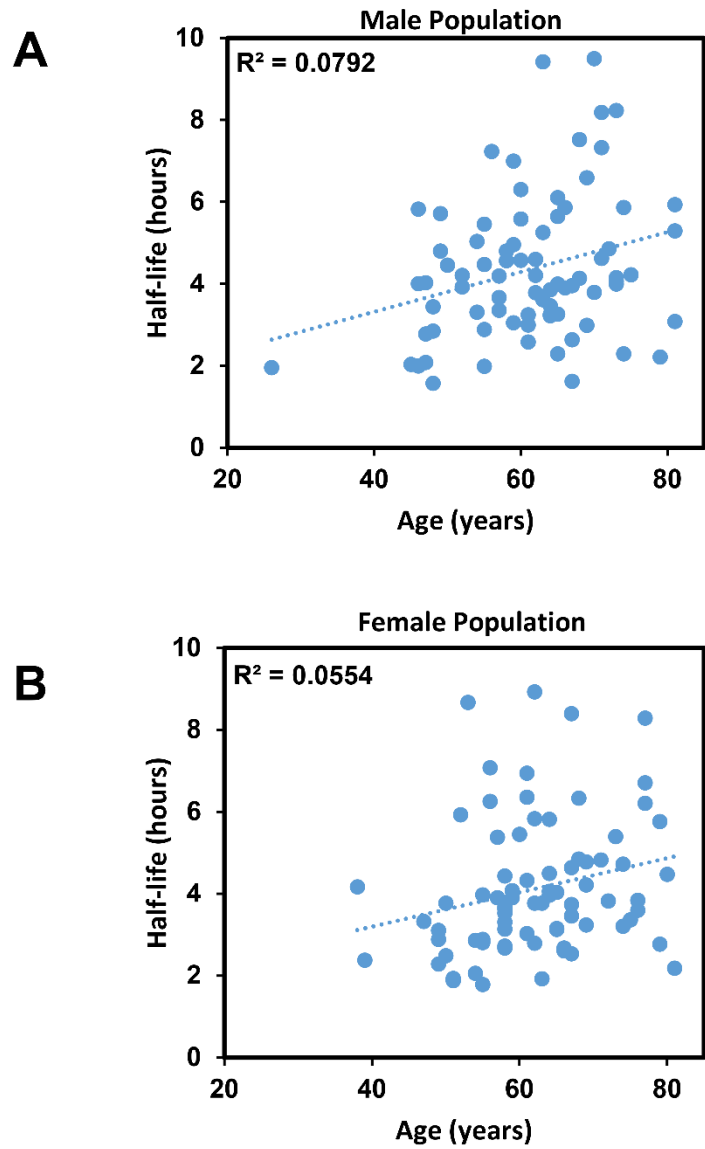
**Figure 6.5** Scatter plots of the distribution of half-life compared to subject age for

male (A) and female (B) subjects in the study population are shown.  Regression

analysis of the scatter plots of subject age verses half-life provided $R^2$ values of

0.0792 (p = 0.013) and 0.0554 (p = 0.038) for males and females respectively,

indicating a significant relationship between age and NER efficiency in both

populations.

**Figure 6.6 The age-dependent decline in NER efficiency depends on**

**smoking status**
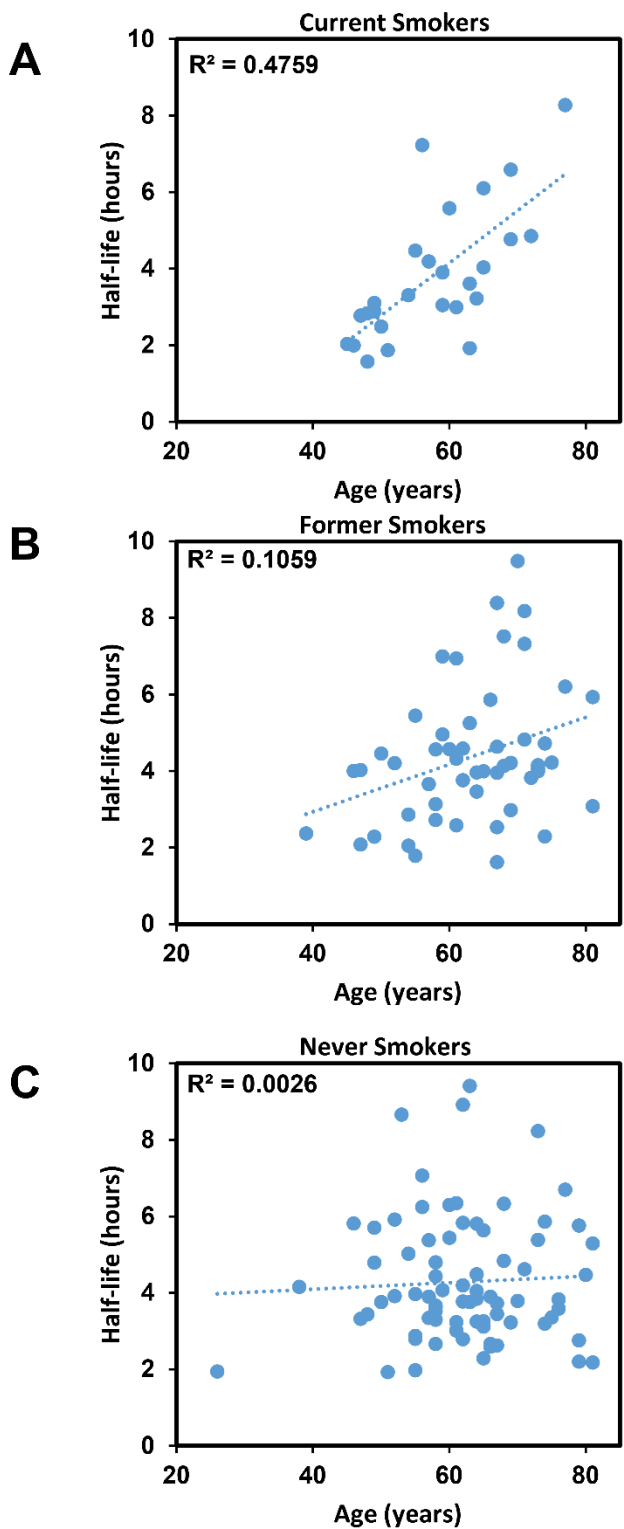
**Figure 6.6** Scatter plots of the distribution of half-life compared to subject age for current (A), former (B), and never smoker (C) subjects in the study population are shown. Regression analysis of the scatter plots of subject age verses half-life provided $R^2$ values for the current and former smoking populations of 0.4759 (p = 0.0001) and 0.1059 (p = 0.02) respectively, indicating a significant relationship between age and NER efficiency in both populations. The regression analysis performed on the non-smoking population, however, produced an $R^2$ value of 0.0026 (p = 0.655), indicating there is no decline in NER efficiency with increasing subject age in the non-smoking population.

**Supplemental Table 6.1 Strength of association between subject age and**

**NER efficiency in different study populations**

| Group | N | R² | P-value |
|---|---|---|---|
| All | 156 | 0.0655 | **0.0013** |
| Male | 78 | 0.0792 | **0.013** |
| Female | 78 | 0.0554 | **0.038** |
| Current smokers | 26 | 0.4759 | **0.0001** |
| Former smokers | 51 | 0.1059 | **0.020** |
| Never smokers | 79 | 0.0026 | 0.655 |
| Male Never smokers | 29 | 0.0356 | 0.327 |
| Female Never smokers | 50 | 0.0025 | 0.730 |

**S. Table 6.1** A linear regression analysis of subject age and NER efficiency was performed for the study population as well as several subpopulations in the study. Both genders overall show a strong age effect on NER efficiency, as do the current and former smoking populations, but the males and females in the never-smoking population do not.

# Chapter 7

# Unanswered questions, a case-control study, and a future

# direction for the NER functional assay

## 7.1 Did reduced XPC expression impair NER?

I have presented my findings indicating that the efficiency of the

Nucleotide Excision Repair pathway can be impacted by environmental factors

such as tobacco smoking and arsenic and by epidemiological factors like age

and possibly even gender.  The variations in NER efficiency observed in the

previous chapters may increase the risk of cancer development. Numerous

reports indicate that defective NER [40, 211, 278], reduction in expression of

NER factors [249, 287], polymorphisms in NER genes [279, 281, 283-286], and

variations in (TC) NER efficiency within a population [259, 312, 317] all contribute

to cancer risk.

We observed a reduction of XPC protein and RNA levels that was

concurrent with a reduction in NER efficiency in cells treated with either CSC or

arsenic.  One important question remains unanswered: Is the inhibition of NER

produced by CSC or arsenic a consequence of the reduction of XPC?  Loss of

XPC eliminates functional GG-NER [245, 318], and evidence suggests that

reduction in expression may also inhibit repair of UV-induced 6-4 PPs [245] and

ultimately increase cancer risk [47, 152], but we need a clear demonstration that

this occurred in the cell lines that were tested.

There are at least two experimental systems that could be implemented to

test the hypothesis that NER function was inhibited in the cell lines examined as

a consequence of XPC inhibition.  NER function could be measured in cells where XPC is artificially reduced in the absence of CSC or arsenic.  Small interfering RNA (siRNA) specific to XPC can knock down XPC expression [151]. Stable transfection of the siRNA with an inducible promoter could be used to reduce the level of XPC to a level that is similar to what was observed in treatments of CSC or arsenic.  If reducing XPC in this manner produces a similar reduction in the removal of the photolesions we studied in the presence of CSC or arsenic, then the hypothesis would be supported.

The shortfall of this experiment, modulating XPC levels in the absence of the chemicals responsible for the observed reductions of XPC, is that any additional mechanisms of NER inhibition that CSC/arsenic may have employed that we have not yet quantified would be lost.  Therefore, if reducing XPC levels in the cell lines examined does not inhibit NER efficiency in the same manner that CSC/arsenic did, we are left not knowing whether this is because the XPC reduction was unrelated to repair inhibition or simply one of multiple changes the chemicals produced that combined to inhibit NER.

Alternatively, overexpressing XPC as a means to protect the cell from the inhibitory effects of CSC or arsenic on NER would also provide evidence that XPC levels are responsible for the inhibition of NER seen in the treated cells. This experiment runs up against the issue of XPC toxicity, however.  As indicated in chapter 3, XPC is a highly regulated protein that has considerable post-transcriptional modifications, including ubiquitylation.  Maintaining low levels of XPC is important, as XPC actively scans chromosomes looking for DNA damage.

Overexpression of XPC and its binding partner HR23B produced significant toxicity, likely by interfering with DNA metabolism by binding to a wide range of DNA structures [162]. Overexpressing XPC sensitizes p53-deficient cells to DNA damage-induced apoptosis [319], which may be a useful tool in combatting therapy resistant tumors, but it complicates the repair assay by potentially exaggerating the apoptotic effect of the DNA damage-inducing UV. Additionally, a conditional, ubiquitous oxerpression of mus210, an XPC homolog, reduced average lifespan in fruit flies by almost 50% [320]. Therefore, overexpressing XPC in a cell containing endogenous XPC poses some challenges.

## 7.2 XPC expression as a biomarker for cancer risk

After determining that NER function and XPC expression could be impacted by CSC and arsenic, both known human lung carcinogens, we explored the hypothesis that XPC expression in human lung tissue may be a biomarker for cancer risk. We proposed that XPC expression may be reduced in lung cancer tissue, relative to non-neoplastic adjacent tissue, and this reduction could be a driving force behind the lung carcinogenesis. XPC expression level reduction could be a result of exposure to environmental agents like tobacco smoke or arsenic, similar to what was observed in our cell culture studies. Alternatively, XPC levels may be impacted by a mutation or an epigenetic change that results in reduced expression of XPC, reducing the ability of the NER pathway to function and perhaps ultimately driving the formation of the cancer. Using immunohistochemistry, lung tissue samples acquired from biopsies of lung cancer patients were stained for XPC protein. In a small pilot

study, nuclear XPC content was reduced in cancer tissue relative to non-neoplastic adjacent tissue in the same slide. However, we were unable to replicate our initial findings, and as such we could not draw any conclusions regarding the relationship between XPC expression and lung cancer.

## 7.3 A more complete biomarker of cancer susceptibility

Using the protein level of a repair gene such as *XPC* as a biomarker for cancer risk would be more easily accomplished from a bio-specimen that is more easily obtained, such as blood. However, it would be prudent to establish a few other things first. As the cancer risk associated with variations in protein expression of certain NER factors has produced inconclusive results thus far (see chapter 5), it may be more productive to evaluate the relationship between the efficiency of the NER pathway in cells isolated from blood and cancer risk first, and then look for a relationship between this efficiency and expression levels of certain proteins like XPC. As an assay that measures NER efficiency is more involved than a simple expression study, it is less feasible to use the efficiency of NER as a predictor of cancer risk in a high-throughput fashion. However, it is critical that the correlation between NER efficiency and cancer risk be established as well as a strong relationship between, for example, XPC expression and NER efficiency, as this increases the strength of association between XPC expression (a more easily measured value than NER efficiency) and cancer risk.

With that in mind, we participated in a large scale case-control study designed to investigate the causes of the large incidence of lung cancer in

Appalachian Kentucky. Several investigators were involved in the study, and several hypotheses were tested to determine what was responsible for this cancer incidence. Previous work done by members of the group had determined that, even though smoking rates are higher in Appalachian Kentucky than the rest of the state, several geographical areas had lung cancer incidences higher than the state average after adjusting for smoking [14]. At the time of this writing, the analyses performed by the other investigators are not complete, and as such I will not include their findings.

Based on our findings from cell culture experiments that the carcinogenic chemicals CSC and arsenic inhibited NER efficiency as indicated in chapters 3 and 4, we hypothesized that variations in NER efficiency between individuals could be a risk factor for cancer development. The study was designed to provide blood samples from lung cancer subjects and age/gender matched control subjects. Lymphocytes would be isolated from the blood samples and a repair assay would be performed. Unfortunately, the number of lung cancer cases that were evaluated (42) was insufficient to produce the power necessary to draw any conclusions about potential differences in repair between the cancer and control groups. Additionally, as we explored the variables that could potentially impact NER efficiency, we observed that age, smoking, and gender all had a role to play in determining individual NER efficiency (as discussed in chapter 6). Since the study was not designed to have equal representation of smokers between cases and controls, the power of the study was even further

reduced, as comparing cases and controls after adjusting for smoking status reduced the sample sizes even further.

Instead, we chose to report the findings of the population distribution of NER efficiency from the control group only, drawing inferences as to how a case-control study should be designed in the future. The study was designed to match lung cancer case and cancer-free control subjects on the basis of age and gender. It did not require that they be smoking matched. The results of our study indicate that the different smoking groups did not have statistically different average half-lives. However, the observed effect of age and NER efficiency was modulated by smoking. This means that an appropriate study design should include an even distributions of smokers between the cases and controls. It also means that smoking incidence should also be evenly distributed across the range of subject ages for both cases and controls, as doing otherwise would produce an uneven effect of subject age on NER efficiency between the populations. To our knowledge, no such adjustment has been made in any population studies that evaluate DNA repair from human lymphocytes.

Even though we did not have a sufficient lung cancer subject sample size to properly compare cases and controls in our Appalachian Kentucky study, we did analyze NER efficiency in 42 lung cancer subjects. One of our most significant findings in the control population, that NER efficiency was significantly reduced with increasing subject age (Figure 6.3), was not significant in the case population (Figure 7.1) or any of the subpopulations (males, females, current or former smokers). This could be a consequence of selection bias – if reduced

NER efficiency is indeed a risk factor for the development of lung cancer, then

lung cancer patients may not show a reduction in NER efficiency with increasing

subject age, as the younger population is already presenting a reduction in NER

efficiency.  If that is the case, then the age-dependent reduction in NER

efficiency seen in the control population could be absent in the case population.

The observation that the lung cancer population did not show a reduction

with NER efficiency with increasing age suggests that comparisons between

cases and controls may be more meaningful at lower age groups, before the

control population experiences an age dependent reduction in NER efficiency.

Indeed, the younger case population (individuals under 50 years old) had a

slower NER efficiency than the younger control population, but the difference

was not significant (Figure 7.2).  As expected, since the case population does not

show a reduction in NER efficiency with age, the NER efficiency in the case

population is actually greater than the control population in the second and third

age groups (although it is only significant in the middle age group).  An important

observation was that all of the lung cancer cases under 50 years old were current

smokers. Since we observed that subject smoking status impacted NER

efficiency via an interaction effect with subject age, we concluded that the

smoking distribution needs to be comparable between groups.  As such, we

compared the NER efficiency of the current smoking case subjects to the current

smokers in the control group (Figure 7.3).  In the youngest age group, the case

subjects had a significantly slower NER efficiency than the youngest current

smoking control subjects.  However, the sample sizes in this limited analysis are

quite small, only 7 cases and 7 controls. A larger study that includes greater

representation of individuals under the age of 50 would be desired to obtain a

more meaningful comparison between cases and controls using this repair

assay. Separating the study participants by age group and smoking status is not

typically done in population studies evaluating DNA repair as a biomarker of

cancer risk, and although these results are from smaller than desired sample

sizes, they suggest that both smoking status and age play a role in individual

NER efficiency. Since we have observed a reduction in NER efficiency with

increasing subject age among the control population, but not the case population,

this assay may most informative in case-control studies on younger individuals.

As cancer is a disease of aging, detecting cancer risk early is a very useful

preventative tool.

Ultimately, the goal of measuring DNA repair, be it by measuring

polymorphisms in DNA repair genes, by measuring levels of proteins involved in

DNA repair, or with an assay to determine individual DNA repair efficiency, is to

use this measurement as a biomarker of cancer susceptibility in a healthy

population. As stated above, there have been considerable efforts to use the

association of repair gene polymorphisms and cancer incidence as a predictor for

cancer risk. The drawback to these sorts of studies is that genetic

polymorphisms do not tell the whole story – they cannot provide a measure of

how efficiently the expressed proteins would interact in a repair process that

requires the concerted efforts of dozens of individual factors. Measuring protein

levels also has this limitation. While there may be statistically significant

associations between the levels of various proteins involved in DNA repair and cancer risk, it is also possible that an individual who has normal levels of the measured proteins has a deficiency elsewhere that increases cancer risk. The best way to capture the whole picture of an individual's cancer risk is to perform a functional assay to measure the ability of their DNA repair machinery to remove DNA damage. We have successfully implemented a functional assay to measure NER efficiency in a human population, and from that study we have discovered insights into the factors that modulate NER efficiency. Primary among these factors is subject age, and related to that is that observation that smoking interacts with subject age to regulate NER efficiency.

Previous studies have observed an age effect on DNA repair (albeit TC-NER via the host-cell reactivation assay) in isolated lymphocytes, and there has been some evidence that smoking status may affect repair as well, but the interplay between these two factors has not, to our knowledge, been observed before. We found that non-smokers, who were (mostly) cancer free in our study, did not show an age dependent reduction in NER efficiency, while current and former smokers did (Table 7.1). This was a rather unexpected finding, since the total study population (n = 198) and the control only population (n = 156) each show a significant reduction in NER efficiency with increasing subject age. In other words, the smoking population, which represented about half of the control subjects and over half of the total subject population, had such a strong age-dependent reduction in NER efficiency that it effectively masked the lack of an age-dependent response in the non-smokers. Additionally, the total female

population did not have a significant age-dependent decline in NER efficiency, but the control female population did.  This was a result of the female case subjects who did not have an age-dependent reduction in NER efficiency and therefore masked the age-dependent reduction in NER efficiency in the control female population when they were grouped together in the "total female" population.  These findings highlight the importance of determining not only what factors effect NER efficiency in a population before a comparison between cases and controls can be done, but additionally the findings suggest that variables can interact with each other to contribute to the distribution of NER efficiency within a population and that these interactions must also be properly understood and corrected for when measuring NER efficiency in a case/control study as a marker of cancer susceptibility in a population.

## 7.4 The relationship between metal exposure and NER efficiency in a human population

Our investigation of the levels of NER efficiency in an at-risk population in Appalachian Kentucky, a population where lung cancer rates are greater than expected based on smoking prevalence, also included collecting environmental exposure data.  Of principle interest was the potential exposure to arsenic, an established human lung carcinogen, and the relationship between any exposure and NER efficiency.  Since we had observed that arsenic, in an acute, high concentration exposure, could reduce NER efficiency in multiple cell culture models, we hypothesized that arsenic exposure in a human population could also result in a reduction in NER efficiency.  A U.S. study found that individuals with

toe-nail arsenic levels above a certain threshold (0.114 µg arsenic/1 gram nail) had a significant association with small-cell and squamous-cell carcinoma of the lung compared with those in the lowest exposure group [83]. A previous study indicated that residents of Appalachian Kentucky had significantly higher toenail arsenic levels than a non-Appalachian reference county, with over a quarter of all samples exceeding the 0.114 µg/g threshold [321]. We ran a similar comparison, looking at lung cancer risk in the highest exposed population compared to the lowest exposure group, and determined that there was no appreciable increase in cancer incidence (odds ratio ). The arsenic exposure levels in the reference study were considerably higher than those seen in the Appalachian Kentucky study, which may make meaningful comparisons based on arsenic nail levels difficult. About 22% of the nail samples in the reference study, across cases and controls, had arsenic nail levels above the threshold indicated above. In our study, only 11% of the subjects whose lymphocytes were analyzed for repair had toenail arsenic levels above the same threshold. Even though the levels of arsenic in the nails were low, we also investigated the relationship between arsenic nail levels and the efficiency of NER (Fig 7.4 A). We found no trend in half-life values with increasing arsenic concentrations in subject toenails. In general, toenail metal analysis revealed that there was little exposure to toxic metals such as arsenic, cadmium, lead, and nickel (data not shown). These low levels, while certainly a positive for the individuals in the region, make any analysis of the effects of heavy metal exposure difficult. We therefore shifted focus to look at the effect of essential trace metals, such as selenium,

manganese, chromium and iron, on the efficiency of NER for any potential

protective effects.  Although we saw no relationship between chromium and

selenium levels in toenails and NER efficiency, we observed a positive

correlation between the efficiency of NER and toenail concentrations of both

manganese and iron in a statistically significant manner (Fig 7.4 B/C).  Iron is an

essential cofactor for all three mammalian DNA polymerases and is required by

the NER factor XPD to facilitate its helicase activities [322].  Manganese is used

by SOD2 to reduce oxidative damage in the mitochondria, where the superoxide

radical is generated during mitochondrial respiration.  Oxidation of NER factors

has been implicated as a mechanism through which oxidative damage may

inhibit NER [323].  It follows then that, within the non-toxic range of

concentrations of these two metals, higher levels of iron and manganese could

promote DNA repair, and indeed we observed an increase in the efficiency of

NER with increasing toenail concentrations of manganese and iron.  We then

chose to examine the relationship between the efficiency of NER and nail metal

exposure between individuals who had the highest exposure of a given metal to

the remainder of the population (Table 7.2).  When comparing the toenail

samples containing the highest concentrations of manganese or iron

(approximately top 10%) to the remainder of the samples, the high metal level

groups had a greater DNA repair rate (reduced half-life) compared to the

remainder of the study population (Fig 7.5 A/B).  In the case of manganese, the

difference was statistically significant.  When subjects were grouped based on

their toenail arsenic levels, on the other hand, there was no difference in NER

efficiency between the two groups (Fig 7.5 C).  In this case, the cutoff used was

close to the level (0.114 µg As/g nail) that produced a significant risk for

developing certain lung cancers [83].

## 7.5 Future directions

The NER pathway is vital to the prevention of DNA mutations that drive

carcinogenesis.  The efficiency of the pathway fluctuates within a population,

both as a consequence of an age-dependent decline in NER efficiency as well as

the existence of many polymorphisms in the genes involved in NER.  Additional

environmental factors may also contribute to variations in NER efficiency.  Less

efficient NER faculties could result in increased mutational burdens, and

individuals who are subject to higher levels of DNA damage (like smokers) would

be at an even higher risk of developing cancer should their NER efficiency be

low.  It is this premise that has drawn us into investigating the NER pathway as

potential biomarker for cancer susceptibility.  Our efforts exploring the causes for

increased lung cancer incidence in Appalachian Kentucky produced mixed

results.  Overall, there was no reduction in NER efficiency in the lung cancer

subjects compared to the control subjects.  After observing that these two

populations showed very different responses to increasing subject age as it

related to NER efficiency, and after concluding that smoking had a dramatic

effect on the observed age-dependent decline in NER efficiency, we determined

that both smoking and age must be controlled for when performing a case-control

study using our measurement of NER efficiency.  Unfortunately, correcting for

age and smoking status reduced the sample size of the only population that

showed any substantial difference in NER efficiency between cases and controls, young smokers.

In addition to the hypothesis that reduced NER efficiency could increase lung cancer risk in the Appalachian Kentucky population, there were two additional hypotheses that were developed after cell culture work determined that environmental factors, namely tobacco smoke (condensate) and arsenic, could inhibit the efficiency of NER. We hypothesed that (1) tobacco smoke or (2) arsenic could cause a reduction in the NER efficiency of individuals within the study population who were exposed to these environmental agents. This then extends the hypothesis that reduced NER efficiency could increase cancer risk to include a possible explanation as to how the NER efficiency of individuals could be reduced. As this was the first time the repair assay was used to measure NER efficiency in a human population, there is no external reference point to compare the measures to, and it is therefore impossible to determine if individuals in Appalachia as a whole have reduced NER efficiency compared to those outside the region. Further studies are required to compare NER efficiency measured in this study to populations outside the region. Rather, the study permitted comparisons of NER efficiency between lung cancer cases and controls, as well as between individuals who were and were not exposed to various environmental agents. The results of the comparison of half-lives between cases and controls has been laid out above. The impact of tobacco smoke on NER efficiency was not what was expected. As detailed in chapter 6, the different smoking populations did not have statistically different average half-

lives, after adjusting for the difference in ages between the populations. The cell culture work suggested that tobacco smoke could potentially inhibit NER efficiency, and as such a hypothesis was established that posited that smokers in Appalachia would have a reduced NER efficiency as a consequence of their smoking. Instead, the current and former smokers had faster rates of repair relative to the non-smoking population, although the difference was not significant, and was likely highly influenced by the disparity in average ages of the smoking populations (See Table 6.1). The other hypothesis cultivated by the cell culture studies that applied to the human population study on NER efficiency was related to arsenic exposure. Arsenic, in cultured skin and lung cells, reduced NER efficiency. It was hypothesized that this would occur in the human population study, that individuals exposed to arsenic would have reduced NER efficiency. This was tied to an additional hypothesis that arsenic levels in the region were sufficiently high to increase lung cancer incidence in exposed individuals.

Neither of these turned out to be the case. Arsenic levels were quite low in the nail samples from individuals in the Appalachian study, and there was no correlation between arsenic levels and either cancer incidence or NER efficiency. Given the low level of exposure, it would not be prudent to conclude that arsenic is not a risk factor for lung cancer development, but rather that this population lacked sufficient exposure to arsenic to observe any affect it might have on lung cancer risk. The same can be said for the impact of arsenic exposure on NER efficiency. There is no comparable study that investigated NER efficiency as it

correlates to arsenic exposure in a human population, so the conclusion that arsenic levels were insufficiently high to perform a comparison with NER efficiency will have to remain based on the fact that arsenic levels were not high enough to increase lung cancer risk.

Rather than try to draw conclusions regarding the impact of heavy metal exposure on NER efficiency when exposure was quite low, another opportunity presented itself within the metal data collected. Essential trace metal exposure data was also collected, and there were sufficient levels of several of these metals to examine their relationship to NER efficiency. As indicated in section 7.4, manganese and iron showed a positive relationship with NER efficiency in a statistically significant manner. While it was not part of the hypothesis relating to lung cancer incidence in the region, the positive association between NER efficiency and manganese and iron levels provides evidence that this measure of NER efficiency can be correlated to environmental exposures. The positive association between essential trace metals and NER efficiency may pave the way for further studies of the relationship between NER efficiency and metal exposure in individuals who have higher levels of exposure to toxic metals like arsenic, cadmium, nickel, and lead.

Ultimately, NER efficiency may be a useful tool in predicting cancer risk, especially among individuals who are already at increased risk due to factors like smoking or excessive sun exposure. The repair assay discussed herein needs to be validated in a large case-control study with appropriate matching conditions. Age, smoking status, and perhaps even gender should all be

considered when establishing a case-control study that measures NER efficiency as an indicator of cancer risk.  As cancer is a disease of aging, and since cases and controls showed dramatically different age-dependent changes in NER efficiency, measuring NER efficiency using the technique described herein would like be most effective in a younger population.  As it is designed to provide a measure of cancer risk, those individuals with poor NER efficiency would be able to take that knowledge and perhaps decide to mitigate their risk by reducing exposure to DNA damaging agents such as sunlight and tobacco smoke. Knowing you are at a high risk for developing cancer may be the incentive needed to make a lifestyle change, avoiding additional risk factors, which will reduce the risk of developing cancer.

**Table 7.1 The effect of subject age on NER efficiency in different populations of the study**

| Population | Factor | N | $R^2$ | P-value |
|---|---|---|---|---|
| Total | All | 198 | 0.0586 | **0.0005** |
| | Male | 95 | 0.0993 | **0.002** |
| | Female | 103 | 0.0293 | 0.084 |
| | Current smokers | 52 | 0.1198 | **0.012** |
| | Former smokers | 65 | 0.1055 | **0.008** |
| | Never smokers | 81 | 0.0021 | 0.685 |
| Control | All | 156 | 0.0655 | **0.0013** |
| | Male | 78 | 0.0792 | **0.013** |
| | Female | 78 | 0.0554 | **0.038** |
| | Current smokers | 26 | 0.4759 | **0.0001** |
| | Former smokers | 51 | 0.1059 | **0.02** |
| | Never smokers | 79 | 0.0062 | 0.491 |
| Case | All | 42 | 0.0496 | 0.1563 |
| | Male | 17 | 0.065 | 0.3234 |
| | Female | 25 | 0.0002 | 0.947 |
| | Current smokers | 26 | 0.0025 | 0.8083 |
| | Former smokers | 14 | 0.0256 | 0.5848 |
| | Never smokers | 2 | -- | -- |

**Table 7.1** Subject age was plotted against half-life for each individual in a given population and a linear regression analysis was performed.  The $R^2$ value and subsequent P-value derived from the $R^2$ value are shown in Table 7.1 for the total, control, and case populations as well as further divisions within each of these groups.

**Table 7.2 Half-life values of individuals in high metal exposure populations compared to the remainder of the population**

| Factor | Subject Type | | | | | | P-value |
|---|---|---|---|---|---|---|---|
| Metal | High exposure | | | Rest of population | | | |
| | N | Cutoff | Mean ± SD | N | Cutoff | Mean ± SD | |
| Arsenic | 18 | > .125 ug/g | 4.38 ± 1.80 | 139 | < .125 ug/g | 4.10 ± 1.81 | 0.537 |
| Manganese | 16 | > 1 ug/g | 3.29 ± 1.36 | 141 | < 1 ug/g | 4.23 ± 1.83 | 0.048 |
| Iron | 15 | > 30 ug/g | 3.47 ± 1.60 | 142 | < 30 ug/g | 4.20 ± 1.81 | 0.133 |

**Table 7.2** Individuals who had toenail levels of arsenic, manganese, or iron above cutoffs representing approximately the 90th percentile for each metal we compared to those below the cutoff for each metal.  The sample size, half-lives, and standard deviation are shown for each population, along with a p-value from a 2-way Student's T-test comparing the half-lives of individuals above and below the cutoff for each metal are shown.

**Figure 7.1 NER efficiency is not affected by subject age in lung cancer subjects**
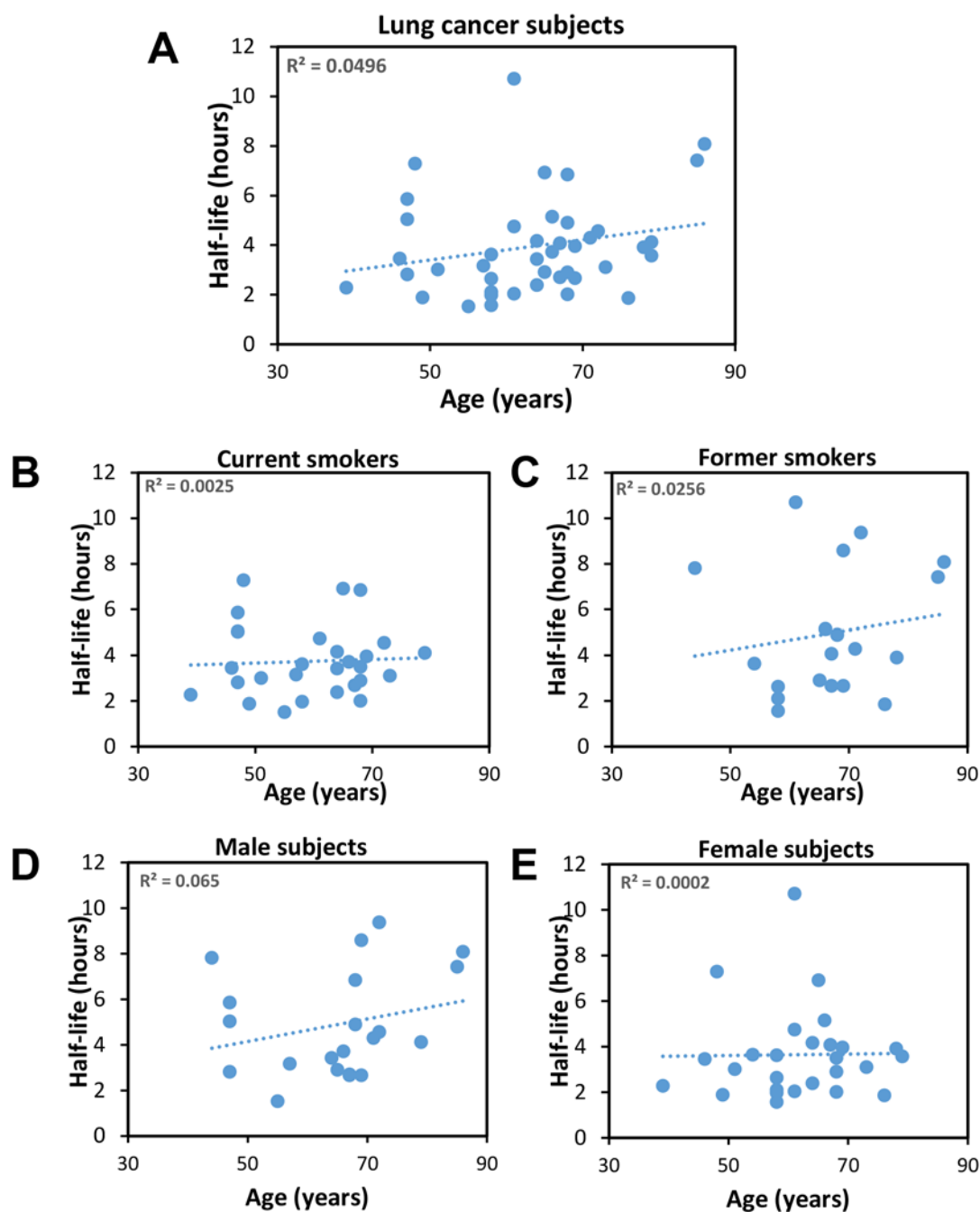
**Figure 7.1** Scatter plots of the distribution of half-lives compared to subject age is shown for (A) the total lung cancer population, (B) the current smokers with lung cancer, (C) the former smokers with lung cancer, (D) the male subjects with lung cancer, and (E) the female subjects with lung cancer. All populations show no significant reduction in NER efficiency with increasing subject age. Regression analysis of the scatter plots provided $R^2$ values of 0.0496, 0.0025, 0.0256, 0.0650, and 0.0002 for the total, current smoking, forming smoking, male, and female populations, respectively.  The associated p-values were all non-significant (see table 7.1).

**Figure 7.2 NER efficiency in cases verses controls separated by age group**



**A**
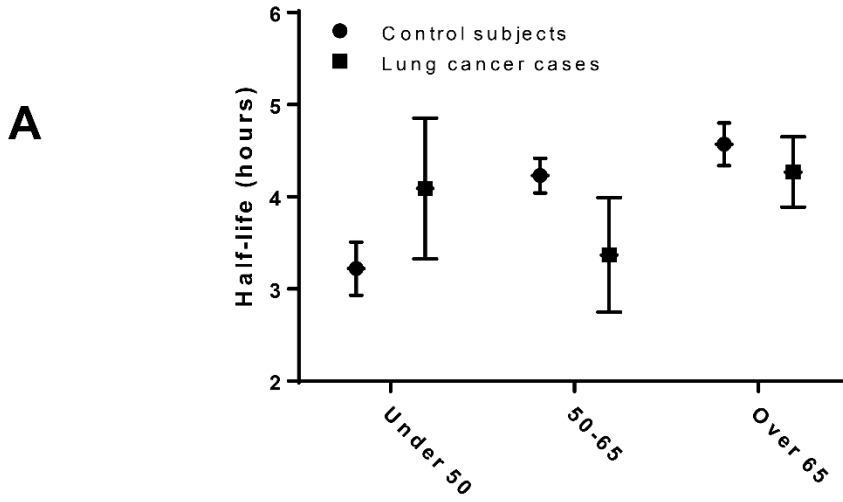
| Factor | Subject Type | | | | P-value |
|---|---|---|---|---|---|
| Age group | Control | | Case | | |
| | N | Mean ± SD | N | Mean ± SD | |
| Under 50 | 19 | 3.22 ± 1.25 | 7 | 4.09 ± 2.02 | 0.0982 |
| 50-65 | 76 | 4.23 ± 1.63 | 14 | 3.37 ± 2.32 | 0.0472 |
| Over 65 | 61 | 4.57 ± 1.80 | 21 | 4.27 ± 1.75 | 0.2545 |

**B**

**Figure 7.2** (A) A graph of the mean +/- standard error for half-life in each age group separated by case and control subjects. (B) A table of the results shown in (A). When comparing the cases to control at each age group, the controls have a greater NER efficiency in the youngest age group (although not significantly), while they have a reduced NER efficiency in the middle and oldest age groups compared to the cases.

**Figure 7.3 NER efficiency in current smoking cases verses controls**

**separated by age group**



**A**

| Factor | Subject Type | | | | P-value |
|---|---|---|---|---|---|
| Age group | Control | | Case | | |
| | N | Mean ± SD | N | Mean ± SD | |
| Under 50 | 7 | 2.45 ± 0.58 | 7 | 4.09 ± 2.02 | 0.0307 |
| 50-65 | 13 | 3.68 ± 1.48 | 9 | 3.12 ± 1.03 | 0.1689 |
| Over 65 | 6 | 5.77 ± 1.72 | 10 | 4.09 ± 1.65 | 0.0316 |

**B**
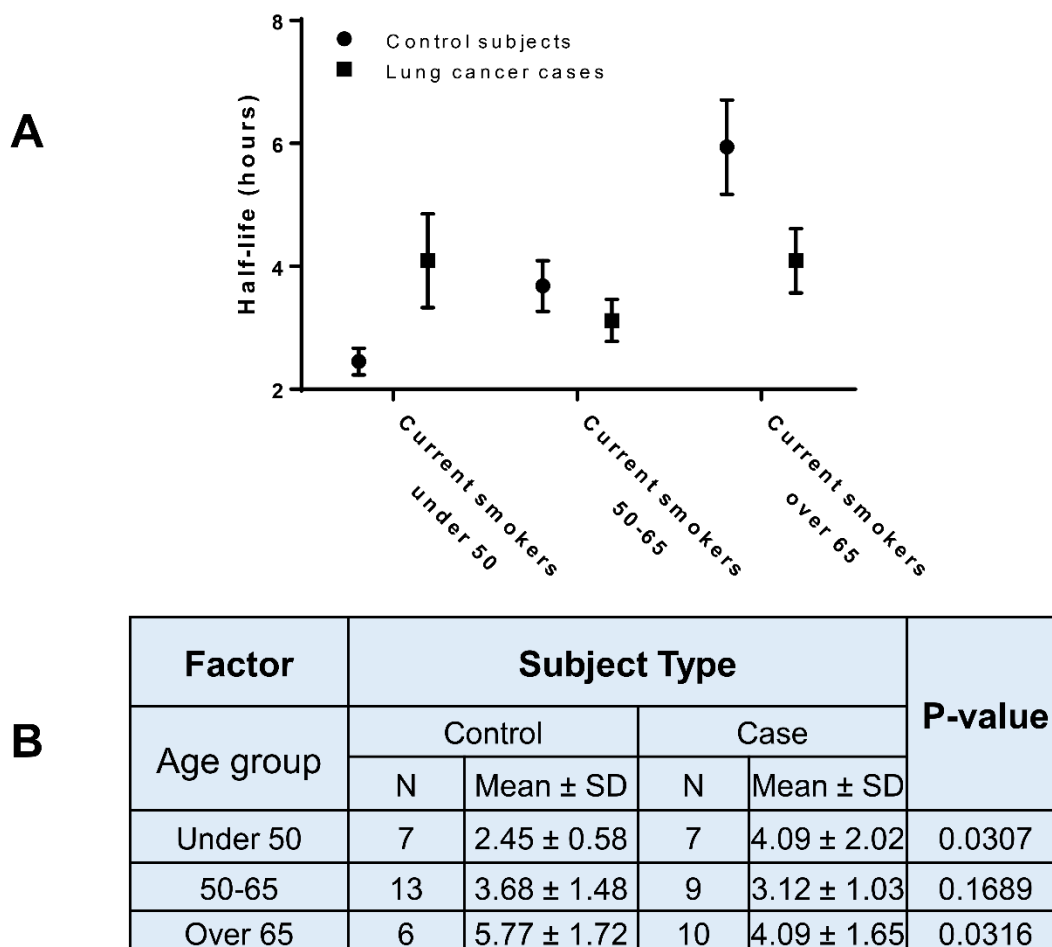
**Figure 7.3** (A) A graph of the mean +/- standard error for half-life of current smokers in each age group separated by case and control subjects. (B) A table of the results shown in (A). When comparing the cases to control at each age group, the controls have a greater NER efficiency in the youngest age group (p = 0.0307), while they have a lower NER efficiency in the middle and oldest age groups compared to the cases (p = 0.1689 and 0.0316 for middle and oldest age groups, respectively).

**Figure 7.4 The relationship between NER efficiency and toenail metal concentration in a human population**
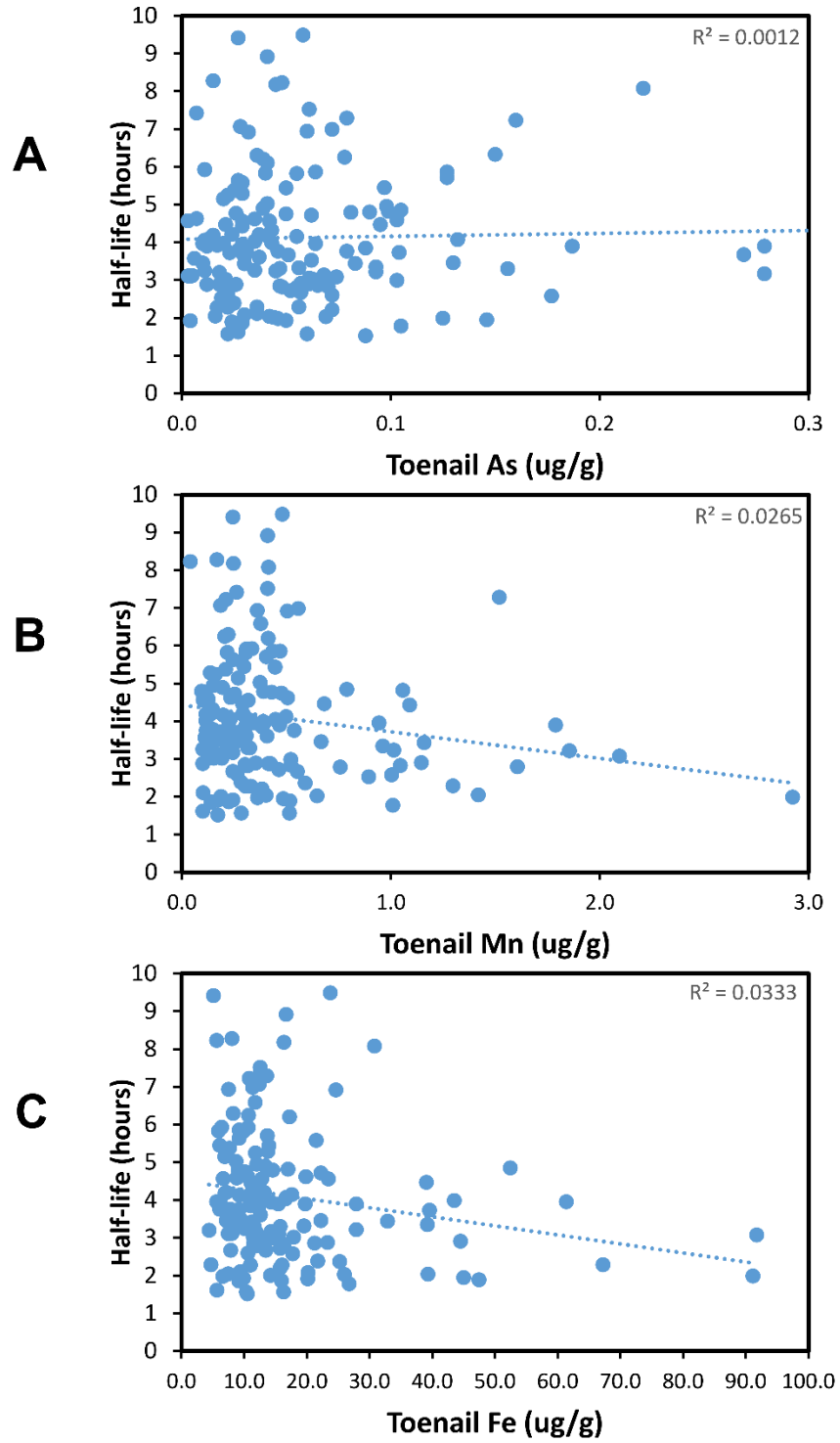
**Figure 7.4** (A) A scatter plot of the half-life vs toenail arsenic levels shows no

correlation.  A linear regression analysis of the data produced an $R^2$ value of

0.0012 (p = 0.667).  (B) A scatter plot of the half-life vs toenail manganese levels

shows a reduction in half-life with increasing manganese concentration.  A linear

regression analysis of the data produced an $R^2$ value of 0.0265 (p = 0.045). (C)

A scatter plot of the half-life vs toenail iron levels shows a reduction in half-life

with increasing iron concentration.  A linear regression analysis of the data

produced an $R^2$ value of 0.0333 (p = 0.026).

**Figure 7.5 NER efficiency of individuals in the study population separated into high and low metal exposure for Mn, Fe, and As**
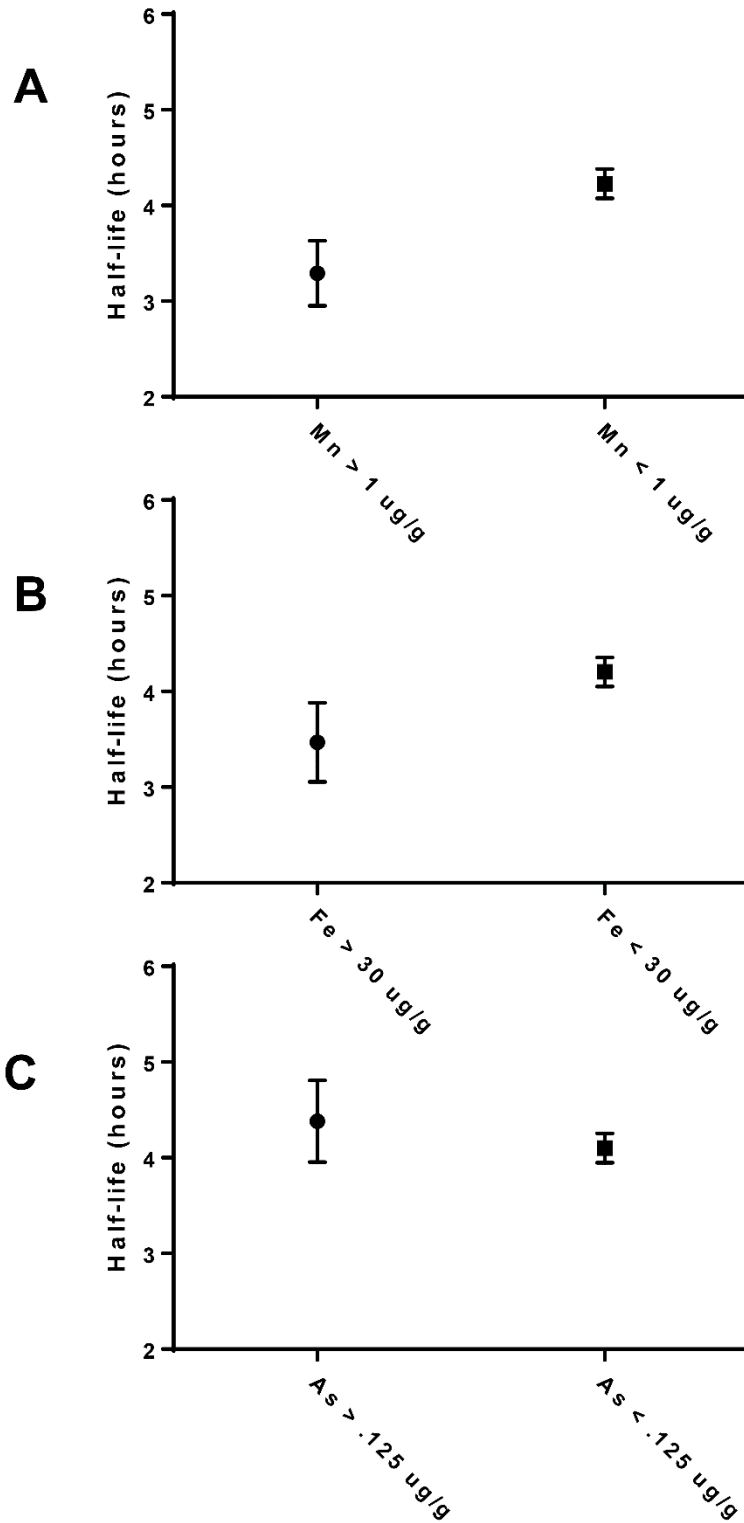
**Figure 7.5** (A) A graph of the average half-life and standard error from individuals above and below 1 µg Manganese / 1 g toenail is shown. (B) A graph of the average half-life and standard error from individuals above and below 30 µg Iron / 1 g toenail is shown. (C) A graph of the average half-life and standard error from individuals above and below 0.125 µg Arsenic/ 1 g toenail is shown. These cutoffs roughly separate the population into a 10% high exposure population and the remaining 90% of the population for each metal exposure.

# Bibliography

1. Prior, I.A., P.D. Lewis, and C. Mattos, *A comprehensive survey of Ras mutations in cancer.* Cancer Res, 2012. **72**(10): p. 2457-67.
2. Rivlin, N., et al., *Mutations in the p53 Tumor Suppressor Gene: Important Milestones at the Various Steps of Tumorigenesis.* Genes Cancer, 2011. **2**(4): p. 466-74.
3. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer.* Cell, 2000. **100**(1): p. 57-70.
4. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. **144**(5): p. 646-74.
5. Siegel, R.L., K.D. Miller, and A. Jemal, *Cancer statistics, 2015.* CA Cancer J Clin, 2015. **65**(1): p. 5-29.
6. Jones, D.S., S.H. Podolsky, and J.A. Greene, *The burden of disease and the changing task of medicine.* N Engl J Med, 2012. **366**(25): p. 2333-8.
7. Crimmins, E.M., *Lifespan and Healthspan: Past, Present, and Promise.* Gerontologist, 2015. **55**(6): p. 901-11.
8. Hoyert, D.L., *75 years of mortality in the United States, 1935-2010.* NCHS Data Brief, 2012(88): p. 1-8.
9. Xu, J., et al., *Deaths: Final Data for 2013.* Natl Vital Stat Rep, 2016. **64**(2): p. 1-119.
10. Kochanek, K.D., et al., *Deaths: Final Data for 2014.* Natl Vital Stat Rep, 2016. **65**(4): p. 1-122.
11. USDHSS, in *The Health Consequences of Smoking-50 Years of Progress: A Report of the Surgeon General.* 2014: Atlanta (GA).
12. Schoenberg, N.E., et al., *Trends in cigarette smoking and obesity in Appalachian Kentucky.* South Med J, 2015. **108**(3): p. 170-7.
13. Wilson, R.J., et al., *Cancer Incidence in Appalachia, 2004-2011.* Cancer Epidemiol Biomarkers Prev, 2016. **25**(2): p. 250-8.
14. Christian, W.J., et al., *Exploring geographic variation in lung cancer incidence in Kentucky using a spatial scan statistic: elevated risk in the Appalachian coal-mining region.* Public Health Rep, 2011. **126**(6): p. 789-96.
15. Hendryx, M., K. O'Donnell, and K. Horn, *Lung cancer mortality is elevated in coal-mining areas of Appalachia.* Lung Cancer, 2008. **62**(1): p. 1-7.
16. Hoeijmakers, J.H., *Genome maintenance mechanisms for preventing cancer.* Nature, 2001. **411**(6835): p. 366-74.
17. Gates, K.S., *An overview of chemical processes that damage cellular DNA: spontaneous hydrolysis, alkylation, and reactions with radicals.* Chem Res Toxicol, 2009. **22**(11): p. 1747-60.
18. Marnett, L.J., J.N. Riggins, and J.D. West, *Endogenous generation of reactive oxidants and electrophiles and their reactions with DNA and protein.* J Clin Invest, 2003. **111**(5): p. 583-93.
19. Gillet, L.C. and O.D. Scharer, *Molecular mechanisms of mammalian global genome nucleotide excision repair.* Chem Rev, 2006. **106**(2): p. 253-76.
20. Smith, M.L. and Y.R. Seo, *p53 regulation of DNA excision repair pathways.* Mutagenesis, 2002. **17**(2): p. 149-56.
21. Adimoolam, S. and J.M. Ford, *p53 and DNA damage-inducible expression of the xeroderma pigmentosum group C gene.* Proc Natl Acad Sci U S A, 2002. **99**(20): p. 12985-90.

22. Amundson, S.A., et al., *A nucleotide excision repair master-switch: p53 regulated coordinate induction of global genomic repair genes.* Cancer Biol Ther, 2002. **1**(2): p. 145-9.

23. Fitch, M.E., I.V. Cross, and J.M. Ford, *p53 responsive nucleotide excision repair gene products p48 and XPC, but not p53, localize to sites of UV-irradiation-induced DNA damage, in vivo.* Carcinogenesis, 2003. **24**(5): p. 843-50.

24. Vermeulen, W. and M. Fousteri, *Mammalian transcription-coupled excision repair.* Cold Spring Harb Perspect Biol, 2013. **5**(8): p. a012625.

25. Scharer, O.D., *Nucleotide excision repair in eukaryotes.* Cold Spring Harb Perspect Biol, 2013. **5**(10): p. a012609.

26. Sugasawa, K., et al., *Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair.* Mol Cell, 1998. **2**(2): p. 223-32.

27. Sugasawa, K., et al., *A molecular mechanism for DNA damage recognition by the xeroderma pigmentosum group C protein complex.* DNA Repair (Amst), 2002. **1**(1): p. 95-107.

28. Yeo, J.E., et al., *The efficiencies of damage recognition and excision correlate with duplex destabilization induced by acetylaminofluorene adducts in human nucleotide excision repair.* Chem Res Toxicol, 2012. **25**(11): p. 2462-8.

29. Luijsterburg, M.S., et al., *Stochastic and reversible assembly of a multiprotein DNA repair complex ensures accurate target site recognition and efficient repair.* J Cell Biol, 2010. **189**(3): p. 445-63.

30. Nishi, R., et al., *UV-DDB-dependent regulation of nucleotide excision repair kinetics in living cells.* DNA Repair (Amst), 2009. **8**(6): p. 767-76.

31. Svejstrup, J.Q., *Mechanisms of transcription-coupled DNA repair.* Nat Rev Mol Cell Biol, 2002. **3**(1): p. 21-9.

32. Spivak, G. and A.K. Ganesan, *The complex choreography of transcription-coupled repair.* DNA Repair (Amst), 2014. **19**: p. 64-70.

33. Riedl, T., F. Hanaoka, and J.M. Egly, *The comings and goings of nucleotide excision repair factors on damaged DNA.* EMBO J, 2003. **22**(19): p. 5293-303.

34. Sugasawa, K., et al., *Two-step recognition of DNA damage for mammalian nucleotide excision repair: Directional binding of the XPC complex and DNA strand scanning.* Mol Cell, 2009. **36**(4): p. 642-53.

35. Orelli, B., et al., *The XPA-binding domain of ERCC1 is required for nucleotide excision repair but not other DNA repair pathways.* J Biol Chem, 2010. **285**(6): p. 3705-12.

36. Shivji, M.K., et al., *Nucleotide excision repair DNA synthesis by DNA polymerase epsilon in the presence of PCNA, RFC, and RPA.* Biochemistry, 1995. **34**(15): p. 5011-7.

37. Ogi, T., et al., *Three DNA polymerases, recruited by different mechanisms, carry out NER repair synthesis in human cells.* Mol Cell, 2010. **37**(5): p. 714-27.

38. Moser, J., et al., *Sealing of chromosomal DNA nicks during nucleotide excision repair requires XRCC1 and DNA ligase III alpha in a cell-cycle-specific manner.* Mol Cell, 2007. **27**(2): p. 311-23.

39. Lehmann, A.R., *DNA repair-deficient diseases, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy.* Biochimie, 2003. **85**(11): p. 1101-11.

40. Bradford, P.T., et al., *Cancer and neurologic degeneration in xeroderma pigmentosum: long term follow-up characterises the role of DNA repair.* J Med Genet, 2011. **48**(3): p. 168-76.

41.     Kraemer, K.H., et al., *Xeroderma pigmentosum, trichothiodystrophy and Cockayne syndrome: a complex genotype-phenotype relationship.* Neuroscience, 2007. **145**(4): p. 1388-96.

42.     Kraemer, K.H. and J.J. DiGiovanna, *Forty years of research on xeroderma pigmentosum at the US National Institutes of Health.* Photochem Photobiol, 2015. **91**(2): p. 452-9.

43.     Hirai, Y., et al., *Heterozygous individuals bearing a founder mutation in the XPA DNA repair gene comprise nearly 1% of the Japanese population.* Mutat Res, 2006. **601**(1-2): p. 171-8.

44.     Moriwaki, S., et al., *Absence of DNA repair deficiency in the confirmed heterozygotes of xeroderma pigmentosum group A.* J Invest Dermatol, 1993. **101**(1): p. 69-72.

45.     Berg, R.J., et al., *Relative susceptibilities of XPA knockout mice and their heterozygous and wild-type littermates to UVB-induced skin cancer.* Cancer Res, 1997. **57**(4): p. 581-4.

46.     Khan, S.G., et al., *Reduced XPC DNA repair gene mRNA levels in clinically normal parents of xeroderma pigmentosum patients.* Carcinogenesis, 2006. **27**(1): p. 84-94.

47.     Yang, M., et al., *Associations between XPC expression, genotype, and the risk of head and neck cancer.* Environ Mol Mutagen, 2005. **45**(4): p. 374-9.

48.     Iyer, R.R., et al., *DNA mismatch repair: functions and mechanisms.* Chem Rev, 2006. **106**(2): p. 302-23.

49.     McCulloch, S.D., L. Gu, and G.M. Li, *Nick-dependent and -independent processing of large DNA loops in human cells.* J Biol Chem, 2003. **278**(50): p. 50803-9.

50.     Li, G.M., *Mechanisms and functions of DNA mismatch repair.* Cell Res, 2008. **18**(1): p. 85-98.

51.     Fukui, K., *DNA mismatch repair in eukaryotes and bacteria.* J Nucleic Acids, 2010. **2010**.

52.     Lynch, H.T., et al., *Review of the Lynch syndrome: history, molecular genetics, screening, differential diagnosis, and medicolegal ramifications.* Clin Genet, 2009. **76**(1): p. 1-18.

53.     Fishel, R. and R.D. Kolodner, *Identification of mismatch repair genes and their role in the development of cancer.* Curr Opin Genet Dev, 1995. **5**(3): p. 382-95.

54.     Lo, Y.L., et al., *Polymorphisms of MLH1 and MSH2 genes and the risk of lung cancer among never smokers.* Lung Cancer, 2011. **72**(3): p. 280-6.

55.     Nogueira, G.A., et al., *Association between genetic polymorphisms in DNA mismatch repair-related genes with risk and prognosis of head and neck squamous cell carcinoma.* Int J Cancer, 2015. **137**(4): p. 810-8.

56.     Zhu, H., et al., *Polymorphisms in mismatch repair genes are associated with risk and microsatellite instability of gastric cancer, and interact with life exposures.* Gene, 2016. **579**(1): p. 52-7.

57.     Weren, R.D., et al., *A germline homozygous mutation in the base-excision repair gene NTHL1 causes adenomatous polyposis and colorectal cancer.* Nat Genet, 2015. **47**(6): p. 668-71.

58.     Al-Tassan, N., et al., *Inherited variants of MYH associated with somatic G:C-->T:A mutations in colorectal tumors.* Nat Genet, 2002. **30**(2): p. 227-32.

59.     van Puijenbroek, M., et al., *The natural history of a combined defect in MSH6 and MUTYH in a HNPCC family.* Fam Cancer, 2007. **6**(1): p. 43-51.

60.     Robertson, A.B., et al., *DNA repair in mammalian cells: Base excision repair: the long and short of it.* Cell Mol Life Sci, 2009. **66**(6): p. 981-93.

61.     Cavalieri, E.L., et al., *Molecular origin of cancer: catechol estrogen-3,4-quinones as endogenous tumor initiators.* Proc Natl Acad Sci U S A, 1997. **94**(20): p. 10937-42.

62.     Ali, M.F., et al., *Prevalence of BER gene polymorphisms in sporadic breast cancer.* Oncol Rep, 2008. **19**(4): p. 1033-8.

63.     Ali, K., et al., *Germline variations of apurinic/apyrimidinic endonuclease 1 (APEX1) detected in female breast cancer patients.* Asian Pac J Cancer Prev, 2014. **15**(18): p. 7589-95.

64.     Jamal, A., et al., *Current Cigarette Smoking Among Adults - United States, 2005-2015.* MMWR Morb Mortal Wkly Rep, 2016. **65**(44): p. 1205-1211.

65.     Hecht, S.S., *Research opportunities related to establishing standards for tobacco products under the Family Smoking Prevention and Tobacco Control Act.* Nicotine Tob Res, 2012. **14**(1): p. 18-28.

66.     Hecht, S.S., *Progress and challenges in selected areas of tobacco carcinogenesis.* Chem Res Toxicol, 2008. **21**(1): p. 160-71.

67.     Hecht, S.S., *Tobacco smoke carcinogens and lung cancer.* J Natl Cancer Inst, 1999. **91**(14): p. 1194-210.

68.     Jiang, H., et al., *Metabolism of benzo[a]pyrene in human bronchoalveolar H358 cells using liquid chromatography-mass spectrometry.* Chem Res Toxicol, 2007. **20**(9): p. 1331-41.

69.     Shou, M., F.J. Gonzalez, and H.V. Gelboin, *Stereoselective epoxidation and hydration at the K-region of polycyclic aromatic hydrocarbons by cDNA-expressed cytochromes P450 1A1, 1A2, and epoxide hydrolase.* Biochemistry, 1996. **35**(49): p. 15807-13.

70.     Eisenstadt, E., et al., *Carcinogenic epoxides of benzo[a]pyrene and cyclopenta[cd]pyrene induce base substitutions via specific transversions.* Proc Natl Acad Sci U S A, 1982. **79**(6): p. 1945-9.

71.     Peltonen, K. and A. Dipple, *Polycyclic aromatic hydrocarbons: chemistry of DNA adduct formation.* J Occup Environ Med, 1995. **37**(1): p. 52-8.

72.     Cheng, S.C., et al., *DNA adducts from carcinogenic and noncarcinogenic enantiomers of benzo[a]pyrene dihydrodiol epoxide.* Chem Res Toxicol, 1989. **2**(5): p. 334-40.

73.     Chiapperino, D., et al., *Preferential misincorporation of purine nucleotides by human DNA polymerase eta opposite benzo[a]pyrene 7,8-diol 9,10-epoxide deoxyguanosine adducts.* J Biol Chem, 2002. **277**(14): p. 11765-71.

74.     Hess, M.T., et al., *Base pair conformation-dependent excision of benzo[a]pyrene diol epoxide-guanine adducts by human nucleotide excision repair enzymes.* Mol Cell Biol, 1997. **17**(12): p. 7069-76.

75.     Hang, B., *Formation and repair of tobacco carcinogen-derived bulky DNA adducts.* J Nucleic Acids, 2010. **2010**: p. 709521.

76.     Zhao, B., et al., *Poleta, Polzeta and Rev1 together are required for G to T transversion mutations induced by the (+)- and (-)-trans-anti-BPDE-N2-dG DNA adducts in yeast cells.* Nucleic Acids Res, 2006. **34**(2): p. 417-25.

77.     Hollstein, M., et al., *p53 mutations in human cancers.* Science, 1991. **253**(5015): p. 49-53.

78.     Denissenko, M.F., et al., *Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in P53.* Science, 1996. **274**(5286): p. 430-2.

79.     Pfeifer, G.P., et al., *Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers.* Oncogene, 2002. **21**(48): p. 7435-51.

80.     Tornaletti, S. and G.P. Pfeifer, *Complete and tissue-independent methylation of CpG sites in the p53 gene: implications for mutations in human cancers.* Oncogene, 1995. **10**(8): p. 1493-9.

81.     Yoon, J.H., et al., *Methylated CpG dinucleotides are the preferential targets for G-to-T transversion mutations induced by benzo[a]pyrene diol epoxide in mammalian cells:*

*similarities with the p53 mutation spectrum in smoking-associated lung cancers.* Cancer Res, 2001. **61**(19): p. 7110-7.

82. Celik, I., et al., *Arsenic in drinking water and lung cancer: a systematic review.* Environ Res, 2008. **108**(1): p. 48-55.

83. Heck, J.E., et al., *Lung cancer in a U.S. population with low to moderate arsenic exposure.* Environ Health Perspect, 2009. **117**(11): p. 1718-23.

84. Rossman, T.G., *Mechanism of arsenic carcinogenesis: an integrated approach.* Mutat Res, 2003. **533**(1-2): p. 37-65.

85. Shiber, J.G., *Arsenic in domestic well water and health in central appalachia, USA.* Water Air Soil Pollut, 2005.

86. Tuttle, M.L.W., G.N. Breit, and M.B. Goldhaber, *Geochemical data from New Albany Shale, Kentucky : a study in metal mobility during weathering of black shales*, in *Open-File Report*. 2003.

87. Kolker, A., et al., *Mode of occurrence of arsenic in four US coals.* Fuel Processing Technology, 2000. **63**(2): p. 167-178.

88. Goldhaber, M.B., et al., *Role of Large Scale Fluid-Flow in Subsurface Arsenic Enrichment*, in *Arsenic in Ground Water: Geochemistry and Occurrence*, A.H. Welch and K.G. Stollenwerk, Editors. 2003, Springer US: Boston, MA. p. 127-164.

89. Tuttle, M.L.W., et al., *Arsenic in rocks and stream sediments of the central Appalachian Basin, Kentucky*, in *Open-File Report*. 2002.

90. Holzman, D.C., *Mountaintop removal mining: digging into community health concerns.* Environ Health Perspect, 2011. **119**(11): p. A476-83.

91. Yuspa, S.H., et al., *Clonal growth of mouse epidermal cells in medium with reduced calcium concentration.* J Invest Dermatol, 1981. **76**(2): p. 144-6.

92. Pillsbury, H.C. and C.C. Bright, *Comparison of aliquot and complete sample procedure for the determination of nicotine in cigarette smoke.* J Assoc Off Anal Chem, 1972. **55**(3): p. 636-8.

93. Riedhammer, C., D. Halbritter, and R. Weissert, *Peripheral Blood Mononuclear Cells: Isolation, Freezing, Thawing, and Culture.* Methods Mol Biol, 2016. **1304**: p. 53-61.

94. Mayer, C., et al., *DNA repair capacity after gamma-irradiation and expression profiles of DNA repair genes in resting and proliferating human peripheral blood lymphocytes.* DNA Repair (Amst), 2002. **1**(3): p. 237-50.

95. Freeman, S.E. and S.L. Ryan, *Excision repair of pyrimidine dimers in human peripheral blood lymphocytes: comparison between mitogen stimulated and unstimulated cells.* Mutat Res, 1988. **194**(2): p. 143-50.

96. Ford, J.M. and P.C. Hanawalt, *Expression of wild-type p53 is required for efficient global genomic nucleotide excision repair in UV-irradiated human fibroblasts.* J Biol Chem, 1997. **272**(44): p. 28073-80.

97. Mellon, I., et al., *Polymorphisms in the human xeroderma pigmentosum group A gene and their impact on cell survival and nucleotide excision repair.* DNA Repair (Amst), 2002. **1**(7): p. 531-46.

98. Islami, F., et al., *Potentially preventable premature lung cancer deaths in the USA if overall population rates were reduced to those of educated whites in lower-risk states.* Cancer Causes Control, 2015.

99. Jemal, A., et al., *Global cancer statistics.* CA Cancer J Clin, 2011. **61**(2): p. 69-90.

100. Jemal, A., et al., *Cancer statistics, 2010.* CA Cancer J Clin, 2010. **60**(5): p. 277-300.

101. Islami, F., L.A. Torre, and A. Jemal, *Global trends of lung cancer mortality and smoking prevalence.* Transl Lung Cancer Res, 2015. **4**(4): p. 327-38.

102. Torre, L.A., et al., *Global Cancer Incidence and Mortality Rates and Trends--An Update.* Cancer Epidemiol Biomarkers Prev, 2016. **25**(1): p. 16-27.

103. Larsen, J.E. and J.D. Minna, *Molecular biology of lung cancer: clinical implications.* Clin Chest Med, 2011. **32**(4): p. 703-40.

104. Hecht, S.S., *Lung carcinogenesis by tobacco smoke.* Int J Cancer, 2012. **131**(12): p. 2724-32.

105. DeMarini, D.M., *Genotoxicity of tobacco smoke and tobacco smoke condensate: a review.* Mutat Res, 2004. **567**(2-3): p. 447-74.

106. Youlden, D.R., S.M. Cramb, and P.D. Baade, *The International Epidemiology of Lung Cancer: geographical distribution and secular trends.* J Thorac Oncol, 2008. **3**(8): p. 819-31.

107. Gibbons, D.L., L.A. Byers, and J.M. Kurie, *Smoking, p53 mutation, and lung cancer.* Mol Cancer Res, 2014. **12**(1): p. 3-13.

108. Talikka, M., et al., *Genomic impact of cigarette smoke, with application to three smoking-related diseases.* Crit Rev Toxicol, 2012. **42**(10): p. 877-89.

109. Gunz, D., M.T. Hess, and H. Naegeli, *Recognition of DNA adducts by human nucleotide excision repair. Evidence for a thermodynamic probing mechanism.* J Biol Chem, 1996. **271**(41): p. 25089-98.

110. Meschini, R., et al., *DNA repair deficiency and BPDE-induced chromosomal alterations in CHO cells.* Mutat Res, 2008. **637**(1-2): p. 93-100.

111. Quan, T. and J.C. States, *Preferential DNA damage in the p53 gene by benzo[a]pyrene metabolites in cytochrome P4501A1-expressing xeroderma pigmentosum group A cells.* Mol Carcinog, 1996. **16**(1): p. 32-43.

112. Quan, T., et al., *Differential mutagenicity and cytotoxicity of (+/-)-benzo[a]pyrene-trans-7,8-dihydrodiol and (+/-)-anti-benzo[a]pyrene-trans-7,8-dihydrodiol-9,10-epoxide in genetically engineered human fibroblasts.* Mol Carcinog, 1995. **12**(2): p. 91-102.

113. Porter, P.C., I. Mellon, and J.C. States, *XP-A cells complemented with Arg228Gln and Val234Leu polymorphic XPA alleles repair BPDE-induced DNA damage better than cells complemented with the wild type allele.* DNA Repair (Amst), 2005. **4**(3): p. 341-9.

114. Hwang, B.J., et al., *Expression of the p48 xeroderma pigmentosum gene is p53-dependent and is involved in global genomic repair.* Proc Natl Acad Sci U S A, 1999. **96**(2): p. 424-8.

115. Barckhausen, C., et al., *Malignant melanoma cells acquire resistance to DNA interstrand cross-linking chemotherapeutics by p53-triggered upregulation of DDB2/XPC-mediated DNA repair.* Oncogene, 2014. **33**(15): p. 1964-74.

116. Coggins, C.R., *A further review of inhalation studies with cigarette smoke and lung cancer in experimental animals, including transgenic mice.* Inhal Toxicol, 2010. **22**(12): p. 974-83.

117. Hecht, S.S., *Carcinogenicity studies of inhaled cigarette smoke in laboratory animals: old and new.* Carcinogenesis, 2005. **26**(9): p. 1488-92.

118. Hutt, J.A., et al., *Life-span inhalation exposure to mainstream cigarette smoke induces lung cancer in B6C3F1 mice through genetic and epigenetic pathways.* Carcinogenesis, 2005. **26**(11): p. 1999-2009.

119. Witschi, H., *Tobacco smoke-induced lung cancer in animals--a challenge to toxicology (?).* Int J Toxicol, 2007. **26**(4): p. 339-44.

120. Balansky, R., et al., *Prevention of cigarette smoke-induced lung tumors in mice by budesonide, phenethyl isothiocyanate, and N-acetylcysteine.* Int J Cancer, 2010. **126**(5): p. 1047-54.

121.     Kwon, M.C. and A. Berns, *Mouse models for lung cancer.* Mol Oncol, 2013. **7**(2): p. 165-77.

122.     Chen, G., et al., *Induction of lacZ mutations in MutaMouse primary hepatocytes.* Environ Mol Mutagen, 2010. **51**(4): p. 330-7.

123.     DeMarini, D.M., et al., *Genotoxicity of 10 cigarette smoke condensates in four test systems: comparisons between assays and condensates.* Mutat Res, 2008. **650**(1): p. 15-29.

124.     Guo, X., et al., *Mutagenicity of 11 cigarette smoke condensates in two versions of the mouse lymphoma assay.* Mutagenesis, 2011. **26**(2): p. 273-81.

125.     Hsu, T.C., et al., *Mitosis-arresting effects of cigarette smoke condensate on human lymphoid cell lines.* Mutat Res, 1991. **259**(1): p. 67-78.

126.     Luo, L.Z., et al., *Cigarette smoke induces anaphase bridges and genomic imbalances in normal cells.* Mutat Res, 2004. **554**(1-2): p. 375-85.

127.     Moktar, A., et al., *Cigarette smoke-induced DNA damage and repair detected by the comet assay in HPV-transformed cervical cells.* Int J Oncol, 2009. **35**(6): p. 1297-304.

128.     Foresta, M., et al., *Accelerated repair and reduced mutagenicity of DNA damage induced by cigarette smoke in human bronchial cells transfected with E.coli formamidopyrimidine DNA glycosylase.* PLoS One, 2014. **9**(1): p. e87984.

129.     Jaiswal, A.S., et al., *Adenomatous polyposis coli-mediated accumulation of abasic DNA lesions lead to cigarette smoke condensate-induced neoplastic transformation of normal breast epithelial cells.* Neoplasia, 2013. **15**(4): p. 454-60.

130.     Moktar, A., et al., *Cigarette smoke condensate-induced oxidative DNA damage and its removal in human cervical cancer cells.* Int J Oncol, 2011. **39**(4): p. 941-7.

131.     Kundu, C.N., et al., *Cigarette smoke condensate-induced level of adenomatous polyposis coli blocks long-patch base excision repair in breast epithelial cells.* Oncogene, 2007. **26**(10): p. 1428-38.

132.     Narayan, S., et al., *Cigarette smoke condensate-induced transformation of normal human breast epithelial cells in vitro.* Oncogene, 2004. **23**(35): p. 5880-9.

133.     Han, S.G., et al., *Bhas 42 cell transformation activity of cigarette smoke condensate is modulated by selenium and arsenic.* Environ Mol Mutagen, 2016.

134.     Curtin, G.M., et al., *Short-term in vitro and in vivo analyses for assessing the tumor-promoting potentials of cigarette smoke condensates.* Toxicol Sci, 2004. **81**(1): p. 14-25.

135.     Gairola, C., *Genetic effects of fresh cigarette smoke in Saccharomyces cerevisiae.* Mutat Res, 1982. **102**(2): p. 123-36.

136.     Walaszek, Z., M. Hanausek, and T.J. Slaga, *The role of skin painting in predicting lung cancer.* Int J Toxicol, 2007. **26**(4): p. 345-51.

137.     Friedberg, E.C., et al., *Defective nucleotide excision repair in xpc mutant mice and its association with cancer predisposition.* Mutat Res, 2000. **459**(2): p. 99-108.

138.     Ide, F., et al., *Mice deficient in the nucleotide excision repair gene XPA have elevated sensitivity to benzo[a]pyrene induction of lung tumors.* Carcinogenesis, 2000. **21**(6): p. 1263-5.

139.     Ishikawa, T., et al., *Importance of DNA repair in carcinogenesis: evidence from transgenic and gene targeting studies.* Mutat Res, 2001. **477**(1-2): p. 41-9.

140.     Hollander, M.C., et al., *Deletion of XPC leads to lung tumors in mice and is associated with early events in human lung carcinogenesis.* Proc Natl Acad Sci U S A, 2005. **102**(37): p. 13200-5.

141. Wang, Y.C., et al., *Evidence from mutation spectra that the UV hypermutability of xeroderma pigmentosum variant cells reflects abnormal, error-prone replication on a template containing photoproducts.* Mol Cell Biol, 1993. **13**(7): p. 4276-83.

142. Moser, J., et al., *The UV-damaged DNA binding protein mediates efficient targeting of the nucleotide excision repair complex to UV-induced photo lesions.* DNA Repair (Amst), 2005. **4**(5): p. 571-82.

143. Sugasawa, K., et al., *UV-induced ubiquitylation of XPC protein mediated by UV-DDB-ubiquitin ligase complex.* Cell, 2005. **121**(3): p. 387-400.

144. Wang, Q.E., et al., *Ubiquitylation-independent degradation of Xeroderma pigmentosum group C protein is required for efficient nucleotide excision repair.* Nucleic Acids Res, 2007. **35**(16): p. 5338-50.

145. Wang, Q.E., et al., *DNA repair factor XPC is modified by SUMO-1 and ubiquitin following UV irradiation.* Nucleic Acids Res, 2005. **33**(13): p. 4023-34.

146. He, J., et al., *Ubiquitin-specific protease 7 regulates nucleotide excision repair through deubiquitinating XPC protein and preventing XPC protein from undergoing ultraviolet light-induced and VCP/p97 protein-regulated proteolysis.* J Biol Chem, 2014. **289**(39): p. 27278-89.

147. Wang, H.T., et al., *Effect of carcinogenic acrolein on DNA repair and mutagenic susceptibility.* J Biol Chem, 2012. **287**(15): p. 12379-86.

148. Tang, M.S., et al., *Acrolein induced DNA damage, mutagenicity and effect on DNA repair.* Mol Nutr Food Res, 2011. **55**(9): p. 1291-300.

149. Feng, Z., et al., *Acrolein is a major cigarette-related lung cancer agent: Preferential binding at p53 mutational hotspots and inhibition of DNA repair.* Proc Natl Acad Sci U S A, 2006. **103**(42): p. 15404-9.

150. Nollen, M., et al., *Impact of arsenic on nucleotide excision repair: XPC function, protein level, and gene expression.* Mol Nutr Food Res, 2009. **53**(5): p. 572-82.

151. Renaud, E., et al., *Differential contribution of XPC, RAD23A, RAD23B and CENTRIN 2 to the UV-response in human cells.* DNA Repair (Amst), 2011. **10**(8): p. 835-47.

152. Qiu, J., et al., *Attenuated NER expressions of XPF and XPC associated with smoking are involved in the recurrence of bladder cancer.* PLoS One, 2014. **9**(12): p. e115224.

153. Muotri, A.R., et al., *Low amounts of the DNA repair XPA protein are sufficient to recover UV-resistance.* Carcinogenesis, 2002. **23**(6): p. 1039-46.

154. Koberle, B., V. Roginskaya, and R.D. Wood, *XPA protein as a limiting factor for nucleotide excision repair and UV sensitivity in human cells.* DNA Repair (Amst), 2006. **5**(5): p. 641-8.

155. Kang, T.H., J.T. Reardon, and A. Sancar, *Regulation of nucleotide excision repair activity by transcriptional and post-transcriptional control of the XPA protein.* Nucleic Acids Res, 2011. **39**(8): p. 3176-87.

156. Matsumoto, S., et al., *Functional regulation of the DNA damage-recognition factor DDB2 by ubiquitination and interaction with xeroderma pigmentosum group C protein.* Nucleic Acids Res, 2015. **43**(3): p. 1700-13.

157. Puumalainen, M.R., et al., *Chromatin retention of DNA damage sensors DDB2 and XPC through loss of p97 segregase causes genotoxicity.* Nat Commun, 2014. **5**: p. 3695.

158. van Cuijk, L., et al., *SUMO and ubiquitin-dependent XPC exchange drives nucleotide excision repair.* Nat Commun, 2015. **6**: p. 7499.

159. Poulsen, S.L., et al., *RNF111/Arkadia is a SUMO-targeted ubiquitin ligase that facilitates the DNA damage response.* J Cell Biol, 2013. **201**(6): p. 797-807.

160. Povlsen, L.K., et al., *Systems-wide analysis of ubiquitylation dynamics reveals a key role for PAF15 ubiquitylation in DNA-damage bypass.* Nat Cell Biol, 2012. **14**(10): p. 1089-98.

161. Bunick, C.G., et al., *Biochemical and structural domain analysis of xeroderma pigmentosum complementation group C protein.* Biochemistry, 2006. **45**(50): p. 14965-79.

162. Ng, J.M., et al., *A novel regulation mechanism of DNA repair by damage-induced and RAD23-dependent stabilization of xeroderma pigmentosum group C protein.* Genes Dev, 2003. **17**(13): p. 1630-45.

163. Wagner, S.A., et al., *A proteome-wide, quantitative survey of in vivo ubiquitylation sites reveals widespread regulatory roles.* Mol Cell Proteomics, 2011. **10**(10): p. M111 013284.

164. Kim, S.Y., et al., *Cigarette smoke induces Akt protein degradation by the ubiquitin-proteasome system.* J Biol Chem, 2011. **286**(37): p. 31932-43.

165. Nik-Zainal, S., et al., *The genome as a record of environmental exposure.* Mutagenesis, 2015. **30**(6): p. 763-70.

166. Greenblatt, M.S., et al., *Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis.* Cancer Res, 1994. **54**(18): p. 4855-78.

167. Hainaut, P. and G.P. Pfeifer, *Patterns of p53 G-->T transversions in lung cancers reflect the primary mutagenic signature of DNA-damage by tobacco smoke.* Carcinogenesis, 2001. **22**(3): p. 367-74.

168. Ruggeri, B., et al., *Benzo[a]pyrene-induced murine skin tumors exhibit frequent and characteristic G to T mutations in the p53 gene.* Proc Natl Acad Sci U S A, 1993. **90**(3): p. 1013-7.

169. Bolt, H.M., *Arsenic: an ancient toxicant of continuous public health impact, from Iceman Otzi until now.* Arch Toxicol, 2012. **86**(6): p. 825-30.

170. IARC, *Arsenic, metals, fibres, and dusts.* IARC Monogr Eval Carcinog Risks Hum, 2012. **100**(Pt C): p. 11-465.

171. Nordstrom, D.K., *Public health. Worldwide occurrences of arsenic in ground water.* Science, 2002. **296**(5576): p. 2143-5.

172. Yager, J.W., T. Greene, and R.A. Schoof, *Arsenic relative bioavailability from diet and airborne exposures: Implications for risk assessment.* Sci Total Environ, 2015. **536**: p. 368-81.

173. Campbell, R.C., W.E. Stephens, and A.A. Meharg, *Consistency of arsenic speciation in global tobacco products with implications for health and regulation.* Tob Induc Dis, 2014. **12**(1): p. 24.

174. Naujokas, M.F., et al., *The broad scope of health effects from chronic arsenic exposure: update on a worldwide public health problem.* Environ Health Perspect, 2013. **121**(3): p. 295-302.

175. IARC, *Some drinking-water disinfectants and contaminants, including arsenic.* IARC Monogr Eval Carcinog Risks Hum, 2004. **84**: p. 1-477.

176. Karagas, M.R., et al., *Drinking Water Arsenic Contamination, Skin Lesions, and Malignancies: A Systematic Review of the Global Evidence.* Curr Environ Health Rep, 2015. **2**(1): p. 52-68.

177. Surdu, S., et al., *Occupational exposure to arsenic and risk of nonmelanoma skin cancer in a multinational European study.* Int J Cancer, 2013. **133**(9): p. 2182-91.

178. Hunt, K.M., et al., *The mechanistic basis of arsenicosis: pathogenesis of skin cancer.* Cancer Lett, 2014. **354**(2): p. 211-9.

179. Hubaux, R., et al., *Molecular features in arsenic-induced lung tumors.* Mol Cancer, 2013. **12**: p. 20.

180. Hopenhayn-Rich, C., M.L. Biggs, and A.H. Smith, *Lung and kidney cancer mortality associated with arsenic in drinking water in Cordoba, Argentina.* Int J Epidemiol, 1998. **27**(4): p. 561-9.

181. Wang, W., S. Cheng, and D. Zhang, *Association of inorganic arsenic exposure with liver cancer mortality: A meta-analysis.* Environ Res, 2014. **135**: p. 120-5.

182. Saint-Jacques, N., et al., *Arsenic in drinking water and urinary tract cancers: a systematic review of 30 years of epidemiological evidence.* Environ Health, 2014. **13**: p. 44.

183. Scheindlin, S., *The duplicitous nature of inorganic arsenic.* Mol Interv, 2005. **5**(2): p. 60-4.

184. Hughes, M.F., et al., *Arsenic exposure and toxicology: a historical perspective.* Toxicol Sci, 2011. **123**(2): p. 305-32.

185. WHO, *Background Document for Development of WHO Guidelines for Drinking-Water Quality. Geneva, Switzerland: World Health Organization; Arsenic in Drinking-Water.* 2011.

186. Gilbert-Diamond, D., et al., *A population-based case-control study of urinary arsenic species and squamous cell carcinoma in New Hampshire, USA.* Environ Health Perspect, 2013. **121**(10): p. 1154-60.

187. Kwong, R.C., et al., *Arsenic exposure predicts bladder cancer survival in a US population.* World J Urol, 2010. **28**(4): p. 487-92.

188. Firkin, F., F. Roncolato, and W.K. Ho, *Dose-adjusted arsenic trioxide for acute promyelocytic leukaemia in chronic renal failure.* Eur J Haematol, 2015. **95**(4): p. 331-5.

189. Wang, Z.Y. and Z. Chen, *Acute promyelocytic leukemia: from highly fatal to highly curable.* Blood, 2008. **111**(5): p. 2505-15.

190. Mi, J.Q., et al., *Synergistic targeted therapy for acute promyelocytic leukaemia: a model of translational research in human cancer.* J Intern Med, 2015. **278**(6): p. 627-42.

191. Wu, F., et al., *Arsenite-induced alterations of DNA photodamage repair and apoptosis after solar-simulation UVR in mouse keratinocytes in vitro.* Environ Health Perspect, 2005. **113**(8): p. 983-6.

192. Okui, T. and Y. Fujiwara, *Inhibition of human excision DNA repair by inorganic arsenic and the co-mutagenic effect in V79 Chinese hamster cells.* Mutat Res, 1986. **172**(1): p. 69-76.

193. Danaee, H., et al., *Low dose exposure to sodium arsenite synergistically interacts with UV radiation to induce mutations and alter DNA repair in human cells.* Mutagenesis, 2004. **19**(2): p. 143-8.

194. Li, J.H. and T.G. Rossman, *Comutagenesis of sodium arsenite with ultraviolet radiation in Chinese hamster V79 cells.* Biol Met, 1991. **4**(4): p. 197-200.

195. Lee, T.C., R.Y. Huang, and K.Y. Jan, *Sodium arsenite enhances the cytotoxicity, clastogenicity, and 6-thioguanine-resistant mutagenicity of ultraviolet light in Chinese hamster ovary cells.* Mutat Res, 1985. **148**(1-2): p. 83-9.

196. Rossman, T.G., *Enhancement of UV-mutagenesis by low concentrations of arsenite in E. coli.* Mutat Res, 1981. **91**(3): p. 207-11.

197. Shen, S., et al., *Arsenite and its mono- and dimethylated trivalent metabolites enhance the formation of benzo[a]pyrene diol epoxide-DNA adducts in Xeroderma pigmentosum complementation group A cells.* Chem Res Toxicol, 2009. **22**(2): p. 382-90.

198. Shen, S., et al., *Elevation of cellular BPDE uptake by human cells: a possible factor contributing to co-carcinogenicity by arsenite.* Environ Health Perspect, 2006. **114**(12): p. 1832-7.

199. Schwerdtle, T., I. Walter, and A. Hartwig, *Arsenite and its biomethylated metabolites interfere with the formation and repair of stable BPDE-induced DNA adducts in human cells and impair XPAzf and Fpg.* DNA Repair (Amst), 2003. **2**(12): p. 1449-63.

200. Shen, S., et al., *Attenuation of DNA damage-induced p53 expression by arsenic: a possible mechanism for arsenic co-carcinogenesis.* Mol Carcinog, 2008. **47**(7): p. 508-18.

201. Chiang, H.C. and T.C. Tsou, *Arsenite enhances the benzo[a]pyrene diol epoxide (BPDE)-induced mutagenesis with no marked effect on repair of BPDE-DNA adducts in human lung cells.* Toxicol In Vitro, 2009. **23**(5): p. 897-905.

202. Ferreccio, C., et al., *Arsenic, tobacco smoke, and occupation: associations of multiple agents with lung and bladder cancer.* Epidemiology, 2013. **24**(6): p. 898-905.

203. Fischer, J.M., et al., *Co-mutagenic activity of arsenic and benzo[a]pyrene in mouse skin.* Mutat Res, 2005. **588**(1): p. 35-46.

204. Evans, C.D., et al., *Effect of arsenic on benzo[a]pyrene DNA adduct levels in mouse skin and lung.* Carcinogenesis, 2004. **25**(4): p. 493-7.

205. Morikawa, T., et al., *Promotion of skin carcinogenesis by dimethylarsinic acid in keratin (K6)/ODC transgenic mice.* Jpn J Cancer Res, 2000. **91**(6): p. 579-81.

206. Motiwale, L., A.D. Ingle, and K.V. Rao, *Mouse skin tumor promotion by sodium arsenate is associated with enhanced PCNA expression.* Cancer Lett, 2005. **223**(1): p. 27-35.

207. Burns, F.J., et al., *Arsenic-induced enhancement of ultraviolet radiation carcinogenesis in mouse skin: a dose-response study.* Environ Health Perspect, 2004. **112**(5): p. 599-603.

208. Rossman, T.G., et al., *Arsenite is a cocarcinogen with solar ultraviolet radiation for mouse skin: an animal model for arsenic carcinogenesis.* Toxicol Appl Pharmacol, 2001. **176**(1): p. 64-71.

209. Shah, P. and Y.Y. He, *Molecular regulation of UV-induced DNA repair.* Photochem Photobiol, 2015. **91**(2): p. 254-64.

210. Feng, Z., et al., *N-hydroxy-4-aminobiphenyl-DNA binding in human p53 gene: sequence preference and the effect of C5 cytosine methylation.* Biochemistry, 2002. **41**(20): p. 6414-21.

211. Kraemer, K.H. and J.J. DiGiovanna, *Xeroderma Pigmentosum*, in *GeneReviews(R)*, R.A. Pagon, et al., Editors. 1993: Seattle (WA).

212. Feltes, B.C. and D. Bonatto, *Overview of xeroderma pigmentosum proteins architecture, mutations and post-translational modifications.* Mutat Res Rev Mutat Res, 2015. **763**: p. 306-20.

213. Marteijn, J.A., et al., *Understanding nucleotide excision repair and its roles in cancer and ageing.* Nat Rev Mol Cell Biol, 2014. **15**(7): p. 465-81.

214. Li, J.H. and T.G. Rossman, *Inhibition of DNA ligase activity by arsenite: a possible mechanism of its comutagenesis.* Mol Toxicol, 1989. **2**(1): p. 1-9.

215. Hartwig, A., et al., *Interaction of arsenic(III) with nucleotide excision repair in UV-irradiated human fibroblasts.* Carcinogenesis, 1997. **18**(2): p. 399-405.

216. Lynn, S., et al., *Arsenite retards DNA break rejoining by inhibiting DNA ligation.* Mutagenesis, 1997. **12**(5): p. 353-8.

217. Hu, Y., L. Su, and E.T. Snow, *Arsenic toxicity is enzyme specific and its affects on ligation are not caused by the direct inhibition of DNA repair enzymes.* Mutat Res, 1998. **408**(3): p. 203-18.

218. Grosskopf, C., et al., *Antimony impairs nucleotide excision repair: XPA and XPE as potential molecular targets.* Chem Res Toxicol, 2010. **23**(7): p. 1175-83.

219. Nguyen, T.A., et al., *The oncogenic phosphatase WIP1 negatively regulates nucleotide excision repair.* DNA Repair (Amst), 2010. **9**(7): p. 813-23.

220. Pan, J., et al., *Nucleotide excision repair deficiency increases levels of acrolein-derived cyclic DNA adduct and sensitizes cells to apoptosis induced by docosahexaenoic acid and acrolein.* Mutat Res, 2016. **789**: p. 33-8.

221. de Waard, H., et al., *Cell-type-specific consequences of nucleotide excision repair deficiencies: Embryonic stem cells versus fibroblasts.* DNA Repair (Amst), 2008. **7**(10): p. 1659-69.

222. Clement, V., I. Dunand-Sauthier, and S.G. Clarkson, *Suppression of UV-induced apoptosis by the human DNA repair protein XPG.* Cell Death Differ, 2006. **13**(3): p. 478-88.

223. Andrew, A.S., M.R. Karagas, and J.W. Hamilton, *Decreased DNA repair gene expression among individuals exposed to arsenic in United States drinking water.* Int J Cancer, 2003. **104**(3): p. 263-8.

224. Miyamoto, I., et al., *Mutational analysis of the structure and function of the xeroderma pigmentosum group A complementing protein. Identification of essential domains for nuclear localization and DNA excision repair.* J Biol Chem, 1992. **267**(17): p. 12182-7.

225. Okuda, Y., et al., *Relative levels of the two mammalian Rad23 homologs determine composition and stability of the xeroderma pigmentosum group C protein complex.* DNA Repair (Amst), 2004. **3**(10): p. 1285-95.

226. Ikehata, H. and T. Ono, *Significance of CpG methylation for solar UV-induced mutagenesis and carcinogenesis in skin.* Photochem Photobiol, 2007. **83**(1): p. 196-204.

227. Cheng, P.S., et al., *Relationship between arsenic-containing drinking water and skin cancers in the arseniasis endemic areas in Taiwan.* J Dermatol, 2016. **43**(2): p. 181-6.

228. Cooper, K.L., J.W. Yager, and L.G. Hudson, *Melanocytes and keratinocytes have distinct and shared responses to ultraviolet radiation and arsenic.* Toxicol Lett, 2014. **224**(3): p. 407-15.

229. Chervona, Y., A. Arita, and M. Costa, *Carcinogenic metals and the epigenome: understanding the effect of nickel, arsenic, and chromium.* Metallomics, 2012. **4**(7): p. 619-27.

230. Riedmann, C., et al., *Inorganic Arsenic-induced cellular transformation is coupled with genome wide changes in chromatin structure, transcriptome and splicing patterns.* BMC Genomics, 2015. **16**: p. 212.

231. Chen, H., et al., *Genetic events associated with arsenic-induced malignant transformation: applications of cDNA microarray technology.* Mol Carcinog, 2001. **30**(2): p. 79-87.

232. Benbrahim-Tallaa, L., et al., *Molecular events associated with arsenic-induced malignant transformation of human prostatic epithelial cells: aberrant genomic DNA methylation and K-ras oncogene activation.* Toxicol Appl Pharmacol, 2005. **206**(3): p. 288-98.

233. Zhao, C.Q., et al., *Association of arsenic-induced malignant transformation with DNA hypomethylation and aberrant gene expression.* Proc Natl Acad Sci U S A, 1997. **94**(20): p. 10907-12.

234. Jensen, T.J., et al., *Arsenicals produce stable progressive changes in DNA methylation patterns that are linked to malignant transformation of immortalized urothelial cells.* Toxicol Appl Pharmacol, 2009. **241**(2): p. 221-9.

235. Treas, J., T. Tyagi, and K.P. Singh, *Chronic exposure to arsenic, estrogen, and their combination causes increased growth and transformation in human prostate epithelial cells potentially by hypermethylation-mediated silencing of MLH1.* Prostate, 2013. **73**(15): p. 1660-72.

236. Xu, Y., E.J. Tokar, and M.P. Waalkes, *Arsenic-induced cancer cell phenotype in human breast epithelia is estrogen receptor-independent but involves aromatase activation.* Arch Toxicol, 2014. **88**(2): p. 263-74.

237. Zhang, A., et al., *Aberrant methylation of nucleotide excision repair genes is associated with chronic arsenic poisoning.* Biomarkers, 2016: p. 1-10.

238. Liu, H., et al., *Arsenite-loaded nanoparticles inhibit PARP-1 to overcome multidrug resistance in hepatocellular carcinoma cells.* Sci Rep, 2016. **6**: p. 31009.

239. Bakhshaiesh, T.O., et al., *Arsenic Trioxide Promotes Paclitaxel Cytotoxicity in Resistant Breast Cancer Cells.* Asian Pac J Cancer Prev, 2015. **16**(13): p. 5191-7.

240. Kryeziu, K., et al., *Synergistic anticancer activity of arsenic trioxide with erlotinib is based on inhibition of EGFR-mediated DNA double-strand break repair.* Mol Cancer Ther, 2013. **12**(6): p. 1073-84.

241. Muenyi, C.S., et al., *Sodium arsenite +/- hyperthermia sensitizes p53-expressing human ovarian cancer cells to cisplatin by modulating platinum-DNA damage responses.* Toxicol Sci, 2012. **127**(1): p. 139-49.

242. Yang, J.L., et al., *Posttreatment with sodium arsenite alters the mutational spectrum induced by ultraviolet light irradiation in Chinese hamster ovary cells.* Environ Mol Mutagen, 1992. **20**(3): p. 156-64.

243. Hsu, C.H., et al., *Mutational spectrum of p53 gene in arsenic-related skin cancers from the blackfoot disease endemic area of Taiwan.* Br J Cancer, 1999. **80**(7): p. 1080-6.

244. Lai, J.P., et al., *The influence of DNA repair on neurological degeneration, cachexia, skin cancer and internal neoplasms: autopsy report of four xeroderma pigmentosum patients (XP-A, XP-C and XP-D).* Acta Neuropathol Commun, 2013. **1**: p. 4.

245. Emmert, S., et al., *The xeroderma pigmentosum group C gene leads to selective repair of cyclobutane pyrimidine dimers rather than 6-4 photoproducts.* Proc Natl Acad Sci U S A, 2000. **97**(5): p. 2151-6.

246. Wang, L., et al., *A modified host-cell reactivation assay to measure repair of alkylating DNA damage for assessing risk of lung adenocarcinoma.* Carcinogenesis, 2007. **28**(7): p. 1430-6.

247. Kleijer, W.J., et al., *Incidence of DNA repair deficiency disorders in western Europe: Xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy.* DNA Repair (Amst), 2008. **7**(5): p. 744-50.

248. Seetharam, S., et al., *Ultraviolet mutational spectrum in a shuttle vector propagated in xeroderma pigmentosum lymphoblastoid cells and fibroblasts.* Mutat Res, 1991. **254**(1): p. 97-105.

249. Cheng, L., et al., *Expression of nucleotide excision repair genes and the risk for squamous cell carcinoma of the head and neck.* Cancer, 2002. **94**(2): p. 393-7.

250. Liu, J., et al., *Nucleotide excision repair related gene polymorphisms and genetic susceptibility, chemotherapeutic sensitivity and prognosis of gastric cancer.* Mutat Res, 2014. **765**: p. 11-21.

251. Camps, C., et al., *Assessment of nucleotide excision repair XPD polymorphisms in the peripheral blood of gemcitabine/cisplatin-treated advanced non-small-cell lung cancer patients.* Clin Lung Cancer, 2003. **4**(4): p. 237-41.

252. Qiao, Y., et al., *Rapid assessment of repair of ultraviolet DNA damage with a modified host-cell reactivation assay using a luciferase reporter gene and correlation with polymorphisms of DNA repair genes in normal human lymphocytes.* Mutat Res, 2002. **509**(1-2): p. 165-74.

253. Jin, B., et al., *Association of XPC polymorphisms and lung cancer risk: a meta-analysis.* PLoS One, 2014. **9**(4): p. e93937.

254. Zou, Y.F., et al., *Association of XPC gene polymorphisms with susceptibility to prostate cancer: evidence from 3,936 subjects.* Genet Test Mol Biomarkers, 2013. **17**(12): p. 926-31.

255. Mocellin, S., D. Verdi, and D. Nitti, *DNA repair gene polymorphisms and risk of cutaneous melanoma: a systematic review and meta-analysis.* Carcinogenesis, 2009. **30**(10): p. 1735-43.

256. Peng, Q., et al., *Current evidences on XPC polymorphisms and gastric cancer susceptibility: a meta-analysis.* Diagn Pathol, 2014. **9**: p. 96.

257. Benhamou, S. and A. Sarasin, *ERCC2/XPD gene polymorphisms and cancer risk.* Mutagenesis, 2002. **17**(6): p. 463-9.

258. Athas, W.F., et al., *Development and field-test validation of an assay for DNA repair in circulating human lymphocytes.* Cancer Res, 1991. **51**(21): p. 5786-93.

259. Wei, Q., et al., *Reduced DNA repair capacity in lung cancer patients.* Cancer Res, 1996. **56**(18): p. 4103-7.

260. Wang, L.E., et al., *DNA repair capacity in peripheral lymphocytes predicts survival of patients with non-small-cell lung cancer treated with first-line platinum-based chemotherapy.* J Clin Oncol, 2011. **29**(31): p. 4121-8.

261. Saha, D.T., et al., *Quantification of DNA repair capacity in whole blood of patients with head and neck cancer and healthy donors by comet assay.* Mutat Res, 2008. **650**(1): p. 55-62.

262. Stoilov, L.M., et al., *Adaptive response to DNA and chromosomal damage induced by X-rays in human blood lymphocytes.* Mutagenesis, 2007. **22**(2): p. 117-22.

263. Gajecka, M., et al., *Reduced DNA repair capacity in laryngeal cancer subjects. A comparison of phenotypic and genotypic results.* Adv Otorhinolaryngol, 2005. **62**: p. 25-37.

264. Li, C., L.E. Wang, and Q. Wei, *DNA repair phenotype and cancer susceptibility--a mini review.* Int J Cancer, 2009. **124**(5): p. 999-1007.

265. Hu, J.J., et al., *Deficient nucleotide excision repair capacity enhances human prostate cancer risk.* Cancer Res, 2004. **64**(3): p. 1197-201.

266. Matta, J., et al., *The association of DNA Repair with breast cancer risk in women. A comparative observational study.* BMC Cancer, 2012. **12**: p. 490.

267. DiGiovanna, J.J. and K.H. Kraemer, *Shining a light on xeroderma pigmentosum.* J Invest Dermatol, 2012. **132**(3 Pt 2): p. 785-96.

268. Holcomb, N., et al., *Exposure of Human Lung Cells to Tobacco Smoke Condensate Inhibits the Nucleotide Excision Repair Pathway.* PLoS One, 2016. **11**(7): p. e0158858.

269. Wang, L.E., et al., *Reduced DNA repair capacity for removing tobacco carcinogen-induced DNA adducts contributes to risk of head and neck cancer but not tumor characteristics.* Clin Cancer Res, 2010. **16**(2): p. 764-74.

270. Bull, M., et al., *Defining blood processing parameters for optimal detection of cryopreserved antigen-specific responses for HIV vaccine trials.* J Immunol Methods, 2007. **322**(1-2): p. 57-69.

271. Mallone, R., et al., *Isolation and preservation of peripheral blood mononuclear cells for analysis of islet antigen-reactive T cell responses: position statement of the T-Cell Workshop Committee of the Immunology of Diabetes Society.* Clin Exp Immunol, 2011. **163**(1): p. 33-49.

272.    McKenna, K.C., et al., *Delayed processing of blood increases the frequency of activated CD11b+ CD15+ granulocytes which inhibit T cell function.* J Immunol Methods, 2009. **341**(1-2): p. 68-75.

273.    Gorlova, O.Y., et al., *DNA repair capacity and lung cancer risk in never smokers.* Cancer Epidemiol Biomarkers Prev, 2008. **17**(6): p. 1322-8.

274.    Gaivao, I., et al., *Comet assay-based methods for measuring DNA repair in vitro; estimates of inter- and intra-individual variation.* Cell Biol Toxicol, 2009. **25**(1): p. 45-52.

275.    Vande Loock, K., et al., *An aphidicolin-block nucleotide excision repair assay measuring DNA incision and repair capacity.* Mutagenesis, 2010. **25**(1): p. 25-32.

276.    Chang, J.L., et al., *DNA damage and repair measurements from cryopreserved lymphocytes without cell culture--a reproducible assay for intervention studies.* Environ Mol Mutagen, 2006. **47**(7): p. 503-8.

277.    Slyskova, J., et al., *DNA damage and nucleotide excision repair capacity in healthy individuals.* Environ Mol Mutagen, 2011. **52**(7): p. 511-7.

278.    Lehmann, A.R., D. McGibbon, and M. Stefanini, *Xeroderma pigmentosum.* Orphanet J Rare Dis, 2011. **6**: p. 70.

279.    Pesz, K.A., et al., *Polymorphisms in nucleotide excision repair genes and basal cell carcinoma of the skin.* Int J Dermatol, 2014. **53**(12): p. 1474-7.

280.    Emmert, S. and K.H. Kraemer, *Do not underestimate nucleotide excision repair: it predicts not only melanoma risk but also survival outcome.* J Invest Dermatol, 2013. **133**(7): p. 1713-7.

281.    Paszkowska-Szczur, K., et al., *Polymorphisms in nucleotide excision repair genes and susceptibility to colorectal cancer in the Polish population.* Mol Biol Rep, 2015. **42**(3): p. 755-64.

282.    Mirecka, A., et al., *Common variants of xeroderma pigmentosum genes and prostate cancer risk.* Gene, 2014. **546**(2): p. 156-61.

283.    Wang, B., et al., *The association of six polymorphisms of five genes involved in three steps of nucleotide excision repair pathways with hepatocellular cancer risk.* Oncotarget, 2016. **7**(15): p. 20357-20367.

284.    Rafiq, R., et al., *Potential risk of esophageal squamous cell carcinoma due to nucleotide excision repair XPA and XPC gene variants and their interaction among themselves and with environmental factors.* Tumour Biol, 2016.

285.    Motovali-Bashi, M., et al., *Association between XPD (Lys751G1n) Polymorphism and Lung Cancer Risk: A Population-Based Study in Iran.* Cell J, 2014. **16**(3): p. 309-14.

286.    Wheless, L., et al., *A community-based study of nucleotide excision repair polymorphisms in relation to the risk of non-melanoma skin cancer.* J Invest Dermatol, 2012. **132**(5): p. 1354-62.

287.    Cheng, L., et al., *Reduced expression levels of nucleotide excision repair genes in lung cancer: a case-control analysis.* Carcinogenesis, 2000. **21**(8): p. 1527-30.

288.    Etemadi, A., et al., *Variation in PAH-related DNA adduct levels among non-smokers: the role of multiple genetic polymorphisms and nucleotide excision repair phenotype.* Int J Cancer, 2013. **132**(12): p. 2738-47.

289.    Spitz, M.R., et al., *Genetic susceptibility to lung cancer: the role of DNA damage and repair.* Cancer Epidemiol Biomarkers Prev, 2003. **12**(8): p. 689-98.

290.    Woollons, A., et al., *The 0.8% ultraviolet B content of an ultraviolet A sunlamp induces 75% of cyclobutane pyrimidine dimers in human keratinocytes in vitro.* Br J Dermatol, 1999. **140**(6): p. 1023-30.

291.    Sugasawa, K., et al., *A multistep damage recognition mechanism for global genomic nucleotide excision repair.* Genes Dev, 2001. **15**(5): p. 507-21.

292.    Szymkowski, D.E., C.W. Lawrence, and R.D. Wood, *Repair by human cell extracts of single (6-4) and cyclobutane thymine-thymine photoproducts in DNA.* Proc Natl Acad Sci U S A, 1993. **90**(21): p. 9823-7.

293.    van der Wees, C., et al., *Nucleotide excision repair in differentiated cells.* Mutat Res, 2007. **614**(1-2): p. 16-23.

294.    Pines, A., et al., *Differential activity of UV-DDB in mouse keratinocytes and fibroblasts: impact on DNA repair and UV-induced skin cancer.* DNA Repair (Amst), 2009. **8**(2): p. 153-61.

295.    Rastogi, R.P., et al., *Molecular mechanisms of ultraviolet radiation-induced DNA damage and repair.* J Nucleic Acids, 2010. **2010**: p. 592980.

296.    Guerrero-Santoro, J., A.S. Levine, and V. Rapic-Otrin, *Co-localization of DNA repair proteins with UV-induced DNA damage in locally irradiated cells.* Methods Mol Biol, 2011. **682**: p. 149-61.

297.    Shen, H., et al., *Smoking, DNA repair capacity and risk of nonsmall cell lung cancer.* Int J Cancer, 2003. **107**(1): p. 84-8.

298.    Deng, L., et al., *Estimation of the effects of smoking and DNA repair capacity on coefficients of a carcinogenesis model for lung cancer.* Int J Cancer, 2009. **124**(9): p. 2152-8.

299.    Wang, L.E., et al., *Repair capacity for UV light induced DNA damage associated with risk of nonmelanoma skin cancer and tumor progression.* Clin Cancer Res, 2007. **13**(21): p. 6532-9.

300.    Wang, L.E., et al., *Genome-wide association study reveals novel genetic determinants of DNA repair capacity in lung cancer.* Cancer Res, 2013. **73**(1): p. 256-64.

301.    Gems, D. and L. Partridge, *Genetics of longevity in model organisms: debates and paradigm shifts.* Annu Rev Physiol, 2013. **75**: p. 621-44.

302.    Moskalev, A.A., et al., *The role of DNA damage and repair in aging through the prism of Koch-like criteria.* Ageing Res Rev, 2013. **12**(2): p. 661-84.

303.    Lopez-Otin, C., et al., *The hallmarks of aging.* Cell, 2013. **153**(6): p. 1194-217.

304.    Diderich, K., M. Alanazi, and J.H. Hoeijmakers, *Premature aging and cancer in nucleotide excision repair-disorders.* DNA Repair (Amst), 2011. **10**(7): p. 772-80.

305.    Gurkar, A.U. and L.J. Niedernhofer, *Comparison of mice with accelerated aging caused by distinct mechanisms.* Exp Gerontol, 2015. **68**: p. 43-50.

306.    Barnhoorn, S., et al., *Cell-autonomous progeroid changes in conditional mouse models for repair endonuclease XPG deficiency.* PLoS Genet, 2014. **10**(10): p. e1004686.

307.    Gorbunova, V., et al., *Changes in DNA repair during aging.* Nucleic Acids Res, 2007. **35**(22): p. 7466-74.

308.    Vijg, J., et al., *UV-induced unscheduled DNA synthesis in fibroblasts of aging inbred rats.* Mutat Res, 1985. **146**(2): p. 197-204.

309.    Hart, R.W. and R.B. Setlow, *DNA repair in late-passage human cells.* Mech Ageing Dev, 1976. **5**(1): p. 67-77.

310.    Goukassian, D., et al., *Mechanisms and implications of the age-associated decrease in DNA repair capacity.* FASEB J, 2000. **14**(10): p. 1325-34.

311.    D'Errico, M., et al., *Factors that influence the DNA repair capacity of normal and skin cancer-affected individuals.* Cancer Epidemiol Biomarkers Prev, 1999. **8**(6): p. 553-9.

312.    Cheng, L., et al., *Reduced DNA repair capacity in head and neck cancer patients.* Cancer Epidemiol Biomarkers Prev, 1998. **7**(6): p. 465-8.

313. Wei, Q., et al., *DNA repair and aging in basal cell carcinoma: a molecular epidemiology study.* Proc Natl Acad Sci U S A, 1993. **90**(4): p. 1614-8.

314. Matta, J.L., et al., *DNA repair and nonmelanoma skin cancer in Puerto Rican populations.* J Am Acad Dermatol, 2003. **49**(3): p. 433-9.

315. Grossman, L. and Q. Wei, *DNA repair and epidemiology of basal cell carcinoma.* Clin Chem, 1995. **41**(12 Pt 2): p. 1854-63.

316. Liu, Z., et al., *Reduced DNA double-strand break repair capacity and risk of squamous cell carcinoma of the head and neck-A case-control study.* DNA Repair (Amst), 2016. **40**: p. 18-26.

317. Wei, Q., et al., *Repair of UV light-induced DNA damage and risk of cutaneous malignant melanoma.* J Natl Cancer Inst, 2003. **95**(4): p. 308-15.

318. Khan, S.G., et al., *XPC initiation codon mutation in xeroderma pigmentosum patients with and without neurological symptoms.* DNA Repair (Amst), 2009. **8**(1): p. 114-25.

319. Wang, Q.E., et al., *Nucleotide excision repair factor XPC enhances DNA damage-induced apoptosis by downregulating the antiapoptotic short isoform of caspase-2.* Cancer Res, 2012. **72**(3): p. 666-75.

320. Shaposhnikov, M., et al., *Lifespan and Stress Resistance in Drosophila with Overexpressed DNA Repair Genes.* Sci Rep, 2015. **5**: p. 15299.

321. Johnson, N., et al., *Concentrations of arsenic, chromium, and nickel in toenail samples from Appalachian Kentucky residents.* J Environ Pathol Toxicol Oncol, 2011. **30**(3): p. 213-23.

322. Zhang, C., *Essential functions of iron-requiring proteins in DNA replication, repair and cell cycle control.* Protein Cell, 2014. **5**(10): p. 750-60.

323. McAdam, E., R. Brem, and P. Karran, *Oxidative Stress-Induced Protein Damage Inhibits DNA Repair and Determines Mutation Risk and Therapeutic Efficacy.* Mol Cancer Res, 2016. **14**(7): p. 612-22.

Vita

1. Place of Birth: Owensboro, Ky

2. Educational History

    University of Kentucky – Undergraduate education (2003-2007)

        Bachelor of Science in Biology (2007)

        Bachelor of Arts in Chemistry (2007)

3. Work History

    Academic Tutor, University of Kentucky Center for Academic and Tutorial
    Services (2004-2008)

    Graduate Researcher, University of Kentucky Department of Toxicology and
    Cancer Biology (2007-2016)

4. Scholastic Honors

    **UK**: Jeffrey Fellowship for tobacco related research, (15,000.00 annually), awarded three consecutive years
    (2011-2013)
    **UK:** National Institute for Environmental Safety and Health - T32 Training Grant in Molecular
    Mechanisms of Toxicology (2009-2011)
    **UK:** Dean's List (4 Semesters from 2003-2007)
    **UK:** National Merit Finalist Scholarship (2003)

5. Publications

    Holcomb N, Goswami M Mellon I et al. (2017) Inorganic arsenic inhibits the nucleotide
    excision repair pathway and reduces the expression of XPC (In Press)
    http://dx.doi.org/10.1016/j.dnarep.2017.02.009

    Holcomb N, Goswami M, Han SG, Clark S, Orren DK, Gairola CG Mellon I, et al. (2016)
    Exposure of Human Lung Cells to Tobacco Smoke Condensate Inhibits the Nucleotide
    Excision Repair Pathway. PLoS ONE 11(7): e0158858. doi:10.1371/journal.pone.0158858

    Tian, L., Hou, C., Tian, K., Holcomb, N. C., Gu, L., Li, G. M. (2009) Mismatch recognition
    protein MutSbeta does not hijack (CAG)n hairpin repair in vitro. J Biol Chem 2009 Jul
    31;284(31):20452-6. doi: 10.1074/jbc.C109.014977

6. Nathaniel Holcomb