The Dissertation Committee for Kevin J. Wolfe Certifies that this is the approved version of the following dissertation

Biological and Functional Consequences of Polymorphisms in the *XPD* gene

Committee:

Sherif Z.Abdel-Rahman, Ph.D.

Jonathan B. Ward, Ph.D.

Mary Treinen-Moslen, Ph.D.

Thomas E. Albrecht, Ph.D.

Randa A. El-Zein, MD. Ph.D.

Dean of the Graduate School

Rodney S. Nairn, Ph.D.

Biological and Functional Consequences of Polymorphisms in the *XPD* gene

by

Kevin J. Wolfe, B.S.

Dissertation

Presented to the Faculty of the Graduate School of
The University of Texas Medical Branch
in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

The University of Texas Medical Branch

December, 2006

To my mother and father,
for their never ending support and belief in what I can accomplish

PREFACE

My reason for this body of work is the same reason that education is so important to me. I wanted to challenge my mind and push my abilities. This study represents a continuation of the knowledge needed to achieve a better understanding of susceptibility among populations. One important unknown for population health is understanding how genetic differences predispose individuals to certain undesirable outcomes, such as cancer. This unknown stimulated my research. The goals of my research are to acquire new insight about populations who are susceptible to increased cancer risk because of inherited genetic differences, and to identify mechanisms responsible for their susceptibility. The hope is that this new insight will facilitate adoption of preventive measures that reduce risks due to susceptibilities.

This research has required me to push my abilities further and further, and could not have been accomplished without the help of numerous very important individuals. I thank Dr. Mary Treinen-Moslen for her support and willingness to help me improve my communication skills. I greatly appreciate the help of both Dr. Rodney Nairn and Dr. Randa El-Zein for their guidance in experimental design. I also thank Dr. El-Zein for taking the time to teach me Fluorescent *In Situ* Hybridization. I give my sincerest thanks to the Molecular Genomics and Recombinant DNA Core Facility, specifically Deborah Prusak and Richard Kanost, for help with mRNA quantitation and genetic sequencing, respectively. I express my appreciation to Eugene Knutson, Dr. Michael Schmiederer and Dr. Thomas Albrecht at the Optical Imagining Core Facility, for their help with all the confocal fluorescent microscopy work, data analysis and general experimental design of that specific study. I am grateful to both Dr. Michael Schmiederer and Dr. Jeffrey Wickliffe for their help with motivation, ideas and experimental design and I thank Ms. Mirtha S. Lopez for carefully scoring chromosome aberrations using classical cytogenetics. I express my gratitude to Dr. Marinel Ammenheuser for working tirelessly with me to help improve my writing skills. I appreciate Courtney Hill, Carla Kinslow and Adele Guerin for their motivation, support and help in lab. Last, but not least, I thank Dr. Sherif Abdel-Rahman for believing in me and giving me the opportunity to study under his tutelage.

Biological and Functional Consequences of Polymorphism in the *XPD* gene

Publication No. _____

Kevin James Wolfe, Ph.D.

The University of Texas Medical Branch at Galveston, September 2006

Supervisor: Sherif Abdel-Rahman

Epidemiological studies have documented many associations between single nucleotide polymorphisms (SNPs) in the nucleotide excision repair gene *XPD* (*ERCC2*) and cancer risk. Little is known, however, about the underlying mechanisms for these associations. I used lymphocytes from healthy subjects to explore novel mechanisms which could explain the reported risk-modifying effects on disease susceptibility of three SNPs in the XPD gene, the synonymous C156A SNP in exon 6 and the nonsynonymous SNPs Asp312Asn in exon 10 and Lys751Gln in exon 23. Baseline and NNK-induced chromosomal aberrations (Ca) were assessed by cytogenetic analysis and then P values were calculated as estimates of sub-population differences for increased frequencies of CAs associated with the three *XPD* polymorphisms. I found significant elevation in baseline frequencies of CAs among smokers with the variant 312Asn polymorphism (P=0.028). Elevations in NNK-induced aberrations were found among younger subjects (age<39 years) with the 156A or the 751Gln polymorphism, (P=0.025 and P=0.037, respectively) and in females compared to males with a combination of the 312Asn and the 751Gln polymorphisms (P=0.045). Application of real time PCR showed that each SNP, alone and in combination, significantly decreased constitutive *XPD* mRNA levels (P<0.003) in lymphocytes. Decreases in XPD mRNA levels were significantly higher in older subjects and in smokers. Localized Mfold structure analysis of the mRNA sequence surrounding the studied SNPs were predicted to alter mRNA secondary structure, which indicated that these SNPs potentially affect local folding and mRNA stability. UVC treatment of cells with wild type XPD produced a significant increase (P=0.03) in XPD protein levels by 30 min, which surprisingly coincided with a decrease in *XPD* mRNA transcript copy number (P=0.0002). Fluorescent confocal microscopy demonstrated that this increase in XPD protein appeared largely due to an increase in nuclear localization of XPD, which was evident at 30 min and persistent at 6 hrs. New observations from this project provide possible mechanistic explanations for the association of polymorphisms in XPD with increased genetic damage (CAs) and cancer risk.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

INTRODUCTION

DNA repair is a process that is critical for the survival of cells and organisms. The mechanism of DNA repair takes on many forms, and the type and extent of genetic damage that has occurred largely dictates which repair mechanism(s) or pathway(s) will be used for the repair of that damage. In Yu *et al*. (1999), DNA repair is defined in a general sense as, "a range of cellular responses associated with restoration of the genetic instructions as provided by the normal primary DNA sequence". Research in the field of DNA repair has been pursued for many decades. However, human DNA repair has been most extensively studied through research devoted to understanding the cause of many human diseases associated with defects in DNA repair (Yu *et al*., 1999). One of the most studied of these human diseases, which has led to numerous breakthroughs in the understanding of DNA repair, is Xeroderma pigmentosum (XP). XP was the first human DNA repair deficiency disease identified (Jung and Bantle, 1971). Two other human diseases closely related to XP that have been studied are Trichothiodystrophy (TTD) (Rebora and Crovato, 1987) and Cockayne syndrome (Poon *et al*., 1975).

People are constantly under threat from agents in their environment that can potentially damage their genomes. DNA damaging agents, however, are not merely external. Our own bodies produce byproducts from metabolism and other cellular processes that can also cause damage to our genomes. Internal agents include oxygen and nitrogen free radicals, and external sources of damage include ionizing radiation, UV radiation, and a diverse number of chemicals. For these reasons mammalian cells have a number of mechanisms to combat DNA damaging agents and maintain genome stability. Some of these mechanisms include, (1) apoptosis to rid the body of extensively damaged cells, (2) initiation of cell cycle check points to halt the progress of DNA synthesis and replication until damage can be repaired, (3) transcriptional changes, to respond to the cell's damaged state by manufacturing the necessary proteins to insure repair and survival, and (4) initiation of the appropriate DNA repair mechanism for the removal of DNA damage. These processes do not act individually, but collectively, and are responsible for the fate of the cell (Madhusudan and Middleton, 2005).

There are a number of different DNA repair pathways, and these pathways have overlapping responsibility for the repair of DNA damage. Therefore, an understanding of all the mechanisms of DNA repair is crucial. For example all DNA repair mechanisms (direct reversal repair, mismatch repair, homologous recombination, non-homologous end joining, base excision repair and nucleotide excision repair) play a role in repairing the damage from alkylating agents or secondary damage from compounds such as tobacco-specific nitrosamines (Drablos *et al*., 2004). In the following sections each of the DNA repair pathways will be discussed with regard to its mechanism of action, related diseases, and issues of genetic susceptibility associated with single nucleotide polymorphisms in DNA repair genes. The last section is devoted to the nucleotide excision repair pathway and to the *XPD* gene, which is the subject of the current dissertation project.

The focus of this dissertation is on understanding the functional and biological significance of single nucleotide polymorphisms (SNPs). SNPs are changes in the genetic sequence of a gene that occur in the general population at a prevalence greater then 1%. This dissertation focuses on three of the major SNPs that occur in the coding region of the *XPD* gene. As will be discussed in the last section of this introduction, the XPD protein is involved in the nucleotide excision repair pathway, which is important in DNA repair of large bulky adducts. The central hypothesis is that polymorphisms in the *XPD* gene result in structural or functional changes that could lead to altered protein function and, therefore, altered nucleotide excision repair (NER) capacity.

DIRECT REVERSAL DNA REPAIR.

Direct reversal DNA repair (DRR) in humans is largely centered on a single alklytransferase, O^6-methylguanine-DNA-methyl transferase, (MGMT). This repair pathway utilizing MGMT is the main pathway for the repair of damage induced by alkylating agents. Such exogenous compounds, through reacting with DNA, can add alkyl groups, such as methyl and ethyl groups, to the bases in DNA (Figure 1). Simple alkylating agents can also transfer such unmodified groups as methyl and ethyl groups to

nucleophilic sites on proteins and RNA (Margison and Santibanez-Koref, 2002). These alkylating agents can cause mutations, carcinogenesis, cell death, cell cycle arrest, and teratogenicity. One of the most common adducts formed from alkylating agents is O^6-methylguanine (O^6-meG). This adduct has the potential to mismatch with adenine or thymine, resulting in a GC-AT transition which is implicated in cellular oncogenesis (Margison and Santibanez-Koref, 2002).

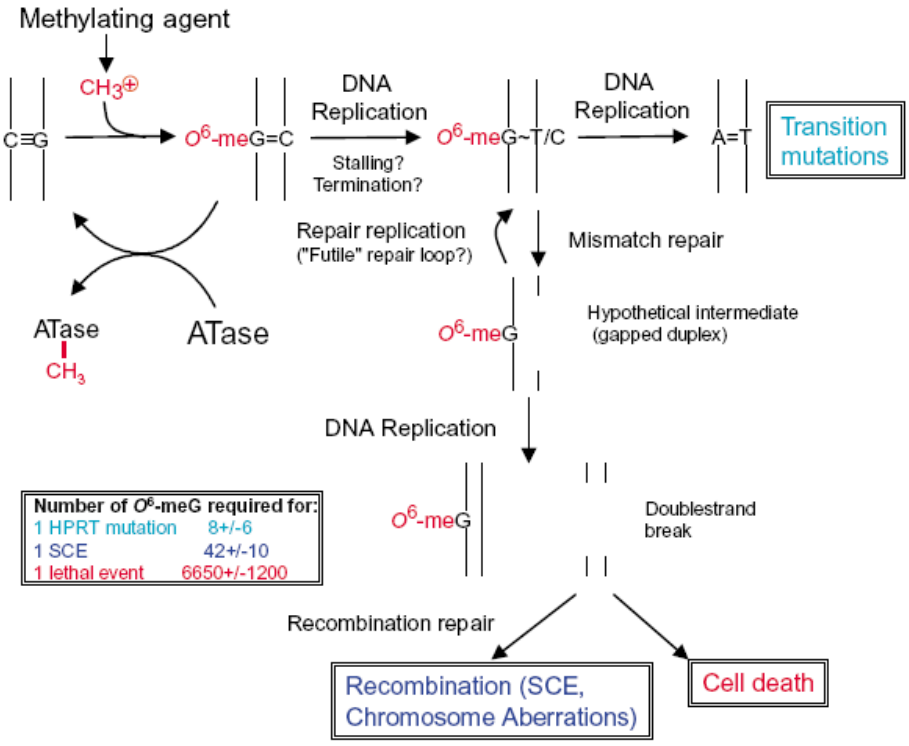


Figure 1. Mechanisms of DNA damage

Mechanism by which DNA damage is caused by alkylating agents, and the resulting genotoxic effects that may occur, such as GC→AT transitions and, DNA double strand breaks. These may or may not be repaired later by the mismatch repair mechanism or the double strand break repair mechanism, respectively. The left box shows the number of alkylating event required for genomic damage to occur as estimated from studies in Chinese hamster ovary cells. Reprinted from BioEssays, vol 24(3), Margison GP., Santibanez-Koref MF., O6-alkylguanine-DNA alkyltransferase: role in carcinogenesis and chemotherapy, 255-266, copyright 2002, with permission from Wiley-Liss INC.

Mechanisms of direct reversal repair

A direct reversal repair system, such as that involving MGMT, was first theorized when it was observed that O^6-meG adducts were removed faster than cell turnover following treatment of rats with the hepatocarcinogen N,N-dimethylnitrosamine (O'Connor *et al*., 1973). Since that time, more extensive research has been conducted on the mechanism of action and importance of direct reversal on genome stability. MGMT is a protein that belongs to a class of enzymes called alkyltransferases. Alkyltransferases, such as MGMT, are important in the repair of alkylation damage in an irreversible reaction. The alkyl group from the damaged site is directly transferred to a cysteine residue in the active site of MGMT (Hazra *et al*., 1997). Once MGMT is used to transfer the alkyl group, it is inactivated and degraded via the ubiquitin proteolytic pathway (Srivenugopal *et al*., 1996). Therefore, any further repair of DNA-alkyl adducts would require the continued production of active MGMT. Alkyltransferases bind to DNA through a helix-turn-helix tertiary structure in the C-terminal. The second helix of this structure houses an arginine that acts like a finger to flip out the alkylated base into the active site of the protein. A histidine within the conserved active site and a proximal glutamic acid residue form a network of hydrogen bonds, with the active site cysteine thereby effectively transferring the alkyl group to the cysteine of MGMT. This reaction results in an alteration of the DNA binding domain upon which the newly alkylated MGMT is released from the DNA and targeted for degradation by ubiquitination (Daniels and Tainer, 2000). The cost of this repair pathway is great, when the production of this 22 kDa protein and the high energy consuming degradation pathway is taken into account. Although other less energy expensive means exists to repair alkylation damage, they are, however, slower (Yu *et al*., 1999). Along with O^6-methylguanine, MGMT also repairs O^6-ethylguanine, O^6-butylguanine and O^4-methylthymine but at a lower efficiency (Sancar, 1995). Recently, it was also shown that MGMT repairs other bulky adducts, such as pyridyloxobutyl adducts which are produced from tobacco specific nitrosomines such as 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) (Wang *et al*., 1997; Mijal *et al*., 2004)

There are enzymes other then MGMT that repair DNA using a direct reversal mechanism, although most are not as active in the mammalian system. AlkB, a prokaryotic enzyme, was shown to repair 1-methyladenine and 3-methylcytosine with a direct reversal mechanism using molecular oxygen, 2-oxogulerate and Fe(II) to oxidize the methyl group resulting in direct repair and release of formaldehyde (Figure 2). Interestingly, eight human AlkB homologues (hABH) have recently been found. Similar to AlkB, hABH2 and hABH3 can also repair 1-methyladenine and 3-methylcytosine in a direct reversal 2-oxogluterate-dependent manner (Drablos *et al*., 2004), hABH2 and hABH3 also function in the repair of 1-ethyladenine DNA adducts, resulting in the release of acetaldehyde (Duncan *et al*., 2002).

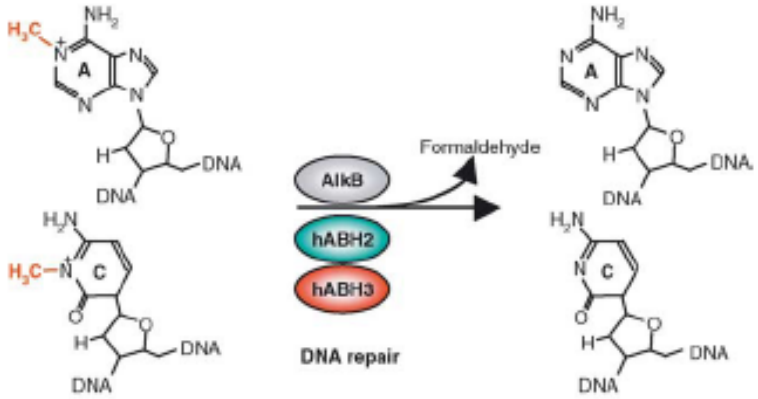


Figure 2. DNA repair by ALKB and hABH
The E. coli ALKB and human homologous hABH2 and hABH3 repair O6-methyladenine and O6-methylcytosine in a direct reversal mechanism that results in the release of formaldehyde. Reprinted from DNA Repair vol 3, Drablos F., Feyzi E., Aas PA., Vaagbo CB., Kavli B., Bratlie MS., Pena-Diaz J., Otterlei M., Slupphaug G., Krokan HE. Alkylation damage in DNA and rna-repair mechanisms and medical significance. 1389-1407, copyright 2004 with permission from Elsevier.

Polymorphisms and cancer susceptibility in direct reversal repair genes

Two important polymorphisms that have been discovered in the *MGMT* gene have been shown to be associated with genetic damage. An *in vitro* study from our laboratory showed that polymorphisms in *MGMT*, specifically Leu48Phe and Ile143Val may increase the risk of chromosomal aberrations due to exposure to tobacco-specific nitrosamine carcinogens like 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK)

(Hill *et al*., 2005a). Although the mechanism by which these polymorphisms lead to increased genetic damage is not well understood, it has been shown that loss or lowered expression of *MGMT* is linked to increased cancer risk and heightened sensitivity to methylating agents (Gerson, 2004). With this knowledge, one could hypothesize that the possible consequences of these polymorphisms are either lowered enzymatic activity level, through misfolding of the active site of the MGMT protein, or possible decreased protein stability, resulting in less active protein available for damage repair. A study of oral cancer prognosis showed that *MGMT* gene expression was significantly decreased in groups of patients with non-epithelial dysplasis, epithelial dysplasis and oral squamous cell carcinoma. This study also showed a significant loss of *MGMT* gene expression in smokers with oral squamous cell carcinoma (Rodriguez *et al*., 2006).

MISMATCH REPAIR

Mismatch repair (MMR) is a post-replicative DNA repair pathway and, as its name implies, repairs mismatched bases, those not conforming to the A:T, G:C base pairing of DNA. This repair pathway also repairs small insertion/deletion mutations (Scherer *et al*., 2005). Mismatched base pairs can arise through a number of mechanisms, such as misincorporations during DNA replication, formation of hetroduplexes (Bishop *et al*., 1985), and deamination of 5-methylcytosine to uracil resulting in an G:T mismatch. (Shen *et al*., 1994; Yu *et al*., 1999). During DNA synthesis it has been shown that DNA polymerase has the tendency to incorporate an adenine where it doesn't belong, resulting in an A/G mismatch (Kutchta *et al*., 1988; Eger and Benkovic, 1992). Perhaps it is because of this, and the mutagenic effect of the resulting G:C→A:T transition, that mismatch repair glycosylases, during the repair of either an A:G or T:G mismatch, have evolved to remove the adenine and thymine keeping the guanine intact (Wiebauer and Jiricny, 1989; Yeh *et al*., 1991).

Mechanisms of mismatch repair.

MMR involves two distinct mechanisms: long patch and short patch mismatch repair. Long patch repair removes a section of one of the DNA strands from an incision point 90 to 170 nucleotides beyond the mismatch (Fang and Modrich, 1993). In short patch repair, only one or a few nucleotides are excised. Three known enzymes are involved in this repair pathway (Wiebauer and Jiricny, 1989; Yeh *et al*., 1991). Each of these enzymes has a specific cutting pattern for MMR. The T/G specific enzyme, G/T-specific thymine DNA gylcosylase, removes the thymine (Nedderman *et al*., 1996) which creates an AP site opposite the G (Wiebauer and Jiricny, 1989). The A/G specific enzyme cuts at both the 5' and 3' phosphodiester bond flanking the mismatch pair. On the other hand the third enzyme, also known as the "all type enzyme", only cuts the phosphodiester bond 5' to the mismatch. (Nedderman *et al*., 1996). Following the excision step by any of these enzymes, the resulting gap is filled by polymerase β (Yu *et al*., 1999).

During long patch repair, MMR is initiated by the recognition of mispaired bases by MutS proteins followed by the recruitment of MutL proteins involved in base removal and resynthesis (Scherer *et al*., 2005). In humans there are three different MutS homologs: MSH2, MSH3, and MSH6. These proteins function as heterodimer complexes; MSH2-MSH6 are responsible for initiating repair of mispaired bases, and MSH2-MSH3 are responsible for initiating repair of larger insertion/deletion mispairs of 2 to 4 bases (Marsischly *et al*., 1996; Genschel *et al*., 1998; Edelmann *et al*., 2000). Humans also have four MutL homologs: MLH1, PMS1, PMS2, and MLH3. These, like the MutS homologs, also form heterodimeric complexes of MLH1-MLH3, MLH1-PMS1, and MLH1-PMS2. The primary response in MMR is for the MLH1-PMS2 complex to interact with the MSH2-MSH6 or MSH2-MSH3 complex and, therefore, drive further MMR steps (Prolla *et al*., 1994; Li and Modrich, 1995; Wang *et al*., 1999).

Unique among the DNA repair pathways, MMR does not repair the damage in modified bases. Therefore, the difficult task in repairing DNA mismatches is determining which of the paired bases is the incorrect one that needs to be replaced. During MMR the base to be replaced is always on the newly synthesized strand. During long patch repair, MMR determines the newly synthesized strand by searching for the presence of gaps between Okazaki fragments on the lagging strand or by a free 3'-terminus on the leading strand (Stojic *et al*., 2004). During the process of MMR, mismatched bases must first be recognized by the human MSH2/MSH6 (hMSH2/hMSH6) heterodimer (Scherer *et al*., 2005). An ATP dependent conformational change transforms the hMSH2/hMSH6 heterodimer into a sliding clamp structure, which has the capability to translocate along the DNA strand (Blackwell *et al*., 1998; Gradia *et al*., 1999). Another heterodimer complex, hMLH1/hPMS2, is recruited to the mismatch site prior to clamp migration, in another ATP dependent reaction (Li and Modrich, 1995). This complex, hMSH2/hMSH6/ hMLH1/hPMS2, is now capable of traversing the DNA in the 5'

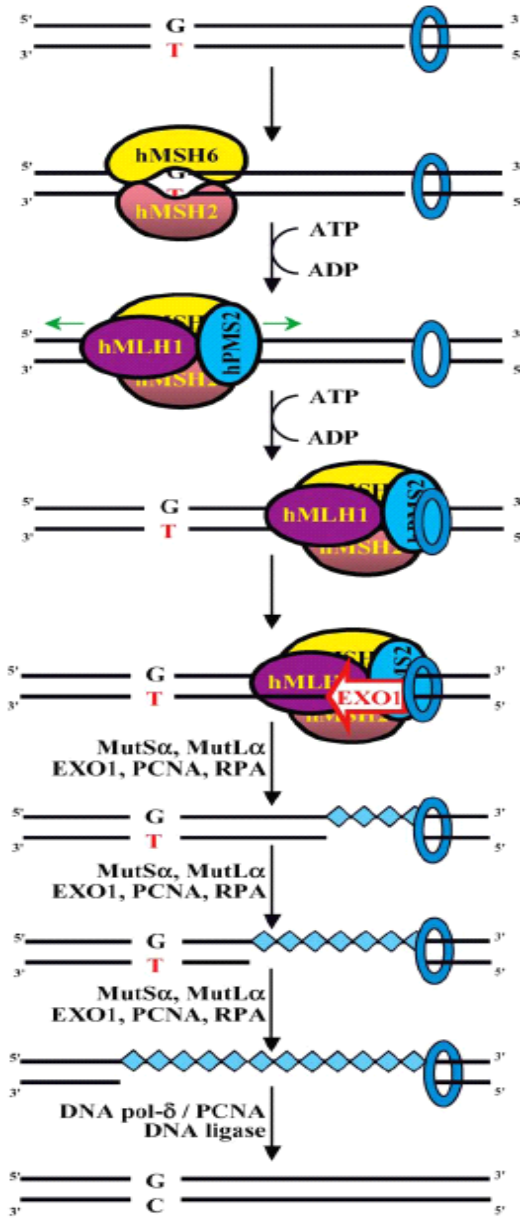


Figure 3. Long patch MMR
Schematic representation of the protein complexes and mechanisms involved in long patch MMR. Reprinted from DNA Repair, vol. 3, Stojic L., Brun R., Jiricny J. Mismatch repair and DNA damage signaling, 1091-1101, copyright 2004, with permission from Elsevier

direction in search of markers that identify the newly synthesized strand. Once this strand discontinuity has been identified, the clamp complex recruits exonuclease 1 (EXO1) 5' to the mismatch on the new strand (Stojic *et al*., 2004)). EXO1 has 5'→3' exonuclease activity and begins to degrade that newly synthesized strand at the point of strand continuity moving in the 3' direction towards the mismatch site. EXO1 requires constant stimulation from newly organized clamp complexes until it has reached the mismatch site. At this point no more clamp complex can be organized, and the stimulation of EXO1 ceases. During strand degradation, replication protein A is recruited to help stabilized the single stranded DNA (Genschel and Modrich, 2003). At this point, DNA polymerase and PCNA can resynthesize the DNA strand, repairing the mismatch and leaving DNA ligase to close the final nick following synthesis (Stojic *et al*., 2004) (Figure 3).

Mutations, polymorphisms, and cancer susceptibility in MMR genes

Components of the MMR pathways not only play a role in DNA repair but have also been shown to be sensors for DNA damage that may regulate cell cycle progression and apoptosis (Bellacosa, 2001; Brown *et al*., 2003). Because of its role in both DNA repair and cell cycle progression, cells defective in MMR show many signs of genome instability such as slippage mutations, chromosomal rearrangements, gene amplification, radio resistant DNA synthesis, and mutation in DNA repeat regions, which is known as microsatellite instability (MSI) (Surtees *et al*., 2004). MSI has been associated with increased cancer risk. The current mechanistic hypothesis is that the increased level of mutations that result in MSI is a genome-wide increase in mutation rate that, not only affects microsatellite regions, but also affects exonic coding regions of genes that play a role in cancer suppression (Chao and Lipkin, 2006). Hofseth *et al*. (2003) have shown that over-expression of the base excision repair gene *APE1* causes an increase in MSI. In addition, Chang *et al*. (2005) have shown that the over-expression of *APE1*, which inhibits MMR activity, results in increased MSI. Specifically, high levels of APE1

decrease protein levels of hMSH6, a key member of the MMR pathway that forms the hMUTSα complex (Chang *et al*., 2005).

Mutations in MMR genes have been identified in a large number of hereditary non-polyposis colon cancer (HNPCC) patients who are heterozygous mutant for at least one of the MMR genes (Lucci-Cordisco, 2003). As many as 400 mutations in MMR genes may exist (Peltomaki, 1997). Identification of these mutations may be especially important in cancer research and understanding cancer susceptibility. HNPCC is an autosomal dominant cancer characterized by early onset of colorectal cancer, and it is estimated to be responsible for 5% of all colorectal cancers. Some patients may also develop stomach, intestinal, ovarian, and endometrial cancers (Lynch, 1996). The main sites of mutations are in the *MSH1* and *MSH2* genes, resulting in microsatellite instability, and reduced MMR activity (Aaltonen, 1993; Yuan *et al*., 2002). These two genes account for most of the mutations identified in MMR. The remaining mutations are accounted for by mutations in the *MSH6*, *PMS1*, and *PMS2* genes (Peltomaki, 2001; Peltomaki, 2003). Recent studies have also shown that each year, in the United States approximately 17,000 sporadic colorectal cancer cases (Gryfe *et al*., 2000; Gryfe and Gallinger, 2001), and approximately 10,000 endometrial cancer cases (Lynch, 1999) are attributed to MMR somatic mutations.

It is not only the case that these mutational events alter MMR function, but polymorphisms in some MMR genes have also been associated with modifying MMR activity and increasing susceptibility to cancer, such as breast cancer. Three polymorphisms in the MMR gene *PMS2* (*i.e.* Pro511Lys, Thr597Ser, and Met622Ile), have been shown to decrease the protein-protein interaction between PMS2 and MLH1. Although these polymorphisms do not lie within the MLH1 binding domain of PMS2, they do result in altering the protein structure to such a degree as to interrupt normal protein-protein interactions (Yuan *et al*., 2002). The Gly322Asp polymorphism in the MMR gene *hMSH2* may be a modifying factor in breast cancer occurrence. Poplawisk *et al*. (2005) recently found that the Gly allele showed a strong association with increased risk of breast cancer. To demonstrate the importance of proper mismatch repair,

Campbell *et al*., (2006) reported that cells in an *in vitro* system that are deficient in MMR, specifically *Msh2*-/- cells, showed a significant increase in chromosomal abnormalities, mainly resulting in chromosomal aneuploidy. This same group showed that this was due to an increase in the number of centrosomes formed during mitosis, resulting in multiple spindle formation which causes mis-segregation of chromosomes and, therefore, chromosomal aneuploidy (Campell *et al*., 2006).

DOUBLE STRAND BREAK REPAIR

Double strand break repair (DSB) can arise from a number of circumstances, such as somatic recombination, overlapping excision repair (Lieber, 1991), ionizing radiation, alkylating agents and oxidative stress (Dizdaroglu, 1992; Demple and Harand, 1994). Due to the constant insult on the genome, it is estimated that double strand breaks occur at a rate of about 10 breaks per cell per day (Shin *et al*., 2004). With this number of breaks accumulating in DNA of the human body during any given day, it is easy to understand why DNA repair is so vital to survival. Two main pathways of double strand break repair exist, homologous recombination (HR) and non-homologous end joining (NHEJ). These will be discussed in greater detail below.

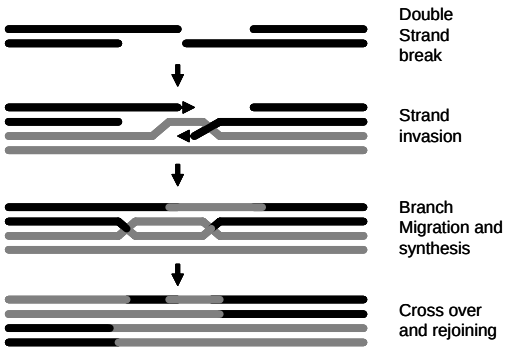


Figure 4. Double strand break repair by homologous recombination
Double strand break (black) is followed by strand invasion. The lost sequence is then resynthesized using the invaded strand as a template. Finally the homologous strand undergoes cross over and rejoining. Modified from Cytognet Genome Res, vol 104, Griffin CS., Thacker J. The role of homologous recombination repair in the formation of chromosome aberrations, 21-27, copyright 2004, with permission from S. Karger AG

Mechanisms of double strand break repair

As mentioned above, double strand break repair consists of two main pathways, HR and NHEJ. Under the umbrella of HR there are two distinct pathways, a more conservative error-free pathways that is Rad51 dependent (Figure 4) and a relatively non-conservative pathways that does not require Rad51 (Figure 5). The conservative pathway, also referred to as gene conversion, incorporates a homology search and strand invasion to repair DSB by copying an undamaged sequence. This repair pathway forms the well-known Holliday junctions leading to HR. That is the combining of the resected DNA ends with the homologous double-strand DNA template. Random cutting from nucleases involved in this mechanism may result in an equal number of crossover and non-crossover events, potentially resulting in DNA-sequence exchanges (Griffin and Thacker 2004). However, mammalian cells have been shown to have a negligible frequency of crossover events (Richardson *et al*., 1998).

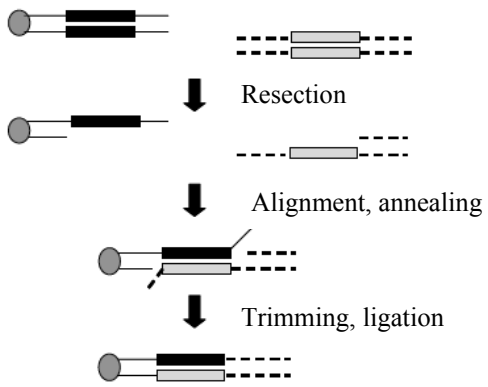


Figure 5. Double strand break repair by single strand annealing

For double strand breaks that occur within or near repeat sequences (black and gray boxes), hetrologous DNA strands can be used for direct annealing and ligation. Reprinted from Cytognet Genome Res, vol. 104, Griffin CS., Thacker J., The role of homologous recombination repair in the formation of chromosome aberrations, 21-27, copyright 2004, with permission from S. Karger AG.

The non-conservative pathway is also known as single strand annealing (SSA). This pathway requires the direct annealing of repeat sequences adjacent to the double strand break. Since the broken strands are directly annealed, following a double strand break some microdeletions may arise (Griffin and Thacker 2004). Therefore, SSA involves a homology search but no strand invasion. Rad52 plays a role in promoting the annealing of DNA strands central to the process of SSA (Pfeiffer *et al*., 2004). Sugawara *et al*. (2000) have shown that the repair efficiency is directly related to the length of the

homologous sequence. They have determined that the optimum sequence length is approximately 200 base pairs.

Rad51-dependent HR involves three steps or phases. Phase I is referred to as pre-synapsis, and involves the preparation of DNA ends for proficient recombination, specifically the creation of a 3' single strand over-hang. This phase requires the Mre11 complex (Mre11, Rad50, and NBS1) for DNA end processing. Also required is RPA, a single stranded binding protein; Rad52 and Rad51, for recombinase loading; and Rad54, for the formation of a recombinase nucelo-protein filament and for stabilization. Phase II is synapsis; this phase is characterized by the formation of holiday junctions. Phase II requires Rad51 and Rad54 for strand invasions and holiday junction formation. Phase III, post-synapsis and resolution, involves the synthesis of newly repaired DNA and separation of the recombined DNA molecules. Phase III involves the proteins Rad54 for branch migration, and Mus8/EmwI and Rad51C/XRCC3 for resolution of combined DNA strands (Wyman *et al*., 2004). In Rad51-dependent HR, a number of Rad51 monomers bind to the 3'-end of each single strand generated by a DSB, to form a nucleoprotein responsible for homologous duplex invasion (Eggler *et al*., 2002). For this reason the catalytic activity of Rad51 is essential for initiating strand exchange (Pfeiffer *et al*., 2004). Brca2 is another protein involved in HR and DSB repair. This protein forms an inhibitory complex with Rad51, preventing Rad51 from binding to the DNA forming the nucleoprotein complex. In this way Brca2 regulates the availability of Rad51 and functions to prevent HR during periods of normal cell metabolism (Davies *et al*., 2001).

During Rad51-dependent HR in E. coli, the RecBCD complex is responsible for generating the 3' tail necessary for strand invasion and replication, by providing a priming site (Shin *et al*., 2004). However, there is no eukaryotic homologue for this complex. In eukaryotes, as depicted in Figure 6, Mre11, which contains N-terminal phosphoesterase motifs, as well as two DNA binding motifs, has a DNA 3'→5' exonuclease, a flap endonuclease, and hair-pin nuclease activities (Hopfner *et al*., 2001; D'Amours *et al*., 2002). Rad50, which functions along side Mre11, contains an ABC

ATPase with a coiled coil structure (Hopfner *et al*., 2000). Upon the binding of ATP, the two Rad50 ATPase ends form a homodimer, effectively separating the DNA strands cut by Mre11 (Hopfner *et al*., 2000). These two proteins, working in conjunction, allow for the cutting and separation of the damaged DNA strand for invasion (Figure 6). A common essential complex required for HR, SSA, NHEJ, telomere maintenance, meiotic DSB repair, G2/M cell cycle regulation, and chromatin structure is the Rad50-Mre11-Nbs1p complex (Haber *et al*., 1998; Kuschel *et al*., 2002). The role of Mre11 in this complex is as a nuclease involved in the generation of long DNA single strands from double-strand breaks which is essential for HR and SSA (Pfeiffer *et al*., 2004)

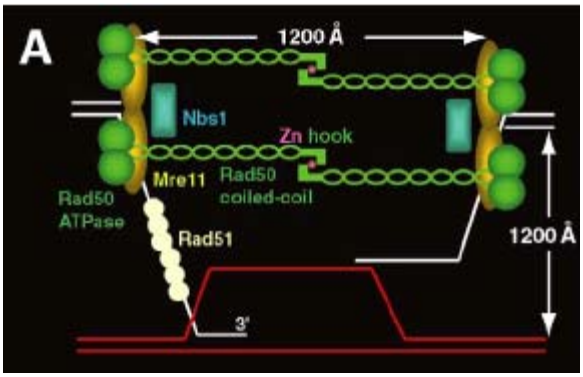


Figure 6. DNA strand separation during double strand break repair
One homodimer of Mre11 on both ends attracts two Rad 50 proteins, effectively separating the DNA ends twice the length of the coiled coil region of Rad 50 (1200Å). Following separation Rad 51 initiates strand invasion. Reprinted from DNA Repair, Vol. 3, Shin DS., Chahwan C., Huffman JL., Tainer JA, Structure and function of the double-strand break repair machinery, 863-873, copyright 2004, with permission from Elsevier

Unlike Rad51-dependent HR, NHEJ is able to rejoin broken DNA ends directly. This can be accomplished in two ways. Direct re-ligation of complementary DNA strands results in the restoration of the original sequence. As an alternative option, if two non-complementary ends are used they must first be processed by a polymerase and nuclease to create ligatable ends. However, the use of non-complementary ends leads to small alterations on the order of a few bases and may result in mutations and chromosomal alterations (Pfeiffer *et al*., 2004). During NHEJ repair of double strand breaks a single Ku70/80 hetrodimer must bind to both free ends of the broken DNA strand. This occurs because Ku70/80 recognizes blunt or short overhang DNA ends (Critchlow and Jackson, 1998). This complex, plus the addition of the catalytic subunit of DNA dependent protein kinase (DNA-PKcs) that is activated by Ku70/80, forms a bridging structure between the two DNA ends (Smith and Jackson, 1999). The addition

of DNA-PKcs recruits the Artemis protein, and endo- and exo-nucleases, which in return primes the damage site and recruits the heterodimer of XRCC4/DNA ligase IV to complete the repair process (Grawunder *et al*., 1997). Essential proteins for NHEJ include XRCC4, XRCC7, and LIG4 (Jeggo *et al*., 1998). Gene products XRCC5, 6, and 7, along with the hetrodimer Ku70/Ku80, form the DNA-dependent protein kinase (DNA-PK) essential for NHEJ (Featherston and Jackson, 1999a, b) The enzymatic subunit of DNA-PK (DNA-PKcs) functions in signal transduction following DNA damage and recruitment of further repair proteins to the DSB site (Jackson, 2002).

It has been shown that HR is predominantly active during specific cell cycle stages. Cox (2001) has shown that this cell cycle specific activity is largely during the S and G2 phase. This is because during these cell cycle phases the sister chromatid is available as a template. This does not mean, however, that HR is not active during other cell cycle phases. For example, during G1 when no sister chromatid is available as a template, the homologous chromosome or sequence repeats on heterologous chromosomes, which make up approximately 50% of the human genome, can be utilized as a template (Griffin and Thacker,. 2004). On the other hand, NHEJ is the predominant form of DSB repair during cell cycle phase G0/G1, and requires the heterodimer Ku70/80, the DNA-dependent protein kinase catalytic subunit (DNA-PKcs), and DNA ligase IV/XRCC4 (Takata *et al*., 1998; Essers *et al*., 2000; Rothkamm *et al*., 2003)

A study by Iliakis *et al*. (2004) that examined DSB repair, following infrared radiation in DNA-PK deficient and wild-type cells, found that, although repair was slower in DNA-PK deficient cells, after 24 hr the total repair reached levels close to the wild-type cells, suggesting close to complete repair. This fast component of DSB repair is both a DNA-PK and DNA ligase IV dependent mechanism of NHEJ. Based on kinetic studies, approximately 78% of DSBs induced by IR are repaired through this DNA-PK-dependent fast NHEJ pathway, and only about 22% are repaired through the slow mechanism of end joining, suggested to be a backup mechanism of NHEJ (Iliakis, 2004). It is hypothesized that homologous-dependent recombination (Rad51-dependent) is used

in higher eukaryotes only to repair critical DNA sequences at a later time after NHEJ has taken care of the majority of double strand breaks (Iliakis, 2004).

Polymorphisms and cancer susceptibility in DSB repair genes

DNA repair is a vital part of maintaining genome stability. Not only do inherited mutations in important DNA repair genes result in specific diseases such as atexia telangiectasia, atexia telangiectasia-like disorder, Nijmegen breakage disorder, and ligase IV syndrome, but some single nucleotide polymorphism in these genes may also result in cancer susceptibility. In the pathway of HR, polymorphisms in the *XRCC3* gene have been associated with breast cancer risk (Kuschel *et al*., 2002; Benhamou *et al*., 2004). The *XRCC3* IVS5 A17893G heterozygous and homozygous G allele show a significant protective effect over the A allele. However, another polymorphism, *XRCC3* Thr241Met, is associated with a significant increase in breast cancer incidence with inheritance of the homozygous 241Met allele (Kuschel *et al*., 2002). In a study of 250 oral cancer cases, the *XRCC3* Thr241Met polymorphism has been associated with a reduced risk of supraglottic cancer. In this same study the *XRCC2* Arg188His polymorphism has been associated with an increased risk of pharyngeal cancer (Benhamou *et al*., 2004), and the inheritance of the heterozygous 188His polymorphism presents a nearly 3-fold increase in the risk of non-small cell lung cancer (Zienolddinh *et al*., 2006). Another study showed that the *XRCC2* Arg188His polymorphism may have a dose response protective association with risk of epithelial ovarian cancer, in that individuals heterozygous for 188His had a 20% risk reduction, and individuals homozygous for 188His had a greater than 50% risk reduction (Auranen *et al*., 2005).

Other genes in the NHEJ repair pathway have also been implicated in cancer susceptibility. For example, Roddam *et al*. (2002) has shown from a case-control study, that the presence of the heterozygous genotype (CT) of either *LIG4* Ala3Val or Thr9Ile polymorphisms confers a 1.5-fold reduced risk of multiple myeloma. Inheriting the homozygous variant genotype (TT) confers a four-fold risk reduction. Another polymorphism in *LIG4*, Ile658Val, has also been associated with cancer risk. Sakiyama

et al. (2005) have shown this polymorphism to be significantly protective against squamous cell carcinoma of the lung. Both the *Ku70* C61G and *XRCC4* T1394G intron polymorphisms showed significant increases in the risk of breast cancer incidence (Fu *et al*., 2003). When polymorphic genes were examined as a complex, more significant odds ratios were observed. The DSB end-binding complex, consisting of Ku70, Ku80 and DNA-PKcs, showed a significant increase in breast cancer prevalence with two or more polymorphic variants. The same was observed when NHEJ genes (Ku70, Ku80, DNA-PKcs, LIG4, and XRCC4) were examined together (Fu *et al*., 2003). Another polymorphic gene involved in NHEJ that has also been associated with cancer is *ATR*. The Thr211Met polymorphism in this gene has been associated with a significant decrease in risk of non-small cell lung cancer (Zienolddiny *et al*., 2006).

The key defining gene in HR, *Rad51*, is a suspected modifier of breast cancer risk. The *Rad51* G135C 5'-UTR polymorphism, when assessed alone, does not show an association with increased breast cancer risk. However, the *Rad51* 135C allele does modulate the effect in *BRCA2* carriers. Kadouri *et al*. (2004) has shown that among Ashkenazi Jews carrying the *BRCA2* mutation the inheritance of the *Rad51* 135C polymorphism increases the risk of breast cancer development two-fold and lowers the age of onset significantly by 7 years. It has also been shown that the *Rad51* 135C polymorphism may modify the risk of breast cancer associated with the presence of the *BRAC2* Met1915Thr polymorphism (Sliwinski *et al*., 2005).

BASE EXCISION REPAIR

Base excision repair (BER) is a multi-protein, multi-step pathway responsible for the repair of so called non-bulky adducts, such as those formed from methylation, oxidation, reduction, or the results of ionizing radiation or oxidative stress (Yu *et al*., 1999; Tuteja and Tuteja, 2001). Consequently, with its role in repair of methylation adducts it overlaps to some degree with the repair mechanism of MGMT. The basic steps involved in BER include: (1) the release of the damaged base from DNA by a DNA glycosylase, (2) following the removal of the damaged base, AP lyase incises 3' to the

damage site while AP hydrolase incises 5' of the apurinic site (formerly the damaged site), (3) DNA polymerase then fills in the gap due to the removal of the damaged base beginning at the exposed 3'OH end, (4) The final step involves the sealing of the new base within the DNA structure by DNA ligase.

Mechanisms of base excision repair.

The BER pathway can be divided into four major steps: (1) base removal by a DNA glycosylase; (2) creation of the 3'-OH end; (3) nucleotide synthesis to fill in the gap; and (4) ligation to seal the DNA strands (Figure 7, pathway II). The first step is identifying and removing damaged bases by DNA glycosylases. Because of the varying substrates that the BER pathway is responsible for repairing, it is no surprise that there are many DNA glycosylases with different substrate specificities. All characterized glycosylases are small monomeric proteins that do not require factors for either of their two functions: DNA damage recognition or enzymatic activity (Yu *et al*., 1999).

A highly mutagenic and common reactive oxygen species generated adduct is 8-oxoguanine (8-oxoG). It can cause GC→TA or TA→GC transversions if not repaired (Tajiri *et al*., 1995). In humans the enzyme responsible for cleaving this damaged base is 8-oxoG DNA glycosylase (OGG1) (Figure 7). Once the enzyme has cleaved the glycosylic bond at 8-oxoG, the protein can either be released, acting as a mono-functional glycosylase generating an AP site (Figure 7, pathway II, step 1), or it can further process the damage site acting as a bifunctional glycosylase and generating a 3'-α, β-unsaturated aldehyde end (3'-PUA) (Izumi *et al*., 2003; Wiederhold *et al*., 2004) (Figure 7, pathway I, step 1). Since OGG1 is inhibited by its own β-elimination products, OGG1 has been shown to mainly generate AP sites at 8-oxoG (Hill *et al*., 2001). Another important DNA glycosylase is E. coli endonuclease III homolog 1 (NTH1). NTH1 is a bifunctional glycosylase that removes thymine glycol damage with an associated AP lyase activity that will remove the resulting AP site by β-elimination generating the 3'-PUA end (Bailly *et al*., 1989) (Figure 7, pathway I, step 1). A third unique family of DNA glycosylases are the E. coli nei homologs, NEIL1, NEIL2 and NEIL3, of which NEIL1 and NEIL2

have been most extensively studied. NEIL1 mainly excises 5'-hydroxyuracil,

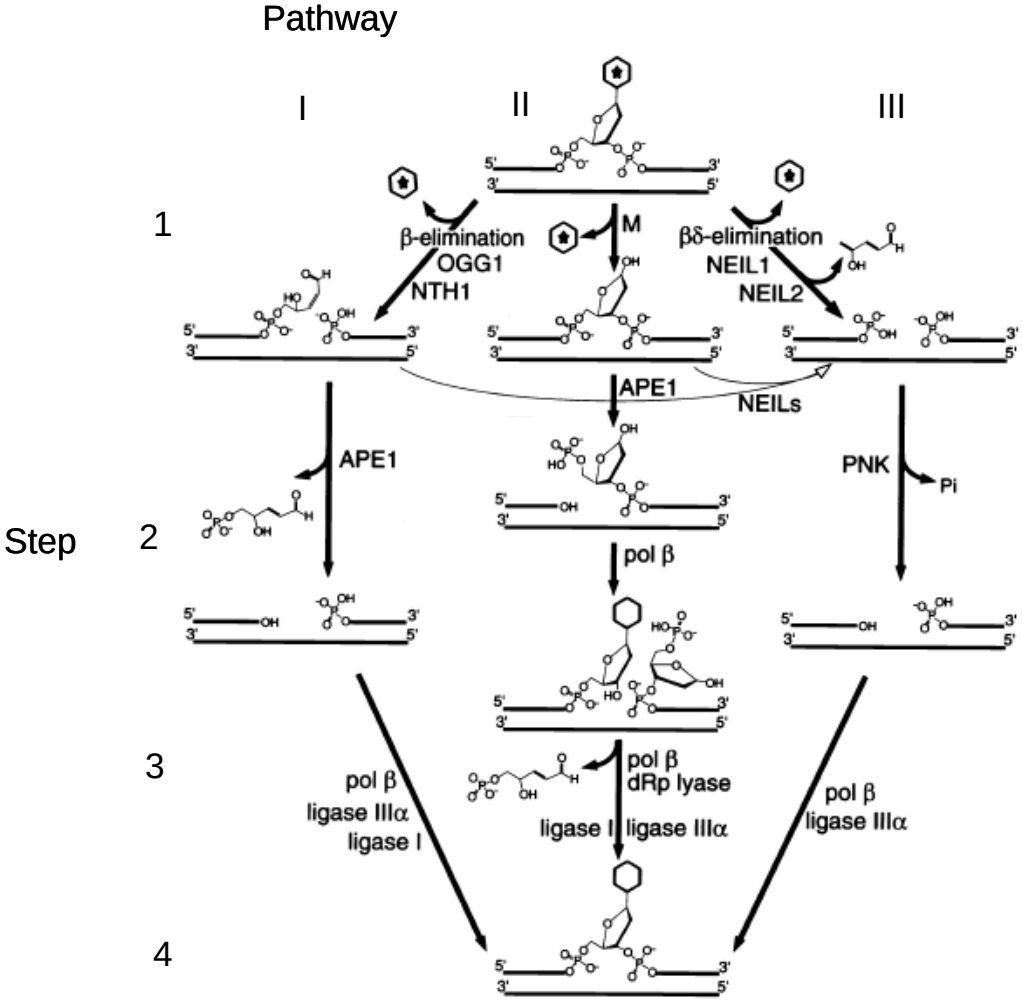


Figure 7. Three pathways of BER
Pathway I shows the repair of DNA damage by BER using bi-functional glycosylases, such as OOG1 and NTH1. Pathway II shows the pathway of BER from damage repaired by mono-functional glycosylases. Pathway III shows BER by way of NEIL1 and NEIL2. Reprinted from Molecular Cell, vol. 15, Wiederhold L., Leppard JB., Kedar P., Karimi-Busheri F., Rasouli-Nia A., Weinfeld M., Tomkinson AE., Izumi T., Prasad R., Wilson SH, AP endonuclease-independent DNA base excision repair in human cells, 209-220, copyright 2004, with Elsevier.

5'-hydroxycystosine, 5,5'-dihydroxyuracil, and formamidopyrimidines (Hazra *et al*., 2000a). Whereas NEIL2 cleaves mainly 5'-hydroxyuracil, and 5'-hydroxycystosine adducts (Hazra *et al*., 2000b). NEIL1 and NEIL2 differ form the other mentioned

glycosylases because, unlike them, they do not require APE1 to create the 3'OH end from an AP site or require a 3'PUA for DNA polymerase action. Both NEIL1 and NEIL2 have β,δ-elimination activities that generate 3' and 5'-phosphate ends that DNA polymerase β can synthesize past (Izumi *et al*., 2003; Wiederhold *et al*., 2004) (Figure 7, pathway III, step 1).

Thus, the apurinic sites, or lesions formed from the removal of the damaged base, are usually removed by one of two classes of endonucleases. The class I endonucleases, such as NTH1 have a dual function both as a glycosylase and an AP lyase, and they cleave the bond 3' to the abasic sugar. Class II AP, endonucleases such as APE, cleave the bond 5' to the abasic sugar (Yu *et al*., 1999). AP-endonucleases, of which only one has been characterized in humans (APE1), generates the 3'OH end from both AP sites and 3'PUA ends, following removal of the damaged base by a glycosylase (Izumi *et al*., 2003) (Figure 7, pathway I and II, step 2).

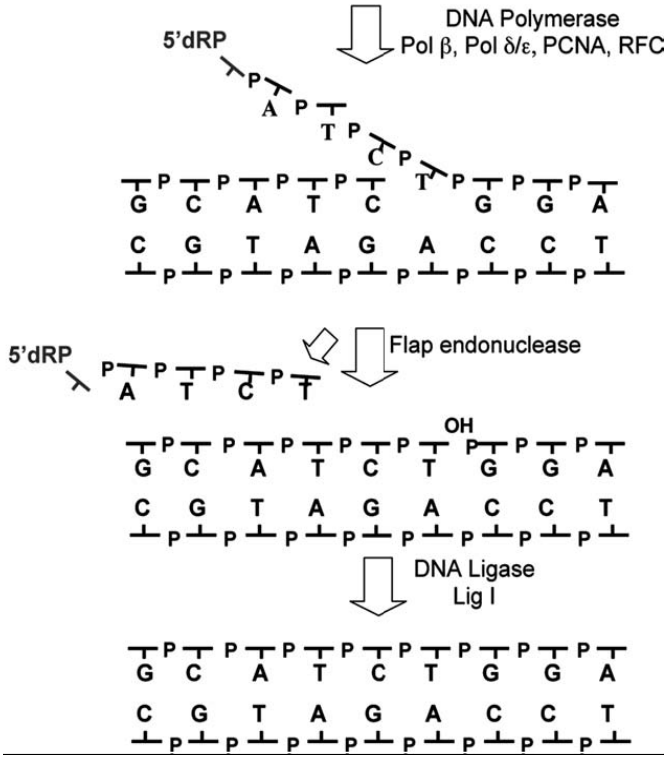


Figure 8. Long patch BER

Long patch BER can be initiated following cleavage of the phosphodiester and generation of the 3'OH end. It is characterized by an overhanging flap of 2-7 nucleotides that must be cleaved by FEN1. Modified from Biochimie, vol 85, Fortini P., Pascucci B., Parlanti E., D'Errico M., Simonelli V., Dogliotti E, The base excision repair: Mechanisms and its relevance for cancer susceptibility, 1053-1071, copyright 2003, with permission from Elsevier.

Following cleavage of the phosphodiester bond the BER pathway may take one of two routes, short patch repair or long patch repair. The result of adduct removal through the BER pathway is the

generation of a 1-7 nucleotide gap. In short patch repair, resynthesis of a single nucleotide by pol β takes place. DNA polymerase β is the optimal polymerase to fill in these short gaps (Randahl *et al*., 1988) (Figure 7, step 3). Polymerase β has an intrinsic AP lyase activity that can remove the remaining 5'-abasic site following DNA glycosylase action, prior to DNA synthesis (Srivastava *et al*., 1998). In long patch repair the resynthesis of 2 to 7 nucleotides occurs by either Pol β or Pols δ/ε (Figure 8). This route requires the use of flap endonuclease 1 (FEN1) and some accessory proteins not required for short patch repair to cleave the nucleotide flap that is generated by synthesis of more than one nucleotide (Fortini *et al*., 2003). Flap endonuclease 1 (FEN1) actually cleaves the 5'-phosphodeoxyribose end 2-6 nucleotides down from the AP site to generate a larger gap which can then be filled in by pol β (Kim *et al*., 1998). This FEN1 pathway is more important for the repair of oxidized or reduced AP sites (Xu *et al*., 2003). The concluding step in BER is sealing the DNA strand (Figure 7, step 4) using DNA ligase, of which there are four types (DNA ligase I, II, III, and IV). Although they show different substrate specificity, all have a common conserved lysine residue at the active site (Tomkinson *et al*., 1991).

Proper BER function also requires many accessory proteins that do not have a specific enzymatic activity. XRCC1 has been shown to interact with pol β, DNA lig IIIα, PNK, APE1, and PARP (Caldecott *et al*., 1996; Kubota *et al*., 1996; Whitehouse *et al*., 2001), as well as with NTH1 and NEIL2 (Campalans *et al*., 2005). Like XRCC1, XPG is another structural or scaffolding protein involved in BER that interacts with many proteins, such as NTH1 and APE1 (Izumi *et al*., 2003). Two other accessory proteins are poly(ADP-ribose) polymerase 1 (PARP1) and BRCA1. It has been suggested that PARP1's affinity for SSBs may play a role in sequestering SSBs until other BER proteins are recruited to the site, and that BRCA1 may be the basis for forming a platform for the recruitment of other BER proteins at the site of a stalled RNA pol II complex located on the damaged strand (Izumi *et al*., 2003).

Polymorphisms and cancer susceptibility in BER genes

Since the human genome experiences thousands of damaging events per day that require repair initiated by DNA glycosylases, it may stand to reason that deficiencies in these DNA glycosylases may be modifiers for the development of human disease, such as cancer. Some evidence of this is shown by BER failure in E. coli (Duncon and Weiss, 1982), yeast (Impellizzeri *et al*., 1991) and rodents (Klungland *et al*., 1995) that have shown an increase in mutations. But deficiencies in DNA glycosylases are not the only concern among BER genes. For example, DNA ligase deficiency has been shown in both lymphoblastic leukemic cells (Rusquet *et al*., 1988) and in Fanconi's anemia (Schwaiger *et al*., 1982). Also, over-expression of *APE* can cause frame-shift mutations in human cell lines (Hofseth *et al*., 2003) and has been shown in Chinese hamster ovary (CHO) cells to result in a 40% increase in micronuclei frequency and a 30% increase in sister chromatid exchanges (Sossou *et al*., 2005). Three polymorphisms in *APE1* (Leu104Arg, Glu126Asp, Arg237Ala) have also been extensively studied and have been shown to exhibit a 40-60% decrease in AP-endonuclase activity (Hadi *et al*., 2000).

In a large meta-analysis of previous case control studies, Hung *et al*. (2005) found that the Ser326Cys polymorphism in the *OGG1* gene is a risk factor for lung cancer. Subjects inheriting the homozygous 326Cys variant genotype had a significantly increased risk of lung cancer, which is consistent with previous work that showed decreased BER capacity with this genotype (Kohno *et al*., 1998). Zienolddiny *et al*. (2006) also showed that the *OGG1* Ser326Cys polymorphism and the *PCNA* A1876G intron polymorphism were associated with an increased risk of non-small cell lung cancer (NSCLC). Interestingly, Zienolddiny *et al*. (2006) also presented evidence that the *APE1* Ile63Val polymorphism was significantly protective against NSCLC. However, Ito *et al*. (2004) have shown that the *APE1* homozygous 148Glu polymorphism in combination with smoking significantly increases the risk of lung cancer. It was also observed that subjects who have the *XRCC1* Arg194Trp genotype had lowered sensitivity to mutagens using a mutagen sensitivity assay in which their lymphocytes were challenged with bleomycin and benzo[a]pyrene-diol-epoxide (BPDE) (Wang *et al*., 2003), which may

explain the observed association with tobacco-related cancers (Hung *et al*., 2005). Another polymorphism, *XRCC1* IV10+141G>A, in the intronic region of *XRCC1* has been shown to be associated with altered gene expression, apoptosis and a significant increase in familial breast cancer risk (Bu *et al*., 2006). Interestingly, the *XRCC1* Arg399Gln polymorphism has been shown to be a significant risk modifier for prostate cancer (Chen *et al*., 2006), colorectal carcinoma (Abdel-Rahman *et al*., 2000; Skjelbred *et al*., 2006), and alcohol induced liver cirrhosis (Rossit *et al*., 2002). It has also been shown that the 399Gln polymorphism in XRCC1 is associated with increased genetic damage as measured by chromosome aberration frequencies following challenge with NNK (Abdel-Rahman and El-Zein, 2000).

NUCLEOTIDE EXCISION REPAIR

Similar to BER, nucleotide excision repair (NER) is a multi-protein repair system. Unlike BER and the other repair pathways, most of the steps in NER require a significant amount of protein interaction within large protein complexes. NER is the major repair pathway for the repair of large bulky adducts, such as cisplatin-guanine, thymine dimmers, cyclobutane pyrimidine dimers and 6-4 photoproducts, that are accumulated due to exposure to UV irradiation or exogenous chemicals (Figure 9A). The NER pathway involves four steps: (1) damage recognition by a complex of bound proteins including XPC; (2) unwinding of the DNA by XPD and XPB in the transcription factor IIH (TFIIH) complex; (3) removal of the damaged single-stranded fragment by the ERCC1/XPF complex; and (4) synthesis of new nucleotides by DNA polymerases (see above and Friedberg, 2001 for details on NER).

Mechanisms of nucleotide excision repair.

NER consists of two distinct pathways: (1) transcription coupled repair (TCR), which preferentially repairs DNA damage that has occurred in the transcribed strand of actively transcribed genes to eliminate transcription blockage, and (2) global genome

repair (GGR), which repairs DNA damage throughout the entire genome including the non-transcribed strand of actively transcribed genes (Hoogervorst *et al*., 2005). Within TCR, damage recognition and removal is initiated by the presence of a stalled RNA polymerase II that cannot bypass the lesion, in conjunction with CSA and CSB proteins. Damage recognition in GGR is accomplished by the XPC-hHR23B and XPE-DDB complexes. This damage recognition step in NER is believed to be the rate limiting step (Thoma and Vasquez, 2003) (Figure 9B). In the most widely accepted and documented model of NER, termed the sequential assembly model (Figure 10), it is hypothesized that protein components are recruited to the site of damage based on their task and then, upon completion of that task, are either removed from the area or remain if their participation in repair is further required (Guzder *et al*., 1996; Thoma and Vasquez, 2003).

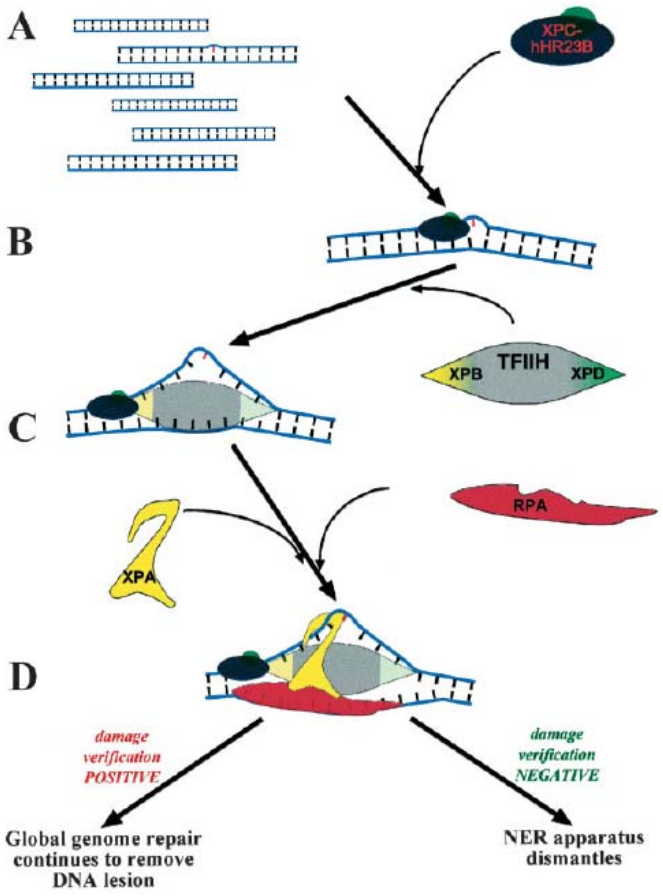


Figure 9. NER recognition/preincision and verification
A) occurrence of DNA damage. B) Damage recognition by XPC-hHR23B. C) Recruitment of TFIIH complex and further unwinding of DNA through XPB and XPD helicase activities. D) Recruitment of and verification of actual damage by XPA and RPA. Reprinted from Molecular Carcinogenesis, vol. 38, Thoma BS., Vasquez KM, Critical DNA damage recognition functions of XPC-hHR23B and XPS-RPA in nucleotide excision repair, 1-13, copyright 2003, with permission from Wiley-Liss INC.

It was initially hypothesized that the damage recognition role in NER is played by the complex of replication

protein A (RPA) and XPA (He *et al*., 1995; Matsuda *et al*., 1995) or XPC/hHR23B (Thoma and Vasquez, 2003). XPA is a 40kDA zinc metalloprotein that forms a complex with RPA, a single stranded DNA binding protein made up of three subunits. XPC is a 106kDa protein that complexes with hHR23B, a 43 kDa homolog of the yeast RAD23 protein (Thoma and Vasquez, 2003). Both of these protein complexes have been shown to initially bind to damaged DNA, however, this binding does not insure repair. For instance, XPA has been shown to bind to N-acetyl-2-aminofluorine adducts, but no subsequent repair was observed (Sugasawa *et al*., 1998). To help determine the true initiating complex in GGR one might examine the binding affinity of damaged DNA. Lao *et al*. (2000) have estimated the binding affinity of XPA-RPA to UV-damaged DNA at approximately 4.2×10^{-7} M and, in comparison, Reardon *et al*. (1996) have observed the binding capacity of XPC-hHR23B to be about 100 times greater at even lower doses of UV irradiation. However, it would seem that the specificity of binding to damaged DNA over undamaged DNA is stronger with the XPA-RPA complex that can be 10 to 1000 times more specific for damaged DNA, depending on the source of damage (Asahina *et al*., 1994; Lao *et al*., 2000; Hey *et al*., 2001; Vasquez *et al*., 2002).

On the contrary, the specificity of the XPC-hHR32B complex may be only 2 to 10 times higher for damaged DNA over undamaged DNA (Thoma and Vasquez, 2003). Sugasawa *et al*. (1998), using a series of competition experiments, found that XPC-hHR32B initiation results in faster repair than if initiated by XPA-RPA. These data, and that from Evans *et al*. (1997), showing the requirement of XPA-RPA in NER, support the model that, in GGR, XPC-hHR32B is the damage recognition complex and the XPA-RPA complex provides a sort of secondary checkpoint to verify that there is actual damage. Without this verification by XPA-RPA, NER will not continue (Evans *et al*., 1997; Sugasawa *et al*., 1998). Support of this model also comes from protein interaction studies in which XPC-hHR23B, but not XPA-RPA, is required for the effective recruitment of the transcription factor IIH (TFIIH) complex (Li *et al*., 1998; Yokoi *et al*., 2000). Binding of the XPC-hHR23B complex to the site of damage induces a further structural change by distorting the DNA (Susgsawa *et al*., 1998; Susgasawa *et al*., 2002).

This binding of XPC-hHR23B and the further conformation change as a result, recruits the TFIIH complex, which will further unwind the DNA through its helicases, XPB and XPD. The initiation steps of NER, that is the binding of XPC-hHR23B and the recruitment of TFIIH, occur independently of ATP (Riedl *et al*., 2003) (Figure 9C). However, ATP is required, following the recruitment of TFIIH, for XPD and XPB helicase activity to further open the DNA in the proximity of the damage (Schaeffer *et al*., 1993; Evans *et al*., 1997). The increased strand opening as a result of helicase activity allows greater accessibility for XPA and RPA to bind to the site (Aboussekhra *et al*., 1995; Mu *et al*., 1995; Missura *et al*., 2001), and hence verifying the presence of DNA damage that must be repaired or resulting in the dismantling of the NER proteins in the event there is no damage (Figure 9D). Following the binding of XPA and RPA, XPG is then recruited, which results in the exclusion of XPC-hHR23B from the damage site (Riedl *et al*., 2003). This may induce further structural change that helps recruit XPF-ERCC1 (Evans *et al*., 1997).

XPG and XPF-XRCC1 are the DNA endonucleases responsible for the dual incision of damaged DNA during NER. XPG specifically cuts 3' to the single stranded DNA bubble created as a result of the NER complex (Hoy *et al*., 1985; O'Donovan *et al*., 1994). XPF-XRCC1 specifically cuts 5' to the bubble and in a 5' to 3' direction (Sijber *et al*., 1996). The recruitment of XPF-XRCC1 to the damage site signals the release of the TFIIH complex, which is now free to join another cycle of DNA repair or participate in RNA transcription (Riedl *et al*., 2003). This dual incision step is carried out in an ATP-dependent manner in which single-strand incision at the 6^{th} phosphodiester bond 3' to the adduct and the 22^{th} phosphodiester bond 5' to the adduct are made. The exact size of the excision may depend on the specific adduct being repaired (Mayne *et al*., 1982; Huang *et al*., 1992). Also with the arrival of XPF-XRCC1 and the dual incision of the DNA, Riedl *et al*., (2003) observed a release of XPA along with TFIIH. The resulting gap is filled, starting at the exposed 3'OH end, by DNA polymerases. Polymerase δ and ε are the DNA polymerases involved in filling the gap created by excision with the help of PCNA (Pan *et al*., 1995). The final step of sealing the nick is carried out by DNA

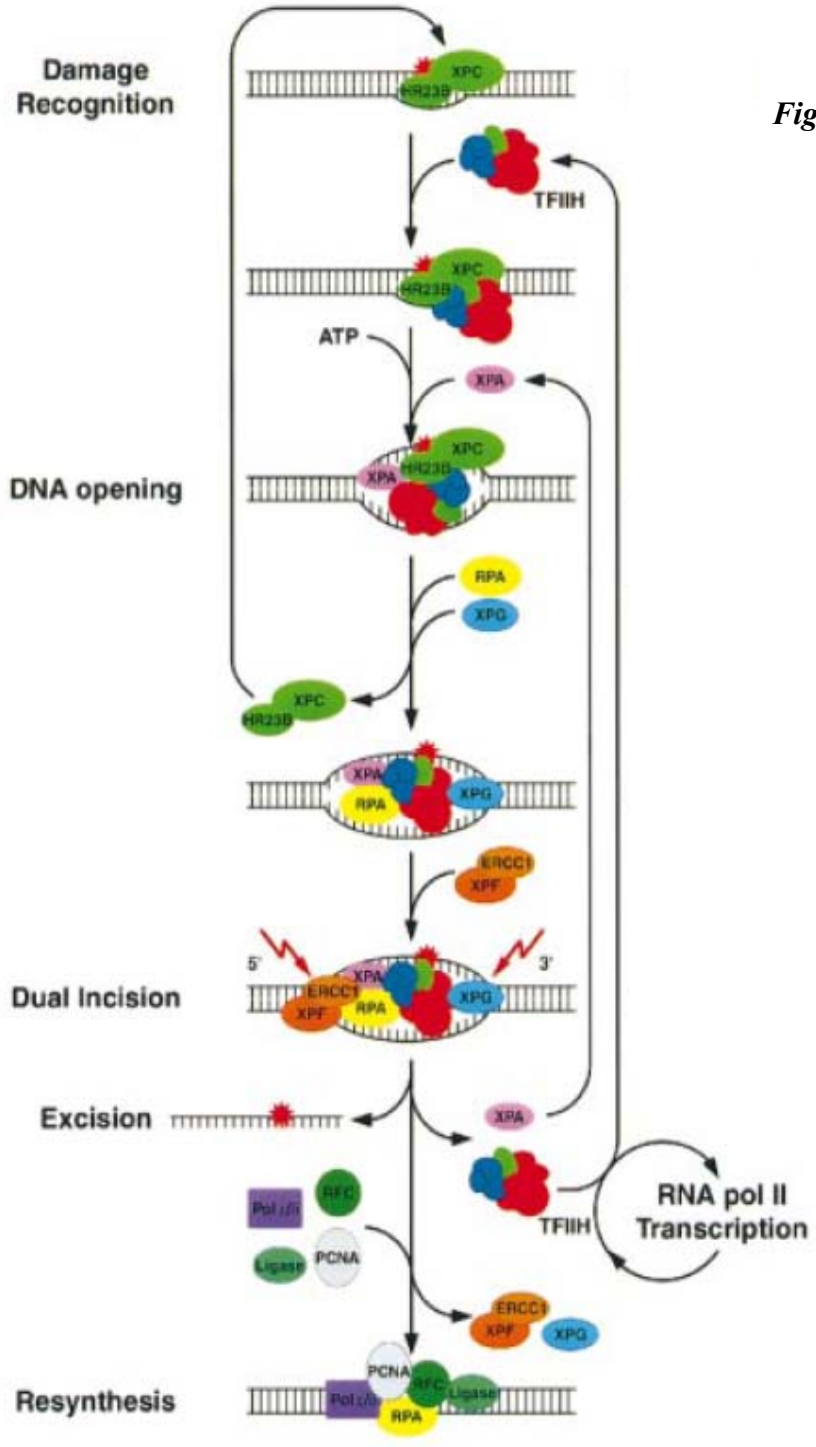

Damage Recognition
XPC
HR23B
TFIIH
ATP
XPA
DNA opening
RPA
XPG
ERCC1
XPF
5'
3'
Dual Incision
Excision
RNA pol II Transcription
RFC
PCNA
Ligase
Resynthesis

Figure 10.

Figure 10. NER: sequential assembly model
During damage recognition XPC-hHR23B binds to the site of damage and initiates NER by the recruitment of TFIIH. During an ATP dependent activity XPB and XPD of TFIIH unwind the DNA allowing XPA and RPA to bind and verify damage. Following damage verification dual excision is carried out 3' by XPG and 5' by XPF-ERCC. These repair complexes are then released and Pol δ, PCNA, Lig III are recruited for resynthesis and ligation. Reprinted by permission from Macmillan Publishers Ltd: The EMBO Journal (Riedl T, Hanaoka F, Egly JM. 2003. The coming and goings of nucleotide excision repair factors on damaged DNA. EMBO 22(19); 5293-5303), copyright (2003).

ligase. Figure 10 summarizes the mechanism of NER and the order in which proteins and protein complexes are recruited to the damaged site.

Initially the TFIIH complex was thought to consist of the 6 subunit core (XPB, XPD, p62, p52, p44, and p34) plus the 3 subunit CAK complex (Mat1, cdk7, and cyclin H) (Araujo *et al*., 2000). However, recently Giglia-Mari *et al*., (2004) discovered a tenth subunit that belongs to the core TFIIH complex. This group found that the NER defect in cells from individuals with Trichothiodystrophy (TTD) could be corrected using purified TFIIH; however, none of the originally identified nine genes in TFIIH were mutated in the TTD cells. This suggests that another protein was involved with the TTD phenotype. Using protein screening techniques, they were able to identify a 8kDa protein in yeast TFIIH and its subsequent human homolog p8. Through antibody labeling, Giglia-Mari *et al*. (2004), were able to show that p8 is part of the core TFIIH complex, and they suggest a role for p8 as a stabilizing protein that protects TFIIH from degradation. They also showed that p8 stimulated the ATPase activity of XPB through a joint interaction with p52. This ATPase activity, stimulated by p8, is suggested to catalyze a conformational change that allows for XPC-hHR23B repositioning and DNA unwinding by XPD (Coin *et al*., 2006). Electron microscopy imaging suggests that the TFIIH complex is a ring-like structure approximately 16 x 12.5 x 7.5 nm in size with a single protein domain extruding from one side and a central hole 2.5 to 3.5 nm in diameter (Schultz *et al*., 2000). From immuno-labeling experiments, Schultz *et al*. (2000) were also able to show that the protein extrusion contains the CAK complex with MAT1 located closest to the ring and interacting with XPD. On the other side of the bulge from XPD is p44 and XPB (Figure 11). The TFIIH complex, as previously mentioned briefly, functions both in DNA repair (GGR and TCR), as well as in RNA transcription. It has been suggested by

Coin *et al*., (2004) that the regulation of TFIIH between NER and RNA transcription results from the phosphorylation of a single amino acid, S751, of XPB. The phosphorylation of this residue does not prevent TFIIH binding to the damaged site nor does it prevent DNA unwinding. Interestingly, phosphorylation of S751 prevents complete dual incision by blocking 5' incision by XPF-XRCC1. However, the phosphorylation of S751 does not prevent TFIIH's role in RNA transcription (Coin *et al*., 2004).

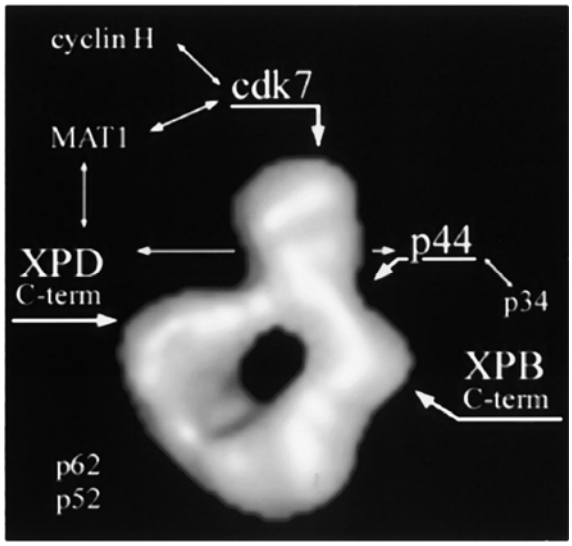


Figure 11. Electron microscope image of the TFIIH complex.
TFIIH is a ring structure 16 x 12.5 x 7.5 nm in size with a 2.5 to 3.5nm diameter central hole. The protein protrusion contains the CAK complex immediately adjacent to XPD on one side of the ring and p44 on the other. Opposite from XPD is XPB, in close proximity with p62 and p52 which are on the ring opposite the CAK complex. Reprinted from Cell, vol. 102, Schultz P., Fribourg S., Poterszman A., Mallouh V., Moras D., Egly JM, Molecular structure of human TFIIH, 599-607, copyright 2000, with permission from Elsevier.

Many accessory proteins may also be involved in NER. For example, p53, a tumor suppressor gene with involvement in cell cycle checkpoint regulation (Kastan *et al*., 1993), DNA damage (Lee *et al*., 1995), gene expression (Livingstone *et al*., 1992), and apoptosis (Yonish-Rouach *et al*., 1991), may be a regulatory element in the NER pathway. P53 may modulate the repair pathway through transcriptional activation, and it has been shown that p53 can regulate the expression levels of XPC and XPE (Adimoolam *et al*., 2002; Hwang *et al*., 1999). It is also possible that p53 acts as a direct DNA repair factor because of its direct interaction with a number of NER proteins, including XPB and XPD, by interaction of its carboxyl terminal domain with the helicase III motif of XPB (Wang *et al*., 1995). Wang *et al*. (1996) showed that p53-induced apoptosis can be modulated through its interaction with XPB and XPD, and they showed that p53 apoptosis is deficient in the cells of xeroderma pigmentosum (XP) patients with mutated nonfunctional XPD or XPB. In these cells, normal phenotype can

be rescued by expression of wild type XPD and XPB. It is suspected that p53 may modulate NER by binding to single-strand DNA resulting from DNA damage and inhibiting XPB and XPD helicase activity thereby preventing completion of NER and driving the cell towards apoptosis (Wang *et al*., 1995).

Diseases from defective nucleotide excision repair

Xeroderma pigmentosum (XP) is a rare autosomal recessive disease. This disease is characterized by early development of basal cell and squamous cell carcinomas, sensitivity to sun resulting in degenerative alterations in the skin and eyes, and a 1000-fold increase in both non-melanoma and melanoma skin cancers (Tuteja and Tuteja, 2001) with a mean age of onset of two years (Kraemer, 1997). There are two major forms of XP; one form is a result of a defect in nucleotide excision repair, and the second form is a result of a defect in repair of replicative DNA damage (Tuteja and Tuteja, 2001). Some XP patients also show symptoms associated with De Sanctis-Cacchione syndrome (DSC), which is characterized by immature sexual development, growth retardation, and mental retardation (Tuteja and Tuteja, 2001). XP patients may also present with neurological abnormalities, such as intellectual deterioration, and loss of speaking ability, which may be due to premature neuronal death (Robbins *et al*., 1991). The symptoms of XP can be quite variable, ranging from mild forms of skin abnormalities to the most severe forms of skin cancer and mental retardation. It is possible that the variability may be influenced by the specific mutations that each patient has that resulted in the XP phenotype. In each of the seven XP groups (XP-A through XP-G) a corresponding NER gene (XPA through XPG) has been mutated (Vermeulen *et al*., 1991; Hoeijmakers, 1993). For example, XP patient XP11BE has a mutation that results in a structural alteration in the carboxy-terminal of XPB and reduces TFIIH activity to only 15% of normal (Coin *et al*., 1999). This was shown by Evans *et al*., (1997) to be due the reduced ability of XPF-XRCC1 to carry out 5' DNA incision. Mutations in XP-B, XP-A, and XP-G can be found in patients that have combined symptoms of both XP and Cockayne's Syndrome (CS).

Cockayne's Syndrome is characterized by the early onset of cachectic dwarfism, microcephaly, deterioration of adipose tissue, mental retardation, retinal atrophy, cataracts, sun sensitivity, and growth retardation (Leech *et al*., 1985; Tuteja and Tuteja, 2001). The average age of onset of CS is between ages 3 and 5. Patients with CS typically die from pneumonia and respiratory infections by the age of 12 (Nance and Berry, 1992). The symptoms of CS are a result of a deficiency or loss of DNA repair (Lehmann *et al*., 1979) or RNA synthesis (Mayne and Lehmann, 1982) following UV irradiation. This was shown to be due to an impaired preference for TCR over GGR following exposure to DNA damaging agents (Venema *et al*., 1991; van Hoffen *et al*., 1993). There are three identified groups of CS: A, B and C (Tanka *et al*., 1981; Lehmann, 1982). CS groups A and B comprise patients with classical CS, and group C are individuals with combined features of CS and XP. Consequently CS-A and CS-B patients show no signs of defects in NER global genome repair, but are deficient in TCR induced by exposure to ionizing radiation, such as UV radiation from sunlight (Leadon and Cooper, 1993). CS is largely a result of mutations in two genes, CSA and CSB, that play a role as initiating factors for transcription repair of DNA damage (Troelstra *et al*., 1992), resulting in defective TCR. Mutations in these two genes are the cause of over 90% of all CS cases (Tuteja and Tuteja, 2001). CSA interacts directly with components of the TFIIH complex, whereas CSB will interact with CSA and XPG (Henning *et al*., 1995). Recent evidence suggests that CS is a disease state that results from increased and excessive cellular apoptosis (Hanawalt, 2000).

Trichothiodystrophy (TTD), like XP, is a rare autosomal recessive disorder. TTD is characterized by mental retardation, impaired sexual development, and sulfur-deficient brittle hair and nails as a result of decreased expression of cysteine-rich matrix proteins in cuticle cells (Itin and Pittelkow, 1990; Tuteja and Tuteja, 2001). Approximately 50% of TTD patients present with UV sensitivity, but there is no association with cancer, and the majority of these patients have a defect in NER that is indistinguishable from XP-D patients (Tuteja and Tuteja, 2001). TTD patients fall into three groups, depending on which gene is mutated. These are XP-D (Stefanini *et al*., 1993a), XP-B (Vermeulen *et*

al., 1994) and TTD-A (Stefanini *et al*., 1993b), later identified as p8 (Giglia-Mari *et al*., 2004). TTD-A can result in a decrease in cellular concentration of TFIIH by up to 70% (Botta *et al*., 2002). It has been shown that most UV-sensitive TTD patients have only about 15 to 25% of unscheduled DNA synthesis, compared to normal cells, and are defective in the repair of UV-induced cyclobutane pyrimidine dimmers, but are more proficient in the repair of 6-4 photoproducts (Eveno *et al*., 1995). In fact these cyclobutane pyrimidine dimmers were shown to be the predominate mutagenic lesion in TTD cells (Marionnet *et al*., 1998).

Although these three diseases are linked, XP is largely a result of deficiencies in GGR, CS is mainly a result of TCR deficiencies, and TTD as been associated with transcriptional problems (Zurita and Merino, 2003). Those patients that have mutations in XPD have mainly mutations that affect the helicase activity of XPD. This results in NER defects that are largely due to structural changes in the TFIIH complex. In those patients that present with symptoms of TTD, the mutation in XPD also affects transcription (Zurita and Merino, 2003). Table 1 summarizes the similarities and difference between these three disease states.

Table I. Characteristics of XP, CS, and TTD

Genes	**XP** XPA-XPG	**CS** CSA, CSB, XPG	**XP/CS** XPD, XPB, XPG	**TTD** XPD, p8
Disease symptoms				
Photosensitivity	mild-severe	mild	moderate-severe	mild-moderate
skin cancer	mild-severe	NA	mild-severe	NA
brittle hair	NA	NA	NA	yes
mental retardation	mild-severe	mild-servere	mild-severe	mld
growth retardation	mild	mild	mild	NA
Cellular abnormalities				
GG-NER activity (% of normal)	0-70	100	0-40	Oct-45
transcriptional recovery after UV exposure	low-normal	low	low	low
UV sensitivity (% of normal)	1.5-10x	2.5-5X	4-10X	1.5-5X
TFIIH level (% of normal)	60-100	100	75	30-50

Polymorphisms in nucleotide excision repair genes and cancer susceptibility

Polymorphisms in NER genes have been studied extensively, and associations with a variety of cancers and biomarkers of exposure have been found. A 2004 review of cancer susceptibility stated that prostate cancer accounted for about 33% of incident cancer cases, resulting in approximately 30,000 deaths annually (Jemal *et al*., 2004). NER is suspected to be an important player in the pathogenesis of prostate cancer for several reasons: (1) NER is known to repair genetic damage resulting from exposure to a number of suspected prostate cancer carcinogens, such as polycyclic aromatic hydrocarbons and hetrocyclic aromatic amines, and (2) both of these types of carcinogens can be activated by prostate cells (Martin *et al*., 2002). In one study, the *XPD* Asp312Asn and *XRCC1* Arg399Gly polymorphisms in combination were shown to be associated with increased risk of prostate cancer (Rybicki *et al*., 2002). This may be partly explained by the observations of Spitz *et al*. (2001) who found that homozygous variants at the 312 and 751 loci in *XPD* reduced NER capacity. Individuals with these polymorphisms were also shown to be at higher risk for cutaneous malignant melanoma (Baccarelli *et al*., 2004). The NER initiation/verification gene *XPA* was also shown to result in decreased NER capacity with the inheritance of the polymorphic G allele in the 5' non-coding region upstream of the start codon (Wu *et al*., 2003)

Mutations in the N-terminal region of p44 or in the C-terminal region of *XPD* that result in decreased interaction between those two proteins as members of the TFIIH complex result in sub-optimal XPD helicase activity, and thus reduced NER capacity and TFIIH destabilization (Tremeau-Bravard *et al*., 2001). In two large case control skin cancer studies, it was observed that skin cancer patients had a 19%-42% decrease in NER capacity compared to non-skin cancer controls. (Matta *et al*., 2003; Wei *et al*., 2003). The same trend was seen in a number of studies of lung cancer patients, where a 12%-55% decrease in NER capacity was observed (Wei *et al*., 1996; Wei *et al*., 2000; Spitz *et al*., 2001; Shen *et al*., 2003; Spitz *et al*., 2003). NER capacity has also been shown to be decreased in patients with head and neck cancer by about 31% (Cheng *et al*., 1998), breast cancer by 18%-36% (Ramos *et al*., 2004; Shi *et al*., 2004), and even in patients

with prostate cancer by 20% (Hu *et al*., 2004). The *ERCC1* Asn118Asn, *ERCC1* C15310G and *XPD* Lys751Gln polymorphisms have recently been associated with increased risk of non-small cell lung cancer (NSCLC). However, the *XPA* A23G and *XPG* His46His polymorphisms have been associated with a significant decrease in NSCLC (Zienolddiny *et al*., 2006). Another polymorphism in *XPG*, the His1104Asp, has also been associated with reduced cancer risk, specifically squamous cell carcinoma of the lung (Jeon, 2003). *XPD* 312Asn and *XPA* (-4AA) polymorphisms, in combination, have been shown to be associated with increased risk of squamous cell carcinoma of the lung (Popanda *et al*., 2004). While the *XPD* 751Gln homozygous genotype has been shown to reduce DNA repair capacity (Spitz *et al*., 2001), the inheritance of the *XPA* -4G allele has been associated with increased DNA repair capacity (Butkiewicz *et al*., 2004). An intrinsic poly (AT) polymorphism in the *XPC* gene has been shown to contribute to an increased risk of squamous cell carcinoma of the head and neck (Shen *et al*., 2001).

The two most commonly studied polymorphisms in the *XPD* gene, 1012G→A and 2329A→C, were first discovered by Broughton *et al*. (1996) and later renamed according to the change in protein to the more commonly recognized Asp312Asn and Lys751Gln. Both of these polymorphisms are highly prevalent in the population with the variant Asn allele at position 312 occurring at a prevalence of 24%. For the polymorphism at position 751 the variant Gln allele occurs in 21% of studied individuals (Clarkson and Wood, 2005). In a host cell reactivation study, lymphocytes with the XPD homozygous Gln genotype at position 751 showed a decreased ability to repair the normal phenotype following UV irradiation compared to cells with the homozygous Lys genotype or the heterozygous genotype (Qiao *et al*., 2002a).

Importance of the XPD protein and its role in the NER pathway and cell-cycle regulation

As mentioned earlier the focus of this dissertation is on understanding the functional and biological significance of SNPs that occur in the coding region of the *XPD* gene. The XPD protein is a major player in the NER pathway and is essential for life.

This is demonstrated by the embryo-lethality of *xpd* knockout mice (Boer and Hoeijmakers, 1999). NER system defects, due to mutations in the *XPD* gene, are responsible for the DNA repair defect syndromes trichothiodystrophy (TTD), Cockayne's syndrome (CS) and xeroderma pigmentosum (XP) (Lehmann, 2000; Lehmann, 2003) as discussed above. Although *XPD* mutation-related diseases represent an extreme defect in the function of the XPD protein, variation in the activity of this protein in the general population due to inheritance of SNPs in the gene, for example, is highly plausible and probably implicated in some types of inherited cancer susceptibility. The XPD protein plays a vital role in the NER pathway, which as stated earlier is the major DNA pathway responsible for the repair of bulky DNA adducts resulting from exposure to a variety of environmental carcinogens. In this pathway, the XPD protein plays a critical role as a member of the TFIIH transcription complex necessary for lesion demarcation and unwinding (Goodie *et al*., 2002).

TFIIH (Figure 11), as discussed above, is a multiprotein complex consisting of a ring-like structure made of XPD, XPB, p34, p44, p62, p8 and p52, and a protrusion made up of the CAK complex (Ronen and Glickman, 2001). XPD has both an ATPase and a 5' → 3' helicase activity that is stimulated by its binding to p44 in the TFIIH complex (Sung *et al*., 1993). This gives XPD the responsibility for unwinding the damaged DNA structure to allow the rest of the repair complex to enter the site of damage and repair the damaged nucleotide.

In addition to its critical role in NER, the XPD protein also plays another important role in cell cycle regulation as a part of the cdk-activating kinase (CAK) complex (Coin *et al*., 1999). The CAK complex is a protein aggregate made up of cdk7, cyclin H, and MAT1. This protein complex is responsible for activating cdk1, cdk2, and cdk4, which play an important role in cell cycle progression. CAK also phosphorylates RNA polymerase II, thereby initiating promoter clearance and transcription elongation (Rossignol, 1997). There is extensive evidence that the XPD protein is the regulatory unit of the CAK complex (Chen *et al*., 2003). Down-regulation of XPD contributes to the up-regulation of CAK mitotic activity. It thus regulates cell cycle progression and may also

function in mitotic silencing of basal transcription. Alternatively, XPD up-regulation can result in decreased Cdk phosphorylation resulting in mitotic defects (Chen *et al*., 2003).

Despite the documented associations between SNPs in the *XPD* gene and increased susceptibility to adverse health effects resulting from environmental exposures (discussed briefly above), the effect of these SNPs on the ATPase and helicase activities of the XPD protein, which are critical for NER, have not yet been determined. In addition, the modifying effect of these SNPs on the interaction of XPD with the other protein members of the TFIIH and CAK complexes (e.g. p44 and MAT1) has not yet been elucidated. If the ATPase or helicase activities of the XPD protein are impaired (due to altered binding of XPD with p44 or with MAT1) NER capacity could ultimately be altered. In turn, this would lead to alteration in the levels and types of genetic damage observed in individuals exposed to carcinogens. Our study design, encompassing both *in vitro* and *in vivo* approaches, will address these important gaps in knowledge.

Cellular localization of the XPD protein

In eukaryotic cells, proteins involved in the management of genes are translated in the cytoplasm and then transported into the nucleus through the nuclear pore complexes. Most nuclear proteins are actively translocated into the nucleus by several mechanisms that involve, among others, the recognition of specific amino acid sequences, known as nuclear localization signals (NLSs), by specific receptors (reviewed by Silver, 1991; Garcia-Bustos *et al*., 1991; Jans *et al*., 2000). Unlike other DNA repair genes (e.g. *XPB* and *XPC*), and even other members of the TFIIH complex, which rapidly accumulate in the nucleus (in agreement with the predicted presence of classic NLSs in their amino acid sequences), the nuclear translocation of XPD appears to be partial and time-dependent. Thus translocation of XPD does not rely on the presence of a putative NLS (Santagati, *et al*., 2001), which implies that XPD may be imported into the nucleus in association with a NLS-containing protein. This protein is hypothesized to be one of the other components of the TFIIH complex (Boulikas, 1997; Santagati *et al*., 2001). This also suggests that structural alterations (such as those leading to interaction modifications of XPD with the

other members of the TFIIH complex) would result in an altered ability for XPD to translocate into the nucleus. The effect of the SNPs in the *XPD* gene on the ability of the resulting variant protein to localize to the nucleus is currently unknown. It is hypothesized that SNPs in the *XPD* gene may alter the nuclear translocation of the XPD protein by altering its interaction with other protein members of the TFIIH complex.

SNPS in the XPD gene

The *XPD* gene consists of 23 exons and spans ~54.3 kb at 19q13.3. The cDNA has a size of 2400 nucleotides. The *XPD* gene contains six common exon SNPs, of which four result in amino acid changes (non-synonymous) and two are silent (synonymous). Table 2 depicts the location, base change, the resulting amino acid change (if any) and the reported frequencies of these SNPs, which range from 0.04 to 0.42 in human populations. We have selected three SNPs to be studied (Figure 12). These are the two coding SNPs, Asp312Asn in exon 10 and Lys751Gln in exon 23, and the Arg156Arg SNP in exon 6. There are several reasons behind the selection of these SNPs: (1) Two of these SNPs (the Asp312Asn and the Lys751Gln) result in amino acid substitutions. As such they could potentially affect the resulting protein structure and function; (2) The Lys751Gln SNP is located within the p44 binding domain of the XPD protein, and could thus affect its interaction with p44 and, consequently, its helicase activity; (3) The Arg156Arg SNP in exon 6, although a C→A silent SNP, is associated with increased cancer risk (Sturgis *et al*., 2003; Yin *et al*., 2005), and is located within the MAT1 binding domain of XPD (Sandrock *et al*., 2001); (4) Each SNP occurs at a relatively high frequency (lowest frequency over 0.25), thus affecting a large segment of

Table II. SNPs in the coding region of the XPD gene

Codon	Base Change	Amino Acid Change	Variant Allele Frequency
156	C→A	None	0.25
199	C→G	Met →Ile	0.04
201	C→T	Tyr →His	0.04
312	G→A	Asp→Asn	0.42
711	C→T	None	0.25
751	A→C	Lys→Gln	0.29

the general population; and (5) Each of these SNPs has been shown to modulate DNA repair capacity and the risk of environmentally-associated cancers (as discussed above).

Due to the importance of *XPD* as a DNA repair gene, many epidemiological studies have focused on determining the association between polymorphisms in this gene and cancer risk. Several studies have shown that SNPs in the *XPD* gene modulate the risk of basal cell carcinoma of the skin (Dybdahl *et al*., 1999; Vogel *et al*., 2001; Lovatt *et al*., 2005), and cancers of the head and neck (Sturgis *et al*., 2002), and lung (Liang *et al*., 2003). Sturgis *et al*., (2002) reported an association between the synonymous 156C>A (Arg156Arg) polymorphism in exon 6 and squamous cell carcinoma of the head and neck. Even though this SNP does not result in amino acid substitution, Yin *et al*., (2005) also reported that the variant 156A was associated with increased risk of adenocarcinoma of the lung. The nonsynonymous variant 312Asn allele in *XPD* was associated with increased risk for lung cancer in a large study by Zhou *et al*., (2002), but this association was not detected in a smaller study by Butkiewicz *et al*., (2001). When the effect of 312Asn and another nonsynonymous variant allele, 751Gln, was considered together, the risk for lung cancer was higher in individuals carrying both of these *XPD* variant alleles (Spitz *et al*., 2001). In other studies, the Lys751Gln polymorphism was not

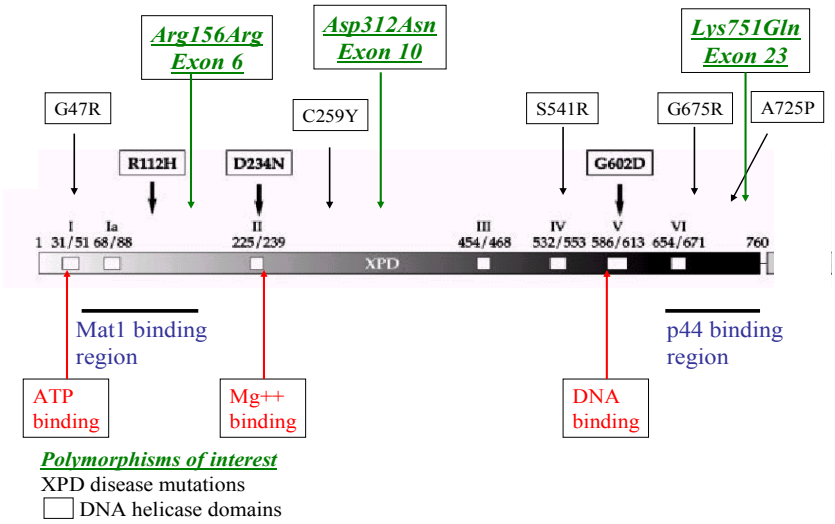


Figure 12. XPD protein map
XPD mutations, the location of selected SNPs studied, and binding regions are presented along the linear map of XPD. Modified from Benhamou S., Sarasin A. *ERCC2/XPD* gene polymorphisms and cancer risk. Mutagenesis. 2002; 17(6): 463-469, with permission from Oxford University Press.

associated with an elevated risk of lung cancer (Benhamou and Sarasin, 2005), but, in a study by Sturgis *et al*., (2003), this Lys751Gln polymorphism was found to be associated with a higher risk of head and neck cancer. These studies, and others, with sometimes conflicting results (summarized in Benhamou and Sarasin, 2005), have documented associations between cancer risk and *XPD* polymorphisms, but the exact mechanism(s) for these associations is unknown.

CHROMOSOME ABERRATION (CA) AS A BIOMARKER OF EFFECT

Chromosomal aberrations are established biomarkers of early biological effects preceding cancer in exposed populations. The importance of CAs as biomarkers of effects associated with cancer risk stems from the results of three separate large-scale epidemiological studies conducted on Nordic, Italian and Czech cohorts. These studies clearly indicate that a high frequency of CA in circulating peripheral blood lymphocytes (PBLs) is associated with an approximately two-fold increase in risk of developing cancer (reviewed by Hagmar *et al*., 2004; Norppa, 2004). This would correspond to an attributable proportion of cancer of 0.25, meaning that 25% of cancer cases could be prevented if factors leading to high CA frequencies were avoided (Hagmar *et al*., 2004). In fact, it is now possible to detect specific cytogenetic alterations in surrogate cells or tissues of exposed individuals at a very early stage of the carcinogenic process, long before the development of tumors. Furthermore, El-Zein *et al*. (2000) have shown that the levels of CAs in circulating lymphocytes of lung cancer patients can reflect the grade of the carcinoma.

Chromosome aberrations and tobacco smoking

Several studies have found that smokers have higher levels of CA in their PBLs (Obe *et al*., 1984; Galloway *et al*., 1986; Milillo *et al*., 1996). In fact, Gazdar and Minna, (1999) have shown that normal lung epithelial cells from smokers contain the specific types of chromosomal aberrations that are frequently found in histologically confirmed

lung tumors. In recent years, the sensitivity of cytogenetic techniques has been enhanced by the introduction of the fluorescence in situ hybridization (FISH) assay. This assay was shown to be highly sensitive in detecting chromosomal breakage, exchanges, and aneuploidy in both metaphase and interphase cells from populations exposed to low levels of environmental toxins (Rupa *et al*., 1995). Using the FISH assay, investigators have consistently observed a significant increase in CA in smokers compared to non-smokers (Conforti-Froes *et al*., 1997; Abdel-Rahman *et al*., 1998). To confirm a preliminary study on *XPD* polymorphisms (Affatato *et al*., 2004) conducted in our laboratory, we used both conventional and molecular FISH cytogenetics to increase the sensitivity for detecting genetic damage associated with exposure to tobacco carcinogens. Specific probes for chromosome 1 and 3 were utilized in the current project because: (1) studies with human lymphocytes have shown an elevated frequency of breakage in the heterochromatin region 1q12 after exposure to a variety of environmental clastogens (Pinkel *et al*., 1986; Eastmond *et al*., 1995; Conforti-Froes *et al*., 1997); (2) The chromosome 3p21 locus is reported to be sensitive to tobacco carcinogens such as the polycyclic aromatic hydrocarbon benzo(a)pyrene (Wu *et al*., 1998); and (3) the 3p21 deletion is a common finding in smokers who develop lung cancer (Roche *et al*., 1996; Sozzi *et al*., 1996).

USE OF THE MUTAGEN-SENSITIVITY ASSAY TO ELUCIDATE THE BIOLOGICAL SIGNIFICANCE OF SNPS IN SUSCEPTIBILITY GENES

The mutagen-sensitivity assay is a well-established technique that has been used to evaluate an individual's response to the mutagenic effects of genotoxic compounds using cells from that individual, such as Peripheral blood lymphocytes (PBLs) (reviewed in Wu *et al*., 2004). This assay, originally developed by T.C Hsu in 1984, reflects an individual's sensitivity to the mutagen used in the assay, as well as the DNA repair capacity of that individual. A number of studies have shown that PBLs from subjects with tobacco-related cancer have higher mutagen-sensitivity when compared to the cells of control subjects who do not have cancer (Wu *et al*., 2000; Shen *et al*., 2003). Using this

approach, Blasiak *et al*. (2004) reported that PBLs of breast cancer patients have reduced DNA repair kinetics and increased DNA damage following exposure to hydrogen peroxide and doxorubicin. The mutagen-sensitivity assay has also been used successfully to evaluate the effect of SNPs in susceptibility genes. For example, it has been shown that PBLs from individuals with the *CYP1A1* polymorphism in exon 7 and the *GSTM1* null allele exhibit higher DNA adduct levels following exposure to benzo(a)pyrene diolepoxide (Lodovici *et al*., 2004). Our laboratory has used this approach successfully to evaluate the effect of several SNPs in susceptibility genes on genetic damage induced by environmental carcinogens (Abdel-Rahman *et al*., 1999; Abdel-Rahman and El-Zein, 2000; Affatato *et al*., 2004; Hill *et al*., 2005). For example, in a preliminary study we observed an increased sensitivity to the mutagenic effects of NNK in PBLs with the 312Asn polymorphism in the *XPD* gene compared to PBLs with the homozygous wild-type Asp312Asp genotype (Affatato *et al*., 2004).

There are several strengths for using the mutagen-sensitivity assay: 1) it uses surrogate target cells such as PBLs which are easy to obtain and culture, 2) the information can be obtained from human cells as opposed to animal models, and 3) the effects of single and a complex mixture of compounds can be addressed. A weakness of this technique is that there is still a debate as to the relevance of effects in PBLs as opposed to the actual target tissue of the test compound. However, for this study sufficient evidence has shown that results with the biological marker, CAs, used in this research with PBLs closely resembles the results seen in target tissues (see above).

Why tobacco smokers constitute a suitable model for populations exposed to environmental carcinogens

Tobacco contains more than 3000 chemicals, at least 40 of which are known carcinogens (Bartsch *et al*., 2000). Although smoking is strongly linked to lung cancer, not all smokers will develop this cancer, suggesting that other factors, such as certain genetic polymorphisms may contribute to lung cancer development in smokers. Because of this, tobacco smokers constitute an eminently suitable model for populations exposed to

environmental carcinogens. We decided to continue to use tobacco smokers in this CA study of XPD polymorphisms because of the following reasons: (1) smoking has been linked to several diseases, including chronic obstructive pulmonary disease (COPD), cardiovascular disease, lung cancer and stroke (reviewed in Zaher *et al*., 2004), making smoking a public health concern; (2) smoking has been shown to affect CA frequency and other types of genetic damage (reviewed in DeMarini, 2004), making smokers a relevant population for studies of genetic damage induced by environmental exposure; (3) smoking is highly prevalent and smokers are easily accessible subjects. This provides a suitable exposed population that can be recruited easily in a short period of time.

As discussed above, tobacco smoke has been linked to increased genetic damage and the development of many diseases. The cause of these diseases may be due to exposure to the many carcinogens within the tobacco smoke. There are many classes of chemicals that are released in tobacco smoke, such as polycyclic aromatic hydrocarbons (PAHs), nitrosamines, aromatic amines, aldehydes, and phenols, as well as organic and inorganic compounds (Wogan *et al*., 2004). Among these classes of chemicals, two stand out as major concerns in cancer susceptibility, PAHs and nitrosamines (Wang *et al*., 2003; Wogan *et al*., 2004). For this research we decided to continue to use the tobacco-specific nitrosamine NNK as a model for tobacco carcinogen.

NNK metabolism and induced genetic damage

NNK can be metabolized by five different pathways; carboxyl reduction, pyridine oxidation, α-hydroxylation, denitrosation, and ADP adduct formation (Hecht, 1998). Of these pathways only α-hydroxylation produces reactive intermediates that result in the formation of bulky DNA adducts that are repaired by NER. This α-hydroxylation pathway will be the only one discussed further. Metabolic activation of NNK by α-hydroxylation forms two distinct reactive intermediates, 4-(3-pyridyl)-4-oxobutane-1-diazohydroxide and methane diazohydroxide. α-Methylene hydroxylation, which produces the reactive intermediate methane diazohydroxide, and α-methyl hydroxylation, which produces the reactive intermediate 4-(3-pyridyl)-4-oxobutane-1-diazohydroxide

(Figure 13), are by-products of NNK metabolism. The cytochrome

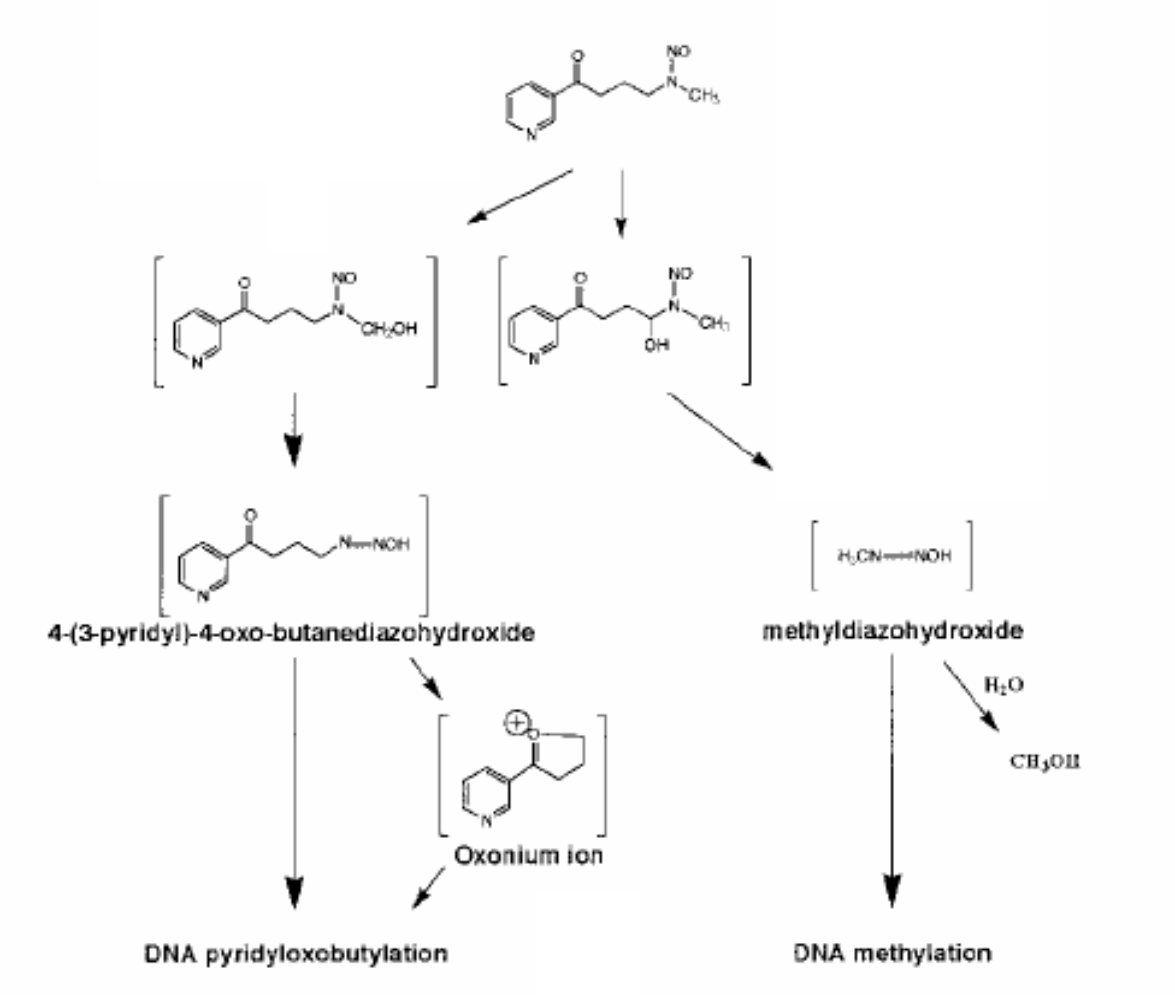


Figure 13. The α-hydroxylation metabolic pathway of NNK
On the right α-methylene hydroxylation that produces the reactive intermediate methane diazohydroxide resulting in DNA methyl adducts. On the left, α-methyl hydroxylation that produces the reactive intermediate 4-(3-pyridyl)-4-oxobutane-1-diazohydroxide resulting in DNA pyridyloxobutyl adducts. Modified from J. Mol. Biol., vol 313, Cloutier JF, Drouin R., M., O'Connor TR., Castonguay A. Characterization and mapping of DNA damage induced by reactive metabolites of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) at nucleotide resolution in human genomic DNA, 539-557, copyright 2001, with permission from Elsevier.

P450s 1A2, 2A1, 2A6 and 2B1 are involved in this metabolism (Hecht, 1998; Fujita and Kamataki, 2001). These two intermediates cause DNA-pyridyloxobutylated and DNA-methylated adducts, respectively (Hecht, 1998). The only characterized pyridyloxobutyl

(POB) adduct from NNK metabolism is POB-O^6-guanine (Cloutier *et al*., 2001) (Figure 14), whereas NNK-induced DNA methylation forms N^7-methylguanine and O^6-methylguanine adducts (Hecht, 1999). These adducts, are produced as a by-product of NNK metabolism by α-hydroxylation, especially the pyridyloxobutyl adducts, are considered bulky adducts.

QUANTITATIVE REAL-TIME REVERSE TRANSCRIPTION (qRT)-PCR

This technology utilizes the amplification of genetic material to accurately quantitate the relative amount of the starting material. This effectively allows a researcher to quantitate the abundance of mRNA of a specific gene. This has the potential to greatly enhance genomic research. Quantitative real-time (RT)-PCR (qRT-PCR) is similar to traditional PCR except for an additional step. With qRT-PCR the starting material is mRNA that must first be reverse transcribed into cDNA (Mocellin *et al*., 2003; Bustin *et al*., 2005; Bustin and Mueller, 2005). The cDNA product is then run through a thermo-cycling series just like regular PCR, except that a fluorescent probe is used along with the primer set to allow for detection as PCR products accumulate. The fluorescent scheme used in the *XPD* mRNA research discussed later is called the extension-phase method (Figure 15). In this scheme the generated cDNA is hybridized with a set of primers and a probe that contains a reporter dye on the 5' end and a quencher on the 3' end. When the probe is intact, the quencher suppresses the fluorescent property of the reporter dye. However, during the

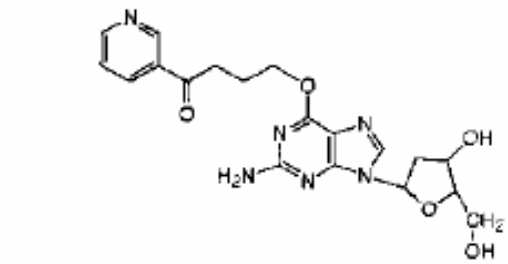


Figure 14. Alkyl-O^6 guanine
The POB-O^6-guanine DNA adduct.

extension phase of PCR the DNA polymerase will displace the

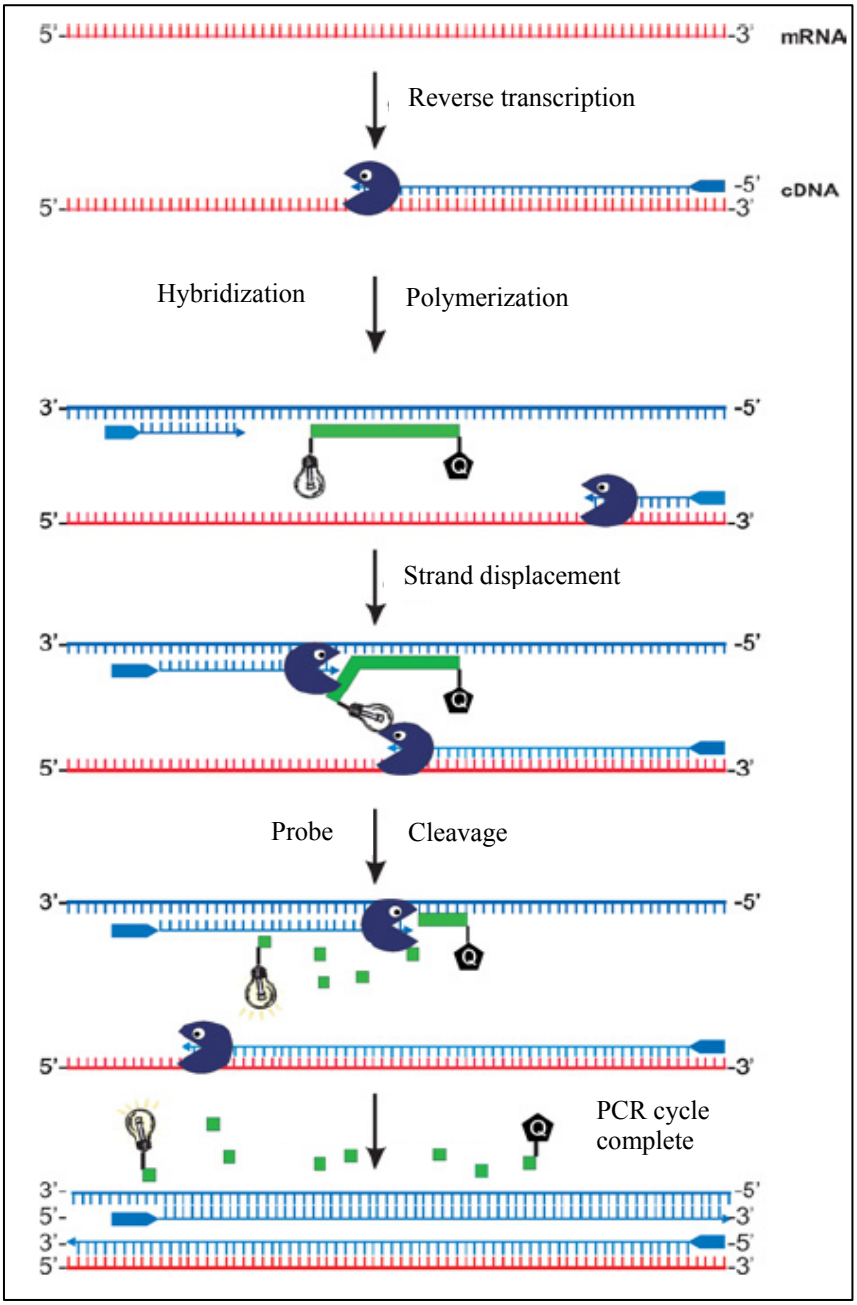


Figure 15. Extension phase fluorescent qRT-PCR scheme
During the RT step the RNA is reverse transcribed into cDNA which is then hybridized with a probe and strand primers. During strand elongation the probe is displaced from the cDNA and the reporter dye is released from the quencher. The amount of fluorescence is proportional to the number of cDNA copies. Reproduced with permission, from Bustin SA., Mueller R. 2005. Clin Sci (Lond). 109: 365-379. © the Biochemical Society

probe starting at the 5' end which contains the reporter dye, at which time it will become fluorescent (Mocellin *et al*., 2003; Bustin and Mueller, 2005).

Quantitation of RT-PCR requires the use of these fluorescent probes. This is done by initially determining the background level of fluorescence and setting a threshold value approximately 100-times greater than the background fluorescence. The number of PCR cycles needed for a sample to reach this threshold value is termed the threshold cycle (Ct), and it is inversely proportional to the amount of starting material (Figure 16). The more starting material that is used the fewer the number of cycles that are required to reach the threshold value. There are two common ways to normalize or quantitate the data. The first method, termed relative quantitation, involves the use of an internal standard, such as a house keeping gene like GAPDH, β-actin, or 18S rRNA. This allows the researcher to normalize the amount of gene of interest to an internal standard that, in theory, should be universally expressed. The second commonly used method is termed absolute quantitation. This method utilizes a standard curve generated from a serial dilution of a vector carrying a copy of the gene of interest. Using the standard curve, a sample Ct value can be correlated back to the initial gene copy number of the sample.

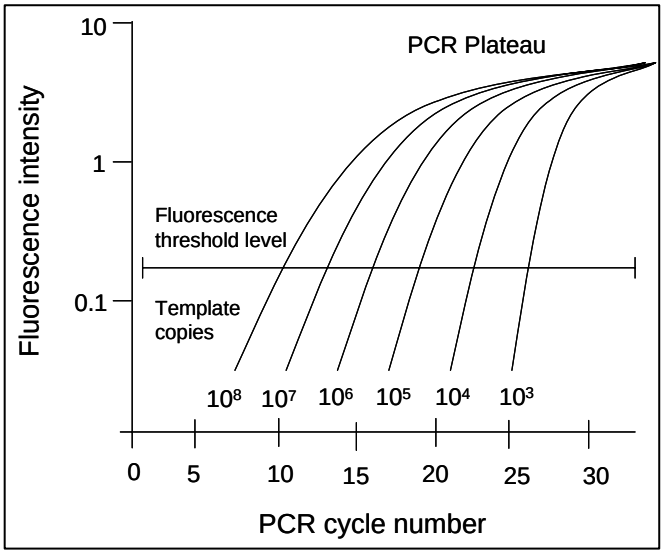


Figure 16. Threshold curve for quantitative real-time RT-PCR
The theory of qRT-PCR, as shown, is that the greater the amount of starting material the fewer PCR cycles that are required to reach the fluorescent threshold level.

Limitations of qRT-PCR

Although this technique is very useful in quantitating gene copy number, like most technologies, it has some inherent pitfalls that must be discussed. In the first step of RT-PCR there are three ways in which reverse transcription may be accomplished. These three ways only differ with regard to the type of primers that are used. One could use a set of random primers, oligo-dT primers, or target-specific primers. Each has its own advantages and disadvantages. The use of random primers results in the priming of multiple origins along every RNA template. The problem is that, since more than one cDNA target is produced per RNA template, genes of low copy number may not be reverse transcribed proportionally, and therefore, its quantitation may not be reliable (Bustin *et al*., 2005). Oligo-dT priming is considered the best method to obtain a good representation of the cDNA. However, this method requires long stretches of RNA and may fail if the RNA is not of good quality (Bustin *et al*., 2005). Thus, it is important to verify the quality of the RNA prior to reverse transcription if using oligo-dTs.

The third method for reverse transcription is the use of target-specific primers and, although they are very specific for a given gene of interest, this is wasteful. The RNA sample can be used only once, and if more experiments must be done additional samples must be taken. This will add another level of variability when trying to quantitate the sample. The PCR reaction itself may introduce some limitations on quantitation of gene expression. For example many biological samples contain substances such as hemoglobin, urea, organic and phenolic compounds, humic acids, and heavy metals, that may inhibit either the RT or PCR steps (Bustin and Mueller, 2005; Bustin *et al*., 2005). However, if care is taken in extracting the RNA sample, absolute quantitation has been shown to produce results that are biologically relevant (Tricarico *et al*., 2002). Lastly, there could exist a problem with normalizing to an internal standard. The use of the above mentioned internal standards is traditional. However, for qRT-PCR it is assumed that these internal standard genes are constituently expressed and do not vary between tissue types. Since this may not be the case, gene quantitation between

different tissues using an internal standard may not be a reliable approach for gene expression quantitation.

Although this technique only provides information as to the total abundance of a gene, and no information is obtained as to the rate of transcription or mRNA degradation, it is, however, useful in the context of this project. It is, first, important to examine whether possible changes in gene abundance occur before addressing whether the rate of transcription or the rate of mRNA degradation is the source of this change in gene abundance. The strength of this technique, its ease of use and rapid data acquisition time, far outway the fact that no information is given as to the half-life of mRNA copies. Those techniques are far more complicated and time consuming to be useful as an initial analysis of gene expression.

FLUORESCENT CONFOCAL MICROSCOPY

The concept involved in confocal microscopy was first theorized by Marvin Minsky in 1955. By focusing light to a single plane in the z-axis, Minsky was able to avoid collecting much of the out of focus light from above and below the ideal focal plane (Minsky, 1988). Most of the confocal microscopes used today image either by collecting the transmitted light off of a specimen or by stimulating the fluorescence from a conjugated dye or fluorophore (Semwogerere and Weeks, 2005). Although microscopes have evolved over the years, most of the advances in microscopy have been focused on improving the signal to noise ratio (Lichtman and Conchello, 2005). The strengths of the confocal microscope are: First, the design of the scope itself, which uses a confocal pinhole to ensure that a single plane about the z-axis is imaged. This pinhole restricts the amount of light that is detected by a photomultiplier tube. By removing from the captured image the information that is increasingly out of focus above and below the z plane, the signal to noise ratio is vastly improved thus increasing the quality of the information obtained for each "Z-section". Second, by changing the focal plane up and down, additional high signal-to-noise ratio planes (Z-sections) can be imaged and subsequently combined to generate a stack of Z-sections or 3-D images. Finally, since

the stack of Z-sections is information-rich and the information is of a high quality, numerous evaluations, both qualitative and quantitative, and manipulations of the images may be undertaken with modern image processing and evaluation software based on validated statistical algorithms.

A weakness of confocal fluorescent microscopy is that it is relatively slow and expensive. However, for the purposes of this research, its sensitivity and the quantitative data that can be obtained from the information-rich images significantly outweigh the weaknesses.

Principles of fluorescence

The principle of fluorescent microscopy requires that an object or specimen fluoresce. Although many biological samples auto-fluoresce, that is they have an innate weak fluorescence, in most cases strongly fluorescent dyes are required. There are typically two ways in which fluorescent dyes are used. The first way is to conjugate, or chemically bond the dye to an antibody that is then incubated with the specimen and then bonds to the protein specific to that antibody. The second way is to create genetically engineered proteins that, when translated, contain a fluorescent dye, such as the green fluorescent protein (GFP), which is tagged to either the N- or C-terminal. In either case the results are the same; the fluorescent signal is imaged by the microscope, thus increasing the signal to noise ratio. The unique property of fluorophores that allows them to fluoresce is that they emit light of a longer wavelength than the one they absorb. So by blocking the wavelength of the absorbed light and collecting the emitted light, a signal specific to the desired fluoraphore can be obtained (Lichtman and Conchello, 2005). All fluorophores in their relaxed state are referred to as being in their ground state, meaning that they are able to absorb light energy, or photons. If the energy absorbed is enough, an electron in the outer orbital of the molecule may move to an orbit that is further away from the nucleus, and thus, the molecule is now referred to as being in an excited state (Lichtman and Conchello, 2005). The property of fluorescence comes when the molecule transitions back to the ground state. Most of the useful fluorophores will transition back

to the ground state in a matter of nanoseconds. This loss of energy is typically in the form of fluorescence, giving off light of longer wavelength than what was needed to excite the molecule in the first place. Most fluorescent compounds take advantage of conjugated double bonds, such as pi bonds found in ring structures. These characteristics allow the molecule to be easily excited by low energy photons, and the wave length difference between the exciting light and the emitted light is small. This translates to more red in the visible spectrum of the emitted light (Lichtman and Conchello, 2005). For example, a molecular compound created by Molecular Probes Inc. (Eugene, OR) known as Alexa Fluor 594 (Figure 17) has many double-bonded ring structures and an excitation wavelength of 590 nm and an emission wavelength of 617 nm. The effectiveness of a fluorophore, which is the probability that the molecule will absorb light and produce a fluorescent signal, is referred to as its molar extinction coefficient. Most useful fluorophores have an extinction coefficient between 25,000 and 200,000 (Lichtman and Conchello, 2005). The molar extinction coefficient of Alexa Fluor 594 is 73,000. The higher the extinction coefficient, the lower the excitation light energy needs to be.

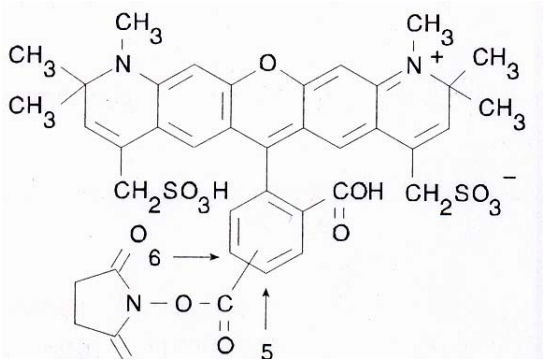


Figure 17. Alexa Fluor 594

STUDY RATIONALE AND HYPOTHESIS

Tobacco smoking is one of the major environmental causes of cancer. In fact, tobacco smoke contains over 40 known carcinogens, and has been shown to be a major contributor to increased cancer risk, most notably oral cancer and lung cancer. Following heart disease, lung cancer has the second highest mortality rate in the US, with an estimated 174,470 new cases and 162,460 deaths projected for 2006 (American

Cancer Society, 2006). Although exposure to tobacco smoke is believed to be an important cause of cancer, only a fraction of smokers (about 10%) actually develop the disease. This suggests that genetic as well as other host factors may modulate susceptibility to tobacco carcinogenesis. SNPs in DNA repair genes can plausibly contribute to the well-documented interindividual susceptibility to tobacco-related cancer. These SNPs can affect the structure and/or function of the resulting DNA repair proteins, and thus can alter DNA repair capacity. Consequently, these inherited traits may significantly influence the level of genetic damage resulting from exposure to mutagens in tobacco smoke, which is an early and critical step in the cascade of events leading to cancer.

Recent molecular epidemiological studies have linked several SNPs in DNA repair genes to a higher risk of smoking-related cancers. For example, SNPs in the *XPD* gene, involved in the nucleotide excision repair pathway (NER), were reported to be associated with increased risk of head and neck and lung cancers. However, the functional significance (i.e. the effect of these SNPs on the structure and functions of the expressed XPD protein) has not yet been characterized. In addition, the biological significance of these SNPs (i.e. their effect on the levels and types of genetic damage induced by specific tobacco smoke carcinogens) has not been fully elucidated. The central hypothesis tested in the current project is that SNPs in the *XPD* gene alter the structure and function of the translated XPD protein resulting in altered NER capacity. This, in turn, will result in an increased accumulation of genetic damage in response to exposure to environmental carcinogens. This hypothesis is tested in the current project using both *in vivo* and *in vitro* approaches. Both classical and FISH cytogenetics are used to address the relationship between polymorphisms in *XPD* and smoking. Reverse transcription real-time PCR is utilized to examine the effects of genetic polymorphisms on *XPD* gene expression. Lastly, confocal microscopy is used to study the effect of *XPD* polymorphisms on XPD protein expression, and cellular translocation.

Materials and Methods

STUDY PARTICIPANTS

This study involved a subset of a cohort composed of 800 healthy volunteers who were recruited, regardless of age, sex or ethnicity, from the staff and students attending the University of Texas Medical Branch (UTMB) in Galveston, Texas and surrounding areas. These individuals had responded to posted notices and advertisements requesting volunteers for genetic studies of smokers and non-smokers. Non-smokers were defined as individuals who had smoked less than 100 cigarettes during their lifetime. Smokers were defined as current smokers who have been smoking at least 10 cigarettes per day for at least three years prior to enrollment in the study. All volunteers were residents of the Houston-Galveston metropolitan area. The study followed a protocol approved by the UTMB Institutional Review Board. Participants were asked to sign a consent form and to fill out a questionnaire that provided background demographic, occupational, personal and medical information. The questionnaire also included detailed information on smoking-related behavior and tobacco usage (e.g., number of cigarettes smoked per day; brand; use of smokeless tobacco, pipes or cigars; former tobacco use; and duration of smoking).

Criteria for exclusion of volunteers included current or recent acute bacterial or viral infections or major chronic illnesses (such as present or previous cancer, or autoimmune diseases), recent blood transfusion, or medical treatment with potentially mutagenic agents (such as chemotherapeutic drugs or radiotherapy). Exclusion criteria also included excessive consumption of alcoholic beverages, and employment (present or past) involving exposure to toxic or mutagenic chemicals (such as benzene, nitrosamines, and polycyclic aromatic hydrocarbons) or radiation. After informed consent was obtained, a blood sample (50-70 ml) was collected from each volunteer. A portion of the blood was used for the cytogenetic cultures and an additional portion of blood was used to obtain DNA for genotype analysis. The remaining portion was used for lymphocyte

isolation and cryopreservation and stored in liquid nitrogen for use in RNA isolation and mRNA quantitation using qRT-PCR.

From this cohort, 290 subjects were genotyped for *XPD* and used for the analysis of classical cytogenetics. Cells from these individuals were randomly chosen from the entire sample population for classical cytogentic analysis. A subpopulation of 38 subjects out of the 290 were chosen for fluorescent *in situ* hybridization (FISH) analysis, based on genotype results to obtain samples that represented all XPD haplotype groups. From the study population (n=800), a subset of 110 individuals was selected for mRNA analysis; similar to those used for the FISH analysis, based on their previously predetermined *XPD* genotypes. Blood samples from fourteen female study participants, chosen because of specific *XPD* genotypes were used for experiments involving exposure to UV radiation for protein quantitation and localization experiments.

XPD GENOTYPE ANALYSIS

The genotype analysis of the C→A polymorphism at codon 156 (C156A) of the *XPD* gene was determined by PCR assay as described by Dybdahl *et al*. (1999). Briefly, PCR primers flanking codon 156 were used to generate a 652-bp product containing the polymorphic site. The PCR products were subjected to restriction digestion for 16 hrs at 65^{o}C by *Tfi*I (New England Biolabs, Beverly, MA), which recognizes the variant allele, and resolved on 2% high-resolution NuSieve® agarose gels (FMC Bioproducts, Rockland, ME). The wild-type C allele is identified by the presence of a band at 596-bp (indicative of the absence of the *Tfi*I cutting site), while the variant A allele is identified by the presence of a band at 482- and 114-bp. In addition to these bands, a 56-bp band, resulting from an additional invariant cutting site for *Tfi*I in the 652-bp amplified fragment, is always present and serves as an internal control to insure complete *Tfi*I digestion.

Genotype analysis of the G→A polymorphism in codon 312 of the *XPD* gene (Asp312Asn) was determined in genomic DNA by PCR assay as described previously (Spitz *et al*., 2001; Affatato *et al*., 2004). Briefly, PCR primers flanking codon 312 were

used to generate a 744-bp product containing the polymorphic site. The PCR product was then subjected to restriction digestion by *Sty*I (New England Biolabs, Beverly, MA), which recognizes the variant allele. The digestion product was then resolved on 2% high-resolution NuSieve® agarose gels (BMA products, Rockland, ME). The wild-type G allele is identified by the presence of a band at 507-bp (indicative of the absence of the *Sty*I cutting site), while the variant allele is identified by the presence of a band at 474-bp. In addition to these bands, a 244-bp band, resulting from an additional invariant cutting site is always present and serves as an internal control for complete *Sty*I digestion.

Genotype analysis of the A→C polymorphism in codon 751 (Lys751Gln) of the *XPD* gene was determined in genomic DNA by a PCR assay as described by Dybdahl *et al*. (1999). Briefly, PCR primers flanking codon 751 were used to generate a 324-bp product containing the polymorphic site. The PCR products were then subjected to restriction digestion by *Pst*I (New England Biolabs, Beverly, MA), which recognizes the variant allele. The digestion product was then resolved on 2% high-resolution NuSieve® agarose gels (BMA products, Rockland, ME). The wild-type C allele is identified by the presence of a band at 224-bp (indicative of the absence of the *Pst*I cutting site), while the variant allele is identified by the presence of a band at 158- and 66-bp. In addition to these bands, a 100-bp band, resulting from an additional invariant cutting site for *Pst*I in the 324-bp amplified fragment, is always present and serves as an internal control for complete *Pst*I digestion.

CYTOGENETIC ASSAYS

Cytogenetic cultures were set up and harvested following standard cytogenetic procedures as described in earlier publications (Abdel-Rahman *et al*., 1998; El-Zein *et al*., 2000; Affatato *et al*., 20040. Briefly, 1 ml of blood collected on the same day was added in a 15 ml tubes to 9 ml of incomplete medium consisting of RPMI 1649 containing 10% fetal calf serum, 90 units/ml penicillin G, and 90 ug/ml streptomycin.

Then 100 ul of 200 mM L-glutamine and 200 ul (12.5 ug/ml) phytohaemagglutinin (PHA) (Murex Diagnostics, Norcross, GA) was added. The cultures were mixed well and incubated, with tubes horizontal, at 37°C. Cultures were allowed to incubate for 48 hr before colcemid is added for 1 hr to arrest the cell cycle. The samples were then processed for Cytogenetic analysis.

For the mutagen sensitivity assay samples were taken as described above in duplicate (Figure 18). Following the initial 40 hr incubation time this samples were treat for 1 hr with 0.24 mM NNK. Forty-six hours following stimulation, cells were separated from the culture fluids by centrifugation at 200 x g, and the supernatant fluids were reserved. The PBLs were resuspended in 5 ml of RPMI-1640 medium, without serum supplement, containing 0.24 mM NNK (National Cancer Institute, Midwest Carcinogen Repository) and incubated at 37°C for 1 hour.

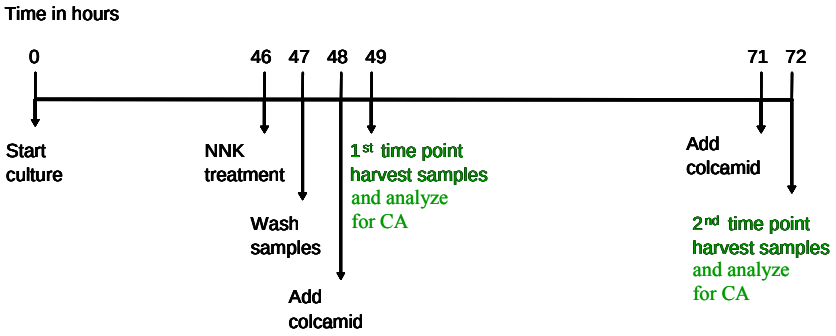


Figure 18. Time-line for the treatment of cultures with NNK
Whole blood samples were cultured in triplicate and initially cultured for 46 hrs then challenged with NNK or sham-challenged with RPMI for hr. Following NNK challenge they were washed and treated with colcemid for 1hr to arrest the cell cycle. The sham-challenged and one NNK challenged sample were harvested 2 hrs after challenge and processed for CAs. The third sample was harvested 24 hrs after challenge and processed for CAs.

Following NNK treatment, the cells were washed twice in serum-free RPMI-1640 and transferred into clean tubes. Following centrifugation at 200 x g, the cells were then resuspended in the reserved growth medium until harvested at 2 or 24 hrs after treatment. For each subject, a non-treated control culture was prepared to provide an *in vivo* baseline (background) level of CAs as discussed above. For the remainder of this discussion, samples will be referred to as *in vivo* or baseline if no NNK treatment was carried out, 2 hrs following NNK treatment for those samples that were harvested at 49 hrs, and 24 hrs following NNK treatment for those samples harvested at 72 hrs (Figure 18).

Conventional and FISH cytogenetics:

One of the advantages of using cytogenetics as an endpoint is the assortment of available assays. A large data set can be generated from the same initial culture by determining damage according to conventional cytogenetics, as well as by FISH to target specific sites. Therefore, both conventional and FISH assays were used in the current study to increase the sensitivity for detecting genetic damage. Preparation of cells for cytogenetic analysis was performed as follows: One hour prior to harvesting, colcemid was added to the baseline, 2 hr and 24 hr hour cultures, the samples were then spun at 200 x g for 10 min, the supernatant was removed and the cells were resuspended and incubated for 30 min at 37^{o}C in 10 ml of hypotonic solution containing potassium chloride. Then 1 ml of fixative (75% methanol and 25% acetic acid) was added to the samples, mixed and then spun at 200 x g for 10 min. Following centrifugation the supernatant was removed and the pellet resuspended in 10 ml of fixative. The sample was then placed at 4^{o}C for at least 15 min, then spun down and resuspended in 10 ml fixative, and stored at 4^{o}C until processed. For long term storage, the samples were centrifuged one more time and resuspended in 1 ml of fixative and stored at -80^{o}C until spread on slides for cytogenetic and/or FISH analysis.

For conventional cytogenetics: Slides were prepared from fixed cells and stained with Giemsa. This solid staining method allows the identification of chromosome aberrations such as dicentrics, breaks, gaps and exchanges. For conventional cytogenetic assays, 50 cells were scored from coded slides. Slides were scored by Ms. Mirtha S. Lopez in Dr. El-Zein's laboratory at the MD Anderson Cancer Center in Houston, TX. The adequacy and the reliability of scoring 50 metaphases in the conventional cytogenetic assay are well documented (Lee *et al*., 1996).

For the FISH assay: Slides were prepared from fixed lymphocytes by adding two drops of the sample to the middle of a non-frosted fluorescence microscope slide and the slide was then placed on a 42^{o}C slide warmer to dry. Each slide was then aged in 2xSSC at 37^{o}C for 30 min, and then washed in 70%, 85%, and 100% ice-cold ethanol for 2 min

each. The slides were washed in a 70% formamide/ 2X SSC solution at 73°C for 5 min, and then washed in 70%, 85%, 100% room temperature ethanol for 1 min each. During this time the two FISH probes from Cytocell that hybridize to the regions chromosome 1q12 and chromosome 3p11.1-q11.1 were prepared by mixing 3 ul of each probe with 4 ul of hybridization solution. The probes were then denatured at 73°C for 5 min. Then 10 ul of probe solution was added to each slide, a cover-slip was added and sealed with rubber cement. The slides were then placed in a humidified hybridization oven overnight at 42°C. The next day the slides were removed from the hybridization oven, the cover-slips removed and the slides were washed three times in 55% formamide at 42°C for 10 min, then 2xSSC for 10 min and finally in 2xSSC/0.1% NP-40 for 5 min. Following the final wash the slides were allowed to dry and 10 ul of antifade DAPI and a cover-slip was added. The slides were then stored at -80°C until read. Chromosome strand breaks and stretches in the probed region of chromosome 1 and 3 were recorded using a Nikon fluorescent scope equipped with a triple bypass filter.

REAL-TIME QUANTITATIVE PCR (qRT-PCR) FOR *XPD* GENE EXPRESSION ANALYSIS

The procedure for *XPD* gene expression analysis and RT-PCR is described in detail in our publication (Wolfe *et al.*, in press). Total RNA was isolated from frozen lymphocytes previously collected from the study participants using a microisolation method (RNAqueous® Kit, Ambion, Austin, TX). Total isolated RNA was then assessed for both quality and quantity on an Agilent 2100 Bioanalyzer. A primer and TaqMan® probe set was designed from the published cDNA sequence for *XPD* (GenBank Accession NM_000400). These 2 primers and the TaqMan® minor groove binding (MGB) probe were designed using the Applied Biosystems Primer Express program (ABI) to specifically target the *XPD* transcript. To estimate the expression of *XPD* in our samples, rqPCR was carried out using the ABI Universal Master Mix on an ABI PRISM® 7000 Sequence Detection System (SDS). To reliably estimate *XPD* expression, Absolute quantitation was used on the resulting threshold cycle (C_T) values generated on

the SDS. Absolute quantitation was carried out by establishing a standard curve of *XPD* expression from a known starting amount of *XPD* cDNA from a vector carrying an *XPD* invert. Relative quantitation, utilizing both human β Actin and 18S rRNA, was used on a limited number of replicated samples as a control to determine the effect of UV exposure on normal housekeeping genes. The Recombinant DNA Core at the University of Texas Medical Branch was instrumental in performing both the RT-PCR reactions and generated the XPD expression vector and standard curve for absolute quantitation.

PREDICTION OF mRNA SECONDARY STRUCTURE

To evaluate the effect of *XPD* polymorphisms on mRNA structure, the mRNA secondary structure of *XPD* (GenBank NM_000400) was predicted using Mfold version 3.2 (Vogel *et al*., 2000) as described in detail in our publication Wolfe *et al*., (in press). This web-server program is an internet-based RNA folding program (http://www.bioinfo.rpi.edu/applications/mfold/old/rna/) (Zuker 2003; Mathews *et al*., 1999). Partial *XPD* mRNA sequences were entered as 141 nucleotide sequences including flanking sequences (70 nucleotides) on either side of the *XPD* Arg156Arg, the Asp312Asn and the Lys751Gln polymorphisms. The corresponding wild-type mRNA sequence was used as a control for comparison with the respective polymorphic structures studied. For each wild-type and polymorphic sequence, the amount of free energy required for folding (dG) and the average amount of single stranded RNA was calculated. Qualitative assessments of the folded structures were also made.

CELL CULTURES FOR CONFOCAL MICROSCOPIC DETERMINATION OF XPD PROTEIN CONCENTRATION AND LOCALIZATION

Human lymphocytes were isolated from fresh whole blood collected from study participants as follows: Whole blood was layered onto Histopaque-1077 for gradient separation by centrifugation at 400 x g for 30 min. The middle lymphocyte layer was then transferred to a new tube and washed with PBS. The washed pellet was then

resuspended in 30 ml of RPMI and $15x10^6$ cells per sample were then resuspended in 20 ml of stimulation media (68.5% RPMI, 20.0% HL-1, 10.0% FBS, 1.0% L-glutamine, and 0.5% PHA-16) and incubated at 37°C for 40 hrs.

UV-exposure and cell culture

Confocal microscopy was used for the quantatition and nuclear translocation of the XPD protein in cultured isolated lymphocytes after UVC irradiation. Following a 40 hr incubation at 37°C as discussed previously, lymphocytes were resuspended in 5 ml of PBS and a 4 ml aliquot from each sample was transferred to a 100x20 mm Petri dish. UV irradiation was preformed as described by Albrecht et al. (1974). Briefly, a General Electric G8T5 'germicidal' bulb emitting light at 254nm was used and the distance from the light source was controlled to deliver a dose rate of UVC irradiation of 400 $\mu J/cm^2$. The dose rate was measured by a Blak-ray ultra-violet intensity meter (Ultra-violet Products Inc. San Gabriel, California). One ml was set aside before UV irradiation to serve as a non-treated control.

Following exposure, each treated aliquot was washed in PBS and resuspended in 8 ml of stimulation media (69% RPMI, 20.0% HL-1, 10.0% FBS, and 1.0% L-glutamine), the non-treated control sample was washed in PBS and resuspended in 2 ml of stimulation media, and both aliquots were then cultured at 37°C. Forty µl from each sample was removed at 30 min, 2 hrs, 4 hrs and 6 hrs following UV exposure. The non-treated control was processed along with the 30 min treated aliquots of cells. The 40 µl aliquots, representing the four different time points after treatment and the non-treated control, were each spread onto a 12 mm circular coverslip and allowed to dry. The coverslips were then placed in separate wells of a 24-well plate and fixed with formalin overnight at 4°C. Following fixation, the samples were washed with PBS; incubated with XPD C-20 a primary goat polyclonal antibody (Santa Cruz Biotech, CA), at a dilution of 1:75 for 2 hrs; and then, following a second wash, the samples were incubated for 2 hrs with 10ug/ul of secondary chicken anti-goat IgG labeled with Alexa Fluor 594 with fluorescence in the red spectrum (Molecular Probes, OR). The samples were then

washed twice with PBS and once with PBS/0.01% tween 20 and allowed to dry. A drop of DAPI stain was added and the coverslip was then inverted onto a slide and sealed.

Fluorescent confocal microscopy for the determination of XPD protein concentration and cellular localization

XPD protein quantitation and localization was performed at the Optical Imaging Core Facility at UTMB, using a Zeiss LSM 510 UV META laser scanning confocal fluorescent microscope with a 63x/1.4 water immersion planapochromat objective. Fixed lymphocytes stained with DAPI were visualized using UV excitation (351-364 nm) and a 385-470 nm band-pass emission filter. XPD protein labeled with Alexa Fluor 594 was detected using excitation at 543 nm and a 560 nm long-pass emission filter. An average of 100 cells per time-point per sample was imaged using the LSM510 software. Actual protein quantitation and localization was later performed on these images using Metamorph imaging software (Molecular Devices, PA) for protein quantitation or the LSM 510 software for protein localization.

STATISTICAL ANALYSIS

For the polymorphisms evaluated in this study, heterozygous individuals were combined with homozygous variant individuals to increase statistical power. Age classification was achieved by stratifying the population into two groups, using the median age as the cutoff, as was done in previous studies (Spitz *et al*., 2001; Affatato *et al*., 2004). The median, rather than the mean, was used to account for the influence of outliers. The students t-test was used to compare the mean chromosome aberration frequency between groups, the mean *XPD* gene expression levels between groups, and the mean protein expression levels between groups. A two-sided P value less than 0.05 was used as the criterion for significance. All statistical analyses were performed using the SPSS 14.0 software program.

Results

CHROMOSOME ABERRATION ANALYSIS

The working hypothesis of this section was that individuals with the variant *XPD* alleles are more likely to exhibit phenotypes reflected by increased frequencies of CAs. The goal was to delineate the effect of the Arg156Arg, Asp312Asn and Lys751Gln SNPs in the *XPD* gene on: (1) *in vivo* baseline levels of CAs and, (2) the sensitivity to NNK using the mutagen sensitivity assay. The approach involved the collection of blood samples and a detailed risk factor questionnaire from study participants. These blood samples were used for the following procedures:

1. Genotype analysis for SNPs in the *XPD* gene, using PCR-based methodologies, on DNA isolated from PBLs from all study subjects who participated in this study.
2. Determination of CA using conventional cytogenetic techniques and the FISH assay in cultured PBLs from all study participants and evaluation genotype effects on CA frequencies
3. Challenge of PBLs with NNK followed by cytogenetic analysis for the evaluation of the influence of XPD genotypes on NNK-induced CA.

The CAs were correlated with the presence of *XPD* variant alleles and with other important information obtained from the questionnaire (smoking intensity and duration, age and sex), as well as other genotype data available to evaluate the effect of these parameters on the in vivo baseline levels of genetic damage observed. Based on previous studies (Affatato *et al*., 2004) a power calculation conducted revealed that a total sample size of 178 subjects would be adequate to give a power of 0.80 at 95% confidence and allow the detection of an effect size of 0.375 (Faul and Erdfelder 1992).

Demographics of the population studied for the CA evaluation

The population used for the evaluation of the effect of *XPD* polymorphisms on levels of CA frequencies was composed of 290 individuals. There were 88 males (30.3%)

and 202 females (69.7%). The majority (n=196), were White non-Hispanics (67.6%). There were also 31 African-American subjects (10.7%), 41 Hispanics (14.1%), and 20 Asians (6.9%). Of the 290 volunteers who participated in this study, 90 (31%) were smokers, 56 (19.3%) were ex-smokers, and 144 (49.1%) were non-smokers. Non-smokers were defined as individuals who had smoked less than 100 cigarettes during their lifetime. Smokers were defined as current smokers who had been smoking at least 10 cigarettes per day for at least three years prior to enrollment in the study. Ex-smokers were defined as individuals who had quit smoking for at least 1 yr prior to enrollment into this study. For this study and for statistical analysis smokers and ex-smokers were analyzed both as a combined group and as separate groups. For the smokers, the mean ± S.D number of accumulated pack years (defined as cigarettes smoked per day multiplied by the number of years of smoking divided by 20) was 18.1 ± 2.2, for ex-smokers the mean accumulated pack years was 20.6 ± 3.5, and the number of accumulated pack years for the combined group of smokers and ex-smokers was 19.1 ± 1.9. The median number of pack years was 11. This median value was used to categorize smokers and ex-smokers into heavy smokers (>11 pack years) and light smokers (≤11 pack years) for subsequent stratified analysis for the effect of smoking on chromosomal aberrations. The age of the participants ranged from 19 to 84 years, with a median of 39 years. The study subjects were divided into two age groups based on the population median: younger individuals (≤39 years old) and older individuals (>39 years old) for subsequent stratified analysis for the effect of age on induction of chromosomal aberrations.

In vivo and NNK-induced CA frequencies

Classical cytogenetics

The mean values (± SE) of the *in vivo* (baseline) and the NNK-induced CA frequencies are presented in Table III. There was no significant difference in baseline *in vivo* CA frequencies between non-smokers vs. smokers or ex-smokers and between older

(>39 years) compared to younger (<39 years) individuals. In addition, the difference in CA frequencies was not statistically significant when comparing the CA frequencies among each of the groups and the levels observed in all the study subjects. In all cases the P value was greater than 0.05. The same observation was also true when lymphocytes were treated with NNK, except a statistically significant increase in CA was observed among ex-smokers 24 hrs following treatment compared to the same time point in non-smokers (P=0.025). The frequencies of CA observed among the various strata (*i.e.* smokers vs. non smokers and older vs. younger individuals) were not significantly different from each other and no significant difference was observed between the frequencies of breaks among these groups compared to the frequencies observed in all study subjects. In all cases the P value was greater than 0.05.

Table III. In vivo and NNK-induced CA frequencies

No. of subjects	In vivo (baseline) CA frequency (mean ± SE)	NNK -induced CA frequency (mean ± SE) 2hr following treatment	NNK -induced CA frequency (mean ± SE) 24hr following treatment
All subjects (n=290)	0.75 ± 0.06	2.61 ± 0.11	4.71 ± 0.14
Stratified by sex			
male (n=88)	0.80 ± 0.12	2.33 ± 0.18	4.69 ± 0.29
female (n=202)	0.72 ± 0.07	2.73 ± 0.15	4.72 ± 0.16
Stratified by smoking			
non-smokers (n=144)	0.68 ± 0.08	2.44 ± 0.16	4.47 ± 0.19
smokers (n=90)	0.69 ± 0.10	2.62 ± 0.21	4.72 ± 0.29
ex-smoker (n=56)	1.00 ± 0.17	3.02 ± 0.25	5.29 ± 0.30*
combined smokers (n=146) and ex-smokers	0.81 ± 0.09	2.77 ± 0.16	4.94 ± 0.21
Stratified by ethnicity			
Caucasian (n=196)	0.75 ± 0.08	2.75 ± 0.14	4.70 ± 0.18
Afican american (n=31	0.61± 0.14	2.52 ± 0.46	5.37 ± 0.52
Asian (n=20)	0.47 ± 0.16	2.16 ± 0.42	4.74 ± 0.61
Hispanic (n=41)	0.90 ± 0.15	2.26 ± 0.21	4.27 ± 0.30
Stratified by age			
>39 years old (n=140)	0.74 ± 0.07	2.80 ± 0.18	4.99 ± 0.22
<39 years old (n=150)	0.75 ± 0.10	2.43 ± 0.14	4.44 ± 0.18

* Significantly different relative to non-smokers, p=0.025

FISH analysis

Fluorescent In Situ Hybridization was performed on a subset composed of 38 samples by scoring 1000 cells per-subject for chromosome breaks using chromosome specific probes regions of chromosomes 1q12 and chromosome 3p11.1-q11.1 (Figure 19). The mean values (± SE) of the *in vivo* and the NNK-induced CA frequencies for both chromosome1 and chromosome 3 showed no significant difference in baseline *in vivo* CA frequencies between non-smokers (n=21) vs. smokers (combined smokers and ex-smokers n=17) and between older (>39 years, n=21) vs. younger (<39 years, n=17) individuals or between males (n=7) and females (n=31) (Table IV). In addition, the difference in CA frequencies was not statistically significant when comparing the CA frequencies among each of the groups and the levels observed in all the study subjects. In all cases the P value was greater than 0.05. The same observation was also true when lymphocytes were treated with NNK (Table IV). The

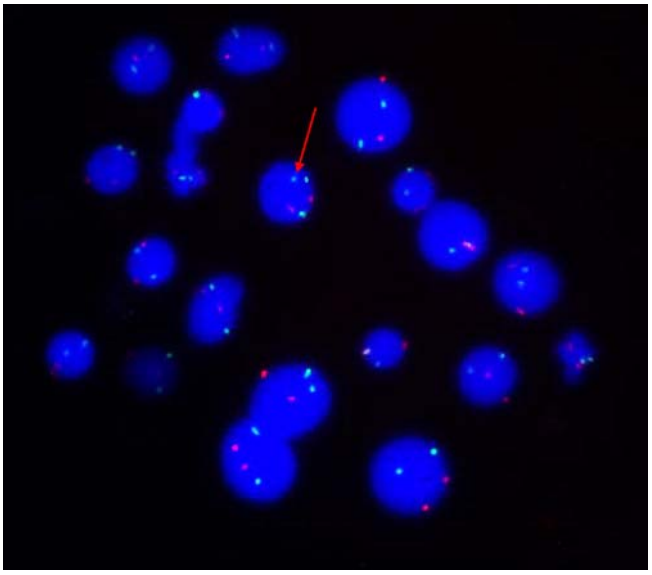


Figure 19. FISH analysis
Microscopic image of a FISH slide when the DAPI stained nuclei of lymphocytes are probed for chromosome 1 (green) and chromosome 3 (red). The presence of a chromosome break in chromosome 1 is indicated by the red arrow.

Table IV. In vivo and NNK-induced CA frequencies using FISH

	Baseline in vivo CA		NNK-induced CA	
No. of subjects	Chromosome 1	Chromosome 3	Chromosome 1	Chromosome 3
All subjects (n=38)	2.52 ± 0.19	2.71 ± 0.22	3.34 ± 0.18	4.57 ± 0.29
Stratified by sex				
male (n=7)	2.57 ± 0.53	2.57 ± 0.37	3.43 ± 0.20	4.86 ± 0.86
female (n=31)	2.51 ± 0.22	2.74 ± 0.26	3.32 ± 0.21	4.52 ± 0.31
Stratified by smoking				
non-smokers (n=21)	2.71 ± 0.33	2.93 ± 0.31	3.38 ± 0.28	5.05 ± 0.38
combined smokers (n=17) and ex-smokers	2.29 ± 0.17	2.41 ± 0.30	3.29 ± 0.20	4.00 ± 0.43
Stratified by age				
>39 years old (n=21)	2.58 ± 0.21	2.95 ± 0.38	3.42 ± 0.24	4.74 ± 0.46
<39 years old (n=17)	2.47 ± 0.34	2.47 ± 0.22	3.26 ± 0.26	4.42 ± 0.38

frequencies of CA observed among the various strata (*i.e.* smokers vs. non-smokers and older vs. younger individuals) were not significantly different from each other, and no significant difference was observed between the frequencies of breaks among these groups compared to the frequencies observed in all study subjects. In all cases the P value was greater than 0.05. In all cases the frequency of CA was significantly higher 2 hrs following NNK treatment than at baseline.

XPD genotypes among the study population

The distribution of *XPD* exon 6 (codon 156), exon 10 (codon 312) and exon 23 (codon 751) genotypes among the study population is presented in Table V. Overall, no significant difference was observed in the genotypic distribution among all the subjects and the subgroups with respect to any of the three polymorphisms considered. The variant 156A allele frequency in exon 6 was 0.40 for all subjects. This allelic frequency was slightly elevated in smokers (0.45), compared to ex-smokers (0.40), and non-smokers (0.37) as well as in older subjects (>39 years, 0.39) and in younger subjects (<39 years, 0.42). The 156A exon 6 allelic frequency was also similar in males (0.41) and females (0.40). The variant 312Asn allele frequency in exon 10 was 0.25 for all subjects. This frequency was similar in smokers (0.26) and in non-smokers (0.23), but slightly over-represented in ex-smokers (0.30). In older subjects the 312Asn allele frequency was 0.26 and in younger subjects it was 0.23. In males it was 0.23 and in females 0.25. With respect to the variant 751Gln allele in exon 23, the frequency of 0.30 observed in the total population studied was also similar in smokers (0.29), ex-smokers (0.32) and non-smokers (0.30), as well as in older individuals (0.29) and younger individuals (0.28) and in males (0.29) and females (0.30). The allele frequencies are consistent with those previously reported in studies conducted on healthy subjects (Hemminki *et al*., 2001; Spitz *et al*., 2001) indicating that the subjects in this study are representative of the general population.

Table V. XPD genotype distribution among the study population

	Exon 6			Exon10			Exon23		
No. of subjuects	C/C n (%)	C/A n (%)	A/A n (%)	Asp/Asp n (%)	Asp/Asn n (%)	Asn/Asn n (%)	Lys/Lys n (%)	Lys/Gln n (%)	Gln/Gln n (%)
All subjects n=290	107 (36.9)	114 (39.3)	52 (17.9)	142 (49.0)	125 (43.1)	5 (1.7)	139 (47.9)	123 (42.4)	23 (7.9)
Stratified by smoking									
non-smokers (n=144)	56 (38.9)	59 (41.0)	20 (13.9)	73 (50.7)	59 (41.0)	1 (0.7)	70 (48.6)	58 (40.3)	13 (9.0)
smokers (n=90)	30 (33.3)	33 (36.7)	21 (23.3)	46 (51.1)	38 (42.2)	2 (2.2)	44 (48.9)	39 (43.3)	5 (5.5)
ex-smokers (n=56)	21 (37.5)	22 (39.3)	11 (19.6)	23 (41.1)	28 (50.0)	2 (3.6)	25 (44.6)	26 (46.4)	5 (8.9)
combined smokers (n=146) and ex-smokers	51 (34.9)	55 (37.7)	32 (21.9)	69 (47.3)	66 (45.2)	4 (2.7)	69 (47.3)	65 (44.5)	10 (6.8)
Stratified by age									
>39 years old (n=140)	54 (38.6)	52 (37.1)	24 (17.1)	66 (47.1)	60 (42.8)	5 (3.6)	73 (52.1)	51 (36.4)	14 (10.0)
<39 years old (n= 150)	53 (35.3)	62 (34.7)	28 (18.7)	76 (50.7)	65 (43.3)	0 (0.0)	66 (44.0)	72 (48.0)	9 (6.0)
Stratified by sex									
males (n=88)	33 (37.5)	31 (35.2)	18 (20.4)	45 (51.1)	36 (40.9)	1 (1.14)	45 (51.1)	35 (39.8)	7 (7.9)
females (n=202)	74 (36.6)	83 (41.1)	34 (16.8)	97 (48.0)	89 (44.0)	4 (2.0)	94 (46.5)	88 (43.6)	16 (7.9)

Not all genotype categories add up to the total number of subjects because genotype determination was not 100% successful

Effect of *XPD* polymorphisms on *in vivo* CA frequencies

Classical cytogenetics

Evaluation of the *in vivo* baseline CA frequencies in PBLs of all study subjects revealed a range of 0 to 7 breaks/50 cells with a mean of 0.75 breaks/50 cells. Students t-test was used to evaluate the difference between baseline CA frequencies between wild-type individuals and individuals with the variant alleles for the three polymorphisms (exon 6, exon 10, and exon 23) in *XPD*. No significant differences were observed

Table VI. Relationship between XPD genotype and in vivo CAs among smokers compared to non-smokers

	non-smoker (n=144)		smokers (n=90)		ex-smoker (n=56)		combined smoker and ex-smokers (n=146)	
No. of subjects	CA frequency[a]	P[b] value	CA frequency[a]	P[b] value	CA frequency[a]	P[b] value	CA frequency[a]	P[b] value
Stratified by XPD exon 6 C156A polymorphism								
homozygous wild-type (n=107	0.78 ± 0.15	ref.	0.70 ± 0.20	0.75	0.86 ± 0.19	0.78	0.76 ± 0.14	0.93
hetrozygous variant (n=114)	0.61 ± 0.12	ref.	0.63 ± 0.15	0.96	1.34 ± 0.33	**0.015***	0.91 ± 0.17	0.16
homozygous variant (n=52)	0.42 ± 0.14	ref.	0.85 ± 0.22	0.11	0.82 ± 0.38	0.25	0.84 ± 0.19	0.13
Stratified by XPD exon 10 Asp312Asn polymorphism								
homozygous wild-type (n=142)	0.72 ± 0.13	ref.	0.58 ± 0.12	0.47	0.83 ± 0.23	0.70	0.67 ± 0.11	0.77
hetro- or homozygous variant (n=130)	0.58 ± 0.11	ref.	0.82 ± 0.28	0.22	1.17 ± 0.25	**0.015***	0.67 ± 0.15	**0.045***
Stratified by XPD exon 23 Lys751Gln polymorphism								
homozygous wild-type (n=139)	0.87 ± 0.13[C]	ref	0.61 ± 0.12	0.19	1.12 ± 0.33	0.40	0.80 ± 0.15	0.74
hetro- or homozygous variant (n=148)	0.45 ± 0.10	ref.	0.75 ± 0.17	0.12	0.90 ± 0.14	**0.014***	0.81 ± 0.12	**0.029***

a Chromosomal aberrations were reported as mean number of breaks/50 cells ± standard error
b Two-sided P-value (Students t-test)
c Statistically significant compared to homo- or heterozygous variant carriers (p=0.03)
* Statistically significant (P<0.05)

between any of the variant genotypes compared to the corresponding homozygous wild-type genotype when all subjects were analyzed or when combined analysis between two or more polymorphisms were examined. When stratified by smoking, ex-smokers, but not smokers, with the *XPD* 156A polymorphism in exon 6 had a significantly higher base line CA frequency than non-smokers with the same genotype (P=0.015) (Table VI). This observation was also seen among ex-smokers with the *XPD* 312Asn polymorphism in exon 10 compared to non-smokers with the same genotype (P=0.015). This significant difference also held up when ex-smokers and current smokers were combined as one group (P=0.045). The same observation was seen with the *XPD* 751Gln polymorphism

in exon 23 (for ex-smokers P=0.014, and combined smokers and ex-smokers P=0.030) (Table VI). Among non-smokers, *XPD* exon 23 wild-type individuals had significantly more CAs than individuals carrying one or two polymorphic exon 23 alleles (P=0.022). Interestingly, ex-smokers with only one exon 6 polymorphic allele had significantly higher CAs (1.34 ± 0.33) than did current smokers with the same genotype (0.63 ± 0.15, P=0.041) (Table VI).

When stratified by gender, cells from males homozygous wild-type for *XPD* exon 10 had significantly higher CAs frequencies than cells from females with the same genotype (P=0.015) (Table VII). Among the females, those with the XPD 312Asn exon 10 polymorphism had significantly higher CAs frequencies than the wild-type females (P=0.020). No significant differences were seen between any of the polymorphisms when stratified by age.

Table VII. Relationship between XPD genotype and in vivo CA's with respect to gender

	Males (n=88)		Females (n=202)	
No. of subjects	CA frequency [a]	P[b] value	CA frequency [a]	P[b] value
Stratified by XPD exon 6				
C156A polymorphism				
homozygous wild-type (n=107	0.66 ± 0.23	ref.	0.82 ± 0.12	0.46
hetrozygous varaint (n=114)	0.87 ± 0.18	ref.	0.71 ± 0.12	0.51
homozygous variant (n=52)	1.00 ± 0.28	ref.	0.50 ± 0.12	0.07
Stratified by XPD exon 10				
Asp312Asn polymorphism				
homozygous wild-type (n=142)	1.00 ± 0.21	ref.	0.55 ± 0.07	**0.015***
hetro- or homozygous varaint (n=130)	0.55 ± 0.12	ref.	0.89 ± 0.12[c]	0.11
Stratified by XPD exon 23				
Lys751Gln polymorphism				
homozygous wild-type (n=139)	1.02 ± 0.21	ref.	0.75 ± 0.10	0.20
hetrozygous varaint (n=123)	0.64 ± 0.14	ref.	0.67 ± 0.10	0.84
homozygous varaint (n=23)	0.14 ± 0.14	ref.	0.81 ± 0.28	0.14

a Chromosomal aberrations were reported as mean number of breaks/50 cells ± standard error
b Two-sided P-value (Students t-test)
c Statistically significant compared to exon 10 homozygous wild-type carriers (p=0.02)
* Statistically significant (P<0.05)

FISH analysis

Evaluation of the *in vivo* baseline CA frequencies in PBLs of all (n=38) study subjects revealed a range of 0 to 5 breaks/1000 cells with a mean of 2.56 breaks/1000 cells on chromosome 1 and a range of 1 to 6 breaks/1000 cells with a mean of 2.71 breaks/1000 cells on chromosome 3. Comparisons were performed between wild-type individuals and individuals variant for three polymorphisms (exon 6, exon 10, and exon 23) in *XPD*. No significant differences were observed between cells from wild-type individuals (n=9) for the three polymorphisms compared to cells from individuals variant for the three polymorphisms (exon 6, n=6); exon 10, n=13; and exon 23, n=17) in *XPD*. When stratified by smoking (ex-smokers and current smokers combined) cells from individuals who smoked and inherited the 751Gln polymorphism (n=10) had significantly higher chromosome 3 aberration frequencies (3.14 ± 0.51) than did cells from individuals who smoked but carried the homozygous wild-type 751Lys allele (n=7) (1.9 ± 0.28; P=0.035). There were no significant differences when stratified by age. Due to the small number of male subjects (n=7) involved in the FISH study, it was not possible to stratify this subpopulation with regard to gender.

Effect of *XPD* polymorphisms on NNK-induced CA frequencies after 2 hrs

Classical cytogenetics

Using the mutagen-sensitivity approach, 2 hrs after treatment with NNK, CA frequencies were evaluated in lymphocytes of the study subjects with respect to the 156A, the 312Asn and the 715Gln variant alleles of the *XPD* gene. Evaluation of NNK-induced CA frequencies revealed a range of 0 to 11 breaks/50 cells with a mean of 2.61 breaks/50 cells. No significant differences were observed between subjects with the wild-type *XPD* genotype and subjects with variant genotypes when the effect of the three polymorphisms was analyzed alone or in combination with each other. When stratified by smoking no significant differences were observed across the smoking categories or among the different genotypes. When stratified by gender, females with the *XPD* C156 exon 6

wild-type allele (n=74) had significantly higher frequencies of CAs than males with the *XPD* exon 6 wild-type allele (n=33, P=0.046) (Table VIII). Females with the homozygous *XPD* 751Gln exon 23 variant had a significantly higher frequency of CAs than males with the same genotype (P=0.005) (Table VIII). Among the females, those with the *XPD* 751Gln homozygous variant had significantly higher frequencies of CAs than females wild-type for Lys761Gln (P=0.022) and heterozygous 751Gln females (P=0.024). No significant differences were seen among any of the polymorphisms when stratified by age.

Table VIII. Relationship between XPD genotype and NNK-induced CAs with respect to gender 2 hrs after NNK challenge.

	Males (n=88)		Females (n=202)	
No. of subjects	CA frequency [a]	P [b] value	CA frequency [a]	P [b] value
Stratified by XPD exon 6				
C156A polymorphism				
homozygous wild-type (n=107	2.00 ± 0.30	ref.	2.82 ± 0.23	**0.046***
hetrozygous varaint (n=114)	2.87 ± 0.30	ref.	2.62 ± 0.22	0.55
homozygous variant (n=52)	2.19 ± 0.41	ref.	2.62 ± 0.37	0.49
Stratified by XPD exon 10				
Asp312Asn polymorphism				
homozygous wild-type (n=142)	2.43 ± 0.27	ref.	2.51 ± 0.22	0.83
hetro- or homozygous varaint (n=130)	2.22 ± 0.27	ref.	2.90 ± 0.20	0.069
Stratified by XPD exon 23				
Lys751Gln polymorphism				
homozygous wild-type (n=139)	2.48 ± 0.27	ref.	2.63 ± 0.21 [c]	0.68
hetrozygous varaint (n=123)	2.41 ± 0.28	ref.	2.64 ± 0.22 [c]	0.56
homozygous varaint (n=23)	1.14 ± 0.40	ref.	3.94 ± 0.57	**0.005***

a Chromosomal aberrations were reported as mean number of breaks/50 cells ± standard error
b Two-sided P-value (Students t-test)
c Statistically significant compared to female homozygous variant carriers ($p<0.05$)
* Statistically significant ($P<0.05$)

FISH analysis

Using the mutagen-sensitivity approach, 2 hrs after treatment with NNK, CA frequencies were evaluated on chromosomes 1 and 3 in lymphocytes of 38 study subjects

in relation to the inheritance of the 156A, the 312Asn and the 715Gln variant alleles of the *XPD* gene. Evaluation of NNK-induced CA frequencies revealed a range of 1 to 6 breaks/1000 cells with a mean of 3.34 breaks/1000 cells on chromosome 1 and a range of 1 to 8 breaks/1000 cells with a mean of 4.58 breaks/1000 cells on chromosome 3. Cells from individuals who inherited either one or two 312Asn variant alleles had a significantly higher chromosome 3 aberration frequency (5.38 ± 0.51) than did cells from individuals homozygous for the wild-type 312Asp allele (4.16 ± 0.34; P=0.048). Interestingly, the 751Gln polymorphism seems to provide a protective effect. Cells with the XPD 751Gln polymorphisms were found to have significantly lower chromosome 1 aberration frequencies (2.94 ± 0.26) than cells with the homozygous wild-type 751Lys allele (3.67 ± 0.22; P=0.040). This observation was also true when the results were stratified by age, where cells from younger individuals (age≤39) carrying the 751Lys wild-type allele had significantly higher CA frequencies (3.88 ± 0.35) than cells from younger individuals with the 751Gln polymorphism (2.7 ± 0.30; P=0.019). In addition, cells from smokers with the 751Lys wild-type allele had significantly higher CA frequencies (3.7 ± 0.21) on chromosome 1 than cells from smokers with the 751Gln polymorphism (2.71 ± 0.28; P=0.013).

Effect of *XPD* polymorphisms on NNK-induced CA frequencies after 24 hrs using classical cytogenetics

Using the mutagen-sensitivity approach, 24 hrs after treatment with NNK CA frequencies were evaluated in PBLs of the subjects in association with the inheritance of the 156A, the 312Asn and the 715Gln variant alleles of the *XPD* gene. Evaluation of NNK-induced CA frequencies revealed a range of 0 to 12 breaks/50 cells with a mean of 4.71 breaks/50 cells. No significant differences were observed between any of the subjects with the wild-type genotypes and subjects with the variant genotypes when analyzed alone or in combination. When stratified by smoking, non-smoker individuals heterozygous for the exon 6 polymorphism, 156C>A, had significantly lower frequencies of CAs than non-smokers homozygous for the same polymorphism (P=0.046). Both ex-

smokers and smokers (current smokers and ex-smokers) who were heterozygous for the exon 6 polymorphism, had significantly higher CAs frequencies than non-smokers with the same genotype (P=0.005 and 0.011 respectively). Ex-smokers with polymorphic alleles in exon 10 and in exon 23 also showed significantly higher frequencies of CAs than non-smokers with the same genotype (P=0.007 and P=0.02, respectively) (Table IX). No significant differences were observed between males and female with regard to any of the polymorphisms. Likewise, no significant differences were seen between any of the polymorphisms when stratified by age.

Table IX. Relationship between XPD genotype and 24 hour NNK-induced CA's among smokers

	non-smokers (n=144)		smokers (n=90)		ex-smokers (n=56)		combined smokers and ex-smokers (n=146)	
No. of subjects	CA frequency[a]	P[b] value	CA frequency[a]	P[b] value	CA frequency[a]	P[b] value	CA frequency[a]	P[b] value
Stratified by XPD exon 6 C156A polymorphism								
homozygous wild-type (n=107	4.53 ± 0.31	ref.	4.83 ± 0.57	0.61	5.43 ± 0.48	0.13	5.08 ± 0.39	0.27
hetrozygous variant (n=114)	4.03 ± 0.29	ref.	4.91 ± 0.47	0.10	5.67 ± 0.46	**0.005***	5.21 ± 0.34	**0.011***
homozygous variant (n=52)	5.21 ± 0.46 [c]	ref.	4.21 ± 0.56	0.19	4.27 ± 0.77	0.28	4.23 ± 0.44	0.16
Stratified by XPD exon 10 Asp312Asn polymorphism								
homozygous wild-type (n=142)	4.51 ± 0.29	ref.	4.76 ± 0.43	0.93	4.72 ± 0.54	0.66	4.77 ± 0.34	0.59
hetro- or homozygous variant (n=130)	4.39 ± 0.26	ref.	4.65 ± 0.42	0.57	5.63 ± 0.36	**0.007***	5.07 ± 0.30	0.096
Stratified by XPD exon 23 Lys751Gln polymorphism								
homozygous wild-type (n=139)	4.51 ± 0.30	ref.	4.75 ± 0.41	0.65	5.04 ± 0.55	0.39	4.86 ± 0.33	0.45
hetro- or homozygous variant (n=148)	4.44 ± 0.24	ref.	4.79 ± 0.43	0.51	5.50 ± 0.31	**0.02***	5.07 ± 0.30	0.14

a Chromosomal aberrations were reported as mean number of breaks/50 cells ± standard error
b Two-sided P-value (Fishers exact test)
c Statistically significant compared to heterozygous 156A carriers (p=0.046)
* Statistically significant (P<0.05)

Effect of *XPD* haplotype on *in vivo* and NNK-induced CAs using classical cytogenetics

Since the three polymorphisms in this study, C156A, Asp312Asn, and Lys751Gln are at 70-80% linkage disequilibrium (information obtained from the web-based program http://www.chgb.org.cn/lda/lda.htm) (Ding et al., 2003) it was prudent to analyze the associations as haplotypes in a manner that allows the true association of one polymorphism or a combination to be addressed. This was done by grouping the subjects

in such a way that two of the polymorphisms were homozygous wild-types and, therefore, the effect of the third polymorphism could be established.

When stratified by smoking, Figure 20, it was observed that smokers (current smokers and ex-smokers combined) who were heterozygous for all three polymorphisms had significantly higher frequencies of CAs in vivo (0.82 ± 0.18) compared to smokers homozygous for the wild-type allele for all three polymorphisms (0.20 ± 0.13, P=0.053). This was also observed among smokers with the variant 156A polymorphism, but homozygous wild-type for the 312 and 751 alleles (0.89 ± 0.27, P= 0.053), and more significantly for smokers heterozygous for the 312Asn polymorphism and homozygous wild-type for the C156 and Lys751 alleles (0.89 ± 0.36, P=0.028) compared to smokers homozygous wild-type for all three polymorphisms (Figure 20). When this population was stratified by age using the median age (39 years old), all older individuals homozygous for the wild-type allele for all three polymorphisms had significantly higher frequencies of CAs than younger individuals with the same genotype (P=0.017) 24 hrs following NNK exposure. At 24 hrs post NNK treatment, we also observed significantly more CAs in younger individuals heterozygous variant for the C156A polymorphism (5.21 ± 0.57, P=0.025), and for the Lys751Gln polymorphism (5.37 ± 0.68, P=0.037), and in younger individuals with polymorphic alleles at both positions 312 and 751, but homozygous wild-type at position 156 (4.96 ± 0.41, P=0.020) compared to younger individuals homozygous wild-type for all three polymorphisms (3.08 ± 0.68) (Figure 21).

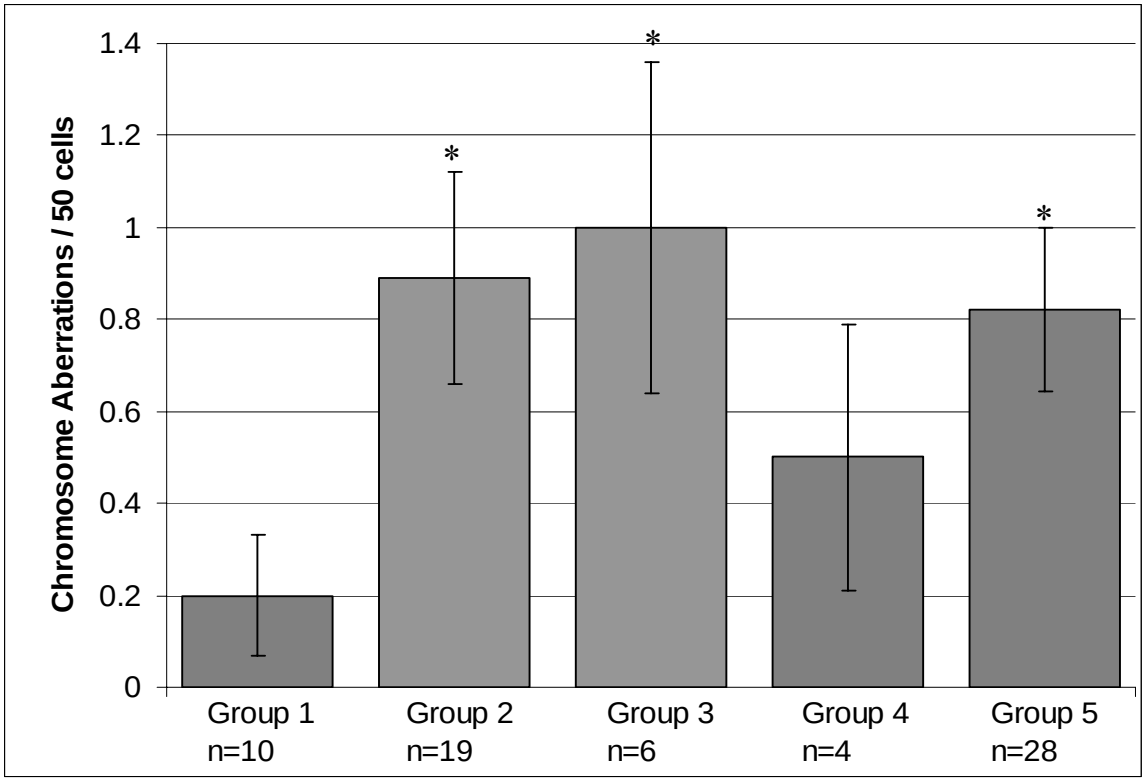


Figure 20. Effect of smoking and genotype on in vivo CAs using classical cytogenetics.
Group 1: C156C, Asp312Asp, and Lys751Lys (wild-type), Group 2: C156A, Asp312Asp, and Lys751Lys Group 3: C156C, Asp312Asn, and Lys751Lys, Group 4: C156C, Asp312Asp, and Lys751Gln, Group 5: C156A, Asp312Asn, and Lys751Gln.
* statistically significant $P<0.05$ compared to group 1
See appendix Table XII for more information

As stated previously an age effect was also observed, older individuals (>39 years old), who were wild-type for all three polymorphisms, had significantly higher CA frequencies (6.09 ± 0.97, P=0.017) than younger individuals with the same genotype (3.08 ± 0.68). This was both at 2 hrs (3.26 ± 0.35 for age>39 and 2.14 ± 0.29 for age<39; P=0.019) and 24 hrs (P=0.017, Figure 21) post-NNK challenge. When stratified by gender we observed that males (n=20) homozygous or heterozygous for variant allele at both positions 312 and 751, but homozygous wild-type at position 156, had significantly less CAs than females (n=36) with the same genotype both *in vivo* (0.36 ± 0.13 for the

polymorphic subjects and 0.94 ± 0.19 for the wild-type subjects; P=0.045) and 24 hrs following NNK treatment (1.94 ± 0.39 for the polymorphic subjects and 3.03 ± 0.33 for the wild-type subjects; P=0.05).

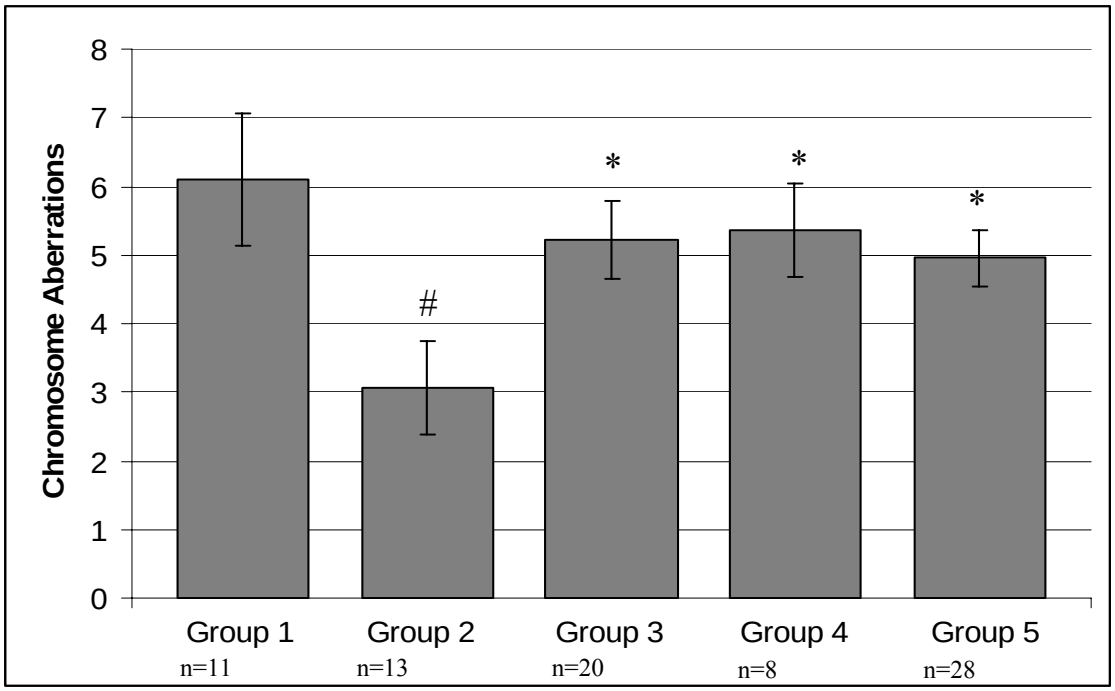


Figure 21. Effect of age and genotype on CAs 24 hrs following NNK treatment using classical cytogenetics.
Group 1: old (age> 39 years) C156C, Asp312Asp, and Lys751Lys (wild-type), Group 2: young (age ≤ years) C156C, Asp312Asp, and Lys751Lys (wild-type), Group 3: young C156A, Asp312Asp, and Lys751Lys, Group 4: young C156C, Asp312Asp, and Lys751Gln, Group 5: young C156C, Asp312Asn, and Lys751Gln. #, statistically significant P<0.05 compared to group 1. * statistically significant P<0.05 compared to group 2. See appendix Table XIII for more information

XPD mRNA ANALYSIS [this work is described by Wolfe *et al.* (in press)]

The exact mechanism(s) by which SNPs in the *XPD* gene may alter DNA repair, and increase the frequency of CAs in response to environmental carcinogens (see above) is still unknown. Changes in DNA repair capacity that result in this observed increase in CAs could be a result of a combination of factors that include decreased *XPD* mRNA

levels, or decreased XPD protein abundance and/or function. Consistent with this hypothesis, Vogel *et al*. (2000) have shown that both *XPD* (*ERCC2*) and *ERCC1* mRNA levels correlate positively with DNA repair capacity. In the current study, we used peripheral blood lymphocytes (PBLs), from human subjects to test the hypothesis that SNPs in the *XPD* gene alter *XPD* gene expression, and thus influence DNA repair capacity, leading potentially to the accumulation of genetic damage. The goal was to delineate the impact of the Arg156Arg, Asp312Asn and Lys751Gln SNPs in the *XPD* gene, independently and in combination with each other, on altering base-line *in vivo* *XPD* gene expression. Demographic and genotype data generated previously guided the selection of lymphocytes from representative subgroups of smokers and non-smokers who were homozygous or heterozygous for the polymorphic variants of the alleles and who exhibited different levels of baseline genetic damage. PBLs were also selected from individuals who were wild-type for the three SNPs studied. These subjects served as the reference group in analysis of the influence of the *XPD* polymorphisms

Demographics of the population used in the mRNA analysis study

110 individuals participated in this study. This population consisted of 21 males (19%) and 89 females (81%). The majority (n=68), were White non-Hispanics (62%). There were also 21 African American subjects (19%), 11 Hispanics (10%), and 10 Asians (9%). Of the 110 volunteers who participated in this part of the study, 54 (49%) were smokers, and 56 (51%) were non-smokers. Non-smokers, and smokers were defined as previously described. For the smokers, the mean ± S.D. number of cigarettes smoked per day was 17.4 ± 10.7. The mean number of accumulated pack years was 16.5 ± 16. The median number of pack years was 12. This median value was used to categorize smokers into heavy smokers (>12 pack years) and light smokers (≤12 pack years) for subsequent stratified analysis for the effect of smoking on *XPD* mRNA transcript copy number. The age of the participants ranged from 21 to 73 years, with a median of 40 years. The study subjects were divided into two age groups based on the median age: younger individuals

(≤40 years old) and older individuals (>40 years old) for subsequent stratified analysis for the effect of age on *XPD* mRNA transcript copy number.

Determination of mRNA copy number in the study population

The mean (± SE) mRNA transcript copy number in the total study population, is presented in Table X. When smoking status was considered (yes or no), there was a slight, but not statistically significant, decrease in mRNA copy number observed in smokers compared to non-smokers. However, the intensity of smoking (expressed as pack years) significantly affected the mRNA transcript copy numbers. Heavy smokers who smoked >12 pack-years had a statistically significant decrease (1.4-fold; P<0.002) in mRNA copy number compared to non-smokers. Subjects who smoked >12 pack-years also had a statistically significant decrease in mRNA copy number (1.5-fold; P<0.002) compared to light smokers who smoked ≤ 12 pack-years. Ethnicity, age and gender did not significantly affect mRNA transcript copy number (Table X).

Table X. Mean mRNA Transcript copy number among population subgroups

No. of subjects	mRNA copy number[a]	P value[b]
All subjects (n=110)	17407 ± 730	
Stratified by smoking		
non-smokers (n=56)	18382 ± 1063	reference
smokers (n=54)	16296 ± 1002	0.160
light smokers (n=26)	19121 ± 1612	reference
heavy smokers (n=27)	12994 ± 951	0.002*
Stratified by race		
non-Hispanic Whites (n=68)	16707 ± 891	reference
African Americans (n=21)	18702 ± 1995	0.300
Hispanics (n=11)	15690 ± 1933	0.660
Asians (n=9)	21548 ± 3378	0.075
Stratified by age		
greater than 40 yrs (n=55)	16173 ± 988	reference
younger than 40 yrs (n=54)	18787 ± 1067	0.075
Stratified by gender		
Males (n=21)	17272 ± 1765	reference
Females (n=89)	17439 ± 805	0.093

[a] Values are expressed as mean mRNA transcript copy numbers ± standard error.
[b] The reference group in each category is compared to the corresponding test group(s).
* Statistically significant at p<0.05.

Reprinted from Pharmacogenetics and Genomics, Wolfe KJ., Wickliffe JK., Hill CE., Ammenheuser MM., Abdel-Rahman SZ., in press.

Effect of polymorphisms in *XPD* on the mRNA transcript copy number

When the effect of *XPD* polymorphisms on the copy number of the mRNA transcript was evaluated, we found that the variant 156A, 312Asn and 751Gln alleles were each individually associated with a highly significant decrease in mRNA copy number. As shown in Figure 22, there was a 1.5-fold decrease ($P<0.003$) in transcript copy number in cells from group 2 individuals, who had only the 156A polymorphism, but were homozygous wild-type for both the Asp312Asp and Lys751Lys alleles (group 2, mean ± SE mRNA transcript copy number =16595 ± 845), compared to cells from individuals homozygous wild-type for all three SNPs studied (reference category, group 1, 24427 ± 2483). Similarly, individuals with only the 312Asn polymorphism (*i.e.* homozygous wild-type for C156C and Lys751Lys genotypes) had a significant 1.8-fold

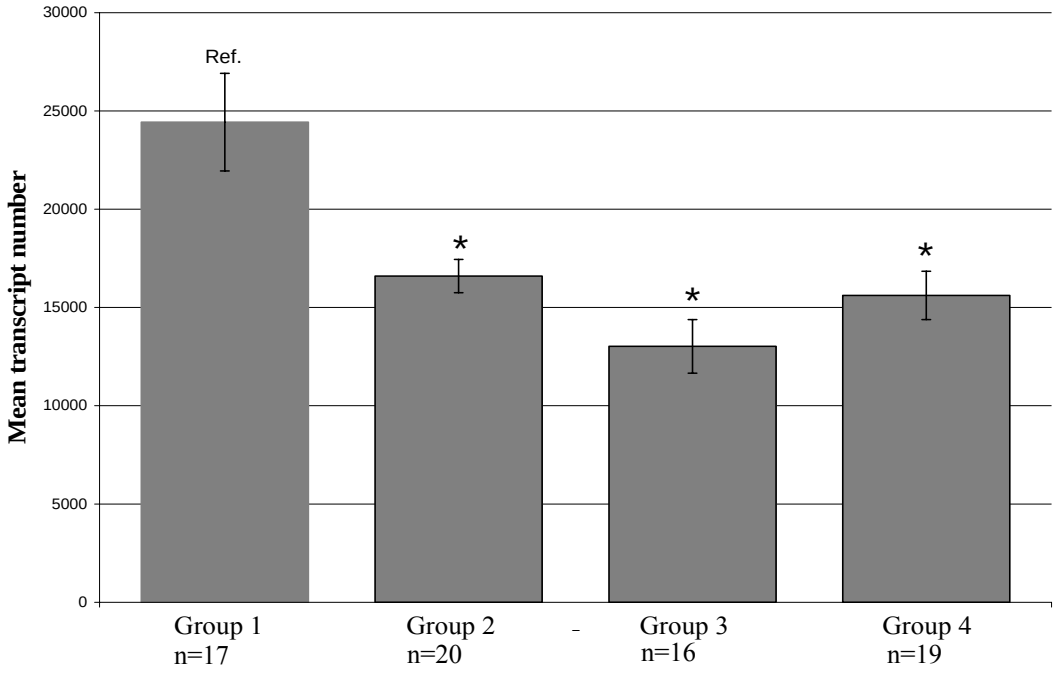


Figure 22. Influence of XPD polymorphisms on mRNA transcript copy number.
Genotypes are Group 1: C156C, Asp312Asp, and Lys751Lys individuals (wild-type), Group 2: C156A, Asp312Asp, and Lys751Lys individuals, Group 3: C156C, Asp312Asn, and Lys751Lys individuals, Group 4: C156C, Asp312Asp, and Lys751Gln individuals. * Statistically significant, $P<0.05$. For more information see appendix Table XIV. Reprinted from Pharmacogenetics and Genomics, Wolfe KJ., Wickliffe JK., Hill CE., Ammenheuser MM., Abdel-Rahman SZ., in press.

decrease (P=0.0004) in transcript copy number (group 3, 13017 ± 1361), compared to the reference category (group 1). The same observation was true in cells from group 4 individuals with the 751Gln polymorphism (*i.e.* homozygous wild-type for the other two SNPs studied), where a significant 1.5-fold decrease (P<0.002) in transcript copy number (15610 ± 1230) was observed, compared to the reference group (Figure 22).

The interaction between the individual SNPs and smoking, age and gender was then examined. No significant interaction between the individual SNPs studied and gender on the *XPD* transcript copy number was observed. However, a significant interaction between smoking and age and SNPs was observed. As shown in Figure 23, among both smokers and nonsmokers, SNPs induced a decrease in *XPD*

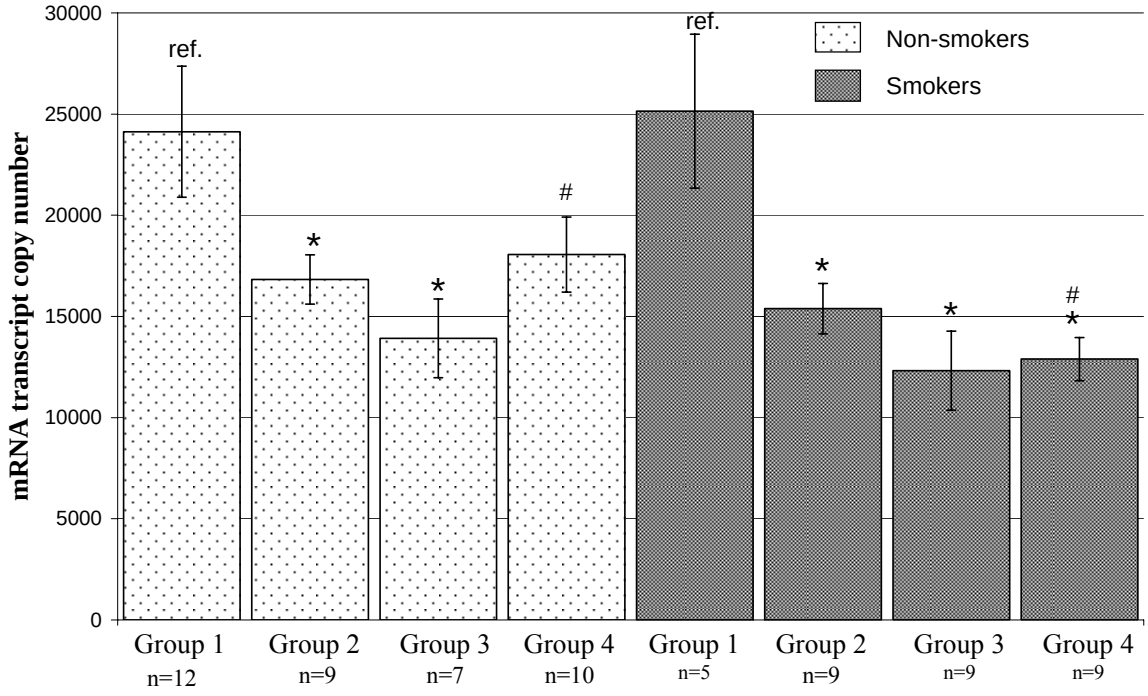


Figure 23. XPD genotype as a modifying factor in mRNA transcript copy number among smokers and nonsmokers

Genotypes are Group 1: C156C, Asp312Asp, and Lys751Lys individuals (wild-type), Group 2: C156A, Asp312Asp, and Lys751Lys individuals, Group 3: C156C, Asp312Asn, and Lys751Lys individuals, Group 4: C156C, Asp312Asp, and Lys751Gln individuals. * Statistically significant, P<0.05. # Comparatively statistically significant, P<0.05. For more information see appendix Table XV.

mRNA transcript copy number. However, among cells from subjects with the 751Gln (group 4) polymorphism, smokers had a significantly greater decrease (1.4-fold) in the *XPD* mRNA transcript copy number than cells from nonsmokers with the same genotype. Also, as shown in Figure 24, significant effects of the SNPs studied were observed in cells from older subjects (> 40 years old). A 1.5-fold decrease in *XPD* transcript copy number was observed among older individuals with the 156A variant (P<0.03), a 1.8-fold decrease with the 312Asn polymorphism (P <0.01) and a 1.7-fold decrease with the 751Gln polymorphism (P<0.02) compared to older individuals homozygous for the wild-type alleles at all three loci (group 1, reference

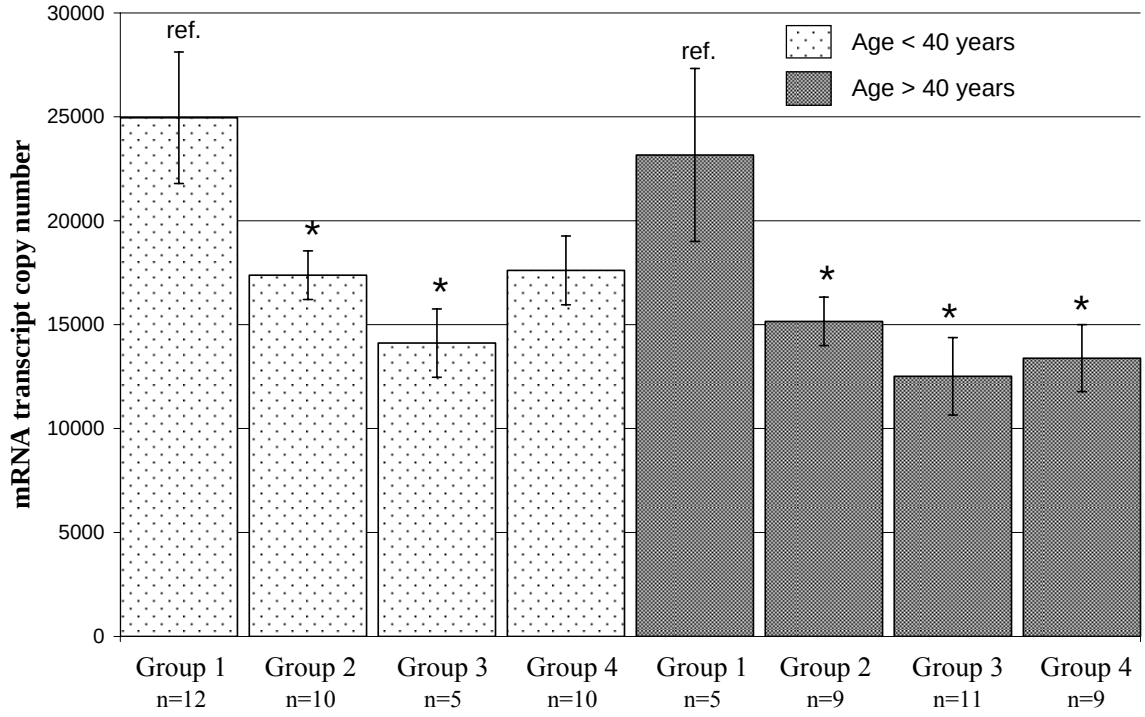


Figure 24. XPD genotype as a modifying factor in mRNA transcript copy number among younger (age <40 years) and older (age >40 years) individuals.
Genotypes are Group 1: C156C, Asp312Asp, and Lys751Lys individuals (wild-type), Group 2: C156A, Asp312Asp, and Lys751Lys individuals, Group 3: C156C, Asp312Asn, and Lys751Lys individuals, Group 4: C156C, Asp312Asp, and Lys751Gln individuals. * Statistically significant, P<0.05. For more information see appendix Table XVI. Reprinted from Pharmacogenetics and Genomics, Wolfe KJ., Wickliffe JK., Hill CE., Ammenheuser MM., Abdel-Rahman SZ., in press.

category, mean ± SE mRNA transcript copy number of 23164 ± 4159). In younger subjects (<40 years old), however, the effect of each polymorphism studied was less pronounced. Compared to younger individuals homozygous for the wild-type alleles at all three loci (group 1, 24953 ± 3162), a 1.4-fold decrease in *XPD* transcript copy number was observed among younger individuals with the 156Adenine polymorphism (P<0.05), and a 1.7-fold decrease with the Asp312Asn polymorphism (P <0.05), but no significant effect was observed with the Lys751Gl polymorphism.

When the combined effect of all three SNPs was considered, we found that in cells from subjects who had all of the polymorphisms present (*i.e.* 156Adenine, 312Asn and 751Gln), there was a significant 38% decrease (P<0.001) in mRNA transcript copy number compared to cells homozygous wild-type at all three loci (Table XI). The results

Table XI. Effect of age and XPD polymorphisms on mRNA transcript copy number

Genotype (No. of subjects)	mRNA copy number[a]	P value[b]
All subjects homozygous wild-type at all three loci (n=17)	24427 ± 2483	ref.
All subjects with variant alleles at all three loci (n=18)	15214 ± 1051	0.001*
Younger subjects (age<40) homozygous wild-type at all three loci (n=12)	24953 ± 3162	ref.
Younger subjects (age<40) with variant alleles at all three loci (n=10)	17397 ± 1171	0.042*
Older subjects (age>40) homozygous wild-type at all three loci (n=5)	23164 ± 4159	ref.
Older subjects (age>40) with variant alleles at all three loci (n=8)	13242 ± 1203	0.017*

[a] Values are expressed as mean mRNA transcript copy numbers ± standard error.
[b] Comparison between the reference category (homozygous wild-type at all three loci) and the test group (variant at all three loci)
* Statistically significant at p<0.05.
Reprinted from Pharmacogenetics and Genomics, Wolfe KJ., Wickliffe JK., Hill CE., Ammenheuser MM., Abdel-Rahman SZ., in press.

presented in Table XI also indicate an effect of age in cells from subjects with all three polymorphisms in the *XPD* gene. In older subjects (> 40 years old), cells with all three polymorphisms (156Adenine, 312Asn, and 751Gln) had a 43% lower mRNA transcript copy number compared to cells from older subjects homozygous wild-type for all three alleles (P=0.017). In younger subjects (<40 years old), cells with all three polymorphisms in the *XPD* gene had only a 30% lower mRNA transcript copy number compared to cells from younger subjects homozygous wild-type for all three alleles (P=0.042).

Mfold analysis of *XPD* polymorphisms on mRNA secondary structure

To test whether the silent 156C→A substitution (Arg156Arg), the G→A substitution (Asp312Asn) or the A→C substitution (Lys751Gln) change the secondary structure of mRNA, we conducted Mfold analysis (Vogel *et al*., 2000) of *XPD* mRNA. We selected the mRNA sequences surrounding each SNP (70 nucleotides on each side) for analysis and used the corresponding wild-type sequences for comparing the predicted effect on the secondary mRNA structure. Our results, presented in Figure 25, indicate that both the 156C>A polymorphism in exon 6 and the Lys751Gln polymorphism in exon 23 result in an apparent change in the mRNA secondary structure. Substitution of C at the 156 position with A or substitution of A with C at position 751 (Lys751Gln) alters apparent mRNA secondary structure and the amount of single strandedness (Figure 25). In contrast, the G→A substitution corresponding to the Asp312Asn amino acid change (the exon 10 polymorphism) did not change the apparent secondary structure of mRNA or its single strandedness. Nucleotide substitutions corresponding to the SNPs in exon 6 and exon 23 resulted in a slightly lower energy requirement for mRNA folding (dG) as compared to the dG of the wild type. In contrast, substitution of G with A at position 312, corresponding to the Asp312Asn polymorphism in exon 10, exhibited a slightly higher dG for mRNA folding (Figure 25).

Figure 25.

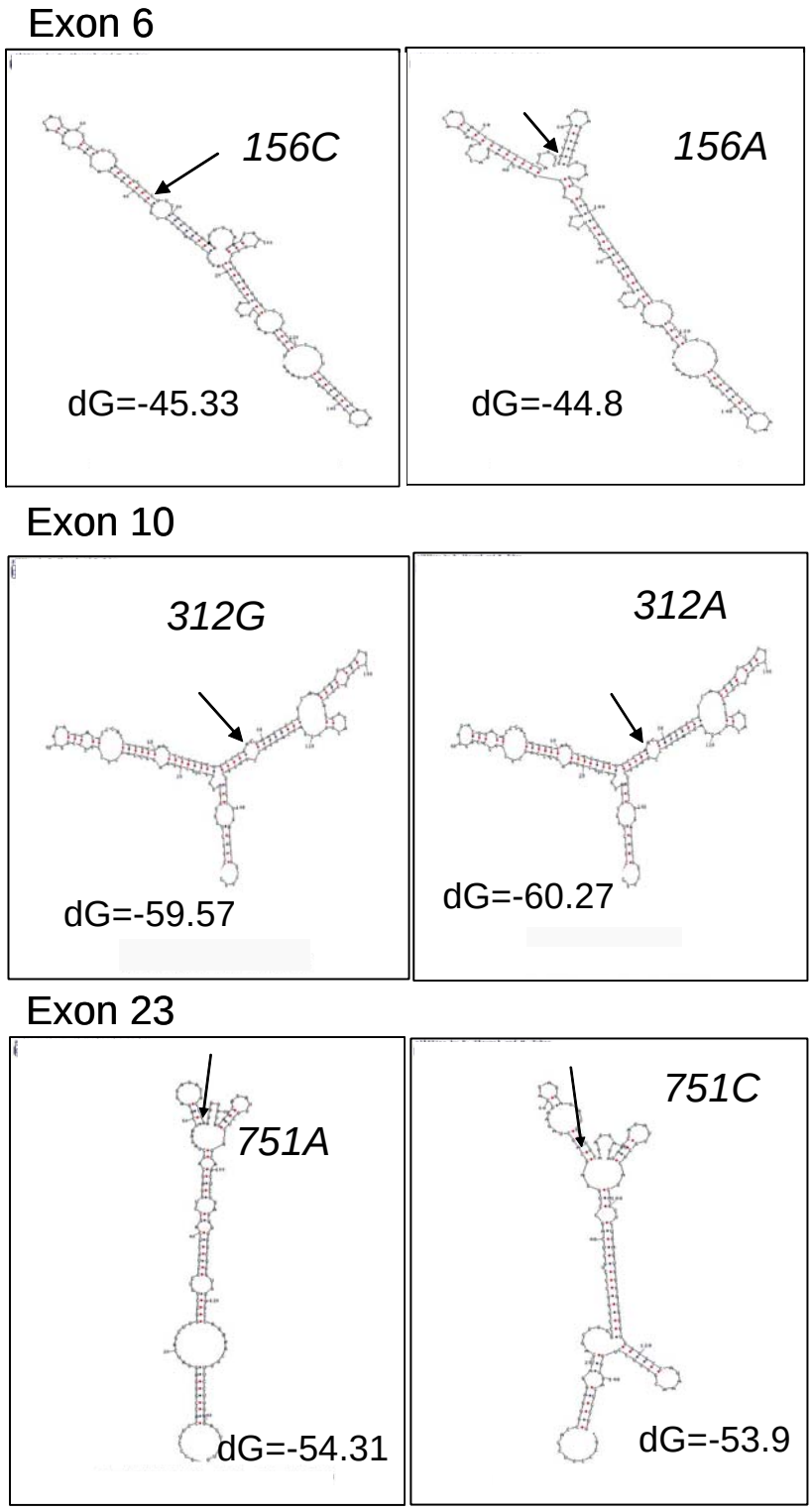

Exon 6
156C
dG=-45.33
156A
dG=-44.8
Exon 10
312G
dG=-59.57
312A
dG=-60.27
Exon 23
751A
dG=-54.31
751C
dG=-53.9

Figure 25. Mfold predictions of the mRNA structures of the XPD polymorphisms
XPD exon 6 Arg156Arg (C→A), **XPD** exon 10 Asp312Asn (G→A) and *XPD* exon 23 Lys751Gln (A→C) polymorphisms and their corresponding folding energy (dG). Partial sequence of *XPD* mRNA (140 nt) surrounding each SNP under study was used to predict secondary structure by Mfold as described in the Methods. Reprinted from Pharmacogenetics and Genomics, Wolfe KJ., Wickliffe JK., Hill CE., Ammenheuser MM., Abdel-Rahman SZ., in press.

XPD PROTEIN EXPRESSION AND LOCALIZATION

The effect of SNPs in the *XPD* gene on the nuclear translocation of the resulting protein has not been studied before. Because nuclear translocation of the XPD protein seems to occur via its interaction with proteins forming the TFIIH complex (Boulikas, 1997; Santagati *et al*., 2001), structural alterations (due to SNPs, as we propose) could lead to interaction modifications between these proteins. This could result in the altered ability of XPD to translocate into the nucleus, and could, consequently, lead to accumulation of genetic damage. The studies described in this chapter were designed to test this hypothesis, and to provide a clearer understanding of the underlying mechanisms involved in this process

We hypothesized that SNPs in the *XPD* gene alter DNA repair capacity through perturbing XPD protein expression and interrupting its nuclear translocation, thus leading to accumulation of genetic damage. The goal of this section is to delineate the impact of the Asp312Asn and Lys751Gln SNPs in the *XPD* gene, independently and in combination with each other on altering XPD protein expression, and altering the cellular localization of XPD following DNA damage due to exposure to UV irradiation. To achieve our goals we used PBLs collected from study participants. Endpoints and genotype data generated previously guided the selection of lymphocytes from representative individuals who were homozygous or heterozygous for the polymorphic variants of the alleles, and who exhibited different levels of baseline genetic damage. Also, selected were PBLs from individuals who were wild-type for the 312Asp and 751Lys SNPs studied. These served as the reference group for comparisons in the analysis of the influence of the *XPD* polymorphisms. This study was intended to determine the impact of the SNPs on altering XPD protein expression using fluorescent

confocal microscopy. In addition, another goal for this study was to determine the effect of the SNPs on altering the cellular localization of the XPD protein.

Demographics of the population used in the protein expression/localization study

The population for this study consisted of 14 females, recruited from the total population (n=800). The majority (n=10), were White non-Hispanics (71.4%) and the other 4 were African American (28.6%). Of the 14 volunteers who participated in this study, 2 (14.3%) were smokers, 4 (28.6%) were ex-smokers, and 8 (57.1%) were non-smokers. Non-smokers, and smokers were defined as previously described. The mean number of accumulated pack-years (defined as cigarettes smoked per day multiplied by the number of years of smoking divided by 20) for the combined smokers and ex-smokers was 2.57 ± 1.24. The pack-years of the two smokers were 8 and 15. The age of the participants ranged from 20 to 59 years, with a median of 36 years and a mean of 36.9 ± 3.6.

Cell survival following UV exposure

In these studies, we used UVC radiation as the challenging mutagenic agent to more specifically target the nucleotide excision repair pathway. To determine the appropriate UV dosage to be used in the exposure experiments, a dose response relationship was established at a dose rate of 400μJ/cm^2 for 33 sec, 100 sec, and 333 sec. A dose rate of 400 μJ/cm^2 for 100 sec was established as the optimum experimental condition because it resulted in a significant XPD expression with no significant changes in cell survival compared to untreated cells. Percentage of cell death was determined by bromophenol blue (BPB) exclusion. Processing and manipulation of the cells and exposure to UV resulted in some initial cell death in the first 30 mins that disappeared thereafter, as shown in Figure 26. Thirty minutes following UV exposure, BPB-exclusion dropped to 80.5%. However, 2 hrs following UV exposure there was no significant difference in BPB-exclusion which was above 95% out to 6 hrs after exposure.

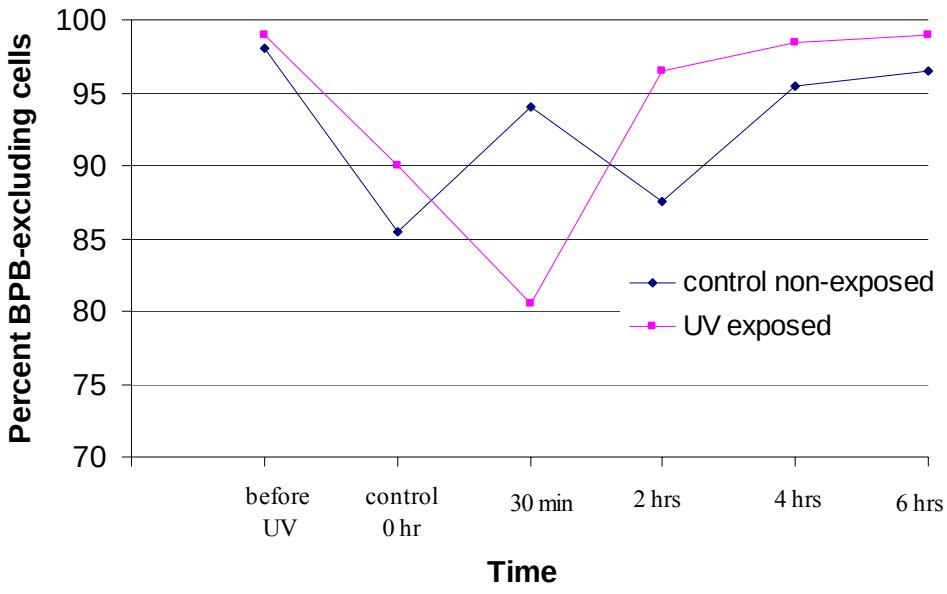
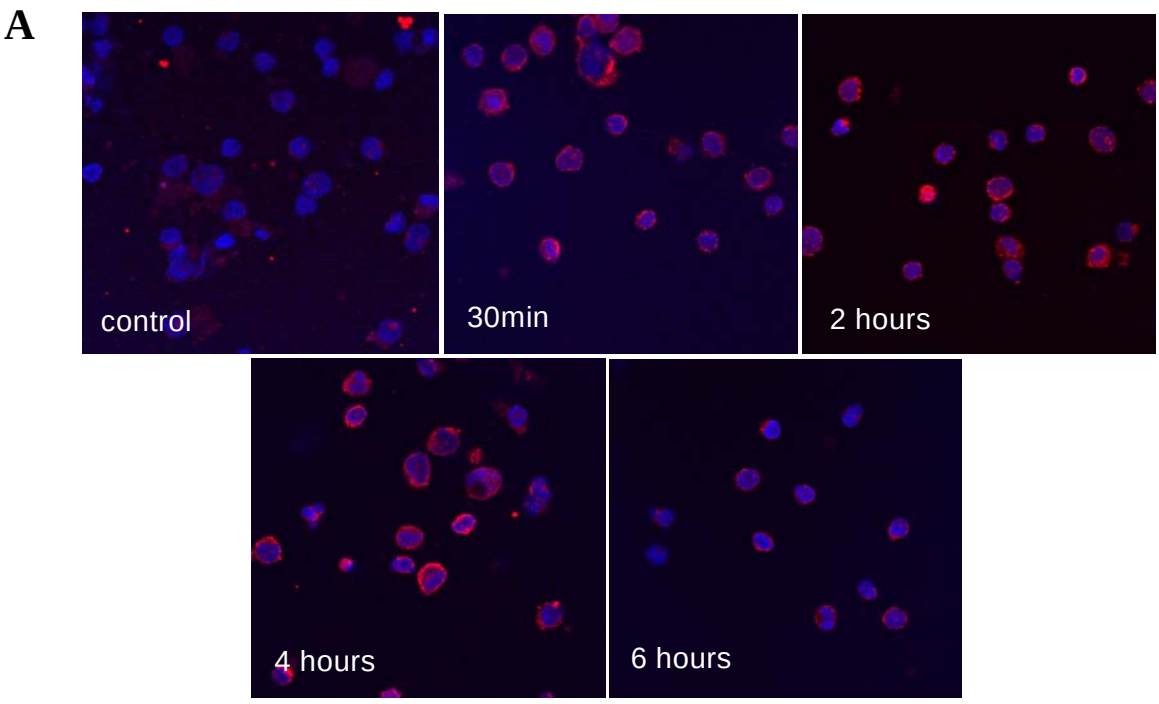


Figure 26. Percent of BPB-excluding cells following UV irradiation
Replicate samples were sham-irradiated or UV irradiated at a dose rate of 400 μJ/cm^2 for 100 sec, and allowed to incubate from 0hr, 30min, 2hr, 4hr or 6hr. Percent of bromo-phenol blue excluding cells were calculated by light microscopy. For more information see appendix Table XVII.

Fluorescent microscopy for XPD protein concentration

XPD protein expression was measured in the lymphocytes of the 14 study participants using the XPD specific antibody (C-20) (Molecular Probes) and a secondary antibody conjugated with Alexa Fluor 594 (Molecular Probes). The amount of fluorescence was visualized and quantitated using fluorescent confocal microscopy and Metamorph imagining software (Molecular Devices, PA). As shown in Figure 27A, a significant increase in fluorescence intensity over time after exposure to UV, depicted as an increase in the red fluorescent signal, was observed. A secondary control sample composed of cells incubated with only the secondary conjugated fluorescent antibody was used to determine the auto-fluorescence and establish the parameters for measuring the change in fluorescence intensity. XPD protein abundance was measured in terms of arbitrary units measuring fluorescent intensity. Shown in Figure 27B, the fluorescent signal, which directly corresponding to the amount of XPD protein, increased

significantly ($P<0.05$) over the control level (0 hr) by 1.14-, 1.24-, 1.34-, and 1.27-fold at 30 min, 2 hrs, 4 hrs, and 6 hrs, respectively.



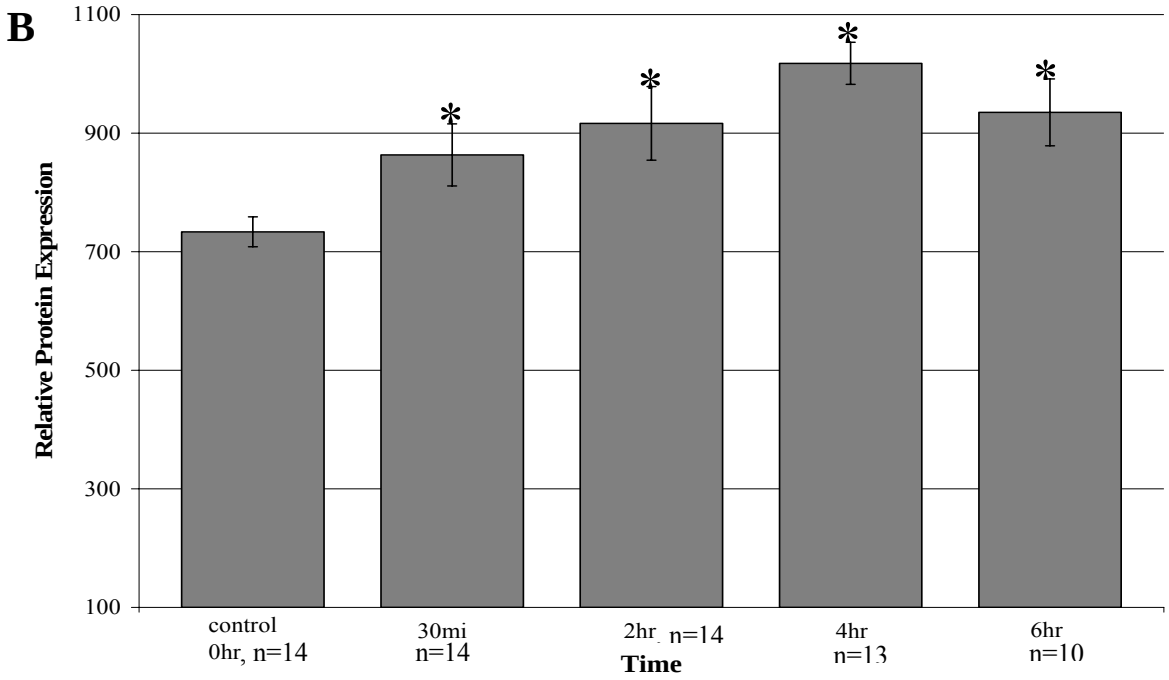


Figure 27. XPD protein abundance following UV irradiation

XPD protein expression (red) within individual cells (DAPI nuclear stain) over time following UV irradiation using a specific XPD C-20 goat polyclonal antibody (Snata Cruz Biotech, CA) and a chicken anti-goat polyclonal antibody conjugated with Alex Fluor 594 (Molecular Probes, OR). Cells were irradiated at a dose rate of 400 $\mu J/cm^2$ for 100 sec, A) visualized using a Zeiss LSM UV meta laser scanning confocal microscope for protein abundance. B) Graphical representation of XPD protein abundance from an average of 100 cells from each time point calculated using Metamorph imaging software. *P<0.05 when compared to the control (0 hr). For more information see appendix Table XVIII.

To investigate the effect of the *XPD* genotype on protein expression following UV irradiation, a haplotype analysis was conducted. Protein expression from cells from individuals who inherited the 312Asn exon 10 SNP (n=3), cells from individuals with the 751Gln exon 23 SNP (n=3), and cells from individuals who inherited both the 312Asn exon 10 and the 751Gln exon 23 SNPs (n=4) were compared to cells homozygous wild-type for both the exon 10 and exon 23 SNPs (n=4). As shown in Figure 28, protein expression was highly variable between samples within the same genotype groups and between the different genotypes. Variation in protein expression was also observed over time. With the small numbers of samples in each category, it was difficult to determine how XPD protein expression is influenced by the different genotypes. However, based on the results observed with this small sample, there appears to be a significant 30% decrease in protein expression in the cells from individuals with the 751Gln SNP in exon 23 compared to cells homozygous wild-type for all three XPD SNPs (P=0.02) at 30 min following UV treatment. It also appears that cells with both the 312Asn SNP in exon 10 and the 751Gln SNP in exon 23 may have a faster response time. Figure 28 shows that, XPD protein expression from cells with both the 312Asn SNP in exon 10 and the 751Gln SNP in exon 23 SNPs at 2 hrs following UV irradiation showed a significant 1.4-fold increase compared to protein expression in cells homozygous wild-type for all three *XPD* SNPs (P=0.04). These data were obtained with cells from only four subjects, however. Also cells from subjects with either the exon 10 or the exon 23 polymorphisms alone did not show any change in protein expression. To be able to make any further inferences as to the differences in protein expression following DNA damage, a larger sample size will need to be analyzed.

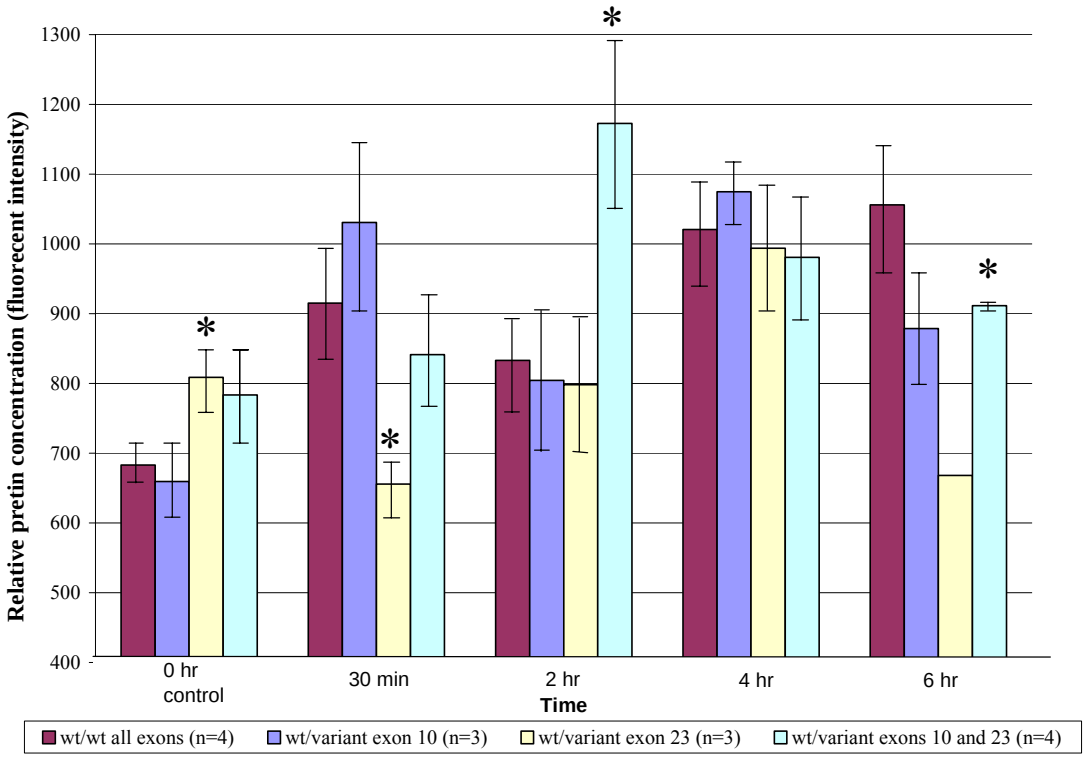


Figure 28. Genotype variation on XPD protein abundance measured by immunofluorescence

Graphical representation of XPD protein abundance averaged from each genotype group over time following UV irradiation, and calculated from fluorescent confocal micrograph images using Metamorph imaging software. * $P < 0.001$ when compared to cells from individuals wt/wt in all exons. For more information see appendix Table XIX.

Real-time quantitative PCR (rqPCR) for *XPD* gene expression

To evaluate the relationship between *XPD* gene expression and protein expression following exposure to UV irradiation, cells from the above study were used to determine *XPD* gene expression by rqPCR, following the protocol outlined earlier. As shown in Figure 29, the observed increase in protein expression was coupled with a significant decrease in gene expression following UV irradiation up to 6 hrs with a 43% decrease at 30 min ($p=0.0002$), a 44% decrease at 2 hrs ($P=0.0001$), a 45% decrease at 4 hrs ($P=0.0005$), and a 51% decrease at 6 hrs ($P<0.0001$). As shown in Figure 30 this

decrease in mRNA gene copy number was observed regardless of the genotype. Gene expression from cells with the 312Asn SNP in exon 10, the 751Gln SNP in

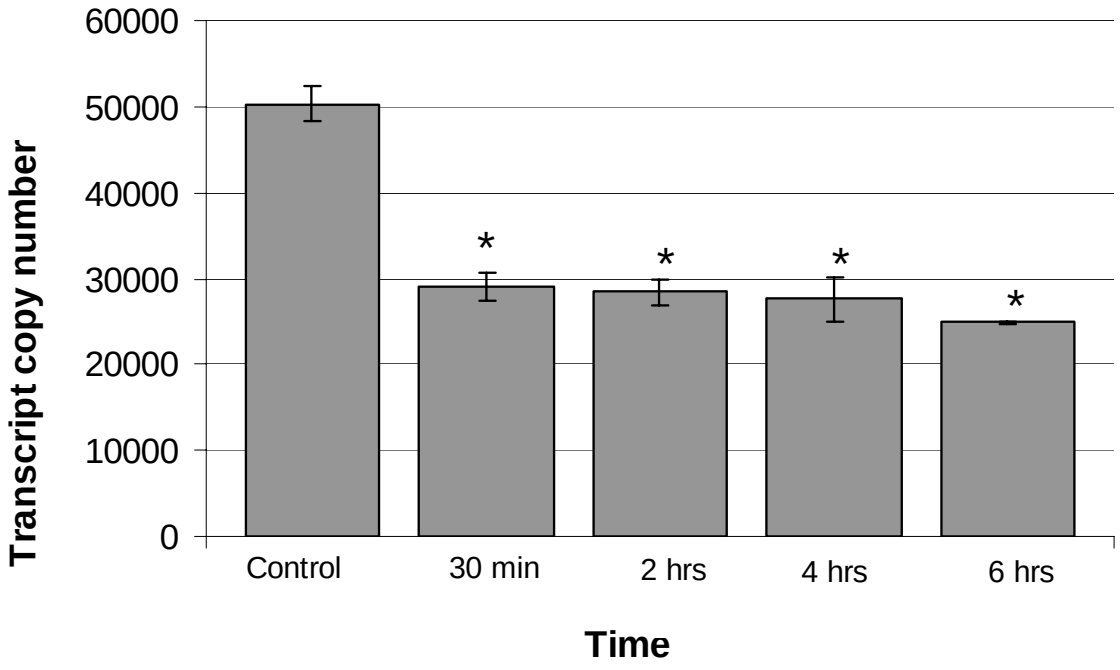


Figure 29. UV- induced XPD gene expression
XPD mRNA transcript copy number over time following UV irradiation measured by reverse transcription real-time PCR using absolute quntitation. * $P < 0.001$ when compared to control. For more information see appendix Table XX.

exon 23, the 312Asn and the 751Gln SNPs, and cells homozygous wild-type for all three *XPD* SNPs (156c, 312Asp, and 751Lys) showed a decrease in gene expression within 30 min following UV irradiation. The decrease in gene expression at 30 min was significantly more pronounced for cells with the 312Asn polymorphism (a 50% decrease, P=0.003) and cells with both the 312Asn and 751Gln polymorphisms (a 51% decrease, P=0.003). All genotypes showed a significant decrease by 6 hrs (a 54% decrease for all wild-type, P=0.0004; a 54% decrease for 312Asn, P=0.003; a 47% decrease for 751Gln, P=0.01; a 55% decrease for both 312Asn and 751 Gln, P=0.001) (Figure 30).

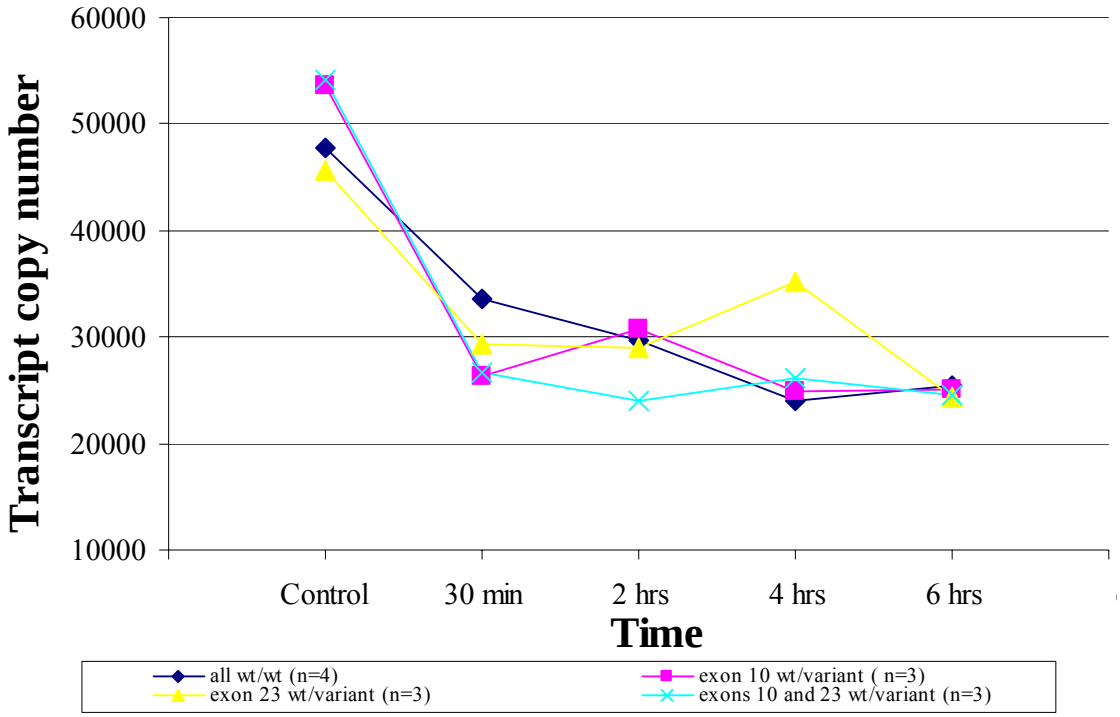


Figure 30. Genotype variation in XPD gene expression following UV exposure
XPD mRNA transcript copy number of cells from polymorphic individuals over time following UV irradiation measured by reverse transcription real-time PCR using absolute quntitation. For more information see appendix Table XXI.

In order to further explore the observed decrease in *XPD* gene expression following UV exposure, a second experiment was performed to examine whether this decrease is confined exclusively to XPD or to other genes. The house-keeping gene, β-actin was used, for that purpose, in which the abundance of β-actin following UV exposure was examined. In this experiment, relative quantitation was used with 18S rRNA as the internal reference standard. As shown in Figure 31, expression of β-actin mRNA was also found to be decreased following UV exposure. From this observation it seems that UV exposure does not have a specific effect on XPD, exclusively, but can also result in a decrease in mRNA transcription in other genes.

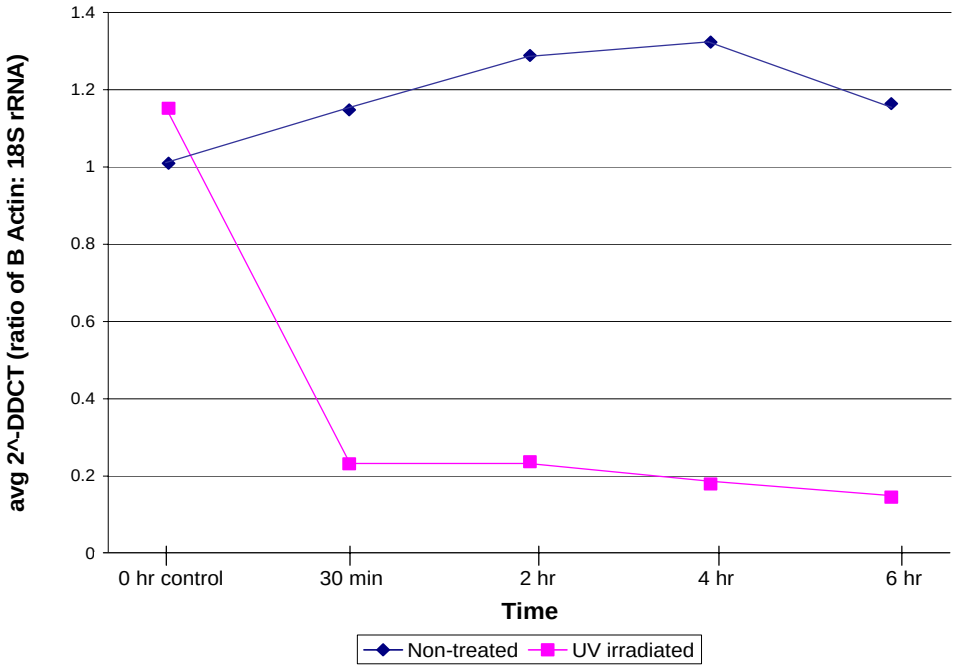


Figure 31. ß-actin mRNA expression following UV exposure
Relative β-actin mRNA transcript copy number from samples over time, following sham-exposure or UV irradiation at a dose rate of 400 μJ/cm^2 for100 sec. Relative standardization using 18S RNA as an internal standard was used. Form additional information see appendix Table XXII.

XPD Localization

Fluorescent confocal microscopy and LSM 510 imagining software (Zeiss) were used to visually and semi-quantitatively study the abundance and movement of the XPD protein following UV irradiation expressed in terms of arbitrary units measuring fluorescent intensity. As shown in Figure 32, the fluorescent intensity, corresponding to protein concentration, was graphed along a vector through the center of the cell. From these micrographs, XPD can be seen increasing in concentration and being localized within the nucleus of the cell. From these images and their analysis (Figure 32 and Figure 33) it appears that both XPD migration and XPD expression increases between 30 min and 4 hrs following UV irradiation with a maximum nuclear localization at 2 hrs following UV irradiation (Figure 33). This is evident from the increased fluorescence of the red signal (XPD protein) within the boundaries of the blue signal (DAPI) identifying the edges of the nucleus.

A

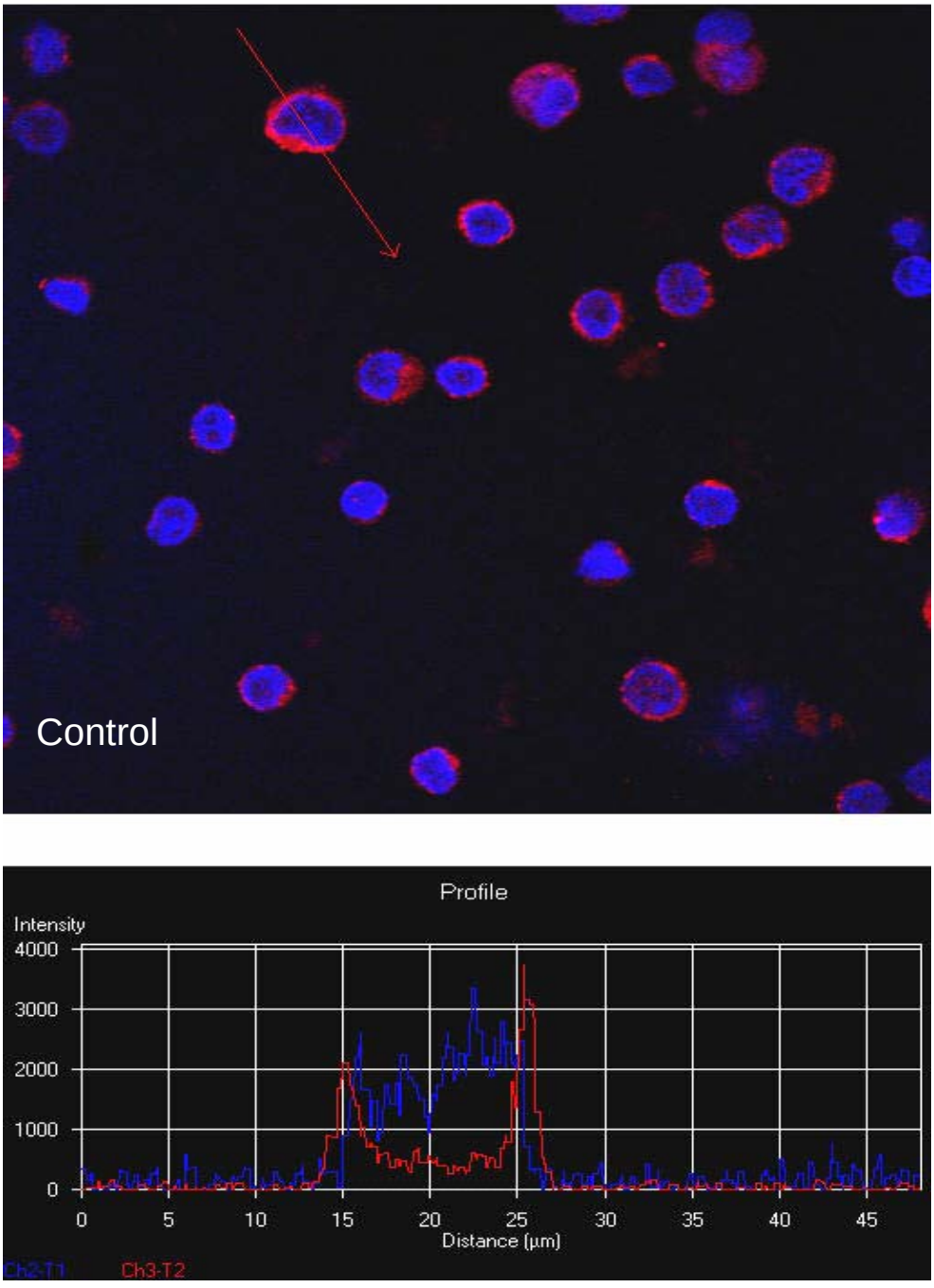


Figure 32. Fluorescent image of XPD localization

XPD protein localization using confocal fluorescent microscopy and LSM 510 analysis software. A vector through the cell is used to graph the DAPI (nuclear) and Alexa Fluor 954 (XPD) intensity. A) XPD nuclear intensity following sham-irradiation.

B

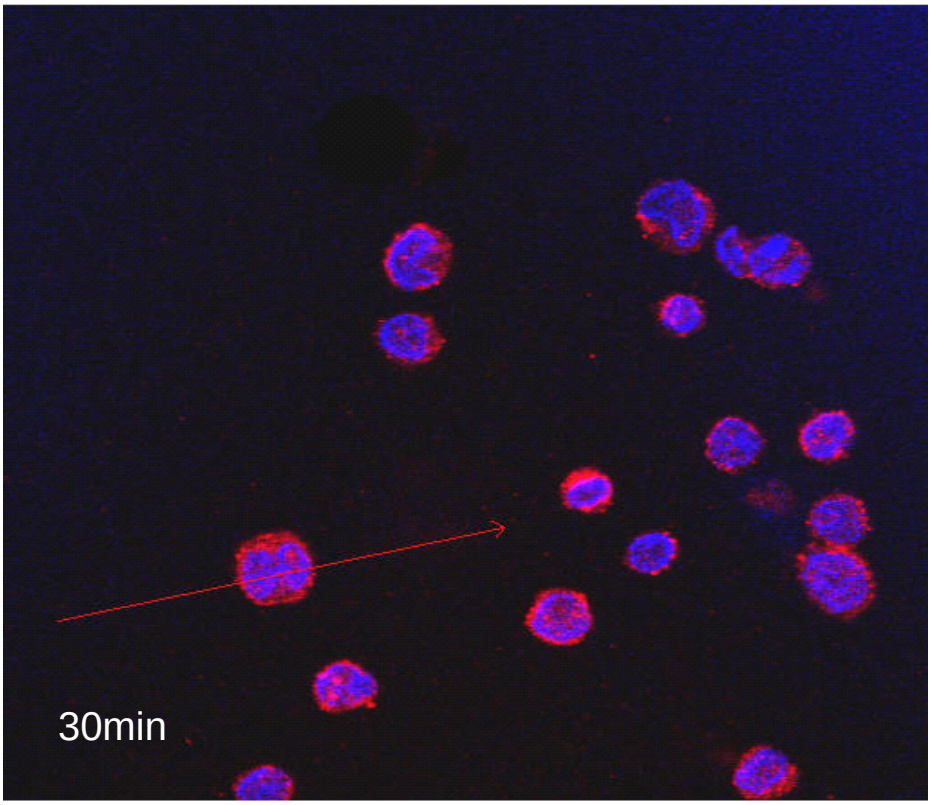


Figure 32. Fluorescent image of XPD localization (continued)
B) XPD nuclear intensity 30 min after UV irradiation at a dose rate of 400 µJ/cm^2 for 100 sec.

C

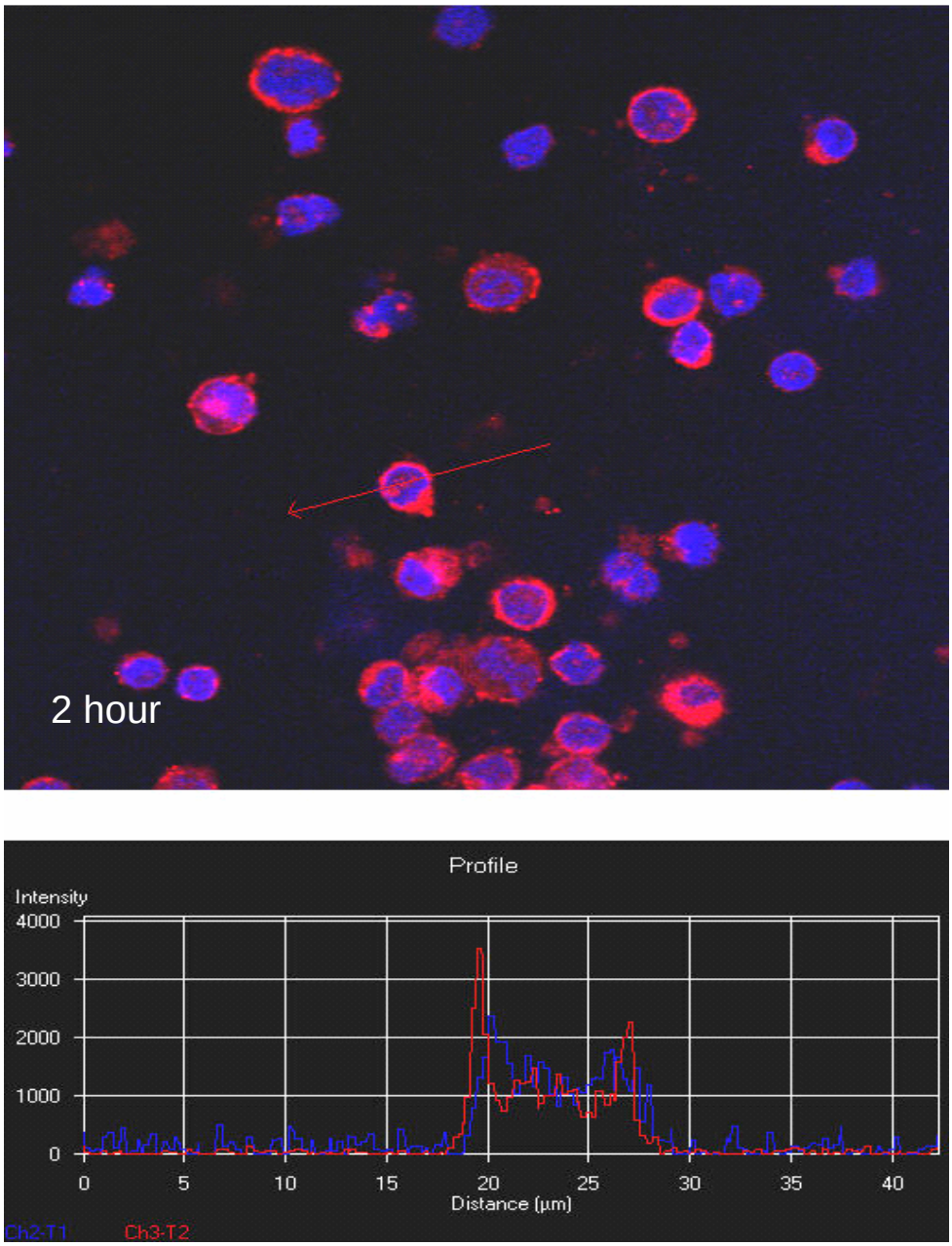


Figure 32. Fluorescent image of XPD localization (continued)
C) XPD nuclear intensity 2 hrs after UV irradiation at a dose rate of 400 µJ/cm^2 for 100 sec.

D

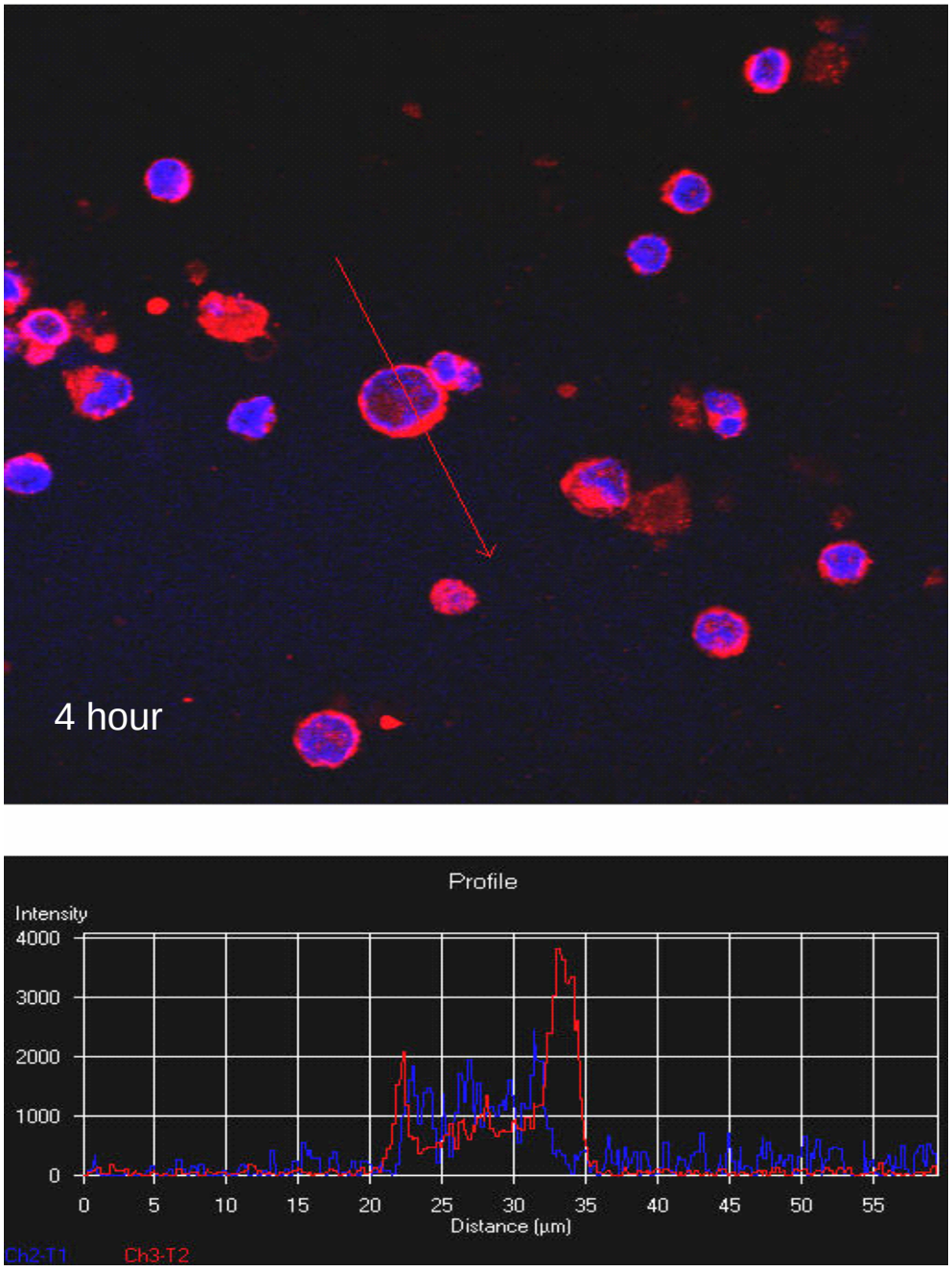
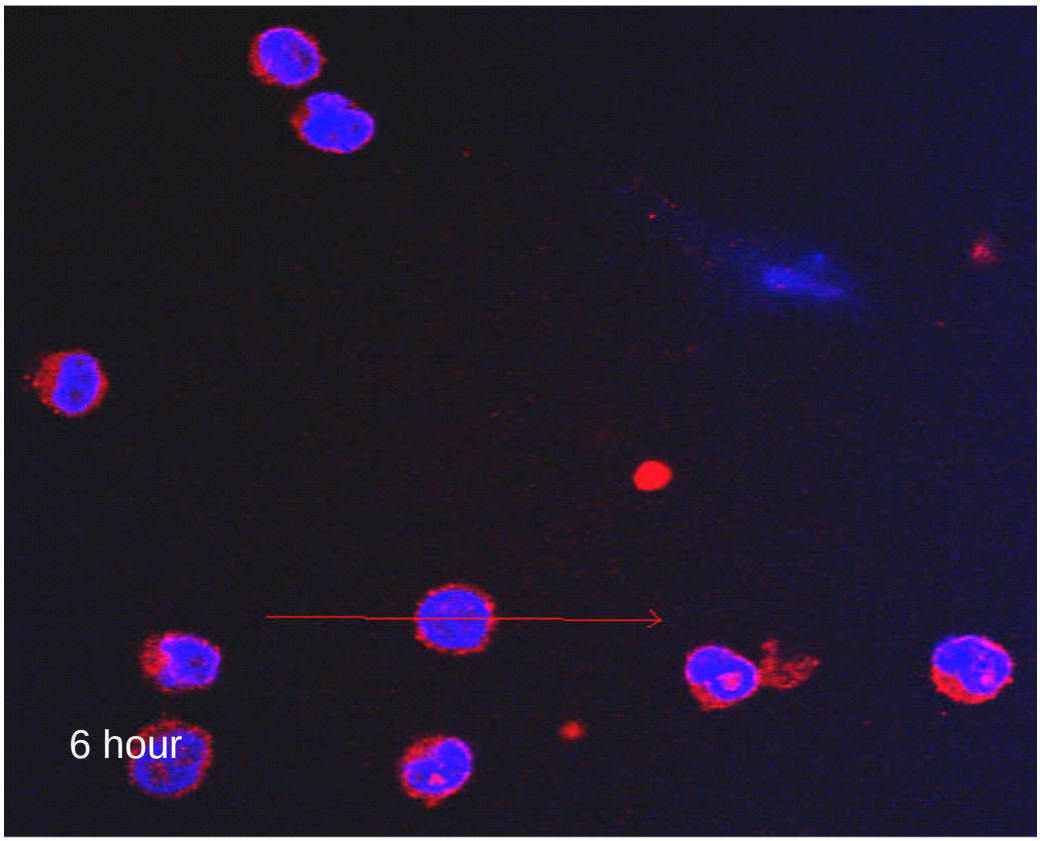


Figure 32. Fluorescent image of XPD localization (continued)
D) XPD nuclear intensity 4 hrs after UV irradiation at a dose rate of 400 µJ/cm^2 for 100 sec.

E



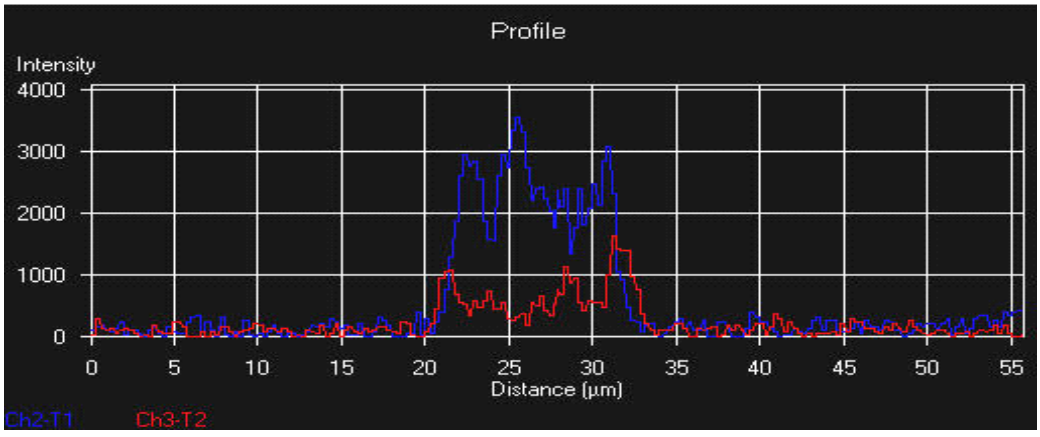


Figure 32. Fluorescent image of XPD localization (continued)
E) XPD nuclear intensity 6 hrs after UV irradiation at a dose rate of 400 µJ/cm^2 for 100 sec.

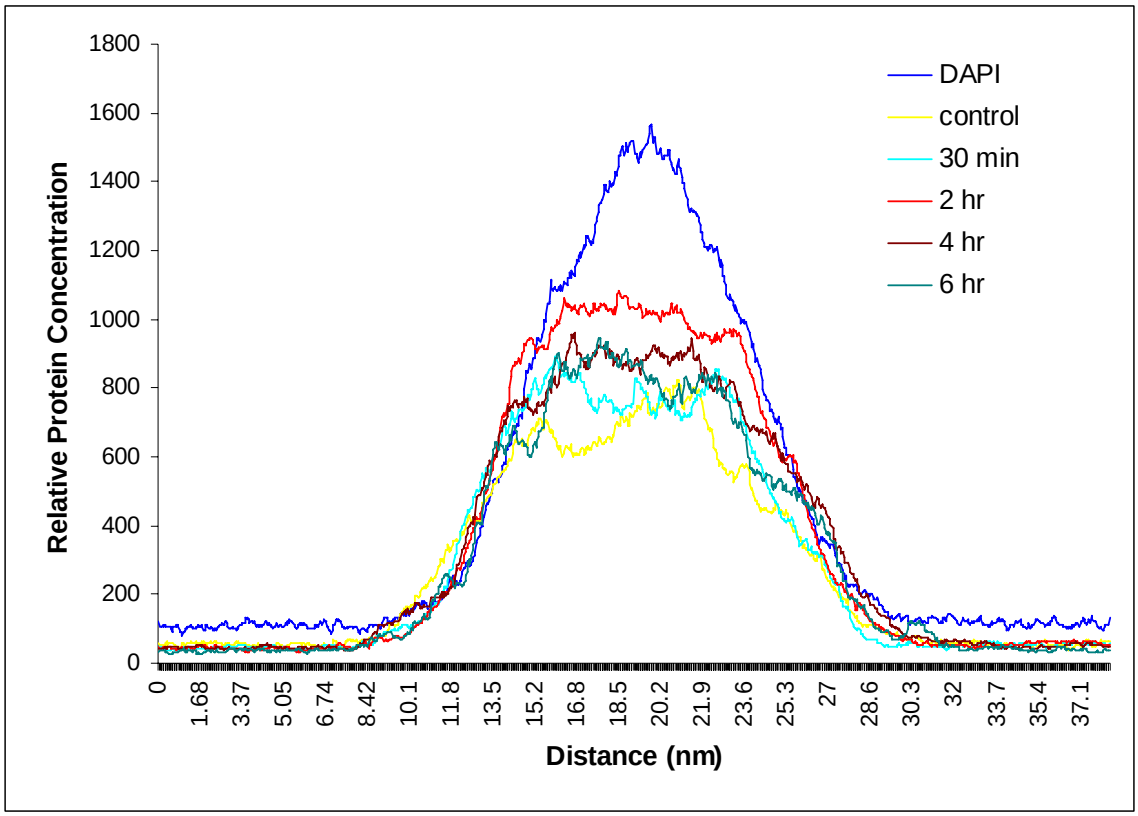


Figure 33. XPD Protein translocation into the nucleus following UV irradiation
Time course of XPD protein nuclear abundance following UV exposure at a dose rate of 400 μJ/cm^2 for 100 sec. Average of images from 100 cells from all samples, calculated using LSM 510 software.

To investigate whether the *XPD* genotype influences the protein localization and the nuclear abundance of XPD, the 14 samples were analyzed separately by genotype. Cells from four individuals homozygous wild-type for 156C, 312Asp, and 751Lys, three individuals heterozygous variant for 312Asn and homozygous wild-type for the 156C and 751Lys polymorphisms, three individuals heterozygous variant for the 751Gln polymorphism and wild-type for the 156C and 312Asp polymorphisms, and four individuals, who were heterozygous variant for both the 312Asn and the 751Gln polymorphisms and homozygous wild-type for the 156C polymorphism were studied. Cells with the variant alleles for both the 312Asn and the 751Gln polymorphisms (Figure 34B) showed an approximate 2-fold increase in nuclear XPD protein concentration at 2

hrs following UV treatment compared to cells from individuals homozygous wild-type for all three polymorphisms (reference group) (Figure 34A). No significant differences were observed with either the 312Asn or 751Gln polymorphisms each independently.

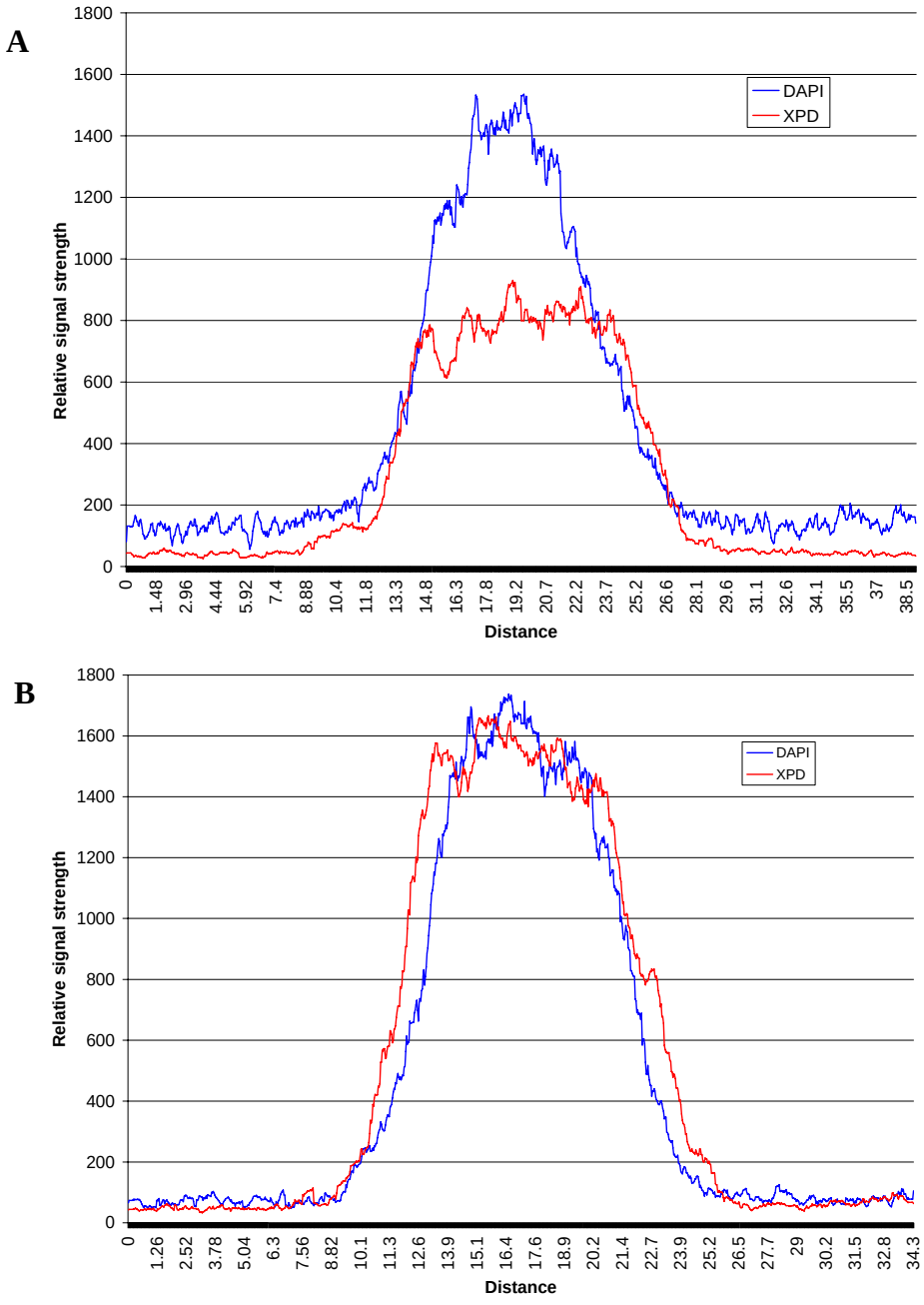


Figure 34. Genotype-dependent differences in XPD protein localization 2 hrs after UV exposure

A. XPD localization in homozygous wild-type cells 2 hr following UV irradiation. B. XPD localization in 312Asn and 751Gln polymorphic cells 2 hr following UV irradiation. Note the increase in XPD protein (red) within the nuclear region outlined by the increase above base line of DAPI (blue)

Discussion

Most of the epidemiological studies that addressed effects of *XPD* polymorphisms on disease risk have focused on the synonymous SNP in codon 156 in exon 6 or the nonsynonymous SNPs in codons 312 in exon 10 and 751 in exon 23 of the gene because of the high prevalence of these polymorphisms (Benhamou and Sarasin, 2002). Several studies have reported associations between these three SNPs and alterations in DNA repair capacity and/or susceptibility to cancer at different sites, but these associations were not evident in other studies (reviewed in Benhamou and Sarasin, 2005). Inconsistencies in the results observed in the different studies could be explained by several factors, including differences in study design, variability in the populations studied, and potential confounding effects of the linkage disequilibrium known to exist among the exon 6, 10 and 23 polymorphisms (Zhou *et al*., 2002; Benhamou and Sarasin, 2005). Little is known about the mechanisms for the observed associations.

CHROMOSOME ABERRATION ANALYSIS

One of the hypotheses tested in this dissertation was that polymorphisms in the DNA repair gene *XPD* influence an individual's ability to repair genetic damage produced by environmental insults. This will manifest as an increase in chromosome aberration frequency. Chromosome aberrations (CAs) were used as a marker of early biological effects preceding cancer, since increased frequency of CAs is a well-established early biomarker for cancer risk, according to large human cohort studies (Hagmar *et al*., 2004; Norppa, 2004). Our research group has frequently used CAs as a biomarker of genetic damage for the evaluation of the effects of polymorphisms in

susceptibility genes (Abdel-Rahman *et al*., 1998; Abdel-Rahman *et al*., 1999; Affatato *et al*., 2004, Hill *et al*., 2005a). The sensitivity of CAs as an early biomarker of genetic damage and cancer risk has also been established by many other investigators. El-Zein *et al*. (2000) have also shown that CAs are good predictors of the clinical stage of smoking-associated lung cancer.

Baseline CAs

In order to evaluate my hypothesis, the influence of smoking as a representative external insult was first investigated by using CA as an indicator of genetic damage. Ex-smokers with the 156A, the 312Asn, or the 751Gln polymorphism in *XPD* were found to have significantly higher frequencies of baseline CAs than non-smokers. This was, however, not true for current smokers. This inconsistency could not be explained by the demographic differences between ex-smoker and current smoker groups. In both groups the average number of pack years was similar (9.5), and both groups had approximately the same average age (mean = 41 years old). In addition, both groups had a comparable male to female ratio, which needed to be assessed since the female gender was found to be more susceptible to NNK-induced CAs in a prior study (Hill *et al*., 2005b). The increased frequency of baseline CA in ex-smokers may, however, be explained at least in part by genetic influences. For example, the inheritance of the variant 312Asn genotype in a large percentage of the ex-smokers might be responsible for the observed greater increase in CA frequency in ex-smokers compared to smokers. To test this hypothesis, we examined the 312Asn genotype frequency among the ex-smokers and current smokers. A higher percentage (53%) of the 312Asn genotype was found among ex-smokers than smokers (45%). Therefore, it is possible that the higher percentage of 312Asn polymorphisms among the ex-smoker group influenced the mean frequency of CAs detected, thus resulting in the observed higher CA frequencies in the ex-smokers when compared to both smokers and non-smokers.

When baseline CAs were examined with regard to gender, no significant differences were observed. However, when stratified by genotype, males homozygous

wild-type for exon 10 had significantly higher CA frequencies than females with the same genotype. This could be explained by the age effect since the average age of the males was significantly higher than that of the females. Although no overall age effect was found with respect to a significant difference in CA frequency, age is known to be a factor in the accumulation of genetic damage. For example, age-related decreases have been reported for DNA repair capacity (Wei *et al*., 1993; Barnett and King, 1995; Moriwaki *et al*., 1996; Wei 1998).

In an attempt to investigate how the three polymorphisms in the *XPD* gene, alone and in combination with each other, could influence genetic damage, we investigated the effects of the different haplotypes on CA frequency. The study population was stratified according to several genotype combinations. These combinations were such that one polymorphism was present at a time. In doing so, we found that cells from smokers (current smokers combined with ex-smokers) with either the 156A or 312Asn *XPD* polymorphism alone showed increased baseline CA frequencies compared to cells from smokers homozygous wild-type for all three polymorphisms. This again suggests that the increase in CAs found in ex-smokers, could be attributed to the presence of more subjects in the group with the 156A or the 312Asn polymorphisms. This increase in CAs was also observed among cells from smokers with all three *XPD* polymorphisms, but not from smokers with only the 751Gln polymorphism. Thus, our observations indicate several gene-environment interactions due to the presence of certain polymorphisms in the *XPD* gene.

The mutagen-sensitivity assay

In order to assess effects of *XPD* polymorphisms on sensitivity to genetic damage induced by exposure to specific tobacco carcinogens, we used the mutagen-sensitivity assay with NNK as the test mutagen. This assay is a well-established technique for defining individual susceptibility towards carcinogens (Hsu *et al*., 1989; Wei *et al*., 1996; Landi *et al*., 1998; Wang *et al*., 1998). This technique has previously been used successfully by our group (Abdel-Rahman *et al*., 1999; Affatato *et al*., 2004; Hill *et al*.,

2005a) and others (Kelsey *et al*., 1997; Norppa *et al*., 1997) to address the biological significance of polymorphisms in metabolic and DNA repair genes following genotoxic exposure. In the current project, the tobacco smoke-specific nitrosamine NNK was used to challenge cultured human lymphocytes *in vitro* in order to assess the effects of *XPD* polymorphisms on exposure to a representative DNA damaging agent. Two time points were used. The first time point of 2 hrs following NNK treatment was used to represent initial genetic damage due to exposure to NNK. The second time point of 24 hrs following NNK treatment was used to represent persisting genetic damage or subsequent repair following exposure to NNK. Following NNK exposure, an increase in CA frequency was observed in ex-smokers with either the 312Asn or the 751Gln polymorphism compared to non-smokers with the same genotype. Our consistent observation of increased CA frequency among ex-smokers with the 312Asn polymorphism, both after NNK challenge and at baseline, indicates that this *XPD* polymorphism increases vulnerability to environmental insults. In addition, Seker et al. (2001) reported that lymphoblastoid cell lines with the 312Asn polymorphism exhibited a diminished apoptotic response following UV or ionizing radiation. It is possible that diminished apoptosis following NNK exposure in cells with the 312Asn polymorphism could allow the selective survival of genetically damaged cells, and thus explain further our observations of increased CAs in cells with this genotype.

Contrary to our observations of higher baseline CA frequency in males, we found that females with the 156A or the 751Gln polymorphism were more sensitive to the genotoxic effects of NNK. This increase in sensitivity to NNK among females is consistent with our previous observation (Hill *et al*. 2005b) that lymphocytes from females are more sensitive to the genotoxic effects of NNK than lymphocytes from males. Within the same gender, genetic influences could also affect the frequencies of CAs. In fact, we observed significantly higher NNK-induced CA frequencies in cells from females homozygous variant for the 751Gln polymorphism than in cells from females homozygous for the wild-type allele at codon 751. These observations indicate a gene-dose response and that the inheritance of two variant alleles at the *XPD* 751 site

confers an increased risk of genetic damage than the inheritance of one allele. In support of this observation, Qiao *et al*. (2002b) have shown, using the host cell reactivation assay, that cells with the homozygous 751Gln polymorphism had a reduced ability to reactivate a UV-irradiated plasmid than did cells with the homozygous wild-type. Collectively, our findings and the findings of Qiao *et al*. (2002b) could explain the observation that the homozygous 751Gln polymorphism confers an increased risk of lung squamous cell carcinoma (Liang *et al*., 2003).

To tease out more directly how these polymorphisms could individually influence CA frequency, influences of specific haplotype combinations were analyzed for response to NNK challenge in the mutagen-sensitivity assay after stratification for age. We observed that cells from younger individuals with only the 156A or the 751Gln polymorphisms had increased NNK-induced CA frequencies compared to cells from individuals in the same age group who were homozygous wild-type for all three *XPD* polymorphisms. These observations suggest that the 156A and the 751Gln polymorphisms are more influential among younger individuals in the sensitivity to NNK than the 312Asn polymorphism despite the observed influence of the 312Asn polymorphism on the response to smoking. Similar to our baseline CA analysis, we found that cells from older individuals homozygous wild-type for all three polymorphisms had significantly higher CA frequencies when challenged with NNK than cells from younger individuals with the same genotype. This increase in CA frequency in older individuals could be related to the possible accumulation of genetic damage over time. Because this was only found when both the older and younger populations were genetically similar, age appears to influence CA frequency among a genetically homogenous population.

However, among the same age group, *XPD* haplotype could influence the accumulation of genetic damage. Such age-related increases in CA frequencies could be due to the documented decrease in cellular DNA repair capacity in older individuals that is related to increased levels of DNA damage. This, consequently, confers an increased risk of developing cancers associated with environmental exposures (Wei *et al*., 1993; Barnett and King, 1995; Moriwaki *et al*., 1996; Wei, 1998). Collectively, the results of

this part of this project show that sensitivity to NNK was higher in smokers and in older subjects, and that *XPD* genotype played a role in further modifying the associations between age and smoking and increased CA frequencies.

FISH analysis

To increase the sensitivity of detecting chromosome aberrations, we also used fluorescence *in situ* hybridization (FISH). Two molecular probes, one hybridized to chromosome 1 and the other hybridized to chromosome 3, were used on a subset of the total population studied. These probes specifically target chromosome 1q12 and chromosome 3p11.1-q11.1 and were chosen for the following three reasons. (1) Chromosome 1 covers the largest portion of genetic material for a single chromosome. (2) Chromosome 3 has been implicated as a preferential target for tobacco carcinogens (Wu *et al*., 1998), and regional deletions in chromosome 3 are found in cells of individuals with smoking-associated lung cancer (Roche *et al*., 1996; Sozzi *et al*., 1996). (3) Many essential genes are located on chromosome 3, such as fragile histone triad (FRIT) whose deletion increases tumorigenesis (Zanesi *et al*., 2001), and OGG1 whose allelic deletion has been associated with lung cancer (Wikman *et al*., 2000). The FISH assay is a sensitive technique for detection of DNA breaks because it relies on unique probes that target specific chromosomal regions. This technique has been well-established by numerous laboratories for its ability to detect genetic damage in exposed individuals (Lohani *et al*., 2002; El-Zein and Abdel-Rahman, 2003; Wu *et al*., 2006).

In contrast to some previous investigations (Conforti-Froes *et al*., 1997; Abdel-Rahman *et al*., 1998), an increase in chromosome breaks in smokers compared to non-smokers was not observed in a small number of samples evaluated in the current project. This is most likely due to the small size of the groups evaluated in the current project. However, we did find a preferential increase in CA frequency on chromosome 3 compared to chromosome 1, both at baseline and following NNK treatment. This observation could be explained by the possibility that chromosome 3 posses genomic areas that are more prone to genetic damage than chromosome 1. This argument is

substantiated by earlier reports showing increased sensitivity of chromosome 3 to tobacco carcinogens (Wu *et al*., 1998), and the specific chromosome 3 deletions found at increased frequency in smokers with lung cancer (Roche *et al*., 1996; Sozzi *et al*., 1996).

Limitation of the CA analysis

A limitation of this part of the project was the small sample size for both haplotype analysis and FISH analysis. Future studies on this topic could be more definitive if group sizes were larger and if more detailed information was collected from the subjects.

mRNA ANAYLSIS

In an attempt to better understand the underlying mechanisms linking SNPs in susceptibility genes to health effects, researchers have recently tried to predict the phenotypic consequences of nonsynonymous polymorphisms on protein function using web-server based computer tools, such as the Sorting Intolerant From Tolerant (SIFT) (http://blocks.fhcrc.org/sift/SIFT.html) and PolyPphen (http://www.bork.embl-heidelberg.de/PolyPhen) programs. Using these programs, many SNPs in DNA repair genes were predicted by Xi *et al*. (2004), to affect protein function, but neither the Asp312Asn nor the Lys751Gln polymorphism in *XPD* was predicted to have an appreciable effect. Therefore, these types of predictions do not explain the epidemiological findings linking these two SNPs to altered DNA repair capacity and risk for disease (Qiao *et al*., 2002a; Rybicki *et al*., 2002; Spitz *et al*., 2002). Furthermore, these programs are not suitable for addressing potential effects of synonymous SNPs, such as the 156C>A polymorphism in exon 6 of the *XPD* gene.

In the current project, a novel mechanism that could explain the risk-modifying effects of these polymorphisms on disease susceptibility was explored. Specifically, we examined the relationships between SNPs in codons 156, 312 and 751 of *XPD* and their mRNA levels. Alterations in mRNA levels provide a potential mechanism by which a

polymorphism could ultimately alter DNA repair capacity and affect disease risk. The approach was to investigate effects of each of three prevalent polymorphisms on mRNA levels by choosing cells which had only one of these studied polymorphisms with the other two in their wild-type form. This approach allowed determining the effects of each SNP on mRNA levels independently. Also by studying cells with all three variants, we could determine the combined effect of these SNPs on mRNA levels. Results shown in Figure 22 indicate that all three SNPs studied, alone and in combination, altered mRNA levels.

Regulation of gene expression by *cis*-acting polymorphisms in promoter regions of genes has been studied extensively as a prevalent mechanism for altering mRNA levels (Wang and Sadee, 2006). In addition, SNPs in the coding regions of many genes have recently been reported to alter mRNA expression levels (Duan *et al*., 2003; Wang *et al*., 2005; Zhang *et al*., 2005). Results from this part of the current project provide the first information that *XPD* mRNA abundance is altered by both the synonymous SNP 156C→A (Arg156Arg in exon 6) and the non-synonymous SNPs G312A (Asp312Asn in exon 10) and A751C (Lys751Gln in exon 23). Our results are consistent with recent reports indicating that both synonymous and nonsynonymous SNPs in coding regions of other genes can alter mRNA levels by affecting mRNA stability, mRNA processing, or mRNA maturation (Duan *et al*., 2003; Johnson *et al*., 2005).

Polymorphism-associated alterations in mRNA abundance, as observed in the current project, could ultimately affect protein levels, and may thus impact protein activity and functions. Support for this assumption is provided by Duan *et al*. (2003), who recently reported that the synonymous 957T substitution in the human dopamine receptor *D2* gene (*DRD2*), rather than being "silent", led to decreased mRNA stability and resulted in a 50% decrease in levels of protein synthesis, an indication of decreased translation. Multiple other nonsynonymous SNPs have been shown to alter gene expression levels. For example, Malarstig *et al*. (2005) recently reported that the 5466A→G polymorphism in the tissue factor (*TF*) gene was associated with a significant decrease in mRNA expression levels and decreased TF protein activity. Similarly, Zhang

et al. (2005) reported that the 118A→G SNP in the human mu opioid receptor gene (*OPRM1*) caused significantly reduced yields of mRNA and receptor protein. These reports clearly indicate that both synonymous and nonsynonymous SNPs can exert potential pathophysiological and pharmacological effects by affecting mRNA transcription levels.

There are several mechanisms by which synonymous and nonsynonymous SNPs in the *XPD* gene could affect mRNA levels. One of these mechanisms is that SNPs alter the secondary mRNA structure. This structural alteration was reported to result in altered gene expression (van Leeuwen *et al*., 2005; Zhang *et al*., 2005). To investigate this issue, we performed a localized Mfold structure analysis. The results from our analysis indicated that the synonymous 156C→A (exon 6 polymorphism) and the nonsynonymous 751A→C (exon 23 polymorphism) SNPs resulted in significant changes in mRNA secondary structure and a change in single strandedness. These results indicate that both of these SNPs have the potential to affect local folding and, hence, mRNA stability. The changes we observed in the secondary structure of mRNA with the 156C→A and the 751A>C SNPs could affect mRNA maturation, mRNA half-life, or protein translation.

It is noteworthy that, while the *XPD* 312G→A polymorphism (Asp312Asn in exon 10) was associated with reduced mRNA levels, our Mfold structural analysis revealed that this SNP did not result in an apparent change in mRNA secondary structure. As such, it is possible that this polymorphism alters mRNA expression by a different mechanism. For example, the 312G→A polymorphism could affect mRNA stability, similar to the coding SNPs in other genes that were reported to exert this effect (Wang *et al*., 2005). In this context, Wang *et al*. (2005) has recently shown that the 3435C→T SNP in the multidrug resistance gene *MDR1 (ABC1)* decreased mRNA levels by affecting mRNA stability. Further studies are needed to clarify the underlying mechanisms by which SNPs in the *XPD* gene alter mRNA expression transcription levels.

Observations from this project indicated that the intensity of smoking altered *XPD* gene expression levels. Specifically, heavy smokers (defined in this project as smokers with greater than 12 pack yrs) had a statistically significant decrease in mRNA copy

number compared to non-smokers or light smokers. These findings are consistent with recent studies indicating that exposure of human peripheral blood mononuclear cells to cigarette smoke condensate or to the carcinogenic tobacco-specific nitrosamine NNK *in vitro* decreased mRNA levels of the DNA repair genes *XRCC1* and *ERCC5* (van Leeuwen *et al*., 2005). Although we found that *XPD* polymorphisms were associated with reduced mRNA levels, and smoking was also associated with reduced mRNA levels, we did not observe a three-way interaction between SNPs, smoking, and mRNA levels. This could be attributed to the small number of samples available for our stratified analysis. Future studies with a larger sample size are needed to identify potential three-way interactions.

An inverse relationship was found between age and levels of *XPD* mRNA. An interaction between age and the SNPs studied was also found, where a significant effect of each SNP was evident in the older group of subjects (>40 years old), but was less pronounced in younger subjects (<40 years old). These findings could be explained by the well-known age-related decreases in cellular DNA repair capacity reported in many studies (Wei *et al*., 1993; Barnett and King, 1995; Moriwaki *et al*., 1996; Wei *et al*., 1998; Carlini *et al*., 2001). For example, Hazane *et al*. (2006), have shown that cells from older individuals repaired DNA damage induced by UV irradiation much more slowly (60% slower) than in cells from younger individuals. Our observed age-related decrease in *XPD* mRNA levels could provide a mechanistic explanation for our prior finding that older subjects with the 312Asn allele have a higher background (baseline) level of chromosome aberrations (Affatato *et al*., 2004). Collectively our observation of interactions between age and decreased *XPD* mRNA in cells from older subjects with either single or multiple *XPD* polymorphisms indicated that any one of the three polymorphisms studied could exert a significant effect on *XPD* gene expression in older subjects.

Limitations of mRNA analysis

A limitation associated with the analysis of mRNA transcript copy number is an innate limitation of the technique itself. The use of reverse transcription real-time PCR to study mRNA copy number can only measure the total abundance of mRNA. No information is provided about the rate of mRNA transcription or the rate of mRNA degradation. Future evaluation of the rates of *XPD* transcription and degradation could more specifically explain why *XPD* polymorphisms were linked in this project, to decreases in mRNA abundance.

XPD PROTEIN ANALYSIS

The third major question addressed by this project was whether XPD expression is stimulated following exposure to UVC irradiation in a manner which can be quantitated by confocal fluorescence microscopy. We needed a reliable technique for this parameter to test our hypothesis that *XPD* polymorphisms influence protein expression as well as localization into the nucleus. Irradiation with UVC was used because UVC induces NER repairable DNA damage, such as 6,4-photoproducts and cyclobutane pyrimidine dimers at a greater frequency than UVB irradiation (Li et al., 2005). UVC also produces DNA damage that is more specifically repaired by NER than does challenge with NNK (Cipollini et al., 2006). Confocal fluorescent microscopy was used to address this question because this approach allows for a specific protein target to be visualized within a single cell or cellular compartment. Many recent studies have documented the versatility and quantitative powers of confocal fluorescent microscopy (Belichenko *et al*., 2005; Voss *et al*., 2005; Goentoro *et al*., 2006 Nadrigny *et al.*, 2006). For example, Bhattacharya et al, (2004) used laser scanning confocal fluorescent microscopy to visualize and quantitate the co-localization of various HIV-1 sub-type viral proteins with DNA. In the current project, UVC irradiation of cells was found to stimulate an increase in total XPD protein abundance by 30 min, with continued increase

up to 4 hrs after treatment. This increase in XPD protein levels after UVC exposure is consistent with the important role of XPD in NER.

To investigate whether the observed increase in XPD protein expression was mainly due to increased translation, mRNA levels were also measured at the same time points. Surprisingly, we found a significant drop in *XPD* mRNA gene copy number starting at 30 mins following UVC irradiation and still evident at 6 hrs after treatment. One possible explanation is that increased production of mRNA transcripts is not required for increased protein production. It is possible that protein translation of existing mRNA was increased, while new RNA transcription was decreased in an attempt to save cellular energy. These findings are in accordance with recent studies that have shown that mRNA levels and protein production are not always well-correlated with each other following cellular stress (Gygi *et al*., 1999; Ideker *et al*., 2001). For example, Hitzl *et al*. (2004) found no correlation between levels of mRNA expression of the MDR1 gene and expression of its corresponding gene product, P-glycoprotein. In order to determine whether the UVC irradiation decreased expression of other genes, we also assessed levels of mRNA for the constitutively expressed house-keeping gene β-actin. We found that the β-actin mRNA transcript copy number was also decreased following UV irradiation. This could suggest a generalized decrease in mRNA production for other genes after UV irradiation. Additional studies are still needed, however, to substantiate or refute this hypothesis.

Since XPD is known not to possess its own nuclear localization signal, the XPD protein was most likely carried into the nucleus by piggy-backing on another TFIIH protein that does have a characteristic nuclear localization signal (Boulikas, 1997; Santagati *et al*., 2001). This piggy-backing likely involves either p44 or MAT1 proteins with which XPD has direct contact. Our confocal microscopy analysis of XPD localization showed an increase in XPD protein within the nucleus from 30 min to 4 hrs following UV irradiation. The observed increase of XPD protein within the nucleus following UV damage suggests that the small amount of XPD protein initially within the nucleus is not sufficient for the greater workload of NER following UV damage. This is,

however, in disagreement with some previous reports. For example, in a report employing the host cell reactivation assay, human lymphocytes over-expressing XPD showed no discernable correlation between XPD protein levels and NER activity (Vogel *et al*., 2000; Xu *et al*., 2002). Another report by Santagati *et al*. (2001) used the expression of a GFP-XPD fusion protein to reveal that XPD appears to be primarily absent from the nucleus, and is evenly distributed throughout the cytoplasm of the cell, with only 5-30% of the fused XPD located within the nucleus. These investigators did not detect an increase in XPD localization to the nucleus after UV irradiation at 20 J/m^2, whereas XPB, XPC, p62, and p44 were all localized exclusively to the nucleus. These reports about a lack of correlation between XPD protein levels and NER activity indicate that XPD protein levels are not a determining factor in nucleotide excision repair capacity.

These reports did not, however, take into account the molar ratio of XPD to that of the other TFIIH complex proteins. As discussed previously, XPD activity and nuclear translocation are likely regulated by the interaction of XPD with p44 or Mat1 (Sung *et al*., 1993; Boulikas, 1997; Santigati *et al*., 2001). Therefore, it is possible that the over-expression of XPD will not increase its nuclear translocation or NER capacity if all the available p44 and Mat1 proteins are saturated. It is also possible that the use of tumor cell lines and assays, such as the host cell reactivation assay which does not rely on freshly isolated human cells, and which is known to be prone to interindividual variability, do not accurately convey what occurs in normal human cells. An additional possibility is that different cell types have varying degrees of baseline nuclear localized XPD that is sufficient for DNA repair. For example, Xu *et al*. (2002) found that XPD protein levels vary greatly, based on western blot analysis, by over 200-fold in 60 studied human tumor cell lines, with some cell lines having a barely detectable level of XPD. Therefore, cells or tissues with low baseline levels of XPD likely need to induce XPD synthesis and translocation into the nucleus in order to cope with a DNA-damaging event.

Our investigation of the influence of polymorphisms on nuclear localization of XPD found a nearly 50% increase in the nuclear localization of XPD protein after UVC

irradiation in cells that carried a variant allele for both the 312Asn and the 751Gln polymorphisms. This result, which was obtained with a small sample size (n=4), is counterintuitive, given the multiple reports of associations between the 312Asn or 751Gln polymorphisms and increased cancer risk. For example, Costa *et al*. (2006) reported that both the 315Asn and the 751Gln polymorphisms confer an increased risk for ovarian cancer. Although we observed higher levels of nuclear XPD protein in cells with a specific polymorphic genotype, (*i.e.* the combination of 312Asn and 751Gln), this high localization may not equate to greater NER capacity. It is conceivable that the combination of the 312Asn and 751Gln polymorphisms in XPD confers decreased protein function by affecting the structure or enzymatic activity of the resulting protein. Therefore, more of this poorly functional protein should accumulate in the nucleus when DNA damage has occurred. Future functional studies with larger sample sizes are needed to clarify the reasons behind the observed increases in XPD protein and to determine whether higher nuclear accumulation of the polymorphic XPD protein actually results in greater DNA repair activity.

Limitations of XPD protein analysis

A limitation in the analysis of protein expression and translocation was the small sample size. Our initial goal was to reduce sample variability by only using cells from female non-smokers for the protein evaluations. However, due to haplotype restrictions and subject recruitment difficulties, 6 of the 14 subjects used were either smokers or ex-smokers. Thus, smoking is likely a confounding factor that could have influenced the expression and translocations levels of the XPD protein. Future studies should reduce the number of confounding factors in order to enhance the likelihood of linking protein alterations to the presence of a single XPD polymorphism.

Fluorescent confocal microscopy was shown in this project to be very sensitive in its ability to visualize and quantitate both XPD protein expression and translocation. However, this technique is not well-suited for investigations with large numbers of samples because confocal microscopy is very time consuming and expensive. In

addition, instrumental variation from day to day confounds evaluations that are made on different days. The best way to overcome this limitation is to drastically reduce the sample variation. This could be achieved by switching from cells provided by human subjects, which inherently exhibit wide genetic variability, to genetically identical human cell lines that are transfected with specific *XPD* haplotypes of interest.

CONCLUSIONS

Both synonymous (C156A) and nonsynonymous (Asp312Asn and Lys751Gln) SNPs in the coding region of the *XPD* gene were found to be associated with three different measurable biological consequences. Polymorphisms in the *XPD* gene were associated with increased CAs at baseline. New information was also obtained indicating that SNPs in the *XPD* gene decrease mRNA levels. Such decreases could ultimately affect protein function and subsequent DNA repair activity. Our investigation provides mechanistic explanations for epidemiological studies indicating that these three SNPs are potential modifiers of DNA repair capacity and disease risk. The results also illustrate the importance of utilizing a comprehensive approach to investigate effects of other synonymous and nonsynonymous SNPs in DNA repair genes and in other genes, before arriving at definitive conclusions about the potential pathological consequences of such polymorphisms. UVC exposure was found to stimulate XPD protein expression and the increased translocation of this protein into the nucleus, where it could be directly involved in NER. However, this monitoring of XPD abundance and presence in the nucleus does not directly document its functional integrity. To further address the mechanism(s) responsible for the biological consequences of the XPD polymorphisms, future studies need to examine the ATPase and helicase activities of XPD polymorphisms. In addition, interactions of polymorphic XPD with p44 and Mat1 should also be examined, since these two proteins are directly involved in the control of XPD helicase activity.

Appendix

Table XII. Effect of smoking and genotype on in vivo CAs

	In vivo CAs	P-value
Group 1 (n=10)	0.20 ± 0.13	ref.
Group 2 (n=19)	0.89 ± 0.23	0.05
Group 3 (n=6)	0.89 ± 0.36	0.028
Group 4 (n=4)	0.50 ± 0.29	0.30
Group 5 (n=28)	0.82 ± 0.18	0.05

Group 1: C156C, Asp312Asp, and Lys751Lys (wild-type), Group 2: C156A, Asp312Asp, and Lys751Lys Group 3: C156C, Asp312Asn, and Lys751Lys, Group 4: C156C, Asp312Asp, and Lys751Gln, Group 5: C156A, Asp312Asn, and Lys751Gln.

Table XIII. Effect of age and genotype on CAs 24 hrs following NNK treatment

	24 hrs NNK mean CAs	P-value
Group 1 (n=11)	6.09 ± 0.97	0.017
Group 2 (n=13)	3.08 ± 0.68	ref.
Group 3 (n=20)	5.21 ± 0.57	0.025
Group 4 (n=8)	5.37 ± 0.68	0.037
Group 5 (n=28)	4.96 ± 0.41	0.020

Group 1: old (age> 39 years) C156C, Asp312Asp, and Lys751Lys (wild-type), Group 2: young (age ≤ years) C156C, Asp312Asp, and Lys751Lys (wild-type), Group 3: young C156A, Asp312Asp, and Lys751Lys, Group 4: young C156C, Asp312Asp, and Lys751Gln, Group 5: young C156C, Asp312Asn, and Lys751Gln. #, statistically significant P<0.05 compared to group 1.

Table XIV. Influence of XPD polymorphisms on mRNA transcript copy number

	Mean mRNA levels	P-value
Group 1 (n=17)	24427.41 ± 2483	ref.
Group 2 (n=20)	16595.55 ± 845	0.003
Group 3 (n=16)	13017.46 ± 1361	0.0004
Group 4 (n=19)	15610.13 ± 1230	0.002

Genotypes are Group 1: C156C, Asp312Asp, and Lys751Lys individuals (wild-type), Group 2: C156A, Asp312Asp, and Lys751Lys individuals, Group 3: C156C, Asp312Asn, and Lys751Lys individuals, Group 4: C156C, Asp312Asp, and Lys751Gln individuals.

Table XV. XPD genotype as a modifying factor in mRNA transcript copy number among smokers and nonsmokers

	Mean mRNA levels	P-value	
Non-smokers			
Group 1 (n=12)	24127.43 ± 3239	ref	
Group 2 (n=9)	16827.14 ± 1217	0.062	
Group 3 (n=7)	13915.52 ± 1950	0.038	
Group 4 (n=10)	18056.07 ± 1855	0.139	
Smokers			
Group 1 (n=5)	25147.39 ± 3800	ref	P=0.032
Group 2 (n=9)	15385.48 ± 1244	0.010	
Group 3 (n=9)	12318.97 ± 1954	0.0057	
Group 4 (n=9)	12892.42 ± 1067	0.0019	

Genotypes are Group 1: C156C, Asp312Asp, and Lys751Lys individuals (wild-type), Group 2: C156A, Asp312Asp, and Lys751Lys individuals, Group 3: C156C, Asp312Asn, and Lys751Lys individuals, Group 4: C156C, Asp312Asp, and Lys751Gln individuals.

Table XVI. XPD genotype as a modifying factor in mRNA transcript copy number among younger (age<40 yrs) and older (age>40 yrs) individuals

	Mean mRNA levels	P-value
Age < 40 yrs		
Group 1 (n=12)	24953.51 ± 3162	ref
Group 2 (n=9)	17379.55 ± 1172	0.050
Group 3 (n=7)	14112.07 ± 1643	0.050
Group 4 (n=10)	17610.75 ± 1654	0.067
Age > 40 yrs		
Group 1 (n=5)	23164.78 ± 4159	ref
Group 2 (n=9)	15159.02 ± 1163	0.034
Group 3 (n=9)	12519.91 ± 1863	0.016
Group 4 (n=9)	13387.22 ± 1612	0.022

Genotypes are Group 1: C156C, Asp312Asp, and Lys751Lys individuals (wild-type), Group 2: C156A, Asp312Asp, and Lys751Lys individuals, Group 3: C156C, Asp312Asn, and Lys751Lys individuals, Group 4: C156C, Asp312Asp, and Lys751Gln individuals.

Table XVII. Percent of BPB-excluding cells

	control non-exposed (%)	UV exposed (%)	P- value
before uv	98 ± 2	99 ± 1	0.698489
control	85.5 ± 0.5	90 ± 8	0.631036
30min	94 ± 2	80.5 ± .5	0.022534
2hr	87.5 ± 5.5	96.5 ± 1.5	0.255155
4hr	95.5 ± 1.5	98.5 ± 1.5	0.292893
6hr	96.5 ± 3.5	99 ± 1	0.563148

Replicate samples were sham-irradiated or UV irradiated at a dose rate of 400 μJ/cm^2 for 100 sec, and allowed to incubate from 0hr, 30min, 2hr, 4hr or 6hr. Percent of bromo-phenol blue excluding cells were calculated by light microscopy

Table XVIII. XPD protein abundance following UV irradiation

	Relative protein expression	P-value
control	733.55 ± 25.3	ref.
30min	863.28 ± 52.3	0.034
2hr	916.45 ± 62.1	0.011
4hr	1017.67 ± 35.5	6.75E-07
6hr	935.15 ± 56.4	0.0016

XPD protein expression over time following UV irradiation using a specific XPD C-20 goat polyclonal antibody (Snata Cruz Biotech, CA) and a chicken anti-goat polyclonal antibody conjugated with Alex Fluor 594 (Molecular Probes, OR). Cells were. Cells were irradiated at a dose rate of 400 μJ/cm^2 for 100 sec. XPD protein abundance from the average of 100 cells from each time point calculated using Metamorph imaging software. *P>0.05 when compared to the control (0 hr).

Table XIX. Genotype variation on XPD protein abundance measured by immunofluorescence

	Relative protein expression	P-value
Control		
all wt/wt all exons	683.09 ± 20.63	ref
wt/variant exon 10	659.29 ± 39.61	0.58
wt/variant exon 23	808.61 ± 35.76	0.023
wt/variant exons 10 & 23	783.42 ± 58.52	0.16
30 min		
all wt/wt all exons	915.15 ± 82.61	ref
wt/variant exon 10	1030.72 ± 112.27	0.43
wt/variant exon 23	655.86 ± 36.69	0.052
wt/variant exons 10 & 23	841.41 ± 93.89	0.58
2 hrs		
all wt/wt all exons	832.98 ± 63.40	ref
wt/variant exon 10	804.34 ± 101.23	0.81
wt/variant exon 23	798.02 ± 101.93	0.77
wt/variant exons 10 & 23	1172.84 ± 115.51	0.042
4 hrs		
all wt/wt all exons	1020.51 ± 78.54	ref
wt/variant exon 10	1074.76 ± 44.44	0.61
wt/variant exon 23	993.75 ± 86.2	0.83
wt/variant exons 10 & 23	980.72 ± 92.11	0.75
6 hrs		
all wt/wt all exons	1056.07 ± 90.55	ref
wt/variant exon 10	878.71 ± 88.71	0.23
wt/variant exon 23	668.15	N/A
wt/variant exons 10 & 23	911.47 ± 6.58	0.35

XPD protein abundance averaged from each genotype group over time following UV irradiation, and calculated from fluorescent confocal micrograph images using Metamorph imaging software. N/A only one sample in subgroup

Table XX. UV-induced XPD gene expression

	Transcript copy number	P-value
Control	50324.97 ± 2117	ref
30 min	28988.14 ± 1689	0.0029
2 hrs	28387.25 ± 1493	0.0021
4 hrs	27564.49 ± 2593	0.0029
6 hrs	24852.33 ± 236	0.0019

XPD mRNA transcript copy number over time following UV irradiation measured by reverse transcription real-time PCR using absolute quntitation.

Table XXI. Genotype variation in XPD gene expression following UV exposure

	Transcript copy number	P-value
all wt/wt		
Control	47803.14 ± 3187.86	ref
30 min	33628.95 ± 6215.60	0.089
2 hrs	29748.09 ± 2751.25	0.0052
4 hrs	24001.36 ± 2582.61	0.0011
6 hrs	25355.82 ± 665.24	0.00046
exon 10 wt/variant		
Control	53643.48 ± 4334.54	ref
30 min	26329.59 ± 931.00	0.0035
2 hrs	30762.48 ± 10773.65	0.12
4 hrs	24950.65 ± 2126.55	0.0040
6 hrs	25135.90 ± 506.73	0.0028
exon 23 wt/variant		
Control	45678.92 ± 2554.11	ref
30 min	29348.42 ± 6494.71	0.079
2 hrs	28997.49 ± 358.81	0.015
4 hrs	35240.43 ± 1132.07	0.054
6 hrs	24346.73 ± 4143.09	0.012
exon 10 and 23 wt/variant		
Control	54174.35 ± 5542.53	ref
30 min	26645.63 ± 1383.16	0.0029
2 hrs	24040.95 ± 1791.87	0.0021
4 hrs	26065.54 ± 1836.78	0.0029
6 hrs	24570.86 ± 1182.74	0.0019

XPD mRNA transcript copy number of cells from polymorphic individuals over time following UV irradiation measured by reverse transcription real-time PCR using absolute quntitation

Table XXII. β-actin mRNA expression following UV exposure

	mean 2^-□□CT vlaues		P-vale
	Non-exposed	UV-exposed	
Control	1.009 ± 0.01	1.152 ± 0.03	0.017
30 min	1.148 ± 0.01	0.231 ± 0.01	8.83E-07
2 hrs	1.288 ± 0.05	0.237 ± 0.04	2.50E-05
4 hrs	1.323 ± 0.03	0.179 ± 0.01	5.59
6 hrs	1.165 ± 0.04	0.145 ± 0.00	1.00E-05

Relative β-actin mRNA transcript copy number from samples over time, following sham-exposure or UV irradiation at a dose rate of 400 μJ/cm^2 for100 sec. Relative standardization using 18S RNA as an internal standard.

References

Aaltonen LA., Peltomaki P., Leach FS., Sistonen P., Pylkkanen L., Mecklin JP., Jarvinen H., Powell SM., Jen J., Hamilton SR. Clues to the pathogenesis of familial colorectal cancer. Science. 1993; 260: 812–816.

Abdel-Rahman SZ., El-Zein RA., Zwischenberger JB. and Au WW. Association of the NAT1*10 genotype with increased chromosome aberrations and higher lung cancer risk in cigarette smokers. Mutat. Res. 1998; 398: 43-54.

Abdel-Rahman SZ., Salama AS., Au WW., Hamada FA. Role of CYP2E1 and CYP2D6 allelic variants in NNK-induced chromosome aberration. Pharmacogenetics. 1999; 10: 1-11.

Abdel-Rahman SZ., El-Zein RA. The 399Gln polymorphism in the DNA repair gene XRCC1 modulates the genotoxic response induced in human lymphocytes by the tobacco-specific nitrosamine NNK. Cancer Lett. 2000; 159: 63-71.

Abdel-Rahman SZ., Soliman AS., Bondy ML., Omar S., El-Badawy SA., Khaled HM., Seifeldin IA., Levin B. Inheritance of the 194Trp and the 399Gln variant alleles of the DNA repair gene XRCC1 are associated with increased risk of early-onset colorectal carcinoma in Egypt. Cancer Lett. 2000; 159: 79-86.

Aboussekhra A., Biggerstaff M., Shivji MK., Vilpo JA., Moncollin V., Podust VN., Protic M., Hubscher U., Egly JM., Wood RD. Mammalian DNA nucleotide excision repair reconstituted with purified protein components. Cell. 1995; 80: 859–868.

Adimoolam S., Ford JM. p53 and DNA-damage-inducible expression of the xeroderma pigmentosum group C gene. Proc. Natl. Acad. Sci. U.S.A. 2002; 99: 12985–12990.

Affatato AA., Wolfe KJ., Lopez MS., Hallberg C., Ammenheuser MM., and Abdel-Rahman SZ. Effect of XPD/ERCC2 polymorphisms on chromosome aberration frequencies in smokers and on sensitivity to the mutagenic tobacco-specific nitrosamine NNK. Enviro. Mol. Mutagen. 2004; 44: 65-73.

Albrecht T., St. Jeor SC., Funk FD., Rapp F. Multiplicity reactivation of human cytomegalovirus inactivated by ultra-violet light. Int. J. Radiat. Bio. 1974; 26: 445-454.

American Cancer Society. Cancer reference information: What are the key statistics for lung cancer. http://www.cancer.org/docroot/CRI/content/CRI_2_4_1x_What_Are_the_Key_Statistic_About_Lung_Cancer_15.asp?sitearea=. 6/14/2006.

Araujo SJ., Tirode F., Coin F., Pospiech H., Syvaoja JE., Stucki M., Hubscher U., Egly JM., Wood RD. Nucleotide excision repair of DNA with recombinant human proteins: Definition of the minimal set of factors, active forms of TFIIH, and modulation by CAK. Genes Dev. 2000; 14: 349-359.

Asahina H., Kuraoka I., Shirakawa M., Morita EH., Miura N., Miyamoto I., Ohtsuka E., Okada Y., Tanaka K. The XPA protein is a zinc metalloprotein with an ability to recognize various kinds of DNA damage. Mutat. Res 1994; 315: 229–237.

Auranen A., Song H., Waterfall C., DiCioccio RA., Kuschel B., Kjaer SK., Hogdall E., Hogdall C., Stratton J., Whittemore AS., Easton DF., Ponder BAJ., Novik KL., Dunning AM., Gayther S., Pharoah PDP. Polymorphisms in DNA repair genes and epithelial ovarian cancer risk. Int. J. Cancer. 2005; 117: 611-618.

Baccarelli A., Calista D., Minghetti P., Marinelli B., Albetti B., Tseng T., Hedayati M., Grossman L., Landi G., Struewing JP., Landi MT. XPD gene polymorphism and host characteristics in the association with cutaneous malignant melanoma risk. Br. J. Cancer. 2004; 90: 497–502.

Bailly V., Sente B., Verly W. Bacteriophage-T4 and *Micrococcus luteus* UV endonucleases are not endonucleases but beta-elimination and sometimes beta delta-elimination catalysts. Biochem. J. 1989; 259: 751–759.

Barnett YA., King CM. An investigation of antioxidant status, DNA repair capacity and mutation as a function of age in humans. Mutat Res. 1995; 338: 115-128.

Bartsch H., Nair U., Risch A., Rojas M., Wikman H., Alexandrov K. Genetic polymorphism of CYP genes, alone or in combination, as a risk modifier of tobacco-related cancers. Cancer Epidemiol. Biomarkers Prev. 2000; 9: 3-28.

Belichenko PV., Dickson PI., Passage M., Jungles S., Mobley WC., Kakkis ED. Penetration, diffusion, and uptake of recombinant human alpha-L-iduronidase after intraventricular injection into the rat brain. Mol Genet Metab. 2005; 86: 141-149.

Bellacosa A. Functional interactions and signaling properties of mammalian DNA mismatch repair proteins. Cell Death Differ. 2001; 8: 1070-1092

Benhamou S., Sarasin A. *ERCC2*/*XPD* gene polymorphisms and cancer risk. Mutagenesis. 2002; 17: 463-469.

Benhamou S., Tuimala J., Bouchardy C., Dayer P., Sarasin A., Hirvonen A. DNA repair gene XRCC2 and XRCC3 polymorphisms and susceptibility to cancers of the upper aerodigestive tract. Int J Cancer. 2004; 112(5): 901-904.

Benhamou S., Sarasin A. ERCC2 /XPD gene polymorphisms and lung cancer: a HuGE review. Am J Epidemiol. 2005; 161: 1-14.

Bhattacharya B., Weiss RA., Davis C., Holmes H., Hockley D., Fassati A. Detection and quantitation of human immunodeficiency virus type-1 particles by confocal microscopy. J Virol Methods. 2004; 120: 13-21.

Bishop JO., Selman GG., Hickman J., Black L., Saunders RD., Clark AJ. The 45-kb unit of major urinary protein gene organization is a gigantic imperfect palindrome. Mol Cell Biol. 1985; 5: 1591–1600.

Blackwell LJ., Martik D., Bjornson KP., Bjomson ES., Modrich P. Nucleotide-promoted release of hMutSalpha from heteroduplex DNA is consistent with an ATP-dependent translocation mechanism. J. Biol. Chem. 1998; 273: 32055–32062.

Blasiak J., Arabski M., Krupa R., Wozniak K., Rykala J., Kolacinska A., Morawiec Z., Drzewoski J., Zadrozny M. Basal, oxidative and alkylative DNA damage, DNA repair efficacy and mutagen sensitivity in breast cancer. Mutat. Res. 2004; 554: 139-148.

Boer J., Hoeijmakers JH. Cancer from the outside, aging from the inside: mouse models to study the consequences of defective nucleotide excision repair. Biochemie. 1999; 81: 127-137.

Botta E., Nardo T., Lehmann AR., Egly JM., Pedrini AM., Stefanini M. Reduced level of the repair/transcription factor TFIIH in trichothiodystrophy. Hum Mol Genet. 2002; 11: 2919-2928.

Boulikas T. Nuclear import of DNA repair proteins. Anticancer Res. 1997; 17: 843-863.

Broughton BC., Steingrimsdottir H., Lehmann AR. Five polymorphisms in the coding sequence of the xeroderma pigmentosum group D gene. Mutat. Res. 1996; 362: 209–211.

Brown KD., Rathi A., Kamath R., Beardsley DI., Zhan Q., Mannino JL., Baskaran R. The mismatch repair system is required for S-phase checkpoint activation. Nat. Genet. 33: 2003; 80–84.

Bu D., Tomlinson G., Lewis C., Zhang C., Kildebeck E., Euhus D. An intronic polymorphism associated with increased xrcc1 expression, reduced apoptosis and familial breast cancer. Breast Cancer Research and Treatment. 2006: 1-9.

Bustin SA., Benes V., Nolan T., Pfaffl MW. Quantitative real-time rt-pcr - a perspective. J Mol Endocrinol. 2005; 34: 597-601.

Bustin SA., Mueller R. Real-time reverse transcription pcr (qrt-pcr) and its potential use in clinical diagnosis. Clin Sci (Lond). 2005; 109: 365-379.

Butkiewicz D., Rusin M., Enewold L., Shields PG., Chorazy M., Harris CC. Genetic polymorphisms in DNA repair genes and risk of lung cancer. Carcinogenesis. 2001; 22: 593–597.

Butkiewicz D., Popanda O., Risch A., Edler L., Dienemann H., Schulz V., Kayser K., Drings P., Bartsch H. and Schmezer P. Association between the risk for lung adenocarcinoma and a (-4) G-to-A polymorphism in the XPA gene. Cancer Epidemiol. Biomarkers Prev. 2004; 13: 2242–2246.

Caldecott K., Aoufouchi S., Johnson P., Sharll S. XRCC1 polypeptide interacts with DNA polymerase beta and possibly poly(ADP-ribose) polymerase, and DNA ligase III is a novel molecular 'nick-sensor' *in vitro*. Nucl. Acids Res. 1996; 24: 4387–4394.

Campalans A., Marsin S., Nakabeppu Y., O'connor TR., Boiteux S., Radicella JP. XRCC1 interactions with multiple DNA glycosylases: a model for its recruitment to base excision repair. DNA Repair (Amst). 2005; 4: 826-35.

Carlini DB., Chen Y., Stephan W. The relationship between third-codon position nucleotide content, codon bias, mRNA secondary structure and gene expression in the drosophilid alcohol dehydrogenase genes Adh and Adhr. Genetics. 2001; 159: 623-33.

Chang IY., Kim SH., Cho HJ., Lee DY., Kim MH., Chung MH., You Hj. Human AP endonuclease suppresses DNA mismatch repair activity leading to microsatellite instability. Nucl. Acids Res. 2005; 33: 5073–5081.

Chao EC., Lipkin SM. Molecular models for the tissue specificity of DNA mismatch repair-deficient carcinogenesis. Nucl. Acids Res. 2006; 34: 840-852.

Chen J., Larochelle S., Li X., and Suter, B. Xpd/Ercc2 regulates CAK activity and mitotic progression. Nature. 2003; 424: 228-232.

Chen L., Ambrosone CB., Lee J., Sellers TA., Pow-Sang J., Park JY. Association between polymorphisms in the DNA repair genes xrcc1 and ape1, and the risk of prostate cancer in white and black americans. The Journal of Urology. 2006; 175: 108-112.

Cheng L., Eicher SA., Guo Z., Hong WK., Spitz MR., Wei Q. Reduced DNA repair capacity in head and neck cancer patients. Cancer Epidemiol. Biomarkers Prev. 1998; 7: 465-468.

Cipollini M, He J, Rossi P, Baronti F, Micheli A, Rossi AM, Barale R. Can individual repair kinetics of UVC-induced DNA damage in human lymphocytes be assessed through the comet assay? Mutat Res. 2006; 601: 150-161

Clarkson SG., Wood RD. Polymorphisms in the human xpd (ercc2) gene, DNA repair capacity and cancer susceptibility: An appraisal. DNA Repair. 2005; 4: 1068-1074.

Cloutier JF., Drouin R., Weinfeld M., O'Connor TR., Castonguay A. Characterization and mapping of DNA damage induced by reactive metabolites of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) at nucleotide resolution in human genomic DNA. J. Mol. Biol. 2001; 313: 539-557.

Coin F., Bergmann E., Tremeau-Bravard A., Egly JM. Mutations in XPB and XPD helicases found in xeroderma pigmentosum patients impair the transcription function of TFIIH. EMBO J. 1999; 18: 1357-1366.

Coin F., Auriol J., Tapias A., Clivio P., Vermeulen W., Egly JM. Phosphorylation of XPB helicase regulates TFIIH nucleotide excision repair activity. EMBO J. 2004; 23: 4835-4846.

Coin F., De Santis LP., Nardo T., Zlobinskaya O., Stefanini M., Egly J-M. P8/ttd-a as a repair-specific tfiih subunit. Mol. Cell. 2006; 21: 215-226.

Conforti-Froes N., El-Zein RA., Abdel-Rahman SZ., Zwischenberger JB. and Au WW. Predisposing genes and increased chromosome aberrations in lung cancer cigarette smokers. Mutat. Res. 1997; 379: 53-59.

Costa S., Pinto D., Pereira D., Vasconcelos A., Afonso-Lopes C., Osorio T., Lopes C., Medeiros R. Importance of xeroderma pigmentosum group D polymorphisms in susceptibility to ovarian cancer. Cancer Lett. 2006; 3: [Epub ahead of print].

Cox MM. Historical overview: searching for replication help in all of the rec places. Proc. Natl. Acad. Sci. USA. 2001; 98: 8173–8180.

Critchlow SE., Jackson SP. DNA end-joining: from yeast to man. Trends Biochem. Sci. 1998; 23(10): 394-398.

D'Amours D., Jackson SP. The Mre11 complex: at the crossroads of DNA repair and checkpoint signaling. Nat. Rev. Mol. Cell Biol. 2002; 3: 317–327.

Daniels DS., Tainer JA. Conserved structural motifs governing the stoichiometric repair of alkylated DNA by O6-alkylguanine-DNA alkyltransferase. Mutat. Res. 2000; 460: 151-163.

Davies AA., Masson JY., McIlwraith MJ., Stasiak AZ., Stasiak A., Venkitaraman AR., West SC. Role of BRCA2 in control of the RAD51 recombination and DNA repair protein. Mol. Cell. 2001; 7: 273–282.

DeMarini DM. Genotoxicity of tobacco smoke and tobacco smoke condensate: a review. Mutat Res. 2004; 567: 447-74.

Demple B., Harrison L. Repair of oxidative damage to DNA: Enzymology and biology. Annu. Rev. Biochem. 1994; 63: 915-948.

Ding K., Zhou K., He F., Shen Y. LDA--a java-based linkage disequilibrium analyzer. Bioinformatics. 2003; 19: 2147-2148.

Dizdaroglu M. Measurement of radiation-induced damage to DNA at the molecular level. Int. J. Radiat. Bio. 1992; 61: 175–183.

Drablos F., Feyzi E., Aas PA., Vaagbo CB., Kavli B., Bratlie MS., Pena-Diaz J., Otterlei M., Slupphaug G., Krokan HE. Alkylation damage in DNA and rna--repair mechanisms and medical significance. DNA Repair. 2004; 3: 1389-1407.

Duan J., Wainwright MS., Comeron JM., Saitou N., Sanders AR., Gelernter J., Gejman PV. Synonymous mutations in the human dopamine receptor d2 (drd2) affect mRNA stability and synthesis of the receptor. Hum. Mol. Genet. 2003; 12: 205-216.

Duncan BK., Weiss B. Specific mutator effects of ung (uracil-DNA glycosylase) mutations in Escherichia coli. J Bacteriol. 1982; 151: 750–755.

Duncan T., Trewick SC., Koivisto P., Bates PA., Lindahl T., Sedgwick B. Reversal of DNA alkylation damage by two human dioxygenases. PNAS. 2002; 99: 16660-16665.

Dybdahl M., Vogel U., Frentz G., Wallin H., Nexo BA. Polymorphisms in the DNA repair gene xpd: Correlations with risk and age at onset of basal cell carcinoma. Cancer Epidemiol. Biomarkers Prev. 1999; 8: 77-81.

Eastmond DA., Schuler M., Rupa DS. Advantages and limitations of using fluorescence in situ hybridization for the detection of aneuploidy in interphase human cells. Mutat. Res.1995; 348: 153-162.

Edelmann W., Umar A., Yang K., Heyer J., Kucherlapati M., Lia M., Kneitz B., Avdievich E., Fan K., Wong E., Crouse G., Kunkel T., Lipkin M., Kolodner RD., Kucherlapati R. The DNA mismatch repair genes Msh3 and Msh6 cooperate in intestinal tumor suppression. Cancer Res. 2000; 60: 803-7.

Eger BT., Benkovic SJ. Minimal kinetic mechanism for misincorporation by DNA polymerase I (Klenow fragment). Biochemistry. 1992; 31: 9227–9236.

Eggler AL., Inman RB., Cox MM. The Rad51-dependent pairing of long DNA substrates is stabilized by replication protein A. J. Biol. Chem. 2002; 277: 39280–39288.

El-Zein R., Abdel-Rahman SZ., Conforti-Froes N., Alpard SK., Zwischenberger JB. Chromosome aberrations as a predictor of clinical outcome for smoking associated lung cancer. Cancer Lett. 2000; 158: 65-71.

El-Zein RA., Abdel-Rahman SZ. Detection of chromosomal aberrations in lung tissue and peripheral blood lymphocytes using interphase fluorescence in situ hybridization (FISH). Methods Mol Med. 2003; 75: 145-61.

Essers J., van Steeg H., de Wit J., Swagemakers SM., Vermeij M., Hoeijmakers JH., Kanaar R. Homologous and non-homologous recombination differentially affect DNA damage repair in mice. EMBO J. 2000; 19: 1703–1710.

Evans E., Moggs JG., Hwang JR., Egly JM., Wood RD. Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. EMBO J. 1997; 16: 6559–6573.

Eveno E., Bourre F., Quilliet X., Chevallier- Lagente O., Roza L., Eker A., Kleijer W., Nikaido O., Stefanini M., Hoeijmakers JHJ., Bootsma D., Cleaver JE., Sarasin A., Mezzina, M. Different removal of ultraviolet photoproducts in genetically related xeroderma pigmentosum and trichothiodystrophy diseases. Cancer Res. 1995; 55: 4325–4332.

Fang WH., Modrich P. Human strand-specific mismatch repair occurs by a bidirectional mechanism similar to that of the bacterial reaction. J. Biol. Chem. 1993; 268: 11838 – 11844.

Faul F., Erdfelder E. GPower: A priori, post-hoc, and compromise power analyses for MS-DOS [computer program]. Bonn FRG: Bonn University, Dept of Psychology, 1992.

Featherstone C., Jackson SP. DNA-dependent protein kinase gets a break: its role in repairing DNA and maintaining genomic integrity. Br. J. Cancer. 1999a; 80: 14–19.

Featherstone C., Jackson SP. Ku, a DNA repair protein with multiple cellular functions? Mutat. Res. 1999b; 434: 3– 15.

Fortini P., Pascucci B., Parlanti E., D'Errico M., Simonelli V., Dogliotti E. The base excision repair: Mechanisms and its relevance for cancer susceptibility. Biochimie. 2003; 85: 1053-1071.

Fu YP., Yu JC., Cheng TC., Lou MA., Hsu GC., Wu CY., Chen ST., Wu HS., Wu PE., Shen CY. Breast cancer risk associated with genotypic polymorphism of the nonhomologous end-joining genes: A multigenic study on cancer susceptibility. Cancer Res. 2003; 63: 2440-2446.

Fujita K., Kamataki T. Predicting the mutagenicity of tobacco-related N-nitrosamines in humans using 11 strains of *salmonella typhimurium* yg7108, each coexpressing a form of human cytochrome p450 along with nadph-cytochrome p450 reductase. Environ. Mol.Mutagen. 2001; 38: 339-346.

Galloway SM., Berry PK., Nichols WW., Wolman SR., Soper KA, Stolley PD., Archer P. Chromosome aberrations in individuals occupationally exposed to ethylene oxide, and in a large control population. Mutat. Res. 1986; 170: 55-74.

Garcia-Bustos J., Heitman J., Hall MN. Nuclear protein localization. Biochim Biophys Acta. 1991; 1071: 83-101.

Gazdar AF., Minna JD. Molecular detection of early lung cancer. J. National Cancer Inst. 1999; 91: 299-301.

Genschel J., Littman SJ., Drummond JT., Modrich P. Isolation of MutSbeta from human cells and comparison of the mismatch repair specificities of MutSbeta and MutSalpha. J. Biol. Chem. 1998; 273: 19895–19901.

Genschel J., Modrich P. Mechanism of 5_-directed excision in human mismatch repair. Mol. Cell. 2003; 12: 1077–1086.

Gerson SL. MGMT: its role in cancer aetiology and cancer therapeutics. Nat. Rev. Cancer. 2004; 4: 296–307.

Giglia-Mari G., Coin F., Ranish JA., Hoogstraten D., Theil A., Wijgers N., Jaspers NGJ., Raams A., Argentini M., van der Spek PJ., Botta E., Stefanini M., Egly J-M., Aebersold R., Hoeijmakers JHJ., Vermeulen W. A new, tenth subunit of tfiih is responsible for the DNA repair syndrome trichothiodystrophy group a. Nat. Genet. 2004; 36: 714-719.

Goentoro LA., Yakoby N., Goodhouse J., Schupbach T., Shvartsman SY. Quantitative analysis of the GAL4/UAS system in Drosophila oogenesis. Genesis. 2006; 44: 66-74.

Goode EL., Ulrich CM., Potter JD. Polymorphisms in DNA repair genes and associations with cancer risk. Cancer Epidemiol. Biomarkers Prev. 2002; 11: 1513-1530.

Gradia S., Subramanian D., Wilson T., Acharya S., Makhov A., Griffith J., Fishel R. hMSH2-hMSH6 forms a hydrolysis-independent sliding clamp on mismatched DNA. Mol. Cell. 1999; 3: 255–261.

Grawunder U., Wilm M. Wu X., Kulesza P., Wilson TE., Mann M., Lieber MR. Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. Nature. 1997; 388: 492–495.

Griffin CS., Thacker J. The role of homolgous recombination repair in the formation of chromosome aberrations. Cytognet Genome Res. 2004; 104: 21-27.

Gryfe R., Kim H., Hsieh ET., Aronson MD., Holowaty EJ., Bull SB., Redston M., Gallinger S. Tumor microsatellite instability and clinical outcome in young patients with colorectal cancer. N. Engl. J. Med. 2000; 342: 69-77.

Gryfe R., Gallinger S. Microsatellite instability, mismatch repair deficiency, and colorectal cancer. Surgery. 2001; 130: 17-20.

Guzder SM., Sung P., Prakash L., Prakash S. Nucleotide excision repair in yeast is mediated by sequential assembly of repair factors and not by a pre-assembled repairosome. J. Biol. Chem. 1996; 271: 8903-8910.

Gygi SP., Rochon Y., Franza BR., Aebersold R. Correlation between protein and mRNA abundance in yeast. Mol. Cell Biol. 1999;1:1720-30.

Haber JE. The many interfaces of Mre11. Cell. 1998; 95: 583–586.

Hadi MZ., Coleman MA., Fidelis K., Mohrenweiser HW., Wilson DM. Functional characterization of Ape1 variants identified in the human population. Nucl. Acids Res. 2000; 28: 3871–3879.

Hagmar L., Brogger A., Hansteen IL., Heim S., Hogstedt B., Knudsen L., Lambert B., Linnainmaa K., Mitelman F., Nordenson I. Cancer risk in humans predicted by increased levels of chromosomal aberrations in lymphocytes: Nordic study group on the health risk of chromosome damage. Cancer Res. 1994; 54: 2919-2922.

Hanawalt PC. DNA repair, The bases for Cockayne syndrome. Nature. 2000; 405: 415–416.

Hazane F., Sauvaigo S., Douki T., Favier A., Beani JC. Age-dependent DNA repair and cell cycle distribution of human skin fibroblasts in response to uva irradiation. Journal of Photochemistry and Photobiology Biology. 2006; 82: 214-23.

Hazra TK., Roy R., Biswas T., Grabowski DT., Pegg AE., Mitra S. Specific recognition of O6-methylguanine in DNA by active site mutants of human O6-methylguanine-DNA methyltransferase. Biochemistry. 1997; 36: 5769-5776.

Hazra T., Izumi T., Boldogh I., Imhoff B., Kow Y., Jaruga P., Dizdaroglu M., Mitra S. Identification and characterization of a human DNA glycosylase for repair of modified bases in oxidatively damaged DNA. Proc. Natl. Acad. Sci. USA. 2000a; 99: 3523–3528.

Hazra T., Kow Y., Hatahet Z., Imhoff,B., Boldogh I., Mokkapati S., Mitra S., Izumi T., Identification and characterization of a novel human DNA glycosylase for repair of cytosine-derived lesions. J. Biol. Chem. 2000b; 277: 30417–30420.

He Z., Henricksen LA., Wold MS., Ingles CJ. RPA involvement in the damage-recognition and incision steps of nucleotide excision repair. Nature. 1995; 374: 566 – 569.

Hecht SS. Biochemistry, biology, and carcinogenicity of tobacco-specific *N*-nitrosamines. Chem. Res. Toxicol. 1998; 11: 559–603.

Hecht SS. DNA adduct formation from tobacco-specific N-nitrosamines. Mutat. Res. 1999; 424: 127–142.

Hemminki K., Xu G., Angelini S., Snellman E., Jansen CT., Lambert B., Hou S-M. Xpd exon 10 and 23 polymorphisms and DNA repair in human skin in situ. Carcinogenesis. 2001; 22: 1185-1188.

Henning KA., Li L., Iyer N., McDaniel L., Reagan MS., Legerski R., Schultz RA., Stefanini M., Lehmann AR., Mayne LV., Friedberg EC. The Cockayne syndrome group A gene encodes a WD repeat protein that interacts with CSB protein and a subunit of RNA polymerase II TFIIH. Cell. 1995; 82: 555–564.

Hey T., Lipps G., Krauss G. Binding of XPA and RPA to damaged DNA investigated by fluorescence anisotropy. Biochemistry. 2001; 40: 2901–2910.

Hill JW., Hazra TK., Izumi T., Mitra S. Stimulation of human 8-oxoguanine-DNA glycosylase by AP-endonuclease: potential coordination of the initial steps in base excision repair. Nucl. Acids Res. 2001; 29: 430–438.

Hill CE., Wickliffe JK., Wolfe KJ., Kinslow CJ., Lopez MS., Abdel-Rahman SZ. The L84F and the I143V polymorphisms in the O6-methylguanine-DNA-methyltransferase (MGMT) gene increase human sensitivity to the genotoxic effects of the tobacco-specific nitrosamine carcinogen NNK. Pharmacogenet Genomics. 2005a; 15: 571-578.

Hill CE., Affatato AA., Wolfe KJ., Lopez MS., Hallberg CK., Canistro D., Abdel-Rahman SZ. Gender differences in genetic damage induced by the tobacco-specific nitrosamine NNK and the influence of the Thr241Met polymorphism in the XRCC3 gene. Environ. Mol. Mutagen. 2005b; 46: 22-29.

Hitzl M., Schaeffeler E., Hocher B., Slowinski T., Halle H., Eichelbaum M., Kaufmann P., Fritz P., Fromm MF., Schwab M. Variable expression of P-glycoprotein in the human placenta and its association with mutations of the multidrug resistance 1 gene (MDR1, ABCB1). Pharmacogenetics. 2004; 14: 309-318.

Hoeijmakers JH. Nucleotide excision repair. II: From yeast to mammals. Trends Genet. 1993; 9: 211-217.

Hofseth LJ., Khan MA., Ambrose M., Nikolayeva O., Xu-Welliver M., Kartalou M., Hussain SP., Roth RB., Zhou X., Mechanic LE., Zurer I., Rotter V., Samson LD., Harris CC. The adaptive imbalance in base excision-repair enzymes generates microsatellite instability in chronic inflammation. J. Clin. Invest. 2003; 112: 1887–1894.

Hoogervorst EM., van Steeg H., de Vries A. Nucleotide excision repair- and p53-deficient mouse models in cancer research. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis. 2005; 574: 3-21.

Hopfner KP., Karcher A., Shin DS., Craig L., Arthur LM., Carney JP., Tainer JA. Structural biology of Rad50 ATPase: ATP-driven conformational control in DNA double-strand break repair and the ABC-ATPase superfamily. Cell. 2000; 101: 789–800.

Hopfner KP., Karcher A, Craig L., Woo TT., Carney JP., Tainer JA. Structural biochemistry and interaction architecture of the DNA double-strand break repair Mre11 nuclease and Rad50-ATPase. Cell. 2001; 105: 473–485.

Hoy CA., Thompson LH., Mooney CL., Salazar EP. Defective DNA cross-link removal in Chinese hamster cell mutants hypersensitive to bifunctional alkylating agents. Cancer Res. 1985; 45: 1737–1743.

Hsu TC., Johnson DA., Cherry LM., Ramkisson D., Schantz SP., Jessup JM., Winn JR., Shirley L., Furlong C. Sensitvity to genotoxic effects of bleomycin in humans: possible relationship to environmental carcinogenesis. Int. J. Cancer. 1989; 43: 403-409.

Hu JJ., Hall MC., Grossman L, Hedayati M., McCullough DL., Lohman K., Case LD. Deficient nucleotide excision repair capacity enhances human prostate cancer risk, Cancer Res. 2004; 64: 1197–1201.

Hung RJ., Hall J., Brennan P., Boffetta P. Genetic polymorphisms in the base excision repair pathway and cancer risk: A huge review. Am. J. Epidemiol. 2005; 162: 925-942.

Hwang BJ., Ford JM., Hanawalt PC., Chu G. Expression of the p48 xeroderma pigmentosum gene is p53-dependent and is involved in global genomic repair. Proc. Natl. Acad. Sci. USA. 1999; 96: 424–428.

Ideker T., Thorsson V., Ranish JA., Christmas R., Buhler J., Eng JK., Bumgarner R., Goodlett DR., Aebersold R., Hood L. Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. Science. 2001; 292: 929-934.

Iliakis G., Wang H., Perrault AR., Boecker W., Rosidi B., Windhofer F., Wu W., Guan J., Terzoudi G., Pantelias G. Mechanisms of DNA double strand break repair and chromosome aberration formation. Cytogen. Genome Res. 2004; 104: 14-20.

Impellizzeri KJ., Anderson B., Burgers PM. The spectrum of spontaneous mutations in a *Saccharomyces cerevisiae* uracil-DNA-glycosylase mutant limits the function of this enzyme to cytosine deamination repair. J. Bacteriol. 1991; 173: 6807–6810.

Itin PH., Pittelkow MR. Trichothiodystrophy: review of sulfur-deficient brittle hair syndromes and association with the ectodermal dysplasias. J. Amer. Acad. Dermat 1990; 22: 705–717.

Ito H., Matsuo K., Hamajima N., Mitsudomi T., Sugiura T., Saito T., Yasue T., Lee K-M., Kang D., Yoo K-Y., Sato S., Ueda R., Tajima K. Gene-environment interactions between the smoking habit and polymorphisms in the DNA repair genes, ape1 asp148glu and xrcc1 arg399gln, in Japanese lung cancer risk. Carcinogenesis. 2004; 25: 1395-1401.

Izumi T., Wiederhold LR., Roy G., Roy R., Jaiswal A., Bhakat KK., Mitra S., Hazra TK. Mammalian DNA base excision repair proteins: Their interactions and role in repair of oxidative DNA damage. Toxicology. 2003; 193: 43-65.

Jackson SP. Sensing and repairing DNA double-strand breaks. Carcinogenesis. 2002; 23: 687–696.

Jans DA., Xiao CY., Lam MH. Nuclear targeting signal recognition: a key control point in nuclear transport? Bioessays. 2000; 22: 532-544.

Jeggo PA. Identification of genes involved in repair of DNA double-strand breaks in mammalian cells. Radiat. Res. 1998; 150: S80–S91

Jemal A., Tiwari RC., Murray T., Ghafoor A., Samuels A., Ward E., Feuer EJ., Thun MJ. American Cancer Society. Cancer statistics. CA Cancer J. Clin. 2004; 54: 8–29.

Jeon HS., Kim KM., Park SH., Lee SY., Choi JE., Lee GY., Kam S., Park RW., Kim IS., Kim CH., Jung TH., Park JY. Relationship between *XPG* codon 1104 polymorphism and risk of primary lung cancer. Carcinogenesis. 2003; 24: 1677–1681.

Johnson AD., Wang D., Sadee W. Polymorphisms affecting gene regulation and mRNA processing: broad implications for pharmacogenetics. Pharmacol Ther. 2005; 106: 19-38.

Jung EG., Bantle K. Xeroderma pigmentosum and pigmented xerodermoid. Birth Defects Orig. Artic. Ser. 1971; 7: 125-128.

Kadouri L., Kote-Jarai Z., Hubert A., Durocher F., Abeliovich D., Glaser B., Hamburger T., Eeles RA., Peretz T. A single-nucleotide polymorphism in the RAD51 gene modifies breast cancer risk in BRAC2 carriers, but not in BRAC1 carriers or noncarriers. Brit. J Cancer. 2004; 90: 2002-2005.

Kastan MB., Zhan Q., El-Deiry WS., Carrier F., Jacks T., Walsh WV., Plunkett BS., Vaogelstein B. A mammalian cell cycle checkpoint pathways utilizing p53 and GADD45 is defective in Ataxia-Telangiectasia. Cell. 1992; 71: 587-597.

Kelsey KT., Christiani DC., Little JB., Enhancement of benzo[a]pyrene-induced suister chromatid exchanges in lymphocytes from cigarette smokers occupationally exposed to asbestos. J. Natl. Cancer Inst. 1986; 77: 321-327.

Kim K., Biade S., Matsumoto Y. Involvement of flap endonuclease 1 in base excision DNA repair. J. Biol. Chem. 1998; 273: 8842–8848.

Klungland A., Laake K., Hoff E., Seeberg E. Spectrum of mutations induced by methyl and ethyl methanesulfonate at the hprt locus of normal and tag expressing Chinese hamster fibroblasts. Carcinogenesis. 1995; 16: 1281–1285.

Kohno T., Shinmura K., Tosaka M., Tani M., Kim SR., Sugimura H., Nohmi T., Kasai H., Yokota J. Genetic polymorphisms and alternative splicing of the *hOGG1* gene, that is

involved in the repair of 8-hydroxyguanine in damaged DNA. Oncogene 1998; 16: 3219–3225

Kraemer KH. Sunlight and skin cancer: another link revealed. Proc. Natl. Acad. Sci. USA. 1997; 94: 11–14.

Kubota Y., Nash RA., Klungland A., Schar P., Barnes DE., Lindahl T. Reconstitution of DNA base excision-repair with purified human proteins: interaction between DNA polymerase beta and the XRCC1 protein. EMBO J. 1996; 15: 6662–6670.

Kuchta RD., Benkovic P., Benkovic SJ. Kinetic mechanism whereby DNA polymerase I (Klenow) replicates DNA with high fidelity. Biochemistry. 1988; 27: 6716–6725.

Kuschel B., Auranen A., McBride S., Novik KL., Antoniou A., Lipscombe JM., Day NE., Easton DF., Ponder BAJ., Pharoah PDP., Dunning A. Variants in DNA double-strand break repair genes and breast cancer susceptibility. Hum. Mol. Genet. 2002; 11: 1399-1407.

Landi S., Norppa H., Frenzilli G., Cipollini G., Ponzanelli I., Barale R., Hirvonen A. Individual sensitivity to Cytogenetic effects of 1,2:3,4-diepoxybutane in cultured human lymphocytes: influence of glutathione S-transferase M1, P1 and T1 genotypes. Pharmacogenetics. 1998; 8: 461-471.

Lao Y., Gomes XV., Ren Y., Taylor JS., Wold MS. Replication protein A interactions with DNA. III. Molecular basis of recognition of damaged DNA. Biochemistry. 2000; 39: 850–859.

Leadon SA., Cooper PK. Preferential repair of ionizing-radiation induced damage in the transcribed strand of an active human gene is defective in Cockayne syndrome. Proc. Natl. Acad. Sci. USA. 1993; 90: 10499–10503.

Lee S., Elenbase B., Levine A., Griffith J. p53 and its 14kDa C-terminal domain recognize primary DNA damage in the form of insertion/deletion mismatchs. Cell. 1995; 81: 1013-1020.

Lee JJ., Trizna Z., Hsu TC., Spitz MR., Hong WK. A statistical analysis of the reliability and classification error in application of the mutagen sensitivity assay. Cancer Epidemiol Biomarkers Prev. 1996; 5: 191-197.

Leech RW., Brumback RA., Miller RH., Otsuka F., Tarrone RE., Robbins JH. Cockayne syndrome: clinicopathologic and tissue culture studies of affected siblings. J. Neuropathol. Exp. Neurol. 1985; 44: 507–519.

Lehmann AR., Kirk-Bell S., Mayne L. Abnormal kinetics of DNA synthesis in ultraviolet light irradiated cells from patients with Cockayne syndrome. Cancer Res. 1979; 3: 4237-4241

Lehmann AR. Three complementation groups in Cockayne syndrome. Mutat. Res. 1982; 106: 347–356.

Lehmann AR.. The xeroderma pigmentosum group D (XPD) gene: one gene, two functions, three diseases. Genes and Development. 2001; 15: 15-23.

Lehmann AR. DNA repair-deficient diseases, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. Biochimie. 2003; 85: 1101–1111.

Li GM., Modrich P. Restoration of mismatch repair to nuclear extracts of H6 colorectal tumor cells by a heterodimer of human MutL homologs. Proc. Natl. Acad. Sci. USA. 1995; 92: 1950– 1954.

Li RY., Calsou P., Jones CJ., Salles B. Interactions of the transcription/DNA repair factor TFIIH and XP repair proteins with DNA lesions in a cell-free repair assay. J. Mol. Biol. 1998; 281: 211–218.

Li H, Chang TW, Tsai YC, Chu SF, Wu YY, Tzang BS, Liao CB, Liu YC. Colcemid inhibits the rejoining of the nucleotide excision repair of UVC-induced DNA damages in Chinese hamster ovary cells. Mutat Res. 2005; 588: 118-128.

Liang G., Xing D., Miao X., Tan W., Yu C., Lu W., Lin D. Sequence variations in the DNA repair gene xpd and risk of lung cancer in a chinese population. Int. J. Cancer. 2003; 105: 669-673.

Lichtman JW., Conchello JA. Fluorescence microscopy. Nature Methods. 2005; 2: 910-919.

Lieber MR. Site-specific recombination in the immune system. FASEB J. 1991; 5: 2934 –2944.

Livingstone LR., White A., Sprouse J., Livanos E., Jacks T., Tlsty TD. Altered cell cycle arrest and gene amplification potential accompany loss of wild-type p53. Cell. 1992; 70: 932-935.

Lodovici M., Luceri C., Guglielmmi F., Bacci C., Akpan V., Fonnesu M.L., Boddi V., Dolara P. Benzo(a)pyrene diolepoxide (BPDE)-DNA adduct levels in leukocytes of smokers in relation to polymorphisms of CYP1A1, GSTM1, GSTP1, GSTT1, and mEH. Cancer Epidemiol. Biomarkers Prev. 2004; 13: 1342-1348.

Lohani M., Dopp E., Becker HH., Seth K., Schiffmann D., Rahman Q. Smoking enhances asbestos-induced genotoxicity, relative involvement of chromosome 1: a study using multicolor FISH with tandem labeling. Toxicol Lett. 2002; 136: 55-63.

Lovatt T., Alldersea J., Lear JT., Hoban PR., Ramachandran S., Fryer AA., Smith AG., Strange RC. Polymorphism in the nuclear excision repair gene ercc2/xpd: Association between an exon 6-exon 10 haplotype and susceptibility to cutaneous basal cell carcinoma. Human Mutation. 2005; 25: 353-359.

Lucci-Cordisco E., Zito I., Gensini F., Genuardi M. Hereditary nonpolyposis colorectal cancer and related conditions. Am. J. Med. Genet. A. 2003; 122: 325-334.

Lynch HT. Smyrk T. Hereditary nonpolyposis colorectal cancer (Lynch syndrome). An updated review. Cancer. 1996; 78: 1149-1167.

Lynch HT. Inherited susceptibility to cancer: clinical, predictive and ethical perspectives. Gut. 1999; 44: 765B-765.

Madhusudan S., Middleton MR. The emerging role of DNA repair proteins as predictive, prognostic and therapeutic targets in cancer. Cancer Treatment Reviews. 2005; 31: 603-617.

Malarstig A., Tenno T., Johnston N., Lagerqvist B., Axelsson T., Syvanen AC., Wallentin L., Siegbahn A. Genetic variations in the tissue factor gene are associated with clinical outcome in acute coronary syndrome and expression levels in human monocytes. Arteriosclerosis. Thrombo. & Vas. Biol. 2005; 25: 2667-2672.

Margison GP., Santibanez-Koref MF. O6-alkylquanine-DNA alkyltransferase: role in carcinogenesis and chemotherapy. BioEssays. 2002: 24: 255-266.

Marionnet C., Armier J., Sarasin A., Stary A. Cyclobutane pyrimidine dimmers are the main mutagenic DNA photoproducts in DNA repair-deficient trichothiodystrophy cells. Cancer Res. 1998; 58: 102–108.

Marsischky GT., Filosi N., Kane MF., Kolodner R. Redundancy of Saccharomyces cerevisiae MSH3 and MSH6 in MSH2-dependent mismatch repair. Genes Dev. 1996; 10: 407–420.

Martin FL., Cole KJ., Muir GH., Kooiman GG., Williams JA., Sherwood RA., Grover PL., Phillips DH. Primary cultures of prostate cells and their ability to activate carcinogens. Prostate Cancer Prostatic Dis. 2002; 5: 96–104.

Mathews DH., Sabina J., Zuker M., Turner DH. Expanded Sequence Dependence of Thermodynamic Parameters Improves Prediction of RNA Secondary Structure
J. Mol. Biol. 1999; 288: 911-940.

Matsuda T., Saijo M., Kuraoka I., Kobayashi T., Nakatsu Y., Nagai A., Enjojim T., Masutani C., Sugasawa K., Hanaoka F. DNA repair protein XPA binds replication protein A (RPA). J Biol Chem. 1995; 270: 4152– 4157.

Matta JL., Villa JL., Ramos JM., Sanchez J., Chompre G., Ruiz A., Grossman L. DNA repair nonmelanoma skin cancer in Puerto Rican populations, J. Am. Acad. Dermatol. 2003; 49: 433–439.

Mayne LV., Lehmann AR. Failure of RNA synthesis to recover after UV irradiation: An early defect in cells from individuals with Cockayne's syndrome and xeroderma pigmentosum. Cancer Res. 1982; 42: 1473–1478.

Mijal RS., Thomson NM, Fleischer NL., Pauly GT., Moschel RC., Kanugula S., Fang Q., Pegg AE., Peterson LA. The repair of the tobacco specific nitrosamine derived adduct O6-[4-Oxo-4-(3-pyridyl)butyl]guanine by O6-alkylguanine-DNA alkyltransferase variants. Chem Res Toxicol. 2004; 17: 424-434.

Milillo C.P., Gemignani F., Sbrana I., Carrozzi L., Viegi G., Barale R. Chromosome aberrations in humans in relation to site of residence. Mutat. Res. 1996; 360: 173-179.

Minsky M. Memoir on inventing the confocal microscope. Scanning. 1988; 10: 128–138.

Missura M., Buterin T., Hindges R., Hubscher U., Kasparkova J., Brabec V., Naegeli H. Double-check probing of DNA bending and unwinding by XPA-RPA: An architectural function in DNA repair. EMBO J. 2001; 20:3554–3564.

Mocellin S., Rossi CR., Pilati P., Nitti D., Marincola FM. Quantitative real-time PCR: A powerful ally in cancer research. Trends in Molecular Medicine. 2003; 9: 189-195.

Moriwaki S., Ray S., Tarone RE., Kraemer KH., Grossman L. The effect of donor age on the processing of UV-damaged DNA by cultured human cells: reduced DNA repair capacity and increased DNA mutability. Mutat Res. 1996; 364: 117-123.

Mu D., Park CH., Matsunaga T., Hsu DS., Reardon JT., Sancar A. Reconstitution of human DNA repair excision nuclease in a highly defined system. J Biol Chem. 1995; 270: 2415–2418.

Nance MA., Berry SA. Cockayne syndrome: review of 140 cases. Am. J. Med. Genet. 1992; 42: 68–84.

Nadrigny F., Rivals I., Hirrlinger PG., Koulakoff A., Personnaz L., Vernet M., Allioux M., Chaumeil M., Ropert N., Giaume C., Kirchhoff F., Oheim M. Detecting fluorescent protein expression and co-localisation on single secretory vesicles with linear spectral unmixing. Eur Biophys J. 2006; 35: 533-47.

Neddermann P., Gallinari P., Lettieri T., Schmid D., Truong O., Hsuan JJ., Wiebauer K., Jiricny J. Cloning and expression of human G/T mismatch-specific thymine-DNA glycosylase. J Biol Chem. 1996; 271: 12767–12774.

Norppa H. Cytogenetic markers of susceptibility: influence of polymorphic carcinogen-metabolizing enzymes. Environ. Health Perspect. 1997; 105: 829-835.

Norppa H. Cytogenetic Biomarkers. IARC Sci. Publ. 2004; 157: 179-205.

Obe G., Heler W.D., Vogt H.J. Mutagenic activity of cigarette smoke, G. Obe (ed). Mutation in Man, Springer-Verlag, Berlin, Germany, 1984; 223-246.

O'Connor PJ., Capps MJ., Greig AW. Comparative studies of the hepatocarcinogen N,N-dimethylnitrosamine in vivio: reaction sites in rat liver DNA and the significance of their relative stabilities. Br J. Cancer. 1973; 27: 153-166

O'Donovan A., Davies AA., Moggs JG., West SC., Wood RD. XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair. Nature. 1994; 371: 432– 435.

Pan ZQ., Reardon JT., Li L., Flores-Rozas H., Legerski R., Sancar A., Hurwitz J. Inhibition of nucleotide excision repair by the cyclindependent kinase inhibitor p21. J. Biol. Chem. 1995; 270: 22008 –22016.

Peltomaki P., Vasen HF. Mutations predisposing to hereditary nonpolyposis colorectal cancer: database and results of a collaborative study. The International Collaborative Group on Hereditary Nonpolyposis Colorectal Cancer. Gastroenterology. 1997; 113: 1146-1158.

Peltomaki P. Deficient DNA mismatch repair: a common etiologic factor for colon cancer. Hum. Mol. Genet. 2001; 10: 735-740.

Peltomaki P. Role of DNA mismatch repair defects in the pathogenesis of human cancer. J. Clin. Oncol. 2003; 21:1174-1179.

Pfeiffer P., Goedecke W., Kuhfittig-Kulle S., Obe G. Pathways of DNA double-strand break repair and their impact on the prevention and formation of chromosomal aberrations. Cytogen. Gen. Res. 2004; 104: 7-13.

Pinkel D., Straume T., Gray J.W. Cytogenetic analysis using quantitative, high-sensitivity, Fluorescence hybridization. Proc. Natl. Acad. Sci. USA. 1986; 83: 2934-2938.

Poon PK., Parker JW., O'Brien RL. Faulty DNA repair following ultraviolet irradiation in Fanconi's anemia. Basic Life Sci. 1975; 5B: 821-824.

Popanda O., Schattenberg T., Phong CT., Butkiewicz D., Risch A., Edler L., Kayser K., Dienemann H., Schulz V., Drings P., Bartsch H., Schmezer P. Specific combinations of DNA repair gene variants and increased risk for non-small cell lung cancer. Carcinogenesis. 2004; 25: 2433–2441.

Poplawski, T., Zadrozny M., Kolacinska A., Rykala J., Morawiec Z., Blasiak j. Polymorphisms of the DNA mismatch repair gene hMSH2 in breast cancer occurrence and progression. Breast Cancer Res. Treat. 2005; 94: 199-204.

Prolla TA., Pang Q., Alani E., Kolodner RD., Liskay RM. MLH1, PMS1, and MSH2 interactions during the initiation of DNA mismatch repair in yeast. Science. 1994; 265: 1091-1093.

Qiao Y., Spitz MR., Shen H., Guo Z., Shete S., Hedayati M., Grossman L., Mohrenweiser H., Wei Q. Modulation of repair of ultraviolet damage in the host-cell reactivation assay by polymorphic XPC and XPD/ERCC2 genotypes. Carcinogenesis. 2002a; 23: 295–299.

Qiao Y., Spitz MR., Guo Z., Hadeyati M., Grossman L., Kraemer KH., Wei Q. 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. 2002b; 509: 165-174.

Ramos JM., Ruiz A., Colen R., Lopez D., Grossman L., Matta JL. DNA repair breast carcinoma susceptibility in women. Cancer. 2004; 100: 1352–1357.

Reardon JT., Mu D., Sancar A. Overproduction, purification, and characterization of the XPC subunit of the human DNA repair excision nuclease. J. Biol. Chem. 1996; 271: 19451–19456.

Rebora A., Crovato F. PIBI(D)S syndrome--trichothiodystrophy with xeroderma pigmentosum (group D) mutation. J. Am. Acad. Dermatol. 1987; 16: 940-947.

Richardson C., Moynahan ME., Jasin M. Double-strand break repair by interchromosomal recombination: suppression of chromosome translocations. Genes Dev. 1998; 12: 3831-3842.

Riedl T., Hanaoka F., Egly JM. The comings and goings of nucleotide excision repair factors on damaged DNA. EMBO J. 2003; 22: 5293-303.

Robbins JH., Brumback RA., Mendiones M., Barrett SF., Carl JR., Cho S., Denckla MB., Ganges MB., Berger LH., Guthrie RA., Meer J., Moshell AN., Polinsky RJ., Ravin PD., Sonies BC., Tarone RE. Neurological disease in xeroderma pigmentosum. Documentation of a late onset type of the juvenile onset form. Brain. 1991; 114: 1335–1361.

Roche J., Boldog F., Robuinson M., Robinson L., Varella-Garcis M., Swanton M., Waggoner B., Fishel R., Frannklin W., Gemmill R., Drabkin H. Distinct 3p21.3 deletions in lung cancer and identification of a new human semaphorin. Oncogene. 1996; 12: 1289-1297.

Roddam PL., Rollinson S., O'Driscoll M., Jeggo PA., Jack A., Morgan GJ. Genetic variants of nhej DNA ligase IV can affect the risk of developing multiple myeloma, a tumour characterised by aberrant class switch recombination. J. Med. Genet. 2002; 39: 900-905.

Rodriguez MJ., Acha A., Ruesga MT., Rodriguez C., Rivera JM., Aguirre JM. Loss of expression of DNA repair enzyme mgmt in oral leukoplakia and early oral squamous cell carcinoma. A prognostic tool? Cancer Letters. 2006; [Epub ahead of print].

Ronen A., Glickman BW. Human DNA repair genes. Environ. Mol. Mutagen. 2001; 37: 241-283.

Rossignol M., Kolb-Cheynel I., Egly JM. Substrate specificity of the cdk-activating kniase (CAK) is altered upon association with TFIIH. EMBO J. 1997; 16: 1628-1637.

Rossit AR., Cabral IR., Hackel C., da Silva R., Froes ND., Abdel-Rahman SZ. Polymorphisms in the DNA repair gene XRCC1 and susceptibility to alcoholic liver cirrhosis in older Southeastern Brazilians. Cancer Lett. 2002; 180: 173-182.

Rothkamm K., Kruger I., Thompson LH., Lobrich M. Pathways of DNA double-strand break repair during the mammalian cell cycle. Mol. Cell Biol. 2003; 23: 5706–5715.

Rupa DS., Reddy PP., Reddi OS. Frequencies of chromosomal aberrations in smokers exposed to pesticides in cotton fields. Mutat. Res. 1989; 222: 37-41.

Rupa DS., Hasegawa L., Eastmond DA. Detection of chromosomal breakage in the 1cen-1q12 region of Interphase human lymphocytes using multicolor fluorescence in situ hybridization with tandem DNA probes. Cancer Res. 1995; 55: 640-645.

Rusquet RM., Feon SA., David JC. Association of a possible DNA ligase deficiency with T-cell acute leukemia. Cancer Res. 1988; 48: 4038– 4044.

Rybicki AB., Conti DV., Moreira A., Cicek M., Casey G., Witte JS. DNA repair gene XRCC1 and XPD polymorphisms and risk of prostate cancer. Cancer Epidemiol. Biomarkers Prev. 2004; 13: 23–29.

Sakiyama T., Kohno T., Mimaki S., Ohta T., Yanagitani N., Sobue T., Kunitoh H., Saito R., Shimizu K., Hirama C., Kimura J., Maeno G., Hirose H., Eguchi T., Saito D., Ohki M., Yokota J. Association of amino acid substitution polymorphisms in DNA repair genes tp53, poli, rev1 and lig4 with lung cancer risk. Int. J. Cancer. 2005; 114: 730-737.

Sancar A., Tang MS. Nucleotide excision repair. Photochem Photobiol. 1993; 57: 905-921.

Sancar A. DNA repair in humans. Annu. Rev. Genet. 1995; 29: 69 –105.

Sandrock B., Egly J.M. A yeast four-hybrid system identifies Cdk-activating kniase as a regulator of the XPD helicase, a subunit of transcription factor IIH. J. Bio. Chem. 2001; 276: 35328-35333.

Santagati F., Botta E., Stefanini M., Pedrini AM. Different dynamics in nuclear entry of subunits of the repair/transcription factor TFIIH. Nucl. Acids Res. 2001; 29: 1574-1581.

Schaeffer L., Roy R., Humbert S., Moncollin V., Vermeulen W., Hoeijmakers JHJ., Chambon P., Egly JM. DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. Science. 1993; 260: 58-63.

Scherer SJ., Avdievich E., Edelmann W. Functional consequences of DNA mismatch repair missense mutations in murine models and their impact on cancer predisposition. Biochem. Soc. Trans. 2005; 33: 689–693.

Schultz P., Fribourg S., Poterszman A., Mallouh V., Moras D., Egly JM. Molecular structure of human tfiih. Cell. 2000; 102: 599-607.

Schwaiger H., Hirsch-Kauffmann M., Schweiger M. UV-repair is impaired in fibroblasts from patients with Fanconi's anemia. Mol. Gen. Genet. 1982; 185: 454–456.

Seker H, Butkiewicz D, Bowman ED, Rusin M, Hedayati M, Grossman L, Harris CC. Functional significance of XPD polymorphic variants: attenuated apoptosis in human lymphoblastoid cells with the XPD 312 Asp/Asp genotype. Cancer Res. 2001 Oct 15; 61(20): 7430-4.

Semwogerere D., Weeks ER. Confocal Microscopy. Encyclopedia of Biomaterials and Biomedical Engineering. Taylor and Francis. 2005

Shen JC., Rideout WM., Jones PA.. The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. Nucl. Acids Res. 1994; 22: 972–976.

Shen H., Spitz M.R., Qiao Y., Guo Z., Wang L., Bosken C., Amos C.I. Smoking, DNA repair capacity and risk of non small cell lung cancer. Int. J. Cancer. 2003; 107: 84-88.

Shi Q., Wang LE., Bondy ML., Brewster A, Singletary SE., Wei Q. Reduced DNA repair of benzo(a)pyrene diol epoxideinduced adducts and common XPD polymorphisms in breast cancer patients. Carcinogenesis. 2004; 25: 1695–1700.

Shin DS., Chahwan C., Huffman JL., Tainer JA. Structure and function of the double-strand break repair machinery. DNA Repair. 2004; 3: 863-873.

Sijber AM., Laat WL., Ariza RR., Bifferstaff M., Wei YF., Moggs JG., Carter KC., Shell BK., Evans E., de Jong MC., Rademakers S., de Rooji J., Jaspers NG., Hoeijmakers JH., Wood Rd. Xeroderma pigmentosum group F caused by defect in a structure-specific DNA repair endonuclease. Cell. 1996; 86: 811-822.

Silver PA. How proteins enter the nucleus. Cell. 1991; 64: 489-497.

Skjelbred CF., Saebo M., Wallin H., Nexo BA., Hagen PC., Lothe IMB., Aase S., Johnson E., Hansteen IL., Vogel U., Kure E. 2006. Polymorphisms of the XRCC1, XRCC3 and XPD genes and risk of colorectal adenoma and carcinoma, in a Norwegian cohort: a case control study. BMC Cancer. 2006; 6: 67-76.

Sliwinski T., Krupa R., Majsterek I., Rykala J., Kolacinska A., Morawiec Z., Drzewoski J., Zadrozny M., Blasiak J. Polymorphisms of the *brca2* and *rad51* genes in breast cancer. Breast Cancer Res. Treat.. 2005; 94: 105-109.

Sossou M., Flohr-Beckhaus C., Schulz I., Daboussi F., Epe B., Radicella JP. APE overexpression in XRCC1-deficient cells complements the defective repair of oxidative single strand breaks but increases genomic instability. Nucl. Acids Res. 2005; 33: 298–306.

Sozzi G., Veronese M.L.., Negrini M., Baffa R., Cotticelli M.G., Inoue H., Tornielli S., Pilotti S., Gregorio L.D., Pastorino U., Pierotti M.A., Ohta M., huebner K., Croce C.M. The FHIT gene at 3p14.2 is abnormal in lung cancer. Cell. 1996; 85: 17-26.

Spitz MR., Wu X., Wang Y., Wang LE., Shete S., Amos CI., Guo Z., Lei L., Mohrenweiser H., Wei Q. Modulation of nucleotide excision repair capacity by XPD polymorphisms in lung cancer patients. Cancer Res. 2001; 61: 1354-1357.

Spitz MR., Wei Q, Dong Q, Amos CI., Wu X. Genetic susceptibility to lung cancer: the role of DNA damage and repair. Cancer Epidemiol. Biomarkers Prev. 2003; 12: 689–698.

Srivenugopal KS., Yuan XH., Friedman HS., Ali-Osman F. Ubiquitination-dependent proteolysis of O6-methylguanine-DNA methyltransferase in human and murine tumor cells following inactivation with O6-benzylguanine or 1,3-bis(2-chloroethyl)-1-nitrosourea. Biochemistry. 1996; 35: 1328 –1334.

Srivastava D., Berg B., Prasad R., Molina J., Beard W., Tomkinson A., Wilson S. Mammalian abasic site base excision repair. Identification of the reaction sequence and rate-determining steps. J. Biol. Chem. 1998; 273: 21203–21209.

Stefanini M., Lagomarisini P., Gilliani S., Nardo T., Botta E., Peserico A., Kleyer WJ., Lehmann AR., Sarasin A. Genetic heterogeneity of the excision repair defect associated with trichothiodystrophy. Carcinogenesis. 1993a; 14: 1101–1105.

Stefanini M., Vermeulen W., Weeda G., Giliani S., Nardo T., Mezzina M., Sarasin A., Harper JL., Arlett CF., Hoeijmakers JHJ., Lehmann AR. A new nucleotide-excision-repair gene associated with the disorder trichothiodystrophy. Am. J. Hum. Genet. 1993b; 53: 817–821.

Stojic L., Brun R., Jiricny J. Mismatch repair and DNA damage signalling. DNA Repair. 2004; 3: 1091-1101.

Sturgis EM., Dahlstrom KR., Spitz MR., Wei Q. DNA repair gene ERCC1 and ERCC2/XPD polymorphisms and risk of squamous cell carcinoma of the head and neck. Arch Otolaryngol Head Neck Surg. 2002; 128: 1084-8.

Sturgis EM., Zheng R., Li L., Castillo EJ., Eicher SA., Chen Chen M., Strom SS., Spitz MR., Wei Q. XPD/ERCC2 polymorphisms and risk for the head and neck cancer: a case-control analysis. Carcinogenesis. 2003; 21: 2219–2222.

Sugawara AT., Grzegorz I., Haber JE. DNA length dependence of the single-strand annealing pathways and the role of Saccharomyces cerevisiae RAD59 in double-strand break repair. Mol. Cell Biol. 2000; 20: 833-835.

Sugasawa K., Ng JM., Masutani C., Iwai S., van der Spek PJ., Eker AP., Hanaoka F., Bootsma D., Hoeijmakers JH. Xeroderma igmentosum group C protein complex is the initiator of global genome nucleotide excision repair. Mol. Cell. 1998; 2: 223–232.

Sugasawa K., Shimizu Y., Iwai S., Hanaoka F. A molecular mechanism for DNA damage recognition by the xeroderma pigmentosum group C protein complex. DNA Repair (Amst). 2002; 1: 95–107.

Sung P., Bailly V., Weber C., Thompson LH., Prakash L., Prakash S. Human xeroderma pigmentosum group d gene encodes a DNA helicase. Nature. 1993; 365: 852-855.

Surtees JA., Argueso JL., Alani E. Mismatch repair proteins: Key regulators of genetic recombination. Cytogenetic & Genome Research. 2004; 107: 146-159.

Tajiri T., Maki H., Sekiguchi,M. Functional cooperation of MutT, MutM and MutY proteins in preventing mutations caused by spontaneous oxidation of guanine nucleotide in *Escherichia coli*. Mutat. Res. 1995; 336: 257–267.

Takata M., Sasaki MS., Sonoda E., Morrison C., Hashimoto M., Utsumi H, Yamaguchi-Iwai Y., Shinohara A., Takeda S. Homologous recombination and non- homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. EMBO J. 1998; 17: 5497–5508.

Tanaka K., Kawai Y., Kumahara Y., Ikenaga M., Okada Y. Genetic complementation groups in Cockayne syndrome. Som. Cell Genet. 1981; 7: 445–455.

Thoma BS., Vasquez KM. Critical DNA damage recognition functions of XPC-hHR23B and XPS-RPA in nucleotide excision repair. Molecular Carcinogenesis. 2003; 38: 1-13.

Tomkinson AE., Totty NF., Ginsburg M., Lindahl T. Location of theactive site for enzyme-adenylate formation in DNA ligases. Proc. Natl. Acad. Sci. USA. 1991; 88: 400–404.

Tremeau-Bravard A., Perez C., Egly JM. A role of the C-terminal part of p44 in the promoter escape activity of transcription factor IIH. J. Biol. Chem. 2001; 276: 27693-12697.

Tricarico C., Pinzani P., Bianchi S., Paglierani M., Distante V., Pazzagli M., Bustin SA., Orlando C. Quantitative real-time reverse transcription polymerase chain reaction: Normalization to rRNA or single housekeeping genes is inappropriate for human tissue biopsies. Analytical Biochemistry. 2002; 309: 293-300.

Troelstra C., van Gool A., de Wit J., Vermeulen W., Bootsma D., Hoeijmakers JHJ. ERCC6, a member of a subfamily of putative helicases, is involved in Cockayne syndrome and preferential repair of active genes. Cell. 1992; 71: 939–953.

Tuteja N., Tuteja R. Unraveling DNA repair in human: molecular mechanisms and consequences of repair defect. Crit. Rev. Biochem Mol. Biol. 2001; 36: 261-290.

van Leeuwen DM., Gottschalk RWH., van Herwijnen MH.., Moonen EJ., Kleinjans JCS., van Delft JHM. Differential gene expression in human peripheral blood mononuclear cells induced by cigarette smoke and its constituents. Toxicol Sci. 2005; 86: 200-210.

van Hoffen A., Natarajan AT., Mayne LV., van Zeeland AA., Mullenders LHF., Venema J. Defecient repair of the transcribed strand of active genes in Cockayne syndrome. Nucl. Acids Res. 1993; 21: 5890–5895.

Vasquez KM., Christensen J., Li L., Finch RA., Glazer PM. Human XPA and RPA DNA repair proteins participate in specific recognition of triplex-induced helical distortions. Proc. Natl. Acad. Sci. USA. 2002; 99: 5848–5853.

Venema J., Mullenders LHF., Natarajan AT., van Zeeland AA., Mayne LV. The genetic defect in Cockayne syndrome is associated with a defect in repair of UV-induced DNA damage in transcriptionally active DNA. Proc. Natl. Acad. Sci. USA. 1991; 87: 4707–4711.

Vermeulen W., Stefanini M., Giliani S., Hoeijmakers JHJ., Bootsma D.Xeroderma pigmentosum complementation group H falls into complementationgroup D. Mutat. Res. 1991; 255: 201–208.

Vermeulen W., van Vuuren AJ., Chipoulet M., Schaeffer L., Appeldoorn E., Weeda G., Jaspers NG., Priestley A., Arlett CF., Lehmann AR., Stefanini M., Mezzina M., Sarasin A., Bootsma D., Egly J.–M., Hoeijmakers JHJ. Three unusual repair deficiencies associated with transcription factor BTF2(TFIIH): evidence for the existence of a transcription syndrome. Cold Spring Harb. Symp. Quant. Biol. 1994; 59: 317–329.

Vogel U., Dybdahl M., Frentz G., Nexo BA. DNA repair capacity: Inconsistency between effect of over-expression of five ner genes and the correlation to mrna levels in primary lymphocytes. Mutation Research/DNA Repair. 2000; 461: 197-210.

Vogel U., Hedayati M., Dybdahl M., Grossman L., Nexo BA. Polymorphisms of the DNA repair gene xpd: Correlations with risk of basal cell carcinoma revisited. Carcinogenesis. 2001; 22: 899-904.

Voss TC., Demarco IA., Day RN. Quantitative imaging of protein interactions in the cell nucleus. Biotechniques. 2005; 38: 413-424.

Wang XW., Yeh H., Schaeffer L., Roy R., Moncollin V., Egly JM., Wang Z., Friedberg EC., Evans MK., Taffe BG., Bohr JH., Hoeijmakers JH., Forrester K., Harris CC. p53 modulation of TFIIH associated nucleotide excision repair activity. Nature Genet. 1995; 10: 188-195.

Wang XW., Vermeulen W., Coursen JD., Gibson M., Lupold SE., Forrester K., Xu G., Elmore L., Yeh H., Hoeijmakers JHJ., Harris CC. The XPB and XPD DNA helicases sre components of the p53 apoptosis pathway. Genes Dev.. 1996; 10: 1219-1232.

Wang L., Spratt TE., Liu XK., Hecht SS., Pegg AE., Peterson LA. Pyridyloxobutyl adduct O6-[4-oxo-4-(3-pyridyl)butyl]guanine is present in 4(acetoxymethylnitrosamino)-1-(3-pyridyl)-1-butanone-treated DNA and is a substrate for O6-alkylguanine-DNA alkyltransferase. Chem. Res. Toxicol. 1997; 10: 562-567.

Wang LE., Sturgis EM., Eicher SA., Spitz MR., Hong WK., Wei Q. Mutagen sensitivity to benzo(a)pyrene diol epoxide and the risk od sqamous cell carcinoma of the head and neck. Clin. Cancer Res. 1998; 4: 1773-1778.

Wang TF., Kleckner N.. Hunter, N. Functional specificity of MutL homologs in yeast: evidence for three Mlh1-based heterocomplexes with distinct roles during meiosis in recombination and mismatch correction. Proc. Natl. Acad. Sci. USA. 1999; 96: 13914–13919

Wang Y., Spitz MR., Zhu Y., Dong Q., Shete S. Wu X. From genotype to phenotype: correlating *XRCC1* polymorphisms with mutagen sensitivity. DNA Repair (Amst) 2003; 2: 901–908.

Wang D., Johnson AD., Papp AC., Kroetz DL., Sadee W. Multidrug resistance polypeptide 1 (MDR1, ABCB1) variant 3435C>T affects mRNA stability. Pharmacogenet Genomics. 2005; 15: 693-704.

Wang D., Sadee W. Searching for polymorphisms that affect gene expression and mRNA processing: example ABCB1 (MDR1). AAPS J. 2006; 8: E515-E520.

Wei Q., Matanoski GM., Farmer ER., Hedayati MA., Grossman L. DNA repair and aging in basal cell carcinoma: a molecular epidemiology study. Proc. Natl. Acad. Sci. USA. 1993; 90: 1614-1618.

Wei Q., Gu J., Cheng L., Bondy ML., Jiang H., Hong WK., Spitz MR. Benzo(a)pyrene diol epoxide-induced chromosomal aberrations and risk of lung cancer. Cancer Res. 1996; 56: 3975-3979.

Wei Q. Effect of aging on DNA repair and skin carcinogenesis: a minireview of population-based studies. J. Investig. Dermatol. Symp. Proc. 1998; 3:19-22.

Wei Q., Eicher SA., Guan Y., Cheng L., Xu J., Young LN., Saunders KC., Jiang H., Hong WK., Spitz MR., Strom SS. Reduced expression of hMLH1 and hGTBP/hMSH6: a risk factor for head and neck cancer. Cancer Epidemiol. Biomarkers Prev. 1998; 7: 309-314.

Wei Q., Cheng L., Amos CI., Wang LE., Guo Z., Hong WK., Spitz MR. Repair of tobacco carcinogen-induced DNA adducts and lung cancer risk: a molecular epidemiologic study. J. Natl. Cancer Inst. 2000; 92: 1764–1772.

Wei Q., Lee JE., Gershenwald JE., Ross MI., Mansfield PF., Strom SS., Wang LE., Guo Z., Qiao Y., Amos CI., Spitz MR., Duvic M. Repair of UV light-induced DNA damage and risk of cutaneous malignant melanoma. J. Natl. Cancer. Inst. 2003; 95: 308–315.

Whitehouse C., Taylor,R., Thistlethwaite A., Zhang H., Karimi-Busheri F., Lasko D., Weinfeld M., Caldecott K. XRCC1 stimulates human polynucleotide kinase activity at damaged DNA termini and accelerates DNA single-strand break repair. Cell. 2001; 104: 107–117.

Wiebauer K., Jiricny J. *In vitro* correction of G.T mispairs to G.C pairs in nuclear extracts from human cells. Nature. 1989; 339: 234 –236.

Wiederhold L., Leppard JB., Kedar P., Karimi-Busheri F., Rasouli-Nia A., Weinfeld M., Tomkinson AE., Izumi T., Prasad R., Wilson SH. AP endonuclease-independent DNA base excision repair in human cells. Mol. Cell. 2004; 15: 209-220.

Wikman H., Risch A., Klimek F., Schmezer P., Spiegelhalder B., Dienemann H., Kayser K., Schulz V., Drings P., Bartsch H. hOGG1 polymorphism and loss of heterozygosity (LOH): significance for lung cancer susceptibility in a caucasian population. Int J Cancer. 2000; 88(6): 932-7.

Wogan GN., Hecht SS., Felton JS., Conney AH., Loeb LA. Environmental and chemical carcinogenesis. Semin Cancer Biol. 2004; 14: 473-86.

Wolfe KJ., Wickliffe JK., Hill CE., Ammenheuser MM., Abdel-Rahman SZ. Single nucleotide polymorphisms of the DNA repair gene *XPD/ERCC2* alter mRNA expression. (Pharmacogenetics and Genomics (in press)

Wu X., Zhao Y., Honn S.E., Tomlinson G.E., Minna J.D., Ki Hong W., Spitz M.R. Benzo[a]pyrene diol epoxide-induced 3p21.3 aberrations and genetic predisposition to lung cancer. Cancer Res. 1998; 58: 1605-1608.

Wu X., Spitz M.R., de Andrande M., Benowitz N.L., Swan G.E. Genetic influence on mutagen sensitivity: a twin study. Proc Amer Assoc Cancer Res. 2000; 41: 437.

Wu X., Zhao H., Wei Q., Amos CI., Zhang K., Guo Z., Qiao Y., Hong WK., Spitz MR. XPA polymorphism associated with reduced lung cancer risk and a modulating effect on nucleotide excision repair capacity. Carcinogenesis. 2003; 24: 505–509.

Wu X., Zheng TL., Hsu TC. Mutagen-induced chromatid breakage as a marker of cancer risk. Methods Molecular Biology. 2004; 291: 59-68.

Wu YP., Yang YL., Yang GZ., Wang XY., Luo ML., Zhang Y., Feng YB., Xu X., Han YL., Cai Y., Zhan QM., Wu M., Dong JT., Wang MR. Identification of chromosome aberrations in esophageal cancer cell line KYSE180 by multicolor fluorescence in situ hybridization. Cancer Genet Cytogenet. 2006; 170: 102-107.

Wyman C., Ristic D., Kanaar R. Homologous recombination-mediated double-strand break repair. DNA Repair. 2004; 3: 827-833.

Xi T., Jones IM., Mohrenweiser HW. Many amino acid substitution variants identified in DNA repair genes during human population screenings are predicted to impact protein function. Genomics. 2004; 83: 970-979.

Xu Z., Chen ZP., Malapetsa A., Alaoui-Jamali M., Bergeron J., Monks A., Myers TG., Mohr G., Sausville EA., Scudiero DA., Aloyz R., Panasci LC. DNA repair protein levels vis-a-vis anticancer drug resistance in the human tumor cell lines of the National Cancer Institute drug screening program. Anticancer Drugs. 2002; 13: 511-519.

Xu Y., DeMott M., Hwang J., Greenberg M., Demple B. Action of human apurinic endonuclease (Ape1) on C1-oxidized deoxyribose damage in DNA. DNA Repair. 2003; 2: 175– 185.

Yeh YC., Chang DY., Masin J., Lu AL. Two nicking enzyme systems specific for mismatch-containing DNA in nuclear extracts from human cells. J. Biol. Chem. 1991; 266: 6480–6484.

Yin J., Lin J., MA Y., Guo L., Wang H., Vogel U. The DNA repair gene *ERCC2/XPD* polymorphism Arg 156Arg (A22541C) and risk of lung cancer in a Chinese population. Cancer Lett. 2005; 223: 219-226.

Yokoi M., Masutani C., Maekawa T., Sugasawa K., Ohkuma Y., Hanaoka F. The xeroderma pigmentosum group C protein complex XPC-HR23B plays an important role in the recruitment of transcription factor IIH to damaged DNA. J. Biol. Chem. 2000; 275: 9870–9875.

Yonish-Rouach E., Resnitzky D., Lotem J., Sachs L., Kimchi A., Oren M. Wild type p53 induces apoptosis of myeloid leukaemic cells that is inhibited by interleukin 6. Nature. 1991; 352: 345-347.

Yuan ZQ., Gottlieb B., Beitel LK., Wong N., Gordon PH., Wang Q., Puisieux A., Foulkes WD., Trifiro M. Polymorphisms and hnpcc: Pms2-mlh1 protein interactions diminished by single nucleotide polymorphisms. Human Mutation. 2002; 19: 108-113.

Yu Z., Chen J., Ford BN., Brackley ME., Glickman BW. Human DNA repair systems: An overview. Environ Mol Mutagen. 1999; 33: 3-20.

Zaher C., Halbert R., Dubois R., George D., Nonikov D. Smoking-related diseases: the importance of COPD. Int. J. Tuberc. Lung Dis. 2004; 8: 1423-1428.

Zanesi N., Fidanza V., Fong LY., Mancini R., Druck T., Valtieri M., Rudiger T., McCue PA., Croce CM., Huebner K. The tumor spectrum in FHIT-deficient mice. Proc Natl Acad Sci U.S.A. 2001; 98: 10250-10255.

Zhang Y., Wang D., Johnson AD., Papp AC., Sadee W. Allelic expression imbalance of human mu opioid receptor (OPRM1) caused by variant a118g. J. Biol. Chem. 2005; 280: 32618-32624.

Zhou W., Liu G., Miller DP., Thurston SW., Xu LL., Wain JC., Lynch TJ., Su L., Christiani DC. Gene–environment interaction for the ERCC2 polymorphisms and cumulative cigarette smoking exposure in lung cancer. Cancer Res. 2002; 62: 1377–1381.

Zienolddiny S., Campa D., Lind H., Ryberg D., Skaug V., Stangeland L., Phillips DH., Canzian F., Haugen A. Polymorphisms of DNA repair genes and risk of non-small cell lung cancer. Carcinogenesis. 2006; 27: 560-567.

Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucl. Acid Res. 2003; 31: 3406-3415.

Zurita M., Merino C. The transcriptional complexity of the TFIIH complex. Trends Genet. 2003; 19: 578-584.

VITA

Kevin Wolfe was born on December 17th, 1978 to James and Roswitha Wolfe. After attending High Point University for his Bachelor's degree, Kevin matriculated at the University of Texas Medical Branch. While at graduate school, Kevin received several honors. From 2002 to 2005, Kevin was awarded the NIEHS Toxicology Training Grant scholarship, and in 2005 he received the Society of Toxicology Travel award to the 44th annual Society of Toxicology meeting. Kevin gained significant teaching experience while at the University of Texas Medical Branch. Twice he participated as an instructor at Texas A&M University at Galveston, in the class entitled MARB 489: Special Topics in Toxicology

Kevin can be contacted through his parents at 17321 Chiswell Rd., Poolesville, MD 20837.

Education

B.S., May 2001, High Point University, High Point, NC

Publications

Affatato AA, **Wolfe KJ**, Lopez MS, Hallberg C, Ammenheuser MM and Abdel-Rahman, SZ. Effect of XPD/ERCC2 polymorphisms on chromosome aberration frequencies in smokers and on sensitivity to the mutagenic tobacco-specific nitrosamine NNK. Environ. Mol. Mutagen 2004; 44: 65-73. 2004.

Hill CE, Affatato AA, **Wolfe KJ**, Lopez MS, and Abdel-Rahman SZ. Gender Differences in Mutagenic Sensitivity to the Tobacco-Specific Nitrosamine NNK and the Influence of the Thr241Met Polymorphism in the XRCC3 Gene. Environ Mol Mutagen. 2005 Jul; 46(1): 22-9.

Hill CE, Wickliffe JK, **Wolfe KJ**, Kinslow CJ, Lopez MS, Abdel-Rahman SZ. The L84F and the I143V polymorphisms in the O6-methylguanine-DNA-methyltransferase

(MGMT) gene increase human sensitivity to the genotoxic effects of the tobacco-specific nitrosamine carcinogen NNK. Pharmacogenet Genomics. 2005; 15(8): 571-8.

Wolfe KJ, Wickliffe JK, Hill CE, Ammenheuser MM and Abdel-Rahman SZ. Single nucleotide polymorphisms of the DNA repair gene *XPD/ERCC2* alter mRNA expression. (Pharmacogenetics and Genomics (in press)