Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature: _____ _____

Ludmila Vladimirova Petrova Date

Context-Dependent Roles of Transcriptional Mutagenesis in Oncogene Activation and Its Phenotypic Consequences

By

Ludmila Vladmirova Petrova, B.S.

Graduate Division of Biological and Biomedical Science
Genetics and Molecular Biology

_____
Paul W. Doetsch
Advisor

_____
Guy M. Benian
Committee Member

_____
Gray F. Crouse
Committee Member

_____
Carlos S. Moreno
Committee Member

_____
Wei Zhou
Committee Member

Accepted:

_____
Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

_____ Date

Context-Dependent Roles of Transcriptional Mutagenesis in Oncogene Activation and Its Phenotypic Consequences

By

Ludmila Vladmirova Petrova, B.S., Georgia State University, 2008

Advisor: Paul W. Doetsch, Ph.D.

An abstract of a dissertation submitted to the Faculty of the James T. Laney School of Graduate Studies of Emory University in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Graduate Division of Biological and Biomedical Science, Genetics and Molecular Biology, 2016.

Abstract

Context-Dependent Roles of Transcriptional Mutagenesis in Oncogene Activation and Its Phenotypic Consequences

By

Ludmila Vladimirova Petrova

Regardless of its proliferation state, each mammalian cell acquires thousands of chemically diverse DNA damage lesions per day due to exposure to a variety of endogenous and exogenous DNA damaging agents. Each lesion could differentially compromise the fidelity of genetic information transfer at the level of both DNA replication as well as transcription, and lead to deleterious biological endpoints. While the mechanisms and consequences of replicative mutagenesis have been more thoroughly investigated, our knowledge of the *in vivo* molecular and phenotypic consequences of the encounters of RNA polymerase with DNA damage remains very limited. Whether mutagenic RNA polymerase-catalyzed bypass of DNA damage, or transcriptional mutagenesis (TM), occurs and results in phenotypic change *in vivo* is likely dependent on the identity of the lesion, its position and the context in which it occurs, and its timely repair.

The genetic fate and biological consequences of DNA damage are lesion- and context-dependent. However, the study of defined DNA damage *in vivo* requires technically challenging approaches for the targeted introduction of DNA damage lesions into relevant DNA sequences of interest. I have developed a highly efficient and reliable methodology for the production of mammalian expression vectors containing site-specific base modifications in any position of interest and DNA strand of choice. Employing this method, I show that the cytosine-derived oxidative lesions 5-hydroxyuracil (5-OHU) and dihydrouracil (DHU) are transcriptionally mutagenic *in vivo*, and when placed in the G12D mutational hotspot of the proto-oncogene K-Ras, they can result in sustained activation of more than one Ras effector pathways, including ERK and AKT. Results employing mouse cells deficient in Neil1, Neil2, or both, suggest that Neil2 is the primary glycosylase repairing 5-OHU and DHU *in vivo* and that the transcription status of DNA containing lesions may be an important factor influencing DNA repair in vivo. Further studies employing the tools and systems developed in this dissertation will help address whether base excision repair DNA glycosylases may not be entirely redundant, but may be influenced by the transcription or replication status of the affected DNA, potentially influencing the occurrence and biological outcomes of transcriptional or replicative mutagenesis, respectively.

Context-Dependent Roles of Transcriptional Mutagenesis in Oncogene Activation and Its Phenotypic Consequences

By

Ludmila Vladimirova Petrova, B.S., Georgia State University, 2008

Advisor: Paul W. Doetsch, Ph.D.

A dissertation submitted to the Faculty of the James T. Laney School of Graduate Studies of Emory University in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Graduate Division of Biological and Biomedical Science, Genetics and Molecular Biology, 2016.

Table of Contents

List of Figures

List of Tables

1. INTRODUCTION: REPAIR, MOLECULAR AND PHENOTYPIC CONSEQUENCES OF BASE DAMAGE IN MAMMALIAN CELLS

DNA Damage in the Context of Human Disease

Each cell, regardless of its proliferation state, acquires thousands of chemically diverse DNA damage lesions per day (1), each of which could differentially compromise the fidelity of genetic information transfer at the level of both DNA replication as well as transcription and lead to deleterious phenotypic consequences (2,3). Exposure to a variety of endogenous (e.g. reactive oxygen species) and exogenous (e.g. ultraviolet light) DNA damaging sources can result in the generation of DNA lesions that can lead to mutagenesis when they are bypassed by DNA or RNA polymerases, or programmed cell death when they block DNA replication or transcription. In proliferating cells, translesion synthesis by DNA polymerases during DNA replication can result in heritable mutations that can lead to diseases, such as cancer. Genomic instability is arguably one of the most fundamental hallmarks of cancer as it may allow cells to acquire any of the other hallmarks and characteristics of cancer, such as sustained proliferation, by the acquisition of activating mutations in oncogenes, or evasion of growth suppressive signaling, by the acquisition of loss-of-function mutations in tumor suppressors (4). While not all cells in an organism undergo DNA replication frequently, and a variety of cells that can serve as targets of tumorigenesis or other pathologies exist in a largely quiescent state, all cells continuously undergo transcription. RNA polymerase (RNAP) can also misincorporate bases opposite DNA damage lesions during transcription, equivalent to those incorporated by DNA polymerase during replication, and result in transcriptional mutagenesis (TM), as illustrated in Figure 1.2 (2,3,5). Since TM is a mutagenesis mechanism that is not limited only to proliferating cells, but could affect every single cell in the human body, it potentially has broad implications to human health. However, the

overwhelming majority of DNA damage studies have focused on DNA damage from the perspective of replicative mutagenesis. Few DNA damage lesions have been studied *in vivo* in the context of TM, and the *in vivo* biological consequences and significance of TM in the initiation or progression of disease has not yet been established. Technical challenges and development of systems, as will be further described in the following sections, pose one of the main obstacles for the study of the consequences of defined DNA damage and transcriptional mutagenesis *in vivo*.

Base excision repair (BER) and nucleotide excision repair (NER) are the two main pathways that repair base damage and they are highly evolutionarily conserved from prokaryotes to humans. Defects in repair promote the persistence of DNA damage and replicative or transcriptional mutagenesis. Inherited mutations in NER genes are associated with predisposition to cancer, developmental and neurological disorders, as well as premature aging (6,7). While defects in BER occur less frequently, they are associated with similar diseases, and mutations and genetic polymorphisms in BER components are associated with increased cancer risk (8-10).

The majority of DNA damage repair studies to date have employed exposure to DNA damaging agents that create a plethora of diverse lesions at levels significantly exceeding endogenously occurring, physiological DNA damage. Moreover, as DNA damaging agents induce lesions at random positions throughout the genome, the repair and relative contributions of each particular lesion to the development of a DNA damage-induced phenotype cannot be delineated. While we have gained a wealth of knowledge by employing such systems, delineating the consequences of each individual lesion in physiologically relevant conditions, without the confounding factor of extensive DNA

damage requires further investigation using well-defined systems. Studying the DNA repair, molecular and phenotypic consequences of defined DNA damage *in vivo*, from the perspective of not only replicative but also transcriptional mutagenesis, has broad implications for the study of human health. It will allow us to define the factors that influence DNA repair *in vivo* and understand the mechanisms by which DNA damage can initiate diseased phenotypes in both proliferating and non-proliferating cells. A broader mechanistic understanding of the impact of DNA damage on human health will allow us to design more effective strategies for the prevention and treatment of diseases associated with DNA damage and deficiencies in its repair.

Base Excision Repair of DNA Base Damage in Mammalian Cells

The genome of a single cell can acquire several thousand chemically diverse, potentially mutagenic base lesions per day due to oxidation, alkylation, deamination, depurination/depyrimidination or hydrolysis following exposure to a variety of endogenous and exogenous DNA damaging agents (1,11-13). Some forms of damage can also occur spontaneously at low levels (13). BER is the main pathway by which such non-bulky lesions are repaired. It is initiated by distinct DNA glycosylases that recognize specific damages. Monofunctional DNA glycosylases excise the damaged base by hydrolyzing the *N*-glycosidic bond between the sugar and the base, leaving an apurinic/apyrimidinic (AP) site, while bifunctional glycosylases further catalyze the incision of the resulting AP site by cleaving the DNA backbone 3' to the lesion via a β- or β,δ-elimination reaction (14). AP sites are further processed by one out of two BER subpathways: short patch repair, involving only a single nucleotide, or long patch repair

whereas 2-13 nucleotides are removed from the original damage site (9). AP endonucleases cleave the phosphodiester backbone 5' to the AP site, leaving a single strand break. The gap is then filled by DNA polymerase β (short patch repair) or DNA polymerase β, δ, or ε (long patch repair) and ligated by DNA ligase III in complex with XRCC1 (Figure 1.1).

Currently, eleven DNA glycosylases with overlapping specificities have been described in humans (15). DNA glycosylases UNG (uracil-DNA glycosylase), MUTYH (mutY DNA glycosylase), TDG (thymine-DNA glycosylase), SMUG1 (single-strand-selective monofunctional uracil-DNA glycosylase 1), MBD4 (methyl-CpG binding domain 4 DNA glycosylase), and MPG (N-methylpurine DNA glycosylase) repair deaminated or alkylated bases. OGG1 (8-oxoguanine DNA glycosylase), which removes purine-derived lesions such as 8-oxoguanine (8-oxoG), as well as NEIL1, NEIL2, NEIL3 (Nei endonuclease VIII-like 1, 2, and 3), and NTHL1 (Nth endonuclease III-like) (16), which remove pyrimidine-derived lesions such as 5-hydroxyuracil (5-OHU), are involved in the repair of oxidative base lesions (Table 1.1). Oxidative damage is one of the most frequently occurring type of damage and arises due to the generation of reactive oxygen species (ROS) as byproducts of normal cellular metabolism (17), inflammation and host defenses against pathogenic microorganisms (18), exposure to ionizing or ultraviolet radiation, and various environmental pollutants (19). While ROS serve important biological functions in signal transduction, oxidative damage is associated with diseases, such as cancer, and plays a key role in the theory of aging (14).

While the substrate specificities of DNA glycosylases have been characterized more extensively *in vitro* employing short synthetic oligonucleotides in defined enzymatic

reactions (Table 1.1), very few studies have focused on studying repair of defined DNA damages in intact, live cells. However, several major factors present in live eukaryotic cells can potentially influence DNA repair, modulating the impact of DNA damage *in vivo*. First, the gene expression of repair enzymes can vary between different individuals, tissue types (20), and can also be cell cycle dependent, as is the case for NEIL1 (21), or independent, as is the case for NEIL2 (22). DNA glycosylase expression likely requires precise regulation. Deficiencies in repair components are associated with increased mutagenesis, and ablation of some BER components (e.g. APEX1, XRCC1) can even be embryonic lethal. However, overexpression of repair enzymes, such as human MPG, APEX1, or DNA polymerase β, can also induce frameshift mutagenesis and microsatellite instability (23-25). Second, the *in vivo* activity, protein levels and sub-cellular localization of base excision repair proteins can be regulated via a variety of post-translational modifications, such as SUMOylation, ubiquitylation, phosphorylation, and acetylation (15). Emerging evidence indicates that mechanisms modulating the repair capacity in response to genotoxic insults exist in mammalian cells (15). Moreover, DNA in cells exists in the context of chromatin, which can influence the accessibility of DNA damage to repair factors, and chromatin remodeling may accompany and serve importation functions in facilitating BER (26), as well as NER (6,27,28). Whether the damage occurs in replicating single-stranded DNA, double-stranded DNA, transcription bubble, or quadruplex DNA, such as that found at telomeres, is also likely to be important as the activity of DNA glycosylases varies depending on the structure and strandedness of the affected DNA (Table 1.1). Repair of specific lesions may also be influenced by the presence of additional damage in close proximity to the lesion, such as single-strand

breaks (29). Such clustered DNA damage in the form of two or more lesions, including base lesions only or base damage in the proximity of a single-strand or double-strand break, may occur following exposure to certain types of DNA damaging agents, such as ionizing radiation (30,31).

Nucleotide Excision Repair of DNA Base Damage

Bulky, helix-distorting DNA lesions, such as those generated due to exposure to ultraviolet (UV) light or DNA cross-linking agents such as chemotherapeutic drugs, are repaired by NER. NER consists of two subpathways, global genome repair (GG-NER) and transcription-coupled repair (TC-NER), that are involved in the repair of a variety of structurally unrelated bulky lesions. The two subpathways differ in the mechanisms by which DNA damage substrates are recognized. GG-NER is involved in the repair of lesions in both transcribed and untranscribed DNA strands that distort the DNA structure, such as the UV-generated 6-4 photoproducts. The XPC protein, in association with RAD23B and CETN2, serves as the main sensor that detects the presence of helix-distorting lesions and initiates GG-NER (32). TC-NER is involved in the repair of lesions, such as cyclobutane pyrimidine dimers (CPDs), that stall RNA polymerase. RNA polymerase stalling recruits the Cockayne syndrome proteins CSA and CSB, initiating TC-NER (6). The two subpathways then converge for the steps of DNA damage verification, dual 5' and 3' incision, gap filling and ligation.

Base lesions that are not repaired by neither BER nor NER can be repaired via direct repair. Some alkylated DNA bases can be repaired by direct repair mechanisms involving O^6-methylguanine methyltransferase (MGMT), which directly removes the aberrant

methyl group from O^6-methylguanine (O^6-mG) lesions, or the AlkB family of dioxygenases via hydroxylation of the methyl group, that then spontaneously leaves as formaldehyde (33). The mismatch excision repair (MMR) pathway can also be involved in the response to mispairs caused by the presence of some lesions, including O^6-mG:T mispairs (34).

Encounters of RNA Polymerase with DNA Damage and Transcriptional Mutagenesis

The phenomenon of TM was first reported in 1993 using an *in vitro* system employing SP6 and *E. coli* RNA polymerases and transcription templates containing site-specific abasic sites or single-strand breaks (35). Since then, the RNAP bypass efficiencies and transcriptional mutagenicity for a variety of BER and NER substrate lesions have been determined *in vitro*, and for a limited number of lesions *in vivo*, using phage, yeast, bacterial, and mammalian RNA polymerases, as illustrated in Table 1.2. Structural studies are beginning to elucidate the mechanisms by which base damages can result in RNAP stalling or translesion synthesis, and lead to the generation of normal transcripts or various levels of transcriptional mutagenesis. While some common themes have emerged, the outcome of the encounter of RNAP with DNA damage appears to be lesion-specific.

For BER substrate non-bulky lesions, mutagenesis depends on the mispairing properties of each specific lesion, which can vary substantially between lesions (Table 1.2). For example, while thymine glycol can result in negligible levels of transcriptional mutagenesis (2%), 5-hydroxyuracil leads to transcriptional mutations in almost all

transcripts (98%) (36). Notably, for the majority of lesions studied to date, the base misincorporated opposite each lesion by RNAP is equivalent to that misincorporated by DNA polymerase during DNA replication. Also, similar to the "A-rule" for DNA polymerase (37), RNA polymerase also incorporates adenine opposite non-informative sites, including abasic sites and single-strand breaks (38).

While the majority of TM studies have been carried out *in vitro*, they may not reliably predict all molecular consequences of unrepaired DNA damage *in vivo*. For example, while oxidative lesions including 8-oxoguanine, thymine glycol, and 5-hydroxyuracil stall RNAP *in vitro* to various extents, transcription elongation factors present in HeLa nuclear extracts can allow RNAP to bypass them, albeit mutagenically (36,39). 8-oxoG is also mutagenically bypassed *in vivo* in mammalian cells (40,41), and *E. coli* (42). Likewise, single-strand breaks stall prokaryotic RNA polymerases *in vitro (35)*, but are mutagenically bypassed in *E. coli* (38). However, once BER substrates are bypassed by RNAP, the types of mutations occurring *in vivo* are generally consistent to those observed *in vitro*. What other factors, such as sequence context, presence of clustered DNA damage, or gene expression levels, may influence TM remains to be determined. Evidence suggests that, at least in the case of 8-oxoG, TM does not appear to depend on gene expression levels but the levels of TM can be affected by the sequence context in which the damage occurs (41).

While structurally distorting NER substrates typically serve as a block to transcription elongation, some bulky lesions do not entirely stall RNA polymerase and can be bypassed *in vivo* to various extents. Defects in NER can modulate the bypass efficiency of bulky lesions occurring in the template DNA strand, and translesion

synthesis past some NER substrates may also lead to TM in a substantial proportion of the resulting transcripts. Using a recently developed assay based on quantitative liquid chromatography-tandem mass spectrometry (43), the relative bypass efficiency and transcriptional mutagenicity of several bulky lesions, including oxidatively-induced (43), thiopurine chemotherapeutic drug-induced (44), carboxymethylated (45), and ethylated (46), have been determined. The transcriptional mutagenicity of each lesion for NER substrates can also be highly variable (Table 1.2).

In vitro studies using *Saccharomyces cerevisiae* RNA polymerase are beginning to elucidate the mechanisms of RNA polymerase stalling or bypass at helix-distorting lesions. If the lesion cannot enter the active site, RNAP non-templated adenosine monophosphate (AMP) incorporation following the A-rule can also occur for NER substrate lesions, such as CPDs and 1,2-d(GpG) intrastrand crosslinks induced by the chemotherapeutic drug cisplatin (47-49). For lesions such as CPDs, this can be non-mutagenic. With the lesion still outside the active site, yeast RNA polymerase incorporates a non-templated AMP opposite the 3' T following the A-rule, and a templated incorporation of AMP opposite the 5' T allows CPDs to be bypassed (47). Also, only very low frequencies (1.3-5.8%, depending on NER capacity) of CPD-mediated transcriptional mutations were observed in human cells (50). While uridine monophoshate (UMP) can also be incorporated at the 5' T and result in a T:U mismatch, UMP misincorporation results in irreversible RNA polymerase stalling (47,51). While AMP incorporation at sites of cisplatin lesions results in G:A mismatches, RNA polymerase stalling is independent of the G:A mismatch but occurs due to the cisplatin lesion being unable to pass a translocation barrier (49).

AMP incorporation by RNA polymerase following the A-rule can also occur 5' of the lesion. While yeast RNA polymerase incorporates UMP opposite oxidative 8,5′-cyclo-2'-deoxyadenosine (CydA) lesions *in vitro*, impaired loading of the template base 5' of the lesion and misincorporation of AMP following the A-rule occurs for the base 5' of the lesion (48). While lesion bypass is significantly impaired, addition of the transcription factor TFIIF can facilitate the bypass of CydA and result in the production of transcripts containing both cognate U or non-cognate A opposite the base 5' of the lesion (48). CydA also gives rise to 5' A mutations in human cells (43). Thus, since the mechanisms of RNA polymerase stalling or bypass can differ for each specific lesion studied so far, further studies are needed to determine the mechanisms of RNA polymerase stalling or bypass for a broader variety of lesions. Moreover, further investigation is needed to elucidate each lesion's molecular and phenotypic consequences *in vivo,* in the presence of all cellular components that may influence the outcomes of its encounter with RNA polymerase.

The Phenotypic Consequences of Transcriptional Mutagenesis

The first study to demonstrate the *in vivo* phenotypic consequences of TM employed an expression construct containing uracil placed in a stop codon of the firefly luciferase gene, such that when TM occurs the active form of luciferase would be produced instead of the inactive form in *E. coli* cells (52). In uracil DNA glycosylase (*ung*) knockout cells a significant increase in luciferase signal was present, unlike in cells not lacking *ung*, indicating that the DNA repair capacity of a cell is an important factor influencing the consequences of TM. Similar systems, also using luciferase as a reporter,

were later employed establishing that other lesions such as 8-oxoG, through the RNA polymerase misincorporation of adenine (42), and abasic sites and strand breaks (38), can also result in transcriptional mutagenesis in *E. coli*.

Systems using luciferase or fluorescent proteins as reporters of TM *in vivo* have also been adapted for the study of defined DNA damage lesions in mammalian cells. Using such systems, it was demonstrated that, similar to *E. coli*, 8-oxoG can also be transcriptionally mutagenic in human cells (41). In addition, the alkylated lesion O^6-methylguanine can result in the misincorporation of uridine in 58% of the transcripts in cells where MGMT activity was compromised using the MGMT inhibitor O^6-benzylguanine, and result in the reactivation of red fluorescent protein (53).

While these studies have demonstrated that TM can result in the production of altered proteins *in vivo*, the question of whether, and under what context, transient production of mutant proteins via a TM mechanism can induce disease or physiologically relevant phenotypic change remains unanswered. The area of TM was brought to disease relevance in a study in which 8-oxoG was placed in the Q61 mutational hotspot of the proto-oncogene H-Ras, such that when TM occurs, the oncogenic Q61K mutant would be produced. In $Ogg1^{-/-}$ mouse embryonic fibroblasts, 8-oxoG-mediated TM resulted in the production of the Q61K mutant and activation of the Ras downstream effector ERK (40). Since the transcriptional mutations are equivalent to those during replicative mutagenesis for the same lesion, it has been proposed that such transcriptional mutations, that can potentially switch a quiescent cell into a pro-growth state, can result in the permanent establishment of the original transcriptional mutation and phenotype into one of the daughter cells following DNA replication by a process termed retromutagenesis (RM)

(54). RM could result in a population of mutant cells. DNA replication may not be necessary if the transcriptional mutation induces an irreversible phenotype, however the phenotype would be confined to the cells where the transcriptional mutation occurred.

Technical Challenges for the Study of Defined DNA Damages *in vivo* and Development of Systems

While the potential role of TM in several disease-associated contexts, including oncogenic transformation (54), neurodegeneration (5), and bacterial antibiotic resistance (38), have been proposed, to date, no studies have demonstrated a role for TM in the initiation of such phenotypes and no systems have been developed to directly test these hypotheses. One of the most significant obstacles for the study of the repair and phenotypic consequences arising from defined DNA damage lesions in mammalian cells, including transcriptional mutagenesis, are the lack of efficient and reliable methods for the generation of large quantities of vectors containing DNA damage lesions in any position in any sequence of interest, including protein-coding genes. One of the main goals of this dissertation, as will be further described in Chapter 2, was to develop an efficient and reliable protocol for the high-yield production of vectors containing DNA damage lesions at defined positions of any sequence of interest and identify factors that determine protocol reliability. The main steps of the protocol, as illustrated in Figure 1.3, involve the large-scale production of single-stranded phagemid DNA followed by annealing of synthetic oligodeoxynucleotides containing the lesion of interest, *in vitro* second strand synthesis and ligation using T4 DNA polymerase and T4 DNA ligase, respectively.

I have also developed an experimental system to determine the *in vivo* transcriptional mutagenesis and phenotypic consequences of defined base lesions, such as 5-OHU and DHU, placed in defined positions of protein coding genes, identify the DNA glycosylases involved in the *in vivo* repair of oxidative lesions and monitor the long-term repair dynamics and time-course of their repair. Results employing immortalized BER-deficient MEFs indicate that, when 5-OHU and DHU are placed in the G12 mutational hotspot of the proto-oncogene K-Ras, TM can result in the production of the constitutively active oncogenic G12D mutant and can activate more than one Ras effector pathways, including Raf-MEK-ERK and PI3K-AKT-mTOR. The Raf-MEK-ERK pathway regulates cell cycle progression and cell migration, while PI3K-AKT-mTOR regulates cell survival by inhibition of several tumor suppressors (55). Thus, sustained TM-mediated oncogene activation could mediate several different oncogenic cellular processes. While Neil1, Neil2, Neil3, and UNG have incision activity towards 5-OHU *in vitro* (Table 1.1), preliminary results suggest that Neil2 is the primary glycosylase involved in the repair of 5-OHU *in vivo* in mammalian cells. Deficiency in Neil2 promotes the persistence of the lesion in double-stranded, transcribed DNA and sustained TM-mediated phenotypic change. Likewise, while Neil1, Neil2, and Nthl1 all have incision activity towards DHU *in vitro*, it appears that Neil2 is also the primary glycosylase involved in the repair of DHU *in vivo*.

Cells can be oncogenically transformed by a limited set of defined mutations. The role of constitutively activated Ras in tumor initiation is well established (56), however Ras-driven tumorigenesis often requires the presence of a cooperating mutation, such as one leading to the loss of a tumor suppressor gene (57). Expression of mutant Ras alone,

or other oncogenes such as BRAF, can lead to increases in the intracellular ROS, the accumulation of extensive DNA damage in the form of double-strand breaks and base damage, activation of the DNA damage response (DDR), and ultimately oncogene-induced senescence (OIS) (58-61). OIS is a stable growth arrest prevalent in pre-malignant tumors that serves as a barrier to oncogenic transformation, and pre-malignant cells able to evade it become oncogenically transformed (62). By employing primary MEFs, with intact DDR, and deficient in Neil2 and expression constructs containing site-specific 5-OHU or DHU lesions in codon 12 of K-Ras, the role of TM in the initiation of phenotypes beyond signaling changes in the context of OIS can be studied. This well-defined system obviates unsolved technical challenges with pre-existing systems as it 1) does not rely on establishing a population of entirely non-proliferating mammalian cells in cell culture, such that the confounding factor of replicative mutagenesis can be excluded, 2) every transfected cell in the population contains an identical lesion at identical position of the gene of interest, and 3) the phenotypic read-outs involve well-established and well-defined phenotypes and common laboratory techniques.

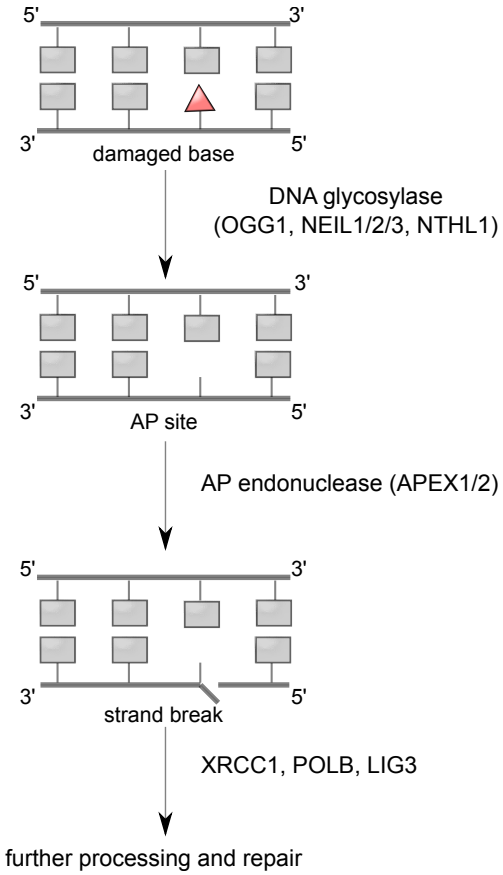


Figure 1.1: Base excision repair of oxidative DNA damage.

BER of oxidative base damage is initiated by DNA glycosylases that recognize and excise specific DNA damage lesions. The resulting AP site is then processed into a single-strand break by an AP endonuclease. XRCC1, a DNA polymerase and ATP-dependent DNA ligase III further process the single-strand break, leading to the replacement of the damaged base with a normal base. Other types of damage, such as alkylation damage, are repaired by the same mechanism, however repair is initiated by a different set of specialized glycosylases as described in the main text.

Enzyme	Lesion	Species	DNA	Experiment	Notes	Ref.
NEIL1	5-OHU	Human, Mouse	ds, ss, bubble	*In vitro*	Plasmid digest fragment; bubble, ss, and duplex oligos	(63-67)
	DHU	Human, Mouse	ds	*In vitro*	Duplex oligo	(21,65,67)
	FapyG	Human, Mouse	ds	*In vitro*	Irradiated DNA; duplex oligo	(21,65)
	FapyA	Human		*In vitro*	Irradiated DNA	(21)
	8-oxoG	Human, Mouse	ds	*In vitro*	Duplex oligo; low activity	(21,66)
	Tg	Mouse	ds, ss	*In vitro*	Duplex and ss oligo	(65,66)
	Sp	Human	ds, quadruplex	*In vitro*	Duplex, telomeric quadruplex oligos	(67-69)
	Gh	Human	ds, quadruplex	*In vitro*	Duplex; telomeric, Na^+ promoter quadruplex oligos	(67-69)
NEIL2	5-OHU	Human	ds, ss, bubble	*In vitro*	Bubble and ss oligos > duplex oligo	(22,63,64)
	DHU	Human	ds	*In vitro*	Low activity	(22)
	Gh	Human	quadruplex	*In vitro*	Na^+ promoter quadruplex oligos	(69)
NEIL3	Sp	Mouse, Human	ds, ss, quadruplex	*In vitro, cell extract*	Duplex and ss oligos; $Neil3^{-/-}$ NSPCs; telomeric quadruplex	(68-71)
	Gh	Mouse, Human	ds, ss, quadruplex	*In vitro, cell extract*	Duplex and ss oligos; $Neil3^{-/-}$ NSPCs; telomeric, Na^+ promoter quadruplex	(68-71)
	5-OHC	Mouse, Human	ds, ss	*In vitro*	ss > duplex oligo	(68,70)
	5-OHU	Mouse, Human	ds, ss	*In vitro*	ss > duplex oligo	(68,70)
NTHL1	DHU	Human	ds	*In vitro*	Duplex oligo	(21,72)
	Tg	Human, Mouse	ds	*In vitro*	Duplex oligo; 5R6S Tg stereoisomer	(65,72)
	FapyA	Mouse	ds	*Cell extract*	Liver mitochondria, liver nuclei, duplex oligo	(73)
OGG1	8-oxoG	Human, Mouse	ds	*In vitro*	Duplex DNA, no activity on ss and bubble	(64)
	FapyG	Mouse	ds	*Cell extract*	Liver mitochondria, liver nuclei, duplex oligo	(73)
UNG	5-OHU					(74)

Table 1.1: Incision activity of oxidative damage DNA glycosylases.

Lesions for which a glycosylase has incision activity are indicated in the table, as well as the origin of the glycosylase, structure of the lesion-containing substrate DNA employed,

including double-stranded (ds), single- stranded (ss), *in vitro* transcription bubble oligonucleotide model containing several mismatched bases, or quadruplex DNA. 5-OHU, 5-hydroxyuracil; DHU, dihydrouracil; FapyG, 2,6-diamino-4-oxo-formamidopyrimidine; FapyA, 4,6-diamino-5-formamidopyrimidine; 8-oxoG, 8-oxoguanine; Tg, thymine glycol; Sp, spiroiminodihydantoin; Gh, guanidinohydantoin; 5-OHC, 5-hydroxycytidine.

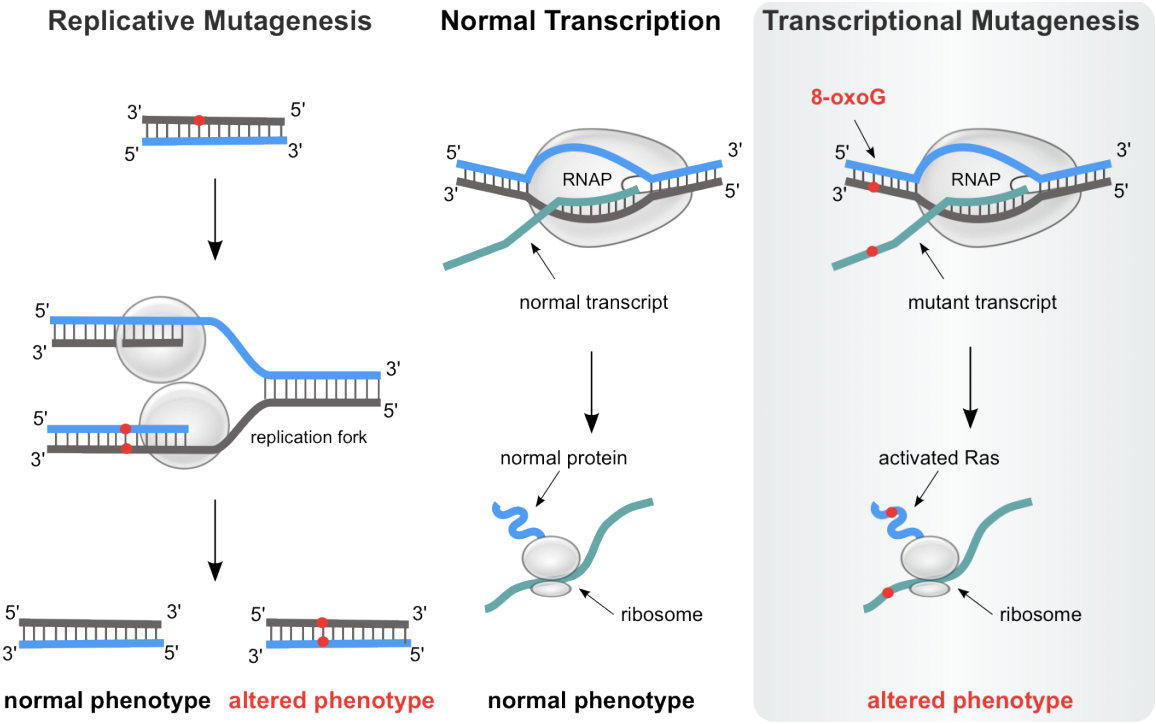


Figure 1.2: Consequences of unrepaired DNA damage.

DNA polymerase bypass of DNA damage can be mutagenic and lead to phenotypic change with potentially deleterious consequences in one of the daughter cells following cellular division. However, similarly to DNA polymerase, RNA polymerase can also misincorporate bases opposite the lesion during transcription, which can result in the production of mutant proteins that could also induce phenotypic change in proliferating as well as quiescent cells. Modified from (2).

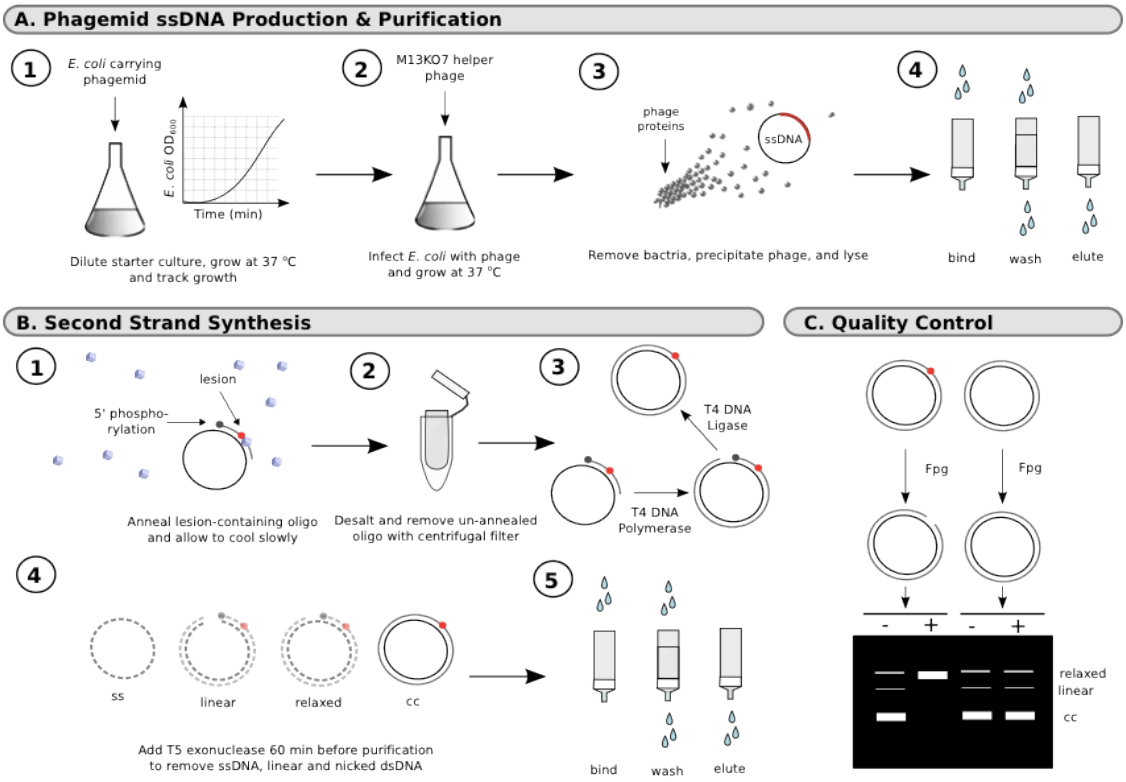


Figure 1.3: Method for production of mammalian expression vectors containing site-specific DNA damages.

General steps of the method include large-scale cultures of DH12S *E. coli* carrying a phagemid of interest, which are infected with M13KO7 helper phage in order to produce circular single-stranded DNA (ssDNA) for the strand of choice. The bacteria are pelleted and filtered out, phage are precipitated from the medium and lysed, and ssDNA is purified by anion-exchange columns. Synthetic lesion-containing oligos are annealed to the ssDNA, salt and unannealed oligos are removed using molecular weight cut-off spin columns and the second strand is synthesized using T4 DNA polymerase, and ligated using T4 DNA ligase. The double-stranded product is optionally treated with T5 exonuclease to digest nicked and linear product, and then it is purified using anion-exchange columns. The presence of a variety of lesions can be confirmed using a nicking assay employing glycosylases, such as *E. coli* formamidopyrimidine DNA glycosylase (Fpg).

Lesion	RNAP	Nucleotide	Experiment Notes	Ref.
Direct Repair and Base Excision Repair Substrates and Intermediates				
O^6-meG	Human	U (3-58%)	*In vivo*, -/+ O^6-benzylguanine	(53)
	Human	U (25%)	*In vitro*	(75)
	Phage (T7)	U (47%), A (6%)	*In vitro*	(75)
8-oxoG	Mouse	A (12%), del (2%)	*In vivo,* Ogg1$^{-/-}$ MEFs	(40)
	E. coli	C (41%), A (33%), del (26%)	*In vivo, mutM*$^-$*mfd*$^-$ *E. coli*	(42)
	Human	A (8%)	*In vitro*	(36)
	Phage (T7)	A, C	*In vitro*	(76)
5-OHU	Human	A (98%)	*In vitro*	(36)
DHU	*E. coli*	A	*In vitro*	(77)
	Phage (SP6, T7)	A	*In vitro*	(78)
Tg	Human	A (2%)	*In vitro*	(36)
AP site	*E. coli*	A (100%)	*In vivo, xth*$^-$*nfo*$^-$*nfi*$^-$ *E. coli*	(38)
	E.coli	A	*In vitro*	(35)
	Phage (SP6)	A	*In vitro*	(35)
SSB	*E. coli*	A (40-45%)	*In vivo, xth*$^-$*nfi*$^-$ *E. coli*	(38)
	E. coli	Stalled	*In vitro*	(35)
	Phage (SP6)	Stalled	*In vitro*	(35)
U	*E. coli*	A	*In vivo, ung*$^-$*mug*$^-$ *E. Coli*	(42)
	Human	A, G	*In vitro*	(79)
Nucleotide Excision Repair Substrates				
BPDE-A(-)	Phage (T7)	A, G, del	*In vitro*	(80)
BPDE-A(+)	Phage (T7)	A, G, del	*In vitro*	(80)
BPDE-G	Phage (T7)	C	*In vitro,* elongation block	(81)
1,2-d(GpG)	Yeast	AC	*In vitro*, elongation block	(49)
CPD	Human	GA (1.3-5.8%)	*In vivo*, -/+ NER deficiency	(50)
	Yeast	Not mutagenic	*In vitro*	(47)
*N*3-CMdT	Human*	U (40-65%)	*In vivo,* -/+ NER deficiency	(45)
O^4-CMdT	Human*	G (2-3%)	*In vivo,* -/+ NER deficiency	(45)
*N*3-EtdT	Human*	T (20-34%), C (11-18%)	*In vivo*, -/+ NER deficiency	(46)
O^2-EtdT	Human*	G, T, C (2-5%)	*In vivo*, -/+ NER deficiency	(46)
O^4-EtdT	Human*	G (34%)	*In vivo*, -/+ NER deficiency	(46)
^{S}G	Human*	Not mutagenic	*In vivo*	(44)
S^6-mG	Human*	U (20-75%)	*In vivo*, -/+ NER deficiency	(44)
CydA	Human*	A (21%)	*In vivo*, 5' mutation, NER-deficient cells	(43)
	Yeast	A	*In vitro*, 5' mutation	(48)
CydG	Human*	A (32%)	*In vivo*, 5' mutation, NER-deficient cells	(43)
N^2-CEdG	Human*	Not mutagenic	*In vivo*	(43)
AF-G	T7	Not mutagenic	*In vitro*	(76)
AAF-G	T7	Not mutagenic	*In vitro*	(76)

Table 1.2: Transcriptionally mutagenic DNA damage lesions.

Lesions studied in the context of transcriptional mutagenesis. RNAP indicates the polymerase employed and nucleotide indicates base misincorporated opposite the lesion with approximate frequencies where available, depending on repair defect if applicable, as percent of total transcript indicated in brackets. O^6-meG, O^6-methylguanine; 8-oxoG, 8-oxoguanine; 5-OHU, 5-hydroxyuracil; DHU, dihydrouracil; Tg, thymine glycol; AP site, abasic site; SSB, single-strand break; U, uracil; BPDE, benzo[a]pyrene diol epoxide; CPD, cyclobutane pyrimidine dimer; N^3-CMdT, N^3-carboxymethylthymidine; O^4-CMdT , O^4-carboxymethylthymidine; *N*3-EtdT, N3-ethylthymidine; O^2-EtdT, O^2-ethylthymidine; O^4-EtdT, O^4-ethylthymidine; ^{S}G, 6-thioguanine; S^6-mG, S^6-methylthioguanine; CydA, 8,5'-cyclo-2'-deoxyadenosine; CydG, 8,5'-cyclo-2'-deoxyguanosine; N^2-CEdG, N^2-(1-carboxymethyl)-2'-deoxyguanosine; AF, 2-aminofluorene; AAF, N-acetyl-2-aminofluorene. *The transcriptional mutagenicity was also determined following *in vitro* human RNAPII bypass in HeLa nuclear extracts and/or T7 DNA polymerase in the same study.

References

1. Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature*, **362**, 709-715.
2. Morreall, J.F., Petrova, L. and Doetsch, P.W. (2013) Transcriptional mutagenesis and its potential roles in the etiology of cancer and bacterial antibiotic resistance. *J Cell Physiol*, **228**, 2257-2261.
3. Bregeon, D. and Doetsch, P.W. (2011) Transcriptional mutagenesis: causes and involvement in tumour development. *Nat Rev Cancer*, **11**, 218-227.
4. Hanahan, D. and Weinberg, R.A. (2011) Hallmarks of cancer: the next generation. *Cell*, **144**, 646-674.
5. Saxowsky, T.T. and Doetsch, P.W. (2006) RNA polymerase encounters with DNA damage: transcription-coupled repair or transcriptional mutagenesis? *Chem Rev*, **106**, 474-488.
6. Marteijn, J.A., Lans, H., Vermeulen, W. and Hoeijmakers, J.H. (2014) Understanding nucleotide excision repair and its roles in cancer and ageing. *Nat Rev Mol Cell Biol*, **15**, 465-481.
7. Cleaver, J.E., Lam, E.T. and Revet, I. (2009) Disorders of nucleotide excision repair: the genetic and molecular basis of heterogeneity. *Nat Rev Genet*, **10**, 756-768.
8. Hung, R.J., Hall, J., Brennan, P. and Boffetta, P. (2005) Genetic polymorphisms in the base excision repair pathway and cancer risk: a HuGE review. *Am J Epidemiol*, **162**, 925-942.

9. Wallace, S.S., Murphy, D.L. and Sweasy, J.B. (2012) Base excision repair and cancer. *Cancer Lett*, **327**, 73-89.

10. Weren, R.D., Ligtenberg, M.J., Kets, C.M., de Voer, R.M., Verwiel, E.T., Spruijt, L., van Zelst-Stams, W.A., Jongmans, M.C., Gilissen, C., Hehir-Kwa, J.Y. *et al*. (2015) A germline homozygous mutation in the base-excision repair gene NTHL1 causes adenomatous polyposis and colorectal cancer. *Nat Genet*, **47**, 668-671.

11. Helbock, H.J., Beckman, K.B., Shigenaga, M.K., Walter, P.B., Woodall, A.A., Yeo, H.C. and Ames, B.N. (1998) DNA oxidation matters: the HPLC-electrochemical detection assay of 8-oxo-deoxyguanosine and 8-oxo-guanine. *Proc Natl Acad Sci U S A*, **95**, 288-293.

12. Lindahl, T. and Nyberg, B. (1972) Rate of depurination of native deoxyribonucleic acid. *Biochemistry*, **11**, 3610-3618.

13. Gates, K.S. (2009) An overview of chemical processes that damage cellular DNA: spontaneous hydrolysis, alkylation, and reactions with radicals. *Chem Res Toxicol*, **22**, 1747-1760.

14. Wilson, D.M., 3rd and Bohr, V.A. (2007) The mechanics of base excision repair, and its relationship to aging and disease. *DNA Repair (Amst)*, **6**, 544-559.

15. Dianov, G.L. and Hubscher, U. (2013) Mammalian base excision repair: the forgotten archangel. *Nucleic Acids Res*, **41**, 3483-3490.

16. Bauer, N.C., Corbett, A.H. and Doetsch, P.W. (2015) The current state of eukaryotic DNA base damage and repair. *Nucleic Acids Res*, **43**, 10083-10101.

17. Cooke, M.S., Evans, M.D., Dizdaroglu, M. and Lunec, J. (2003) Oxidative DNA damage: mechanisms, mutation, and disease. *Faseb J*, **17**, 1195-1214.

18. Fang, F.C. (2004) Antimicrobial reactive oxygen and nitrogen species: concepts and controversies. *Nat Rev Microbiol*, **2**, 820-832.

19. Maynard, S., Schurman, S.H., Harboe, C., de Souza-Pinto, N.C. and Bohr, V.A. (2009) Base excision repair of oxidative DNA damage and association with cancer and aging. *Carcinogenesis*, **30**, 2-10.

20. Wilson, D.M., 3rd and McNeill, D.R. (2007) Base excision repair and the central nervous system. *Neuroscience*, **145**, 1187-1200.

21. Hazra, T.K., Izumi, T., Boldogh, I., Imhoff, B., Kow, Y.W., Jaruga, P., Dizdaroglu, M. and Mitra, S. (2002) Identification and characterization of a human DNA glycosylase for repair of modified bases in oxidatively damaged DNA. *Proc Natl Acad Sci U S A*, **99**, 3523-3528.

22. Hazra, T.K., Kow, Y.W., Hatahet, Z., Imhoff, B., Boldogh, I., Mokkapati, S.K., Mitra, S. and Izumi, T. (2002) Identification and characterization of a novel human DNA glycosylase for repair of cytosine-derived lesions. *J Biol Chem*, **277**, 30417-30420.

23. Klapacz, J., Lingaraju, G.M., Guo, H.H., Shah, D., Moar-Shoshani, A., Loeb, L.A. and Samson, L.D. (2010) Frameshift mutagenesis and microsatellite instability induced by human alkyladenine DNA glycosylase. *Mol Cell*, **37**, 843-853.

24. Chan, K., Houlbrook, S., Zhang, Q.M., Harrison, M., Hickson, I.D. and Dianov, G.L. (2007) Overexpression of DNA polymerase beta results in an increased rate of frameshift mutations during base excision repair. *Mutagenesis*, **22**, 183-188.

25. Hofseth, L.J., Khan, M.A., Ambrose, M., Nikolayeva, O., Xu-Welliver, M., Kartalou, M., Hussain, S.P., Roth, R.B., Zhou, X., Mechanic, L.E. *et al*. (2003) The adaptive imbalance in base excision-repair enzymes generates microsatellite instability in chronic inflammation. *J Clin Invest*, **112**, 1887-1894.

26. Odell, I.D., Wallace, S.S. and Pederson, D.S. (2013) Rules of engagement for base excision repair in chromatin. *J Cell Physiol*, **228**, 258-266.

27. Ura, K., Araki, M., Saeki, H., Masutani, C., Ito, T., Iwai, S., Mizukoshi, T., Kaneda, Y. and Hanaoka, F. (2001) ATP-dependent chromatin remodeling facilitates nucleotide excision repair of UV-induced DNA lesions in synthetic dinucleosomes. *Embo J*, **20**, 2004-2014.

28. Gong, F., Fahy, D. and Smerdon, M.J. (2006) Rad4-Rad23 interaction with SWI/SNF links ATP-dependent chromatin remodeling with nucleotide excision repair. *Nat Struct Mol Biol*, **13**, 902-907.

29. Parsons, J.L., Kavli, B., Slupphaug, G. and Dianov, G.L. (2007) NEIL1 is the major DNA glycosylase that processes 5-hydroxyuracil in the proximity of a DNA single-strand break. *Biochemistry*, **46**, 4158-4163.

30. Shikazono, N., Noguchi, M., Fujii, K., Urushibara, A. and Yokoya, A. (2009) The yield, processing, and biological consequences of clustered DNA damage induced by ionizing radiation. *J Radiat Res*, **50**, 27-36.

31. Eccles, L.J., O'Neill, P. and Lomax, M.E. (2011) Delayed repair of radiation induced clustered DNA damage: friend or foe? *Mutat Res*, **711**, 134-141.

32. Sugasawa, K., Ng, J.M., Masutani, C., Iwai, S., van der Spek, P.J., Eker, A.P., Hanaoka, F., Bootsma, D. and Hoeijmakers, J.H. (1998) Xeroderma pigmentosum

group C protein complex is the initiator of global genome nucleotide excision repair. *Mol Cell*, **2**, 223-232.

33. Fu, D., Calvo, J.A. and Samson, L.D. (2012) Balancing repair and tolerance of DNA damage caused by alkylating agents. *Nat Rev Cancer*, **12**, 104-120.

34. Yoshioka, K., Yoshioka, Y. and Hsieh, P. (2006) ATR kinase activation mediated by MutSalpha and MutLalpha in response to cytotoxic O6-methylguanine adducts. *Mol Cell*, **22**, 501-510.

35. Zhou, W. and Doetsch, P.W. (1993) Effects of abasic sites and DNA single-strand breaks on prokaryotic RNA polymerases. *Proc Natl Acad Sci U S A*, **90**, 6601-6605.

36. Charlet-Berguerand, N., Feuerhahn, S., Kong, S.E., Ziserman, H., Conaway, J.W., Conaway, R. and Egly, J.M. (2006) RNA polymerase II bypass of oxidative DNA damage is regulated by transcription elongation factors. *Embo J*, **25**, 5481-5491.

37. Strauss, B.S. (1991) The 'A rule' of mutagen specificity: a consequence of DNA polymerase bypass of non-instructional lesions? *Bioessays*, **13**, 79-84.

38. Clauson, C.L., Oestreich, K.J., Austin, J.W. and Doetsch, P.W. (2010) Abasic sites and strand breaks in DNA cause transcriptional mutagenesis in Escherichia coli. *Proc Natl Acad Sci U S A*, **107**, 3657-3662.

39. Kuraoka, I., Suzuki, K., Ito, S., Hayashida, M., Kwei, J.S., Ikegami, T., Handa, H., Nakabeppu, Y. and Tanaka, K. (2007) RNA polymerase II bypasses 8-oxoguanine in the presence of transcription elongation factor TFIIS. *DNA Repair (Amst)*, **6**, 841-851.

40. Saxowsky, T.T., Meadows, K.L., Klungland, A. and Doetsch, P.W. (2008) 8-Oxoguanine-mediated transcriptional mutagenesis causes Ras activation in mammalian cells. *Proc Natl Acad Sci U S A*, **105**, 18877-18882.

41. Bregeon, D., Peignon, P.A. and Sarasin, A. (2009) Transcriptional mutagenesis induced by 8-oxoguanine in mammalian cells. *PLoS Genet*, **5**, e1000577.

42. Bregeon, D., Doddridge, Z.A., You, H.J., Weiss, B. and Doetsch, P.W. (2003) Transcriptional mutagenesis induced by uracil and 8-oxoguanine in Escherichia coli. *Mol Cell*, **12**, 959-970.

43. You, C., Dai, X., Yuan, B., Wang, J., Brooks, P.J., Niedernhofer, L.J. and Wang, Y. (2012) A quantitative assay for assessing the effects of DNA lesions on transcription. *Nat Chem Biol*, **8**, 817-822.

44. You, C., Dai, X., Yuan, B. and Wang, Y. (2012) Effects of 6-thioguanine and S6-methylthioguanine on transcription in vitro and in human cells. *J Biol Chem*, **287**, 40915-40923.

45. You, C., Wang, J., Dai, X. and Wang, Y. (2015) Transcriptional inhibition and mutagenesis induced by N-nitroso compound-derived carboxymethylated thymidine adducts in DNA. *Nucleic Acids Res*, **43**, 1012-1018.

46. You, C., Wang, P., Dai, X. and Wang, Y. (2014) Transcriptional bypass of regioisomeric ethylated thymidine lesions by T7 RNA polymerase and human RNA polymerase II. *Nucleic Acids Res*, **42**, 13706-13713.

47. Walmacq, C., Cheung, A.C., Kireeva, M.L., Lubkowska, L., Ye, C., Gotte, D., Strathern, J.N., Carell, T., Cramer, P. and Kashlev, M. (2012) Mechanism of

translesion transcription by RNA polymerase II and its role in cellular resistance to DNA damage. *Mol Cell*, **46**, 18-29.

48. Walmacq, C., Wang, L., Chong, J., Scibelli, K., Lubkowska, L., Gnatt, A., Brooks, P.J., Wang, D. and Kashlev, M. (2015) Mechanism of RNA polymerase II bypass of oxidative cyclopurine DNA lesions. *Proc Natl Acad Sci U S A*, **112**, E410-419.

49. Damsma, G.E., Alt, A., Brueckner, F., Carell, T. and Cramer, P. (2007) Mechanism of transcriptional stalling at cisplatin-damaged DNA. *Nat Struct Mol Biol*, **14**, 1127-1133.

50. Nagel, Z.D., Margulies, C.M., Chaim, I.A., McRee, S.K., Mazzucato, P., Ahmad, A., Abo, R.P., Butty, V.L., Forget, A.L. and Samson, L.D. (2014) Multiplexed DNA repair assays for multiple lesions and multiple doses via transcription inhibition and transcriptional mutagenesis. *Proc Natl Acad Sci U S A*, **111**, E1823-1832.

51. Brueckner, F., Hennecke, U., Carell, T. and Cramer, P. (2007) CPD damage recognition by transcribing RNA polymerase II. *Science*, **315**, 859-862.

52. Viswanathan, A., You, H.J. and Doetsch, P.W. (1999) Phenotypic change caused by transcriptional bypass of uracil in nondividing cells. *Science*, **284**, 159-162.

53. Burns, J.A., Dreij, K., Cartularo, L. and Scicchitano, D.A. (2010) O6-methylguanine induces altered proteins at the level of transcription in human cells. *Nucleic Acids Res*, **38**, 8178-8187.

54. Doetsch, P.W. (2002) Translesion synthesis by RNA polymerases: occurrence and biological implications for transcriptional mutagenesis. *Mutat Res*, **510**, 131-140.

55. Berndt, N., Hamilton, A.D. and Sebti, S.M. (2011) Targeting protein prenylation for cancer therapy. *Nat Rev Cancer*, **11**, 775-791.

56. Schubbert, S., Shannon, K. and Bollag, G. (2007) Hyperactive Ras in developmental disorders and cancer. *Nat Rev Cancer*, **7**, 295-308.

57. Serrano, M., Lin, A.W., McCurrach, M.E., Beach, D. and Lowe, S.W. (1997) Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. *Cell*, **88**, 593-602.

58. Bartkova, J., Rezaei, N., Liontos, M., Karakaidos, P., Kletsas, D., Issaeva, N., Vassiliou, L.V., Kolettas, E., Niforou, K., Zoumpourlis, V.C. *et al*. (2006) Oncogene-induced senescence is part of the tumorigenesis barrier imposed by DNA damage checkpoints. *Nature*, **444**, 633-637.

59. Tu, Z., Aird, K.M., Bitler, B.G., Nicodemus, J.P., Beeharry, N., Xia, B., Yen, T.J. and Zhang, R. (2011) Oncogenic RAS regulates BRIP1 expression to induce dissociation of BRCA1 from chromatin, inhibit DNA repair, and promote senescence. *Dev Cell*, **21**, 1077-1091.

60. Lee, A.C., Fenster, B.E., Ito, H., Takeda, K., Bae, N.S., Hirai, T., Yu, Z.X., Ferrans, V.J., Howard, B.H. and Finkel, T. (1999) Ras proteins induce senescence by altering the intracellular levels of reactive oxygen species. *J Biol Chem*, **274**, 7936-7940.

61. Michaloglou, C., Vredeveld, L.C., Soengas, M.S., Denoyelle, C., Kuilman, T., van der Horst, C.M., Majoor, D.M., Shay, J.W., Mooi, W.J. and Peeper, D.S. (2005) BRAFE600-associated senescence-like cell cycle arrest of human naevi. *Nature*, **436**, 720-724.

62. Collado, M. and Serrano, M. (2010) Senescence in tumours: evidence from mice and humans. *Nat Rev Cancer*, **10**, 51-57.

63. Hegde, M.L., Hegde, P.M., Bellot, L.J., Mandal, S.M., Hazra, T.K., Li, G.M., Boldogh, I., Tomkinson, A.E. and Mitra, S. (2013) Prereplicative repair of oxidized bases in the human genome is mediated by NEIL1 DNA glycosylase together with replication proteins. *Proc Natl Acad Sci U S A*, **110**, E3090-3099.

64. Dou, H., Mitra, S. and Hazra, T.K. (2003) Repair of oxidized bases in DNA bubble structures by human DNA glycosylases NEIL1 and NEIL2. *J Biol Chem*, **278**, 49679-49684.

65. Rosenquist, T.A., Zaika, E., Fernandes, A.S., Zharkov, D.O., Miller, H. and Grollman, A.P. (2003) The novel DNA glycosylase, NEIL1, protects mammalian cells from radiation-mediated cell death. *DNA Repair (Amst)*, **2**, 581-591.

66. Takao, M., Kanno, S., Kobayashi, K., Zhang, Q.M., Yonei, S., van der Horst, G.T. and Yasui, A. (2002) A back-up glycosylase in Nth1 knock-out mice is a functional Nei (endonuclease VIII) homologue. *J Biol Chem*, **277**, 42205-42213.

67. Prakash, A., Carroll, B.L., Sweasy, J.B., Wallace, S.S. and Doublie, S. (2014) Genome and cancer single nucleotide polymorphisms of the human NEIL1 DNA glycosylase: activity, structure, and the effect of editing. *DNA Repair (Amst)*, **14**, 17-26.

68. Liu, M., Bandaru, V., Bond, J.P., Jaruga, P., Zhao, X., Christov, P.P., Burrows, C.J., Rizzo, C.J., Dizdaroglu, M. and Wallace, S.S. (2010) The mouse ortholog of NEIL3 is a functional DNA glycosylase in vitro and in vivo. *Proc Natl Acad Sci U S A*, **107**, 4925-4930.

69. Zhou, J., Fleming, A.M., Averill, A.M., Burrows, C.J. and Wallace, S.S. (2015) The NEIL glycosylases remove oxidized guanine lesions from telomeric and promoter quadruplex DNA structures. *Nucleic Acids Res*, **43**, 4039-4054.

70. Krokeide, S.Z., Laerdahl, J.K., Salah, M., Luna, L., Cederkvist, F.H., Fleming, A.M., Burrows, C.J., Dalhus, B. and Bjoras, M. (2013) Human NEIL3 is mainly a monofunctional DNA glycosylase removing spiroimindiohydantoin and guanidinohydantoin. *DNA Repair (Amst)*, **12**, 1159-1164.

71. Sejersted, Y., Hildrestrand, G.A., Kunke, D., Rolseth, V., Krokeide, S.Z., Neurauter, C.G., Suganthan, R., Atneosen-Asegg, M., Fleming, A.M., Saugstad, O.D. *et al*. (2011) Endonuclease VIII-like 3 (Neil3) DNA glycosylase promotes neurogenesis induced by hypoxia-ischemia. *Proc Natl Acad Sci U S A*, **108**, 18802-18807.

72. Ikeda, S., Biswas, T., Roy, R., Izumi, T., Boldogh, I., Kurosky, A., Sarker, A.H., Seki, S. and Mitra, S. (1998) Purification and characterization of human NTH1, a homolog of Escherichia coli endonuclease III. Direct identification of Lys-212 as the active nucleophilic residue. *J Biol Chem*, **273**, 21585-21593.

73. Hu, J., de Souza-Pinto, N.C., Haraguchi, K., Hogue, B.A., Jaruga, P., Greenberg, M.M., Dizdaroglu, M. and Bohr, V.A. (2005) Repair of formamidopyrimidines in DNA involves different glycosylases: role of the OGG1, NTH1, and NEIL1 enzymes. *J Biol Chem*, **280**, 40544-40551.

74. Dizdaroglu, M., Karakaya, A., Jaruga, P., Slupphaug, G. and Krokan, H.E. (1996) Novel activities of human uracil DNA N-glycosylase for cytosine-derived products of oxidative DNA damage. *Nucleic Acids Res*, **24**, 418-422.

75. Dimitri, A., Burns, J.A., Broyde, S. and Scicchitano, D.A. (2008) Transcription elongation past O6-methylguanine by human RNA polymerase II and bacteriophage T7 RNA polymerase. *Nucleic Acids Res*, **36**, 6459-6471.

76. Chen, Y.H. and Bogenhagen, D.F. (1993) Effects of DNA lesions on transcription elongation by T7 RNA polymerase. *J Biol Chem*, **268**, 5849-5855.

77. Liu, J. and Doetsch, P.W. (1998) Escherichia coli RNA and DNA polymerase bypass of dihydrouracil: mutagenic potential via transcription and replication. *Nucleic Acids Res*, **26**, 1707-1712.

78. Liu, J., Zhou, W. and Doetsch, P.W. (1995) RNA polymerase bypass at sites of dihydrouracil: implications for transcriptional mutagenesis. *Mol Cell Biol*, **15**, 6729-6735.

79. Kuraoka, I., Endou, M., Yamaguchi, Y., Wada, T., Handa, H. and Tanaka, K. (2003) Effects of endogenous DNA base lesions on transcription elongation by mammalian RNA polymerase II. Implications for transcription-coupled DNA repair and transcriptional mutagenesis. *J Biol Chem*, **278**, 7294-7299.

80. Remington, K.M., Bennett, S.E., Harris, C.M., Harris, T.M. and Bebenek, K. (1998) Highly mutagenic bypass synthesis by T7 RNA polymerase of site-specific benzo[a]pyrene diol epoxide-adducted template DNA. *J Biol Chem*, **273**, 13170-13176.

81. Choi, D.J., Roth, R.B., Liu, T., Geacintov, N.E. and Scicchitano, D.A. (1996) Incorrect base insertion and prematurely terminated transcripts during T7 RNA polymerase transcription elongation past benzo[a]pyrenediol epoxide-modified DNA. *J Mol Biol*, **264**, 213-219.

2. EFFICIENT AND RELIABLE PRODUCTION OF VECTORS FOR THE STUDY OF THE REPAIR, MUTAGENESIS, AND PHENOTYPIC CONSEQUENCES OF DEFINED DNA DAMAGE LESIONS IN MAMMALIAN CELLS

Lucy Petrova[1,2], Christine Gran[3], Magnar Bjoras[3,4] and Paul W. Doetsch[2,5,6,7*]

1. Program in Genetics and Molecular Biology, Emory University, Atlanta, GA
2. Department of Biochemistry, Emory University School of Medicine, Atlanta, GA
3. Department of Microbiology, Oslo University Hospital and University of Oslo, Norway
4. Department of Cancer Research and Molecular Medicine, Norwegian University of Science and Technology, Trondheim, Norway
5. Department of Radiation Oncology, Emory University School of Medicine, Atlanta, GA
6. Department of Hematology and Medical Oncology, Emory University School of Medicine,
7. Atlanta, GA Winship Cancer Institute, Emory University School of Medicine, Atlanta, GA

*To whom correspondence should be addressed. Tel: 404 727 0409 Fax: 404 727 3231; Email: medpwd@emory.edu

Keywords: DNA damage, site-specific base modifications, mammalian expression vectors, transcriptional mutagenesis

Published in PLoS One, doi: 10.1371/journal.pone.0158581.

Abstract

Mammalian cells are constantly and unavoidably exposed to DNA damage from endogenous and exogenous sources, frequently to the detriment of genomic integrity and biological function. Cells acquire a large number of chemically diverse lesions per day, and each can have a different genetic fate and biological consequences. However, our knowledge of how and when specific lesions are repaired or how they may compromise the fidelity of DNA replication or transcription and lead to deleterious biological endpoints in mammalian cells is limited. Studying individual lesions requires technically challenging approaches for the targeted introduction of defined lesions into relevant DNA sequences of interest. Here, we present a systematic analysis of factors influencing yield and an improved, efficient and reliable protocol for the production of mammalian expression phagemid vectors containing defined DNA base modifications in any sequence position of either complementary DNA strand. We applied our improved protocol to study the transcriptional mutagenesis-mediated phenotypic consequences of the common oxidative lesion 5-hydroxyuracil, placed in the G12 mutational hotspot of the KRAS oncogene. 5-OHU induced sustained oncogenic signaling in Neil1$^{-/-}$Neil2$^{-/-}$ mouse cells. The resulting advance in technology will have broad applicability for investigation of single lesion DNA repair, mutagenesis, and DNA damage responses in mammalian cells.

Introduction

Mammalian cells are continuously exposed to DNA damage, which can be detrimental to health and is associated with cancer, neurodegenerative disease and aging (1-3). Exposure to a variety of unavoidable endogenous sources, such as cellular respiration, and ubiquitous exogenous sources, such as UV light, X-rays and chemical agents, can result in a vast array of known and unknown lesions (4-6). At the molecular level, DNA lesions can cause variable levels of DNA or RNA polymerase stalling or mutagenic bypass during replication or transcription (7). Similarly to DNA polymerase, when RNA polymerase encounters DNA lesions, it can misincorporate incorrect nucleotides opposite to the lesions resulting in transcriptional mutagenesis (TM). At the cellular level, DNA lesions can induce senescence or cell death (8), and at the organismal level they can lead to disease (3). The consequences of DNA damage have been extensively studied in systems where populations of cells are exposed to DNA damaging agents. Such treatments result in, often extensive, DNA damage in the form of a variety of known and unknown lesions, at random genomic positions unique for each cell. Much can be understood regarding DNA damage and repair using these systems. However, delineating the fates and consequences of each specific lesion, in physiological contexts with or without the potentially confounding factor of additional damage that can induce phenotypes related to the levels of damage rather than any individual lesion, requires a more targeted approach. DNA and RNA polymerase bypass efficiencies, lesion stability and replicative or transcriptional mutagenicity, as well as the type of mutation introduced when mutagenesis occurs differ for each specific lesion, and different lesions are recognized and repaired by different components of the DNA repair pathways. All of

these factors, as well as the specific sequence position in which a lesion occurs, can determine biological outcomes. Each particular lesion, and the unique context in which it occurs, could differentially contribute to human disease. For example, a highly mutagenic lesion occurring in an oncogene mutational hotspot and evading DNA repair is likely to result in detrimental biological consequences, in contrast to one that is rarely mutagenic or quickly repaired.

DNA repair of defined lesions can be studied *in vitro* using radiolabeled synthetic oligodeoxynucleotides (ODNs) containing the desired modification, however this precludes the study of any potential phenotypic consequences, and biochemical characterizations may not be reproducible *in vivo*. For example, in a reconstituted *in vitro* transcription system, the common oxidative lesion 8-oxoguanine (8-oxoG) hinders transcription, nonetheless in HeLa nuclear extracts (9) and *in vivo* in bacterial (10) and mammalian (11,12) cells it is a miscoding lesion, efficiently bypassed by RNA polymerase, inducing transcriptional mutagenesis and phenotypic change. However, the ability of other oxidative lesions, such as 5-hydroxyuracil (5-OHU) or dihydrouracil (DHU) to compromise the fidelity of transcription *in vivo* and induce phenotypic change have not been investigated. If left unrepaired, they have the potential to induce more pervasive and long-lasting phenotypic change due to significantly higher levels of mutagenesis following RNA polymerase (RNAP) bypass that have been demonstrated *in vitro* (9,13).

Several methods, with different advantages and limitations, have been developed for the study of individual lesions *in vivo*. One strategy, the gapped duplex method, involves the digestion of a plasmid using sequence-specific nicking enzymes and

exchanging the excised ODN with one containing the lesion of interest (14). However, the applications of this approach are limited, as it requires the presence of two tandem nicking endonuclease recognition sequences adjacent to the lesion site. A more versatile approach involves the annealing of a synthetic lesion-containing ODN to a single-stranded vector followed by enzymatic complementary strand synthesis. While the single-stranded (ss) M13 phage genome can also be used (15), the use of ss phagemid vectors (16,17) permits the use of any mammalian expression vector, containing any feature and gene or sequence of interest, so long as it also contains the f1 phage origin of replication. Such mammalian expression vectors containing site-specific base modifications in any position or sequence of interest are an enabling technology for the study of the consequences of defined DNA damage lesions occurring in targeted positions of protein-coding regions of genes, such as oncogene mutational hotspots, in mammalian cells (12). Also, when the lesion of interest is strategically placed in a reporter gene, such as a fluorescent protein from which the normal fluorescent sequence is transcribed only when transcriptional mutagenesis occurs, these vectors allow for the study of the mechanisms and regulation of DNA damage repair as well as monitor repair capacities and repair dynamics of known lesions of interest in live mammalian cells (18). The use of these tools complements and can vastly extend our understanding of DNA damage and repair in contexts relevant to human health. However, their construction poses significant technical challenges. The necessary protocols are complex, time-consuming and laborious, they can require the use of expensive or toxic chemicals, and can result in low or unpredictable yields. Mammalian transfection requires large

quantities of highly pure DNA, however the factors that determine yield and protocol reproducibility are not well characterized.

In order to simplify the protocol and determine factors influencing yield, reproducibility, and achieve highest final product quality, we performed systematic analyses of the steps for the production and purification of single-stranded phagemid DNA and double-stranded vectors containing site-specific DNA damage lesions for applications in mammalian cell culture systems. We present an optimized protocol yielding large quantities of ultra-pure, double-stranded, lesion-containing vectors well suited for mammalian transfection. In order to improve protocol reproducibility between different preparations and experimenters, we provide reliable predictors of yield. We applied this improved protocol to study the phenotypic consequences resulting from mutagenic transcriptional bypass of 5-OHU placed the G12 mutational hotspot of KRAS, such that when TM occurs due to the misincorporation of adenine opposite to 5-OHU, it would result in the production of constitutively active $KRAS^{G12D}$ transcripts and proteins activating downstream effectors of Ras. Incorporation of guanine opposite this cytosine-derived lesion, due to it not being transcriptionally mutagenic or repaired, would result in wild type transcripts and proteins. We find that in $Neil1^{-/-}Neil2^{-/-}$ (Nei endonuclease VIII-like 1 and 2) DNA glycosylase deficient mouse embryonic fibroblasts, 5-OHU induces sustained transcriptional mutagenesis-mediated oncogene activation, implicating Neil1 and/or Neil2 in the repair of 5-OHU *in vivo*. Such oncogene activation via a TM mechanism, in comparison to that mediated by 8-oxoG shown previously (12), is sustained much longer than previously known. Moreover, we show TM activates AKT – a pathway downstream of Ras previously not known to be activated via a TM

mechanism. Thus, the continuous time course of transcriptional mutagenesis-mediated changes in cellular signaling and the activation of multiple downstream effectors of Ras can potentially be significant through inducing a variety of detrimental physiological consequences.

Materials and Methods

Plasmids and Cloning Procedures. The backbone for all vectors used was pcDNA3.1(+) (Invitrogen). pcDNA3.1(+) vectors contain the f1 origin of replication in the forward direction, such that the non-transcribed strand is the produced ssDNA. EGFP was sub-cloned from pEGFP-N1 into pcDNA3.1(+) by restriction digestion using BamHI and NotI and standard cloning procedures. The $HRAS^{WT}$ and $KRAS^{WT}$ plasmids containing the human sequences were from Guthrie cDNA Resource Center. The $HRAS^{Q61K}$ plasmid has been previously described (12). The $KRAS^{G12D}$ mutant was produced using the $KRAS^{WT}$ plasmid and the QuickChange Site-Directed Mutagenesis Kit (Agilent Technologies) according to the manufacturer's instructions and primers with the sequence CTCTTGCCTACGCCATCAGCTCCAACTACC (forward) and GGTAGTTGGAGCTGATGGCGTAGGCAAGAG (reverse).

Escherichia coli culture and M13KO7 phage. DH12S *E. coli* cells (Invitrogen Cat. #18312-017), which are endA+ in order to mimimize dsDNA production, were transformed with pcDNA3.2(+) plasmids containing the insert of interest and grown in LB-Miller medium containing 100 μg/mL carbenicillin at 37 ºC and 225 rpm. On the day of phage infection, cultures were diluted 1:500 in 2X-YT medium with carbenicillin. For large-scale cultures, 200 mL of culture in 2 L baffled flasks were used, and for small-scale cultures, used for systematic analysis of phage infection conditions, 20 mL in 250

mL baffled flasks. Cells were infected with M13KO7 phage stock at > 1 x 10^{11} pfu/mL (Invitrogen Cat. #18311-019) at the MOI indicated for each sample. The *E. coli* density was determined by measuring OD_{600} and assuming OD_{600} of 1.0 = 8 x 10^8 cells/mL. The cultures were then incubated for 30 minutes without shaking and an additional 1.5 hours with shaking before adding kanamycin to a final concentration of 75 μg/mL and incubating with shaking overnight. The next day, bacteria were pelleted by centrifugation and supernatants filtered through a low protein binding 0.22 μm filter (Corning, Cat. #431097).

SDS Phage Lysis and Anion-Exchange ssDNA Purification. PEG-8000 was purchased from Sigma. Phage were precipitated by the addition of 0.2 volumes solution M1 (3 M NaCl and 30% (w/v) PEG-8000) and incubating at 4 ºC for ≥ 1 hr. Phage were pelleted at ≥ 10,000 x g for 20 minutes at 4 ºC, supernatants decanted and pellets drip-dried for approximately 5 minutes. The pellets from one to two 200 mL starting cultures were purified using a single midi column from the PureLink® HiPure Plasmid Midiprep Kit (Invitrogen, Cat. # K2100-04). Briefly, each pellet was resuspended in 3 mL buffer M2 (100 mM Tris-HCl, pH 8.0 and 25 mM EDTA), then 3 mL solution M3 (4% SDS) was added, samples were mixed by inversion and then incubated at 70 ºC for 20 min. Then, 3 mL buffer N3 were added, samples were again mixed by inversion, and centrifuged at 14,000 x g for 10 min at room temperature. The supernatants were applied to columns pre-equilibrated with buffer EQ1 and DNA purified according to the manufacturer's instructions, except that the elution buffer (E4) was pre-warmed to 50 ºC before use. Pellets were resuspended in TE buffer, yields quantified using NanoDrop spectrophotometry, and a small sample of each ssDNA visualized on a 0.7% agarose gel

containing ethidium bromide to ensure purity.

For the silica spin column purifications, phage were precipitated and ssDNA purified using QIAprep spin M13 kit (Qiagen, Cat. #27704) according to the manufacturer's instructions and samples were pooled.

Determination of Phage Yield by Proteinase K Digestion. In order to determine amounts of ssDNA, buffer M1 precipitated phage were pelleted at 14,000 x g on a table-top microcentrifuge at 4 ºC for 20 min. Pellets from each mL of culture were resuspended in 50 μL TE pH 8.0 containing 100 μg proteinase K (Invitrogen, Cat. #25530-015) and 0.1% SDS and incubated at 42 ºC for 1 hr. The samples were then resolved on 0.7% agarose TBE gels containing ethidium bromide. Phagemid ssDNA band intensities were quantified using the ImageQuant TL software.

Oligodeoxynucleotide Annealing. PAGE-purified, 5' phosphorylated lesion-free and 8-oxoguanine-containing ODNs, with sequences as indicated in Table S1, were purchased from Eurofins MWG Operon, and those containing 5-hydroxyuracil or dihydrouracil were purchased from Midland Certified Reagent Company. 80 pmole of ODN, from 100 μM stock in TE buffer, were added for every 10 μg of ssDNA and annealed in 1X saline-sodium citrate (SSC) buffer, in a final volume of 50 μL, at 75 ºC for 10 minutes in a sterile microcentrifuge tube placed in a water beaker in a 37 ºC water bath, after which they were allowed to cool slowly to room temperature. Amicon Ultra-0.5 30K centrifugal filter units (Millipore, Cat. # UFC503024), which are recommended for efficient removal of primers ranging from 10-48 bases, were used according to the manufacturer's instructions to concentrate the products and remove salts and unannealed ODNs. Briefly, for each sample, PCR grade water was added to a final volume of 500

μL, and then the concentration centrifugation step was carried out at 14,000 x g for 10 minutes and the elution step at 1,000 x g for 2 minutes.

Second Strand Synthesis. PEG-8000 was dissolved in nuclease-free water and filtered through a 0.45 μm filter. T4 DNA ligase (Cat. #15224-017) and T4 DNA polymerase (Cat. #18005-017) were from Invitrogen. All other reagents were purchased from New England Biolabs. Second strand synthesis was performed overnight as previously described (16), scaling up or down as necessary. Polymerization reactions starting with 80 μg of ssDNA annealed with ODN were carried out in 1X NEBuffer 2 (10 mM Tris-HCl, 10 mM MgCl2, 50 mM NaCl, 1 mM DTT, pH 7.9), containing 7.5% PEG-8000, 50 μg/mL BSA, 600 μM each dNTP, 1 mM ATP, 80 U T4 DNA ligase, and 40 U T4 DNA polymerase, in a total volume of 600 μL. The samples were incubated for 5 min on ice, then 5 min at room temperature, before incubating at 37 ºC overnight. In order to enzymatically digest nicked, linear and single-stranded DNA, T5 exonuclease (New England Biolabs, Cat. #M0363) was added for the samples indicated in the text, directly to the second strand synthesis reaction or to purified dsDNA product in buffer NEBuffer 2 at 5 units per μg of starting ssDNA or dsDNA, respectively, and incubated for an hour at 37 ºC. Reactions were stopped by the addition of EDTA to a final concentration of 11 μM and DNA purified.

dsDNA Purification Using Anion-Exchange Columns. Double-stranded products were purified using anion-exchange column kits (Qiagen, Cat. #12123 and 12243), using the products of no more than 150 μL second-strand synthesis reaction per Tip-20 column with capacity of 20 μg and 600 μL per Tip-100 column. Each sample was resuspended in at least 10 volumes of 750 mM NaCl, 50 mM MOPS pH 7.0 resuspension buffer, then

applied to a Tip-20 or Tip-100 column pre-equilibrated with buffer QBT. The remainder of the protocol was according to the manufacturer's instructions for plasmid mini- or midiprep procedures. Purified DNA pellets were resuspended in buffer TE, pH 8.0 and yields were determined using NanoDrop spectrophotometry.

Fpg and Nth Nicking Assays. 250 ng of each construct were digested with Formamidopyrimidine DNA glycosylase (Fpg, New England Biolabs, Cat. #M0240S) using 1 μL of Fpg (8 units) in the presence of BSA, according to the manufacturer's instructions, in 1X NEBuffer 1 for 1 hr at 37 ºC. Products were separated and visualized on 0.7% agarose TBE gels containing ethidium bromide. For the Endonuclease III (Nth) assay, Endonuclease III (New England Biolabs, Cat. # M0268S) in 1X Endonuclease III reaction buffer was used instead.

Alkaline Gel Electrophoresis. T5 exonuclease treated or Fpg nicked control constructs were purified using a PureLink® PCR Purfication kit (Invitrogen, Cat. #K3100-01) as per the manufacturer's instructions, and eluted in nuclease-free dH_2O. All restriction enzymes were obtained from New England Biolabs. Constructs were digested with SmaI for 90 min at 25 ºC in 1X CutSmart® Buffer, NdeI was added and the samples were incubated at 37 ºC for 90 minutes, and then enzymes were heat inactivated at 65 ºC for 20 min. Samples were ethanol-precipitated and resuspended in 1X alkaline gel loading buffer, then heated for 10 minutes at 75 ºC, and cooled on ice for 3 minutes before loading on 0.6% alkaline gels. Alkaline agarose gel electrophoresis was performed as previously described (19). After neutralization, gels were stained four times, 15 minutes each, with 0.5 μg/mL ethidium bromide in TAE buffer and destained in TAE buffer without ethidium bromide.

Cell Culture and Mammalian Transfection. Primary mouse embryonic fibroblasts (MEFs) were generated from 13.5 days old C57BL/6 Neil1$^{-/-}$/2$^{-/-}$ embryos (Gran *et al*, manuscript in preparation). Limbs were removed from embryos, the tissue was chopped into small pieces and cell suspension was made by pipetting vigorously. MEFs were grown in DMEM medium (Gibco) supplemented with 10% fetal bovine serum (FBS, Sigma), 2 mM glutamine (GlutaMAX, Gibco) and 1x penicillin/streptomycin (Gibco). Cells grown for 4-5 days were frozen at passage 2 in DMEM medium with 20% FBS/10% DMSO. Neil1$^{-/-}$Neil2$^{-/-}$ MEFs displayed the same proliferation rate as wild type MEFs. Experimental procedures were approved by the Norwegian Animal Research Authority. MEFs were immortalized by frequent passaging, using the 3T3 protocol as described previously (20). Neil1$^{-/-}$Neil2$^{-/-}$ and Ogg1$^{-/-}$ MEFs, described previously (12), were cultured in a humidified incubator at 10% CO_2 in high-glucose DMEM containing GlutaMAX (Invitrogen) supplemented with 10% fetal bovine serum (GE Healthcare). MEFs were electroporated using an Amaxa Nucleofector 2B device and MEF 1 Nucleofector® kit (Lonza, Cat. #VPD-1004) using the T-020 setting as per the manufacturer's instructions and 3-4 μg of DNA per 1.0 x 10^6 to 1.5 x 10^6 cells.

Western Blot Analysis. Cells were washed with PBS, and switched to serum-free DMEM 1.5 hours before lysis. Samples were lysed using RIPA (150 mM NaCl, 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris pH 8.0) freshly supplemented with cOmplete protease inhibitor cocktail tablets (Roche, Cat. #04693159001) and Halt phosphatase inhibitor cocktail (ThermoFisher Scientific, Cat. #78420). Ten to twenty μg of protein were resolved using 10% NuPAGE Bis-Tris gels (Life Technologies) and transferred onto PVDF membranes for two hours at 80V.

Antibodies against phospho-AKT (Ser473, Cat. #4060), AKT (pan, Cat. #4691), phospho-ERK1/2 (Thr202/Tyr204, Cat. #9106) and ERK (Cat. #9102) were from Cell Signaling Technology, and K-Ras was from Santa Cruz Biotechnology (Cat.# sc-30). Membranes were blocked for 1 hour at room temperature in 2% ECL Prime blocking reagent (GE Healthcare, Cat. # RPN418) diluted in TBST, and antibodies were diluted 1:3,000 (pAKT and AKT), 1:200 (pERK1/2 and ERK), or 1:500 (K-Ras) in blocking buffer. Secondary horseradish peroxidase-conjugated antibodies were from Promega and diluted 1:5,000 (anti-mouse) or 1:10,000 (anti-rabbit) in blocking buffer. Three washes in TBST were carried out after each antibody incubation, and the membranes were developed for ECL and exposed to film. Membranes were cut and blotted for pAKT or pERK and K-Ras, then stripped and re-probed for total AKT or ERK. Film was scanned, images quantified using ImageQuant TL, the ratio of (pAKT/AKT)/K-Ras was determined for each sample and expressed as percent of the mutant positive control which was set to 100%.

Composition of buffers and media not described in Materials and Methods can be found in Table S2.

Results

Reliable Predictors of Single-Stranded Phagemid Yield

In order to produce phagemid single-stranded DNA (ssDNA), we infected log-phase DH12S *E. coli* cultures with the M13KO7 derivative of the M13 phage that preferentially packages phagemid ssDNA containing the f1 origin of replication, rather than its own genome, which it packages in the absence of phagemid (21). Which specific complementary strand of the phagemid is replicated and packaged depends on the

orientation of the f1 origin. For the study of transcriptional mutagenesis, we employed a vector, pcDNA3.1(+), containing f1 in the forward direction and for which the non-transcribed strand undergoes single stranded replication, allowing the annealing of a lesion-containing ODN which would become part of the transcribed strand. For investigations of lesions on the non-transcribed strand, a vector containing f1 in the opposite orientation, such as pcDNA3.1(-), can be used instead.

Mammalian transfection requires microgram quantities of DNA and thus high-yield, large-scale single-stranded phagemid production. However, one of the most significant obstacles to the reproducibility of phagemid production is that small variations in the experimental conditions can result in large differences in ssDNA yield. The growth medium, levels of aeration, multiplicity of infection (MOI) and growth stage of the target *E. coli* cells are all factors that could affect ssDNA yield. Different protocols provide different instructions for the timing of M13KO7 infection, including growing the cultures to early log-phase (21), for 3 hours (16), to an optical density at 600 nm (OD_{600}) of 0.1 (17), or 0.05 (22). Previous studies also recommend different MOIs, including 2-10 (21), 5 (17), or may not specify MOI (16). In order to identify optimal infection conditions and reliable predictors of ssDNA yield, we performed systematic analysis of the culture and M13KO7 infection conditions. We infected DH12S *E. coli* with M13KO7 phage at an identical MOI (> 2.5) and varying cell densities as well as at identical densities but varying MOI (Figure 2.1A). We found that the DH12S *E. coli* culture density at the time of M13KO7 infection is a reliable predictor of ssDNA yield. ssDNA yield increases as OD_{600} increases, before plateauing at OD_{600} of approximately 0.6, yielding a several-fold increase compared to the ssDNA yields at OD_{600} of approximately 0.05 or 0.1 (Figure

2.1B and Figure 2.1C). In cultures infected at the same OD_{600}, we find that increasing the MOI does not increase yield (Figure 2.1D and Figure 2.1E). Moreover, if cultures that otherwise produce high quantities of phagemid ssDNA are diluted before overnight incubation (Figure 2.1F) or after overnight incubation (Supplementary Figure 2.1), no ssDNA is produced. Thus, it appears that following the initial infection, the replication of M13KO7 may be insufficient for sustained phagemid production in the progeny of infected DH12S *E. coli* cells after multiple rounds of cell division. While the possibility that other factors may influence ssDNA yield in other *E. coli* strains cannot be excluded, the number of initially transduced DH12S cells is a reliable predictor of phagemid ssDNA yield.

Consistent with the systematic analysis of culture density at the time of infection, we also observed an increase in ssDNA yield with increasing OD_{600} at the time of M13KO7 infection in large-scale (200 mL) ssDNA preparations following anion-exchange column purification (Figure 2.2). By following these optimizations, it is possible to prepare more than 300 μg of highly pure phagemid ssDNA, with OD_{260}/OD_{280} ratio of 1.83 ± 0.02 s.d. (n = 27), from a single 200 mL culture. If even larger quantities of ssDNA are desired, the number of cultures can be increased, since increasing the culture volume reduces the culture aeration and thus ssDNA yield. The ssDNA can be stored long-term at -20 °C. We have successfully used over a year old ssDNA for second strand synthesis reactions.

Starting ssDNA yield and purity ultimately determine second strand synthesis product yield. In order to determine the yield of covalently closed double-stranded product, we used ssDNA purified by phenol:chloroform:isoamyl alcohol extraction

(PCIA), silica spin columns, and anion-exchange columns for the second strand synthesis reaction. ssDNA purified only by PCIA did not yield any dsDNA product (data not shown). While both anion-exchange columns and silica spin columns yielded dsDNA product, that of the silica column-purified ssDNA contained mostly nicked and linear vector while anion-exchange column purified ssDNA yielded the highest quantities of covalently closed product. Hence, purification by anion-exchange columns results in ssDNA highly suitable for second strand synthesis.

Purification of Highly Pure, Covalently-Closed, Double-Stranded Vectors.

Several options exist for the purification of double-stranded products that can yield varying levels of DNA recovery, purity, adventitious background damage, such as oxidation or UV damage, as well as amount of covalently closed plasmid. Background DNA damage can confound mutagenesis studies and vector purity as well as the amount of covalently closed plasmid can significantly affect transfection efficiencies. The general strategy employed here for second strand synthesis and construct purification is depicted in Figure 2.3A and Figure 2.3B (8-oxoguanine example). The second strand synthesis reactions do not result in all covalently closed product, but also contain the nicked and linear form of plasmid. The presence of nicked and linear product could be due to incomplete ligation by T4 DNA ligase, in which case the breaks would occur all in the same position and can be a confounding factor as single strand breaks can also be mutagenic, or at random positions. In order to determine whether the nicks occur all in the same position, we performed alkaline gel electrophoresis of constructs after digestion with restriction enzymes (SmaI and NdeI), selected such that each of these scenarios can be distinguished due to the different fragmentation patterns (Figure 2.3C, Figure 2.3D,

and Supplementary Figure 2.2) they would produce in denatured, single-stranded DNA. We find that while the positive control, Fpg nicked 8-oxoG construct, produces the expected lower molecular weight bands due to fragmentation of the nicked transcribed strand, the construct not treated with Fpg does not, indicating that nicks occur at random positions (Figure 2.3D), likely due to adventitious background single-strand break damage occurring during DNA manipulation procedures.

A previously used method, purification from low melting point agarose gels using β-agarase, allows for the isolation of initially closed circular product from the second strand synthesis reaction, but can result in significant levels of oxidation and single strand breaks (Supplementary Figure 2.3). Constructs purified by anion-exchange columns contain similar levels of nicked and linear form but produce significantly higher yield (36.1 μg ± 8.8 s.d., n = 38, versus about 15 μg (16) per 40 μg of starting ssDNA). Also, anion-exchange column purification can easily be scaled-up to hundreds of micrograms, without an increase in time or effort, by employing columns with larger capacity. We observe corresponding increases in product yields when scaling up second strand synthesis reactions (Figure 2.4A) and similar yields for lesion-free versus lesion-containing constructs (Figure 2.4B).

We performed second strand synthesis using ODNs containing three different oxidative lesions, including 8-oxoguanine, 5-hydroxyuracil (5-OHU), and dihydrouracil (DHU), and confirmed the presence of each lesion using the *E. coli* Formamidopyrimidine DNA glycosylase (Fpg) or Endonuclease III (Nth) nicking assays (Figure 2.5A and Figure 2.5B). Fpg cleaves oxidative lesions such as 8-oxoG and 5-OHU (23), leaving a single strand break, in this case resulting in nicked plasmid that can be

visualized on an agarose gel due to its altered migration pattern (Figure 2.5C). All constructs with second strands synthesized using the 8-oxoG or 5-OHU oligodeoxynucleotides were completely converted into the nicked form, indicating the presence of the lesions in these constructs as well as absence of detectable lesion-free dsDNA contamination, while those synthesized using lesion-free ODNs were not (Figure 2.5C). Dihydrouracil is a substrate for Nth (24), and the construct containing DHU was cut by Nth. While DHU has been described as a non-specific substrate for Fpg, DHU can be recognized by *B. stearothermophilus* Fpg in a manner similar to 8-oxoG (25), and can also be excised by *E. coli* Fpg from ODNs, albeit less efficiently than 8-oxoG from 8-oxoG:C pairs (26). We find that the DHU-containing construct is completely converted into nicked form in the Fpg cleavage assay (Figure 2.5C), indicating that DHU is also a substrate of Fpg and Fpg is also a suitable enzyme for determining the presence of DHU.

The presence of nicked and linear product could affect transfection efficiencies. In order to determine whether the presence of nicked vector affects transfection efficiency, we compared EGFP constructs purified using anion-exchange columns with or without enzymatic digestion of nicked, linear, and ssDNA using T5 exonuclease (15,27), and EGFP bacterial maxiprep. T5 exonuclease treatment followed by anion-exchange column purification results in highly pure closed circular product (Figure 2.6A and Figure 2.6B), albeit at the cost of reduction in yield. Higher yields can be obtained if the T5 exonuclease treatment is performed directly in the second strand synthesis reaction (Figure 2.6C). We find that treatment with T5 exonuclease does not improve transfection efficiencies, and EGFP constructs treated or not treated with the enzyme result in similar efficiencies (Figure 2.6D). We also found that bacterial EGFP plasmid purified using the

same method and of identical purity results in higher efficiencies than both types of constructs, likely due to differences in the plasmid coiling.

Lesion-Containing Constructs for the Study of the Phenotypic Consequences of Transcriptional Mutagenesis

8-oxoG can induce transcriptional mutagenesis and induce significant increases in extracellular-signal-regulated kinases 1 and 2 (ERK1/2) phosphorylation at 6 hours post-nucleofection in MEFs deficient in Ogg1, while it is almost immediately repaired in wild type (WT) cells (12). Oncogenic mutant Ras can activate multiple downstream signaling cascades, including the Raf-MEK-ERK pathway, PI3K-AKT-mTOR, and RalGEF-Ral. and regulate a variety of cellular processes and cancer hallmarks, including cellular proliferation, survival, and angiogenesis (28). However, downstream effectors of Ras in addition to ERK have not been previously studied in the context of TM. As the levels and duration of TM may vary, TM may differentially influence a variety of biological processes, each of which may have a different time course.

We predict that lesion transcriptional mutagenicity and persistence, due to repair deficiency, influence the robustness and longevity of signaling and its biological outcomes. Thus, we tested 5-OHU, which has been shown to be more highly transcriptionally mutagenic than other lesions such as 8-oxoG *in vitro*, for its ability to mediate oncogene activation *in vivo*. We placed 5-OHU in the G12 mutational hotspot of K-Ras, such that if TM occurs *in vivo*, similarly to *in vitro*, it would induce the production of oncogenic $KRAS^{G12D}$ mutant transcripts and proteins, activating pathways downstream of Ras (Figure 2.7A). DNA glycosylases, including Neil1, Neil2, Neil3, and uracil-DNA glycosylase (UNG) have incision activity towards 5-OHU *in vitro* (29-32).

Thus, we employed cells doubly deficient in Neil1 and Neil2, Neil1$^{-/-}$Neil2$^{-/-}$ MEFs (Materials and Methods) in order to determine whether 5-OHU can induce TM and phenotypic change *in vivo*.

We observed increases in both AKT and ERK phosphorylation at 24 hours, demonstrating that TM-mediated signaling can last longer than previously known. These results also indicate that in cells with combined deficiency in Neil1 and Neil2 , 5-OHU persists unrepaired long enough to induce TM and phenotypic change (Figure 2.7B and Figure 2.7C), implicating Neil1 and/or Neil2 in the repair of 5-OHU *in vivo*. Moreover, TM can induce the activation of more than a single Ras effector pathway, which could have important implications *in vivo* due to the aberrant activation of more than a single cellular process. The increases in AKT phosphorylation are statistically significant, and while the increases in ERK phosphorylation are not, as both downstream pathways are dependent on mutant Ras, these phenotypic changes are likely to be biologically significant and may influence a variety of downstream biological consequences.

Discussion

Here, we have presented a systematic analysis of factors influencing protocol reliability and yield of vectors containing site-specific base modifications in any position and sequence of interest. We further identified optimal conditions for reliable large-scale production of ultra-pure vectors highly suitable for applications in mammalian cell culture systems. We employed our improved protocol to study the phenotypic consequences of 5-OHU in cells deficient in both Neil1 and Neil2 DNA glycosylases and found that TM can induce sustained oncogenic signaling and activate more than one

downstream effectors of Ras. It is likely that such sustained TM-mediated oncogene activation of multiple pathways downstream of Ras is sufficient for and can lead to phenotypic consequences beyond biochemical signaling such as induction of proliferation, acquisition of a permanent DNA mutation and oncogenic transformation via retromutagenesis, or increases in DNA damage (7), activation of the DNA damage response and oncogene-induced senescence (33). Deficiency in both Neil1 and Neil2 is sufficient to allow the occurrence of 5-OHU-mediated TM *in vivo*, implicating Neil1 and/or Neil2 in the repair of 5-OHU *in vivo*. However, the relative contributions of Neil1 and Neil2, as well as other enzymes known to have activity towards 5-OHU *in vitro,* in the repair of 5-OHU remain to be determined.

This streamlined protocol should prove useful for the study of the mutagenicity, physiological consequences, and repair of individual lesions in a variety of contexts in basic and translational research, including emerging areas that have not yet been thoroughly investigated. For example, while a great proportion of the physiologically diverse cells in mammals exist in a non-proliferative state, very little is known how DNA damage present in non-dividing cells, such as quiescent stem cells, pre-malignant senescent cells, or terminally differentiated neurons and glia, contributes to the aging process or the development and pathology of cancer, neurodegenerative disease or other illnesses associated with DNA damage. Since reporters of transcriptional mutagenesis do not require DNA replication, they constitute a tool suitable for the study of DNA damage not only in dividing but also non-dividing cells. Due to the method's versatility and ability to position defined lesions in any sequence and reporter of interest, some of its potential applications in translational research include high-content screening of anti-

cancer compounds targeting DNA repair as well as monitoring therapeutic responses in cultured patient tumour samples. While our focus has been mammalian systems, it is important to note that the applications of this technique are not limited to only mammalian cells but also, after an appropriate choice of vector, other systems such as bacteria. Also, while the vectors we have used are non-replicating in the absence of the SV40 large T antigen, in order to avoid the confounding factor of replicative mutagenesis, as opposed to transcriptional mutagenesis, transfection into cell lines that contain it or other episomally replicating vectors could be used to study replicative mutagenesis.

This and other similar methods rely on prior knowledge of the modified base, its stability in experimental procedures, and its successful chemical synthesis into an ODN. A variety of oxidative base lesions (e.g. 8-oxoguanine, 5-hydroxyuracil, dihydrouracil, thymine glycol, spiroiminohydantion), alkylating DNA damage lesions (e.g. O6-methylguanine), those produced by reactive nitrogen species, UV, chemotherapeutic drugs (e.g. cisplatin) or other DNA damaging compounds (e.g. aflatoxin) have already been successfully incorporated into ODNs. Future advances in endogenous DNA damage detection and characterization methods and nucleic acid synthetic chemistry will likely further expand the growing number DNA damage lesions that can be studied using this and similar techniques. A further improvement of the present system would be the development of efficient and reliable technologies for the targeted introduction of DNA damage lesions into genomic DNA.

Acknowledgements

We would like to thank Brian Petrich for sharing the Nucleofector 2B device, and Guy Benian for helpful discussions and encouragement.

Funding

This work was supported by NIH ES011163 and CA 120288 to P.W.D.

Author Contributions

Conceived and designed the experiments: LP PWD. Performed the experiments: LP. Analyzed the data: LP PWD. Contributed reagents/materials/analysis tools: LP CG MB. Wrote the paper: LP PWD.

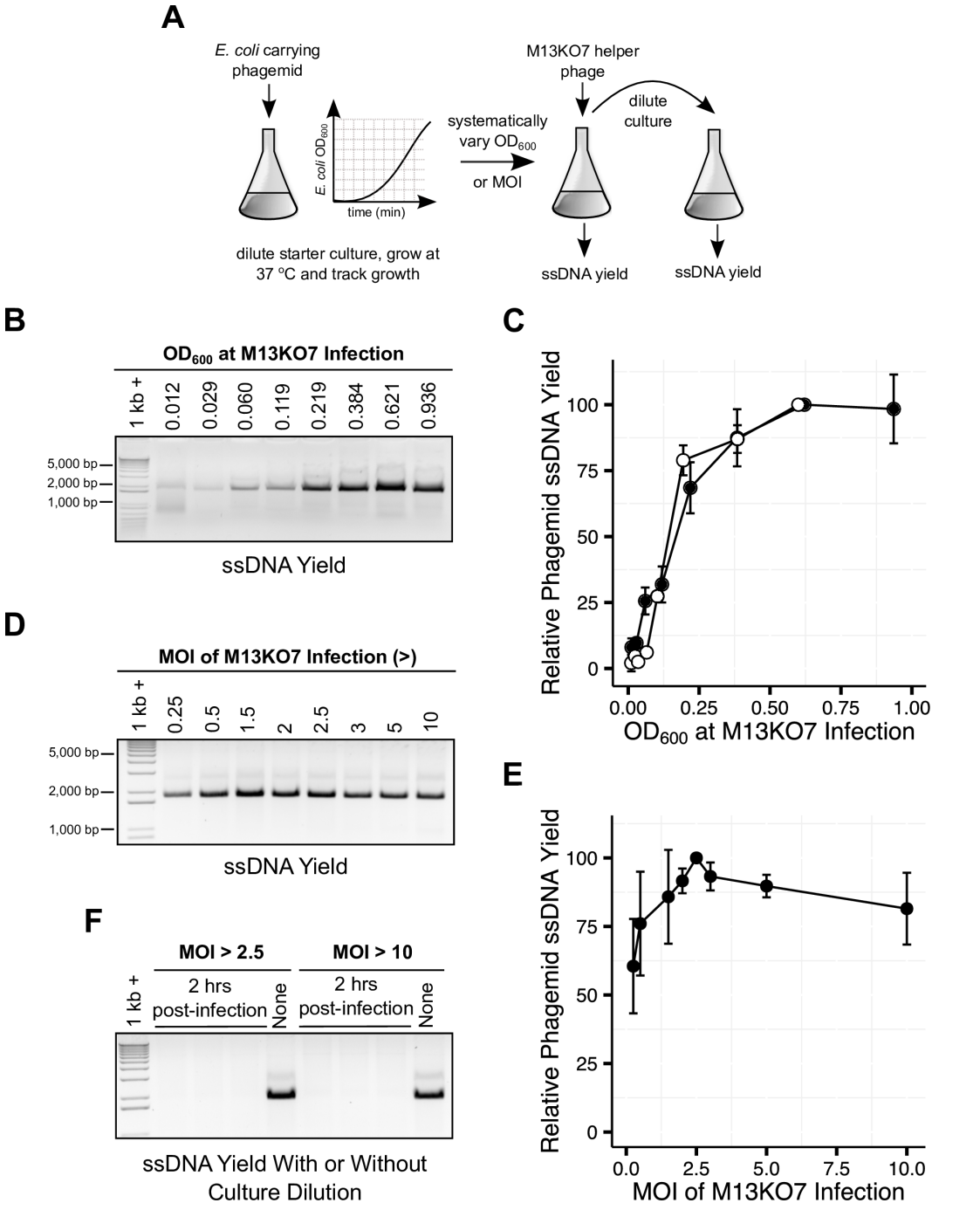


Figure 2.1: Optimization of the DH12S *E. coli* culture and phage infection conditions for phagemid ssDNA yield.

(A) Schematic of the experimental design for the systematic analysis of the effect of culture OD_{600} at phage infection and multiplicity of infection (MOI) on ssDNA yield. Cultures were infected at varying OD_{600} or MOI and ssDNA yields were determined by proteinase K digestion of precipitated phage particles, followed by agarose gel electrophoresis. (B) OD_{600} at phage infection predicts ssDNA yield. (C) Increasing MOI beyond that which is necessary to infect all cells does not improve ssDNA yield. (D) Gel

quantification of two biological replicates, each containing three technical replicates, of cultures infected at varying OD_{600} and identical MOI. (E) Quantification of ssDNA yields from cultures infected at varying MOI. Averages include two biological replicates, each containing three technical replicates. Error bars indicate the standard deviation. (F) Cultures that otherwise produce high yields of phagemid ssDNA, infected at MOI > 2.5 or 10, do not yield ssDNA when diluted 2 hours following phage-infection.

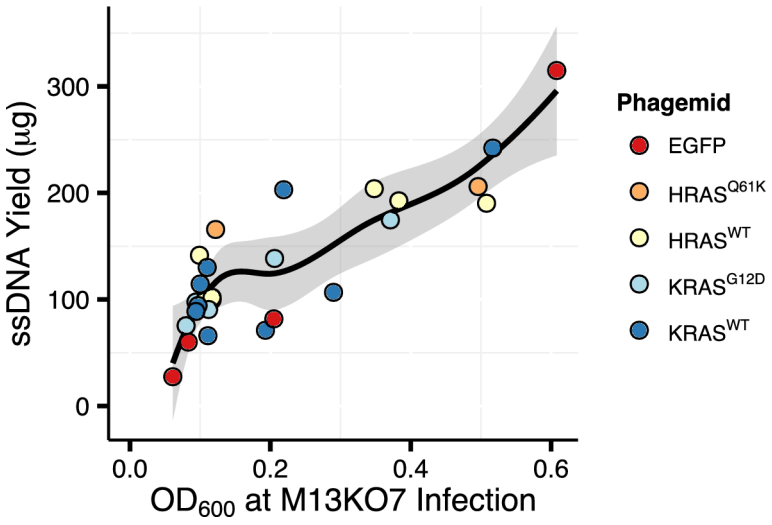


Figure 2.2: Culture density at time of helper phage infection predicts ssDNA yield.

Scatter plot of large-scale (200 mL) ssDNA preparations infected at various MOI greater than one, following anion-exchange column purification. Black line represents the local polynomial regression (loess) curve and grey area the 95% confidence interval, as determined by R software.

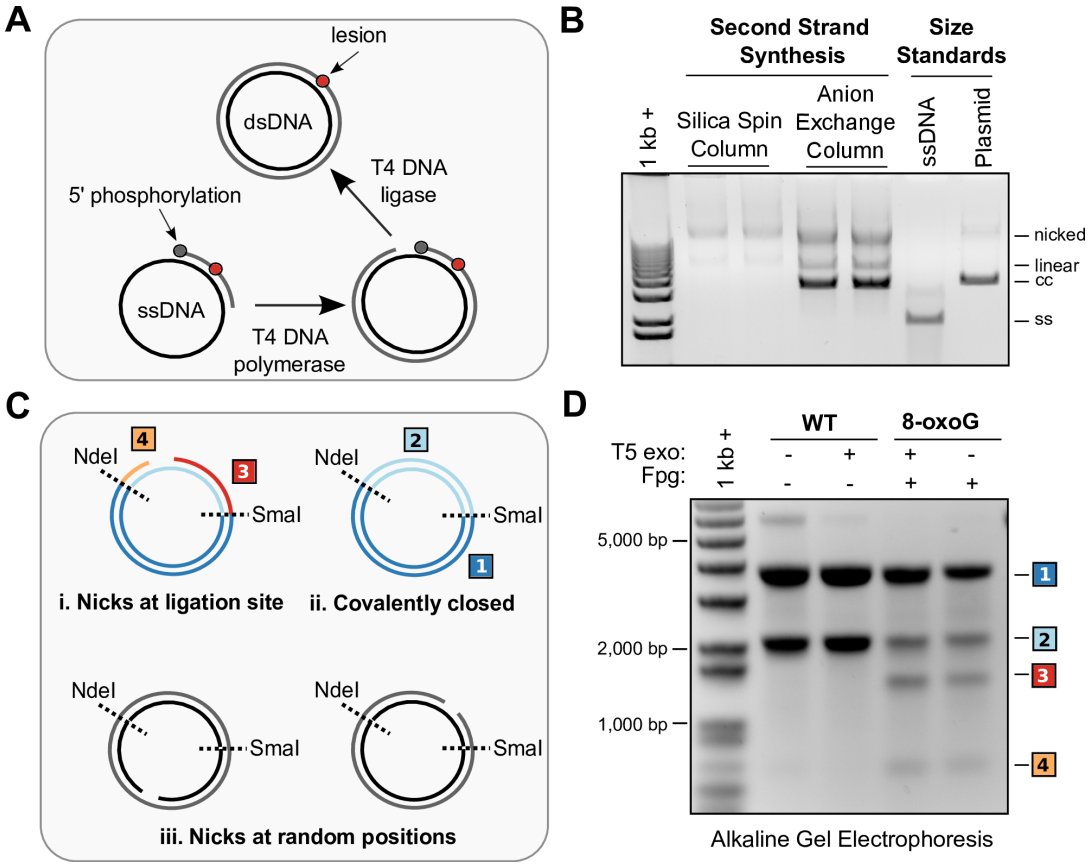


Figure 2.3: Optimizations for second strand synthesis.

(A) Schematic of the second strand synthesis procedure. Synthetic 5' phosphorylated ODNs containing the lesion of interest are annealed to phagemid single-stranded DNA, complimentary strands are synthesised by T4 DNA polymerase, and ligated by T4 DNA ligase. (B) Second strand synthesis of HRAS construct using ssDNA purified by silica spin columns or anion-exchange columns. ssDNA purified by anion-exchange column produces high yields of covalently closed product. (C) Schematic of the alkaline gel analysis of the construct nicks positions. Double-digest of pcDNA3.1(+)-HRAS with SmaI and NdeI produces two fragments (labelled 1 and 2). If the synthetic ODN that becomes part of the transcribed strand is not ligated, the transcribed strand fragment 2 produces two smaller fragments (3 and 4). (D) Alkaline gel analysis of HRAS constructs. Negative control $HRAS^{WT}$ T5 exonuclease (T5 exo) treated, covalently closed construct produces only two bands and positive control Fpg nicked $HRAS^{8\text{-}oxoG}$ constructs, treated and not treated with T5 exonuclease, produce the expected four bands. The anion-exchange purified $HRAS^{WT}$ construct produces only two bands, indicating the nicks following second strand synthesis occur at random positions.

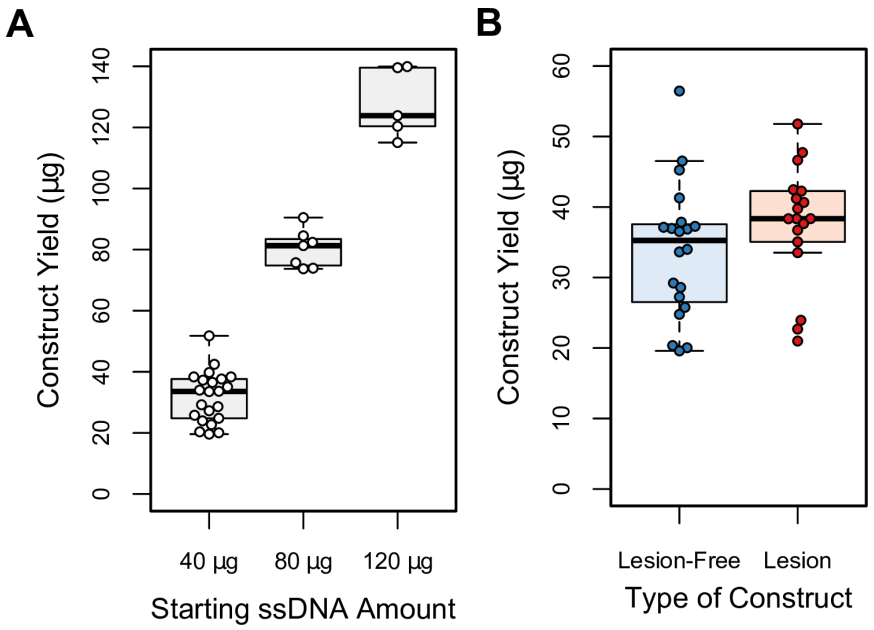


Figure 2.4: Second strand synthesis yields.

(A) Box plots of second strand syntheses yields by starting amount of ssDNA indicate scalable and reliable yields. (B) Box plots comparing yields of lesion-free versus lesion-containing (8-oxoG, 5-OHU, or DHU) constructs, per 40 µg of ssDNA, or as expected per 40 µg of ssDNA if actual starting ssDNA amount was less (20-30 µg). Similar yields are obtained for lesion-free and lesion-containing constructs. Whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles, which are indicated by box limits. Center lines indicate the medians and circles represent the values for individual preparations. Determined by R software.

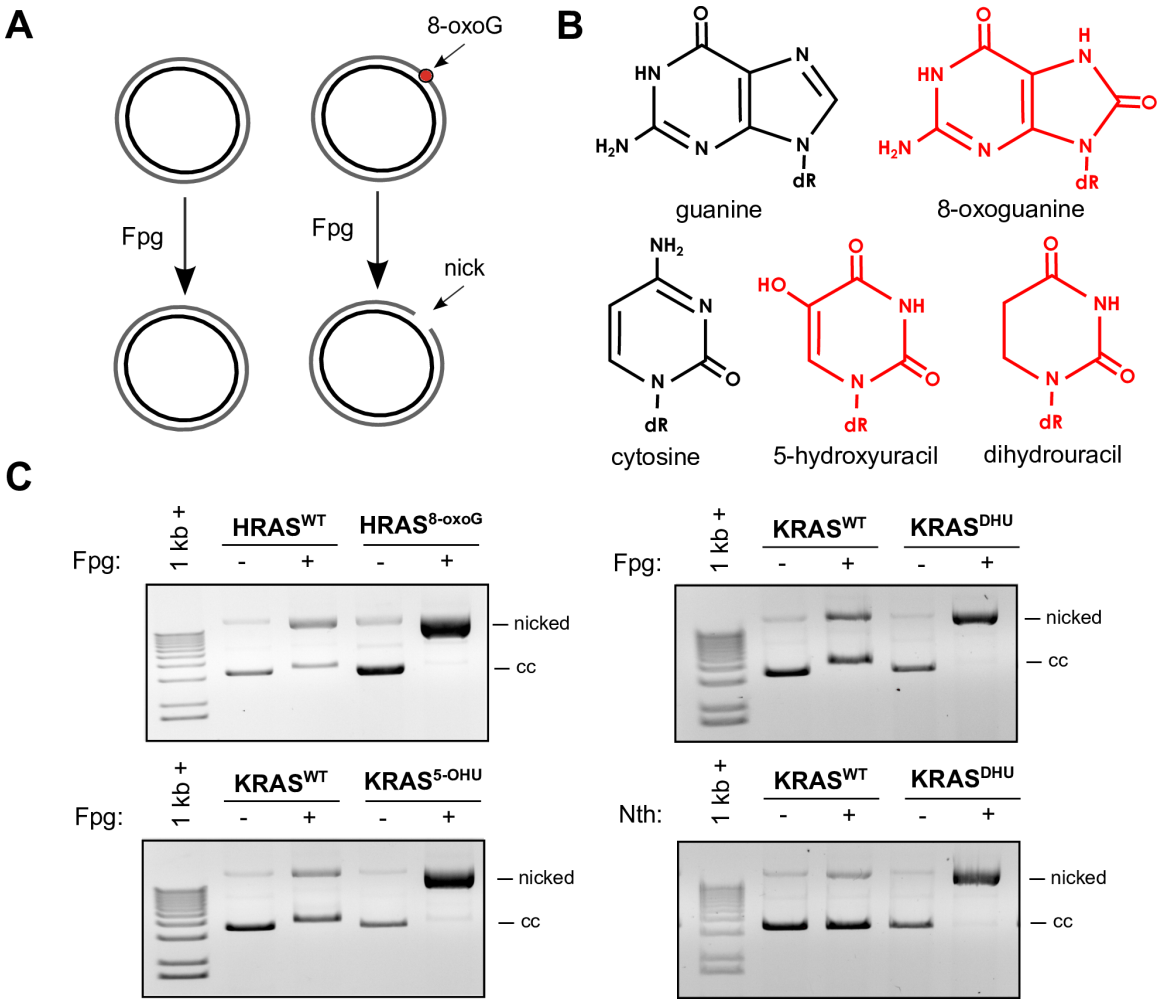


Figure 2.5: Lesion-containing construct quality controls.

(A) Schematic of the Fpg nicking assay. Fpg cleaves damages, such as 8-oxoG and 5-OHU, leaving a single-strand break, converting the construct from covalently closed (cc) to nicked form. (B) Lesion structures. (C) Representative images of Fpg and Nth nicked T5 exonuclease-treated lesion-containing and lesion-free control constructs. Fpg cleaves 8-oxoG, 5-OHU, and DHU, nicking the lesion-containing constructs almost entirely, but not the lesion-free controls, and Nth cleaves DHU.

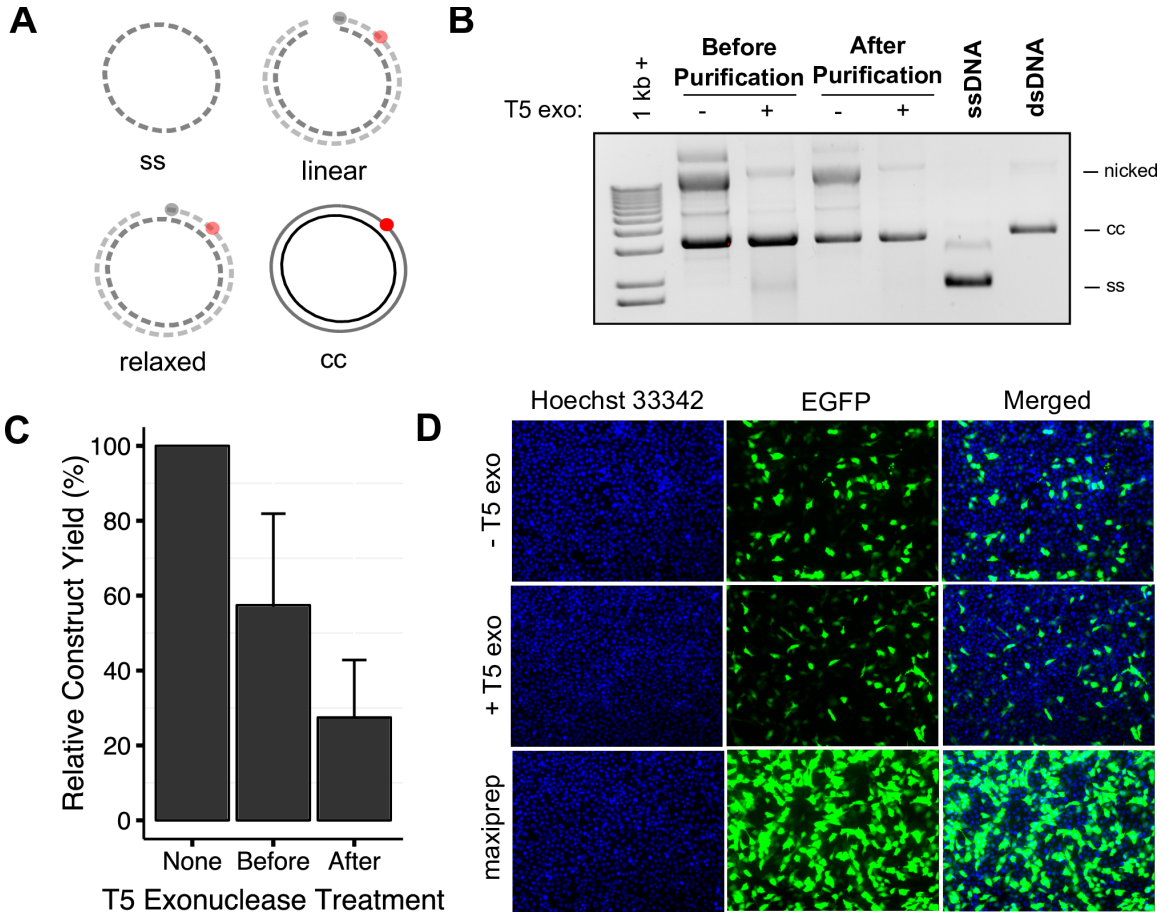


Figure 2.6: Optimization for DNA integrity and mammalian transfection.

(A) Schematic representing T5 exonuclease digestion of nicked, linear, and ssDNA. (B) Representative gel electrophoresis of a construct with and without T5 exonuclease treatment prior to purification and after purification. (C) Construct yields after T5 exonuclease treatment after initial purification (after) or directly in the second strand synthesis reaction (before), relative to non-T5 exonuclease treated construct (none). Error bars represent the standard deviation. (D) Live cell images of Ogg1$^{-/-}$ MEFs nucleofected with EGFP construct treated or not treated with T5 exonuclease or EGFP bacterial plasmid maxiprep, and stained with Hoechst 33342 dye. T5 exonuclease digestion of nicked and linear construct does not improve transfection efficiencies.

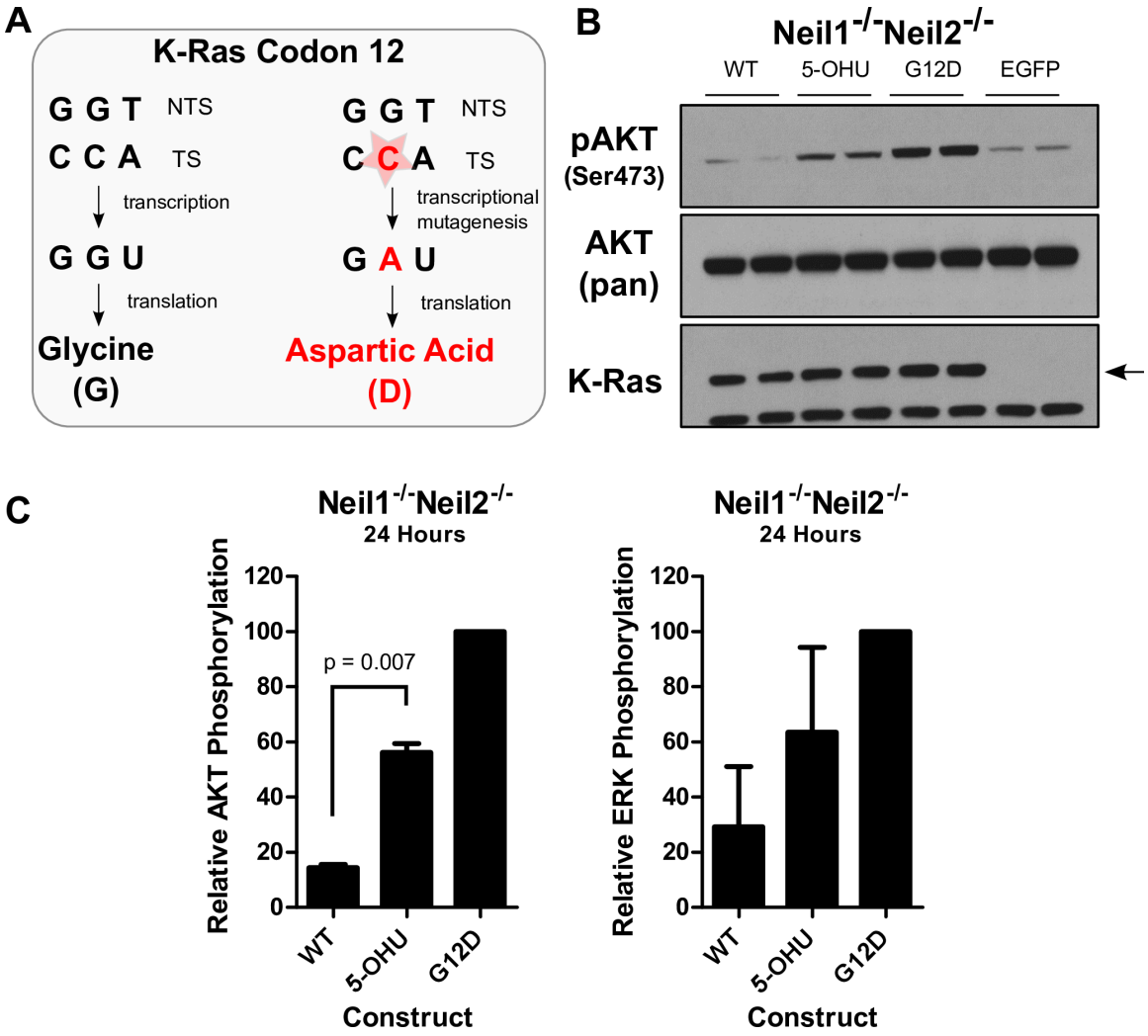


Figure 2.7: Phenotypic consequences of 5-hydroxyuracil in Neil1$^{-/-}$Neil2$^{-/-}$ MEFs.

(A) Schematic of the construct, designed such that mutagenic bypass by RNA polymerase of 5-OHU produces the G12D mutant of K-Ras, leading to the activation of downstream oncogenic signaling. (B) Representative Western blot showing sustained increase in AKT phosphorylation at 24 hours post-nucleofection, with each sample loaded twice serving as a technical replicate. (C) Western blot quantification of two biological replicates, each containing two technical replicates, of pAKT and pERK relative to G12D positive control. The increase in AKT phosphorylation is statistically significant, as determined by a t-test. Error bars represent the standard error of the mean.

References

1. Hanahan, D. and Weinberg, R.A. (2011) Hallmarks of cancer: the next generation. *Cell*, **144**, 646-674.
2. Lu, T., Pan, Y., Kao, S.Y., Li, C., Kohane, I., Chan, J. and Yankner, B.A. (2004) Gene regulation and DNA damage in the ageing human brain. *Nature*, **429**, 883-891.
3. Jackson, S.P. and Bartek, J. (2009) The DNA-damage response in human biology and disease. *Nature*, **461**, 1071-1078.
4. Barnes, D.E. and Lindahl, T. (2004) Repair and genetic consequences of endogenous DNA base damage in mammalian cells. *Annu Rev Genet*, **38**, 445-476.
5. De Bont, R. and van Larebeke, N. (2004) Endogenous DNA damage in humans: a review of quantitative data. *Mutagenesis*, **19**, 169-185.
6. Friedberg, E.C. (2003) DNA damage and repair. *Nature*, **421**, 436-440.
7. Bregeon, D. and Doetsch, P.W. (2011) Transcriptional mutagenesis: causes and involvement in tumour development. *Nat Rev Cancer*, **11**, 218-227.
8. d'Adda di Fagagna, F. (2008) Living on a break: cellular senescence as a DNA-damage response. *Nat Rev Cancer*, **8**, 512-522.
9. Charlet-Berguerand, N., Feuerhahn, S., Kong, S.E., Ziserman, H., Conaway, J.W., Conaway, R. and Egly, J.M. (2006) RNA polymerase II bypass of oxidative DNA damage is regulated by transcription elongation factors. *Embo J*, **25**, 5481-5491.

10. Bregeon, D., Doddridge, Z.A., You, H.J., Weiss, B. and Doetsch, P.W. (2003) Transcriptional mutagenesis induced by uracil and 8-oxoguanine in Escherichia coli. *Mol Cell*, **12**, 959-970.

11. Bregeon, D., Peignon, P.A. and Sarasin, A. (2009) Transcriptional mutagenesis induced by 8-oxoguanine in mammalian cells. *PLoS Genet*, **5**, e1000577.

12. Saxowsky, T.T., Meadows, K.L., Klungland, A. and Doetsch, P.W. (2008) 8-Oxoguanine-mediated transcriptional mutagenesis causes Ras activation in mammalian cells. *Proc Natl Acad Sci U S A*, **105**, 18877-18882.

13. Liu, J. and Doetsch, P.W. (1998) Escherichia coli RNA and DNA polymerase bypass of dihydrouracil: mutagenic potential via transcription and replication. *Nucleic Acids Res*, **26**, 1707-1712.

14. Luhnsdorf, B., Kitsera, N., Warken, D., Lingg, T., Epe, B. and Khobta, A. (2012) Generation of reporter plasmids containing defined base modifications in the DNA strand of choice. *Anal Biochem*, **425**, 47-53.

15. Shivji, M.K., Moggs, J.G., Kuraoka, I. and Wood, R.D. (1999) Dual-incision assays for nucleotide excision repair using DNA with a lesion at a specific site. *Methods Mol Biol*, **113**, 373-392.

16. Bregeon, D. and Doetsch, P.W. (2004) Reliable method for generating double-stranded DNA vectors containing site-specific base modifications. *Biotechniques*, **37**, 760-762, 764, 766.

17. Bregeon, D. and Doetsch, P.W. (2006) Assays for transcriptional mutagenesis in active genes. *Methods Enzymol*, **409**, 345-357.

18. Nagel, Z.D., Margulies, C.M., Chaim, I.A., McRee, S.K., Mazzucato, P., Ahmad, A., Abo, R.P., Butty, V.L., Forget, A.L. and Samson, L.D. (2014) Multiplexed DNA repair assays for multiple lesions and multiple doses via transcription inhibition and transcriptional mutagenesis. *Proc Natl Acad Sci U S A*, **111**, E1823-1832.

19. Sambrook, J. and Russell, D.W. (2006) Alkaline agarose gel electrophoresis. *CSH Protoc*, **2006**.

20. Todaro, G.J. and Green, H. (1963) Quantitative studies of the growth of mouse embryo cells in culture and their development into established lines. *J Cell Biol*, **17**, 299-313.

21. Vieira, J. and Messing, J. (1987) Production of single-stranded plasmid DNA. *Methods Enzymol*, **153**, 3-11.

22. Zhou, B., Dong, Q., Ma, R., Chen, Y., Yang, J., Sun, L.Z. and Huang, C. (2009) Rapid isolation of highly pure single-stranded DNA from phagemids. *Anal Biochem*, **389**, 177-179.

23. Hatahet, Z., Kow, Y.W., Purmal, A.A., Cunningham, R.P. and Wallace, S.S. (1994) New substrates for old enzymes. 5-Hydroxy-2'-deoxycytidine and 5-hydroxy-2'-deoxyuridine are substrates for Escherichia coli endonuclease III and formamidopyrimidine DNA N-glycosylase, while 5-hydroxy-2'-deoxyuridine is a substrate for uracil DNA N-glycosylase. *J Biol Chem*, **269**, 18814-18820.

24. Venkhataraman, R., Donald, C.D., Roy, R., You, H.J., Doetsch, P.W. and Kow, Y.W. (2001) Enzymatic processing of DNA containing tandem dihydrouracil by endonucleases III and VIII. *Nucleic Acids Res*, **29**, 407-414.

25. Fromme, J.C. and Verdine, G.L. (2003) DNA lesion recognition by the bacterial repair enzyme MutM. *J Biol Chem*, **278**, 51543-51548.

26. Kuznetsov, N.A., Koval, V.V., Zharkov, D.O., Vorobjev, Y.N., Nevinsky, G.A., Douglas, K.T. and Fedorova, O.S. (2007) Pre-steady-state kinetic study of substrate specificity of Escherichia coli formamidopyrimidine--DNA glycosylase. *Biochemistry*, **46**, 424-435.

27. Kiss-Toth, E., Dower, S.K. and Sayers, J.R. (2001) A method for enhancing the transfection efficiency of minipreps obtained from plasmid cDNA libraries. *Anal Biochem*, **288**, 230-232.

28. Schubbert, S., Shannon, K. and Bollag, G. (2007) Hyperactive Ras in developmental disorders and cancer. *Nat Rev Cancer*, **7**, 295-308.

29. Dizdaroglu, M., Karakaya, A., Jaruga, P., Slupphaug, G. and Krokan, H.E. (1996) Novel activities of human uracil DNA N-glycosylase for cytosine-derived products of oxidative DNA damage. *Nucleic Acids Res*, **24**, 418-422.

30. Dou, H., Mitra, S. and Hazra, T.K. (2003) Repair of oxidized bases in DNA bubble structures by human DNA glycosylases NEIL1 and NEIL2. *J Biol Chem*, **278**, 49679-49684.

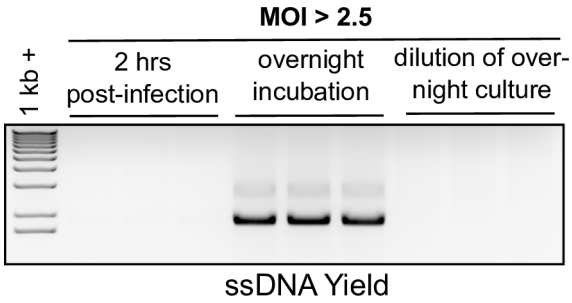
31. Hegde, M.L., Hegde, P.M., Bellot, L.J., Mandal, S.M., Hazra, T.K., Li, G.M., Boldogh, I., Tomkinson, A.E. and Mitra, S. (2013) Prereplicative repair of oxidized bases in the human genome is mediated by NEIL1 DNA glycosylase together with replication proteins. *Proc Natl Acad Sci U S A*, **110**, E3090-3099.

32. Liu, M., Bandaru, V., Bond, J.P., Jaruga, P., Zhao, X., Christov, P.P., Burrows, C.J., Rizzo, C.J., Dizdaroglu, M. and Wallace, S.S. (2010) The mouse ortholog of

NEIL3 is a functional DNA glycosylase in vitro and in vivo. *Proc Natl Acad Sci U S A*, **107**, 4925-4930.

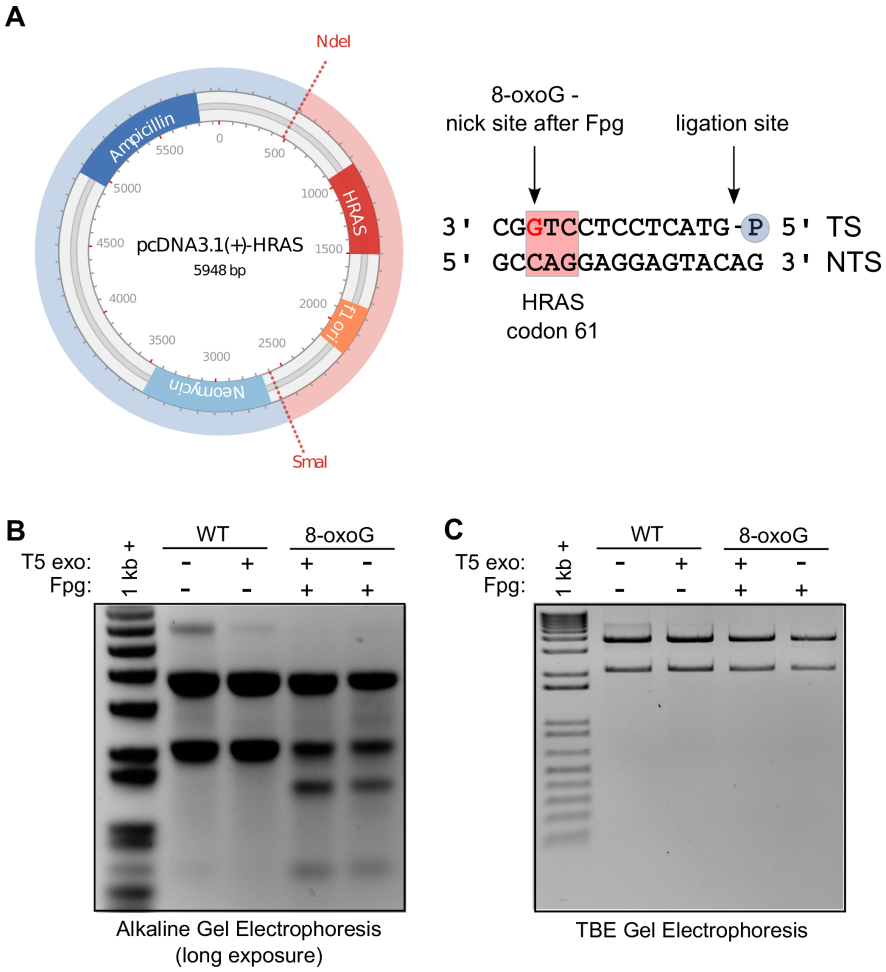
33. Bartkova, J., Rezaei, N., Liontos, M., Karakaidos, P., Kletsas, D., Issaeva, N., Vassiliou, L.V., Kolettas, E., Niforou, K., Zoumpourlis, V.C. *et al*. (2006) Oncogene-induced senescence is part of the tumorigenesis barrier imposed by DNA damage checkpoints. *Nature*, **444**, 633-637.

Supplementary Information



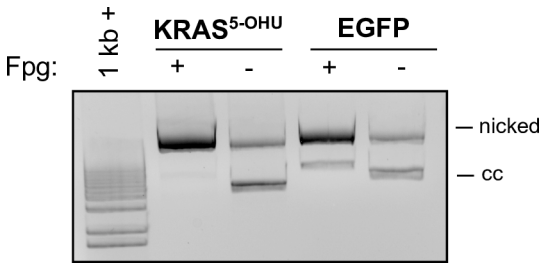
Supplementary Figure 2.1: ssDNA yields of diluted cultures.

ssDNA yields determined by proteinase K digestion of precipitated phage, followed by gel electrophoresis, from infected *E. Coli* cells 2 hours post-infection, after overnight incubation, and after dilution of the first overnight culture and a second overnight incubation. High yields of ssDNA are only present in undiluted cultures after an overnight incubation.



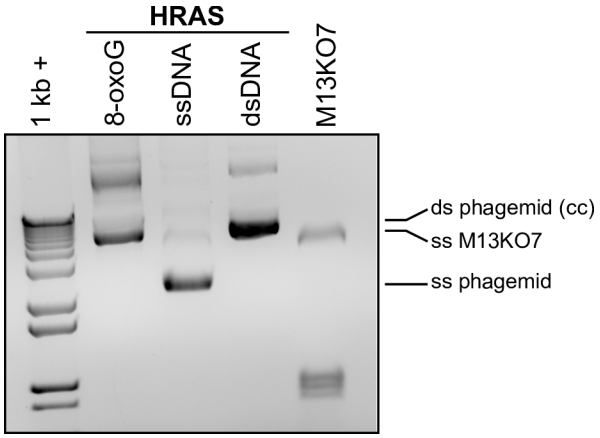
Supplementary Figure 2.2: Determination of the second strand synthesis product nicks positions.

(A) pcDNA3.1(+)-HRAS plasmid map generated using Angular Plasmid (http://angularplasmid.vixis.com/) and sequence surrounding the 8-oxoG lesion and ligation site for second strand synthesis. (B) Overexposure of the alkaline gel electrophoresis. (C) The same samples separated on non-denaturing agarose gel in TBE buffer.



Supplementary Figure 2.3: Gel purification of constructs.

Low melting point agarose (LMP) and β-agarase purification of constructs. Covalently closed forms of $KRAS^{5\text{-}OHU}$ and EGFP maxiprep were purified from SeaPlaque GTG LMP agarose using β-agarase (Lonza) as per the manufacturer's instruction and digested with Fpg as described in Materials and Methods. LMP purification can result in nicking and high levels of oxidation.



Supplementary Figure 2.4: M13KO7 preparation.

DH12S *E. coli* not containing phagemid were infected with M13KO7 phage as per the same protocol for phagemid production and ssDNA was purified using PCIA extraction. M13KO7 $HRAS^{8\text{-}oxoG}$ second strand synthesis reaction, $HRAS^{WT}$ ssDNA, $HRAS^{WT}$ plasmid maxiprep and M13KO7 ssDNA preparation were resolved on an agarose gel to compare sizes. The faint upper band in the ssDNA preparation has the same migration pattern as M13KO7 ssDNA. While we do not observe significant M13KO7 ssDNA contamination in purified constructs not treated with T5 exonuclease (Figure 2.6B), treatment with T5 exonuclease can be employed if minimizing ssDNA contamination is preferred.

ODN	Sequence
$HRAS^{WT}$	5′ P-GTACTCCTCCTGGCCGGCGGTATCCAGGATGT 3′
$HRAS^{8\text{-}oxoG}$	5′ P-GTACTCCTCCT(8-oxoG)GCCGGCGGTATCCAGGATGT 3′
$HRAS^{Q61K}$	5′ P-GTACTCCTCCTTGCCGGCGGTATCCAGGATGT 3′
$KRAS^{WT}$	5′ P-CTCTTGCCTACGCCACCAGCTCCAACTACC 3′
$KRAS^{5\text{-}OHU}$	5′ P-CTCTTGCCTACGCCA(5-OHU)CAGCTCCAACTACC 3′
$KRAS^{DHU}$	5′ P-CTCTTGCCTACGCCA(DHU)CAGCTCCAACTACC 3′
$KRAS^{G12D}$	5′ P-CTCTTGCCTACGCCATCAGCTCCAACTACC 3′
EGFP	5′ P-GTCGTCCTTGAAGAAGATGGTGCGCTC 3′

Supplementary Table 1: ODN sequences.

Sequences of oligodeoxynucleotides containing 5' phosphorylation (P), 8-oxoguanine (8-oxoG), 5-hydroxyuracil (5-OHU), or dihydrouracil (DHU), used for second strand synthesis.

Buffer/Medium	Composition
Prepared Buffers/Media	
LB-Miller	10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl
2X-YT	15 g/L tryptone, 10 g/L yeast extract, 5 g/L NaCl
20X SSC	3 M NaCl, 300 mM sodium citrate, pH 7.0
10X TE	200 mM Tris, 10 mM EDTA, pH 8.0
10X TBE	1 M Tris, 1 M boric acid, 0.02 M EDTA
50X TAE	2M Tris, 50 mM EDTA, pH 8.3
TBST	20 mM Tris, 150 mM NaCl, 0.1% Tween-20, pH 7.5
Commercial Buffers	
N3	3.1 M potassium acetate, pH 5.5
EQ1	0.1 M sodium acetate, 0.6 M NaCl, 0.15% (v/v) Triton X-100, pH 5.0
W8	0.1 M sodium acetate, 825 mM NaCl, pH 5.0
E4	100 mM Tris-HCl, 1.25 M NaCl, pH 8.5
TE	10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0
NEBuffer 1	10 mM Bis Tris propane-HCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.0
NEBuffer 2	10 mM Tris-HCl, 10 mM $MgCl_2$, 50 mM NaCl, 1 mM DTT, pH 7.9
1X Endonuclease III	20 mM Tris-HCl, 1 mM EDTA, 1 mM DTT, pH 8
CutSmart	50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 100 μg/mL BSA, pH 7.9

Supplementary Table 2: Supplementary list of buffers and media.

Composition of commercial buffers used and buffers and media prepared.

3. NEIL2-MEDIATED REPAIR OF 5-HYDROXYURACIL AND DIHYDROURACIL FROM TRANSCRIBED DNA PROTECTS MAMMALIAN CELLS FROM SUSTAINED TRANSCRIPTIONAL MUTAGENESIS AND ITS PHENOTYPIC CONSEQUENCES

Abstract

Oxidative DNA damage can compromise genomic integrity, and is associated with the development of diseases, such as cancer. Transcription is critical for cellular function, and DNA damage can also compromise the fidelity of transcription. However, our knowledge of the *in vivo* transcriptional mutagenesis (TM) mediated by oxidative DNA damage is very limited. Using a well-defined system employing mammalian expression constructs containing the common oxidative lesion dihydrouracil (DHU) in a mutational hotspot of the transcribed strand of K-Ras, we show that DHU is transcriptionally mutagenic *in vivo* and can cause aberrant, sustained activation of the oncogenic Raf-MEK-ERK and PI3K-AKT-mTOR pathways, similar to what we previously described for 5-hydroxyuracil (5-OHU). Several DNA glycosylases with overlapping specificity display activity towards DHU and 5-OHU *in vitro*. However, the relative contribution of each glycosylase towards the repair 5-OHU and DHU *in vivo*, and its influence on the occurrence, timecourse and phenotypic consequences of TM remain unclear. Our results implicate Neil2 as the primary glycosylase repairing 5-OHU and DHU *in vivo* from transcribed DNA. Neil2 deficiency promotes sustained transcriptional mutagenesis and its phenotypic consequences. Deficiencies in Neil2-mediated base excision repair from transcribed regions may have important implications for the occurrence of transcriptional mutagenesis and its downstream biological consequences.

Introduction

Cells acquire thousands of diverse potentially mutagenic oxidative base lesions per day due to exposure to reactive oxygen species (ROS) generated endogenously during normal cellular metabolism, inflammation and as a defense against pathogenic microorganisms, or due to exposure to exogenous DNA damaging agents, such as ionizing or ultraviolet radiation (1-3). The ability of DNA damage to compromise the fidelity of DNA replication and lead to heritable mutations, and its association with the development of cancer (4,5), neurodegenerative disease (6), and premature aging (7), have been well established. However, a large variety of cell types in the human body exist in largely quiescent state, rarely undergoing DNA replication yet remaining transcriptionally active. DNA transcription is critical for cellular function and survival throughout all stages of the cell cycle, and the fidelity of transcription can also be compromised by DNA damage due to misincorporation of incorrect nucleotides opposite DNA damage lesions by RNA polymerase (RNAP), equivalent to those misincorporated by DNA polymerase during replication (8-10). This transcriptional mutagenesis (TM) may also have broad implications to human health as it could occur and potentially lead to deleterious phenotypic consequences in every single cell of the human body, irrespective of proliferation status. However, the relative contributions of TM in disease development are unknown. Very few studies have investigated the *in vivo* transcriptional mutagenesis and transcriptional mutagenesis-mediated phenotypic consequences of common, endogenously occurring oxidative base lesions and how defects in repair associated with disease development may impact the persistence of DNA damage that can compromise the fidelity of transcription. The transcriptionally mutagenic potential of

variety of oxidative lesions, such as dihydrouracil (DHU), and their capacity to induce TM-mediated phenotypic change in intact mammalian cells has not yet been established.

Base excision repair (BER) is a highly evolutionarily conserved pathway responsible for the repair of non-bulky, non-helix-distorting lesions, past which efficient translesion synthesis by DNA and RNA polymerases can occur, and it is initiated by distinct DNA glycosylases with overlapping substrate specificities. Currently, eleven DNA glycosylases have been identified in mammalian cells, five of which are responsible for the repair of oxidative base damage (11). 8-oxoguanine DNA glycosylase (Ogg1) is responsible for the repair of the common replicatively and transcriptionally mutagenic lesion 8-oxoguanine (8-oxoG) *in vivo* (12), and has activity towards other purine-derived lesions, such as 2,6-diamino-4-oxo-formamidopyrimidine (FapyG) (13). The Nei endonuclease VIII-like family of glycosylases, including Neil1, Neil2, and Neil3 as well as Nth endonuclease III-like 1 (Nthl1) have enzymatic activity primarily towards pyrimidine-derived oxidative lesions, such as 5-hydroxyuracil (5-OHU) or DHU.

Neil1 (14-18), Neil2 (14,15,19), and Neil3 (20,21) all have incision activity towards 5-OHU and Neil1 (16,18,22), Neil2 (19), Nthl1 (22,23) have incision activity towards DHU *in vitro* and/or in cellular extracts. However, the relative contribution of each glycosylase towards the repair 5-OHU and DHU *in vivo*, and its influence on the occurrence, timecourse and phenotypic consequences of TM mediated by 5-OHU and DHU remain unclear. Deficiencies in repair promote the persistence of DNA damage, and thus deficiencies in repair components responsible for the repair of DNA damage at transcribed regions would promote transcriptional mutagenesis and its phenotypic consequences. The overlapping specificity of oxidative damage DNA glycosylases may

not entirely serve as a mechanism of redundancy in DNA repair, and emerging evidence suggests that DNA glycosylases may have more specialized functions modulated by additional factors, such as cell cycle phase or the transcription status of DNA containing substrate lesions. For example, Neil2$^{-/-}$ mice accumulate oxidative damage in transcribed regions (24), and human NEIL2 associates with human RNAPII and heterogeneous nuclear ribonucleoprotein-U in cells and repairs 5-OHU preferentially from the transcribed strand *in vitro* (25), while NEIL1 mediates prereplicative repair of oxidized bases occurring at replication forks (14), and its expression is transcriptionally regulated by cell cycle checkpoint protein Rad9 in human cells and post-translationally in mouse embryonic stem cells (26).

8-oxoguanine-mediated TM can induce activation of the Ras effector ERK six hours post nucleofection (12), and in Chapter 2 we showed that 5-OHU is transcriptionally mutagenic *in vivo* and 5-OHU-mediated TM can induce sustained activation of more than one Ras effector pathways that regulate a variety of cancer hallmarks such as proliferation, survival, and cell migration. The *in vivo* timecourse of DNA repair of transcriptionally mutagenic lesions may have important implications for the phenotypic consequences of TM. The timecourse of physiological processes, such as cellular proliferation, requires sustained signaling. The strength of TM-mediated oncogenic signaling, determined by the *in vivo* RNAP bypass of a lesion and its transcriptional mutagenicity, likely also influences the occurrence of downstream biological processes. In this study, we sought to investigate the *in vivo* transcriptional mutagenesis of DHU and identify the DNA glycosylases primarily responsible for the repair of 5-OHU and DHU *in vivo* using a well-defined system employing mammalian expression vectors containing

lesions site-specifically placed in the G12 mutational hotspot of the proto-oncogene K-Ras. Results indicate that DHU is transcriptionally mutagenic *in vivo*, and deficiency in Neil2 promotes the persistence of 5-OHU and DHU in non-replicating, transcribed DNA, and transcriptional mutagenesis and its phenotypic consequences in mammalian cells. While replication of the experiments and determining repair of 5-OHU and DHU placed on the non-transcribed strand will be necessary to further support these observations, our results suggest that Neil2 is the primary glycosylase repairing 5-OHU and DHU *in vivo* from transcribed DNA.

Materials and Methods

Preparation of Site-Specific DNA Damage-Containing Constructs. $HRAS^{WT}$, $HRAS^{8\text{-}oxoG}$, $HRAS^{Q61K}$, $KRAS^{WT}$, $KRAS^{5\text{-}OHU}$, $KRAS^{DHU}$, $KRAS^{G12D}$, and EGFP constructs were prepared as described in Chapter 2. Second strand synthesis reactions were purified using anion-exchange columns without T5 exonuclease treatment.

Cell Culture and Mammalian Transfection. Primary $Neil2^{-/-}$ and $Neil1^{-/-}Neil2^{-/-}$ mouse embryonic fibroblasts (MEFs) were provided by Christine Gran and Magnar Bjoras (Gran *et al*, manuscript in preparation). Experimental procedures were approved by the Norwegian Animal Research Authority. $Neil2^{-/-}$ cells were prepared using the same protocols as for $Neil1^{-/-}Neil2^{-/-}$, as described in Chapter 2, and immortalized using the 3T3 protocol, as described previously (27). E1A immortalized $Neil1^{-/-}$ MEFs were a gift from Stephen Lloyd, and $Ogg1^{-/-}$ MEFs immortalized by frequent passaging have been described previously (12). All MEFs were cultured in a humidified incubator at 10% CO_2 in high-glucose DMEM containing GlutaMAX™ (Invitrogen) supplemented with 10%

fetal bovine serum (GE Healthcare). MEFs were electroporated as per the manufacturer's instructions using an Amaxa Nucleofector 2B device and MEF 1 Nucleofector® kit (Lonza, Cat.# VPD-1004) or Ingenio® Electroporation Kit (Mirus Bio, Cat.# MIR 50112) using the T-020 setting and 3-4 μg of DNA per 1.0 x 10^6 to 1.5 x 10^6 cells. TransIT-2020 (Mirus Bio, Cat. #MIR-5404) transfections were carried out as per the manufacturer's instructions in 6-well plates, using 4 μg of DNA per sample.

Western Blot Analysis. Cells were washed with PBS, and switched to serum-free DMEM 1.5 hours before lysis. Cells were lysed as described in Chapter 2, and Western blots were done also as described in Chapter 2. Antibodies against phospho-AKT (Ser473, Cat. #4060), AKT (pan, Cat. #4691), phospho-ERK1/2 (Thr202/Tyr204, Cat. #9106) and ERK (Cat. #9102) were from Cell Signaling Technology, K-Ras was from Santa Cruz Biotechnology (Cat. #sc-30), and H-Ras antibodies were from Abcam (Cat.# ab97488) or Santa Cruz Biotechnology (Cat. #sc-520). Film was scanned, images quantified using ImageQuant TL, the ratios of (pAKT/AKT)/K-Ras, (pERK/ERK)/K-Ras or (pERK/ERK)/H-Ras were determined for each sample and expressed as percent of the mutant positive control which was set to 100%. Experiments with low transfection efficiencies for which pERK or pAKT did not reach above background levels were excluded from analysis.

Results

Development of Systems for the Study of Transcriptional Mutagenesis-Mediated Oncogene Activation

The $KRAS^{G12D}$ mutation is one of the most frequently occurring mutations in human cancers, including pancreatic (28), lung (29), colorectal (30), as well as several others. It leads to constitutive activation of KRAS and persistent stimulation of downstream pathways controlling a variety of cellular processes and cancer hallmarks, including proliferation, apoptosis, and cell migration, and drives tumorigenesis in mouse models of several different types of cancers (31). RNA polymerase misincorporates adenine opposite 5-OHU and DHU *in vitro* (32-34). Thus, we employed our recently developed efficient and reliable method for the large-scale production of mammalian expression vectors containing site-specific base modifications, described in Chapter 2, to generate KRAS constructs containing 5-OHU or DHU in the transcribed strand of codon 12 of KRAS instead of unmodified cytosine, such that when TM occurs it would produce the constitutively active G12D mutant (Figure 3.1). We also generated constructs containing cytosine or thymidine, encoding glycine (G) or aspartic acid (D), respectively, as negative and positive controls. Employing the same backbone as for all other constructs, pcDNA3.1(+), we also generated constructs encoding EGFP in order to monitor nucleofection efficiencies. We then employed this set of constructs in immortalized mouse embryonic fibroblasts (MEFs) deficient in Neil1, Neil2, or both.

Dihydrouracil Causes Transcriptional Mutagenesis *in vivo*, Induces Oncogene Activation, and is Repaired by Neil2

Bypass of DHU by *E. coli*, SP6 and T7 RNA polymerases *in vitro* is highly mutagenic and results in misincorporation of adenine opposite the lesion (33,34). Moreover, translesion synthesis by DNA polymerase also results in adenine misincorporation (33). However, *in vitro* results may not entirely recapitulate *in vivo* events, even when employing RNA polymerase from the same organism. For example,

single-strand breaks stall *E. coli* RNA polymerase *in vitro* (35), but are bypassed, causing TM and luciferase reactivation in intact *E. coli* cells (36). Likewise, 5-OHU, thymine glycol, and 8-oxoG stall human RNA polymerase *in vitro* to various extents, but transcription elongation factors present in HeLa nuclear extracts allow RNA polymerase to mutagenically bypass them (32). Thus, we sought to investigate the *in vivo* transcriptional mutagenesis and phenotypic consequences mediated by DHU in intact mammalian cells.

We employed DHU-containing constructs and controls in Neil2$^{-/-}$ and Neil1$^{-/-}$ Neil2$^{-/-}$ MEFs. We observe a substantial increase in ERK phosphorylation in Neil2$^{-/-}$ cells, as well as increases in ERK and AKT phosphorylation in Neil1$^{-/-}$Neil2$^{-/-}$ cells 24 hours post-nucleofection, indicating that DHU is transcriptionally mutagenic *in vivo* and DHU-mediated TM can induce oncogene activation (Figure 3.2). Neil2 deficiency is sufficient to promote the persistence of the lesion, TM and phenotypic change observable 24 hours post-nucleofection, in contrast to the rapid timecourse of base excision repair in wild type cells. For example, unlike in Ogg1$^{-/-}$ MEFs, 8-oxoG does not induce TM-mediated phenotypic change in repair proficient cells 6 hours post-nucleofection and it is almost entirely repaired by 24 hours (12). Thus, our results indicate that Neil2 appears to play a major role in the repair of DHU from transcribed DNA *in vivo*. Repeating these experiments in Neil2$^{-/-}$ and Neil1$^{-/-}$Neil2$^{-/-}$ cells, as well as in wild type and Neil1$^{-/-}$ MEFs would help further support these observations and help determine whether Neil2 may be the primary glycosylase repairing DHU from transcribed DNA *in vivo*. In order to establish whether Nthl1 has slow back-up incision activity towards DHU, cells deficient

in Nthl1 can be employed or TM during longer time points, such as 48 and 72 hours, in Neil2$^{-/-}$ cells can be determined.

Neil2 Appears to be the Main DNA Glycosylase Repairing 5-Hydroxyuracil from Transcribed, Non-Replicating DNA *in vivo*

We showed in Chapter 2 that 5-OHU is transcriptionally mutagenic *in vivo* and combined deficiency of Neil1 and Neil2 promotes the persistence of the lesion and TM-mediated phenotypic change. In order to determine the relative contribution of Neil1 and Neil2 towards the repair of 5-OHU occurring in transcribed DNA *in vivo*, we determined the TM-mediated phenotypic consequences in the context of Ras activation in Neil1$^{-/-}$, Neil2$^{-/-}$, and Neil1$^{-/-}$Neil2$^{-/-}$ MEFs. While we observe an increase in ERK phosphorylation at 6 hours post-nucleofection in Neil1$^{-/-}$ MEFs (Figure 3.4), no increase in ERK phosphorylation is present in Neil1$^{-/-}$ MEFs at 24 hours, indicating subsequent complete repair of 5-OHU in the absence of Neil1 (Figure 3.3). We also find evidence of complete repair of 5-OHU 24 hours post-transfection in Ogg1$^{-/-}$ MEFs, in which we find evidence of sustained 8-oxoG-mediated H-Ras activation at timepoints longer than previously known (Supplementary Figure 3.2). In contrast to Neil1$^{-/-}$ MEFs, we observe substantial increases in AKT and ERK phosphorylation in cells deficient in Neil2 only at 24 hours, similar to the results observed in Neil1$^{-/-}$Neil2$^{-/-}$ MEFs. In the absence of Neil2, 5-OHU-mediated TM can still induce sustained activation of more than one oncogenic Ras effector pathways. Thus, these results indicate that Neil2 is the primary glycosylase repairing 5-OHU *in vivo* from transcribed DNA.

Discussion

Since only Neil2 appears to have strong repair activity towards 5-OHU and DHU placed in a transcribed gene in non-replicating constructs *in vivo*, these data indicate that oxidative DNA damage glycosylases may not be generally redundant where they overlap in enzymatic specificity, but factors such as the transcription status of affected DNA may influence *in vivo* glycosylase activity. Whether a cell is proficient in the repair of lesions, such as 5-OHU and DHU, from transcribed or replicating DNA should have important implications for the occurrence and biological consequences of transcriptional or replicative mutagenesis, and thus transient/non-heritable versus permanent/heritable phenotype, respectively. Deficiencies in repair that promote the persistence of DNA damage in transcribed DNA promote TM and its phenotypic consequences. If TM initiates an irreversible phenotype, such as neuronal cell death, defects in repair that promote TM could be sufficient to promote deleterious biological consequences. However, if TM initiates a reversible phenotype, permanent establishment of the phenotype would require permanent establishment of the original transcriptional mutation via replicative mutagenesis. Since the bases misincorporated by RNA polymerase opposite the majority of DNA lesions studied so far are equivalent to those misincorporated by DNA polymerase, if TM induces a phenotype, such as aberrant cell cycle entry, the original transcriptional mutation could become a permanent DNA mutation via replicative mutagenesis. This process has been termed retromutagenesis (37), and evidence implicating retromutagenesis as a mechanism for adaptive mutation has been observed in *E. coli* (38). Glycosylases such as Neil1 could be important in the repair of replicating DNA and they could have important implications for the concept of

retromutagenesis in mammalian cells. Moreover, in cells deficient only in Neil2, higher levels of replicative mutagenesis could be anticipated due to the accumulation of excessive DNA damage in transcribed regions before cell cycle entry and thus more lesions that may evade repair prior to DNA replication. Functionally impaired polymorphic NEIL2 variants are associated with an increased risk of lung cancer (39), squamous cell carcinoma of the oral cavity and oropharynx (40), and genomic analyses indicate that homozygous NEIL2 deletions can occur in several different types of cancers (41), including 10.5% of prostate tumors (42), 9.4% of bladder tumors (43), and in lung tumors at a frequency of 6.1% (44). Thus, transcriptional mutagenesis and retromutagenesis may contribute towards the development and progression of such cancers.

Neil1, Neil2, Neil3, and Nthl1 overlap in substrate specificity for additional lesions other than 5-OHU and DHU, such as guanidinohydantoin. While the possibility that factors other than the transcription or replication status of the affected DNA may influence repair of such lesions, the collective evidence suggests that the factors that influence the repair of 5-OHU and DHU *in vivo* may influence repair more generally. NEIL2 expression is independent of the cell cycle (19), and as mentioned in the introduction, it binds human RNA polymerase, and Neil2$^{-/-}$ mice accumulate DNA damage in transcribed regions. Mouse Neil3 displays preference towards single-stranded oligonucleotide substrates containing spiroiminodihydantoin lesions (45), and human NEIL3 expression is cell cycle-dependent, induced during early S phase.(46,47) Evidence indicates that NTHL1 and NEIL1 expression are also cell cycle-dependent. It has been reported that NTHL1 is upregulated during S phase (47,48), expressed during G1/S (47),

and that NEIL1 expression is also upregulated during S phase (22). Since we did not observe significant back-up activity of Neil1, Neil3, or Nthl1 in the absence of Neil2 towards at least two lesions out of their common repertoire of substrates, placed in a transcribed gene in non-replicating constructs, the choice of DNA glycosylase may be, at least in part, determined by transcription versus replication and cell cycle phase.

Replication of the experiments described here employing well-defined *in vivo* systems, as well as determining the repair and transcriptional mutagenesis of these lesions occurring in the non-transcribed strand of Ras or in replicating constructs would help address these questions.

Acknowledgements

We would like to thank Stephen Lloyd for providing the Neil1$^{-/-}$ MEFs, as well as Christine Gran and Magnar Bjoras for providing the Neil2$^{-/-}$ and Neil1$^{-/-}$Neil2$^{-/-}$ MEFs.

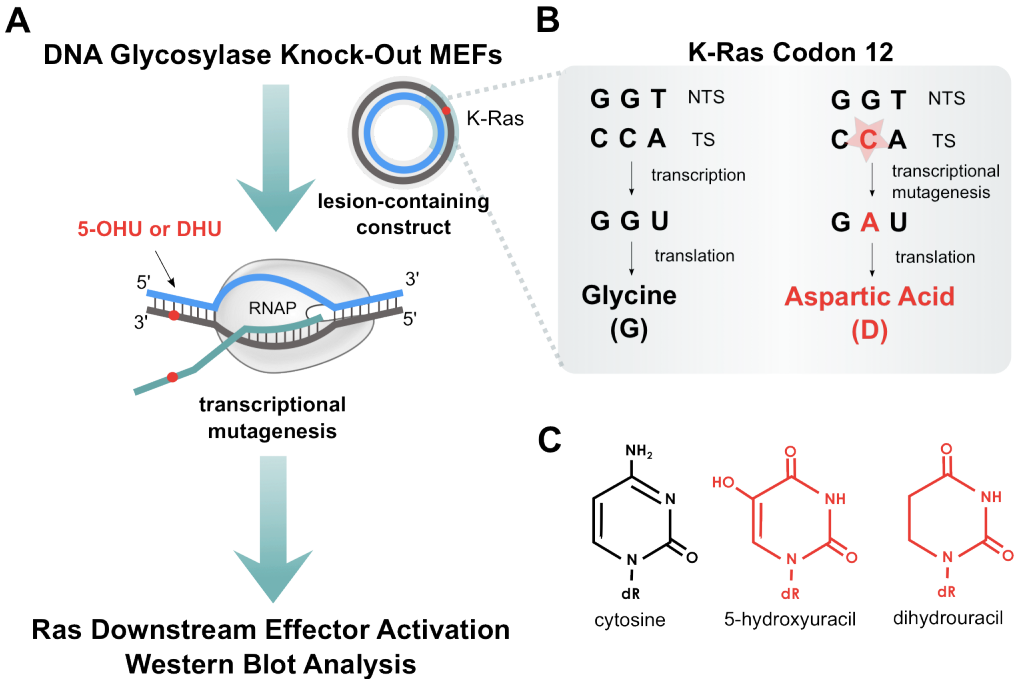


Figure 3.1: Systems for the Study of 5-Hydroxyuracil and Dihydrouracil Repair, Transcriptional Mutagenesis, and Its Phenotypic Consequences *in vivo*.

(A) Schematic of the experimental setup for determining the *in vivo* repair, TM and phenotypic consequences of 5-OHU and DHU. Mammalian expression vectors containing DNA damage lesions in the transcribed strand of Ras are delivered in glycosylase-deficient mouse embryonic fibroblasts using nucleofection and transcriptional mutagenesis due to the presence of unrepaired damage is determined by determining the downstream phenotypic effects of mutant Ras. (B) DNA sequence of codon 12 of K-Ras. Transcription past cytosine-containing constructs would result in the normal glycine-containing form of K-Ras and transcriptional mutagenesis past 5-OHU or DHU would result in the production of the constitutively active G12D mutant due to misincorporation of adenine opposite the lesion by RNA polymerase. (C) Chemical structures of cytosine and the cytosine-derived oxidative lesions 5-hydroxyuracil and dihydrouracil.

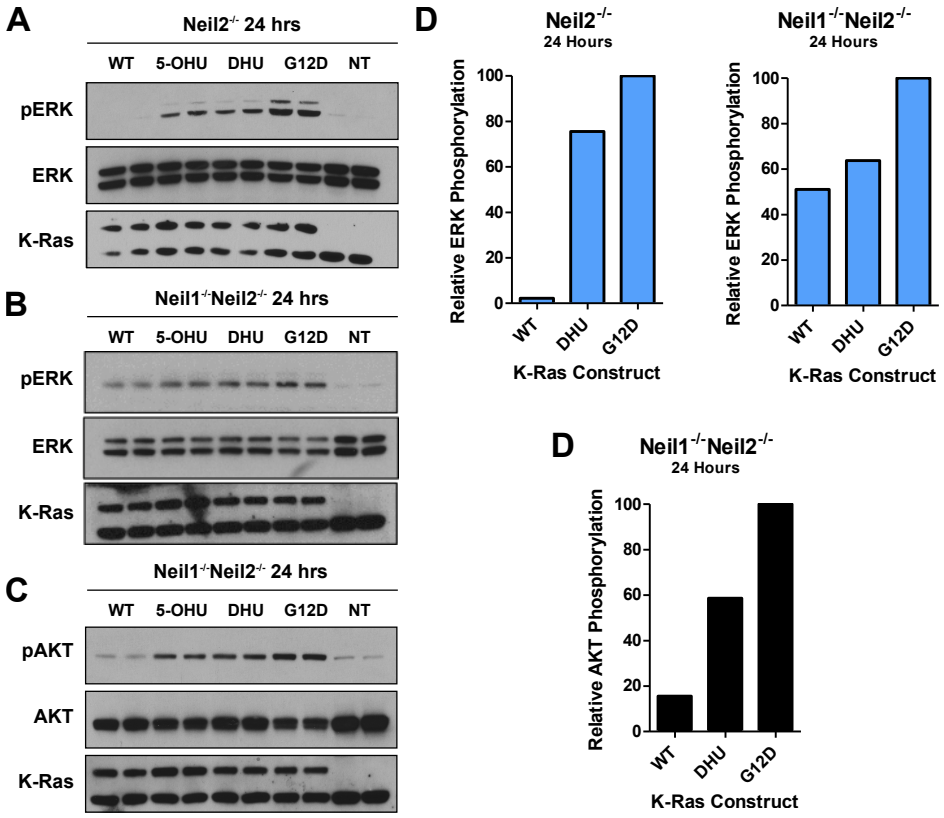


Figure 3.2: Dihydrouracil is Transcriptionally Mutagenic *in vivo*, Induces Oncogene Activation, and Neil2 is Implicated as a Major Glycosylase Repairing DHU from Transcribed DNA.

(A) Western blot analysis of phospho-ERK1/2 (Thr202/Tyr204), total ERK and K-Ras in Neil2$^{-/-}$ MEFs at 24 hours post-nucleofection with wild type (WT), 5-OHU, DHU, and G12D constructs and non-treated control (NT). Two lanes for each sample represent two technical replicates. (B) Same as (A), but using Neil1$^{-/-}$Neil2$^{-/-}$ MEFs. (C) Western blot analysis of AKT phosphorylation at Ser473 in Neil1$^{-/-}$Neil2$^{-/-}$ MEFs. (D) Quantification of the DHU and control samples from the Western blots shown above. DHU is transcriptionally mutagenic *in vivo*, and Neil2 appears to play a major role in its repair from transcribed DNA *in vivo*.

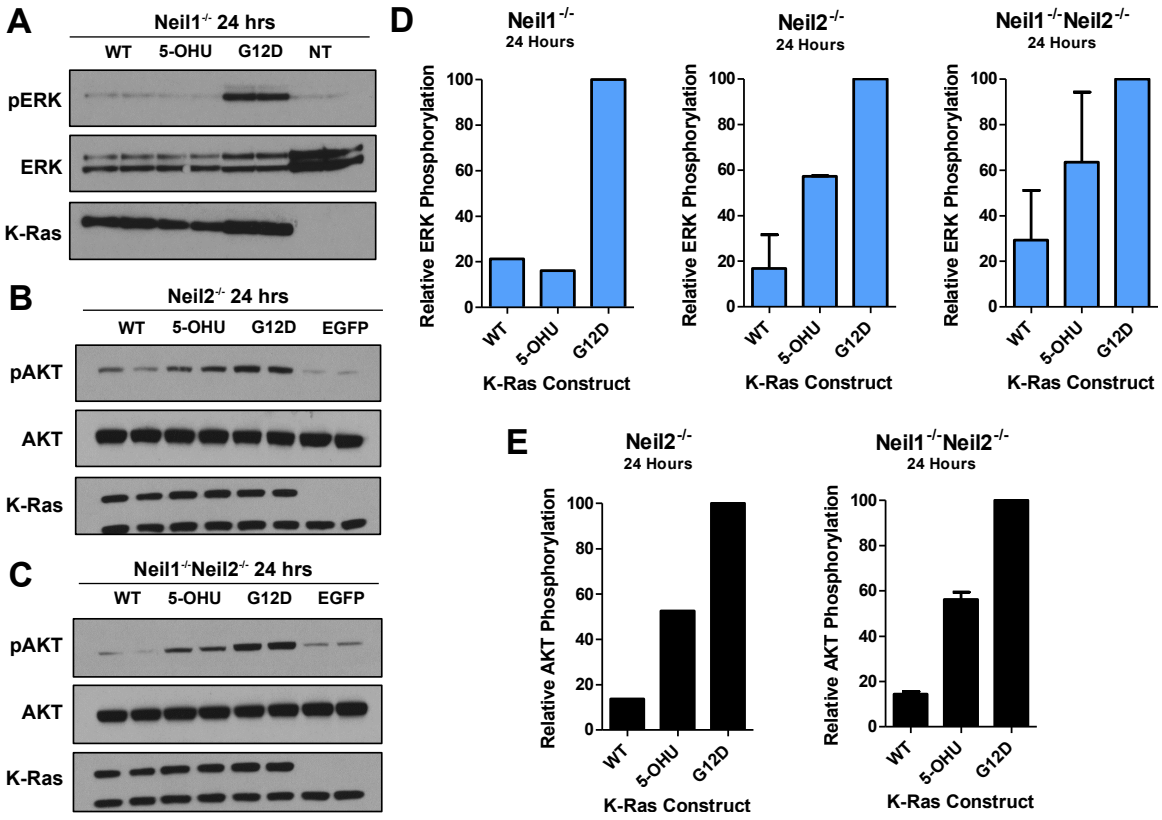


Figure 3.3: Neil2 is a Critical Glycosylase Repairing 5-OHU from Transcribed DNA *in vivo*, and Protects Mammalian Cells from Transcriptional Mutagenesis and Its Phenotypic Consequences.

(A) Western blot TM-mediated ERK phosphorylation analysis in Neil1$^{-/-}$ cells. (B) Western blot TM-mediated AKT phosphorylation analysis in Neil2$^{-/}$ MEFs. (C) Same as (B), but in Neil1$^{-/-}$Neil2$^{-/-}$ MEFs. (D) Western blot quantification of ERK phosphorylation in Neil1$^{-/-}$, Neil2$^{-/-}$, and Neil1$^{-/-}$Neil2$^{-/-}$ MEFs at 24 hours post-nucleofection. (E) AKT phosphorylation Western blot quantification in Neil2$^{-/-}$ and Neil1$^{-/-}$Neil2$^{-/-}$ MEFs. Error bars represent the standard error of the mean of two biological replicates. Neil2 deficiency was sufficient to promote the persistence of 5-OHU in mammalian cells, inducing TM and activation of more than one oncogenic pathways downstream of Ras at 24 hours post-nucleofection. 5-OHU appears to be fully repaired in the absence of Neil1 at 24 hours, and Neil1 deficiency does not appear to increase TM-mediated oncogenic signaling in cells deficient in both Neil1 and Neil2 compared to cells deficient only in Neil2. Neil1$^{-/-}$Neil2$^{-/-}$ 5-OHU data was also presented in Chapter 2.

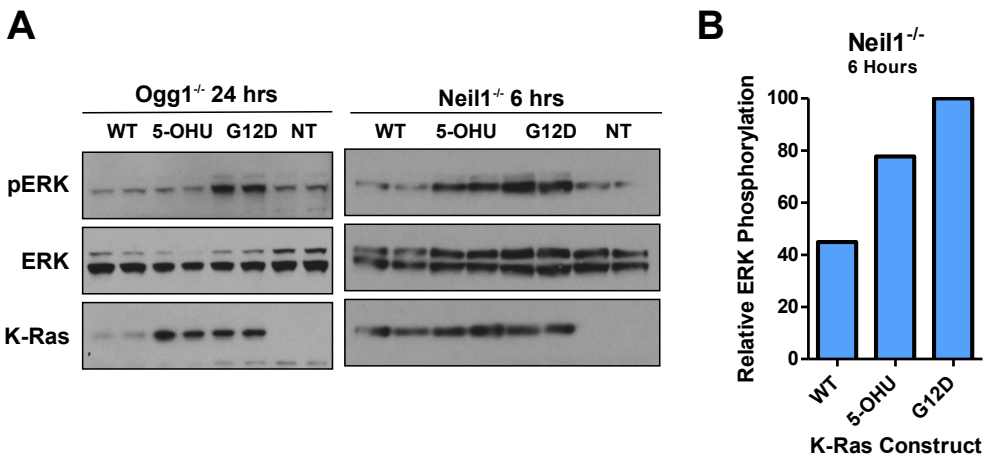


Figure 3.4: 5-OHU Appears to be Repaired at 24 hours in Ogg1$^{-/-}$ MEFs, but not in Neil1$^{-/-}$ MEFs at 6 hours.

(A) We observed 5-OHU-mediated Ras activation in Neil1$^{-/-}$ MEFs at 6 hours post-nucleofection. 5-OHU appeared to be repaired 24 hours following transfection with TransIT-2020 in Ogg1$^{-/-}$ MEFs, not inducing Ras activation. (B) Quantification of the Neil1$^{-/-}$ Western blot.

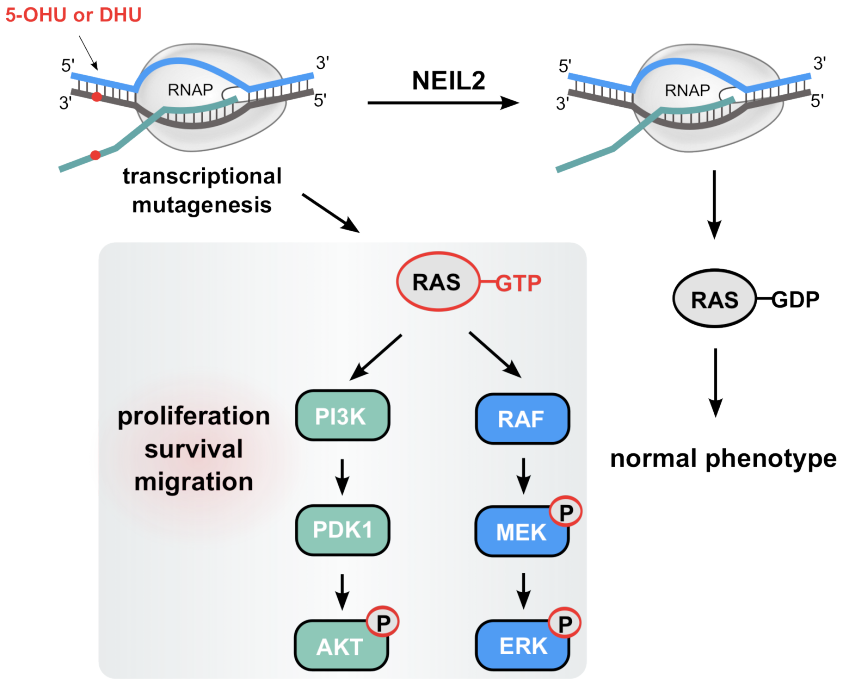


Figure 3.5: Graphical Summary of Results.

5-OHU and DHU are transcriptionally mutagenic in mammalian cells. If left unrepaired, when they occur in the G12 mutational hotspot of K-Ras, they can induce sustained activation of Ras effector pathways regulating a variety of oncogenic processes, such as proliferation, apoptosis, cellular migration and metastasis. Neil2 appears to be the primary DNA glycosylase repairing 5-OHU and DHU *in vivo* from transcribed DNA, and Neil2 deficiency promotes the occurrence of transcriptional mutagenesis and its downstream phenotypic consequences.

Literature Cited

1. Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature*, **362**, 709-715.
2. Gates, K.S. (2009) An overview of chemical processes that damage cellular DNA: spontaneous hydrolysis, alkylation, and reactions with radicals. *Chem Res Toxicol*, **22**, 1747-1760.
3. Cooke, M.S., Evans, M.D., Dizdaroglu, M. and Lunec, J. (2003) Oxidative DNA damage: mechanisms, mutation, and disease. *Faseb J*, **17**, 1195-1214.
4. Wallace, S.S., Murphy, D.L. and Sweasy, J.B. (2012) Base excision repair and cancer. *Cancer Lett*, **327**, 73-89.
5. Scott, T.L., Rangaswamy, S., Wicker, C.A. and Izumi, T. (2014) Repair of oxidative DNA damage and cancer: recent progress in DNA base excision repair. *Antioxid Redox Signal*, **20**, 708-726.
6. Wilson, D.M., 3rd and McNeill, D.R. (2007) Base excision repair and the central nervous system. *Neuroscience*, **145**, 1187-1200.
7. Wilson, D.M., 3rd and Bohr, V.A. (2007) The mechanics of base excision repair, and its relationship to aging and disease. *DNA Repair (Amst)*, **6**, 544-559.
8. Bregeon, D. and Doetsch, P.W. (2011) Transcriptional mutagenesis: causes and involvement in tumour development. *Nat Rev Cancer*, **11**, 218-227.
9. Morreall, J.F., Petrova, L. and Doetsch, P.W. (2013) Transcriptional mutagenesis and its potential roles in the etiology of cancer and bacterial antibiotic resistance. *J Cell Physiol*, **228**, 2257-2261.

10. Saxowsky, T.T. and Doetsch, P.W. (2006) RNA polymerase encounters with DNA damage: transcription-coupled repair or transcriptional mutagenesis? *Chem Rev*, **106**, 474-488.

11. Dianov, G.L. and Hubscher, U. (2013) Mammalian base excision repair: the forgotten archangel. *Nucleic Acids Res*, **41**, 3483-3490.

12. Saxowsky, T.T., Meadows, K.L., Klungland, A. and Doetsch, P.W. (2008) 8-Oxoguanine-mediated transcriptional mutagenesis causes Ras activation in mammalian cells. *Proc Natl Acad Sci U S A*, **105**, 18877-18882.

13. Hu, J., de Souza-Pinto, N.C., Haraguchi, K., Hogue, B.A., Jaruga, P., Greenberg, M.M., Dizdaroglu, M. and Bohr, V.A. (2005) Repair of formamidopyrimidines in DNA involves different glycosylases: role of the OGG1, NTH1, and NEIL1 enzymes. *J Biol Chem*, **280**, 40544-40551.

14. Hegde, M.L., Hegde, P.M., Bellot, L.J., Mandal, S.M., Hazra, T.K., Li, G.M., Boldogh, I., Tomkinson, A.E. and Mitra, S. (2013) Prereplicative repair of oxidized bases in the human genome is mediated by NEIL1 DNA glycosylase together with replication proteins. *Proc Natl Acad Sci U S A*, **110**, E3090-3099.

15. Dou, H., Mitra, S. and Hazra, T.K. (2003) Repair of oxidized bases in DNA bubble structures by human DNA glycosylases NEIL1 and NEIL2. *J Biol Chem*, **278**, 49679-49684.

16. Rosenquist, T.A., Zaika, E., Fernandes, A.S., Zharkov, D.O., Miller, H. and Grollman, A.P. (2003) The novel DNA glycosylase, NEIL1, protects mammalian cells from radiation-mediated cell death. *DNA Repair (Amst)*, **2**, 581-591.

17. Takao, M., Kanno, S., Kobayashi, K., Zhang, Q.M., Yonei, S., van der Horst, G.T. and Yasui, A. (2002) A back-up glycosylase in Nth1 knock-out mice is a functional Nei (endonuclease VIII) homologue. *J Biol Chem*, **277**, 42205-42213.

18. Prakash, A., Carroll, B.L., Sweasy, J.B., Wallace, S.S. and Doublie, S. (2014) Genome and cancer single nucleotide polymorphisms of the human NEIL1 DNA glycosylase: activity, structure, and the effect of editing. *DNA Repair (Amst)*, **14**, 17-26.

19. Hazra, T.K., Kow, Y.W., Hatahet, Z., Imhoff, B., Boldogh, I., Mokkapati, S.K., Mitra, S. and Izumi, T. (2002) Identification and characterization of a novel human DNA glycosylase for repair of cytosine-derived lesions. *J Biol Chem*, **277**, 30417-30420.

20. Liu, M., Bandaru, V., Bond, J.P., Jaruga, P., Zhao, X., Christov, P.P., Burrows, C.J., Rizzo, C.J., Dizdaroglu, M. and Wallace, S.S. (2010) The mouse ortholog of NEIL3 is a functional DNA glycosylase in vitro and in vivo. *Proc Natl Acad Sci U S A*, **107**, 4925-4930.

21. Krokeide, S.Z., Laerdahl, J.K., Salah, M., Luna, L., Cederkvist, F.H., Fleming, A.M., Burrows, C.J., Dalhus, B. and Bjoras, M. (2013) Human NEIL3 is mainly a monofunctional DNA glycosylase removing spiroimindiohydantoin and guanidinohydantoin. *DNA Repair (Amst)*, **12**, 1159-1164.

22. Hazra, T.K., Izumi, T., Boldogh, I., Imhoff, B., Kow, Y.W., Jaruga, P., Dizdaroglu, M. and Mitra, S. (2002) Identification and characterization of a human DNA glycosylase for repair of modified bases in oxidatively damaged DNA. *Proc Natl Acad Sci U S A*, **99**, 3523-3528.

23. Ikeda, S., Biswas, T., Roy, R., Izumi, T., Boldogh, I., Kurosky, A., Sarker, A.H., Seki, S. and Mitra, S. (1998) Purification and characterization of human NTH1, a homolog of Escherichia coli endonuclease III. Direct identification of Lys-212 as the active nucleophilic residue. *J Biol Chem*, **273**, 21585-21593.

24. Chakraborty, A., Wakamiya, M., Venkova-Canova, T., Pandita, R.K., Aguilera-Aguirre, L., Sarker, A.H., Singh, D.K., Hosoki, K., Wood, T.G., Sharma, G. *et al*. (2015) Neil2-null Mice Accumulate Oxidized DNA Bases in the Transcriptionally Active Sequences of the Genome and Are Susceptible to Innate Inflammation. *J Biol Chem*, **290**, 24636-24648.

25. Banerjee, D., Mandal, S.M., Das, A., Hegde, M.L., Das, S., Bhakat, K.K., Boldogh, I., Sarkar, P.S., Mitra, S. and Hazra, T.K. (2011) Preferential repair of oxidized base damage in the transcribed genes of mammalian cells. *J Biol Chem*, **286**, 6006-6016.

26. Panigrahi, S.K., Hopkins, K.M. and Lieberman, H.B. (2015) Regulation of NEIL1 protein abundance by RAD9 is important for efficient base excision repair. *Nucleic Acids Res*, **43**, 4531-4546.

27. Todaro, G.J. and Green, H. (1963) Quantitative studies of the growth of mouse embryo cells in culture and their development into established lines. *J Cell Biol*, **17**, 299-313.

28. Witkiewicz, A.K., McMillan, E.A., Balaji, U., Baek, G., Lin, W.C., Mansour, J., Mollaee, M., Wagner, K.U., Koduru, P., Yopp, A. *et al*. (2015) Whole-exome sequencing of pancreatic cancer defines genetic diversity and therapeutic targets. *Nat Commun*, **6**, 6744.

29. Rizvi, N.A., Hellmann, M.D., Snyder, A., Kvistborg, P., Makarov, V., Havel, J.J., Lee, W., Yuan, J., Wong, P., Ho, T.S. *et al*. (2015) Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science*, **348**, 124-128.

30. (2012) Comprehensive molecular characterization of human colon and rectal cancer. *Nature*, **487**, 330-337.

31. Pylayeva-Gupta, Y., Grabocka, E. and Bar-Sagi, D. (2011) RAS oncogenes: weaving a tumorigenic web. *Nat Rev Cancer*, **11**, 761-774.

32. Charlet-Berguerand, N., Feuerhahn, S., Kong, S.E., Ziserman, H., Conaway, J.W., Conaway, R. and Egly, J.M. (2006) RNA polymerase II bypass of oxidative DNA damage is regulated by transcription elongation factors. *Embo J*, **25**, 5481-5491.

33. Liu, J. and Doetsch, P.W. (1998) Escherichia coli RNA and DNA polymerase bypass of dihydrouracil: mutagenic potential via transcription and replication. *Nucleic Acids Res*, **26**, 1707-1712.

34. Liu, J., Zhou, W. and Doetsch, P.W. (1995) RNA polymerase bypass at sites of dihydrouracil: implications for transcriptional mutagenesis. *Mol Cell Biol*, **15**, 6729-6735.

35. Zhou, W. and Doetsch, P.W. (1993) Effects of abasic sites and DNA single-strand breaks on prokaryotic RNA polymerases. *Proc Natl Acad Sci U S A*, **90**, 6601-6605.

36. Clauson, C.L., Oestreich, K.J., Austin, J.W. and Doetsch, P.W. (2010) Abasic sites and strand breaks in DNA cause transcriptional mutagenesis in Escherichia coli. *Proc Natl Acad Sci U S A*, **107**, 3657-3662.

37. Doetsch, P.W. (2002) Translesion synthesis by RNA polymerases: occurrence and biological implications for transcriptional mutagenesis. *Mutat Res*, **510**, 131-140.

38. Morreall, J., Kim, A., Liu, Y., Degtyareva, N., Weiss, B. and Doetsch, P.W. (2015) Evidence for Retromutagenesis as a Mechanism for Adaptive Mutation in Escherichia coli. *PLoS Genet*, **11**, e1005477.

39. Dey, S., Maiti, A.K., Hegde, M.L., Hegde, P.M., Boldogh, I., Sarkar, P.S., Abdel-Rahman, S.Z., Sarker, A.H., Hang, B., Xie, J. *et al*. (2012) Increased risk of lung cancer associated with a functionally impaired polymorphic variant of the human DNA glycosylase NEIL2. *DNA Repair (Amst)*, **11**, 570-578.

40. Zhai, X., Zhao, H., Liu, Z., Wang, L.E., El-Naggar, A.K., Sturgis, E.M. and Wei, Q. (2008) Functional variants of the NEIL1 and NEIL2 genes and risk and progression of squamous cell carcinoma of the oral cavity and oropharynx. *Clin Cancer Res*, **14**, 4345-4352.

41. Gao, J., Aksoy, B.A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S.O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E. *et al*. (2013) Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal*, **6**, pl1.

42. (2015) The Molecular Taxonomy of Primary Prostate Cancer. *Cell*, **163**, 1011-1025.

43. (2014) Comprehensive molecular characterization of urothelial bladder carcinoma. *Nature*, **507**, 315-322.

44. (2014) Comprehensive molecular profiling of lung adenocarcinoma. *Nature*, **511**, 543-550.

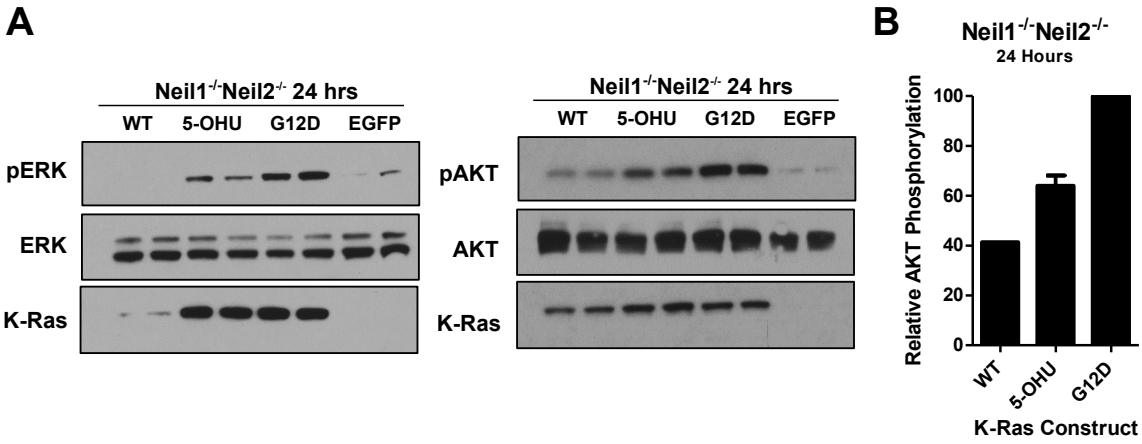
45. Liu, M., Imamura, K., Averill, A.M., Wallace, S.S. and Doublie, S. (2013) Structural characterization of a mouse ortholog of human NEIL3 with a marked preference for single-stranded DNA. *Structure*, **21**, 247-256.

46. Neurauter, C.G., Luna, L. and Bjoras, M. (2012) Release from quiescence stimulates the expression of human NEIL3 under the control of the Ras dependent ERK-MAP kinase pathway. *DNA Repair (Amst)*, **11**, 401-409.

47. Mjelle, R., Hegre, S.A., Aas, P.A., Slupphaug, G., Drablos, F., Saetrom, P. and Krokan, H.E. (2015) Cell cycle regulation of human DNA repair and chromatin remodeling genes. *DNA Repair (Amst)*, **30**, 53-67.

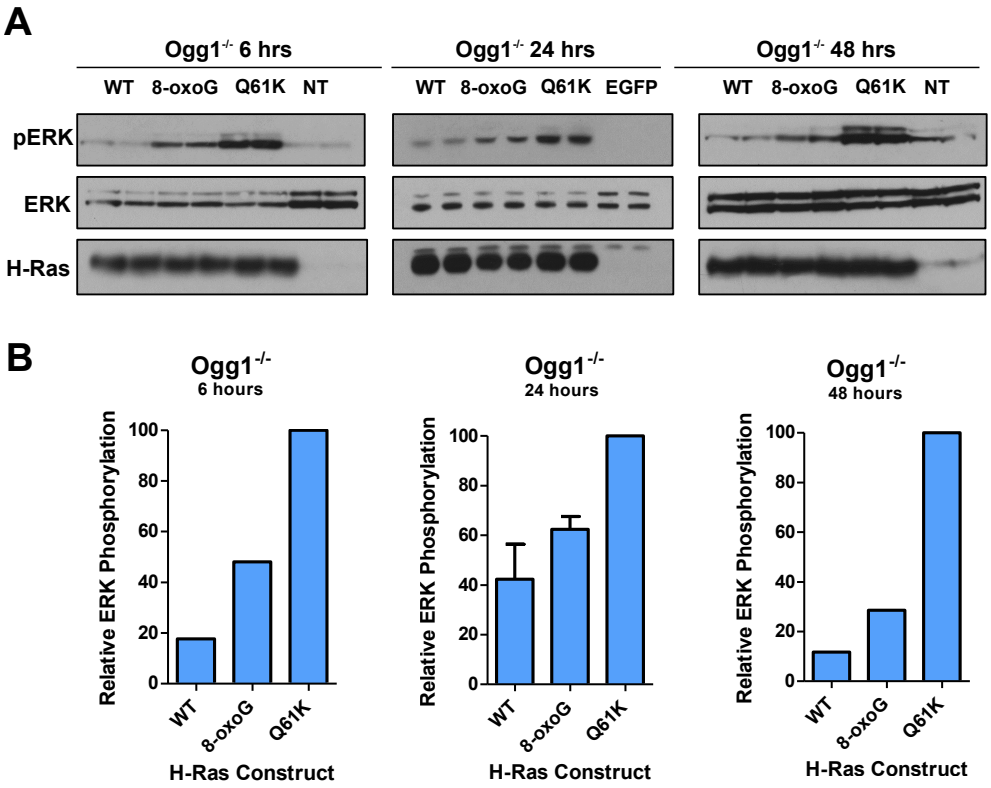
48. Luna, L., Bjoras, M., Hoff, E., Rognes, T. and Seeberg, E. (2000) Cell-cycle regulation, intracellular sorting and induced overexpression of the human NTH1 DNA glycosylase involved in removal of formamidopyrimidine residues from DNA. *Mutat Res*, **460**, 95-104.

Supplementary Information



Supplementary Figure 3.1: 5-OHU-Mediated Ras Activation at 24 hours Following Magnetofection of Neil1$^{-/-}$Neil2$^{-/-}$ MEFs.

(A) Western blot ERK and AKT phosphorylation analysis in Neil1$^{-/-}$Neil2$^{-/-}$ MEFs at 24 hours following magnetofection using LipoMag kit (OZ Biosciences), as per the manufacturer's instructions. (B) Phospho-AKT Western blot quantification. Error bar indicates standard error of the mean of two biological replicates. The wild type sample for one replicate was not quantified due to failure of transfection for this sample. Results employing magnetofection are consistent with those following nucleofection.



Supplementary Figure 3.2: Sustained 8-oxoguanine-Mediated Oncogene Activation in Ogg1$^{-/-}$ MEFs.

In order to determine the *in vivo* timecourse of 8-oxoG repair and TM-mediated oncogene activation, we generated constructs containing 8-oxoG in the transcribed strand of codon 61 of HRAS, such that mutagenic translesion synthesis by RNA polymerase would result in the constitutively active Q61K mutant (12), as well as control non-mutant and Q61K mutant HRAS constructs. In order to study 8-oxoG, we employed MEFs deficient in Ogg1. (A) ERK phosphorylation Western blot` analyses of Ogg1$^{-/-}$ samples nucleofected with H-Ras construct set. (B) Quantification of the Ogg1$^{-/-}$ and 8-oxoG Western blots. Error bars for the 24 hours set represent standard error of the mean of two biological replicates. 8-oxoG-mediated TM may be sustained longer than previously known.

4. Discussion and Future Directions

Introduction

The genome of each mammalian cell acquires thousands of diverse DNA damage lesions per day, each of which can differentially contribute to transcriptional or replicative mutagenesis and their phenotypic consequences, and emerging evidence indicates that the activity of each DNA repair pathway component may be differentially modulated by the context in which DNA damage occurs. However, the *in vivo* transcriptional mutagenesis (TM) mediated by a variety of defined DNA lesions, the *in vivo* factors that may influence their repair, and the relative contributions of transcriptional or replicative mutagenesis to the development of potentially deleterious cellular phenotypes or disease, remain largely unexplored areas. Better understanding of the molecular and phenotypic consequences and the factors that influence repair of defined DNA damage *in vivo* requires further investigation and the development of efficient and reliable methodologies, as well as novel systems and approaches to address these questions. In this dissertation, I developed a reliable and efficient protocol for the large-scale production of mammalian expression phagemid vectors containing site-specific base modifications. I identified the *E. coli* culture density at the time of M13KO7 helper phage infection as a predictor of single-stranded DNA (ssDNA) yield and developed a protocol from which more than 300 μg of ultra-pure ssDNA, suitable for second-strand synthesis, can be produced from a single preparation. I also developed a protocol for the large-scale production of double-stranded, site-specific lesion-bearing vectors, also with significantly improved yields. Employing the tools and systems developed, we provided novel insights on the *in vivo* transcriptional mutagenesis induced by the oxidative cytosine lesions 5-hydroxyuracil (5-OHU) and dihydrouracil (DHU) and

their phenotypic consequences in oncogene activation, as well as their *in vivo* repair and the factors that may influence base excision repair of these lesions, and thus transcriptional or replicative mutagenesis.

Transcriptional Mutagenesis and Defined DNA Damage Repair Studies

In Chapters 2 and 3, we presented evidence that 5-OHU and DHU are transcriptionally mutagenic *in vivo* and suggesting that Neil2 is the primary glycosylase repairing 5-OHU and DHU from transcribed, non-replicating DNA in mammalian cells. The phenotypic read-out phospho-AKT for 5-OHU in Neil1$^{-/-}$Neil2$^{-/-}$ MEFs reached statistical significance, and we have observed a similar phenotype employing an alternative method of transfection (magnetofection). However, the remaining experiments should be repeated in order to confirm the observations that DHU is transcriptionally mutagenic *in vivo* and confirm that Neil2 is the primary DNA glycosylase repairing 5-OHU and DHU *in vivo* from transcribed DNA. In addition, the systems described here, with minor modifications as described below, can be employed to rigorously test the role of Neil2 in mediating transcription-coupled repair of 5-OHU and DHU and determine the role of Neil1, Neil3, and Nthl1 for the repair of these lesions when they occur in replicating or non-transcribed DNA. We propose that Neil2 is the primary DNA glycosylase protecting cells from TM mediated by 5-OHU and DHU occurring in transcribed regions, and Neil2 deficiency can result in the initiation of potentially deleterious phenotypic consequences, such as senescence or cell death, *in vivo* due to TM even in the absence of proliferation. Retromutagenesis and permanent establishment of a phenotype mediated by 5-OHU or DHU, such as aberrant cell cycle entry due to

oncogene activation, may be promoted by deficiency in pre-replicative repair mediated by Neil1, Neil3, and/or Nthl1 (Figure 4.1).

In order to confirm the absence of Neil2-mediated repair of 5-OHU and DHU from non-transcribed DNA *in vivo*, constructs containing these lesions on the non-transcribed strand of codon 12 of K-Ras can be employed. This can be achieved by making ssDNA using pcDNA3.1(-), which contains the f1 origin of replication in the reverse direction such that ssDNA containing the transcribed strand of K-Ras can be produced. Alternatively, pcDNA3.1(+) can be employed with lesion-containing oligonucleotides annealing outside of the coding region. Lesion repair can then be determined by purification of the constructs from transfected cells using Hirt extract protocol, followed by Fpg digestion and Southern blot analysis (1). In order to determine the relative contribution of Neil1, Neil3 or Nthl1 towards the repair of of 5-OHU or DHU occurring in non-transcribed DNA *in vivo*, similar experiments can be performed employing cells deficient in Neil1, Neil3, and/or Nthl1.

Mouse polyoma virus-based vectors can replicate autonomously in murine cells as episomes (2), and they have been previously used to study translesion synthesis by specialized polymerases (3,4), as well as replication-associated repair of A:8-oxoG mismatches by *E. coli* MutY homolog MYH in MEFs (5). In order to determine the relative contribution of Neil1, Neil2, Neil3, or Nthl1 towards the repair of 5-OHU and DHU from replicating DNA, replicating site-specific lesion-containing constructs can be prepared using a mouse polyoma virus-based backbone, also containing the f1 origin or replication for ssDNA production (5). Lesion repair can then be determined as described above. ERK and AKT phosphorylation analysis following a round of cell division in

these experiments would indicate the occurrence of replicative mutagenesis, which can also be confirmed by sequencing analysis of re-isolated constructs. These well-defined systems will allow us to precisely delineate the relative contributions of DNA glycosylases towards the repair, replicative or transcriptional mutagenesis, and biological consequences of a diverse variety of DNA damage lesions in mammalian cells. Since these systems allow us to deliver expression vectors containing a single DNA damage lesion, they also allow us to study DNA damage repair in more physiological conditions without the addition of excess DNA damage that may affect DNA repair pathway function and also alter cell viability and phenotype. Development of systems that allow the targeted introduction of a single defined lesion at a defined position of genomic DNA would represent a further improvement of the systems described here.

The tools and systems described here have broad applications for the study of DNA repair of defined DNA damage and transcriptional or replicative mutagenesis, and can be employed for the study of DNA damage in a variety of contexts, including largely unexplored, emerging areas. For example, employing TM as a reporter of DNA repair in a variety of studies that do not necessarily address the biological consequences of TM poses a unique advantage in that it is a convenient tool amenable for the study of DNA damage repair in quiescent or terminally differentiated cells, such as neurons. DNA repair in non-proliferating cells, such as rarely dividing stem cells or post-mitotic neurons, has high relevance to human disease development. Few studies have addressed DNA repair specifically in largely quiescent or post-mitotic cells, however the mechanisms of DNA repair and the consequences of DNA damage may differ in non-dividing cells (6,7). For example, the expression of some BER components can vary throughout the cell cycle

(8,9), BER is attenuated in terminally differentiated muscle cells compared to their less-differentiated counterparts (10), and may also be modulated during neuronal differentiation (11). The activity of transcription-coupled nucleotide excision repair of bulky, helix-distorting damage can differ compared to global genome nucleotide excision repair in terminally differentiated cells (6). Whether BER components can be similarly separated into transcription-coupled and non-transcription-coupled, or replication-coupled and non-replication-coupled, and how the activity of each could vary in non-dividing cells that undergo transcription but not replication remains to be determined and the tools and systems described here can be employed to address these questions.

Phenotypic Consequences of Transcriptional Mutagenesis: Beyond Biochemical Signaling

Tumors can be initiated by the acquisition of a limited set of defined oncogenic mutations, and the Ras family of oncogenes are one of the most frequently mutated genes in human tumors. The role of activating Ras mutations in tumorigenesis is well-established, however tumorigenesis can require the presence of a cooperating mutation, such as one inactivating a tumor suppressor (12,13). In the absence of a cooperating mutation, activation of oncogenes, such as Ras and others, can lead to increases in reactive oxygen species (ROS), extensive DNA damage, and activation of the DNA damage response (DDR), ultimately resulting in a growth arrest, termed oncogene-induced senescence (OIS) (14,15). OIS is present in human pre-malignant lesions (16,17), serving as a barrier that restricts malignant progression (18).

Thus, depending on the context, TM-mediated oncogene activation may result in tumor initiation or OIS (Figure 4.2). By employing primary Neil2$^{-/-}$ MEFs with intact

DDR and the 5-OHU or DHU constructs described here, the phenotypic readouts for TM can be extended beyond sustained oncogenic signaling and the role of TM-mediated oncogene activation in inducing OIS phenotypes can be determined (Figure 4.3). OIS-related phenotypic changes, such as double-strand breaks can be determined using immunocytochemistry employing antibodies against DSB markers such as γH2A.X or 53BP1 or by the comet assay (19). Activation of the DDR can be determined by Western blot analysis employing antibodies against DDR components, such as p53 (20).

Conclusions

The work in this dissertation provides major improvements towards the development of methodology and systems for the study of defined DNA damage repair in mammalian cells and brings novel insights into the repair of 5-OHU and DHU *in vivo* and their phenotypic consequences due to their *in vivo* transcriptional mutagenicity. The significantly improved, reliable and predictable methodology for the production of mammalian expression vectors containing site-specific DNA damages and the systems described here represent a major advance for the study of defined DNA damage repair, transcriptional mutagenesis, and its phenotypic consequences. The majority of DNA damage studies have employed *in vitro* systems due to their simplicity and ease of use, and we have gained a wealth of knowledge by employing such systems. However, the relevance of such systems to *in vivo* physiological conditions and disease development is somewhat limited. Previous protocols for the production of damage-containing constructs have employed protocols that are greatly influenced by a variety of uncharacterized factors. In this dissertation, we have identified the factors that determine protocol

reproducibility and reliability and have developed an improved technique for the generation of hundreds of micrograms of constructs at a time, as opposed to about 15 micrograms per preparation (21). The ability to reliably and reproducibly generate large quantities of highly pure, site-specific damage-bearing constructs enable the study of defined damage repair in the context of transcriptional or replicative mutagenesis and their phenotypic consequences *in vivo*. The ability to study defined damage repair *in vivo* allows us to delineate the consequences of each individual damage that may differentially compromise transcription or replication and differentially contribute towards the development of disease. Broadening the study of DNA damage to the contexts of not only replicative but also transcriptional mutagenesis will allow us to gain a better mechanistic understanding on the impact of DNA damage to human health.

The application of the tools and systems described here would allow for a greater understanding of the mechanisms of base excision repair and the *in vivo* factors that influence its efficiency. While further investigation is required, the results here suggest that DNA glycosylases may not be entirely redundant, but may have more specialized functions depending on the transcription or replication status of the damaged DNA. Such specialization may influence the biological outcomes of DNA damage, transcriptional or replicative mutagenesis, and the development of diseases such as cancer or neurodegeneration. Unrepaired DNA damage in rapidly proliferating cells may lead to replicative mutagenesis and tumor development, while unrepaired damage in non-dividing cells may lead to transcriptional mutagenesis and senescence, cell death, or cell cycle entry and permanent establishment of the original transcriptional mutation via retromutagenesis. The systems described here should prove useful for future studies of

defined DNA damage repair and the differential contribution each specific lesion, DNA repair component, and mutagenesis mechanism towards the development of human disease.

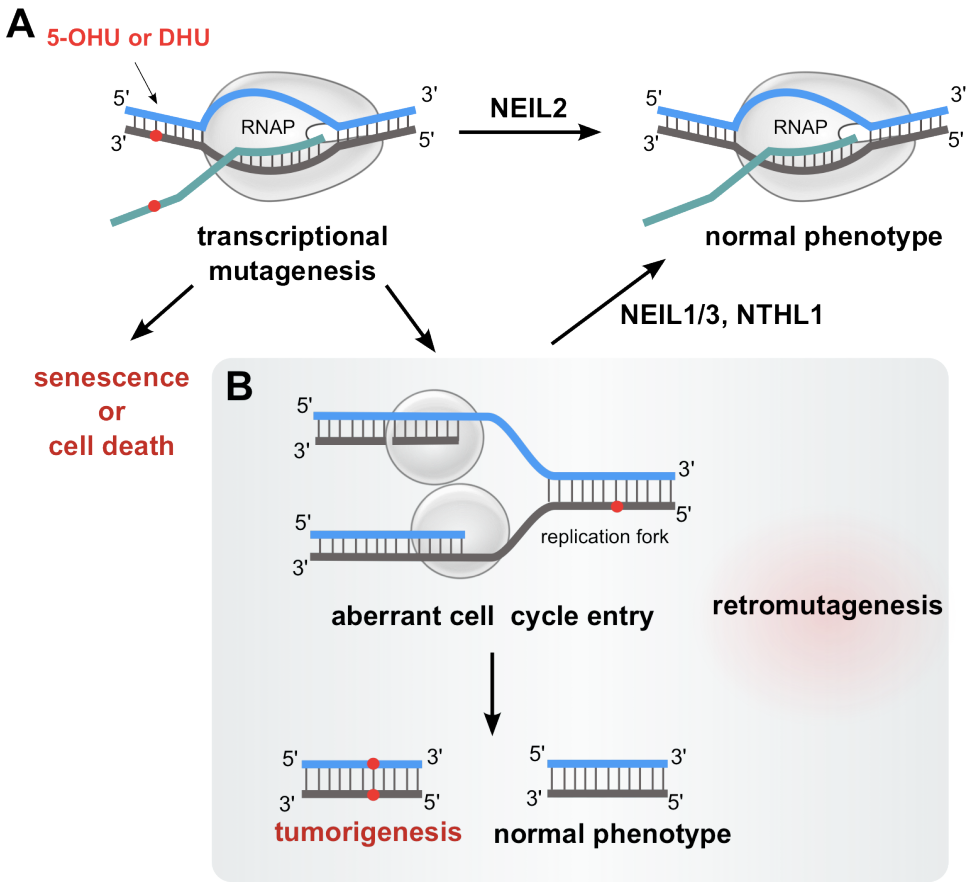


Figure 4.1: Model for the Potential Consequences of Transcriptional Mutagenesis Depending on Base Excision Repair Capacity.

(A) We propose TM mediated by lesions, such as 5-OHU or DHU, in the presence of a Neil2 deficiency may induce a phenotype such as senescence or cell death even in the absence of cell division. (B) A transcriptional mutation may induce oncogene activation and cell cycle entry. Permanent establishment of the original 5-OHU- or DHU-mediated transcriptional mutation via retromutagenesis may be promoted by a deficiency in additional enzymes such as Neil1, Neil3 or Nthl1.

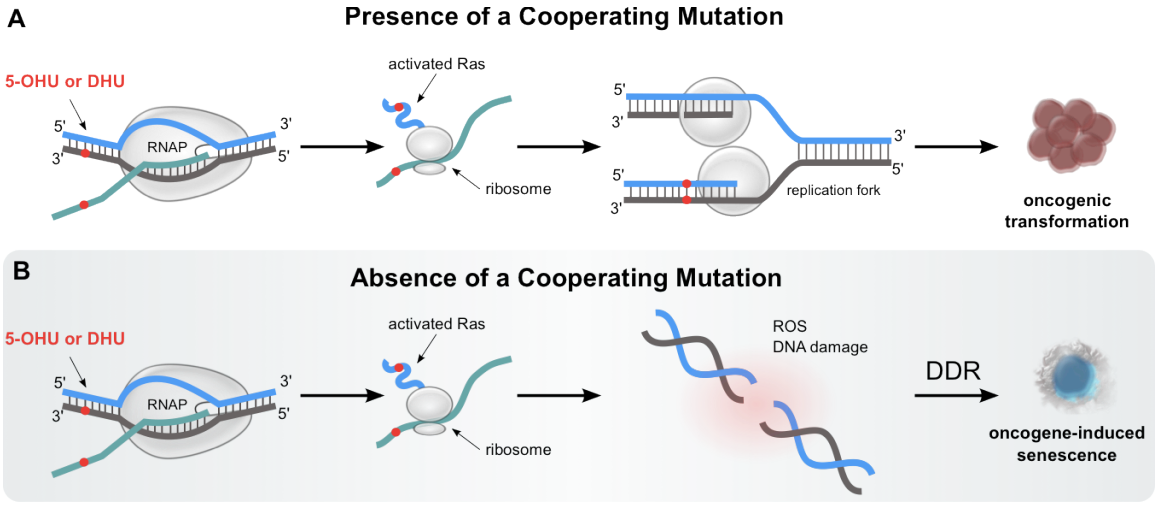


Figure 4.2: Potential Consequences of Oncogene Activation Mediated by Transcriptional Mutagenesis.

(A) In the presence of a cooperating mutation, TM-mediated oncogene activation may result in cell cycle entry, permanent establishment of the original transcriptional mutation, and oncogenic transformation. (B) In the absence of a cooperating mutation, TM may result in OIS. Modified from (22). DDR, DNA damage response; ROS, reactive oxygen species.

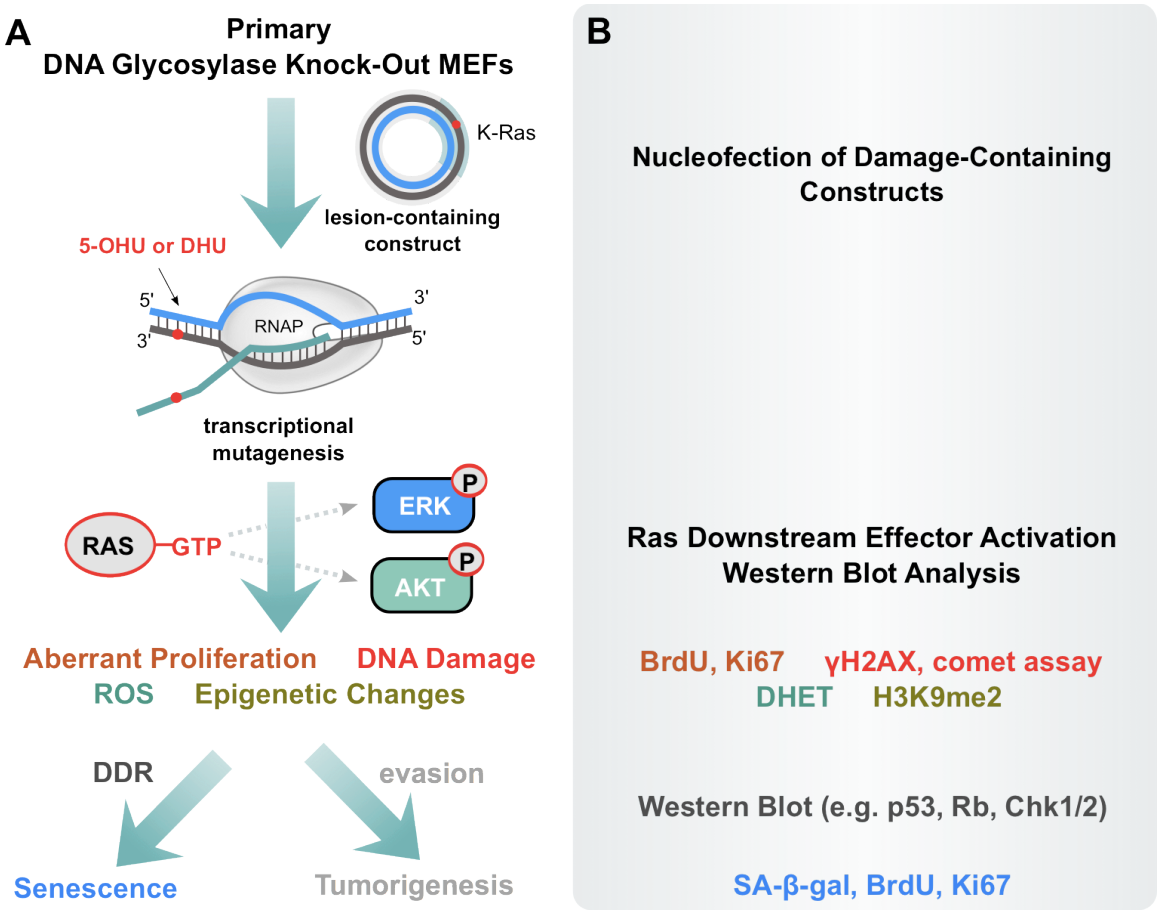


Figure 4.3: An Experimental System for the Study of Transcriptional Mutagenesis in the Context of Oncogene-Induced Senescence.

(A) Model system for the study of TM-mediated oncogene activation in the context of oncogene-induced senescence. Primary MEFs with intact DDR can be nucleofected with lesion-containing constructs. Aberrant oncogene activation in primary MEFs can induce increases in ROS, additional DNA damage, activation of the DDR and ultimately OIS. (B) Experimental read-outs for TM-mediated phenotypes related to oncogene-induced senescence, with assays color-coded depending on the phenotype they measure in (A). BrdU, 5-bromo-2'-deoxyuridine; DHET, dihydroxyeicosatrienoic acid; H3K9me2, dimethylation of histone H3 at lysine 9; SA-β-gal, senescence-associated β-galactosidase.

References

1. Saxowsky, T.T., Meadows, K.L., Klungland, A. and Doetsch, P.W. (2008) 8-Oxoguanine-mediated transcriptional mutagenesis causes Ras activation in mammalian cells. *Proc Natl Acad Sci U S A*, **105**, 18877-18882.
2. Gassmann, M., Donoho, G. and Berg, P. (1995) Maintenance of an extrachromosomal plasmid vector in mouse embryonic stem cells. *Proc Natl Acad Sci U S A*, **92**, 1292-1296.
3. Yang, I.Y., Hashimoto, K., de Wind, N., Blair, I.A. and Moriya, M. (2009) Two distinct translesion synthesis pathways across a lipid peroxidation-derived DNA adduct in mammalian cells. *J Biol Chem*, **284**, 191-198.
4. Yuan, B., You, C., Andersen, N., Jiang, Y., Moriya, M., O'Connor, T.R. and Wang, Y. (2011) The roles of DNA polymerases kappa and iota in the error-free bypass of N2-carboxyalkyl-2'-deoxyguanosine lesions in mammalian cells. *J Biol Chem*, **286**, 17503-17511.
5. Hayashi, H., Tominaga, Y., Hirano, S., McKenna, A.E., Nakabeppu, Y. and Matsumoto, Y. (2002) Replication-associated repair of adenine:8-oxoguanine mispairs by MYH. *Curr Biol*, **12**, 335-339.
6. Nouspikel, T. and Hanawalt, P.C. (2002) DNA repair in terminally differentiated cells. *DNA Repair (Amst)*, **1**, 59-75.
7. Iyama, T. and Wilson, D.M., 3rd. (2013) DNA repair mechanisms in dividing and non-dividing cells. *DNA Repair (Amst)*, **12**, 620-636.

8. Neurauter, C.G., Luna, L. and Bjoras, M. (2012) Release from quiescence stimulates the expression of human NEIL3 under the control of the Ras dependent ERK-MAP kinase pathway. *DNA Repair (Amst)*, **11**, 401-409.

9. Mjelle, R., Hegre, S.A., Aas, P.A., Slupphaug, G., Drablos, F., Saetrom, P. and Krokan, H.E. (2015) Cell cycle regulation of human DNA repair and chromatin remodeling genes. *DNA Repair (Amst)*, **30**, 53-67.

10. Narciso, L., Fortini, P., Pajalunga, D., Franchitto, A., Liu, P., Degan, P., Frechet, M., Demple, B., Crescenzi, M. and Dogliotti, E. (2007) Terminally differentiated muscle cells are defective in base excision DNA repair and hypersensitive to oxygen injury. *Proc Natl Acad Sci U S A*, **104**, 17010-17015.

11. Sykora, P., Yang, J.L., Ferrarelli, L.K., Tian, J., Tadokoro, T., Kulkarni, A., Weissman, L., Keijzers, G., Wilson, D.M., 3rd, Mattson, M.P. *et al*. (2013) Modulation of DNA base excision repair during neuronal differentiation. *Neurobiol Aging*, **34**, 1717-1727.

12. Serrano, M., Lin, A.W., McCurrach, M.E., Beach, D. and Lowe, S.W. (1997) Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. *Cell*, **88**, 593-602.

13. Bennecke, M., Kriegl, L., Bajbouj, M., Retzlaff, K., Robine, S., Jung, A., Arkan, M.C., Kirchner, T. and Greten, F.R. (2010) Ink4a/Arf and oncogene-induced senescence prevent tumor progression during alternative colorectal tumorigenesis. *Cancer Cell*, **18**, 135-146.

14. Bartkova, J., Rezaei, N., Liontos, M., Karakaidos, P., Kletsas, D., Issaeva, N., Vassiliou, L.V., Kolettas, E., Niforou, K., Zoumpourlis, V.C. *et al*. (2006)

Oncogene-induced senescence is part of the tumorigenesis barrier imposed by DNA damage checkpoints. *Nature*, **444**, 633-637.

15. Tu, Z., Aird, K.M., Bitler, B.G., Nicodemus, J.P., Beeharry, N., Xia, B., Yen, T.J. and Zhang, R. (2011) Oncogenic RAS regulates BRIP1 expression to induce dissociation of BRCA1 from chromatin, inhibit DNA repair, and promote senescence. *Dev Cell*, **21**, 1077-1091.

16. Collado, M., Gil, J., Efeyan, A., Guerra, C., Schuhmacher, A.J., Barradas, M., Benguria, A., Zaballos, A., Flores, J.M., Barbacid, M. *et al*. (2005) Tumour biology: senescence in premalignant tumours. *Nature*, **436**, 642.

17. Michaloglou, C., Vredeveld, L.C., Soengas, M.S., Denoyelle, C., Kuilman, T., van der Horst, C.M., Majoor, D.M., Shay, J.W., Mooi, W.J. and Peeper, D.S. (2005) BRAFE600-associated senescence-like cell cycle arrest of human naevi. *Nature*, **436**, 720-724.

18. Halazonetis, T.D., Gorgoulis, V.G. and Bartek, J. (2008) An oncogene-induced DNA damage model for cancer development. *Science*, **319**, 1352-1355.

19. Olive, P.L. and Banath, J.P. (2006) The comet assay: a method to measure DNA damage in individual cells. *Nat Protoc*, **1**, 23-29.

20. Kastan, M.B., Onyekwere, O., Sidransky, D., Vogelstein, B. and Craig, R.W. (1991) Participation of p53 protein in the cellular response to DNA damage. *Cancer Res*, **51**, 6304-6311.

21. Bregeon, D. and Doetsch, P.W. (2004) Reliable method for generating double-stranded DNA vectors containing site-specific base modifications. *Biotechniques*, **37**, 760-762, 764, 766.

22. Morreall, J.F., Petrova, L. and Doetsch, P.W. (2013) Transcriptional mutagenesis and its potential roles in the etiology of cancer and bacterial antibiotic resistance. *J Cell Physiol*, **228**, 2257-2261.