DNA glycosylases as modulators of chemotherapeutic response

by

David Svilar

B.S. Biochemistry, Case Western Reserve University, 2006

Submitted to the Graduate Faculty of

University of Pittsburgh School of Medicine in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2012

UNIVERSITY OF PITTSBURGH

SCHOOL OF MEDICINE-MOLECULAR PHARMACOLOGY

This dissertation was presented

by

David Svilar

It was defended on

March 16th, 2012

and approved by

Chairperson: Richard A. Steinman, M.D., Ph.D., Associate Professor, Department of Medicine

Christopher Bakkenist, Ph.D., Assistant Professor, Department of Radiation Oncology

Hideho Okada, M.D., Ph.D., Associate Professor, Department of Neurological Surgery

Patrick Pagano, Professor, Department of Pharmacology and Chemical Biology

Dissertation Advisor: Robert W. Sobol, Ph.D., Associate Professor, Department of Pharmacology & Chemical Biology

DNA glycosylases as modulators of chemotherapeutic response

David Svilar, PhD

University of Pittsburgh, 2012

Glioblastoma multiforme (GBM) is the most common and aggressive primary brain tumor. Median survival is less than two years due to several factors, including challenges in surgical removal and chemotherapy resistance, underlining the need for more effective therapeutic options. To identify genes that contribute to chemotherapy resistance, we conducted a synthetic lethal screen in a chemotherapy-resistant GBM derived cell line (T98G) with the clinical alkylator temozolomide (TMZ) and an siRNA library tailored towards "druggable" targets. This screen for TMZ-sensitizing genes indicated that a subset of genes that were over-expressed in GBM cells increased the cell's sensitivity to TMZ when knocked down.

An ubiquitin ligase, UBE3B, and a DNA glycosylase, UNG, were among the TMZ-sensitizing genes identified using the siRNA library. We demonstrate that UBE3B and UNG are sensitizing genes in the screen validation studies using unique siRNA and shRNA sequences. Although UNG is one of four human DNA glycosylases that remove uracil lesions, UNG was the only uracil removing glycosylase to sensitize GBM cells in the validation studies. Notably, analysis of archived transcription datasets revealed that over-expression of UNG was correlated with poor outcomes in glioma patients.

In order to uncover functional groupings of TMZ-sensitizing proteins, we conducted *in situ* pathway analysis of gene candidates for synthetic lethal functions from our screen. This analysis discovered statistically significant enrichment of ontogeny clusters related to base

excision repair (BER), response to DNA damage, cellular proliferation and protein modification. Interestingly, this pathway topography overlapped with TMZ-sensitizing genes identified from similar experiments in yeast and bacteria.

In order to facilitate rapid *in vitro* identification of lesion-specific repair activity in cancer cells, we developed a novel fluorescent assay that extends the state of the art. The molecular beacon assay measures real-time DNA repair rates of specific DNA lesions by defined DNA repair proteins. These studies reveal that GBM up-regulates several TMZ-sensitizing genes that correlate with poor patient survival and inhibiting these genes may increase TMZ cytotoxicity in a tumor specific manner. These TMZ-sensitizing genes are not only potential targets for adjuvant therapy, but also represent potential biomarkers to predict TMZ response.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

Acknowledgements

I would like to thank the multitude of people who have made this journey possible and made the journey a lot more fun if I ever could have accomplished this alone. Using a chronological order for the plethora of people who have been unwavering stanchions in my life is easier than trying to justify who made the most important contributions, since there are so many and all of them have been life changing important interactions. First my family made a tremendous impact on my life and taught me many of the life skills I still employ today. Foremost for my scientific training they encouraged questions of the natural world and came up with some amazing explanations for a young impressionable child to understand. I do not know many children who get explanations on Rayleigh scattering to explain the color of the sky during the day and why it changes at night or about how when you cut a chocolate Easter bunny it undergoes similar cleavage, dimple rupture and decohesive rupture to metals when they are cut. We had a "Fabled Chalkboard" in the garage that was named by childhood friends because there would be explanations of all kinds on it, such as diagrams of how siphons work. My parents taught me not only to ask the why questions, but also how to answer them and learn more about myself and surroundings constantly. My siblings have had an undeniable impact on my life because I spent many of my formative years running around with them. They have undoubtedly

shaped my personality beyond my penchant for devouring food quickly. I attribute much of my determination and analytical skills from the many fights and arguments in the family.

The person who really pushed me off the edge and into the research hole was Dr. Irene Lee who stimulated my interest by foremost being incredibly supportive of a student who would not stop trying to poke holes in her explanations or models. Further, she was never afraid to admit that she did not know the answer or if anyone had even tried to answer the question. Her calm demeanor and acknowledgement that people still had many important questions to answer relating to basic biological mechanisms was refreshing as she encouraged me to learn more and potentially study methods to determine the answers. She also introduced me to Dr. Lawrence Sayre, my undergraduate mentor. He was an inspiration to work with as he always had time for an undergraduate student even though he chaired the Department of Chemistry and Department of Modern Languages and Literatures. He demanded scientific rigor and taught me that a properly controlled experiment may not give us the answer we wanted or predicted, but it will still yield useful data to learn about what actually occurred.

My current mentor, Dr. Robert Sobol, has been incredibly influential on my scientific training. He has stressed scientific rigor and been incredibly supportive. His kind words and time were invaluable to promote my efforts. He pushed me to expand my abilities, grow as a scientist and most importantly he cared about me. I am very thankful for all the help and knowledge he has given me. His passion for research is contagious and the sense of community in the lab makes it a fantastic learning environment. The lab and its members have been a source of help, pride and family away from home. They have been instrumental in helping me learn many techniques or help me on my projects. The time spent in lab is easier working among friends who believe in teamwork for the common good of the lab. I will miss the camaraderie of

the lab. My thesis committee has been instrumental in my completion of my PhD. They have offered many helpful suggestions and been incredibly supportive. Without their guidance I would not have a complete project.

Finally I would like to thank Shannon for her hard work and unfailing love. She has been the greatest supporter of me throughout. Her unwavering faith did not feel justified, but was much appreciated. She is my motivation and light for I would be lost without her. I cannot find words to thank her enough.

Abbreviations

GBM	Glioblastoma multiforme
BER	Base Excision Repair
p16	cyclin-dependent kinase inhibitor 2A
PTEN	Phosphatase and tensin homologue
EGFR	Epidermal growth factor receptor
p53	Tumor protein p53
PDGFR	Platelet derived growth factor receptor
TMZ	Temozolomide
Gy	Gray
MTIC	Monomethyl triazeno imidazole carboxamide
MMS	Methyl methanesulfonate
MNNG	methylnitronitrosoguanidine
N7-MeG	N7-methyl-guanine
N3-MeA	N3-methyl-adenine
O^6-Me-G	O^6-methyl-guanine

MGMT	O^6-methylguanine-DNA methyltransferase
MSH6	mutS homolog 6
MPG	methylpurine DNA glycosylase
APE1	APEX nuclease (multifunctional DNA repair enzyme) 1
PARP1	Poly (ADP-ribose) polymerase 1
XRCC1	X-ray repair cross complementing group 1
Polß	DNA polymerase ß
UNG	uracil DNA glycosylase
MUTYH	mutY homolog
8-oxoG	8-Oxo-7,8-dihydroguanine
SMUG1	Single-strand-selective monofunctional uracil-DNA glycosylase 1
MBD4	Methyl-CpG binding domain protein 4
NEIL1	nei endonuclease VIII-like 1 (*E. coli*)
TDG	Thymine DNA glycosylase
OGG1	8-oxoguanine DNA glycosylase
MYC	v-myc myelocytomatosis viral oncogene homolog
LSD1	lysine (K)-specific demethylase 1A
H_2O_2	Hydrogen Peroxide
AID	Activation-Induced Deaminase
DSB	double strand break
Ugi	uracil-DNA glycosylase inhibitor
CENP-A	centromere protein A
siRNA	short-interfering RNA

shRNA	short-hairpin RNA
miRNA	micro RNA
DROSHA	drosha, ribonuclease type III
DICER1	dicer 1, ribonuclease type III
RISC	RNA-Induced Silencing Complex
FBS	Fetal Bovine Serum
DMSO	dimethyl sulfoxide
DAVID	NIH DAVID
IPA	Ingenuity Pathway Analysis
GO	Gene Ontology
UBE3B	ubiquitin protein ligase E3B
ICMT	isoprenylcysteine carboxyl methyltransferase
B4GALT7	xylosylprotein beta 1,4-galactosyltransferase, polypeptide 7
CHRM3	cholinergic receptor, muscarinic 3
PADI1	peptidyl arginine deiminase, type I
ß-actin	actin, beta
PCNA	proliferating cell nuclear antigen
THF	tetrahydrofuran
Hx	Hypoxanthine
6-FAM	carboxyfluorescein
Dabcyl	4-(4’-dimethylaminophenylazo) benzoic acid
FRET	Förster Resonance Energy Transfer
EDTA	Ethylenediaminetetraacetic acid

DTT	Dithiothreitol
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
KOH	Potassium Hydroxide
KCl	Potassium Chloride
NAD^+	Nicotinamide adenine dinucleotide
NADH	Nicotinamide adenine dinucleotide hydride
$NADP^+$	Nicotinamide adenine dinucleotide phosphate
NADPH	Nicotinamide adenine dinucleotide phosphate hydride
-KD	Knockdown
alka	3-methyl-adenine DNA glycosylase II
MAG1	DNA-3-methyladenine glycosylase
IR	Ionizing Radiation
TCGA	The Cancer Genome Atlas
REMBRANDT	REpository for Molecular BRAin Neoplasia DaTa
FACS	Fluorescence Activated Cell Sorting
HDAC	histone deacetylase

1.0 INTRODUCTION

1.1 GLIOBLASTOMA MULTIFORME

1.1.1 Incidence, pathogenesis and prognosis

Cancer is the second leading cause of death in the United States, is the leading cause of death among men and women under the age of 85 years, the second most common cause of death in children (ages 1-14), and is one of the top five leading causes of death in any age group (1). The lifetime risk of developing an invasive cancer is approximately 44% for men and 38% for women (1). Glioblastoma multiforme (GBM) comprises over 22% of all brain tumors and is the most common and aggressive primary brain tumor (2, 3). The incidence of GBM tumors is approximately 2-4 cases per 100,000 people per year (2, 3). There are approximately 10,000 newly diagnosed GBM tumors annually in the United States. GBM is a grade four tumor as determined by the World Health Organization and confers a very poor prognosis, with the best therapy yielding a median survival of 14.6 months and a seven percent three year survival rate (2, 3). The peak age of diagnosis is between the ages of 45-75 with a slight male predominance (3 males: 2 females). GBM is diagnosed by a neuropathologist who rates several different criteria, including mitotic index, cellular appearance, and cellular differentiation that help to

determine tumor type and grade. Vascular (endothelial) proliferation and a necrotic tumor center are hallmark features of GBM and are required to differentiate between grade III and IV tumors.

Glioblastomas arise from either a primary glioblastoma or as secondary tumors from lower grade anaplastic astrocytoma. A lower grade anaplastic astrocytoma can evolve over several years to develop into a glioblastoma. The two mechanisms for glioblastoma development creates two different types of glioblastomas with primary glioblastomas having an older age of onset (median age of onset 64 years old), while glioblastomas derived from secondary to lower grade gliomas tend to affect a younger population (median age of onset 45 years old) (2). There are genetic differences between the two glioblastoma subtypes. Primary glioblastomas are characterized by a loss of heterozygosity of chromosome 10q, deletion of cyclin-dependent kinase inhibitor 2A (p16) and phosphatase and tensin homologue (PTEN) and amplification and mutations of epidermal growth factor receptor (EGFR) (2). This differs from secondary glioblastomas that typically affect younger patients and contain mutations to tumor protein p53 (p53), dysfunction of the retinoblastoma 1 (RB) and cyclin-dependent kinase inhibitor 2A (p16) pathways and over-expression of the platelet derived growth factor receptor (PDGFR) (2). The two subtypes share a similarity, in that both primary and secondary glioblastomas frequently lose heterozygosity of chromosome 10q (2).

1.1.2 Treatment overview

Temozolomide (TMZ), radiation and surgery are currently used for treatment of GBM, yet median survival is still less than 2 years, with a three year survival rate of 7% (3-5). Chemotherapy resistance and difficulties in surgical removal contribute to poor prognosis (6). If possible, maximal surgical debulking is recommended and improves patient symptoms by

relieving tumor mass effects. Further, the surgery allows placement of carmustine laced wafers which also improve survival and the surgical material removed can be used for staging, typing and research studies to improve future outcomes (2). Unfortunately, the highly infiltrative nature of glioblastomas preclude complete surgical removal, but the improvements from decreased mass effects still benefit the patient (2).

Radiotherapy is central to treating glioblastoma and has shown to increase survival from 3-4 months to 7-12 months (2). The radiation is fractionated up to 2 gray (Gy) doses for 5 days a week, up to a cumulative total dose of 60 Gy. Although radiation improves survival, approximately 90% of tumors recur at the original tumor site, suggesting that primary tumor cells survive and grow to form the recurrent tumor (2). Simultaneous treatment of TMZ with radiation improves median survival from 12.1 to 14.6 months compared to radiation alone (5). Further, adjuvant TMZ improved the two-year survival rate from 10.4% to 26.5% compared to radiation alone (5). Finally, the addition of adjuvant biodegradable carmustine wafers to the tumor location during surgical resection can improve median survival from 11.6 months to 13.9 months compared to placebo wafers (2). The basis of the wafers is to deliver a highly localized dose of carmustine directly to tumor cells that remain after surgical resection. This localized dose is released over several weeks as the wafers biodegrade.

However, treating the tumor is only a subsection of treating the patient as a whole, because the brain tumor affects the normal brain tissue. Therefore, there are many medical interventions necessary for patient management, which are not designed to decrease tumor burden. Some common complications include seizures, peritumeral edema, thromboembolism, and cognitive dysfunction (2). Patients may be treated with antiepileptic drugs, corticosteroids,

and low-molecular-weight heparin to combat these complications which directly impact their care and outcomes (2).

1.1.3 Temozolomide

Temozolomide (TMZ) is an alkylating agent and like all alkylating agents, TMZ can modify protein, lipids, RNA and DNA in cells (7). It was developed at Aston University by Malcolm Stevens and colleagues (8). TMZ is an imidazotetrazine pro-drug which undergoes spontaneous degradation at physiological pH to yield the intermediate, monomethyl triazeno imidazole carboxamide (MTIC). MTIC is also an intermediate formed by the chemotherapy dacarbazine. However, unlike TMZ, dacarbazine is converted to MTIC by Cyp450 enzymes in the liver (9). MTIC itself is not the clinically active compound for TMZ or dacarbazine. MTIC spontaneously breaks down to 4-amino-5-imadizole-carboxamide and a methyldiazonium ion, with the methyldiazonium ion being the active alkylating agent from both TMZ and dacarbazine (9). Thus TMZ undergoes two spontaneous chemical reactions to yield the active alkylating agent, the methyldiazonium ion. One benefit of the *in vivo* conversion process is the inherent hydrophobicity and relative stability of TMZ compared to the reactive diazonium ion allowing TMZ to distribute throughout the body yielding good bioavailability (8-10). TMZ easily penetrates the blood-brain barrier, with a concentration of 30-40% of peak plasma levels (10). TMZ is used as an oral medication and is stable in the stomach's low pH before being intestinally absorbed with peak plasma levels occurring ninety minutes after an oral dose (9, 10). TMZ is primarily excreted in urine with a smaller fraction fecally eliminated (11). The most common side effects involve the gastrointestinal tract and include nausea, vomiting, constipation

and diarrhea. Other common general side effects include fatigue and headache. The most prevalent serious adverse effect that may stop TMZ treatment is myelosupression (12).

Temozolomide's cytotoxicity is mediated through the DNA lesions created by TMZ. The main DNA lesions induced by TMZ are N7-methylguanine, N3-methyladenine and O^6-methylguanine (13). TMZ has a similar alkylation spectrum to the other alkylating agents used in these studies, including methyl methanesulfonate (MMS) and methylnitronitrosoguanidine (MNNG). Their lesion spectrum is impacted by both the electronegativity of the methyl acceptor and the mechanism of action of the methyl donor (14). The N7 atom of guanine is the most electronegative atom of a DNA base and is therefore the most modified base of electrophilic alkylating agents. All three alkylating agents used are electrophiles that react by donating a methyl group to an electron rich methyl acceptor. However, MMS's mechanism of action differs from TMZ and MNNG and because of this difference, yields a slightly different lesion spectrum. Both TMZ and MNNG modify DNA via a S_N1 mechanism (7, 14). The S_N1 reaction allows more modifications by the O^6 position of guanine compared to alkylators that use a S_N2 mechanism (7, 14). In contrast to TMZ and MNNG, MMS uses a S_N2 mechanism to alkylate DNA and similar to other alkylators with a S_N2 mechanism, MMS yields significantly lower O^6-methylguanine lesions compared to S_N1 type alkylators (7, 14). All three alkylators are used in this study and their most common DNA lesions are found below (**Table 1**) (13, 14).

Table 1. Lesion spectrum induced by the DNA alkylators used in this study

	DNA Lesion (%)		
Alkylator	N3-methyladenine	N7-methylguanine	O^6-methylguanine
TMZ (13)	9.2	>70	5
MMS (14)	10.4 - 11.3	81 – 83	0.3
MNNG (14)	12.0	67.0	7.0

1.1.4 Tumor resistance to therapies

Radiation, temozolomide and carmustine can damage protein, fats, RNA and DNA in cells, with the damaged DNA playing a major role in their cytotoxicity (7, 15). The tumor cells are preferentially killed compared to normal tissue due to their high replication rate. However, like normal cells, many tumors contain the ability to repair DNA damage and develop resistance to the therapies. Temozolomide and carmustine are alkylating agents that damage DNA by the addition of alkyl groups to DNA.

The main DNA lesions produced by TMZ are the N7-methylguanine (N7-MeG), N3-methyladenine (N3-MeA) and the O^6-methylguanine (O^6-MeG) (13). The N7-MeG and N3-MeA DNA lesions account for approximately 80% of the DNA lesions induced by TMZ, yet result in negligible clinical toxicity due to repair by the BER pathway (13). Virtually all of TMZ's clinical cytotoxicity is attributable to the O^6-methylguanine (O^6-MeG) lesion, which accounts for approximately 5% of TMZ induced lesions (13). The O^6-MeG lesion is repaired via a direct reversal mechanism by the protein O^6-methylguanine-DNA methyltransferase (MGMT), which transfers the O^6-methyl group from the guanine base onto a Cys residue in the MGMT

protein (16). If the O^6-MeG lesion is not removed by MGMT, during cellular replication the mis-pairing of O^6-MeG with thymine is detected by the mismatch repair (MMR) enzymes, triggering apoptosis signaling and cytotoxicity (13, 17, 18). However, 5-year survival rates still remain low in TMZ treated patients (4, 5), and TMZ resistance by recurrence of chemotherapy resistant tumors is common. Resistant cells can harbor mutations in mismatch repair proteins such as the mutS homolog 6 (MSH6) (19) or have elevated expression of MGMT (as in the T98G cell line) (20). Earlier endeavors to enhance TMZ efficacy by using MGMT inhibitors to prevent the repair of O^6-MeG lesions have not shown an increase in sensitivity or efficacy in clinical trials (21-23), especially in TMZ-resistant GBM (24, 25).

1.1.5 Base Excision Repair Overview

The N7-MeG and N3-MeA lesions are predominantly repaired by the BER pathway, a DNA repair mechanism that involves as many as 20 different proteins (26). BER proteins are responsible for the recognition and repair of small base lesions, abasic sites and single strand breaks (27). There are a myriad of different alkylation and especially oxidative DNA lesions recognized and removed by BER, and they may arise from endogenous or exogenous sources (15). Although other DNA repair mechanisms may have substrate overlap with some BER lesions, BER is thought to be the primary mechanism to remove small base lesions. While the small base lesions do not look as impressive as the large adducts or double strand breaks that are repaired by other DNA repair pathways, BER and its associated proteins are used to repair approximately 10,000 of the estimated 20,000 endogenous DNA lesions in a diploid mammalian cell every day (7, 27). This is an astounding number and reinforces the importance of BER in

maintenance of the mammalian genome, especially considering the several trillion cells in a human that encounter endogenous DNA damage.

Although BER is initiated by one of eleven lesion specific DNA glycosylases, there are several steps that are constant throughout any BER process (28). All BER processes involve lesion recognition, lesion removal, strand scission, gap tailoring, DNA synthesis, and DNA ligation (15). The proteins that fulfill these functions are dependent on the lesion type that partially dictates the specific protein that recognizes and removes the lesion. This is because downstream steps are determined by the specific DNA chemistries in the lesion and repair intermediates. The lesion alone does not always dictate what repair intermediates occur, since some lesions can be removed by multiple proteins yielding different products and thus different repair intermediates. The typical BER pathway for TMZ DNA lesions is initiated by methylpurine DNA glycosylase (MPG) which recognizes and removes the N7-MeG and N3-MeA lesions that are induced by TMZ (15). The resulting abasic site is a substrate for APEX nuclease (multifunctional DNA repair enzyme) 1 (APE1), which cleaves the sugar phosphate backbone, leaving a 3'OH and a 5'-deoxyribose-phosphate moiety (26). Poly (ADP-ribose) polymerase 1 (PARP1) recognizes and binds the nick or single strand break (SSB), which activates PARP1 to recruit downstream proteins such as X-ray repair cross complementing group 1 (XRCC1) and DNA polymerase ß (Polß) (26). Polß is recruited by XRCC1 and PARP1 to hydrolyze the 5'-deoxyribose-phosphate and fill the nucleotide gap, followed by ligation mediated by the XRCC1/DNA ligase III heterodimer or DNA ligase 1 (26, 29, 30).

1.1.5.1 DNA glycosylases

The eleven lesion specific DNA glycosylases initiate BER by recognizing and removing the DNA lesion. Many of the glycosylases are evolutionarily conserved and uracil DNA

glycosylase can even be found in some DNA viruses (31). The glycosylases can be classified into three different categories based on their lesion removal mechanism: mono-functional, bi-functional with associated β-elimination, and bi-functional with associated β,δ-elimination (15). The mono-functional glycosylases have only one function, to remove the base lesion but leave the DNA backbone intact. The bi-functional glycosylases differ in their ability to both remove the offending base lesion, but also to initiate strand cleavage by either a β-elimination or β,δ-elimination mechanism, thereby leaving different repair intermediates at the DNA gap (15). Although the strand scission by bi-functional glycosylases occur in biochemical experiments, the prevalence of the strand scission in live cell BER by DNA glycosylases is uncertain (32). This is because several bi-functional DNA glycosylases, including NTHL1 and OGG1, are inhibited because they bind tightly to their abasic product (32). The addition of APE1 increased the bi-functional glycosylase catalytic activity, but did so while generating predominantly 5'-lyase products, not 3'-lyase products (32). The bi-functional glycosylase would yield a 3'-lyase product, while APE1 strand cleavage results in a 5'-lyase product, indicating under physiological conditions of abundant APE1, some bi-functional glycosylases may act as mono-functional glycosylases (32).

The DNA glycosylases have a wide range of DNA substrates, from multiply oxidized DNA bases to the unique case of mutY homolog (MUTYH) which can remove the normal adenine across from an 8-oxoG DNA lesion (15). A more complete list of glycosylase names, subcellular localization, and substrates can be found below (**Table 2, Table 3, and Table 4**).

Table 2. Summary of human bi-functional DNA glycosylases (with associated β-elimination)

Gene Symbol	Gene Name	Gene ID	Uniprot Accession Number	Organelle expressed	Known Substrate[a]
OGG1	8-oxoguanine DNA glycosylase	4968	O15527	Nucleus and Mitochondria	8-oxoG:C/T/G; me-FapyG:C; FapyG:C (32) 8-oxoA:C (33) urea (34)
NTHL1 (NTH1)	Nth endonuclease III-like 1 (*E.coli*)	913	P78549	Nucleus and Mitochondria	T or C-glycol; FapyA (32) 5,6-dihydro-U:G/A (34) 5-formyl-U (34) 5,6-dihydroxy-C (34) 5,6-dihydro-T (34) urea (34) 5-OH-U:G (34) 5-OH-C:G>A (34) 5-hydroxy-5,6,-dihydro-T (34) 8-oxoG:G (35)
NEIL3	nei endonuclease VIII-like 3 (*E. coli*)	55247	Q8TAT5	Nucleus	spiroiminodihydantoin (Sp):C (36) guanidinohydantoin (Gh):C (36) FapyA (36) FapyG (36) 5-OH-U (36) 5-OH-C (36) Tg (36)

[a]Target base on left in mismatches.

Table 3. Summary of human monofunctional DNA glycosylases

Gene Symbol	Gene Name	Gene ID	Uniprot Accession Number	Organelle expressed	Known Substrate[a]
UNG	Uracil DNA glycosylase	7374	P13051	Nucleolus (UNG2) and Mitochondria (UNG1)	ssU; U:G; U:A; 5-fluorouracil (32) 5,6-dihydroxy-U:G (34) 5-OH-U:G (34, 37) Isodialuric acid; Alloxan (37)
SMUG1	Single-strand-selective monofunctional uracil-DNA glycosylase 1	23583	Q53HV7	Nucleolus	ssU; U:G; U:A (32) 5-chlorouracil:G, 5-fluorouracil:G, (38) 5-carboxyuracil:G (38) 5-hydroxyuracil; 5-formyl-U (34) 5-(hydroxymethyl)-U (34, 39)
TDG	Thymine DNA glycosylase	6996	Q13569	Nucleus	U:G; T:G; ethenoC:G (32) 5-fluorouracil; 5-fluorouracil (ss) (40) 5-hydroxymethyluracil; 5-bromouracil (40) hypoxanthine:G; εC:A (40) 5-formyl-U (34) Tg:G (41)
MBD4	Methyl-CpG binding domain protein 4	8930	O95243	Nucleus	U or T in U/TpG; 5-meCpG (32) 5-formyluracil; 5-(hydroxymethyl)-U (34) Tg:G (41)
MPG (AAG)	N-methyl DNA glycosylase	4350	P29372	Cytoplasm and Nucleus	3-meA; 7-meA; 3-meG; 7-meG; hypoxanthine; ethenoA; ethenoG (32) 1,N2-εG:C; U:G; ethanoadenine (42) 1-methylguanine; etheno-A(ss) (42) hypoxanthine(ss); ssU (42) 8-oxoG:C (Mouse) (43) cyanuric acid:CT>GA (34)
MUTYH (MYH)	mutY homolog (*E. coli*)	4595	Q9UIF7	Nucleus and Mitochondria	A:G; A:8-oxoG; C:A; 2-OH-A (32) 8-oxoA:G (44)

[a]Target base on left in mismatches.

Table 4. Summary of human bi-functional DNA glycosylases (with associated ß,δ-elimination)

Gene Symbol	Gene Name	Gene ID	Uniprot Accession Number	Organelle expressed	Known Substrate[a]
NEIL1	nei endonuclease VIII-like 1 (*E. coli*)	79661	Q96FI4	Nucleus, Cytoplasm and Mitochondria (Mouse)	TgG; 5-OH-C; 5-OH-U:AT>G 12(32) guanidinohydantoin (45) guanidinohydantoin (ss) (45) iminoallantoin (45) iminoallantoin (ss) (45) spiroiminodihydantoin (45) spiroiminodihydantoin (ss) (45) 5,6-dihydro-T (34) 5,6-dihydro-U:G/C/A>T (34) fapyG:C (34) 8-oxo-G:C/G>T>A (34) fapyA:T (34) (5'R)-8,5'-cyclo-2'-deoxyadenosine (46) (5'S)-8,5'-cyclo-2'-deoxyadenosine (46) 8-oxo-A:C (47)
NEIL2	nei endonuclease VIII-like 2 (*E. coli*)	252969	Q969S2	Nucleus and Cytoplasm	5-OH-U:G>T>A; 5-OH-C (32) 5,6-dihydro-U:G/A (34) 8-oxo-G:C/A (34) 5,6-dihydrothymine (34) guanidinohydantoin (45) guanidinohydantoin (ss) (45) iminoallantoin (45) iminoallantoin (ss) (45) spiroiminodihydantoin (ss) (45)

[a]Target base on left in mismatches.

The glycosylases are most well known for their role in initiating BER by recognizing and removing DNA lesions. However, recently there have been a series of reports on several glycosylases demonstrating essential functions of outside repair of DNA damage to prevent DNA mutations or cytotoxic DNA lesions. The glycosylases MBD4 and TDG have been implicated as having important roles in organism development by affecting DNA methylation, particularly 5-methyl-cytosine (48, 49). The mouse knockout of TDG was embryonic lethal, but likely not due to its role in BER from exogenous damage (48). The authors presented a model for TDG and other glycosylases, particularly SMUG1, to be responsible for DNA methylation status of 5-methyl-cytosine (48). The model predicts that 5-methyl-cytosine is present in nascent DNA and is specifically damaged by an oxidant-producing complex that includes the DNA glycosylase, TDG. This local damage is then recognized by the TDG present in the complex. The lesion is then removed and replaced by a normal cytosine (48). The specific types of damage to the 5-methyl-cytosine would dictate which glycosylases are required for methylation removal. This is not the first time that investigators have hypothesized local endogenous DNA damage is essential for normal DNA metabolism and function. Several reports have documented that the DNA glycosylase OGG1 is essential for transcriptional activation of v-myc myelocytomatosis viral oncogene homolog (MYC) and lysine (K)-specific demethylase 1A (KDM1A or LSD1) targeted proteins (50-52). The proposed OGG1-dependent model is similar to the model described for TDG removal of 5-methyl-cytosine from DNA. LSD1 demethylates histone H3 using its flavin containing catalytic site (50-52). Removal of the methyl marks generates H_2O_2 as a byproduct of the demethylation reaction (50-52). The locally produced H_2O_2 can damage the nearby DNA, creating oxidative DNA lesions. These oxidative DNA lesions are recognized by the DNA glycosylases and of particular importance are the lesions

recognized by OGG1. OGG1 primarily removes one of the most studied oxidative DNA lesions, 8-oxoguanine (8-oxodG) (15). Removal of the 8-oxodG lesion by OGG1 and APE1 dependent cleavage of the DNA backbone yields a single strand break in the DNA, with the cleavage of the DNA backbone removing the physical constraints on the DNA giving access to the transcriptional machinery and initiating transcriptional activation of LSD1 target genes (50, 51).

OGG1, TDG, and MBD4 are not the only DNA glycosylases with functions outside of BER. Uracil-DNA glycosylase (UNG) plays a critical role in class switch recombination and somatic hypermutation (53). Class switch recombination (also called isotype switching or immunoglobulin class switching) is the process by which a B-cell changes the constant version of the heavy chain (for example IgM to IgG), while keeping the same variable section of the heavy chain. The switching of the isotype/class of antibodies from IgM to IgG yields the process's name. Although not fully explained, UNG is essential for this process and UNG knockout mice do not undergo class switch recombination and only express IgM antibodies (53). It was originally thought that Activation-Induced Deaminase (AID), which is also required for this process, deaminates many cytosines on the DNA where the class switch is to occur (53). Deaminating cytosines yields 2'-deoxyuridine as the product base recognized and removed from the DNA by UNG. The resulting cluster of abasic sites are then hydrolyzed by APE1 leaving a double strand break (DSB). The DSB is a substrate for repair by non-homologous end joining, which completes the class switch recombination by fusing the new constant heavy chain to the old variable heavy chain.

This paradigm has been challenged by a series of elegant experiments demonstrating UNG's catalytic activity is not required for class switch recombination and instead an unknown function of the glycosylase is required for the event (54). To test if UNG's glycosylase function

is required the uracil-DNA glycosylase inhibitor (Ugi), a specific peptide inhibitor of UNG from the bacteriophage *Bacillus subtilis*, was used to prevent uracil removal and DNA binding of UNG in human B-cells (54, 55). These cells did have reduced ability to perform class switch recombination, but DNA cleavage was not prevented, as determined by the appearance of gamma-H2AX foci, suggesting an important aspect of UNG not related to its DNA binding and glycosylase activity is essential for the recombination (54). The most convincing experiments were a series of separation of function point mutants that contained no glycosylase activity, but still had robust class switch recombination (54, 56). There were several point mutants with undetectable catalytic activity, but still retained rates of class switch recombination comparable to wild-type protein (54, 56). Further there were separation of function mutants with some UNG mutants which maintained glycosylase activity, but had the same class switch recombination rates as knockout cells (56). Thus the N-terminal portion of UNG is essential for class switch recombination to fulfill an unknown function, while the glycosylase activity of UNG is dispensable, demonstrating an important purpose of UNG outside of BER (56).

Another glycosylase BER independent function of UNG was reported to promote changes in chromatin structure (57). The protein centromere protein A (CENP-A) is a histone 3 variant, which is an essential protein due to its role in mitotic segregation. CENP-A assembly was prevented by inhibiting UNG or reducing protein levels, while increasing UNG protein levels promoted CENP-A assembly (57). While reducing UNG protein levels is not a lethal event, because UNG knockout mice are viable, it diminishes cellular replication, potentially due to delay in CENP-A assembly (57). UNG also co-localizes to sites of DNA damage, gamma-H2AX, and CENP-A, which is not surprising because it is a DNA repair protein (47). It is intriguing to speculate that UNG can sense the DNA damage, since it has an affinity to damaged

DNA, and signal downstream repair via chromatin interactions. The increase in CENP-A assembly and chromatin modifications could be the requirement of UNG for class switch recombination as both processes are independent of glycosylase activity. The double strand break ends may require UNG dependent chromatin modification signaling to either signal for repair and recruitment of proteins or allow access of proteins to the recombination site. While the DNA glycosylases are most well known for their function in BER, it is possible that all the glycosylases have other unknown functions important for organism survival independent of BER and discovering these functions remains an area of great interest.

1.1.6 siRNA-Mediated Knockdown

The ability to specifically target genes with interfering RNA to mediate knockdown has drastically changed our scientific capabilities. There are several possible interfering RNA approaches such as short-interfering RNA (siRNA) or short-hairpin RNA (shRNA) that are exogenously introduced, or microRNA (miRNA) that are endogenously made by the cell. The different RNA interference approaches have different starting points, but eventually merge pathways to share a common mechanism with several conserved steps. For endogenously produced interfering RNAs, the initial pri-miRNA is transcribed before being processed by drosha, ribonuclease type III (DROSHA) to a shorter pre-miRNA in the nucleus. This pre-miRNA is then exported to the cytoplasm. Once in the cytoplasm the pre-miRNA steps merge with that of exogenously introduced shRNAs. The pre-miRNA or shRNA is then processed by dicer 1, ribonuclease type III (DICER1) to yield a twenty-one nucleotide double-stranded RNA that no longer contains a hairpin. After this step, the pre-miRNA and shRNA steps merge with that of exogenously introduced siRNAs. The subsequent siRNA is then incorporated into the

RNA-Induced Silencing Complex (RISC), which incorporates one of the siRNA strands into the complex. Using the single stranded RNA, the RISC searches for complementary mRNAs to either inhibit their translation or to degrade the mRNAs. An mRNA is targeted for degradation if it contains perfect complimentary sequence, while only partial complementation of the siRNA to the mRNA will inhibit the mRNA translation without being degraded. Both mRNA degradation and inhibition of translation will decrease protein levels of the targeted gene. Different mRNAs can be targeted by varying the siRNA sequence. By using an siRNA containing 21 nucleotides and having 4 possible bases at each site yields a possible 4^{21} targeting sequences. Although lack of perfect complimentary sequence still inhibits translation, it is still possible to create siRNAs targeting specific genes to limit off target possibilities and effects due to sequence specificity created by the siRNA sequence. Thus large siRNA libraries targeting the entire genome can be designed and used to query if knockdown of specific genes impact a studied response. To ensure off-target effects are not responsible for the phenotypic effects of siRNA-mediated knockdown, it is common to use at least two different sequences targeting the same gene and to generate the same phenotype. The likelihood that two different siRNA sequences contain the same off-target gene to generate the desired phenotype is small. Another method to eliminate studying siRNA off-target effects is to rescue the siRNA knockdown cells by expressing an siRNA resistant transcript to express the protein of interest. This should diminish the phenotype observed by knockdown of the protein and demonstrate that specific knockdown of the gene is causing the studied phenotype.

2.0 MATERIAL AND METHODS

2.1.1 Chemicals and Reagents

DharmaFECT 2 transfection reagent and the siGENOME Non-Targeting siRNA #1 were from Dharmacon (Lafayette, CO). CellTiter-Blue Cell Viability Assay and CellTiter 96 AQueous One Solution Cell Proliferation Assay were from Promega (Madison, WI). The Silencer Druggable Genome siRNA Library (Version 1.1) and 5x siRNA resuspension buffer were from Ambion (Austin, TX). Tissue culture-treated 384-well microtiter plates were from Greiner Bio-One (GmbH, Frickenhausen, Germany). OptiMEM, EMEM, phosphate buffered saline (PBS) and Hoechst 33342 were from InVitrogen (Carlsbad, CA). The shRNA vectors used for stable KD cell creation were obtained in glycerol stocks from Sigma-Aldrich. Temozolomide was from National Cancer Institute Developmental Therapeutics Program (Bethesda, MD). Puromycin was from Clontech Laboratories (Mountain View, CA) and gentamicin was obtained from Irvine Scientific (Santa Ana, CA). We used the following primary antibodies: UNG antibody (#MBS200056) was from MyBioSource (San Diego, CA) and PCNA antibody (#sc-56) was from Santa Cruz (Santa Cruz, CA). Secondary antibodies: GAM-HRP conjugates were from Bio-Rad (Hercules, CA). Signal generation substrates were from Bio-Rad (Hercules, CA) and Thermo Fisher Scientific (Waltham, MA). All electrophoresis reagents were from Bio-Rad (Hercules, CA).

2.1.2 Cell lines and culture conditions

Cell line and culture conditions were as previously described (58). Briefly, T98G cells were cultured in EMEM with 10% heat-inactivated FBS, non-essential amino acids, sodium pyruvate, antibiotic/antimycotic and gentamicin. The LN428 cells were cultured in alpha MEM (InVitrogen, Carlsbad, CA), 10% fetal bovine serum (Atlanta Biologics, Lawrenceville, GA), antibiotic/antimycotic (Sigma), gentamicin, L-glutamine (Sigma). The LN428/XRCC1-KD and the LN428/GFP glioblastoma cell lines were cultured in the same media as the LN428 cells described above supplemented with 1 µg/mL puromycin (Sigma). The LN428/MPG cell line (LN428 cells modified for elevated expression of MPG) was cultured in the same media as the LN428 cells and supplemented with 600 µg/mL Geneticin (InVitrogen) and the LN428/MPG/XRCC1-KD glioblastoma cell line (LN428 cells modified for elevated expression of MPG and loss of XRCC1 expression) was cultured in the same media as the LN428 cells and supplemented with 600 µg/mL Geneticin (InVitrogen) plus 1 µg/mL puromycin (Sigma). The HCT116 and SAOS-2 cells were cultured in McCoy's 5A media (CellGro #10-050-CV) supplemented with 10% FBS and penicillin/streptomycin. The MDA-MB-231 cells were cultured in RPMI 1640 (BW #12-702F) supplemented with 10% FBS and gentamicin. The MCF-7 cells were cultured in EMEM (Invitrogen #11095-080) supplemented with 10% FBS, sodium pyruvate (Invitrogen), insulin (Invitrogen) and non-essential amino acids (Invitrogen). The U2OS cells were cultured in DMEM (CellGro# 10-013-CV) supplemented with 5% FBS, penicillin/streptomycin and L-glutamine. The A-172, M059J and M059K cells were cultured in D-MEM-glucose (Invitrogen #11960-044) with 10% FBS, penicillin/streptomycin and Glutamax (Invitrogen# 35050-061). The DBTRG5-MG cells were cultured in RPMI (CellGro # MT10-

040-CV) supplemented with 10% heat-inactivated FBS, glutamine, penicillin/streptomycin, Hepes buffer solutions (Invitrogen # 15630-080) and non-essential amino acids.

2.1.3 Screen siRNA transient transfection and temozolomide treatment

Transient transfection by siRNA was as previously described (58, 59). Briefly, T98G cells were wet-reverse transfected with the Ambion Silencer Druggable Genome siRNA library. The library consisted of siRNAs targeting 5,520 different genes. Three unique siRNA duplexes targeting the same gene were pooled into a single well. Therefore, the siRNA knockdown targeted one gene in each well. The siRNA were prepared with DharmaFECT2 and OptiMEM and split evenly into two, 384-well plates. T98G cells were added directly to the siRNA complexes. The plates were incubated for five hours at 37ºC with 5% CO_2 before the media was removed and replaced with fresh complete media.

Cells were incubated at 37ºC with 5% CO_2 for 48 hr after siRNA transfection in a humidified incubator to allow for gene silencing before addition of TMZ or vehicle. After incubation, the media was removed from the cells and replaced with fresh media containing either dimethyl sulfoxide (DMSO) vehicle (final concentration 1.0%) or the approximate EC_{10} of TMZ (final concentration 1 mM in 1% DMSO). Cells were incubated at 37ºC with 5% CO_2 in the presence of TMZ for 48 hrs in a humidified incubator. Cell viability was measured 96 hours after siRNA transfection with the CellTiter-Blue viability assay according to manufacturer's protocol.

2.1.4 Data analysis

The siRNA screen was performed 3 times over 6 separate weeks. Fluorescence units from each well were normalized to plate negative controls (scrambled siRNA) enabling cell viability comparisons between different plates. Data was analyzed by creating a high confidence hit list as described previously (60). Briefly, cell viabilities for each targeting gene from the three screening replicates were averaged and data was analyzed using two statistical analysis methods. For each gene, a two-sample t-test was performed to determine if there was a sensitization effect on cellular survival between 'siRNA plus vehicle' treated cells as compared to 'siRNA plus TMZ' treated cells. We selected targeting siRNAs with a p-value less than or equal to this threshold.

Also, viability ratios (ratio of survival of 'siRNA plus TMZ' to 'siRNA plus vehicle') were calculated for each targeting siRNA to determine the magnitude of response between the two screening conditions (**Equation 1**).

Equation 1: ***(siRNA+TMZ)/(siRNA+vehicle) = viability ratio***

The viability ratios were sorted in descending order and targeting siRNAs with a viability ratio in the lowest 5th percentile were selected for further analysis. The targeting siRNAs identified by both methods (p-value ≤ 0.05 and viability ratio in the lowest 5th percentile) were included in the final high-confidence hit list for further analysis and validation. These methods produced an overlap of approximately 10%.

2.1.5 Pathway analysis and network visualization

The hit list created from the screening data was further analyzed for pathway enrichment and network visualization. Several different programs were used to analyze and visualize the data including NIH DAVID (DAVID), Princeton GO term finder and Ingenuity Pathway Analysis (IPA). DAVID and Princeton GO term finder both classify genes based on their gene ontology (GO) (61-63). Each gene can have multiple classifications that are dependent on the gene's molecular function, biological process, or cellular compartment and are referred to as "GO terms" (62). These classifications enable one to determine if certain GO terms are over represented or enriched in the hit list (61, 62). Groups of genes with the associated GO terms that are enriched are more likely to be true hits because multiple genes affecting the same process all yield sensitization and lend insight into which pathways are most affected since many genes with similar functions impact response (61, 64). The gene enrichment calculations used the 5,520 genes screened as the background, due to the relatively small number of genes screened, instead of the entire human genome. The analysis with NIH DAVID was performed using multiple classification stringencies, gene ontology hierarchies, protein interaction databases and pathways to obtain insight into the biological pathways involved in TMZ resistance. The hit list was also analyzed with IPA (Ingenuity® Systems, www.ingenuity.com), utilizing a proprietary, manually curated, interaction database. The Functional Analysis identified the biological functions that were most significant to the data set. Right-tailed Fisher's exact test was used to calculate a p-value determining the probability that each biological function assigned to that data set is due to chance alone. Because IPA utilizes a different scheme and algorithm to organize proteins into networks not based on GO terms, it was used as a separate method to analyze the hit list for enriched protein functions and networks. Fisher's exact test was used with $\alpha = 0.05$ to compute

the probability of correct functional assignment for the genes in the hit list. The enriched networks created by IPA were visualized with genes serving as nodes and edges representing known interactions. Genes colored green modulate toxicity to alkylating agents as determined in the screen.

2.1.6 Generation of lentiviral single knockdown cell lines

The shuttle vectors for expression of shRNA were from Sigma (St. Louis, MO). Lentiviruses were prepared in collaboration with the UPCI Lentiviral facility. Lentiviral particles were generated by co-transfection of 4 plasmids [the shuttle vector plus three packaging plasmids: pMD2.g(VSVG), pVSV-REV and PMDLg/pRRE] into 293-FT cells (65, 66) using FuGene 6 Transfection Reagent (Roche, Indianapolis, IN), as described previously (67). Forty-eight hours after transfection, lentivirus-containing supernatant was collected and passed through 0.45μM filters to isolate the viral particles. Lentiviral transduction was performed as described earlier (67). Briefly, 6.0×10^4 cells were seeded into a 6-well plate 24 hours before transduction. Cells were transduced for 18 hours at 32˚C and then cultured for 72 hours at 37˚C. Cells were then selected by culturing in growth media with 1.0 μg/mL puromycin, as previously described (67).

2.1.7 Generation of lentiviral double knockdown cell lines

The dual targeting of UNG and UBE3B by different lentiviral shRNA was done by first creating a stable UBE3B knockdown cell lines as described above. This cell line was then transduced using the above procedure, with a lentivirus expressing GFP or a lentivirus

expressing GFP and shRNA against UNG. The cells were incubated for 4 days after transduction before selecting for cells expressing GFP or GFP with UNG shRNA using Fluorescence Activated Cell Sorting (FACS). The isolated GFP-expressing cells were incubated at 37°C for 48 hours before being analyzed for UBE3B and UNG mRNA expression.

To create double knockdowns using UBE3B knockdown single cell clones, the population of the UBE3B knockdown cell line was single cell cloned by limiting dilution. UBE3B knockdown cells were seeded at a concentration of 3 cells per mL of media into 96 well plates at a volume of 100 µL of cell solution per well. The plates were incubated at 37°C for a week before single cell colonies were determined by visual inspection. Single cell colonies were expanded before quantifying UBE3B mRNA levels. Clones with low levels of UBE3B were transduced with lentivirus GPF control or GFP and UNG shRNA virus as described above. Cells were FACS sorted for GFP expression and incubated for 48 hours before quantification of UNG and UBE3B mRNA levels.

2.1.8 Quantitative RT-PCR analysis

Expression of mRNA for each of ten glycosylases and five protein modification genes (OGG1, SMUG1, MBD4, UNG, MYH, NTHL1, MPG, NEIL1, NEIL2, NEIL3, UBE3B, ICMT, B4GALT7, CHRM3, and PADI1) after shRNA-mediated knockdown was measured by quantitative RT-PCR (qRT-PCR) using an Applied Biosystems StepOnePlus system as described previously (67). Applied Biosystems TaqMan® Gene Expression Assays used are as follows: human OGG1: Hs00213454_m1; human SMUG1: Hs00204820_m1; human MBD4: Hs00187498_m1; human UNG: Hs00422172_m1; human MYH: Hs01014856_m1; human NTHL1: Hs00267385; human MPG: Hs01012594_m1; human NEIL1: Hs0022637_m1; human

NEIL2: Hs00376746_m1; human NEIL3: Hs00217387_m1; human UBE3B: Hs00296200_m1; Human ICMT: Hs00202655_m1; Human B4GALT7: Hs01011258_m1; Human CHRM3: Hs00265216_s1; Human PADI1: Hs00203458_m1. Gene expression of each gene was normalized to the expression of human ß-actin (part #4333762T). When comparing qRT-PCR data across multiple cell lines, the mRNA was normalized to T98G or T98G/GFP mRNA expression.

2.1.9 Cell extract preparation and immunoblot

Nuclear extracts were prepared and protein concentrations were determined as described previously (67). Fifteen micrograms of protein was loaded on a pre-cast 4-20% Tris-Glycine gel (InVitrogen, Carlsbad, CA). The following primary antibodies were used in immunoblot assays: anti-human UNG from MyBioSource.com (#MBS200056), anti-human MPG (Mab; clone 506-3D) (67), anti-XRCC1 (Bethyl Labs; Montgomery, TX) and anti-human proliferating cell nuclear antigen (PCNA; Santa Cruz Biotechnology). Immunoblot membranes were stripped before re-probing for PCNA. The membranes were stripped with Restore PLUS Western Blot Stripping buffer (#46430) from Thermo Fisher Scientific (Waltham, MA) per the manufacturer's instructions. Immunoblots quantified by densitometry analysis were scanned and quantified using NIH ImageJ and the associated analysis software package, normalizing the expression across the four cell lines to the LN428/MPG cell line and to the expression of PCNA within each sample.

2.1.10 DNA glycosylase molecular beacon activity assay and design

All oligodeoxyribonucleotides in the molecular beacon assay were purchased from Integrated DNA Technologies (Coralville, USA) as shown below (**Table 5**, **Table 6, Table 7**).

Table 5. Initial Design of Molecular Beacons

Name	3' Modifier	Sequence	Modified Base	5' Modifier	Target
FD-Con	6-FAM	GCACTATTGAATTGACACGCCATG TCGATCAATTCAATAGTGC	-	Dabcyl	Control Oligo
FD-UNG1	6-FAM	GCACT**X**AAGAATTGACACGCCATG TCGATCAATTCTT**A**AGTGC	2'-deoxyUridine	Dabcyl	UNG
FD-MPG1	6-FAM	GCACT**X**TTGAATTGACACGCCATG TCGATCAATTCAA**T**AGTGC	Ethenoadenine (εA)	Dabcyl	MPG
FD-MPG2	6-FAM	GCACT**X**TTGAATTGACACGCCATG TCGATCAATTCAA**T**AGTGC	Hypoxanthine (Inosine)	Dabcyl	MPG
FD-THF	6-FAM	GCACT**X**TTGAATTGACACGCCATG TCGATCAATTCAA**T**AGTGC	THF, tetrahydrofuran (Abasic site mimic)	Dabcyl	APE1

6-FAM is carboxyfluorescein and Dabcyl is 4-(4'-dimethylaminophenylazo) benzoic acid. Target base is labeled **bold and underlined** and base opposite target base is labeled ***with bold and italics***. MYH removes A opposite 8oxoG while OGG1 removes the 8oxoG lesion.

Table 6. Molecular beacons designed for measuring DNA repair rates *in vivo*.

Name	3' Modifier	Sequence	Modified Base	5' Modifier	Target
FD-Con3-Cy5	6-FAM	CCACTATTGAATTGACACGCC/Cy5 /ATGTCGATCAATTCAATAGTGG	-, also contains a Cy5 in hairpin	Dabcyl	Control Oligo
FD-dU/A-Cy5	6-FAM	CCACT**X**TTGAATTGACACGCC/Cy5 /ATGTCGATCAATTCAA**A**AGTGG	2'-deoxyUridine, also contains a Cy5 in hairpin	Dabcyl	UNG, SMUG1

6-FAM is carboxyfluorescein and Dabcyl is 4-(4'-dimethylaminophenylazo) benzoic acid. The target base is **bold** while the base opposite the target base is **bold and underlined**. MYH removes A opposite 8oxoG while OGG1 removes the 8oxoG lesion.

Table 7. Improved Molecular Beacon design and structure.

Name	3' Modifier	Sequence	Modified Base	5' Modifier	Target
FD-Con2	6-FAM	CCACTATTGAATTGACACGCCATGT CGATCAATTCAATAGTGG	-	Dabcyl	Control Oligo
FD-THF2	6-FAM	CCACT**X**TTGAATTGACACGCCATGT CGATCAATTCAA**T**AGTGG	THF, tetrahydrofuran (Abasic site mimic); opposite a 'T'	Dabcyl	APE1
FD-THF2/G	6-FAM	CCACT**X**TTGAATTGACACGCCATGT CGATCAATTCAA**G**AGTGG	THF, tetrahydrofuran (Abasic site mimic); opposite a 'G'	Dabcyl	APE1
FD-THF2/A	6-FAM	CCACT**X**TTGAATTGACACGCCATGT CGATCAATTCAA**A**AGTGG	THF, tetrahydrofuran (Abasic site mimic); opposite a 'A'	Dabcyl	APE1
FD-THF2/C	6-FAM	CCACT**X**TTGAATTGACACGCCATGT CGATCAATTCAA**C**AGTGG	THF, tetrahydrofuran (Abasic site mimic); opposite a C'	Dabcyl	APE1
FD-8oxoG/C	6-FAM	CCACT**X**TTGAATTGACACGCCATGT CGATCAATTCAA**C**AGTGG	8-oxoguanine (8-oxo-7,8-dihydro-2'-deoxyguanosine); opposite C	Dabcyl	OGG1
FD-8oxoG/A	6-FAM	CCACT**X**TTGAATTGACACGCCATGT CGATCAATTCAA**A**AGTGG	8-oxoguanine (8-oxo-7,8-dihydro-2'-deoxyguanosine); opposite A	Dabcyl	OGG1 and MYH
FD-C/8oxoG	6-FAM	CCACT**C**TTGAATTGACACGCCATGT CGATCAATTCAA**X**AGTGG	C opposite 8-oxoguanine (8-oxo-7,8-dihydro-2'-deoxyguanosine)	Dabcyl	OGG1
FD-A/8oxoG	6-FAM	CCACT**A**TTGAATTGACACGCCATGT CGATCAATTCAA**X**AGTGG	A opposite 8-oxoguanine (8-oxo-7,8-dihydro-2'-deoxyguanosine)	Dabcyl	OGG1 and MYH
FD-Tg/A	6-FAM	CCACT**X**TTGAATTGACACGCCATGT CGATCAATTCAA**A**AGTGG	Thymine glycol (5,6-dihydroxy-5,6-dihydrothymine); opposite A	Dabcyl	NEIL1
FD-5FU	6-FAM	CCACT**X**TTGAATTGACACGCCATGT CGATCAATTCAA**A**AGTGG	5-fluoro-uracil (5FU)	Dabcyl	UNG,TDG, SMUG1
FD-5HMDU	6-FAM	CCACT**X**TTGAATTGACACGCCATGT CGATCAATTCAA**A**AGTGG	5-hydroxymethyl-2'-deoxyuridine	Dabcyl	SMUG1
FD-Hx	6-FAM	CCACT**X**TTGAATTGACACGCCATGT CGATCAATTCAA**T**AGTGG	Hypoxanthine (Inosine)	Dabcyl	MPG
FD-dU/A	6-FAM	CCACT**X**TTGAATTGACACGCCATGT CGATCAATTCAA**A**AGTGG	2'-deoxyUridine	Dabcyl	UNG, SMUG1

6-FAM is carboxyfluorescein and Dabcyl is 4-(4'-dimethylaminophenylazo) benzoic acid. Target base is labeled **bold and underlined** and base opposite target base is labeled ***with bold and italics***. MYH removes A opposite 8oxoG while OGG1 removes the 8oxoG lesion.

The molecular beacon design underwent several changes during optimization and testing. The basic design was similar, including the carboxyfluorescein (6-FAM) fluorophore and 4-(4'-dimethylaminophenylazo) benzoic acid (Dabcyl) quencher used. All the molecular beacons are designed to form a stem-loop structure containing a 13-nucleotide loop and a 15 base pair stem forcing the 6-FAM fluorophore (at the 5' end) and the Dabcyl quencher (at the 3' end) into close proximity. The close proximity of 6-FAM and Dabcyl enables quenching of the 6-FAM signal in a non-fluorescent manner via Förster Resonance Energy Transfer (FRET) (68, 69). However, if the base lesion (in **Figure 1** this is ethenoadenine) is removed by a DNA glycosylase (in **Figure 1** this is MPG) and the DNA backbone is hydrolyzed by APE1, the 6-FAM containing oligonucleotide (5 bases in length) will dissociate from the hairpin at 37˚C (**Figure 1**). The 6-FAM dissociation from the DNA hairpin prevents the quenching by Dabcyl. The increase in 6-FAM-mediated fluorescence is proportional to base lesion (e.g. ethenoadenine) removal. Any increase in fluorescence in control beacon containing a normal base is the result of non-specific DNA backbone cleavage.

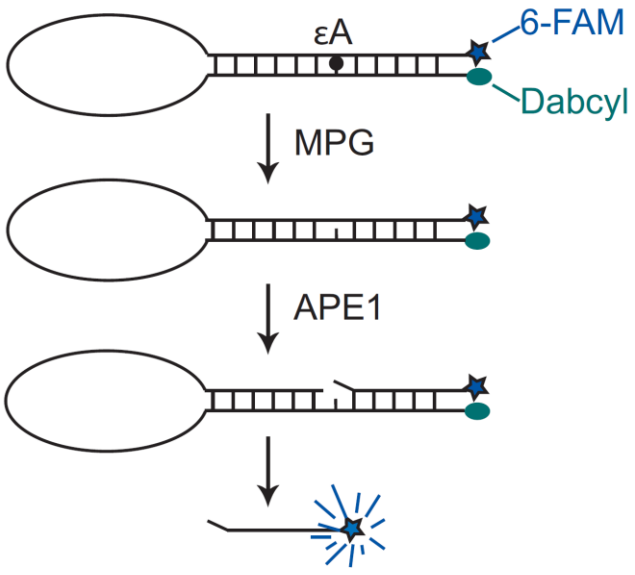


Figure 1. Model of Molecular Beacon mechanism.

The molecular beacon forms a DNA hairpin bringing the 6-FAM fluorophore near the Dabcyl quencher, enabling efficient quenching of 6-FAM. The DNA lesion can be recognized and removed by a DNA glycosylase leaving an abasic site. The abasic site is a substrate for APE1 that cleaves the DNA backbone enabling dissociation of the 6-FAM fluorophore containing oligonucleotide from the other DNA strand. The increase in distance of 6-FAM from Dabcyl eliminates the quenching of Dabcyl. The increase in 6-FAM fluorescence is measured and is proportional to the amount of DNA lesion removed by the DNA glycosylase.

To ensure that the beacons correctly adopted a stem-loop structure, each was incubated at 95°C for 3 min. The beacons were removed from the heat and allowed to slowly cool overnight to room temperature in an insulated container. Once the hairpin was formed, minimal measurable fluorescence was detected (**Figure 2**) and the hairpin was stable at 37°C for greater than 120 min. However, when heated, the hairpin unfolds, resulting in maximum fluorescence intensity (**Figure 3**). Nuclear protein extracts were prepared as described above.

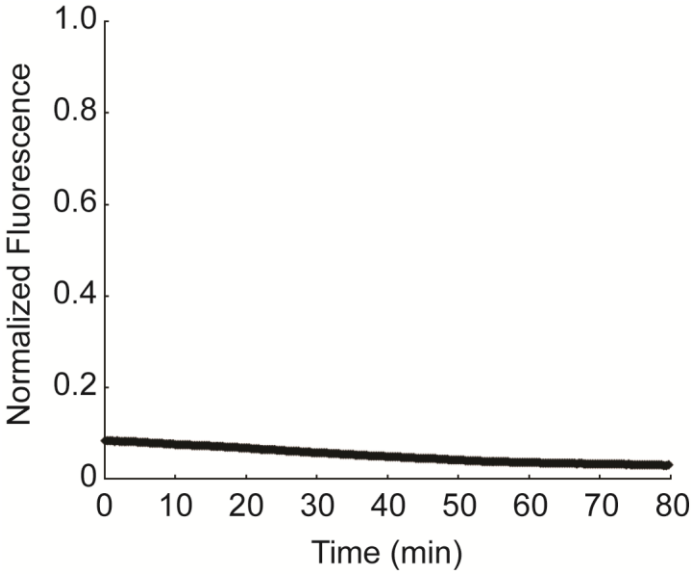


Figure 2. Background normalized fluorescence values for molecular beacons stay constant.
After annealing the beacons overnight, we incubated the beacons in the absence of lysate at 37°C and graphed normalized fluorescence values. The fluorescence values vary little during the duration of the experiment, demonstrating the stability of both the DNA structure and the dye/quencher pair contained in the beacons at 37°C. There is some photobleaching at later time points, but this is expected when taking rapid fluorescent measurements and why we subtract out this background from the experimental wells.

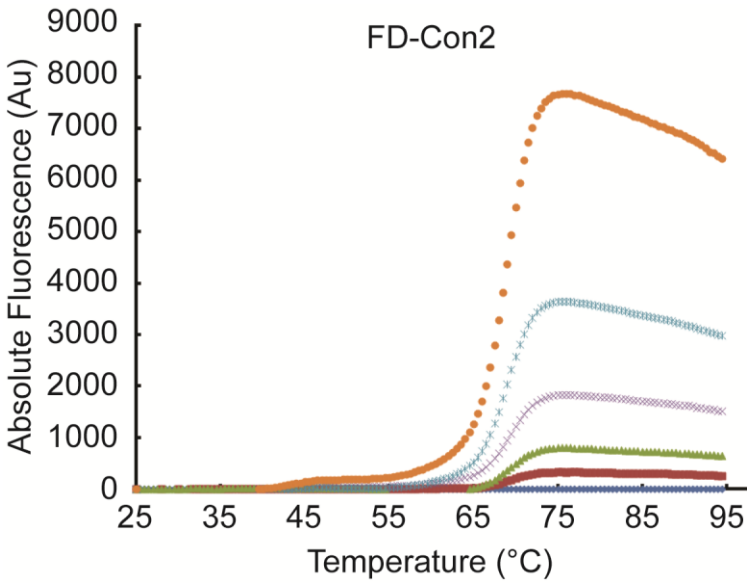


Figure 3. Melt curve data for FD-Con2 molecular beacon.
After annealing the beacons overnight, we performed at melt curve on the beacons to ensure they adopted the same secondary structure. The melt curve was done by heating various concentrations of the beacons in 0.5°C increments and measuring 6-FAM fluorescence at each 0.5°C step. The 6-FAM was measured because as the beacon is heated the DNA will dissociate from its stem-loop structure, increasing the distance between the 6-FAM fluorophore and Dabcyl quencher yielding an increase in fluorescence. The maximum fluorescence is obtained when all the beacons are no longer in a stem-loop structure. The beacons have a temperature dependent decrease in fluorescence. This is seen after 75°C for this beacon and is believed to be from the high heats affecting the dye, not from stem-loop formation at higher temperatures.

Approximately 500 µL of nuclear protein extracts were dialyzed twice using the 7,000 molecular weight cut-off Slide-A-Lyzer Dialysis Cassette (Pierce; Rockford, IL). The samples were dialyzed for 90 min at 4°C in the following buffer: 50 mM Hepes pH7.5, 100 mM KCl, 0.5 mM EDTA, 20% Glycerol and 1 mM DTT. The reaction solution, samples and beacons were chilled to 4°C and mixed on ice to prevent initiation of the reaction before beginning to record data. Reactions were performed at 37°C using 10 µg of dialyzed protein extract and beacon substrate (final conc. = 40 nM) in the following buffer: 25 mM HEPES-KOH pH7.8, 150 mM KCl, 0.5 mM EDTA, 1% Glycerol, 0.5 mM DTT. For initial experiments, fluorescence was measured every 20 seconds for 60 minutes, using a StepOnePlus real-time PCR system and expressed as arbitrary units (AU). When normalizing the molecular beacons, several steps were added. After measuring fluorescence every 20 seconds for 60, 90 or 120 min at 37°C to

determine DNA repair rate, the beacons were incubated sequentially at 60°C, 65°C, 70°, 75°C, 80°C for 5 min at each temperature. The fluorescence was measured every 20 sec at each step to determine the maximum possible fluorescence for each beacon in each individual well. The temperatures selected for these measurements encompassed the maximum fluorescence possible for each beacon as determined by previous melting assay experiments.

2.1.11 Molecular Beacon Data Analysis

2.1.11.1 Initial Analysis Methodology

The fluorescence data were analyzed to enable comparisons across cell lines and for comparison of control and lesion-containing BER beacons. We eliminated the background fluorescence due to incubation of the beacon alone by subtracting the fluorescence values of a control well containing no protein extract from all wells using that molecular beacon. An example of the background fluorescence values is seen in **Figure 2**. To enable comparisons across different cell lines, molecular beacons, and trials, we selected the fluorescence value of the 5-min time point as the zero value for each well. We subtracted this value from all other time points in that well so all graphs begin from zero AU and 5 min after initiating the reaction. Five minutes was selected as the point from which to begin comparisons, because time points earlier than 4 min contained variations in absolute fluorescence measurements independent of the molecular beacon and cell line (not shown). Five minutes was selected to eliminate the variable measurements and to facilitate valid comparisons between trials and conditions. The mean of 3 separate trials was plotted, with error bars representing the standard error of the mean.

2.1.11.2 Molecular Beacon Analysis by normalizing each well

After measuring fluorescence every 20 seconds for 60, 90 or 120 min to determine DNA repair rates, the beacons were incubated sequentially at 60°C, 65°C, 70°, 75°C, 80°C for 5 min at each temperature. The average fluorescence at each temperature was determined. The maximum average fluorescence intensity in each well was used to normalize the beacon fluorescence in each corresponding well to account for pipetting error, well-to-well variability of measurements in the machine and machine-to-machine variability in fluorescence measurements. Under the likely assumption that maximum fluorescence intensity corresponds to the maximal possible fluorescence value when fully incised, these normalized data represent % free FAM (= % BER incised beacon). The normalized fluorescence values were plotted as the mean of three independent experiments with error bars representing the SEM.

2.1.12 Cell cytotoxicity assays

2.1.12.1 Short-term MTS assay

TMZ induced cytotoxicity was determined as described previously (12). Briefly, cells were seeded into 96-well plates at a density of 2,000 cells per well. Cells were treated with TMZ for 48 hours at 37˚C before determining the relative amount of metabolically active cells by an MTS assay. Results were the average of three separate experiments and normalized to vehicle treated control cells with error bars representing the standard error of the mean.

MMS induced cytotoxicity was determined essentially as previously described (70). Briefly, cells were seeded into a 96-well plate at a density of 2,000 cells per well. Cells were treated with MMS for 4 hours at 37°C. MMS containing media was then removed and replaced with fresh media. Cells were incubated for a total of 48 hours at 37˚C before determining the

relative amount of metabolically active cells by an MTS assay. Results were the average of three separate experiments and normalized to vehicle treated control cells with error bars representing the standard error of the mean.

2.1.12.2 Long-term CyQuant assay

Cells were grown until approximately 50-75% confluence before being trypsinized and counted using a CASY counter per the manufacturer's instructions. Cells were seeded into 96-well plates at a density of 125cells/well and incubated at 37˚C for 24 hrs. The cells were treated with vehicle, TMZ or MNNG and incubated for nine days at 37˚C. Plates were removed and fluorescence was determined using the CyQuant kit (InVitrogen, #C7026) following the manufacturer's instructions. Results were the average of two separate experiments and normalized to vehicle treated control cells with error bars representing the standard error of the mean.

2.1.13 Cross species analysis of alkylation screens

The enriched biological process information associated with our TMZ screen was compared to other alkylation screens reported for *S. cerevisiae* and *E. coli*. Our cross species analysis was performed similar to as reported (71). Briefly, to determine if the same gene ontology biological processes were enriched in the alkylation sensitive hit lists from all three organisms, representing X number of genes, we aligned GO-terms from each organism. To identify GO-terms significantly enriched for alkylation sensitive genes from all three organisms we randomly sampled X genes from the complete search space for each screen and identified associated GO functional terms. Random sampling was performed over two hundred iterations

and the average number of hits in each GO category and the standard deviation were determined. These values were than compared to the actual values to identify GO terms significantly ($P < 0.05$) enriched for alkylation sensitive genes from all three organisms. The molecular functions and the corresponding genes, which were significantly enriched in all three alkylation screen data sets, were then visualized using Cytoscape (72). Genes which modulate alkylator toxicity from *E. coli* (dark blue), *S. cerevisiae* (light blue) *and* human (light green), all served as protein nodes, with enriched GO terms serving as central nodes. The edges between GO terms and genes demarcate which genes are annotated to that biological process.

2.1.14 Generation of yeast knockout and viability analysis

Media preparation and other yeast manipulations were performed using standard methods. Mutants were made using a G418 knockout cassette from the *S. cerevisiae* Gene Deletion Project and were selected on Yeast Peptone Dextrose (YPD) plates containing G418 (200 mg/ml). A *mag1Δ* mutant was also made using a *URA3* based strategy, with selection occurring on Synthetic Defined media lacking uracil (SD-URA). Mutants were confirmed by PCR. Plate based MMS viability studies were performed as previously reported (73).

2.1.15 Transient siRNA knockdown for validation studies

We performed a wet-reverse transfection of the T98G cells in six-well plates using siPORT NeoFX (Applied Biosystems) per the manufacturer's instruction. Briefly, the siRNA was diluted to a working concentration of 2μM before being used. The siPORT NeoFX was mixed with OptiMEM and incubated at room temperature for 15 min. The siRNA was then

mixed with the OptiMEM and siPORT NeoFX solution and incubated at room temperature for another 10 min. The siRNA mixture was then plated into the wells. Cells were added to the plates for a concentration of 300,000 cells per well and an siRNA final concentration of 90nM. The cells were incubated for 24 hrs at 37°C before replacing transfection media with fresh media. Forty-eight hours after transfection cells were trypsinized and seeded for mRNA quantification and MTS cytotoxicity assays as described above. For knockdown of UNG and UBE3B the following Silencer Select siRNAs from Applied Biosystems were used: UNG: s14679; UBE3B: s40200; Silencer Negative control v2: am4613.

2.1.16 Cloning of UBE3B

The LN428 cells were seeded at $2x10^6$ cells per 100 mm dish and incubated until they reached approximately 70% confluence. The RNA was purified using the RNeasy Mini Kit (Qiagen, #217004) as per the manufacturer's instructions. The RNA concentration was measured using a NanoDrop spectrophotometer. Three μg of RNA was reverse transcribed to create cDNA using the SuperScript III first-Strand synthesis kit (Invitrogen, #18080-400) and an Oligo$(dT)_{20}$ primer in a 20 μl reaction following the manufacturer's instructions. The product cDNA was diluted tenfold before being used for cloning.

The primers were designed using the published UBE3B sequence from PubMed and guidelines provided in the pENTR cloning kit. The primers were tested for self-annealing, hairpin formation, and primer-dimers using the MacVector program. The PCR reaction was run as per the manufacturer's instructions. The PCR product was quantified using a NanoDrop spectrophotometer, before being run on a 1% agarose gel. PCR products with an expected base pair size (3200 base pairs) were cloned into the pENTR plasmid. The resulting plasmid was then

used to transform *E. coli* bacteria that were plated onto kanamycin plates. Four clones were picked and grown to purify the plasmid DNA. The plasmid DNA purified using a Qiagen miniprep kit from all four clones were then sent for double coverage sequencing using sequencing primers that were independent of initial cloning primers. The sequences obtained were then compared to the published PubMed sequence. The sequence of the primers used for cloning and sequencing can be found below (**Table 8**).

Table 8. Cloning and Sequencing Oligonucleotides for UBE3B

Cloning Oligonucleotides	
Oligo Name	Sequence
UBE3B-1-37	CACCATGTTCACCCTGTCTCAGACCTCGAGAGCATGG
UBE3B-3192-3155-R	CTAGGAGAGTTCAAAGCCCGTGTTCATGCTGATGGCG
Sequencing Oligonucleotides	
UBE3B-503-522	CCTTCACAGACACTTCAACG
UBE3B-983-1002	AGGAGGAGACAGATGGGTTC
UBE3B-1478-1497	CCAAACTGTGGGCATTTATC
UBE3B-1736-1754	TTGTAGAGAACGCCAAGGG
UBE3B-2251-2274	AAGACAACCAGTGGGGATGAGAGG
UBE3B-2714-2732	CCATTATCAAACCCGAGTG
UBE3B-354-334-R	CACATACCACACCTTAGGCTC
UBE3B-772-753-R	TGATGTGGATGAGGAACGGC
UBE3B-1112-1094-R	TCCACAGATTGGGAGAACC
UBE3B-1477-1457-R	GAAGCAGGTCATCAAGGTAAG
UBE3B-1999-1980-R	TCTCCTTCTCCTTGGTAACC
UBE3B-2415-2397-R	AAGCAGTTGGCTCAGGAAG
UBE3B-2840-2821-R	TGGAAACCACCGTAGTAGAC

3.0 SYNTHETIC LETHAL SCREEN OF T98G CELLS USING TEMOZOLOMIDE AND SIRNA

3.1 INTRODUCTION

Glioblastoma multiforme (GBM) is the most common and aggressive primary brain tumor (6). Temozolomide (TMZ), radiation and surgery are currently used for treatment of GBM, yet median survival is still less than 2 years (3-5). Chemotherapy resistance and difficulties in surgical removal contribute to poor prognosis (6). TMZ methylates several bases in DNA, including the O^6 position of guanine (O^6-MeG), the N7 position of guanine (N7-MeG) and the N3 position of adenine (N3-MeA). Two of the lesions (N7-MeG and N3-MeA) account for greater than 80% of the DNA lesions induced by TMZ, yet result in negligible clinical toxicity due to robust repair mechanisms (13). These lesions are predominantly repaired by the base excision repair (BER) pathway, a DNA repair mechanism that involves as many as 20 different proteins (26). BER is initiated by one of eleven lesion specific DNA glycosylases (28). For example, methylpurine DNA glycosylase (MPG) recognizes and removes the N7-MeG and N3-MeA lesions that are induced by TMZ (15). The resulting abasic site is a substrate for AP endonuclease (APE1), which cleaves the sugar phosphate backbone, leaving a 3'OH and a 5'-deoxyribose-phosphate moiety (15). Poly (ADP-ribose) polymerase 1 (PARP1) recognizes and binds the nick or single strand break (SSB), which activates PARP1 to recruit downstream

proteins such as X-ray repair cross complementing group 1 (XRCC1) and DNA polymerase ß (Polß) (26). Polß is recruited by XRCC1 and PARP1 to hydrolyze the 5'-deoxyribose-phosphate and fill the nucleotide gap, followed by ligation mediated by the XRCC1/DNA ligase III heterodimer (26) or DNA ligase I (74).

Virtually all of TMZ's clinical cytotoxicity is attributable to the O^6-methylguanine (O^6-MeG) lesion, which accounts for approximately 5% of TMZ induced lesions (13). The O^6-MeG lesion is repaired via a direct reversal mechanism by the protein O^6-methylguanine-DNA methyltransferase (MGMT), which transfers the O^6-methyl group from the guanine base onto a Cys residue in the MGMT protein (16). If the O^6-MeG lesion is not removed by MGMT, during cellular replication the mis-pairing of O^6-MeG with thymine is detected by the mismatch repair enzymes, triggering apoptosis signaling and cytotoxicity (13, 17, 18). However, 5-year survival rates still remain low in TMZ treated patients (4, 5), and TMZ resistance and/or recurrence with chemotherapy resistant tumors is common. Resistant cells can harbor mutations in mismatch repair proteins such as the mutS homolog 6 (MSH6) (19) or have elevated expression of MGMT (as in the T98G cell line) (20). Earlier endeavors to enhance TMZ efficacy by using MGMT inhibitors to prevent the repair of O^6-MeG lesions have not shown an increase in sensitivity or efficacy in clinical trials (21, 23), especially in TMZ-resistant GBM (22, 24, 25).

Because GBM tumors can be resistant to TMZ therapy and recurring tumors may acquire resistance to the O^6-MeG lesion, we explored novel mechanisms of TMZ resistance in tumor cells independent of the O^6-MeG lesion (17, 75). To discover genes that sensitize cells to TMZ, we undertook a synthetic lethal screen using Ambion's Silencer® Human Druggable Genome siRNA library, the chemotherapeutic agent TMZ and the T98G cell line, a TMZ resistant cell line derived from a GBM tumor (76). We determined that several biological

pathways were enriched for TMZ sensitizing genes, including proteins involved in BER, response to DNA damage, cellular proliferation, and protein modification. Many similar biological processes were also highly enriched in DNA alkylation screens performed in *Saccharomyces cerevisiae* and *Escherichia coli* (77, 78). We combined the selected hits from all three species and performed a functionome analysis to identify significantly enriched biological processes conserved across all three organisms (71). We created a cross-species network representing the shared alkylation response from bacteria, yeast and humans and suggest that this network represents an *Alkylation Functionome* that includes many novel proteins not previously thought to impact TMZ or MMS resistance. Our identification of evolutionarily conserved mechanisms affecting TMZ sensitivity suggests that although many proteins and processes impact sensitivity to alkylators, several critical survival pathways can be targeted to improve chemotherapy efficacy. Simultaneous inhibition of both DNA repair and protein modification processes in yeast yielded greater sensitivity to alkylating agents than inhibition of either process alone. However, knockdown of both protein modification and DNA repair genes did not have an increased effect on alkylation-induced toxicity in human cells. This potential epistatic interaction between the ubiquitin protein ligase E3B (UBE3B) and the DNA repair protein uracil-DNA glycosylase (UNG) suggests they may be in the same survival pathway. Many DNA repair genes are regulated by ubiquitylation and several crucial DNA repair proteins are E3 ligases such as BRCA1 (79-82). UBE3B previously has not been implicated in DNA repair or alkylation survival and elucidating its role and substrates will be important to discover how it improves alkylation survival. The highly conserved functionome enriches for essential genes in critical biological processes, thus also enriching for possible gene interactions in these crucial processes. Importantly, the biological processes and corresponding genes identified in our

functionome analysis represent novel potential drug targets to increase TMZ efficacy independent of O^6-MeG mediated toxicity as well as potential epistatic relationships that would provide mechanistic insight into the function of specific activities.

3.2 RESULTS

3.2.1 Synthetic lethal siRNA screen and generation of high confidence hit list

To identify genes that when silenced confer sensitivity to the alkylating agent TMZ we conducted a synthetic lethal screen (in triplicate) on a TMZ-resistant GBM cell line (T98G) using TMZ and an siRNA library (**Figure 4A**). A concentration of 1 mM TMZ was selected due to minimal (~10%) toxicity measured at 48 hrs via an MTS assay. Further, T98G cells were treated with varying concentrations of DMSO to determine the impact of DMSO on cell survival (**Figure 4B**). No toxicity was observed at 1% DMSO after 48 hrs; this was the percentage of DMSO used for both vehicle and TMZ treatment in the screen.

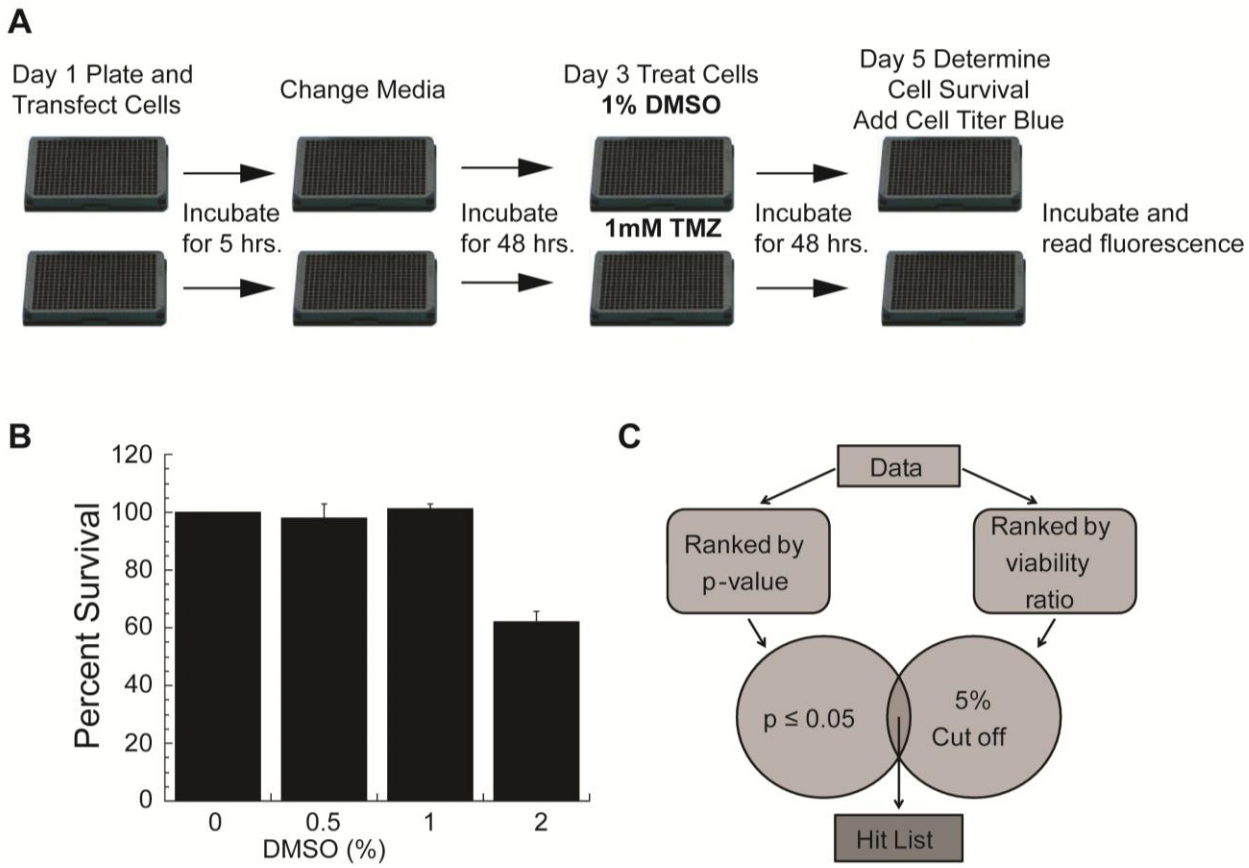


Figure 4. Screen methodology, vehicle effects to verify screening conditions, and selection criteria. **(A)** A wet-reverse transfection was used to knockdown genes with siRNA. Transfection consisted of 3 siRNAs targeting the same gene in every well. Transfection reagents were mixed and split into two different plates before the addition of T98G cells to the plates. After 48hrs one plate was treated with 1mM TMZ in DMSO (1% DMSO, final concentration) while the other was treated with 1% DMSO. **(B)** The vehicle (1%DMSO) is nontoxic after 48hrs of incubation before determining growth inhibition by an MTS assay. C) Methodology used to create high confidence hit list from two distinct methods

The transfection conditions used in the screen maximized knockdown, similar to previous screens using these conditions (58). We first calculated viability ratios (ratio of survival of 'siRNA plus TMZ' to 'siRNA plus vehicle') and normalized to vehicle-treated scrambled siRNA, enabling cross plate comparisons. We then statistically analyzed the viability ratios and determined a group of gene targets that sensitized cells to TMZ by performing a two-tail sample t-test on each gene to determine the effect of siRNA and TMZ on cellular survival. Targeting siRNAs were selected with a p-value of less than or equal to ≤ 0.05. The high confidence hit list

contained targets that satisfied both criteria: a p-value ≤ 0.05 and a viability ratio in the lowest 5 percent (**Figure 4C**). By using these two methods, we created a hit list of 172 genes (**Dataset S1**). While most of the viability ratios on the hit list were between 0.4 and 0.65, these genes significantly modulated the toxicity of TMZ, yielding many new potential targets to increase response to alkylation chemotherapy. All hits were analyzed for pathway and network enrichment to determine the potential biological pathways that modulated alkylation toxicity.

3.2.2 Analysis of screen results for pathway and network enrichment

The genes contained in the hit list were analyzed with NIH DAVID, Princeton GO term finder and Ingenuity Pathways Analysis (Ingenuity® Systems, www.ingenuity.com) (IPA) for pathway and network enrichment. The analysis with DAVID and Princeton GO term finder was performed to search for gene enrichment in molecular functions, cellular compartments, or biological processes based on the associated gene ontology terms. Each gene is linked to different terms based on their gene ontology in a hierarchical manner and gene enrichment was determined based on the probability of selecting proteins with similar terms. By focusing on the 5,520 "druggable" genes we inherently expected a small but presumably more relevant and smaller gene enrichment dataset compared to probing the entire human genome as background. Several different thresholds for viability ratio percentiles were initially analyzed for gene enrichment, including 2.5%, 5%, 7.5% and 10% percentiles (**Datasets S2-S4**). The 5% group was selected for a more detailed analysis because it contained many of the gene enrichment groups seen in both the 7.5% and 10% datasets, but was more focused with higher gene enrichment scores.

The enriched genes from the NIH DAVID analysis contained several different groups including DNA repair, response to DNA damage stimulus, cell proliferation, amino acid glycosylation and biopolymer glycosylation (**Table 9**). The most enriched groups have specific DNA repair functions including BER activity, hydrolase activity and DNA N-glycosylase activity. Surprisingly, four genes belonging to the three most enriched groups were DNA glycosylases that do not recognize alkylation damage (15, 27, 28, 34, 37, 40, 45). Interestingly, these genes (TDG, OGG1, NEIL1 and UNG) are known to recognize and repair multiple types of oxidative DNA damage (**Table 10**).

Table 9. Top gene enrichment groups determined by NIH DAVID

Gene Ontology Term	p-value	Genes	Fold Enrichment	FDR
Base-excision repair	2.92E-05	OGG1, NEIL1, POLD1, TDG, UNG	14.7	0.054348
Hydrolase activity, Hydrolyzing N-glycosyl compounds	4.71E-05	CD38, OGG1, TDG, UNG, NEIL1	13.48214	0.083124
Glycosidase	9.54E-05	SPAM1, SMPDL3A, OGG1, TDG, TREH, CTBS, UNG, NEIL1	5.992216	0.148149
DNA N-glycosylase activity	1.07E-04	OGG1, TDG, UNG, NEIL1	17.47685	0.188687
DNA repair	1.73E-04	LIG4, OGG1, TDG, ATR, TREX1, LIG1, UNG, NEIL1	5.522238	0.269107
DNA damage	1.99E-04	LIG4, OGG1, TDG, ATR, TREX1, LIG1, UNG, NEIL1	5.416041	0.309503

Table 10. Sensitizing DNA glycosylases, their known substrates and viability ratios

Gene Symbol	Gene Name	Known Substrate*	Screen Viability Ratio
OGG1	8-oxoguanine DNA glycosylase	8-oxoG:C=T=G; me-FapyG:C; FapyG:C; 8-oxoA:C; urea	0.648
UNG	Uracil DNA glycosylase	ssU; U:G; U:A; 5-fluorouracil; 5,6-Dihydroxy-U:G; 5-OH-U:G; Isodialuric acid; Alloxan	0.446
TDG	Thymine DNA glycosylase	U:G; T:G and ethenoC:G; 5-Fluorouracil; 5-fluorouracil (ss); 5-Hydroxymethyluracil; hypoxanthine:G; 5-bromouracil; εC:A Tg:G; 5-formyl-U	0.541
NEIL1	Nei endonuclease VIII-like 1 (E. coli)	TgG; 5-OH-C; 5-OH-U:AT>G; Guanidinohydantoin; guanidinohydantoin (ss); Iminoallantoin; Iminoallantoin (ss); Spiroiminodihydantoin; Spiroiminodihydantoin (ss); 5,6-Dihydro-T; 5,6-Dihydro-U:G=C=A>T; FapyG:C; 8-Oxo-G:C=G>T>A; FapyA:T; (5'R)-8,5'-Cyclo-2'-deoxyadenosine; (5'S)-8,5'-Cyclo-2'-deoxyadenosine; 8-Oxo-A:C	0.559
* Target base on left in mismatches			

There were large overlaps of enriched genes from DAVID and Princeton GO term finder analyses (**Figure 5 and Dataset S5**). The oxidative DNA glycosylases, which sensitized cells to TMZ, were identified in the most significantly enriched networks determined by both DAVID and Princeton GO term finder (**Table 9 and Dataset S6**). Data were also analyzed through the use of IPA. IPA uses a manually curated proprietary interaction database and a right-tailed Fisher's exact test to calculate a p-value determining the probability that each biological function assigned to that network is due to chance alone. By using a distinct database to evaluate gene enrichment, which also returned similar biological processes, pathways and functions, we

validated our results from the previous analysis methods. IPA identified several of the DNA glycosylases in a statistically significant network, with a network value of twenty-four (**Figure 6**). The network score is the negative logarithm of the p-value (p-value $=10^{-24}$). This network is very similar to the biological processes and corresponding genes overrepresented in the DAVID and Princeton GO analysis. Further, IPA returned several highly enriched molecular functions including DNA repair (p-value=5.31 x 10^{-4}), excision repair (p-value=8.87 x 10^{-4}), ligation of a DNA fragment (p-value= 2.84 x 10^{-3}), ligation of DNA (p-value= 9.07 x 10^{-3}) and nicking of DNA (p-value=1.83 x 10^{-2}), very similar to the functions returned by DAVID.

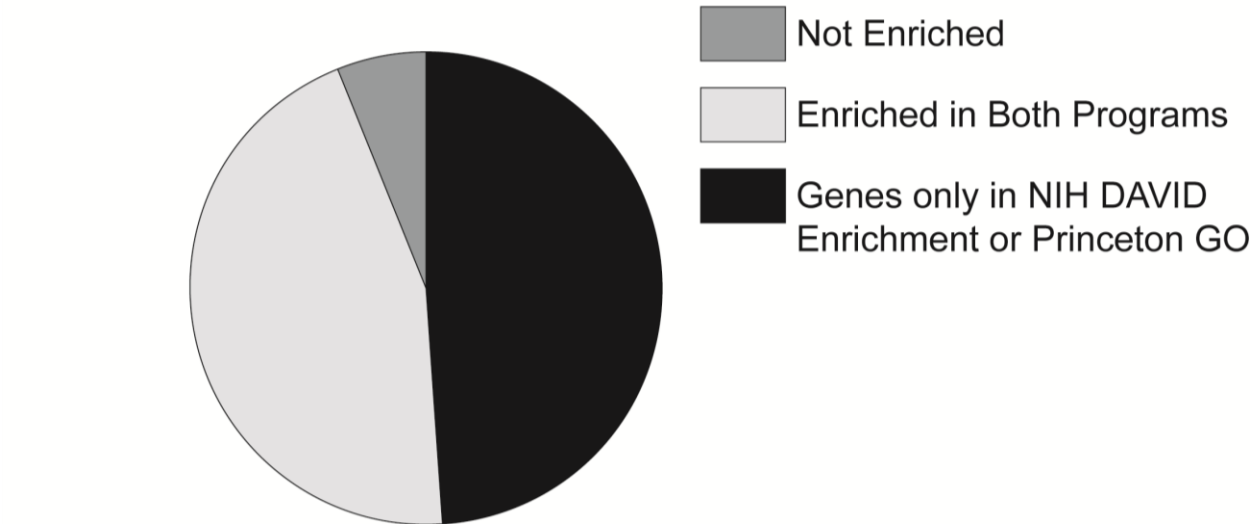


Figure 5. Different methods of analysis yield similar gene enrichments.
The diagram demonstrates the large overlap of gene enrichment groups determined by NIH DAVID and Princeton GO term finder.

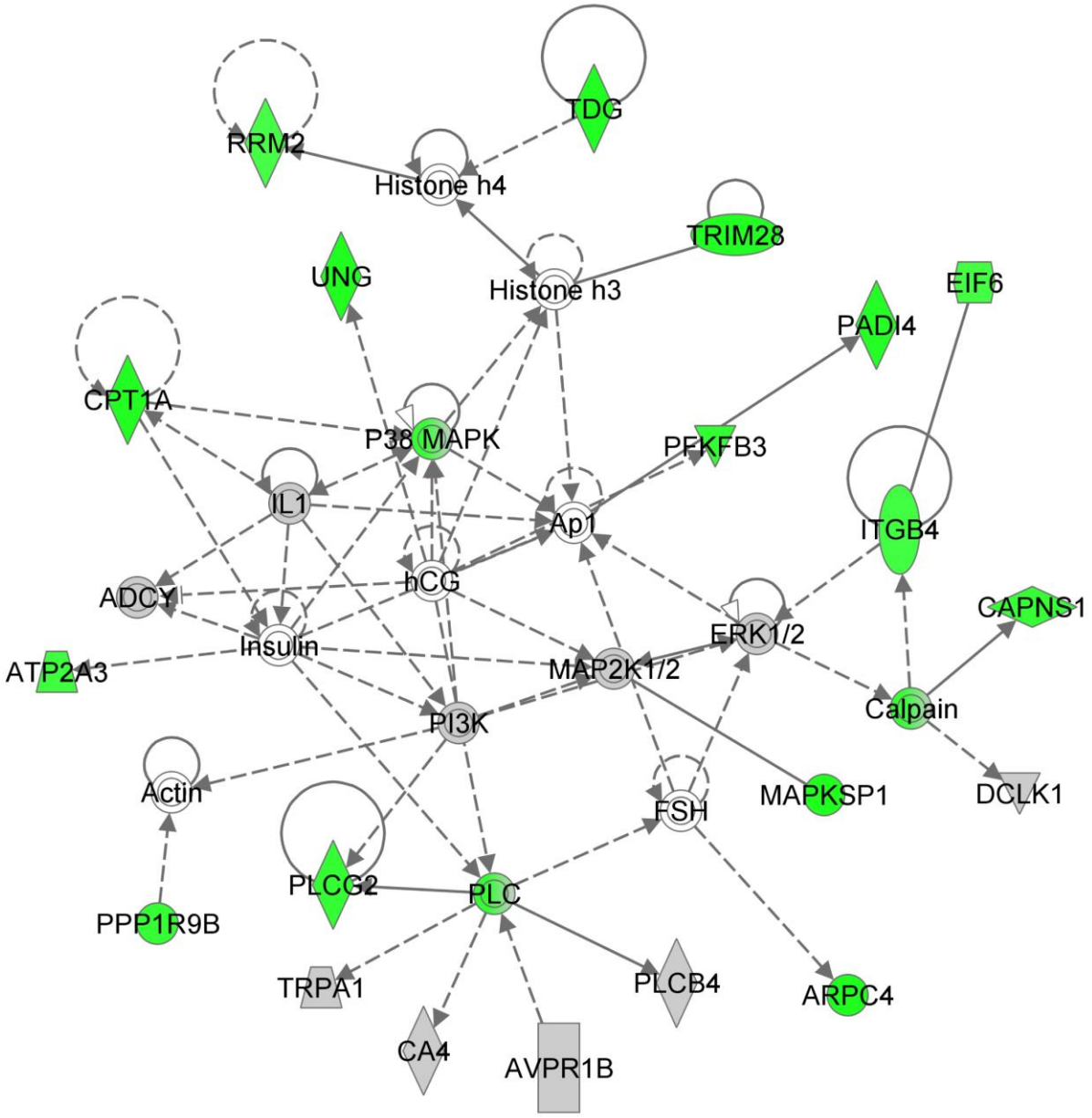


Figure 6. Analysis with NIH DAVID, Princeton GO term finder and Ingenuity Pathway Analysis yield similar gene enrichments, biological processes and pathways.
The high confidence hit list was also analyzed by IPA for biological pathway enrichment. One highly significant network generated by Ingenuity Pathway Analysis containing similar biological processes with other analysis methods is shown. The network is composed of many overlapping genes with genes in the high confidence hit list colored green. The top functions in the network are nucleic acid metabolism, small molecule biochemistry and metabolic disease with a network score of 24. The network score is the negative logarithm of the p-value.

Other highly enriched molecular functions included post-translational modifications (p-value= 2.98×10^{-3} - 3.12×10^{-2}), carbohydrate metabolism (p-value= 5.56×10^{-3} - 4.57×10^{-2}), nucleic acid metabolism (p-value= 9.65×10^{-4} - 3.12×10^{-2}), and small molecule biochemistry

(p-value= 9.49 x 10^{-4} - 4.7 x 10^{-2}). Within these different functional groups, a conserved theme was the ability of the genes to modify DNA or proteins. The modifications of DNA most significantly contained the excision of uracil (p-value = 9.65 x 10^{-4}), while the many protein modifications included metabolism of proteoglycan (p-value = 5.56 x 10^{-3}), metabolism of polysaccharide (p-value = 7.29 x 10^{-3}), metabolism of carbohydrate (p-value = 8.05 x 10^{-3}), generation of diacylglycerol (p-value = 5.56 x 10^{-3}) and the release of acetylcholine (p-value = 9.07 x 10^{-3}). The "Protein Modification" theme was in our DAVID and Princeton GO results, and included protein amino acid glycosylation, bipolymer glycosylation, glycoprotein biosynthetic process, glycoprotein metabolic process and cellular carbohydrate metabolic process. These processes contained genes whose corresponding proteins are involved in protein modifications and we were intrigued by our result that when silenced the corresponding cells are sensitive to TMZ. However, because of the highly significant enrichment of DNA repair processes in all three programs, genes associated with the DNA Repair pathway category were selected for initial biological validation.

3.3 DISCUSSION

To discover "druggable" targets that contribute to TMZ sensitivity, independent of the O^6-MeG lesion, we conducted a synthetic lethal siRNA screen against 5,520 genes in a TMZ-resistant cell line that has elevated expression of MGMT (T98G), uncovering enriched biological processes independent of MGMT and O^6-MeG lesion induced cell death. The vehicle control did not have any cytotoxicity effects observed in the dose or duration used in the screening assay, thereby limiting the cytotoxicity results to that of TMZ and siRNA interactions. The synthetic

lethal siRNA screen used similar screening conditions and selection criteria to obtain maximal knockdown of T98G mRNA and to determine positive screening hits (58, 60). After selection of genes that significantly sensitize T98G cells to TMZ when knocked down, these hits were analyzed for pathway, network and gene enrichment using several programs. By using NIH DAVID, Princeton GO term finder and IPA, with distinct methods of categorizing genes, organizing them into groups and determining probability of gene enrichment, we are more confident the biological processes and their corresponding genes enriched in all three programs are true hits with biological impact on survival from alkylation damage. Importantly, the three programs contained considerable overlap in not only the genes that were enriched, but they also corresponded to similar biological processes and functions. The increase in biological processes and functions enable interrogation of related genes not used in the screen that contain similar biological processes and functions to determine if they are also suitable targets. By performing the enrichment analysis, we not only focus our workup efforts on the most likely positive hits, but also uncover the essential processes and functions for alkylation survival. These processes possibly can be inhibited at several different steps. Learning the essential alkylation survival processes yields not only many possible drug targets, but also a potential increase in therapeutic window for TMZ treatment. This can result from inhibition of alkylation survival pathways that may be partially compromised in tumors due to spontaneous mutations or tumorigenisis. An example of potential tumor selectivity by inhibiting alkylation survival pathways was demonstrated in a recent publication from our lab (70). Although DNA alkylator treatment is not selective, by concurrently targeting the NAD^+ biosynthesis pathway it may be possible to selectively induce tumor cytotoxicity from alkylation damage due to prevalent defects in tumor NAD^+ biosynthesis (70). Therefore, it is crucial to not only discover genes essential for

alkylation survival, but to also uncover if cancers are selectively defective in these pathways. With the defects in DNA repair and cellular responses to the many mutations that cancers undergo during transformation, it is possible to selectively target tumor DNA repair pathways. The recent development of PARP inhibitors and their synthetic lethal interaction with homologous recombination deficient cells suggests other synthetic lethal DNA repair defect combinations are possible (83, 84). Determining how normal and cancer cells differ in their repair and response to TMZ genotoxic lesions is central to uncovering potential synthetic lethal interactions and cancer specific repair mechanisms to target to selectively kill tumors.

The initial computational analysis contained some expected gene enrichment groups and networks, particularly those that focused on BER, as it is known to repair TMZ induced DNA lesions. When mining deeper into the role of BER in TMZ survival, we discovered that several BER proteins not previously thought to be involved in alkylation repair or survival greatly sensitized the cells to TMZ. Of particular interest were the four sensitizing DNA glycosylases (UNG, TDG, OGG1 and NEIL1) that remove oxidative DNA damage and not alkylation DNA lesions (15). Further, cancer cells appear to be under greater oxidative stress compared to normal cells and impairment of oxidative DNA damage repair may yield tumor selective cytotoxicity (85). With this initial pathway, network and gene enrichment determined, we then biologically validated several genes with the most significant enrichment findings from the three programs.

4.0 VALIDATION OF SIRNA SCREEN

4.1 INTRODUCTION

In the previous chapter the screening protocol was vetted and the screen was carried out. There was initial computational analysis of the selected genes to determine pathway, network and gene enrichment to increase our understanding of the important biological processes in alkylation survival, but to also increase the likelihood of identifying a true positive result. After computational analysis, we sought to biologically validate our screen using distinct methods to eliminate the possibilities of siRNA off-target effects in the screening procedure and verify our results. Further, we wanted to determine how our screen in human cells corresponded with that of previous alkylation screens done in other organisms. Overlap of conserved proteins or functions would support our screen results and lend importance to the conserved proteins and functions in alkylation survival. These conserved proteins and functions would be of great interest due to their conserved nature and likely fulfill essential roles in alkylation survival. With these functions uncovered, we could focus our efforts on novel essential mechanisms for alkylation survival with the associated proteins as new potential therapeutic targets.

4.2 RESULTS

4.2.1 Knockdown of oxidative DNA glycosylases sensitize cells to TMZ

In our analysis, we discovered that the DNA repair category was enriched for TMZ sensitizing genes. This category contained four DNA glycosylases (UNG, OGG1, TDG and NEIL1), each specific for the repair of oxidative DNA damage (27, 28, 34, 86). To determine if other DNA glycosylases also sensitize cells to TMZ and to validate the siRNA synthetic lethal screen results, seven of the eleven DNA glycosylases were knocked down using a lentiviral system to create stable T98G-derived cell lines. Knockdown of glycosylase mRNA was validated by qRT-PCR (**Figure 7A**). The stable cell lines were then tested for sensitization to 1 mM TMZ in a 96-well plate format using a modified MTS assay (67). As confirmation of our siRNA screen analysis, the newly developed UNG-KD cell line was also sensitive to TMZ (**Figure 7B**). Furthermore, knockdown of the DNA glycosylases MYH and MPG also sensitized cells to TMZ (**Figure 7B**).

Figure 2

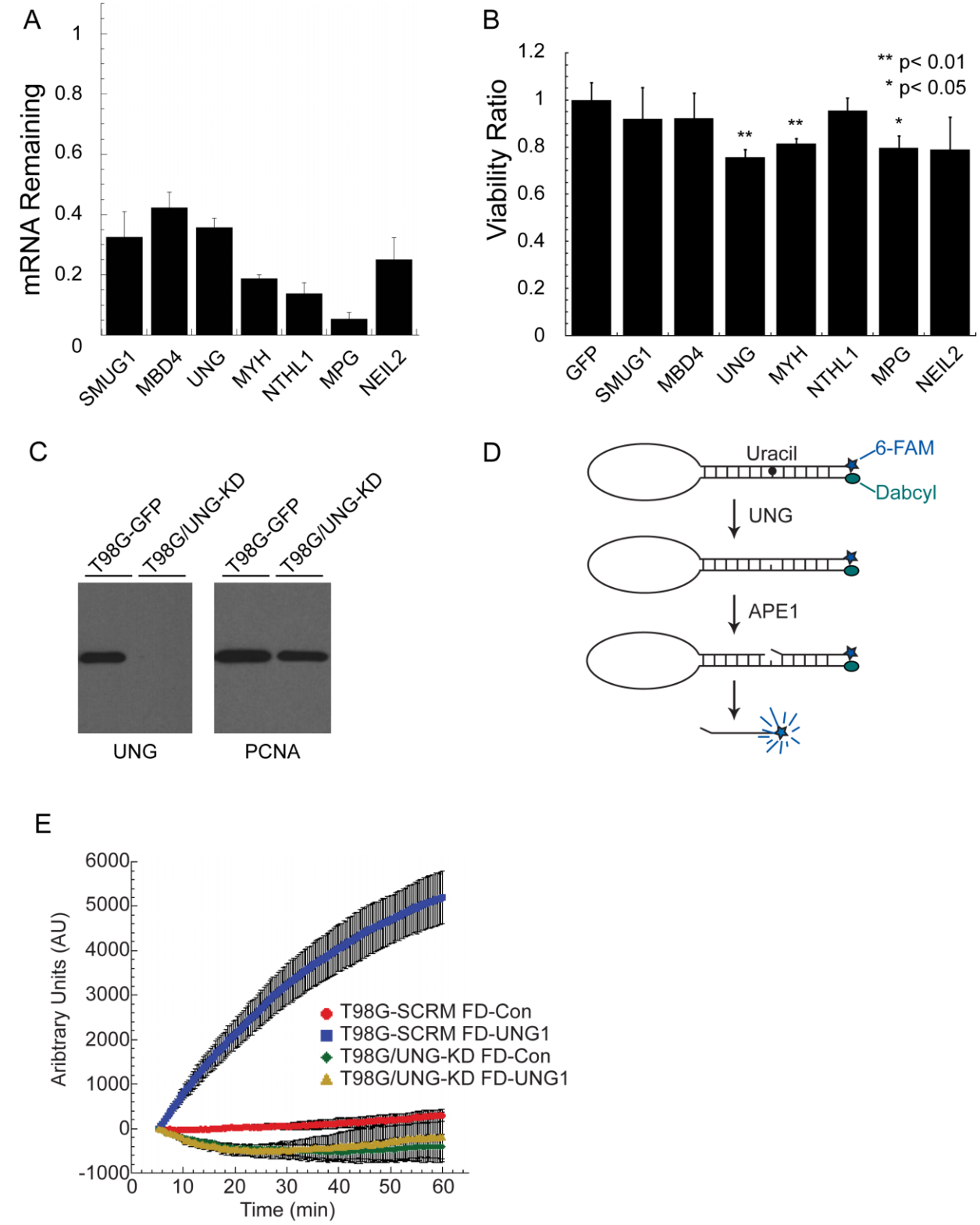


Figure 7. Glioma cells with shRNA knockdown of DNA glycosylases are more sensitive to the clinical alkylator TMZ.

(A) Quantification of DNA glycosylase mRNA knockdown in cells as determined by qRT-PCR. TaqMan probes were used to quantify mRNA levels on an Applied Biosystems StepOnePlus machine. The qRT-PCR data was analyzed using the ΔΔCt method and was normalized to GFP infected plate controls. Gene expression of each gene was normalized to the expression of human ß-actin. The mean of three independent experiments is plotted ± SEM. **(B)** Validation of TMZ sensitization with knockdown of specific DNA glycosylases. T98G DNA glycosylase knockdown cell line sensitivity to TMZ was determined by an MTS assay 48 hr after exposure to 1 mM TMZ. The viability ratio is double normalized to account for both vehicle treated shRNA mediated growth defects and toxicity of control cells to TMZ. The mean viability ratio of three independent experiments is plotted ± SEM. **(C)** Knockdown of UNG mRNA corresponds with decreased UNG protein levels. Control and UNG-KD cell line nuclear extracts were resolved in a 4-20% SDS/PAGE gel and immunoblotted for UNG. The blot was stripped and re-probed for PCNA, which was used as a loading control. **(D)** Molecular beacon model for real-time detection of uracil removal at 37°C in nuclear extracts. **(E)** Knockdown of UNG abolishes ability to remove uracil from DNA as determined by DNA Glycosylase Molecular Beacon Activity Assay. Specific activity of DNA glycosylase activity for the uracil lesion was measured in nuclear extracts from T98G-SCR (control-beacon, red circles; uracil-beacon, blue squares) and T98G/UNG-KD (control-beacon, green diamonds; uracil-beacon, yellow triangles). The mean fluorescence response unit of three experiments is plotted ± SEM.

4.2.2 UNG knockdown eliminates removal of uracil in glioma cells

Although UNG mRNA knockdown sensitized the T98G cells to TMZ in the validation study, we were uncertain if the mRNA knockdown affected UNG protein levels and DNA repair activity because other DNA glycosylases have a similar lesion spectrum (**Table 3**). First, we tested UNG protein expression via immunoblot and determined that UNG protein levels were decreased in T98G/UNG-KD cells as compared to T98G-GFP control cells (**Figure 7C**). However, we were interested in ascertaining if the decrease in UNG protein levels affects the DNA repair capacity of the cells, because there is a large functional overlap of UNG with the three DNA glycosylases SMUG1, TDG and MBD4 (27, 28, 34, 86). We were concerned that depletion of one of these DNA glycosylases would not have a functional DNA repair defect due to compensation by the other glycosylases. We therefore developed a DNA glycosylase molecular beacon assay to quantify the functional loss of uracil removal by UNG knockdown, similarly to what we have previously described for the analysis of MPG activity (87, 88). The assay uses a molecular beacon composed of a single stranded DNA molecule with a 5' fluorophore (6-FAM) and a 3' quencher (Dabcyl). The oligonucleotide also contains a uracil

lesion, a substrate for UNG. We selected uracil as the lesion because it is removed by UNG, SMUG1, TDG and MBD4 (27, 28, 34, 86). Therefore, we could determine if UNG-KD impacts uracil removal and if compensation by other DNA repair proteins occurred. A similar oligonucleotide with a normal base was used as a control substrate. Removal of the uracil lesion by UNG and hydrolysis of the DNA backbone by APE1 results in separation of the 6-FAM fluorophore from the Dabcyl quencher and the increase in 6-FAM fluorescence is proportional to uracil removal (**Figure 7D**). The T98G-SCR lysate incubated with control beacon (**Figure 7E**, red circles) had a minimal increase in fluorescence, signifying the control beacon is intact. However, the T98G-SCR lysate incubated with the beacon containing uracil (**Figure 7E**, blue squares) exhibited a large increase in fluorescence (17.69 fold at 60 min) compared to the control beacon, indicative of robust uracil removal from the molecular beacon. The T98G/UNG-KD lysates incubated with either the control beacon (**Figure 7E**, green diamonds) or the uracil-containing beacon (**Figure 7E**, yellow triangles) contained no difference in fluorescence and exhibited low fluorescence, supportive of loss of UNG activity.

These results support our conclusion that the T98G/UNG-KD cells have impaired uracil removal due to UNG knockdown as compared to the T98G-SCR control. The scrambled shRNA control was used instead of the GFP control for the activity assay due to the interference of GFP fluorescence with the molecular beacon fluorophore 6-FAM. Although uracil can also be removed by SMUG1, TDG and MBD4, UNG-KD alone was sufficient to deplete uracil removal activity (27, 28, 34, 86).

4.2.3 Cross-species functionome analysis creates a network of conserved processes important for survival after alkylation damage

After biological validation, we then compared our T98G siRNA screen results to other alkylation screens conducted in *S. cerevisiae* and *E. coli*, comparing the outcome data using GO molecular function terms (77, 78). This was done by compiling the sensitizing genes from all the screens and linking them to their specific Gene Ontology Biological Process. The list of sensitizing genes from *E. coli, S. cerevisiae* and human was then computationally analyzed to identify GO-functional categories over-represented with genes from human and either *E. coli* or *S. cerevisiae* (**Table 11**). The molecular functions enriched in all three alkylation screen datasets were then visualized using Cytoscape to create a cross species functionome of proteins which modulate toxicity to alkylating agents in *S. cerevisiae*, *E. coli* and human (**Figure 8**). This network of proteins has several "GO-hubs" of proteins from all three species clustered around a conserved biological process (node) that significantly impacts the survival of correspondingly depleted cells after exposure to alkylating agents. These include proteins involved in global processes such as response to drugs or changes in pH. There was also a collection of processes involved in macromolecule biosynthesis and modifications including fatty acid biosynthetic processes, transcription, regulation of transcription, transcription initiation, negative regulation of translation, protein modification process, protein processing and protein targeting to membrane. These results suggest that synthesis of new RNA, protein, and fatty acids are essential for survival to alkylation.

Table 11. Significantly conserved biological processes as determined by cross-species Functionome Analysis.

GO Category	Human		Bacteria		Yeast		Combined		
	Number Sensitive	Gene	Number Sensitive	Gene	Number Sensitive	Gene	Total Proteins	# Alkylation Sensitive	Z-score
Response to DNA damage stimulus	1	ATR	11	ada, alka, reca, recb, recc, recn, reco, ruva, ruvb, ruvc, yjiw	56	MMS4, DCC1, RAD18, RAD59, DUN1, NUP84, RAD57, RAD55, HPR1, RAD9, XRS2, MUS81, RAD51, MAG1, RAD24, RPB9, MMS2, RAD54, WSS1, CTF8, MET18, CSM2, REV7, MPH1, CTK2, SRS2, ASF1, RPB4, POL32, MGM101, DEF1, RAD27, CTK1, RTT109, RAD5, SLX4, BUR2, TOP3, MMS22, PSY3, BDF1, RAD52, CTK3, HOF1, NPL6, SGS1, YAF9, RAD50, CKB2, REV1, RAD17, RMI1, REV3, DDC1, CTF4, MMS1	297	68	17.9
DNA repair	6	LIG1, RFC5, UNG, ATR, POLD1, TREX1	10	ada, alka, reca, recb, recc, recn, reco, ruva, ruvb, ruvc	45	MMS4, MSI1, RAD18, RAD59, DUN1, RAD57, RAD55, RAD9, XRS2, MUS81, RAD51, MAG1, RAD24, RPB9, RAD54, WSS1, CTF8, MET18, CSM2, REV7, MPH1, SRS2, ASF1, SPT10, RPB4, POL32, MGM101, RAD27, RTT109, RAD5, SLX4, TOP3, MMS22, PSY3, BDF1, RAD52, SGS1, YAF9, RAD50, REV1, RAD17, REV3, DDC1, CTF4, MMS1	316	62	16.2
DNA replication	9	LIG1, RFC5, RFC5, LIG4, ATR, POLD1, RRM2, TREX1, TREX1	5	dnag, dnak, dnat, holc, pria	21	CCR4, DCC1, RAD9, RAD24, RRM3, WSS1, RTT107, CTF8, SRS2, ASF1, POL32, RAD27, SLX4, TOP3, RAD52, SGS1, REV1, RAD17, RMI1, REV3, CTF4	203	35	10.6
DNA recombination	4	LIG1, RAG2, TREX1, TREX1	7	reca, recn, reco, ruva, ruvb, ruvc, xerc	11	MMS4, RAD59, HPR1, MUS81, ERG28, RAD51, SHU1, SLX4, CDC73, RAD52, SGS1	158	22	9.4
DNA metabolic process	1	LIG1	2	dnag, reca	6	MMS4, RAD57, RAD55, MUS81, RAD51, TOP3	36	9	8.2
Base-excision repair	4	OGG1, TDG, UNG, NEIL1, MPG*	1	alka	4	XRS2, MAG1, POL32, RAD50	36	10	6.9
DNA dealkylation	1	MPG*	1	alka	1	MAG1	7	4	5.1
Protein modification process	6	B4GALT7, ICMT, CHRM3, PADI1, UBE3B, PADI4	0		4	PBY1, STP22, AIM22, LIP2	122	10	4.6
Transcription initiation	2	KLP1, TAF1L	0		2	SWC3, DEP1, CCR4, ROX3, SWC5, HPR1, UME6, ADR1, ESC2, IES6, MOT2, RPB9, RTF1, OPI1, RIM101, STB5, MET18, CTK2, ASF1, RPB4, CTK1, RTT109, ARP6, IOC2, SWI6, BUR2, RSC2, BDF1, CDC73, YAP1, CTK3, NPL6, YAF9, CSE2, HTZ1, UAF30, LGE1, TAF14, HFI1	29	4	4.3
Protein targeting to membrane	2	ICMT, APG4C	0		2	STP22, VPS28	18	4	4.0
Transcription	1	GRLF1	10	ada, arcb, cadc, cynr, dnag, fis, mara, oxyr, slya, uidr	39	SWC3, DEP1, CCR4, ROX3, SWC5, HPR1, UME6, ADR1, ESC2, IES6, MOT2, RPB9, RTF1, OPI1, RIM101, STB5, MET18, CTK2, ASF1, RPB4, CTK1, RTT109, ARP6, IOC2, SWI6, BUR2, RSC2, BDF1, CDC73, YAP1, CTK3, NPL6, YAF9, CSE2, HTZ1, UAF30, LGE1, TAF14, HFI1	901	50	3.8
Regulation of transcription	4	KLP1, POFUT1, PADI4, TRIM28	2	mara, uidr	34	UME6, STB5, SSK1, YAP1, YAF9, TAF14	650	40	3.6
Nucleotide-excision repair	2	OGG1, NEIL1	0		4	RAD9, RAD24, MET18, POL32	49	6	3.4
Negative regulation of DNA replication	1	ATR	0		1	RAD9	12	2	2.9
Mismatch repair	3	TDG, TREX1, TREX1	1	dam	1	POL32	48	5	2.4
Protein processing	1	CASP1	0	NA	3	RIM101, RIM13, DFG16	34	4	2.4
Fatty acid biosynthetic process	2	PTGIS, THEDC1	0	NA	4	OAR1, ERG3, HFA1, SCS7	79	6	2.4
Response to drug	2	HMGCS2, RPS6KB1	0	NA	12	SWC3, CCR4, UBP14, PPH3, UME6, RPB9, STB5, SOD1, EAP1, PEP3, MMS22, YAP1	260	14	2.0
Response to pH	1	STS	0	NA	1	RIM101	18	2	2.0
Negative regulation of translation	1	HRI	0	NA	1	EAP1	16	2	1.8
Positive regulation of gene-specific transcription	1	TRIM28	0	NA	1	SWI6	33	2	1.6

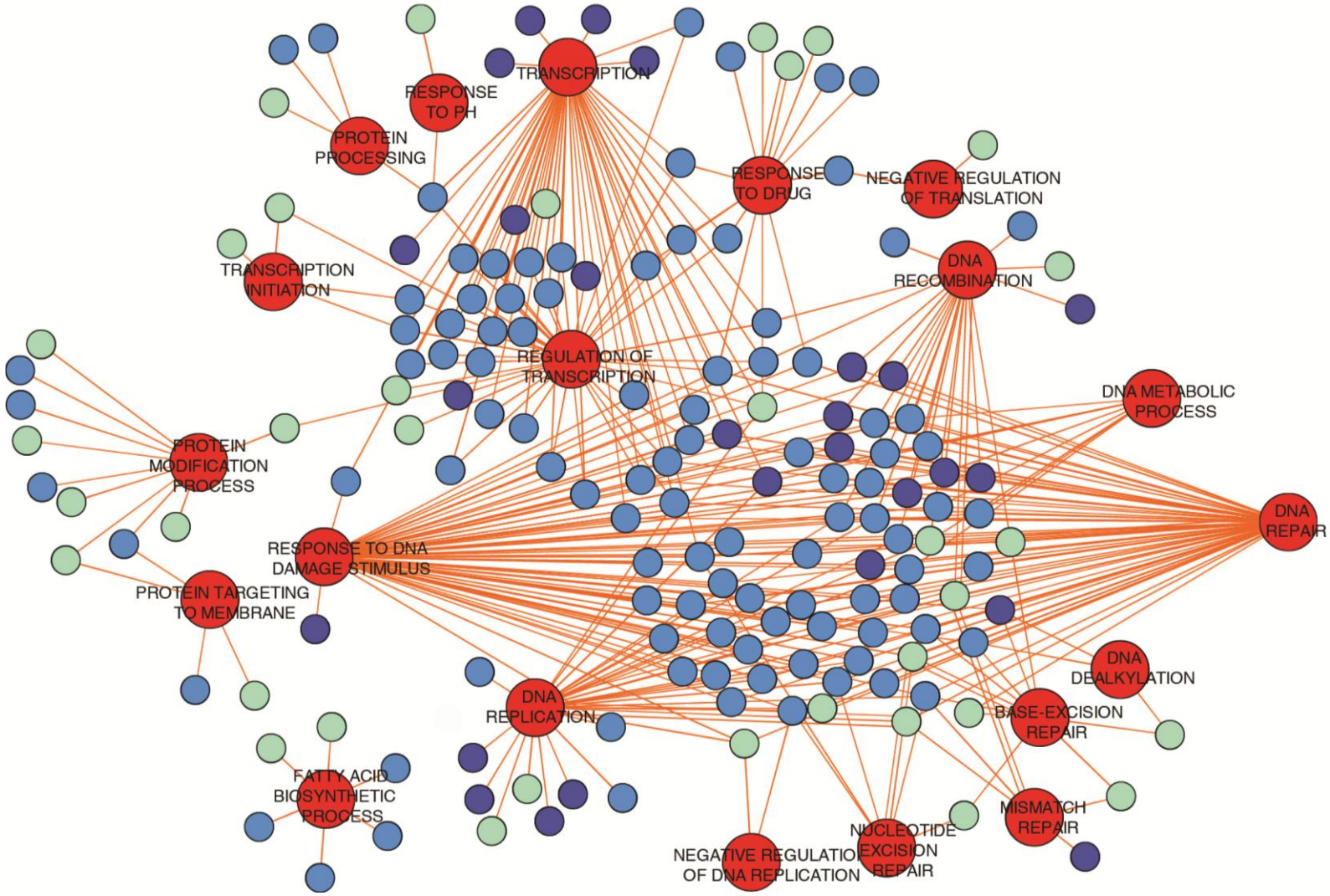


Figure 8. Biological processes necessary for survival after alkylation damage are conserved and when compiled generate a cross-species functionome.
The conserved biological processes with their corresponding genes were visualized using Cytoscape. Red nodes denote conserved biological processes. The remaining nodes are alkylation-modulating proteins belonging to *E. coli* (dark blue), *S. cerevisiae* (light blue) and humans (light green). The edges connect specific biological processes to each protein belonging to the processes. Many proteins are involved in more than one biological process which affects alkylation exposure survival and therefore are connected to several biological processes nodes.

4.3 DISCUSSION

A conserved theme involved DNA metabolism and included the GO hubs of DNA metabolic process, DNA replication, negative regulation of DNA replication, response to DNA damage stimulus, DNA repair, DNA recombination, mismatch repair, nucleotide excision repair, base excision repair and DNA dealkylation. The conserved processes affect several different layers of DNA metabolism with many genes involved in DNA replication and repair. The prevalence of DNA replication and DNA recombination genes emphasize the importance of new DNA synthesis and repairing damaged DNA via recombination mechanisms. Many genes involved in other DNA repair pathways including nucleotide excision repair, mismatch repair and base excision repair are present in the DNA repair hub. The base excision repair subset from all three organisms was predominantly composed of DNA glycosylases that recognize and remove base lesions. This subset contained several known alkylation resistance genes such as the 3-methyl-adenine DNA glycosylases *alka* and MAG1, in *E. coli* and *S. cerevisiae*, respectively. Also included were human DNA glycosylases that recognize and repair oxidative base lesions including UNG, OGG1, TDG and NEIL1 (27, 28, 34, 86). The *alkA* gene product has diverse substrate specificity and also removes oxidative lesions such as xanthine, oxanine, 5-formyluracil and hypoxanthine (89-92). The ability of AlkA to remove oxidation induced lesions correlates with the results seen in the human screen and may affect the sensitivity of *alkA* mutants to alkylation damage. We obtained complimentary results in a different study using another glioblastoma cell line (LN428) with the DNA alkylating agent MMS, where decreased DNA repair activity as measured by the molecular beacon assay and comet assay inversely

correlated with cell survival (**Page 112, Figure 22**) (88). The emphasis on proteins that recognize and remove DNA lesions, and correlation to cell survival, suggests the DNA lesions themselves are cytotoxic, with removal of both alkylation and oxidative lesions being important for survival following alkylation exposure.

Unfortunately, we could not create stable knockdowns of three of the DNA glycosylase hits, likely because the knockdown of OGG1, TDG and NEIL1 created cells with a growth disadvantage compared to cells expressing normal levels of the corresponding protein. Cells lacking oxidative DNA glycosylases may be more sensitive to reactive oxygen species (ROS) because repair of oxidative lesions is compromised. For example, cells lacking NEIL1 are more sensitive to gamma-irradiation (93) and $NEIL1^{-/-}$ mice have decreased expansion of germinal center B cells (94). However, we could effectively reduce UNG mRNA, protein and the cell's ability to repair uracil lesions in the T98G/UNG-KD cell lines. Further, the UNG-KD sensitization was validated using a separate shRNA targeting sequence and experimental design, eliminating the role of RNAi off-target effects or screening artifacts in UNG-KD sensitization.

Although UNG has not been reported to repair alkylation damage, knockdown of UNG sensitizes T98G cells to TMZ. It is possible that TMZ induces toxic oxidative lesions by increasing ROS levels as seen with MMS treatment in yeast (95, 96). TMZ may directly alkylate the electron transport chain proteins and mitochondrial DNA, which could impair electron flow through the chain by direct alkylation damage of the protein and decreased functional protein expression due to replication blocking lesions or mutated DNA. It is possible that mitochondrial alkylation damage induces an increase in ROS formation. Mitochondrial ROS production can increase from defects in electron transport chain proteins, such as complex I (97). This increase in ROS creates lethal oxidative lesions that UNG and other oxidative DNA glycosylases must

repair to prevent cell death after alkylation exposure. NAD kinase, which phosphorylates NAD to form $NADP^+$, was also a hit in our screen. Cells maintain $NADP^+$ in its reduced state, NADPH, to use NADPH's reducing potential in many different pathways including glutathione regeneration for oxidant defense, reduction of RNA to DNA, and synthesis of fatty and amino acids. New synthesis and repair of macromolecules are essential for alkylation survival (73, 78, 98). NAD kinase may affect alkylation survival via NADPH production by its positive effect on DNA and RNA synthesis and also for its central role in creating reducing equivalents for antioxidant defense. Support for a potential role of ROS in alkylation sensitivity was also shown by others using mouse embryonic fibroblasts and several human cancer cell lines, which were treated with the alkylating agent MNNG and contained greater ROS production from both NADPH oxidase and mitochondria sources (99, 100). MNNG cytotoxicity could be completely abrogated by pretreatment with N-acetylcysteine, demonstrating that increased ROS formation may contribute to alkylation toxicity and is a likely source of cytotoxic DNA lesions (100). Although N-acetylcysteine could also prevent MNNG cytotoxicity by being an alkylation target of MNNG and thereby reducing the effective MNNG dose when used as a pretreatment, multiple organisms and cell lines suggest that alkylation exposure can increase ROS formation. Cells expressing UNG may be less sensitive to TMZ because the increase in ROS produces cytotoxic oxidative DNA lesions removed by UNG, promoting cellular survival.

UNG was not the only DNA glycosylase to sensitize cells to alkylation exposure in the validation experiments. MPG and MYH knockdown also sensitized T98G cells to TMZ. This was unexpected because MPG did not sensitize in the siRNA screen, perhaps due to lack of knockdown. Our results are consistent with previous reports from several labs, as MPG-KD led to either an increase or decrease in cell death after alkylation exposure depending on the system

under study (101-104). Further, MPG status alone does not adequately predict response to alkylators, but instead the balance of the entire BER pathway must be investigated to predict sensitivity to DNA damage (87, 88, 105). The ability of MYH-KD to sensitize cells to alkylators was not anticipated since expression of MYH had been previously shown to promote cell death after alkylator exposure (106). In that study, however, the authors used a clonogenic cell survival assay to interrogate cell death related to the O^6-MeG lesion, measuring surviving cells after at least two replication cycles (106). In contrast, in our initial validation experiments, cell survival was determined after 48 hours post TMZ treatment, too short for two replication cycles to occur (107). Therefore, the role of MYH in O^6-MeG mediated cell death is likely very different than its role in cell survival after 48hr exposure to TMZ as described in this study. MYH is primarily known for the removal of the mismatched A opposite 8-oxodG lesions (34) and MYH deficiency does not sensitize cells to H_2O_2, IR, or *cis*-platinum (34, 86, 108). Therefore, the sensitization of MYH depletion to TMZ treatment may be the result of an increase in ROS induced lesions, although less likely due to MYH depletion insensitivity to H_2O_2, and instead could be the result of inhibiting a novel glycosylase function.

After biological validation of the screen, we conducted a cross-species functionome analysis that illustrated that many biological processes are conserved in *E. coli*, *S. cerevisiae* and humans, with multiple proteins from each organism belonging to the different processes. We have used this approach to discover a network of proteins that constitutes a cross-species functionome of evolutionarily important processes essential for survival after alkylation exposure. The functionome analysis affirms that our TMZ screen results from human cells are consistent with previous alkylation studies in other species. Importantly, the analysis distils the crucial processes and corresponding proteins for alkylation survival in humans. The targets

found in the screen may not be possible targets for adjuvant therapy due to adverse cytotoxicity or difficulties in drug development. However, the essential pathways contain other proteins that were not tested in the screen and may also be viable targets to improve sensitivity to alkylating agents. Mining the pathways for genes essential for the biological processes may yield the best targets for improving DNA alkylator therapy. Additionally, the conserved processes and corresponding genes are excellent candidates to investigate using the new large cancer datasets such as the Cancer Genome Atlas and REMBRANDT to determine if patient tumors have varying expression of the genes and if the gene expression correlates with treatment outcome. The differential expression of these crucial genes may serve as important biomarkers to predict outcome of alkylation treatment and which patients would derive most or least benefit from treatment. It would yield possible personalized medicine and potential tumor-specific effects in certain cancers due to decreased activity and expression of these genes. This network links seemingly disparate genes that may prevent cell death after alkylation damage through their crucial biological processes and represent novel targets for adjuvant chemotherapy.

5.0 TARGETING NOVEL BIOLOGICAL PROCESSES TO IMRPOVE TEMOZOLOMIDE SENSITIVITY

5.1 RESULTS

5.1.1 Disruption of protein modification processes sensitizes human cells to alkylators

Based on the functionome network we identified, we chose to further validate another node of proteins to improve sensitivity to alkylators. The protein modifications node was chosen due to enrichment in the cross-species analysis, good sensitization in the screening studies and the observation that this node had multiple genes in both human and yeast cells. The human knockdown cell lines were generated using lentivirus shRNA vectors as described above. Knockdown levels of mRNA were determined using qRT-PCR (**Figure 9A**). The cell lines with mRNA knockdown were then tested for sensitivity to TMZ treatment using a modified MTS assay. The knockdown of UBE3B and ICMT significantly sensitized the glioblastoma cells to TMZ resulting in viability ratios of 0.6 and 0.7, respectively (**Figure 10B**), as compared to the control. We then tested each knockdown cell line for sensitivity to TMZ in a long-term assay similar to a clonogenic cell survival assay. The CyQuant assay utilizes a highly sensitive fluorescent intercalating DNA dye to determine relative DNA content and cell number nine days after TMZ or vehicle treatment. The UBE3B-KD cell lines were approximately 45% more

sensitive compared to control cells in the CyQuant assay at concentrations as low as 50μM TMZ (**Figure 10C**).

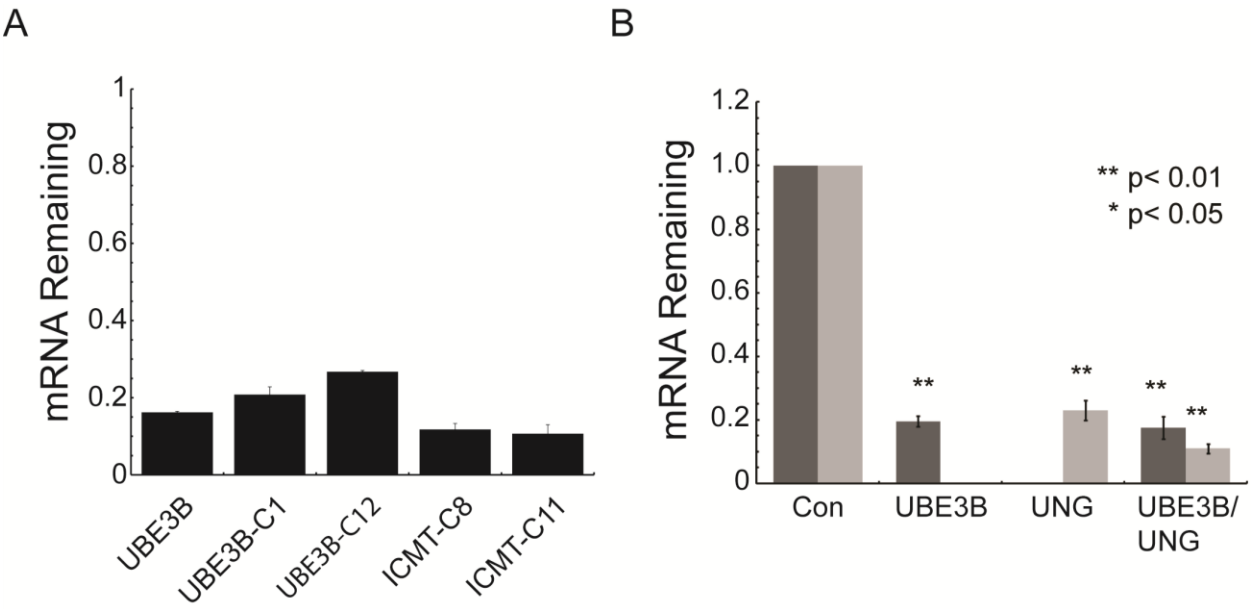


Figure 9. Lentivirus and siRNA mediated knockdown of UNG, UBE3B, and ICMT.
The mRNA was quantified by qRT-PCR. A) Quantification of protein modification processes mRNA knockdown in lentiviral-infected cells as determined by qRT-PCR. TaqMan probes were used to quantify mRNA levels on an Applied Biosystems StepOnePlus machine. The qRT-PCR data was analyzed using the $_{\Delta\Delta}C_t$ method and was normalized to GFP infected plate controls. Gene expression of each gene was normalized to the expression of human ß-actin. The mean of three independent experiments is plotted ± SEM. B) Quantification of UBE3B and UNG mRNA knockdown in siRNA transfected cells as determined by qRT-PCR. Dark gray bars denote UBE3B mRNA levels while light gray bars denote UNG mRNA levels. TaqMan probes were used to quantify mRNA levels as described above. The qRT-PCR data was normalized to scrambled siRNA controls.

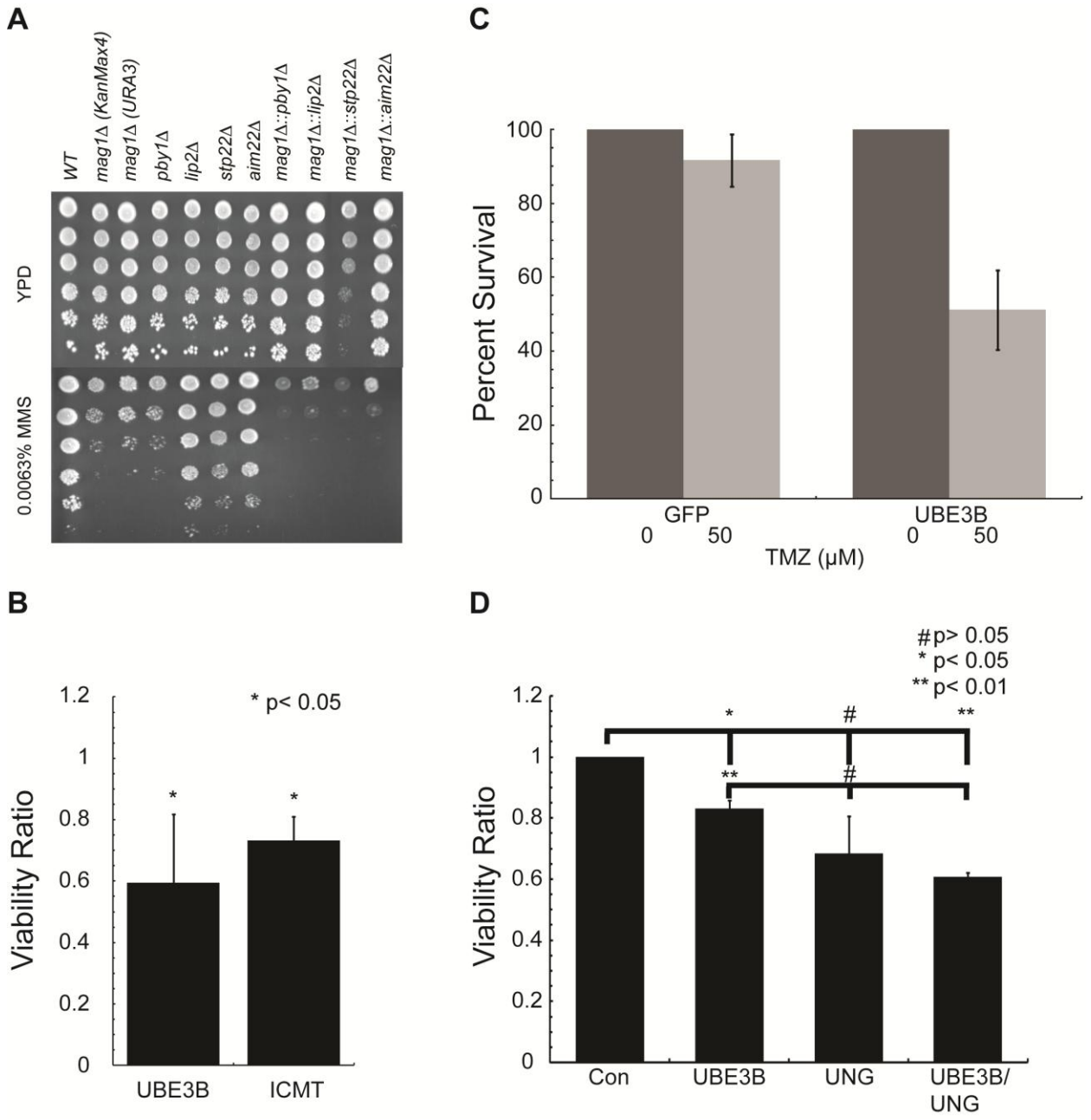


Figure 10. Knockout or depletion of protein modification genes sensitizes *S. cerevisiae* and human glioma cells to alkylating agents.

(A) A dilution series of yeast single and double knockouts were tested in YPD and YPD + MMS containing media. **(B)** Knockdown of protein modification genes sensitizes glioma cells to TMZ. The T98G cell lines with protein modification gene knockdown were treated with 1mm TMZ to determine sensitivity by an MTS assay 48 hours after exposure. The mean viability ratio of three independent experiments is plotted ± SEM. **(C)** Knockdown of UBE3B sensitizes glioma cells to TMZ at clinically achievable doses. The T98G/UBE3B-KD cell line was treated with 50μM of TMZ and survival was determined by CyQuant assay 9 days after exposure. The percent survival of two independent experiments is plotted ± SEM. **(D)** Sensitivity to TMZ via knockdown of UNG and UBE3B are not independent. T98G cells transiently transfected with siRNA that was scrambled, targeting UNG, UBE3B or both UNG and UBE3B. Forty-eight hours after transfection cells were seeded to determine sensitivity to TMZ by MTS assay. Both individual knockdowns of UNG and UBE3B sensitized glioma cells to TMZ, but dual knockdown did

not have a greater effect than either single knockdown, suggesting an epistatic relationship between UNG and UBE3B.

5.1.2 Protein modification gene knockout sensitizes yeast to alkylators, while simultaneous knockout of protein modification genes and an alkylbase DNA glycosylase yields the greatest sensitivity

Our collaborators have previously reported that *aim22Δ, lip22Δ, pby1Δ* and *stp22Δ* cells from *S. cerevisiae,* which belong to the protein modification node, are sensitive to the alkylating agent MMS (77). In addition, it has been firmly established that *mag1Δ* cells are sensitive to MMS because of their alkylbase DNA glycosylase deficiency and inability to repair damaged DNA. We assayed all individual knockouts to further validate our previously reported screening results (**Figure 10A**) and demonstrate decreased growth after MMS treatment for all five mutants. Based on our functionome results, we reasoned that there would be increased alkylation sensitivity when deficiencies in protein modification and base excision repair were combined. We generated double knockouts in the four protein modification associated mutants, using a *mag1Δ* deletion cassette. We demonstrate that there is increased sensitivity in all of the double mutants tested, relative to the individual parent or mag1Δ knockout strains. The protein modification genes correspond to activities involved in the modification of mitochondrial enzymes by the attachment of lipoic acid groups (Lip2), a protein that works with Lip2 (Aim22), a tubulin tyrosine ligase associated with P-bodies (Pby1) and a component of the ESCRT complex that is involved in ubiquitin-dependent sorting of proteins into the endosome (Stp22). We note that Pby1 is homologous to the mouse and human Tsg101 tumor susceptibility genes, which are homologs of ubiquitin-conjugating enzymes implicated in being involved in cell cycle regulation and genome maintenance. The *mag1Δstp22Δ* double mutant is the most sensitive to

MMS, but there also appears to be a growth defect in these strains on untreated media, suggesting some synthetic lethal interaction outside of alkylation damage. Nonetheless, our results with combined protein modification and BER mutants in budding yeast supports the idea that similar dual knockouts in humans may confer increased sensitivity to TMZ.

5.1.3 Dual disruption of both DNA repair and protein modification nodes reveals a potential epistatic relationship between UNG and UBE3B

Although the UNG-KD and UBE3B-KD cell lines were both more sensitive to TMZ when compared to control cells, we were interested in determining if the double-KD cells would have an increased TMZ sensitivity, as was seen for the double KO lines in *S. cerevisiae* (**Figure 10A**). To this end, we investigated if simultaneous inhibition of both DNA repair and protein modification genes increased sensitivity compared to either inhibition alone. Our initial strategy for dual inhibition of DNA repair and protein modification genes used a dual lentivirus approach. The stable UBE3B-KD cell lines were transduced with either a control GFP virus or a virus expressing both GFP and shRNA specific for UNG. Four days after transduction the cells were sorted for GFP expression using Fluorescence Activated Cell Sorting (FACS). The isolated cells were expanded then tested for UBE3B and UNG mRNA expression. Although UBE3B was knocked down at earlier passage numbers, the double knockdown cell lines contained reduced levels of UNG mRNA, but did not have a reduction in UBE3B mRNA (**Figure 11**).

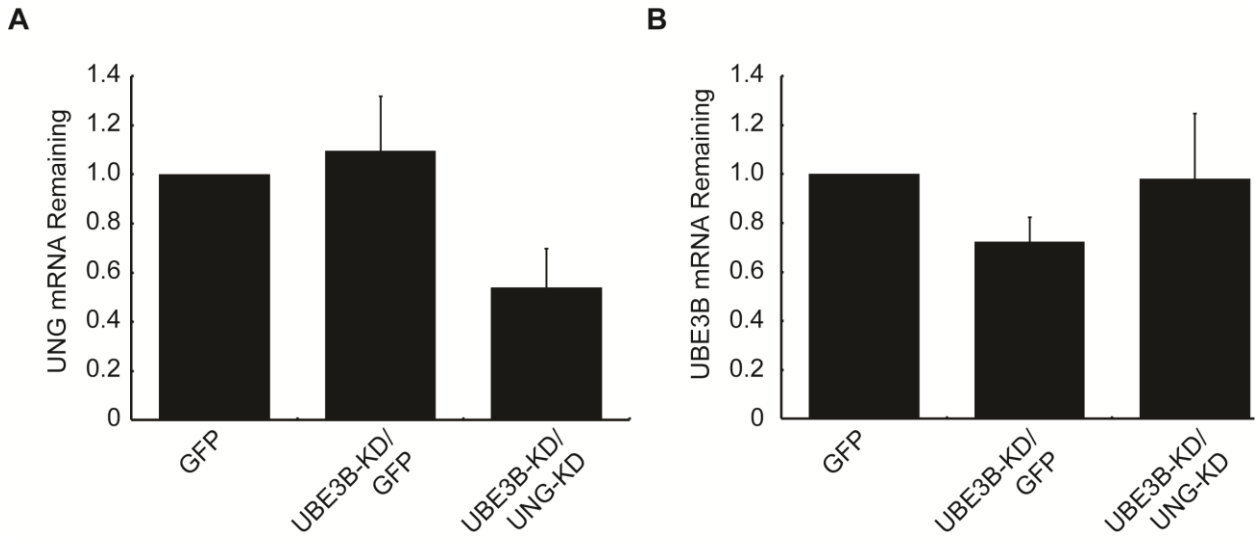


Figure 11. UNG mRNA levels are reduced in the cells containing GFP expression and shRNA against UNG, while UBE3B knockdown is absent in double lentivirus infected cell lines.
(A) Quantification of UNG and **(B)** UBE3B in the double lentivirus infected cell lines as determined by qRT-PCR described above.

However, we noticed during long-term passaging of the UBE3B-KD cell line that sensitivity and knockdown levels diminished, suggesting outgrowth of WT cells in a pooled population. To prevent the loss of UBE3B knockdown we single cell cloned the UBE3B-KD population by limiting dilution. The cells were seeded into 96-well plates and incubated for a week before determining colonies derived from single cells by visual inspection. The clones were expanded before being tested for UBE3B mRNA expression levels. Two clones (clones 1 and 12, shown in gray) had significantly low levels of UBE3B and were selected for lentivirus-mediated knockdown of UNG (**Figure 12**).

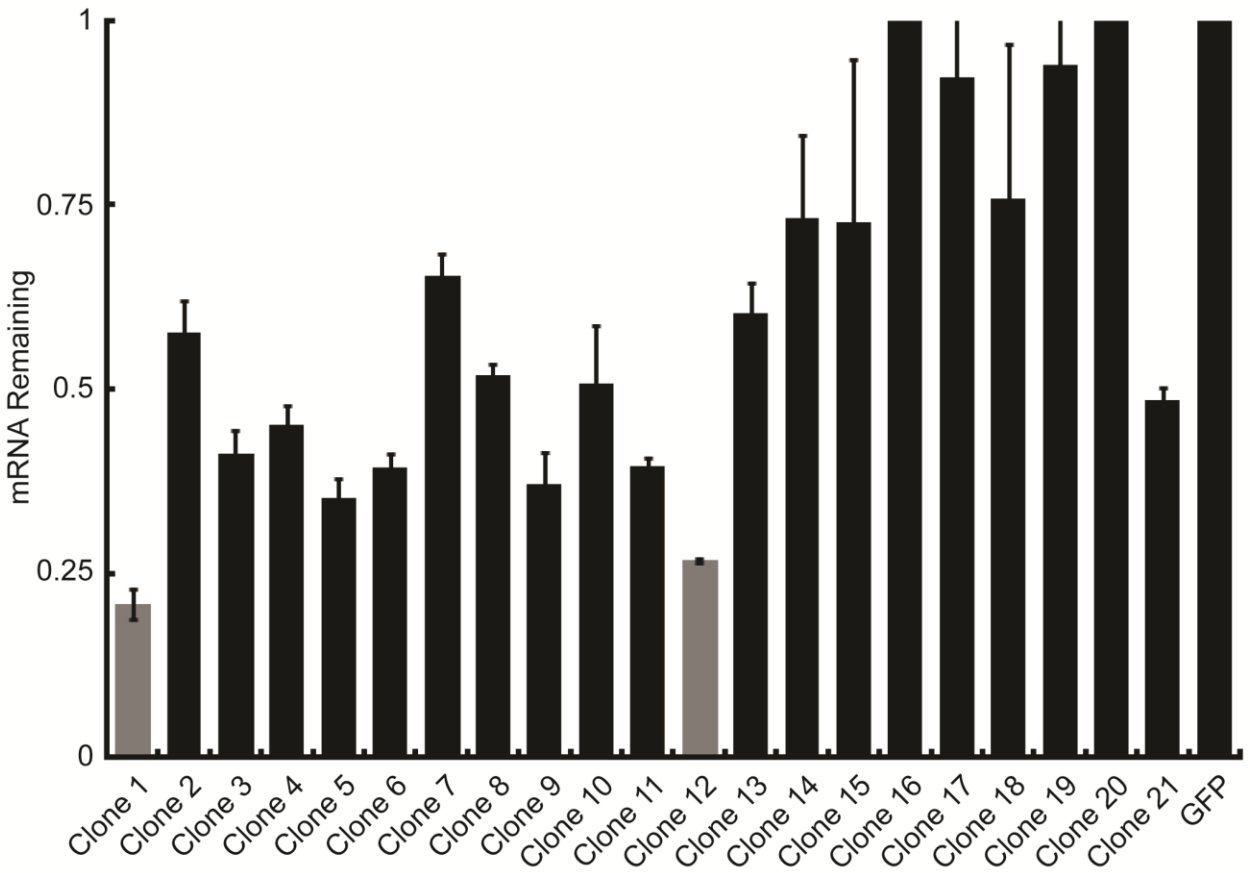


Figure 12. Single cell clones contain significant knockdown of UBE3B compared to GFP controls. The T98G/UBE3B-KD cell line was cloned by single cell dilution and tested for UBE3B mRNA expression levels as previously described. Clones 1 and 12 (gray bars) were selected for double infection with UNG shRNA lentivirus.

The two clones that were double infected were then tested for both UBE3B and UNG mRNA levels. Surprisingly, even though the T98G/UBE3B-KD clones had measurable knockdown before UNG lentivirus transduction, the mRNA validation studies again showed loss of UBE3B-KD (**Figure 13**).

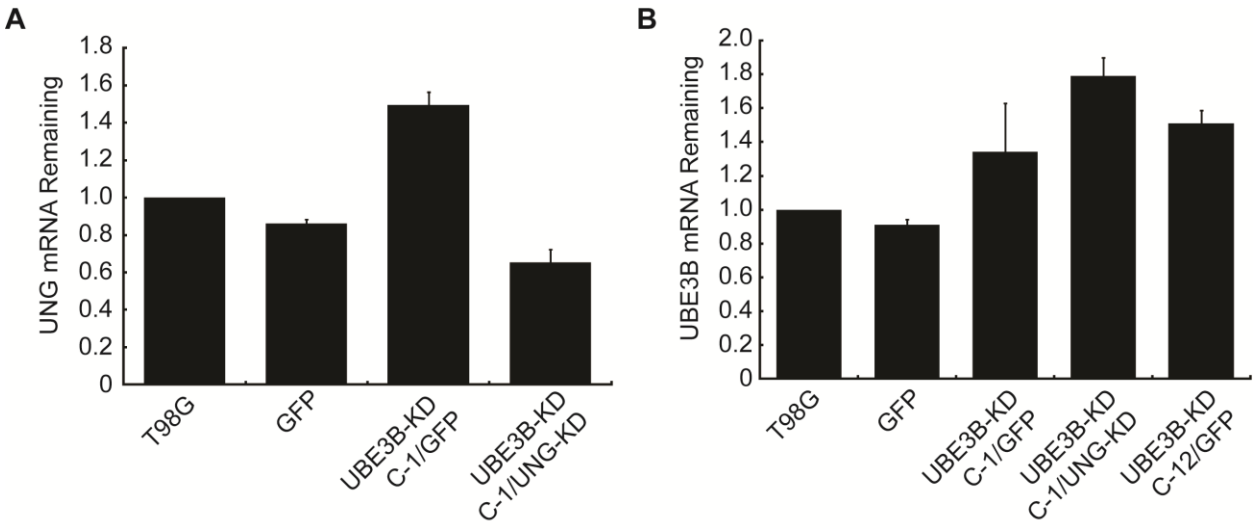


Figure 13. T98G/UBE3B-KD clones lose knockdown of UBE3B with time in culture.
(A) Quantification of UNG and **(B)** UBE3B in the T98G/UBE3B-KD clones 1 and 12 infected with GFP or shRNA against UNG, as determined by qRT-PCR described above.

Because long-term culture of UBE3B-KD cells lost UBE3B knockdown levels, even when derived from single cell clones, we returned to using an siRNA transient transfection approach to knockdown both UNG and UBE3B. We reasoned that only short-term loss of UBE3B is possible in the T98G cell lines since previous methods for stable knockdown had failed. Therefore, we transfected parental T98G cells with a scrambled control siRNA, UNG siRNA, UBE3B siRNA, or UNG plus UBE3B siRNA before determining mRNA levels by qRT-PCR (**Figure 9B**). Cells were seeded for qRT-PCR and sensitivity determination by MTS assay at the same time. Wells with less than 35% mRNA for either UNG or UBE3B remaining were analyzed for sensitivity to TMZ. We expected the dual knockdown of UNG and UBE3B to confer at least an additive effect to TMZ sensitivity, suggesting that the sensitivity observed after knockdown of the proteins resulted from independent mechanisms. However, although UNG and UBE3B knockdown alone both confer sensitivity to TMZ, when they are simultaneously knocked down there is no significant increase in sensitivity to TMZ below the level of either single knockdown (**Figure 10D**). Thus, TMZ sensitivity to knockdown of UNG and UBE3B are

not independent events, demonstrating a possible epistatic relationship between the UNG and UBE3B genes. The conserved biological processes also enrich for possible alkylation sensitivity gene interactions as seen with UNG and UBE3B. The activity of many DNA glycosylases are regulated by post-translational modifications (26). SUMOylation of TDG promotes its catalytic activity by increasing catalytic turnover by decreasing TDG's affinity for the abasic site product (109). Although it is more likely UBE3B has an indirect role in UNG sensitivity, it is possible that ubiquitylation of UNG by UBE3B can modify protein localization, abundance or activity as it has been suggested that UNG is targeted for ubiquitin-mediated proteolysis during the S phase of the cell cycle (110, 111).

5.2 DISCUSSION

Many genes in yeast and humans whose corresponding activities are involved in protein modifications are required for protection from TMZ or methyl methanesulfonate (MMS) exposure, suggesting conservation of gene activation via post-translational modifications to alkylation damage. Several different types of protein modifications including ubiquitylation, methylation, glycosylation and deimination are present in the hub (**Figure 8**). These include two of the four known human peptidyl arginine deiminases, type I and type IV (PADI1 and PADI4) (112). Petidyl arginine deiminases catalyzes the post-translational deimination of proteins by converting arginine residues into citrullines (112). The PADs have distinct substrate specificities and tissue-specific expression patterns (113, 114). The enzymatic reaction is highly dependent on pH and calcium concentrations, with the different PAD isoforms containing different activity profiles (113). The differential expression patterns, substrates and activity profiles all contribute

to the distinct functions of the PAD genes. Currently, the most prevalent function of PADI1 is its role in epidermal differentiation, where it deiminates filaggrin and keratin (113, 114). However, its role in DNA repair or survival to alkylation exposure has not been determined. In contrast, PADI4 is likely involved in DNA repair and cell death signaling as it can modulate p53 signaling and p53 gene target expression levels (115, 116). PADI4 can also antagonize histone methylation by arginine deimination to citruline, thus removing the methylation mark (117). PADI4 also regulates gene expression by associating with histone deacetylase 2 (HDAC2) (118). HDAC2 is known to modulate the DNA damage response and cells depleted of HDAC2 are hypersensitive to DNA damaging agents (119). Further, glioblastoma cells pretreated with the HDAC inhibitor MS275 were sensitized to TMZ (120). These data suggests that PADI4 sensitivity may result from interference with multiple steps in the DNA damage response, likely through deregulation of HDAC2, but PADI4's interaction with p53 likely contributes to alkylation survival as well.

One of the human genes in the node is isoprenylcysteine carboxyl methyltransferase (ICMT), a gene that is ubiquitously expressed and encodes the last of three steps for isoprenylation (121, 122). Methylation of isoprenylcysteine allows movement of the protein from the endoplasmic reticulum to the plasma membrane (122). ICMT has been of interest to cancer researchers because ICMT inhibition can disrupt RAS signaling in several model systems (123-125). Knockout of ICMT prevented oncogenic transformation of cells with oncogenic K-RAS or B-RAF in soft agar and mice models (124). Although there is interest in creating potent and selective ICMT inhibitors, there has been no link for ICMT to DNA repair or response to alkylation exposure (126). ICMT may play a role downstream of the DNA damage cascade and

regulate signaling for cell death, as it has been linked with both autophagic and apoptotic cell death (127).

Another gene in the protein modification group is the human ubiquitin protein ligase E3B (UBE3B), which may play a role in the repair of alkylated DNA. The activity of many DNA repair proteins can be affected by ubiquitylation such as FANCD2 (79) and PCNA (80). Monoubiquitylation of FANCD2 is used as a marker of intra-strand crosslink (ICL) repair activity, while ubiquitylation of PCNA determines polymerase switching (81). Furthermore, the E3 ubiquitin ligase activity of BRCA1, CUL4A and CUL4B is critical for their function in DNA repair (82, 128). The E3 activity of BRCA1 is crucial for its role in cell cycle checkpoint activation and sensitivity to DNA damage, while CUL4A and CUL4B E3 activity is required to regulate chromatin structure and access of DNA repair proteins to the DNA in nucleotide excision repair (NER) (82, 128, 129). The potential role of UBE3B and its substrates in DNA repair are of great interest, but have yet to be defined.

We sought to create greater sensitivity to alkylators by simultaneously knockdown of both UNG and UBE3B. We expected an additive effect on sensitivity to TMZ that would denote independent effects of UNG and UBE3B knockdown. To our surprise, the double knockdown cell line had no increase in sensitivity compared to the single knockdown cell lines, suggesting that UNG and UBE3B knockdown do not have independent effects on TMZ sensitivity and represent an epistatic interaction. This is likely an indirect interaction with downstream targets of UBE3B modulating TMZ cytotoxicity through a UNG dependent pathway. However, a direct interaction is possible, since UNG is a known target of ubiquitylation (111, 130, 131), although these data would not implicate UBE3B in the cell cycle-regulated proteolysis of UNG. Further, the activity of many DNA repair proteins are modulated by ubiquitylation and several essential

DNA repair proteins are E3 ubiquitin ligases (79-82, 128). The epistatic interaction between the two genes supports the possibility of UBE3B playing a role in modulating DNA repair or survival to DNA damage. Determining how UBE3B and UNG regulate alkylation survival and their mechanism of interaction is an area of interest.

Any kind of functional interaction between UBE3B and UNG was unexpected because the probability of selecting two genes from the alkylation functionome that have a functional interaction to impact alkylation survival appeared small. The lack of independence on UNG and UBE3B mediated alkylation cytotoxicity is likely due to the enrichment analysis that selected for proteins that could be members of the same essential pathway for alkylation survival, thereby eliminating independent events. The analysis of the three screens revealed that many biological processes that modulate survival after alkylation damage are conserved through evolution and supports the validity of the screen, which suggests that novel processes are involved in repairing alkylation damage. The conserved biological processes importance to alkylation survival is underscored by the sensitivity of the cell lines depleted of various protein modification genes. These yeast and human cell lines were more sensitive to alkylating agents in both short-term and long-term clonogenic-like assays. Both DNA repair and protein modification gene depletion enhanced TMZ efficacy in glioma cells; however, the greatest alkylator toxicity in yeast cells were generated by dual inhibition of both processes. With multiple avenues to develop alkylation resistance, simultaneous inhibition may be required to achieve a clinically significant sensitization for chemotherapies.

6.0 UBIQUITIN LIGASE E3B AND URACIL DNA GLYCOSYLASE AS TARGETS FOR CANCER THERAPEUTICS

6.1 RESULTS

6.1.1 UBE3B is differentially expressed in various tumor types

Expression of UBE3B affected GBM sensitivity to TMZ and may impact response of various cancers to other chemotherapeutic alkylating agents such as cyclophosphamide, chlorabucil and dacarbazine. With little known on the function or expression of UBE3B in normal and cancerous tissue, we wanted to determine if UBE3B expression levels vary in different cancer cell lines. To determine if UBE3B is differentially expressed in different cancer types we quantified relative expression levels of UBE3B mRNA by qRT-PCR. We discovered that UBE3B mRNA expression varies by approximately 7-8 fold in the tumor cell lines tested (**Figure 14**). Although this is a small sample size, it reveals that levels of UBE3B can vary dramatically in different cancers and may lead to an increase in therapeutic index in certain cancers. We were then interested in determining of variations in UBE3B and UNG also vary in patient primary tumors and if these levels correlated with outcome or response to therapy.

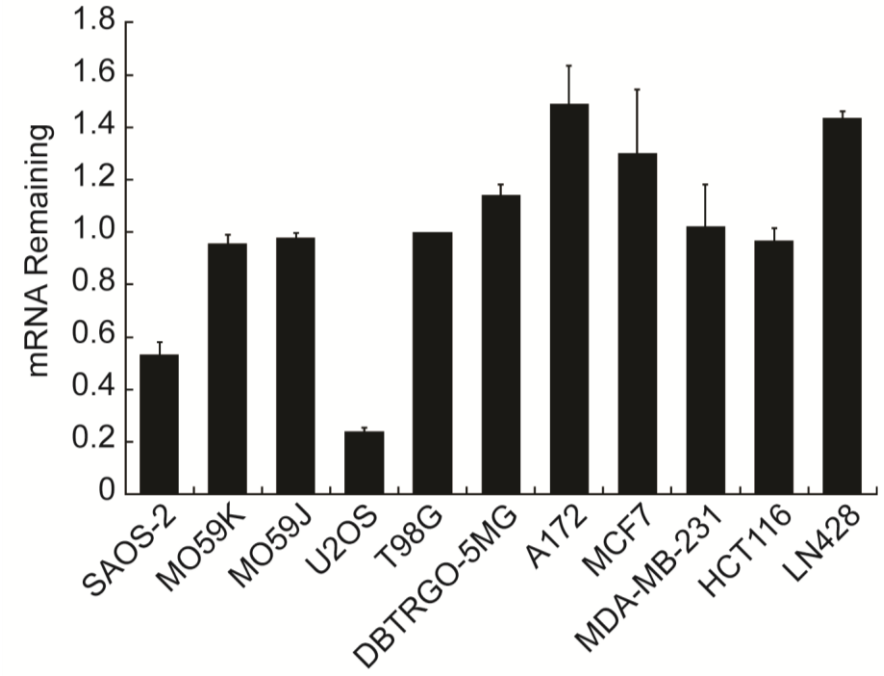


Figure 14. UBE3B is differentially expressed with a seven-fold difference in cancer cell lines.
The UBE3B mRNA levels were determined in the different cancer cell lines using qRT-PCR as described above (**Page 24**). The UBE3B mRNA levels from the cancer cell lines were normalized to the T98G cell line. The cell lines used were a collection of glioblastoma (MO59K, MO59J, T98G, DBTRGO-5MG, A172, LN428), osteosarcoma (SAOS-2, U2OS), breast cancer (MCF7, MDA-MB231) and colon cancer (HCT116) cell lines.

6.1.2 UNG is over expressed in GBM, while UBE3B maintains similar levels of expression in GBM compared to normal controls

The change in UBE3B expression in our different cancer cell lines intrigued us and we sought to determine if UBE3B and UNG levels vary in the LN428 and T98G cancer cell lines and in clinical samples when compared to normal controls. In a separate project, we have run Affymetrix microarrays on our different cell lines to measure mRNA expression levels and we determined our GBM cell lines had levels of UNG which varied considerably from that of normal primary human astrocytes, while the UBE3B mRNA levels were similar to that of the

normal astrocytes (**Table 12**). The increased UNG expression for both LN428 and T98G cell lines was significant, as evidenced by very small p-values.

Table 12. Relative expression of UNG and UBE3B mRNA in cell lines used compared to normal human astrocytes using an Affymetrix microarray.

	Relative Gene Expression	
Cells	UNG	UBE3B
Normal Human Astrocytes	1	1
LN428	5.17776 p-value = 3.44972×10^{-7}	0.816376 p-value = 0.216969255
T98G	3.354535 p-value = 2.84847×10^{-6}	0.85976429 p-value = 0.302796239

We were interested in determining if similar increases in UNG expression levels are present in primary patient tumors. We investigated this by interrogating The Cancer Genome Atlas (TCGA) dataset on GBM. GBM was one of the pilot cancers for the TCGA and is completed with publically available data (132). The frequency of two-fold decrease or increase in expression of UNG and UBE3B is shown in **Table 13**. UNG expression is two-fold greater in over 14% of the GBMs studied in the TCGA while UBE3B has two-fold less expression in greater than 19% of samples. Interestingly, there are no samples with either a decrease in UNG expression or an increase in UBE3B expression in TCGA files. We also queried the Rembrandt database for UNG and UBE3B gene expression data and found that similar to the TCGA, the levels of UNG tended to be higher in glioblastoma compared to normal controls and that UBE3B tended to be lower in GBM compared to normal controls (133, 134). These results from a separate database reinforced our conclusion that expression levels of our proteins of interest varied significantly in tumors. Importantly, the directionality of the changes are important as it appears over-expression of UNG is beneficial for GBM, particularly in the absent of any samples with decreased expression in either tumor database. We also graphed the change in UNG expression for all different brain tumors and observed that all brain tumors had higher median

expression of UNG, with GBM having the highest expression of all brain tumors, compared to normal controls (**Figure 15**). The expression data on UBE3B is less impressive due to the smaller amount of samples that deviate from the normal tissue in Rembrandt, although several samples do contain significantly less UBE3B compared to controls in TGCA.

Table 13. Frequency of UNG and UBE3B relative expression differences compared to normal controls in The Cancer Genome Atlas and REMBRANDT databases.
The frequency of 2-fold decrease or increase in UNG and UBE3B gene expression data is recorded below. This data was collected from the Rembrandt Database (133).

	UNG		UBE3B	
	2x Decrease	2x Increase	2x Decrease	2x Increase
The Cancer Genome Atlas	0%	14%	19%	0%
Rembrandt Database	0%	33.77%	1.1%	0%

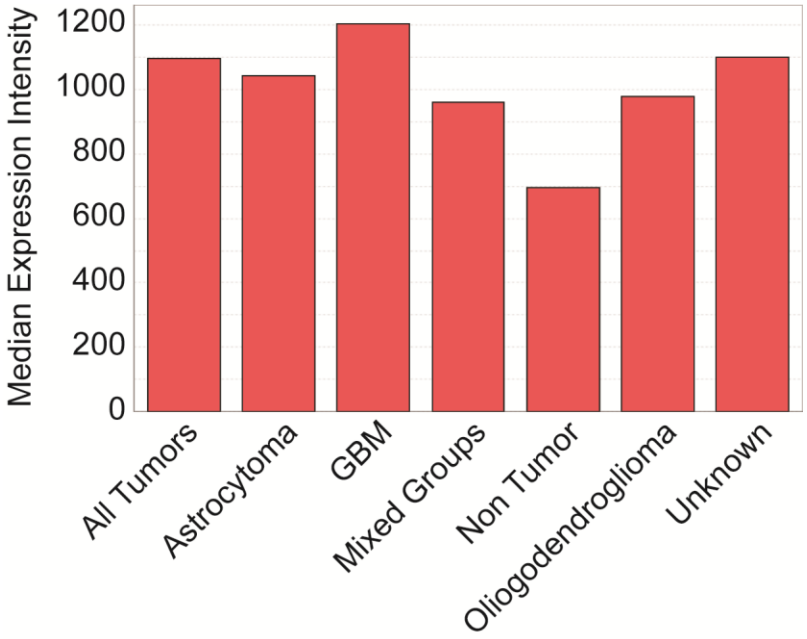


Figure 15. UNG is expressed at higher levels in all brain tumors, but particularly in GBM, when compared to non-tumor controls.
The Rembrandt brain tumor database expression data was queried for expression of UNG based on tumor types and normal controls. The median expression intensity was plotted against the cell of origin and all brain tumors contained higher levels of UNG median expression intensity, with GBM the highest of all measured tumor types. This data was collected from the Rembrandt Database (133).

Knowing that the expression levels of UNG and UBE3B fluctuate in primary tumor samples, we desired to determine if expression levels of UNG and UBE3B correlated with

response to therapy or survival. We would predict the increase in UNG expression levels would correlate with reduced survival due to the ability of the cells to repair their DNA more efficiently, especially TMZ induced lesions.

6.1.3 UNG expression correlates with a worse prognosis in brain tumors

We then queried the Rembrandt database for patients with GBM tumors and over-expression of UNG and plotted a Kaplan-Meier survival curve to determine if over-expression of UNG correlated with decreased survival (**Figure 16**). Surprisingly, the over-expression of UNG did not correlate with decreased survival as expected, although we believe this was due to the small sample size of GBM tumors in the database, as there appeared to be a trend of decreased survival, particularly around 2000-4000 days in the study. Therefore, we plotted a Kaplan-Meier survival curve of all glial tumors over-expressing UNG compared to tumors containing an intermediate level of UNG (**Figure 17**). We determined that two-fold over-expression of UNG correlated with a decrease in probability of survival compared to intermediate expression of UNG, with a log-rank p-value, calculated using the Mantel- Haenszel procedure, of 1.0177×10^{-4}. This correlation matches our prediction based on our *in vitro* data, suggesting that increased expression of UNG negatively impacts brain tumor survival, possibly from increased resistance to chemotherapeutic treatment.

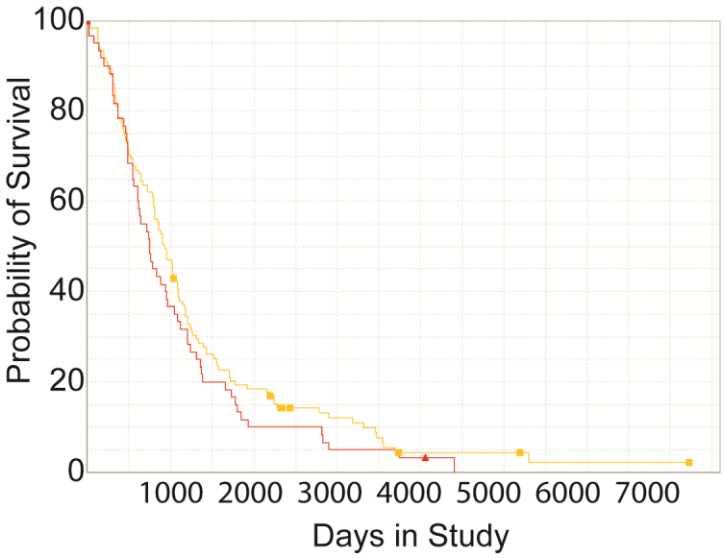


Figure 16. The over-expression of UNG in GBM tumors does not significantly impact patient survival.
The Kaplan-Meier survival curve in GBM tumors reveals reduced survival for patients with over expression of UNG (Red trace) compared to that of intermediate expression of UNG (Yellow trace) with a log-rank p-value, calculated using the Mantel- Haenszel procedure, of 0.2389423205. This data was collected from the Rembrandt Database (133).

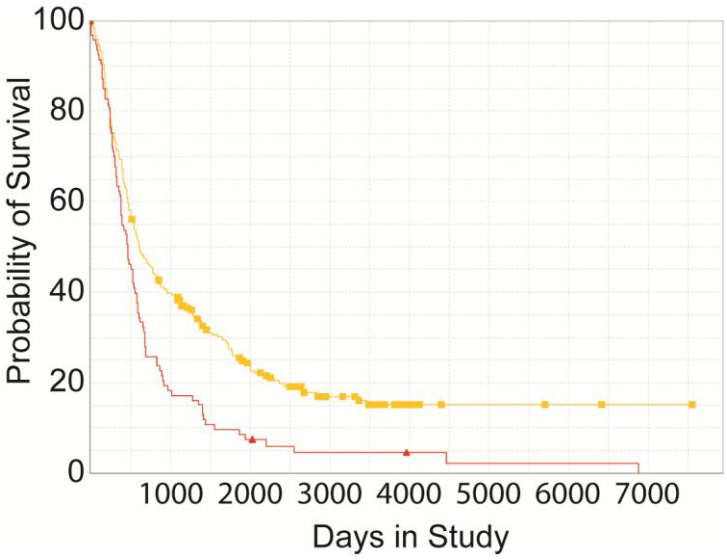


Figure 17. The over-expression of UNG negatively impacts survival of all glioma tumor patients.
The Kaplan-Meier survival curve in all glial tumors reveals reduced survival for patients with over expression of UNG (Red trace) compared to that of intermediate expression of UNG (Yellow trace) with a log-rank p-value, calculated using the Mantel- Haenszel procedure, of 1.0177×10^{-4}. This data was collected from the Rembrandt Database (133).

6.1.4 Cloning of UBE3B

To learn more about the function of these proteins in response to alkylation treatment we began to clone them from our LN428 glioblastoma cell line. We were successfully able to isolate four different sequences of UBE3B. One of the clones matched the wild-type published sequence found in PubMed. However, the other three clones contained point mutations, with some of them conserved across the different clones, suggesting the point mutants were not the result of the cloning reaction. One of the clones also contained a stop codon fifty amino acids into the HECT domain of UBE3B. The HECT domain contains the E3 ubiquitin ligase domain of UBE3B and deletion of almost the entire HECT domain likely creates an inactive protein. The other mutations have unknown effects due to the little information known on UBE3B structure and function. The sequences of the different UBE3B sequences that we cloned are found in Appendix B (**Page 120**). Complementation experiments with inactive UBE3B would be possible using the clone missing the HECT domain and would determine if the ubiquitin ligase activity is required for UBE3B's role in alkylation survival.

6.2 DISCUSSION

We have demonstrated that expression of UBE3B and UNG correlate with tumor cell line survival when treated with the clinical alkylator TMZ. We were interested in determining if these two proteins are differentially expressed in different cancers. We found that expression of UBE3B can vary as much as eight fold in the cancer cell lines we studied. Using the Cancer Genome Atlas (TCGA) and the Repository for Molecular Brain Neoplasia Data

(REMBRANDT) data sets we discovered that UBE3B and UNG expression vary in GBM when compared to normal controls. Although, UBE3B was only found to be either down regulated or contain similar expression to controls, the expression pattern of UNG in brain tumors supported it as a potential target in GBM. UNG was over expressed in brain tumors in general when compared to normal controls. However, it was interesting to note that over expression of UNG was highest in GBM when being compared to even other brain tumor types. Over expression of UNG correlated with a decrease survival in all glioma, suggesting that UNG expression is deleterious to brain tumor treatment or increases tumor growth. These clinical data correlate with the screening results and pre-clinical testing.

How UNG is playing a role in patient survival is still unknown. The pre-clinical data suggests it may be due to increased resistance to chemotherapeutics such as TMZ. However, this is not certain. How UNG impacts alkylation survival is still uncertain. UNG appears to play a role in cellular proliferation because of its involvement with CENP-A in mitosis. It is possible that UNG over-expression enables increased tumor growth as decreased expression of UNG decreased proliferation in cell models (57). While UNG's role in repair of oxidative DNA lesions is a plausible explanation for contributing to alkylation survival, there are still experiments needed to prove this relationship. Adding back separation of function point mutants, such as those used in studying class-switch recombination, to the knockdown cell lines and testing sensitivity on these cells would be very informative. The mutants used would be deficient in DNA glycosylase activity to determine if repair of DNA is crucial, or have mutations in the N-terminal domain as that is required for class-switch recombination for unknown reasons. If the unknown function of the N-terminal domain is required, a new function potentially relating UNG and double strand break repair would be an interesting avenue to pursue.

7.0 MEASUREING REAL-TIME DNA REPAIR RATES WITH MOLECULAR BEACONS

7.1 RESULTS

7.1.1 Using different DNA lesions to probe catalytic activity of specific DNA glycosylases

After validating the molecular beacons with either over-expression (MPG) (Figure 20, Figure 21) or knockdown of glycosylases (UNG) (Figure 7), we were interested in determining if a variety of well studied base lesions were repaired in unaltered control cells. We used our new beacon design to improve the fluorescence signal (Table 7) and varied the base opposite the THF lesion, to determine if the base opposite the abasic mimic affected catalytic rate in the T98G nuclear extracts previously prepared. To our surprise, the small differences in beacon design repeatedly revealed that the base opposing the THF lesion can drastically affect strand scission by APE1 (Figure 18). The strand scission by APE1 was very similar when thymine, adenine and guanine bases were opposite the THF lesion, with a normalized maximum fluorescence approaching 1.0. Strand scission was markedly reduced in molecular beacons with a cytosine opposite the THF lesion, as it obtained a maximum fluorescence of approximately 0.7 of normalized maximum fluorescence. Although APE1 is not the rate limiting step for BER (135), the bases opposite the lesion may affect glycosylase lesion removal rates and must be considered

when comparing DNA repair rates and beacon designs. The stand scission step of THF opposite cytosine bases in LN428 nuclear extract was similarly slower compared to adenine, thymine and guanine.

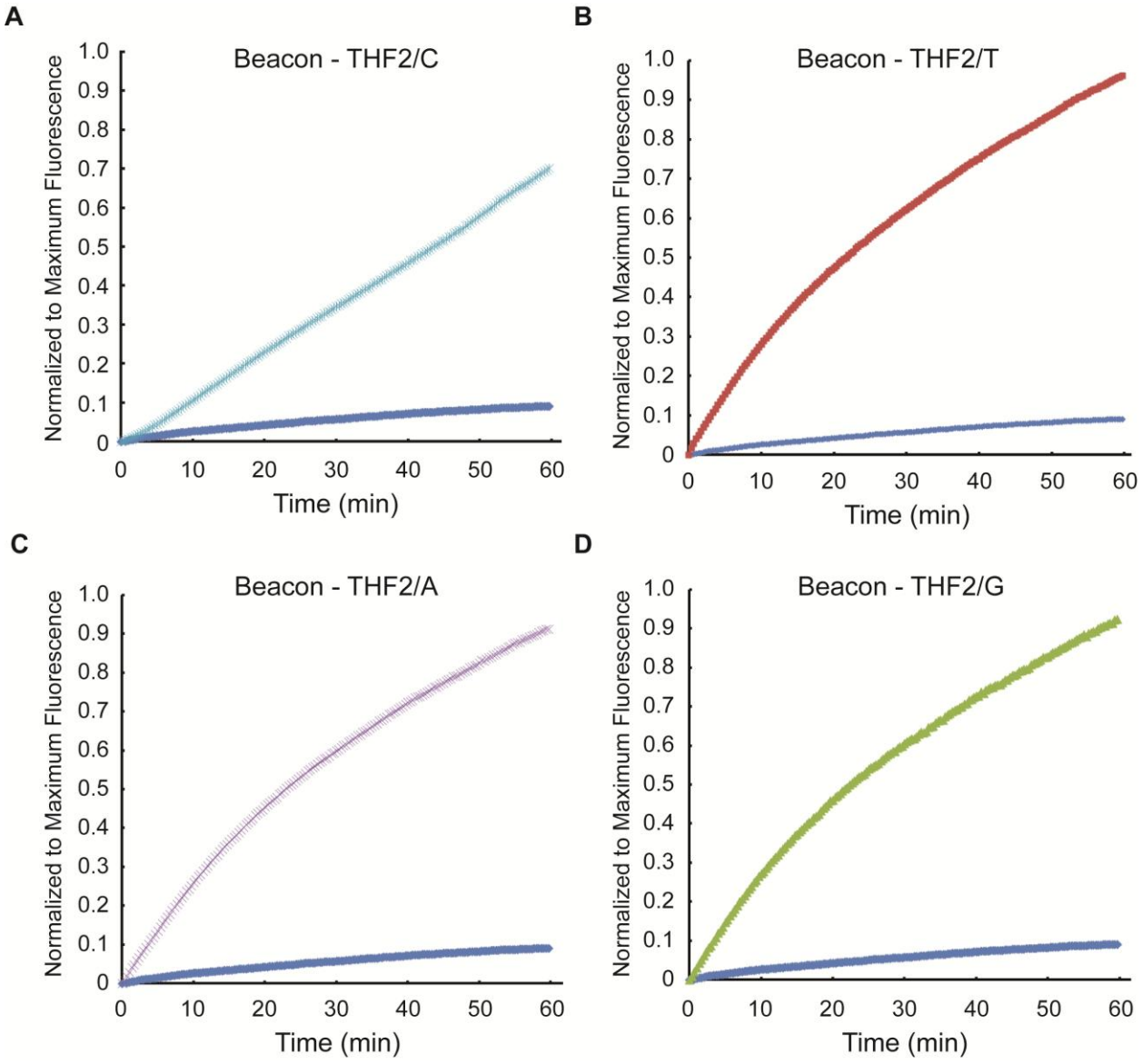


Figure 18. The base opposite THF DNA lesions can affect APE1 strand scission.
Using the molecular beacon assay we can sensitively measure DNA repair rates of several different enzymes. The strand scission step mediated by APE1 is affected by the base opposite the lesion. The molecular beacon assay was performed and analyzed as described above (**Pages 26-33**). The control beacon with no lesion was used as a negative control and is the blue trace in all four graphs. The positive control beacon containing the THF lesion was opposite a **(A)** cytosine, **(B)** thymine, **(C)** adenine or **(D)** guanine nucleotide. The strand scission by APE1 was very similar when thymine, adenine and guanine bases were opposite the THF lesion, with a normalized maximum fluorescence approaching 1.0. Strand scission was markedly reduced in molecular beacons with a cytosine opposite the THF lesion, as it obtained a maximum fluorescence of approximately 0.7 of normalized maximum fluorescence.

We then tested the repair capacity of the T98G nuclear extracts using all of our currently available molecular beacons (**Table 7**) to determine if we could measure repair activity to each lesion and if we could create a baseline to compare with the previously made glycosylase knockdown cell lines (**Figure 19**).

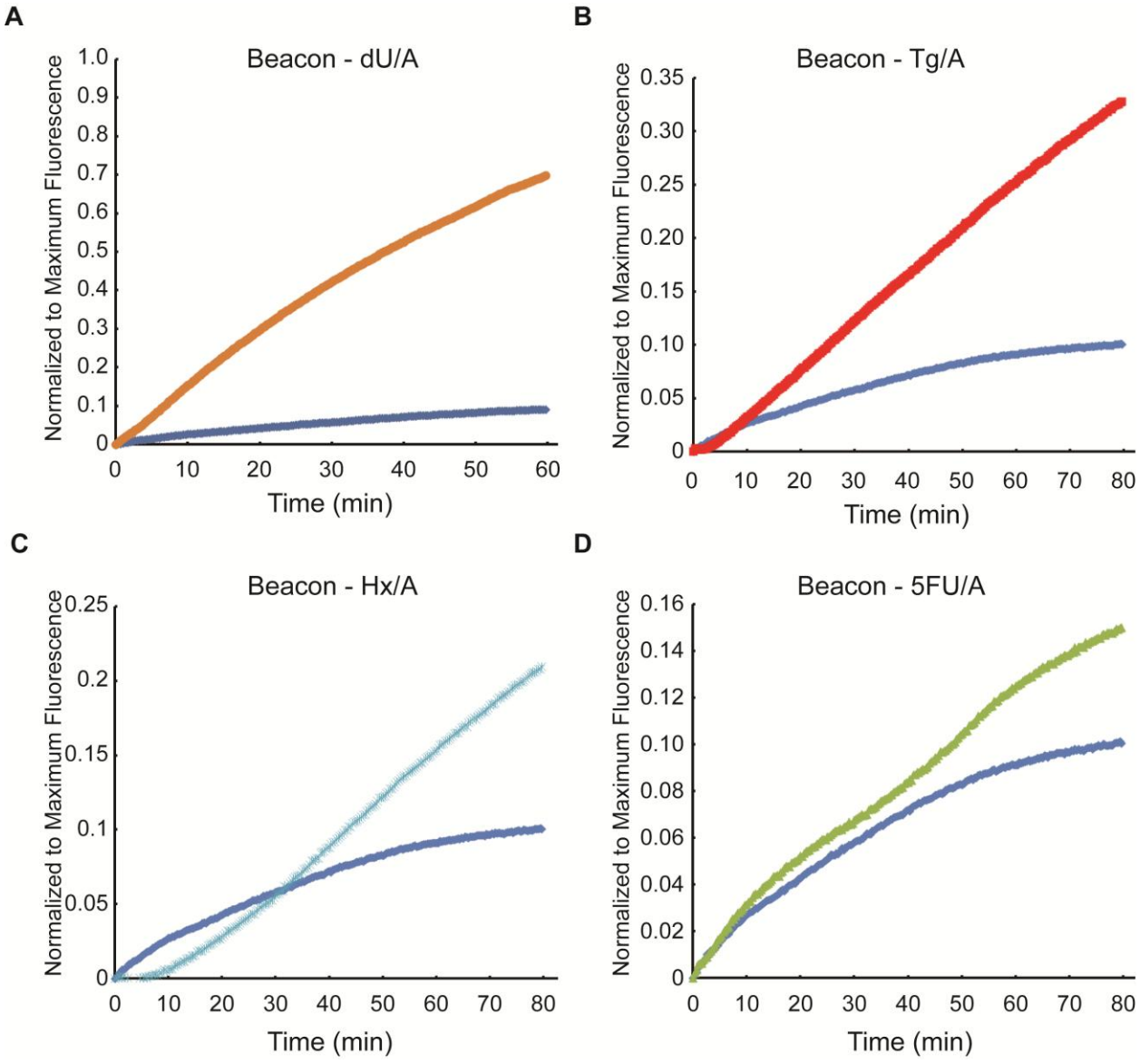


Figure 19. T98G nuclear extracts endogenous proteins can remove several oxidative DNA lesions. The T98G nuclear extracts were used to test beacon design of several oxidative DNA lesions and endogenous protein levels were sufficient to remove the DNA lesions, but at drastically different rates. **(A)** The removal of 2'-deoxyuridine (orange) was the fastest lesion removed tested that required a DNA glycosylase. The rates of **(B)** thymine glycol (red), **(C)** hypoxanthine (light blue), and **(D)** 5-fluorouracil (green) were much slower, but did appear to differ from that of the control beacon (blue trace in each panel), signifying repair of the molecular beacon.

Excluding the THF beacons that measure APE1 activity, the 2'-deoxyuridine beacon was the fastest repaired beacon. Although uracil can be removed by UNG, SMUG1, TDG and MBD4, the uracil removal activity seen in the assay is likely from UNG as demonstrated previously using the T98G and T98G/UNG-KD cell lines (**Figure 7**). The removal of thymine glycol (Tg) was not as robust as that of 2'-deoxyuridine yet it was still clearly elevated when compared to the control beacon. The hypoxanthine and 5-fluorouracil had the lowest removal rates. A low level of hypoxanthine repair was expected because we cannot detect MPG expression in the T98G cell lines by immunoblot. However, the slow removal of 5-fluoruracil (5-FU) was interesting because several DNA glycosylases are able to remove that DNA lesion including UNG, SMUG1 and TDG (**Table 3**). Although 5-FU was repaired, its repair was drastically slower than other beacon repair rates, even though 2-deoxyuridine is avidly removed. This is likely related to the rate of 5-FU removal in a cellular extract, as previous *in vitro* assays have incubated DNA incorporated with 5-FU for 20 hours before determining DNA lesion removal via electrophoresis (136). Those experiments still had substantial substrate remaining after 20 hour incubating with 20 μg of nuclear proteins, which is 17 times longer and twice as much protein we used in the molecular beacon assay (136).

8.0 CONCLUSIONS AND DISCUSSION

8.1 DISCOVERY OF TMZ SENSITIZING GENES BY SIRNA SCREEN

GBMs are the most common and aggressive primary brain tumor with the current standard of care consisting of maximum surgical removal, radiation and TMZ (3-5). Many tumors are refractory to TMZ treatment with very poor outcomes. Difficulties in treating GBM make it a candidate for drug research since an increase in therapeutic options is necessary to improve survival rate. Because TMZ clinical efficacy is due to the O^6-MeG lesion, much of the alkylation damage sensitization research has focused on this lesion. Unfortunately, previous attempts to improve efficacy of O^6-MeG lesion cytotoxicity have been futile. The clinical trials testing the effectiveness of MGMT inhibitors (21, 23) do not show an increase in efficacy, particularly in TMZ-resistant GBM (22, 24, 25). Due to the inherent difficulties and resistance of GBM to chemotherapies, any increase in cancer sensitization could impact patient outcomes.

In these studies we investigated both new ways to improve cancer chemotherapy efficacy and ways to measure DNA repair which correlate with therapy resistance. We sought to improve efficacy of TMZ, independent of the O^6-MeG lesion. By using a synthetic lethal siRNA screen toward "druggable" targets we biased the potential targets based on what genes are known to bind drug-like molecules, but otherwise we allowed the biology to direct our search for novel genes that modulate TMZ toxicity. This enabled us to discover genes that impact DNA alkylator

survival, but have no known function related to DNA repair. The approach allowed us to find the epistatic interaction between UNG and UBE3B that would not have been found by a candidate gene approach.

Although we have learned much about essential biological processes for alkylation survival that are conserved from bacteria to humans, we have learned from the difficulties in the screen and would likely have modified our approach after obtaining more practical experience. One important adjustment would be the dose and duration of TMZ treatment. Knowing more about alkylation survival, we would use a lower TMZ dose more likely to be found in patients, such as 100 μM, and measure cell survival after a longer duration such as the nine days used in the CyQuant assay. The short time frame used in the screen precluded observation of many alkylation cytotoxicity effects as alkylators generally take several replication cycles to produce cytotoxic events. This would be more challenging to undertake in the screen using 384-well plates due to the small area for growth, but modifications such as using stable cell lines to reduce days needed for siRNA transfection incubation or increasing well size by using 96-well plates could have made this possible. One benefit of using the 384-well plates compared to 96-well plates was our ability to carry out the screen in triplicate to generate more reliable data for selection of positive hits, due to the lower cost of reagents when using a 384-well plate. The increased replicates increased our statistical power and made our analysis more likely to represent true results. The inclusion of positive and negative controls for both transfection and TMZ dosing would have been helpful, since the screen only contained negative controls for both siRNA transfection and TMZ. By transfecting a lethal siRNA we could have obtained a positive control for transfection efficiency. Unfortunately, we were not aware of any genes that would

improve alkylation toxicity in the short-term when we began the screen, but by using a longer exposure to TMZ we could knockdown MGMT and increase T98G sensitivity to TMZ.

To address the need for cancer selectivity, screening a normal control cell line in parallel to the T98G cell line would have been beneficial. Although this would have doubled the cost of the screen, discovering the important genes and biological processes used in normal cells to survive alkylators is crucial. With this information, essential sensitizing genes and biological processes from the T98G cell line could be compared to the normal control cell line, with overlapping genes and processes removed from further study. This would focus target selection on genes that preferentially kill GBM cells when treated with TMZ or genes that selectively promote normal cell survival. This would likely result in lower possible TMZ doses and enhanced tumor selectivity, which would reduce dose-limiting side effects and improve TMZ efficacy. Selection of the correct normal control line for the comparison would be essential. TMZ's cytotoxic effects on bone marrow cells causing myelosuppression, thrombocytopenia and neutropenia, can limit TMZ dosing (137). Therefore, we propose that selecting a human bone marrow cell line, for the normal control would be critical for the comparison. If practical constraints limit the control cell line to an immortalized cell line a myeloid precursor cell line such as the 32D cell line (138). If the 32D cell line is not amenable to the screening procedure, Jurkat cells would be a possible, although less ideal control. Jurkat cells are immortalized T-cells derived from an acute T-cell leukemia patient and would yield baseline knowledge of bone marrow derived cells TMZ resistance mechanisms (139). Current efforts to increase bone marrow resistance to TMZ induced O^6-MeG lesion toxicity includes transplanting bone marrow overexpressing mutant MGMT protein (140). The approach involves creation of bone marrow cells that stably express the MGMT protein with the P140K mutation, which render MGMT

insensitive to O^6-benzyl-guanine inhibition (141). Although the gene therapy methods have shown promise in pre-clinical trials, they have not yet entered clinical trials (142). Studying TMZ resistance mechanisms in bone marrow-derived cell lines would yield potential targets to improve bone marrow survival and enable specific targeting of tumor resistance mechanisms, both resulting in increased TMZ efficacy.

Even though the screen was not ideally prepared, we still discovered several important genes that contribute to TMZ sensitivity and uncovered some mechanistic insight into their survival mechanisms. The genes we discovered to modulate DNA alkylator cytotoxicity were validated by using distinct shRNA vectors against the genes of interest and retesting sensitivity to TMZ. This eliminated off-target effects of the siRNA and shRNA vectors and ensured we were pursuing true positive results. We found that many genes and processes not previously thought to play a role in TMZ toxicity significantly impact alkylation sensitivity. For example, UNG was previously known to remove uracil-based DNA lesions, interact with CENP-A and is required for class switch recombination. However, UNG's role in DNA alkylator survival was previously unknown. Due to the multiple functions and pathways UNG is involved in, it is impossible to definitively determine from our experiments what function of UNG is responsible for alkylator sensitivity when UNG is depleted. It is possible that the DNA repair activity of UNG is involved in removing cytotoxic oxidative lesions induced by TMZ and lack of UNG increases the amount of genotoxic lesions yielding increased cell death. Also UNG is essential for class switch recombination, although this function is independent of UNG's DNA glycosylase activity. UNG is not required for stand break formation, but instead appears to be involved in resolution of the breaks to complete class switch recombination (54, 56). Perhaps a mutant UNG can also confer a worse patient survival. Over-expression of wild-type UNG may

not be essential, but instead UNG mutants that contain only glycosylase function, interaction with CENP-A or ability to perform class-switch recombination can also confer worse patient survival. If this is true, it would emphasize one of the pathways UNG is involved for future study to determine what is important in this pathway that correlates with decreased patient survival with UNG over-expression. Re-expression of shRNA resistant separation of function mutants in the T98G/UNG-KD cell line could delineate if UNG's glycosylase activity or N-terminal domain are required for alkylation sensitivity. Determining the function of UNG in alkylation survival would be important to correlate with the decreased survival of brain tumor patients with tumors over-expressing UNG.

We also determined that UNG expression varies in our tumor cell lines compared to normal human astrocytes, with UNG being highly over-expressed in both the T98G and LN428 glioblastoma cell lines. Although the over-expression varied in the two GBM cell lines, there was a 3-5-fold increase in mRNA expression when compared to normal human astrocyte controls. This over-expression was statistically significant with p-values of 3.44972×10^{-7} for the LN428 cells and 2.84847×10^{-6} for the T98G cells. To investigate whether clinical tumors also contain this over-expression, we mined TCGA and REMBRANDT data sets and discovered that UNG is upregulated in all brain tumors types compared to control tissues, with GBM having the highest expression level of all brain tumor types (133). The frequency of two-fold UNG over-expression varied from 14-33% of all GBM tumors in the two data sets. Thus, UNG over-expression occurs naturally in patient brain tumors. Importantly, brain tumor over-expression of UNG has a strong negative correlation with patient survival when compared to brain tumors containing intermediate expression of UNG with a p-value of 1.0177×10^{-4} (133). Why UNG up-regulation is correlated with negative patient survival is not known. UNG may have no direct

role in determining patient survival, but is merely a bystander marker useful for prognosis. Conversley, UNG may play a central role in patient survival and determining the important function of UNG in brain tumor patient survival would be important for brain tumor biomarker development. Currently we assume that the DNA repair activity of UNG likely impacts cytotoxicity from TMZ treatment creating more resistant tumors. However, UNG already has two known functions unrelated to its glycosylase activity and they could also be responsible for decreased patient survival. Sequencing of UNG in the different patient tumors would lend insight into the required functions, with several "separation-of-function" mutants already described. Testing of new mutants discovered in patient tumors for glycosylase activity, class switch recombination and assembly of CENP-A could easily be done in cell culture. If mutants are not identified, we could also determine if UNG glycosylase activity is an important predictor of survival by using the molecular beacon assay. Using the sensitive new molecular beacon design it may be possible to use small tumor samples to determine UNG catalytic rate. This could be compared to normal controls and define the role of UNG's glycosylase activity in patient survival and tumor resistance. It would be interesting to determine if over-expression of UNG contributes to worse patient survival due to increased tumor resistance to alkylation treatment as predicted by our pre-clinical work. However, much more work is required to determine how UNG over-expression impacts patient survival to brain tumor treatment.

Although the discovery of UNG's role in TMZ sensitivity in cell culture and its correlation to patient survival is exciting, other important observations made from the data analysis yield greater potential targets and biomarkers for TMZ response. This stems from the analysis of our human screen results to similar alkylator sensitivity screens done in *S. cerevisiae* and *E. coli* (77, 78). Many of the same biological processes significantly enriched in the human

screen were enriched in both screens from yeast and bacteria. The conservation of these processes through evolution suggests these are the essential processes required for alkylation survival in bacteria, yeast and humans with non-essential or secondary processes lost during evolution. There are a diverse collection of biological processes, ranging from DNA repair to modulation of transcription, translation, protein modifications and fatty acid biosynthesis that appear to be essential for survival after alkylation exposure. The many processes necessary for cellular survival after alkylation exposure likely stems from the alkylator's ability to modify all macromolecules in a cell, thereby requiring repair of not only DNA, but also of RNA, proteins and fatty acids. Contained in this diverse set of processes are the conserved processes in alkylator survival that could be potentially targeted anywhere in the pathway. It is important to determine if inhibition of these processes by ablating other important proteins within the pathway also increase sensitivity to TMZ. If more "druggable" genes in the conserved pathways are also essential for TMZ survival, this would yield a plethora of protein targets with potential tumor specific effects and generate a cadre of possible biomarkders to predict TMZ response and patient side-effect toxicity. With personalized medicine being increasingly developed, determining the possible biomarkers to predict tumor response to therapy would be crucial. We could potentially pre-determine who would derive the most benefit from treatment, diminish the costs of misguided treatments and reduce patient side effects in unresponsive patients.

One of the enriched biological processes was protein modification processes that were validated using both the human and yeast system. The knockdown of the human protein UBE3B significantly sensitized human cells to TMZ by approximately 45% in both short and long-term assays. Notably, the concentration of TMZ used in the long-term assay was only 50 μM, well within the clinically achievable dose of TMZ. This was unexpected because the mechanisms for

short and long term survival to DNA alkylators are thought to be related, yet distinct. Our lab has shown that short-term survival to alkylating agents is dependent on both BER status and NAD^+ levels in the cell (67, 70). Long-term survival to alkylators is usually dictated by resistance to the O^6-MeG lesion since it is the most cytotoxic DNA lesion created by alkylators. However, over-expression of the MGMT protein in the T98G cells likely removes many of the O^6-MeG lesions. There are a myriad of methods for UBE3B to play a role to increase cell survival in response to TMZ, with several possibilities illustrated below. The first is simplest, for UBE3B may promote MGMT expression or activity and lack of UBE3B decreases MGMT activity. MGMT protein levels could be tested by immunoblot and MGMT activity can be measured using a fluorescence-based assay (143). This mechanism is not likely because there is an increase in cell death in the short-term assay as well, which as previously described does not measure O^6-MeG toxicity. If this does occur, a separate function of UBE3B would have to promote cell survival in the short-term assay. Second, UBE3B could modulate the MMR response and signaling to apoptosis for cell death mediated by O^6-MeG lesions. This could be determined by investigating UBE3B knockdown cells for MMR activity by measuring microsatellite instability. This is also not likely because there is an increase in cell death in the short-term when UBE3B expression is inhibited. If this does occur, a separate function of UBE3B would have to promote cell survival in the short-term assay. Third, UBE3B could play a role in repair of BER substrates by modulating protein activity. There is precedence for BER protein activity to be regulated by post-translational modifications such as phosphorylation, acetylation, methylation, ubiquitylation and SUMOylation (26, 144). Many other DNA repair proteins are regulated by ubiquitylation. Measuring BER activity in the wild-type and UBE3B-KD cell lines could test this hypothesis. BER activity can be measured using the molecular

beacon assay or traditional BER assays using ^{32}P end labeling and denaturing gels (145, 146). A fourth possibility is that UBE3B is regulating protein quality and turnover which impacts cell survival. Fifth, UBE3B is involved in regulating DNA damage-mediated cell death. This appears to be more likely since there is the epistatic relationship between UBE3B and UNG, another DNA repair protein, suggesting they are part of the same alkylation survival pathway.

Investigating the epistatic relationship between UNG and UBE3B in DNA damage mediated cell death pathways to uncover potential drug targets would be an interesting next step. This relationship would be particularly interesting to study not only because UBE3B's functions are essentially unknown, but also because it would identify other proteins in the survival pathway that may be potentially targeted to increase TMZ efficacy. As mentioned previously, the known role of UBE3B in cellular processes is minimal and simple experiments such as tagging UBE3B with GFP to visualize sub-cellular localization is beneficial. Using a fluorescently tagged UBE3B protein, it is possible to determine whether UBE3B changes locations, forms foci or associates with chromatin after TMZ treatment and would add to our basic understanding of the protein. This would also improve our ability to predict potential interaction points between UBE3B and UNG in their survival pathway. For example, if UBE3B is only located in the cytoplasm and never traffics to the nucleus, it likely is not involved in direct DNA repair like other ubquitin ligases. Instead it would more likely be involved in cell death signaling. UBE3B likely has an indirect interaction with UNG that mediates their epistatic relationship and may regulate DNA repair activity in cells. The BER activity can be measured by the molecular beacon assay or traditional repair assay methods. Tagging of UBE3B for immunoprecipitation to discover possible substrates is also possible. Another method for discovering UBE3B targets is transfecting tagged ubiquitin into wild-type and UBE3B

knockdown cells and immunoprecipitating the tagged ubiquitin. Separating the tagged ubiquitin on a 2D gel to determine what proteins are ubiquitylated in the wild-type cells compared to the knockdown cell line. We are also interested in determining if UBE3B knockdown sensitizes cells to other alkylating agents, such as cyclophosphamide, chlorabucil and dacarbazine, and other types of DNA damage induced by other chemotherapeutics, such as carboplatin, doxorubicin and etoposide. This would help determine if UBE3B is involved in DNA repair and signaling of other genotoxic agents or sensitizes to TMZ in an alkylator specific manner.

Critical to any TMZ adjuvant treatment would be introducing cancer selectivity to TMZ toxicity. There are several methods available to confer tumor-selective killing. One method is the creation of a compound that can enter all cells, but selectively kill tumor cells. An example of a drug that employs this mechanism could be found in many of the targeted therapies such as the PARP inhibitor ABT-888 (Veliparib) or the Bcr-Abl inhibitor imatinib (Gleevec). A second method would be to have a uniformly toxic compound, but selectively target the compound to cancer cells to minimize normal tissue exposure. An example of this method is administration of radioactive iodine-131 for the treatment of thyroid tumors or the DNA alkylator streptozotocin for the treatment of metastatic pancreatic beta cell cancer. The third method to selectively kill tumor cells would be local delivery of a toxic compound to tumor tissue, thereby limiting contact with normal tissue. This is an extension of the second method to limit normal tissue exposure to a toxic agent. An example of this method is the placement of Gliadel wafers in the tumor bed of GBM patients or tumor focused tumor radation therapy instead of whole body irradation.

TMZ is a DNA alkylator that indiscriminately modifies macromolecules in all cells and is not easily adapted to the second and third methods stated above. Although it may be possible to target TMZ specifically to cancers, as seen with the DNA alkylator streptozotocin, any

modifications to TMZ may disrupt the spontaneous breakdown required for conversion to the active methyldiazonium ion. Local delivery of TMZ via wafer therapy is also possible, but GBM tumor resections are not always feasible, limiting frequency of wafer treatment. Also, invasion of the tumor cells outside the tumor bed may not be effectively treated by the TMZ released via wafers. The most promising methodology to improve TMZ selectivity is to target the resistance mechanisms that differ in cancer compared to normal tissue. For example, PARP inhibitors selectively target cells and tumors defective in homologous recombination (83, 84). Although the PARP inhibitor mechanism of action in homologous recombination defective cells is controversial, the key idea is a specific mechanism of DNA repair is compromised in cancer cells and modulating parallel DNA repair pathways selectively induces cancer cell death (84, 147).

A deep understanding of all the required biological processes that modulate TMZ sensitivity will enable tumor specific targeting if sensitivity pathways or associated pathways are defective in cancers. Discovering the overlapping TMZ sensitivity pathways in normal tissue and cancer will facilitate tumor selective target acquisition. The large scale cancer sequencing and expression databases are beginning to uncover and quantify the genes and pathways disrupted in specific cancers. The large datasets will generate targets for cancer therapy because of dysregulation of specific genes and pathways. Equally important will be mapping chemotherapy sensitivity pathways to the datasets to discern if specific cancers or subsets of cancers are defective in repair pathways making them susceptible to current therapy or alternative treatment options. For example, our lab has shown that although DNA alkylator treatment is not selective, by concurrently targeting the NAD^+ biosynthesis pathway it may be possible to selectively induce tumor cytotoxicity from alkylation damage due to prevalent defects

in tumor NAD^+ biosynthesis (70). Likewise, it may be possible to inhibit TMZ sensitivity pathways and increase efficacy in a subset of glioma cases. The UNG is up-regulated two-fold in approximately one-third of all GBM patients in the REMBRANDT database and UNG overexpression correlates with worse patient survival suggesting that UNG may be contributing to poor response (133). Inhibiting UNG activity may increase TMZ efficacy in all tissue. However, the overexpression of UNG in the cancer tissue may make it more susceptible to chemical inhibition and yield a larger increase in TMZ sensitivity in tumors compared to normal tissue. Chemical inhibition of UNG may increase TMZ's specificity and efficacy. More studies are necessary to characterize UNG's role TMZ sensitivity and patient survival in GBM and glioma to validate UNG as an adjuvant drug target.

In summary, these studies suggest that alkylation resistance mechanisms are evolutionarily conserved. The collection of conserved biological processes in *E. coli*, *S. cerevisiae* and humans composes an *Alkylation Functionome* that includes many novel proteins not previously thought to impact TMZ or MMS resistance. We can now begin to appreciate the multiple processes that are required for cellular survival after alkylation damage and form an unbiased approach to discovering targets for adjuvant chemotherapy. This is an important next step, so as to determine if genes that were not tested in the screen, but share conserved biological pathways, are also possible targets to enhance TMZ response. Because the analysis enriched for the most essential pathways and genes, there may be functional overlap of the conserved genes and pathways. The highly conserved nature of these biological processes should generate significant interest into studying potential gene and pathway interactions. Investigating the relationships between these genes and biological processes is important in determining pathway redundancy. Importantly, some of the targets studied have already been shown to be

significantly up-regulated in cancers and correlate with worse patient survival. If targeted by adjuvant chemotherapies, this would potentially create tumor specific sensitivity, increase the therapeutic window and reduce side effects. However, due to the plethora of resistance mechanisms it is possible that depletion or inhibition of one gene in a conserved biological process cannot overcome the drug resistance due to compensation by another gene or biological process. By investigating multiple pathways and processes we can determine important interactions that promote tumor survival to be targeted for improved chemotherapy response.

8.2 MOLECULAR BEACON ASSAY EXPERIMENTS

The molecular beacon assay is a very versatile assay enabling sensitive and precise real-time measurements of DNA repair activity in nuclear extracts. One of the interests in the lab has been to move the assay into live cells using a modified beacon design (**Table 6**), which incorporates a Cy5 fluorophore in the loop of the hairpin. This secondary fluorophore does not overlap with the absorbance/emission spectra of the 6-FAM dye or the Dabcyl quencher. The Cy5 would be used as a normalization fluorophore. The modified molecular beacon would be transfected into cells, similar to siRNA due to the similar structure and size, and the appearance of 6-FAM fluorescence would be monitored as molecular beacon repair rates and would be normalized to the Cy5 fluorophore. The extra Cy5 not only eliminates any temperature adjustments to melt the hairpin after DNA repair readings, but also allows us to normalize for transfection efficiency in every cell measured. We could use a confocal microscope to monitor both dyes simultaneously. Furthermore, the development of our Microwell Comet Assay

dovetails nicely with measuring the beacons in real-time cells (88, 148). The Microwell Comet Assay enables separation of each cell into single wells in defined locations on a 96-well plate. The gel works by having small wells, roughly 20 microns, which can only fit one cell in each well. The well locations are preprogrammed into the computer that uses the individual cell location to quickly scan the plate for all cells to increase throughput of individual cell images without having to manually scan for isolated cells for measurements. This would allow a large sample size of individual cells to be scanned for Cy5 dye, which would correspond to beacon uptake, and followed for 6-FAM fluorescence, corresponding with DNA repair activity. Furthermore, the ability to scan individual cells quickly and monitor their location and fluorescence would yield an interesting ability to correlate real-time DNA repair rates with genomic DNA damage. For example, we could use the Microwell Comet Assay to separate our cells and treat them with molecular beacons to determine DNA repair rates of the individual cells. Since these cells are now already in the wells we could then carryout a comet assay on the same wells. With this in place we could monitor an individual cell's ability to repair DNA and correlate that to incorporated genomic DNA damage. This could be used with different DNA repair deficient knockdown cell lines, such as the isogenic cell lines being developed in the Sobol Lab, or known repair deficient cancer cell lines. This would enable a direct correlation of how DNA repair rates in cells affect genomic lesions, because all the experiments are individually normalized and controlled. This could be further used to discover if repair rates of alkylation damage correlates with genomic DNA damage. For example, the Microwell Comet Assay and molecular beacons can be used to determine DNA repair rates of the individual cells and record their exact location. Cells are then washed to remove the molecular beacons before treating the cells with a DNA alkylator such as MNNG. The cell's genomic DNA damage

measured via a Comet Assay would be measured. Using the level of DNA repair measured by the molecular beacons and genomic DNA damage measured by comet assay, we could determine if robust DNA repair was inverselty correlated with levels of genomic DNA damage measured in a time dependent manner. An alternative would be to plate cells in the Microwell Comet Assay, measure molecular beacon repair rates, treat with MNNG and monitor apoptosis or other forms of cell death. By measuring DNA repair rates and cell death the correlation of increased DNA repair promotes or prevents cell survival after alkylation or any type of DNA damage exposure in different cell types would be elucidated. This would determine if over-expression or knockdown of DNA repair proteins actually impact the rate of DNA repair, cell survival signaling or other methods of DNA damage tolerance that do not actually repair the DNA to the original sequence. Our lab has already optimized MNNG cell treatment before measuring genomic DNA damage using the Microwell Comet Assay. Although adding the molecular beacon step would increase the difficulty of the experiment, another washing procedure is possible before determining DNA damge via the Microwell Comet Assay. By combining several novel methods in live cells we could learn mechanistic detail into not only rates of DNA repairs, but also have an easy method to determine what genes actually impact the rate of repair and correlate these to survival.

The role of functional biomarkers would be another important avenue to investigate. Due to the predominance of genotoxic chemotherapeutic drugs and cancers containing inherent defects in DNA repair, the development of functional DNA repair assays that can be used to monitor cancer biomarkers to predict outcome of treatment and predisposition to cancer would be important clinical tools. The molecular beacon assay enables sensitive and precise measurements of DNA glycosylase activity. Our lab has shown that MPG activity impacts

cancer cell line sensitivity to TMZ (148). MPG expression and sensitivity to DNA alkylators has been a controversial topic. Several papers show that in different biological systems, MPG expression can either increase or decrease sensitivity to alkylating agents (101, 105, 149). However the expression of MPG alone cannot accurately predict survival to DNA alkylators, because MPG is involved in a pathway which uses over 20 different proteins (26). DNA polymerase β (Polβ) also plays a critical role in alkylating agent toxicity and is likely a confounding factor in some studies investigating MPG expression as many groups do not interrogate Polβ expression or activity. Polβ expression and more importantly activity can compensate for high expression of MPG and the balance of BER needs to be considered when studying alkylation damage. The need for functional DNA repair assays to ameliorate this confusion is critical.

MPG clinical tumor cell expression has been reported to correlate with worse survival in clinical cases (149). However, they used an antibody that does not recognize their protein of interest as their immunoblot recognized a 24 kDa protein, when MPG is actually a 34 kDa protein. To clarify the outcomes, a functional assay would enable groups to accurately determine if MPG and Polβ tumor activity dictate alkylation sensitivity more accurately than expression. Towards these ends our lab is developing molecular beacons to accurately measure MPG activity in tumor lysates or primary tumor cells. However, we have also been developing a similar molecular beacon approach to quantify Polβ polymerase activity in cell lysates and subsequently in live cells. This approach would take advantage of pyrene deoxyriboside's ability to be quenched by thymine and uracil DNA bases (150). A single-stranded DNA that forms a hairpin structure, containing internal pyrene deoxyriboside fluorophore adjacent to a THF site would enable efficient detection of fluorescent signal (150). However, when incubated with nuclear

lysates, the active proteins would recognize and repair the abasic site leaving an undamaged DNA strand. The newly incorporated DNA base would efficiently quench the fluorescence and loss of signal would correspond with repaired DNA. The oligonucleotide will be larger than that already used for the molecular beacons to ensure the oligo that is cut by APE1 does not dissociate quickly from the other DNA strand. Based on previous BER assay work using ^{32}P end-labeling, a stem of 30 bases should be sufficient to maintain the stem-loop structure and allow BER proteins to correctly process the THF and subsequent DNA intermediates (146). This would enable us to accurately measure Polβ polymerase activity in real-time and not only enable us to screen for inhibitors of Polβ, which is a wide interest in the field (151), but to also determine if the relationship between MPG and Polβ activity can accurately predict sensitivity to alkylating agents in clinical tumor samples. The production of a functional biomarker assay would be a boon to the field as the current best assays only measure expression or promoter methylation to predict sensitivity to TMZ. Polβ is known to have many mutations in cancer, but the functionality of many of these mutations is unknown (151). Also the expression levels may not accurately predict protein activity, which would undermine correlative studies. Thus, development of both DNA glycosylase and Polβ functional assays in clinical tumor samples is an important step to learn how to most accurately predict sensitivity to TMZ and to enable us to accurately measure the pathways to manipulate the proteins via chemical inhibitors for increased clinical benefit.

APPENDIX A. OPTIMIZATION OF MOLECULAR BEACON AND CYQUANT ASSAYS

A.1 MOLECULAR BEACON DEVELOPMENT AND OPTIMIZATION

In the course of the project we needed to determine if DNA repair rates correlate with response to alkylating agents. Therefore, we developed and optimized a DNA repair assay that enables us to measure real-time repair rates of nuclear cell extracts with enough accuracy to determine enzymatic kinetic parameters and to validate or screen inhibitors. This highly sensitive and quantitative assay is also applicable to the analysis of purified proteins, enabling studies on substrate specificity by altering the molecular beacons accordingly. Finally, the assay is applicable to analysis of tumor and tissue lysates, providing an opportunity to measure functional DNA repair endpoints as biomarkers of response or therapeutic efficacy, as we have suggested for evaluating tumors for PARP1 inhibitor potentiation of the alkylating agent temozolomide (148).

We first designed the molecular beacons as described above (**Page 25, Figure 1**) to determine rates of the DNA glycosylase MPG, because we had previously characterized repair rates of MPG in a glioblastoma cell line (LN428) with low levels of endogenous MPG and in the LN428 cell line over-expressing MPG (LN428/MPG) using the standard ^{32}P-endlabeling methodology previously (152). Using these cell lines we could prepare nuclear extracts from

both cell lines as described above (**Page 25**) and determine DNA repair rates of a control beacon (**FD-Con, Table 5**) and the MPG specific DNA lesion ethenoadenine (**FD-MPG1, Table 5**) (15, 148). The same BER-beacon structure with a normal adenine was used as the control substrate. Following removal of 1A by MPG and subsequent DNA strand excision by APE1 5′ to the AP site, the fluorophore 6-FAM is separated from the quencher (Dabcyl) and the increase in fluorescence signal (517 nm) is proportional to the level of MPG activity (**Figure 1**). The LN428 lysate incubated with the control beacon (**Figure 20D**, red filled squares) had a minimal (if any) increase in fluorescence, indicating the control beacon is largely intact. The LN428 lysate has little or no endogenous MPG activity, since when incubated with the beacon containing the MPG-specific substrate 1A, there was no observable change in fluorescence (**Figure 20D**, green open squares). The LN428/MPG lysate also did not have a negligible increase in fluorescence when incubated with the control beacon (**Figure 20D**, orange open circles), indicating that MPG over-expression does not increase cleavage of normal DNA. However, the LN428/MPG lysate exhibited robust MPG activity visible with a large increase in fluorescence when incubated with the molecular beacon containing the MPG substrate 1A (**Figure 20D**, filled blue circles). This corresponded to an overall 7.9-fold increase in MPG activity (measured at 60 min), as compared with the LN428 cells and an estimated rate of repair (based on the slope of the curve from 15 to 30 min) of 107.00 AU/min, whereas the background rate of repair in the LN428 cell lysate was similar to the background signal using the control beacon (14.64 AU/ min). This demonstrates that the LN428/MPG cell line has increased functional MPG and does not recognize normal DNA as a substrate. These data are in-line with our earlier report showing that over-expression of MPG results in an increase in DNA glycosylase activity (148).

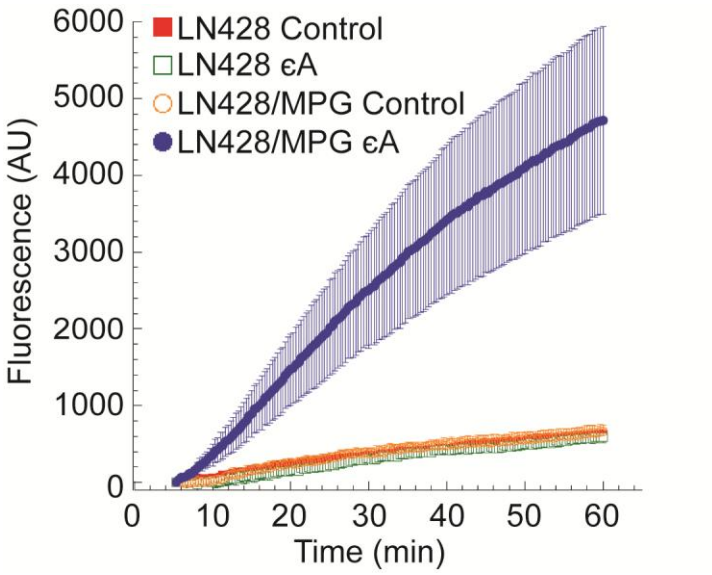


Figure 20. Increased DNA glycosylase activity in MPG over-expressing LN428/MPG cells as determined by the DNA Glycosylase Molecular Beacon Activity Assay.
DNA glycosylase activity specific for removal of the MPG substrate 1A was measured in nuclear lysates from the control cell line (LN428) and the MPG over-expression cell line (LN428/MPG), as described in the Materials and Methods section. Each lysate was analyzed using either the control-beacon or the 1A-beacon: LN428 lysates (control-beacon, red filled squares; 1A-beacon, green open squares) and LN428/MPG lysates (control-beacon, orange open circles; 1A-beacon, filled blue circles). Results are reported as the mean fluorescence response unit of 3 independent experiments + S.E.M.

With this method in hand, we wanted to determine if DNA glycosylase repair rates correlated with cellular survival, if different lesions for the same glycosylase yields similar results and if knockdown of essential BER proteins, such as XRCC1, affected the DNA repair rate. Therefore, we tested four different glioblastoma cell lines that had endogenous levels of protein (LN428), over-expressed MPG (LN428/MPG), had reduced XRCC1 via shRNA knockdown (LN428/XRCC1-KD), and both over-expressed MPG and had reduced XRCC1 (LN428/MPG/XRCC1-KD), for sensitivity to the alkylator MMS and real-time repair rates of two DNA lesions (εA and Hx), which are substrates of MPG (15, 88). Furthermore, εA inhibits DNA replication and is thus a toxic base lesion (153, 154). We first determined that the MPG over-expression cell line did have elevated levels of MPG as the LN428 and LN428/XRCC1-KD cells had levels of MPG lower than our ability to detect (**Figure 21A**). To compare the

expression levels of XRCC1 and MPG among the four cell lines, the immunoblot was scanned and quantified using NIH Image J and the associated analysis software package, normalizing the expression across the four cell lines to the MPG cells and to the expression of PCNA within each sample. As indicated in the plot of relative protein expression (**Figure 21B**), the MPG expression levels are very similar when comparing the LN428/MPG and LN428/MPG/XRCC1-KD cells. Similarly, the expression of XRCC1 is similar when comparing the LN428 and LN428/MPG cells. Importantly, no expression of XRCC1 was detected in the LN428/XRCC1-KD and LN428/MPG/XRCC1-KD cells.

We assessed the ability of cell extracts from LN428WT, XRCC1-KD, MPG and MPG/XRCC1-KD cells to excise εA, Hx or tetrahydrofuran (THF), an abasic site analog (155). We observed minimal fluorescence when extracts were incubated with a control molecular beacon that contained no lesion (**Figure 21C**). Since the signal (fluorescence) is generated following base lesion removal and abasic site hydrolysis by APE1, we also tested each lysate using a THF beacon to ensure that all four lysates could similarly cleave an abasic site. As expected, a similar increase in fluorescence was observed by all four of the cell extracts when evaluated for APE1 activity using the THF beacon (**Figure 21D**). It should be noted that we did not observe any difference in activity between the XRCC1 expressing cells (LN428WT and MPG) and the XRCC1-KD cells (XRCC1-KD and MPG/XRCC1-KD); indicating that cleavage of the AP site by APE1 is not influenced by XRCC1. An earlier report using purified enzymes suggested that XRCC1 stimulates APE1-mediated 3'–5' exonuclease activity but it is not likely that this activity of APE1 is measured in this assay (156). However, our data are in contrast to an earlier report suggesting that XRCC1 and APE1 interact and that XRCC1 stimulates APE1

activity (157). Clearly, such an XRCC1-dependent stimulation of APE1 activity is not observed here, at least for the hydrolysis of the THF abasic site analog.

Conversely, a molecular beacon containing Hx resulted in a time-dependent increase in fluorescence when incubated with extracts from LN428WT, XRCC1-KD, MPG and MPG/XRCC1-KD cells, which is indicative of base release by the MPG glycosylase and cleavage of the resulting AP site (**Figure 21E**). As expected, cells over-expressing MPG (MPG) yielded the highest increase in fluorescence, which is consistent with these cells being the most efficient at excising the Hx lesion. However, when XRCC1 expression is knocked down in the MPG cells (MPG/XRCC1-KD), there is a decrease in the observed fluorescence (**Figure 21E**). The observed 25% reduction in glycosylase activity in the XRCC1-deficient cells (when measured at 40 min) indicates that XRCC1 facilitates MPG excision of Hx, which is consistent with a previous report using purified recombinant proteins [39]. In contrast to the cells over-expressing MPG, the molecular beacon assay yields very little signal for the LN428WT or the XRCC1-KD cell extracts. The nominal increase in fluorescence is probably due to the exceptionally low amounts of MPG in the LN428WT parental cell line. Indeed, the levels of MPG in the LN428WT parental cell line and the XRCC1-KD cell line are so low as to preclude noticeable detection by western blot (**Figure 21A**). Thus, additional expression of MPG leads to a significant increase in the excision of εA and Hx.

To learn if these observations are can be generalized to other MPG substrates we queried whether εA excision by MPG could be facilitated by XRCC1. We incubated cell extracts from the glioblastoma cells with a molecular beacon containing εA. Similar to the Hx molecular beacon, a time-dependent increase in fluorescence was observed following incubation with lysates from LN428WT, XRCC1-KD, MPG and MPG/XRCC1-KD cells. Of the cell lines

examined, the MPG cells yielded the greatest increase in fluorescence and the MPG/XRCC1-KD cells yielded a significant (45% at 40 min) reduction in the kinetics of BER initiation (**Figure 21F**). The difference in the observed fluorescence between the MPG cells and the MPG/XRCC1-KD cells indicates that XRCC1 does indeed facilitate MPG excision of εA. Collectively, these data indicate that XRCC1 is capable of stimulating the MPG-mediated excision of two inflammation-associated base lesions, Hx and εA, *in vitro*.

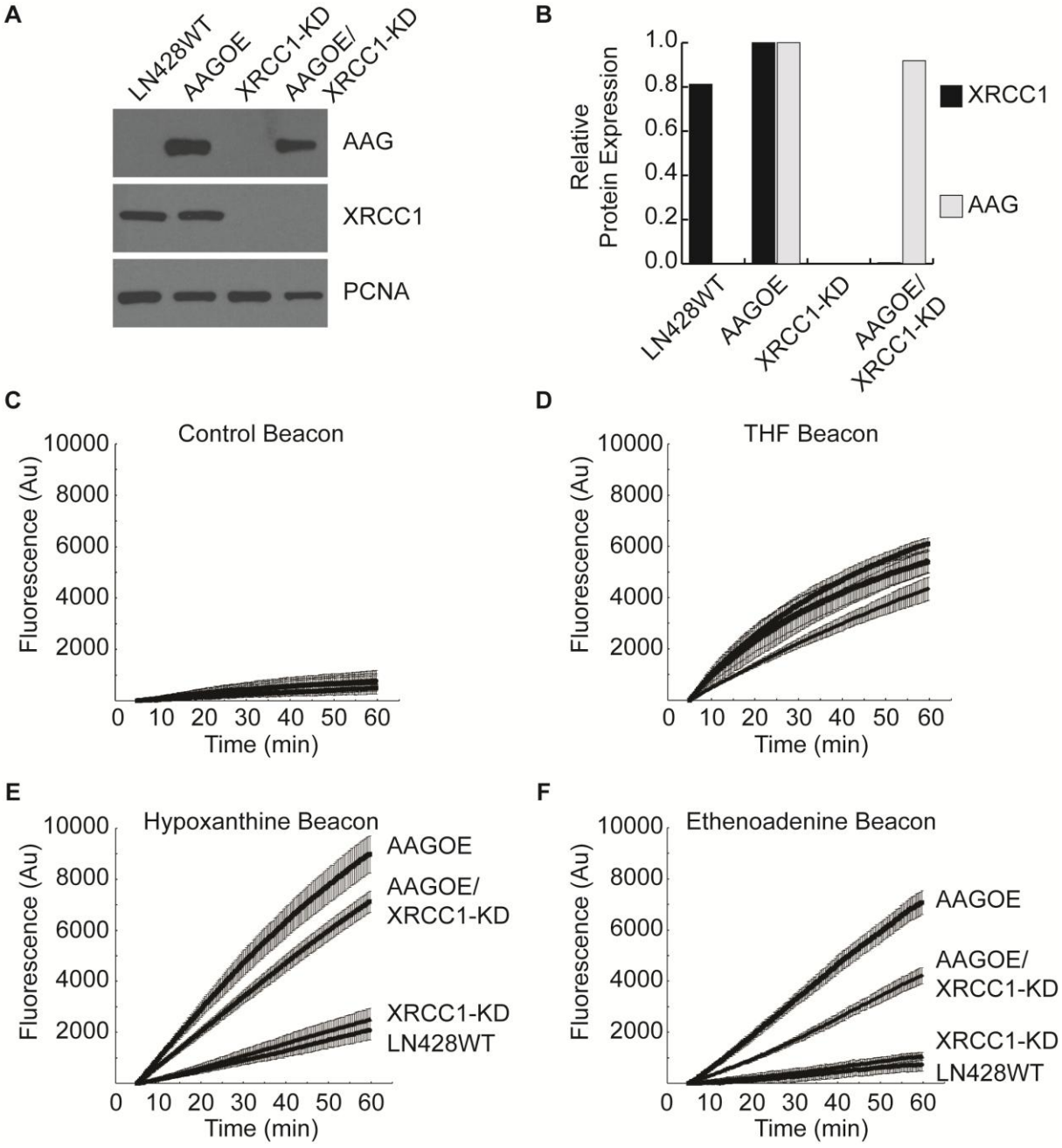


Figure 21. XRCC1 facilitates MPG-mediated NO•-induced base lesion removal.

(A) Immunoblot showing expression of MPG and XRCC1 protein in wild type (LN428), MPG over-expressing cells (MPG), XRCC1 knockdown cells (XRCC1-KD), and MPG over-expressing cells with XRCC1 knocked down (MPG/XRCC1-KD). **(B)** Bar graph showing the relative protein expression of MPG and XRCC1 in LN428, MPG, XRCC1-KD and MPG/XRCC1-KD cells. The immunoblot in (A) was scanned and quantified using NIH Image J and the associated analysis software package, normalizing the expression across the four cell lines to the MPG cells and to the expression of PCNA within each sample. **(C–F**) Fluorescent molecular beacon assay indicating XRCC1 facilitated excision of NO•-induced base lesions by MPG from DNA double-stranded oligonucleotides containing **(C)** no lesion, **(D)** an abasic analog THF, **(E)** hypoxanthine (Hx) or **(F)** 1,N6-ethenoadenine (εA).

A.1.1 Molecular beacon DNA repair rates correlates with genomic DNA damage and survival to alkylating agents

We wanted to confirm that XRCC1-KD not only affected DNA repair rates, but also impacted cell survival to alkylators. In collaboration with Dr. Bevin Engleward, the genomic DNA damage of the LN428, LN428/MPG, LN428/XRCC1-KD and LN428/MPG/XRCC1-KD cells was determined using an alkaline buffer comet assay to measure not only DNA strand breaks, but also abasic sites, a BER intermediate. They determined that the LN428/MPG/XRCC1-KD cells contained significantly more genomic DNA damage 4 hours after treatment with MMS when compared to controls. This data fit with the decreased DNA repair rates seen in the molecular beacon assay and we sought to determine if survival to alkylators correlates with decreased DNA repair rates and increased genomic DNA damage. We then treated the four different cell lines with vehicle or various doses of MMS for four hours, the same time frame that contained increased genomic damage in the comet assay, and measured survival via an MTS assay (**Figure 22**). We found that as previously reported the LN428 cells were not sensitive; while LN428/MPG cells are sensitive to alkylating agents (67, 70, 148). However, the LN428 XRCC1-KD cell line was not sensitive, likely due to the undetectable level of MPG to initiate BER. The most sensitive cell line was LN428/MPG/XRCC1-KD cell line, which was significantly more sensitive to MMS than the all three other cell lines, suggesting that once repair is initiated XRCC1 performs a critical survival function, likely related to the defect in DNA repair observed using the molecular beacon and comet assays. The strong correlation between DNA repair rates, genomic DNA damage and cellular survival after alkylation treatment encouraged us to develop better methods to measure DNA repair rate, with the goal to measure DNA repair *in vivo*.

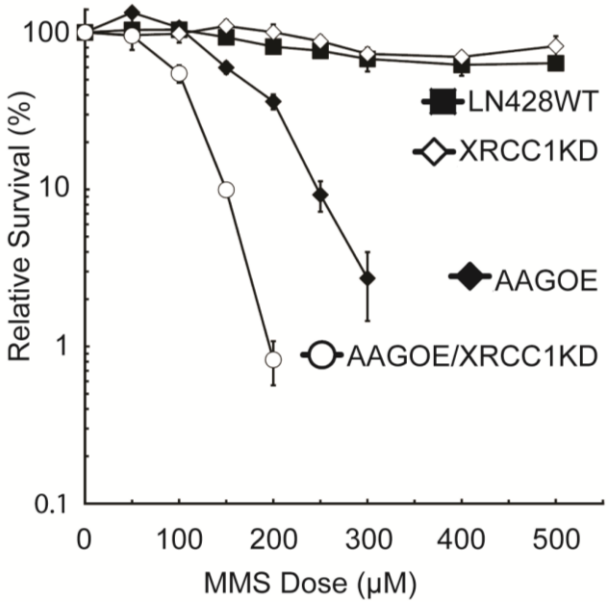


Figure 22. Role of XRCC1 and MPG in human glioblastoma cells following exposure to MMS.
MMS induced cytotoxicity in LN428, LN428/MPG, LN428/XRCC1-KD and LN428/MPG/XRCC1-KD cells. After treatment (4 h MMS followed by 48 h in normal media), viable cells were determined using a modified MTT assay. Plots show the % viable cells as compared to untreated (control) cells. Means are calculated from six values in each experiment. Results indicate the mean ± S.E. of three independent experiments.

A.1.2 New molecular beacon design improves signal detection

Upon further research into molecular beacon design, we discovered that 6-FAM fluorescence is partially quenched when in close proximity to a guanine base. Our original design created a beacon with guanine adjacent to the 6-FAM fluorophore (**Table 5**). We designed new molecular beacons by replacing the guanine abutting the 6-FAM with a cytosine to take advantage of the potential increase in fluorescence (**Table 7**). The new THF (FD-THF2) beacon was tested in parallel with our old beacon designs (FD-Con and FD-THF) using nuclear lysates from the LN428 and LN428/MPG cell lines. We expected identical results from the two different cell lines, as seen in **Figure 21,** due to the same levels of APE1 in each lysate. However, we did expect a modest increase in fluorescence with the new beacon design when

compared to the old design. To our surprise, the new beacon design yielded over a 300% increase in signal when compared to the old design (**Figure 23**). This new design was incorporated into all our new molecular beacons and the increased signal potentially enables live cell imaging of the beacons to determine *in vivo* DNA repair rates. Further, we determined that pre-chilling all the solutions and mixing the reactions on ice drastically reduced the initial variability of the measurements. This reduces the amount of beacon that is cleaved while mixing the plate and before beginning measurement readings.

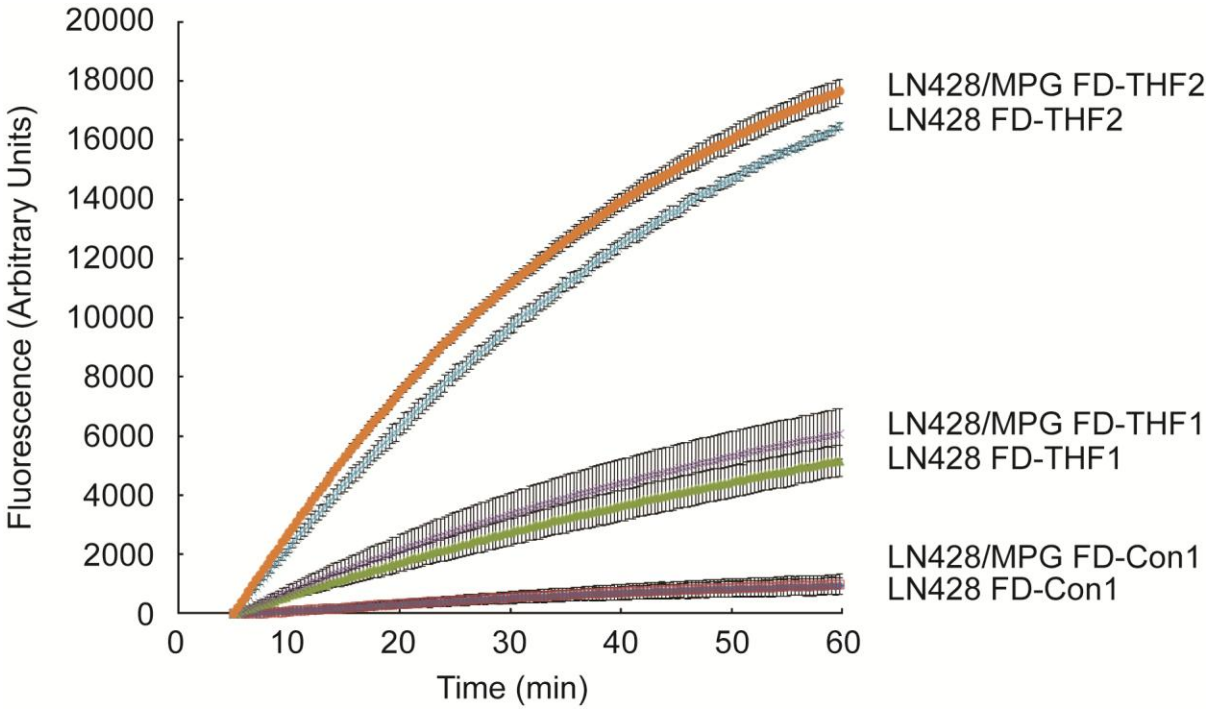


Figure 23. The nucleotide adjacent to 6-FAM fluorophore impacts fluorescence signal.
The substitution of a cytosine for the guanine adjoining the 6-FAM fluorophore yields a 300% increase in maximum signal, enabling more precise measurements of DNA repair rates. The large increase in signal can be seen in both the LN428 and LN428/MPG cell lines when comparing the FD-THF1 to the FD-THF2 beacon. As expected, using different cell lysates did not impact removal of the THF lesion, since both lysates have equal amounts of APE1. Results are the average of two independent experiments with error bars representing ± SEM.

A.1.3 Normalization analysis of molecular beacons

Although the background levels of fluorescence were low and the results were very consistent, we noticed there was high variability in absolute fluorescence when using different qRT-PCR machines. The relationships to the different beacons were the same, but plotting the absolute fluorescence was not possible if using different machines. To enable comparisons across machines, we developed a different analysis method to normalize for well-to-well and machine differences. This was done by adding several steps at the end of the beacon protocol to collect absolute fluorescence data that could be used for normalization.

We determined DNA repair rates by measuring fluorescence every 20 seconds for 60, 90 or 120 min. The beacons were then incubated sequentially at 60°C, 65°C, 70°, 75°C, 80°C for 5 min at each temperature. The average fluorescence value at each temperature was determined. The maximum average fluorescence value in each well was used to normalize the beacon fluorescence in each corresponding well to account for pipetting error, well to well variability of measurements in the machine and machine to machine variability in fluorescence measurements (**Figure 24**).

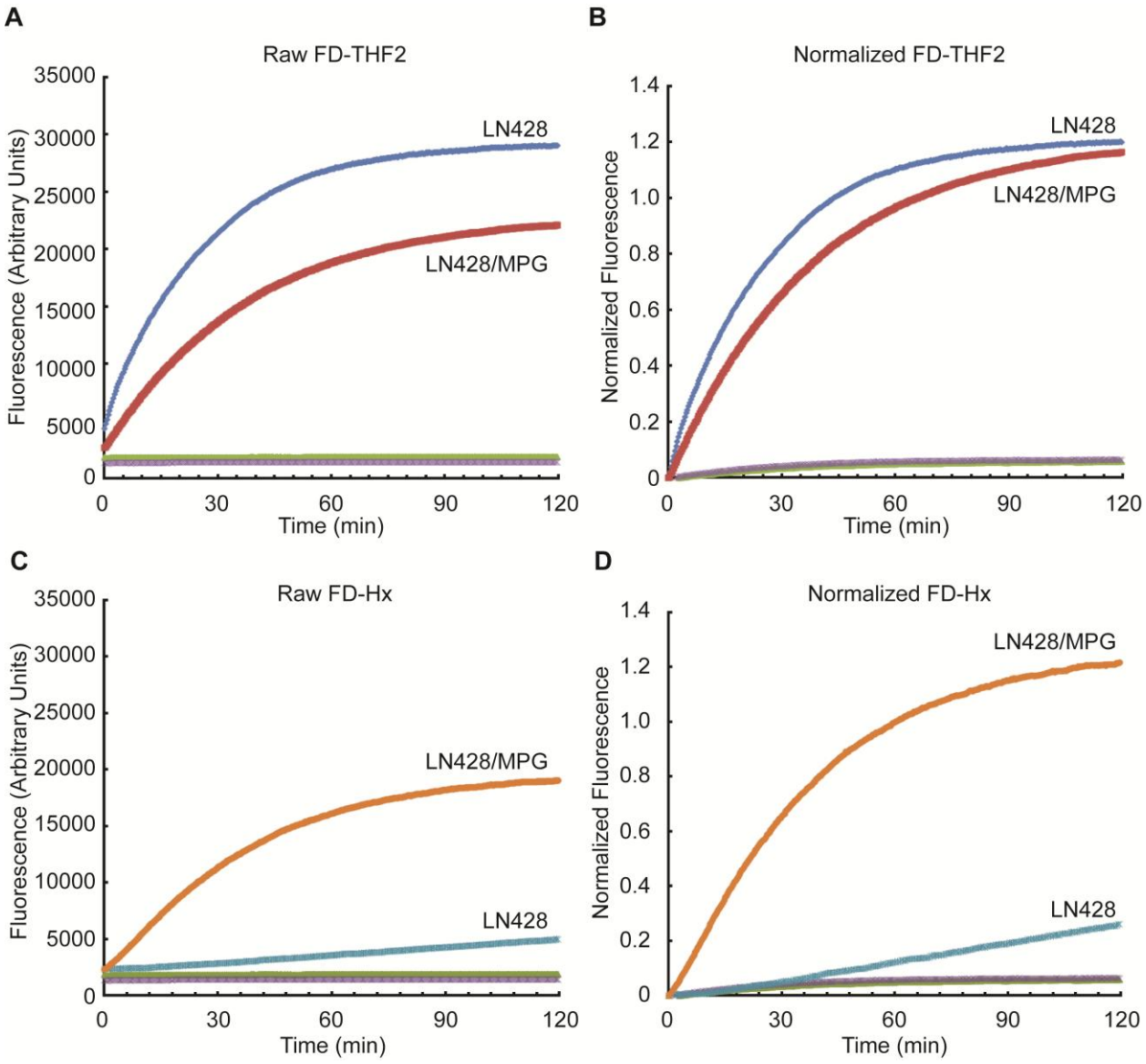


Figure 24. Raw and normalized data for LN428 and LN428/MPG cells extracts.

AP endonuclease activity specific for hydrolysis of the abasic site analog tetrahydrofuran (THF) was measured in nuclear lysates from the control cell line (LN428) and the MPG over-expression cell line (LN428/MPG). Each lysate was analyzed using either the BER Molecular Beacon Control (LN428, green; LN428/MPG, purple) or the BER THF Molecular Beacon (LN428, blue; LN428/MPG, red). Results are reported as **(A)** the mean fluorescence response units and **(B)** the same data normalized as described above (**Page 33**) (THF = tetrahydrofuran). DNA glycosylase activity specific for removal of the MPG substrate hypoxanthine (Hx) was measured in nuclear lysates from the control cell line (LN428) and the MPG over-expression cell line (LN428/MPG). Each lysate was analyzed using either the BER Molecular Beacon Control (LN428, green; LN428/MPG, purple) or the BER Hx Molecular Beacon (LN428, blue; LN428/MPG, red). Results are reported as **(C)** the mean fluorescence response units and **(D)** the same data normalized as described above (**Page 33**) (Hx = hypoxanthine).

A.2 CYQUANT ASSAY OPTIMIZATION AND DEVELOPMENT

The CyQuant assay yields similar results to that of the clonogenic survival assay. The clonogenic survival assay is the standard for cytotoxicity studies, but is dependent on the cells understudy to form good colonies that can be counted. Because our cell lines studies do not for colonies, we needed to develop the CyQuant assay to generate long-term survival results. The CyQuant assay uses a sensitive DNA intercalating dye, similar to ethidium bromide that fluoresces when bound to DNA. Using the CyQuant dye, we can accurately measure the small amounts of DNA, which is proportional to the number of cells present. Because the CyQuant dye is more sensitive than ethidium bromide, we can determine small number of cells, yielding similar results to that of a clonogenic survival assay.

We first determined the correct number of cells to be seeded per well. We seeded different concentrations of LN428 cells and measured the fluorescence at 8, 9 and 10 days after mock treatments. We sought a concentration of cells that were within the linear range of the experiment (**Figure 25**). We selected 9 days of treatment to maintain a minimum of five doubling times, to increase the dynamic range of the experiment, and because it enabled greater than 50 cells per well which decreases pipetting errors and increases the precision of the experiment. We chose a concentration of 70 cells per well since the 100 cells per well appeared to have growth restriction, by comparing day 8 to day 9.

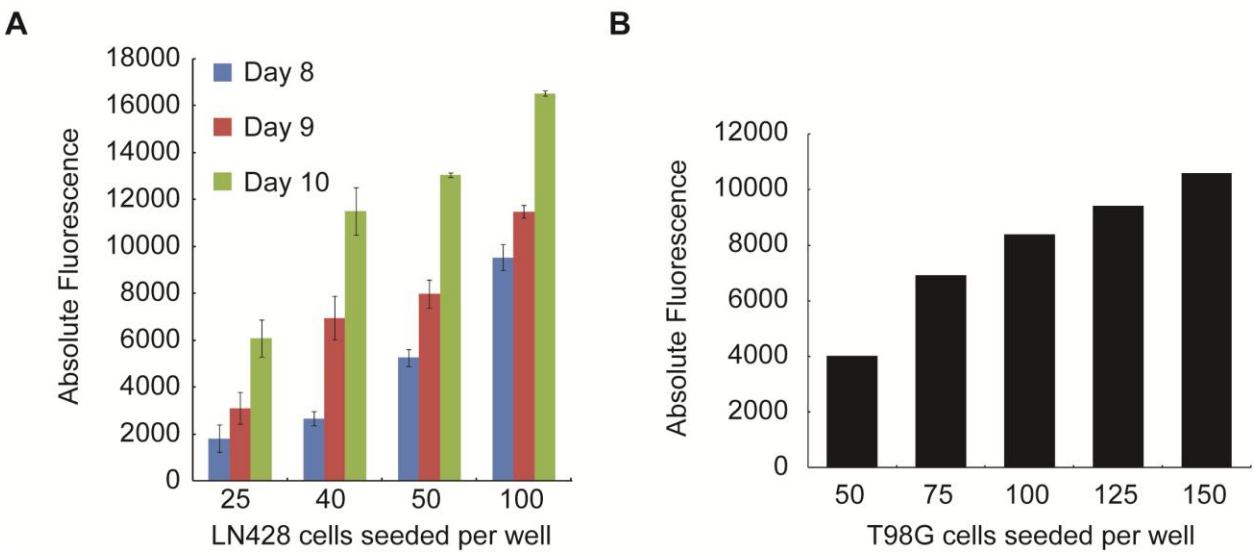


Figure 25. CyQuant cell seeding optimization for LN428 and T98G cells.
The CyQuant cell seeding experiments were done for the LN428 and T98G cell lines due to their different doubling times. **(**A**)** The LN428 cell line was seeded at different cell densities and had their fluorescence measured either 8, 9 or 10 days after mock treatment. A cell density of 70 cells per well and 9 days of treatment was selected as seeding density to maintain fluorescence measurements in the linear range. **(B)** The T98G cell line was seeded at different cell densities and measured 9 days after mock treatment. The 9 days were chosen to stay consistent with the LN428 cells. A final concentration of 125 cells per well was selected for the T98G cells to maintain measurements in the linear range.

We then validated the assay by using the LN428, LN428/MLH1-KD, LN428/MSH2-KD and LN428/MSH6-KD cell lines treated with the alkylating agent MNNG. The parental LN428 cells are resistant to alkylators in short-term cytotoxicity assays, but are sensitive to alkylators in long-term survival assays due to an intact mismatch repair and lack of MGMT expression (67, 70, 148). However, the LN428/MLH1-KD, and LN428/MSH2-KD and LN428/MSH6-KD cells are deficient in mismatch repair and are therefore resistant to alkylators (17, 19, 70). The CyQuant assay was validated as the LN428 cells were very sensitive to MNNG at very low doses, while all three MMR defective cell lines maintained very high survival rates, indicative of very resistant cell lines (**Figure 26**). With the CyQuant assay we can clearly determine alkylation survival similar to that of a clonogenic survival assay.

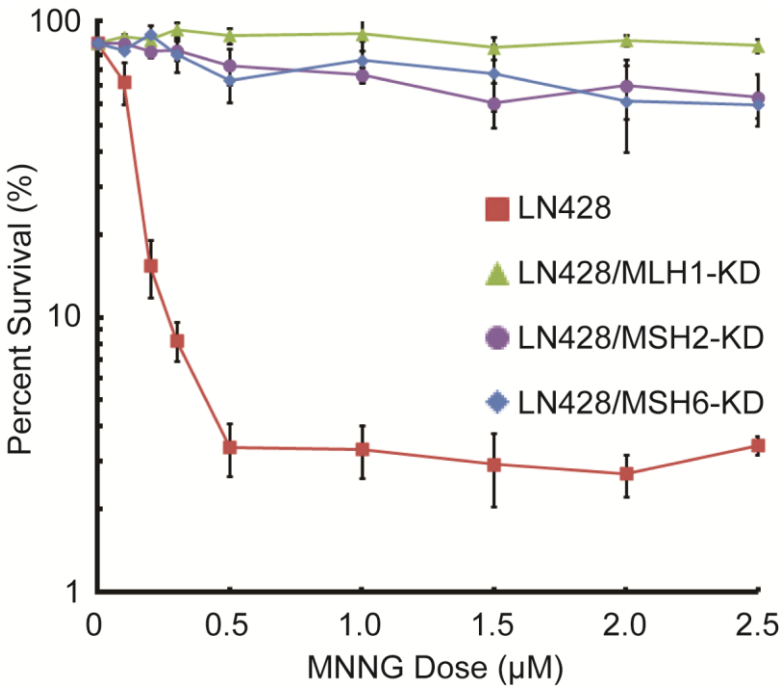


Figure 26. Parental LN428 cell line is sensitive to MNNG while cells defective in mismatch repair are resistant to DNA alkylating agents in long-term CyQuant assay.
Cells were seeded at 70 cells per well and treated with various doses of MNNG. The parental LN428 cells are sensitive to alkylators in a long-term assay due to a functional MMR pathway and lack of MGMT expression. The LN428 knockdown lines tested are defective in key enzymes in MMR and are therefore resistant to alkylation cytotoxicity.

APPENDIX B. SEQUENCE OF UBE3B CLONED FROM LN428

B.1 CLONE 2 – WILD-TYPE

Clone 2 contained wild-type sequence and was identical to the published PubMed sequence. The DNA sequence is below.

ATGTTCACCCTGTCTCAGACCTCGAGAGCATGGTTCATCGATAGAGCCCGTCAGGCACGAGAAGAAAGGCTTGTGCA
GAAGGAACGGGAGCGGGCAGCTGTTGTGATCCAGGCCCATGTCCGGAGTTTTCTCTGTCGGAGTCGACTGCAGAGAG
ATATCAGGAGAGAGATTGATGACTTTTTTAAAGCAGATGACCCTGAGTCCACTAAAAGAAGTGCACTTTGTATTTTC
AAGATTGCCAGGAAACTGCTGTTCCTATTCAGAATCAAAGAGGATAATGAGAGATTTGAGAAGTTGTGTCGCAGCAT
CCTGAGCAGCATGGATGCTGAGAATGAGCCTAAGGTGTGGTATGTGTCCCTGGCTTGTTCTAAGGACCTCACCCTCC
TTTGGATTCAACAGATCAAGAACATTTTGTGGTACTGCTGTGATTTTCTCAAGCAGCTCAAGCCTGAAATCCTGCAG
GACTCCCGACTCATCACCCTGTACCTCACGATGCTTGTCACCTTCACAGACACTTCAACGTGGAAAATTCTTCGGGG
AAAAGGTGAAAGTCTTCGACCAGCGATGAACCACATTTGTGCAAATATAATGGGACATCTCAACCAGCATGGATTTT
ATTCTGTGCTGCAGATATTGTTAACCCGTGGCCTGGCAAGACCCCGTCCTTGTCTATCCAAAGGCACTTTAACAGCA
GCTTTTTCTCTAGCGTTACGCCCTGTGATTGCTGCACAGTTCTCAGACAATCTGATTCGGCCGTTCCTCATCCACAT
CATGTCTGTGCCTGCTCTGGTGACTCATCTCAGCACAGTGACCCCTGAGCGCCTCACTGTTTTAGAATCCCATGACA
TGCTTCGTAAATTCATCATATTTTTAAGAGACCAAGATCGATGCCGTGATGTATGTGAAAGTTTAGAAGGATGCCAT
ACGCTTTGTCTAATGGGCAACCTCCTACACTTGGGCTCCCTCAGCCCCAGAGTGTTAGAGGAGGAGACAGATGGGTT
CGTGAGTTTGCTCACCCAGACGCTGTGCTACTGTCGGAAGTATGTGTCTCAGAAGAAGTCCAACCTGACCCACTGGC
ATCCTGTCCTTGGCTGGTTCTCCCAATCTGTGGATTATGGCCTTAACGAGTCAATGCACTTGATCACCAAACAGCTG
CAGTTCTTGTGGGGGGTGCCTCTGATCCGGATCTTCTTCTGTGACATCCTGAGCAAGAAGCTACTGGAGAGCCAGGA
GCCAGCCCACGCACAGCCAGCATCCCCTCAGAATGTGCTCCCAGTGAAGAGTCTCCTAAAGCGTGCTTTTCAAAAGT
CGGCATCAGTCCGGAATATTCTCAGGCCTGTCGGGGGTAAACGGGTCGACTCTGCAGAAGTCCAGAAGGTTTGCAAC
ATCTGTGTCCTCTACCAGACCTCGCTGACAACTCTCACACAGATTCGGCTGCAGATACTCACAGGTCTCACTTACCT
TGATGACCTGCTTCCCAAACTGTGGGCATTTATCTGTGAGCTCGGGCCCCACGGAGGGTTAAAGCTCTTCTTGGAAT
GCCTGAACAATGACACTGAAGAGTCCAAGCAACTCTTGGCCATGCTGATGCTGTTCTGTGACTGTTCGCGGCACCTC
ATCACAATCCTTGATGACATTGAAGTTTATGAAGAACAGATTTCATTCAAACTGGAAGAGCTGGTCACTATCTCCTC
TTTCCTGAATTCTTTTGTGTTTAAGATGATCTGGGATGGAATTGTAGAGAACGCCAAGGGTGAGACCTTGGAGCTGT
TCCAGTCTGTCCACGGGTGGCTTATGGTGCTGTACGAGCGGGACTGCCGGCGGCGCTTCACCCCCGAGGACCACTGG
CTGCGAAAGGATCTCAAACCTAGCGTGCTCTTCCAAGAACTCGACAGGGACAGAAAACGGGCACAGTTGATCCTGCA
GTACATCCCACATGTCATCCCTCACAAAAACAGAGTTCTACTGTTTCGAACCATGGTTACCAAGGAGAAGGAGAAAC
TGGGGCTGGTGGAAACCAGCTCTGCCTCCCCGCATGTCACTCACATCACCATCCGCCGGTCCAGGATGCTGGAGGAC
GGCTACGAGCAGCTTAGGCAGCTCTCCCAGCACGCCATGAAGGGGGTCATCCGTGTGAAGTTTGTCAATGACCTCGG
GGTGGACGAAGCAGGGATTGATCAAGACGGTGTTTTTAAGGAGTTCTTGGAAGAGATCATCAAGAGAGTTTTTGACC
CAGCACTCAATCTGTTCAAGACAACCAGTGGGGATGAGAGGCTGTACCCCTCACCCACATCCTACATCCATGAGAAT
TACCTGCAGCTCTTCGAGTTTGTGGGGAAGATGCTGGGGAAGGCTGTGTATGAGGGAATTGTGGTGGACGTGCCATT
TGCATCCTTCTTCCTGAGCCAACTGCTTGGGCACCACCACAGCGTCTTCTATAGCTCGGTGGATGAACTGCCTTCTC
TGGACTCCGAGTTCTATAAAAACCTCACCTCCATCAAGCGCTATGATGGGGACATCACTGACCTGGGCCTGACGCTG
TCTTACGACGAGGACGTCATGGGTCAGCTTGTTTGCCATGAACTGATTCCTGGAGGGAAGACCATTCCTGTTACAAA

TGAAAATAAAATTAGCTACATCCATCTGATGGCACATTTTCGAATGCACACTCAAATAAAAAACCAAACAGCTGCCC
TCATTAGCGGATTCCGTTCCATTATCAAACCCGAGTGGATCCGAATGTTCTCAACTCCTGAACTGCAGCGTCTCATC
TCTGGCGACAATGCTGAGATTGATCTGGAAGATTTAAAGAAGCACACAGTCTACTACGGTGGTTTCCATGGAAGTCA
CAGAGTCATCATCTGGCTCTGGGATATTCTGGCCTCCGACTTCACACCGGATGAGAGAGCTATGTTTCTGAAGTTCG
TGACCAGCTGCTCCAGACCCCCGCTCCTGGGATTCGCCTACCTCAAGCCTCCATTCTCCATCCGCTGCGTGGAGGTG
TCGGACGATCAGGACACCGGGGACACTCTGGGCAGCGTCCTCCGGGGCTTCTTCACCATCCGCAAGCGGGAGCCAGG
CGGCCGCCTGCCCACCTCCTCCACCTGCTTCAACCTGCTCAAGCTGCCCAACTACAGCAAGAAGAGCGTCCTCCGCG
AGAAGCTGCGCTACGCCATCAGCATGAACACGGGCTTTGAACTCTCCTAG

B.2 CLONE 4

Clone 4 contains two point mutants which change the amino acid sequence as follows Arg215Cys, Ile1059Phe. The DNA sequence is below.

ATGTTCACCCTGTCTCAGACCTCGAGAGCATGGTTCATCGATAGAGCCCGTCAGGCACGAGAAGAAAGGCTTGTGCA
GAAGGAACGGGAGCGGGCAGCTGTTGTGATCCAGGCCCATGTCCGGAGTTTTCTCTGTCGGAGTCGACTGCAGAGAG
ATATCAGGAGAGAGATTGATGACTTTTTTAAAGCAGATGACCCTGAGTCCACTAAAAGAAGTGCACTTTGTATTTTC
AAGATTGCCAGGAAACTGCTGTTCCTATTCAGAATCAAAGAGGATAATGAGAGATTTGAGAAGTTGTGTCGCAGCAT
CCTGAGCAGCATGGATGCTGAGAATGAGCCTAAGGTGTGGTATGTGTCCCTGGCTTGTTCTAAGGACCTCACCCTCC
TTTGGATTCAACAGATCAAGAACATTTTGTGGTACTGCTGTGATTTTCTCAAGCAGCTCAAGCCTGAAATCCTGCAG
GACTCCCGACTCATCACCCTGTACCTCACGATGCTTGTCACCTTCACAGACACTTCAACGTGGAAAATTCTTCGGGG
AAAAGGTGAAAGTCTTCGACCAGCGATGAACCACATTTGTGCAAATATAATGGGACATCTCAACCAGCATGGATTTT
ATTCTGTGCTGCAGATATTGTTAACCtGTGGCCTGGCAAGACCCCGTCCTTGTCTATCCAAAGGCACTTTAACAGCA
GCTTTTTCTCTAGCGTTACGCCCTGTGATTGCTGCACAGTTCTCAGACAATCTGATTCGGCCGTTCCTCATCCACAT
CATGTCTGTGCCTGCTCTGGTGACTCATCTCAGCACAGTGACCCCTGAGCGCCTCACTGTTTTAGAATCCCATGACA
TGCTTCGTAAATTCATCATATTTTTAAGAGACCAAGATCGATGCCGTGATGTATGTGAAAGTTTAGAAGGATGCCAT
ACGCTTTGTCTAATGGGCAACCTCCTACACTTGGGCTCCCTCAGCCCCAGAGTGTTAGAGGAGGAGACAGATGGGTT
CGTGAGTTTGCTCACCCAGACGCTGTGCTACTGTCGGAAGTATGTGTCTCAGAAGAAGTCCAACCTGACCCACTGGC
ATCCTGTCCTTGGCTGGTTCTCCCAATCTGTGGATTATGGCCTTAACGAGTCAATGCACTTGATCACCAAACAGCTG
CAGTTCTTGTGGGGGGTGCCTCTGATCCGGATCTTCTTCTGTGACATCCTGAGCAAGAAGCTACTGGAGAGCCAGGA
GCCAGCCCACGCACAGCCAGCATCCCCTCAGAATGTGCTCCCAGTGAAGAGTCTCCTAAAGCGTGCTTTTCAAAAGT
CGGCATCAGTCCGGAATATTCTCAGGCCTGTCGGGGGTAAACGGGTCGACTCTGCAGAAGTCCAGAAGGTTTGCAAC
ATCTGTGTCCTCTACCAGACCTCGCTGACAACTCTCACACAGATTCGGCTGCAGATACTCACAGGTCTCACTTACCT
TGATGACCTGCTTCCCAAACTGTGGGCATTTATCTGTGAGCTCGGGCCCCACGGAGGGTTAAAGCTCTTCTTGGAAT
GCCTGAACAATGACACTGAAGAGTCCAAGCAACTCTTGGCCATGCTGATGCTGTTCTGTGACTGTTCGCGGCACCTC
ATCACAATCCTTGATGACATTGAAGTTTATGAAGAACAGATTTCATTCAAACTGGAAGAGCTGGTCACTATCTCCTC
TTTCCTGAATTCTTTTGTGTTTAAGATGATCTGGGATGGAATTGTAGAGAACGCCAAGGGTGAGACCTTGGAGCTGT
TCCAGTCTGTCCACGGGTGGCTTATGGTGCTGTACGAGCGGGACTGCCGGCGGCGCTTCACCCCCGAGGACCACTGG
CTGCGAAAGGATCTCAAACCTAGCGTGCTCTTCCAAGAACTCGACAGGGACAGAAAACGGGCACAGTTGATCCTGCA
GTACATCCCACATGTCATCCCTCACAAAAACAGAGTTCTACTGTTTCGAACCATGGTTACCAAGGAGAAGGAGAAAC
TGGGGCTGGTGGAAACCAGCTCTGCCTCCCCGCATGTCACTCACATCACCATCCGCCGGTCCAGGATGCTGGAGGAC
GGCTACGAGCAGCTTAGGCAGCTCTCCCAGCACGCCATGAAGGGGGTCATCCGTGTGAAGTTTGTCAATGACCTCGG
GGTGGACGAAGCAGGGATTGATCAAGACGGTGTTTTTAAGGAGTTCTTGGAAGAGATCATCAAGAGAGTTTTTGACC
CAGCACTCAATCTGTTCAAGACAACCAGTGGGGATGAGAGGCTGTACCCCTCACCCACATCCTACATCCATGAGAAT
TACCTGCAGCTCTTCGAGTTTGTGGGGAAGATGCTGGGGAAGGCTGTGTATGAGGGAATTGTGGTGGACGTGCCATT
TGCATCCTTCTTCCTGAGCCAACTGCTTGGGCACCACCACAGCGTCTTCTATAGCTCGGTGGATGAACTGCCTTCTC
TGGACTCCGAGTTCTATAAAAACCTCACCTCCATCAAGCGCTATGATGGGGACATCACTGACCTGGGCCTGACGCTG
TCTTACGACGAGGACGTCATGGGTCAGCTTGTTTGCCATGAACTGATTCCTGGAGGGAAGACCATTCCTGTTACAAA
TGAAAATAAAATTAGCTACATCCATCTGATGGCACATTTTCGAATGCACACTCAAATAAAAAACCAAACAGCTGCCC

TCATTAGCGGATTCCGTTCCATTATCAAACCCGAGTGGATCCGAATGTTCTCAACTCCTGAACTGCAGCGTCTCATC
TCTGGCGACAATGCTGAGATTGATCTGGAAGATTTAAAGAAGCACACAGTCTACTACGGTGGTTTCCATGGAAGTCA
CAGAGTCATCATCTGGCTCTGGGATATTCTGGCCTCCGACTTCACACCGGATGAGAGAGCTATGTTTCTGAAGTTCG
TGACCAGCTGCTCCAGACCCCCGCTCCTGGGATTCGCCTACCTCAAGCCTCCATTCTCCATCCGCTGCGTGGAGGTG
TCGGACGATCAGGACACCGGGGACACTCTGGGCAGCGTCCTCCGGGGCTTCTTCACCATCCGCAAGCGGGAGCCAGG
CGGCCGCCTGCCCACCTCCTCCACCTGCTTCAACCTGCTCAAGCTGCCCAACTACAGCAAGAAGAGCGTCCTCCGCG
AGAAGCTGCGCTACGCCtTCAGCATGAACACGGGCTTTGAACTCTCCTAG

B.3 CLONE 5

Clone 5 contains 1 point mutant which change the amino acid sequence as follows, Arg346Gln. The DNA sequence of the clone is below.

ATGTTCACCCTGTCTCAGACCTCGAGAGCATGGTTCATCGATAGAGCCCGTCAGGCACGAGAAGAAAGGCTTGTGCA
GAAGGAACGGGAGCGGGCAGCTGTTGTGATCCAGGCCCATGTCCGGAGTTTTCTCTGTCGGAGTCGACTGCAGAGAG
ATATCAGGAGAGAGATTGATGACTTTTTTAAAGCAGATGACCCTGAGTCCACTAAAAGAAGTGCACTTTGTATTTTC
AAGATTGCCAGGAAACTGCTGTTCCTATTCAGAATCAAAGAGGATAATGAGAGATTTGAGAAGTTGTGTCGCAGCAT
CCTGAGCAGCATGGATGCTGAGAATGAGCCTAAGGTGTGGTATGTGTCCCTGGCTTGTTCTAAGGACCTCACCCTCC
TTTGGATTCAACAGATCAAGAACATTTTGTGGTACTGCTGTGATTTTCTCAAGCAGCTCAAGCCTGAAATCCTGCAG
GACTCCCGACTCATCACCCTGTACCTCACGATGCTTGTCACCTTCACAGACACTTCAACGTGGAAAATTCTTCGGGG
AAAAGGTGAAAGTCTTCGACCAGCGATGAACCACATTTGTGCAAATATAATGGGACATCTCAACCAGCATGGATTTT
ATTCTGTGCTGCAGATATTGTTAACCCGTGGCCTGGCAAGACCCCGTCCTTGTCTATCCAAAGGCACTTTAACAGCA
GCTTTTTCTCTAGCGTTACGCCCTGTGATTGCTGCACAGTTCTCAGACAATCTGATTCGGCCGTTCCTCATCCACAT
CATGTCTGTGCCTGCTCTGGTGACTCATCTCAGCACAGTGACCCCTGAGCGCCTCACTGTTTTAGAATCCCATGACA
TGCTTCGTAAATTCATCATATTTTTAAGAGACCAAGATCGATGCCGTGATGTATGTGAAAGTTTAGAAGGATGCCAT
ACGCTTTGTCTAATGGGCAACCTCCTACACTTGGGCTCCCTCAGCCCCAGAGTGTTAGAGGAGGAGACAGATGGGTT
CGTGAGTTTGCTCACCCAGACGCTGTGCTACTGTCAGAAGTATGTGTCTCAGAAGAAGTCCAACCTGACCCACTGGC
ATCCTGTCCTTGGCTGGTTCTCCCAATCTGTGGATTATGGCCTTAACGAGTCAATGCACTTGATCACCAAACAGCTG
CAGTTCTTGTGGGGGGTGCCTCTGATCCGGATCTTCTTCTGTGACATCCTGAGCAAGAAGCTACTGGAGAGCCAGGA
GCCAGCCCACGCACAGCCAGCATCCCCTCAGAATGTGCTCCCAGTGAAGAGTCTCCTAAAGCGTGCTTTTCAAAAGT
CGGCATCAGTCCGGAATATTCTCAGGCCTGTCGGGGGTAAACGGGTCGACTCTGCAGAAGTCCAGAAGGTTTGCAAC
ATCTGTGTCCTCTACCAGACCTCGCTGACAACTCTCACACAGATTCGGCTGCAGATACTCACAGGTCTCACTTACCT
TGATGACCTGCTTCCCAAACTGTGGGCATTTATCTGTGAGCTCGGGCCCCACGGAGGGTTAAAGCTCTTCTTGGAAT
GCCTGAACAATGACACTGAAGAGTCCAAGCAACTCTTGGCCATGCTGATGCTGTTCTGTGACTGTTCGCGGCACCTC
ATCACAATCCTTGATGACATTGAAGTTTATGAAGAACAGATTTCATTCAAACTGGAAGAGCTGGTCACTATCTCCTC
TTTCCTGAATTCTTTTGTGTTTAAGATGATCTGGGATGGAATTGTAGAGAACGCCAAGGGTGAGACCTTGGAGCTGT
TCCAGTCTGTCCACGGGTGGCTTATGGTGCTGTACGAGCGGGACTGCCGGCGGCGCTTCACCCCCGAGGACCACTGG
CTGCGAAAGGATCTCAAACCTAGCGTGCTCTTCCAAGAACTCGACAGGGACAGAAAACGGGCACAGTTGATCCTGCA
GTACATCCCACATGTCATCCCTCACAAAAACAGAGTTCTACTGTTTCGAACCATGGTTACCAAGGAGAAGGAGAAAC
TGGGGCTGGTGGAAACCAGCTCTGCCTCCCCGCATGTCACTCACATCACCATCCGCCGGTCCAGGATGCTGGAGGAC
GGCTACGAGCAGCTTAGGCAGCTCTCCCAGCACGCCATGAAGGGGGTCATCCGTGTGAAGTTTGTCAATGACCTCGG
GGTGGACGAAGCAGGGATTGATCAAGACGGTGTTTTTAAGGAGTTCTTGGAAGAGATCATCAAGAGAGTTTTTGACC
CAGCACTCAATCTGTTCAAGACAACCAGTGGGGATGAGAGGCTGTACCCCTCACCCACATCCTACATCCATGAGAAT
TACCTGCAGCTCTTCGAGTTTGTGGGGAAGATGCTGGGGAAGGCTGTGTATGAGGGAATTGTGGTGGACGTGCCATT
TGCATCCTTCTTCCTGAGCCAACTGCTTGGGCACCACCACAGCGTCTTCTATAGCTCGGTGGATGAACTGCCTTCTC
TGGACTCCGAGTTCTATAAAAACCTCACCTCCATCAAGCGCTATGATGGGGACATCACTGACCTGGGCCTGACGCTG

TCTTACGACGAGGACGTCATGGGTCAGCTTGTTTGCCATGAACTGATTCCTGGAGGGAAGACCATTCCTGTTACAAA
TGAAAATAAAATTAGCTACATCCATCTGATGGCACATTTTCGAATGCACACTCAAATAAAAAACCAAACAGCTGCCC
TCATTAGCGGATTCCGTTCCATTATCAAACCCGAGTGGATCCGAATGTTCTCAACTCCTGAACTGCAGCGTCTCATC
TCTGGCGACAATGCTGAGATTGATCTGGAAGATTTAAAGAAGCACACAGTCTACTACGGTGGTTTCCATGGAAGTCA
CAGAGTCATCATCTGGCTCTGGGATATTCTGGCCTCCGACTTCACACCGGATGAGAGAGCTATGTTTCTGAAGTTCG
TGACCAGCTGCTCCAGACCCCCGCTCCTGGGATTCGCCTACCTCAAGCCTCCATTCTCCATCCGCTGCGTGGAGGTG
TCGGACGATCAGGACACCGGGGACACTCTGGGCAGCGTCCTCCGGGGCTTCTTCACCATCCGCAAGCGGGAGCCAGG
CGGCCGCCTGCCCACCTCCTCCACCTGCTTCAACCTGCTCAAGCTGCCCAACTACAGCAAGAAGAGCGTCCTCCGCG
AGAAGCTGCGCTACGCCATCAGCATGAACACGGGCTTTGAACTCTCCTAG

B.4 CLONE 6

Clone 6 contains two point mutants that would change the amino acid sequence as follows, Arg346Gln, Glu757Stop. DNA sequence is below.

ATGTTCACCCTGTCTCAGACCTCGAGAGCATGGTTCATCGATAGAGCCCGTCAGGCACGAGAAGAAAGGCTTGTGCA
GAAGGAACGGGAGCGGGCAGCTGTTGTGATCCAGGCCCATGTCCGGAGTTTTCTCTGTCGGAGTCGACTGCAGAGAG
ATATCAGGAGAGAGATTGATGACTTTTTTAAAGCAGATGACCCTGAGTCCACTAAAAGAAGTGCACTTTGTATTTTC
AAGATTGCCAGGAAACTGCTGTTCCTATTCAGAATCAAAGAGGATAATGAGAGATTTGAGAAGTTGTGTCGCAGCAT
CCTGAGCAGCATGGATGCTGAGAATGAGCCTAAGGTGTGGTATGTGTCCCTGGCTTGTTCTAAGGACCTCACCCTCC
TTTGGATTCAACAGATCAAGAACATTTTGTGGTACTGCTGTGATTTTCTCAAGCAGCTCAAGCCTGAAATCCTGCAG
GACTCCCGACTCATCACCCTGTACCTCACGATGCTTGTCACCTTCACAGACACTTCAACGTGGAAAATTCTTCGGGG
AAAAGGTGAAAGTCTTCGACCAGCGATGAACCACATTTGTGCAAATATAATGGGACATCTCAACCAGCATGGATTTT
ATTCTGTGCTGCAGATATTGTTAACCCGTGGCCTGGCAAGACCCCGTCCTTGTCTATCCAAAGGCACTTTAACAGCA
GCTTTTTCTCTAGCGTTACGCCCTGTGATTGCTGCACAGTTCTCAGACAATCTGATTCGGCCGTTCCTCATCCACAT
CATGTCTGTGCCTGCTCTGGTGACTCATCTCAGCACAGTGACCCCTGAGCGCCTCACTGTTTTAGAATCCCATGACA
TGCTTCGTAAATTCATCATATTTTTAAGAGACCAAGATCGATGCCGTGATGTATGTGAAAGTTTAGAAGGATGCCAT
ACGCTTTGTCTAATGGGCAACCTCCTACACTTGGGCTCCCTCAGCCCCAGAGTGTTAGAGGAGGAGACAGATGGGTT
CGTGAGTTTGCTCACCCAGACGCTGTGCTACTGTCaGAAGTATGTGTCTCAGAAGAAGTCCAACCTGACCCACTGGC
ATCCTGTCCTTGGCTGGTTCTCCCAATCTGTGGATTATGGCCTTAACGAGTCAATGCACTTGATCACCAAACAGCTG
CAGTTCTTGTGGGGGGTGCCTCTGATCCGGATCTTCTTCTGTGACATCCTGAGCAAGAAGCTACTGGAGAGCCAGGA
GCCAGCCCACGCACAGCCAGCATCCCCTCAGAATGTGCTCCCAGTGAAGAGTCTCCTAAAGCGTGCTTTTCAAAAGT
CGGCATCAGTCCGGAATATTCTCAGGCCTGTCGGGGGTAAACGGGTCGACTCTGCAGAAGTCCAGAAGGTTTGCAAC
ATCTGTGTCCTCTACCAGACCTCGCTGACAACTCTCACACAGATTCGGCTGCAGATACTCACAGGTCTCACTTACCT
TGATGACCTGCTTCCCAAACTGTGGGCATTTATCTGTGAGCTCGGGCCCCACGGAGGGTTAAAGCTCTTCTTGGAAT
GCCTGAACAATGACACTGAAGAGTCCAAGCAACTCTTGGCCATGCTGATGCTGTTCTGTGACTGTTCGCGGCACCTC
ATCACAATCCTTGATGACATTGAAGTTTATGAAGAACAGATTTCATTCAAACTGGAAGAGCTGGTCACTATCTCCTC
TTTCCTGAATTCTTTTGTGTTTAAGATGATCTGGGATGGAATTGTAGAGAACGCCAAGGGTGAGACCTTGGAGCTGT
TCCAGTCTGTCCACGGGTGGCTTATGGTGCTGTACGAGCGGGACTGCCGGCGGCGCTTCACCCCCGAGGACCACTGG
CTGCGAAAGGATCTCAAACCTAGCGTGCTCTTCCAAGAACTCGACAGGGACAGAAAACGGGCACAGTTGATCCTGCA
GTACATCCCACATGTCATCCCTCACAAAAACAGAGTTCTACTGTTTCGAACCATGGTTACCAAGGAGAAGGAGAAAC
TGGGGCTGGTGGAAACCAGCTCTGCCTCCCCGCATGTCACTCACATCACCATCCGCCGGTCCAGGATGCTGGAGGAC
GGCTACGAGCAGCTTAGGCAGCTCTCCCAGCACGCCATGAAGGGGGTCATCCGTGTGAAGTTTGTCAATGACCTCGG
GGTGGACGAAGCAGGGATTGATCAAGACGGTGTTTTTAAGGAGTTCTTGGAAGAGATCATCAAGAGAGTTTTTGACC
CAGCACTCAATCTGTTCAAGACAACCAGTGGGGATtAGAGGCTGTACCCCTCACCCACATCCTACATCCATGAGAAT
TACCTGCAGCTCTTCGAGTTTGTGGGGAAGATGCTGGGGAAGGCTGTGTATGAGGGAATTGTGGTGGACGTGCCATT
TGCATCCTTCTTCCTGAGCCAACTGCTTGGGCACCACCACAGCGTCTTCTATAGCTCGGTGGATGAACTGCCTTCTC
TGGACTCCGAGTTCTATAAAAACCTCACCTCCATCAAGCGCTATGATGGGGACATCACTGACCTGGGCCTGACGCTG
TCTTACGACGAGGACGTCATGGGTCAGCTTGTTTGCCATGAACTGATTCCTGGAGGGAAGACCATTCCTGTTACAAA

TGAAAATAAAATTAGCTACATCCATCTGATGGCACATTTTCGAATGCACACTCAAATAAAAAACCAAACAGCTGCCC
TCATTAGCGGATTCCGTTCCATTATCAAACCCGAGTGGATCCGAATGTTCTCAACTCCTGAACTGCAGCGTCTCATC
TCTGGCGACAATGCTGAGATTGATCTGGAAGATTTAAAGAAGCACACAGTCTACTACGGTGGTTTCCATGGAAGTCA
CAGAGTCATCATCTGGCTCTGGGATATTCTGGCCTCCGACTTCACACCGGATGAGAGAGCTATGTTTCTGAAGTTCG
TGACCAGCTGCTCCAGACCCCCGCTCCTGGGATTCGCCTACCTCAAGCCTCCATTCTCCATCCGCTGCGTGGAGGTG
TCGGACGATCAGGACACCGGGGACACTCTGGGCAGCGTCCTCCGGGGCTTCTTCACCATCCGCAAGCGGGAGCCAGG
CGGCCGCCTGCCCACCTCCTCCACCTGCTTCAACCTGCTCAAGCTGCCCAACTACAGCAAGAAGAGCGTCCTCCGCG
AGAAGCTGCGCTACGCCATCAGCATGAACACGGGCTTTGAACTCTCCTAG

APPENDIX C. SIRNA SCREEN DATASETS

Dataset 1.

Accession No	Gene Symbol	Avg POC Ratio	P-Value	FDR
NM_018324	THEDC1	0.429536424	0.0209992	0.1273799
NM_004666	VNN1	0.429964133	0.0240549	0.1341246
NM_080911	UNG	0.445720958	0.0154252	0.1129269
XM_165973	USP24	0.454884119	0.0314827	0.1476503
NM_021970	MAP2K1IP1	0.468591817	0.0403303	0.1595864
NM_002933	RNASE1	0.469939834	0.0133807	0.109913
NM_153189	SPAM1	0.486478025	0.0415279	0.1605281
NM_138340	ABHD3	0.486532905	0.0067143	0.0919669
NM_013251	TAC3	0.491722626	0.0279567	0.1404196
NM_006327	TIMM23	0.501591174	0.0036578	0.0810883
NM_003841	TNFRSF10C	0.512219543	0.0140413	0.1100968
NM_172236	POFUT1	0.521360956	0.0160409	0.1146967
NM_003799	RNMT	0.521827669	0.0118005	0.106785
NM_003839	TNFRSF11A	0.521925008	0.0147432	0.111789
NM_001025243	IRAK1	0.524453336	0.0053168	0.0860671
NM_005718	ARPC4	0.526481464	0.0136158	0.1102038
NM_002309	LIF	0.53479142	0.0026576	0.081955
NM_173353	TPH2	0.535499705	0.0315239	0.1475927
NM_012125	CHRM5	0.535566907	0.0201234	0.1250917
NM_012360	OR1F1	0.540526776	0.0168029	0.1166694
NM_003211	TDG	0.541857164	0.0034462	0.0806063
NM_130767	CACH-1	0.544235702	0.0455433	0.1651768
NM_005762	TRIM28	0.545059748	0.028319	0.1409565
NM_001876	CPT1A	0.54638003	0.0023785	0.0805481
NM_033181	CNR1	0.548978297	0.0084768	0.0972802
NM_017686	GDAP2	0.549407967	0.0317187	0.1478776
NM_145910	NEK11	0.551522047	0.0006802	0.0682706
NM_014465	SULT1B1	0.552146948	0.0173686	0.118951
NM_012387	PADI4	0.552576488	0.0187445	0.1220159
NM_001702	BAI1	0.556730459	7.658E-05	0.0528393
NM_199040	NUDT4	0.557359244	0.0072567	0.0949217
NM_024608	NEIL1	0.559400704	0.0324636	0.1483438

NM_003227	TFR2	0.559841619	0.0059989	0.0878351
NM_012214	MGAT4A	0.560188918	0.0103896	0.1038964
NM_001184	ATR	0.5610753	0.0363013	0.1534327
NM_003268	TLR5	0.563577132	0.0043176	0.085424
NM_003071	SMARCA3	0.564168075	0.0291905	0.1438677
NM_002934	RNASE2	0.564991995	0.001158	0.0718247
NM_002572	PAFAH1B2	0.565111742	0.0373269	0.1552708
NM_001866	COX7B	0.565778198	0.0222086	0.1302777
NM_003272	TM7SF1	0.569465132	0.0455148	0.1652907
NM_005508	CCR4	0.569924369	0.0006306	0.0740637
NM_006714	SMPDL3A	0.571885333	0.0153518	0.1132911
NM_153446	GALGT2	0.57287254	0.0011754	0.0705242
NM_001024592	UNG2	0.573614706	5.972E-05	0.0470927
NM_001002019	PUS1	0.574720416	0.0017126	0.0807981
NM_024342	GRLF1	0.575297825	0.0034254	0.080804
NM_000814	GABRB3	0.575759866	0.0415214	0.1606154
NM_000536	RAG2	0.575804467	0.0125883	0.1080676
NM_007255	B4GALT7	0.575961939	0.0005847	0.0733561
NM_177552	SULT1A3	0.577010219	0.000668	0.0695751
NM_003650	CST7	0.579687602	0.0174241	0.1190359
NM_023018	FLJ13052	0.581448317	0.0020006	0.0824126
NM_000961	PTGIS	0.58160867	0.0199218	0.1245393
NM_000150	FUT6	0.582777658	0.0218182	0.129641
NM_001775	CD38	0.583916315	0.0497526	0.1706863
NM_170706	NMNAT2	0.586556927	0.015877	0.114116
NM_014058	DESC1	0.586739966	0.0331931	0.1496942
NM_000926	PGR	0.587029255	9.715E-05	0.0446905
NM_138578	BCL2L1	0.587481551	0.0149647	0.1123878
NM_145207	SPATA5	0.588113844	0.0077668	0.0959127
NM_004566	PFKFB3	0.588865827	0.0136729	0.1098611
NM_004443	EPHB3	0.589677553	0.0187482	0.1218965
NM_002661	PLCG2	0.590020743	0.035635	0.152959
NM_006894	FMO3	0.590684268	0.0300392	0.145071
NM_001667	ARL2	0.590825053	0.0015146	0.0760068
NM_032236	USP48	0.590977992	0.0436502	0.1640225
NM_018010	ESRRBL1	0.596206886	0.0084695	0.0973997
NM_007171	POMT1	0.596482185	0.0442913	0.1639757
NM_130806	LGR8	0.596729273	0.0263943	0.1373202
NM_019092	KIAA1164	0.597090987	0.014906	0.1125597
NM_002754	MAPK13	0.597218471	0.0029406	0.0780399
NM_032595	PPP1R9B	0.597285454	0.019346	0.1237426
NM_001003962	CAPNS1	0.597680159	0.0254831	0.1364369
NM_174971	SIAT6	0.59826402	0.0461521	0.1657512
NM_000234	LIG1	0.599034365	0.0003014	0.063993
NM_003137	SRPK1	0.599583268	0.0075878	0.0960651

NM_004117	FKBP5	0.600816686	0.0066917	0.0921147
NM_000855	GUCY1A2	0.60113068	0.0110279	0.1058683
NM_004154	P2RY6	0.601188399	0.0498261	0.1707263
NM_000351	STS	0.601222343	0.0092447	0.0994753
NM_198951	TGM2	0.601758521	0.0009684	0.0685309
NM_002021	FMO1	0.602065855	0.0229544	0.1310323
NM_144581	C14orf149	0.602300115	0.0116555	0.1061689
NM_000320	QDPR	0.602997311	0.0083105	0.0969856
NM_003161	RPS6KB1	0.603063439	0.0273024	0.1390307
NM_003918	GYG2	0.603714359	0.011583	0.1062095
NM_014256	B3GNT3	0.604006836	0.0011821	0.0686863
NM_001056	SULT1C1	0.604916863	0.0017337	0.0797486
NM_022098	LOC63929	0.605243397	0.0055014	0.0867646
NM_013358	PADI1	0.606979719	0.0090657	0.0990943
NM_000741	CHRM4	0.607206284	0.0405204	0.1596521
NM_181578	RFC5	0.607822106	0.0015698	0.0780657
NM_032582	USP32	0.608711231	0.0489132	0.1692796
XM_371285	LOC128102	0.609257966	0.0073032	0.0950794
NM_130384	TREX1	0.609335989	0.0464234	0.1660772
NM_018413	CHST11	0.609436442	0.002225	0.0802728
NM_016616	TXNDC3	0.60962619	0.0115702	0.1062685
NM_002663	PLD2	0.610103282	0.0038365	0.0820837
NM_147173	NUDT2	0.61043487	0.030907	0.1470748
NM_005518	HMGCS2	0.611468472	0.0268483	0.13812
NM_020378	KLP1	0.612693702	0.0017788	0.0804841
NM_003390	WEE1	0.613803508	0.0118051	0.1066516
NM_138608	MPPE1	0.613884159	0.0324861	0.1483237
NM_000353	TAT	0.61573104	0.0020768	0.080167
NM_153002	GPR156	0.616499941	0.0403791	0.1594369
NM_004388	CTBS	0.616778083	0.0065255	0.0905043
NM_000199	SGSH	0.618577267	0.0091277	0.0995745
NM_153809	TAF1L	0.619586798	0.0494498	0.1699644
NM_033294	CASP1	0.620480218	0.0035705	0.0807748
NM_182547	HNLF	0.621360183	0.0044515	0.0856168
NM_003986	BBOX1	0.622305367	0.0239427	0.1340401
NM_025227	BPIL1	0.625613404	0.0049732	0.0852548
NM_007041	ATE1	0.626081973	0.0115264	0.1062202
NM_016347	CML2	0.626721943	0.0111975	0.1060213
NM_005309	GPT	0.626783436	0.0011814	0.069373
NM_005173	ATP2A3	0.627154941	0.0140974	0.1100674
NM_020960	GPR107	0.627281238	0.0033474	0.0796453
NM_197941	ADAMTS6	0.627476229	0.0445869	0.1639703
NM_183415	UBE3B	0.62761405	0.0092349	0.0995638
NM_172234	IL17RB	0.629036343	0.0372194	0.1549404
NM_002014	FKBP4	0.629862888	0.0003133	0.0617624

NM_004054	C3AR1	0.630710213	0.0154231	0.1130619
NM_138327	TRAR1	0.631346468	0.0054855	0.0867625
NM_014275	MGAT4B	0.631640548	0.0023032	0.0804675
NM_000875	IGF1R	0.633037344	0.0121453	0.1072677
NM_001355	DDT	0.633879169	0.0124526	0.108249
NM_001445	FABP6	0.634237146	0.0336435	0.1508629
NM_012460	TIMM9	0.635270062	0.004009	0.0831943
NM_170705	ICMT	0.635701849	0.003514	0.0801539
NM_005373	MPL	0.636111958	0.0293957	0.1438515
NM_005605	PPP3CC	0.636322741	0.0003141	0.0597816
NM_001683	ATP2B2	0.637090269	0.0351494	0.1524154
NM_001008505	OPRM1	0.637171567	0.0042776	0.0852424
NM_001005619	ITGB4	0.637353948	0.0133707	0.1101584
NM_007180	TREH	0.637408937	0.000708	0.0685671
NM_013382	POMT2	0.637409881	0.0389863	0.1574281
NM_020438	DOLPP1	0.638044755	0.0426131	0.1616661
NM_022120	OXCT2	0.638353099	0.0071393	0.09428
NM_006610	MASP2	0.638566567	0.0236021	0.1334871
NM_198066	GNPNAT1	0.639193548	0.0135816	0.1102507
NM_005471	GNPDA1	0.639644785	0.0373323	0.1551763
NM_014911	AAK1	0.640408077	0.0452324	0.1649161
NM_002103	GYS1	0.640761325	0.0010296	0.0710457
NM_024681	FLJ12242	0.641444177	0.0355712	0.1530422
NM_002312	LIG4	0.641576109	0.0122672	0.1074837
NM_001679	ATP1B3	0.642017327	0.007819	0.0959135
NM_181468	ITGB4BP	0.642270855	0.0270808	0.1384132
NM_022341	PDF	0.642691534	0.0395572	0.1583435
NM_003654	CHST1	0.644429746	0.0044646	0.0855709
NM_153699	GSTA5	0.644761056	0.0044309	0.0858199
NM_000797	DRD4	0.644792449	0.0358351	0.1529851
NM_000481	AMT	0.645664635	0.0008662	0.0654976
NM_054014	FKBP1A	0.64576482	0.0003291	0.0567634
NM_003595	TPST2	0.645769416	0.0063442	0.0897951
NM_000740	CHRM3	0.645843857	0.0076131	0.0959463
NM_177524	MEST	0.645866015	0.0428969	0.1622965
NM_002676	PMM1	0.64604996	0.0002546	0.0610985
NM_004624	VIPR1	0.646071697	0.0051226	0.0854283
NM_198974	PTK9	0.646260808	0.001041	0.0709427
NM_002691	POLD1	0.646310661	0.0396336	0.1584197
NM_000750	CHRNB4	0.646724844	0.0410758	0.160466
NM_002630	PGC	0.646801181	0.0429236	0.1622865
NM_018060	FLJ10326	0.64683082	0.0338063	0.1508573
NM_014413	HRI	0.648032791	0.0048498	0.0858047
NM_002542	OGG1	0.64860652	0.0020099	0.0821818
NM_178221	APG4C	0.649558894	0.0110738	0.1059404

NM_197975	BTNL3	0.649616816	0.0130022	0.1089109
NM_002410	MGAT5	0.650583395	0.0077468	0.0960949
NM_022481	ARAP3	0.650767829	0.0156172	0.1137292
NM_001034	RRM2	0.651065915	0.0416831	0.1606777
NM_152942	TNFRSF8	0.651437898	0.0193805	0.1238198

Dataset 2.

Accession No	Gene Symbol	Avg_POC_Ratio	P-Value	FDR
NM_018324	THEDC1	0.429536424	0.0209992	0.1273799
NM_004666	VNN1	0.429964133	0.0240549	0.1341246
NM_080911	UNG	0.445720958	0.0154252	0.1129269
XM_165973	USP24	0.454884119	0.0314827	0.1476503
NM_021970	MAP2K1IP1	0.468591817	0.0403303	0.1595864
NM_002933	RNASE1	0.469939834	0.0133807	0.109913
NM_153189	SPAM1	0.486478025	0.0415279	0.1605281
NM_138340	ABHD3	0.486532905	0.0067143	0.0919669
NM_013251	TAC3	0.491722626	0.0279567	0.1404196
NM_006327	TIMM23	0.501591174	0.0036578	0.0810883
NM_003841	TNFRSF10C	0.512219543	0.0140413	0.1100968
NM_172236	POFUT1	0.521360956	0.0160409	0.1146967
NM_003799	RNMT	0.521827669	0.0118005	0.106785
NM_003839	TNFRSF11A	0.521925008	0.0147432	0.111789
NM_001025243	IRAK1	0.524453336	0.0053168	0.0860671
NM_005718	ARPC4	0.526481464	0.0136158	0.1102038
NM_002309	LIF	0.53479142	0.0026576	0.081955
NM_173353	TPH2	0.535499705	0.0315239	0.1475927
NM_012125	CHRM5	0.535566907	0.0201234	0.1250917
NM_012360	OR1F1	0.540526776	0.0168029	0.1166694
NM_003211	TDG	0.541857164	0.0034462	0.0806063
NM_130767	CACH-1	0.544235702	0.0455433	0.1651768
NM_005762	TRIM28	0.545059748	0.028319	0.1409565
NM_001876	CPT1A	0.54638003	0.0023785	0.0805481
NM_033181	CNR1	0.548978297	0.0084768	0.0972802
NM_017686	GDAP2	0.549407967	0.0317187	0.1478776
NM_145910	NEK11	0.551522047	0.0006802	0.0682706
NM_014465	SULT1B1	0.552146948	0.0173686	0.118951
NM_012387	PADI4	0.552576488	0.0187445	0.1220159
NM_001702	BAI1	0.556730459	7.658E-05	0.0528393
NM_199040	NUDT4	0.557359244	0.0072567	0.0949217
NM_024608	NEIL1	0.559400704	0.0324636	0.1483438

NM_003227	TFR2	0.559841619	0.0059989	0.0878351
NM_012214	MGAT4A	0.560188918	0.0103896	0.1038964
NM_001184	ATR	0.5610753	0.0363013	0.1534327
NM_003268	TLR5	0.563577132	0.0043176	0.085424
NM_003071	SMARCA3	0.564168075	0.0291905	0.1438677
NM_002934	RNASE2	0.564991995	0.001158	0.0718247
NM_002572	PAFAH1B2	0.565111742	0.0373269	0.1552708
NM_001866	COX7B	0.565778198	0.0222086	0.1302777
NM_003272	TM7SF1	0.569465132	0.0455148	0.1652907
NM_005508	CCR4	0.569924369	0.0006306	0.0740637
NM_006714	SMPDL3A	0.571885333	0.0153518	0.1132911
NM_153446	GALGT2	0.57287254	0.0011754	0.0705242
NM_001024592	UNG2	0.573614706	5.972E-05	0.0470927
NM_001002019	PUS1	0.574720416	0.0017126	0.0807981
NM_024342	GRLF1	0.575297825	0.0034254	0.080804
NM_000814	GABRB3	0.575759866	0.0415214	0.1606154
NM_000536	RAG2	0.575804467	0.0125883	0.1080676
NM_007255	B4GALT7	0.575961939	0.0005847	0.0733561
NM_177552	SULT1A3	0.577010219	0.000668	0.0695751
NM_003650	CST7	0.579687602	0.0174241	0.1190359
NM_023018	FLJ13052	0.581448317	0.0020006	0.0824126
NM_000961	PTGIS	0.58160867	0.0199218	0.1245393
NM_000150	FUT6	0.582777658	0.0218182	0.129641
NM_001775	CD38	0.583916315	0.0497526	0.1706863
NM_170706	NMNAT2	0.586556927	0.015877	0.114116
NM_014058	DESC1	0.586739966	0.0331931	0.1496942
NM_000926	PGR	0.587029255	9.715E-05	0.0446905
NM_138578	BCL2L1	0.587481551	0.0149647	0.1123878
NM_145207	SPATA5	0.588113844	0.0077668	0.0959127
NM_004566	PFKFB3	0.588865827	0.0136729	0.1098611
NM_004443	EPHB3	0.589677553	0.0187482	0.1218965
NM_002661	PLCG2	0.590020743	0.035635	0.152959
NM_006894	FMO3	0.590684268	0.0300392	0.145071
NM_001667	ARL2	0.590825053	0.0015146	0.0760068
NM_032236	USP48	0.590977992	0.0436502	0.1640225
NM_018010	ESRRBL1	0.596206886	0.0084695	0.0973997
NM_007171	POMT1	0.596482185	0.0442913	0.1639757
NM_130806	LGR8	0.596729273	0.0263943	0.1373202
NM_019092	KIAA1164	0.597090987	0.014906	0.1125597
NM_002754	MAPK13	0.597218471	0.0029406	0.0780399
NM_032595	PPP1R9B	0.597285454	0.019346	0.1237426

NM_001003962	CAPNS1	0.597680159	0.0254831	0.1364369
NM_174971	SIAT6	0.59826402	0.0461521	0.1657512
NM_000234	LIG1	0.599034365	0.0003014	0.063993
NM_003137	SRPK1	0.599583268	0.0075878	0.0960651
NM_004117	FKBP5	0.600816686	0.0066917	0.0921147
NM_000855	GUCY1A2	0.60113068	0.0110279	0.1058683
NM_004154	P2RY6	0.601188399	0.0498261	0.1707263
NM_000351	STS	0.601222343	0.0092447	0.0994753
NM_198951	TGM2	0.601758521	0.0009684	0.0685309
NM_002021	FMO1	0.602065855	0.0229544	0.1310323
NM_144581	C14orf149	0.602300115	0.0116555	0.1061689
NM_000320	QDPR	0.602997311	0.0083105	0.0969856
NM_003161	RPS6KB1	0.603063439	0.0273024	0.1390307
NM_003918	GYG2	0.603714359	0.011583	0.1062095
NM_014256	B3GNT3	0.604006836	0.0011821	0.0686863
NM_001056	SULT1C1	0.604916863	0.0017337	0.0797486
NM_022098	LOC63929	0.605243397	0.0055014	0.0867646
NM_013358	PADI1	0.606979719	0.0090657	0.0990943
NM_000741	CHRM4	0.607206284	0.0405204	0.1596521
NM_181578	RFC5	0.607822106	0.0015698	0.0780657
NM_032582	USP32	0.608711231	0.0489132	0.1692796

Dataset 3.

Accession No	Gene Symbol	Avg_POC_Ratio	P-Value	FDR
NM_018324	THEDC1	0.429536424	0.0209992	0.1273799
NM_004666	VNN1	0.429964133	0.0240549	0.1341246
NM_080911	UNG	0.445720958	0.0154252	0.1129269
XM_165973	USP24	0.454884119	0.0314827	0.1476503
NM_021970	MAP2K1IP1	0.468591817	0.0403303	0.1595864
NM_002933	RNASE1	0.469939834	0.0133807	0.109913
NM_153189	SPAM1	0.486478025	0.0415279	0.1605281
NM_138340	ABHD3	0.486532905	0.0067143	0.0919669
NM_013251	TAC3	0.491722626	0.0279567	0.1404196
NM_006327	TIMM23	0.501591174	0.0036578	0.0810883
NM_003841	TNFRSF10C	0.512219543	0.0140413	0.1100968
NM_172236	POFUT1	0.521360956	0.0160409	0.1146967
NM_003799	RNMT	0.521827669	0.0118005	0.106785
NM_003839	TNFRSF11A	0.521925008	0.0147432	0.111789
NM_001025243	IRAK1	0.524453336	0.0053168	0.0860671

NM_005718	ARPC4	0.526481464	0.0136158	0.1102038
NM_002309	LIF	0.53479142	0.0026576	0.081955
NM_173353	TPH2	0.535499705	0.0315239	0.1475927
NM_012125	CHRM5	0.535566907	0.0201234	0.1250917
NM_012360	OR1F1	0.540526776	0.0168029	0.1166694
NM_003211	TDG	0.541857164	0.0034462	0.0806063
NM_130767	CACH-1	0.544235702	0.0455433	0.1651768
NM_005762	TRIM28	0.545059748	0.028319	0.1409565
NM_001876	CPT1A	0.54638003	0.0023785	0.0805481
NM_033181	CNR1	0.548978297	0.0084768	0.0972802
NM_017686	GDAP2	0.549407967	0.0317187	0.1478776
NM_145910	NEK11	0.551522047	0.0006802	0.0682706
NM_014465	SULT1B1	0.552146948	0.0173686	0.118951
NM_012387	PADI4	0.552576488	0.0187445	0.1220159
NM_001702	BAI1	0.556730459	7.658E-05	0.0528393
NM_199040	NUDT4	0.557359244	0.0072567	0.0949217
NM_024608	NEIL1	0.559400704	0.0324636	0.1483438
NM_003227	TFR2	0.559841619	0.0059989	0.0878351
NM_012214	MGAT4A	0.560188918	0.0103896	0.1038964
NM_001184	ATR	0.5610753	0.0363013	0.1534327
NM_003268	TLR5	0.563577132	0.0043176	0.085424
NM_003071	SMARCA3	0.564168075	0.0291905	0.1438677
NM_002934	RNASE2	0.564991995	0.001158	0.0718247
NM_002572	PAFAH1B2	0.565111742	0.0373269	0.1552708
NM_001866	COX7B	0.565778198	0.0222086	0.1302777
NM_003272	TM7SF1	0.569465132	0.0455148	0.1652907
NM_005508	CCR4	0.569924369	0.0006306	0.0740637
NM_006714	SMPDL3A	0.571885333	0.0153518	0.1132911
NM_153446	GALGT2	0.57287254	0.0011754	0.0705242
NM_001024592	UNG2	0.573614706	5.972E-05	0.0470927
NM_001002019	PUS1	0.574720416	0.0017126	0.0807981
NM_024342	GRLF1	0.575297825	0.0034254	0.080804
NM_000814	GABRB3	0.575759866	0.0415214	0.1606154
NM_000536	RAG2	0.575804467	0.0125883	0.1080676
NM_007255	B4GALT7	0.575961939	0.0005847	0.0733561
NM_177552	SULT1A3	0.577010219	0.000668	0.0695751
NM_003650	CST7	0.579687602	0.0174241	0.1190359
NM_023018	FLJ13052	0.581448317	0.0020006	0.0824126
NM_000961	PTGIS	0.58160867	0.0199218	0.1245393
NM_000150	FUT6	0.582777658	0.0218182	0.129641
NM_001775	CD38	0.583916315	0.0497526	0.1706863

NM_170706	NMNAT2	0.586556927	0.015877	0.114116
NM_014058	DESC1	0.586739966	0.0331931	0.1496942
NM_000926	PGR	0.587029255	9.715E-05	0.0446905
NM_138578	BCL2L1	0.587481551	0.0149647	0.1123878
NM_145207	SPATA5	0.588113844	0.0077668	0.0959127
NM_004566	PFKFB3	0.588865827	0.0136729	0.1098611
NM_004443	EPHB3	0.589677553	0.0187482	0.1218965
NM_002661	PLCG2	0.590020743	0.035635	0.152959
NM_006894	FMO3	0.590684268	0.0300392	0.145071
NM_001667	ARL2	0.590825053	0.0015146	0.0760068
NM_032236	USP48	0.590977992	0.0436502	0.1640225
NM_018010	ESRRBL1	0.596206886	0.0084695	0.0973997
NM_007171	POMT1	0.596482185	0.0442913	0.1639757
NM_130806	LGR8	0.596729273	0.0263943	0.1373202
NM_019092	KIAA1164	0.597090987	0.014906	0.1125597
NM_002754	MAPK13	0.597218471	0.0029406	0.0780399
NM_032595	PPP1R9B	0.597285454	0.019346	0.1237426
NM_001003962	CAPNS1	0.597680159	0.0254831	0.1364369
NM_174971	SIAT6	0.59826402	0.0461521	0.1657512
NM_000234	LIG1	0.599034365	0.0003014	0.063993
NM_003137	SRPK1	0.599583268	0.0075878	0.0960651
NM_004117	FKBP5	0.600816686	0.0066917	0.0921147
NM_000855	GUCY1A2	0.60113068	0.0110279	0.1058683
NM_004154	P2RY6	0.601188399	0.0498261	0.1707263
NM_000351	STS	0.601222343	0.0092447	0.0994753
NM_198951	TGM2	0.601758521	0.0009684	0.0685309
NM_002021	FMO1	0.602065855	0.0229544	0.1310323
NM_144581	C14orf149	0.602300115	0.0116555	0.1061689
NM_000320	QDPR	0.602997311	0.0083105	0.0969856
NM_003161	RPS6KB1	0.603063439	0.0273024	0.1390307
NM_003918	GYG2	0.603714359	0.011583	0.1062095
NM_014256	B3GNT3	0.604006836	0.0011821	0.0686863
NM_001056	SULT1C1	0.604916863	0.0017337	0.0797486
NM_022098	LOC63929	0.605243397	0.0055014	0.0867646
NM_013358	PADI1	0.606979719	0.0090657	0.0990943
NM_000741	CHRM4	0.607206284	0.0405204	0.1596521
NM_181578	RFC5	0.607822106	0.0015698	0.0780657
NM_032582	USP32	0.608711231	0.0489132	0.1692796
XM_371285	LOC128102	0.609257966	0.0073032	0.0950794
NM_130384	TREX1	0.609335989	0.0464234	0.1660772
NM_018413	CHST11	0.609436442	0.002225	0.0802728

NM_016616	TXNDC3	0.60962619	0.0115702	0.1062685
NM_002663	PLD2	0.610103282	0.0038365	0.0820837
NM_147173	NUDT2	0.61043487	0.030907	0.1470748
NM_005518	HMGCS2	0.611468472	0.0268483	0.13812
NM_020378	KLP1	0.612693702	0.0017788	0.0804841
NM_003390	WEE1	0.613803508	0.0118051	0.1066516
NM_138608	MPPE1	0.613884159	0.0324861	0.1483237
NM_000353	TAT	0.61573104	0.0020768	0.080167
NM_153002	GPR156	0.616499941	0.0403791	0.1594369
NM_004388	CTBS	0.616778083	0.0065255	0.0905043
NM_000199	SGSH	0.618577267	0.0091277	0.0995745
NM_153809	TAF1L	0.619586798	0.0494498	0.1699644
NM_033294	CASP1	0.620480218	0.0035705	0.0807748
NM_182547	HNLF	0.621360183	0.0044515	0.0856168
NM_003986	BBOX1	0.622305367	0.0239427	0.1340401
NM_025227	BPIL1	0.625613404	0.0049732	0.0852548
NM_007041	ATE1	0.626081973	0.0115264	0.1062202
NM_016347	CML2	0.626721943	0.0111975	0.1060213
NM_005309	GPT	0.626783436	0.0011814	0.069373
NM_005173	ATP2A3	0.627154941	0.0140974	0.1100674
NM_020960	GPR107	0.627281238	0.0033474	0.0796453
NM_197941	ADAMTS6	0.627476229	0.0445869	0.1639703
NM_183415	UBE3B	0.62761405	0.0092349	0.0995638
NM_172234	IL17RB	0.629036343	0.0372194	0.1549404
NM_002014	FKBP4	0.629862888	0.0003133	0.0617624
NM_004054	C3AR1	0.630710213	0.0154231	0.1130619
NM_138327	TRAR1	0.631346468	0.0054855	0.0867625
NM_014275	MGAT4B	0.631640548	0.0023032	0.0804675
NM_000875	IGF1R	0.633037344	0.0121453	0.1072677
NM_001355	DDT	0.633879169	0.0124526	0.108249
NM_001445	FABP6	0.634237146	0.0336435	0.1508629
NM_012460	TIMM9	0.635270062	0.004009	0.0831943
NM_170705	ICMT	0.635701849	0.003514	0.0801539
NM_005373	MPL	0.636111958	0.0293957	0.1438515
NM_005605	PPP3CC	0.636322741	0.0003141	0.0597816
NM_001683	ATP2B2	0.637090269	0.0351494	0.1524154
NM_001008505	OPRM1	0.637171567	0.0042776	0.0852424
NM_001005619	ITGB4	0.637353948	0.0133707	0.1101584
NM_007180	TREH	0.637408937	0.000708	0.0685671
NM_013382	POMT2	0.637409881	0.0389863	0.1574281
NM_020438	DOLPP1	0.638044755	0.0426131	0.1616661

NM_022120	OXCT2	0.638353099	0.0071393	0.09428
NM_006610	MASP2	0.638566567	0.0236021	0.1334871
NM_198066	GNPNAT1	0.639193548	0.0135816	0.1102507
NM_005471	GNPDA1	0.639644785	0.0373323	0.1551763
NM_014911	AAK1	0.640408077	0.0452324	0.1649161
NM_002103	GYS1	0.640761325	0.0010296	0.0710457
NM_024681	FLJ12242	0.641444177	0.0355712	0.1530422
NM_002312	LIG4	0.641576109	0.0122672	0.1074837
NM_001679	ATP1B3	0.642017327	0.007819	0.0959135
NM_181468	ITGB4BP	0.642270855	0.0270808	0.1384132
NM_022341	PDF	0.642691534	0.0395572	0.1583435
NM_003654	CHST1	0.644429746	0.0044646	0.0855709
NM_153699	GSTA5	0.644761056	0.0044309	0.0858199
NM_000797	DRD4	0.644792449	0.0358351	0.1529851
NM_000481	AMT	0.645664635	0.0008662	0.0654976
NM_054014	FKBP1A	0.64576482	0.0003291	0.0567634
NM_003595	TPST2	0.645769416	0.0063442	0.0897951
NM_000740	CHRM3	0.645843857	0.0076131	0.0959463
NM_177524	MEST	0.645866015	0.0428969	0.1622965
NM_002676	PMM1	0.64604996	0.0002546	0.0610985
NM_004624	VIPR1	0.646071697	0.0051226	0.0854283
NM_198974	PTK9	0.646260808	0.001041	0.0709427
NM_002691	POLD1	0.646310661	0.0396336	0.1584197
NM_000750	CHRNB4	0.646724844	0.0410758	0.160466
NM_002630	PGC	0.646801181	0.0429236	0.1622865
NM_018060	FLJ10326	0.64683082	0.0338063	0.1508573
NM_014413	HRI	0.648032791	0.0048498	0.0858047
NM_002542	OGG1	0.64860652	0.0020099	0.0821818
NM_178221	APG4C	0.649558894	0.0110738	0.1059404
NM_197975	BTNL3	0.649616816	0.0130022	0.1089109
NM_002410	MGAT5	0.650583395	0.0077468	0.0960949
NM_022481	ARAP3	0.650767829	0.0156172	0.1137292
NM_001034	RRM2	0.651065915	0.0416831	0.1606777
NM_152942	TNFRSF8	0.651437898	0.0193805	0.1238198
NM_000928	PLA2G1B	0.651517841	0.036521	0.1540076
NM_002529	NTRK1	0.651678416	0.0449509	0.1647603
NM_006039	MRC2	0.652198766	0.0050108	0.0851064
NM_199283	LOC220686	0.653181345	0.0464494	0.165955
NM_198584	CA13	0.653337445	0.0388229	0.157113
NM_000476	AK1	0.654123518	0.0056323	0.0878264
NM_017784	OSBPL10	0.65425603	8.99E-05	0.0496239

NM_012330	MYST4	0.654294965	0.0403303	0.1594722
NM_005299	GPR31	0.656339085	0.0295267	0.1441092
NM_173490	LOC134285	0.656569671	0.0023369	0.0811287
NM_001004056	GRK4	0.656952826	0.0349507	0.1523917
NM_052937	LOC115294	0.657658683	0.019758	0.1242188
NM_152910	DGKH	0.657893406	0.0110532	0.1059266
NM_004129	GUCY1B2	0.658032167	0.0028169	0.0805659
NM_002501	NFIX	0.658226633	0.0034335	0.0806513
NM_002736	PRKAR2B	0.658436006	0.0373593	0.1550553
NM_173200	NR4A3	0.659488791	0.0120367	0.107512
NM_016341	PLCE1	0.659502531	0.0123266	0.1076626
NM_138424	KIF12	0.660339152	0.0380171	0.1561415
NM_014893	NLGN4Y	0.66085573	0.0053907	0.0867548
NM_017900	AKIP	0.661031432	0.0207875	0.1273549
NM_004979	KCND1	0.6619906	0.0375173	0.1548956
NM_146421	GSTM1	0.662236555	0.0063191	0.0899012
NM_002601	PDE6D	0.662351041	0.0328282	0.1487781
NM_004309	ARHGDIA	0.662460906	0.0106407	0.1035923
NM_020126	SPHK2	0.66257518	0.016916	0.117307
NM_020299	AKR1B10	0.662863804	0.0442284	0.164073
NM_002633	PGM1	0.662880128	0.0031137	0.0781268
NM_032049	AGTR1	0.663180261	0.0139065	0.1099772
NM_018339	RFK	0.663441715	0.0209674	0.1273271
NM_003129	SQLE	0.663545089	0.0059708	0.0878907
NM_012413	QPCT	0.663632443	0.0312606	0.1479917
NM_017827	SARS2	0.663705726	0.0075042	0.095445
NM_001248	ENTPD3	0.663725011	0.0301049	0.1451347
NM_006849	PDIP	0.663977507	0.0138641	0.1104326
NM_183386	PTE1	0.664435297	0.0187156	0.1219716
NM_000738	CHRM1	0.664979911	0.0237905	0.1340035
XM_070277	OC90	0.665013566	0.0301148	0.1450555
NM_021615	CHST6	0.665560402	0.0108472	0.1046789
NM_001461	FMO5	0.666405064	0.0049815	0.0848696
NM_004441	EPHB1	0.66685553	0.0252007	0.1361132
NM_006182	DDR2	0.666943324	0.022541	0.1309749
NM_183239	GSTO2	0.667527337	0.0120436	0.1073998
NM_000024	ADRB2	0.667755257	0.0190932	0.1231245
NM_032294	CAMKK1	0.667877264	0.0083602	0.0969498
NM_005135	SLC12A6	0.66817077	0.0039956	0.0832287
NM_003313	TSTA3	0.668746602	0.0251078	0.1362783
NM_000936	PNLIP	0.669420476	0.0178941	0.1198728

NM_145042	MGC16703	0.669503075	0.0398338	0.1589897
NM_198400	NEDD4	0.669910495	0.0210623	0.1273428
NM_014725	STARD8	0.670114672	0.0135495	0.1103149
NM_012474	UCK2	0.670168468	0.0143033	0.1102709
NM_001270	CHD1	0.670232554	0.0295748	0.1439619
NM_005107	ENDOGL1	0.670276508	0.030445	0.1461359
NM_019886	CHST7	0.670424203	0.0061752	0.089703
NM_033266	ERN2	0.6704522	0.0133505	0.1103218
NM_181573	RFC4	0.670500351	0.001226	0.0683604
NM_138335	GNPDA2	0.671617032	0.008562	0.0978517
NM_000486	AQP2	0.671975352	0.015375	0.113311
NM_014064	AD-003	0.672678758	0.0062567	0.0894747
NM_145798	OSBPL7	0.672814525	0.012691	0.1081088
NM_198319	HRMT1L2	0.672839254	0.0009354	0.0670598
NM_001911	CTSG	0.673144369	0.0428208	0.16212
NM_024917	CXorf34	0.673284608	0.0094973	0.1004317
NM_003052	SLC34A1	0.674121665	0.0120129	0.1078229
NM_006570	RRAGA	0.674154049	0.0402031	0.1596554

Dataset 4.

Accession No	Gene Symbol	Avg_POC_Ratio	P-Value	FDR
NM_018324	THEDC1	0.429536424	0.0209992	0.1273799
NM_004666	VNN1	0.429964133	0.0240549	0.1341246
NM_080911	UNG	0.445720958	0.0154252	0.1129269
XM_165973	USP24	0.454884119	0.0314827	0.1476503
NM_021970	MAP2K1IP1	0.468591817	0.0403303	0.1595864
NM_002933	RNASE1	0.469939834	0.0133807	0.109913
NM_153189	SPAM1	0.486478025	0.0415279	0.1605281
NM_138340	ABHD3	0.486532905	0.0067143	0.0919669
NM_013251	TAC3	0.491722626	0.0279567	0.1404196
NM_006327	TIMM23	0.501591174	0.0036578	0.0810883
NM_003841	TNFRSF10C	0.512219543	0.0140413	0.1100968
NM_172236	POFUT1	0.521360956	0.0160409	0.1146967
NM_003799	RNMT	0.521827669	0.0118005	0.106785
NM_003839	TNFRSF11A	0.521925008	0.0147432	0.111789
NM_001025243	IRAK1	0.524453336	0.0053168	0.0860671
NM_005718	ARPC4	0.526481464	0.0136158	0.1102038
NM_002309	LIF	0.53479142	0.0026576	0.081955
NM_173353	TPH2	0.535499705	0.0315239	0.1475927

NM_012125	CHRM5	0.535566907	0.0201234	0.1250917
NM_012360	OR1F1	0.540526776	0.0168029	0.1166694
NM_003211	TDG	0.541857164	0.0034462	0.0806063
NM_130767	CACH-1	0.544235702	0.0455433	0.1651768
NM_005762	TRIM28	0.545059748	0.028319	0.1409565
NM_001876	CPT1A	0.54638003	0.0023785	0.0805481
NM_033181	CNR1	0.548978297	0.0084768	0.0972802
NM_017686	GDAP2	0.549407967	0.0317187	0.1478776
NM_145910	NEK11	0.551522047	0.0006802	0.0682706
NM_014465	SULT1B1	0.552146948	0.0173686	0.118951
NM_012387	PADI4	0.552576488	0.0187445	0.1220159
NM_001702	BAI1	0.556730459	7.658E-05	0.0528393
NM_199040	NUDT4	0.557359244	0.0072567	0.0949217
NM_024608	NEIL1	0.559400704	0.0324636	0.1483438
NM_003227	TFR2	0.559841619	0.0059989	0.0878351
NM_012214	MGAT4A	0.560188918	0.0103896	0.1038964
NM_001184	ATR	0.5610753	0.0363013	0.1534327
NM_003268	TLR5	0.563577132	0.0043176	0.085424
NM_003071	SMARCA3	0.564168075	0.0291905	0.1438677
NM_002934	RNASE2	0.564991995	0.001158	0.0718247
NM_002572	PAFAH1B2	0.565111742	0.0373269	0.1552708
NM_001866	COX7B	0.565778198	0.0222086	0.1302777
NM_003272	TM7SF1	0.569465132	0.0455148	0.1652907
NM_005508	CCR4	0.569924369	0.0006306	0.0740637
NM_006714	SMPDL3A	0.571885333	0.0153518	0.1132911
NM_153446	GALGT2	0.57287254	0.0011754	0.0705242
NM_001024592	UNG2	0.573614706	5.972E-05	0.0470927
NM_001002019	PUS1	0.574720416	0.0017126	0.0807981
NM_024342	GRLF1	0.575297825	0.0034254	0.080804
NM_000814	GABRB3	0.575759866	0.0415214	0.1606154
NM_000536	RAG2	0.575804467	0.0125883	0.1080676
NM_007255	B4GALT7	0.575961939	0.0005847	0.0733561
NM_177552	SULT1A3	0.577010219	0.000668	0.0695751
NM_003650	CST7	0.579687602	0.0174241	0.1190359
NM_023018	FLJ13052	0.581448317	0.0020006	0.0824126
NM_000961	PTGIS	0.58160867	0.0199218	0.1245393
NM_000150	FUT6	0.582777658	0.0218182	0.129641
NM_001775	CD38	0.583916315	0.0497526	0.1706863
NM_170706	NMNAT2	0.586556927	0.015877	0.114116
NM_014058	DESC1	0.586739966	0.0331931	0.1496942
NM_000926	PGR	0.587029255	9.715E-05	0.0446905

NM_138578	BCL2L1	0.587481551	0.0149647	0.1123878
NM_145207	SPATA5	0.588113844	0.0077668	0.0959127
NM_004566	PFKFB3	0.588865827	0.0136729	0.1098611
NM_004443	EPHB3	0.589677553	0.0187482	0.1218965
NM_002661	PLCG2	0.590020743	0.035635	0.152959
NM_006894	FMO3	0.590684268	0.0300392	0.145071
NM_001667	ARL2	0.590825053	0.0015146	0.0760068
NM_032236	USP48	0.590977992	0.0436502	0.1640225
NM_018010	ESRRBL1	0.596206886	0.0084695	0.0973997
NM_007171	POMT1	0.596482185	0.0442913	0.1639757
NM_130806	LGR8	0.596729273	0.0263943	0.1373202
NM_019092	KIAA1164	0.597090987	0.014906	0.1125597
NM_002754	MAPK13	0.597218471	0.0029406	0.0780399
NM_032595	PPP1R9B	0.597285454	0.019346	0.1237426
NM_001003962	CAPNS1	0.597680159	0.0254831	0.1364369
NM_174971	SIAT6	0.59826402	0.0461521	0.1657512
NM_000234	LIG1	0.599034365	0.0003014	0.063993
NM_003137	SRPK1	0.599583268	0.0075878	0.0960651
NM_004117	FKBP5	0.600816686	0.0066917	0.0921147
NM_000855	GUCY1A2	0.60113068	0.0110279	0.1058683
NM_004154	P2RY6	0.601188399	0.0498261	0.1707263
NM_000351	STS	0.601222343	0.0092447	0.0994753
NM_198951	TGM2	0.601758521	0.0009684	0.0685309
NM_002021	FMO1	0.602065855	0.0229544	0.1310323
NM_144581	C14orf149	0.602300115	0.0116555	0.1061689
NM_000320	QDPR	0.602997311	0.0083105	0.0969856
NM_003161	RPS6KB1	0.603063439	0.0273024	0.1390307
NM_003918	GYG2	0.603714359	0.011583	0.1062095
NM_014256	B3GNT3	0.604006836	0.0011821	0.0686863
NM_001056	SULT1C1	0.604916863	0.0017337	0.0797486
NM_022098	LOC63929	0.605243397	0.0055014	0.0867646
NM_013358	PADI1	0.606979719	0.0090657	0.0990943
NM_000741	CHRM4	0.607206284	0.0405204	0.1596521
NM_181578	RFC5	0.607822106	0.0015698	0.0780657
NM_032582	USP32	0.608711231	0.0489132	0.1692796
XM_371285	LOC128102	0.609257966	0.0073032	0.0950794
NM_130384	TREX1	0.609335989	0.0464234	0.1660772
NM_018413	CHST11	0.609436442	0.002225	0.0802728
NM_016616	TXNDC3	0.60962619	0.0115702	0.1062685
NM_002663	PLD2	0.610103282	0.0038365	0.0820837
NM_147173	NUDT2	0.61043487	0.030907	0.1470748

NM_005518	HMGCS2	0.611468472	0.0268483	0.13812
NM_020378	KLP1	0.612693702	0.0017788	0.0804841
NM_003390	WEE1	0.613803508	0.0118051	0.1066516
NM_138608	MPPE1	0.613884159	0.0324861	0.1483237
NM_000353	TAT	0.61573104	0.0020768	0.080167
NM_153002	GPR156	0.616499941	0.0403791	0.1594369
NM_004388	CTBS	0.616778083	0.0065255	0.0905043
NM_000199	SGSH	0.618577267	0.0091277	0.0995745
NM_153809	TAF1L	0.619586798	0.0494498	0.1699644
NM_033294	CASP1	0.620480218	0.0035705	0.0807748
NM_182547	HNLF	0.621360183	0.0044515	0.0856168
NM_003986	BBOX1	0.622305367	0.0239427	0.1340401
NM_025227	BPIL1	0.625613404	0.0049732	0.0852548
NM_007041	ATE1	0.626081973	0.0115264	0.1062202
NM_016347	CML2	0.626721943	0.0111975	0.1060213
NM_005309	GPT	0.626783436	0.0011814	0.069373
NM_005173	ATP2A3	0.627154941	0.0140974	0.1100674
NM_020960	GPR107	0.627281238	0.0033474	0.0796453
NM_197941	ADAMTS6	0.627476229	0.0445869	0.1639703
NM_183415	UBE3B	0.62761405	0.0092349	0.0995638
NM_172234	IL17RB	0.629036343	0.0372194	0.1549404
NM_002014	FKBP4	0.629862888	0.0003133	0.0617624
NM_004054	C3AR1	0.630710213	0.0154231	0.1130619
NM_138327	TRAR1	0.631346468	0.0054855	0.0867625
NM_014275	MGAT4B	0.631640548	0.0023032	0.0804675
NM_000875	IGF1R	0.633037344	0.0121453	0.1072677
NM_001355	DDT	0.633879169	0.0124526	0.108249
NM_001445	FABP6	0.634237146	0.0336435	0.1508629
NM_012460	TIMM9	0.635270062	0.004009	0.0831943
NM_170705	ICMT	0.635701849	0.003514	0.0801539
NM_005373	MPL	0.636111958	0.0293957	0.1438515
NM_005605	PPP3CC	0.636322741	0.0003141	0.0597816
NM_001683	ATP2B2	0.637090269	0.0351494	0.1524154
NM_001008505	OPRM1	0.637171567	0.0042776	0.0852424
NM_001005619	ITGB4	0.637353948	0.0133707	0.1101584
NM_007180	TREH	0.637408937	0.000708	0.0685671
NM_013382	POMT2	0.637409881	0.0389863	0.1574281
NM_020438	DOLPP1	0.638044755	0.0426131	0.1616661
NM_022120	OXCT2	0.638353099	0.0071393	0.09428
NM_006610	MASP2	0.638566567	0.0236021	0.1334871
NM_198066	GNPNAT1	0.639193548	0.0135816	0.1102507

NM_005471	GNPDA1	0.639644785	0.0373323	0.1551763
NM_014911	AAK1	0.640408077	0.0452324	0.1649161
NM_002103	GYS1	0.640761325	0.0010296	0.0710457
NM_024681	FLJ12242	0.641444177	0.0355712	0.1530422
NM_002312	LIG4	0.641576109	0.0122672	0.1074837
NM_001679	ATP1B3	0.642017327	0.007819	0.0959135
NM_181468	ITGB4BP	0.642270855	0.0270808	0.1384132
NM_022341	PDF	0.642691534	0.0395572	0.1583435
NM_003654	CHST1	0.644429746	0.0044646	0.0855709
NM_153699	GSTA5	0.644761056	0.0044309	0.0858199
NM_000797	DRD4	0.644792449	0.0358351	0.1529851
NM_000481	AMT	0.645664635	0.0008662	0.0654976
NM_054014	FKBP1A	0.64576482	0.0003291	0.0567634
NM_003595	TPST2	0.645769416	0.0063442	0.0897951
NM_000740	CHRM3	0.645843857	0.0076131	0.0959463
NM_177524	MEST	0.645866015	0.0428969	0.1622965
NM_002676	PMM1	0.64604996	0.0002546	0.0610985
NM_004624	VIPR1	0.646071697	0.0051226	0.0854283
NM_198974	PTK9	0.646260808	0.001041	0.0709427
NM_002691	POLD1	0.646310661	0.0396336	0.1584197
NM_000750	CHRNB4	0.646724844	0.0410758	0.160466
NM_002630	PGC	0.646801181	0.0429236	0.1622865
NM_018060	FLJ10326	0.64683082	0.0338063	0.1508573
NM_014413	HRI	0.648032791	0.0048498	0.0858047
NM_002542	OGG1	0.64860652	0.0020099	0.0821818
NM_178221	APG4C	0.649558894	0.0110738	0.1059404
NM_197975	BTNL3	0.649616816	0.0130022	0.1089109
NM_002410	MGAT5	0.650583395	0.0077468	0.0960949
NM_022481	ARAP3	0.650767829	0.0156172	0.1137292
NM_001034	RRM2	0.651065915	0.0416831	0.1606777
NM_152942	TNFRSF8	0.651437898	0.0193805	0.1238198
NM_000928	PLA2G1B	0.651517841	0.036521	0.1540076
NM_002529	NTRK1	0.651678416	0.0449509	0.1647603
NM_006039	MRC2	0.652198766	0.0050108	0.0851064
NM_199283	LOC220686	0.653181345	0.0464494	0.165955
NM_198584	CA13	0.653337445	0.0388229	0.157113
NM_000476	AK1	0.654123518	0.0056323	0.0878264
NM_017784	OSBPL10	0.65425603	8.99E-05	0.0496239
NM_012330	MYST4	0.654294965	0.0403303	0.1594722
NM_005299	GPR31	0.656339085	0.0295267	0.1441092
NM_173490	LOC134285	0.656569671	0.0023369	0.0811287

NM_001004056	GRK4	0.656952826	0.0349507	0.1523917
NM_052937	LOC115294	0.657658683	0.019758	0.1242188
NM_152910	DGKH	0.657893406	0.0110532	0.1059266
NM_004129	GUCY1B2	0.658032167	0.0028169	0.0805659
NM_002501	NFIX	0.658226633	0.0034335	0.0806513
NM_002736	PRKAR2B	0.658436006	0.0373593	0.1550553
NM_173200	NR4A3	0.659488791	0.0120367	0.107512
NM_016341	PLCE1	0.659502531	0.0123266	0.1076626
NM_138424	KIF12	0.660339152	0.0380171	0.1561415
NM_014893	NLGN4Y	0.66085573	0.0053907	0.0867548
NM_017900	AKIP	0.661031432	0.0207875	0.1273549
NM_004979	KCND1	0.6619906	0.0375173	0.1548956
NM_146421	GSTM1	0.662236555	0.0063191	0.0899012
NM_002601	PDE6D	0.662351041	0.0328282	0.1487781
NM_004309	ARHGDIA	0.662460906	0.0106407	0.1035923
NM_020126	SPHK2	0.66257518	0.016916	0.117307
NM_020299	AKR1B10	0.662863804	0.0442284	0.164073
NM_002633	PGM1	0.662880128	0.0031137	0.0781268
NM_032049	AGTR1	0.663180261	0.0139065	0.1099772
NM_018339	RFK	0.663441715	0.0209674	0.1273271
NM_003129	SQLE	0.663545089	0.0059708	0.0878907
NM_012413	QPCT	0.663632443	0.0312606	0.1479917
NM_017827	SARS2	0.663705726	0.0075042	0.095445
NM_001248	ENTPD3	0.663725011	0.0301049	0.1451347
NM_006849	PDIP	0.663977507	0.0138641	0.1104326
NM_183386	PTE1	0.664435297	0.0187156	0.1219716
NM_000738	CHRM1	0.664979911	0.0237905	0.1340035
XM_070277	OC90	0.665013566	0.0301148	0.1450555
NM_021615	CHST6	0.665560402	0.0108472	0.1046789
NM_001461	FMO5	0.666405064	0.0049815	0.0848696
NM_004441	EPHB1	0.66685553	0.0252007	0.1361132
NM_006182	DDR2	0.666943324	0.022541	0.1309749
NM_183239	GSTO2	0.667527337	0.0120436	0.1073998
NM_000024	ADRB2	0.667755257	0.0190932	0.1231245
NM_032294	CAMKK1	0.667877264	0.0083602	0.0969498
NM_005135	SLC12A6	0.66817077	0.0039956	0.0832287
NM_003313	TSTA3	0.668746602	0.0251078	0.1362783
NM_000936	PNLIP	0.669420476	0.0178941	0.1198728
NM_145042	MGC16703	0.669503075	0.0398338	0.1589897
NM_198400	NEDD4	0.669910495	0.0210623	0.1273428
NM_014725	STARD8	0.670114672	0.0135495	0.1103149

NM_012474	UCK2	0.670168468	0.0143033	0.1102709
NM_001270	CHD1	0.670232554	0.0295748	0.1439619
NM_005107	ENDOGL1	0.670276508	0.030445	0.1461359
NM_019886	CHST7	0.670424203	0.0061752	0.089703
NM_033266	ERN2	0.6704522	0.0133505	0.1103218
NM_181573	RFC4	0.670500351	0.001226	0.0683604
NM_138335	GNPDA2	0.671617032	0.008562	0.0978517
NM_000486	AQP2	0.671975352	0.015375	0.113311
NM_014064	AD-003	0.672678758	0.0062567	0.0894747
NM_145798	OSBPL7	0.672814525	0.012691	0.1081088
NM_198319	HRMT1L2	0.672839254	0.0009354	0.0670598
NM_001911	CTSG	0.673144369	0.0428208	0.16212
NM_024917	CXorf34	0.673284608	0.0094973	0.1004317
NM_003052	SLC34A1	0.674121665	0.0120129	0.1078229
NM_006570	RRAGA	0.674154049	0.0402031	0.1596554
NM_001837	CCR3	0.674582808	0.0293203	0.144121
NM_000206	IL2RG	0.67501986	0.0002131	0.058805
NM_018290	PGM2	0.675094928	0.039837	0.1588876
NM_152889	CHST13	0.675142325	0.0253295	0.1361431
NM_003549	HYAL3	0.675143604	0.0103898	0.10371
NM_006528	TFPI2	0.675356169	0.0413905	0.1603335
NM_016591	C2GNT3	0.675675628	0.012223	0.1074379
NM_183393	CADPS	0.675737764	0.0175031	0.1189867
NM_001527	HDAC2	0.675783842	0.0295379	0.1440363
NM_147132	GALT	0.675967957	0.0319167	0.1480504
NM_007202	AKAP10	0.676405388	0.0203815	0.1261275
NM_000466	PEX1	0.676679603	0.0026693	0.0818578
NM_023915	GPR87	0.676777159	0.0442707	0.1640096
NM_032456	PCDH7	0.677443398	0.0044667	0.0853159
NM_001152	SLC25A5	0.677486857	0.0147472	0.1116658
NM_012397	SERPINB13	0.677552697	0.0436609	0.1639511
NM_001139	ALOX12B	0.677585878	0.0242424	0.1342206
NM_007181	MAP4K1	0.677873893	0.0103818	0.1040068
NM_003465	CHIT1	0.677923282	0.0027708	0.0800774
NM_002317	LOX	0.677985718	0.0057022	0.0879227
NM_004217	AURKB	0.678181929	0.0232251	0.1317602
NM_004897	MINPP1	0.67859448	0.0472034	0.1668136
NM_201636	TBXA2R	0.678785356	0.0474841	0.1671635
NM_001607	ACAA1	0.67989441	0.0110782	0.1057991
NM_001877	CR2	0.680159969	0.012545	0.1083698
NM_153343	ENPP6	0.680181749	0.0274232	0.1390046

NM_021939	FKBP10	0.680302848	0.0280914	0.1407121
NM_006169	NNMT	0.681028202	0.0020821	0.079812
NM_182644	EPHA3	0.681468171	0.0420049	0.1607956
NM_018641	CHST12	0.681800646	0.0027291	0.0801303
NM_002943	RORA	0.681805343	0.03523	0.1526451
NM_032846	RAB2B	0.682578515	0.0456172	0.1652278
NM_002185	IL7R	0.682596728	0.0142672	0.1106107
NM_145752	CDIPT	0.68323425	0.0023688	0.0807154
NM_000180	GUCY2D	0.683264171	0.0138959	0.1100504
NM_004535	MYT1	0.68346942	0.0016479	0.0790968
NM_004164	RBP2	0.683511985	0.0007462	0.0686511
NM_004863	SPTLC2	0.683802243	0.0058195	0.0875311
NM_001001928	PPARA	0.683814274	0.0022452	0.0799576
NM_007023	RAPGEF4	0.684168328	0.0365103	0.1541981
NM_004832	GSTO1	0.684232703	0.0008383	0.0661096
NM_173708	MTND1	0.684517616	0.0153399	0.1133554
NM_031934	RAB34	0.684691005	0.0204471	0.1262506
NM_016322	RAB14	0.684741587	0.0360728	0.1529355
NM_001396	DYRK1A	0.684898784	0.0172747	0.1186026
NM_014317	TPRT	0.685355855	0.0051061	0.0854109
NM_020474	GALNT1	0.685914779	0.0028542	0.0791725
NM_005296	GPR23	0.68604417	0.0350014	0.1524925
NM_018960	GNMT	0.686477977	0.0373668	0.1549698
NM_005026	PIK3CD	0.686686576	0.0238126	0.1339912
NM_001800	CDKN2D	0.686802251	0.0478851	0.1672949
NM_000848	GSTM2	0.68698218	0.000973	0.0679882
NM_018988	GFOD1	0.687326737	0.0006792	0.069427
NM_024619	FN3KRP	0.687522757	0.0307653	0.1472894
NM_012407	PRKCABP	0.687595722	0.0089856	0.0994001
XM_291266	OPLAH	0.687876265	0.0002016	0.0585618
NM_178155	FUT8	0.688311891	0.0069482	0.0930919
NM_000054	AVPR2	0.688355889	0.0041551	0.0837092
NM_006657	FTCD	0.688382932	0.0230971	0.131439
NM_017886	FLJ20574	0.688909325	0.0236008	0.1336167
NM_002029	FPR1	0.689110661	0.0296266	0.1439603
NM_003554	OR1E2	0.689288079	0.0333178	0.1501342
NM_175067	TRAR4	0.690127638	0.0169436	0.1172037
NM_005911	MAT2A	0.690322097	0.0011085	0.0746213
NM_021734	SLC25A19	0.690346362	0.0463213	0.1661428
NM_020683	ADORA3	0.690355948	0.0317327	0.1478183
NM_002310	LIFR	0.690460554	0.0249986	0.1359527

NM_002566	P2RY11	0.690653847	0.0155367	0.1132927
XM_291241	IPLA2(GAMMA)	0.690793543	0.0007564	0.0673409

Dataset 5.

Enriched genes identified by NIH DAVID and Princeton GO Term Finder

AAK1
ARPC4
ATG4C
B3GNT3
B4GALNT2
B4GALT7
BAI1
BTNL3
C3AR1
CAPNS1
CCNO
CENTD3
CHRM4
CHRM5
CHST1
CNR1
DDT
DOLPP1
EPHB3
FKBP4
FKBP5
FMO3
FUT6
GNPDA1
GNPNAT1
GPR156
GPT
GSTA5
HMGCS2
ICMT
IFT57
IL17RB
LIF

MAP2K1IP1
MGAT4A
MGAT5
NAT14
NAT8B
NEIL1
NEK11
NMNAT2
NUDT2
NUDT4
OPRM1
OR1F1
PADI1
PADI4
PLCG2
PMM1
POFUT1
POLD1
POMT2
PPP1R9B
PTGIS
PUS1
RAG2
RNASE2
RNMT
RRM2
SGSH
SMPDL3A
SPATA5
ST3GAL3
STS
SULT1B1
SULT1C2
TAF1L
TAT
TDG
TFR2
TIMM9
TMPRSS11E
TNFRSF11A
TREH

TREX1
TXNDC3
UBE3B
UNG
USP48
VNN1
WEE1

Dataset 6.

Gene Ontology term	Corrected P-value	FDR	Genes annotated to the term
protein amino acid glycosylation	1.41E-06	0.00%	FUT6, MGAT4A, GALGT2, POMT2, DOLPP1, B3GNT3, MGAT5, SIAT6, POFUT1
biopolymer glycosylation	1.41E-06	0.00%	FUT6, MGAT4A, GALGT2, POMT2, DOLPP1, B3GNT3, MGAT5, SIAT6, POFUT1
glycosylation	1.41E-06	0.00%	FUT6, MGAT4A, GALGT2, POMT2, DOLPP1, B3GNT3, MGAT5, SIAT6, POFUT1
glycoprotein biosynthetic process	2.40E-06	0.00%	FUT6, MGAT4A, GALGT2, POMT2, DOLPP1, B3GNT3, MGAT5, SIAT6, POFUT1
carbohydrate metabolic process	4.08E-06	0.00%	MGAT4A, GALGT2, GPT, MGAT5, POFUT1, GNPDA1, B4GALT7, CHST1, SGSH, FUT6, PMM1, POMT2, DOLPP1, TREH, B3GNT3, SIAT6
glycoprotein metabolic process	9.40E-06	0.00%	FUT6, MGAT4A, GALGT2, POMT2, DOLPP1, B3GNT3, MGAT5, SIAT6, POFUT1
metabolic process	0.00014	0.00%	GALGT2, PLCG2, MGAT5, UBE3B, EPHB3, WEE1, PMM1, POMT2, NEK11, FKBP4, B3GNT3, GSTA5, NUDT2, TXNDC3, GNPNAT1, ITGB4BP, POLD1, TDG, GNPDA1, STS, CHST1, SGSH, PTGIS, SMPDL3A, HMGCS2, UNG2, NUDT4, TREH, NEIL1, TREX1, NMNAT2, SULT1C1, CML2, ESRRBL1, RNMT, TAF1L, APG4C, POFUT1, DDT, DOLPP1, USP48, BTNL3, TFR2, ICMT, RAG2, MGAT4A, FKBP5, TAT, RNASE2, PUS1, GPT, PPP1R9B, SULT1B1, UNG, RRM2, MAP2K1IP1, AAK1, FMO3, B4GALT7, PADI4, DESC1, FUT6, VNN1, USP24, KLP1, LIG4, PADI1, SIAT6, LIF

cellular metabolic process	0.00023	0.00%	APG4C, TAF1L, GALGT2, PLCG2, MGAT5, UBE3B, EPHB3, WEE1, POFUT1, PMM1, POMT2, DDT, DOLPP1, NEK11, USP48, FKBP4, B3GNT3, NUDT2, TFR2, ICMT, TXNDC3, MGAT4A, RAG2, FKBP5, ITGB4BP, TAT, RNASE2, POLD1, GPT, PUS1, UNG, SULT1B1, PPP1R9B, RRM2, MAP2K1IP1, AAK1, TDG, GNPDA1, PADI4, B4GALT7, STS, CHST1, SGSH, PTGIS, DESC1, FUT6, USP24, VNN1, KLP1, SMPDL3A, LIG4, HMGCS2, PADI1, UNG2, NUDT4, TREH, NEIL1, TREX1, NMNAT2, SIAT6, SULT1C1, ESRRBL1, RNMT, LIF
primary metabolic process	0.00024	0.00%	APG4C, TAF1L, GALGT2, PLCG2, MGAT5, UBE3B, EPHB3, WEE1, POFUT1, PMM1, POMT2, DDT, DOLPP1, NEK11, USP48, FKBP4, B3GNT3, NUDT2, BTNL3, TFR2, ICMT, TXNDC3, MGAT4A, RAG2, FKBP5, ITGB4BP, TAT, RNASE2, POLD1, GPT, PUS1, UNG, SULT1B1, PPP1R9B, RRM2, MAP2K1IP1, AAK1, TDG, GNPDA1, PADI4, B4GALT7, STS, CHST1, SGSH, PTGIS, DESC1, FUT6, USP24, KLP1, SMPDL3A, LIG4, HMGCS2, PADI1, UNG2, NUDT4, TREH, NEIL1, TREX1, NMNAT2, SIAT6, ESRRBL1, RNMT, LIF
cellular process	0.00039	0.00%	GALGT2, PLCG2, MGAT5, UBE3B, EPHB3, WEE1, ARL2, GABRB3, PMM1, POMT2, NEK11, FKBP4, B3GNT3, NUDT2, P2RY6, TXNDC3, ITGB4BP, POLD1, TNFRSF11A, TDG, GNPDA1, STS, CHST1, OPRM1, SGSH, PTGIS, SMPDL3A, HMGCS2, UNG2, NUDT4, TREH, NEIL1, TREX1, SPATA5, NMNAT2, GPR156, CNR1, SULT1C1, CHRM4, ESRRBL1, ARPC4, RNMT, APG4C, TAF1L, TAC3, CAPNS1, BAI1, TIMM9, POFUT1, IL17RB, DDT, DOLPP1, USP48, TNFRSF10C, TFR2, ICMT, RAG2, MGAT4A, FKBP5, TAT, RNASE2, PUS1, GPT, SULT1B1, UNG, PPP1R9B, RRM2, MAP2K1IP1, TRAR1, AAK1, ARAP3, B4GALT7, PADI4, OR1F1, DESC1, FUT6, VNN1, USP24, C3AR1, KLP1, LIG4, CHRM5, PADI1, SIAT6, LGR8, LIF
cellular carbohydrate metabolic process	0.00081	0.00%	FUT6, PMM1, GALGT2, GPT, TREH, POFUT1, B4GALT7, GNPDA1, CHST1, SGSH
base-excision repair	0.00253	0.00%	NEIL1, TDG, UNG2, UNG

nitrogen compound metabolic process	0.00326	0.00%	VNN1, TAT, DDT, GPT, PADI1, SULT1B1, SULT1C1, PADI4, B4GALT7, CHST1, SGSH
cellular response to stimulus	0.00339	0.00%	APG4C, POLD1, LIG4, NEK11, PPP1R9B, UNG2, UNG, NEIL1, TREX1, TDG
biopolymer modification	0.00452	0.13%	MGAT4A, GALGT2, PUS1, MGAT5, UBE3B, EPHB3, MAP2K1IP1, AAK1, WEE1, POFUT1, B4GALT7, PADI4, FUT6, POMT2, DOLPP1, NEK11, USP48, PADI1, B3GNT3, SIAT6, ICMT, LIF
protein modification process	0.00864	0.12%	MGAT4A, GALGT2, MGAT5, UBE3B, EPHB3, MAP2K1IP1, AAK1, WEE1, POFUT1, PADI4, B4GALT7, FUT6, POMT2, DOLPP1, NEK11, USP48, PADI1, B3GNT3, SIAT6, ICMT, LIF

APPENDIX D. ALKYLATION SENSITIVITY SCREENS REVEAL A CONSERVED CROSS-SPRECIES FUNCTIONOME

Submitted to Molecular Cancer Research and compiled paper is below.

Title: **Alkylation sensitivity screens reveal a conserved cross-species functionome**

Authors
David Svilar[1,2,3], Madhu Dyavaiah[4], Ashley R. Brown[2], Jiang-bo Tang[2,5,7], Jianfeng Li[1,2], Peter R. McDonald[1,8], Tong Ying Shun[1], Xiao-hong Wang[2], John S. Lazo[1,2,9], Ian F. Pollack[2,6], Thomas J. Begley[4] and Robert W. Sobol[1,2,5]*

Affiliations
[1]Department of Pharmacology & Chemical Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15213.
[2]University of Pittsburgh Cancer Institute, Hillman Cancer Center, Pittsburgh, PA 15213.
[3]Medical Scientist Training Program, University of Pittsburgh/Carnegie Mellon University, Pittsburgh, PA 15261.
[4]Department of Nanobiosciences and Systems Toxicology Laboratory, College of Nanoscale Science and Engineering, University at Albany, Rensselaer, NY 12203.
[5]Department of Human Genetics, University of Pittsburgh Graduate School of Public Health, Pittsburgh, PA 15213.
[6]Department of Neurosurgery, Children's Hospital of Pittsburgh, University of Pittsburgh School of Medicine, Pittsburgh, PA 15224.

*Corresponding author: Robert W. Sobol
Hillman Cancer Center
University of Pittsburgh Cancer Institute
Research Pavilion, Suite 2.6a
5117 Centre Avenue
Pittsburgh, Pennsylvania 15213-1863
Office phone 412-623-7764
rws9@pitt.edu

[7]Current Address: Department of Cancer Biology, Abramson Family Cancer Research Institute, University of Pennsylvania School of Medicine, Philadelphia PA 19104.
[8]Current Address: High Throughput Screening Laboratory, University of Kansas, Lawrence KS 66047.
[9]Departments of Pharmacology & Chemistry, University of Virginia, School of Medicine, Charlottesville, VA 22908.

Running title: A cross-species alkylation functionome

Subject Categories: Metabolic and regulatory networks, Molecular Biology of Disease

Key words: Temozolomide/chemotherapy resistance/alkylation response/siRNA/DNA Repair

Abstract

To identify genes that contribute to chemotherapy resistance in glioblastoma, we conducted a synthetic lethal screen in a chemotherapy-resistant glioblastoma derived cell line with the clinical alkylator temozolomide (TMZ) and an siRNA library tailored towards “druggable” targets. We compared the human TMZ sensitizing genes identified in our screen with those identified from alkylator screens previously conducted in *E. coli* and *S. cerevisiae.* The conserved biological processes across all three species composes an *Alkylation Functionome* that includes many novel proteins not previously thought to impact alkylator resistance and our validation studies suggested additive or epistatic relationships between corresponding processes. The conserved processes of base excision repair (BER) and protein modification were dual targeted and yielded additive sensitization to alkylators in yeast. In contrast, dual targeting of BER and protein modification genes in human cells did not increase sensitivity, suggesting an epistatic relationship. Importantly, these studies provide potential new targets to overcome alkylating agent resistance and also provide potential mechanistic insight for the regulation of DNA repair and/or DNA damage response proteins.

Introduction

Glioblastoma multiforme (GBM) is the most common and aggressive primary brain tumor (1). Temozolomide (TMZ), radiation and surgery are currently used for treatment of GBM, yet median survival is still less than 2 years (2-4). Chemotherapy resistance and difficulties in surgical removal contribute to poor prognosis (1). TMZ methylates several bases in DNA, including the O^6 position of guanine (O^6-MeG), the N7 position of guanine (N7-MeG) and the N3 position of adenine (N3-MeA). Two of the lesions (N7-MeG and N3-MeA) account for greater than 80% of the DNA lesions induced by TMZ, yet result in negligible clinical toxicity due to robust repair mechanisms (5). These lesions are predominantly repaired by the base excision repair (BER) pathway (6), initiated by one of eleven lesion specific DNA glycosylases (7).

Virtually all of TMZ's clinical cytotoxicity is attributable to the O^6-methylguanine (O^6-MeG) lesion, which accounts for approximately 5% of TMZ induced lesions (5). The O^6-MeG lesion is repaired via a direct reversal mechanism by the protein O^6-methylguanine-DNA methyltransferase (MGMT), which transfers the O^6-methyl group from the guanine base onto a Cys residue in the MGMT protein (8). If the O^6-MeG lesion is not removed by MGMT, during cellular replication the mis-pairing of O^6-MeG with thymine is detected by the mismatch repair enzymes, triggering apoptosis signaling and cytotoxicity (9). However, 5-year survival rates still remain low in TMZ treated patients (3, 4), and TMZ resistance and/or recurrence with chemotherapy resistant tumors is common. Resistant cells can harbor mutations in mismatch repair proteins such as mutS homolog 6 (MSH6) (10) or have elevated expression of MGMT (as in the T98G cell line) (11). Earlier endeavors to enhance TMZ efficacy by using MGMT inhibitors to prevent the repair of O^6-MeG lesions have not shown an increase in sensitivity or efficacy in clinical trials (12), especially in TMZ-resistant GBM (13).

Because GBM tumors can be resistant to TMZ therapy and recurring tumors may acquire resistance to the O^6-MeG lesion (14, 15), we explored novel mechanisms of TMZ resistance in

tumor cells independent of the O^6-MeG lesion. To discover genes that sensitize cells to TMZ, we undertook a synthetic lethal screen using Ambion's Silencer® Human Druggable Genome siRNA library, the chemotherapeutic agent TMZ and the T98G cell line, a TMZ resistant cell line derived from a GBM tumor (16). We determined that several biological pathways were enriched for TMZ sensitizing genes, including proteins involved in BER, response to DNA damage, cellular proliferation, and protein modification. Many similar biological processes were also highly enriched in DNA alkylation screens performed in *S. cerevisiae* and *E. coli* (17, 18). We combined the selected hits from all three species and performed a functionome analysis to identify significantly enriched biological processes conserved across all three organisms (19). This allowed us to create a cross-species network representing the shared alkylation response from bacteria, yeast and humans and suggest that this network represents an *Alkylation Functionome* that includes many novel proteins not previously thought to impact alkylation resistance. Our identification of evolutionarily conserved mechanisms affecting TMZ sensitivity suggests that although many proteins and processes impact sensitivity to alkylators, several critical survival pathways can be targeted to improve chemotherapy efficacy. Simultaneous inhibition of both DNA repair and protein modification processes in yeast yielded greater sensitivity to alkylating agents than inhibition of either process alone. However, knockdown of both protein modification and DNA repair genes did not have an increased effect on alkylation-induced toxicity in human cells. This potential epistatic interaction between the ubiquitin protein ligase E3B (UBE3B) and the DNA repair protein uracil-DNA glycosylase (UNG) suggests they may be in the same survival pathway. Many DNA repair genes are regulated by ubiquitinylation and several crucial DNA repair proteins are E3 ligases such as BRCA1 (20). UBE3B previously has not been implicated in DNA repair or alkylation survival and elucidating its role and substrates will be important to discover how it improves alkylation survival. Importantly, the biological processes and corresponding genes identified in our functionome analysis represent

novel potential drug targets to increase TMZ efficacy independent of O^6-MeG mediated toxicity as well as potential epistatic relationships that would provide mechanistic insight into the function of specific activities.

Materials and methods

Chemicals and reagents

DharmaFECT 2 transfection reagent and the siGENOME Non-Targeting siRNA #1 were from Dharmacon (Lafayette, CO). CellTiter-Blue Cell Viability Assay and CellTiter 96 AQueous One Solution Cell Proliferation Assay were from Promega (Madison, WI). The Silencer Druggable Genome siRNA Library (Version 1.1) and 5x siRNA resuspension buffer were from Ambion (Austin, TX). Tissue culture-treated 384-well microtiter plates were from Greiner Bio-One (GmbH, Frickenhausen, Germany). OptiMEM, EMEM, phosphate buffered saline (PBS) and Hoechst 33342 were from InVitrogen (Carlsbad, CA). The shRNA vectors used for stable KD cell creation were obtained as glycerol stocks from Sigma-Aldrich. Temozolomide was from the National Cancer Institute Developmental Therapeutics Program (Bethesda, MD). Puromycin was from Clontech Laboratories (Mountain View, CA) and gentamicin was obtained from Irvine Scientific (Santa Ana, CA). We used the following primary antibodies: UNG antibody (#MBS200056) was from MyBioSource (San Diego, CA) and PCNA antibody (#sc-56) was from Santa Cruz (Santa Cruz, CA). Secondary antibodies: GAM-HRP conjugates were from Bio-Rad (Hercules, CA). Signal generation substrates were from Bio-Rad (Hercules, CA) and Thermo Fisher Scientific (Waltham, MA). All electrophoresis reagents were from Bio-Rad (Hercules, CA).

Cell lines and culture conditions

Cell line and culture conditions were as previously described (21-23). Briefly, T98G cells were cultured in EMEM with 10% heat-inactivated FBS, non-essential amino acids, sodium pyruvate, antibiotic/antimycotic and gentamicin. The LN428 cells were cultured in alpha MEM with 10% heat-inactivated FBS, antibiotic/antimycotic and gentamicin.

Pathway analysis and network visualization

Several different programs were used to analyze the hit list for pathway enrichment and network visualization, including NIH DAVID (DAVID), Princeton GO term finder and Ingenuity Pathway Analysis (IPA). DAVID and Princeton GO term finder both classify genes based on their gene ontology (GO) (24-26). Each gene is classified based on it's molecular function, biological process, or cellular compartment and are referred to as "GO terms" (25). Enriched GO terms and their associated genes are more likely to be true hits because multiple genes affecting the same process all yield sensitization (24, 27). The gene enrichment calculations used the 5,520 genes screened as the background, due to the relatively small number of genes screened, instead of the entire human genome. The analysis with NIH DAVID was performed using multiple classification stringencies, gene ontology hierarchies, protein interaction databases and pathways. The hit list was also analyzed with IPA (Ingenuity® Systems, www.ingenuity.com), utilizing a proprietary, manually-curated, interaction database. Because IPA utilizes a different scheme and algorithm to organize proteins into networks not based on GO terms, it was used as a distinct method to analyze the hit list for enriched protein functions and networks. The Functional Analysis identified the biological functions that were most significant to the data set. Right-tailed Fisher's exact test was used to calculate a p-value determining the probability that each biological function assigned to that data set is due to chance alone. Fisher's exact test was used with $\alpha = 0.05$ to compute the probability of correct functional assignment for the genes in the hit list. The enriched networks created by IPA were visualized with genes serving as nodes and edges representing known interactions. Genes colored green modulate toxicity to alkylating agents as determined in the screen.

Generation of lentiviral knockdown cell lines

The shuttle vectors for expression of shRNA were from Sigma (St. Louis, MO). Lentiviruses were prepared in collaboration with the UPCI Lentiviral facility. Lentiviral particles were generated by co-transfection of 4 plasmids [the shuttle vector plus three packaging plasmids: pMD2.g(VSVG), pVSV-REV and PMDLg/pRRE] into 293-FT cells (28, 29) using FuGene 6 Transfection Reagent (Roche, Indianapolis, IN), as described previously (22). Lentiviral transduction was performed as described earlier (22). Briefly, 6.0×10^4 cells were seeded into a 6-well plate 24 hours before transduction. Cells were transduced for 18 hours at 32˚C and then cultured for 72 hours at 37˚C. Cells were then selected by culturing in growth media with 1.0 µg/mL puromycin, as previously described (22).

Quantitative RT-PCR analysis

Expression of mRNA for each of ten glycosylases and five protein modification genes (OGG1, SMUG1, MBD4, UNG, MYH, NTHL1, MPG, NEIL1, NEIL2, NEIL3, UBE3B, ICMT, B4GALT7, CHRM3, and PADI1) after shRNA-mediated knockdown was measured by quantitative RT-PCR (qRT-PCR) using an Applied Biosystems StepOnePlus system as described previously (22). Applied Biosystems TaqMan® Gene Expression Assays used are as follows: human OGG1: Hs00213454_m1; human SMUG1: Hs00204820_m1; human MBD4: Hs00187498_m1; human UNG: Hs00422172_m1; human MYH: Hs01014856_m1; human NTHL1: Hs00267385; human MPG: Hs01012594_m1; human NEIL1: Hs0022637_m1; human NEIL2: Hs00376746_m1; human NEIL3: Hs00217387_m1; human UBE3B: Hs00296200_m1; Human ICMT: Hs00202655_m1; Human B4GALT7: Hs01011258_m1; Human CHRM3: Hs00265216_s1; Human PADI1: Hs00203458_m1. Gene expression of each gene was normalized to the expression of human ß-actin (part #4333762T).

Cell extract preparation and immunoblot

Nuclear extracts were prepared and protein concentrations were determined as described previously (22). Fifteen micrograms of protein was loaded on a pre-cast 4-20% Tris-Glycine gel (InVitrogen, Carlsbad, CA). The following primary antibodies were used in immunoblot assays: anti-human UNG from MyBioSource.com (#MBS200056) and anti-human proliferating cell nuclear antigen (PCNA; Santa Cruz Biotechnology). Immunoblot membranes were stripped before re-probing for PCNA. The membranes were stripped with Restore PLUS Western Blot Stripping buffer (#46430) from Thermo Fisher Scientific (Waltham, MA) per the manufacturer's instructions.

DNA glycosylase molecular beacon activity assay and data analysis

All oligodeoxyribonucleotides in the molecular beacon assay were purchased from Integrated DNA Technologies (Coralville, USA) including the following: FD-Con, 6-FAM-dGCACTATTGAATTGACACGCCATGTCGATCAATTCAATAGTGC-Dabcyl, where 6-FAM is carboxyfluorescein and Dabcyl is 4-(4'-dimethylaminophenylazo) benzoic acid; FD-UNG1, 6-FAM-dGCACTUAAGAATTGACACGCCATGTCGATCAATTCTTAAGTGC-Dabcyl, where U is 2'-deoxyuridine. All DNA glycosylase molecular beacon activity assay experiments and data analysis were as done as previously described (23, 30).

Briefly, the molecular beacons form a stem-loop structure containing a 13-nucleotide loop and a 15 base pair stem forcing 6-FAM fluorophore at the 5' end and the Dabcyl quencher at the 3' end into close proximity. When in a stem-loop structure, the 6-FAM fluorescence is efficiently quenched by Dabcyl in a non-fluorescent manner via F□rster Resonance Energy Transfer (FRET) (31, 32). However, if the uracil is removed by UNG and the DNA backbone is hydrolyzed by APE1, the 6-FAM containing oligonucleotide (5 bases in length) will dissociate from the hairpin at 37˚C (**Fig. 2D**). This will decrease Dabcyl quenching of 6-FAM and the

increase in fluorescence is proportional to uracil removal. Any increase in fluorescence in a control beacon with no base lesion is the result of non-specific DNA backbone cleavage.

Dialysis of nuclear lysates for molecular beacon assay were done as previously described (23, 30). Briefly, nuclear protein extracts were dialyzed twice for 90 min at 4°C using the 7,000 molecular weight cut-off Slide-A-Lyzer Dialysis Cassette (Pierce; Rockford, IL). Molecular beacon reactions were performed at 37°C using 2 μg of dialyzed protein extract and beacon substrate (final conc. = 40 nM). Fluorescence was measured every 20 seconds for 60 minutes, using a StepOnePlus real-time PCR system and expressed as arbitrary units (AU).

Cell cytotoxicity assays. Short-term MTS assay

TMZ induced cytotoxicity was determined as described previously (33). Briefly, cells were seeded into 96-well plates at a density of 2,000 cells per well. Cells were treated with TMZ for 48 hours at 37˚C before determining the relative amount of metabolically active cells by an MTS assay. Results were the average of three separate experiments and normalized to vehicle treated control cells with error bars representing the standard error of the mean. **Long-term CyQuant assay.** Cells were grown until approximately 50-75% confluence before being trypsinzed and counted using a CASY counter per the manufacturer's instructions. Cells were seeded into 96-well plates at a density of 120cells/well and incubated at 37˚C for 24 hrs. The cells were treated with vehicle or TMZ and incubated for nine days at 37˚C. Plates were removed and fluorescence was determined using the CyQuant kit (Invitrogen, #C7026) following the manufacturer's instructions. Results were the average of two separate experiments and normalized to vehicle treated control cells with error bars representing the standard error of the mean.

Cross species analysis of alkylation screens

The enriched biological process information associated with our TMZ screen was compared to other alkylation screens reported for *S. cerevisiae* and *E. coli.* Our cross species analysis was performed similar to as reported (19). Briefly, to determine if the same gene ontology biological processes were enriched in the alkylation sensitive hit lists from all three organisms, representing X number of genes, we aligned the GO-terms from each organism. To identify GO-terms significantly enriched for alkylation sensitive genes from all three organisms we randomly sampled X genes from the complete search space for each screen and identified associated GO functional terms. Random sampling was performed over two hundred iterations and the average number of hits in each GO category and the standard deviation were determined. These values were than compared to the actual values to identify GO terms significantly ($P < 0.05$) enriched for alkylation sensitive genes from all three organisms. The molecular functions and the corresponding genes, which were significantly enriched in all three alkylation screen data sets, were then visualized using Cytoscape (34). Genes which modulate alkylator toxicity from *E. coli* (dark blue)*, S. cerevisiae* (light blue) *and* human (light green), all served as protein nodes, with enriched GO terms serving as central nodes. The edges between GO terms and genes demarcate which genes are annotated to that biological process.

Generation of yeast knockout and viability analysis

Media preparation and other yeast manipulations were performed using standard methods. Mutants were made using a G418 knock out cassette from the *S. cerevisiae* Gene Deletion Project and were selected on Yeast Peptone Dextrose (YPD) plates containing G418 (200 mg/ml). A *mag1Δ* mutant was also made using a *URA3* based strategy with selection occurring on Synthetic Defined media lacking uracil (SD-URA). Mutants were confirmed by PCR. Plate based MMS viability studies were performed as previously reported (35).

Screen siRNA transient transfection and temozolomide treatment

Transient transfection by siRNA was as previously described (21, 36). Briefly, T98G cells were wet-reverse transfected with the Ambion Silencer Druggable Genome siRNA library. The library consisted of siRNAs targeting 5,520 different genes. Three unique siRNA duplexes targeting the same gene were pooled into a single well with one-gene target per well. The siRNA were prepared with DharmaFECT2 and OptiMEM and split evenly into two, 384-well plates. T98G cells were added directly to the siRNA complexes. The plates were incubated for five hours at 37ºC with 5% CO_2 before replacement with fresh media.

Cells were incubated at 37ºC with 5% CO_2 for 48 hr after siRNA transfection to allow for gene silencing before addition of TMZ or vehicle. The media was removed and the cells were treated with media containing either dimethyl sulfoxide (DMSO) vehicle (final concentration 1.0%) or the approximate EC_{10} of TMZ (final concentration 1 mM in 1% DMSO) for 48 hrs. Cell viability was measured 96 hours after siRNA transfection with the CellTiter-Blue viability assay according to manufacturer's protocol.

Druggable genome siRNA screen data analysis

The siRNA screen was performed 3 times over 6 separate weeks. Fluorescence units from each well were normalized to plate negative controls (scrambled siRNA) enabling cell viability comparisons between different plates. Data was analyzed by creating a high confidence hit list as described previously (37). Briefly, cell viabilities for each targeting gene from the three screening replicates were averaged and data was analyzed using two statistical analysis methods. For each gene, a two-sample t-test was performed to determine if there was a sensitization effect on cellular survival between 'siRNA plus vehicle' treated cells as compared to 'siRNA plus TMZ' treated cells. We selected targeting siRNAs with a p-value less than or equal to this threshold.

Also, viability ratios (ratio of survival of 'siRNA plus TMZ' to 'siRNA plus vehicle') were calculated for each targeting siRNA to determine the magnitude of response between the two screening conditions. The viability ratios were sorted in descending order and targeting siRNAs with a viability ratio in the lowest 5th percentile were selected for further analysis. The targeting siRNAs identified by both methods (p-value ≤ 0.05 and viability ratio in the lowest 5th percentile) were included in the final high-confidence hit list for further analysis and validation.

Transient siRNA knockdown for validation studies

T98G cells were wet-reverse transfected in six-well plates using siPORT NeoFX (Applied Biosystems) per the manufacturer's instruction. Briefly, the siRNA was diluted to a working concentration of 2μM before being used. Cells were plated at a concentration of 300,000 cells per well and an siRNA final concentration of 90nM. The cells were incubated for 24 hrs at 37°C before replacing transfection media with fresh media. Forty-eight hours after transfection cells were trypsinized and seeded for mRNA quantification and MTS cytotoxicity assays as described above. For knockdown of UNG and UBE3B the following Silencer Select siRNAs from Applied Biosystems were used: UNG: s14679; UBE3B: s40200; Silencer Negative control v2: am4613.

Preparation of cDNA for human astrocyte qRT-PCR

Total RNA was isolated and purified from T98G and LN428 glioblastoma cell lines and normal human astrocytes using Qiazol Lysis Reagent and MiRNeasy spin columns (Qiagen, Valencia, CA) according to the manufacturer's instructions. RNA quality was assessed using an Agilent Bioanalyzer (Agilent Technologies, Palo Alto, CA). RNAs with an RNA integrity number (RIN) above 9 were used for studies. Three μg RNA was reverse transcribed into cDNA using the SuperScript III first-Strand synthesis kit (Invitogen, #18080-400) according to the

manufacturer's instructions with Olig$(dT)_{20}$ primer in a 20 µl reaction. The cDNA library was then diluted 10 fold for qRT-PCR as described above.

REMBRANDT and TCGA queries

The REMBRANDT database was interrogated in February 2012 using the simple search functions for gene expression of UNG and UBE3B with related Kaplan-Meier survival curves (38, 39). The Cancer Genome Atlas GBM database was queried using the gene expression tool for two-fold greater or less expression of UNG and UBE3B compared to controls. Percentages of GBM overexpressing or underexpressing UNG or UBE3B were recorded.

Results

Synthetic lethal siRNA screen and generation of high confidence hit list

To identify genes that when silenced confer sensitivity to the alkylating agent TMZ, we conducted a synthetic lethal screen (in triplicate) on a TMZ-resistant GBM cell line (T98G) using TMZ and an siRNA library (**Fig. S1A**). A concentration of 1 mM TMZ was selected due to minimal (~10%) toxicity measured at 48 hrs via an MTS assay. Further, T98G cells were treated with varying concentrations of DMSO to determine the impact of DMSO on cell survival (**Fig. S1B**). No toxicity was observed at 1% DMSO after 48 hrs, the % DMSO used for both vehicle and TMZ treatment in the screen.

The transfection conditions used in the screen maximized knockdown, similar to previous screens using these conditions (21). We first calculated viability ratios (ratio of survival of 'siRNA plus TMZ' to 'siRNA plus vehicle') and normalized to vehicle-treated scrambled siRNA, enabling cross plate comparisons. We then statistically analyzed the viability ratios and determined a group of gene targets that sensitized cells to TMZ by performing a two-tail sample t-test on each gene to determine the effect of siRNA and TMZ on cellular survival. Targeting siRNAs were selected with a p-value of less than or equal to 0.05. The high confidence hit list contained targets that satisfied both criteria: a p-value ≤ 0.05 and a viability ratio in the lowest 5 percent (**Fig. S1C**). By using these two methods, we created a hit list of 172 genes (**Dataset S1**). While most of the viability ratios on the hit list were between 0.4 and 0.65, these genes significantly modulated the toxicity of TMZ, yielding many new potential targets to increase response to alkylation chemotherapy. All hits were analyzed for pathway and network enrichment to determine the potential biological pathways that modulated alkylation toxicity

Analysis of screen results for pathway and network enrichment

The genes contained in the hit list were analyzed with NIH DAVID, Princeton GO term

finder and Ingenuity Pathways Analysis (Ingenuity® Systems, www.ingenuity.com) (IPA) for pathway and network enrichment. The analysis with DAVID and Princeton GO term finder was performed to search for gene enrichment in molecular functions, cellular compartments, or biological processes based on the associated gene ontology terms. Each gene is linked to different terms based on their gene ontology in a hierarchical manner and gene enrichment was determined based on the probability of selecting proteins with similar terms. By focusing on 5,520 druggable genes, we expected a small but presumably more relevant gene enrichment dataset compared to probing the entire human genome as background. Several different thresholds for viability ratio percentiles were initially analyzed for gene enrichment, including 2.5%, 5%, 7.5% and 10% (**Datasets S2-S4**). The 5% group was selected for a more detailed analysis because it contained many of the gene enrichment groups seen in both the 7.5% and 10% datasets, but was more focused with higher gene enrichment scores.

The enriched genes from the NIH DAVID analysis contained several different groups including DNA repair, response to DNA damage stimulus, cell proliferation, amino acid glycosylation and biopolymer glycosylation (**Table 1**). The most enriched groups have specific DNA repair functions including BER activity, hydrolase activity and DNA N-glycosylase activity. Surprisingly, four genes belonging to the three most enriched groups were DNA glycosylases that do not recognize alkylation damage (7). Interestingly, these genes (TDG, OGG1, NEIL1 and UNG) are known to recognize and repair multiple types of oxidative DNA damage (**Table 2**) (7).

There were large overlaps of enriched genes from DAVID and Princeton GO term finder analyses (**Fig. S2 and Dataset S5**). The oxidative DNA glycosylases, which sensitized cells to TMZ, were identified in the most significantly enriched networks determined by both DAVID and Princeton GO term finder (**Table 1 and Dataset S6**). Data were also analyzed through the use of IPA. IPA uses a manually curated interaction database and a right-tailed Fisher's exact test to calculate a p-value, determining the probability that each biological function assigned to that

network is due to chance alone. Analysis using IPA identified several of the DNA glycosylases in a statistically significant network, with a network value of twenty-four (**Fig. 1**). The network score is the negative logarithm of the p-value (p-value =10^{-24}). This network is very similar to the biological processes and corresponding genes overrepresented in the DAVID and Princeton GO analysis. Further, IPA returned several highly enriched molecular functions including DNA repair (p-value=5.31 x 10^{-4}), excision repair (p-value=8.87 x 10^{-4}), ligation of a DNA fragment (p-value= 2.84 x 10^{-3}), ligation of DNA (p-value= 9.07 x 10^{-3}) and nicking of DNA (p-value=1.83 x 10^{-2}), very similar to the functions returned by DAVID.

Other highly enriched molecular functions included post-translational modifications (p-value= 2.98 x 10^{-3} - 3.12 x 10^{-2}), carbohydrate metabolism (p-value= 5.56 x 10^{-3} - 4.57 x 10^{-2}), nucleic acid metabolism (p-value= 9.65 x 10^{-4} - 3.12 x 10^{-2}), and small molecule biochemistry (p-value= 9.49 x 10^{-4} - 4.7 x 10^{-2}). Within these different functional groups, a conserved theme was the ability of the genes to modify DNA or proteins. The modifications of DNA most significantly involved the excision of uracil (p-value = 9.65 x 10^{-4}), while the many protein modifications included metabolism of proteoglycan (p-value = 5.56 x 10^{-3}), metabolism of polysaccharide (p-value = 7.29 x 10^{-3}), metabolism of carbohydrate (p-value = 8.05 x 10^{-3}), generation of diacylglycerol (p-value = 5.56 x 10^{-3}) and the release of acetylcholine (p-value = 9.07 x 10^{-3}). The “Protein Modification” theme was also observed in the DAVID and Princeton GO analyses, and included protein amino acid glycosylation, bipolymer glycosylation, glycoprotein biosynthetic process, glycoprotein metabolic process and cellular carbohydrate metabolic process. These processes contained genes whose corresponding proteins are involved in protein modifications and when silenced, the corresponding cells are sensitive to TMZ. However, because of the highly significant enrichment of DNA repair processes in all three programs, genes associated with the DNA Repair pathway category were selected for initial biological validation.

Knockdown of oxidative DNA glycosylases sensitize cells to TMZ

In our analysis, we discovered that the DNA repair category was enriched for TMZ sensitizing genes. This category contained four DNA glycosylases (UNG, OGG1, TDG and NEIL1), each specific for the repair of oxidative DNA damage (7). To determine if other DNA glycosylases also sensitize cells to TMZ and to validate the siRNA synthetic lethal screen results, seven of the eleven DNA glycosylases were knocked down using a lentiviral system to create stable T98G-derived cell lines. Knockdown of glycosylase mRNA was validated by qRT-PCR (**Fig. 2A**). The stable cell lines were then tested for sensitization to 1 mM TMZ in a 96-well plate format using a modified MTS assay (22). As confirmation of our siRNA screen analysis, the newly developed UNG-KD cell line was also sensitive to TMZ (**Fig. 2B**). Furthermore, knockdown of the DNA glycosylases MYH and MPG also sensitized cells to TMZ, as did knockdown of NEIL2 (**Fig. 2B**).

UNG knockdown eliminates removal of uracil in glioma cells

Although UNG mRNA knockdown sensitized the T98G cells to TMZ in the validation study, we were uncertain if the mRNA knockdown affected UNG protein levels and DNA repair activity. First, we tested UNG protein expression via immunoblot and determined that UNG protein levels were decreased in T98G/UNG-KD cells as compared to T98G-GFP control cells (**Fig. 2C**). However, we were interested in ascertaining if the decrease in UNG protein levels affects the DNA repair capacity of the cells, because there is a large functional overlap of UNG with the three DNA glycoslylases SMUG1, TDG and MBD4 (7). We were concerned that depletion of one of these DNA glycosylases would not have a functional DNA repair defect due to compensation by the other glycosylases. We therefore developed a DNA glycosylase molecular beacon assay to quantify the functional loss of uracil removal by UNG knockdown, essentially as we have described previously for the analysis of MPG activity (23). The assay

uses a molecular beacon composed of a single stranded DNA molecule with a 5' fluorophore (6-FAM) and a 3' quencher (Dabcyl). The oligonucleotide also contains a uracil lesion, a substrate for UNG. We selected uracil as the lesion because it is removed by UNG, SMUG1, TDG and MBD4 (7). Therefore, we could determine if UNG-KD impacts uracil removal and if compensation by other DNA repair proteins occurred. A similar oligonucleotide with a normal base was used as a control substrate. Removal of the uracil lesion by UNG and hydrolysis of the DNA backbone by APE1 results in separation of the 6-FAM fluorophore from the Dabcyl quencher and the increase in 6-FAM fluorescence is proportional to uracil removal (**Fig. 2D**). The T98G-SCR cell lysate (from cells expressing a scrambled shRNA) incubated with control beacon (**Fig. 2E**, red circles) had a minimal increase in fluorescence, signifying the control beacon is intact. However, the T98G-SCR lysate incubated with the beacon containing uracil (**Fig. 2E**, blue squares) exhibited a large increase in fluorescence (17.69 fold at 60 min) compared to the control beacon, indicative of robust uracil removal. The T98G/UNG-KD lysates incubated with either the control beacon (**Fig. 2E**, green diamonds) or the uracil-containing beacon (**Fig. 2E**, yellow triangles) contained no difference in fluorescence.

These results support our conclusion that the T98G/UNG-KD cells have impaired uracil removal due to UNG knockdown as compared to the T98G-SCR control, with no evidence for repair compensation by SMUG1, TDG or MBD4. The scrambled shRNA control was used instead of the GFP control for the activity assay due to the interference of GFP fluorescence with the molecular beacon fluorophore 6-FAM. Although uracil can also be removed by SMUG1, TDG and MBD4, UNG-KD alone was sufficient to deplete uracil removal activity (7).

Cross-species functionome analysis creates a network of conserved processes important for survival after alkylation damage

After biological validation, we then compared our T98G siRNA screen results to other alkylation screens conducted in *S. cerevisiae* and *E. coli*, using GO molecular function terms (17, 18). This was done by compiling the sensitizing genes from all the screens and linking them to their specific Gene Ontology Biological Process. The list of sensitizing genes from *E. coli, S. cerevisiae* and human was then computationally analyzed to identify GO-functional categories over-represented with genes from human and either *E. coli* or *S. cerevisiae* (**Supplemental Table S1**). The molecular functions enriched in all three alkylation screen datasets were then visualized using Cytoscape to create a cross species functionome of proteins which modulate toxicity to alkylating agents in *S. cerevisiae*, *E. coli* and human (**Fig. 3**). This network of proteins has several "GO-hubs" of proteins from all three species clustered around a conserved biological process (node) that significantly impacts the survival of correspondingly depleted cells after exposure to alkylating agents. These include proteins involved in global processes such as response to drugs or changes in pH. There was also a collection of processes involved in macromolecule biosynthesis and modifications including fatty acid biosynthetic processes, transcription, regulation of transcription, transcription initiation, negative regulation of translation, protein modification process, protein processing and protein targeting to membrane. These results suggest that synthesis of new RNA, protein, and fatty acids are essential for survival to alkylation.

Protein Modifications are essential for yeast and human cell survival to TMZ

Many genes in yeast and humans whose corresponding activities are involved in protein modification are required for protection from alkylation damage mediated by TMZ or methyl methanesulfonate (MMS) exposure. Several different types of protein modifications including

ubiquitinylation, methylation, glycosylation and deimination are present in the hub (**Fig. 3**). These include two of the four known human peptidyl arginine deiminases, type I and type IV (PADI1 and PADI4) (40). Petidyl arginine deiminases catalyzes the post-translational deimination of proteins by converting arginine residues into citrullines (40). The PADs have distinct substrate specificities and tissue-specific expression patterns (41, 42). The enzymatic reaction is highly dependent on pH and calcium concentrations, with the different PAD isoforms containing different activity profiles (41). The differential expression patterns, substrates and activity profiles all contribute to the distinct functions of the PAD genes. Currently, the most prevalent function of PADI1 is its role in epidermal differentiation, where it deiminates filaggrin and keratin (41, 42). However, its role in DNA repair or survival to alkylation exposure has not been determined. In contrast, PADI4 is likely involved in DNA repair and cell death signaling as it can modulate p53 signaling and p53 gene target expression levels (43, 44). PADI4 can also antagonize histone methylation by arginine deimination to citruline, thus removing the methylation mark (45). PADI4 also regulates gene expression by associating with histone deacetylase 2 (HDAC2) (46). HDAC2 is known to modulate the DNA damage response and cells depleted of HDAC2 are hypersensitive to DNA damaging agents (47). Further, glioblastoma cells pretreated with the HDAC inhibitor MS275 were sensitized to TMZ (48). These data suggest that PADI4 sensitivity may result from interference with multiple steps in the DNA damage response, likely through deregulation of HDAC2, but PADI4's interaction with p53 likely contributes to alkylation survival as well.

One of the human genes in this sensitization node is isoprenylcysteine carboxyl methyltransferase (ICMT), a gene that is ubiquitously expressed and encodes the last of three steps for isoprenylation (49, 50). Methylation of isoprenylcysteine allows movement of the protein from the endoplasmic reticulum to the plasma membrane (50). ICMT has been of interest to cancer researchers because ICMT inhibition can disrupt RAS signaling in several

model systems (51-53). Knockout of ICMT prevented oncogenic transformation of cells with oncogenic K-RAS or B-RAF in soft agar and mice models (52). Although there is interest in creating potent and selective ICMT inhibitors, there has been no link for ICMT to DNA repair or response to alkylation exposure (54). ICMT may play a role downstream of the DNA damage cascade and regulate signaling for cell death, as it has been linked with both autophagic and apoptotic cell death (55).

Another gene in the protein modification group is the human ubiquitin protein ligase E3B (UBE3B), suggested herein to play a role in the repair of or cellular response to alkylation damage. The activity of many DNA repair proteins can be affected by ubiquitinylation such as FANCD2 (56) and PCNA (57). Monoubiquitinylation of FANCD2 is used as a marker of intra-strand crosslink (ICL) repair activity, while ubiquitinylation of PCNA determines polymerase switching (58). Furthermore, the E3 ubiquitin ligase activity is critical for the function of several crucial DNA repair proteins, for example BRCA1, CUL4A and CUL4B (20, 59). The E3 activity of BRCA1 is esential for its role in cell cycle checkpoint activation and sensitivity to DNA damage, while CUL4A and CUL4B regulates chromatin structure and access of DNA repair proteins to the DNA in nucleotide excision repair (NER) (20, 59, 60). The potential role of UBE3B and its substrates in DNA repair are of great interest, but has yet to be defined.

DNA repair defects increase sensitivity to alkylating agents across species

Another conserved theme involved DNA metabolism and included the GO hubs of DNA metabolic process, DNA replication, negative regulation of DNA replication, response to DNA damage stimulus, DNA repair, DNA recombination, mismatch repair, nucleotide excision repair, BER and DNA dealkylation. The prevalence of DNA replication and DNA recombination genes emphasizes the importance of new DNA synthesis and repairing damaged DNA via recombination mechanisms. Many genes involved in other DNA repair pathways including

nucleotide excision repair, mismatch repair and BER are present in the DNA repair hub. The BER subset from all three organisms was predominantly composed of DNA glycosylases that recognize and remove base lesions. This subset contained several known alkylation resistance genes such as the 3-methyl-adenine DNA glycosylases *alka* and MAG1, in *E. coli* and *S. cerevisiae,* respectively. Also included were human DNA glycosylases that recognize and repair oxidative base lesions including UNG, OGG1, TDG and NEIL1 (7). The *alkA* gene product has diverse substrate specificity and also removes oxidative lesions such as xanthine, oxanine, 5-formyluracil and hypoxanthine (61-64). The ability of AlkA to remove oxidation induced lesions correlates with the results seen in the human screen and may affect the sensitivity of *alkA* mutants to alkylation damage. The emphasis on proteins that recognize and remove DNA lesions further demonstrates that the DNA lesions themselves are cytotoxic, with removal of both alkylation and oxidative lesions being important for survival.

Many of the biological processes identified from our cross-species functionome analysis are conserved, containing multiple proteins from each organism. We have used this approach to discover a network of proteins that constitutes a cross-species functionome of evolutionarily important processes essential for survival after alkylation exposure. The functionome analysis affirms that our TMZ screen results from human cells are consistent with previous alkylation studies in other species. This network links seemingly disparate genes that may prevent cell death after alkylation damage through their crucial biological processes and represents novel targets for adjuvant chemotherapy.

Disruption of protein modification processes sensitizes human cells to alkylators

Based on the functionome network we identified, we chose to further validate another node of proteins to improve sensitivity to alkylators. The protein modifications node was chosen due to enrichment in the cross-species analysis, good sensitization in the screening studies and

the observation that this node had multiple genes in both human and yeast cells. Human knockdown cell lines were generated using lentivirus shRNA vectors as described above. Knockdown levels of mRNA were determined using qRT-PCR (**Fig. S3A**). The cell lines with mRNA knockdown were then tested for sensitivity to TMZ treatment using a modified MTS assay. The knockdown of UBE3B and ICMT significantly sensitized the glioblastoma cells to TMZ resulting in viability ratios of 0.6 and 0.7, respectively (**Fig. 4B**), as compared to the control. We then tested each knockdown cell line for sensitivity to TMZ in a long-term assay similar to a clonogenic cell survival assay. The CyQuant assay utilizes a highly sensitive fluorescent intercalating DNA dye to determine relative DNA content and cell number nine days after TMZ or vehicle treatment. The UBE3B-KD cell lines were approximately 45% more sensitive compared to control cells in the CyQuant assay at concentrations as low as 50μM TMZ (**Fig. 4C**).

Protein modification gene knockout sensitizes yeast to alkylators, while simultaneous knockout of protein modification genes and an alkylbase DNA glycosylase yields the greatest sensitivity

We have previously reported that *aim22Δ, lip22Δ, pby1Δ* and *stp22Δ* cells from *S. cerevisiae,* which belong to the protein modification node*,* are sensitive to the alkylating agent MMS (17). In addition, it has been firmly established that *mag1Δ* cells are sensitive to MMS because of their alkylbase DNA glycosylase deficiency and inability to repair damaged DNA. We assayed all individual knockouts to further validate our previously reported screening results (**Fig. 4A**) and demonstrate decreased growth after MMS treatment for all five mutants. Based on our functionome results, we reasoned that there would be increased alkylation sensitivity when deficiencies in protein modification and BER were combined. We generated double knockouts in the four protein modification associated mutants, using a *mag1Δ* deletion cassette. We

demonstrate that there is increased sensitivity in all of the double mutants tested, relative to the individual parent or mag1Δ knockout strains. The protein modification genes correspond to activities involved in the modification of mitochondrial enzymes by the attachment of lipoic acid groups (Lip2), a protein that works with Lip2 (Aim22), a tubulin tyrosine ligase associated with P-bodies (Pby1) and a component of the ESCRT complex that is involved in ubiquitin-dependent sorting of proteins into the endosome (Stp22). We note that Pby1 is homologous to the mouse and human Tsg101 tumor susceptibility genes, which are homologs of ubiquitin-conjugating enzymes implicated in being involved in cell cycle regulation and genome maintenance. The *mag1Δstp22Δ* double mutant is the most sensitive to MMS, but there also appears to be a growth defect in these strains on untreated media, suggesting some synthetic lethal interaction outside of alkylation damage. Nonetheless, our results with combined protein modification and BER mutants in budding yeast supports the idea that similar dual knockouts in humans may confer increased sensitivity to TMZ.

Dual disruption of both DNA repair and protein modification nodes reveals a potential epistatic relationship between UNG and UBE3B

Although the UNG-KD and UBE3B-KD cell lines were both more sensitive to TMZ when compared to control cells, we were interested in determining if the double-KD cells would have an increased TMZ sensitivity, as was seen for the double KO lines in *S. cerevisiae* (**Fig. 4A**). To this end, we investigated if simultaneous inhibition of both DNA repair and protein modification genes increased sensitivity compared to either inhibition alone. However, we noticed during long-term passaging of the UBE3B-KD cell line that sensitivity and knockdown levels diminished, suggesting outgrowth of WT cells in a pooled population.

To prevent the loss of mRNA knockdown, we returned to using an siRNA transient transfection approach to knockdown both UNG and UBE3B. We transfected parental T98G cells

with a scrambled control siRNA, UNG siRNA, UBE3B siRNA, or UNG plus UBE3B siRNA before determining mRNA levels by qRT-PCR (**Fig. S3B**). Cells were seeded for qRT-PCR and sensitivity determination by MTS assay at the same time. Wells with less than 35% mRNA for either UNG or UBE3B remaining were analyzed for sensitivity to TMZ. We expected the dual knockdown of UNG and UBE3B to confer at least an additive effect to TMZ sensitivity, suggesting that the sensitivity observed after knockdown of the proteins resulted from independent mechanisms. However, although UNG and UBE3B knockdown alone both confer sensitivity to TMZ, when they are simultaneously knocked down there is no significant increase in sensitivity to TMZ below the level of either single knockdown (**Fig. 4D**). Thus, TMZ sensitivity to knockdown of UNG and UBE3B are not independent events, demonstrating a possible epistatic relationship between the UNG and UBE3B genes. The conserved biological processes also enrich for possible alkylation sensitivity gene interactions as seen with UNG and UBE3B. The activity of many DNA glycosylases are regulated by post-translational modifications (6). SUMOylation of TDG promotes its catalytic activity by increasing catalytic turnover by decreasing TDG's affinity for the abasic site product (65). Although it is more likely UBE3B has an indirect role in UNG sensitivity, it is possible that ubiquitinylation of UNG by UBE3B can modify protein localization, abundance or activity as it has been suggested that UNG is targeted for ubiquitin-mediated proteolysis during the S phase of the cell cycle (66, 67).

UNG and UBE3B are differentially expressed in various cancer cell lines

Expression of UBE3B affected GBM sensitivity to TMZ and may impact response of various cancers to other chemotherapeutic alkylating agents such as cyclophosphamide, chlorabucil and dacarbazine. With little known on the function or expression of UBE3B in normal and cancerous tissue, we wanted to determine if UBE3B expression levels vary in different cancer cell lines. To determine if UBE3B is differentially expressed in different cancer types we

quantified relative expression levels of UBE3B mRNA by qRT-PCR. We discovered that UBE3B mRNA expression varies by approximately 7-8 fold in the tumor cell lines tested (**Figure 5A).** UBE3B expression fluctuates in our different cancer cell lines and we sought to determine if UBE3B and UNG levels vary in the LN428 and T98G cell lines when compared to normal human astrocyte controls. We determined relative mRNA levels of UNG and UBE3B in LN428 and T98G cell lines compared to normal human astrocyte controls using qRT-PCR (**Figure 5B**). Although UBE3B levels did not fluctuate in the three cell lines, there was a 2-fold increase in UNG expression in both cancer cell lines when compared to human astrocyte controls.

Discussion

GBMs are the most common and aggressive primary brain tumor; the current standard of care consists of maximum surgical removal, radiation and TMZ (2-4). Many tumors are refractory to TMZ treatment, resulting in very poor outcomes. Difficulties in treating GBM make it a candidate for drug research since an increase in therapeutic options is necessary to improve survival rate. Because TMZ clinical efficacy is primarily due to the toxicity of the O^6-MeG DNA lesion, much of the alkylation damage sensitization research has focused on this lesion. Unfortunately, previous attempts to improve efficacy of O^6-MeG lesion-mediated cytotoxicity have been futile. The clinical trials testing the effectiveness of MGMT inhibitors do not show an increase in efficacy or sensitivity (12), particularly in TMZ-resistant GBM (13). Due to the inherent difficulties and resistance of GBM to chemotherapies, any increase in cancer sensitization could greatly impact patient outcomes.

To discover “druggable” targets that contribute to TMZ sensitivity, independent of the O^6-MeG lesion, we conducted a synthetic lethal siRNA screen against 5,520 genes in a TMZ-resistant cell line that has elevated expression of MGMT (T98G), uncovering enriched biological processes independent of MGMT and O^6-MeG lesion induced cell death that included 172 genes (**Dataset S1**), many corresponding to similar or over-lapping biological processes as determined by NIH DAVID, Princeton GO term finder and IPA. After enrichment analysis, increased TMZ-mediated sensitization in cells depleted of UNG (UNG-KD) was validated using a separate shRNA targeting sequence and an independent experimental design. Thus, cell sensitization to alkylating agents with UNG-KD is not due to RNAi off-target effects or artifacts of the screening protocol. We could not create stable knockdowns of the other DNA glycosylase hits, likely because the knockdown of OGG1, TDG and NEIL1 created cells with a growth disadvantage compared to cells expressing normal levels of the corresponding protein.

Although UNG has not been reported to repair alkylation damage, knockdown of UNG sensitizes T98G cells to TMZ. It is possible that TMZ induces toxic oxidative lesions by increasing reactive oxygen species (ROS) levels as seen with MMS in yeast (68). TMZ may directly alkylate the electron transport chain proteins and mitochondrial DNA, which could impair electron flow through the chain by direct alkylation damage of the protein and decreased functional protein expression due to replication blocking lesions or mutated DNA. It is possible that mitochondrial alkylation damage induces an increase in ROS formation. Mitochondrial ROS production can increase from defects in electron transport chain proteins, such as complex I (69). Cells lacking oxidative DNA glycosylases may be more sensitive to ROS because repair of oxidative lesions is compromised. For example, cells lacking NEIL1 are more sensitive to gamma-irradiation (70) and NEIL1$^{-/-}$ mice have decreased expansion of germinal center B cells (71). Thus, an increase in ROS may create lethal oxidative lesions that UNG (e.g., 5,6-dihydro-2'-deoxyuridine) and other oxidative DNA glycosylases must repair to prevent cell death after alkylation exposure (7). NAD kinase, which phosphorylates NAD to form NADP$^+$, was also a hit in our screen. Cells keep NADP$^+$ in its reduced state, NADPH, to use NADPH's reducing potential in many different pathways including glutathione regeneration for oxidant defense, reduction of RNA to DNA, and synthesis of fatty and amino acids. New synthesis and repair of macromolecules is essential for alkylation survival (18, 35, 72). NAD kinase may affect alkylation survival via NADPH production by its positive effect on DNA and RNA synthesis and also for its central role in creating reducing equivalents for antioxidant defense. Support for a potential role of ROS in alkylation sensitivity was also shown by others using mouse embryonic fibroblasts and several human cancer cell lines, which were treated with the alkylating agent MNNG and contained greater ROS production from both NADPH oxidase and mitochondria sources (73, 74). In the report by Chiu et at, MNNG cytotoxicity could be completely abrogated by pretreatment with N-acetylcysteine, demonstrating that increased ROS formation may

contribute to alkylation toxicity and is a likely source of cytotoxic DNA lesions (74). Cells expressing UNG may be less sensitive to TMZ because the increase in ROS produces cytotoxic oxidative DNA lesions removed by UNG, promoting cellular survival.

UNG was not the only DNA glycosylase to sensitize cells to alkylation exposure in the validation experiments. MPG and MYH knockdown (KD) also sensitized T98G cells to TMZ. This was unexpected because MPG did not sensitize in the siRNA screen, perhaps due to lack of knockdown. However, our results are consistent with previous reports from several labs, as MPG-KD led to either an increase or decrease in cell death after alkylation exposure depending on the system under study (75-78). Further, MPG status alone does not adequately predict response to alkylators, but instead the balance of the entire BER pathway must be investigated to predict sensitivity to DNA damage (23, 30, 79). The ability of MYH-KD to sensitize cells to alkylators was not anticipated since expression of MYH had been previously shown to promote cell death after alkylator exposure (80). In that study, however, the authors used a clonogenic cell survival assay that interrogates cell death related to the O^6-MeG lesion because the cells can undergo two or more replication cycles (80). In contrast, in our initial validation experiments, cell survival was determined after 48 hours post TMZ treatment, too short for two replication cycles to occur (81). Therefore, the role of MYH in O^6-MeG mediated cell death is likely very different than its role in cell survival after 48hr exposure to TMZ as described in this study. MYH is primarily known for the removal of the mismatched A opposite 8-oxoG lesions (82, 83) and MYH deficiency does not sensitize cells to H_2O_2, IR, or *cis*-platinum (7). Therefore, the sensitization of MYH depletion to TMZ treatment may be the result of an increase in ROS induced lesions.

After biological validation was completed, the results were compared to similar alkylator sensitivity screens done in *S. cerevisiae* and *E. coli* (17, 18). Many of the same biological processes significantly enriched in the human screen were enriched in both screens from yeast

and bacteria. There are a diverse collection of biological processes, ranging from DNA repair to modulation of transcription, translation, protein modifications and fatty acid biosynthesis that appear to be essential for survival after alkylation exposure. The many processes necessary for cellular survival after alkylation exposure likely stem from the alkylator's ability to modify all macromolecules in a cell, thereby requiring repair of not only DNA, but also of RNA, proteins and fatty acids. In the human siRNA screen, these data suggests a novel epistatic relationship between UNG and UBE3B. This is likely an indirect interaction with downstream targets of UBE3B modulating TMZ cytotoxicity through a UNG dependent pathway. However, we find UBE3B expression to be highly variable when comparing many cancer cell lines but expression is lower than that found in cultured astrocytes (**Fig. 5**). A direct interaction between UBE3B and UNG may be exist, since UNG is a known target of ubiquitinylation (67, 84, 85), although the data described herein would not implicate UBE3B in the cell cycle-regulated proteolysis of UNG.

The analysis of the three screens revealed that many biological processes modulate survival after alkylation damage. The greatest alkylator toxicity in yeast cells was generated by dual inhibition (KO) of both BER and the protein modification processes, suggesting that simultaneous inhibition may be required to achieve a clinically significant sensitization for chemotherapies, provided this approach can provide tumor selectivity. In that regard, we evaluated the range of UNG expression levels in primary patient tumors by interrogating the Cancer Genome Atlas (TCGA) dataset on GBM. UNG expression is two-fold greater in approximately 14% of the GBMs studied in the TCGA. Interestingly, there are no samples with a two-fold decrease in UNG expression in the TCGA database. We also queried the Rembrandt database for UNG gene expression data and found that greater than 33% of GBMs contained 2-fold or greater expression of UNG (39). These results, from a separate database, reinforced our hypothesis that expression levels of UNG varied significantly in tumors. The directionality of the changes were consistent and it appears that over-expression of UNG is beneficial for GBM

tumor growth or survival, with less than 1% of tumors containing a two-fold decrease in UNG expression in the REMBRANDT database. This trend also extended to other brain tumor types such as astrocytomas, which had greater than 14% of tested tumors containing two-fold greater expression of UNG and no samples with a two-fold down regulation of UNG (39).

Knowing that the expression levels of UNG fluctuate in primary tumor samples, we asked if expression levels of UNG and UBE3B correlated with response to therapy or survival. We would predict the increase in UNG expression levels would correlate with reduced survival due to the ability of the cells to repair their DNA more efficiently, especially TMZ induced lesions. We also queried the REMBRANDT database for clinical outcomes related to UNG overexpression compared to intermediate expression in all glioma: two-fold over-expression of UNG correlated with a decrease in the probability of survival compared to intermediate expression of UNG, with a log-rank p-value, calculated using the Mantel-Haenszel procedure, of 1.0177×10^{-4}. This correlation matches our prediction based on our cell culture data, suggesting that increased expression of UNG negatively impacts brain tumor patient survival, possibly from increased resistance to chemotherapeutic treatment.

In summary, these studies suggest that alkylation resistance mechanisms are evolutionarily conserved. The collection of conserved biological processes in *E. coli*, *S. cerevisiae* and humans composes an Alkylation Functionome that includes many novel proteins not previously thought to impact alkylation resistance (**Figure 6**). We can begin to appreciate the multiple processes that are required for cellular survival after alkylation damage and form an unbiased approach to discover targets for adjuvant chemotherapy. This is an important next step, so as to determine if genes that were not tested in the screen, but share conserved biological pathways, are also possible targets to enhance TMZ response. Because the analysis has also been enriched for the most essential pathways and genes, there may be functional overlap of the conserved genes and pathways. The highly conserved nature of these biological

processes lends mechanistic insight into potential gene and pathway interactions. Determining the relationships between these genes and biological processes is important to determine pathway redundancy. Due to the plethora of resistance mechanisms, it is possible that depletion or inhibition of one gene in a conserved biological process cannot overcome the drug resistance due to compensation by another gene or biological process. By investigating multiple pathways and processes we can determine important interactions that promote tumor survival to be targeted for improved chemotherapy response.

Disclosure of Potential Conflicts of interest

RWS is a scientific consultant for Trevigen, Inc. The remaining authors state that there is no conflict of interest.

Grant Support

This work was supported by grants from the National Brain Tumor Society and the National Institutes of Health (NIH) [GM087798; CA148629; ES019498] to RWS and TJB (ES01225101) and a NYSTAR James Watson Award to TJB. Support was also provided by the University of Pittsburgh Department of Pharmacology and Chemical Biology with a Molecular Pharmacology Fellowship to DS. Support for the UPCI Lentiviral Facility was provided by the Cancer Center Support Grant from the National Institutes of Health [CA047904].

References

1. Holland EC. Glioblastoma multiforme: the terminator. Proc Natl Acad Sci U S A. 2000;97:6242-4.
2. Cohen MH, Johnson JR, Pazdur R. Food and Drug Administration Drug approval summary: temozolomide plus radiation therapy for the treatment of newly diagnosed glioblastoma multiforme. Clin Cancer Res. 2005;11:6767-71.
3. Stupp R, Hegi ME, Mason WP, van den Bent MJ, Taphoorn MJ, Janzer RC, et al. Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. Lancet Oncol. 2009;10:459-66.
4. Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJ, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. New England Journal of Medicine. 2005;352:987-96.
5. Sobol RW. Temozolomide. In: Schwab M, editor. Encyclopedia of Cancer. 2nd ed. Berlin, Heidelberg, New York: Springer; 2009.
6. Almeida KH, Sobol RW. A unified view of base excision repair: lesion-dependent protein complexes regulated by post-translational modification. DNA Repair. 2007;6:695-711.
7. Svilar D, Goellner EM, Almeida KH, Sobol RW. Base Excision Repair and lesion-dependent sub-pathways for repair of oxidative DNA damage. Antioxid Redox Signal. 2011;14:2491-507.
8. Gerson SL. MGMT: its role in cancer aetiology and cancer therapeutics. Nat Rev Cancer. 2004;4:296-307.
9. Fu D, Calvo JA, Samson LD. Balancing repair and tolerance of DNA damage caused by alkylating agents. Nat Rev Cancer. 2012;12:104-20.
10. Cahill DP, Levine KK, Betensky RA, Codd PJ, Romany CA, Reavie LB, et al. Loss of the mismatch repair protein MSH6 in human glioblastomas is associated with tumor progression during temozolomide treatment. Clin Cancer Res. 2007;13:2038-45.
11. Sarkaria JN, Kitange GJ, James CD, Plummer R, Calvert H, Weller M, et al. Mechanisms of chemoresistance to alkylating agents in malignant glioma. Clin Cancer Res. 2008;14:2900-8.
12. McMurry TB. MGMT inhibitors--The Trinity College-Paterson Institute experience, a chemist's perception. DNA Repair (Amst). 2007;6:1161-9.
13. Quinn JA, Jiang SX, Reardon DA, Desjardins A, Vredenburgh JJ, Rich JN, et al. Phase II trial of temozolomide plus o6-benzylguanine in adults with recurrent, temozolomide-resistant malignant glioma. J Clin Oncol. 2009;27:1262-7.
14. Yip S, Miao J, Cahill DP, Iafrate AJ, Aldape K, Nutt CL, et al. MSH6 mutations arise in glioblastomas during temozolomide therapy and mediate temozolomide resistance. Clin Cancer Res. 2009;15:4622-9.
15. Hunter C, Smith R, Cahill DP, Stephens P, Stevens C, Teague J, et al. A hypermutation phenotype and somatic MSH6 mutations in recurrent human malignant gliomas after alkylator chemotherapy. Cancer Res. 2006;66:3987-91.
16. Stein GH. T98G: an anchorage-independent human tumor cell line that exhibits stationary phase G1 arrest in vitro. Journal of cellular physiology. 1979;99:43-54.
17. Said MR, Begley TJ, Oppenheim AV, Lauffenburger DA, Samson LD. Global network analysis of phenotypic effects: protein networks and toxicity modulation in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A. 2004;101:18006-11.

18. Rooney JP, George AD, Patil A, Begley U, Bessette E, Zappala MR, et al. Systems based mapping demonstrates that recovery from alkylation damage requires DNA repair, RNA processing, and translation associated networks. Genomics. 2009;93:42-51.
19. Rooney JP, Patil A, Joseph F, Endres L, Begley U, Zappala MR, et al. Cross-species Functionome analysis identifies proteins associated with DNA repair, translation and aerobic respiration as conserved modulators of UV-toxicity. Genomics. 2011;97:133-47.
20. Ruffner H, Joazeiro CA, Hemmati D, Hunter T, Verma IM. Cancer-predisposing mutations within the RING domain of BRCA1: loss of ubiquitin protein ligase activity and protection from radiation hypersensitivity. Proc Natl Acad Sci U S A. 2001;98:5134-9.
21. Thaker NG, Zhang F, McDonald PR, Shun TY, Lewen MD, Pollack IF, et al. Identification of survival genes in human glioblastoma cells by small interfering RNA screening. Mol Pharmacol. 2009;76:1246-55.
22. Tang J, Goellner EM, Wang XW, Trivedi RN, St. Croix CM, Jelezcova E, et al. Bioenergetic Metabolites Regulate Base Excision Repair-Dependent Cell Death in Response to DNA Damage. Molecular Cancer Research. 2010;8:67-79.
23. Tang JB, Svilar D, Trivedi RN, Wang XH, Goellner EM, Moore B, et al. N-methylpurine DNA glycosylase and DNA polymerase beta modulate BER inhibitor potentiation of glioma cells to temozolomide. Neuro-oncology. 2011;13:471-86.
24. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature protocols. 2009;4:44-57.
25. Dennis G, Jr., Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. 2003;4:P3.
26. Boyle EI, Weng S, Gollub J, Jin H, Botstein D, Cherry JM, et al. GO::TermFinder--open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. Bioinformatics. 2004;20:3710-5.
27. Cho RJ, Huang M, Campbell MJ, Dong H, Steinmetz L, Sapinoso L, et al. Transcriptional regulation and function during the human cell cycle. Nat Genet. 2001;27:48-54.
28. Zufferey R, Dull T, Mandel RJ, Bukovsky A, Quiroz D, Naldini L, et al. Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. J Virol. 1998;72:9873-80.
29. Zufferey R, Nagy D, Mandel RJ, Naldini L, Trono D. Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. Nat Biotechnol. 1997;15:871-5.
30. Mutamba JT, Svilar D, Prasongtanakij S, Wang XH, Lin YC, Dedon PC, et al. XRCC1 and base excision repair balance in response to nitric oxide. DNA Repair (Amst). 2011;10:1282-93.
31. Yaron A, Carmel A, Katchalski-Katzir E. Intramolecularly quenched fluorogenic substrates for hydrolytic enzymes. Anal Biochem. 1979;95:228-35.
32. Clegg RM. Fluorescence resonance energy transfer and nucleic acids. Methods Enzymol. 1992;211:353-88.
33. Trivedi RN, Almeida KH, Fornsaglio JL, Schamus S, Sobol RW. The Role of Base Excision Repair in the Sensitivity and Resistance to Temozolomide Mediated Cell Death. Cancer Res. 2005;65:6394-400.
34. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13:2498-504.
35. Begley U, Dyavaiah M, Patil A, Rooney JP, DiRenzo D, Young CM, et al. Trm9-catalyzed tRNA modifications link translation to the DNA damage response. Mol Cell. 2007;28:860-70.

36. Kitchens CA, McDonald PR, Shun TY, Pollack IF, Lazo JS. Identification of chemosensitivity nodes for vinblastine through small interfering RNA high-throughput screens. J Pharmacol Exp Ther. 2011;339:851-8.
37. Whitehurst AW, Bodemann BO, Cardenas J, Ferguson D, Girard L, Peyton M, et al. Synthetic lethal screen identification of chemosensitizer loci in cancer cells. Nature. 2007;446:815-9.
38. Madhavan S, Zenklusen JC, Kotliarov Y, Sahni H, Fine HA, Buetow K. Rembrandt: helping personalized medicine become a reality through integrative translational research. Mol Cancer Res. 2009;7:157-67.
39. NCI. REMBRANDT home page (http://rembrandt.nci.nih.gov). 2005 [cited 2007 February 9, 2012]; Available from: http://rembrandt.nci.nih.gov
40. Guerrin M, Ishigami A, Mechin MC, Nachat R, Valmary S, Sebbag M, et al. cDNA cloning, gene organization and expression analysis of human peptidylarginine deiminase type I. Biochem J. 2003;370:167-74.
41. Mechin MC, Enji M, Nachat R, Chavanas S, Charveron M, Ishida-Yamamoto A, et al. The peptidylarginine deiminases expressed in human epidermis differ in their substrate specificities and subcellular locations. Cell Mol Life Sci. 2005;62:1984-95.
42. Nachat R, Mechin MC, Takahara H, Chavanas S, Charveron M, Serre G, et al. Peptidylarginine deiminase isoforms 1-3 are expressed in the epidermis and involved in the deimination of K1 and filaggrin. J Invest Dermatol. 2005;124:384-93.
43. Li P, Yao H, Zhang Z, Li M, Luo Y, Thompson PR, et al. Regulation of p53 target gene expression by peptidylarginine deiminase 4. Mol Cell Biol. 2008;28:4745-58.
44. Tanikawa C, Ueda K, Nakagawa H, Yoshida N, Nakamura Y, Matsuda K. Regulation of protein Citrullination through p53/PADI4 network in DNA damage response. Cancer Res. 2009;69:8761-9.
45. Lee YH, Coonrod SA, Kraus WL, Jelinek MA, Stallcup MR. Regulation of coactivator complex assembly and function by protein arginine methylation and demethylimination. Proc Natl Acad Sci U S A. 2005;102:3611-6.
46. Wang Y, Wysocka J, Sayegh J, Lee YH, Perlin JR, Leonelli L, et al. Human PAD4 regulates histone arginine methylation levels via demethylimination. Science. 2004;306:279-83.
47. Miller KM, Tjeertes JV, Coates J, Legube G, Polo SE, Britton S, et al. Human HDAC1 and HDAC2 function in the DNA-damage response to promote DNA nonhomologous end-joining. Nat Struct Mol Biol. 2010;17:1144-51.
48. Bangert A, Hacker S, Cristofanon S, Debatin KM, Fulda S. Chemosensitization of glioblastoma cells by the histone deacetylase inhibitor MS275. Anti-cancer drugs. 2011;22:494-9.
49. Choy E, Chiu VK, Silletti J, Feoktistov M, Morimoto T, Michaelson D, et al. Endomembrane trafficking of ras: the CAAX motif targets proteins to the ER and Golgi. Cell. 1999;98:69-80.
50. Dai Q, Choy E, Chiu V, Romano J, Slivka SR, Steitz SA, et al. Mammalian prenylcysteine carboxyl methyltransferase is in the endoplasmic reticulum. J Biol Chem. 1998;273:15030-4.
51. Kramer K, Harrington EO, Lu Q, Bellas R, Newton J, Sheahan KL, et al. Isoprenylcysteine carboxyl methyltransferase activity modulates endothelial cell apoptosis. Mol Biol Cell. 2003;14:848-57.
52. Bergo MO, Gavino BJ, Hong C, Beigneux AP, McMahon M, Casey PJ, et al. Inactivation of Icmt inhibits transformation by oncogenic K-Ras and B-Raf. J Clin Invest. 2004;113:539-50.

53. Winter-Vann AM, Kamen BA, Bergo MO, Young SG, Melnyk S, James SJ, et al. Targeting Ras signaling through inhibition of carboxyl methylation: an unexpected property of methotrexate. Proc Natl Acad Sci U S A. 2003;100:6529-34.
54. Go ML, Leow JL, Gorla SK, Schuller AP, Wang M, Casey PJ. Amino derivatives of indole as potent inhibitors of isoprenylcysteine carboxyl methyltransferase. Journal of medicinal chemistry. 2010;53:6838-50.
55. Wang M, Hossain MS, Tan W, Coolman B, Zhou J, Liu S, et al. Inhibition of isoprenylcysteine carboxylmethyltransferase induces autophagic-dependent apoptosis and impairs tumor growth. Oncogene. 2010;29:4959-70.
56. Meetei AR, de Winter JP, Medhurst AL, Wallisch M, Waisfisz Q, van de Vrugt HJ, et al. A novel ubiquitin ligase is deficient in Fanconi anemia. Nat Genet. 2003;35:165-70.
57. Vlachostergios PJ, Patrikidou A, Daliani DD, Papandreou CN. The ubiquitin-proteasome system in cancer, a major player in DNA repair. Part 1: post-translational regulation. J Cell Mol Med. 2009;13:3006-18.
58. Smogorzewska A, Matsuoka S, Vinciguerra P, McDonald ER, 3rd, Hurov KE, Luo J, et al. Identification of the FANCI protein, a monoubiquitinated FANCD2 paralog required for DNA repair. Cell. 2007;129:289-301.
59. Kapetanaki MG, Guerrero-Santoro J, Bisi DC, Hsieh CL, Rapic-Otrin V, Levine AS. The DDB1-CUL4ADDB2 ubiquitin ligase is deficient in xeroderma pigmentosum group E and targets histone H2A at UV-damaged DNA sites. Proc Natl Acad Sci U S A. 2006;103:2588-93.
60. Guerrero-Santoro J, Kapetanaki MG, Hsieh CL, Gorbachinsky I, Levine AS, Rapic-Otrin V. The cullin 4B-based UV-damaged DNA-binding protein ligase binds to UV-damaged chromatin and ubiquitinates histone H2A. Cancer Res. 2008;68:5014-22.
61. Terato H, Masaoka A, Asagoshi K, Honsho A, Ohyama Y, Suzuki T, et al. Novel repair activities of AlkA (3-methyladenine DNA glycosylase II) and endonuclease VIII for xanthine and oxanine, guanine lesions induced by nitric oxide and nitrous acid. Nucleic Acids Res. 2002;30:4975-84.
62. Masaoka A, Terato H, Kobayashi M, Honsho A, Ohyama Y, Ide H. Enzymatic repair of 5-formyluracil. I. Excision of 5-formyluracil site-specifically incorporated into oligonucleotide substrates by alka protein (Escherichia coli 3-methyladenine DNA glycosylase II). J Biol Chem. 1999;274:25136-43.
63. Saparbaev M, Laval J. Excision of hypoxanthine from DNA containing dIMP residues by the Escherichia coli, yeast, rat, and human alkylpurine DNA glycosylases. Proc Natl Acad Sci U S A. 1994;91:5873-7.
64. Terato H, Masaoka A, Kobayashi M, Fukushima S, Ohyama Y, Yoshida M, et al. Enzymatic repair of 5-formyluracil. II. Mismatch formation between 5-formyluracil and guanine during dna replication and its recognition by two proteins involved in base excision repair (AlkA) and mismatch repair (MutS). J Biol Chem. 1999;274:25144-50.
65. Hardeland U, Steinacher R, Jiricny J, Schar P. Modification of the human thymine-DNA glycosylase by ubiquitin-like proteins facilitates enzymatic turnover. EMBO Journal. 2002;21:1456-64.
66. Fischer JA, Caradonna S. Analysis of nuclear uracil-DNA glycosylase (nUDG) turnover during the cell cycle. Methods Mol Biol. 2011;761:137-49.
67. Fischer JA, Muller-Weeks S, Caradonna S. Proteolytic degradation of the nuclear isoform of uracil-DNA glycosylase occurs during the S phase of the cell cycle. DNA Repair (Amst). 2004;3:505-13.

68. Griffiths LM, Swartzlander D, Meadows KL, Wilkinson KD, Corbett AH, Doetsch PW. Dynamic compartmentalization of base excision repair proteins in response to nuclear and mitochondrial oxidative stress. Mol Cell Biol. 2009;29:794-807.
69. Lazarou M, Thorburn DR, Ryan MT, McKenzie M. Assembly of mitochondrial complex I and defects in disease. Biochim Biophys Acta. 2009;1793:78-88.
70. Rosenquist TA, Zaika E, Fernandes AS, Zharkov DO, Miller H, Grollman AP. The novel DNA glycosylase, NEIL1, protects mammalian cells from radiation-mediated cell death. DNA Repair. 2003;2:581-91.
71. Mori H, Ouchida R, Hijikata A, Kitamura H, Ohara O, Li Y, et al. Deficiency of the oxidative damage-specific DNA glycosylase NEIL1 leads to reduced germinal center B cell expansion. DNA Repair (Amst). 2009;8:1328-32.
72. Begley TJ, Rosenbach AS, Ideker T, Samson LD. Damage recovery pathways in Saccharomyces cerevisiae revealed by genomic phenotyping and interactome mapping. Mol Cancer Res. 2002;1:103-12.
73. Infanger DW, Sharma RV, Davisson RL. NADPH oxidases of the brain: distribution, regulation, and function. Antioxid Redox Signal. 2006;8:1583-96.
74. Chiu LY, Ho FM, Shiah SG, Chang Y, Lin WW. Oxidative stress initiates DNA damager MNNG-induced poly(ADP-ribose)polymerase-1-dependent parthanatos cell death. Biochemical pharmacology. 2011;81:459-70.
75. Sobol RW, Kartalou M, Almeida KH, Joyce DF, Engelward BP, Horton JK, et al. Base Excision Repair Intermediates Induce p53-independent Cytotoxic and Genotoxic Responses. J Biol Chem. 2003;278:39951-9.
76. Elder RH, Jansen JG, Weeks RJ, Willington MA, Deans B, Watson AJ, et al. Alkylpurine-DNA-N-glycosylase knockout mice show increased susceptibility to induction of mutations by methyl methanesulfonate. Molecular and Cellular Biology. 1998;18:5828-37.
77. Roth RB, Samson LD. 3-Methyladenine DNA glycosylase-deficient Aag null mice display unexpected bone marrow alkylation resistance. Cancer Res. 2002;62:656-60.
78. Meira LB, Moroski-Erkul CA, Green SL, Calvo JA, Bronson RT, Shah D, et al. Aag-initiated base excision repair drives alkylation-induced retinal degeneration in mice. Proc Natl Acad Sci U S A. 2009;106:888-93.
79. Trivedi RN, Wang XH, Jelezcova E, Goellner EM, Tang J, Sobol RW. Human methyl purine DNA glycosylase and DNA polymerase ß expression collectively predict sensitivity to temozolomide. Molecular Pharmacology. 2008;74:505-16.
80. Fry RC, Svensson JP, Valiathan C, Wang E, Hogan BJ, Bhattacharya S, et al. Genomic predictors of interindividual differences in response to DNA damaging agents. Genes Dev. 2008;22:2621-6.
81. Quiros S, Roos WP, Kaina B. Processing of O6-methylguanine into DNA double-strand breaks requires two rounds of replication whereas apoptosis is also induced in subsequent cell cycles. Cell Cycle. 2010;9:168-78.
82. Sobol RW. For MutY, it's all about the OG. Chem Biol. 2012;In Press.
83. Brinkmeyer MK, Pope MA, David SS. Catalytic Contributions of Key Residues in the Adenine Glycosylase MutY Revealed by pH-dependent Kinetics and Cellular Repair Assays. Chem Biol. 2012;19:276-86.
84. Ahn J, Vu T, Novince Z, Guerrero-Santoro J, Rapic-Otrin V, Gronenborn AM. HIV-1 Vpr loads uracil DNA glycosylase-2 onto DCAF1, a substrate recognition subunit of a cullin 4A-ring E3 ubiquitin ligase for proteasome-dependent degradation. J Biol Chem. 2010;285:37333-41.

85. Hagen L, Kavli B, Sousa MM, Torseth K, Liabakk NB, Sundheim O, et al. Cell cycle-specific UNG2 phosphorylations regulate protein turnover, activity and association with RPA. Embo J. 2008;27:51-61.

Figure Legends

Figure 1. Analysis with NIH DAVID, Princeton GO term finder and Ingenuity Pathway Analysis yield similar gene enrichments, biological processes and pathways. The high confidence hit list was also analyzed by IPA for biological pathway enrichment. One highly significant network generated by Ingenuity Pathway Analysis containing similar biological processes with other analysis methods is shown. The network is composed of many overlapping genes with genes in the high confidence hit list colored green. The top functions in the network are nucleic acid metabolism, small molecule biochemistry and metabolic disease with a network score of 24. The network score is the negative logarithm of the p-value.

Figure 2. Glioma cells with shRNA knockdown of DNA glycosylases are more sensitive to the clinical alkylator TMZ. **(A)** Quantification of DNA glycosylase mRNA knockdown in cells as determined by qRT-PCR. TaqMan probes were used to quantify mRNA levels on an Applied Biosystems StepOnePlus machine. The qRT-PCR data was analyzed using the $_{\Delta\Delta}C_t$ method and was normalized to GFP infected plate controls. Gene expression of each gene was normalized to the expression of human ß-actin. The mean of three independent experiments is plotted ± SEM. **(B)** Validation of TMZ sensitization with knockdown of specific DNA glycosylases T98G DNA glycoylase knockdown cell line sensitivity to TMZ was determined by an MTS assay 48 hr after exposure to 1 mM TMZ. The viability ratio is double normalized to account for both vehicle treated shRNA mediated growth defects and toxicity of control cells to TMZ. The mean viability ratio of three independent experiments is plotted ± SEM. **(C)** Knockdown of UNG mRNA corresponds with decreased UNG protein levels. Control and UNG-KD cell line nuclear extracts were resolved in a 4-20% SDS/PAGE gel and immunoblotted for UNG. The blot was stripped and re-probed for PCNA, which was used as a loading control. **(D)** Molecular beacon model for

real-time detection of uracil removal at 37°C in nuclear extracts. **(E)** Knockdown of UNG abolishes ability to remove uracil from DNA as determined by DNA Glycosylase Molecular Beacon Activity Assay. Specific activity of DNA glycosylase activity for the uracil lesion was measured in nuclear extracts from T98G-SCR (control-beacon, red circles; uracil-beacon, blue squares) and T98G/UNG-KD (control-beacon, green diamonds; uracil-beacon, yellow triangles). The mean fluorescence response unit of three experiments is plotted ± SEM.

Figure 3. Biological processes necessary for survival after alkylation damage are conserved and when compiled generate a cross-species functionome. The conserved biological processes with their corresponding genes were visualized using Cytoscape. Red nodes denote conserved biological processes. The remaining nodes are alkylation-modulating proteins belonging to *E. coli* (dark blue), *S. cerevisiae* (light blue) and humans (light green). The edges connect specific biological processes to each protein belonging to the processes. Many proteins are involved in more than one biological process to affect alkylation exposure survival and therefore are connected to several biological processes nodes.

Figure 4. Knockout or depletion of protein modification genes sensitizes *S. cerevisiae* and human glioma cells to alkylating agents. **(A)** A dilution series of yeast single and double knockouts were tested in YPD and YPD + MMS containing media. **(B)** Knockdown of protein modification genes sensitize glioma cells to TMZ. The T98G cell lines with protein modification gene knockdown were treated with 1mm TMZ to determine sensitivity by an MTS assay 48 hours after exposure. The mean viability ratio of three independent experiments is plotted ± SEM. **(C)** Knockdown of UBE3B sensitizes glioma cells to TMZ at clinically achievable doses.

The T98G/UBE3B-KD cell line was treated with 50μM of TMZ and survival was determined by CyQuant assay 9 days after exposure. The percent survival of two independent experiments is plotted ± SEM**(D)** Sensitivity to TMZ via knockdown of UNG and UBE3B are not independentT98G cells transiently transfected with siRNA that was scrambled, targeting UNG, UBE3B or both UNG and UBE3B. Forty-eight hours after transfection cells were seeded to determine sensitivity to TMZ by MTS assay. Both individual knockdowns of UNG and UBE3B sensitized glioma cells to TMZ, but dual knockdown did not have a greater effect than either single knockdown, suggesting an epistatic relationship between UNG and UBE3B.

Figure 5. UBE3B and UNG expression fluctuate in cancer cell lines. **(A)** UBE3B is differentially expressed with a seven-fold difference in cancer cell lines. Relative mRNA expression of UBE3B in cells various cancer cell lines as determined by qRT-PCR. Gene expression of UBE3B was normalized to T98G (gray bar) across cell lines and normalized to expression of human ß-actin within each cell line. The relative expression of UBE3B in the T98G cell line is shown as a dashed red line. The mean of three independent experiments is plotted ± SEM. **(B)** UNG is expressed at higher levels in brain tumor cell lines, when compared to non-tumor controls. Quantification of UNG and UBE3B mRNA by qRT-PCR compared to human astrocyte controls. Gene expression of each gene was normalized to the expression of human ß-actin. The mean of three independent experiments is plotted ± SEM.

Figure 6. Graphical depiction of the Alkylation Functionome. The “Alkylation Functionome” is a gene list that encompasses a diverse collection of similar biological processes crucial to survival of human, bacteria and yeast cells following alkylation damage.

Supporting information

Supplemental Figure 1

Supplemental Figure 2

Supplemental Figure 3

Supplemental Figure 4

Supplemental Table S1

Datasets S1-S6:

Dataset 1
Genes of interest from the siRNA screen that satisfied both selection criteria: a p-value less than 0.05 and a viability ratio in the lowest 5%

Dataset 2 - Second group of genes from the siRNA screen analyzed for gene enrichment using selection criteria of a p-value less than 0.05 and a viability ratio in the lowest 2.5%

Dataset 3 - Third group of genes from the siRNA screen analyzed for gene enrichment using selection criteria of a p-value less than 0.05 and a viability ratio in the lowest 7.5%

Dataset 4 - Fourth group of genes from the siRNA screen analyzed for gene enrichment using selection criteria of a p-value less than 0.05 and a viability ratio in the lowest 10%

Dataset 5 - Enriched genes identified by both NIH DAVID and Princeton GO Term Finder

Dataset 6 - Enriched biological processes and associated genes as determined by Princeton GO Term Finder

Figure 1

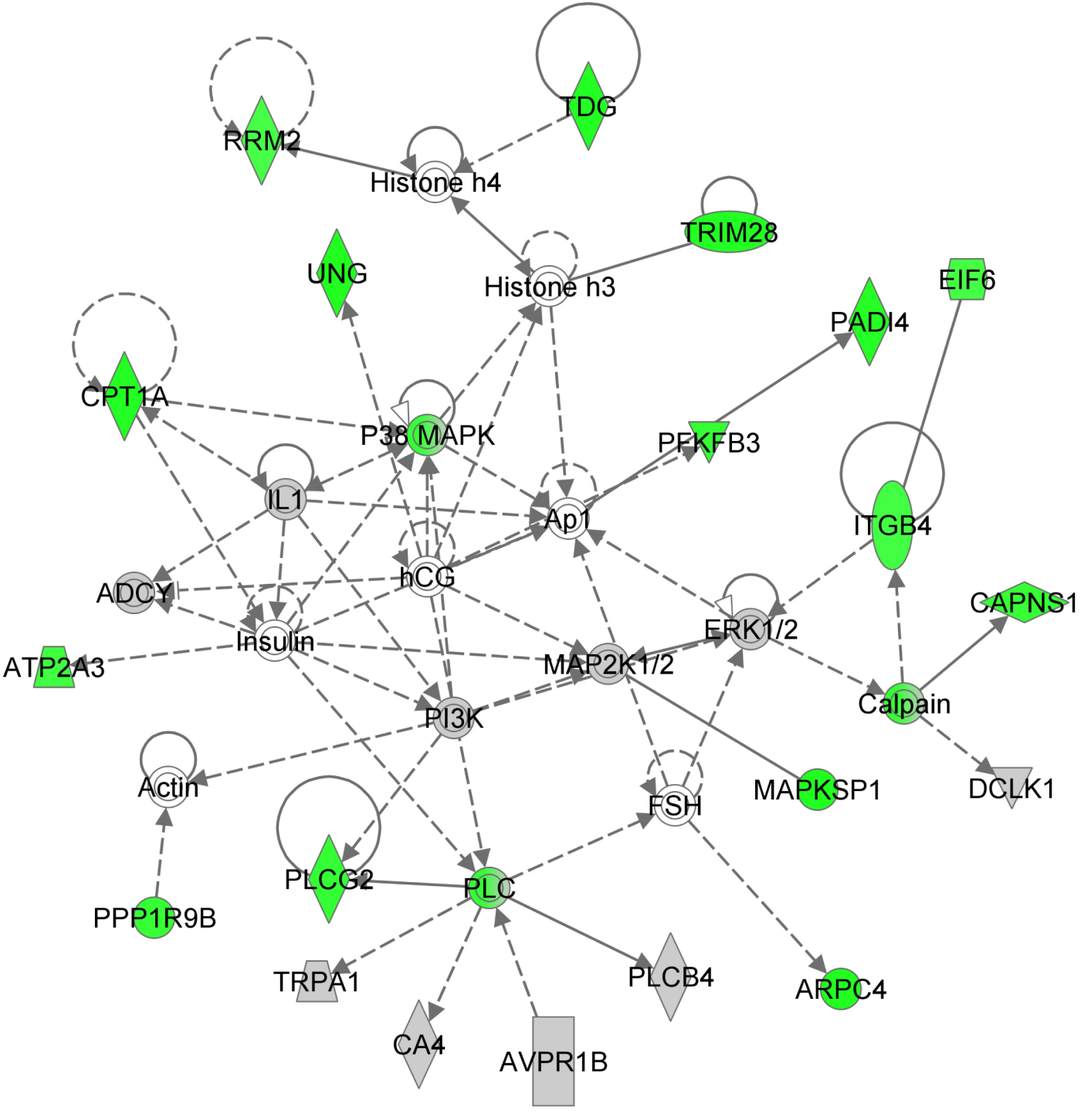
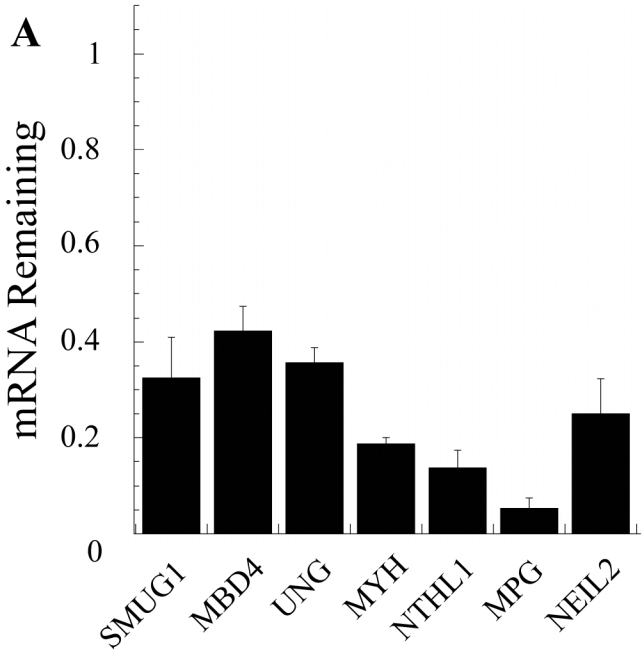
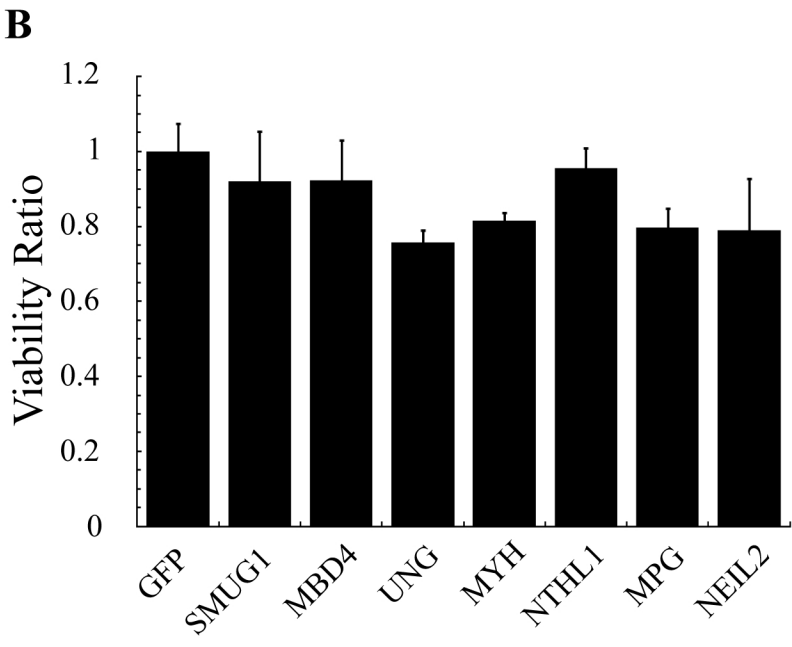
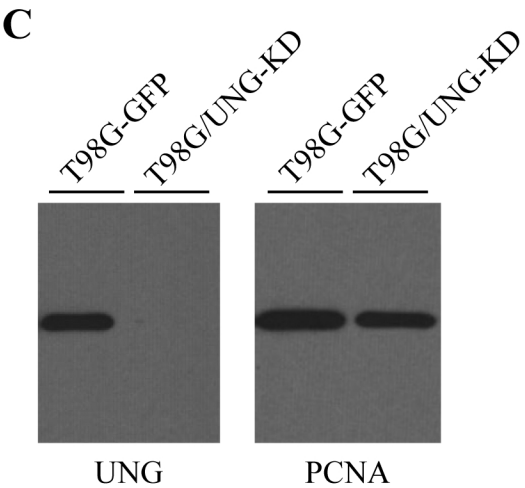
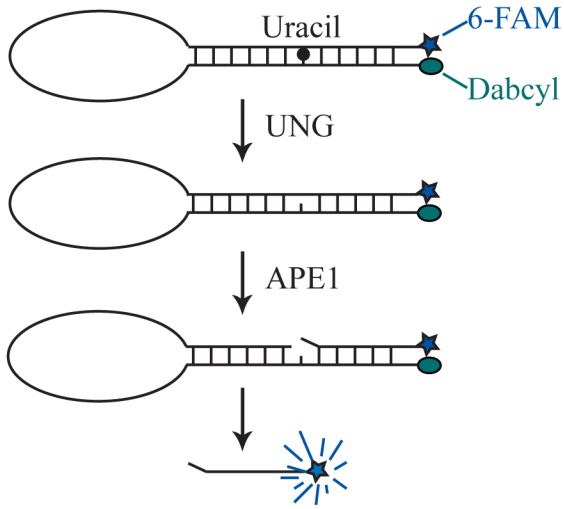

TDG
RRM2
Histone h4
TRIM28
UNG
Histone h3
EIF6
PADI4
CPT1A
P38 MAPK
PFKFB3
IL1
Ap1
ITGB4
ADCY
hCG
CAPNS1
Insulin
ERK1/2
ATP2A3
MAP2K1/2
Calpain
PI3K
Actin
MAPKSP1
DCLK1
FSH
PLCG2
PLC
PPP1R9B
TRPA1
PLCB4
ARPC4
CA4
AVPR1B

Figure 2


A
mRNA Remaining
1
0.8
0.6
0.4
0.2
0
SMUG1
MBD4
UNG
MYH
NTHL1
MPG
NEIL2


B
Viability Ratio
1.2
1
0.8
0.6
0.4
0.2
0
GFP
SMUG1
MBD4
UNG
MYH
NTHL1
MPG
NEIL2


C
T98G-GFP
T98G/UNG-KD
T98G-GFP
T98G/UNG-KD
UNG
PCNA


D
Uracil
6-FAM
Dabcyl
UNG
APE1

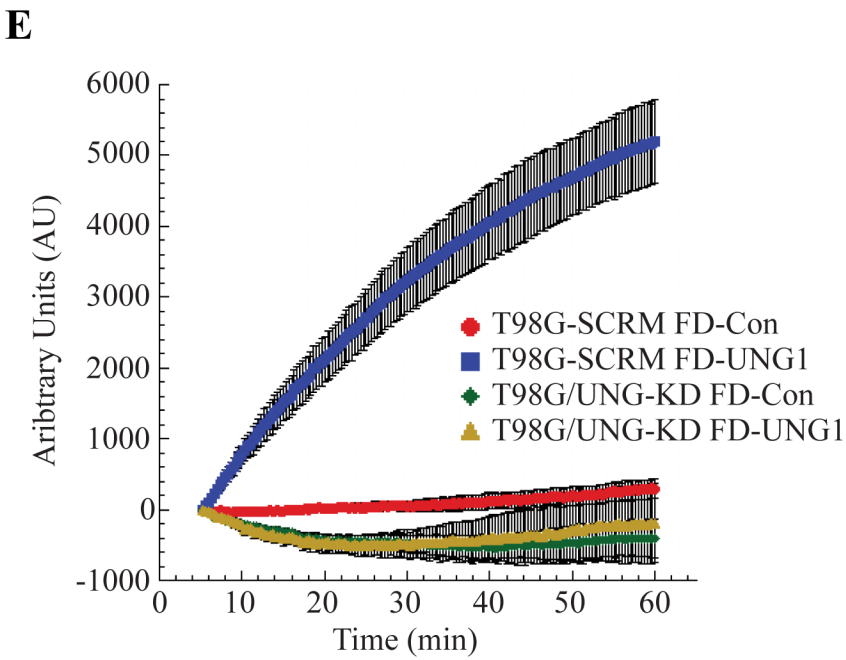

E
Aribtrary Units (AU)
6000
5000
4000
3000
2000
1000
0
-1000
0
10
20
30
40
50
60
Time (min)
T98G-SCRM FD-Con
T98G-SCRM FD-UNG1
T98G/UNG-KD FD-Con
T98G/UNG-KD FD-UNG1

Figure 3

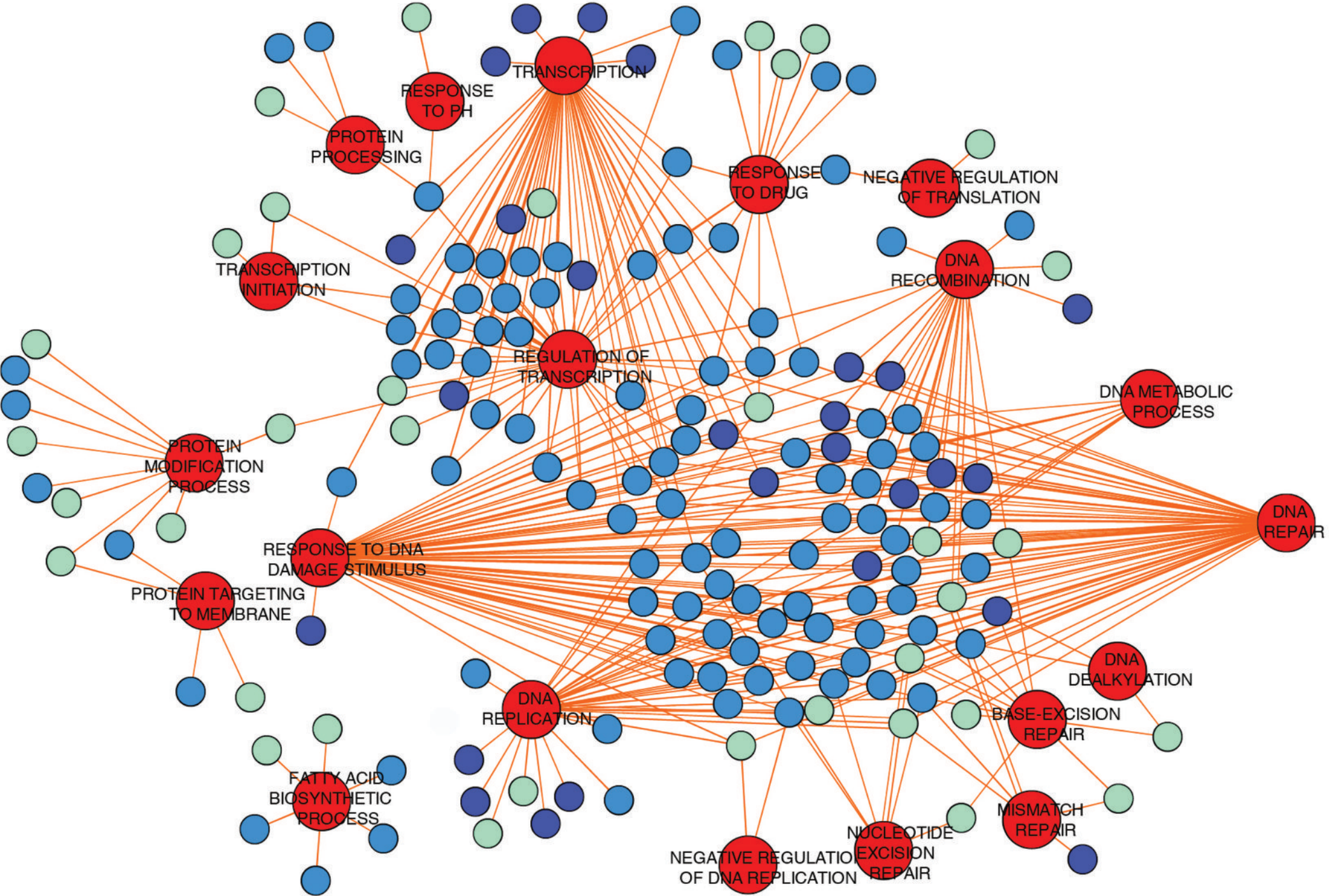

TRANSCRIPTION
RESPONSE TO PH
PROTEIN PROCESSING
RESPONSE TO DRUG
NEGATIVE REGULATION OF TRANSLATION
TRANSCRIPTION INITIATION
DNA RECOMBINATION
REGULATION OF TRANSCRIPTION
DNA METABOLIC PROCESS
PROTEIN MODIFICATION PROCESS
DNA REPAIR
RESPONSE TO DNA DAMAGE STIMULUS
PROTEIN TARGETING TO MEMBRANE
DNA DEALKYLATION
DNA REPLICATION
BASE-EXCISION REPAIR
FATTY ACID BIOSYNTHETIC PROCESS
MISMATCH REPAIR
NUCLEOTIDE EXCISION REPAIR
NEGATIVE REGULATION OF DNA REPLICATION

Figure 4

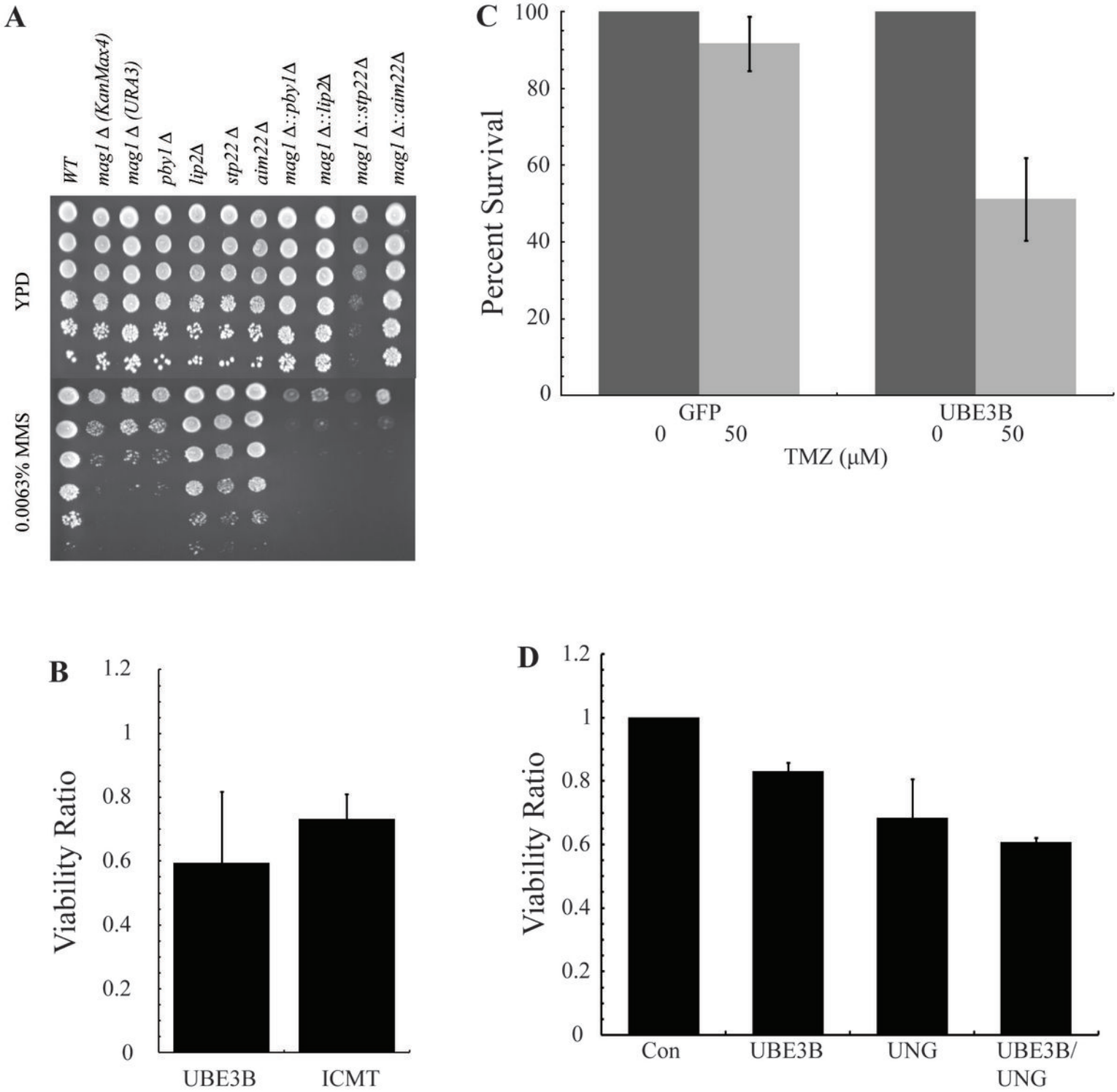

A
WT
mag1 Δ (KanMax4)
mag1 Δ (URA3)
pby1 Δ
lip2Δ
stp22 Δ
aim22 Δ
mag1 Δ::pby1Δ
mag1 Δ::lip2Δ
mag1 Δ::stp22Δ
mag1 Δ::aim22Δ
YPD
0.0063% MMS
C
Percent Survival
100
80
60
40
20
0
GFP
UBE3B
0
50
0
50
TMZ (μM)
B
Viability Ratio
1.2
1
0.8
0.6
0.4
0.2
0
UBE3B
ICMT
D
Viability Ratio
1.2
1
0.8
0.6
0.4
0.2
0
Con
UBE3B
UNG
UBE3B/
UNG

Figure 5

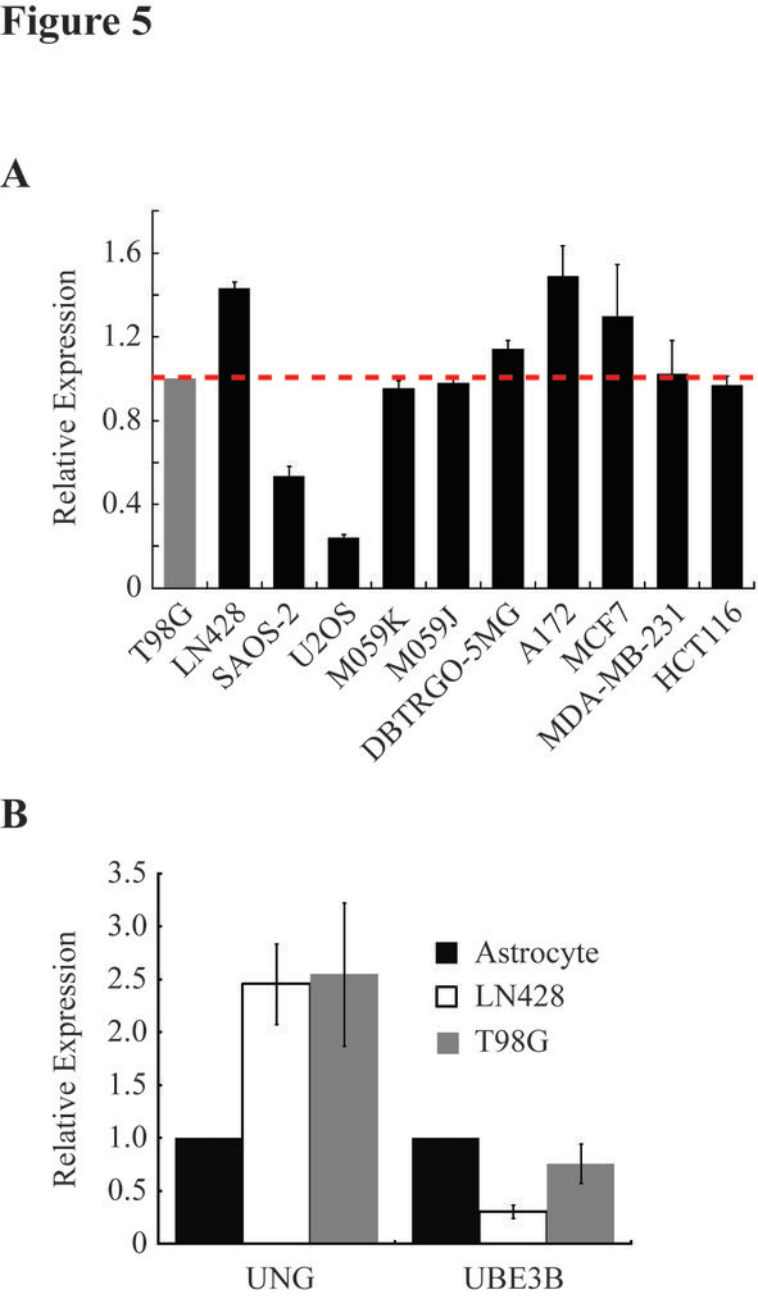

A
Relative Expression
1.6
1.2
0.8
0.4
0
T98G
LN428
SAOS-2
U2OS
M059K
M059J
DBTRGO-5MG
A172
MCF7
MDA-MB-231
HCT116
B
Relative Expression
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
Astrocyte
LN428
T98G
UNG
UBE3B

Figure 6

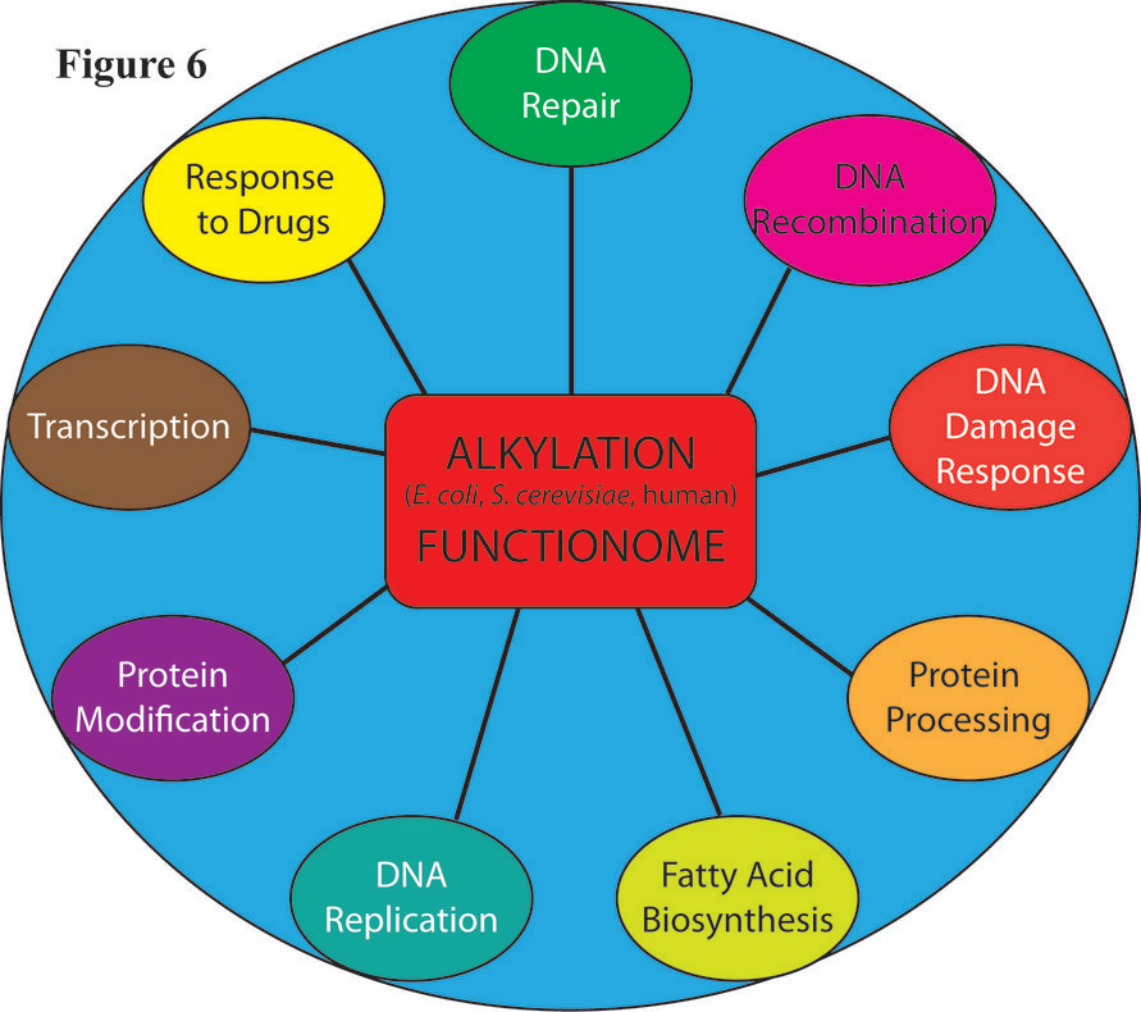

DNA Repair
Response to Drugs
DNA Recombination
Transcription
ALKYLATION
(E. coli, S. cerevisiae, human)
FUNCTIONOME
DNA Damage Response
Protein Modification
Protein Processing
DNA Replication
Fatty Acid Biosynthesis

Table 1. Top gene enrichment groups determined by NIH DAVID

Gene Ontology Term	p-value	Genes*	Fold Enrichment	FDR
Base excision repair	2.92E-05	OGG1, NEIL1, POLD1, TDG, UNG	14.7	0.054348
Hydrolase activity, Hydrolyzing N-glycosyl compounds	4.71E-05	CD38, OGG1, TDG, UNG, NEIL1	13.48214	0.083124
Glycosidase	9.54E-05	SPAM1, SMPDL3A, OGG1, TDG, TREH, CTBS, UNG, NEIL1	5.992216	0.148149
DNA N-glycosylase activity	1.07E-04	OGG1, TDG, UNG, NEIL1	17.47685	0.188687
DNA repair	1.73E-04	LIG4, OGG1, TDG, ATR, TREX1, LIG1, UNG, NEIL1	5.522238	0.269107
DNA damage	1.99E-04	LIG4, OGG1, TDG, ATR, TREX1, LIG1, UNG, NEIL1	5.416041	0.309503

*CCNO is incorrectly labeled in NIH DAVID as UNG2. UNG2 is a DNA glycosylase so it originally appeared in our analysis. However, CCNO does not have a similar function to UNG2 and was removed from further analysis.

Table 2: Sensitizing DNA glycosylases and their known substrates

Gene Symbol	Gene Name	Known Substrate*	Screen Viability Ratio
OGG1	8-oxoguanine DNA glycosylase	8-oxoG:C=T=G; me-FapyG:C; FapyG:C; 8-oxoA:C; urea	0.648
UNG	Uracil DNA glycosylase	ssU; U:G; U:A; 5-fluorouracil; 5,6-Dihydroxy-U:G; 5-OH-U:G; Isodialuric acid; Alloxan	0.446
TDG	Thymine DNA glycosylase	U:G; T:G and ethenoC:G; 5-Fluorouracil; 5-fluorouracil (ss); 5-Hydroxymethyluracil; hypoxanthine:G; 5-bromouracil; εC:A Tg:G; 5-formyl-U	0.541
NEIL1	Nei endonuclease VIII-like 1 (E. coli)	TgG; 5-OH-C; 5-OH-U:AT>G; Guanidinohydantoin; guanidinohydantoin (ss); Iminoallantoin; Iminoallantoin (ss); Spiroiminodihydantoin; Spiroiminodihydantoin (ss); 5,6-Dihydro-T; 5,6-Dihydro-U:G=C=A>T; FapyG:C; 8-Oxo-G:C=G>T>A; FapyA:T; (5'R)-8,5'-Cyclo-2'-deoxyadenosine; (5'S)-8,5'-Cyclo-2'-deoxyadenosine; 8-Oxo-A:C	0.559
*For mismatched base pairs, the repaired or removed base is configured on the left in each case.			

BIBLIOGRAPHY

1. Jemal A, Siegel R, Xu J, Ward E. Cancer Statistics, 2010. CA Cancer J Clin. 2010;60(5):277-300.
2. Wen PY, Kesari S. Malignant gliomas in adults. N Engl J Med. 2008;359(5):492-507.
3. Cohen MH, Johnson JR, Pazdur R. Food and Drug Administration Drug approval summary: temozolomide plus radiation therapy for the treatment of newly diagnosed glioblastoma multiforme. Clin Cancer Res. 2005;11(19 Pt 1):6767-71.
4. Stupp R, Hegi ME, Mason WP, van den Bent MJ, Taphoorn MJ, Janzer RC, et al. Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. Lancet Oncol. 2009;10(5):459-66.
5. Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJ, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. New England Journal of Medicine. 2005;352(10):987-96.
6. Holland EC. Glioblastoma multiforme: the terminator. Proc Natl Acad Sci U S A. 2000;97(12):6242-4. PMCID: 33993.
7. Drablos F, Feyzi E, Aas PA, Vaagbo CB, Kavli B, Bratlie MS, et al. Alkylation damage in DNA and RNA--repair mechanisms and medical significance. DNA Repair (Amst). 2004;3(11):1389-407.
8. Stevens MF, Hickman JA, Langdon SP, Chubb D, Vickers L, Stone R, et al. Antitumor activity and pharmacokinetics in mice of 8-carbamoyl-3-methyl-imidazo[5,1-d]-1,2,3,5-tetrazin-4(3H)-one (CCRG 81045; M & B 39831), a novel drug with potential as an alternative to dacarbazine. Cancer Res. 1987;47(22):5846-52.
9. Saleem A, Brown GD, Brady F, Aboagye EO, Osman S, Luthra SK, et al. Metabolic activation of temozolomide measured in vivo using positron emission tomography. Cancer Res. 2003;63(10):2409-15.
10. Marchesi F, Turriziani M, Tortorelli G, Avvisati G, Torino F, De Vecchis L. Triazene compounds: mechanism of action and related DNA repair systems. Pharmacol Res. 2007;56(4):275-87.
11. Baker SD, Wirth M, Statkevich P, Reidenberg P, Alton K, Sartorius SE, et al. Absorption, metabolism, and excretion of 14C-temozolomide following oral administration to patients with advanced cancer. Clin Cancer Res. 1999;5(2):309-17.
12. Marosi C. Complications of chemotherapy in neuro-oncology. Handb Clin Neurol. 2012;105:873-85.
13. Sobol RW. Temozolomide. In: Schwab M, editor. Encyclopedia of Cancer. 2nd ed. Berlin, Heidelberg, New York: Springer; 2009.

14. Beranek DT. Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents. Mutation Research. 1990;231(1):11-30.
15. Svilar D, Goellner EM, Almeida KH, Sobol RW. Base excision repair and lesion-dependent subpathways for repair of oxidative DNA damage. Antioxid Redox Signal. 2011;14(12):2491-507. PMCID: 3096496.
16. Gerson SL. MGMT: its role in cancer aetiology and cancer therapeutics. Nat Rev Cancer. 2004;4(4):296-307.
17. Yip S, Miao J, Cahill DP, Iafrate AJ, Aldape K, Nutt CL, et al. MSH6 mutations arise in glioblastomas during temozolomide therapy and mediate temozolomide resistance. Clin Cancer Res. 2009;15(14):4622-9.
18. O'Brien V, Brown R. Signalling cell cycle arrest and cell death through the MMR System. Carcinogenesis. 2006;27(4):682-92.
19. Cahill DP, Levine KK, Betensky RA, Codd PJ, Romany CA, Reavie LB, et al. Loss of the mismatch repair protein MSH6 in human glioblastomas is associated with tumor progression during temozolomide treatment. Clin Cancer Res. 2007;13(7):2038-45.
20. Sarkaria JN, Kitange GJ, James CD, Plummer R, Calvert H, Weller M, et al. Mechanisms of chemoresistance to alkylating agents in malignant glioma. Clin Cancer Res. 2008;14(10):2900-8.
21. McMurry TB. MGMT inhibitors--The Trinity College-Paterson Institute experience, a chemist's perception. DNA Repair (Amst). 2007;6(8):1161-9.
22. Ranson M, Hersey P, Thompson D, Beith J, McArthur GA, Haydon A, et al. Randomized trial of the combination of lomeguatrib and temozolomide compared with temozolomide alone in chemotherapy naive patients with metastatic cutaneous melanoma. J Clin Oncol. 2007;25(18):2540-5.
23. Hassel JC, Sucker A, Edler L, Kurzen H, Moll I, Stresemann C, et al. MGMT gene promoter methylation correlates with tolerance of temozolomide treatment in melanoma but not with clinical outcome. Br J Cancer. 2010.
24. Tentori L, Graziani G. Recent approaches to improve the antitumor efficacy of temozolomide. Curr Med Chem. 2009;16(2):245-57.
25. Quinn JA, Jiang SX, Reardon DA, Desjardins A, Vredenburgh JJ, Rich JN, et al. Phase II trial of temozolomide plus o6-benzylguanine in adults with recurrent, temozolomide-resistant malignant glioma. J Clin Oncol. 2009;27(8):1262-7. PMCID: 2667825.
26. Almeida KH, Sobol RW. A unified view of base excision repair: lesion-dependent protein complexes regulated by post-translational modification. DNA Repair. 2007;6(6):695-711.
27. Friedberg EC, Walker GC, Siede W, Wood RD, Schultz RA, Ellenberger T. DNA Repair and Mutagenesis, 2nd Edition. Washington, D.C.: ASM Press; 2006.
28. Wood RD, Mitchell M, Sgouros J, Lindahl T. Human DNA repair genes. Science. 2001;291(5507):1284-9.
29. Simsek D, Furda A, Gao Y, Artus J, Brunet E, Hadjantonakis AK, et al. Crucial role for DNA ligase III in mitochondria but not in Xrcc1-dependent repair. Nature. 2011;471(7337):245-8. PMCID: 3261757.
30. Gao Y, Katyal S, Lee Y, Zhao J, Rehg JE, Russell HR, et al. DNA ligase III is critical for mtDNA integrity but not Xrcc1-mediated nuclear DNA repair. Nature. 2011;471(7337):240-4. PMCID: 3079429.

31. Krokan HE, Drablos F, Slupphaug G. Uracil in DNA--occurrence, consequences and repair. Oncogene. 2002;21(58):8935-48.
32. Almeida KH, Sobol RW. Increased Specificity and Efficiency of Base Excision Repair through Complex Formation. In: Siede W, Doetsch PW, Kow YW, editors. DNA Damage Recognition. New York: Marcel Dekker Inc.; 2005. p. 33-64.
33. Jensen A, Calvayrac G, Karahalil B, Bohr VA, Stevnsner T. Mammalian 8-oxoguanine DNA glycosylase 1 incises 8-oxoadenine opposite cytosine in nuclei and mitochondria, while a different glycosylase incises 8-oxoadenine opposite guanine in nuclei. J Biol Chem. 2003;278(21):19541-8.
34. Bjelland S, Seeberg E. Mutagenicity, toxicity and repair of DNA base damage induced by oxidation. Mutat Res. 2003;531(1-2):37-80.
35. Matsumoto Y, Zhang QM, Takao M, Yasui A, Yonei S. Escherichia coli Nth and human hNTH1 DNA glycosylases are involved in removal of 8-oxoguanine from 8-oxoguanine/guanine mispairs in DNA. Nucleic Acids Res. 2001;29(9):1975-81. PMCID: 37258.
36. Liu M, Bandaru V, Bond JP, Jaruga P, Zhao X, Christov PP, et al. The mouse ortholog of NEIL3 is a functional DNA glycosylase in vitro and in vivo. Proc Natl Acad Sci U S A. 2010;107(11):4925-30. PMCID: 2841873.
37. Dizdaroglu M, Karakaya A, Jaruga P, Slupphaug G, Krokan H. Novel activities of human uracil DNA N-glycosylase for cytosine-derived products of oxidative DNA damage. Nucl Acids Res. 1996;24(3):418-22.
38. Darwanto A, Theruvathu JA, Sowers JL, Rogstad DK, Pascal T, Goddard W, 3rd, et al. Mechanisms of base selection by human single-stranded selective monofunctional uracil-DNA glycosylase. J Biol Chem. 2009;284(23):15835-46. PMCID: 2708880.
39. Wibley JE, Waters TR, Haushalter K, Verdine GL, Pearl LH. Structure and specificity of the vertebrate anti-mutator uracil-DNA glycosylase SMUG1. Mol Cell. 2003;11(6):1647-59.
40. Hardeland U, Bentele M, Jiricny J, Schar P. The versatile thymine DNA-glycosylase: a comparative characterization of the human, Drosophila and fission yeast orthologs. Nucleic Acids Res. 2003;31(9):2261-71. PMCID: 154230.
41. Yoon JH, Iwai S, O'Connor TR, Pfeifer GP. Human thymine DNA glycosylase (TDG) and methyl-CpG-binding protein 4 (MBD4) excise thymine glycol (Tg) from a Tg:G mispair. Nucleic Acids Research. 2003;31(18):5399-404.
42. Lee CY, Delaney JC, Kartalou M, Lingaraju GM, Maor-Shoshani A, Essigmann JM, et al. Recognition and processing of a new repertoire of DNA substrates by human 3-methyladenine DNA glycosylase (AAG). Biochemistry. 2009;48(9):1850-61.
43. Bessho T, Roy R, Yamamoto K, Kasai H, Nishimura S, Tano K, et al. Repair of 8-hydroxyguanine in DNA by mammalian N-methylpurine-DNA glycosylase. Proc Natl Acad Sci U S A. 1993;90(19):8901-4. PMCID: 47468.
44. Parker AR, Sieber OM, Shi C, Hua L, Takao M, Tomlinson IP, et al. Cells with pathogenic biallelic mutations in the human MUTYH gene are defective in DNA damage binding and repair. Carcinogenesis. 2005;26(11):2010-8.
45. Hailer MK, Slade PG, Martin BD, Rosenquist TA, Sugden KD. Recognition of the oxidized lesions spiroiminodihydantoin and guanidinohydantoin in DNA by the mammalian base excision repair glycosylases NEIL1 and NEIL2. DNA Repair (Amst). 2005;4(1):41-50.
46. Jaruga P, Xiao Y, Vartanian V, Lloyd RS, Dizdaroglu M. Evidence for the involvement of DNA repair enzyme NEIL1 in nucleotide excision repair of (5'R)- and (5'S)-8,5'-cyclo-2'-deoxyadenosines. Biochemistry. 2010;49(6):1053-5. PMCID: 2817919.

47. Grin IR, Dianov GL, Zharkov DO. The role of mammalian NEIL1 protein in the repair of 8-oxo-7,8-dihydroadenine in DNA. FEBS Lett. 2010;584(8):1553-7.
48. Cortellino S, Xu J, Sannai M, Moore R, Caretti E, Cigliano A, et al. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. Cell. 2011;146(1):67-79. PMCID: 3230223.
49. Rai K, Huggins IJ, James SR, Karpf AR, Jones DA, Cairns BR. DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. Cell. 2008;135(7):1201-12. PMCID: 2629358.
50. Perillo B, Ombra MN, Bertoni A, Cuozzo C, Sacchetti S, Sasso A, et al. DNA oxidation as triggered by H3K9me2 demethylation drives estrogen-induced gene expression. Science. 2008;319(5860):202-6.
51. Amente S, Bertoni A, Morano A, Lania L, Avvedimento EV, Majello B. LSD1-mediated demethylation of histone H3 lysine 4 triggers Myc-induced transcription. Oncogene. 2010.
52. Amente S, Lania L, Avvedimento EV, Majello B. DNA oxidation drives Myc mediated transcription. Cell Cycle. 2010;9(15):3002-4.
53. Imai K, Slupphaug G, Lee WI, Revy P, Nonoyama S, Catalan N, et al. Human uracil-DNA glycosylase deficiency associated with profoundly impaired immunoglobulin class-switch recombination. Nat Immunol. 2003;4(10):1023-8.
54. Begum NA, Kinoshita K, Kakazu N, Muramatsu M, Nagaoka H, Shinkura R, et al. Uracil DNA glycosylase activity is dispensable for immunoglobulin class switch. Science. 2004;305(5687):1160-3.
55. Mol CD, Arvai AS, Sanderson RJ, Slupphaug G, Kavli B, Krokan HE, et al. Crystal structure of human uracil-DNA glycosylase in complex with a protein inhibitor: protein mimicry of DNA. Cell. 1995;82(5):701-8.
56. Begum NA, Stanlie A, Doi T, Sasaki Y, Jin HW, Kim YS, et al. Further evidence for involvement of a noncanonical function of uracil DNA glycosylase in class switch recombination. Proc Natl Acad Sci U S A. 2009;106(8):2752-7. PMCID: 2650371.
57. Zeitlin SG, Chapados BR, Baker NM, Tai C, Slupphaug G, Wang JY. Uracil DNA N-glycosylase promotes assembly of human centromere protein A. PLoS ONE.6(3):e17151. PMCID: 3047565.
58. Thaker NG, Zhang F, McDonald PR, Shun TY, Lewen MD, Pollack IF, et al. Identification of survival genes in human glioblastoma cells by small interfering RNA screening. Mol Pharmacol. 2009;76(6):1246-55. PMCID: 2784725.
59. Kitchens CA, McDonald PR, Shun TY, Pollack IF, Lazo JS. Identification of chemosensitivity nodes for vinblastine through small interfering RNA high-throughput screens. J Pharmacol Exp Ther. 2011;339(3):851-8. PMCID: 3226368.
60. Whitehurst AW, Bodemann BO, Cardenas J, Ferguson D, Girard L, Peyton M, et al. Synthetic lethal screen identification of chemosensitizer loci in cancer cells. Nature. 2007;446(7137):815-9.
61. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature protocols. 2009;4(1):44-57.
62. Dennis G, Jr., Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. 2003;4(5):P3.
63. Boyle EI, Weng S, Gollub J, Jin H, Botstein D, Cherry JM, et al. GO::TermFinder--open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. Bioinformatics. 2004;20(18):3710-5.

64. Cho RJ, Huang M, Campbell MJ, Dong H, Steinmetz L, Sapinoso L, et al. Transcriptional regulation and function during the human cell cycle. Nat Genet. 2001;27(1):48-54.
65. Zufferey R, Dull T, Mandel RJ, Bukovsky A, Quiroz D, Naldini L, et al. Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. J Virol. 1998;72(12):9873-80.
66. Zufferey R, Nagy D, Mandel RJ, Naldini L, Trono D. Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. Nat Biotechnol. 1997;15(9):871-5.
67. Tang J, Goellner EM, Wang XW, Trivedi RN, St. Croix CM, Jelezcova E, et al. Bioenergetic Metabolites Regulate Base Excision Repair-Dependent Cell Death in Response to DNA Damage. Molecular Cancer Research. 2010;8(1):67-79. PMCID: 2808464.
68. Yaron A, Carmel A, Katchalski-Katzir E. Intramolecularly quenched fluorogenic substrates for hydrolytic enzymes. Anal Biochem. 1979;95(1):228-35.
69. Clegg RM. Fluorescence resonance energy transfer and nucleic acids. Methods Enzymol. 1992;211:353-88.
70. Goellner EM, Grimme B, Brown AR, Lin YC, Wang XH, Sugrue KF, et al. Overcoming temozolomide resistance in glioblastoma via dual inhibition of NAD+ biosynthesis and base excision repair. Cancer Res. 2011;71(6):2308-17. PMCID: 3077901.
71. Rooney JP, Patil A, Joseph F, Endres L, Begley U, Zappala MR, et al. Cross-species Functionome analysis identifies proteins associated with DNA repair, translation and aerobic respiration as conserved modulators of UV-toxicity. Genomics. 2011;97(3):133-47. PMCID: 3053583.
72. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498-504. PMCID: 403769.
73. Begley U, Dyavaiah M, Patil A, Rooney JP, DiRenzo D, Young CM, et al. Trm9-catalyzed tRNA modifications link translation to the DNA damage response. Mol Cell. 2007;28(5):860-70. PMCID: 2211415.
74. Prasad R, Singhal RK, Srivastava DK, Molina JT, Tomkinson AE, Wilson SH. Specific interaction of DNA polymerase ß and DNA ligase I in a multiprotein base excision repair complex from bovine testis. J Biol Chem. 1996;271(27):16000-7.
75. Hunter C, Smith R, Cahill DP, Stephens P, Stevens C, Teague J, et al. A hypermutation phenotype and somatic MSH6 mutations in recurrent human malignant gliomas after alkylator chemotherapy. Cancer Res. 2006;66(8):3987-91.
76. Stein GH. T98G: an anchorage-independent human tumor cell line that exhibits stationary phase G1 arrest in vitro. Journal of cellular physiology. 1979;99(1):43-54.
77. Said MR, Begley TJ, Oppenheim AV, Lauffenburger DA, Samson LD. Global network analysis of phenotypic effects: protein networks and toxicity modulation in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A. 2004;101(52):18006-11. PMCID: 539745.
78. Rooney JP, George AD, Patil A, Begley U, Bessette E, Zappala MR, et al. Systems based mapping demonstrates that recovery from alkylation damage requires DNA repair, RNA processing, and translation associated networks. Genomics. 2009;93(1):42-51.
79. Meetei AR, de Winter JP, Medhurst AL, Wallisch M, Waisfisz Q, van de Vrugt HJ, et al. A novel ubiquitin ligase is deficient in Fanconi anemia. Nat Genet. 2003;35(2):165-70.
80. Vlachostergios PJ, Patrikidou A, Daliani DD, Papandreou CN. The ubiquitin-proteasome system in cancer, a major player in DNA repair. Part 1: post-translational regulation. J Cell Mol Med. 2009;13(9B):3006-18.

81. Smogorzewska A, Matsuoka S, Vinciguerra P, McDonald ER, 3rd, Hurov KE, Luo J, et al. Identification of the FANCI protein, a monoubiquitinated FANCD2 paralog required for DNA repair. Cell. 2007;129(2):289-301.
82. Ruffner H, Joazeiro CA, Hemmati D, Hunter T, Verma IM. Cancer-predisposing mutations within the RING domain of BRCA1: loss of ubiquitin protein ligase activity and protection from radiation hypersensitivity. Proc Natl Acad Sci U S A. 2001;98(9):5134-9. PMCID: 33176.
83. Bryant HE, Schultz N, Thomas HD, Parker KM, Flower D, Lopez E, et al. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. Nature. 2005;434(7035):913-7.
84. Farmer H, McCabe N, Lord CJ, Tutt AN, Johnson DA, Richardson TB, et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. Nature. 2005;434(7035):917-21.
85. Trachootham D, Alexandre J, Huang P. Targeting cancer cells by ROS-mediated mechanisms: a radical therapeutic approach? Nat Rev Drug Discov. 2009;8(7):579-91.
86. Krokan HE, Standal R, Slupphaug G. DNA glycosylases in the base excision repair of DNA. Biochem J. 1997;325 (Pt 1):1-16. PMCID: 1218522.
87. Tang JB, Svilar D, Trivedi RN, Wang XH, Goellner EM, Moore B, et al. N-methylpurine DNA glycosylase and DNA polymerase beta modulate BER inhibitor potentiation of glioma cells to temozolomide. Neuro-oncology. 2011;13(5):471-86. PMCID: 3093332.
88. Mutamba JT, Svilar D, Prasongtanakij S, Wang XH, Lin YC, Dedon PC, et al. XRCC1 and base excision repair balance in response to nitric oxide. DNA Repair (Amst). 2011;10(12):1282-93. PMCID: In Process.
89. Terato H, Masaoka A, Asagoshi K, Honsho A, Ohyama Y, Suzuki T, et al. Novel repair activities of AlkA (3-methyladenine DNA glycosylase II) and endonuclease VIII for xanthine and oxanine, guanine lesions induced by nitric oxide and nitrous acid. Nucleic Acids Res. 2002;30(22):4975-84. PMCID: 137176.
90. Masaoka A, Terato H, Kobayashi M, Honsho A, Ohyama Y, Ide H. Enzymatic repair of 5-formyluracil. I. Excision of 5-formyluracil site-specifically incorporated into oligonucleotide substrates by alka protein (Escherichia coli 3-methyladenine DNA glycosylase II). J Biol Chem. 1999;274(35):25136-43.
91. Saparbaev M, Laval J. Excision of hypoxanthine from DNA containing dIMP residues by the Escherichia coli, yeast, rat, and human alkylpurine DNA glycosylases. Proc Natl Acad Sci U S A. 1994;91(13):5873-7. PMCID: 44099.
92. Terato H, Masaoka A, Kobayashi M, Fukushima S, Ohyama Y, Yoshida M, et al. Enzymatic repair of 5-formyluracil. II. Mismatch formation between 5-formyluracil and guanine during dna replication and its recognition by two proteins involved in base excision repair (AlkA) and mismatch repair (MutS). J Biol Chem. 1999;274(35):25144-50.
93. Rosenquist TA, Zaika E, Fernandes AS, Zharkov DO, Miller H, Grollman AP. The novel DNA glycosylase, NEIL1, protects mammalian cells from radiation-mediated cell death. DNA Repair. 2003;2(5):581-91.
94. Mori H, Ouchida R, Hijikata A, Kitamura H, Ohara O, Li Y, et al. Deficiency of the oxidative damage-specific DNA glycosylase NEIL1 leads to reduced germinal center B cell expansion. DNA Repair (Amst). 2009;8(11):1328-32.

95. Griffiths LM, Swartzlander D, Meadows KL, Wilkinson KD, Corbett AH, Doetsch PW. Dynamic compartmentalization of base excision repair proteins in response to nuclear and mitochondrial oxidative stress. Mol Cell Biol. 2009;29(3):794-807.
96. Rowe LA, Degtyareva N, Doetsch PW. DNA damage-induced reactive oxygen species (ROS) stress response in Saccharomyces cerevisiae. Free Radical Biology and Medicine. 2008;45(8):1167-77.
97. Lazarou M, Thorburn DR, Ryan MT, McKenzie M. Assembly of mitochondrial complex I and defects in disease. Biochim Biophys Acta. 2009;1793(1):78-88.
98. Begley TJ, Rosenbach AS, Ideker T, Samson LD. Damage recovery pathways in Saccharomyces cerevisiae revealed by genomic phenotyping and interactome mapping. Mol Cancer Res. 2002;1(2):103-12.
99. Infanger DW, Sharma RV, Davisson RL. NADPH oxidases of the brain: distribution, regulation, and function. Antioxid Redox Signal. 2006;8(9-10):1583-96.
100. Chiu LY, Ho FM, Shiah SG, Chang Y, Lin WW. Oxidative stress initiates DNA damager MNNG-induced poly(ADP-ribose)polymerase-1-dependent parthanatos cell death. Biochemical pharmacology. 2010.
101. Sobol RW, Kartalou M, Almeida KH, Joyce DF, Engelward BP, Horton JK, et al. Base Excision Repair Intermediates Induce p53-independent Cytotoxic and Genotoxic Responses. J Biol Chem. 2003;278(41):39951-9.
102. Elder RH, Jansen JG, Weeks RJ, Willington MA, Deans B, Watson AJ, et al. Alkylpurine-DNA-N-glycosylase knockout mice show increased susceptibility to induction of mutations by methyl methanesulfonate. Molecular and Cellular Biology. 1998;18(10):5828-37.
103. Roth RB, Samson LD. 3-Methyladenine DNA glycosylase-deficient Aag null mice display unexpected bone marrow alkylation resistance. Cancer Res. 2002;62(3):656-60.
104. Meira LB, Moroski-Erkul CA, Green SL, Calvo JA, Bronson RT, Shah D, et al. Aag-initiated base excision repair drives alkylation-induced retinal degeneration in mice. Proc Natl Acad Sci U S A. 2009;106(3):888-93.
105. Trivedi RN, Wang XH, Jelezcova E, Goellner EM, Tang J, Sobol RW. Human methyl purine DNA glycosylase and DNA polymerase ß expression collectively predict sensitivity to temozolomide. Molecular Pharmacology. 2008;74(2):505-16.
106. Fry RC, Svensson JP, Valiathan C, Wang E, Hogan BJ, Bhattacharya S, et al. Genomic predictors of interindividual differences in response to DNA damaging agents. Genes Dev. 2008;22(19):2621-6.
107. Quiros S, Roos WP, Kaina B. Processing of O6-methylguanine into DNA double-strand breaks requires two rounds of replication whereas apoptosis is also induced in subsequent cell cycles. Cell Cycle. 2010;9(1):168-78.
108. Xie Y, Yang H, Miller JH, Shih DM, Hicks GG, Xie J, et al. Cells deficient in oxidative DNA damage repair genes Myh and Ogg1 are sensitive to oxidants with increased G2/M arrest and multinucleation. Carcinogenesis. 2008;29(4):722-8.
109. Hardeland U, Steinacher R, Jiricny J, Schar P. Modification of the human thymine-DNA glycosylase by ubiquitin-like proteins facilitates enzymatic turnover. EMBO Journal. 2002;21(6):1456-64.
110. Fischer JA, Caradonna S. Analysis of nuclear uracil-DNA glycosylase (nUDG) turnover during the cell cycle. Methods Mol Biol. 2011;761:137-49.

111. Fischer JA, Muller-Weeks S, Caradonna S. Proteolytic degradation of the nuclear isoform of uracil-DNA glycosylase occurs during the S phase of the cell cycle. DNA Repair (Amst). 2004;3(5):505-13.
112. Guerrin M, Ishigami A, Mechin MC, Nachat R, Valmary S, Sebbag M, et al. cDNA cloning, gene organization and expression analysis of human peptidylarginine deiminase type I. Biochem J. 2003;370(Pt 1):167-74. PMCID: 1223146.
113. Mechin MC, Enji M, Nachat R, Chavanas S, Charveron M, Ishida-Yamamoto A, et al. The peptidylarginine deiminases expressed in human epidermis differ in their substrate specificities and subcellular locations. Cell Mol Life Sci. 2005;62(17):1984-95.
114. Nachat R, Mechin MC, Takahara H, Chavanas S, Charveron M, Serre G, et al. Peptidylarginine deiminase isoforms 1-3 are expressed in the epidermis and involved in the deimination of K1 and filaggrin. J Invest Dermatol. 2005;124(2):384-93.
115. Li P, Yao H, Zhang Z, Li M, Luo Y, Thompson PR, et al. Regulation of p53 target gene expression by peptidylarginine deiminase 4. Mol Cell Biol. 2008;28(15):4745-58. PMCID: 2493360.
116. Tanikawa C, Ueda K, Nakagawa H, Yoshida N, Nakamura Y, Matsuda K. Regulation of protein Citrullination through p53/PADI4 network in DNA damage response. Cancer Res. 2009;69(22):8761-9.
117. Lee YH, Coonrod SA, Kraus WL, Jelinek MA, Stallcup MR. Regulation of coactivator complex assembly and function by protein arginine methylation and demethylimination. Proc Natl Acad Sci U S A. 2005;102(10):3611-6. PMCID: 553305.
118. Wang Y, Wysocka J, Sayegh J, Lee YH, Perlin JR, Leonelli L, et al. Human PAD4 regulates histone arginine methylation levels via demethylimination. Science. 2004;306(5694):279-83.
119. Miller KM, Tjeertes JV, Coates J, Legube G, Polo SE, Britton S, et al. Human HDAC1 and HDAC2 function in the DNA-damage response to promote DNA nonhomologous end-joining. Nat Struct Mol Biol. 2010;17(9):1144-51. PMCID: 3018776.
120. Bangert A, Hacker S, Cristofanon S, Debatin KM, Fulda S. Chemosensitization of glioblastoma cells by the histone deacetylase inhibitor MS275. Anti-cancer drugs. 2011;22(6):494-9.
121. Choy E, Chiu VK, Silletti J, Feoktistov M, Morimoto T, Michaelson D, et al. Endomembrane trafficking of ras: the CAAX motif targets proteins to the ER and Golgi. Cell. 1999;98(1):69-80.
122. Dai Q, Choy E, Chiu V, Romano J, Slivka SR, Steitz SA, et al. Mammalian prenylcysteine carboxyl methyltransferase is in the endoplasmic reticulum. J Biol Chem. 1998;273(24):15030-4.
123. Kramer K, Harrington EO, Lu Q, Bellas R, Newton J, Sheahan KL, et al. Isoprenylcysteine carboxyl methyltransferase activity modulates endothelial cell apoptosis. Mol Biol Cell. 2003;14(3):848-57. PMCID: 151564.
124. Bergo MO, Gavino BJ, Hong C, Beigneux AP, McMahon M, Casey PJ, et al. Inactivation of Icmt inhibits transformation by oncogenic K-Ras and B-Raf. J Clin Invest. 2004;113(4):539-50. PMCID: 338259.
125. Winter-Vann AM, Kamen BA, Bergo MO, Young SG, Melnyk S, James SJ, et al. Targeting Ras signaling through inhibition of carboxyl methylation: an unexpected property of methotrexate. Proc Natl Acad Sci U S A. 2003;100(11):6529-34. PMCID: 164480.

126. Go ML, Leow JL, Gorla SK, Schuller AP, Wang M, Casey PJ. Amino derivatives of indole as potent inhibitors of isoprenylcysteine carboxyl methyltransferase. Journal of medicinal chemistry. 2010;53(19):6838-50.
127. Wang M, Hossain MS, Tan W, Coolman B, Zhou J, Liu S, et al. Inhibition of isoprenylcysteine carboxylmethyltransferase induces autophagic-dependent apoptosis and impairs tumor growth. Oncogene. 2010;29(35):4959-70.
128. Kapetanaki MG, Guerrero-Santoro J, Bisi DC, Hsieh CL, Rapic-Otrin V, Levine AS. The DDB1-CUL4ADDB2 ubiquitin ligase is deficient in xeroderma pigmentosum group E and targets histone H2A at UV-damaged DNA sites. Proc Natl Acad Sci U S A. 2006;103(8):2588-93.
129. Guerrero-Santoro J, Kapetanaki MG, Hsieh CL, Gorbachinsky I, Levine AS, Rapic-Otrin V. The cullin 4B-based UV-damaged DNA-binding protein ligase binds to UV-damaged chromatin and ubiquitinates histone H2A. Cancer Res. 2008;68(13):5014-22.
130. Ahn J, Vu T, Novince Z, Guerrero-Santoro J, Rapic-Otrin V, Gronenborn AM. HIV-1 Vpr loads uracil DNA glycosylase-2 onto DCAF1, a substrate recognition subunit of a cullin 4A-ring E3 ubiquitin ligase for proteasome-dependent degradation. J Biol Chem. 2010;285(48):37333-41. PMCID: 2988339.
131. Hagen L, Kavli B, Sousa MM, Torseth K, Liabakk NB, Sundheim O, et al. Cell cycle-specific UNG2 phosphorylations regulate protein turnover, activity and association with RPA. Embo J. 2008;27(1):51-61. PMCID: 2147998.
132. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature. 2008;455(7216):1061-8. PMCID: 2671642.
133. Repository of Molecular Brain Neoplasia Database (REMBRANDT) v1.5.5 [database on the Internet]. National Cancer Institute. 2005 [cited February 9, 2012]. Available from: http://rembrandt.nci.nih.gov.
134. Madhavan S, Zenklusen J-C, Kotliarov Y, Sahni H, Fine HA, Buetow K. Rembrandt: Helping Personalized Medicine Become a Reality through Integrative Translational Research. Molecular Cancer Research. 2009;7(2):157-67.
135. Visnes T, Akbari M, Hagen L, Slupphaug G, Krokan HE. The rate of base excision repair of uracil is controlled by the initiating glycosylase. DNA Repair (Amst). 2008;7(11):1869-81.
136. Kunz C, Focke F, Saito Y, Schuermann D, Lettieri T, Selfridge J, et al. Base excision by thymine DNA glycosylase mediates DNA-directed cytotoxicity of 5-fluorouracil. PLoS Biol. 2009;7(4):e91. PMCID: 2671560.
137. Stupp R, Gander M, Leyvraz S, Newlands E. Current and future developments in the use of temozolomide for the treatment of brain tumours. Lancet Oncol. 2001;2(9):552-60.
138. Valtieri M, Tweardy DJ, Caracciolo D, Johnson K, Mavilio F, Altmann S, et al. Cytokine-dependent granulocytic differentiation. Regulation of proliferative and differentiative responses in a murine progenitor cell line. J Immunol. 1987;138(11):3829-35.
139. Schneider U, Schwenk HU, Bornkamm G. Characterization of EBV-genome negative "null" and "T" cell lines derived from children with acute lymphoblastic leukemia and leukemic transformed non-Hodgkin lymphoma. Int J Cancer. 1977;19(5):621-6.
140. Milsom MD, Jerabek-Willemsen M, Harris CE, Schambach A, Broun E, Bailey J, et al. Reciprocal relationship between O6-methylguanine-DNA methyltransferase P140K expression level and chemoprotection of hematopoietic stem cells. Cancer Res. 2008;68(15):6171-80.

141. Davis BM, Roth JC, Liu L, Xu-Welliver M, Pegg AE, Gerson SL. Characterization of the P140K, PVP(138-140)MLK, and G156A O6-methylguanine-DNA methyltransferase mutants: implications for drug resistance gene therapy. Hum Gene Ther. 1999;10(17):2769-78.
142. Larochelle A, Choi U, Shou Y, Naumann N, Loktionova NA, Clevenger JR, et al. In vivo selection of hematopoietic progenitor cells and temozolomide dose intensification in rhesus macaques through lentiviral transduction with a drug resistance gene. J Clin Invest. 2009;119(7):1952-63. PMCID: 2701865.
143. Kreklau EL, Limp-Foster M, Liu N, Xu Y, Kelley MR, Erickson LC. A novel fluorometric oligonucleotide assay to measure O(6)-methylguanine DNA methyltransferase, methylpurine DNA glycosylase, 8-oxoguanine DNA glycosylase and abasic endonuclease activities: DNA repair status in human breast carcinoma cells overexpressing methylpurine DNA glycosylase. Nucleic Acids Res. 2001;29(12):2558-66. PMCID: 55735.
144. Ulrich HD. Ubiquitin and SUMO in DNA repair at a glance. J Cell Sci. 2012;125(2):249-54.
145. Sobol RW, Prasad R, Evenski A, Baker A, Yang XP, Horton JK, et al. The lyase activity of the DNA repair protein ß-polymerase protects from DNA-damage-induced cytotoxicity. Nature. 2000;405(6788):807-10.
146. Das BB, Dexheimer TS, Maddali K, Pommier Y. Role of tyrosyl-DNA phosphodiesterase (TDP1) in mitochondria. Proc Natl Acad Sci U S A. 2010;107(46):19790-5. PMCID: 2993338.
147. Patel AG, Sarkaria JN, Kaufmann SH. Nonhomologous end joining drives poly(ADP-ribose) polymerase (PARP) inhibitor lethality in homologous recombination-deficient cells. Proceedings of the National Acadamy of Sciences USA. 2011;108(8):3406-11. PMCID: 3044391.
148. Tang JB, Svilar D, Trivedi RN, Wang XH, Goellner EM, Moore B, et al. N-methylpurine DNA glycosylase and DNA polymerase {beta} modulate BER inhibitor potentiation of glioma cells to temozolomide. Neuro-oncology. 2011;13(5):471-86.
149. Agnihotri S, Gajadhar AS, Ternamian C, Gorlia T, Diefes KL, Mischel PS, et al. Alkylpurine-DNA-N-glycosylase confers resistance to temozolomide in xenograft models of glioblastoma multiforme and is associated with poor survival in patients. J Clin Invest. 2011.
150. Ono T, Wang S, Koo CK, Engstrom L, David SS, Kool ET. Direct fluorescence monitoring of DNA base excision repair. Angew Chem Int Ed Engl. 2012;51(7):1689-92.
151. Goellner EM, Svilar D, Almeida KH, Sobol RW. Targeting DNA Polymerase beta for therapeutic intervention. Curr Mol Pharmacol. 2011.
152. Trivedi RN, Almeida KH, Fornsaglio JL, Schamus S, Sobol RW. The Role of Base Excision Repair in the Sensitivity and Resistance to Temozolomide Mediated Cell Death. Cancer Research. 2005;65(14):6394-400.
153. Tolentino JH, Burke TJ, Mukhopadhyay S, McGregor WG, Basu AK. Inhibition of DNA replication fork progression and mutagenic potential of 1, N6-ethenoadenine and 8-oxoguanine in human cell extracts. Nucleic Acids Res. 2008;36(4):1300-8. PMCID: 2275085.
154. Bartsch H, Nair J. Oxidative stress and lipid peroxidation-derived DNA-lesions in inflammation driven carcinogenesis. Cancer Detection & Prevention. 2004;28(6):385-91.
155. Takeshita M, Chang CN, Johnson F, Will S, Grollman AP. Oligodeoxynucleotides containing synthetic abasic sites. Model substrates for DNA polymerases and apurinic/apyrimidinic endonucleases. J Biol Chem. 1987;262(21):10171-9.

156. Dyrkheeva NS, Khodyreva SN, Lavrik OI. Interaction of APE1 and other repair proteins with DNA duplexes imitating intermediates of DNA repair and replication. Biochemistry (Mosc). 2008;73(3):261-72.
157. Vidal AE, Boiteux S, Hickson ID, Radicella JP. XRCC1 coordinates the initial and late stages of DNA abasic site repair through protein-protein interactions. EMBO Journal. 2001;20(22):6530-9.