

The role of NUCKS in DNA repair

A thesis presented for the degree of Master's by Research

Samuel Hume

Supervisors:

Prof. Grigory Dianov

Dr. Arnaud Legrand



St Edmund Hall

CRUK/MRC Oxford Institute for Radiation Oncology,
Department of Oncology,
University of Oxford

Trinity Term 2016



Acknowledgements

Firstly I would like to thank Grigory, for giving me the opportunity to study within a prestigious lab, working on an exciting project and living in a beautiful city. Secondly, I am grateful to Arnaud, for the ambition-provoking motivational words and the indispensable project ideas. I am also very grateful to the other members of the Dianov group: Mattia, Sally, Claudia, Chen, Di and Siobhan.

Declaration

The work and data presented in this thesis are wholly my own. The bioinformatics analysis presented in **Table 5** was a joint effort from my supervisor, Dr. Arnaud Legrand, and I.

Word count

21,000 words inclusive of legends, in-text citations and references.

Abstract

NUCKS is a vertebrate-specific, extensively post-translationally modified transcriptional regulator whose over- and under- expression is associated with cancer and obesity, respectively. We have previously found that NUCKS protein levels increase in DNA-damaged cells, and other studies have reported DNA damage-dependent NUCKS phosphorylation. However, the full impacts of NUCKS depletion on DNA damage repair has not been established. Addressing this question, we show that NUCKS depletion activates some DNA damage response proteins and might be associated with accumulation of DNA damage in cancer cells. Furthermore, the transcriptional targets for NUCKS are poorly defined. To this end, we couple data from a published ChIP-Seq analysis for NUCKS with a bioinformatics approach to identify genes coexpressed with NUCKS in cancer patient samples and cell lines. We identify a number of genes, including many involved in DNA repair. We correspondingly find that NUCKS-depleted cells express lower levels of these genes and exhibit some hallmarks of their deficiency. Finally, we investigate how NUCKS' putative transcriptional targets relate to disease, and therefore how NUCKS might be used as a biomarker for disease.

Contents

Acknowledgements.....	2
Declaration.....	2
Word count.....	2
Abstract.....	3
Contents.....	4
Table of figures	7
Table of tables.....	8
List of abbreviations	9
1. Introduction	14
1.1 Importance of new biomarker discoveries for cancer therapy	14
1.2 Nuclear Ubiquitous Casein and cyclin-dependent Kinase Substrate (NUCKS)	14
1.2.1 Why are we interested in NUCKS?	14
1.2.2 Biochemistry of NUCKS	15
1.2.3 Regulation of NUCKS expression.....	15
1.2.4 NUCKS: a DNA-binding protein	16
1.2.4.1 NUCKS and transcriptional regulation.....	16
1.2.4.2 NUCKS and DNA repair.....	17
1.2.5 NUCKS expression in human diseases.....	19
1.3 DNA damage and repair	20
1.3.1 DNA instability and repair systems	20
1.3.2 DNA mismatch repair	22
1.3.2.1 Biochemistry of DNA mismatch repair.....	22
1.3.2.2 DNA mismatch repair and cancer	25
1.3.2.3 Influence of mismatch repair status on drug sensitivity.....	26
1.3.2.4 Regulation of DNA mismatch repair.....	27
1.3.2.4.1 Control of mismatch repair transcription and mRNA	27
1.3.2.4.2 Post-translational control of mismatch repair	28
1.4 Our project	28
2. Materials and methods	30
2.1 Cell culture	30

2.1.1 Cell culturing and harvesting.....	30
2.1.2 Transfection.....	31
2.1.3 Genotoxic agent treatments	32
2.2 Detection of relative mRNA levels	32
2.2.1 RNA extraction	32
2.2.2 Reverse transcription	33
2.2.3 Quantitative polymerase chain reaction (qPCR).....	33
2.3 Detection of relative protein levels.....	35
2.3.1 Protein extraction.....	35
2.3.2 Polyacrylamide gel electrophoresis (PAGE)	36
2.3.3 Western blotting	37
2.4 Flow cytometry.....	38
2.4.1 Cell cycle analysis	38
2.4.2 Protein levels analysis	39
2.5 Alkaline Comet assay.....	39
2.6 Bioinformatics analysis.....	40
2.6.1 SEEK online gene expression database	40
2.6.2 G-Profiler online disease database	40
2.7. Statistical analysis.....	41
3. Results.....	42
3.1 Identification of NUCKS as a new responder to DNA damage.....	42
3.1.1 NUCKS is a putative ATM target upregulated upon DNA repair deficiency.....	42
3.2 Consequences of NUCKS depletion on the DDR.....	43
3.2.1 NUCKS siRNA sequences validation	43
3.2.2 NUCKS knockdown induces p53 and p21, key DDR proteins.....	44
3.2.3 NUCKS knockdown arrests the cell cycle in G1 phase	46
3.2.4 NUCKS knockdown might be associated with DNA damage in cancer cells	48
3.3 NUCKS and the transcriptional regulation of DNA repair: mismatch repair genes	50
3.3.1 NUCKS correlates with DNA repair, cell cycle and transcription-controlling genes	50
3.3.2 NUCKS is required to maintain expression of mismatch repair genes.....	53
3.3.3 NUCKS' regulation of mismatch repair gene MSH2 is not cell cycle-dependent.....	55
3.3.4 NUCKS-depleted cells display increased resistance to 6-TG, an agent that relies on proficient mismatch repair for lethality.....	57
3.3.5 Genes in NUCKS CHIP-Seq are associated with colorectal cancer.....	59

4. Discussion.....	61
4.1 NUCKS as a DNA damage response protein.....	62
4.2 NUCKS, mismatch repair and colorectal cancer.....	63
4.3 Limitations of our approach.....	65
4.4 Future experiments.....	66
4.4.1 NUCKS and the DDR	66
4.4.2 NUCKS, mismatch repair and colorectal cancer.....	68
4.5 Conclusions.....	69
Appendix.....	71
References	73

Table of figures

1. Introduction.....	14
Figure 1: DNA mismatch repair: schematic	24
3. Results.....	38
Figure 2: Validation of NUCKS siRNA sequences	44
Figure 3: NUCKS knockdown induces p53 and p21.....	45
Figure 4: NUCKS knockdown induces G1 cell cycle arrest.....	48
Figure 5: NUCKS knockdown and DNA damage	49
Figure 6: NUCKS knockdown results in deregulation of transcription	54
Figure 7: Loss of MSH2 after NUCKS depletion is not cell cycle-dependent.....	56
Figure 8: NUCKS-depleted cells display increased resistance to 6-TG	58
4. Discussion.....	57
Figure 9: NUCKS and colorectal cancer	64

Table of tables

2. Materials and methods.....	27
Table 1: siRNA sequences	31
Table 2: qPCR primers.....	35
Table 3: Antibodies.	38
3. Results.....	38
Table 4: Nuclear ATM targets upregulated in XRCC1 SILAC	43
Table 5: Genes in NUCKS ChIP-Seq that correlate with NUCKS expression.	52
Table 6: Genes in the NUCKS ChIP-Seq are associated with colorectal cancer	60
Appendix.....	64
Table 7: Materials and reagents	72

List of abbreviations

5-FU	5-fluorouracil
6-TG	6-thioguanine
8-OxdG	8-hydroxy-2'-deoxyguanosine
8-oxo	8-oxoguanine
53BP1	p53-binding protein 1
AMPK	5' AMP-activated protein kinase
AP	Apurinic/aprimidinic
APC	Adenomatous polyposis coli
APE1	AP endonuclease 1
APS	Ammonium persulfate
ATM	Ataxia telangiectasia mutated
ATR	ATM/Rad3-related
AXIN2	Axis inhibition protein 2
B2M	Beta-2 microglobulin
Bcl	B-Cell CLL/Lymphoma
BER	Base excision repair
BMI	Body mass index
BRAF	V-Raf murine sarcoma viral oncogene homolog B 2
BRCA1/2	Breast cancer type 1/2 susceptibility protein
BSA	Bovine serum albumin
c-Abl	V-Abl abelson murine leukemia viral oncogene homolog 1
CDK	Cyclin-dependent kinase
cDNA	Complementary DNA
ChIP	Chromatin immunoprecipitation
ChIP-Seq	Chromatin immunoprecipitation sequencing
Chk1/2	Checkpoint kinase 1/2
CK2	Casein kinase 2
CPT	Camptothecin
CRC	Colorectal cancer
CRTC1	CREB-regulated transcription coactivator 1
CSDE1	Cold shock domain containing E1
CSE1L	Chromosome segregation 1 like
CSF1	Colony-stimulating factor 1
CT	Cycle threshold
CtIP	CtBP1 interacting protein
CTNNB1	Catenin beta 1
DDR	DNA damage response

D-loop	Displacement-loop
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNA-PK	DNA-protein kinase
dNTPs	Deoxynucleotides
DROSHA	Double-stranded RNA-specific endoribonuclease
ds	Double strand
DSBs	Double strand breaks
DTT	Dithiothreitol
E2F	E2F transcription factor 1
EDTA	Ethylenediaminetetraacetic acid
EGFR	Epidermal growth factor receptor
EXO1	Exonuclease 1
FACS	Fluorescence-activated cell sorting
FBS	Foetal bovine serum
FEN1	Flap endonuclease 1
FOSL2	Fos-related antigen 2
Foxo1	Forkhead box O
G0/G1/G2	Growth
Gadd45a	Growth arrest and DNA damage inducible alpha
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GFP	Green fluorescent protein
H3K36	Histone 3 lysine 36
HCC	Hepatocellular carcinoma
HDAC	Histone deacetylase
HIV	Human immunodeficiency virus
HMG	High mobility group
HNF4A	Hepatocyte nuclear factor 4 alpha
HNPCC	Hereditary nonpolyposis colorectal cancer
HNRNPA1	Heterogeneous nuclear ribonucleoprotein A1
HR	Homologous recombination
hTERT	Human telomerase reverse transcriptase
HU	Hydroxyurea
InDel	Insertion/deletion
ICL	Interstrand crosslink
IRF5	Interferon regulatory factor 5
JAK3	Janus kinase 3
JUN	V-jun avian sarcoma virus 17 oncogene homolog
kDa	Kilodaltons
KPNB1	Karyopherin subunit beta 1

KRAS	V-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog
Ku70	Lupus Ku autoantigen protein p70
Kub5-Hera	Ku70-binding protein 5-Hera
LIG	DNA ligase
LSD1	Lysine demethylase 1
M	Mitosis
MAPK	Mitogen-activated protein kinase
MDM2	Mouse double minute 2
MEFs	Mouse embryonic fibroblasts
miR	Micro-RNA
MLH1	MutL homolog 1
MLLT4	Myeloid/lymphoid or mixed-lineage leukemia; translocated to, 4
MMC	Mitomycin C
MMR	Mismatch repair
MMS	Methyl methanesulfonate
MRE11	Meiotic Recombination 11
MRN	MRE11/Rad50/Nbs1
mRNA	Messenger RNA
MSH2	MutS protein homolog 2
MSH6	MutS protein homolog 6
MSI	Microsatellite instability
MSS	Microsatellite stable
MutS	Mutator S
Myc	v-Myc Avian Myelocytomatosis Viral Oncogene Homolog
NBE1	Normal bronchial epithelial cells
NBS1	Nijmegen breakage syndrome
NC	Negative control
NCL	Nucleolin
NEM	N-Ethylmaleimide
NER	Nucleotide excision repair
NHEJ	Non-homologous end joining
NLS	Nuclear localisation signal
NMR	Nuclear magnetic resonance
NUCKS	Nuclear ubiquitous casein kinase and cyclin-dependent kinase substrate
P21	Tumour protein 21
P53	Tumour protein 53
PAGE	Polyacrylamide gel electrophoresis
PALB2	Partner and localizer of BRCA2
PALLD	Palladin
PAR	Poly-(ADP)-ribose

PARP1	Poly-(ADP)-ribose polymerase 1
PBS	Phosphate-buffered saline
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PCV	Packed cell volume
PDGFRA	Platelet derived growth factor receptor alpha
PINK1	PTEN-induced putative kinase 1
PLK3	Polo-like kinase 3
PMS2	PMS1 homolog 2
PMSF	Phenylmethanesulfonylfluoride
PNISR	PNN-interacting serine/arginine-rich protein
POL β	DNA polymerase beta
PPP1R15A	Protein phosphatase 1 regulatory subunit 15A
PTMA	Prothymosin, Alpha
PTMs	Post-translational modifications
PVDF	Polyvinylidene fluoride
qPCR	Quantitative polymerase chain reaction
RAD51	RecA-like protein 51
RAD51AP1	RAD51 associated protein 1
RCC2	Regulator of chromosome condensation 2
RFC	Replication factor C
RCF	Relative centrifugal force
RELA	v-Rel avian reticuloendotheliosis viral oncogene homolog A
RHAMM	Hyaluronan-mediated motility receptor
RNA	Ribonucleic acid
RNAi	RNA interference
RNAP	RNA polymerase
RNR	Ribonucleoside reductase
ROS	Reactive oxygen species
RPA	Replication protein A
RFC	Replication factor C
RT	Room temperature
S	Synthesis
SBs	Strand breaks
SCID	Severe combined immunodeficiency disorder
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEEK	Search-based exploration of expression compendium
SETD2	SET domain-containing 2
SILAC	Stable isotope labelling by amino acids in cell culture
siRNA	Small interfering RNA

List of abbreviations

SKP2	S-phase kinase-associated protein 2
SMAD4	SMAD family member 4
Sp1	Specificity protein 1
ss	Single strand
SSBR	Single strand break repair
SSBs	Single strand breaks
STATs	Signal transducers and activators of transcription
TBP	TATA-binding protein
TEMED	Tetramethylethylenediamine
TGF	Transforming growth factor
TIG1	Tokyo Metropolitan Institute of Gerontology (Normal human lung fibroblasts)
U2OS	U-2 Osteosarcoma (human bone osteosarcoma epithelial cells)
UV	Ultraviolet
WCE	Whole cell extract
WISP3	WNT1 inducible signaling pathway protein 3
Wnt	Wingless-related integration site
WT	Wildtype
XPO1	Exportin 1
XRCC1	X-ray repair cross-complementing protein 1
YY1	Yin Yang 1

1. Introduction

1.1 Importance of new biomarker discoveries for cancer therapy

In the United States alone, 1.7 million new cancer cases will be diagnosed in 2016, and 600,000 people will die of the disease (Siegel et al., 2016). Nevertheless, rapid developments are being made in diagnosis and treatment, and nearly 2 million people have been saved from cancer deaths in the last 20 years as a result (Siegel et al., 2016). Central to these improvements has been the utilisation of biomarkers, molecules which act as signals for pathology, which can improve screening, prognosis and help to predict response to treatment (Henry and Hayes, 2012). For example, mutation in BRCA1 predicts risk of breast and ovarian cancer (Easton et al., 1995), while KRAS mutation status predicts response to anti-EGFR antibody therapy in colorectal cancer (CRC) (Allegra et al., 2009; Henry and Hayes, 2012). Further identifying and characterising novel biomarkers will be central to the continued improvement of cancer therapy.

1.2 Nuclear Ubiquitous Casein and cyclin-dependent Kinase Substrate (NUCKS)

1.2.1 Why are we interested in NUCKS?

In 2015, our group performed stable isotope labelling with amino acids in cell culture (SILAC) to identify proteins deregulated in base excision repair (BER)-deficient cells (Markkanen et al., 2015); NUCKS was found to be upregulated nearly 2-fold in response to BER loss (Markkanen et al., 2015). Since the functions of NUCKS are poorly defined in the literature, we hypothesised this protein's expression might be induced after DNA damage and might therefore play roles in the DNA damage response. Further investigations, reported in this manuscript, unveiled a more complex role for NUCKS at the crossroad between DNA repair and transcription.

1.2.2 Biochemistry of NUCKS

NUCKS, identified in 1985 as a member of the high mobility group (HMG) family of proteins (Ostvold et al., 1985), is a vertebrate-specific, highly conserved DNA-binding protein (Stvold et al., 2001). Mass spectrometry suggests that NUCKS' molecular weight is 27 kDa. However, NUCKS anomalously migrates at ~50 kDa by SDS-PAGE (Stvold et al., 2001; Ostvold et al., 1985). This difference appears to be because NUCKS – which contains few hydrophobic residues and is heavily phosphorylated - is highly charged. NUCKS also has high proline content. These characteristics alter SDS binding to NUCKS and slow its SDS-PAGE migration (Stvold et al., 2001; Ostvold et al., 1985). NUCKS has a high content of β turns, with few α helices (Stvold et al., 2001). Strikingly, NUCKS undergoes more post-translational modification than any other protein in the proteome (Wiśniewski et al., 2008): overall, more than 10% (37 of 243) of NUCKS' amino acids undergo post-translational modification, 34 of which can be phosphorylated (Wiśniewski et al., 2008). NUCKS is phosphorylated by CK2, DNA-PK, CDK1, several MAPKs, ATM and ATR (Stvold et al., 2001; Matsuoka et al., 2007; Bensimon et al., 2010). The function of these phosphorylations, as well as their provoking stimuli, is poorly defined. However, it is known that phosphorylation of NUCKS at serine 181 by CDK1 during mitosis prevents its binding to DNA (Grundt et al., 2002), and that ATM phosphorylates serine 14 of NUCKS following ionising radiation (Matsuoka et al., 2007).

1.2.3 Regulation of NUCKS expression

NUCKS' promoter contains two initiator elements and two GC boxes but no TATA box (Grundt et al., 2004). E2F1 (Cao et al., 2011) and p53 (Nikulenkov et al., 2012) bind the NUCKS promoter, suggesting its expression might be regulated in a cell cycle- or DNA damage- dependent manner. NUCKS expression can also be regulated by microRNAs (miR): for example, loss of miR-137 increases NUCKS expression, ostensibly leading to increased cell proliferation, migration and survival (Shen et al., 2016).

1.2.4 NUCKS: a DNA-binding protein

1.2.4.1 NUCKS and transcriptional regulation

NUCKS was originally identified as a novel member of the HMG family, DNA-binding proteins that influence transcription by regulating chromatin context (Ostvold et al., 1985; Agresti and Bianchi, 2003). NUCKS is a nuclear DNA-binding protein which contains two nuclear localisation signals and a DNA-binding domain (Grundt et al., 2002, 2007). This DNA-binding domain, which spans residues 183-192 and contains the VGRP DNA-binding consensus, binds chromatin with higher affinity than it does naked DNA (Grundt et al., 2002; Parplys et al., 2015) and NUCKS' phosphorylations can potentiate its DNA-binding capacity (Ostvold et al., 1992). Analysis of NUCKS PTMs in breast tumours revealed that NUCKS undergoes an array of cancer-specific modifications (Wiśniewski et al., 2008). Most strikingly, breast cancer-specific lysine acetylations were identified, including two in the vicinity of the NUCKS binding domain (K184-Ac and K196-Ac). Unfortunately, neither the modifying enzymes – nor the implications - were identified. However, protein acetylation is known to be a potent modulator of DNA binding and transcriptional activity (Das and Kundu, 2005). As such, these modifications might alter NUCKS' DNA binding and transcription of its targets, thereby facilitating breast cancer progression (Wiśniewski et al., 2008).

Only recently, targets for NUCKS' transcriptional regulation have been identified. NUCKS is a tat coactivator that promotes transcription of HIV's viral genes, suggesting that NUCKS might play an important role in HIV infection (Kim et al., 2014). Furthermore, NUCKS promotes transcription of the insulin receptor by opening the chromatin structure at the promoter, promoting RNA polymerase II recruitment to the promoter (Qiu et al., 2014).

Importantly, Qiu et al. published the findings of a chromatin immunoprecipitation sequencing (ChIP-Seq) for NUCKS, globally identifying NUCKS' binding sites in DNA (Qiu et al., 2014). This study reported that NUCKS specifically binds the promoter of 586 genes, whose roles range from nuclear

import to immune system function. This strongly suggests that NUCKS might have unprecedented importance for cellular prosperity, incongruous with its lack of research attention thus far (Qiu et al., 2014).

1.2.4.2 NUCKS and DNA repair

Numerous clues exist hinting towards a role for NUCKS in DNA repair. DNA single strand breaks (SSBs) generated by depletion of the BER scaffold protein XRCC1 induce NUCKS upregulation, suggesting that NUCKS might play a role in the DNA damage response (Markkanen et al., 2015). Other studies have shown that NUCKS is phosphorylated by DNA damage-signalling kinases such as ATM (Stvold et al., 2001; Matsuoka et al., 2007; Bensimon et al., 2010), suggesting NUCKS' activity might be modulated in response to DNA damage.

Recently, a preliminary report proposed a putative role for NUCKS in homologous recombination (HR) (Parplys et al., 2015). This method of DSB repair is initiated by the recruitment of the MRN complex, made up of Mre11, Rad50 and Nbs1 proteins, to the DSB site (D'Amours and Jackson, 2002). This complex, in an interaction with CtIP, facilitates 5'-3' end resection (D'Amours and Jackson, 2002; Sartori et al., 2007). This generates 3' single-stranded overhangs, substrates which RPA molecules bind and stabilise (Wold, 1997). BRCA2 loads RAD51 molecules onto this ssDNA substrate (Jensen et al., 2010), displacing RPA and forming RAD51's nucleofilament structure (Sung et al., 2003). RAD51 subsequently performs a search for homologous regions of DNA to template repair and, once successful, initiates strand invasion into the homologous region, generating a D loop structure (Sung et al., 2003). RAD51's D loop formation is enhanced by PALB2 (Dray et al., 2010) and RAD51AP1 (Wiese et al., 2007). Using the homologous region as a template and the single stranded 3' overhang as a primer, the resected DNA is replaced through DNA synthesis by δ , η , and κ polymerases (McVey et al., 2016). This gives rise to a double Holliday junction structure, which can be resolved by one of three mechanisms: 1) dissolution, by a complex of DNA helicase BLM, topoisomerase Topo III α and RMI1/2 proteins 2) symmetric resolution, by 5' flap

endonuclease GEN1 3) asymmetric resolution, by a complex of structure-specific endonucleases MUS81-EME1 and SLX1-SLX4 (Wyatt and West, 2014).

Parpys et al. identified NUCKS through a search of paralogues of the HR factor RAD51AP1 (Parpys et al., 2015), a protein which interacts with RAD51 and promotes RAD51 D loop formation (Wiese et al., 2007). 28% of NUCKS' amino acids are identical to those of RAD51AP1, and an additional 11% are similar. Accordingly, Parpys et al. tested whether NUCKS might similarly have a role in HR. They were able to show that NUCKS depletion indeed gave characteristics of HR deficiency: 1) S phase-synchronised NUCKS-depleted cells were sensitive to ionising radiation 2) NUCKS-depleted cells were sensitive to mitomycin C, an agent which primarily generates DNA interstrand crosslinks (ICLs) and which relies on HR proteins, including RAD51AP1 (Wiese et al., 2007), for successful repair¹ (Hinz, 2010) 3) NUCKS-depleted cells accumulated more chromatid breaks than control cells 4) NUCKS depletion significantly impaired DNA replication progression 5) NUCKS depletion significantly impaired DSB repair measured using a reporter in which repair of an induced DSB allows functional GFP expression; this DSB repair defect was commensurate with RAD51AP1 depletion (Parpys et al., 2015).

Despite this evidence, Parpys et al. could not provide a mechanism for NUCKS' role in HR. NUCKS did not impact recruitment of RAD51 to DSB sites and did not affect core HR protein expression; the recruitment of NUCKS to DSB sites was not tested. Unlike RAD51AP1, NUCKS did not interact with RAD51 and did not impact RAD51 D-loop formation. Accordingly, NUCKS' role in HR was placed downstream of RAD51 D-loop formation in HR. Therefore, there seem to be two remaining possibilities: 1) NUCKS might be important for DNA synthesis after RAD51 strand invasion (indeed, NUCKS can bind D-loop substrates (Parpys et al., 2015)) or 2) NUCKS might play a role in the

¹ HR is important – but not sufficient – for ICL repair; other repair pathways, including Fanconi anemia, are required (Hinz, 2010). In humans, repair of ICLs is executed during DNA replication, in S phase. The encounter of a replication fork with an ICL causes the replication fork to stall. Subsequent initiation of ICL repair by DNA cleavage at the lesion site generates a one ended DSB intermediate, which is utilised as a substrate by the HR machinery to repair the DSB and restart the replication fork (Hinz, 2010).

resolution of Holliday junctions. NUCKS might achieve these roles by through relaxing chromatin at the DSB site, thereby facilitating recruitment of HR factors to HR intermediates. For example, NUCKS might enable recruitment of DNA polymerases for DNA synthesis or factors necessary for Holliday junction resolution (Parplys et al., 2015).

1.2.5 NUCKS expression in human diseases

NUCKS overexpression has been reported in several carcinomas. NUCKS' chromosome region is amplified in breast cancer (Naylor et al., 2005; Drosos et al., 2009), and NUCKS overexpression is also observed in hepatocellular carcinoma (HCC), suggesting NUCKS might be used as an immunodiagnostic marker for HCC (Cheong et al., 2016). To assess the mechanism underpinning NUCKS' overexpression in HCC, Cheong et al. performed microarrays to assess deregulated genes in NUCKS-depleted HCC cells. Notably, NUCKS-proficient cells expressed higher levels of AMPK signalling genes, a pathway known to promote cell proliferation (Motoshima et al., 2006). Thus, NUCKS might expedite HCC progression through boosting AMPK signalling. High expression of NUCKS is associated with poor survival in cervical cancer (Gu et al., 2014), risk of recurrence in endometrial cancer (Liu et al., 2015), and metastasis in lung (Sargent et al., 2008) and gastric adenocarcinomas (Yang et al., 2014). Furthermore, HCC cells depleted of NUCKS form smaller tumours when injected into mice than their NUCKS-proficient equivalents (Cheong et al., 2016).

The mechanisms underpinning NUCKS overexpression in cancer are poorly understood. However, some lines of evidence suggest that NUCKS might promote cell proliferation by enabling cells to progress from G1 to S phase: 1) NUCKS expression correlates with the proliferation marker Ki67, which is expressed in cycling (G1/S/G2/M) cells, but not in G0 cells (Drosos et al., 2009; Scholzen and Gerdes, 2000) 2) NUCKS expression peaks at the G1/S transition (www.proteinatlas.org/) and 3) NUCKS binds the promoters of genes that are important for entry into S phase, such as cyclin D1 (Baldin et al., 1993), CDK2 (Koff et al., 1992) and SKP2 (Kossatz et al., 2004; Qiu et al., 2014). These

observations together suggest that the overexpression of NUCKS in cancer cells might boost cancer hyperproliferation by enabling G1/S progression.

On the contrary, loss of NUCKS expression is associated with obesity. NUCKS binds the promoter of the insulin receptor gene, opening the chromatin at the gene and permitting access of RNAPII to the promoter (Qiu et al., 2014). As a result, NUCKS-depleted cells express less insulin receptor mRNA, displaying impaired insulin signalling and glucose intolerance as a result. Accordingly, NUCKS expression inversely correlates with BMI in humans and NUCKS knockout mice are obese (Qiu et al., 2014).

To conclude, the level of NUCKS expression seems to be a key determinant for pathology. Overexpression of NUCKS is associated with a wide range of cancers, while loss of NUCKS causes obesity. Curiously, obesity significantly increases risk of cancer (Basen-Engquist and Chang, 2011), suggesting that NUCKS underexpression might also lead to cancer. In the same vein, a recent study reported that loss of NUCKS is associated with lymphomagenesis in mice that concurrently lack p53 (Yue et al., 2016). As such, both aberrant reductions and increases in NUCKS expression might have the potential to lead to cancer.

1.3 DNA damage and repair

1.3.1 DNA instability and repair systems

DNA is an inherently unstable molecule which accumulates thousands of lesions per day from endogenous sources (Lindahl, 1993): ~3,000-11,500 oxidative lesions (Ames et al., 1993), ~2,500-10,000 depurinations and depyrimidations (Lindahl and Nyberg, 1972), ~50,000 SSBs (Tice and Setlow, 1985), ~200 cytosine deaminations (Tice and Setlow, 1985), and ~10-50 DSBs (Vilenchik and Knudson, 2003). DNA is further susceptible to damage by exogenous agents such as UV light, ionising radiation and chemotherapeutics (Friedberg et al., 2004). Defective repair of these lesions

leads to genomic instability and mutation, allowing cancer cells to become aggressive and develop their characteristic hallmarks (Hanahan and Weinberg, 2011).

The majority of endogenously-induced non-helix distorting lesions, including oxidations, alkylations and SSBs, are repaired by BER (Dianov and Hübscher, 2013). Underscoring the importance of BER, mouse knockout models for the core BER proteins Pol β (Gu et al., 1994), ligase III (Puebla-Osorio et al., 2006), XRCC1 (Tebbs et al., 2003) and APE1 (Friedberg and Meira, 2006) result in embryonic lethality.

Bulky, helix-distorting DNA lesions, such as UV induced (6-4) photoproducts, as well as lesions generated by chemotherapeutics such as cisplatin, are recognised and corrected by nucleotide excision repair (NER) (Wood, 1997; Gillet and Scharer, 2006). NER deficiency causes Xeroderma pigmentosum, a genetic disorder characterised by extreme sensitivity to UV light and frequent occurrence of skin cancers (de Boer and Hoeijmakers, 2000).

Cells employ two complementary mechanisms to repair DSBs, particularly toxic lesions that can cause chromosomal aberrations and cell death (Chapman et al., 2012). The majority of DSB repair occurs via non-homologous end joining (NHEJ), an error-prone method of DSB repair where DNA ends on either side of a DSB are directly ligated (Lieber, 2010). NHEJ deficiency causes severe combined immunodeficiency (SCID), an inherited disorder characterised by defective DSB repair, radiosensitivity and microcephaly (O'Driscoll et al., 2001). Cells also offer a high fidelity DSB repair pathway, HR, which intricately templates homologous regions of DNA on sister chromatids during S and G2 phases of the cell cycle (Krejci et al., 2012). Defects in HR, particularly mutations in BRCA1 or BRCA2, are associated with early-onset breast and ovarian cancers (Welch and King, 2001).

The last major DNA repair pathway is mismatch repair (MMR).

1.3.2 DNA mismatch repair

1.3.2.1 Biochemistry of DNA mismatch repair

MMR corrects base-base mismatches and nucleotide insertions/deletions (InDels) arising during DNA replication. Consequently, MMR enhances the fidelity of replication by 50-1000 fold (Kunkel and Erie, 2005) and MMR deficiency is associated with a mutator phenotype characterised by microsatellite instability (MSI) (Loeb, 1994).

Four core MMR proteins coordinate the repair of base-base mismatches and short InDels of one- to two nucleotides: MutS protein homolog 2 (MSH2), MutS protein homolog 6 (MSH6), MutL homolog 1 (MLH1) and PMS1 homolog 2 (PMS2) (Kunkel and Erie, 2015). 90% of cellular MSH2 is found in a complex with MSH6, forming the 260kDa heterodimer MutS α (Jiricny, 2000), while MLH1 and PMS2 associate to form 181 kDa MutL α (Kadyrov et al., 2006). MutS α is an ATPase, which forms a sliding clamp ring structure around DNA, in search of mismatches (Lamers et al., 2000; Obmolova et al., 2000; Warren et al., 2007). MSH6 contains the mismatch-binding domain of MutS α , in which a highly conserved GxFxR motif is located. Crystal structures of MutS proteins in complex with mismatched nucleotides reveal that the central phenylalanine of this motif inserts into the minor groove of DNA at the mismatch and facilitates mismatch recognition (Lamers et al., 2000; Obmolova et al., 2000; Warren et al., 2007). Mismatch binding causes MutS α to hydrolyse ATP and undergo conformational changes culminating in the release of the central phenylalanine from the minor groove (Lee et al., 2014). This converts MutS α to a sliding clamp which recruits and binds the MutL α complex (Friedhoff et al., 2016).

MutL α in solitude has weak endonuclease activity, whose active site is localised to a highly conserved DQHA(X)(2)E(X)(4)E sequence in the PMS2 protein (Kadyrov et al., 2006). MutL α is activated through direct interaction with MutS α and the replication clamp PCNA, a protein loaded onto DNA by RFC (Kadyrov et al., 2006; Pluciennik et al., 2010). Consequently, MutL α generates a

nick in DNA, on the 5' side of the mismatch (Kadyrov et al., 2006). This allows the exonuclease EXO1 to excise the mutant DNA 5'-3', starting at the nascent nick and stretching beyond the mismatch (Schmutte et al., 1998). The resulting gap is filled in by DNA pol δ (Longley et al., 1997) or ϵ (Tran et al., 1999), and the nick is sealed by DNA ligase I (Zhang et al., 2005), resulting in repaired DNA. A (simplified) schematic of the MMR pathway is presented in **Fig. 1**.

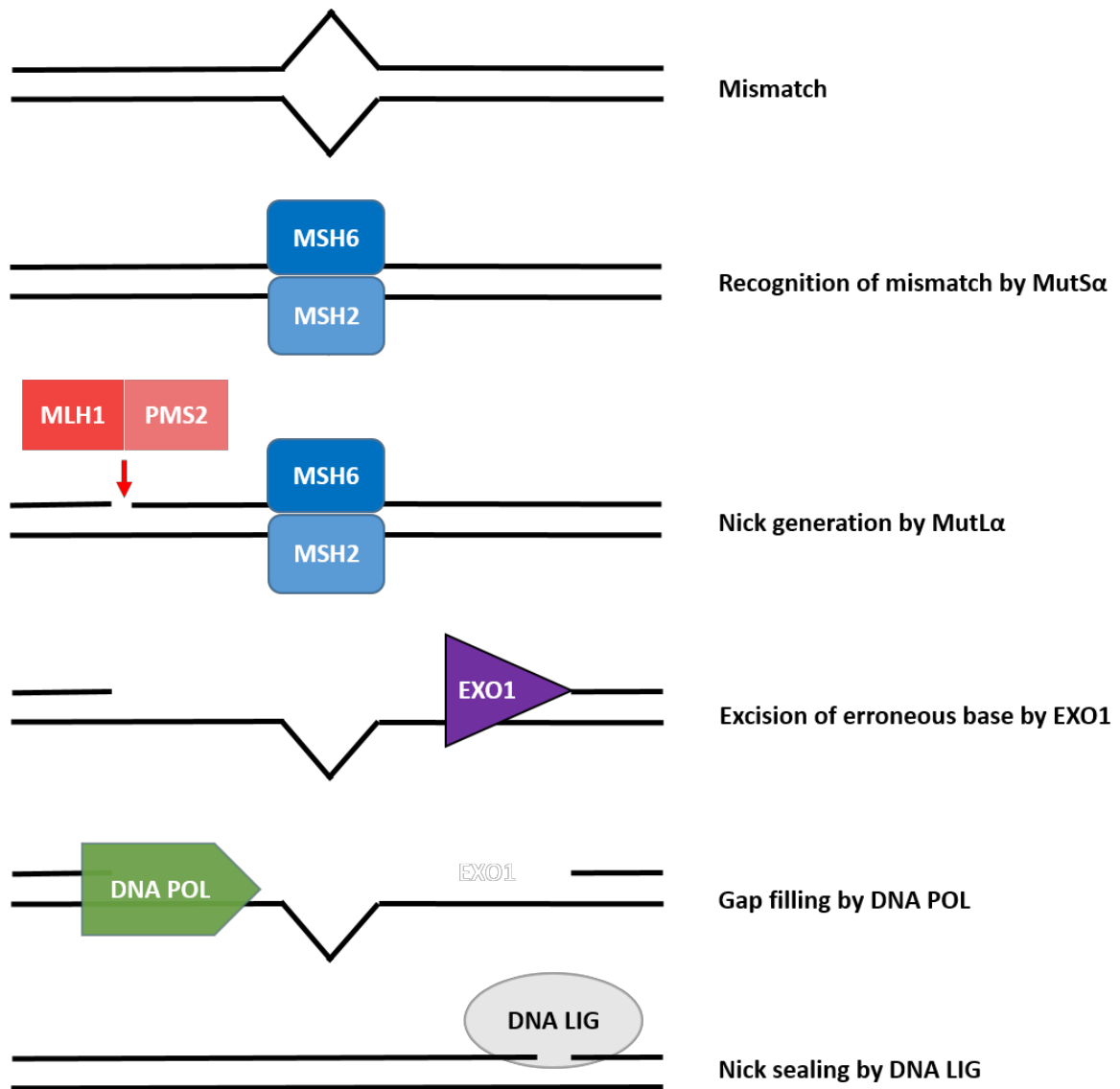


Figure 1: DNA mismatch repair. MutSα recognises a mismatch and recruits MutLα, which generates a nick in the DNA, allowing EXO1 to excise the erroneous base and surrounding bases. The resulting gap is filled by DNA polymerase and sealed by DNA ligase. Adapted from Martin and Scharff, 2002.

1.3.2.2 DNA mismatch repair and cancer

MMR deficiency induces a mutator phenotype characterised by MSI (Loeb, 1994). Microsatellites are tandem repeats of mono-, di-, tri- or tetra- nucleotides, regions upon which DNA polymerases are sensitive to slippage, giving rise to InDels (Ellegren, 2004). MMR-deficient cells cannot repair these InDels, and generate novel microsatellite fragments – MSI - as a result (Loeb, 1994). MSI is associated with a range of cancers, including endometrial, ovarian, breast, skin, glioma, lymphoma and colorectal (Boland et al., 1998). Of these, the contribution of MMR deficiency to CRC is best characterised: of the 1.2 million CRCs diagnosed every year, ~15% display MSI (Brenner et al., 2014; Boland and Goel, 2010). 3% of these MSI-positive CRCs are hereditary nonpolyposis colorectal cancers (HNPCC), the most common type of hereditary CRC which is hallmarked by defective MMR (Muller and Fishel, 2002), while the remaining 12% are sporadic CRC cases (Boland and Goel, 2010). HNPCC cases are largely accounted for by germline mutations in the core MMR proteins MSH2, MLH1, MSH6 and PMS2 (Peltomaki, 2003), which impact MMR functions including DNA-binding and MSH2-MSH6 interaction (Hsieh and Yamane, 2008), while sporadic CRC cases are caused predominantly by MLH1 promoter hypermethylation, abrogating its expression (Boland and Goel, 2010). Alterations in gene expression of core MMR components also contributes to MSI and CRC (Renkonen et al., 2003; Thibodeau et al., 1996).

MSI facilitates cancer cells' acquisition of their characteristic hallmarks by inactivating tumour suppressors and activating oncogenes (Boland and Goel, 2010; Hanahan and Weinberg, 2011). For example, microsatellite sequences in the coding sequence of TGF β 2 cause MSI-positive cell lines to generate a frameshift that abrogates expression of TGF β 2, a growth factor that attenuates proliferation of cells in the colon (Markowitz et al., 1995). BRAF's V599E point mutation is enriched in MSI-positive colorectal tumours, rendering BRAF constitutively active in its activation of the MAPK pathway and promoting uncontrolled cell division (Davies et al., 2002). In MMR-deficient endometrial tumours, InDels generated by MSI are enriched upstream of HR genes, bestowing

these tumours with DSB repair deficiency and potentiating their genomic instability (Zhao et al., 2014). MSI also provokes mutations in genes with a spectrum of other roles, enabling invasion and metastasis (e.g. RHAMM), promoting angiogenesis (e.g. WISP3) and inhibiting apoptosis (e.g. Caspase 5) (Duval and Hamelin, 2002; Boland and Goel, 2010).

1.3.2.3 Influence of mismatch repair status on drug sensitivity

In addition to an essential role in the repair of mismatches, the MMR system initiates apoptosis in response to some genotoxic agents to prevent the immortalisation of mutations after DNA replication (Li, 1999). These agents include 6-thioguanine (6-TG), a purine analogue used to treat leukaemia (Munshi et al., 2014; Hawn et al., 1995), and 5-fluorouracil (5-FU), a thymidylate synthase inhibitor effectively used in adjuvant therapy against CRC (Carrato, 2008). 6-TG is methylated in the cell, forming 6-meTG, which bonds with cytosine or thymine and generates mismatches. These mismatches are recognised by MMR's MutS α complex (Swann et al., 1996) and MMR initiates a c-Abl/p73a/GADD45a signalling pathway culminating in G2/M cell cycle arrest and apoptosis in response (Li et al., 2008). Recognition of 5-FU-generated lesions by MMR similarly results in cell death through this axis (Li et al., 2009). As such, these drugs rely on a proficient MMR system to induce cell death and MMR-deficient cells are resistant to 6-TG and 5-FU (Li, 1999). As clinical consequences, MMR-deficient cells are enriched in 6-TG-treated patients with autoimmune diseases (Karran and Attard, 2008), while MMR-deficient CRC cannot be treated with 5-FU (Jover et al., 2006) despite that 5-FU is considered CRC's Gold Standard treatment option (Wolpin and Mayer, 2008).

In order to circumvent such resistance in MMR-deficient cells, many studies have focused on exploiting defective MMR in cancer. As well as playing an essential role in the repair of mismatches and correction of InDels, MMR is important for the clearance of oxidative lesions, including 8-oxoguanine (8-oxo) and 8-hydroxy-2'-deoxyguanosine (8-OxdG) (Colussi et al., 2002). As a result, cells deficient in MMR or BER alone, both of which contribute to 8-oxo repair, cope with an

accumulation of toxic 8-oxo lesions. However, cells concomitantly deficient in both pathways succumb to the toxic 8-oxo. Accordingly, MSH2 depletion is synthetic lethal with inhibition of POLB, a BER polymerase (Dianov and Hübscher, 2013; Martin et al., 2010). Furthermore, methotrexate is selectively lethal to MSH2-deficient cells: methotrexate is a dihydrofolate reductase inhibitor, treatment with which results in accretion of 8-OxdG lesions. In MSH2-proficient cells, these lesions are easily removed, but in MSH2-deficient cells, the lesions accrue and mediate toxicity (Martin et al., 2009). Finally, PINK1 depletion is synthetic lethal with MMR loss. PINK1's depletion promotes cytochrome C release, causing accumulation of 8-oxo that kills MMR-negative cells (Martin et al., 2011).

1.3.2.4 Regulation of DNA mismatch repair

1.3.2.4.1 Control of mismatch repair transcription and mRNA

MMR proteins' expression must be significantly higher in S and G2 phases than in G1 to ensure the repair of DNA replication-associated errors. To achieve this, transcription of MMR genes is induced at the inception of S phase by cell cycle transcription factors such as E2F (Schroering et al., 2007). Furthermore, cells entering senescence lose MSH2 expression and develop MMR deficiency through reduction in E2F-dependent MSH2 transcription, potentially providing a mechanism by which aging cells accumulate mutations (Chang et al., 2008).

MMR proteins' expression are also responsive to DNA-damaging agents. Upon treatment with cisplatin, p53 transcribes MLH1 and PMS2 (Chen and Sadowski, 2005). In yeast, transcription of Msh2 can be induced through methyl methanesulfonate (Gasch et al., 2001), hydroxyurea (Chang et al., 2002) and camptothecin treatment (Lotito et al., 2009; Tennen et al., 2013). Moreover, after treatment of cancer cells with DNA damaging agents, Bcl-2 promotes mutagenesis by inhibiting E2F-dependent MSH2 transcription (Youn et al., 2005).

Cancer cells are able to hijack regulation of MMR to promote malignancy. In hypoxic regions of tumours, MMR transcription is downregulated in a manner depending on the histone demethylase LSD1, promoting the mutations that permit metastasis (Mihaylova et al., 2003; Koshiji et al., 2005; Lu, 2014). Furthermore, the micro-RNA miR-21 downregulates MSH2 in CRC, contributing to 5-FU resistance (Valeri et al., 2010a) and miR-21 is also activated by TGF- β in breast cancer, promoting mutations (Yu et al., 2010). Another miRNA, miR-155, is overexpressed in CRC and downregulates MSH2, MLH1 and MSH6, contributing to the onset of MSI (Valeri et al., 2010b).

1.3.2.4.2 Post-translational control of mismatch repair

MMR is also regulated by post-translational mechanisms. The Ku70-binding protein Kub5-Hera and the histone deacetylase HDAC6 act antagonistically to maintain healthy levels of MMR proteins. Kub5-Hera promotes basal stability of MLH1 and PMS2 (Patidar et al., 2016), while HDAC6 ubiquitinates MSH2, targeting MSH2 for degradation (Zhang et al., 2014b). Finally, SETD2, which catalyses methylation of lysine 36 of histone 3 (H3K36) to form H3K36me3, is required for the recruitment of MutS α to mismatches in DNA. Tumours deficient in SETD2 are therefore MMR-deficient (Li et al., 2013).

1.4 Our project

Although NUCKS has putative roles in the DDR, the impact of NUCKS on expression of key DDR proteins, cell cycle progression and basal DNA damage repair remains unknown, and we accordingly initiate our study by exploring these ideas. We then transition to explore the role of NUCKS in transcription: we utilise the SEEK coexpression database (<http://seek.princeton.edu/>) to identify genes whose expression correlates with NUCKS in cancer patient samples and cell lines. We filter this list for genes which appear in the NUCKS ChIP-Seq (Qiu et al., 2014) to generate a shortlist of putative NUCKS transcriptional targets. We find that the genes of MMR might be NUCKS targets. Motivated by this finding, we next establish whether NUCKS-depleted cells are deficient in

expression of the core MMR genes MSH2, MLH1, MSH6 and PMS2, and subsequently assess whether NUCKS-depleted cells correspondingly display resistance to the cytotoxic drug 6-TG. Finally, knowing that NUCKS expression is involved in cancer (Qiu et al., 2015) and obesity (Qiu et al., 2014), and that NUCKS might regulate MMR genes, we finish by identifying pathologies associated with genes in NUCKS' ChIP-Seq (Qiu et al., 2014).

2. Materials and methods

Materials used throughout the study are listed in the appendix.

2.1 Cell culture

2.1.1 Cell culturing and harvesting

The normal human foetal lung fibroblast cell line, TIG1, the hTERT-immortalised normal human bronchial epithelial cell line, NBE1, and the human bone osteosarcoma epithelial cell line, U2OS, were used for experiments. These cell lines were purchased from Coriell Cell Culture Repositories. Cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) GlutaMAX™ (Invitrogen) with 15% (TIG1, NBE1) or 10% (U2OS) foetal bovine serum (FBS) in a humidified air incubator (ThermoFisher Scientific) with 5% CO₂ at 37°C. Cells were passaged when they reached 80% confluence by washing with phosphate-buffered saline (PBS), treating with trypsin (0.25% trypsin, 1mM EDTA) for 5 minutes at 37°C and re-seeding at ratios of 1:3-1:12, depending on growth characteristics. When seeding for experiments, cells were counted using a Countess® Automated Cell Counter system (Invitrogen). Cells were seeded at appropriate densities that would lead to ~60-80% confluency when analysed. Cells were not used for experiments if they were older than passage 20.

To harvest cells for RNA extraction, cells were washed in PBS and then lysed and collected in buffer RLT (RNeasy Kit, Qiagen) before RNA extraction (**Section 2.2.1**). To harvest cells for protein extraction, cells were washed in PBS, detached using a cell lifter (Fisherbrand), and centrifuged at 350 relative centrifugal force (rcf) for 3 min at 4°C. Cell pellets were stored at -80°C until protein extraction (**Section 2.3.1**).

2.1.2 Transfection

Proteins levels were depleted using small interfering RNA (siRNA). For this, cells were transfected with siRNA using Lipofectamine RNAiMAX (ThermoFisher Scientific). siRNA sequences used are listed in **Table 1**. For a 10-cm dish, 200 pmol siRNA and 10 µl Lipofectamine RNAiMAX were independently added to two lots of 500 µl serum-free DMEM. These were incubated separately for 5 min at room temperature (RT) and then mixed together. After a 15 min RT incubation, during which culture medium in the 10-cm dish was exchanged for 5 ml fresh medium, the 1 ml transfection mixture was added dropwise to cells. Cells were then incubated for 72 h prior to harvesting for RNA, protein, flow cytometry (**Section 2.4.1**) or Comet assay (**Section 2.5**). Alternatively, if cells were to be treated with genotoxic agents (**Section 2.1.3**), cells were re-seeded into 96-well plates 24 h after transfection and treated at 48 h post-transfection.

Gene	Sequence	Reference
Negative control	Not disclosed	Eurogentec
NUCKS_1	5'-GAGGCGAUCUGGAAAGAAU-3'	Home designed
NUCKS_2	5'-GGCAUCUAAAGCAGCUUCU-3'	Home designed
NUCKS_3	5'-GAGGGUGUUAAGAUUAAU-3'	Home designed
NUCKS_4	5'-GAGGCAGGUAGGAGAUAAA-3'	Home designed
NUCKS_5	5'-AUAGACAAAUAAACUCAUUAU-3'	Home designed
MSH2	5'-TCCAGGCATGCTTGTTGAA-3'	Home designed

Table 1: siRNAs used throughout the study, their respective sequences and how they were designed.

2.1.3 Genotoxic agent treatments

6-TG was purchased from Sigma, prepared in 1 M NaOH and appropriate concentrations for treatments were made in cell culture medium. Cells were incubated in 6-TG-containing medium for 48 h. To assess viability, 10%-well volume of Resazurin (0.1% in PBS) (Sigma) was added for 4 h and fluorescence was read at 620 nm using a plate reader (BMG Labtech, Polar Star Omega). Upon its metabolism, Resazurin is converted to a fluorescent compound with a fluorescence emission maximum at 620 nm. Thus, if more cells are alive, there is more metabolism and a higher intensity of Resazurin fluorescence (O'Brien et al., 2000)

2.2 Detection of relative mRNA levels

2.2.1 RNA extraction

mRNA was extracted from cells using an RNeasy Kit (Qiagen). Total RNA lysates were harvested as described in **Section 2.1.1**. Buffer RLT (600 µl for a 10-cm dish) was mixed with 600 µl 70% ethanol and this mixture was transferred to an RNeasy spin column in a collection tube. Samples were centrifuged for 30 sec at 8000 rcf. The flow-through was discarded and 700 µl buffer RW1 was subsequently added to the column. After another centrifugation for 30 sec at 8000 rcf, the column was washed with two lots of 500 µl RPE buffer supplemented with ethanol. Flow-throughs were discarded and the column was placed into a fresh 1.5 ml microcentrifuge tube. 30 µl RNase-free water was added to the column, allowed to incubate for 5 min and the column was then centrifuged for 1 min at 8000 rcf, to elute the RNA. RNA was quantified using a NanoDrop1000 (ThermoFisher Scientific) at 260 nm and stored at -80°C until reverse transcription.

2.2.2 Reverse transcription

To convert RNA to cDNA for analysis by the quantitative polymerase chain reaction (qPCR), the Superscript II First Strand Synthesis system (Invitrogen) was used. First, 1 µg RNA was calculated, aliquoted and made to a volume of 6 µl with water. This RNA was mixed with 1 µl oligo dT₁₂₋₁₈ (500 µg/ml) and 1 µl dNTPs (10 mM each dNTP) and samples were incubated at 65°C for 5 min in a thermocycler (Eppendorf MasterCycler personal). Subsequently, 4 µl First-Strand Buffer (5X), 2 µl DTT (0.1 M), 2 µl MgCl₂ (5 mM) and 1 µl RNaseOUT™ (40 units/µl) were added per reaction and samples were incubated for 2 min at 42°C. 1 µl (200 units) SuperScript™ RT, the reverse transcriptase for the reaction, was subsequently added and a further incubation for 40 min at 42°C was carried out. The reverse transcriptase was then inactivated at 70°C for 15 min. 1 µl RNaseH (2 units) was added and samples were incubated for a further 20 min at 37°C to degrade residual RNA. cDNA was stored at -20°C before qPCR.

2.2.3 Quantitative polymerase chain reaction (qPCR)

cDNAs were analysed by qPCR, using the Fast SYBR™ Green Master Mix system (ThermoFisher Scientific). Primers used for qPCR are listed in **Table 2**. Per reaction, 0.5 µl forward primer (10 µM), 0.5 µl reverse primer (10 µM), 10 ng cDNA template and 10 µl 2X Fast SYBR™ Green Master Mix solution (containing AmpliTaq® Fast DNA Polymerase, SYBR® Green I dye, dNTPs, uracil-DNA Glycosylase and ROX™ dye) were mixed and made to a final volume of 20 µl with water. Reactions were performed in a high sided, low profile PCR plate (ThermoFisher Scientific) sealed with an optical adhesive cover (Applied Biosystems). The 7500 Fast Real-Time PCR System (Applied Biosystems) was used for analysis. The following cycling conditions were used: 15 min at 95°C; 40 cycles: 15 sec at 95°C, 30 sec at 60°C, 30 sec at 72°C; melting curve analysis: 15 sec at 95°C, 1 min

at 60°C, 30 sec at 95°C and 15 sec at 60°C. The comparative CT method was used to quantify relative mRNA levels.

Primer	Sequence (5'-3')	
NUCKS	F: TCTGATGATGCAGATGAAGATTA	R: CTGCTGAGTGAGAATCATCC
MSH2	F: GGATCAAATGAAGGTTTTACAAGG	R: TTCTGACTTCTCCAAGTTTCAGG
MSH6	F: AGGGGTAACCCTCCATCTTG	R: AAGGCGAAGAACCTCAACG
MLH1	F: AACTTGGTTTGATGCTGTGC	R: TTAATGAGCAGGGACATGAGG
PMS2	F: AACAAAATCAAAGCCATTCTTTCT	R: CAGGGGCAGAGGCTCATAG
CSE1L	F: GGCAGGATCAGGATCAAGTG	R: CGTTTCTGTATCCCCACGAG
PLK3	F: TCTGGAGCCACATAGTTGGG	R: AACATGGAACTGAAGGTGGG
B2M	F: ATGTCTCGCTCCGTGGCCTTA	R: ATCTTGGGCTGTGACAAAGTC
GAPDH	F: AGCCACATCGCTCAGACAC	R: GCCCAATACGACCAAATCC
TBP	F: CGGTTTGCTGCGGTAATCAT	R: TTTCTTGCTGCCAGTCTGGAC

Table 2: qPCR primers used throughout the study, and their respective sequences.

2.3 Detection of relative protein levels

2.3.1 Protein extraction

To perform protein extraction, a modified version of Tanaka's method was used (Tanaka et al., 1992). Tanaka's method employs a buffer I (10 mM Tris-HCl pH 7.8, 200 mM KCl) to resuspend cell pellets followed by a buffer II (10 mM Tris pH 7.8, 600 mM KCl, 2 mM EDTA, 40% glycerol, 0.2% Nonidet P-40) to lyse cells. A protease inhibitor cocktail (1 mM PMSF, 1 mM NEM, 1 µg/ml aprotinin, pepstatin, chymostatin and leupeptin) was added to both buffers. Cell pellets were resuspended in 1 x packed cell volume (PCV) Tanaka buffer I and then 2 x PCVs of Tanaka buffer II were added. Cell suspensions were subsequently incubated for 30 min at 4°C and centrifuged at 16200 rcf for 30 min at 4°C. The supernatant, containing proteins, was stored at -80°C until SDS-PAGE (**Section 2.3.2**). Concentrations of extracted proteins were measured using the

Bio-Rad Protein Assay (Bio-Rad), a version of the Bradford assay (Bradford, 1976). Proteins were mixed with Bio-Rad assay reagent and the optical density of the Coomassie brilliant blue dye, which changes colour upon binding amino acid residues, was measured at 595 nm using an Ultrospec 2000 UV spectrophotometer (Pharmacia). Concentrations were calculated using bovine serum albumin as a reference.

2.3.2 Polyacrylamide gel electrophoresis (PAGE)

Subsequently, proteins from whole cell extracts (WCEs) were separated by SDS-PAGE. Polyacrylamide gels were made using the following recipe, adjusting the acrylamide content depending on the size of proteins of interest. For a 12% gel: 3.77 ml 1 M Tris HCl (pH 8.8), 0.1 ml 10% (w/v) SDS, 40 μ l 0.5 M EDTA, 4 ml 30% acrylamide/0.8% bisacrylamide, 2.02 ml water, 60 μ l (w/v) 10% APS and 7.5 μ l TEMED were mixed, added to gel cassettes (BIO-RAD), and allowed to polymerise. A 5% stacking gel, made up of 0.3125 ml 1M Tris HCl (pH 6.8), 0.025 ml 10% (w/v) SDS, 10 μ l 0.5 M EDTA, 0.415 ml 30% acrylamide/0.8% bisacrylamide, 1.72 ml water, 15 μ l 10% (w/v) APS and 1.875 μ l TEMED was added on top of the polymerised running gel portion. Using the concentrations established from the Bradford assay, 20-30 μ g protein was measured and added to a final volume of 20 μ l with water. 5 μ l Laemmli loading buffer (5X: 60 mM Tris-HCl pH 6.8, 2% SDS, 10% glycerol, 5% β -mercaptoethanol, 0.25% bromophenol blue) was added per sample, samples were boiled at 95°C for 5 min to denature proteins and loaded onto 12% denaturing polyacrylamide gels. Electrophoresis was performed at 125 V for 2 h in running buffer (25 mM Tris, 192 mM glycine, 0.1% (w/v) SDS, pH 8.3). The All-Blue Precision Plus Protein Standards (Bio-Rad) was used as a protein molecular weight marker.

2.3.3 Western blotting

Separated proteins from the polyacrylamide gel were then transferred to Immobilon-FL poly(vinylidene fluoride) (PVDF) membranes: membranes were first activated for 30 s in 100% methanol and then washed in transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol). To facilitate the formation of a tight sandwich and efficient gel-membrane protein transfer, sponges and 3 mm filter paper were assembled along with the gel and activated membrane in transfer apparatus. Transfer was performed at 25 V for 2 h in transfer buffer.

Nonspecific antigens on the freshly-transferred membrane were then blocked using Odyssey Blocker buffer (LI-COR) for 45 min at RT, using a low profile roller (Stovall). Blocked membranes were incubated in the relevant primary antibody, diluted in Odyssey Blocking Buffer (**Table 3**), overnight at 4°C. Membranes were then washed twice in PBS-T (PBS/0.1% Tween-20) and then incubated with secondary antibodies for 1 h at RT (**Table 3**). Membranes were then washed in PBS-T twice, then PBS and finally scanned (LI-COR).

Antibody	Reactivity	Clonality	Dilution	Application	Source
Anti-NUCKS	Rabbit	Polyclonal	1:1000	Flow cytometry	ThermoFisher Scientific
Anti-p53	Mouse	Monoclonal	1:500	Western blot	Santa Cruz
Anti-p21	Rabbit	Monoclonal	1:2000	Western blot	Cell Signaling
Anti- β -actin	Mouse	Monoclonal	1:10000	Western blot	Abcam
AlexaFluor 680/800nm	Rabbit	Polyclonal	1:10000	Western blot	Invitrogen
AlexaFluor 680/800nm	Mouse	Polyclonal	1:10000	Western blot	Invitrogen
Alexa 488	Rabbit	Polyclonal	1:250	Flow cytometry	Invitrogen

Table 3: Antibodies used throughout the study, their reactivity, clonality, dilution ratio, application in this study, and their respective providers.

2.4 Flow cytometry

2.4.1 Cell cycle analysis

72 h after siRNA-mediated knockdown (**Section 2.1.2**), cells were detached using trypsin and resuspended in 15% FBS-containing DMEM. Cells were centrifuged at 250 rcf for 5 min at 4°C, washed in PBS and then centrifuged again. The pellet was fixed by dropwise addition of cold 70% ethanol and cells were then incubated for 30 min on ice. Cells were then centrifuged at 250 rcf for 5 min at 4°C and pellets were resuspended in 100 μ g/ml RNase in PBS and incubated at 37°C for 30 min to degrade RNA, which propidium iodide also stains. 20 μ g/ml propidium iodide was then added directly to the cell suspension for 15 min to stain DNA. Propidium iodide is a fluorescent dye which binds DNA, and its intensity is therefore proportional to DNA content. DNA content in G2/M

phase is double that in G0/G1 phase (4N vs. 2N), while S phase cells have intermediate DNA content. A FACSort flow cytometer (BD) was used to measure fluorescence at 617 nm, and analysis of cell cycle phases was performed using ModFit software.

2.4.2 Protein levels analysis

72 h after siRNA-mediated knockdown, cells were detached using trypsin, washed with PBS and fixed in 4% formaldehyde at RT for 10 min. Cells were then centrifuged for 5 min at 300 rcf and 4°C, washed in cold PBS and centrifuged again. Pellets were resuspended in PBS-1% FBS, centrifuged, and cells permeabilised by resuspension in 100 µl PBS-1% FBS-0.1% Triton-X 100 for 5 min at RT. After another centrifugation, the pellet was resuspended in 100 µl primary antibody and incubated for 1 h at RT (**Table 3**), followed by a PBS wash and then incubation with 100 µl secondary antibody for 1 h at RT (**Table 3**). Finally, fluorescence at 488 nm was read using a FACSort flow cytometer (BD).

2.5 Alkaline Comet assay

To assess apurinic/apyrimidinic sites, SSBs and DSBs in DNA, the alkaline Comet assay was used (Olive and Banath, 2006). Cells treated with siRNA for 72 h were detached using trypsin, resuspended in PBS-FBS 1% and counted as described in **Section 2.1.1**. 50,000 cells were lysed for 1 h at 4°C in lysis buffer (2.5 M NaCl, 100 mM EDTA, 10 mM Tris HCl, pH 8.0) followed by incubation for 30 min in running buffer (300 mM NaOH, 1 mM EDTA, 1 % DMSO, pH 13.0), protected from light. After this, single cell gel electrophoresis was carried out in a Comet-20 electrophoresis tank, in running buffer, at 25 V and 300 mA for 25 min. After neutralisation in neutralisation buffer (500 mM Tris-HCl, pH 8.0), DNA was stained with SYBR gold (ThermoFisher Scientific) for 30 min, protected from light. Imaging and analysis were performed using a Nikon Eclipse 90i microscope and Kinetic Imaging Comet 5.5 software (Andor Technology), respectively. Percentage tail DNA was used to measure DNA damage, and at least 100 cells were analysed per sample.

2.6 Bioinformatics analysis

2.6.1 SEEK online gene expression database

The SEEK coexpression search engine (<http://seek.princeton.edu/>) was used to identify genes whose expression correlates with that of NUCKS (**Table 5**). The SEEK database scrapes data from thousands of RNA microarrays and RNA sequencing studies from the publically-available Gene Expression Omnibus (<https://www.ncbi.nlm.nih.gov/geo/>). SEEK calculates scores for gene coexpression using the standard Pearson correlation (Eisen et al., 1998). SEEK considers coexpression scores above + 0.5 / below - 0.5 to indicate strong positive / negative correlation. SEEK derives p-values by calculating the probability that changes in expression of genes of interest are more reliable than the changes of a randomly-changing background of genes.

2.6.2 G-Profiler online disease database

The G-Profiler online database (<http://biit.cs.ut.ee/gprofiler/index.cgi>) was used to analyse genes in the NUCKS ChIP-Seq (Qiu et al., 2014) (**Table 6**). This database utilises information from Human Phenotype Ontology, which lists genes whose mutation or deregulated expression is associated with pathology. The Human Phenotype Ontology itself derives information from the Online Mendelian Inheritance in Man knowledgebase (Hamosh et al., 2005).

2.7. Statistical analysis

p-values were assessed using a two tailed Student's t-test. In figure legends, ns denotes non-significant, * p-value <0.05, **p-value <0.01 and ***p-value < 0.001.

3. Results

Throughout this study, we perform our most fundamental experiments using a panel of three cell lines. Because the DDR can differ from normal primary cell lines to immortalised or cancer cell lines (e.g. Poletto et al., 2016), we use the normal human foetal lung fibroblast cell line, TIG1, the hTERT-immortalised normal human bronchial epithelial cell line, NBE1, and the human bone osteosarcoma epithelial cell line, U2OS. We also aim to verify our results using at least two sequences of siRNA, in order to reduce the chances that our observed phenotype is due to off-target effects.

3.1 Identification of NUCKS as a new responder to DNA damage

3.1.1 NUCKS is a putative ATM target upregulated upon DNA repair deficiency

We initiated our study by searching our group's SILAC data (Markkanen et al., 2015) for proteins upregulated in the nucleus upon loss of BER that were also phosphorylation targets for the DDR kinase ATM. Since many DDR proteins are induced by DNA damage (e.g. Christmann and Kaina, 2013; Lakin and Jackson, 1999) and since ATM plays a central role in the DDR (Shiloh and Ziv, 2013), we expected these targets might be novel DDR players. This gave us a shortlist of five proteins: PNISR, PALLD, NUCKS, MLLT4 and RCC2 (**Table 4**). We performed literature searches of these proteins and found the NUCKS protein to be the most enticing target: NUCKS has poorly defined functional roles despite its discovery in 1985 (Ostvold et al., 1985), is deregulated in many cancers (Qiu et al., 2015) and is a putative transcriptional regulator with a recently published ChIP-Seq (Qiu et al., 2014).

Protein	Fold upregulation	Function
PNISR	4.08	RNA binding
PALLD	2.09	Cell shape and adhesion
NUCKS	1.80	Transcription
MLLT4	1.55	Cell adhesion
RCC2	1.53	Mitosis

Table 4: Nuclear ATM targets upregulated in XRCC1 SILAC.

The five proteins upregulated in the nucleus after depletion of XRCC1 (Markkanen et al., 2015) which are also phosphorylation targets of ATM. Adapted from Markkanen et al., 2015.

3.2 Consequences of NUCKS depletion on the DDR

3.2.1 NUCKS siRNA sequences validation

First, we aimed to characterise previously unexplored impacts of NUCKS on the DDR. Although NUCKS is known to be modulated in response to DNA damage (Matsuoka et al., 2007; Markkanen et al., 2015) and promotes DSB repair by HR (Parplys et al., 2015), it remains unknown how NUCKS depletion impacts expression of central DDR proteins, cell cycle progression, or basal DNA damage. To start, we screened 5 sequences of NUCKS siRNA (**Table 1**) in U2OS cells and established knockdown efficiency by qPCR (**Fig. 2A**). We were able to achieve highly effective knockdown with sequences #1 and #2 of NUCKS siRNA, and reasonable knockdown with 3rd and 4th sequences. A 5th sequence did not achieve any NUCKS depletion (**Fig. 2A**). In the absence of a suitable NUCKS antibody for use by Western blot (data not shown), we confirmed depletion of NUCKS on the protein level with NUCKS siRNAs #1 and #2 by flow cytometry (**Fig. 2B**). These siRNA sequences achieved 70% and 67% protein depletion relative to NC siRNA-treated cells, respectively, and were selected to be used throughout the study.

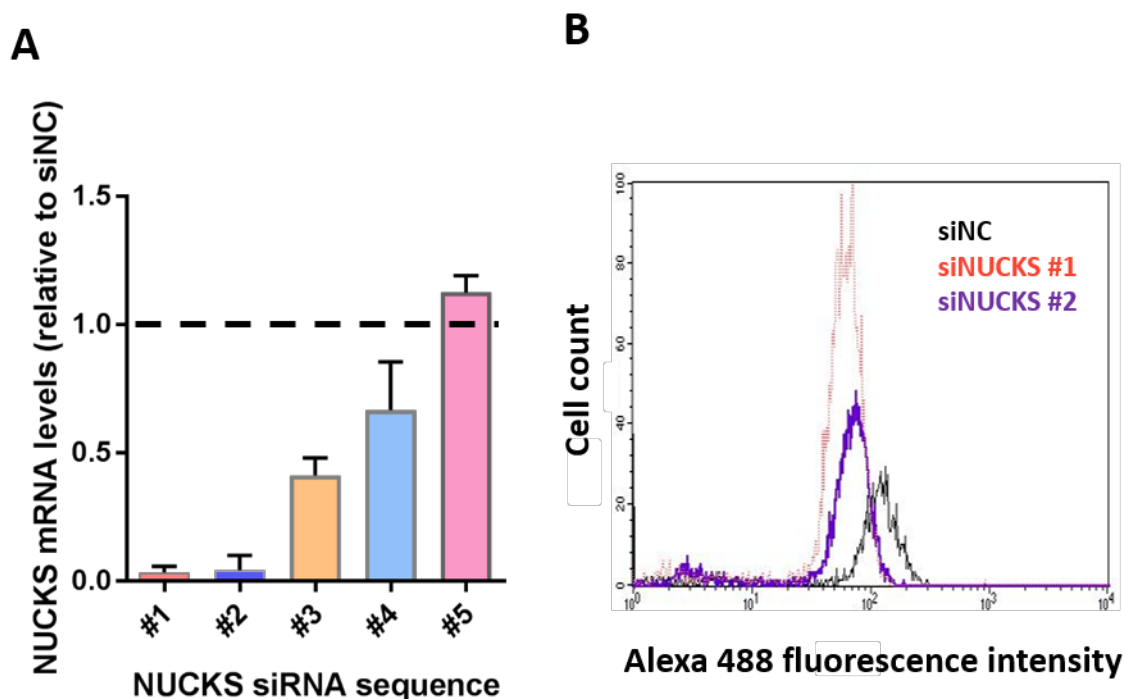


Figure 2: Validation of NUCKS siRNA sequences

A) U2OS cells were treated with sequences #1-#5 of NUCKS siRNA for 72 h and NUCKS mRNA levels, which are presented relative to NC siRNA-treated cells, were assessed by qPCR, using B2M and GAPDH as reference genes. Error bars represent standard deviations of three independent experiments.

B) Flow cytometry, using a NUCKS primary antibody, was used to determine protein depletion after treatment with NUCKS siRNA #1 and #2. The X axis represents fluorescence intensity of the Alexa 488 secondary antibody targeted to the NUCKS primary antibody. The leftward shift of the knockdown relative to the control shows that the knockdown was successful.

3.2.2 NUCKS knockdown induces p53 and p21, key DDR proteins

The DDR is coordinated by an intricate series of interconnecting proteins; these proteins are upregulated or activated in response to DNA damage and cellular stress, preventing transmission of mutated DNA to daughter cells and slowing down the progression of cancer (Zhou and Elledge, 2000). Key DDR components include p53 – “The Guardian of the Genome” – a transcription factor deregulated in the majority of human cancers (Lane, 1992) and p21, a p53 transcriptional target which responds to physiological stress by inhibiting CDKs and halting progression of the cell cycle (Abbas and Dutta, 2009). We established the impact of NUCKS depletion on these DDR players by

treating TIG1, NBE1, and U2OS cells with NUCKS siRNAs and used Western blotting to assess p53 and p21 expression levels (**Fig. 3**). In TIG1, NBE1 and U2OS cells, p53 and p21 were upregulated after NUCKS depletion. Further targets would have to be explored (e.g. p-ATM, PAR, pATR, γ H2AX) to conclude whether NUCKS depletion activates the DDR.

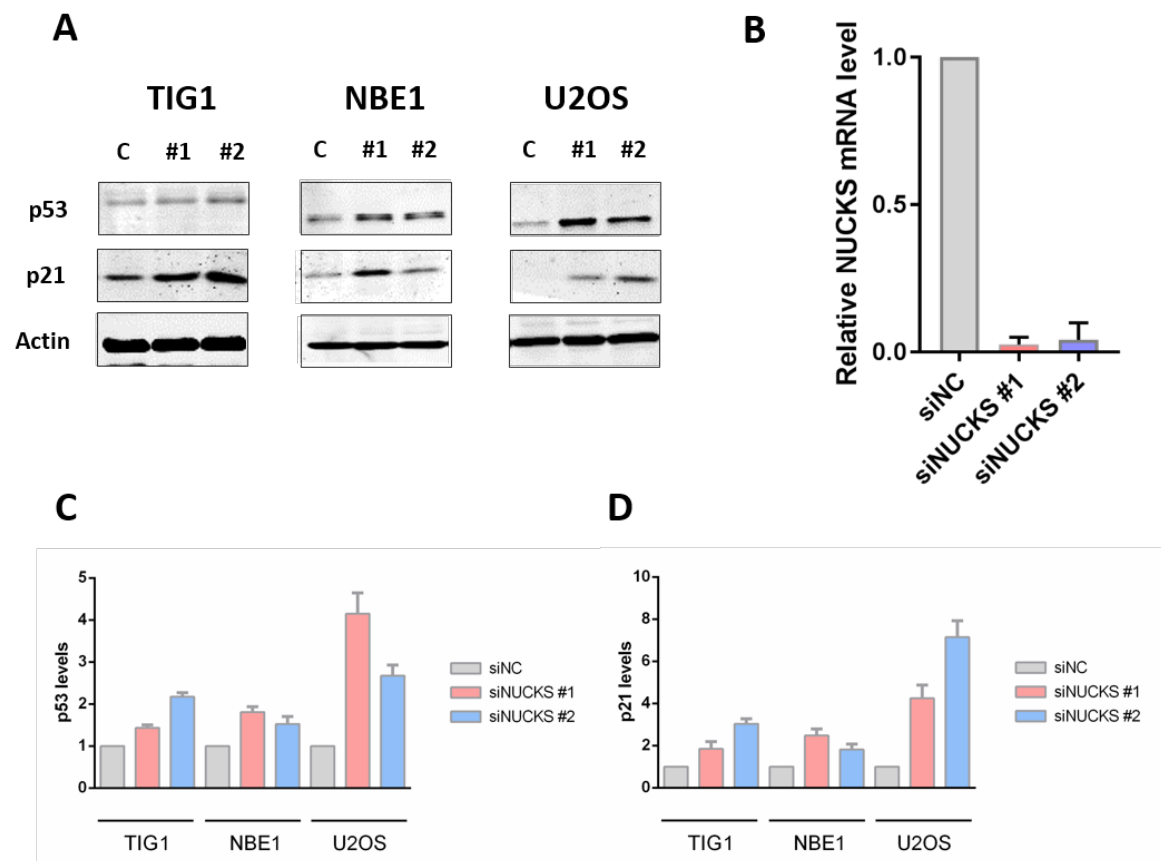


Figure 3: NUCKS knockdown induces p53 and p21

A) TIG1, NBE1 and U2OS cells were treated with NC siRNA or NUCKS siRNA for 72 h. WCEs were loaded on 12% polyacrylamide gels and p53 / p21 levels were analysed by Western blot, using actin as a loading control. Data shown are representative of two independent experiments. **B)** Cell pellets in **A)** were split for RNA extraction and qPCR to confirm NUCKS depletion, using B2M and GAPDH as reference genes (U2OS shown). **C)** and **D)** Quantifications of p53 and p21 levels in Western blots in **A**, normalised to actin expression. Error bars represent standard deviations of two independent experiments.

3.2.3 NUCKS knockdown arrests the cell cycle in G1 phase

One consequence of p53/p21 upregulation is the activation of a G1/S cell cycle checkpoint, whose goal is to prevent cells with damaged DNA from undergoing DNA replication (Kastan and Bartek, 2004). To explore whether NUCKS knockdown activates this checkpoint, we used flow cytometry to assess the cell cycle after NUCKS depletion (**Fig. 4**). In normal cells (TIG1 and NBE1), we found that NUCKS siRNA-treated cells dramatically lost S phase and arrested in G1 phase (**Fig. 4A-C**), consistent with the upregulation of p53 and p21 in these cells (**Fig. 3**). In U2OS cells, sequence #2 of NUCKS siRNA induced a robust G1 arrest, while cells treated with sequence #1 of NUCKS siRNA displayed cell cycle profiles that are comparable with NC siRNA-treated cells (**Fig. 4A, D**). U2OS cells treated with a third sequence of siRNA for NUCKS (#4) also displayed normal cell cycle profiles (**Fig. 4D**), suggesting that loss of NUCKS in cancer cells is less associated with cell cycle arrest. This is in agreement with the G1/S cell cycle checkpoint being among the most deregulated signalling axes in cancer (Sherr and McCormick, 2002).

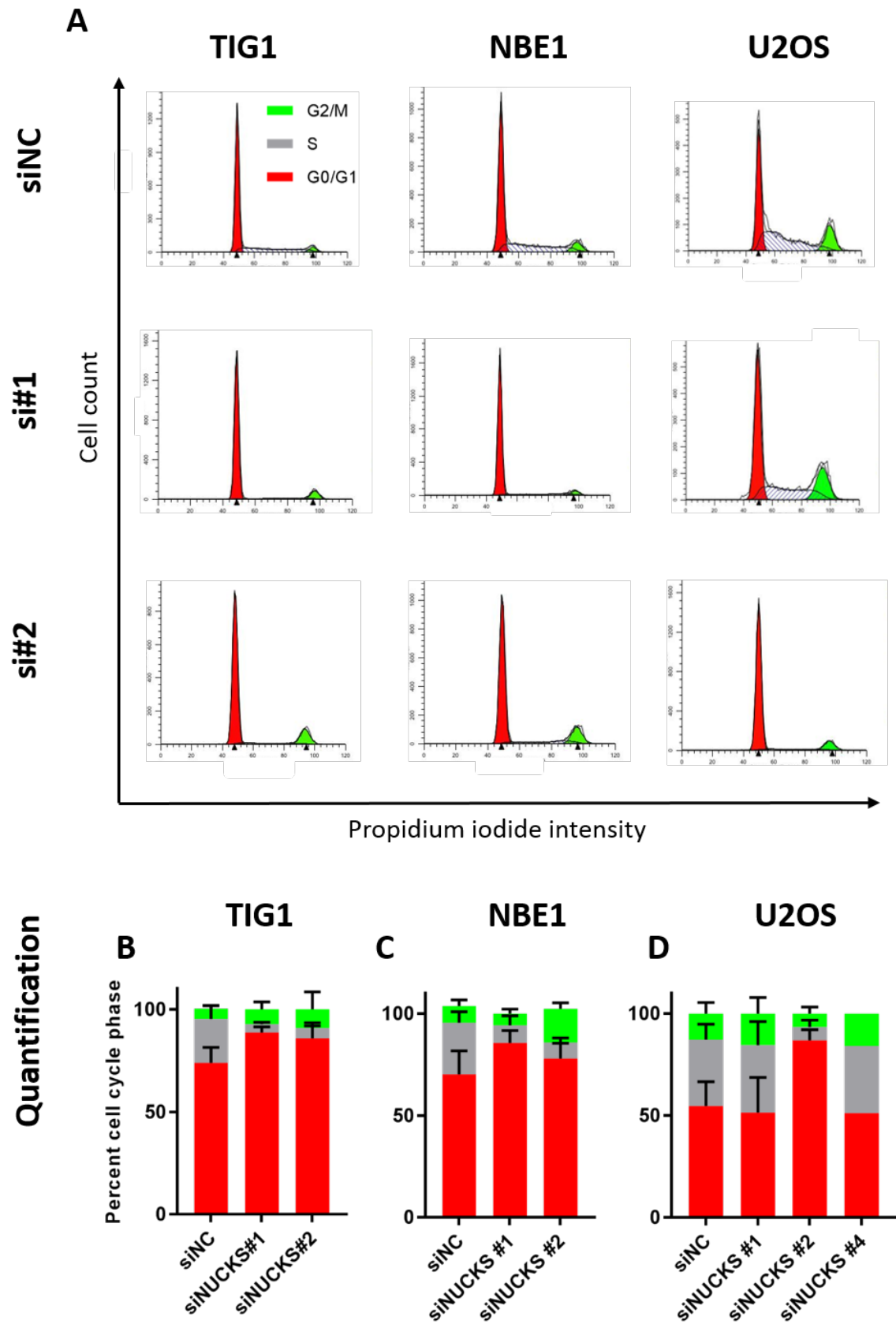


Figure 4: NUCKS knockdown induces G1 cell cycle arrest in normal cells, but not in cancer cells
A) TIG1, NBE1 and U2OS cells were treated with NC siRNA or NUCKS siRNA (#1, 2 or 4) for 72 h and cell cycle phases were analysed by flow cytometry, using propidium iodide staining. Representative cell cycle profiles of three independent experiments are shown. **B), C) and D)** Quantification of flow cytometry data for TIG1, NBE1 and U2OS cells in **A)**, respectively. Error bars represent standard deviations of three independent experiments (U2OS cells treated with NUCKS siRNA #4 is only one experiment).

3.2.4 NUCKS knockdown might be associated with DNA damage in cancer cells

To understand the mechanism underpinning the upregulation of p53 and p21 (**Fig. 3**) and the activation of the G1/S cell cycle checkpoint (**Fig. 4**) in NUCKS-depleted cells, we assessed whether NUCKS depletion is associated with DNA damage. To this end, we utilised the alkaline Comet assay, which detects apurinic/apyrimidinic sites, SSBs and DSBs in DNA (Olive and Banath, 2006). In normal cells (TIG1), we did not find that NUCKS depletion caused an increase in the DNA damage load (**Fig. 5A**). In fact, in this single experiment, NUCKS-depleted TIG1 cells had less DNA damage than control cells – probably because these cells are arrested in G1 phase (**Fig. 4**). However, in U2OS cells, although the error is large, NUCKS depletion seemed to result a moderate increase in DNA lesions (**Fig. 5B**). Although these data require further confirmation through additional repeats and employment of other cell lines, it appears that normal cells, which have accomplished DNA repair systems, stable genomes and functioning cell cycle checkpoints, do not rely on NUCKS to maintain genomic stability. In contrast, in cancer cells, which are more addicted to DNA repair and more prone to DNA damage (Hanahan and Weinberg, 2011), loss of NUCKS might be associated with an increase in DNA damage.

To conclude, NUCKS appears to display many hallmarks of a DDR protein: NUCKS deficiency causes sensitivity to ionising radiation and a defect in DSB repair (Parpys et al., 2015), NUCKS is upregulated (Markkanen et al., 2015) and phosphorylated (Matsuoka et al., 2007) in response to DNA damage, its depletion induces key DDR proteins p53 and p21 (**Fig. 3**), activates the G1/S checkpoint (**Fig. 4**), and could be associated with accumulation of DNA damage in cancer cells (**Fig. 5**). Since we suspected that NUCKS' putative role in the DDR might be coordinated through its

transcriptional targets, we next transitioned to establish what these targets might be, and whether any might have roles in DNA repair.

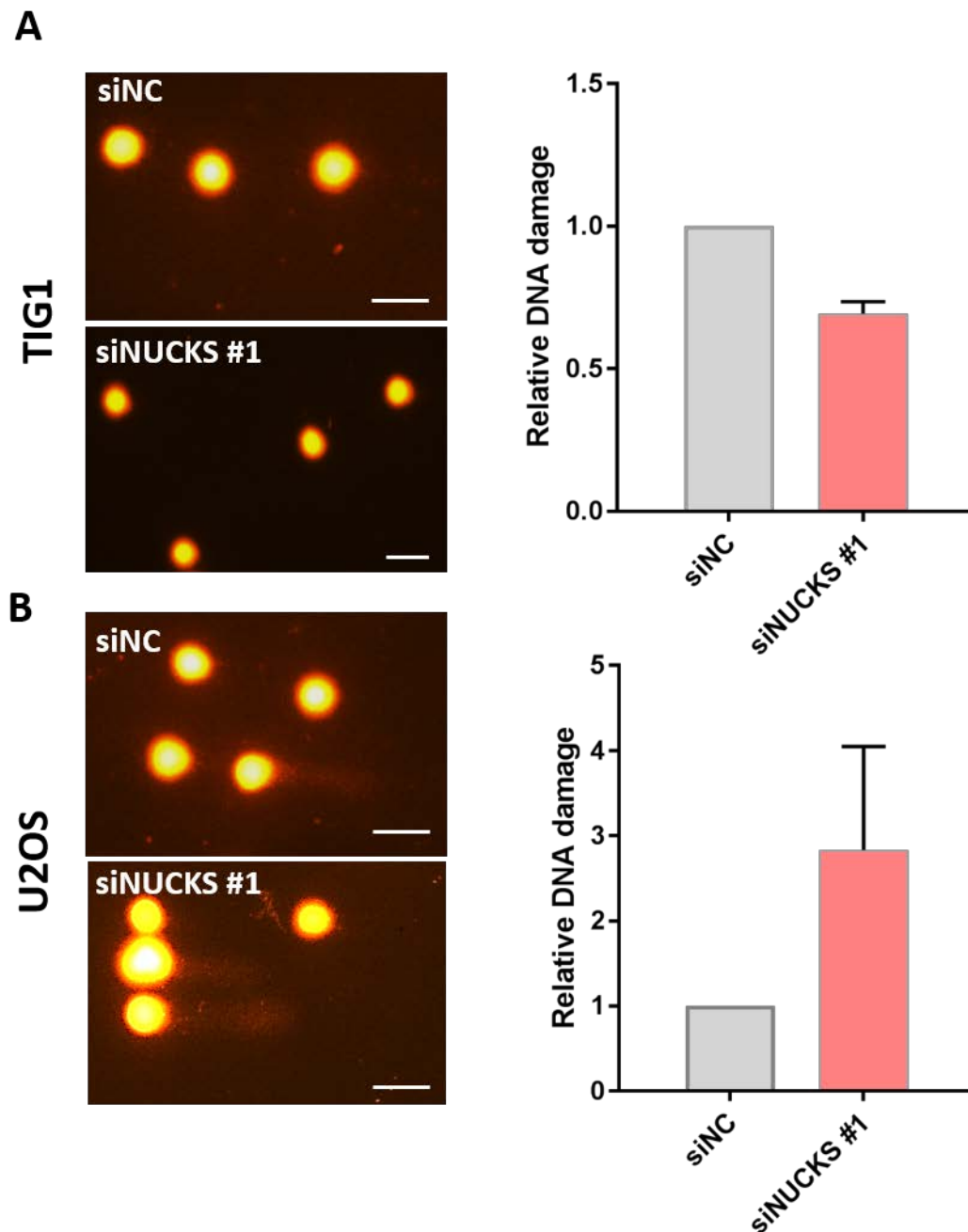


Figure 5: NUCKS knockdown appears to induce a small increase in DNA damage in U2OS cells (A) TIG1 and (B) U2OS cells were treated with NC siRNA or siRNA for NUCKS (#1) and DNA damage was analysed using the alkaline Comet assay. DNA damage was scored on the average Comet tail length of at least 50 cells per sample, using two samples per condition. Left: representative Comets, right: quantifications. Error bars represent standard deviations of two technical replicates. Scale bar, 50 μ m.

3.3 NUCKS and the transcriptional regulation of DNA repair: mismatch repair genes

3.3.1 NUCKS correlates with DNA repair, cell cycle and transcription-controlling genes

Knowing that NUCKS regulates transcription (Qiu et al., 2014; Kim et al., 2014), and motivated by the recent CHIP-Seq report for NUCKS (Qiu et al., 2014), we sought to use a bioinformatics approach to identify potential transcriptional targets for NUCKS with roles in the DDR. To this end, we used the SEEK coexpression database (<http://seek.princeton.edu/>), which contains gene expression data of 4929 mRNA microarray and 281 RNASeq studies from cancer patient samples and cell lines. These data are organised into 16 tissue types, separated generally into data from four categories: patient non-cancer samples, patient cancer samples, patient tumour (cancer + stroma) samples, and cell line samples – totalling 67 datasets. SEEK generates coexpression scores using the standard Pearson correlation (Eisen et al., 1998), ranging from -3 for perfectly negatively-correlating genes to +3 for perfectly positively-correlating genes. A score of zero indicates no detectable correlation. In each of SEEK's 67 datasets individually, we extracted lists of the top 100 genes with the highest and lowest coexpression scores with NUCKS. Next, we filtered the resulting lists for genes appearing in the NUCKS CHIP-Seq (Qiu et al., 2014) and ranked genes by the percentage of the 67 datasets in which they were among the top positively (**Table 5A**) / negatively (**Table 5B**) coexpressed genes. Many of these targets have roles in the DDR (e.g. PPP1R15A), cell cycle regulation (e.g. CSE1L), and transcription (e.g. YY1). The top positive-correlation hit is XPO1, a protein involved in the control of nuclear export (Hutten and Kehlenbach, 2007), while the top negative-correlation hit is CSF1, a cytokine known to induce differentiation of immune cells (Hume and MacDonald, 2012).

However, we were most intrigued by MSH2, an MMR protein appearing among the most 100 positively coexpressed genes in 20% of databases (**Table 5A**). Furthermore, NUCKS is reported to bind the promoters of MLH1 and PMS2, two other core MMR factors (Qiu et al., 2014), and, taking SEEK's 67 datasets altogether, the MMR proteins' coexpression scores with NUCKS are 0.9572 for

MSH2, 0.7188 for MLH1, 0.2339 for PMS2 and 1.0012 for MSH6. This means that these genes have the 94th, 594th, 4612th, and 61st highest coexpression scores with NUCKS, respectively, of 17780 total genes. Motivated by these striking findings, we next sought to explore NUCKS' putative transcriptional regulation of MMR genes.

A

Gene	Function	% databases top 100 (+ve)	All datasets		
			Coexpression score (-3 to +3)	Rank (of 17,780)	P-value
XPO1	Nuclear export	52	1.0774	22	0.0277
HNRNPA1	Pre-mRNA packaging	34	0.9975	64	0.0018
NCL	Ribosome synthesis	30	0.9974	65	0.0148
CSDE1	RNA-binding protein	30	0.9645	89	0.0054
PTMA	Immune function	28	1.0002	62	0.0011
YY1	Transcription factor	28	0.9396	106	0.0023
CSE1L	Cell cycle, DSB repair	25	0.9880	72	0.0637
DROSHA	RNA maturation/decay	23	0.9189	133	0.0298
MSH2	Mismatch repair	20	0.9572	94	0.0482
KPNB1	Nuclear import	19	0.8523	241	0.0269

B

Gene	Function	% databases top 100 (-ve)	All datasets	
			Coexpression score (-3 to +3)	Rank (of 17,780)
CSF1	Immune system	31	-0.6139	19
PLK3	Cell cycle, DSB repair	20	-0.5776	57
IRF5	Immune system	19	-0.5336	186
HNF4A	Transcription factor	13	-0.5731	64
PPP1R15A	DDR	13	-0.5391	161
FOSL2	Proliferation	13	-0.5109	300
JAK3	Immune system	13	-0.4978	390
BCL3	Proliferation	13	-0.4943	425
CRTC1	Transcription	11	-0.5803	50
RELA	Transcription	11	-0.4612	840

Table 5: NUCKS ChIP-Seq genes that correlate with NUCKS expression in patient samples and cell lines.

The online SEEK coexpression database was used to identify genes whose expression correlates with that of NUCKS. Genes, which also appear in the NUCKS ChIP-Seq, were ranked by the percentage of 67 databases in which they were among the top 100 **A)** positively or **B)** negatively correlating genes, giving us well-founded lists of potential NUCKS transcriptional targets. Also reported is a coexpression score across all 67 datasets, and ranks and p-values based on this score (although SEEK does not report a p-value for negatively coexpressed genes).

3.3.2 NUCKS is required to maintain expression of mismatch repair genes

In our exploration of NUCKS' transcriptional regulation of MMR, we first used qPCR to assess the expression levels of MSH2, MLH1, MSH6 and PMS2 after NUCKS siRNA treatment, in TIG1, NBE1 and U2OS cells (**Fig. 6**). Our bioinformatics analysis showed that the expression of these genes correlates positively with NUCKS expression (**Table 5**), suggesting their expression might decrease after NUCKS knockdown. As a proof of principle, we also tested expression of chromosome segregation 1-like (CSE1L), a protein with roles in proliferation, apoptosis (Behrens et al., 2003), and HR repair (Okimoto et al., 2015), which positively correlates with NUCKS expression (**Table 5A**), and polo-like kinase 3 (PLK3), a protein with roles in DSB repair (Barton et al., 2014) and cell cycle regulation (Wang et al., 2002), which negatively correlates with NUCKS expression (**Table 5B**). We found a marked reduction in the mRNA levels of MSH2, MLH1 and MSH6 after NUCKS depletion, while PMS2 levels seemed unaffected (**Fig. 6**) – confirming SEEK data that PMS2 expression poorly correlate with NUCKS. Furthermore, confirming the expectations of our bioinformatics analysis, CSE1L levels were decreased after NUCKS siRNA treatment, while PLK3 levels were increased (**Fig. 6**).

Three lines of evidence, therefore, point strongly towards a role for NUCKS in the transcriptional regulation of MMR: 1) NUCKS binds the promoters of MSH2 and MLH1 (Qiu et al., 2014) 2) NUCKS expression correlates very strongly with the expression of MSH2, MLH1 and MSH6 in patient samples and cell lines (**Table 5**) 3) NUCKS siRNA reproducibly leads to loss of MSH2, MLH1 and MSH6 expression across multiple cell lines (**Fig. 6**).

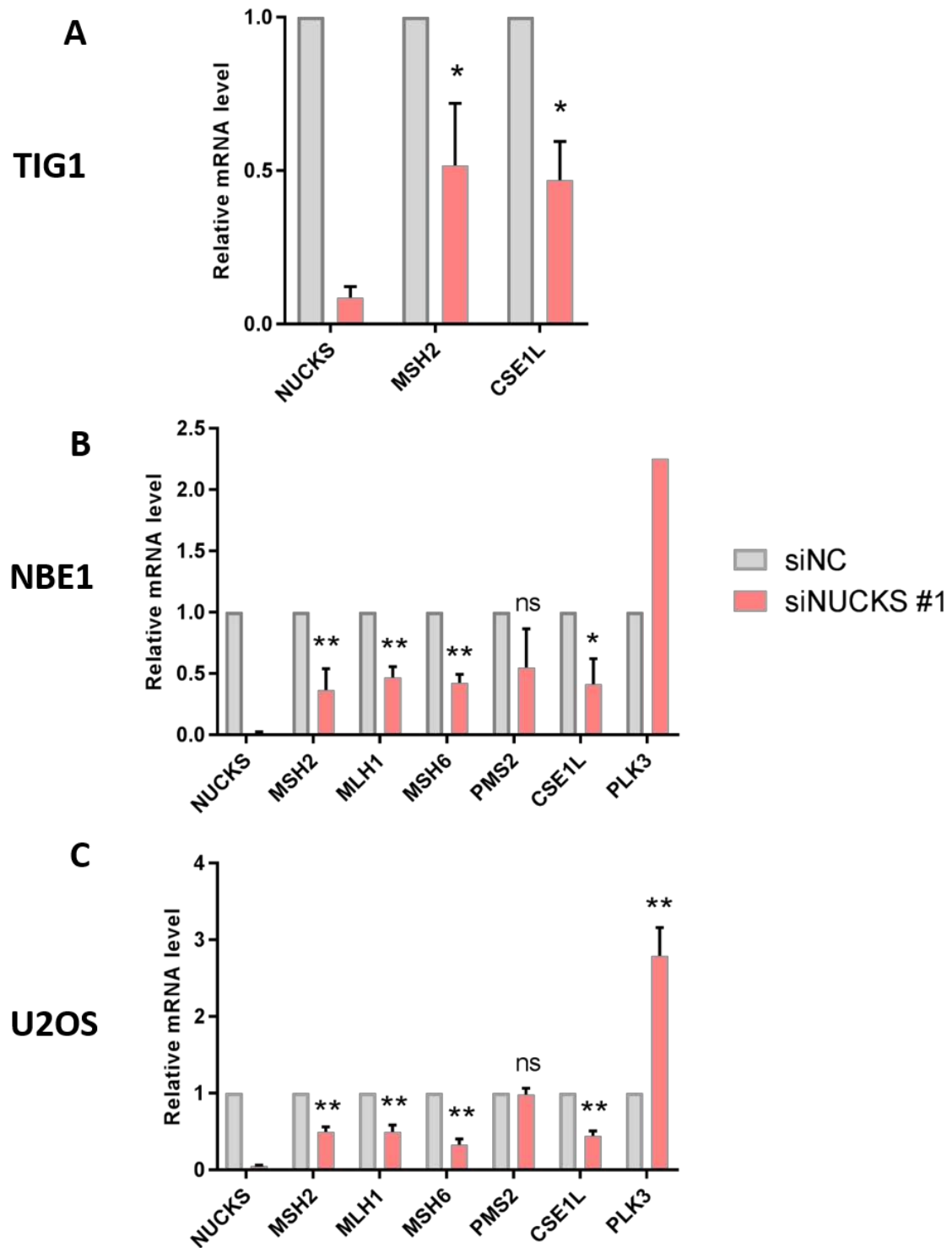


Figure 6: NUCKS knockdown results in deregulated transcription

(A) TIG1, (B) NBE1 and (C) U2OS cells were treated with NC or NUCKS siRNA (#1) for 72 h before analysis of mRNA levels by qPCR, using B2M and GAPDH as reference genes. Error bars represent standard deviations of at least three independent experiments. ns, non significant; *, p-value < 0.05; **, p-value < 0.01.

3.3.3 NUCKS' regulation of mismatch repair gene MSH2 is not cell cycle-dependent

MMR genes, including MSH2, are cell cycle-regulated: MSH2 expression is lower during G1 phase than S and G2 phases, in accordance with a role in the repair of mismatches formed during DNA replication (Jiricny, 2006). Therefore, it could be that MSH2 expression decreases after NUCKS siRNA treatment because NUCKS siRNA induces cell cycle arrest in G1 phase (**Fig. 4**). To address this issue, we synchronised TIG1 and NBE1 cells in G1 phase of the cell cycle by contact inhibition, removing cell cycle effects from the equation (**Fig. 7A-D**). We then treated cells with NC or NUCKS siRNA and found that these G1-arrested cells still lost MSH2 expression commensurate with asynchronous cells (**Fig. 7E and F**). Moreover, we have shown that U2OS cells treated with NUCKS siRNA #1 display cell cycle profiles that are similar to NC siRNA-treated cells (**Fig. 4**), but MSH2, MLH1 and MSH6 remain downregulated after NUCKS siRNA #1 treatment in these cells (**Fig. 6C**). These two lines of evidence provide strong support that NUCKS is a bona fide regulator of MMR transcription, irrespective of cell cycle effects.

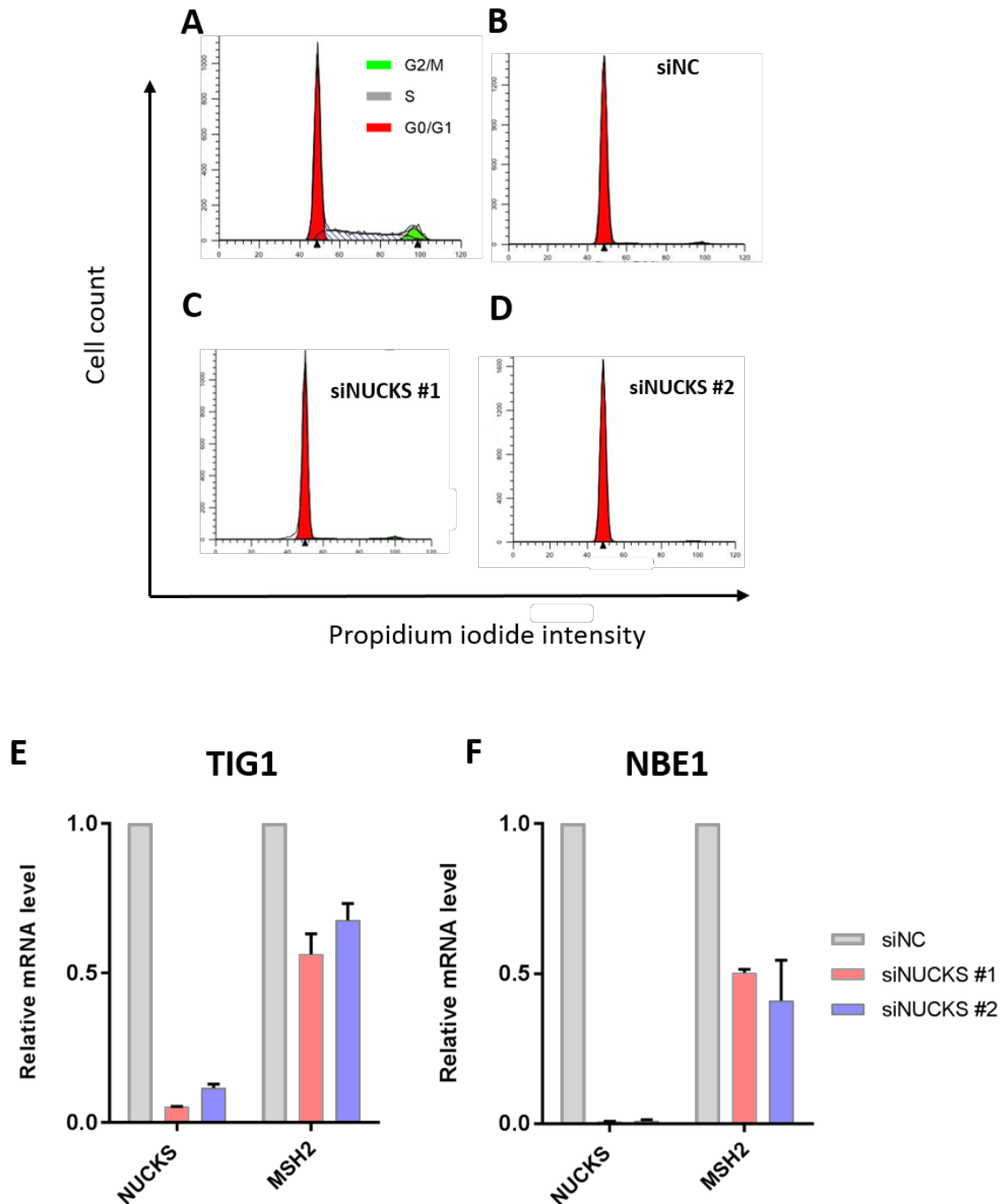


Figure 7: Downregulation of mismatch repair gene MSH2 after treatment with NUCKS siRNA is not cell cycle-dependent

(A) Asynchronous cells (NBE1 shown) grown to confluency, allowed to incubate for three days and then transfected with (B) NC siRNA, (C) NUCKS siRNA #1 or (D) #2 are completely arrested in G1 phase, confirmed by propidium iodide staining with flow cytometry. qPCR was then used to analyse MSH2 expression after NUCKS siRNA treatment in (E) TIG1 and (F) NBE1 cells, respectively, using B2M and TBP as reference genes. Error bars represent standard deviations of two independent experiments.

3.3.4 NUCKS-depleted cells appear to display increased resistance to 6-TG, an agent that relies on proficient mismatch repair for lethality

MMR-deficient cells display resistance to a number of genotoxic agents, including 6-TG (Glaab et al., 1998). The MMR protein complex MutS α recognises 6-TG:T mismatches generated by 6-TG, promoting c-Abl/p73a/GADD45a-dependent G2/M cell cycle arrest and apoptosis in response to treatment (Li et al., 2008; Swann et al., 1996). As such, 6-TG relies on proficient MMR to induce cell death and, accordingly, various studies have employed 6-TG resistance to substantiate novel MMR players (e.g. Chen et al., 2016; Patidar et al., 2016; Wang et al., 2014). To test whether NUCKS-depleted cells displayed such a resistance, we treated NUCKS-depleted and MSH2-depleted (as a positive control) NBE1 cells with 6-TG, and found that NUCKS deficiency indeed provided increased resistance (**Fig. 8**). Although these data require further confirmation through additional independent experiments and employment of additional cell lines, they suggest that NUCKS-depleted cells display functional hallmarks of MMR-deficiency. To further confirm that this resistance in NUCKS-depleted cells is due to MMR deficiency, we could test whether NUCKS is epistatic with MSH2 in 6-TG resistance. Specifically, if NUCKS-depleted and NUCKS + MSH2-depleted cells both displayed similar levels of 6-TG resistance, this would indicate NUCKS and MSH2 act in the same pathway for 6-TG resistance. Alternatively, if we found that NUCKS + MSH2-depleted cells were significantly more resistant to 6-TG than NUCKS-depleted cells alone, this could suggest that NUCKS deficiency generates 6-TG resistance independent of MMR.

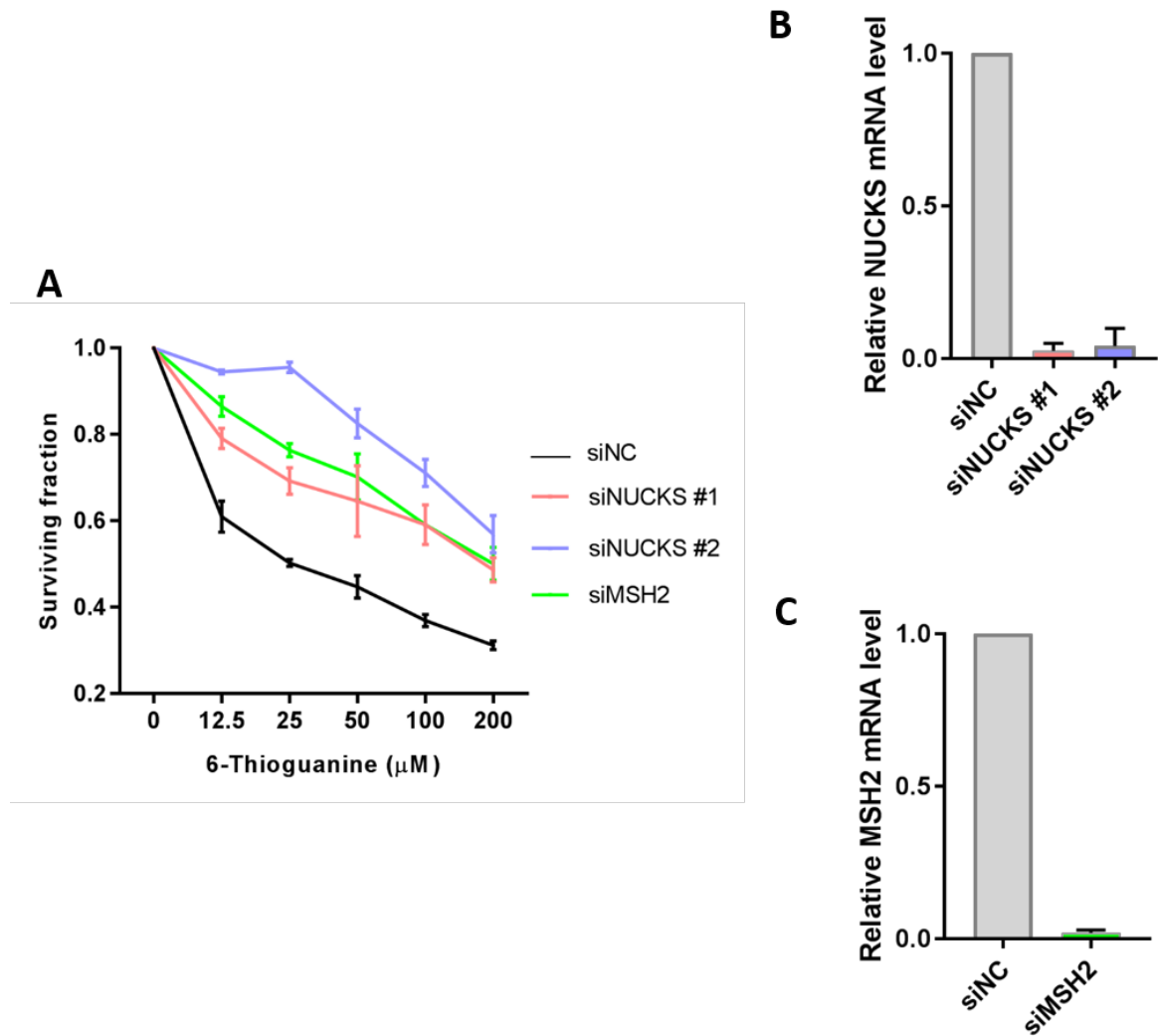


Figure 8: NUCKS-depleted cells display increased resistance to 6-TG, which relies on proficient mismatch repair for lethality

A) NBE1 cells were treated with control, NUCKS siRNA (#1 and #2) or MSH2 siRNA for 48 h and then treated with 6-TG for 48 h. Cell viability was measured using Resazurin. Error bars represent standard deviations of three technical replicates.

B) NUCKS and **C)** MSH2-transfected cells in **A)** were split for RNA extraction and qPCR to confirm knockdowns, using B2M and GAPDH as reference genes.

3.3.5 Genes in NUCKS ChIP-Seq are associated with colorectal cancer

NUCKS regulates MMR (this work) and obesity (Qiu et al., 2014), two major determinants for CRC risk. Deficiencies in MMR result in mutations that inactivate tumour suppressors and activate oncogenes, stimulating CRC pathogenesis (Boland and Goel, 2010), while obesity results in inflammation and alteration of the gut microbiota, aberrantly modulating regulation of p53, DNA repair and cell proliferation genes (Ullman and Itzkowitz, 2011; Zhu et al., 2011; Bardou et al., 2013). Therefore, to explore a potential role for NUCKS in CRC, we wondered whether other putative NUCKS targets identified in NUCKS' ChIP-Seq (Qiu et al., 2014) might be associated with CRC. To this end, we utilised the G-Profiler database (<http://biit.cs.ut.ee/gprofiler/index.cg>), which lists genes known to be mutated or deregulated in human pathology, to examine the 586 genes in the NUCKS ChIP-Seq (Qiu et al., 2014). We found that NUCKS binds the promoters of 10 well-established CRC genes (**Table 6**). These are MMR genes; regulators of PDGF, TGF- β and Wnt signalling, whose deregulation promotes proliferation, angiogenesis and metastasis (Estevez-Garcia et al., 2012; Bellam and Pasche, 2010; Zhang et al., 2014a; Najdi et al., 2011; Yan et al., 2002); and MDM2, a critical regulator of p53 levels (Sugano et al., 2010). Now suspecting an important role for NUCKS in CRC, we complemented these data by using the SEEK database to analyse coexpression with NUCKS specifically in primary colorectal tumours. We found that coexpression scores for MSH2, MLH1 and MSH6 in these tumours were 0.9052, 0.5490 and 1.1464, making these genes the 384th, 2038th and 85th most correlated genes with NUCKS, respectively, of 17863 total genes.

Thus, we present three lines of strong preliminary evidence for NUCKS as a novel transcriptional regulator of MMR in CRC:

- 1) NUCKS is required for MMR gene expression (**Fig. 6, 7**) and function (**Fig. 8**);
- 2) NUCKS' putative transcriptional targets are associated with CRC (**Table 6**);
- 3) in colorectal tumours, NUCKS highly correlates with the expression of MMR genes.

Gene	Function	Contribution to colorectal cancer		References
		Up- or down-regulation in colorectal cancer	Consequence	
MSH2	Mismatch repair	Down	MSI	(Peltomaki, 2001)
MLH1	Mismatch repair	Down	MSI	(Peltomaki, 2001)
PMS2	Mismatch repair	Down	MSI	(Peltomaki, 2001)
PDGFRA	PDGF signalling pathway receptor	Up	Proliferation and angiogenesis	(Estevez-Garcia et al., 2012)
MDM2	Promotes p53 degradation	Up	Loss of p53	(Sugano et al., 2010)
SMAD4	TGF- β signalling pathway receptor	Down	Proliferation and 5-FU resistance	(Bellam and Pasche, 2010, Zhang et al., 2014)
TGFBR2	TGF- β signalling pathway receptor	Down	Proliferation	(Bellam and Pasche, 2010)
CTNNB1 (Beta-catenin)	Promotes Wnt signalling	Up	Proliferation and metastasis	(Najdi et al., 2011)
APC	Represses Wnt signalling	Down	Proliferation and metastasis	(Najdi et al., 2011; Yan et al., 2002)
AXIN2	Represses Wnt signalling	Down	Proliferation and metastasis	(Najdi et al., 2011)

Table 6: Genes in the NUCKS ChIP-Seq are associated with colorectal cancer

Genes in the NUCKS ChIP-Seq associated with colorectal cancer were identified using the G-Profiler database (<http://biit.cs.ut.ee/gprofiler/index.cg>). These genes regulate MMR, p53 status and proliferative and metastatic signalling. Up/downregulations include mutations and expression changes.

4. Discussion

Different pathways of DNA repair, which employ the same cellular resources (e.g. energy and nucleotides) are interconnected and must be tightly balanced to ensure genomic stability and homeostasis. However, how cells achieve this goal is poorly understood. Our laboratory has previously shown that cells deficient in XRCC1, the core coordinator of BER (Caldecott, 2003), downregulate other BER components to prevent the accumulation of SSBs formed during the BER pathway. This alteration depends on two DNA damage-responsive transcription factors, p53 and Sp1 (Poletto et al., 2016). In order to investigate how cells adjust other cellular pathways in response to BER loss, our laboratory performed SILAC in XRCC1-deficient cells (Markkanen et al., 2015). We identified several targets upregulated in the nucleus which were also substrates of the master DDR kinase ATM (Matsuoka et al., 2007) (**Table 4**), the main DDR protein activated upon BER downregulation (Khoronenkova and Dianov, 2015). The most appealing of these targets was NUCKS, a DNA-binding protein (Stvold et al., 2001) with multiple ATM and DNA-PK phosphorylation sites (Wiśniewski et al., 2008), which binds the promoters of several genes linked to DNA repair (Qiu et al., 2014 and this study). Furthermore, NUCKS has a role in promotion of DSB repair by HR (Parplys et al., 2015), suggesting that NUCKS might be a promising candidate as a novel DDR transcription factor – like Sp1 and p53. In this report, we add to the DNA repair inventory of NUCKS, describing a novel role for NUCKS in MMR transcriptional regulation. We find that cells deficient in NUCKS exhibit hallmarks of MMR deficiency, and suggest that this role might link NUCKS deficiency to the development of CRC.

4.1 NUCKS as a DNA damage response protein

To complement implications of NUCKS with the DDR (Parplys et al., 2015; Markkanen et al., 2015; Matsuoka et al., 2007), we explored how NUCKS depletion might impact expression levels of two key components of the DDR: p53 and p21. We found that p53, a stress-responsive transcription factor (Lakin and Jackson, 1999) and p21, a CDK inhibitor whose transcription is activated by p53 (Abbas and Dutta, 2009) are upregulated in response to NUCKS depletion (**Fig. 3**). Although NUCKS depletion was not associated with DNA damage in normal cells (**Fig. 5**), p53 can also be activated in response to a range of other stimuli, including oncogene activation and ribonucleotide depletion (Brooks and Gu, 2010). NUCKS binds the promoters of a number of oncogenes, including Myc (Qiu et al., 2014), and has a role in maintaining replication forks (Parplys et al., 2015), suggesting that p53 activation in NUCKS-depleted cells might be due to oncogene or ribonucleotide pool deregulation, rather than DNA damage. Moreover, NUCKS also binds the promoter of MDM2 (Qiu et al., 2014), which controls p53 levels through ubiquitin-mediated degradation (Moll and Petrenko, 2003). Should NUCKS depletion result in reductions in MDM2 levels (through loss of NUCKS-dependent transcriptional activation), this could also explain the upregulation of p53 in NUCKS-depleted cells. Although this remains to be explored, this would indicate that NUCKS might have a role in control of p53 levels, even in the absence of DNA damage or cellular stress. This would furthermore explain an advantage derived by cancer cells from NUCKS overexpression (Moll and Petrenko, 2003).

Normal cells respond to this upregulation of p53 and p21 by activating a G1/S cell cycle arrest (**Fig. 4**), presumably through p21-mediated CDK inhibition (Abbas and Dutta, 2009). Since NUCKS promotes transcription of cell cycle-promoting proteins such as CSE1L (**Fig. 6, Table 5**, Behrens et al., 2003, Qiu et al., 2014), p53 / p21 upregulation and cell cycle arrest could furthermore be due to loss of such proteins in NUCKS-depleted cells, preventing the onset of S phase. In contrast, two out of three sequences of NUCKS siRNA in cancer cells (U2OS) do not lead to G1 cell cycle arrest

(**Fig. 4D**), despite upregulation of p53 and p21 in these cells (**Fig. 3**). This could be a consequence of cell cycle checkpoint defects in U2OS cells and in cancer cell lines in general (Giono and Manfredi, 2007; Akan et al., 2012; Sherr and McCormick, 2002).

Although we performed only one experiment, we found that U2OS cells, but not TIG1 cells, were associated with a moderate increase in DNA damage after NUCKS depletion (**Fig. 5**). This DNA damage, which we observed by the alkaline Comet assay, does probably not reflect a MMR defect, since MMR deficiency does not lead to the SSBs or DSBs that are detected by this approach (Fairbairn et al., 1995). Instead, it is likely that NUCKS-depleted U2OS cells, whose cell cycle profiles are similar to control cells (**Fig. 4**) and which display a defect in the S phase-associated HR pathway (Parpys et al., 2015), accumulate endogenously-induced DSBs. This is not the case with TIG1 cells, which undergo G1 cell cycle arrest after NUCKS depletion (**Fig. 4**).

To conclude, these data are in agreement with previous studies showing that NUCKS regulates HR and becomes modified after DNA damage (Parpys et al., 2015), and provide further evidence that NUCKS might be important for the DDR. These data also expose future avenues for NUCKS research.

4.2 NUCKS, mismatch repair and colorectal cancer

CRC affects around 1.2 million people worldwide every year, contributing 8.5% of total cancer diagnoses (Brenner et al., 2014). To facilitate early detection, improve diagnosis and predict response to chemotherapy, the identification of novel CRC biomarkers is essential.

Numerous lines of evidence support a potential role for NUCKS in CRC, a disease often caused by defective MMR (Loeb, 1994). NUCKS positively regulates transcription of MMR genes (**Fig. 6 and 7**), is highly coexpressed with MSH2 and MSH6 in primary colorectal tumours (<http://seek.princeton.edu/>), and binds the promoters of a number of genes which are found mutated or deregulated in CRC (**Table 6**). The Oncomine database (<https://www.oncomine.org>)

reveals that NUCKS expression tends to be lower in colorectal tumours than tumours from other tissue, and low NUCKS expression is linked with obesity (Qiu et al., 2014). Obesity, in turn, due to alteration of inflammation and gut microbiota in obese patients influencing the expression of cancer-driving genes (Bardou et al., 2013), is linked with higher incidence of CRC (Bardou et al., 2013). Taking these data and observations together, we propose the scheme in **Fig. 9**. We suggest that reductions in NUCKS expression, which lead to obesity (Qiu et al., 2014), will simultaneously lead to reductions in MMR gene expression, as well as, possibly, deregulation of proteins such as APC and CTNNB1 (**Table 6**). We hypothesise that these phenotypes might synergise to promote a favourable environment for the development of CRC.

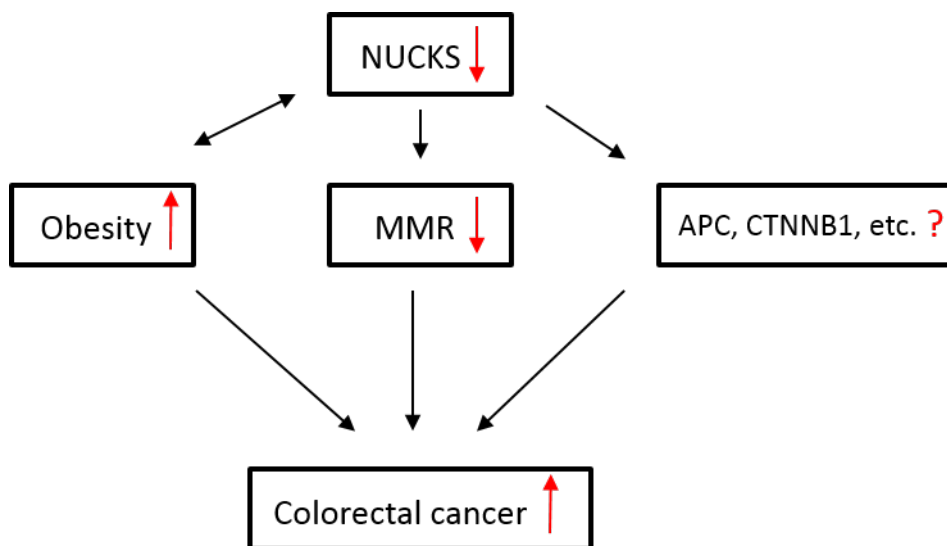


Figure 9: NUCKS and colorectal cancer.

NUCKS downregulation leads to obesity (and vice versa) (Qiu et al., 2014), MMR decline (this study), and potential alterations in APC and other genes associated with CRC (**Table 6**). These phenotypes might associate to promote CRC risk.

Furthermore, we hypothesise that loss of NUCKS expression might contribute to MSI in CRC. The majority of the 15% MSI-positive colorectal tumours are due to mutation in MMR proteins or hypermethylation of the MLH1 promoter (Hampel et al., 2005). However, the mechanism underlying MSI-positive tumours which express WT MMR proteins and lack hypermethylation of the MLH1 promoter is poorly understood (Brenner et al., 2014). Intriguingly, one study suggests that NUCKS expression might be lower in MSI tumours than microsatellite stable (MSS) equivalents (Jorissen et al., 2008), and we therefore propose that loss of NUCKS expression (but not NUCKS mutation, which is rare (<http://cancer.sanger.ac.uk/cosmic>)) might contribute to these poorly characterised MSI-positive tumours.

Finally, we suggest that NUCKS might play a role in resistance to 5-FU treatment, the current Gold Standard for CRC treatment (Wolpin and Mayer, 2008). MMR's MutS α complex binds 5-FU (Tajima et al., 2004) and initiates c-Abl/p73a/GADD45a-dependent apoptosis in response to 5-FU treatment (Li et al., 2009). Accordingly, MMR deficiency has been linked with reduced efficacy of 5-FU treatment in clinical trials (Ribic et al., 2003) and reduction in MSH2 (Valeri et al., 2010a) and MLH1 (Arnold et al., 2003) levels has been linked with resistance to 5-FU treatment. Strikingly, one microarray analysis found that NUCKS and MSH2 levels decreased after 5-FU treatment (Troester et al., 2004), suggesting ambitiously that NUCKS might (passively) orchestrate a novel cellular resistance to 5-FU by downregulating transcription of MMR genes.

4.3 Limitations of our approach

We confirm NUCKS depletion on the mRNA level by qPCR and at the protein level using the (unconventional) flow cytometry approach (**Fig. 2**). In this approach, cells are fixed in formaldehyde, permeabilised with Triton-X 100 and stained with antibodies (**Section 2.4.2**). Although flow cytometry allows us to observe a decrease in protein signal upon knockdown (**Fig. 2**), this approach does permit confirmation, through validation of molecular weight or nuclear/cytoplasmic localisation, that *NUCKS* depletion is the cause of this signal change. As such,

more convincing validations of knockdown could be obtained using Western blotting or immunofluorescence.

Furthermore, we rely heavily on RNA interference using siRNA in this study. However, RNA interference can generate significant undesired off-target effects (Jackson et al., 2006). This can lead to phenotypes that are not specific to the depleted protein, which can also give rise to variability between different siRNA sequences. Indeed, we observe variability between NUCKS siRNA sequences in p53 and p21 induction (**Fig. 3**), in impact on cell cycle (**Fig. 4D**) and in resistance to 6-TG (**Fig. 8**), even despite similar knockdown for these two sequences (**Fig. 2**). As a result, to further confirm that observed phenotypes are due specifically to loss of NUCKS, we could employ a rescue strategy. For example, we could rescue wildtype levels of MSH2 expression in NUCKS-knockout cells by transfecting NUCKS back into these cells.

4.4 Future experiments

4.4.1 NUCKS and the DNA damage response

We have shown that NUCKS depletion induces expression of p53 and p21 (**Fig. 3**). To further characterise a potential role for NUCKS in the DDR, we would start by analysing the impact of NUCKS depletion on expression or activation of numerous DDR components: SSB markers (e.g. PAR), DSB markers (e.g. pATM, γ H2AX, 53BP1), and replication stress markers (e.g. pATR, pChk1, pRPA). Furthermore, we could analyse sensitivity of NUCKS-depleted cells to DNA damaging agents to probe roles for NUCKS in different DNA repair pathways. For example, sensitivity to UV-radiation might indicate a role in NER (Dunkern et al., 2001), sensitivity to MMS could indicate a BER role (Lundin et al., 2005) and sensitivity to formaldehyde could indicate a defect in DNA-protein crosslink (Vaz et al., 2016) or Fanconi anemia repair (de Graaf et al., 2009). Finally, in order to further delineate the role for NUCKS in the DDR, we could assess kinetics of DNA repair in NUCKS-depleted and proficient cells, using different genotoxic agents. For example, we could treat cells

with ionising radiation, which generates a variety of lesions, the most toxic of which is DSBs. To directly measure the kinetics of repair of these DSBs in ionising radiation-damaged NUCKS-proficient and depleted cells, we could use the neutral Comet assay, which specifically detects DSBs (Fairbairn et al., 1995). To complement this approach, we could analyse γ H2AX foci, a marker of DSBs (Kuo et al., 2008), through γ H2AX immunofluorescence analysis.

NUCKS undergoes more posttranslational modification than any other protein, but the functions of these modifications are poorly understood (Wiśniewski et al., 2008). Many of these modifications are focused on the NUCKS DNA-binding domain, which spans NUCKS residues 183-192 (Grundt et al., 2002). At the onset of mitosis, CDK1 phosphorylates NUCKS S181, inhibiting NUCKS' DNA binding (Grundt et al., 2002). Importantly, this is coherent with reduced global gene transcription in mitosis, and with a putative role of NUCKS in G1/S transcription postulated earlier. Furthermore, S181 undergoes ATM-dependent phosphorylation after treatment with neocarzinostatin, a radiomimetic (Bensimon et al., 2010). Assuming this phosphorylation similarly inhibits NUCKS DNA binding, this might implicate NUCKS in G1/S checkpoint activation after DNA damage: DNA damage would inhibit NUCKS DNA-binding and might accordingly prevent NUCKS-dependent S phase transcription, consequently preventing S phase entry. ATM also phosphorylates S14 after ionising radiation (Matsuoka et al., 2007) and S54 after neocarzinostatin treatment (Bensimon et al., 2010). These modifications are targeted away from NUCKS' DNA-binding domain but might influence, for example, NUCKS' protein stability or protein-protein interactions.

The broad influence of NUCKS' PTMs can be elucidated experimentally. To explore the contributions to NUCKS DNA-binding, we could assess DNA-binding capacities of the recombinant NUCKS protein modified by different protein kinases to DNA probes using the EMSA assay. To test contributions to protein stability, we could measure the lifetime of different NUCKS modification-mutants in cyclohexamide-treated cells. Finally, contributions to protein-protein interactions could be assessed by performing immunoprecipitation experiments of differentially modified NUCKS with

known interaction partners. Undoubtedly, these are vital experiments and their results will be central to the complete understanding of the NUCKS protein.

4.4.2 NUCKS, mismatch repair and colorectal cancer

To further explore the role of NUCKS in CRC, we would start by establishing whether NUCKS-dependent MMR regulation is present in MMR-proficient, microsatellite-stable CRC cell lines. In addition, we would establish how the genes listed in **Table 6** respond to NUCKS depletion. For example, in the Wnt signalling pathway, APC promotes degradation of β -catenin, a transcription factor which promotes expression of oncogenes such as Myc (Herbst et al., 2014; Najdi et al., 2011). As such, CRC is often associated with downregulation or mutation of APC and consequent upregulation of β -catenin (Najdi et al., 2011). Should we find that NUCKS promotes APC expression while attenuating that of β -catenin, this would support our hypothesis that low NUCKS levels are prognostic for CRC risk (**Fig. 9**). We would also like to assess correlation of NUCKS expression with that of MSH2 in a panel of primary colorectal tumours, using immunohistochemistry. If the expression of NUCKS correlates with MSH2 in these tumours, this, coupled with our other data, would provide very robust evidence that NUCKS is a transcriptional regulator of MMR in CRC.

Next, we would establish whether cells deficient in NUCKS expression for an extended period of time develop MSI, the quintessential characteristic of MMR-deficient cells (Loeb, 1994). To this end, we would generate NUCKS knockout cell lines, using the CRISPR-Cas9 genome editing technique (Ran et al., 2013). We would then use PCR to amplify a consensus panel of microsatellite regions used to assess MSI (Umar et al., 2004) and separate these microsatellite sequences by denaturing polyacrylamide gel electrophoresis, observing whether novel fragments – indicative of MSI - appear in NUCKS-deficient cells (Umar et al., 2004).

Polymorphisms in the promoters of MSH2 (Shin et al., 2002; Mrkonjic et al., 2007), MLH1 (Samowitz et al., 2008) and MSH6 (Gazzoli and Kolodner, 2003) reduce promoter activity, expedite MSI onset and promote development of CRC. To identify whether these polymorphisms fall within NUCKS' binding regions, we will use ChIP (Das et al., 2004), and to subsequently establish whether they impact NUCKS binding, we will use the electrophoretic mobility shift (Hellman and Fried, 2007) and luciferase (Smale, 2010) assays.

Finally, transcription factor p53 has myriad roles in tumour suppression: from cell cycle arrest, to autophagy, to apoptosis (Lane et al., 1992). To explore whether any putative roles of NUCKS are dependent on p53 activity, we could employ a panel of p53-negative cell lines. For example, the cervical cancer cell line HeLa, colorectal cancer cell line CACO-2 and renal cell carcinoma cell line TK10 all contain p53-inactivating mutations (Leroy et al., 2014). Even more valuable, p53+/+ and p53-/- isogenic cell line pairs, such as those available for colorectal cancer cell line HCT116, could be employed (Sur et al., 2009). For example, it is known that p53 binds the promoters of MSH2 and MLH1 and stimulates their transcription (Nikulenkov et al., 2012). Accordingly, we could hypothesise that the stimulation of MMR transcription by NUCKS might be p53-dependent – because, for example, NUCKS might enable p53 binding to MMR promoters. Such a theory – and many others - could be readily tested using a selection of these cell lines.

4.5 Conclusions

In this study we report the identification of NUCKS as a novel transcriptional regulator of MMR, and provide supporting evidence that NUCKS might play an important role the development of CRC. Given this information, we propose that low levels of NUCKS expression might be a risk factor for CRC due to concomitant MMR deficiency and increased obesity risk (**Fig. 9**), and might furthermore be used to evaluate the risk of CRC, even before the onset of MSI currently used for CRC diagnosis (Umar et al., 2004). Moreover, given the MMR- and HR- promoting activities of NUCKS (this study and Parplys et al., 2015), we suggest that these cancers might display resistance to drugs such as

5-FU or etoposide (Aebi et al., 1997), while displaying sensitivity to treatments such as radiotherapy (Jeggo et al., 2011). We hope that future studies leading from this report might invigorate NUCKS study and inaugurate more comprehensive analyses of NUCKS and DNA repair regulation in cancer.

Appendix

Material	Supplier
0.2 ml Thin-walled tubes	MOLECULAR BIO-PRODUCTS
1 ml syringes	BD
1.1 mm needles	BD
1.5 ml Eppendorf tube	GREINER BIO-ONE
10 cm dishes	GREINER BIO-ONE
15 ml falcon tubes	GREINER BIO-ONE
50 ml falcon tubes	GREINER BIO-ONE
50X TAE	BIO-RAD
6-thioguanine	SIGMA
75 cm ² flasks	GREINER BIO-ONE
96-well flat bottom plates	CORNING
96-well PCR plates	ThermoFisher Scientific
Adhesive PCR films	ThermoFisher Scientific
Agarose	BIO-RAD
Amicon ultra 0.5 ml tubes	MILLIPORE
BioRad Bradford reagent	BIO-RAD
Bromophenol blue powder	SIGMA
BSA	BIO-RAD
Buffer EB	QIAGEN
Cuvettes	FISHER
DEPC-treated H ₂ O	AMBION
DMEM	GIBCO
DMSO	SIGMA
DNA gel cassettes	BIO-RAD
DNA gel tanks	BIO-RAD
dNTPs mix	BIOLINE
DTT 0.1 M	INVITROGEN
EDTA	BDH
SDS-PAGE materials	NOVELL
Ethanol	FISHER
Fast SYBR TM Green Master Mix	ThermoFisher Scientific
FBS	GIBCO
First strand buffer (5X)	INVITROGEN
Gel Loading dye 6X, Blue	NEB
Glycerol	FLUKA
H ₂ O ₂	SIGMA
Heating blocks	EPPENDORF
Immobilon-FL	MILLIPORE
Immobilon-FL PVDF	MILLIPORE
Isopropanol (100%)	FISHER
KCl	BDH
Leupeptin	CALBIOCHEM
Lipofectamine 2000	INVITROGEN
Lipofectamine RNAiMAX	INVITROGEN

Methanol	FISHER
MgCl ₂	SIGMA
NBE1 cells	LONZA
NEM	SIGMA
Nonidet P-40	SIGMA
Odyssey blocker buffer	LI-COR
PBS	SIGMA
Pepstatin	CALBIOCHEM
PMSF	SIGMA
Power suppliers	BIO-RAD
Propidium Iodide	SIGMA
QIAGEN-tip 500	QIAGEN
QIAprep spin column	QIAGEN
RNA extraction kit	QIAGEN
RNase 100 mg/ml	QIAGEN
RNase H enzyme	NEB
RNase-free water	QIAGEN
RNaseOUT (40 U/μl)	INVITROGEN
SDS	FISHER
Sponges for Western blot	NOVELL
SuperScript II RT enzyme	INVITROGEN
TIG1 cells	CORRIEL
Tris/glycine buffer (10X)	BIO-RAD
Tris/glycine/SDS buffer (10X)	BIO-RAD
Triton-X 100	SIGMA
Trizma	SIGMA
Trypan blue	INVITROGEN
Trypsin-EDTA	GIBCO
Tube roller	STOVALL
Tube rotor	LABINCO
U2OS cells	ATCC
Western blot apparatus	NOVELL

Table 7: Materials and reagents used throughout this study, as well as their supplier.

References

- Abbas, T., and A. Dutta. 2009. p21 in cancer: intricate networks and multiple activities. *Nat. Rev. Cancer*. 9:400–414. doi:10.1038/nrc2657.
- Aebi, S., D. Fink, R. Gordon, H.K. Kim, H. Zheng, J.L. Fink, and S.B. Howell. 1997. Resistance to cytotoxic drugs in DNA mismatch repair-deficient cells. *Clin. Cancer Res.* 3:1763–1767.
- Agresti, A., and M.E. Bianchi. 2003. HMGB proteins and gene expression. *Curr. Opin. Genet. Dev.* 13:170–178.
- Akan, P., A. Alexeyenko, P.I. Costea, L. Hedberg, B.W. Solnestam, S. Lundin, J. Hällman, E. Lundberg, M. Uhlén, and J. Lundberg. 2012. Comprehensive analysis of the genome transcriptome and proteome landscapes of three tumor cell lines. *Genome Med.* 4:86. doi:10.1186/gm387.
- Allegra, C.J., J.M. Jessup, M.R. Somerfield, S.R. Hamilton, E.H. Hammond, D.F. Hayes, P.K. McAllister, R.F. Morton, and R.L. Schilsky. 2009. American Society of Clinical Oncology provisional clinical opinion: testing for KRAS gene mutations in patients with metastatic colorectal carcinoma to predict response to anti-epidermal growth factor receptor monoclonal antibody therapy. *J. Clin. Oncol.* 27:2091–2096. doi:10.1200/JCO.2009.21.9170.
- Ames, B.N., M.K. Shigenaga, and T.M. Hagen. 1993. Oxidants, antioxidants, and the degenerative diseases of aging. *Proc. Natl. Acad. Sci. U. S. A.* 90:7915–7922.
- Arnold, C.N., A. Goel, and C.R. Boland. 2003. Role of hMLH1 promoter hypermethylation in drug resistance to 5-fluorouracil in colorectal cancer cell lines. *Int. J. cancer.* 106:66–73. doi:10.1002/ijc.11176.
- Baldin, V., J. Lukas, M.J. Marcote, M. Pagano, and G. Draetta. 1993. Cyclin D1 is a nuclear protein required for cell cycle progression in G1. *Genes Dev.* 7:812–821.
- Bardou, M., A.N. Barkun, and M. Martel. 2013. Obesity and colorectal cancer. *Gut.* 62:933–947. doi:10.1136/gutjnl-2013-304701.
- Barton, O., S.C. Naumann, R. Diemer-Biehs, J. Kunzel, M. Steinlage, S. Conrad, N. Makharashvili, J. Wang, L. Feng, B.S. Lopez, T.T. Paull, J. Chen, P.A. Jeggo, and M. Lobrich. 2014. Polo-like kinase 3 regulates CtIP during DNA double-strand break repair in G1. *J. Cell Biol.* 206:877–894. doi:10.1083/jcb.201401146.
- Basen-Engquist, K., and M. Chang. 2011. Obesity and cancer risk: recent review and evidence. *Curr. Oncol. Rep.* 13:71–76. doi:10.1007/s11912-010-0139-7.
- Behrens, P., U. Brinkmann, and A. Wellmann. 2003. CSE1L/CAS: its role in proliferation and apoptosis. *Apoptosis.* 8:39–44.
- Bellam, N., and B. Pasche. 2010. Tgf-beta signaling alterations and colon cancer. *Cancer Treat. Res.* 155:85–103. doi:10.1007/978-1-4419-6033-7_5.
- Bensimon, A., A. Schmidt, Y. Ziv, R. Elkon, S.-Y. Wang, D.J. Chen, R. Aebersold, and Y. Shiloh. 2010. ATM-dependent and -independent dynamics of the nuclear phosphoproteome after DNA damage. *Sci. Signal.* 3:rs3. doi:10.1126/scisignal.2001034.
- de Boer, J., and J.H. Hoeijmakers. 2000. Nucleotide excision repair and human syndromes. *Carcinogenesis.* 21:453–460.
- Boland, C.R., and A. Goel. 2010. Microsatellite Instability in Colorectal Cancer. *Gastroenterology.* 138:2073–2087.e3. doi:10.1053/j.gastro.2009.12.064.
- Boland, C.R., S.N. Thibodeau, S.R. Hamilton, D. Sidransky, J.R. Eshleman, R.W. Burt, S.J. Meltzer, M.A. Rodriguez-Bigas, R. Fodde, G.N. Ranzani, and S. Srivastava. 1998. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res.* 58:5248–5257.
- Bradford, M.M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72:248–254.

- Brenner, H., M. Kloor, and C.P. Pox. 2014. Colorectal cancer. *Lancet (London, England)*. 383:1490–1502. doi:10.1016/S0140-6736(13)61649-9.
- Brooks, C.L., and W. Gu. 2010. New insights into p53 activation. *Cell Res.* 20:614–621. doi:10.1038/cr.2010.53.
- Caldecott, K.W. 2003. XRCC1 and DNA strand break repair. *DNA Repair (Amst)*. 2:955–969.
- Cao, A.R., R. Rabinovich, M. Xu, X. Xu, V.X. Jin, and P.J. Farnham. 2011. Genome-wide Analysis of Transcription Factor E2F1 Mutant Proteins Reveals That N- and C-terminal Protein Interaction Domains Do Not Participate in Targeting E2F1 to the Human Genome. *J. Biol. Chem.* 286:11985–11996. doi:10.1074/jbc.M110.217158.
- Carrato, A. 2008. Adjuvant Treatment of Colorectal Cancer. *Gastrointest. Cancer Res.* 2:S42-6.
- Chang, I.-Y., M. Jin, S.P. Yoon, C.-K. Youn, Y. Yoon, S.-P. Moon, J.-W. Hyun, J.Y. Jun, and H.J. You. 2008. Senescence-dependent MutS alpha dysfunction attenuates mismatch repair. *Mol. Cancer Res.* 6:978–989. doi:10.1158/1541-7786.MCR-07-0380.
- Chang, M., M. Bellaoui, C. Boone, and G.W. Brown. 2002. A genome-wide screen for methyl methanesulfonate-sensitive mutants reveals genes required for S phase progression in the presence of DNA damage. *Proc. Natl. Acad. Sci. U. S. A.* 99:16934–16939. doi:10.1073/pnas.262669299.
- Chapman, J.R., M.R.G. Taylor, and S.J. Boulton. 2012. Playing the end game: DNA double-strand break repair pathway choice. *Mol. Cell.* 47:497–510. doi:10.1016/j.molcel.2012.07.029.
- Chen, J., and I. Sadowski. 2005. Identification of the mismatch repair genes PMS2 and MLH1 as p53 target genes by using serial analysis of binding elements. *Proc. Natl. Acad. Sci. U. S. A.* 102:4813–4818. doi:10.1073/pnas.0407069102.
- Chen, Z., M. Tran, M. Tang, W. Wang, Z. Gong, and J. Chen. 2016. Proteomic Analysis Reveals a Novel MutS Partner Involved in Mismatch Repair Pathway. *Mol. Cell. Proteomics*. doi:10.1074/mcp.M115.056093.
- Cheong, J.Y., Y.B. Kim, J.H. Woo, D.K. Kim, M. Yeo, S.-J. Yang, K.-S. Yang, S.K. Soon, H.J. Wang, B.W. Kim, J.-H. Park, and S.W. Cho. 2016. Identification of NUCKS1 as a putative oncogene and immunodiagnostic marker of hepatocellular carcinoma. *Gene*. 584:47–53. doi:10.1016/j.gene.2016.03.006.
- Christmann, M., and B. Kaina. 2013. Transcriptional regulation of human DNA repair genes following genotoxic stress: trigger mechanisms, inducible responses and genotoxic adaptation. *Nucleic Acids Res.* 41:8403–8420. doi:10.1093/nar/gkt635.
- Colussi, C., E. Parlanti, P. Degan, G. Aquilina, D. Barnes, P. Macpherson, P. Karran, M. Crescenzi, E. Dogliotti, and M. Bignami. 2002. The mammalian mismatch repair pathway removes DNA 8-oxodGMP incorporated from the oxidized dNTP pool. *Curr. Biol.* 12:912–918.
- D'Amours, D., and S.P. Jackson. 2002. The Mre11 complex: at the crossroads of dna repair and checkpoint signalling. *Nat. Rev. Mol. Cell Biol.* 3:317–327. doi:10.1038/nrm805.
- Das, C., and T.K. Kundu. 2005. Transcriptional regulation by the acetylation of nonhistone proteins in humans -- a new target for therapeutics. *IUBMB Life*. 57:137–149. doi:10.1080/15216540500090629.
- Das, P.M., K. Ramachandran, J. vanWert, and R. Singal. 2004. Chromatin immunoprecipitation assay. *Biotechniques*. 37:961–969.
- Davies, H., G.R. Bignell, C. Cox, P. Stephens, S. Edkins, S. Clegg, J. Teague, H. Woffendin, M.J. Garnett, W. Bottomley, N. Davis, E. Dicks, R. Ewing, Y. Floyd, K. Gray, S. Hall, R. Hawes, J. Hughes, V. Kosmidou, A. Menzies, C. Mould, A. Parker, C. Stevens, S. Watt, S. Hooper, R. Wilson, H. Jayatilake, B.A. Gusterson, C. Cooper, J. Shipley, D. Hargrave, K. Pritchard-Jones, N. Maitland, G. Chenevix-Trench, G.J. Riggins, D.D. Bigner, G. Palmieri, A. Cossu, A. Flanagan, A. Nicholson, J.W.C. Ho, S.Y. Leung, S.T. Yuen, B.L. Weber, H.F. Seigler, T.L. Darrow, H. Paterson, R. Marais, C.J. Marshall, R. Wooster, M.R. Stratton, and P.A. Futreal. 2002. Mutations of the BRAF gene in human cancer. *Nature*. 417:949–954. doi:10.1038/nature00766.

- de Graaf, B., A. Clore, and A.K. McCullough. 2009. Cellular Pathways for DNA Repair and Damage Tolerance of Formaldehyde-Induced DNA-Protein Crosslinks. *DNA Repair (Amst)*. 8:1207–1214. doi:10.1016/j.dnarep.2009.06.007.
- Dianov, G.L., and U. Hübscher. 2013. Mammalian Base Excision Repair: the Forgotten Archangel. *Nucleic Acids Res.* 41:3483–3490. doi:10.1093/nar/gkt076.
- Dray, E., J. Etchin, C. Wiese, D. Saro, G.J. Williams, M. Hammel, X. Yu, V.E. Galkin, D. Liu, M.-S. Tsai, S.M.-H. Sy, D. Schild, E. Egelman, J. Chen, and P. Sung. 2010. Enhancement of RAD51 recombinase activity by the tumor suppressor PALB2. *Nat. Struct. Mol. Biol.* 17:1255–1259. doi:10.1038/nsmb.1916.
- Drosos, Y., M. Kouloukoussa, A. Østfold, K. Grundt, N. Goutas, D. Vlachodimitropoulos, S. Havaki, P. Kollia, C. Kittas, E. Marinou, and V. Aleporou-Marinou. 2009. NUCKS overexpression in breast cancer. *Cancer Cell Int.* 9:19. doi:10.1186/1475-2867-9-19.
- Dunkern, T.R., G. Fritz, and B. Kaina. 2001. Ultraviolet light-induced DNA damage triggers apoptosis in nucleotide excision repair-deficient cells via Bcl-2 decline and caspase-3/-8 activation. *Oncogene*. 20:6026–6038. doi:10.1038/sj.onc.1204754.
- Duval, A., and R. Hamelin. 2002. Mutations at coding repeat sequences in mismatch repair-deficient human cancers: toward a new concept of target genes for instability. *Cancer Res.* 62:2447–2454.
- Easton, D.F., D. Ford, and D.T. Bishop. 1995. Breast and ovarian cancer incidence in BRCA1-mutation carriers. Breast Cancer Linkage Consortium. *Am. J. Hum. Genet.* 56:265–271.
- Eisen, M.B., P.T. Spellman, P.O. Brown, and D. Botstein. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. U. S. A.* 95:14863–14868.
- Ellegren, H. 2004. Microsatellites: simple sequences with complex evolution. *Nat. Rev. Genet.* 5:435–445. doi:10.1038/nrg1348.
- Estevez-Garcia, P., A. Castano, A.C. Martin, F. Lopez-Rios, J. Iglesias, S. Munoz-Galvan, I. Lopez-Calderero, S. Molina-Pinelo, M.D. Pastor, A. Carnero, L. Paz-Ares, and R. Garcia-Carbonero. 2012. PDGFRalpha/beta and VEGFR2 polymorphisms in colorectal cancer: incidence and implications in clinical outcome. *BMC Cancer*. 12:514. doi:10.1186/1471-2407-12-514.
- Fairbairn, D.W., P.L. Olive, and K.L. O'Neill. 1995. The comet assay: a comprehensive review. *Mutat. Res.* 339:37–59.
- Friedberg, E.C., L.D. McDaniel, and R.A. Schultz. 2004. The role of endogenous and exogenous DNA damage and mutagenesis. *Curr. Opin. Genet. Dev.* 14:5–10. doi:10.1016/j.gde.2003.11.001.
- Friedberg, E.C., and L.B. Meira. 2006. Database of mouse strains carrying targeted mutations in genes affecting biological responses to DNA damage Version 7. *DNA Repair (Amst)*. 5:189–209. doi:10.1016/j.dnarep.2005.09.009.
- Friedhoff, P., P. Li, and J. Gotthardt. 2016. Protein-protein interactions in DNA mismatch repair. *DNA Repair (Amst)*. 38:50–57. doi:10.1016/j.dnarep.2015.11.013.
- Gasch, A.P., M. Huang, S. Metzner, D. Botstein, S.J. Elledge, and P.O. Brown. 2001. Genomic expression responses to DNA-damaging agents and the regulatory role of the yeast ATR homolog Mec1p. *Mol. Biol. Cell.* 12:2987–3003.
- Gazzoli, I., and R.D. Kolodner. 2003. Regulation of the Human MSH6 Gene by the Sp1 Transcription Factor and Alteration of Promoter Activity and Expression by Polymorphisms. *Mol. Cell. Biol.* 23:7992–8007. doi:10.1128/MCB.23.22.7992-8007.2003.
- Gillet, L.C.J., and O.D. Scharer. 2006. Molecular mechanisms of mammalian global genome nucleotide excision repair. *Chem. Rev.* 106:253–276. doi:10.1021/cr040483f.
- Giono, L.E., and J.J. Manfredi. 2007. Mdm2 is required for inhibition of Cdk2 activity by p21, thereby contributing to p53-dependent cell cycle arrest. *Mol. Cell. Biol.* 27:4166–4178. doi:10.1128/MCB.01967-06.
- Glaab, W.E., J.I. Risinger, A. Umar, J.C. Barrett, T.A. Kunkel, and K.R. Tindall. 1998. Resistance to 6-thioguanine in mismatch repair-deficient human cancer cell lines correlates with an

- increase in induced mutations at the HPRT locus. *Carcinogenesis*. 19:1931–1937.
- Grundt, K., I. V. Haga, H.S. Huitfeldt, and A.C. Ostvold. 2007. Identification and characterization of two putative nuclear localization signals (NLS) in the DNA-binding protein NUCKS. *Biochim. Biophys. Acta*. 1773:1398–406. doi:10.1016/j.bbamcr.2007.05.013.
- Grundt, K., I.V. Haga, V. Aleporou-Marinou, Y. Drosos, B. Wanvik, and A.C. Ostvold. 2004. Characterisation of the NUCKS gene on human chromosome 1q32.1 and the presence of a homologous gene in different species. *Biochem. Biophys. Res. Commun.* 323:796–801. doi:10.1016/j.bbrc.2004.08.153.
- Grundt, K., L. Skjeldal, H.W. Anthonsen, T. Skauge, H.S. Huitfeldt, and A.C. Østvold. 2002. A putative DNA-binding domain in the NUCKS protein. *Arch. Biochem. Biophys.* 407:168–75.
- Gu, H., J.D. Marth, P.C. Orban, H. Mossmann, and K. Rajewsky. 1994. Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting. *Science*. 265:103–106.
- Gu, L., B. Xia, L. Zhong, Y. Ma, L. Liu, L. Yang, and G. Lou. 2014. NUCKS1 overexpression is a novel biomarker for recurrence-free survival in cervical squamous cell carcinoma. *Tumour Biol*. 35:7831–7836. doi:10.1007/s13277-014-2035-5.
- Hamosh, A., A.F. Scott, J.S. Amberger, C.A. Bocchini, and V.A. McKusick. 2005. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res.* 33:D514-7. doi:10.1093/nar/gki033.
- Hampel, H., W.L. Frankel, E. Martin, M. Arnold, K. Khanduja, P. Kuebler, H. Nakagawa, K. Sotamaa, T.W. Prior, J. Westman, J. Panescu, D. Fix, J. Lockman, I. Comeras, and A. de la Chapelle. 2005. Screening for the Lynch syndrome (hereditary nonpolyposis colorectal cancer). *N. Engl. J. Med.* 352:1851–1860. doi:10.1056/NEJMoa043146.
- Hanahan, D., and R.A. Weinberg. 2011. Hallmarks of cancer: the next generation. *Cell*. 144:646–674. doi:10.1016/j.cell.2011.02.013.
- Hawn, M.T., A. Umar, J.M. Carethers, G. Marra, T.A. Kunkel, C.R. Boland, and M. Koi. 1995. Evidence for a connection between the mismatch repair system and the G2 cell cycle checkpoint. *Cancer Res.* 55:3721–3725.
- Hellman, L.M., and M.G. Fried. 2007. Electrophoretic Mobility Shift Assay (EMSA) for Detecting Protein-Nucleic Acid Interactions. *Nat. Protoc.* 2:1849–1861. doi:10.1038/nprot.2007.249.
- Henry, N.L., and D.F. Hayes. 2012. Cancer biomarkers. *Mol. Oncol.* 6:140–146. doi:10.1016/j.molonc.2012.01.010.
- Herbst, A., V. Jurinovic, S. Krebs, S.E. Thieme, H. Blum, B. Göke, and F.T. Kolligs. 2014. Comprehensive analysis of β -catenin target genes in colorectal carcinoma cell lines with deregulated Wnt/ β -catenin signaling. *BMC Genomics*. 15:74. doi:10.1186/1471-2164-15-74.
- Hingorani, M.M. 2016. Mismatch binding, ADP-ATP exchange and intramolecular signaling during mismatch repair. *DNA Repair (Amst)*. 38:24–31. doi:10.1016/j.dnarep.2015.11.017.
- Hinz, J.M. 2010. Role of homologous recombination in DNA interstrand crosslink repair. *Environ. Mol. Mutagen.* 51:582–603. doi:10.1002/em.20577.
- Hsieh, P., and K. Yamane. 2008. DNA mismatch repair: molecular mechanism, cancer, and ageing. *Mech. Ageing Dev.* 129:391–407. doi:10.1016/j.mad.2008.02.012.
- Hume, D.A., and K.P.A. MacDonald. 2012. Therapeutic applications of macrophage colony-stimulating factor-1 (CSF-1) and antagonists of CSF-1 receptor (CSF-1R) signaling. *Blood*. 119:1810–1820. doi:10.1182/blood-2011-09-379214.
- Hutten, S., and R.H. Kehlenbach. 2007. CRM1-mediated nuclear export: to the pore and beyond. *Trends Cell Biol.* 17:193–201. doi:10.1016/j.tcb.2007.02.003.
- Jackson, A.L., J. Burchard, J. Schelter, B.N. Chau, M. Cleary, L. Lim, and P.S. Linsley. 2006. Widespread siRNA “off-target” transcript silencing mediated by seed region sequence complementarity. *RNA*. 12:1179–1187. doi:10.1261/rna.25706.
- Jeggo, P.A., V. Geuting, and M. Lobrich. 2011. The role of homologous recombination in radiation-induced double-strand break repair. *Radiother. Oncol.* 101:7–12.

- doi:10.1016/j.radonc.2011.06.019.
- Jensen, R.B., A. Carreira, and S.C. Kowalczykowski. 2010. Purified human BRCA2 stimulates RAD51-mediated recombination. *Nature*. 467:678–683. doi:10.1038/nature09399.
- Jiricny, J. 2000. Mediating mismatch repair. *Nat. Genet.* 24:6–8. doi:10.1038/71698.
- Jiricny, J. 2006. The multifaceted mismatch-repair system. *Nat. Rev. Mol. Cell Biol.* 7:335–346. doi:10.1038/nrm1907.
- Jorissen, R.N., L. Lipton, P. Gibbs, M. Chapman, J. Desai, I.T. Jones, T.J. Yeatman, P. East, I.P.M. Tomlinson, H.W. Verspaget, L.A. Aaltonen, M. Kruhoffer, T.F. Orntoft, C.L. Andersen, and O.M. Sieber. 2008. DNA copy-number alterations underlie gene expression differences between microsatellite stable and unstable colorectal cancers. *Clin. Cancer Res.* 14:8061–8069. doi:10.1158/1078-0432.CCR-08-1431.
- Jover, R., P. Zapater, A. Castells, X. Llor, M. Andreu, J. Cubiella, V. Piñol, R.M. Xicola, L. Bujanda, J.M. Reñé, J. Clofent, X. Bessa, J.D. Morillas, D. Nicolás-Pérez, A. Payá, and C. Alenda. 2006. Mismatch repair status in the prediction of benefit from adjuvant fluorouracil chemotherapy in colorectal cancer. *Gut*. 55:848–855. doi:10.1136/gut.2005.073015.
- Kadyrov, F.A., L. Dzantiev, N. Constantin, and P. Modrich. 2006. Endonucleolytic function of MutLalpha in human mismatch repair. *Cell*. 126:297–308. doi:10.1016/j.cell.2006.05.039.
- Karran, P., and N. Attard. 2008. Thiopurines in current medical practice: molecular mechanisms and contributions to therapy-related cancer. *Nat. Rev. Cancer*. 8:24–36. doi:10.1038/nrc2292.
- Kastan, M.B., and J. Bartek. 2004. Cell-cycle checkpoints and cancer. *Nature*. 432:316–323. doi:10.1038/nature03097.
- Khoronenkova, S. V., and G.L. Dianov. 2015. ATM prevents DSB formation by coordinating SSB repair and cell cycle progression. *Proc. Natl. Acad. Sci. U. S. A.* 112:3997–4002. doi:10.1073/pnas.1416031112.
- Kim, H.-Y., B.-S. Choi, S.S. Kim, T.-Y. Roh, J. Park, and C.-H. Yoon. 2014. NUCKS1, a novel Tat coactivator, plays a crucial role in HIV-1 replication by increasing Tat-mediated viral transcription on the HIV-1 LTR promoter. *Retrovirology*. 11:67. doi:10.1186/s12977-014-0067-y.
- Koff, A., A. Giordano, D. Desai, K. Yamashita, J.W. Harper, S. Elledge, T. Nishimoto, D.O. Morgan, B.R. Franza, and J.M. Roberts. 1992. Formation and activation of a cyclin E-cdk2 complex during the G1 phase of the human cell cycle. *Science*. 257:1689–1694.
- Koshiji, M., K.K.-W. To, S. Hammer, K. Kumamoto, A.L. Harris, P. Modrich, and L.E. Huang. 2005. HIF-1alpha induces genetic instability by transcriptionally downregulating MutSalphalpha expression. *Mol. Cell*. 17:793–803. doi:10.1016/j.molcel.2005.02.015.
- Kossatz, U., N. Dietrich, L. Zender, J. Buer, M.P. Manns, and N.P. Malek. 2004. Skp2-dependent degradation of p27kip1 is essential for cell cycle progression. *Genes Dev.* 18:2602–2607. doi:10.1101/gad.321004.
- Krejci, L., V. Altmannova, M. Spirek, and X. Zhao. 2012. Homologous recombination and its regulation. *Nucleic Acids Res.* 40:5795–5818. doi:10.1093/nar/gks270.
- Kunkel, T.A., and D.A. Erie. 2005. DNA mismatch repair. *Annu. Rev. Biochem.* 74:681–710. doi:10.1146/annurev.biochem.74.082803.133243.
- Kunkel, T.A., and D.A. Erie. 2015. Eukaryotic Mismatch Repair in Relation to DNA Replication. *Annu. Rev. Genet.* 49:291–313. doi:10.1146/annurev-genet-112414-054722.
- Kuo, L.J., and L.-X. Yang. 2008. Gamma-H2AX - a novel biomarker for DNA double-strand breaks. *In Vivo*. 22:305–309.
- Lakin, N.D., and S.P. Jackson. 1999. Regulation of p53 in response to DNA damage. *Oncogene*. 18:7644–7655. doi:10.1038/sj.onc.1203015.
- Lamers, M.H., A. Perrakis, J.H. Enzlin, H.H. Winterwerp, N. de Wind, and T.K. Sixma. 2000. The crystal structure of DNA mismatch repair protein MutS binding to a G x T mismatch. *Nature*.

- 407:711–717. doi:10.1038/35037523.
- Lane, D.P. 1992. Cancer. p53, guardian of the genome. *Nature*. 358:15–16. doi:10.1038/358015a0.
- Lee, J.-B., W.-K. Cho, J. Park, Y. Jeon, D. Kim, S.H. Lee, and R. Fishel. 2014. Single-molecule views of MutS on mismatched DNA. *DNA Repair (Amst)*. 20:82–93. doi:10.1016/j.dnarep.2014.02.014.
- Leroy, B., L. Girard, A. Hollestelle, J.D. Minna, A.F. Gazdar, and T. Soussi. 2014. Analysis of TP53 Mutation Status in Human Cancer Cell Lines: A Reassessment. *Hum. Mutat.* 35:756–765. doi:10.1002/humu.22556.
- Li, F., G. Mao, D. Tong, J. Huang, L. Gu, W. Yang, and G.-M. Li. 2013. The Histone Mark H3K36me3 Regulates Human DNA Mismatch Repair through its Interaction with MutS α . *Cell*. 153:590–600. doi:10.1016/j.cell.2013.03.025.
- Li, G.M. 1999. The role of mismatch repair in DNA damage-induced apoptosis. *Oncol. Res.* 11:393–400.
- Li, L.S., J.C. Morales, A. Hwang, M.W. Wagner, and D.A. Boothman. 2008. DNA Mismatch Repair-dependent Activation of c-Abl/p73 α /GADD45 α -mediated Apoptosis. *J. Biol. Chem.* 283:21394–21403. doi:10.1074/jbc.M709954200.
- Li, L.S., J.C. Morales, M. Veigl, D. Sedwick, S. Greer, M. Meyers, M. Wagner, R. Fishel, and D.A. Boothman. 2009. DNA mismatch repair (MMR)-dependent 5-fluorouracil cytotoxicity and the potential for new therapeutic targets. *Br. J. Pharmacol.* 158:679–692. doi:10.1111/j.1476-5381.2009.00423.x.
- Lieber, M.R. 2010. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu. Rev. Biochem.* 79:181–211. doi:10.1146/annurev.biochem.052308.093131.
- Lindahl, T. 1993. Instability and decay of the primary structure of DNA. *Nature*. 362:709–715. doi:10.1038/362709a0.
- Lindahl, T., and B. Nyberg. 1972. Rate of depurination of native deoxyribonucleic acid. *Biochemistry*. 11:3610–3618.
- Liu, T., S. Tan, Y. Xu, F. Meng, C. Yang, and G. Lou. 2015. Increased NUCKS expression is a risk factor for poor prognosis and recurrence in endometrial cancer. *Am. J. Cancer Res.* 5:3659–3667.
- Loeb, L.A. 1994. Microsatellite instability: marker of a mutator phenotype in cancer. *Cancer Res.* 54:5059–5063.
- Longley, M.J., A.J. Pierce, and P. Modrich. 1997. DNA polymerase delta is required for human mismatch repair in vitro. *J. Biol. Chem.* 272:10917–10921.
- Lotito, L., A. Russo, S. Bueno, G. Chillemi, M.V. Fogli, and G. Capranico. 2009. A specific transcriptional response of yeast cells to camptothecin dependent on the Swi4 and Mbp1 factors. *Eur. J. Pharmacol.* 603:29–36. doi:10.1016/j.ejphar.2008.12.002.
- Lu, Y. 2014. Epigenetic silencing of the DNA mismatch repair gene, MLH1, induced by hypoxic stress in a pathway dependent on the histone demethylase, LSD1. 8:501–513. doi:10.1016/j.celrep.2014.06.035.
- Lundin, C., M. North, K. Erixon, K. Walters, D. Jenssen, A.S.H. Goldman, and T. Helleday. 2005. Methyl methanesulfonate (MMS) produces heat-labile DNA damage but no detectable in vivo DNA double-strand breaks. *Nucleic Acids Res.* 33:3799–3811. doi:10.1093/nar/gki681.
- Markkanen, E., R. Fischer, M. Ledentcova, B.M. Kessler, and G.L. Dianov. 2015. Cells deficient in base-excision repair reveal cancer hallmarks originating from adjustments to genetic instability. *Nucleic Acids Res.* 43:3667–3679. doi:10.1093/nar/gkv222.
- Markowitz, S., J. Wang, L. Myeroff, R. Parsons, L. Sun, J. Lutterbaugh, R.S. Fan, E. Zborowska, K.W. Kinzler, and B. Vogelstein. 1995. Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability. *Science*. 268:1336–1338.
- Martin, S.A., M. Hewish, D. Sims, C.J. Lord, and A. Ashworth. 2011. Parallel high-throughput RNA

- interference screens identify PINK1 as a potential therapeutic target for the treatment of DNA mismatch repair-deficient cancers. *Cancer Res.* 71:1836–1848. doi:10.1158/0008-5472.CAN-10-2836.
- Martin, S.A., N. McCabe, M. Mullarkey, R. Cummins, D.J. Burgess, Y. Nakabeppu, S. Oka, E. Kay, C.J. Lord, and A. Ashworth. 2010. DNA Polymerases as Potential Therapeutic Targets for Cancers Deficient in the DNA Mismatch Repair Proteins MSH2 or MLH1. *Cancer Cell.* 17:235–248. doi:10.1016/j.ccr.2009.12.046.
- Martin, S.A., A. McCarthy, L.J. Barber, D.J. Burgess, S. Parry, C.J. Lord, and A. Ashworth. 2009. Methotrexate induces oxidative DNA damage and is selectively lethal to tumour cells with defects in the DNA mismatch repair gene MSH2. *EMBO Mol. Med.* 1:323–337. doi:10.1002/emmm.200900040.
- Matsuoka, S., B. a Ballif, A. Smogorzewska, E.R. McDonald, K.E. Hurov, J. Luo, C.E. Bakalarski, Z. Zhao, N. Solimini, Y. Lerenthal, Y. Shiloh, S.P. Gygi, and S.J. Elledge. 2007. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* (80-.). 316:1160–1166. doi:10.1126/science.1140321.
- Mazur, D.J., M.L. Mendillo, and R.D. Kolodner. 2006. Inhibition of Msh6 ATPase activity by mispaired DNA induces a Msh2(ATP)-Msh6(ATP) state capable of hydrolysis-independent movement along DNA. *Mol. Cell.* 22:39–49. doi:10.1016/j.molcel.2006.02.010.
- McVey, M., V.Y. Khodaverdian, D. Meyer, P.G. Cerqueira, and W.-D. Heyer. 2016. Eukaryotic DNA Polymerases in Homologous Recombination. *Annu. Rev. Genet.* 50:393–421. doi:10.1146/annurev-genet-120215-035243.
- Mendillo, M.L., V. V Hargreaves, J.W. Jamison, A.O. Mo, S. Li, C.D. Putnam, V.L.J. Woods, and R.D. Kolodner. 2009. A conserved MutS homolog connector domain interface interacts with MutL homologs. *Proc. Natl. Acad. Sci. U. S. A.* 106:22223–22228. doi:10.1073/pnas.0912250106.
- Mihaylova, V.T., R.S. Bindra, J. Yuan, D. Campisi, L. Narayanan, R. Jensen, F. Giordano, R.S. Johnson, S. Rockwell, and P.M. Glazer. 2003. Decreased expression of the DNA mismatch repair gene Mlh1 under hypoxic stress in mammalian cells. *Mol. Cell. Biol.* 23:3265–3273.
- Moll, U.M., and O. Petrenko. 2003. The MDM2-p53 interaction. *Mol. Cancer Res.* 1:1001–1008.
- Motoshima, H., B.J. Goldstein, M. Igata, and E. Araki. 2006. AMPK and cell proliferation – AMPK as a therapeutic target for atherosclerosis and cancer. *J. Physiol.* 574:63–71. doi:10.1113/jphysiol.2006.108324.
- Mrkonjic, M., S. Raptis, R.C. Green, N. Monga, D. Daftary, E. Dicks, H.B. Younghusband, P.S. Parfrey, S.S. Gallinger, J.R. McLaughlin, J.A. Knight, and B. Bapat. 2007. MSH2 118T>C and MSH6 159C>T promoter polymorphisms and the risk of colorectal cancer. *Carcinogenesis.* 28:2575–2580. doi:10.1093/carcin/bgm229.
- Muller, A., and R. Fishel. 2002. Mismatch repair and the hereditary non-polyposis colorectal cancer syndrome (HNPCC). *Cancer Invest.* 20:102–109.
- Munshi, P.N., M. Lubin, and J.R. Bertino. 2014. 6-Thioguanine: A Drug With Unrealized Potential for Cancer Therapy. *Oncologist.* 19:760–765. doi:10.1634/theoncologist.2014-0178.
- Najdi, R., R.F. Holcombe, and M.L. Waterman. 2011. Wnt signaling and colon carcinogenesis: Beyond APC. *J. Carcinog.* 10. doi:10.4103/1477-3163.78111.
- Naylor, T.L., J. Greshock, Y. Wang, T. Colligon, Q.C. Yu, V. Clemmer, T.Z. Zaks, and B.L. Weber. 2005. High resolution genomic analysis of sporadic breast cancer using array-based comparative genomic hybridization. *Breast Cancer Res.* 7:R1186-98. doi:10.1186/bcr1356.
- Nikulenkov, F., C. Spinnler, H. Li, C. Tonelli, Y. Shi, M. Turunen, T. Kivioja, I. Ignatiev, A. Kel, J. Taipale, and G. Selivanova. 2012. Insights into p53 transcriptional function via genome-wide chromatin occupancy and gene expression analysis. *Cell Death Differ.* 19:1992–2002. doi:10.1038/cdd.2012.89.
- Obmolova, G., C. Ban, P. Hsieh, and W. Yang. 2000. Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. *Nature.* 407:703–710. doi:10.1038/35037509.
- O'Brien, J., I. Wilson, T. Orton, and F. Pognan. 2000. Investigation of the Alamar Blue (resazurin)

- fluorescent dye for the assessment of mammalian cell cytotoxicity. *Eur. J. Biochem.* 267:5421–5426.
- Okimoto, S., J. Sun, A. Fukuto, Y. Horikoshi, S. Matsuda, T. Matsuda, M. Ikura, T. Ikura, S. Machida, H. Kurumizaka, Y. Miyamoto, M. Oka, Y. Yoneda, Y. Kiuchi, and S. Tashiro. 2015. hCAS/CSE1L regulates RAD51 distribution and focus formation for homologous recombinational repair. *Genes Cells.* 20:681–694. doi:10.1111/gtc.12262.
- Olive, P.L., and J.P. Banath. 2006. The comet assay: a method to measure DNA damage in individual cells. *Nat. Protoc.* 1:23–29. doi:10.1038/nprot.2006.5.
- Ostvold, A.C., J. Holtlund, and S.G. Laland. 1985. A novel, highly phosphorylated protein, of the high-mobility group type, present in a variety of proliferating and non-proliferating mammalian cells. *Eur. J. Biochem.* 153:469–475.
- Ostvold, A.C., I. Hullstein, and S.G. Laland. 1992. The phosphate groups of the high mobility group like protein P1 strengthens its affinity for DNA. *Biochem. Biophys. Res. Commun.* 185:1091–1097.
- Parpys, A.C., K. Kratz, M.C. Speed, S.G. Leung, D. Schild, and C. Wiese. 2014. RAD51AP1-deficiency in vertebrate cells impairs DNA replication. *DNA Repair (Amst).* 24:87–97. doi:http://dx.doi.org/10.1016/j.dnarep.2014.09.007.
- Parpys, A.C., W. Zhao, N. Sharma, T. Groesser, F. Liang, D.G. Maranon, S.G. Leung, K. Grundt, E. Dray, R. Idate, A.C. Ostvold, D. Schild, P. Sung, and C. Wiese. 2015. NUCKS1 is a novel RAD51AP1 paralog important for homologous recombination and genome stability. *Nucleic Acids Res.* doi:10.1093/nar/gkv859.
- Patidar, P.L., E.A. Motea, F.J. Fattah, Y. Zhou, J.C. Morales, Y. Xie, H.R. Garner, and D.A. Boothman. 2016. The Kub5-Hera/RPRD1B interactome: a novel role in preserving genetic stability by regulating DNA mismatch repair. *Nucleic Acids Res.* 44:1718–1731. doi:10.1093/nar/gkv1492.
- Peltomaki, P. 2003. Role of DNA mismatch repair defects in the pathogenesis of human cancer. *J. Clin. Oncol.* 21:1174–1179.
- Pluciennik, A., L. Dzantiev, R.R. Iyer, N. Constantin, F.A. Kadyrov, and P. Modrich. 2010. PCNA function in the activation and strand direction of MutLalpha endonuclease in mismatch repair. *Proc. Natl. Acad. Sci. U. S. A.* 107:16066–16071. doi:10.1073/pnas.1010662107.
- Poletto, M., A.J. Legrand, S.C. Fletcher, and G.L. Dianov. 2016. p53 coordinates base excision repair to prevent genomic instability. *Nucleic Acids Res.* 44:3165–3175. doi:10.1093/nar/gkw015.
- Puebla-Osorio, N., D.B. Lacey, F.W. Alt, and C. Zhu. 2006. Early embryonic lethality due to targeted inactivation of DNA ligase III. *Mol. Cell. Biol.* 26:3935–3941. doi:10.1128/MCB.26.10.3935-3941.2006.
- Qiu, B., W. Han, and V. Tergaonkar. 2015. NUCKS: a potential biomarker in cancer and metabolic disease. *Clin. Sci. (Lond).* 128:715–721. doi:10.1042/CS20140656.
- Qiu, B., X. Shi, E.T. Wong, J. Lim, M. Bezzi, D. Low, Q. Zhou, S.C. Akincilar, M. Lakshmanan, H.L.F. Swa, J.M.L. Tham, J. Gunaratne, K.K.Y. Cheng, W. Hong, K.S.L. Lam, M. Ikawa, E. Guccione, A. Xu, W. Han, and V. Tergaonkar. 2014. NUCKS is a positive transcriptional regulator of insulin signaling. *Cell Rep.* 7:1876–86. doi:10.1016/j.celrep.2014.05.030.
- Rahman, N., S. Seal, D. Thompson, P. Kelly, A. Renwick, A. Elliott, S. Reid, K. Spanova, R. Barfoot, T. Chagtai, H. Jayatilake, L. McGuffog, S. Hanks, D.G. Evans, D. Eccles, D.F. Easton, and M.R. Stratton. 2007. PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nat. Genet.* 39:165–167. doi:10.1038/ng1959.
- Ran, F.A., P.D. Hsu, J. Wright, V. Agarwala, D.A. Scott, and F. Zhang. 2013. Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* 8:2281–2308. doi:10.1038/nprot.2013.143.
- Renkonen, E., Y. Zhang, H. Lohi, R. Salovaara, W.M. Abdel-Rahman, M. Nilbert, K. Aittomaki, H.J. Jarvinen, J.-P. Mecklin, A. Lindblom, and P. Peltomaki. 2003. Altered expression of MLH1, MSH2, and MSH6 in predisposition to hereditary nonpolyposis colorectal cancer. *J. Clin.*

- Oncol.* 21:3629–3637. doi:10.1200/JCO.2003.03.181.
- Ribic, C.M., D.J. Sargent, M.J. Moore, S.N. Thibodeau, A.J. French, R.M. Goldberg, S.R. Hamilton, P. Laurent-Puig, R. Gryfe, L.E. Shepherd, D. Tu, M. Redston, and S. Gallinger. 2003. Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer. *N. Engl. J. Med.* 349:247–257. doi:10.1056/NEJMoa022289.
- Samowitz, W.S., K. Curtin, R.K. Wolff, H. Albertsen, C. Sweeney, B.J. Caan, C.M. Ulrich, J.D. Potter, and M.L. Slattery. 2008. The MLH1 -93 G>A promoter polymorphism and genetic and epigenetic alterations in colon cancer. *Genes. Chromosomes Cancer.* 47:835–844. doi:10.1002/gcc.20584.
- Sargent, L.M., M.X. Ensell, A.-C. Ostvold, K.T. Baldwin, M.L. Kashon, D.T. Lowry, J.R. Senft, A.M. Jefferson, R.C. Johnson, Z. Li, F.L. Tyson, and S.H. Reynolds. 2008. Chromosomal changes in high- and low-invasive mouse lung adenocarcinoma cell strains derived from early passage mouse lung adenocarcinoma cell strains. *Toxicol. Appl. Pharmacol.* 233:81–91. doi:10.1016/j.taap.2008.01.031.
- Sartori, A.A., C. Lukas, J. Coates, M. Mistrik, S. Fu, J. Bartek, R. Baer, J. Lukas, and S.P. Jackson. 2007. Human CtIP promotes DNA end resection. *Nature.* 450:509–514. doi:10.1038/nature06337.
- Schmutte, C., R.C. Marinescu, M.M. Sadoff, S. Guerrette, J. Overhauser, and R. Fishel. 1998. Human exonuclease I interacts with the mismatch repair protein hMSH2. *Cancer Res.* 58:4537–4542.
- Scholzen, T., and J. Gerdes. 2000. The Ki-67 protein: from the known and the unknown. *J. Cell. Physiol.* 182:311–322. doi:10.1002/(SICI)1097-4652(200003)182:3<311::AID-JCP1>3.0.CO;2-9.
- Schroering, A.G., M.A. Edelbrock, T.J. Richards, and K.J. Williams. 2007. The cell cycle and DNA mismatch repair. *Exp. Cell Res.* 313:292–304. doi:10.1016/j.yexcr.2006.10.018.
- Shen, H., L. Wang, X. Ge, C.-F. Jiang, Z.-M. Shi, D.-M. Li, W.-T. Liu, X. Yu, and Y.-Q. Shu. 2016. MicroRNA-137 inhibits tumor growth and sensitizes chemosensitivity to paclitaxel and cisplatin in lung cancer. *Oncotarget.* doi:10.18632/oncotarget.8011.
- Sherr, C.J., and F. McCormick. 2002. The RB and p53 pathways in cancer. *Cancer Cell.* 2:103–112.
- Shiloh, Y., and Y. Ziv. 2013. The ATM protein kinase: regulating the cellular response to genotoxic stress, and more. *Nat. Rev. Mol. Cell Biol.* 14:197–210. doi:10.1038/nrm3546.
- Shin, K.-H., J.-H. Shin, J.-H. Kim, and J.-G. Park. 2002. Mutational analysis of promoters of mismatch repair genes hMSH2 and hMLH1 in hereditary nonpolyposis colorectal cancer and early onset colorectal cancer patients: identification of three novel germ-line mutations in promoter of the hMSH2 gene. *Cancer Res.* 62:38–42.
- Siegel, R.L., K.D. Miller, and A. Jemal. 2016. Cancer statistics, 2016. *CA. Cancer J. Clin.* 66:7–30. doi:10.3322/caac.21332.
- Smale, S.T. 2010. Luciferase assay. *Cold Spring Harb. Protoc.* 2010:pdb.prot5421. doi:10.1101/pdb.prot5421.
- Stvold, A.C., J.H. Norum, S. Mathiesen, B. Wanvik, I. Sefland, and K. Grundt. 2001. Molecular cloning of a mammalian nuclear phosphoprotein NUCKS, which serves as a substrate for Cdk1 in vivo. *Eur. J. Biochem.* 268:2430–2440. doi:10.1046/j.1432-1327.2001.02120.x.
- Sugano, N., T. Suda, T.-I. Godai, K. Tsuchida, M. Shiozawa, H. Sekiguchi, M. Yoshihara, S. Matsukuma, Y. Sakuma, E. Tsuchiya, Y. Kameda, M. Akaike, and Y. Miyagi. 2010. MDM2 gene amplification in colorectal cancer is associated with disease progression at the primary site, but inversely correlated with distant metastasis. *Genes. Chromosomes Cancer.* 49:620–629. doi:10.1002/gcc.20774.
- Sung, P., L. Krejci, S. Van Komen, and M.G. Sehorn. 2003. Rad51 recombinase and recombination mediators. *J. Biol. Chem.* 278:42729–42732. doi:10.1074/jbc.R300027200.
- Sur, S., R. Pagliarini, F. Bunz, C. Rago, L.A.J. Diaz, K.W. Kinzler, B. Vogelstein, and N. Papadopoulos. 2009. A panel of isogenic human cancer cells suggests a therapeutic approach for cancers

- with inactivated p53. *Proc. Natl. Acad. Sci. U. S. A.* 106:3964–3969. doi:10.1073/pnas.0813333106.
- Swann, P.F., T.R. Waters, D.C. Moulton, Y.Z. Xu, Q. Zheng, M. Edwards, and R. Mace. 1996. Role of postreplicative DNA mismatch repair in the cytotoxic action of thioguanine. *Science*. 273:1109–1111.
- Tajima, A., M.T. Hess, B.L. Cabrera, R.D. Kolodner, and J.M. Carethers. 2004. The mismatch repair complex hMutS alpha recognizes 5-fluorouracil-modified DNA: implications for chemosensitivity and resistance. *Gastroenterology*. 127:1678–1684.
- Tanaka, M., J.S. Lai, and W. Herr. 1992. Promoter-selective activation domains in Oct-1 and Oct-2 direct differential activation of an snRNA and mRNA promoter. *Cell*. 68:755–767.
- Tebbs, R.S., L.H. Thompson, and J.E. Cleaver. 2003. Rescue of Xrcc1 knockout mouse embryo lethality by transgene-complementation. *DNA Repair (Amst)*. 2:1405–1417. doi:http://dx.doi.org/10.1016/j.dnarep.2003.08.007.
- Tennen, R.I., J.E. Haye, H.D. Wijayatilake, T. Arlow, D. Ponzio, and A.E. Gammie. 2013. Cell-cycle and DNA damage regulation of the DNA mismatch repair protein Msh2 occurs at the transcriptional and post-transcriptional level. *DNA Repair (Amst)*. 12:97–109. doi:10.1016/j.dnarep.2012.11.002.
- Thibodeau, S.N., A.J. French, P.C. Roche, J.M. Cunningham, D.J. Tester, N.M. Lindor, G. Moslein, S.M. Baker, R.M. Liskay, L.J. Burgart, R. Honchel, and K.C. Halling. 1996. Altered expression of hMSH2 and hMLH1 in tumors with microsatellite instability and genetic alterations in mismatch repair genes. *Cancer Res.* 56:4836–4840.
- Thompson, L.H. 2005. Unraveling the Fanconi anemia-DNA repair connection. *Nat. Genet.* 37:921–922. doi:10.1038/ng0905-921.
- Tice, R., and R. Setlow. 1985. DNA repair and replication in aging organisms and cells. In “Handbook of the Biology of Aging, second edition” C.E. Finch and E.L. Schneider. Van Nostrand Reinhold, New York. 173–224.
- Tran, H.T., D.A. Gordenin, and M.A. Resnick. 1999. The 3′→5′ Exonucleases of DNA Polymerases δ and ϵ and the 5′→3′ Exonuclease Exo1 Have Major Roles in Postreplication Mutation Avoidance in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 19:2000–2007.
- Troester, M.A., K.A. Hoadley, J.S. Parker, and C.M. Perou. 2004. Prediction of toxicant-specific gene expression signatures after chemotherapeutic treatment of breast cell lines. *Environ. Health Perspect.* 112:1607–1613.
- Ullman, T.A., and S.H. Itzkowitz. 2011. Intestinal inflammation and cancer. *Gastroenterology*. 140:1807–1816. doi:10.1053/j.gastro.2011.01.057.
- Umar, A., C.R. Boland, J.P. Terdiman, S. Syngal, A. de la Chapelle, J. Ruschoff, R. Fishel, N.M. Lindor, L.J. Burgart, R. Hamelin, S.R. Hamilton, R.A. Hiatt, J. Jass, A. Lindblom, H.T. Lynch, P. Peltomaki, S.D. Ramsey, M.A. Rodriguez-Bigas, H.F.A. Vasen, E.T. Hawk, J.C. Barrett, A.N. Freedman, and S. Srivastava. 2004. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *J. Natl. Cancer Inst.* 96:261–268.
- Valeri, N., P. Gasparini, C. Braconi, A. Paone, F. Lovat, M. Fabbri, K.M. Sumani, H. Alder, D. Amadori, T. Patel, G.J. Nuovo, R. Fishel, and C.M. Croce. 2010a. MicroRNA-21 induces resistance to 5-fluorouracil by down-regulating human DNA MutS homolog 2 (hMSH2). *Proc. Natl. Acad. Sci. U. S. A.* 107:21098–21103. doi:10.1073/pnas.1015541107.
- Valeri, N., P. Gasparini, M. Fabbri, C. Braconi, A. Veronese, F. Lovat, B. Adair, I. Vannini, F. Fanini, A. Bottoni, S. Costinean, S.K. Sandhu, G.J. Nuovo, H. Alder, R. Gafa, F. Calore, M. Ferracin, G. Lanza, S. Volinia, M. Negrini, M.A. McIlhatton, D. Amadori, R. Fishel, and C.M. Croce. 2010b. Modulation of mismatch repair and genomic stability by miR-155. *Proc. Natl. Acad. Sci. U. S. A.* 107:6982–6987. doi:10.1073/pnas.1002472107.
- Vaz, B., M. Popovic, J.A. Newman, J. Fielden, H. Aitkenhead, S. Halder, A.N. Singh, I. Vendrell, R. Fischer, I. Torrecilla, N. Drobnitzky, R. Freire, D.J. Amor, P.J. Lockhart, B.M. Kessler, G.W.

- McKenna, O. Gileadi, and K. Ramadan. 2016. Metalloprotease SPRTN/DVC1 Orchestrates Replication-Coupled DNA-Protein Crosslink Repair. *Mol. Cell.* 64:704–719. doi:10.1016/j.molcel.2016.09.032.
- Vilenchik, M.M., and A.G. Knudson. 2003. Endogenous DNA double-strand breaks: production, fidelity of repair, and induction of cancer. *Proc. Natl. Acad. Sci. U. S. A.* 100:12871–12876. doi:10.1073/pnas.2135498100.
- Wang, Q., S. Xie, J. Chen, K. Fukasawa, U. Naik, F. Traganos, Z. Darzynkiewicz, M. Jhanwar-Uniyal, and W. Dai. 2002. Cell cycle arrest and apoptosis induced by human Polo-like kinase 3 is mediated through perturbation of microtubule integrity. *Mol. Cell. Biol.* 22:3450–3459.
- Wang, T., J.J. Wei, D.M. Sabatini, and E.S. Lander. 2014. Genetic screens in human cells using the CRISPR-Cas9 system. *Science.* 343:80–84. doi:10.1126/science.1246981.
- Warren, J.J., T.J. Pohlhaus, A. Changela, R.R. Iyer, P.L. Modrich, and L.S. Beese. 2007. Structure of the human MutS α DNA lesion recognition complex. *Mol. Cell.* 26:579–592. doi:10.1016/j.molcel.2007.04.018.
- Welch, P.L., and M.C. King. 2001. BRCA1 and BRCA2 and the genetics of breast and ovarian cancer. *Hum. Mol. Genet.* 10:705–713.
- Wiese, C., E. Dray, T. Groesser, J.S. Filipp, I. Shi, D.W. Collins, M.-S. Tsai, G. Williams, B. Rydberg, P. Sung, and D. Schild. 2007. Promotion of Homologous Recombination and Genomic Stability by RAD51AP1 via RAD51 Recombinase Enhancement. *Mol. Cell.* 28:482–490. doi:10.1016/j.molcel.2007.08.027.
- Wiśniewski, J.R., A. Zougman, S. Krüger, P. Ziolkowski, M. Pudełko, M. Bębenek, and M. Mann. 2008. Constitutive and dynamic phosphorylation and acetylation sites on NUCKS, a hypermodified nuclear protein, studied by quantitative proteomics. *Proteins Struct. Funct. Bioinforma.* 73:710–718. doi:10.1002/prot.22104.
- Wold, M.S. 1997. Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu. Rev. Biochem.* 66:61–92. doi:10.1146/annurev.biochem.66.1.61.
- Wolpin, B.M., and R.J. Mayer. 2008. Systemic Treatment of Colorectal Cancer. *Gastroenterology.* 134:1296–1310. doi:10.1053/j.gastro.2008.02.098.
- Wood, R.D. 1997. Nucleotide excision repair in mammalian cells. *J. Biol. Chem.* 272:23465–23468.
- Wyatt, H.D.M., and S.C. West. 2014. Holliday junction resolvases. *Cold Spring Harb. Perspect. Biol.* 6:a023192. doi:10.1101/cshperspect.a023192.
- Yan, H., Z. Dobbie, S.B. Gruber, S. Markowitz, K. Romans, F.M. Giardiello, K.W. Kinzler, and B. Vogelstein. 2002. Small changes in expression affect predisposition to tumorigenesis. *Nat. Genet.* 30:25–26. doi:10.1038/ng799.
- Yang, M., X. Wang, Q. Zhao, T. Liu, G. Yao, W. Chen, Z. Li, X. Huang, and Y. Zhang. 2014. Combined evaluation of the expression of NUCKS and Ki-67 proteins as independent prognostic factors for patients with gastric adenocarcinoma. *Tumour Biol.* 35:7505–7512. doi:10.1007/s13277-014-1880-6.
- Youn, C.-K., H.-J. Cho, S.-H. Kim, H.-B. Kim, M.-H. Kim, I.-Y. Chang, J.-S. Lee, M.-H. Chung, K.-S. Hahm, and H.J. You. 2005. Bcl-2 expression suppresses mismatch repair activity through inhibition of E2F transcriptional activity. *Nat. Cell Biol.* 7:137–147. doi:10.1038/ncb1215.
- Yu, Y., Y. Wang, X. Ren, A. Tsuyada, A. Li, L.J. Liu, and S.E. Wang. 2010. Context-dependent bidirectional regulation of the MutS homolog 2 by transforming growth factor β contributes to chemoresistance in breast cancer cells. *Mol. Cancer Res.* 8:1633–1642. doi:10.1158/1541-7786.MCR-10-0362.
- Yue, Y., S.G. Leung, Y. Liu, Y. Huang, K. Grundt, A.-C. Ostvold, K.-Y. Jen, D. Schild, J.-H. Mao, and C. Wiese. 2016. Nucks1 synergizes with Trp53 to promote radiation lymphomagenesis in mice. *Oncotarget.* doi:10.18632/oncotarget.11297.
- Zhang, B., B. Zhang, X. Chen, S. Bae, K. Singh, M.K. Washington, and P.K. Datta. 2014a. Loss of Smad4 in colorectal cancer induces resistance to 5-fluorouracil through activating Akt

- pathway. *Br. J. Cancer*. 110:946–957. doi:10.1038/bjc.2013.789.
- Zhang, M., S. Xiang, H.-Y. Joo, L. Wang, K.A. Williams, W. Liu, C. Hu, D. Tong, J. Haakenson, C. Wang, S. Zhang, R.E. Pavlovicz, A. Jones, K.H. Schmidt, J. Tang, H. Dong, B. Shan, B. Fang, R. Radhakrishnan, P.M. Glazer, P. Matthias, J. Koomen, E. Seto, G. Bepler, S. V Nicosia, J. Chen, C. Li, L. Gu, G.-M. Li, W. Bai, H. Wang, and X. Zhang. 2014b. HDAC6 deacetylates and ubiquitinates MSH2 to maintain proper levels of MutSalph. *Mol. Cell*. 55:31–46. doi:10.1016/j.molcel.2014.04.028.
- Zhang, Y., F. Yuan, S.R. Presnell, K. Tian, Y. Gao, A.E. Tomkinson, L. Gu, and G.-M. Li. 2005. Reconstitution of 5'-directed human mismatch repair in a purified system. *Cell*. 122:693–705. doi:10.1016/j.cell.2005.06.027.
- Zhao, H., B. Thienpont, B.T. Yesilyurt, M. Moisse, J. Reumers, L. Coenegrachts, X. Sagaert, S. Schrauwen, D. Smeets, G. Matthijs, S. Aerts, J. Cools, A. Metcalf, A. Spurdle, F. Amant, and D. Lambrechts. 2014. Mismatch repair deficiency endows tumors with a unique mutation signature and sensitivity to DNA double-strand breaks. *Elife*. 3. doi:10.7554/eLife.02725.
- Zhou, B.B., and S.J. Elledge. 2000. The DNA damage response: putting checkpoints in perspective. *Nature*. 408:433–439. doi:10.1038/35044005.
- Zhu, Y., T.M. Luo, C. Jobin, and H.A. Young. 2011. Gut Microbiota and Probiotics in Colon Tumorigenesis. *Cancer Lett*. 309:119–127. doi:10.1016/j.canlet.2011.06.004.