



**A University of Sussex PhD thesis**

Available online via Sussex Research Online:

<http://sro.sussex.ac.uk/>

This thesis is protected by copyright which belongs to the author.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Please visit Sussex Research Online for more information and further details

# INVESTIGATING THE EFFECT OF TDP2 ABSENCE IN CRISPR/CAS9 KNOCK-OUT OR PATIENT-DERIVED CELL LINES

A thesis submitted to the University of Sussex for the degree of  
Doctor of Philosophy

By Guido Zagnoli Vieira

September 2018



## **Declaration**

I hereby declare that this thesis has not been, and will not be, submitted in whole or in part to another University for the award of any other degree.

**Guido Zagnoli Vieira**



## **Acknowledgements**

I would like to thank my supervisor Prof. Keith Caldecott for his guidance and infinite patience. I am also very grateful to all members of the Caldecott Lab, past and present, who were there for me along the way and made a great place to work. Finally, I want to thank all my friends and family for the support (laughs, drinks and tears) shared over the last four years.

UNIVERSITY OF SUSSEX  
DOCTOR OF PHILOSOPHY GENOME STABILITY  
INVESTIGATING TDP2 ABSENCE ON CRISPR-CAS9 *TDP2*<sup>-/-</sup> AND  
SCAR23 PATIENT-DERIVED CELL LINES  
SUMMARY (300 words)

The DNA repair enzyme 5'-Tyrosyl DNA phosphodiesterase 2 (TDP2) removes topoisomerase 2 peptide from the 5'-termini of abortive cleavage complexes. TDP2 activity has been implicated in viral infection, transcriptional regulation, resistance to chemotherapy, and *TDP2* is mutated in Spinocerebellar Ataxia Autosomal Recessive 23 (SCAR23) syndrome. This thesis provides further evidence for the role of TDP2 in the DNA damage response and a better understanding of the molecular phenotypes observed in SCAR23 cells. By employing *TDP2*<sup>-/-</sup> cell lines, generated herein using CRISPR-Cas9, I demonstrate the importance of TDP2 for the repair of topoisomerase 2 induced DNA damage. Furthermore, I identify and characterise two additional isoforms of TDP2, namely TDP $\beta$  and TDP $\gamma$ . TDP2 $\beta$  was found to localise to the cytoplasm and not to be involved in nuclear DNA damage repair. On the other hand, TDP2 $\gamma$  was found to be a pan-cellular protein and to mediate survival and DSB repair following treatment with the topoisomerase 2 'poison', etoposide. In addition, I establish a new tyrosyl DNA phosphodiesterase (TDP) assay using whole cell extracts, to compare TDP activity in different cell types and cell lines and to support ongoing drug discovery efforts by the Sussex Drug Discovery Group. Finally, I identify new human mutations in TDP2 that disrupt DSB repair and resistance to TOP2 induced DSBs and in so doing confirm that TDP2 mutation is associated with spinocerebellar ataxia autosomal recessive 23 (SCAR23).

## Table of Contents

LIST OF ABBREVIATIONS .....	IV
INDEX OF FIGURES .....	XIII
INDEX OF TABLES .....	XVI
<b>1. CHAPTER ONE – INTRODUCTION .....</b>	<b>1</b>
1.1. DNA DAMAGE .....	1
1.2. DNA REPAIR .....	2
1.2.1. Mismatch repair .....	2
1.2.2. Nucleotide excision repair .....	3
1.2.3. Base excision repair & single-strand break repair .....	4
1.2.4. DNA double-strand break (DSB) repair .....	6
1.2.4.1. DNA damage response to DSBs .....	6
1.2.4.2. Homologous recombination .....	9
1.2.4.3. Non-homologous end-joining .....	11
1.2.4.4. Alternative NHEJ .....	13
1.3. TOPOISOMERASES AS SOURCES OF GENOMIC INSTABILITY .....	13
1.3.1. Topoisomerase type IA .....	14
1.3.2. Topoisomerase type IB .....	15
1.3.3. Topoisomerase type IIA .....	17
1.3.4. Topoisomerase type IIB .....	19
1.4. REPAIR OF TOPOISOMERASE CLEAVAGE COMPLEXES BY TYROSYL DNA PHOSPHODIESTERASES .....	20
1.4.1. Tyrosyl DNA phosphodiesterase 1 .....	20
1.4.2. Tyrosyl DNA phosphodiesterase 2 .....	22
1.5. AIM OF THESIS .....	25
<b>2. CHAPTER TWO – MATERIAL AND METHODS .....</b>	<b>26</b>
2.1. MAMMALIAN CELL CULTURE .....	26
2.1.1. siRNA transfection .....	26
2.1.2. Vector transfection .....	27
2.1.2.1. Neon Electroporation .....	27
2.1.2.2. GeneJuice transfection .....	27
2.2. TDP2 $\beta$ SEQUENCE CAPTURE .....	27
2.3. TDP2 <sup>-/-</sup> COMPLEMENTATION .....	28
2.4. SDS-POLYACRYLAMIDE GEL ELECTROPHORESIS (SDS-PAGE) OF PROTEINS .....	29
2.4.1. Western blotting .....	29
2.4.2. Silver staining .....	30
2.5. TDP2 IMMUNOPRECIPITATION .....	30
2.6. CRISPR-Cas9 .....	31
2.7. DNA SEQUENCING .....	31
2.7.1. Sanger sequencing .....	31
2.7.2. Whole exome sequencing .....	32
2.8. MICROSCOPY .....	32
2.8.1. $\gamma$ H2AX .....	32
2.8.2. TDP2 Immunofluorescence signal quantification .....	33
2.8.3. TDP2 isoform cellular localisation .....	33
2.9. CLONOGIC CELL SURVIVAL .....	34
2.10. TYROSYL DNA PHOSPHODIESTERASE ASSAY .....	35
2.11. BIOINFORMATICS .....	35
2.11.1. Sequence alignments .....	35
2.11.2. Cellular localisation .....	36
2.11.3. Kaplan-Meier survival .....	36
2.12. STATISTICS .....	36

<b>3.</b>	<b>CHAPTER THREE – GENERATION AND CHARACTERISATION OF <i>TDP2</i><sup>-/-</sup> CELLS.....</b>	<b>37</b>
3.1.	INTRODUCTION AND AIMS .....	37
3.1.1.	<i>Aim</i> .....	38
3.2.	RESULTS .....	38
3.2.1.	<i>TDP2 gRNA design and CRISPR-Cas9 system</i> .....	38
3.2.2.	<i>CRISPR-Cas9 TDP2<sup>-/-</sup> clone screening</i> .....	39
3.2.3.	<i>TDP2<sup>-/-</sup> characterisation through genome sequencing</i> .....	40
3.2.4.	<i>Reduced rates of DSB repair in the absence of TDP2</i> .....	42
3.2.5.	<i>TDP2<sup>-/-</sup> cells are hypersensitive to TOP2-induced DSBs</i> .....	43
3.3.	CONCLUSIONS AND DISCUSSION .....	44
<b>4.</b>	<b>CHAPTER FOUR – IDENTIFICATION AND CHARACTERISATION OF TDP2 ISOFORMS.....</b>	<b>48</b>
4.1.	INTRODUCTION AND AIMS .....	48
4.1.1.	<i>Aim</i> .....	49
4.2.	RESULTS .....	49
4.2.1.	<i>TDP2 isoform can be detect at 37kDa</i> .....	49
4.2.2.	<i>Deletion of TDP2 isoform leads to increased sensitivity to Topoisomerase 2 induced damage.</i> .....	50
4.2.3.	<i>Origin of the 37 kDa TDP2 isoform.</i> .....	51
4.2.4.	<i>TDP2<math>\beta</math> and TDP2<math>\gamma</math> are present in human cells</i> .....	52
4.2.5.	<i>TDP2 isoforms do not localise exclusively to nucleus</i> .....	53
4.2.6.	<i>Re-expression of TDP2<math>\gamma</math> isoform leads to recovery of DNA repair and survival</i> .....	54
4.3.	CONCLUSIONS AND DISCUSSION .....	55
<b>5.</b>	<b>CHAPTER FIVE – ESTABLISHING AN ASSAY FOR TYROSYL DNA PHOSPHODIESTERASE ACTIVITY IN WHOLE CELL EXTRACTS .....</b>	<b>60</b>
5.1.	INTRODUCTION AND AIMS .....	60
5.1.1.	<i>Aim</i> .....	62
5.2.	RESULTS .....	62
5.2.1.	<i>5'-Tyrosyl DNA phosphodiesterase (5'-TDP) assay</i> .....	63
5.2.2.	<i>3'-Tyrosyl DNA Phosphodiesterase (3'-TDP) Assay</i> .....	64
5.2.3.	<i>Versatility and sensitivity of the TDP assays</i> .....	65
5.2.4.	<i>TDP2 in Cancer</i> .....	66
5.3.	CONCLUSIONS AND DISCUSSION .....	67
<b>6.</b>	<b>CHAPTER SIX – INVESTIGATING THE EFFECT OF TDP2 ABSENCE IN SCAR23 FIBROBLASTS.....</b>	<b>71</b>
6.1.	INTRODUCTION AND AIMS .....	71
6.1.1.	<i>Aim</i> .....	72
6.2.	RESULTS .....	73
6.2.1.	<i>Spinocerebellar Ataxia, Autosomal Recessive 23 (SCAR23) patient of American origin</i> .....	73
6.2.2.	<i>Tyrosyl phosphodiesterase activity in patient fibroblasts</i> .....	74
6.2.3.	<i>DNA repair and cellular sensitivity of SCAR23 fibroblasts following induction of TOP2-induced DSBs</i> .....	75
6.2.4.	<i>SCAR23 patients of Italian origin</i> .....	76
6.3.	CONCLUSIONS AND DISCUSSION .....	77
<b>7.</b>	<b>CHAPTER SEVEN – DISCUSSION AND CONCLUDING REMARKS .....</b>	<b>80</b>
7.1.	ROLE OF TDP2 ISOFORMS.....	80
7.2.	TDP2 IN CANCER AND DRUG DISCOVERY.....	80
7.3.	TDP2 ABSENCE AND DNA DAMAGE REPAIR.....	81
7.4.	TDP2 AND SCAR23.....	83
<b>8.</b>	<b>REFERENCES .....</b>	<b>85</b>
<b>9.</b>	<b>APPENDIX.....</b>	<b>108</b>
9.1.	PUBLISHED WORK.....	108

## List of Abbreviations

5'AMP	5' Adenosine monophosphate
53BP1	Tumour protein P53 Binding protein 1
6-4PPs	6-4 pyrimidine-pyrimidone photoproducts
ADP	Adenosine diphosphate
ANOVA	Analysis of variance
AOA1	Ataxia with oculomotor apraxia type 1
AOA4	Ataxia with oculomotor apraxia type 4
AP	Apurinic or apyrimidinic site
APEX1	AP endo-deoxyribonuclease 1
APLF	Aprataxin and PNKP-like factor
APTX	Aprataxin
ART	Artemis
AT	Ataxia telangiectasia
AT-LD	AT-like disorder
ATM	Ataxia telangiectasia mutated
ATP	Adenosine triphosphate
ATR	ATM and Rad3 – related protein
ATRIP	ATR interacting protein
BCA	Bicinchoninic acid
BCR	Break cluster region
BER	Base excision repair
BHQ	Black hole quencher
BLAST	Basic local alignment search tool
BLM	Bloom Syndrome Helicase
BPDE	Benzo(a)pyrene diolepoxide
BRCA1	Breast cancer type 1 susceptibility protein
BRCA2	Breast cancer type 2 susceptibility protein
BRCT	BRCA1 carboxy-terminal

BSA	Bovine serum albumin
BTR	BLM-TopoIII $\alpha$ -RMI1-RMI2 complex
CDK1	Cyclin Dependent Kinase 1
cDNA	complementary DNA
CDS	Coding sequence
CETN2	Centrin 2
CIP	Alkaline phosphatase
CPDs	Cyclobutene-pyrimidine dimers
CPT	Camptothecin
CRISPR	Clustered regularly interspaced short palindromic repeat
crRNA	CRISPR RNA
CSA	Cockayne syndrome protein A
CSB	Cockayne syndrome protein B
CTD	C-terminal domain
CtIP	Retinoblastoma Binding Protein 8; C-terminal Binding Protein (CtBP) interaction protein
DAPI	4',6-diamidino-2-phenylindole
DDB1	Damage specific DNA binding protein 1
DDB2	Damage specific DNA binding protein 2
DDR	DNA damage response
DMEM	Dulbecco's modified essential medium
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
DNA2	DNA Replication Helicase/Nuclease 2
DNAPK	DNA dependent protein kinase
DNAPKcs	DNA dependent protein kinase catalytic subunit
DRB	5,6-Dichloro-1- $\beta$ -D-ribofuranosylbenzimidazole
dRP	5' deoxyribose phosphate
DSB	Double-strand break
DSBR	Double-strand break repair
dT	oligo deoxythymine

DTT	Dithiothreitol
EAPII	ETS associated protein 2
ECL	Electrochemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EEP	Exonuclease-endonuclease-phosphatase
EGFP	Enhanced GFP
ER	Endoplasmic reticulum
ERCC1	Excision repair cross-complementing group 1
ESR	Oestrogen receptor
ETC	Electron transport chain
ETS1	Erythroblastosis (ETS) proto-oncogene 1
EXO1	Exonuclease 1
FCS	Fetal calf serum
FHA	Forkhead- associated
FISH	Fluorescence in situ hybridisation
FITC	Fluorescein isothiocyanate
FMRP	Fragile X mental retardation protein
FRET	Fluorescence resonance energy transfer
GATK	Genome Analysis Tool Kit
GFP	Green Fluorescent Protein
GG-NER	Global genome nucleotide excision repair
gH2AX	phosphorylated Histone H2A variant X at serine 139
gRNA	guide RNA
H2AX	Histone H2A variant X
HER2	Human epidermal growth factor receptor 2
HIV	Human immunodeficiency virus
HIV IN	human immunodeficiency virus (HIV) integrase
HJ	Holliday junction
HR	Homologous recombination
HRR	Homologous recombination repair
hTERT	human telomerase reverse transcriptase

HTS	High-throughput screening
IF	Immunofluorescence
IgG	Immunoglobulin G
IP	Immunoprecipitation
IR	Ionising radiation
LIGIII	Ligase III
LIGIV	Ligase IV
MAD2L2	Mitotic Arrest Deficient 2 Like 2
MAMs	Mitochondrial-associated endoplasmic reticulum membranes
MCSZ	Microcephaly, seizures, and developmental delay
MDC1	Mediator of DNA damage checkpoint 1
MEFs	Mouse embryonic fibroblasts
MEM	Minimum essential medium
MFU	Mean fluorescence units
MLH1	MutL homologue 1
MLL	mixed lineage leukaemia
MMP1	Matrix metalloproteinase-1
MMR	Mismatch repair
MMS	Methyl methanesulphonate
MPG	3 methyladenine DNA glycosylase
MRE11	Meiotic Recombination 11 homologue
MRI	Mediator of retroviral infection
MRN	MRE11-RAD50-NBN complex
MS	Mass spectrometry
MSH2	MutS homologue 2
MSH6	MutS homologue 6
mtTOP1	Mitochondrial Topoisomerase I
MutL	Mutator L
MutS	Mutator S
MUTYH	mutY DNA glycosylase



NAD <sup>+</sup>	Nicotinamide adenine dinucleotide
NBN	Nibrin
NBS	Nijmegen Breakage Syndrome
NBS1	Nijmegen Breakage Syndrome 1
NEB	New England Biolabs
NER	Nucleotide excision repair
NF- $\kappa$ B	Necrosis-Factor kappa B
NHEJ	Non-homologous end joining
NLS	Nuclear localisation signal
NPPP	4-nitrophenyl phenylphosphonate
NRF1	Nuclear respiratory factor 1
NSCLC	Non-small-cell lung carcinoma
OH	Hydroxyl
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PALB2	Partner and Localizer of BRCA2
PAM	Protospacer adjacent motif
PAM	Proto-spacer adjacent motif
PAR	poly ADP-ribose
PARP1	poly (ADP-ribose) polymerase 1
PARP2	poly (ADP-ribose) polymerase 2
PAXX	Paralogue of XLF and XRCC4
PBM	PAR-binding motif
PBS	Phosphate-buffered saline
PBZ	PAR-binding zinc finger
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PDI	Protein disulphide isomerase
PG	Phosphoglycol
PI3K	Phosphoinositide 3-kinase
PIC	Proteinase inhibitor cocktail

PIKK	Phosphoinositide 3-kinase (PI3K)-related kinases
PLD	Phospholipase D
PMS2	Postmeiotic segregation increased 2
PMSF	Phenylmethanesulfonyl fluoride
PNKP	Polynucleotide 5'-kinase 3'-phosphatase
Pol $\beta$	Polymerase $\beta$
Pol $\mu$	Polymerase $\mu$
PR	Progesterone receptor
PTMs	Post translation modifications
RAD23B	UV excision repair protein RAD23 homolog B
RAD50	Radiation 50 DSB repair protein
RAD51	Radiation 51 Recombinase
RAG1	Recombination activating gene 1
RAG2	Recombination activating gene 2
RIDDLE	Radio-sensitivity, immunodeficiency, dysmorphic features and learning difficulties
RIF1	Replication Timing Regulatory Factor 1
RNA	Ribonucleic acid
RNAPII	RNA polymerase II
RNAse	Ribonuclease
RNAseH	Ribonuclease H
RNF168	Ring finger protein 168
RNF8	Ring finger protein 8
ROS	Reactive oxygen species
RPA	Replication protein A
RPE-1	Retinal pigmented epithelium cell line 1
SB	Sample buffer
SCAN1	Spinocerebellar ataxia with axonal neuropathy 1
SCAR23	Spinocerebellar ataxia, autosomal recessive 23
SDS	Sodium dodecyl sulphate
SDS-PAGE	SDS-Polyacrylamide Gel Electrophoresis

SDSA	Synthesis-dependent strand annealing
SEM	Standard error of the mean
SETX	Senataxin
sgRNA	single guide RNA
SHLD1	Shieldin Complex Subunit 1
SHLD2	Shieldin Complex Subunit 2
SHLD3	Shieldin Complex Subunit 3
SIM	SUMO-interacting motif
siRNA	small interfering Ribonucleic acid
SNP	Single nucleotide polymorphisms
SPO11	Sporulation 11, Initiator of Meiotic Double Stranded Breaks-like
SSB	Single-strand break
SSBR	Single-strand break repair
ssDNA	single-strand DNA
SUMO	Small ubiquitin-like modifier
t-AML	therapy-related Acute Myeloid leukaemia
TADs	Topological associated domains
TAE	Tris-acetate-EDTA
TALENs	Transcription activator-like effector nucleases
TBE	Tris-borate-EDTA
TBNC	Triple negative breast cancers/basal like tumours
TBST	Tris-buffered saline with Tween-20
TC-NER	Transcription-coupled nucleotide excision repair
TDG	Thymine DNA glycosylase
TDP	Tyrosyl DNA phosphodiesterase
TDP1	Tyrosyl-DNA phosphodiesterase 1
TDP2	5' Tyrosyl-DNA phosphodiesterase 2
TDP2 $\beta$	TDP2 isoform $\beta$
TDP2 $\gamma$	TDP2 isoform $\gamma$
TDRD3	Tudor domain containing protein 3

TE	Tris-EDTA
TEMED	Tetramethylethylenediamine
TFIIH	Transcription initiation factor IIH
TGF- $\beta$	Transforming growth factor-b
TOP1	Topoisomerase I
TOP1cc	Topoisomerase I-cleavage complex
TOP2	Topoisomerase II
TOP2A	Topoisomerase II $\alpha$
TOP2B	Topoisomerase II $\beta$
TOP2cc	Topoisomerase II-cleavage complex
TOP3	Topoisomerase III
TOP3A	Topoisomerase III a
TOP3B	Topoisomerase III $\beta$
TP53	Tumour Protein 53
tracrRNA	trans-activating CRISPR RNA
TSS	Transcription start site
TTRAP	Tumour necrosis factor (TNF) and TNF receptor-associated factors (TRAFs), TNF receptor associated protein
TTS	Transcription termination site
UBA	Ubiquitin-associated
UNG	Uracil DNA glycosylase
UV	Ultraviolet irradiation
UV-DDB	UV-DNA damage-binding protein
WB	Western blot
WCE	Whole-cell extract
WES	Whole exome sequencing
WRN	Werner syndrome helicase
WT	Wild-type
XLF	XRCC4-like factor
XP	Xeroderma pigmentosum
XPA	Xeroderma pigmentosum (XP) complementation group A

XPB	Xeroderma pigmentosum (XP) complementation group B
XPC	Xeroderma pigmentosum (XP) complementation group C
XPD	Xeroderma pigmentosum (XP) complementation group D
XPF	Xeroderma pigmentosum (XP) complementation group F
XPG	Xeroderma pigmentosum (XP) complementation group G
XRCC1	X-ray repair cross complementing 1
XRCC2	X-ray repair cross complementing 2
XRCC3	X-ray repair cross complementing 3
XRCC4	X-ray repair cross complementing 4
ZATT	Zinc finger protein associated with TDP2 and TOP2
ZNF	Zinc-finger nuclease

## Index of Figures

<b>Figure 1.1</b>	Endogenous damage lesions on cell per day, frequencies and mutational outcome	1
<b>Figure 1.2</b>	Overview of Nucleotide Excision Repair	3
<b>Figure 1.3</b>	Simplified overview of DNA damage response	8
<b>Figure 1.4</b>	Simplified overview of homologous recombination repair	10
<b>Figure 1.5</b>	Simplified overview of major components of NHEJ	12
<b>Figure 1.6</b>	Topoisomerase type IIA mechanism of action – To release torsional stress on DNA	17
<b>Figure 1.7</b>	TDP1 mediated TOP1cc removal	21
<b>Figure 1.8</b>	5' Tyrosyl DNA Phosphodiesterase 2 (TDP2)	23
<b>Figure 1.9</b>	TDP2 mediated TOP2cc removal	23
<b>Figure 1.10</b>	Overview of TOP2cc repair	24
<b>Figure 3.1</b>	<i>TDP2</i> locus and spCas9 mediated genome editing	39
<b>Figure 3.2</b>	Method for the generation of CRISPR-Cas9 gene edited cell lines	39
<b>Figure 3.3</b>	TDP2 levels in RPE-1 cells transfected with hCas9 and gRNA vectors	40
<b>Figure 3.4</b>	TDP2 protein levels in single clonal isolates of CRISPR-Cas9 gene edited cells.	40
<b>Figure 3.5</b>	CRISPR-Cas9 insertions and deletions at the site of DSB formation in <i>TDP2</i>	40
<b>Figure 3.6</b>	Gating strategy for $\gamma$ H2AX quantification by Olympus ScanR high content imaging	42
<b>Figure 3.7</b>	TDP2 absence impairs DSB repair rates independently of cell cycle	42
<b>Figure 3.8</b>	TDP2 absence impairs DSB repair rates following etoposide treatment	43
<b>Figure 3.9</b>	Loss of TDP2 leads to hypersensitivity to etoposide in clonogenic survival assays	43

<b>Figure 3.10</b>	Complementation of <i>TDP2</i> <sup>-/-</sup> Exon 1 RPE-1 cells with recombinant human TDP2	44
<b>Figure 3.11</b>	Hypersensitivity to etoposide is dependent on TDP2 phosphodiesterase activity	44
<b>Figure 3.12</b>	TDP2 absence leads to increased transcription-dependent chromosomal breakage and translocations	46
<b>Figure 4.1</b>	A putative 37 kDa isoform of TDP2	49
<b>Figure 4.2</b>	CRISPR-Cas9 Exon 4 <i>TDP2</i> <sup>-/-</sup> clones lack the 37 kDa TDP2 isoform	50
<b>Figure 4.3</b>	The 37 kDa isoform of TDP2 promotes resistance to topoisomerase 2 induced DSBs.	50
<b>Figure 4.4</b>	TDP2 isoform is involved in nuclear DSB repair following etoposide	50
<b>Figure 4.5</b>	Loss of TDP2 isoforms does not affect survival or DSB repair following irradiation	51
<b>Figure 4.6</b>	TDP2 isoform beta (TDP2 $\beta$ )	51
<b>Figure 4.7</b>	TDP2 isoform gamma (TDP2 $\gamma$ )	52
<b>Figure 4.8</b>	Mass spectrometry and sequencing analysis of TDP2 isoforms	52
<b>Figure 4.9</b>	TDP2 isoforms cellular localisation	53
<b>Figure 4.10</b>	TDP2 $\beta$ localises mostly to endoplasmic reticulum	54
<b>Figure 4.11</b>	Exogenous expression of TDP2 $\gamma$ -GFP can complement <i>TDP2</i> <sup>-/-</sup> sensitivity and DSB repair delay	54
<b>Figure 4.12</b>	TDP2 $\beta$ and TDP2 $\gamma$ isoforms can mediate viral infection	57
<b>Figure 5.1</b>	Overview of the 5'-Tyrosyl DNA phosphodiesterase (5'-TDP) assay	61
<b>Figure 5.2</b>	5'-TDP activity of TDP2 protein	63
<b>Figure 5.3</b>	5'-TDP activity in whole cell extracts	63
<b>Figure 5.4</b>	Rate of reaction of 5'-TDP activity in whole cell extracts	64
<b>Figure 5.5</b>	5' TDP activity is TDP2 dependent	64
<b>Figure 5.6</b>	3'-TDP assay	64
<b>Figure 5.7</b>	The 5'-TDP assay is versatile and sensitive	65
<b>Figure 5.8</b>	Measuring 5' TDP activity in the presence of a TDP2 inhibitor	66

<b>Figure 5.9</b>	Breast cancer and TDP2 expression levels as possible prognosis marker	67
<b>Figure 5.10</b>	Breast cancer molecular phenotypes and TDP2 expression levels in relapse free survival	67
<b>Figure 5.11</b>	TDP2 addiction and cancer relapse in Luminal A Breast cancers	67
<b>Figure 6.1</b>	Mutation c.425+1G>A in TDP2 in a family from the USA	73
<b>Figure 6.2</b>	SCAR23 fibroblasts lacks detectable TDP2 protein and 5'-TDP activity	74
<b>Figure 6.3</b>	SCAR23 fibroblasts displays normal 3'-TDP activity	74
<b>Figure 6.4</b>	Residual '5'-TDP' activity in 850BR fibroblasts is not the result of residual TDP2	74
<b>Figure 6.5</b>	SCAR23 fibroblasts exhibit slower DSB repair following TOP2-induced DNA damage	75
<b>Figure 6.6</b>	TDP2 exogenous expression complements the DSB repair defect in SCAR23 fibroblasts	75
<b>Figure 6.7</b>	SCAR23 fibroblasts are hypersensitive to TOP2-induced DNA damage	75
<b>Figure 6.8</b>	Intronic homozygous TDP2 c.636+3_636+6del mutation in Italian SCAR23 patients	76
<b>Figure 6.9</b>	The TDP2 mutation c.636+3_636+6del leads to loss of TDP2 protein and reduced DSB repair	76
<b>Figure 6.10</b>	TDP2 mutation c.636+3_636+6del leads to protein loss and defective repair	76



## Index of Tables

<b>Table 3.1</b>	Table of sequencing variants of RPE-1 and TDP2 <sup>-/-</sup> RPE-1 clone #3 cells following whole exome sequencing	40
<b>Table 4.1</b>	Software prediction of TDP2 $\beta$ cellular localisation	54
<b>Table 5.1</b>	TDP2 gene variation according to the Catalogue of Somatic Mutations in Cancer (COSMIC) database	66
<b>Table 6.1</b>	Summary of all reported SCAR23 patients including major clinical and molecular features	78

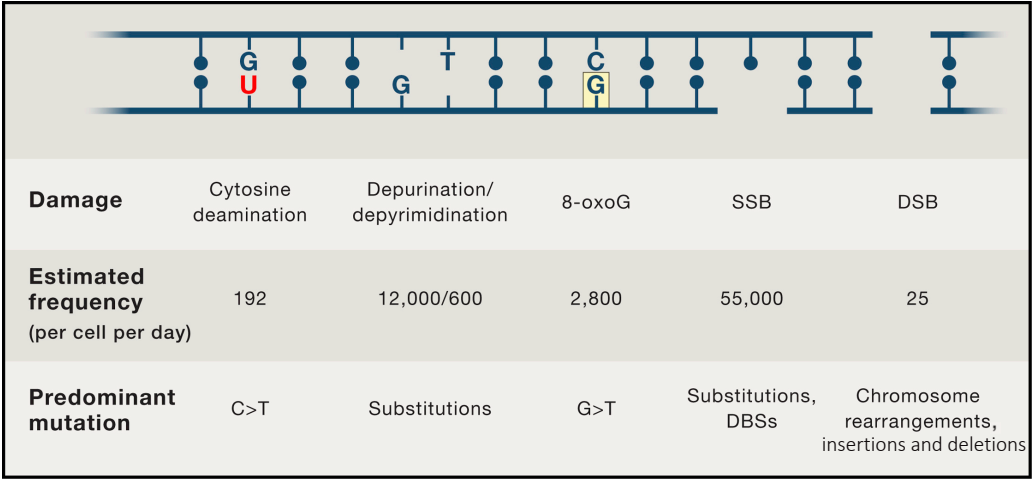
## 1. CHAPTER ONE – INTRODUCTION

### 1.1. DNA DAMAGE

Life on earth is based on deoxyribonucleic acid (DNA), which encodes all of the genetic information required to mediate processes such as transcription and replication. Nevertheless, DNA is intrinsically unstable (Lindahl, 1993; 2016). More than 70,000 DNA lesions can occur per cell per day (Figure 1.1). Most of these lesions will be present as single-strand breaks (SSB) and will arise from endogenous and exogenous sources (Lindahl, 1993; Tubbs and Nussenzweig, 2017)

DNA lesions are diverse and so are the mechanisms of DNA damage response and repair. The mismatch repair (MMR) pathway deals primarily with mis-incorporated nucleotides from replication errors; base excision repair (BER) removes modified nucleotides, such as 8-hydroxyguanine or xanthine, resulting from attack of DNA by reactive oxygen species (ROS) and by base deamination, respectively; nucleotide excision repair (NER) removes helix destabilizing lesions, such as ultraviolet (UV) induced cyclopurines and 6-4 pyrimidine photo-products, in addition to bulky lesions created by tobacco smoke (Jackson and Bartek, 2009; Curtin, 2012). Nevertheless, the major sources of genetic instability are DNA double strand breaks (DSBs), mainly caused by ionising radiation, unrepaired single-strand breaks that lead to replication fork collapse, abortive topoisomerase activity and chemotherapy agents (Deriano and Roth, 2013). Double strand breaks can be processed via homologous recombination (HR), which will take place during S/G2 phase of cell cycle and repairs DSBs by using a sister chromatid as a template and resulting in high fidelity repair (Deriano and Roth, 2013). Alternatively, non-homologous end joining (NHEJ), the second and most common mechanism by which DSBs are repaired, takes place in all cell cycle phases and will repair DSBs by processing and ligating DNA broken ends, potentially resulting in some loss of genetic material (Lieber, 2010).

DNA damage can result in genome instability, one of the emerging hallmarks of carcinogenesis and now considered to be an enabling characteristic of tumours (Negrini *et al.*, 2010; Hanahan and Weinberg, 2011). Genomic damage paves the way to the molecular changes required for the development of abnormal cells; as it give rises to



**Figure 1.1 – Endogenous damage lesions on cell per day, frequencies and mutational outcome** - DNA is under constant damage from endogenous sources. These can be frequent, such as single-strand breaks (SSB), or less frequent, such as double-strand breaks (DSB). In addition, base modifications such as cytosine deamination, oxidation of guanine (8-oxoG) and base loss are also significant sources of damage. **Figure adapted from Tubbs and Nussenzweig, 2017.**

mutations, chromosome translocations, chromosomal aberrations, genes amplifications, deletions, fusions and epigenetic modifications (Negri *et al.*, 2010; Curtin, 2012; Khanna, 2015). In addition, multiple DNA repair proteins are found to be defective in neurodegenerative, immunological and aging disorders highlighting the frequency of genome instability and importance of DNA repair mechanisms (Hoeijmakers, 2009; Woodbine *et al.*, 2014; McKinnon, 2017).

In this chapter, I will give a brief overview of all major DNA damage repair pathways, highlighting how breaks arise and the main players involved in their repair. Particularly, I will focus on DSBs generated by topoisomerase activity, exploring how these lesions arise and their mechanisms of repair. I will highlight the role of the enzyme 5'-Tyrosyl DNA phosphodiesterase 2 (TDP2) on the repair of DNA damage introduced by topoisomerase II and its significance to cell survival and human disease.

## **1.2. DNA REPAIR**

### **1.2.1. Mismatch repair**

The mismatch repair (MMR) pathway excises regions of mis-incorporated nucleotides that do not follow the defined Watson-Crick base pair or small loops of nucleotides that arise from replication slippage (Jiricny, 2013; Kunkel and Erie, 2015; Modrich, 2016). As a result, the absence of this pathway increases mutational load and causes microsatellite instability in highly repetitive regions of the genome that are prone to strand slippage (Jiricny, 2013).

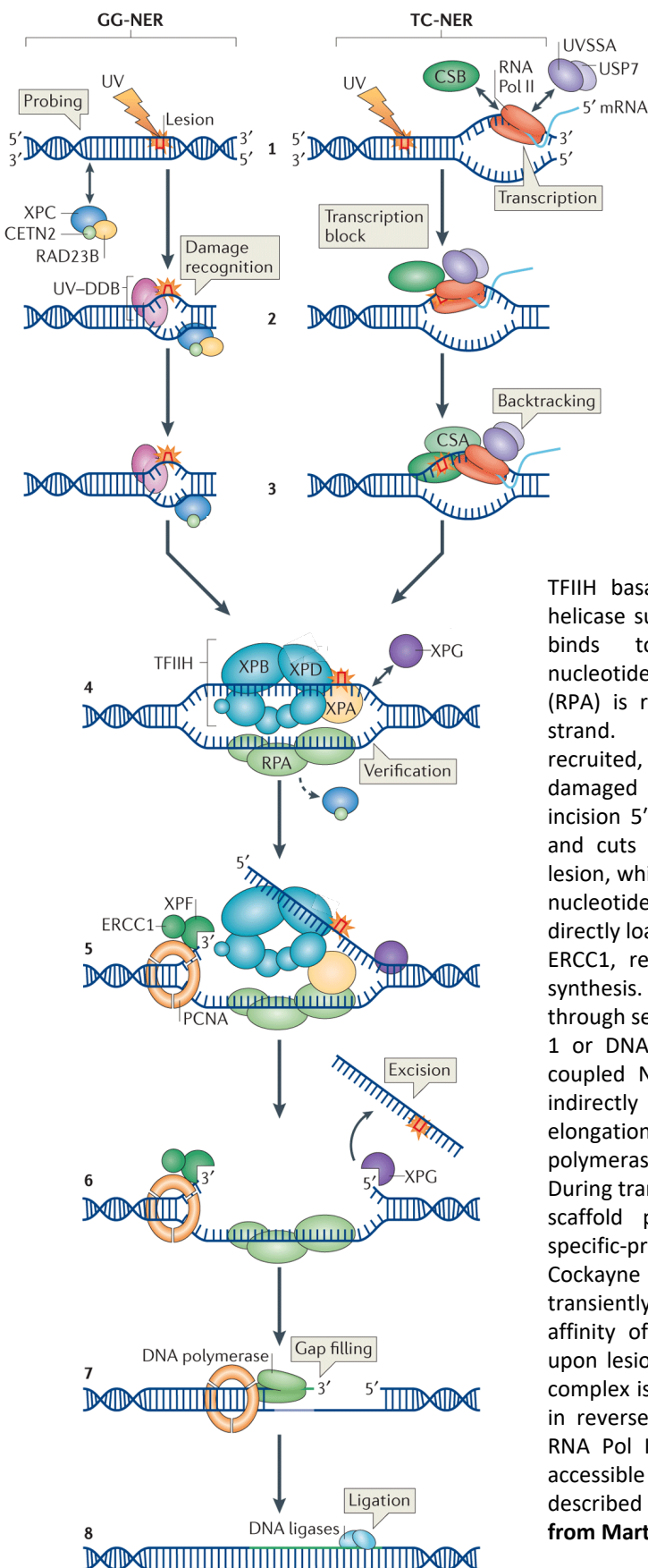
In humans, repair of these lesions is dependent on the type of mismatch present in DNA, but involves mismatch recognition by a Mutator S (MutS) heterodimer, interaction with Mutator L (MutL), strand resection by exonuclease I (EXO1), re-synthesis by polymerase delta and finally ligation (Kunkel and Erie, 2015). The subunits that comprise the ATP dependent DNA clamp MutS, are lesion dependent and thought to give this system its specificity; for example the MutS $\alpha$  heterodimer is formed by MutS homologue 2 (MSH2) and MutS homologue 6 (MSH6) and recognises base mismatches and small loops of 1 or 2 nucleotides (Jiricny, 2013). The MutL enzyme is also a heterodimer with multiple variants. MutL $\alpha$ , the best characterised dimer, is formed by

MutL homologue 1 (MLH1) and postmeiotic segregation increased 2 (PMS2), and mediates ATP dependent endonuclease activity on lesion containing strand (Kunkel and Erie, 2015). MutL $\alpha$  activity is controlled by proliferating cell nuclear antigen (PCNA), which is also a scaffold and interacts with the polymerase delta and ligase I to complete repair (Jiricny, 2013; Kunkel and Erie, 2015). MMR is not only active in post-replicative mismatches, but also in multiple cellular processes, such as meiotic recombination, oxidative damage, homologous recombination and non-homologous end joining (Jiricny, 2013), highlighting the multitude of DNA repair cross-talk and redundancy in the activity of some of these proteins.

### **1.2.2. Nucleotide excision repair**

DNA distorting lesions are repaired by the nucleotide excision repair (NER) pathway. These lesions can include cyclobutene-pyrimidine dimers (CPDs) and 6-4 pyrimidine-pyrimidone photoproducts (6-4PPs) and are caused by UV radiation; chemical adducts such as Benzo(a)pyrene diolepoxide (BPDE) or other polycyclic aromatic hydrocarbons; intra-strand crosslinks, caused by drugs such as cisplatin; and cyclopurines caused by ROS (Marteijn *et al.*, 2014; Spivak, 2015). Like MMR, repair of these lesions involves DNA damage recognition, excision, synthesis and ligation, however the mechanisms of which this occurs is less dependent on the lesions and more on the context in which it happens. The transcription-coupled NER (TC-NER) sub-pathway deals with lesions on template strands encountered by RNA polymerase II (RNAPII) whereas the global genome NER (GG-NER) sub-pathway repairs DNA distortions across the entire genome (Marteijn *et al.*, 2014).

In GG-NER, lesions that cause large helix distortions are recognized directly by Xeroderma pigmentosum (XP) complementation group C (XPC) in conjunction with RAD23B and Centrin 2 (CETN2). While, smaller helix distorting lesions, such as CPDs, are recognised first by the Damage specific DNA binding protein 1 (DDB1) and DDB2 which form the UV-DNA damage-binding protein (UV-DDB) complex prior XPC-RAD23B-CETN2 recruitment (Figure 1.2) (Marteijn *et al.*, 2014). Binding of XPC allows the recruitment of the transcription initiation factor IIH (TFIIH) complex that processes the lesion prior to repair. The DNA helicases, XP complementation group B and D (XPB and XPD), present



**Figure 1.2 – Overview of NER**

- Global genome nucleotide excision repair (GG-NER) sub-pathway, the damage sensor complex XPC/RAD23B/CETN2, probes DNA for helix-distorting lesions, which are recognized with the help of the UV-DDB complex. After damage recognition, the TFIIH complex is recruited to the lesion. In NER, the XPG endonuclease, either associated with TFIIH or separately, binds to the pre-precision NER complex. The helicase activity of TFIIH further opens the double helix around the lesion, and 5'–3' unwinding of the DNA by the

TFIIH basal transcription factor complex helicase subunit XPD and XPB. XPA which binds to the chemically altered nucleotides. The replication protein A (RPA) is recruited to the complimentary strand. XPF–ERCC1 heterodimer is recruited, which is directed to the damaged strand by RPA to create an incision 5' to the lesion. XPG is activated and cuts the damaged strand 3' to the lesion, which excises the lesion within a 30 nucleotide strand. The PCNA ring is directly loaded after the 5' incision by XPF–ERCC1, recruits DNA for gap-filling DNA synthesis. The NER reaction is completed through sealing the final nick by DNA ligase 1 or DNA ligase 3. In the transcription-coupled NER (TC-NER; right), damage is indirectly recognized during transcript elongation by the stalling of RNA polymerase II (RNA Pol II) at a lesion. During transcript elongation UV-stimulated scaffold protein A (UVSSA), ubiquitin-specific-processing protease 7 (USP7) and Cockayne syndrome protein CSB transiently interact with RNA Pol II. The affinity of CSB for RNA Pol II increases upon lesion recognition and the CSA–CSB complex is formed, which probably results in reverse translocation (backtracking) of RNA Pol II that renders the DNA lesion accessible for repair which takes place as described for GG-NER. **Figure adapted from Martejin *et al.* 2014.**

in the TFIIH complex, unwinds DNA around the break site, allowing XP complementation A (XPA) protein to bind and promote association of the replication protein (RPA) with the complementary single-stranded DNA (Figure 1.2) (Spivak, 2015). In addition, TFIIH recruits the ssDNA nuclease XPG, while XPA also recruits the 5' nuclease XPF, which is present in a complex with excision repair cross complementing protein 1 (ERCC1). XPF-ERCC1 initiates damage excision by introducing a ssDNA cut 5' to the DNA lesion, whereas XPG cuts 3' to DNA lesion, resulting in the excision of ~30 nucleotides. Finally, the gap is filled by PCNA, DNA Pol  $\delta$ , DNA Pol  $\epsilon$  or DNA Pol  $\kappa$  and ligated by DNA ligase I or DNA ligase III in complex with x-ray repair cross complementing 1 (XRCC1) (Figure 1.2) (Marteijn *et al.*, 2014; Hu *et al.*, 2017). Alternatively, lesions that are processed by TC-NER are recognised following transcription blockage and RNAPII backtracking (Marteijn *et al.*, 2014). Backtracking is thought to be mediated by the Cockayne syndrome protein B (CSB). CSB is in constant interaction with RNAPII and upon transcription blockage it recruits CSA and alters local DNA conformation promoting TFIIH binding, which results in repair proceeding as described above for GG-NER (Figure 1.2) (Spivak, 2015).

The importance of NER is emphasised by the multitude of disorders arising from mutations in its proteins, leading to cancer predisposition, neurodevelopmental disorders and premature aging (Marteijn *et al.*, 2014). For example, Xeroderma pigmentosum, the best understood of these disorders that arises from any of the XP proteins involved in repair, can lead to more than one-thousand fold increase in cancer incidence and hypersensitivity to UV radiation (Marteijn *et al.*, 2014).

### **1.2.3. Base excision repair & single-strand break repair**

Of the DNA lesions that arise in cells daily, single-strand breaks (SSBs) are the most common (Lindahl, 1993; Tubbs and Nussenzweig, 2017). These SSBs arise not only from direct disintegration of the sugar in DNA, but also indirectly as a result of processivity of damage bases and abasic sites by the base excision repair (BER) pathway and formation of abortive topoisomerase 1 (TOP1) lesions (Caldecott, 2008). The circumstances that lead to the formation of TOP1 lesions and its mechanism of repair will be introduced on section 1.3, while this section will introduce the BER processes that leads to SSBs and common repair mechanisms of single-strand break repair (SSBR).

BER recognises and removes altered bases from DNA (Lindahl, 2016). To do so, it employs DNA glycosylases that cleaves (N-glycosyl) base sugar bonds leaving behind an apurinic or apyrimidinic (AP) site (Lindahl, 1993). The enzyme AP endo-deoxyribonuclease 1 (APEX1) detects and cleaves the AP site generating a 3' hydroxyl (OH) and 5' deoxyribose phosphate (dRP) terminus that requires further processing (Hegde *et al.*, 2008). DNA polymerase  $\beta$  (Pol $\beta$ ), which possesses dRP lyase activity, further processes this lesion to remove the 5'-sugar phosphate residues and convert them to canonical 5'-phosphate termini. Alternatively, some of the many DNA glycosylases present in the cell possess intrinsic AP lyase activity (Hegde *et al.*, 2008). Therefore, following APEX1 and Pol $\beta$  activity or through direct glycosylase activity BER substrates are converted to SSBs.

Single-strand breaks are detected by poly (ADP-ribose) polymerase 1 (PARP1), which uses NAD<sup>+</sup> to modify itself and other proteins with poly-ADP-ribose (PAR) chains (Ray Chaudhuri and Nussenzweig, 2017). Although PARP1 functions are shared and overlap with PARP2 most of the cellular parylation is PARP1 dependent (Hanzlikova *et al.*, 2017). PAR works as a beacon to recruit many proteins to the break site through interaction by PAR-binding motifs. The scaffold protein XRCC1 is one such protein. XRCC1 promotes recruitment of the end processing enzymes Aprataxin (APTX), which removes 5'-AMP moieties from failed ligation reactions; polynucleotide 5'-kinase 3'-phosphatase (PNKP), which phosphorylates 5'-termini and dephosphorylates 3'-termini to re-establish canonical 3'-OH and 5'-P DNA ends; DNA polymerase  $\beta$ , which will promote gap filling at the SSB; and 3'-tyrosyl DNA phosphodiesterase 1 (TDP1) which will remove abortive TOP1 from 3' termini. XRCC1 also interacts with and stabilises DNA ligase III (LIGIII) protein, with which it is present in a constitutive complex (Caldecott, 2008). Following DNA end processing, gap filling is mediated by Pol $\beta$  and DNA ligation is conducted by LIGIII.

Although SSBR operates in all phases of the cell cycle, a recent study has identified increased parylation at unligated Okazaki fragments, suggesting that Okazaki fragment maturation in S phase is as major source of endogenous SSBs (Hanzlikova *et al.*, 2018). In addition, inhibitors of PARP1/2 lead to synthetic lethality in cells with mutations in the homologous recombination proteins Breast cancer type 1 susceptibility



protein (BRCA1) and BRCA2 (Lord and Ashworth, 2017; Ray Chaudhuri and Nussenzweig, 2017). Therefore, defects in SSBR can have a multitude of effects on cells, which is highlighted by neurodegenerative disorders and by the embryonic lethality of mice in which SSBR proteins are mutated or deleted (Caldecott, 2008; McKinnon, 2013; Hoch *et al.*, 2017; Yoon and Caldecott, 2018).

#### **1.2.4. DNA double-strand break (DSB) repair**

DSB repair can arise from multiple sources including ionizing radiation, reactive oxygen species, collapse of replication forks, transcription, and abortive TOP2 cleavage complexes (Goodarzi and Jeggo, 2013; Sallmyr and Tomkinson, 2018). DSBs are thought to be the most dangerous lesions in DNA, if left unrepaired they can lead to cell senescence or apoptosis (Jackson and Bartek, 2009). There are two main pathways governing DSB repair (DSBR); homologous recombination (HR) mediated repair (HRR), which is an error-free pathway, and non-homologous end joining (NHEJ), which can result in insertions or deletions. The choice between these pathways is dependent on the phase of cell cycle and is regulated by the DNA damage response and the presence of epigenetic marks (Her and Bunting, 2018).

##### **1.2.4.1. DNA damage response to DSBs**

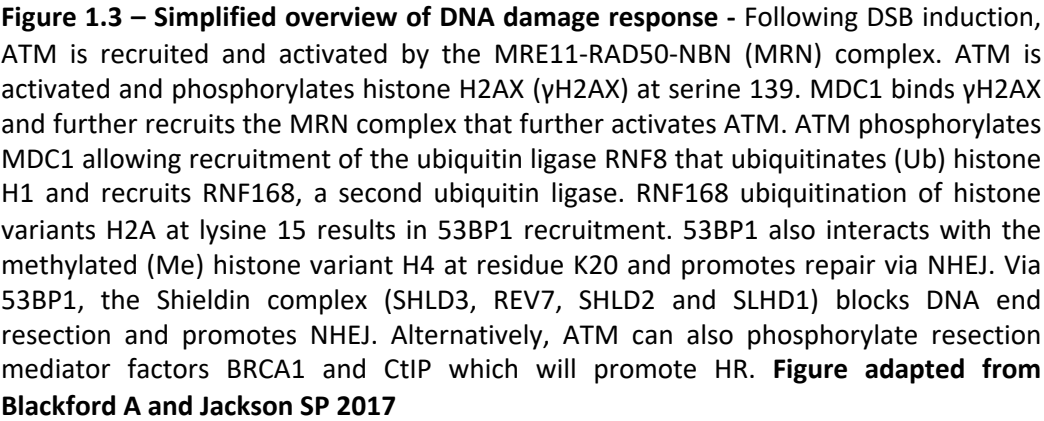
Following double-strand breaks, a cellular signalling cascade is triggered that leads to cell-cycle checkpoint activation, transcriptional regulation, and the modulation of chromatin structure (Jackson and Bartek, 2009). DNA damage signalling is driven by the kinases Ataxia Telangiectasia Mutated (ATM), Ataxia Telangiectasia and Rad3-Related protein (ATR) and the DNA dependent protein kinase (DNA-PK) (Blackford and Jackson, 2017). ATM, ATR and DNAPK are phosphoinositide 3-kinase (PI3K)-related kinases (PIKKs) and preferentially phosphorylate protein targets on serine or threonine residues followed by a glutamic acid (S/T-Q motif). These phosphorylation includes auto and trans-phosphorylation events and also the phosphorylation of hundreds of other protein targets in many cellular processes (Matsuoka *et al.*, 2007; Shiloh and Ziv, 2013; Blackford and Jackson, 2017). The regulation of the activity of these protein kinases is

achieved in part by their selective recruitment to DNA damage sites by protein co-factors. For example, KU70/80 heterodimer recruits DNA-PK catalytic subunit (DNA-PKcs); the protein complex formed by Meiotic recombination 11 homolog (MRE11), RAD50 protein and Nijmegen Breakage Syndrome 1 protein (NBS1 or NBN) (MRN complex), recruits ATM; and ATR Interacting Protein (ATRIP) binds Replication Protein A (RPA) and recruits ATR resulting in the DNA damage response (DDR) activation only in the presence of a DSB or RPA coated ssDNA (Singleton *et al.*, 1999; Zou and Elledge, 2003; Falck *et al.*, 2005).

Of the multiple post translation modifications (PTMs) that happen following DSB induction, the best studied is the phosphorylation of the tail of histone H2 variant X (H2AX) at serine 139 ( $\gamma$ H2AX) (Rogakou *et al.*, 1998; Blackford and Jackson, 2017).  $\gamma$ H2AX phosphorylation is propagated at a DSB to a megabase-region of chromatin flanking the break, resulting in the visualisation of damaged induced focus of  $\gamma$ H2AX (Rogakou *et al.*, 1999). This phosphorylation can be mediated by ATM and DNA-PK following DSB induction and by ATR following DNA replication fork arrest (Georgoulis *et al.*, 2017). A major function of  $\gamma$ H2AX is to recruit Mediator of DNA Damage Checkpoint 1 (MDC1), Tumour protein P53 Binding protein 1 (53BP1), the MRN complex, and Breast cancer type 1 susceptibility protein (BRCA1) to promote DDR signalling and repair (Georgoulis *et al.*, 2017). MDC1 interaction with both  $\gamma$ H2AX, via its BRCA1 carboxy-terminal (BRCT) domain, and with MRN complex, via the forkhead associated (FHA) domain in NBN, results in further activation of ATM, which in turn phosphorylates additional H2AX, promoting signal amplification (Stucki *et al.*, 2005; Spycher *et al.*, 2008). In addition, ATM also phosphorylates MDC1, which results in recruitment of Ring Finger protein 8 (RNF8), which in turn promotes sequential ubiquitination of the linker Histone H1 (Maidland *et al.*, 2007; Thorslund *et al.*, 2015). This ubiquitination further recruits RNF168, a ubiquitin ligase that promotes ubiquitylation of lysine (K) residue 13 and 15 in Histone H2A variants (Mattioli *et al.*, 2012). The scaffold protein 53BP1 binds H2A ubiquitylated at K15 and Histone H4 methylated at K20 and protects DNA ends against resection, promoting NHEJ and inhibiting BRCA1-regulated resection (Fradet-Turcotte *et al.*, 2013; Escribano-Díaz *et al.*, 2013). The mechanism by which 53BP1 inhibits resection was only recently found and still under investigation but it is clear that involves its interaction with Replication Timing Regulatory Factor 1 (RIF1) and the Shieldin

complex, which is formed by Mitotic Arrest Deficient 2 Like 2 (MAD2L2 or REV7) and Shieldin Complex Subunit 1-3 (SHLD1-3) (Escribano-Díaz *et al.*, 2013; Chapman *et al.*, 2013; Dev *et al.*, 2018; Noordermeer *et al.*, 2018; Gupta *et al.*, 2018; Ghezraoui *et al.*, 2018). The interaction of RIF with 53BP1 is dependent on phosphorylation by ATM of 53BP1 N-terminus, while the Shieldin complex interacts directly with RIF1 and promotes NHEJ by binding to ssDNA and blocking the activity of exonucleases (Silverman *et al.*, 2004; Escribano-Díaz *et al.*, 2013; Noordermeer *et al.*, 2018; Dev *et al.*, 2018). Conversely, DSB end resection is an essential committing step for HRR and is initiated by DNA incision mediated by the MRN complex, and the Retinoblastoma Binding Protein 8 (RBBP8 or CtIP) (Syed and Tainer, 2018). CtIP itself is controlled by interaction with BRCA1 protein and Cyclin Dependent Kinase 1 (CDK1) phosphorylation in S/G2 phase of the cell cycle to promote its activity, making this process more active during these stages (Huertas and Jackson, 2009; Cruz-García *et al.*, 2014). A simplified overview of this process is present in Figure 1.3. The mechanism of pathway still not fully understood, however recently the methylation on H4 K20, required for 53BP1 foci formation after damage, was shown to be cell cycle dependent (Saredi *et al.*, 2016; Pellegrino *et al.*, 2017). Leading to a model that resection-dependent repair is promoted only in newly replicated sequences until H4 K20 methylation is re-established in late G2 phase, placing DNA replication as one of the key factors in repair pathway choice (Saredi *et al.*, 2016; Pellegrino *et al.*, 2017; Murray and Carr, 2018).

The DNA damage response is essential to human health, as highlighted by patients with ataxia telangiectasia (AT), AT-like disorder (AT-LD); Nijmegen breakage syndrome (NBS); radio-sensitivity, immunodeficiency, dysmorphic features and learning difficulties (RIDDLE) syndrome; and Seckel syndrome. A-T is caused by mutations in ATM protein and is characterised by cerebellar neurodegeneration and ataxia, telangiectasia, apraxia of eye movements, immunodeficiency, cancer predisposition, sensitivity to ionising radiation, growth retardation and premature aging (Perlman *et al.*, 2012). Patients with mutations in MRE11 present with AT-LD, and although they have a similar phenotype to AT, they lack telangiectasia and immunodeficiency and display slower neurodegeneration (Taylor *et al.*, 2004). NBS patients also share clinical and cellular features with A-T including immunodeficiency, sensitivity to ionizing irradiation, cancer



**Figure 1.3 – Simplified overview of DNA damage response** - Following DSB induction, ATM is recruited and activated by the MRE11-RAD50-NBN (MRN) complex. ATM is activated and phosphorylates histone H2AX ( $\gamma$ H2AX) at serine 139. MDC1 binds  $\gamma$ H2AX and further recruits the MRN complex that further activates ATM. ATM phosphorylates MDC1 allowing recruitment of the ubiquitin ligase RNF8 that ubiquitinates (Ub) histone H1 and recruits RNF168, a second ubiquitin ligase. RNF168 ubiquitination of histone variants H2A at lysine 15 results in 53BP1 recruitment. 53BP1 also interacts with the methylated (Me) histone variant H4 at residue K20 and promotes repair via NHEJ. Via 53BP1, the Shieldin complex (SHLD3, REV7, SHLD2 and SLHD1) blocks DNA end resection and promotes NHEJ. Alternatively, ATM can also phosphorylate resection mediator factors BRCA1 and CtIP which will promote HR. **Figure adapted from Blackford A and Jackson SP 2017**

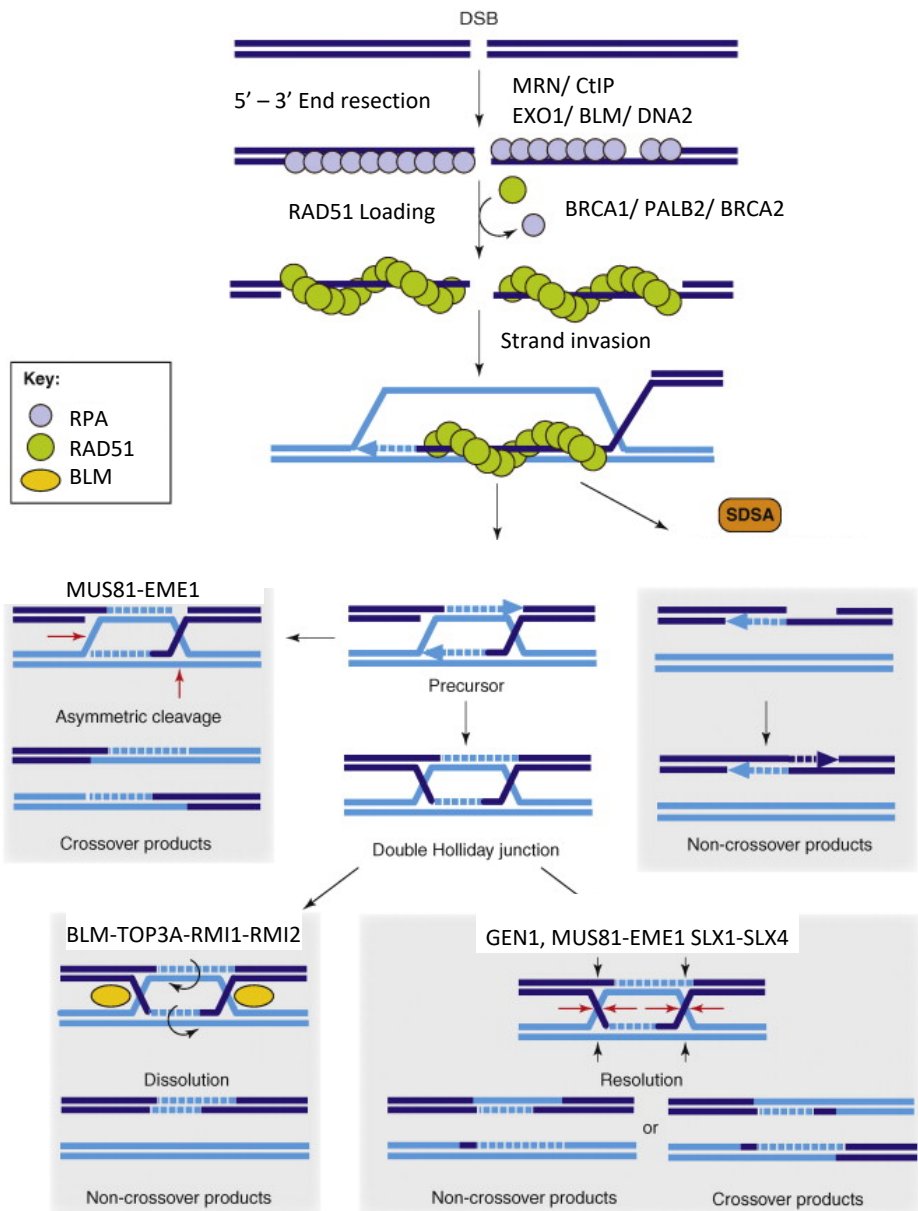
predisposition and growth retardation, and also present with microcephaly and learning difficulties which are usually associated with DNA ligase IV (LIGIV) syndrome (Demuth and Digweed, 2007; Reynolds and Stewart, 2013). In RIDDLE syndrome patients present with lack of immunoglobulin G (IgG) and have learning difficulties, impaired balance and facial dysmorphism, with patient-derived cells displaying impaired 53BP1 recruitment and increased sensitivity to irradiation (Stewart *et al.*, 2007; 2009). Seckel syndrome is caused by mutations in ATR and is associated with growth retardation, skeletal abnormalities, mental retardation and dysmorphic facial features. At the molecular level, Seckel cells possess impaired DDR signalling and decreased cell survival following exposure to UV radiation (O'Driscoll *et al.*, 2003; Kerzendorfer and O'Driscoll, 2009).

#### **1.2.4.2. Homologous recombination**

The homologous recombination repair (HRR) pathway operates during DNA replication and in G2 phase of the cell cycle and meiosis (Wright *et al.*, 2018). In addition, HRR is also implicated in the repair of specific types of DSBs, such as one-ended DSBs arising during DNA replication and those introduced by high linear energy transfer carbon ions, which introduce more complex breaks and requires more extensive end resection (Shibata *et al.*, 2011). MRE11 has endonuclease activity and 3' to 5' exonuclease activity, which are stimulated and controlled by interactions with NBN and CtIP (Deshpande *et al.*, 2016; Paull, 2018). The endonuclease and exonuclease activities of MRN/CtIP create ssDNA tails at DNA breaks, allowing for more extensive resection by the 5' to 3' activity of Exonuclease 1 (EXO1) and a protein complex comprised of Bloom Syndrome Helicase (BLM) and DNA Replication Helicase/Nuclease 2 (DNA2) (Symington and Gautier, 2011). The ssDNA exposed by end resection is bound by RPA which protects DNA and helps to limit resection (Bhat and Cortez, 2018). RPA is then replaced with RAD51 by BRCA2 protein, which is recruited by BRCA1 protein and the bridging factor Partner and Localizer of BRCA2 (PALB2) (Prakash *et al.*, 2015). RAD51 stability is regulated by the RAD51-interacting proteins and paralogues XRCC2, XRCC3, RAD51B, RAD51C, RAD51D and BRCA2 (Wright *et al.*, 2018). RAD51 forms filamentous structures around which the ssDNA is wrapped in order to promote homology search and to initiate strand invasion into the complementary sister chromatid (Figure 1.4)

(Wright *et al.*, 2018). HRR is then completed by extended DNA synthesis as part of either synthesis dependent strand annealing (SDSA) or to facilitate additional heteroduplex and Holliday junction (HJ) formation. HJ's are then resolved in a crossover or non-crossover formation, in the former case resulting in sister chromatid exchange. Holliday junctions can be processed by one of a number of protein complexes including the BTR complex comprised of BLM, Topoisomerase III  $\alpha$  (TOP3A), RMI1, and RMI2; a complex comprised of the SLX1, SLX4, EME1, and MUS81 endonucleases (Wyatt *et al.*, 2017); or the GEN1 Holliday Junction resolvase (Sarbajna and West, 2014). The BTR complex uses BLM helicase activity to migrate the HJ and TOP3A to cut and decatenate DNA strands resulting in HJ dissolution and a non-crossover event (Matos and West, 2014). While the GEN1 resolvase introduces symmetric nick across the junction that can be rejoined by DNA ligases (Wyatt and West, 2014). Endonucleases SLX1-SLX4-MUS81-EME1 cooperatively resolves HJ by asymmetrically nicking resulting in cross-over and non-crossovers (Figure 1.4) (Matos and West, 2014).

The deletion of HRR factors, including RAD51 (Lim and Hasty, 1996), BRCA1 and BRCA2 (Moynahan, 2002) are embryonic lethal in mice, highlighting the importance of this repair pathway during cell proliferation. In addition, heterozygous germline mutations in BRCA1 and BRCA2 proteins lead to an increased risk for breast and ovarian cancers in females and prostate and pancreatic cancers in male individuals (Prakash *et al.*, 2015). During tumour development, loss of heterozygosity results in the cancer cells becoming HRR defective (Lord and Ashworth, 2017). The BRCA defective tumours have increased genome instability and increased sensitivity to types of chemotherapy that induce DNA damage processed during DNA replication or in G2 by HRR, including agents that induce single- and double-strand breaks, DNA crosslinks, and/or replication fork stress such as topoisomerase 'poisons', cisplatin and gemcitabine (Konstantinopoulos *et al.*, 2015). In addition, HRR-defective tumours are hypersensitive to the inhibition of PARP1/2; a phenomenon denoted synthetic lethality (Lord *et al.*, 2015). Synthetic lethality arises, for example, if both of two independent but functionally overlapping pathways are lost. In BRCA cells, HRR is defective and inhibition of PARP1/2 leads to cell death due to its involvement in alternative repair pathways, such as SSBR and alternative NHEJ (discussed at Section 1.2.4.3). The efficacy of PARP inhibitors is enhanced by the inhibitor's mechanism of action, which prevents PARP1 auto-



**Figure 1.4 – Simplified overview of homologous recombination repair** - Following DSB, DNA end resection is mediated by MRN, CtIP, EXO1, BLM and DNA2. 5'–3' resection of the broken ends creates ssDNA tails that are coated by RPA. RPA is replaced by Rad51 to form the nucleoprotein filament by BRCA2, which is recruited to the break via BRCA1-PALB2 interaction. RAD51 will promote strand invasion and recombination. Limited DNA synthesis occurs and the newly synthesised bases can re-anneal to original strand resulting in synthesis dependent strand annealing (SDSA) or result in more than one invasion that will result in a double Holliday junction (dHJ). Processing of this precursor by MUS81-EME1, SLX1-SLX4 generates crossover products. Dissolution of the dHJ via BLM–TopoIIIα-RMI1-RMI2 gives rise to non-crossover products, whereas resolution GEN1 can lead to either crossover or non-crossover products. **Figure adapted from Mimitou E and Symington L 2009.**

parylation, resulting in it becoming trapped on DNA structures requiring HR for repair (Lord and Ashworth, 2017). PARP inhibitors are now the first DNA repair targeted drugs to be approved for clinical use, for relapsed ovarian and breast cancers, and are able to significantly improve patient outcomes (Lord and Ashworth, 2017)

#### **1.2.4.3. Non-homologous end-joining**

The non-homologous DNA end-joining (NHEJ) repair pathway operates throughout the cell cycle, except for mitosis, and still mediates fast repair of most DSBs, even in G2 phase when HRR is present (Beucher *et al.*, 2009; Her and Bunting, 2018). One remarkable characteristic of NHEJ is its ability to ligate various types of DNA ends, but with the potential risk of incurring loss/gain of genetic information at the break and chromosomal translocations (Lieber, 2010).

In NHEJ the break is recognised by KU; a heterodimer comprised of the two subunits KU70 and KU80 (KU70/80), which binds DNA ends with strong affinity (Downs and Jackson, 2004). The C-terminus of KU80 recruits the DNA dependent protein kinase catalytic subunit (DNA-PKcs) and together these enzymes form the DNA-PK holoenzyme (Sibanda *et al.*, 2017). DNA-PK is auto-phosphorylated on the residues S2056 and T2609 prior its activation (Chen *et al.*, 2005). Once active, DNA-PK will phosphorylate many downstream targets and participate in the DNA damage response (discussed in Section 1.2.4.1). In NHEJ, the best known substrates of DNA-PK are DNA-PKcs itself and the endonuclease Artemis (ART), which in conjunction with MRE11 and CtIP promote end processing prior to DNA ligation (Riballo *et al.*, 2004; Goodarzi *et al.*, 2006; Biehs *et al.*, 2017). Additional DNA end processing factors involved in processing of specific types of DNA end are PNKP, Aprataxin, as previously discussed in Section 1.2.3, and the TDP1 enzyme that facilitates repair by removing 3'-phosphoglycolate (3'PG) residues present on 3' overhangs (Li *et al.*, 2017; Kawale *et al.*, 2018). Finally, Tyrosyl DNA Phosphodiesterase 2 (TDP2) can remove covalently linked TOP2 protein from 5'-termini at DSBs arising from abortive TOP2 activity, thereby allowing direct DNA ligation of the break (Ledesma *et al.*, 2009; Zagnoli-Vieira and Caldecott, 2017).

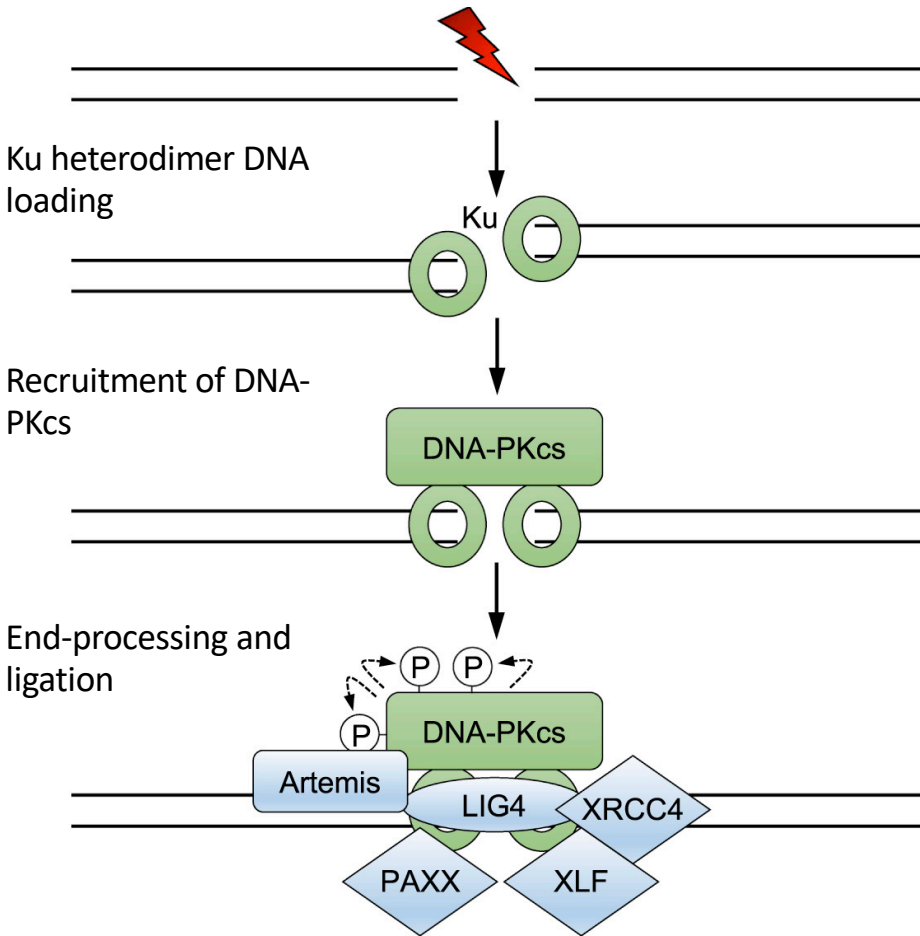
Filling of DNA ends is facilitated by DNA polymerases lambda (Pol $\lambda$ ) and/or mu (Pol $\mu$ ) and DNA ligation is conducted by DNA ligase IV (LIGIV), which is in a complex with



XRCC4 (Grawunder *et al.*, 1997; Lieber, 2010). XRCC4 and the XRCC4-like proteins XLF and PAXX help synapse the broken ends to keep them close allowing processing and ligation (Figure 1.5) (Brouwer *et al.*, 2016; Chang *et al.*, 2017). The KU70/80 heterodimer mediates many of the protein-protein interactions with DNA end processing enzymes and DNA polymerases to facilitate NHEJ (Mahajan *et al.*, 2002; Costantini *et al.*, 2007; Yano *et al.*, 2008). For example, the N-terminal von Willebrand-like domain in KU80 was recently shown to contain a hydrophobic pocket that mediates binding of Ku80 to various proteins including Aprataxin and PNKP like factor (APLF), Werner syndrome protein (WRN) and Mediator of Retroviral Infection (MRI) (Grundy *et al.*, 2013; 2016). The WRN protein suppresses end resection and thereby promotes NHEJ, and the 3'-5' exonuclease activity of WRN is stimulated by interaction with KU80 (Shamanna *et al.*, 2016; Grundy *et al.*, 2016). MRI was shown to be embryonic lethal with XLF and is thought to also participate in break synapse and protein-protein interactions (Hung *et al.*, 2018).

The importance of NHEJ is highlighted by the identification of individuals in which core NHEJ proteins such as DNAPKcs, LIGIV, Artemis and XLF are mutated and which exhibit radiosensitivity, severe combined immunodeficiency, and developmental defects such as microcephaly (Woodbine *et al.*, 2014). The immunodeficiency arises because of the requirement for NHEJ for rejoining of the variable (V), diversity (D) and joining (J) gene segments during the diversification of immunoglobulins in B and T lymphocytes (Alt *et al.*, 2013). This process, known as V(D)J recombination, is initiated by the sequence specific Recombination activating gene 1 (RAG1) and RAG2 recombinases that cleave DNA at their preferred sequence to promote hairpin-containing and blunt ended DSBs (Lieber, 2010). The hairpin structures are subsequently cleaved by Artemis, and further processed during NHEJ to introduce mutations prior to their ligation (Alt *et al.*, 2013). Therefore, NHEJ is important for antibody diversification not only by facilitating immunoglobulin gene rearrangements but also by promoting mutations during these rearrangements. In addition, NHEJ is also required in the development or maintenance of terminally differentiated neurons, as indicated by the neurological pathologies present in patients with mutations in XLF, XRCC4, DNA-PKcs and LIGIV (O'Driscoll *et al.*, 2001; Buck *et al.*, 2006; Woodbine *et al.*, 2013; Guo *et al.*, 2015). Loss of NHEJ is thought to lead to neurological defects due to DSB accumulation

## DNA damage induction



**Figure 1.5 – Simplified overview of major components of NHEJ** - Following DSB induction, Ku70/80 heterodimer is recruited to DSB ends. Ku heterodimer loading promotes the recruitment of DNAPK-cs, its activation and auto-phosphorylation. The end processing enzyme Artemis is also recruited and phosphorylated by DNA-PK. Scaffold protein XRCC4 is recruited together with PAXX and XLF and aid break synapse. In addition XRCC4 brings Ligase IV which re-ligates broken ends following processing. **Figure adapted from Blackford A and Jackson SP 2017**

during neurological development that results in the loss of cell mass (Woodbine *et al.*, 2014). Furthermore, transcription of long neuronal genes generates transcriptional stress that leads to DSB formation and requires NHEJ for its repair (Wei *et al.*, 2016; Alt and Schwer, 2018).

#### **1.2.4.4. Alternative NHEJ**

When NHEJ is deficient, the activity of an alternative NHEJ (altNHEJ) pathway can result in increased resection and ligation between regions with 2 to 20 base-pairs of homology (Chang *et al.*, 2017). This repair pathway requires the activity of PARP1 and the MRN complex rather than KU heterodimer, and in addition the activity of DNA polymerase theta (Pol $\theta$ ) (Sallmyr and Tomkinson, 2018). In addition, the endonuclease XPF-ERCC1 is required to remove single-stranded DNA flaps after the annealing of homologous sequences within single-stranded tails, and XRCC1-DNA ligase III or DNA ligase I for DNA ligation (Sfeir and Symington, 2015). It is currently unclear to what extent altNHEJ is used in a physiologically relevant context. However, Pol $\theta$  does appear to function in an alternative pathway to KU in the repair of DSBs in mouse embryonic stem cells, and this process is indeed mutagenic; introducing large deletions and tandem duplications (Schimmel *et al.*, 2017)

### **1.3. TOPOISOMERASES AS SOURCES OF GENOMIC INSTABILITY**

From the first observations from Watson and Crick that DNA has a helical structure, it was evident that it has to be melted to allow polymerases to bind and unwind DNA allowing polymerase progression (Watson and Crick, 1953). However, DNA is wrapped around histones which in turn forms chromatin, which itself is highly organised in the 3D nucleus space via topological associated domains (TADs) (Yu and Ren, 2017). Therefore, nuclear organisation, coupled with the double helical structure of DNA, creates a topological problem for metabolic processes in the cell such as transcription and replication. These processes result in the accumulation of negative or positive supercoils in DNA that need to be removed (Wang, 1996). The enzymes that

are responsible for the removal of supercoiling and resolution of topological knots are DNA topoisomerases (Wang and Liu, 1979).

Topoisomerase activity was first identified in the *E.coli* protein  $\omega$ , which was shown to resolve negatively supercoiled DNA. This protein was later found to be topoisomerase I (Wang, 1971; Wang and Liu, 1979). From its discovery it was clear that topoisomerase I generated a transient single-strand break during which the enzyme becomes covalently linked to DNA prior to release of topological stress (Wang and Liu, 1979). Similar observations with T4 DNA topoisomerase, which resolves positive and negative supercoils, revealed that this enzyme generates a transient DNA double-strand break (Liu *et al.*, 1980). Topoisomerases that create a transient DNA single-strand break are classified as type I topoisomerases and those that create a DNA double-strand break are classed as type II topoisomerases (Pommier *et al.*, 2016). These classes are further divided into IA, IB and IIA and IIB according to the mechanism by which they release torsional stress. All topoisomerases have a nucleophilic tyrosine residue that facilitates transient cleavage in DNA (Vos *et al.*, 2011). During the cleavage reaction the topoisomerase catalytic tyrosine becomes covalently attached to either the 3' or 5'-termini of DNA through a phosphodiester bond, forming a transient cleavage complex that is normally resolved by topoisomerase-mediated ligation at the end of the enzymes catalytic cycle. The re-ligation of the of DNA breaks could become abortive if transcription or replication machinery collides with the covalently linked topoisomerase cleavage complex creating a DNA lesion linked bulky lesion that requires repair (Vaz *et al.*, 2017).

Taking in account its mechanism of action is evident that topoisomerases are a source of endogenous DNA damage. Here I, will give an overview of each human topoisomerase, how cells require their activity, and deal with such covalently trapped topoisomerase lesions.

### **1.3.1. Topoisomerase type IA**

Type IA topoisomerases in humans are represented by Topoisomerase III $\alpha$  (TOP3A) and III $\beta$  (TOP3B). Type IA enzymes, release torsional stress through a strand passage mechanism, in which a single-strand is broken via a 5'phosphotyrosyl bond

allows a second single-strand to pass through the break and release torsional stress (Chen *et al.*, 2013). These enzymes are thought to relax substrate in steps of one (Champoux, 2001). Both TOP3A and TOP3B act on single-strand supercoils (bubbles) and while TOP3A acts on DNA, TOP3B has been shown to release torsional stress mainly in RNA (Pommier *et al.*, 2016).

TOP3A has been shown to localise to the nucleus and mitochondria (Wang *et al.*, 2002). Its deletion leads to embryonic lethality in mice (Li and Wang, 1998). While in humans it was recently shown that heterozygous amino acid substitutions in the TOP3A catalytic domain lead to mitochondrial DNA segregation failure and symptoms including cerebellar ataxia, cardiac arrhythmias and myopathy (Nicholls *et al.*, 2018). The TOP3A nuclear role on the other hand, has been linked to resolution of Holliday junctions (HJ's), described in section 1.2.4.1, in addition to hemicatenanes and catenates during chromosome segregation that appear as anaphase bridges (Chan *et al.*, 2007).

TOP3B has been found to be mostly cytosolic and to be in a complex with Fragile X mental retardation protein (FMRP) and Tudor domain containing protein 3 (TDRD3) (Xu *et al.*, 2013; Stoll *et al.*, 2013). Deletions or mutations in TOP3B have been linked to mental retardation and schizophrenia in human patients, but incur no overt phenotype in mice other than shortened life span (Stoll *et al.*, 2013; Ahmad *et al.*, 2017). TOP3B was shown to bind RNA directly via its arginine/glycine box (RGG) on the C-terminus and to be present in polyribosomes, indicating that it might remove RNA torsional stress during translation (Stoll *et al.*, 2013). In addition, TOP3B was shown to localise to stress granules after oxidative stress, in which it is thought to mediate mRNA disentanglement (Xu *et al.*, 2013; Ahmad *et al.*, 2017).

So far it is not clear if, similar to TOP2, TOP3A or TOP3B can form abortive cleavage complexes in which the topoisomerase is covalently linked to the 5'-terminus of the break. If so, the abortive DNA breaks created by these enzymes might require processing by TDP2, similar to those created by TOP2 (discussed in section 1.4.2).

### **1.3.2. Topoisomerase type IB**

In humans, type IB topoisomerases are represented by Topoisomerase I (TOP1) and its mitochondrial isoform (mtTOP1) (Pommier *et al.*, 2016). Unlike type IA

topoisomerases, type IB topoisomerases, do not have preference for a single-stranded substrate and the cleavage complexes they form involve linkage of the catalytic tyrosine to 3'-phosphate termini (Champoux, 2001). TOP1 activity involves nicking of a DNA strand and controlled rotation of the intact strand to remove torsional stress, followed by topoisomerase mediated nick ligation (Koster *et al.*, 2005). Deletion of TOP1 is embryonic lethal, while the mtTOP1 gene, although important for resolution of negative supercoiling of mtDNA, is not essential in mice and its function is still under investigation (Morham *et al.*, 1996; Pommier *et al.*, 2016).

The function of TOP1 is mainly linked to the relaxation of positive and negative supercoils that arise during transcription and replication (Kim and Jinks-Robertson, 2017). TOP1 was found to interact with the C-terminal domain (CTD) of RNA polymerase II (RNAPII), and to control the level of negative supercoiling at transcription start sites (TSS) (Baranello *et al.*, 2016). TOP1 is highly active both ahead and behind RNAPII (Baranello *et al.*, 2016). The expression of long genes in neurons (>100kb) was also shown to be dependent on TOP1 and Topoisomerase 2 $\beta$  functions, most likely to remove positive supercoiling ahead of RNAPII (King *et al.*, 2013). In addition, negative supercoiling behind elongating RNAPII can generate R-loops and other RNA:DNA hybrid structures, which are also reduced/prevented by TOP1 activity (Kim and Jinks-Robertson, 2017). In the absence or inhibition of TOP1, R-loops accumulate in cells and can result in the formation of DSBs (Tuduri *et al.*, 2009). These structures are removed by the helicase Senataxin or Ribonuclease H (RNaseH) (Santos-Pereira and Aguilera, 2015). Consistent with this, RNaseH overexpression decreases genetic instability in TOP1-depleted cells (Tuduri *et al.*, 2009).

The transient nick created by TOP1 is normally resealed by the topoisomerase at the end of each catalytic cycle (Champoux, 2001). However, on occasion this does not happen, resulting in TOP1 remaining covalently attached to the 3'-terminus of an 'abortive' cleavage complex (TOP1cc) (Kim and Jinks-Robertson, 2017). An accumulation of these lesions can then occur if genes involved in the repair of such abortive TOP1cc's, such as *TDP1* or *XRCC1*, are mutated (Katyal *et al.*, 2014). The mechanism by which TOP1cc are repaired will be discussed in section 1.4.1. The abortive activity of TOP1 can also lead to DSB formation, because the collision of a TOP1cc with the DNA replication machinery can lead to replication fork collapse and the formation of a single-ended DSB

requiring HRR for repair (Chanut *et al.*, 2016). DSBs can also arise during TOP1 activity independently of DNA replication, if abortive TOP1cc are closely located on opposite strands of the double helix (Pommier *et al.*, 2006)

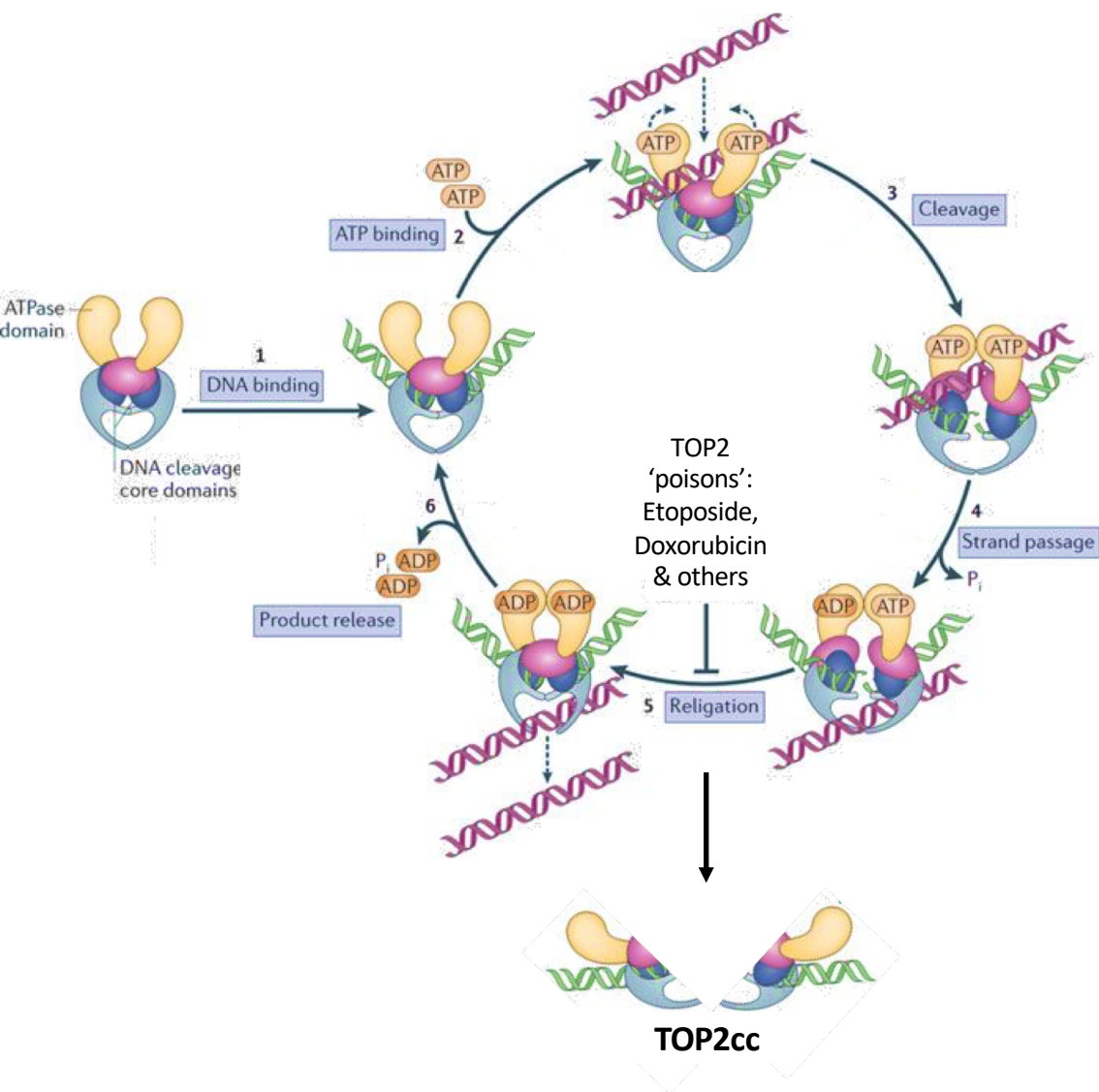
Importantly, the formation of abortive TOP1cc can be induced in cells in the laboratory and in the clinic by incubation with the drug Camptothecin (CPT) and its derivatives (Pommier *et al.*, 2010). CPT acts as a topoisomerase 'poison', trapping the topoisomerase in a TOP1cc and inhibiting the relegation step (Hsiang *et al.*, 1985; Pommier, 2009). Most of the cytotoxicity induced by CPT is in proliferating cells in S phase, resulting from replication fork collapse at abortive TOP1cc (Desai *et al.*, 2001). The cytotoxicity induced in rapidly proliferating cells by this mechanism renders CPT useful in the clinic for treating colorectal, ovarian, and small-cell lung cancers (Pommier, 2006).

### **1.3.3. Topoisomerase type IIA**

Topoisomerase type IIA (TOP2), unlike type IA and IB, catalyses ATP-dependent transport of an intact DNA helix through an enzyme bridged DSB to remove catenates, supercoils, and to facilitate chromosome condensation as well as chromosome segregation (Wang, 2002). Type IIA topoisomerases are homodimeric, with each subunit of the enzyme co-ordinately nicking DNA to create a cleavage complex containing a DSB with a 4-bp overhang in which the catalytic tyrosine is covalently linked to the 5'-terminus of the break (Nitiss, 2009a). Each catalytic cycle is characterised by capture of a DNA segment by the enzyme followed by DNA cleavage to form a TOP2cc and ATP binding. ATP binding and hydrolysis results in conformational changes that close the N-terminal gate, promoting passage of an intact DNA duplex through the enzyme C-terminal gate and followed by topoisomerase-mediated resealing of the break. Hydrolysis of a second ATP completes the reaction by promoting conformational changes that results in the opening of the C-terminal gate to release both intact strands (Figure 1.6) (Berger *et al.*, 1996; Wendorff *et al.*, 2012)

In humans, type IIA enzymes are represented by Topoisomerase II $\alpha$  (TOP2A) and Topoisomerase II $\beta$  (TOP2B). Both enzymes have extensive sequence homology at their N-termini and ATPase domains but are divergent at their C-termini (Austin *et al.*, 1993).

A.



**Figure 1.6 – Topoisomerase type IIA mechanism of action** – To release torsional stress on DNA Type IIA topoisomerase first bind a DNA strand (1). Following DNA binding, ATP is bound by its ATPase domain (2). ATP changes the conformation of the homodimer, promoting gate closure and cleavage of the first captured DNA (3). DNA is cleaved by the formation of a covalent 5'phosphotyrosyl bond between DNA and enzyme. ATP hydrolysis promotes further conformational changes that results in displacement of the broken strands that allow strand passage by the second captured strand to C-terminal gate (4). The broken strands are then re-ligated (5), and the product is released (6). Topoisomerase poisons such as etoposide and doxorubicin inhibit broken DNA strand re-ligation resulting in an abortive TOP2cc. **Figure adapted from Vos *et al.* 2011.**



In mice, TOP2A is highly expressed in proliferative tissues such as bone marrow, spleen and intestine, whereas TOP2B expression is present almost ubiquitously (Capranico *et al.*, 1992). Furthermore, TOP2A expression changes with the cell cycle, and is only present from DNA synthesis to mitosis (Woessner *et al.*, 1991). One of the main roles of DNA topoisomerases is to solve the topological stress that arises during semi-conservative replication, which requires unwinding of DNA strands leading to the accumulation of positively supercoiled DNA ahead of the replication fork and DNA catenates behind the fork (Baxter, 2015). Although TOP1 can release some positive supercoiling ahead of the fork, TOP2A is required to resolve DNA catenation (Carpenter and Porter, 2004). This requirement of TOP2A in resolving catenates makes it essential for DNA replication, and essential for chromosome condensation and segregation (Grue *et al.*, 1998).

TOP2B is dispensable in cells, but is essential in mouse since knock-out animals die postnatally of respiratory failure, with embryos exhibiting defective innervation of the lungs by sensory and motor neurons (Yang *et al.*, 2000). It was also shown that TOP2B is essential for the expression of neurodevelopmentally regulated genes (Lyu *et al.*, 2006). TOP2B activity was also linked to receptor-mediated transcription programmes, such as those driven by oestrogen (Ju *et al.*, 2006), androgen (Haffner *et al.*, 2010), and glucocorticoid (Trotter *et al.*, 2015) receptors, in addition to the expression of early response genes in neuronal cells (Madabhushi *et al.*, 2015). Although the role of TOP2B in transcription is not clear, it has been suggested that the formation of a DSB, mediated by TOP2B, is required to initiate transcription (Madabhushi *et al.*, 2015). TOP2B was shown to locate to TSS, including promoter and enhancer regions and to co-localise with multiple DNA repair proteins including PARP1, Ku70/80 and DNA-PKcs supporting this break-driven transcription mechanism (Madabhushi, 2018).

Because of their role in replication and transcription, the abortive activity of TOP2A and TOP2B is a potential threat to genome instability, since the transient TOP2cc can be converted into an abortive DSB by collision with DNA replication forks, and transcription complexes (Deweese and Osheroff, 2009). In addition, similar to TOP1, TOP2 activity can be catalytic inhibited and/or 'poisoned' (Delgado *et al.*, 2018). Catalytic inhibitors include drugs such as ICRF193 that bind to the enzyme ATPase domain and result in an enzyme that cannot promote strand displacement (Ishida *et al.*,

1991), whereas topoisomerase ‘poisons’ promote the formation of abortive TOP2cc and DSBs (Figure 1.6) (Pommier *et al.*, 2010). Topoisomerase ‘poisons’, such as etoposide, were observed to intercalate between base pairs at site of break and prevent re-ligation of the TOP2cc (Delgado *et al.*, 2018). In cell survival assays, TOP2 poisons were shown to preferentially kill replicating cells, making them useful chemotherapy agents (D’Arpa *et al.*, 1990). The current drawbacks of TOP2 inhibitors are cardiotoxicity and therapy-related Acute myeloid leukaemia (t-AML) (Super *et al.*, 1993; Minotti *et al.*, 2004), both of which have been linked to abortive TOP2B activity (Azarova *et al.*, 2007). A striking feature of t-AML is that the causative chromosomal translocations have been mapped to regions of TOP2B binding in the genome (Cowell *et al.*, 2012). Cardiotoxicity on the other hand, is thought to be caused by an indirect reaction of ROS generated by drug redox mechanisms, which causes abortive TOP2cc of TOP2B as TOP2A is not expressed present in myocytes (Delgado *et al.*, 2018).

#### **1.3.4. Topoisomerase type IIB**

Like topoisomerase type IIA, type IIB enzymes catalyse DNA by introducing a DSB with the catalytic tyrosine linked covalently to the 5'-DNA termini, and in eukaryotes is represented by Sporulation 11, Initiator of Meiotic Double Stranded Breaks-like (SPO11) (Bergerat *et al.*, 1997). Unlike type IIA topoisomerases, SPO11 introduces a DSB with 2-bp 5'-overhang and does not reverse the break introduced in DNA via its catalytic tyrosine (Keeney, 2008). SPO11 is essential for meiotic recombination as it introduces the necessary DSBs that initiate the process that results in genetic diversification via homologous recombination-mediated crossing over (Keeney and Kleckner, 1995; Keeney *et al.*, 1997). The mechanism that leads to the stabilisation of SPO11 towards the covalently protein-DNA complex, instead of break re-ligation, is not well understood. However, it is believed to be stimulated by conserved proteins that help with SPO11 localisation, in addition to protecting DNA ends and promoting repair (Keeney, 2008). In the prophase I stage of meiosis, covalently attached SPO11 is removed from 5' DNA termini by MRE11 and CtIP, which results in DNA end resection and repair through homologous recombination repair, as described above, that will ultimately lead to

Holliday junction formation and resolution and the reshuffling of genetic information between homologous chromosomes (Lam and Keeney, 2014).

#### **1.4. REPAIR OF TOPOISOMERASE CLEAVAGE COMPLEXES BY TYROSYL DNA PHOSPHODIESTERASES**

For many years it was believed that non-specific nucleases mediated the removal of topoisomerase cleavage complexes. It was shown that XPF/ERCC1, XPG, SLX1-SLX4, MUS81-EME1 and MRN nucleases could mediate repair of these lesions through endonucleolytic excision of the damage end with the covalently linked DNA (Pommier *et al.*, 2006). Nevertheless, while these excision mechanisms for topoisomerase lesions are employed, it is now understood that the preferred way of reversal of the phosphodiester bond linkage between DNA and topoisomerase is carried by TDP1 and TDP2 (Kawale and Povirk, 2018).

##### **1.4.1. Tyrosyl DNA phosphodiesterase 1**

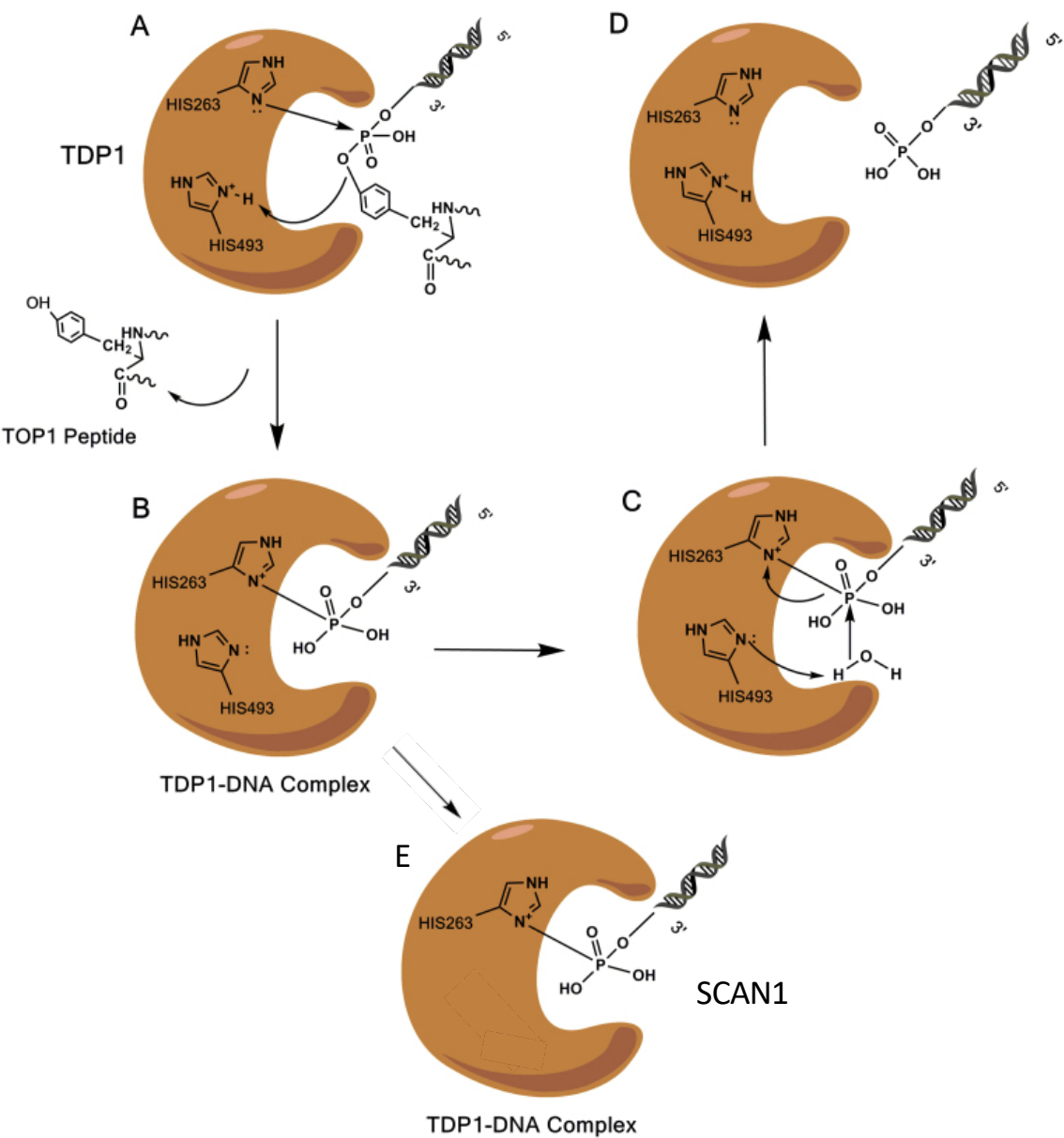
TDP1 was first identified in cellular extracts from yeast cells, where it was shown to promote specific, metal-independent, hydrolysis of the phosphotyrosyl bond linking TOP1 peptide to the 3'-terminus of a SSB, resulting in release of a 3' phosphate terminus (Yang *et al.*, 1996; Pouliot *et al.*, 1999). TDP1 was shown to be able to process TOP1cc containing polypeptides but not intact topoisomerase I (Deb  thune *et al.*, 2002), suggesting that degradation of the covalently linked topoisomerase by the proteasome must occur (Desai *et al.*, 2001; Lin *et al.*, 2009). This process is most likely regulated by ubiquitin and sumo ligases that promote TOP1 ubiquitination, sumoylation and eventual 26S proteasome degradation leaving a polypeptide attached to DNA (Desai *et al.*, 2001).

TDP1 protein belongs to the phospholipase D superfamily family and has been shown to have two conserved histidine-lysine-asparagine (HKN) motifs as part of its active site (Interthal *et al.*, 2001). TDP1 catalysis uses a two-step mechanism involving H493 and H263 for TOP1cc removal (Pommier *et al.*, 2014). In the first step, the catalytic H263 forms a transient covalent phosphoamide bond with DNA releasing the topoisomerase peptide and, in a second step, a water molecule coordinated by H493 is

employed for hydrolysis of the phosphoamide linkage to release DNA and enzyme (Figure 1.7) (Interthal *et al.*, 2001). TDP1 active site and mechanism of action, also allows it to process other DNA end substrates, such as, 3' phosphoglycolates, 3'dRP and 5' phosphotyrosyl bond ((Pommier *et al.*, 2014); (Flett *et al.*, 2018)). The last albeit with much less efficiency than TDP2, however, such redundancy might even allow individuals without TDP2 enzyme, to mediate some repair of TOP2cc lesions, discussed on Chapter 6.

In addition to its role as end processing enzyme, the TDP1 protein also seem to aid repair by multiple protein modifications and interactions. TDP1 is phosphorylated at threonine 81 by ATM and DNAPK following CPT treatment, which promotes TDP1 localisation to  $\gamma$ H2AX foci, interaction with XRCC1-LIGIII and mediates repair and survival (Das *et al.*, 2009; Chiang *et al.*, 2010). In addition, TDP1 is SUMOylated at lysine 111 which facilitates its recruitment to sites of damage (Hudson *et al.*, 2012). Lastly, PARP1 was also shown to PARylate TDP1 in response to camptothecin and promote protein stability and its recruitment to sites of damage together with XRCC1-LIGIII complex (Das *et al.*, 2014).

Following TDP1 processing of the TOP1cc, the SSB formed is processed by PARP1, PNKP and XRCC1-LIGIII proteins, as discussed in section 1.2.3. This role of TDP1 in allowing repair of TOP1cc to be rapidly processed through SSBR seems crucial to neurological function as mutation of TDP1 catalytic residue H493 to arginine (H493R) leads to *spinocerebellar ataxia with axonal neuropathy* (SCAN1). SCAN1 is a rare disease only identified in a single family where multiple individuals present with ataxia, muscle weakness, sensory disturbance and cerebellar atrophy (Takashima *et al.*, 2002). SCAN1 cells cannot mediate removal of TOP1cc from DNA, as it cannot coordinate the water molecule to release the TDP1 enzyme from DNA, becoming abortive and itself a DNA lesion (Figure 1.7) (El-Khamisy *et al.*, 2005; Interthal *et al.*, 2005). However, the TDP1-DNA lesion has a short half-life and the most striking defect in patients cells with the H493R mutation still their inability to mediate SSBR following damage induction with camptothecin (El-Khamisy *et al.*, 2005; Interthal *et al.*, 2005). Interestingly, mice lacking TDP1 do not exhibit any overt phenotype, even though cells derived from this animal were highly sensitive to CPT (Katyai *et al.*, 2007; Hirano *et al.*, 2007).



**Figure 1.7 – TDP1 mediated TOP1cc removal-** TDP1 has two catalytic histidine on its catalytic site. The first histidine (H263) attacks the phosphodiester bond, the second catalytic histidine (H493), donates a proton to the tyrosyl moiety of the leaving group which result in the release of TOP1 peptide and the formation of a TDP1-DNA complex. (B) The TDP1-DNA complex intermediate is attacked by a second nucleophilic attack via an activated water molecule by H493 (C). (D) Generation of a final 3'-phosphate product and free TDP1. (E) The SCAN-1 mutations (H493R) leads to the TDP1-DNA intermediate as second catalytic histidine is absent. **Figure adapted from Kawale AS and Povirk LF 2018.**

#### 1.4.2. Tyrosyl DNA phosphodiesterase 2

Both TDP1 and TDP2 were found by searching for enzymes that can hydrolyse a 3'-phosphotyrosyl bond. TDP2 was discovered somewhat serendipitously, since its primary substrate is 5'-phosphotyrosyl bonds. Cortes-Ledesma *et al.* employed the CPT sensitivity of *Saccharomyces cerevisiae* lacking *tdp1* and *rad1* (homologue of human XPF) to screen a human cDNA library for complementary DNA sequences that promoted survival (Ledesma *et al.*, 2009). From the recovered clones they identified TDP1 and TTRAP ((TRAF) and TNF-associated protein), which was previously shown to exhibit sequence homology with APEX1 endonuclease and to inhibit Necrosis-Factor kappa B (NF- $\kappa$ B) (Rodrigues-Lima *et al.*, 2001; Pype *et al.*, 2000; Ledesma *et al.*, 2009). Cortes-Ledesma *et al.* went on to show that TTRAP was in fact an enzyme that possessed both 3' and 5' tyrosyl phosphodiesterase activity, with the latter 50-fold more active than the former (Ledesma *et al.*, 2009). It was confirmed that the enzyme promoted cellular resistance and DSB repair following treatment with etoposide, and was renamed tyrosyl DNA phosphodiesterase 2 (TDP2) (Ledesma *et al.*, 2009). TDP2 was subsequently confirmed to be the primary source, if not the only, of 5'-tyrosyl DNA phosphodiesterase activity in chicken B cells (DT40) (Zeng *et al.*, 2011). *Tdp2*<sup>-/-</sup> DT40 cell extracts lack detectable 5'-phosphodiesterase activity and were hypersensitive to etoposide but not to either MMS or CPT. In mouse embryonic fibroblasts (MEFs) and chicken cells lacking Tdp1 however, Tdp2 does promote SSB repair and cell survival following treatment with CPT, consistent with the yeast cDNA screen and demonstrating that the weak 3'-tyrosyl DNA phosphodiesterase activity is physiologically relevant if Tdp1 activity is absent (Zeng *et al.*, 2012).

*Tdp2* <sup>$\Delta 1-3$</sup>  mice, in which TDP2 activity is absent, show no overt phenotype unless treated with topoisomerase 'poisons', in which case, spleen, thymus and intestinal atrophy can be observed (Gómez-Herreros *et al.*, 2013). *Tdp2* <sup>$\Delta 1-3$</sup>  MEFs when treated with etoposide exhibit increased chromatid breaks, micronuclei, nucleoplasmic bridges and sister chromatid exchanges, consistent with increased chromosome instability and elevated HRR (Gómez-Herreros *et al.*, 2013). In contrast, in humans, the absence of TDP2 leads to *spinocerebellar ataxia, autosomal recessive 23* (SCAR23). This disorder was first reported in three Irish brothers that lack detectable TDP2 protein and 5'-tyrosyl DNA

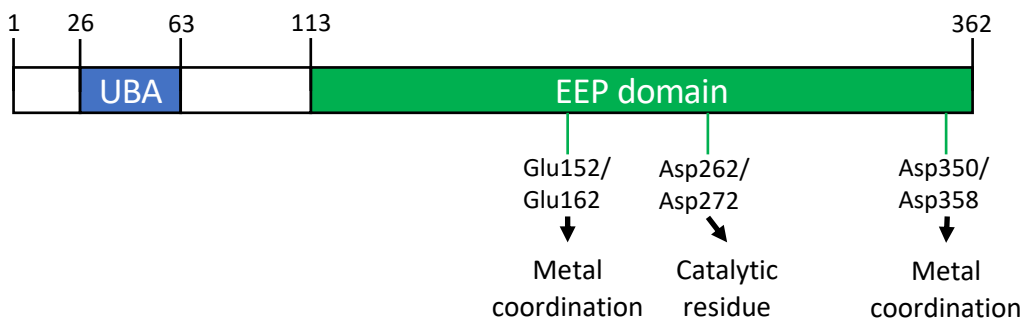
phosphodiesterase activity due to the splicing site mutation c.425+1G>A (Gómez-Herreros *et al.*, 2014). These patients present with seizures, ataxia, dysmorphic facial features and learning disabilities (Gómez-Herreros *et al.*, 2014; Zagnoli-Vieira *et al.*, 2018). Patient lymphoblastoid cells also exhibited the chromosomal instability seen in *Tdp2*<sup>Δ1-3</sup> MEFs following etoposide treatment, and also increased levels of chromosome translocation (Gómez-Herreros *et al.*, 2017). Increased translocations were also observed in *TDP2*<sup>-/-</sup> human cell lines and shown to occur in regions previously linked to TOP2β-mediated transcription, and were greatly reduced or abolished if transcription was prevented during exposure to etoposide (Gómez-Herreros *et al.*, 2017). Evidence in support of a role for TDP2 in the repair of transcription-associated TOP2cc was also found in mouse, in which decreased expression of TOP2β associated genes was observed in *Tdp2*<sup>Δ1-3</sup> brain (Gómez-Herreros *et al.*, 2014).

TDP2 is a protein of 362 amino acids (aa) that encodes an N-terminal ubiquitin-associated (UBA) domain (aa 26-63) and a C-terminal catalytic domain (aa 113-362) (Hornyak *et al.*, 2016) (Figure 1.8.A). The UBA domain was shown to form a four α-helices bundle in *C. elegans* TDP2 (CeTDP2) and to bind polyubiquitin chains linked by lysine (K) 48 or K63 (Rao *et al.*, 2016). Complementation of *Tdp2*<sup>-/-</sup> DT40 cells with human TDP2 lacking the first 100 amino acids failed to fully complement sensitivity to etoposide further suggesting that the UBA domain is important for DNA repair of TOP2cc (discussed on Chapter 4) (Rao *et al.*, 2016). Structural studies of the TDP2 catalytic domain confirmed its similarity to other exonuclease-endonuclease-phosphatase (EEP) domain containing proteins such as APEX1 and the requirement for Mg<sup>2+</sup> for its activity (Schellenberg *et al.*, 2012) (Figure 1.8.B). These studies further revealed that TDP2 grasps its substrate via a binding cleft and promotes hydrolysis by water-mediated nucleophilic attack of the 5'-tyrosyl phosphodiester bond, with Mg<sup>2+</sup> coordinating the 5'-phosphate and resulting in the release of TOP2 peptide and a 5'-phosphate terminus (Figure 1.9) (Schellenberg *et al.*, 2016). This mechanism of action enables TDP2 to process phosphotyrosyl bonds in a variety of substrates including RNA, single-strand DNA, and double-stranded DNA with overhang or blunt ends (Gao *et al.*, 2012; 2014).

Studies using etoposide have previously shown that, following drug treatment, TOP2 is degraded by the proteasome, similar to TOP1 after CPT treatment, and that proteasome inhibition not only stops this degradation but also the DNA damage

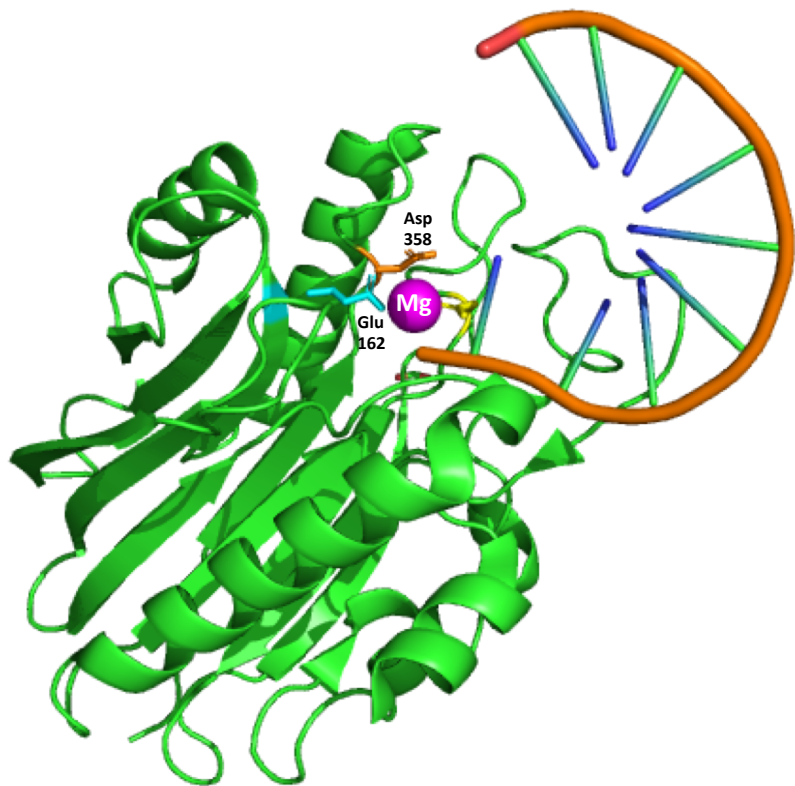
A.

Human TDP2



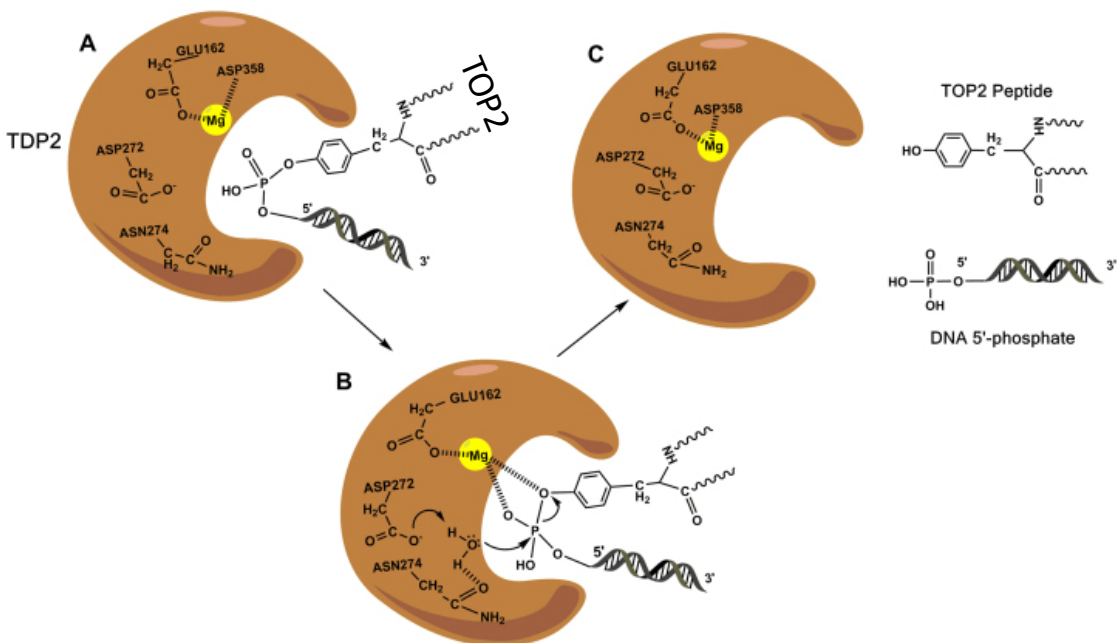
B.

Mouse TDP2



**Figure 1.8 – 5' Tyrosyl DNA Phosphodiesterase 2 (TDP2)** - (A) Diagram showing the human 5' Tyrosyl DNA Phosphodiesterase 2 (TDP2) enzyme with amino acid positions for ubiquitin associated (UBA) domain and the catalytic exonuclease-endonuclease-phosphatase (EEP) domain. In the human protein, glutamic acid (Glu) on position 152 and aspartic acid (Asp) on position 350 are responsible for metal ion coordination on catalytic site, whereas Asp262 is catalytic residue. Additional amino acid coordinates shown indicates correspondent positions on mouse protein. (B) Structure of mouse TDP2 showing DNA interaction with TDP2 catalytic site where a magnesium ion (Mg) is coordinated by Glu162 (cyan) and Asp358 (orange). Diagram was produced in Pymol from structure published by Schellenberg *et al.*, 2016 PDB ID: 4GZ1

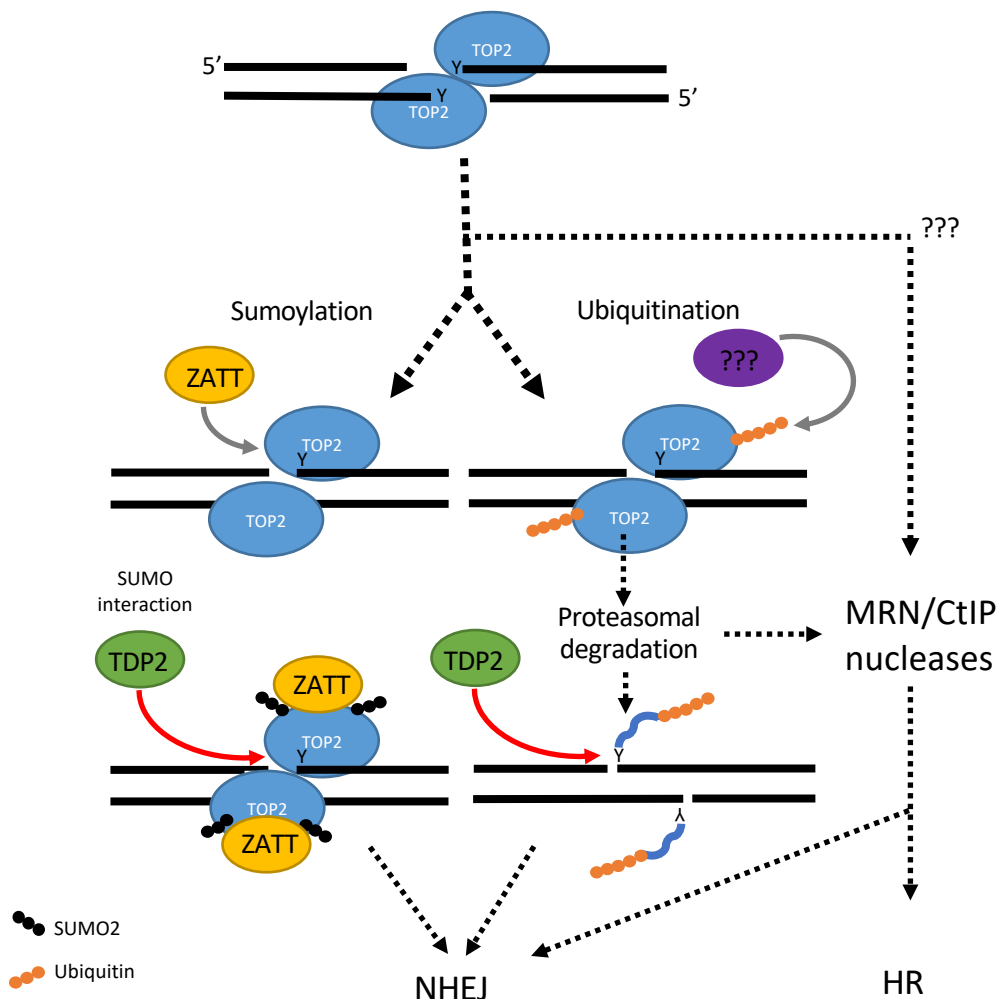




**Figure 1.9 – TDP2 mediated TOP2cc removal** - (A) TDP2 binding cleft accommodates Topoisomerase 2 (TOP2) enzyme (or peptide) that is covalently linked to DNA via a 5'-phosphotyrosyl bond. (B) TDP2 mediates removal of the topoisomerase by coordination of phosphodiester bond through  $Mg^{2+}$  and promotion of nucleophilic attack using a water molecule (C) Cleavage products consisting in the topoisomerase enzyme (top) and the liberated DNA with a 5'-phosphate re-ligatable end are released. **Figure adapted from Kawale AS and Povirk LF 2018.**

response (Zhang *et al.*, 2006). In addition, *in vitro* studies have shown that TDP2 requires prior degradation or denaturation of TOP2 to enable access to the 5'-tyrosyl DNA phosphodiester bond within the TOP2cc (Gao *et al.*, 2014). Whilst this implicates proteasome-mediated degradation of TOP2 in DSB repair, it was reported recently that TDP2 can in fact process intact TOP2cc if TOP2 is first SUMOylated by ZATT (Zinc finger protein associated with TDP2 and TOP2; formerly ZNF451) (Schellenberg *et al.*, 2017). TOP2 SUMOylation appears to alter the conformation of TOP2, rendering the phosphotyrosyl bond accessible to TDP2 (Schellenberg *et al.*, 2017). TDP2 interacts with SUMOylated TOP2 via a C-terminal SUMO-interacting motif (SIM) (Schellenberg *et al.*, 2017; Zagnoli-Vieira and Caldecott, 2017). Although it is not clear how cells differentiate between transient and abortive TOP2cc, it seems that a similar mechanism is also present in fission yeast, in which Rrp2, a Snf2-family DNA translocase, is employed (Wei *et al.*, 2017).

Despite much progress in recent years, the nature of, and interplay between, pathways that repair TOP2cc are still unclear. Whilst genetic epistasis studies in DT40 cell indicating that TDP2 operates as part of Ku-dependent NHEJ (Gómez-Herreros *et al.*, 2013). Currently at least four pathways seems to be involved in repairing TOP2cc lesions; a TDP2-dependent but proteasome-independent NHEJ pathway mediated via ZATT regulated SUMOylation; a TDP2-dependent, proteasome-dependent, NHEJ pathway; an MRE11 nuclease-dependent NHEJ pathway, and an NHEJ-independent pathway that is dependent on HRR (Figure 1.10) (Zagnoli-Vieira and Caldecott, 2017).



**Figure 1.10 – Overview of TOP2cc repair-** *Top*, TOP2 becomes transiently linked to the 5'-terminus of a DSB during its normal catalytic cycle via a phosphotyrosine bond. These TOP2 cleavage complexes (TOP2cc) can become abortive, and so require DSB repair for their removal. **Right**, TOP2 can be removed by a potentially mutagenic mechanism that employs nucleases such as MRE11/RAD50/NBN (MRN) complex and CtIP followed by DSB processing via NHEJ or the error free HR. It is not clear if nuclease processing of TOP2cc requires prior processing via the proteasome. Alternatively, abortive TOP2cc can be repaired by TDP2-dependent mechanisms (left), providing an error free route that does not require DNA degradation. **Middle**, the abortive TOP2cc is sumoylated by ZATT (left), and/or ubiquitinated by an enzyme to be identified (right). **Bottom**, TDP2 is recruited to the abortive TOP2cc by interaction with SUMO (left), and/or possibly ubiquitin (right), releasing TOP2 from the DSB by a hydrolytic mechanism. DSB following TDP2 processing are repaired via NHEJ repair pathway. **Figure adapted from Zagnoli-Vieira G and Caldecott KW 2017.**

### 1.5. AIM OF THESIS

Considering the role of topoisomerase 2 in multiple cellular processes and the potential treat that abortive TOP2cc presents to the genome, the aim of this thesis was to further investigate the role of TDP2 DNA repair and human disease. To do this, I aimed to generate and characterise a variety of human cell lines in which I deleted *TDP2* using CRISPR-Cas9 mediated gene editing (Chapter 3). I then employed these cells to characterise two putative novel TDP2 isoforms and their role in DNA repair (Chapter 4). In addition, I aimed to employ these cells to establish a novel tyrosyl DNA phosphodiesterase assay capable of measuring TDP activity in whole cell extracts (Chapter 5). Finally, I aimed to employ this assay to investigate the impact of *TDP2* mutation in three newly identified TDP2-mutated patients; measuring both their level of 5'-tyrosyl DNA phosphodiesterase activity and their cellular response to TOP2-induced DNA damage (Chapter 6).

## 2. CHAPTER TWO – MATERIAL AND METHODS

### 2.1. Mammalian cell culture

Human A549, U-2-OS, CAL-51 and MCF-7 cells were grown in Dulbecco Modified Eagle Medium (Gibco, ThermoFisher, Waltham, MA) containing 10% fetal calf serum (FCS) (PAN Biotech), 2 mM glutamine, penicillin (100 units/mL), and streptomycin (100 µg/mL). Human fibroblasts were grown in Minimum Essential Media (Gibco) containing 15% FCS, 2 mM glutamine, penicillin (100 units/mL), and streptomycin (100 µg/mL). All cells were grown at 37°C humidified atmosphere containing 5% CO<sub>2</sub>. *TDP2*-mutated patient primary human fibroblasts were established from a patient's skin biopsy and were denoted 850BR by Dr. Limei Ju. The control human fibroblast cell line 1BR.3 (named here 1BR for simplicity) was previously derived from an unrelated normal individual and has been described previously (Hoch *et al.*, 2017). For complementation experiments, we used 1BR cells that were immortalized previously with hTERT (denoted as 1BR hTERT) and a derivative of 850BR immortalized by retroviral-mediated hTERT expression and selected in a medium containing 1 µg/ml puromycin (denoted as 850BR hTERT) by Dr. Elena Korneeva.

#### 2.1.1. siRNA transfection

siRNA transfection was carried using the reverse-transfection method (cells plated with siRNA). *TDP2* siRNA sequence was previously described (Ledesma *et al.*, 2009) and control siRNA sequence was from Non-targeting On-TargetPlus smartpool from Dharmacon (D-001810-10-05). Each siRNA transfection reaction was made of 97.3 µL of OptiMEM (Gibco) medium mixed with 2.7 µL of Lipofectamine RNAiMax reagent (ThermoFisher), which were added to 100 µL of OptiMEM containing 125 nM of siRNA duplex. Once reaction was mixed it was left at room temperature for 5 minutes prior adding it to 800 µL of media containing approximately  $2.5 \times 10^5$  cells. The total 1 mL reaction was plated in 6 well dishes and incubated for 12-16 hours prior supplementation with 1 mL of fresh growth medium. At 48 h post-transfection cells were washed in phosphate-buffered saline (PBS) solution, lysed and assayed for protein depletion via Western-blotting.

### **2.1.2. Vector transfection**

#### **2.1.2.1. Neon Electroporation**

The Neon nucleofection system (ThermoFisher) was used to transfect RPE-1 cells and patient fibroblasts with DNA vectors according to the manufacturer's instructions. For each reaction  $1.2 \times 10^6$  trypsinised and PBS-washed cells were resuspended in 30  $\mu$ L buffer R and combined with a total of 1.5  $\mu$ g plasmid DNA. The mixture was electroporated with the device using the manufacturer recommended setting of 2 pulses of 1350 volts for 20 milliseconds, prior to immediate plating in complete medium. For antibiotic selection, G418 was added to media after 24 hours to a final concentration of 0.5 mg/mL. For stable cell selection cells were kept in antibiotic for 21 days to enrich for transfectants. Population or clonal isolates of positive cells were then used as indicated.

#### **2.1.2.2. GeneJuice transfection**

Transfection using GeneJuice (Merck) was used for vector transfection in CAL-51 and U2OS cell lines. 24 hours prior transfection  $2.5 \times 10^5$  cells were plated in 6 well dishes. On the day of transfection 3  $\mu$ L of GeneJuice was added to 100  $\mu$ L of OptiMEM and mixed thoroughly. This mixture was left in room temperature for 5 minutes prior adding 1.5  $\mu$ g plasmid DNA and mixing by inversion. Final mixture was then incubated for 15 minutes at room temperature before adding to cells dropwise. For antibiotic selection, G418 was added to media 24 hours post-transfection at 0.8 mg/mL for U2OS and 1.0 mg/mL for CAL-51 cells.

## **2.2. TDP2 $\beta$ sequence capture**

Cloning of the TDP2 $\beta$  open reading frame (ORF) sequence was captured in a complementary DNA (cDNA) library prepared by using oligo deoxythymine (dT) from TDP2<sup>-/-</sup> A549 Exon 1 cells by Dr. Fernando Gomez-Herreros. TDP2 $\beta$  sequence was captured using polymerase chain reaction (PCR) with the following the primers

TDP2beta\_FW (5'- GAATTCATGGCAACTTGGAAGTCTCCTTGCT-3') and TDP2beta\_RV (5'- GGATCCTTACAATATTATATCTAAGTTGCACAGAAGAC-3'). Phusion Polymerase (NEB) was used for sequence amplification as per manufacturer instructions. PCR product was sub-cloned into TOPO TA construct according to manufacturers' instructions. Sanger sequencing of single clones of TOPO-TA bacterial transformants were used to confirm TDP2 $\beta$  sequence.

### 2.3. *TDP2*<sup>-/-</sup> complementation

For complementation of *TDP2*<sup>-/-</sup> RPE-1 (Exon 1) clone #3 cells were transfected with either vector (pCD2E) (Caldecott *et al.*, 1992), encoding human *TDP2* or *TDP2*<sup>D262A</sup>, which were generated by PCR amplification of their respective sequences from previously published vectors (Ledesma *et al.*, 2009). The following PCR primers were used TDP2pE\_FW (5'-GAATTCATGGAGTTGGGGAGTTGCCTGGAG-3'); TDP2pE\_RV (5'- AAAGAATTCTTACAATATTATATCTAAGTTGCACAGAAGACC-3'). Both constructs were made by sub-cloning the PCR product into the *EcoRI* sites of pCD2E. Cells were transfected as described above and selected in 0.5 mg/mL of G418 for 21 days prior clonal isolation. Selection of positive clones were done using TDP2 immunofluorescence as described below.

For complementation of *TDP2*<sup>-/-</sup> RPE-1 (Exon 4) with TDP2 isoforms, cells were transfected with either empty eGFP-N1 (Clontech) vector or eGFP-N1 construct encoding GFP-tagged human TDP2 (denoted as TDP2-GFP), TDP2 $\beta$  (TDP2 $\beta$ -GFP) or TDP2 $\gamma$  (TDP2 $\gamma$ -GFP) and stable transfectants selected for 21 days by growth in a medium containing 0.5 mg/mL G418 (Gibco). TDP2-GFP was generated by PCR amplification of the human TDP2 ORF using the primers *TDP2\_FW* (5'- AAAGAATTCATGGAGTTGGGGAGTTGCCTG-3') and *TDP2\_RV* (5'- AAAGGATCCAATATTATATCTAAGTTGCACAGAAGACC-3'). For TDP2 $\beta$ -GFP, the forward primers *TDP2B\_FW* (5'-GAATTCATGGCAACTTGGAAGTCTCCTTGCT-3') were used instead. For TDP2 $\gamma$  construct (TDP2 $\gamma$ -GFP) the following forward primer was used instead *TDP2G\_FW* (5'-AAAGAATTCATGGAAAGGGCTCTGAACTCCTACT-3'). PCR reactions used Phusion HF-DNA Polymerase (NEB). All constructs were made by sub-cloning the PCR product into the *EcoRI/BamHI* sites of eGFP-N1. Cells were transfected

as described above and selected in 0.5 mg/mL of G418 for 21 days. Gene expression was confirmed by immunoblotting as described below.

## **2.4. SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE) of proteins**

All protein samples were prepared in Laemmli buffer (2% sodium dodecyl sulfate (SDS), 10% glycerol, 60mM Tris-HCl pH6.8) and boiled for 10 minutes at 95°C. When preparing cell lysates, protein quantification and normalisation was done using Bicinchoninic acid assay (BCA) working reagent (Pierce) according to manufacturer instructions. Prior loading samples were supplemented with 100 mM dithiothreitol and 0.005% bromophenol blue. Protein samples were resolved in either 10% polyacrylamide as described previously (Sambrook and Russell, 2001) or by using pre-cast 4-20% polyacrylamide gels (BioRad). Samples were loaded along with protein marker (BioRad), and electrophoresis was conducted at 200V in a Mini-PROTEAN Tetra Cell electrophoresis tank (BioRad), using 1X Running Buffer (25 mM Tris-HCl, 250 mM glycine, 0.01% SDS).

### **2.4.1. Western blotting**

After SDS-PAGE, proteins were transferred to nitrocellulose membrane (0.45 µm) (ThermoFisher) at 80 volts for 90 minutes, using 1X Towbin buffer (25 mM Tris-HCl, 192 mM glycine, 10% methanol). After transfer, nitrocellulose membrane was blocked in 5% skimmed milk in TBST (20 mM Tris-HCl, 150 mM NaCl, 0.1% Tween-20) for 1 hour. Then membrane was incubated with primary antibody also present in 5% milk in TBST for 2 hours. Excess antibody was washed for 30 minutes in TBST. Secondary HRP-conjugate antibody was then incubated for 1 hour, followed by further 30 minutes wash with TBST. Finally, the membrane was incubated for 1 minute in enhanced chemiluminescence (ECL) detection reagent (GE Healthcare), prior exposure to autoradiograph film and developing accordingly. Primary antibodies used were Anti-TDP2 (1:5000), Anti-Actin (Protein tech #66009; 1:10000) and Anti-Histone 1.2 (Abcam #17677; 1:20000). Anti-TDP2 polyclonal antibody (#20613) was raised in rabbit against purified human TDP2 and affinity purified as described in Thomson et al. 2013.



Secondary antibodies were Rabbit Anti-mouse HRP (Dako P0260; 1:10000) and Goat Anti-rabbit HRP (Dako P0448; 1:10000).

#### **2.4.2. Silver staining**

After protein resolution using SDS-PAGE as described above, gel membrane was incubated in fixing solution (50% ethanol and 10% acetic acid) for 30 minutes. Membrane was then hydrated for 15 minutes in 1:10 dilution of fixing solution (5% ethanol and 1% acetic acid). It was then washed three times in water for 5 minute per cycle prior incubation with 200 mg/L solution of sodium thiosulphate for 1 minute. Membrane was then rinsed three times with water and incubated for 20 minutes in the staining solution (2 mg/mL  $\text{AgNO}_3$ , 0.0277% formaldehyde). After incubation membrane was rinsed in water three times. Finally, the membrane was placed in developing solution (60 mg/mL sodium carbonate, 0.0185% formaldehyde and 4  $\mu\text{g/mL}$  sodium thiosulphate) till protein bands were visible. Reaction was stopped by replacing developing solution with 5% acetic acid. Membrane was kept in acetic acid prior excision of desired protein bands and mass spectrometry analysis.

#### **2.5. TDP2 Immunoprecipitation**

For TDP2 immunoprecipitation, cells pellets containing  $5.0 \times 10^6$  cells were collected and washed in PBS prior liquid nitrogen freezing and storage in  $-80^\circ\text{C}$ . On the day on the experiment cell pellet was thawed on ice for 30 minutes prior re-suspending by pipetting in 500  $\mu\text{L}$  of IP buffer (50mM Tris-HCl pH8, 150mM NaCl, 1.5mM Ethylenediaminetetraacetic acid (EDTA), 1mM phenylmethylsulfonyl fluoride (PMSF) and 1x EDTA free protease cocktail inhibitor(Sigma)) supplemented with 0.3% Triton detergent. Cell lysate was then sonicated for 10 on/off cycles of 30 seconds in BioRuptor sonicator (Diagenode). Lysate was centrifuged at 15000g for 10 minutes and supernatant was collected. Cleared cell lysate was then incubated at rotating wheel at  $4^\circ\text{C}$  with 10  $\mu\text{L}$  of Dynabeads Protein G solution (Thermofisher) for 30 minutes. After incubation beads were removed using magnetic extraction. 1 mL of IP buffer was then added to lysate prior further incubation with 7.5  $\mu\text{g}$  of TDP2 antibody in rotating wheel

at 4°C for one hour. Next, 40 µL of Dynabeads were added to the lysate and incubated for 1 h. Finally, Dynabeads were captured using magnetic rack and washed in IP buffer supplemented with 1% Triton for three times prior adding Laemmli buffer to beads and boiling at 95°C. Samples were then processed as described in section 2.3 and 2.3.2)

## 2.6. CRISPR-Cas9

*TDP2*<sup>-/-</sup> cells were created as previously reported (Mali *et al.*, 2013a). In brief, we used a 17-bp (minus the PAM) RNA sequence targeting *TDP2* exon 1 (5'-GCGGCGACTTCTGTGTG-3') or *TDP2* exon 4 (5'-GTAGAAATATCACATCT-3'), which were selected using the tool E-CRISP ([e-crisp.org/E-CRISP/](http://e-crisp.org/E-CRISP/)) and cloned into the guide RNA vector #41824 (AddGene). The *TDP2* guide construct was co-transfected with hCas9 expressed from plasmid #41815 (AddGene) method described above. Transfected cells were enriched by G418 (ThermoFisher) selection (concentration for each cell line in section 2.1.2) for 5 days before isolation of single clones and screening for loss of *TDP2* expression by Western blotting.

## 2.7. DNA sequencing

### 2.7.1. Sanger sequencing

DNA was extracted using the DNeasy Blood & Tissue kit (Qiagen, Manchester, UK). PCR reactions used Phusion HF-DNA Polymerase (NEB). For CAL-51 *TDP2*<sup>-/-</sup> #2.6 and RPE-1 *TDP2*<sup>-/-</sup> #3 cells, sequencing was done using the following primers targeting exon 1 *TDP2*E1\_FW (5'-GAATTCATGGAGTTGGGGAGTTGCCTGGAG-3'); *TDP2*E1\_RV (5'-CCTAGGCTGACCGGCCAGAATAACAGGTTTGT-3'). The following primers were used instead when sequencing cells targeted by CRISPR-Cas 9 on exon 4 *TDP2*E4\_FW (5'-GCCAGTGTTGACCTAACCAATGAAGA-3'); *TDP2*E4\_RV (5'-GGGTTCTCTCTCAGTCTGTGTGTGTC-3') (Data not shown). For patient fibroblasts, 850BR, sequencing was carried using the following primers: *TDP2*\_FW (5'-GCCAGTGTTGACCTAACCAATGAAGA-3'); *TDP2*\_RV (5'-CTGTAGAAATATCACATCTGGGCTGTACC-3'); *ZSCAN9*\_FW (5'-ATGAAGTAACCAAGACTGAGGACAGAGAG-3'); and *ZSCAN9*\_RV (5'-

AGACCAGCTCAGCCACTGTGTGGATCT-3'). All PCR products were purified before sequencing using a QIAquick PCR purification kit (Qiagen) prior sequencing.

### **2.7.2. Whole exome sequencing**

DNA was extracted from RPE-1 and *TDP2*<sup>-/-</sup> RPE-1 Clone #3 cells using the DNeasy Blood & Tissue kit (Qiagen) and sent to Source Bioscience for sequencing and variant calling. Exome capture was completed using SureSelect All Exon V5 (Agilent), where molecular probes for 21522 genes and 357999 exons are used. Paired end sequencing was carried on a Nextseq (illumina) with read length of 100 base pairs and median depth of coverage at 60X. Sequencing data was aligned to reference human annotation 19 University of California Santa Cruz using BWA (Li and Durbin, 2010). Variant calling was completed using Genome Analysis Tool Kit (GATK) (McKenna et al., 2010; DePristo et al., 2011). Data was visualised using Integrative Genomics Viewer (Broad Institute).

## **2.8. Microscopy**

### **2.8.1. $\gamma$ H2AX**

Cells were grown on coverslips until confluent and then treated for 30 minutes with 25  $\mu$ M etoposide or irradiated with  $\gamma$ -irradiation (2 Gy). After treatment, cells were rinsed with PBS and fixed for 10 minutes in PBS containing 4% paraformaldehyde at the indicated time points. Cells were permeabilised (20 minutes 0.2% Triton X-100 in PBS), blocked (1 hour in PBS-5% BSA), and incubated with anti- $\gamma$ H2AX (Millipore, 05-636, 1:2500) and anti-CENP-F (Abcam, ab5, 1:2500) antibodies for 3 hours in PBS containing 5% BSA. Cells were then washed (3  $\times$  5 minutes in PBS containing 0.1% Tween-20), incubated for 1 hour with the corresponding Alexa Fluor conjugated secondary antibody (1:1000, 5% BSA), and washed again as described earlier. Finally, cells were counterstained with DAPI (Sigma, Gillingham, UK) and mounted in VECTASHIELD (Vector Labs, Peterborough, UK).

Images were acquired on an automated wide-field microscopy Olympus ScanR system (motorized IX83 microscope) with ScanR Image Acquisition and Analysis

Software, 40×/0.6 (LUCPLFLN 40×PH) dry objectives, and Hamamatsu ORCA-R2 digital CCD camera C10600. Cells were gated to the G1 or S/G2 population according to the DAPI profile of ScanR Image Analysis Software and CENP-F signal as indicated. For patient fibroblast complementation experiments, cells were gated accordingly to eGFP expression on ScanR Analysis Software, in addition to the G1 population according to the DAPI content.

### **2.8.2. TDP2 Immunofluorescence signal quantification**

RPE-1 and *TDP2*<sup>-/-</sup> RPE-1 Exon 1 clone #3 cells complemented with indicated construct were grown on coverslips for at least 24 hours. Cells were rinsed and fixed for 10 minutes in PBS containing 4% paraformaldehyde. Cells were permeabilized (20 minutes in PBS-0.2% Triton X-100), blocked (1 hour in PBS-5% BSA), and incubated with Anti-TDP2 polyclonal antibody (#20613) (1:2000) antibody for 2 hours in PBS containing 5% BSA. Cells were then washed (3 × 5 minutes in PBS containing 0.1% Tween-20), incubated for 1 hour with the corresponding Alexa Fluor conjugated secondary antibody (1:1000, 5% BSA), and washed again as described earlier. Finally, cells were counterstained with DAPI and mounted in VECTASHIELD as previously. Images were acquired on an automated wide-field microscopy Olympus ScanR system (motorized IX83 microscope) with ScanR Image Acquisition and Analysis Software, 40×/0.45 (LUCPLFLN 20×PH) dry objectives, and Hamamatsu ORCA-R2 digital CCD camera C10600. Quantification of TDP2 signal was done in the G1 population according to the DAPI content.

### **2.8.3. TDP2 isoform cellular localisation**

RPE-1, *TDP2*<sup>-/-</sup> RPE-1 Exon 4 and *TDP2*<sup>-/-</sup> RPE-1 Exon 4 cells transfected with TDP2-GFP (TDP2), TDP2β-GFP (TDP2β) or TDP2γ-GFP (TDP2γ) vectors created as above described, were grown on coverslips until for at least 48 hours prior fixation. Cells were rinsed and fixed for 10 minutes in solution of Acetone:Acetic acid (1:1 ratio). Cells were rinsed with PBS, blocked (1 hour 5% BSA in PBS), and incubated with Anti-TDP2 polyclonal antibody (#20613) (1:1000) for 2 hours in PBS containing 5% BSA. Cells were

then washed (3 × 5 minutes in PBS containing 0.1%Tween-20), incubated for 1 hours with the corresponding Alexa Fluor conjugated secondary antibody (1:1000, 5% BSA), and washed again as described earlier. When indicated cells were additionally incubated with anti-PDI (Abcam #2792). Additionally, for experiments using mitochondrial marker, 250 nM of MitoTracker Deep Red (ThermoFisher) dye was added to medium 30 minutes prior fixation. Finally, cells were counterstained with DAPI (Sigma, Gillingham, UK) and mounted in VECTASHIELD (Vector Labs, Peterborough, UK). Images were acquired on a Carl Zeiss AxioObserver Z1 using the objective Plan-Apochromat 63x/1.40 Oil M27 and Hamamatsu ORCA-Flash4.0 LT camera. Image analysis was done in the ZEN blue software (Zeiss).

## 2.9. Clonogenic cell survival

For clonogenic survival treatment, CAL-51 and RPE-1 cells and their respective *TDP2*<sup>-/-</sup> cell lines were plated in 10 cm dishes 4 hours before treatment with the indicated concentrations of etoposide or X-rays irradiation. For acute treatment, RPE-1 cells and *TDP2*<sup>-/-</sup> RPE-1 Clone #3 were plated in 10 cm dishes 4 hours prior treatment with state concentration of etoposide for 1 hour. Dishes were then rinsed with PBS twice prior replenishing dishes with fresh media. All of the cells lines above were incubated for 12 days to allow for macroscopic colony formation. 1BR and 850BR primary fibroblasts cells were plated onto feeder layers and 3 hours later treated with indicated concentrations of etoposide or X-rays irradiation prior incubation for 21 days to allow the formation of macroscopic colonies. For feeder layers, 1BR cells were irradiated (35 Gy) and plated 24 hours before use at  $5 \times 10^4$  cells/10 cm dish.

Following colony formation, cells were fixed in 100% ethanol and dishes rinsed with PBS prior staining with 70% ethanol/1% methylene blue. Dishes were allowed to dry, prior to scoring of colonies of >50 cells using a digital colony counter (Stuart SC6Plus). The surviving fraction at each dose was calculated by dividing the average number of colonies (>50 cells) in treated dishes by the average number in untreated dishes.

## 2.10. Tyrosyl DNA phosphodiesterase assay

Whole-cell extract (WCE) was prepared by resuspension of cell pellets in lysis buffer (40 mM Tris/HCl pH 7.5, 100 mM NaCl, 0.1% Tween-20, 1 mM DTT, 1 mM PMSF, 1x EDTA free protease cocktail inhibitor) to a final volume of 1  $\mu$ L per  $10^4$  cells, followed by 30 minutes of incubation on ice and mild sonication using BioRuptor (Diagenode) for five cycles of 30 seconds on/off. The WCE was clarified by centrifugation for 10 minutes at 4°C at 16000 x *g* in a microfuge and the protein concentration quantified using the bicinchoninic acid (BCA) assay reagent (ThermoFisher). Clarified WCE at stated amounts was incubated with 40 nM of TDP2 substrate (Cy5-5'Tyrosine-ssDNA<sub>19</sub>-BHQ) or TDP1 substrate (BHQ-ssDNA<sub>13</sub>-3'Tyrosine-Cy5) diluted in reaction buffer (50 mM Tris/HCl pH8.0, 10 mM MgCl<sub>2</sub>, 80 mM KCl, and 1 mM DTT, 0.05% Tween-20) in a total volume of 6 $\mu$ L at room temperature. Cy5 fluorescence was measured at 650 nm at the indicated time intervals on a BMG PHERAstar plate reader. When measuring activity in the presence of TDP2 inhibitor, WCE was pre-incubated for 30 minutes at room temperature with either DMSO or state University of Sussex (UoS) 12248 compound dose prior adding substrate and carrying out fluorescence measurement. The purified human TDP2 protein used in all experiments was a kind gift from Dr Antony Oliver and was produced using the codon optimised TDP2 sequence in insect cells.

## 2.11. Bioinformatics

### 2.11.1. Sequence alignments

Protein sequence alignments and phylogeny tree were done in the Clustal Omega (Sievers *et al.*, 2011; Li *et al.*, 2015) using the following sequences from Ensembl (Zerbino *et al.*, 2018) database: Human (*Homo sapiens*) TDP2 ENST00000378198 and ENST00000341060; Cow (*Bos taurus*) ENSBTAT00000000472; Mouse (*Mus musculus*) ENSMUST00000038039; Zebrafish (*Danio rerio*) ENSDART00000103612 ; Frog (*Xenopus laevis*) ENSXETT00000032445 and Chicken (*Gallus gallus*) ENSGALT00000022184.

### **2.11.2. Cellular localisation**

Prediction of TDP2 $\beta$  cellular localisation was done by using DNA sequence present on Ensembl transcript ENST00000341060 on to SignalP or TargetP servers (Emanuelsson *et al.*, 2007; Nielsen, 2017), MitoProtII (Claros and Vincens, 1996) and MitoFates (Fukasawa *et al.*, 2015).

### **2.11.3. Kaplan-Meier survival**

Kaplan-Meier for breast cancer patient were created using the KMPlot tool (Gy  rffy *et al.*, 2010). Data for TDP2 was selected by using probe identifier 202266\_at. Parameters for analysis were modified as indicated on text.

## **2.12. Statistics**

All statistical analysis carried using experimental repeats unless stated otherwise. Analysis of Variance (ANOVA) and t-Test (paired) were used to assess statistical significance unless stated otherwise. Values of calculated probability (p-value) above 0.05 were considered non-significant (ns) and those below 0.05 were considered significant and represented by a star (\*) symbol. \*\* was used to present p-values below 0.01 and \*\*\* for values below 0.001. All analyses were carried in Microsoft Excel Data Analysis package or GraphPad Prism software unless stated otherwise.

### 3. CHAPTER THREE – GENERATION AND CHARACTERISATION OF *TDP2*<sup>-/-</sup> CELLS

#### 3.1. INTRODUCTION AND AIMS

TDP2 function and its role on DNA repair has been studied using exogenous expression systems (Li *et al.*, 2011), shRNA depletion (Ledesma *et al.*, 2009), *TDP2*<sup>-/-</sup> chicken B cells (DT40) (Zeng *et al.*, 2011; 2012), *TDP2*<sup>-/-</sup> mouse embryonic fibroblasts (MEF) (Gómez-Herreros *et al.*, 2013) and patient-derived lymphoblastoid cells (Gómez-Herreros *et al.*, 2014). Therefore, to better study the molecular mechanisms of DSB repair, in addition to the effect of the absence of *TDP2* in human cells, we decided to create cell lines in which *TDP2* has been deleted (*TDP2*<sup>-/-</sup>). Cell lines created using this approach are advantageous, when comparing with the previously used systems, because they are derived from a single cell lineage. Moreover, using an established cell line as the base for investigation, allow us to account for specific cell characteristics, such as expression of particular genes.

Until recently, genome engineering techniques for mammalian systems consisted of employing zinc-finger nucleases (ZNFs) or transcription activator-like effector nucleases (TALENs). These enzymes consisted of designer DNA targeting domains coupled with the FokI nuclease to create DNA site specific breaks (Lander, 2016; Jiang and Doudna, 2017). This approach leads to insertion or deletions (indels) of base pairs if the DSB is processed by the NHEJ repair machinery, or gene targeting if repaired by homologous recombination (Hsu *et al.*, 2014). However, cost and difficulties with the process of protein design limited the utility of these genome engineering systems (Doudna and Charpentier, 2014). With the advent of clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated protein 9 (Cas9) technology, which was demonstrated to be part of an adaptive immune system in bacteria and archaea (Barrangou *et al.*, 2007), genome editing has become more affordable and less technical demanding.

The CRISPR-Cas9 system employs the nuclease Cas9 in conjunction with a guide CRISPR RNA (crRNA), containing a complementary target sequence to create site specific DSB (Jinek *et al.*, 2012). In addition, a trans-activating CRISPR RNA (tracrRNA) that is partially complementary to the crRNA is required to form the RNA secondary structure



that the Cas9 nuclease recognises and binds (Jiang and Doudna, 2017). Importantly, the tracrRNA and crRNA can be fused into a single guide RNA (sgRNA) without affecting DSB cutting efficiency (Jinek *et al.*, 2012). The Cas9 nuclease itself is formed by two endonuclease domains. The Histidine-Asparagine-Histidine (H-N-H) endonuclease domain and a RuvC-like domain, which is homologous to the bacterial RuvC endodeoxyribonuclease, shown to resolve Holliday junctions (Iwasaki *et al.*, 1991; Makarova *et al.*, 2006; Bolotin *et al.*, 2005). Both domains are required to generate the DSB and indeed mutations in these domains either ablate nuclease activity or result on SSBs instead (Jinek *et al.*, 2012). In addition, a short nucleotide sequence downstream of the target sequence, known as the proto-spacer adjacent motif (PAM), was shown to be essential for efficient nuclease target recognition and activity (Gasiunas *et al.*, 2012).

The first use of CRISPR-Cas9 mediated genome engineering in mammalian cells employed a system present in the bacterium *Streptococcus pyogenes* (spCas9), and was shown to work in multiple cell lines with the similar efficiency of TALENs (Mali *et al.*, 2013b; Jinek *et al.*, 2013; Cong *et al.*, 2013). From these pivotal studies many more publications have followed and now the CRISPR-Cas9 technique has evolved into a simple and affordable way to create genome modified cells and organisms.

### **3.1.1. Aim**

Considering the availability of the CRISPR-Cas9 technology, this chapter will focus on the generation and characterisation of human cell lines lacking TDP2. These cells will be characterised for the DSB response and cell survival following DNA damage, and then employed to extend our understanding of the role of TDP2 in the maintenance of genome stability.

## **3.2. RESULTS**

### **3.2.1. TDP2 gRNA design and CRISPR-Cas9 system**

To create a gene knock-out using the CRISPR-Cas9 system, I employed the e-CRISP tool to design appropriate gRNA sequences (Heigwer *et al.*, 2014). I initially

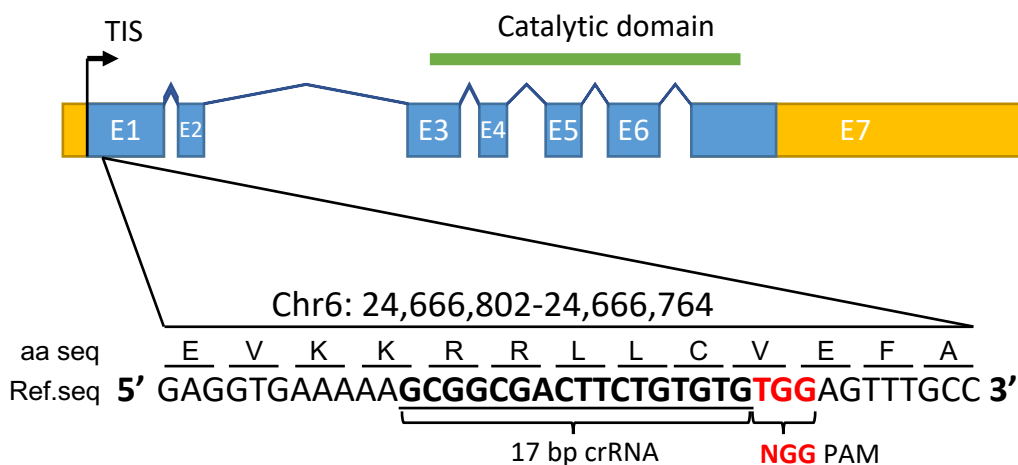
chose a gRNA, targeted against exon 1 of *TDP2*, with the sequence 5' GC GGC GAC TTC TGT GTG 3' (Figure 3.1A). We chose to target exon 1 because the TDP2 catalytic domain is encoded by exons 3-7, and thus any indel created at this site and producing a sequence frame-shift should encode a truncated protein lacking phosphodiesterase activity and, most likely, a mRNA transcript that will be subject to nonsense mediated decay (Figure 3.1.A). The design of the gRNA also took into account that shorter gRNAs (17-19 nucleotides; known as "Tru-guides") reduce the off-target effects of the Cas9 nuclease, while not affecting gene editing efficiency (Fu *et al.*, 2014). The rationale being that on a shorter RNA, fewer mismatches are tolerated before the disruption of RNA:DNA interaction, therefore reducing possible off-target cutting (Fu *et al.*, 2014). We also opted to employed gRNAs adjacent to a PAM of the sequence NGG, rather than NAG, since the former was reported to result in a higher Cas9 cutting efficiency (Gasiunas *et al.*, 2012). Taking these criteria into account, we targeted the sequence shown in Figure 3.1.B. This sequence was predicted to be highly specific to *TDP2* coding sequence, showing off-target sites only in non-coding genomic sequences and intronic regions of coding sequences (data not shown).

### 3.2.2. CRISPR-Cas9 *TDP2*<sup>-/-</sup> clone screening

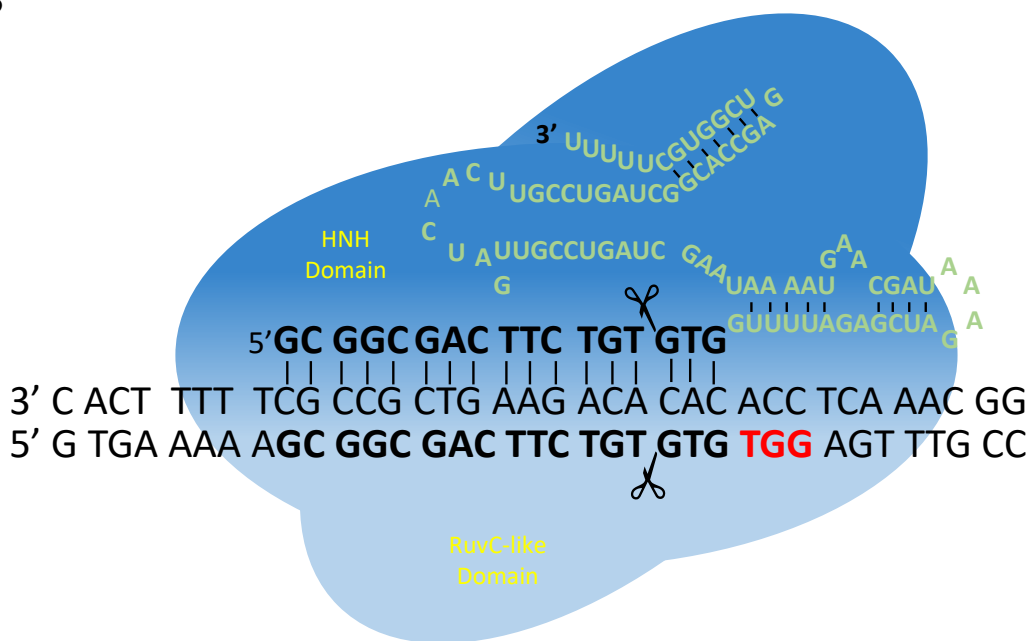
As mentioned above, multiple labs pioneered approaches for genome editing in cell lines using the CRISPR-Cas9 system (Mali *et al.*, 2013b; Jinek *et al.*, 2013; Cong *et al.*, 2013). We opted to follow the method described by the Church Lab (Mali *et al.*, 2013a). This protocol employs a codon-optimised spCas9 for expression in human cells (hCas9), encoded by a vector that also harbours the selection marker Geneticin (G418), allowing for antibiotic selection and enrichment of transfected cells. In addition, the gRNA vector employed in this protocol involves cloning of a single guide RNA, instead of independent crRNA and tracrRNA sequences (Figure 3.2).

I chose for my experiments several different cell lines. I chose hTERT-immortalised diploid Retinal Pigmented Epithelium-1 (RPE-1) cells (Bodnar *et al.*, 1998; Rambhatla *et al.*, 2002; Kuznetsova, Kurinov and Aleksandrova, 2014) as a representative of 'normal' non-transformed, checkpoint proficient cells with a stable diploid karyotype (RPE-1), and both breast cancer-derived Centre Antoine Lacassagne-

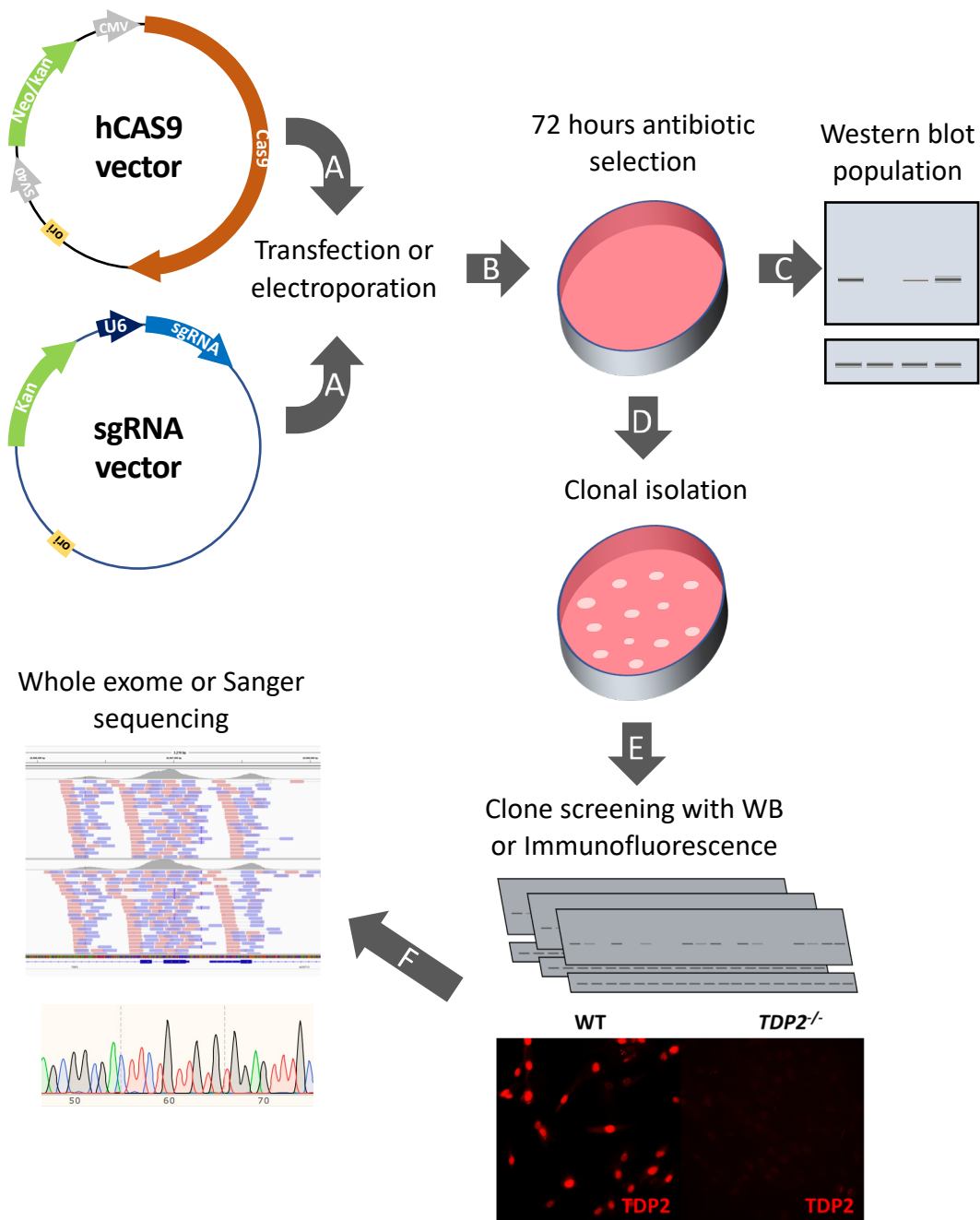
A



B



**Figure 3.1 – *TDP2* locus and spCas9 mediated genome editing** - (A) Schematic of *TDP2* gene structure showing exons and introns. Untranslated regions are indicated in yellow, translation initiation site (TIS) is indicated by directional arrow. The catalytic domain region containing phosphodiesterase activity spanning from exon 3 to 7 is indicated by green bar. Genome editing target of 17 base pairs (bp) on exon 1 is shown in bold, followed by PAM sequence, highlighted in red. Genomic DNA sequence depicted is inverted as *TDP2* is transcribed from anti-sense strand. (B) Model of the spCas9 binding target sequence on *TDP2* locus. Nuclease domains of spCas9 are indicated and also the secondary structure of the sgRNA formed by the fused tracrRNA (sequence in green) and crRNA. The PAM sequence in the complementary genomic sequence is highlighted in red. Scissors depict site of DSB formation (PAM -3bp).



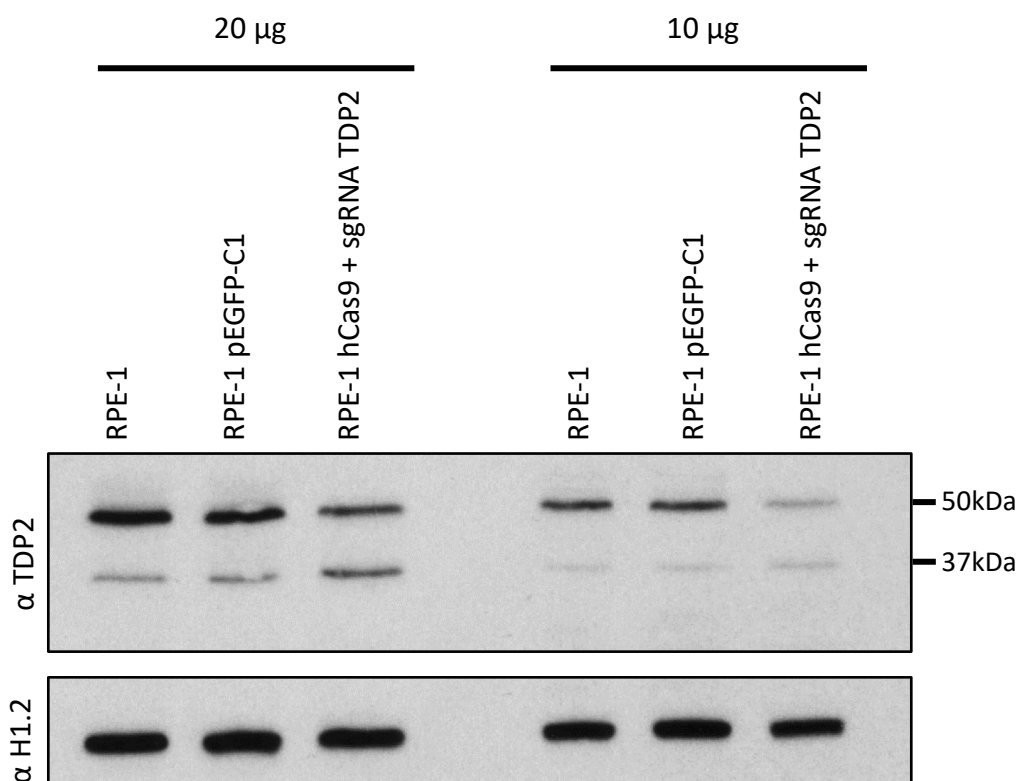
**Figure 3.2 – Method for the generation of CRISPR-Cas9 gene edited cell lines - (A)** Following the protocol described by Mali et al 2013, cells were transfected with the hCas9 and gRNA vectors, **(B)** before plating for antibiotic selection using geneticin. **(C)** After selection the pool of cells were assayed for TDP2 protein levels by western blotting and **(D)** plated for single clonal isolates. **(E)** Once visible colonies formed, single colonies were picked and assayed for TDP2 expression using western blotting (WB) or immunofluorescence with antibody against TDP2. **(F)** Single clones in which TDP2 was absent by western blotting were then sequenced using whole exome or sanger sequencing to further verify the identity of the TDP2 mutation.

51 (CAL-51) (Gioanni *et al.*, 1990) and U2OS human osteosarcoma cell line as representatives of transformed cancer cells.

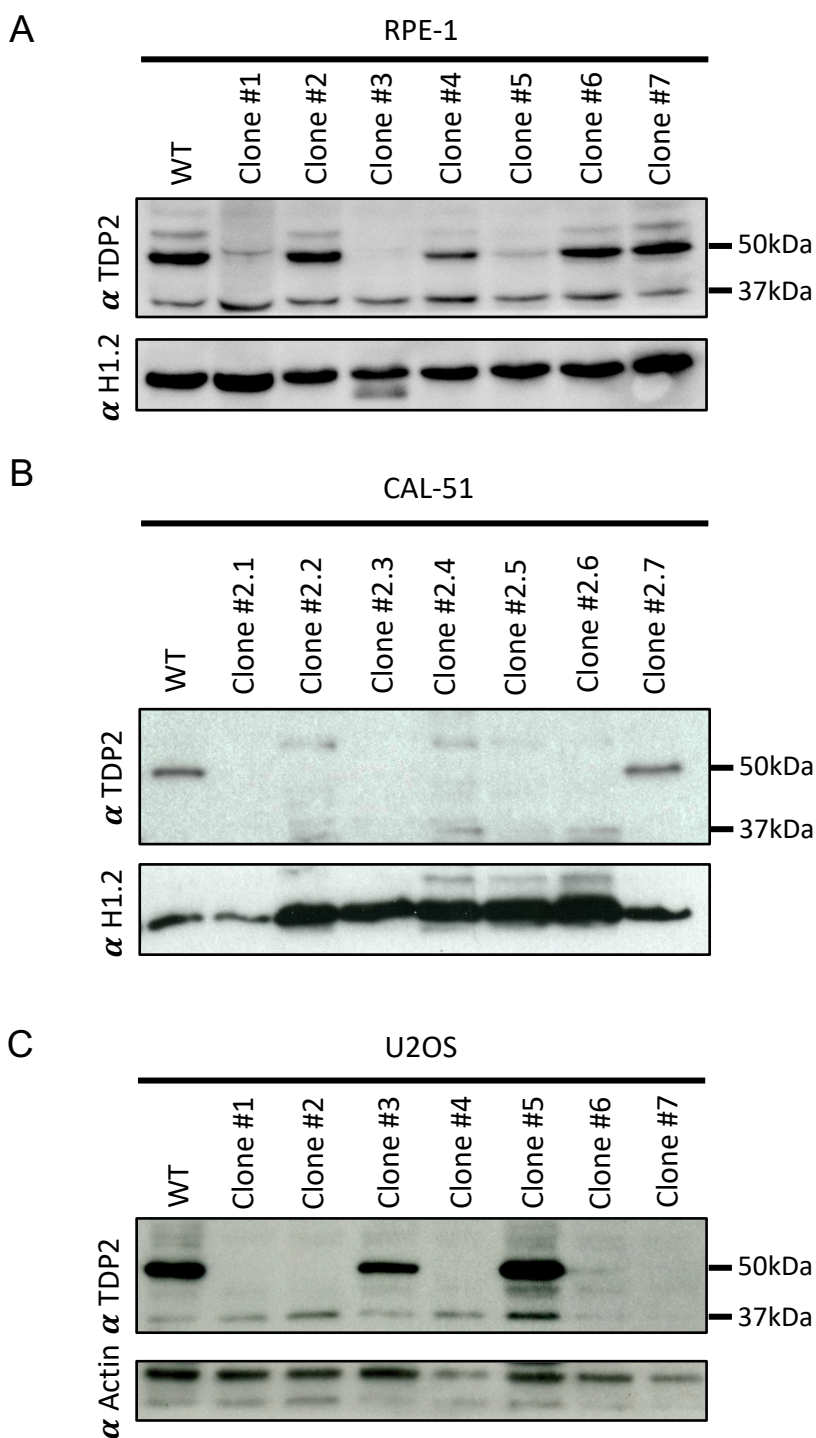
Taking the approach depicted in Figure 3.2, we first electroporated the hCas9 vector and sgRNA construct encoding the hTDP2 guide RNA sequence into RPE-1 cells using the Neon Transfection system. For the CAL-51 and U2OS cell lines, transfection using the GeneJuice (cationic polyamine) was used instead. Following transfection, cells were incubated in medium containing antibiotic selection for 72 hours to enrich for transfected cells. After selection, for RPE-1 cells, TDP2 protein levels were examined in the pooled cell population by Western blotting as an initial assessment of gene editing efficiency (Figure 3.3). Diminished TDP2 protein levels were detected in cells that were transfected with both hCas9 and sgRNA constructs, but not in those transfected with pEGFP-C1 (which also encodes G418 resistance) or with non-transfected RPE-1 cells as negative controls. This reduction suggested that gene editing was successful, and therefore cells were plated at low density to allow colonies to develop from individual single cells. Once colonies of single clonal isolates were large enough to isolate and expand, TDP2 protein levels were measured by western blotting as described above. Multiple clones were detected in which TDP2 was reduced or absent on RPE-1 cells (Figure 3.4A), CAL-51 cells (Figure 3.4B), and U2OS cells (Figure 3.5C).

### 3.2.3. *TDP2*<sup>-/-</sup> characterisation through genome sequencing

To verify that the *TDP2* locus was successfully targeted, the genome of the putative *TDP2*<sup>-/-</sup> clones that we detected by immunoblotting was subject to sequencing. Using PCR followed by Sanger sequencing, a homozygous insertion c.85>+1T in Exon 1 of the putative *TDP2*<sup>-/-</sup> RPE-1 clone #3 was observed on this sample DNA chromatogram (Figure 3.5.A). The insertion was at the site predicted for Cas9 activity, three nucleotides upstream of the PAM, as depicted in Figure 3.1.B. Because only a single mutation was identified by Sanger sequencing of the PCR amplified *TDP2* region, whole exome sequencing (WES) was conducted. Interestingly, in addition to the c.85>+1T insertion, WES identified a large insertion of at least 43 nucleotides in *TDP2* (Figure 3.5.B and Table 3.1), suggesting that both of the *TDP2* alleles in RPE-1 clone #3 cells were mutated. The large insertion was present at the Cas9 predicted cut site, however because the WES

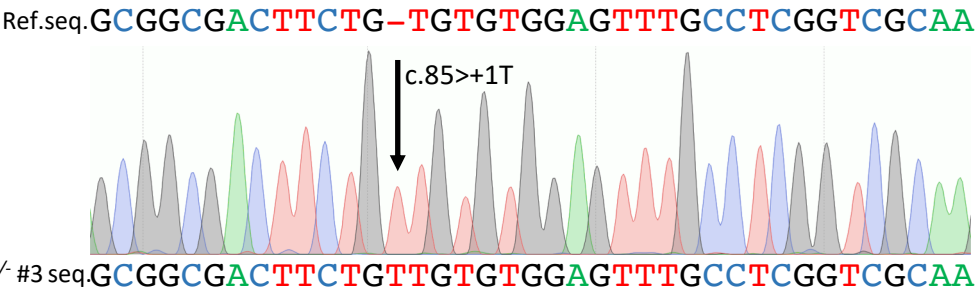


**Figure 3.3 – TDP2 levels in RPE-1 cells transfected with hCas9 and gRNA vectors** . TDP2 and Histone1.2 (loading control) protein levels in 20  $\mu\text{g}$  (left) or 10  $\mu\text{g}$  (right) of whole cell extract protein from pooled cell populations were examined by western blotting, after transfection with either eGFP vector as a negative control or both hCas9 and gRNA vectors targeting TDP2. Transfected populations were enriched by incubating with G418 at 0.5  $\mu\text{g}/\text{mL}$  for five days. Untransfected RPE-1 cells were also examined by western blotting as an additional negative control. Anti-TDP2 antibody (polyclonal) was raised on rabbit against purified full-length TDP2 protein and affinity purified (Thomson *et al.* 2013).



**Figure 3.4. TDP2 protein levels in single clonal isolates of CRISPR-Cas9 gene edited cells.** Representative anti-TDP2 western blots of cell extract from RPE-1 (**A**), CAL-51 (**B**) or U2OS (**C**) clones. Anti-histone1.2 (H1.2) or anti-actin shown as loading control. Anti-TDP2 antibody (polyclonal) was raised on rabbit against purified full-length TDP2 protein and affinity purified (Thomson *et al.* 2013).

A. RPE-1 Sanger Sequencing



B. RPE-1 WES

RPE-1 WT seq	GCGGCGACTTCTG-TGTGTGGAGTTTGCCTCGGTCG
RPE-1 <i>TDP2</i> <sup>-/-</sup> #3 seq#1	GCGGCGACTTCTGT <b>T</b> TGTGTGGAGTTTGCCTCGGTCG
RPE-1 <i>TDP2</i> <sup>-/-</sup> #3 seq#2	GCGGCGACTTCTG-TGTGTGGAGTTTGCCTCGGTCG
RPE-1 <i>TDP2</i> <sup>-/-</sup> #3 seq#3	... <b>TGAGTTTGTATA-C</b> GTGTGGAGTTTGCCTCGGTCG

C. CAL-51 WES

CAL-51 WT seq	GCGGCGACTTCTG-TGTGTGGAGTTTGCCTCGGTCG
CAL-51 <i>TDP2</i> <sup>-/-</sup> #2.6 seq#1	GCGGCGACTTCTGT <b>T</b> TGTGTGGAGTTTGCCTCGGTCG
CAL-51 <i>TDP2</i> <sup>-/-</sup> #2.6 seq#2	GCGGCGACTTCTG- <b>-----</b> TGCCTCGGTCG

D.

Cell Line	Mutation Identified	Protein effect	Likely variant status	TDP2 copy number	Sequencing Method
RPE-1 <i>TDP2</i> <sup>-/-</sup> #3	c.85> +1T	p.75*	Homozygous	2	PCR/Sanger
RPE-1 <i>TDP2</i> <sup>-/-</sup> #3	c.85> +1T	p.75*	Heterozygous	2	WES
CAL-51 <i>TDP2</i> <sup>-/-</sup> #2.6	c.85> +1T	p.75*	Heterozygous	2	WES
CAL-51 <i>TDP2</i> <sup>-/-</sup> #2.6	c.85> - 11TGTGTGGAGTT	p.71*	Heterozygous	2	WES

**Figure 3.5 – CRISPR-Cas9 insertions and deletions at the site of DSB formation in *TDP2*** - **(A)** Representative Sanger sequencing chromatogram of PCR product, amplified using primers for *TDP2* exon 1, on genomic DNA extracted from RPE-1, reference sequence (ref. seq.) or RPE-1 *TDP2*<sup>-/-</sup> clone #3 **(B)** Whole exome sequencing (WES) of RPE-1 wild type and RPE-1 *TDP2*<sup>-/-</sup> clone #3 cell lines. Each line represents mutations identified. Insertions are highlighted in red **(C)** WES of CAL-51 wild type CAL-51 CAL-51 *TDP2*<sup>-/-</sup> clone #2.6 cell line. Each line represents mutations identified. Insertions or deletions are highlighted in red. **(D)** Summary of indels identified using Sanger and whole exome sequencing and their predicted protein effect. TDP2 copy number was obtained from CanSAR database (accessible at cansar.icr.ac.uk).



					RPE-1				TDP2 <sup>-/-</sup> RPE-1 clone #3			
Genomic Location *	Ref. Seq.	Variant sequence	Gene	SNP	Read depth	Read with variant	Mutated (%)	Pre- dicted status	Read depth	Read with variant	Mutated (%)	Pre- dicted status
Chr6: 24476553	G	A	GPLD1	rs760904	38	38	100	Homozygous	57	36	63.16	Heterozygous
Chr6: 24667005	C	CA	TDP2	-	0	0	0	-	165	36	21.82	Heterozygous
Chr6: 24667005	C	CGTATTACAA ACTCACTGCC GACTGGAAA GCGGGCAGT GAGCGC	TDP2	-	0	0	0	-	165	13	7.88	Heterozygous
Chr6: 24667005	C	C	TDP2	-	223	223	100	Homozygous	165	116	70.30	Heterozygous
Chr6: 28198281	A	G	ZSCAN9	rs1233713	111	111	100	Homozygous	136	101	74.26	Heterozygous
Chr6: 29408528	A	G	ORIOC1	rs2074464	177	177	100	Homozygous	229	155	67.69	Heterozygous

**Table 3.1 Table of sequencing variants of RPE-1 and TDP2<sup>-/-</sup> RPE-1 clone #3 cells following whole exome sequencing –** \*Human genome reference sequence and sequenced variants are annotated according to human genome building 19. Variant sequence highlighted in red is homologous to sequence present on hCas9 vector.

protocol uses primers that only capture human exon sequence and sequences only 75bp long fragments (Biesecker and Green, 2014), the whole sequence and length of this insertion could not be resolved. Worryingly, however, WES also identified wild-type *TDP2* allele sequence in over 70% (116 out of 165) of the sequencing reads, whereas the large insertion was only present in 7.88% of reads (Table 3.1).

Because the Sanger and WES sequencing data identified two mutant alleles, and because RPE-1 cells are diploid, it was considered likely that RPE-1 *TDP2*<sup>-/-</sup> clone #3 was *TDP2* null, and that the wild type alleles arose from DNA contamination. Sample contamination on high-throughput sequencing platforms has been linked to poor indexing on pooled samples and to account for up to 3% of data obtained (Vodák *et al.*, 2018; Costello *et al.*, 2018). To try to resolve whether our data was contaminated with foreign sequences, the nearby single nucleotide polymorphisms (SNP) rs760904, rs1233713 and rs2074464 were examined (Table 3.1). These sites were chosen as they are close to *TDP2*'s genome locus and are all homozygous on the RPE-1 cell line, as per data obtained in cells sequenced in parallel with the *TDP2*<sup>-/-</sup> cells, and should therefore be so in the gene edited clones. Interestingly, all three SNPs were found to be heterozygous in the RPE-1 *TDP2*<sup>-/-</sup> clone #3, suggesting that exogenous, contaminating, genomic DNA accounted for ~25% of the sequence reads. To rule out that the large insertion was accounted for by the contaminant DNA, the inserted sequence was examined using NCBI's basic local alignment search tool (BLAST) (Altschul *et al.*, 1990). This analysis revealed that the large inserted sequence was cloning vector sequences (data not shown). Moreover, sequence alignment of the large insertion sequence revealed that 26 out of the inserted 43 nucleotides aligned to hCas9 vector, highlighted in red on Table 3.1. It was thus concluded that the wild-type sequence detected by WES was accounted for by the contaminant DNA, and that the large insertion accounted for the second gene targeted *TDP2* allele.

Whole exome sequencing was also conducted on the putative *TDP2* gene-edited CAL-51 cell clone #2.6 in which *TDP2* was deleted (Figure 3.5.C). These data identified two independent mutations in *TDP2*; c.85>+1T, the same insertion as detected in the *TDP2*<sup>-/-</sup> RPE-1 clone #3, and the large deletion c.85>-11TGTGTGGAGTT. Both of these create a premature stop codon (p.75\* and p71\* respectively; Figure 3.5.D) and are

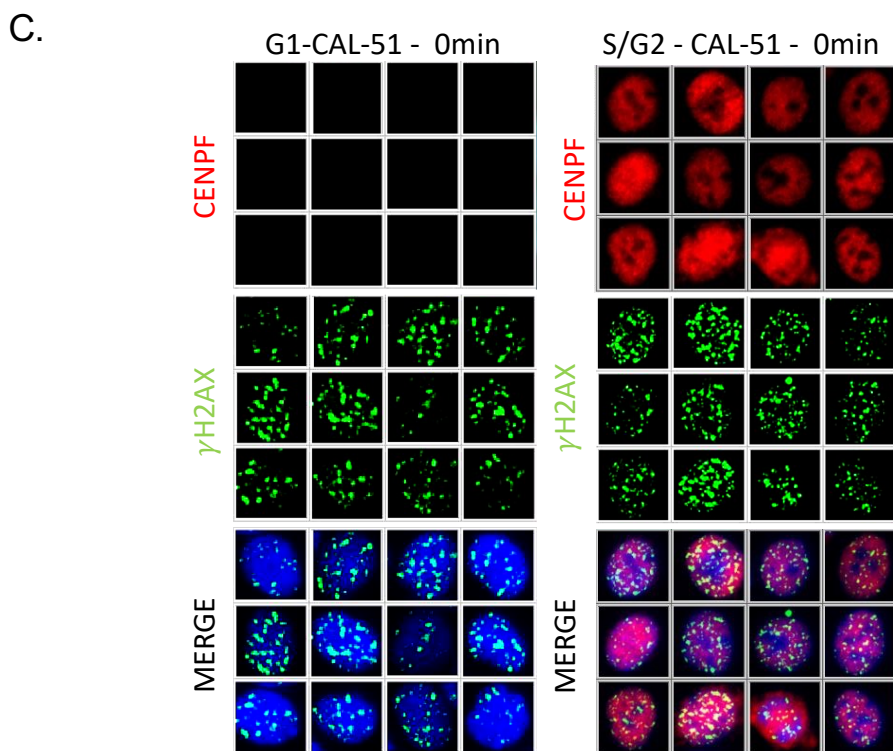
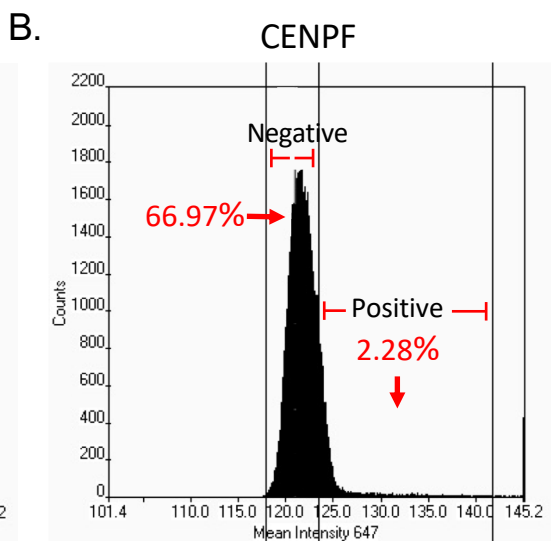
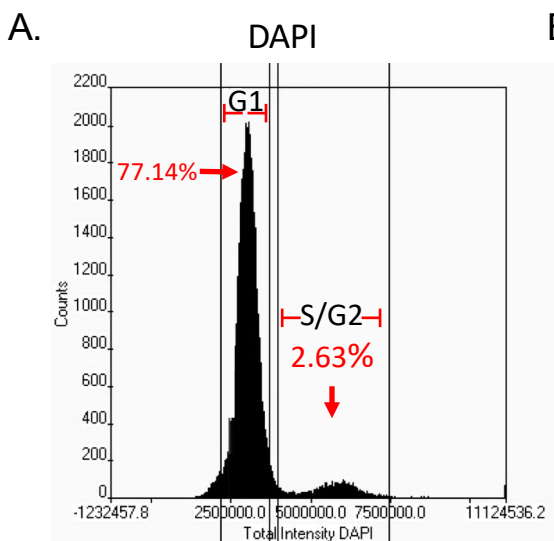
predicted to result in nonsense mediated decay of the *TDP2* mRNA, consistent with the absence of detectable protein by western blotting.

#### 3.2.4. Reduced rates of DSB repair in the absence of TDP2

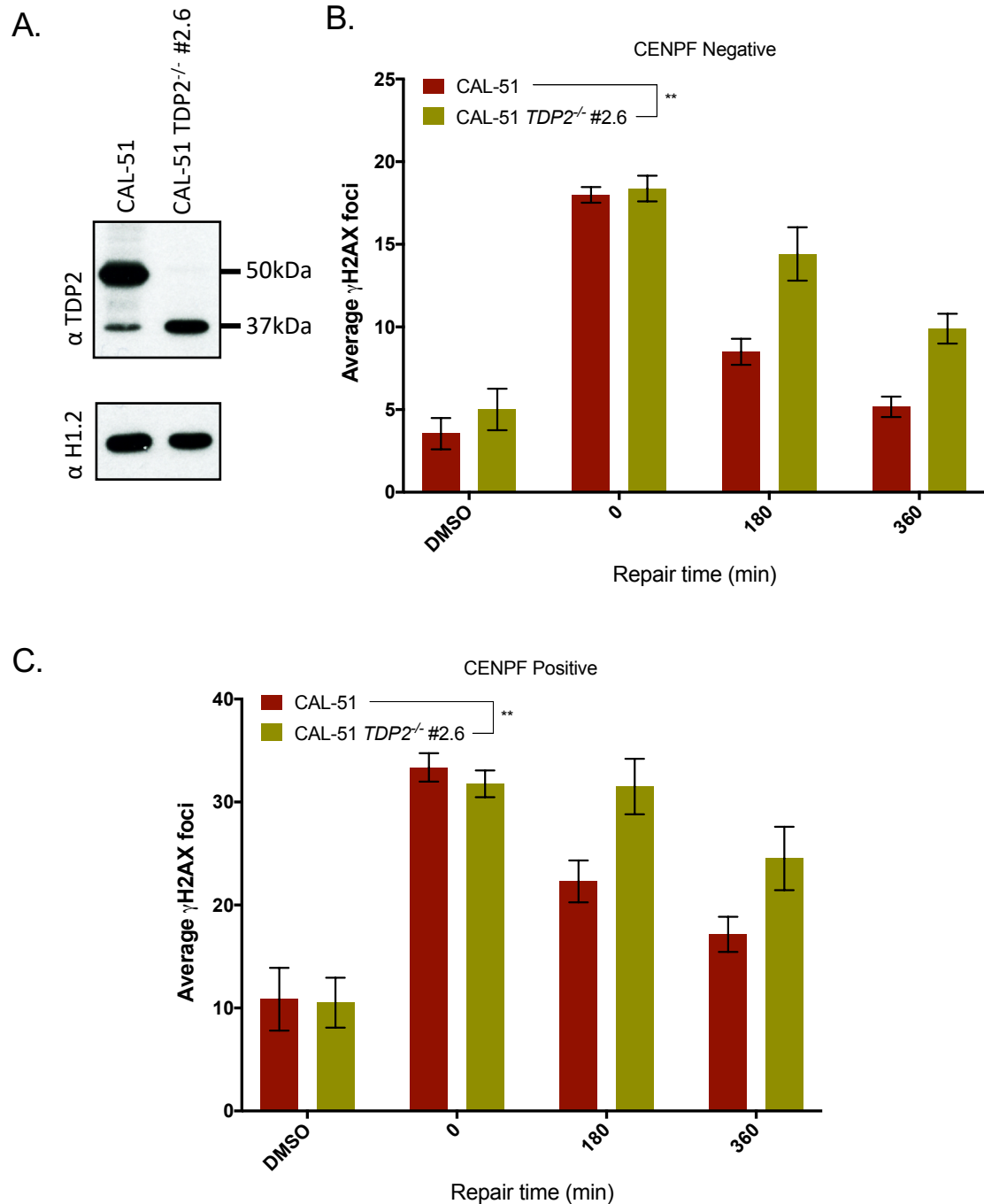
Next, the *TDP2*<sup>-/-</sup> clones were examined for defects in DSB repair following treatment with the topoisomerase “poison”, etoposide. Phosphorylation of histone H2A.X ( $\gamma$ H2AX) on Ser139 was employed as an indirect marker of DSB (Rogakou *et al.*, 1998; 1999). Since the number of DSBs formed and the mechanisms by which they are repaired is affected by cell cycle stage,  $\gamma$ H2AX foci were quantified in cells in specific phases of the cell cycle.

DSB measurements were conducted using the high-throughput microscope Olympus IX71 ScanR system, which has automated image acquisition. The ScanR analysis package used for image processing allows for population gating, analysis of cell morphology, automated foci counting and signal intensity quantification. Therefore, we employed this feature to create a cell cycle profile using the signal intensity of the nuclear counterstain DAPI (Figure 3.6.A), in conjunction with CENPF signal intensity (Figure 3.6.B) to gate cells accordingly to cell cycle phase. CENPF is known to accumulate in early S through G2 phase, but is absent in G1 phase (Liao *et al.*, 1995). This approach, allowed delimitation and quantification of  $\gamma$ H2AX foci in a specific cell cycle phases in etoposide-treated cells (Figure 3.6.C). It was noted that a higher number of DSB were observed in S/G2 gated cells than those gated in G1, consistent with the roles of TOP2 in replication and with the greater amount of DNA in these cells (Figure 3.6.C).

DSB repair rates were first measured in the *TDP2*<sup>-/-</sup> CAL-51 cell clone number #2.6, which I confirmed again lacked detectable TDP2 protein by western blotting (Figure 3.7.A). This clone showed significantly delayed DSB repair following three and six hours recovery in drug-free media, after etoposide treatment, when compared to CAL-51 wild-type cells (Figure 3.7.B). As this defect was present in G1 gated cells, it suggests that TDP2 is required to repair TOP2 induced cleavage complexes (TOP2cc) arising during transcription. This observation is consistent with the similarly reduced rate of DSB repair reported previously *Tdp2* <sup>$\Delta$ 1-3</sup> in mouse astrocytes and confluency arrested *Tdp2* <sup>$\Delta$ 1-3</sup> primary MEFs (Gómez-Herreros *et al.*, 2014; 2013).



**Figure 3.6 – Gating strategy for  $\gamma$ H2AX quantification by Olympus ScanR high content imaging** - Histograms extracted from ScanR Olympus Analysis Software showing **(A)** Cell cycle profile according to total DAPI signal and **(B)** mean signal intensity of CENPF (Alexa-647) in cells in an asynchronous population of CAL-51 cells **(C)** Representative images of CENPF negative (left) and CENPF positive (right) CAL-51 cells following 30 minutes incubation with 20  $\mu$ M etoposide. Cells were labelled with anti-CENPF (Alexa-647) and anti- $\gamma$ H2AX (Alexa-488). DAPI was used as nuclear counterstain. Cells were gated according to DAPI profile (as in A) and CENPF signal (as in B).



**Figure 3.7 – TDP2 absence impairs DSB repair rates independently of cell cycle -**  
**(A)** Western blot against TDP2 in CAL-51 and in the CAL-51 *TDP2*<sup>-/-</sup> clone #2.6. Anti-histone 1.2 (H1.2) shown as loading control. **(B)** Average  $\gamma$ H2AX foci was measured in CAL-51 and the CAL51 *TDP2*<sup>-/-</sup> clone #2.6 cells, following 30 minutes treatment with 20  $\mu$ M etoposide, and recovery in drug-free media for the stated times. DMSO used as vehicle control. The cell population was gated according to CENPF signal and DAPI profile. Only CENPF negative and G1 cells were counted. **(C)** As is (B), however only cells that were positive for CENPF and S/G2 were counted. All data are the mean  $\pm$  SEM of three independent experiments. Statistical significance was determined by two-way ANOVA on GraphPad Prism Software. \*\* $p < 0.01$ .

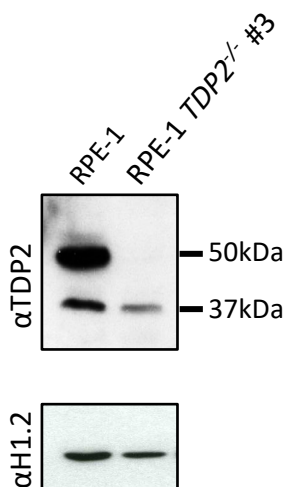
In contrast to G1 cells, S/G2 cells have available to them alternative pathways for DSB repair other than NHEJ, raising the possibility that TDP2 is not required during these cell cycle phases. Indeed, sister chromatid exchanges are elevated following etoposide treatment in *Tdp2*<sup>Δ1-3</sup> primary MEFs, suggesting that homologous recombination mediated repair is operative at a greater level in the absence of TDP2 (Gómez-Herreros *et al.*, 2013). Similarly, Hoa and colleagues reported that MRE11 nuclease is the primary mechanism for removing TOP2 induced DSBs in human lymphoblastoid TK6 cells, even in the presence of TDP2 (Hoa *et al.*, 2016). Therefore, DSB repair rates were measured in *TDP2*<sup>-/-</sup> CAL-51 cells in S/G2 cells. Notably, it was observed that DSB repair was significantly delayed (Figure 3.7.C), suggesting that in CAL-51 breast cancer cells TDP2 is required for the repair of TOP2 induced DSBs in both G1 and S/G2 phases of the cell cycle.

To extend our analyses to non-transformed diploid cells, DSB repair rates were assessed in *TDP2*<sup>-/-</sup> RPE-1. The clone employed was clone number #3, characterised by WES, which was again confirmed to lack detectable TDP2 protein by western blotting (Figure 3.8.A). Once again, rates of DSB repair were delayed in these cells following etoposide treatment, with significantly more foci observed at three and six hours post treatment (Figure 3.8.B).

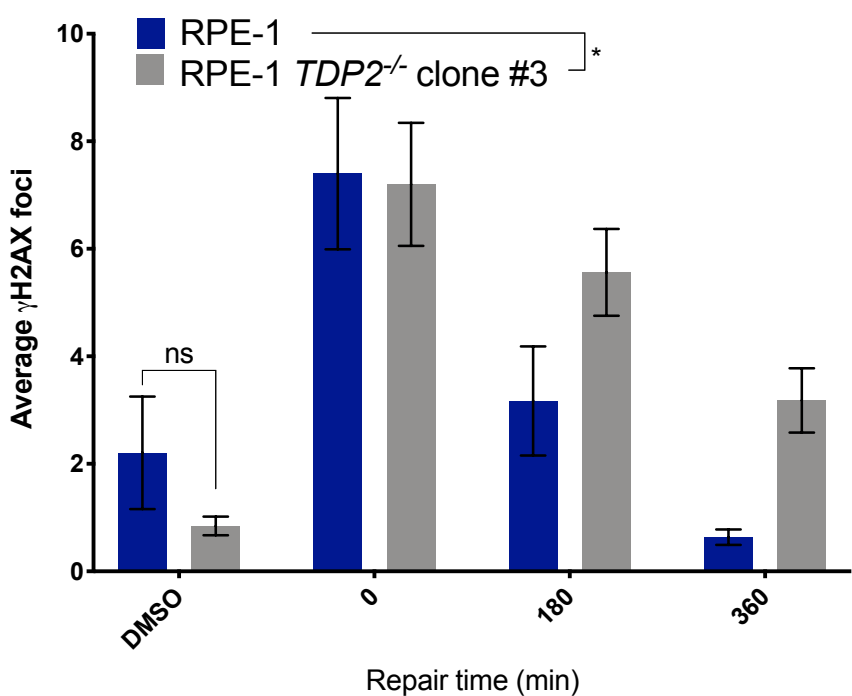
### 3.2.5. *TDP2*<sup>-/-</sup> cells are hypersensitive to TOP2-induced DSBs

Following the delayed DSB repair response in *TDP2*<sup>-/-</sup> cells, it was investigated whether this defect resulted in decreased cell survival following treatment with etoposide. Indeed, by using clonogenic survival assays it was shown that *TDP2*<sup>-/-</sup> RPE-1 clone #3 cells are hypersensitive to a short one hour incubation (Figure 3.9.A). Since chemotherapeutic drugs such as etoposide, are usually administered continuously it was also assessed whether *TDP2*<sup>-/-</sup> cells were hypersensitive to continuous exposure to etoposide. Indeed, continuous treatment with etoposide increased the difference in cell killing between wild type and *TDP2*<sup>-/-</sup> cells, with *TDP2*<sup>-/-</sup> cells being markedly more sensitive (Figure 3.9.B). Next, I confirmed that this hypersensitivity observed in RPE-1 *TDP2*<sup>-/-</sup> cells following chronic exposure to etoposide was also observed in other cell

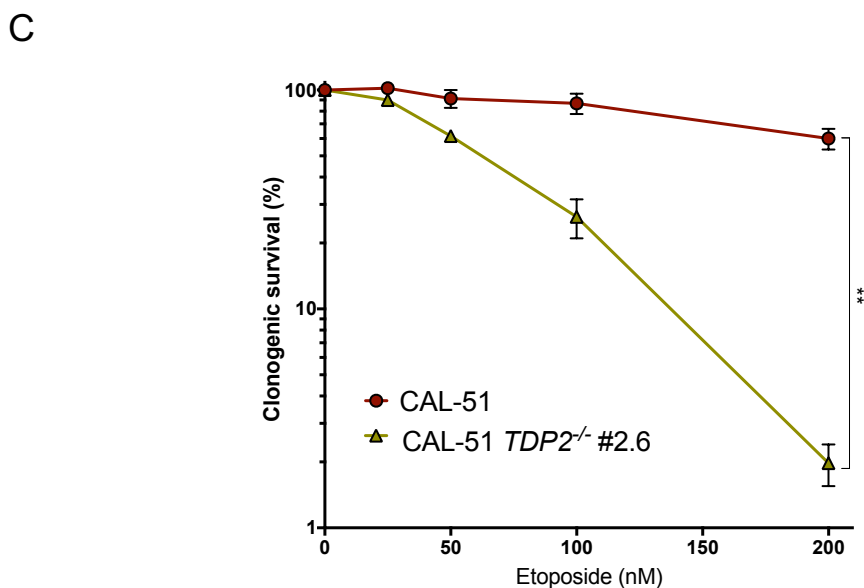
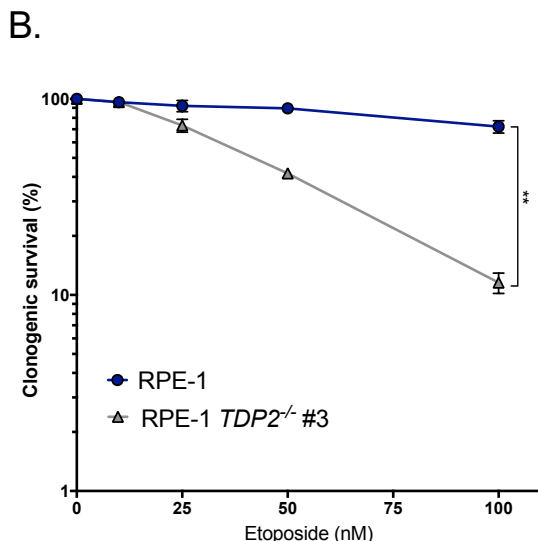
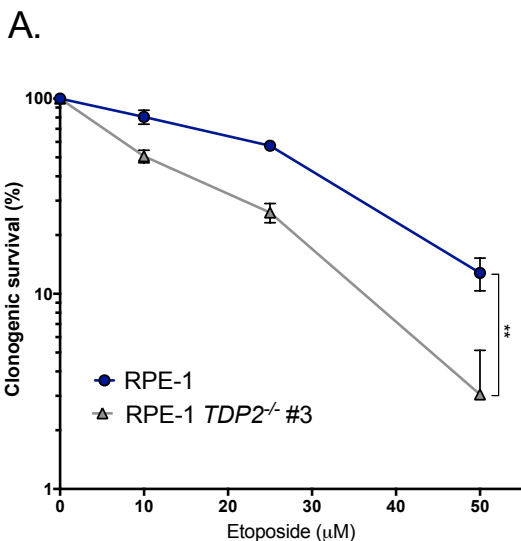
A.



B.



**Figure 3.8 – TDP2 absence impairs DSB repair rates following etoposide treatment**  
- **(A)** Western blot showing TDP2 levels in RPE-1 and RPE1 *TDP2*<sup>-/-</sup> clone #3. Anti-histone 1.2 (H1.2) is shown as a loading control. **(B)**. Average  $\gamma$ H2AX foci count in RPE-1 and RPE1 *TDP2*<sup>-/-</sup> at indicated times following 30 minutes treatment with 20  $\mu$ M etoposide. DMSO used as vehicle control. Population was gated according to CENPF signal and DAPI profile. Only CENPF negative cells (G1) were counted. Data are the mean  $\pm$  SEM of three independent experiments. Statistical significance between cell lines was determined by a two-way ANOVA on GraphPad Prism Software. At DMSO treated statistical significance was given by paired t-Test. ns= not significant, \*p<0.05.



**Figure 3.9 – Loss of TDP2 leads to hypersensitivity to etoposide in clonogenic survival assays-** (A) Clonogenic survival of RPE-1 and RPE-1 *TDP2*<sup>-/-</sup> Exon clone #3, following one hour exposure with the stated doses of etoposide. After treatment cells were washed with PBS and cultivated in drug-free media until colonies were formed (Figure published in Gomez-Herreros et al. 2017)(B) Clonogenic survival as in (A), with stated doses of etoposide incubated continuously in cell media. Drug was added to media four hours after cell plating (C) As in (B), but employing the CAL-51 and CAL-51 *TDP2*<sup>-/-</sup> clone #2.6 cell lines. All data are the mean  $\pm$  SEM of three independent experiments. Statistical significance was determined by two-way ANOVA on GraphPad Prism Software, \*\* $p < 0.01$ .



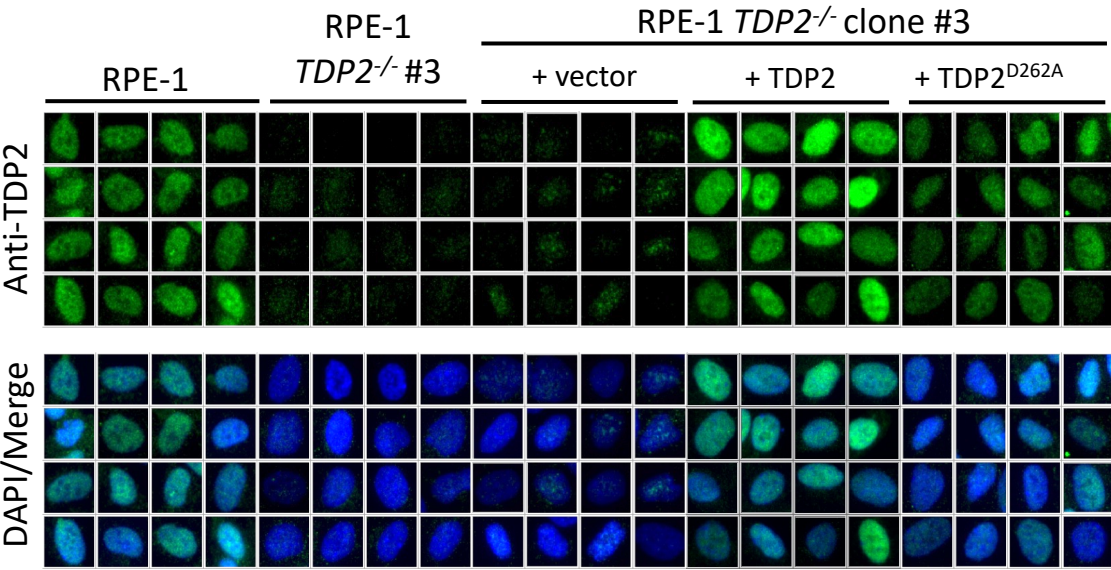
lines. Reassuringly, the CAL-51 *TDP2*<sup>-/-</sup> clone #2.6 cells have also shown hypersensitivity to low doses of etoposide (Figure 3.9.C).

Having demonstrated that the *TDP2*<sup>-/-</sup> cell lines were hypersensitive to etoposide and exhibited decreased DSB repair following treatment with this agent, next I attempted to demonstrate that these phenotypes were complemented by re-expression of exogenous human TDP2 protein. *TDP2*<sup>-/-</sup> RPE-1 cells were therefore transfected with either empty vector alone or with expression vector encoding either wild-type TDP2 or TDP2<sup>D262A</sup>, a mutant which has been shown previously to have no catalytic activity (Ledesma *et al.*, 2009; Gao *et al.*, 2012). Following antibiotic selection, multiple clones were isolated and screened for exogenous TDP2 expression using indirect immunofluorescence. Single clonal isolates were then grown and had their TDP2 expression levels re-accessed using immunofluorescence (Figure 3.10.A). First, using RPE-1 *TDP2*<sup>-/-</sup> clone #3, I confirmed that the signal detected following immunolabelling was specific to TDP2, as these cells presented little to no background (Figure 3.10.A). Consistently, expression was also not detected on cells transfected with vector only, whereas clones stably expressing either TDP2 or TDP2<sup>D262A</sup> showed a strong nuclear signal, similar to which was observed in the RPE-1 wild type cells (Figure 3.10.A). Additionally, using the ScanR high content imaging system, I quantified the signal from the exogenous expression of TDP2. I found that expression of wild type recombinant TDP2 was elevated in the transfectant to a level significantly higher than the endogenous TDP2 signal detected on RPE-1 cells (Figure 3.10.B). While TDP2<sup>D262A</sup> expression was found to be lower than that of endogenous TDP2 in RPE-1 cells but not significantly decreased. Importantly, the hypersensitivity to etoposide observed in *TDP2*<sup>-/-</sup> RPE-1 clone #3 cells was largely, but not entirely, complemented by TDP2 (Figure 3.11). This complementation reflected TDP2 activity because neither empty vector nor TDP2<sup>D262A</sup> complemented these defects.

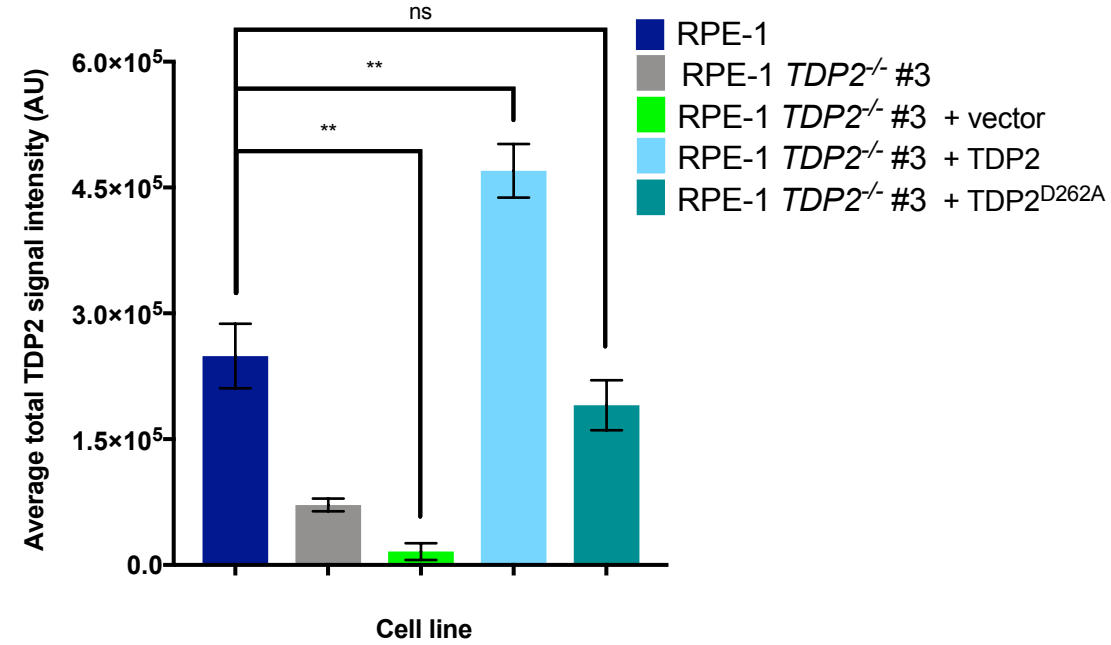
### 3.3. CONCLUSIONS AND DISCUSSION

This first chapter was focused on the generation and characterisation of CRISPR-Cas9 *TDP2*<sup>-/-</sup> in human cell lines. Using the approach described by Mali *et al.* 2013 coupled with usage of shorter sgRNA as suggested by Fu *et al.* 2014 we successfully

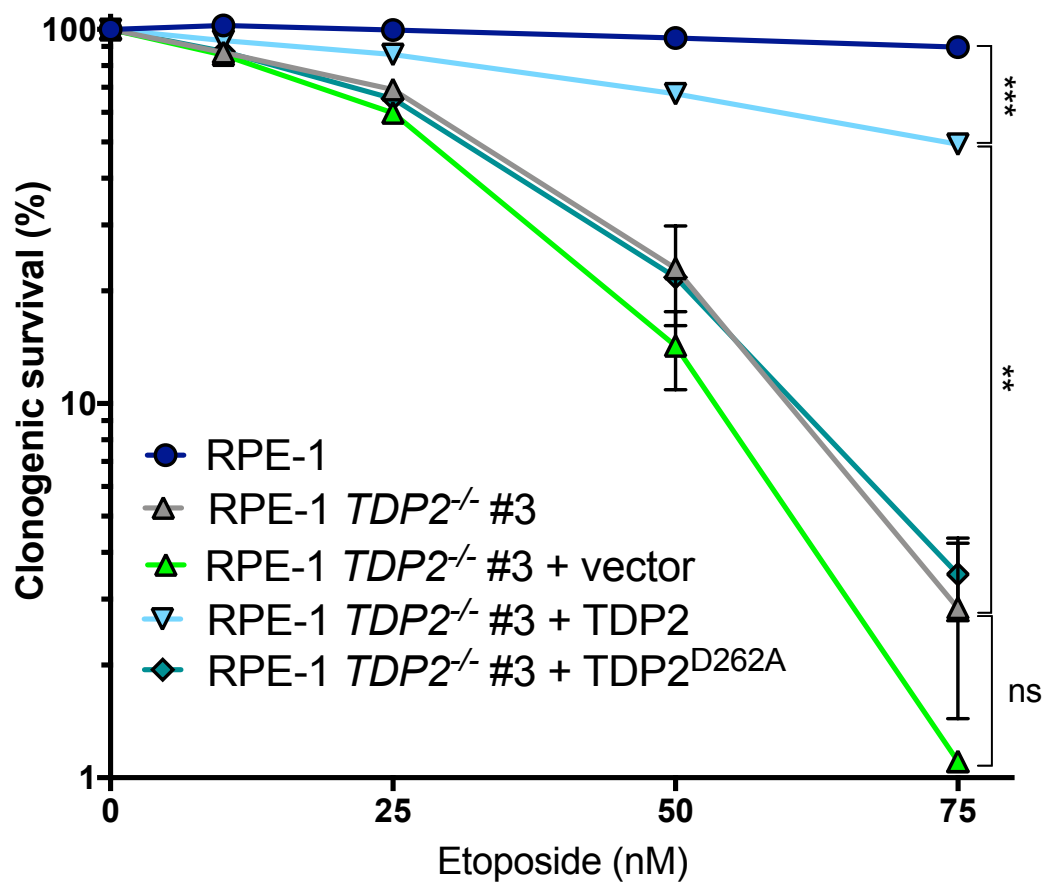
A.



B.



**Figure 3.10 – Complementation of *TDP2*<sup>-/-</sup> Exon 1 RPE-1 cells with recombinant human TDP2** - (A) Representative immunofluorescence image of TDP2 signal in RPE-1, RPE-1 *TDP2*<sup>-/-</sup> clone #3 and RPE-1 *TDP2*<sup>-/-</sup> clone #3 complemented with empty vector or expression vector encoding *TDP2* or *TDP2*<sup>D262A</sup> (catalytic dead). Cells were labelled with anti-TDP2 (Alexa-488), counterstained with DAPI and imaged on an Olympus ScanR microscope (B) Quantification of TDP2 fluorescence intensity in >1000 cells per sample. Data are the mean +/- SEM of four independent experiments. Statistical significance was determined by paired t-test on GraphPad Prism Software. ns=p>0.05 and \*\*p<0.01.



**Figure 3.11 – Hypersensitivity to etoposide is dependent on TDP2 phosphodiesterase activity - (A)** Clonogenic survival of RPE-1 and RPE-1 *TDP2*<sup>-/-</sup> clone #3 complemented with either vector (control), transgene for human *TDP2* or transgene for human *TDP2* with catalytic mutation (*TDP2*<sup>D262A</sup>). Data are the mean +/- SEM of three independent experiments. Statistical significance was determined by two-way ANOVA on GraphPad Prism Software. ns=p>0.05, \*\*p<0.01 and \*\*\*p<0.001.

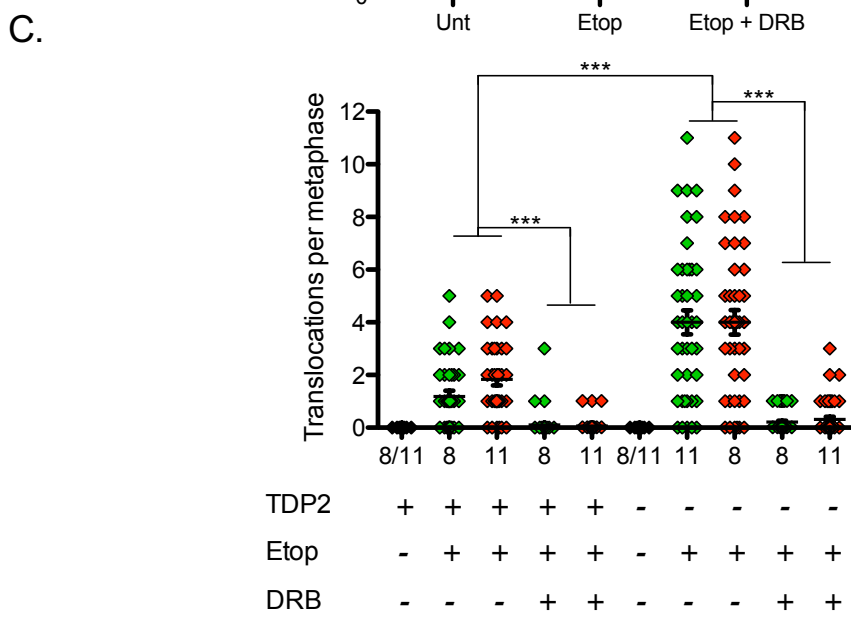
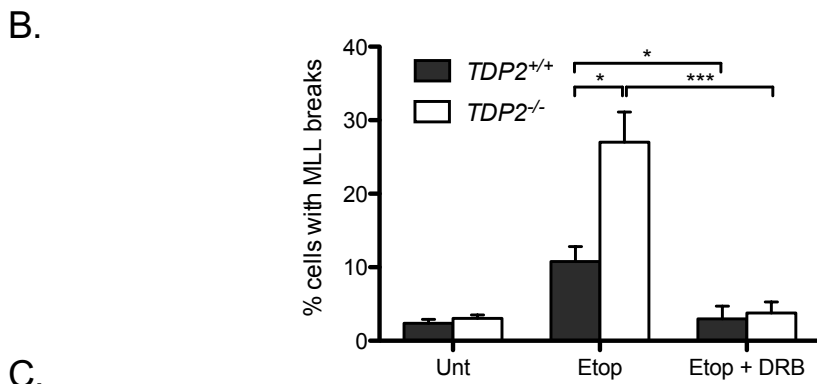
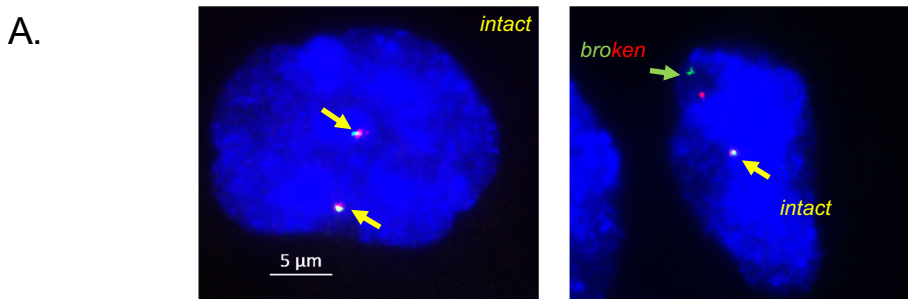
target *TDP2* exon 1 (Figure 3.1, Figure 3.2 and Figure 3.3). The high efficiency obtained with this approach was observed in multiple cell lines independently of genome copy number variation (Figure 3.4.A-C). This may indicate that *TDP2* deletion does not lead to extensive cell cycle checkpoint activation or cell death, at endogenous levels of TOP2-induced breakage at least (Zeng *et al.*, 2011; Gómez-Herreros *et al.*, 2013; 2014).

Following *TDP2*<sup>-/-</sup> clonal isolation, we confirmed gene editing using DNA sequencing. Using first Sanger sequencing of PCR products across *TDP2* exon 1, we identified only a single insertion, c.85>+1T, on RPE 1 *TDP2*<sup>-/-</sup> cells (Figure 3.5.A). This result was later found to be an artefact of PCR amplification as an additional large insertion was also found at same locus following whole exome sequencing (Table 3.1). Our WES highlighted both that contamination on new generation sequencing is an issue that has to be taken in consideration and revealed that the large insertion was found to be from the insertion of the hCas9 vector construct at the site of the Cas9 mediated DSB. Pooling of samples for high-throughput sequence could be the reason of our sample contamination and it has been previously observed (Vodák *et al.*, 2018; Costello *et al.*, 2018). Fortunately, haplotype analysis were sufficient for sample discrimination in our case (Table 3.1), however as NGS is frequently used in the clinical setting for diagnostics, our observations stresses how important it is to further develop of this technique. Additionally, our observation of a large genomic insertion, supports a recent report that CRISPR-Cas9 genomic alterations can induce more than small insertions and deletions (Kosicki *et al.*, 2018).

Our CRISPR-Cas9 *TDP2*<sup>-/-</sup> clones were sensitive to acute or chronic treatment with the topoisomerase ‘poison’ etoposide as expected and observed previously in cells in which *TDP2* was deleted or depleted by other technologies (Figure 3.9.A-C) (Ledesma *et al.*, 2009; Zeng *et al.*, 2011; Gómez-Herreros *et al.*, 2013; 2014; 2017). They also displayed delayed DSB repair kinetics following acute etoposide treatment independently of cell line or cell cycle stage, as delimited by DAPI profile and CENPF signal (Figure 3.6.A-C, Figure 3.7.A-B and Figure 3.8.B). These results indicate that *TDP2* function is required for NHEJ repair both in G1 and S/G2 and is also supported by observation in mouse cells the deletion of *TDP2* is epistatic with Ku70 deletion (Gómez-Herreros *et al.*, 2013). Our data also suggests that *TDP2* catalytic activity is essential for cell survival following etoposide treatment, because complementation of *TDP2*<sup>-/-</sup> cells

with catalytic mutant was not sufficient to rescue cells even though mutant expression was comparable to endogenous TDP2 levels (Figure 3.10.A-B and Figure 3.11). Interestingly, even overexpression of the wild-type TDP2 was not able to fully rescue the cell line sensitivity to etoposide (Figure 3.10.A-B and Figure 3.11). This could reflect genetic drift, and clonal variation in sensitivities in wild type cells (Ben-David *et al.*, 2018), or possibly that overexpression of TDP2 is suboptimal for cellular resistance to TOP2-induced DSBs, in RPE-1 cells at least .

Finally, using the RPE-1 *TDP2*<sup>-/-</sup> cell line established on this chapter, our lab colleague Dr Fernando Gomez-Herreros, was able to investigate the effect of TDP2 on chromosome instability (Gómez-Herreros *et al.*, 2017). It was shown that the absence of TDP2 in patient-derived lymphoblasts leads to higher chromatids breaks and increased micronuclei formation after etoposide (Gómez-Herreros *et al.*, 2017). Furthermore, he was able to reproduce previous observations that breakage occurs on mixed lineage leukaemia (*MLL*) gene following etoposide on RPE *TDP2*<sup>-/-</sup> cell line generated and characterised in this chapter (Figure 3.12.A and B) (Cowell *et al.*, 2012; Gómez-Herreros *et al.*, 2017). The *MLL* locus has been shown to be translocated in therapy-related acute myeloid leukaemia (t-AML) following administration of topoisomerase II ‘poisons’ (Super *et al.*, 1993). These translocations frequently occur in the region between exon 8 and 12, named break cluster region (BCR), which has been shown to be a site of TOP2 $\beta$  binding on ChIP experiments (Cowell *et al.*, 2012). Thus, breakage across this site can be measured using fluorescence in situ hybridisation (FISH) probes located either side of BCR (Gómez-Herreros *et al.*, 2017). RPE *TDP2*<sup>-/-</sup> cells have shown higher breakage across this locus following etoposide when compared to RPE-1 wild-type (Figure 3.12.B). Indicating that TDP2 mediated repair of TOP2cc across this site and prevented chromosome translocations with the *MLL* locus, which are hallmarks of t-AML (Libura *et al.*, 2005) . Furthermore, when repeated in the presence of the transcriptional inhibitor DRB, etoposide no longer generated breaks across the *MLL* locus (Figure 3.12.B). Confirming previous observations that active transcription is required for translocations on this locus to form (Cowell *et al.*, 2012). Lastly, to investigate that TDP2 activity is required to prevent translocations across the genome, whole chromosome FISH probes for chromosome 8 and 11 were used in the presence or absence of etoposide (Figure 3.12.C). TDP2 absence lead to high number of translocations on both of these



**Figure 3.12 – TDP2 absence leads to increased transcription-dependent chromosomal breakage and translocations - (A)** Representative FISH images of MLL loci in RPE-1  $TDP2^{-/-}$  cell line in the untreated (left) or treated (right) cells. Arrows indicate intact or broken MLL loci. **(B)** Detection of broken MLL loci following incubation with 100  $\mu$ M etoposide for 6 hours in serum-starved RPE-1 or RPE-1  $TDP2^{-/-}$  #3 cell lines. For transcription dependency, cell were pre-incubated with 100  $\mu$ M DRB for 3 hours prior etoposide experiment. Data are mean  $\pm$  SEM of three biological experiments. **(C)** Translocation per metaphase spread (48hr post treatment) of serum-starved RPE-1 and RPE-1  $TDP2^{-/-}$  treated with 100  $\mu$ M etoposide for one hour. Translocation was measure by whole chromosome probes for chromosome 8 and 11. When indicated DRB treatment (100  $\mu$ M) occurred 3h prior etoposide. Data are from two independent experiments. All statistical significance was determined by t-test , \* $p$ <0.05 and \*\*\* $p$ <0.001 [Data from Dr. Gomez-Herreros published at Gomez-Herreros et al. 2017]

chromosomes, which was also prevented by pre-incubating cells with DRB. Corroborating, that TDP2 activity is essential for chromosome stability on active transcription sites and consistently, with the role of TDP2 in repair of TOP2 $\beta$  transcriptional breaks in neuronal genes and transcription defect in mouse neurons (Gómez-Herreros *et al.*, 2014; Madabhushi *et al.*, 2015; Madabhushi, 2018). This study raises the question of the role of TDP2 in t-AML and how relevant the observations would translate to *in vivo* settings. Secondary leukemias are common, indicating that levels of damage at the transcription sites are elevated enough to require multiple repair pathways, one of which we speculate to be TDP2. Interestingly, of all haemopoietic cells, TDP2 is found as lowest expressing in haemopoietic stem cells and AML tumours (Bagger *et al.*, 2016), indicating that TDP2 expression could be a hallmark of these tumours.

In summary, this chapter establishes an approach to generate and characterise *TDP2*<sup>-/-</sup> cells. The cells generated here were validated using sequencing, western-blotting and DNA repair assays, and shown to behave as previous orthologous cell lines lacking TDP2.

#### 4. CHAPTER FOUR – IDENTIFICATION AND CHARACTERISATION OF TDP2 ISOFORMS

##### 4.1. INTRODUCTION AND AIMS

Other than its well-established role in DNA repair, as introduced earlier, TDP2 was first identified as an interactor of tumour necrosis factor (TNF) and TNF receptor-associated factors (TRAFs), and consequently was named TRAF and TNF receptor associated protein (TTRAP) (Pype *et al.*, 2000). Its molecular function was found to be repression of necrosis factor (NF) kappa B (NF- $\kappa$ B) activation by interaction with the cytoplasmic domain of the cell receptor CD40 (Pype *et al.*, 2000). Moreover, TTRAP/TDP2 was shown to interact with Transforming growth factor- $\beta$  (TGF- $\beta$ ) cytoplasmic receptors and regulate its signalling (Várady *et al.*, 2011). Later, TDP2 was found to be associated with the Erythroblastosis (ETS) proto-oncogene 1 (ETS1) transcription factor, and was named ETS associated protein 2 (EAPII) (Pei *et al.*, 2003). This work has shown that TDP2 does not inhibit ETS-1 DNA interaction, but it does however regulate its transactivation on matrix metalloproteinase-1 (MMP1) promoter, with overexpression resulting in reduction of cell migration (Pei *et al.*, 2003).

In addition, there are a number of reports that TDP2 is involved in the mechanism of viral infection. TDP2 has been reported to interact with human immunodeficiency virus (HIV) integrase (HIV IN) and Phage $\Phi$ C31 integrase and to facilitate lentiviral vector integration (Zhang *et al.*, 2009; Wang *et al.*, 2010). In addition, TDP2 phosphodiesterase activity was shown to be required for picornaviruses infection, functioning as the VPg unlinkase (Virgen-Slane *et al.*, 2012; Maciejewski *et al.*, 2018). Picornaviruses are positive-strand RNA viruses that use a small viral protein (VPg) as a primer for viral RNA synthesis, resulting in the viral RNA genome being covalently linked to VPg via a O<sup>4</sup>-(5'-uridylyl)tyrosine bond (Virgen-Slane *et al.*, 2012). TDP2 mediates the removal of the VPg cap which facilitates viral infection, although it was shown not to be required for translation and replication of viral genome (Langereis *et al.*, 2014). Nevertheless, that TDP2 is the 38 kDa VPg unlinkase and binds exclusively to viral genome on which VPg is linked strongly suggests that the virus exploits TDP2 to facilitate infection (Virgen-Slane *et al.*, 2012). These reported TDP2 functions are not linked to its established function in nuclear DNA repair, and indeed its proposed role in



cell signalling and viral infection has been suggested to depend on its cytoplasmic localisation (Li *et al.*, 2011; Virgen-Slane *et al.*, 2012) .

#### **4.1.1. Aim**

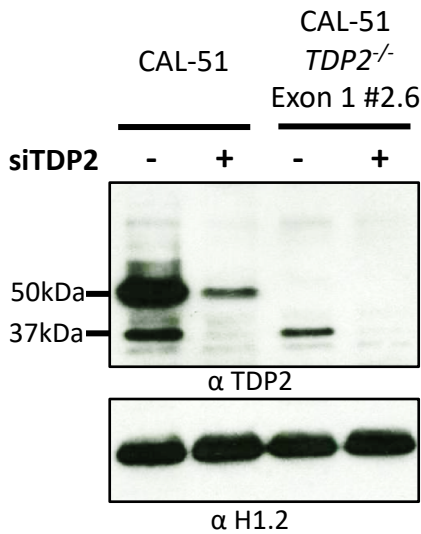
In this chapter, I aim to investigate any putative new isoforms of TDP2 in CRISPR-Cas9 cells in which the canonical TDP2 protein was deleted. To do so, I will analyse a protein of 37 kDa detected by immunoblotting using TDP2 antibody. Any identified isoform of TDP2 will be characterised according to its sequence, evolutionary conservation, cellular localisation and its involvement in the DNA damage response using DSB repair and survival assays.

### **4.2. RESULTS**

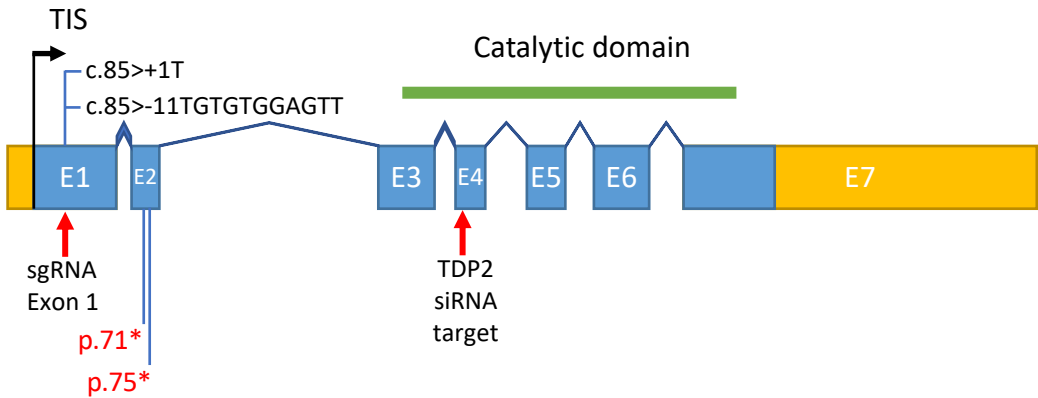
#### **4.2.1. TDP2 isoform can be detect at 37kDa**

TDP2 protein is detected as a band of approximately 50 kDa in cell extracts from CAL-51 cells by western blotting, using a polyclonal antibody raised against a fragment of TDP2 encoding exons 3-7 and including the catalytic active site (Thomson *et al.*, 2013)(Figure 4.1.A; lane 1). As discussed in Chapter 3, targeting exon 1 of TDP2 with CRISPR-Cas9 ablates this 50 kDa protein and renders both CAL-51 and RPE-1 cells hypersensitive to etoposide (Figure 3.7.A and 3.9.C). However, it was noted that these cells possess a faster migrating protein with an apparent molecular mass of ~ 37 kDa that was detected by the anti-TDP antibody in both wild type and CAL-51 and *TDP2*<sup>-/-</sup> CAL-51 cells (Figure 4.1.A; lanes 1 & 3). This lower protein band was previously observed while screening for *TDP2*<sup>-/-</sup> clones in Chapter 1 in RPE-1 and U2OS cells in addition to the isolated RPE-1 *TDP2*<sup>-/-</sup> clone #3 (Figure 3.4.A, 3.4.C and 3.8.A respectively). Intrigued by the possibility that this band could be an isoform of TDP2, wild type and *TDP2*<sup>-/-</sup> CAL-51 cells were transfected with siRNA against exon 4 of TDP2. To our surprise, TDP2 siRNA resulted in depletion of both the 50 kDa and 37 kDa proteins in wild type CAL-51 cells, and depletion of the 37 kDa protein in *TDP2*<sup>-/-</sup> CAL-51 cells (Figure 4.1.A; lanes 2 & 4). This result strongly suggests that the protein detected at 37 kDa is an isoform of TDP2, or contains a very similar catalytic domain (Figure 4.1.B).

A



B



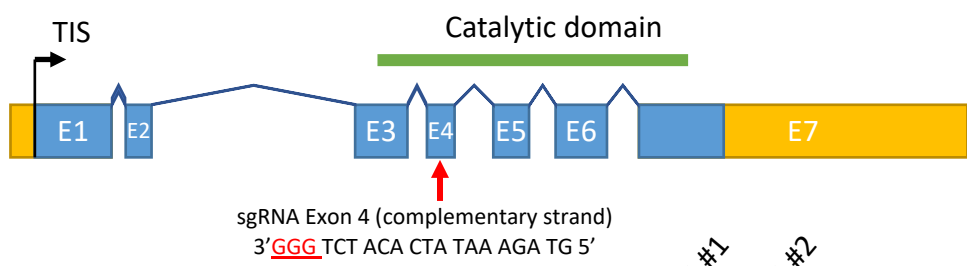
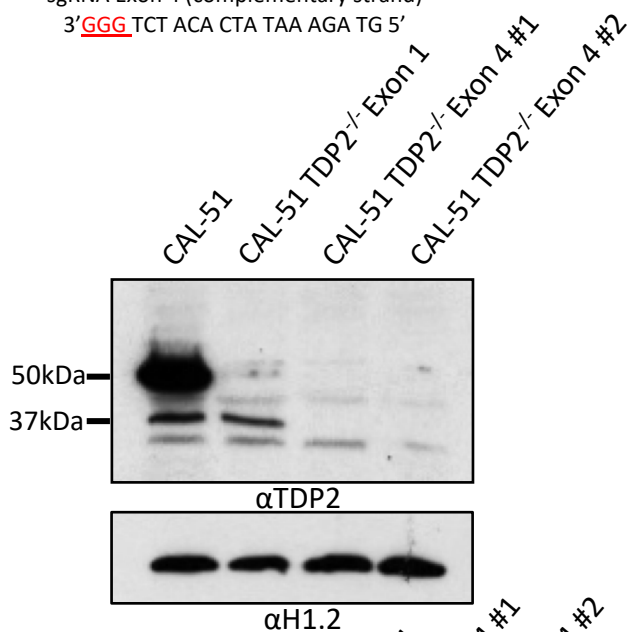
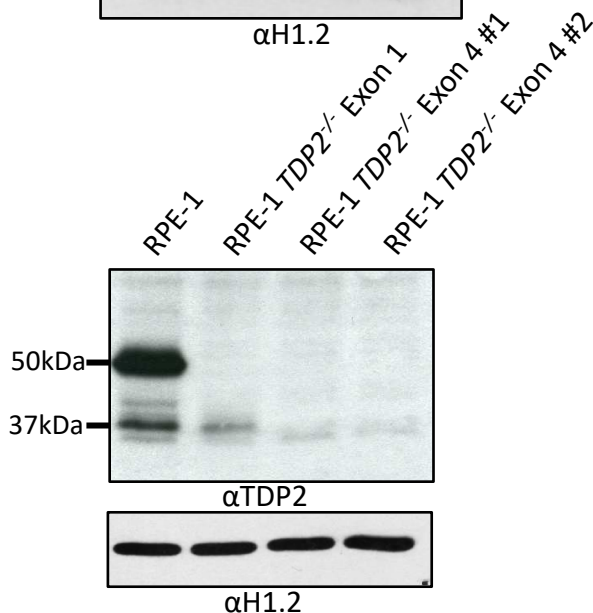
**Figure 4.1 – A putative 37 kDa isoform of TDP2. (A)** Western blot against TDP2 in CAL-51 and *TDP2*<sup>-/-</sup> CAL51 cells (targeted by CRISPR/Cas9 in exon 1) following transfection with non-targeting siRNA (-) or TDP2 siRNA (+). Immunoblotting with anti-histone 1.2 (H1.2) was employed as a loading control. **(B)** Gene structure of *TDP2* depicting the mutations identified in exon 1 of *TDP2*<sup>-/-</sup> CAL-51 cells described in Chapter 3, by WES and Sanger sequencing (the predicted impact on protein sequence is indicated in red). Red arrows indicate the position of the CRISPR-Cas9 sgRNA targets in exon 1 and the siRNA target in exon 4. Exon coding sequence is shown in blue and untranslated regions are shown in yellow.

#### **4.2.2. Deletion of TDP2 isoform leads to increased sensitivity to Topoisomerase 2 induced damage.**

I considered it likely that the 37 kDa protein detected by anti-TDP2 antibody and depleted by TDP2 siRNA is a novel isoform of TDP2, which presumably lacked the exon 1 sequences targeted by the guide RNA employed in Chapter 3. The indels identified by sequencing in *TDP2*<sup>-/-</sup> CAL-51 cells predicted a frameshift resulting in premature stop codons at p.71\* and p.75\* respectively (Figure 4.1.B). Taking in account the observations with the siRNA experiment, I designed a new sgRNA against exon 4 of TDP2 with the sequence 5' GTA GAA ATA TCA CAT CT 3' (Figure 4.2.A). Using the same approach as described in Chapter 3, I isolated single clones and verified TDP2 expression using western blot against TDP2 (data not shown). In CAL-51 cells generated using this approach, two independent clones, CAL-51 *TDP2*<sup>-/-</sup> Exon 4 clone #1 and #2, displayed loss of TDP2 at 50 kDa and 37 kDa, confirming that the lower band encodes an isoform of TDP2 (Figure 4.2.B). I applied the same method to RPE-1 cells and obtained similar results; whereas the expression of the lower isoform was reduced in RPE-1 *TDP2*<sup>-/-</sup> Exon 1 cells, it was completely lost in two independent RPE-1 *TDP2*<sup>-/-</sup> Exon 4 clones (Figure 4.2.C).

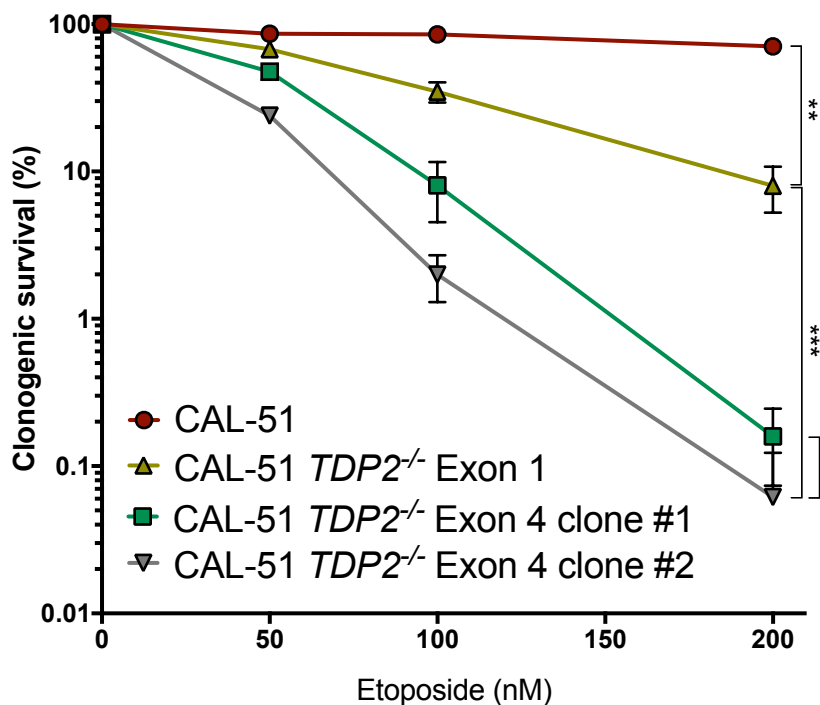
Considering that the putative new TDP2 isoform is predicted to contain the catalytic domain, it was considered possible that it might play a role in the repair of TOP2-induced DSBs. I therefore employed clonogenic survival assays to compare the sensitivity of CRISPR-Cas9 *TDP2*<sup>-/-</sup> cells targeted at exon 1 and exon 4 to etoposide. Two independent *TDP2*<sup>-/-</sup> CAL-51 clones targeted in exon 4 were significantly more sensitive than the CAL-51 cell clone in which TDP2 was targeted in exon 1 (CAL-51 *TDP2*<sup>-/-</sup> Exon 1) (Figure 4.3.A). The same was true when we repeated the clonogenic survival assays using RPE-1 cells, indicating that the putative new TDP2 isoform does indeed promote resistance to TOP2-induced DSBs (Figure 4.3.B).

Next, the rate of repair of TOP2-induced DSBs was measured using immunostaining of  $\gamma$ H2AX foci. As observed previously (Chapter 3), CAL-51 and RPE-1 cells in which TDP2 was gene edited in exon 1 exhibited reduced rates of DSB repair (Fig.4.4 A & B). However, CAL-51 and RPE-1 cells in which TDP2 was targeted in exon 4

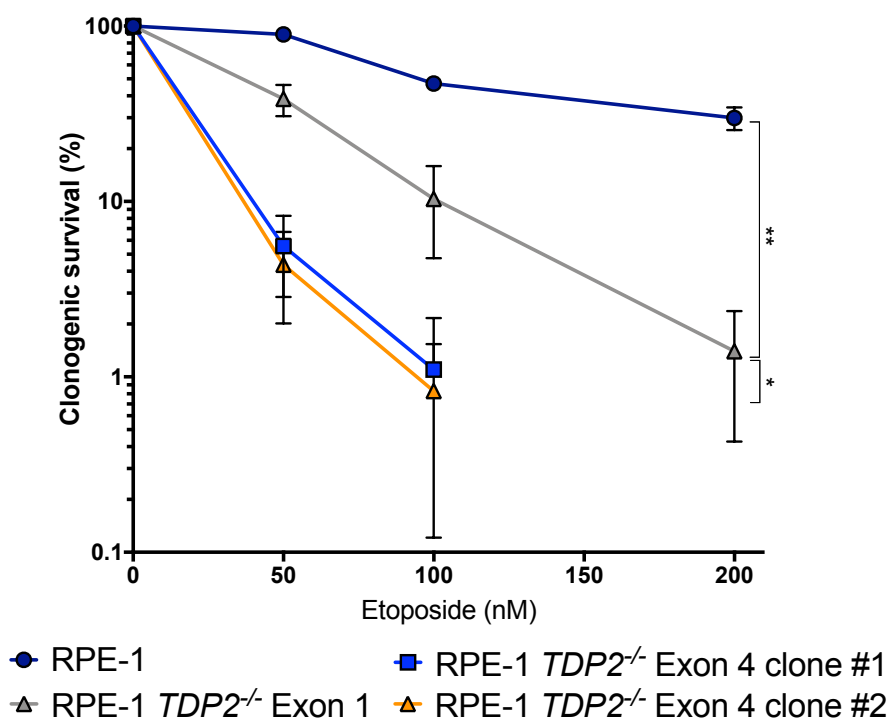
**A****B****C**

**Figure 4.2 – CRISPR-Cas9 Exon 4 *TDP2*<sup>-/-</sup> clones lack the 37 kDa TDP2 isoform - (A)** *TDP2* gene structure indicating CRISPR-Cas9 target site (sgRNA) in exon 4, with the PAM sequence highlighted in red. Exon coding sequence is shown in blue and untranslated regions are shown in yellow. **(B)** Western blot against TDP2 in CAL-51, CAL-51 *TDP2*<sup>-/-</sup> Exon 1 and two independent clones generated by CRISPR-Cas9 targeting of Exon 4 (*TDP2*<sup>-/-</sup> Exon 4 #1 and #2). Anti-histone 1.2 (H1.2) used as loading control **(C)** Same as (B), but employing wild type RPE-1 and *TDP2*<sup>-/-</sup> RPE-1 cells instead.

A

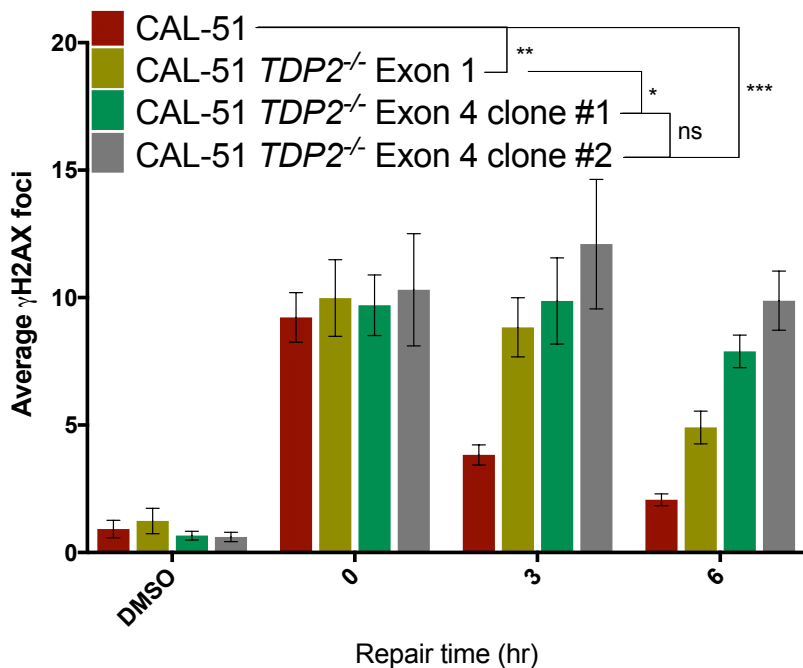


B

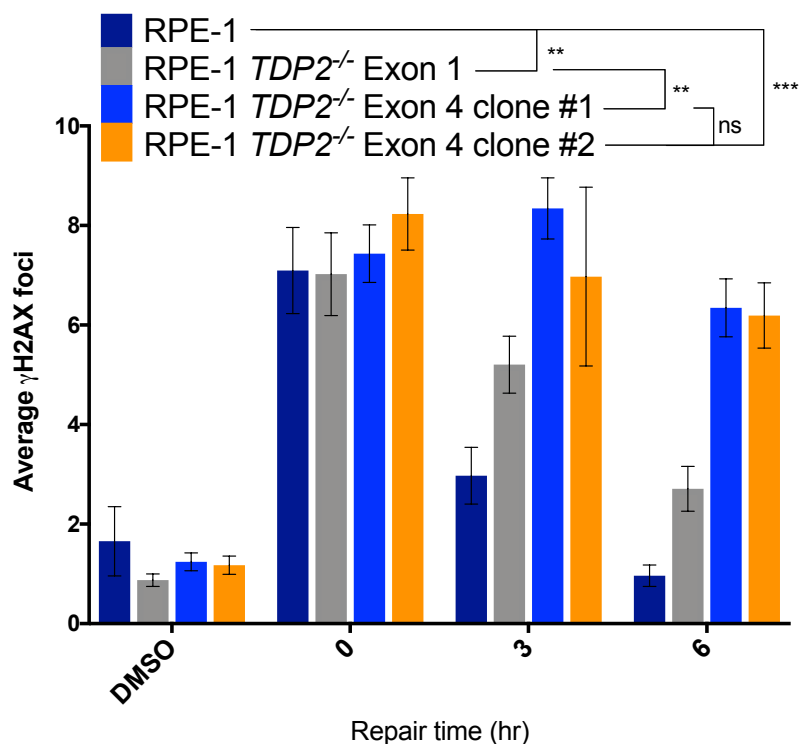


**Figure 4.3 – The 37 kDa isoform of TDP2 promotes resistance to topoisomerase 2 induced DSBs. (A)** Clonogenic survival following continuous exposure to the indicated concentrations of etoposide in CAL-51, CAL-51 *TDP2*<sup>-/-</sup> Exon 1 and two independent clones of CAL-51 *TDP2*<sup>-/-</sup> Exon 4 (#1 and #2). **(B)** As in (A) except that RPE-1, RPE-1 *TDP2*<sup>-/-</sup> Exon 1 and two independent clones *TDP2*<sup>-/-</sup> Exon 4 (#1 and #2) were employed. All data are the mean  $\pm$  SEM of three independent experiments. Statistical significance was determined by two-way ANOVA on GraphPad Prism software. \*p<0.05, \*\*p<0.01 and \*\*\*p<0.001

A



B



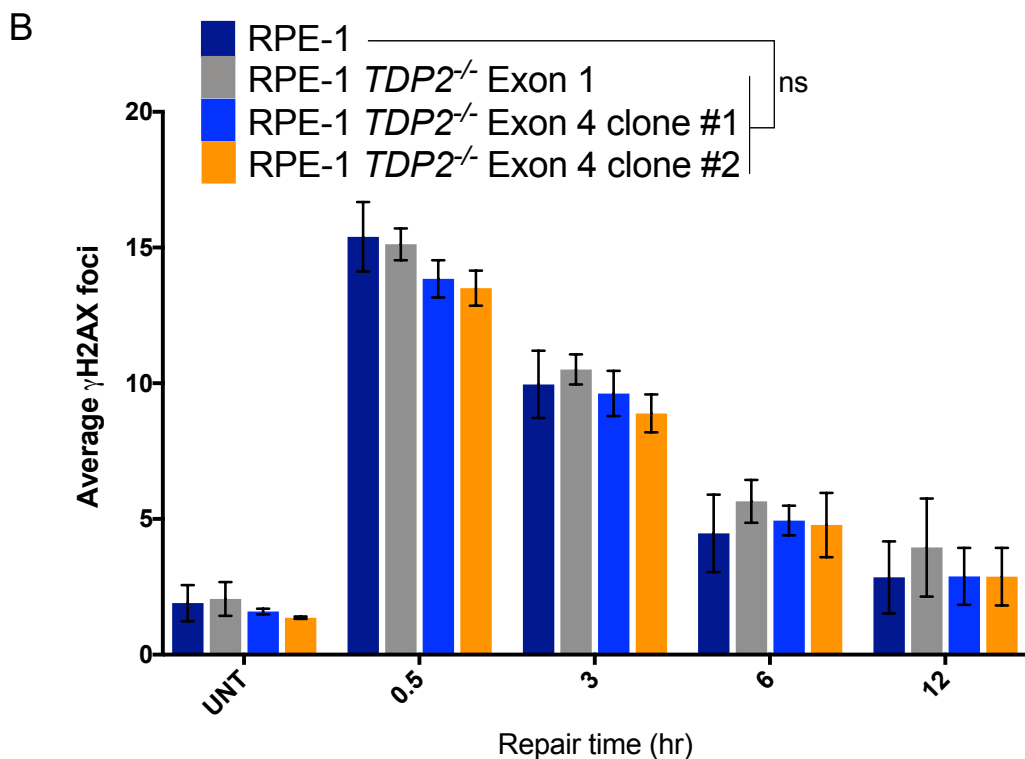
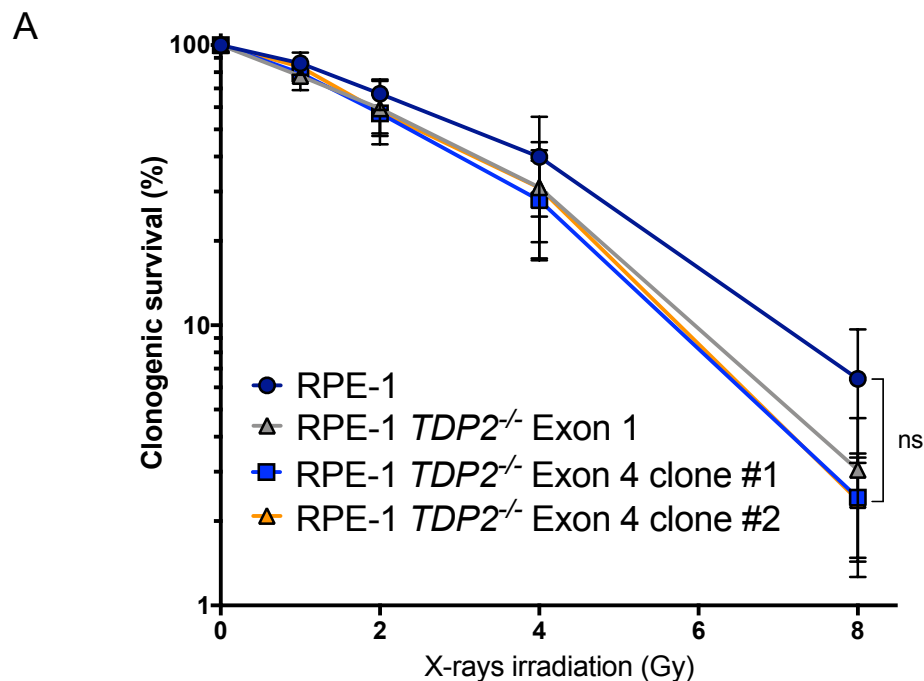
**Figure 4.4 –TDP2 isoform is involved in nuclear DSB repair following etoposide - (A)** DSB was measured by  $\gamma$ H2AX foci counts in in CAL-51, CAL-51 *TDP2*<sup>-/-</sup> Exon 1 and two independent clones of CAL-51 *TDP2*<sup>-/-</sup> Exon 4 (#1 and #2), following 30min incubation with 25 $\mu$ M etoposide and the indicated repair time in drug-free medium. **(B)** Same as in (A) however RPE-1, RPE-1 *TDP2*<sup>-/-</sup> Exon 1 and two independent clones *TDP2*<sup>-/-</sup> Exon 4 (#1 and #2) were used instead. All cells were gated to G1 population according to DAPI profile and CENPF expression. All data are the mean  $\pm$  SEM of three technical replicas. All statistical significance was determined by ANOVA on GraphPad Prism software. ns= $p>0.05$ , \* $p<0.05$ , \*\* $p<0.01$  and \*\*\* $p<0.001$ .

exhibited even slower rates of DSB repair (Figure 4.4.A & B), consistent with the results of the clonogenic survival assays. As these results indicated that the TDP2 isoform played a role in nuclear DSB repair, we assessed its involvement in survival and repair following ionising radiation. Unlike etoposide induced DSB, most DSBs induced by irradiation should not require TDP2 activity because the breaks are not associated with topoisomerase cleavage complexes. Consistent with this idea, the *TDP2*<sup>-/-</sup> RPE-1 cells were not significantly hypersensitive to X-ray irradiation (Figure 4.5.A) and exhibited normal DSB repair kinetics as measured by  $\gamma$ H2AX immunostaining following 2 Gy of  $\gamma$ -irradiation (Figure 4.5.B).

#### 4.2.3. Origin of the 37 kDa TDP2 isoform.

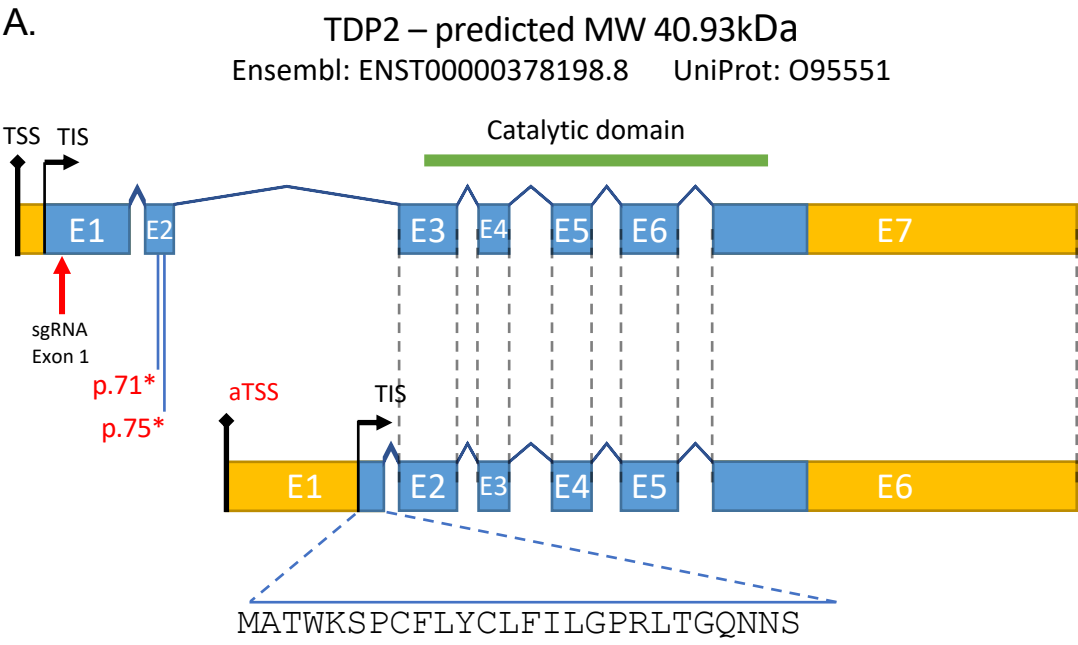
Initially, the Ensembl (Zerbino *et al.*, 2018) and UniProt (UniProt Consortium, 2018) databases were interrogated for DNA sequences or proteins with strong TDP2 sequence similarity. In Ensembl, we found the transcript ENST00000341060.3, and in UniProt the protein X6R5A3. Both entries were related to the same protein, which was listed as a predicted isoform of TDP2 lacking biological confirmation, and based on cDNA sequence from brain, cerebellum and lung (Thierry-Mieg and Thierry-Mieg, 2006). The coding sequence for this TDP2 isoform, which we named TDP2 beta (TDP2 $\beta$ ), suggested an alternative transcription start site (aTSS) located in the large intron 2 on TDP2 genomic sequence (Figure 4.6.A). The novel 5' untranslated region (5'UTR) and exon 1, replaced TDP2 exon 1 and 2 with a sequence composed of 26 amino-acids, resulting in a protein with molecular weight of approximately 34.4 kDa (Figure 4.6.A). As TDP2 $\beta$  exon 1 is spliced into TDP2 exon 3, this new protein would share the catalytic active phosphodiesterase domain of TDP2, and may thus be enzymatically functional. Consistent with this idea, alignment of TDP2 genomic sequence from multiple organisms indicated that the TDP2 $\beta$  start codon (ATG) was conserved in mammals, although it was not present in the genome sequence of *Gallus gallus* (chicken) (Figure 4.6.B), *Danio rerio* (Zebrafish), or *Xenopus laevis* (frog) (data not shown).

Additional to the putative TDP2 $\beta$  isoform identified by database searches, I also attempted to predict the existence of an additional relevant isoform by sequence analysis of the *TDP2* cDNA sequence. This analysis suggested that an isoform of TDP2



**Figure 4.5 – Loss of TDP2 isoforms does not affect survival or DSB repair following irradiation - (A)** Clonogenic survival following exposure to the indicated doses of X-rays irradiation in RPE-1, RPE-1 *TDP2*<sup>-/-</sup> Exon 1 and two independent clones *TDP2*<sup>-/-</sup> Exon 4 (#1 and #2) **(B)** DSB repair was measured by  $\gamma$ H2AX foci counts in same cells of (A), following 2Gy  $\gamma$ -irradiation and the indicated repair time. Untreated cells (UNT) were used as negative control. All cells were gated to G1 population according to DAPI profile and CENPF expression. All data are the mean  $\pm$  SEM of three technical replicas. All statistical significance was determined by two-way ANOVA on GraphPad Prism software. ns= $p>0.05$





**B.**

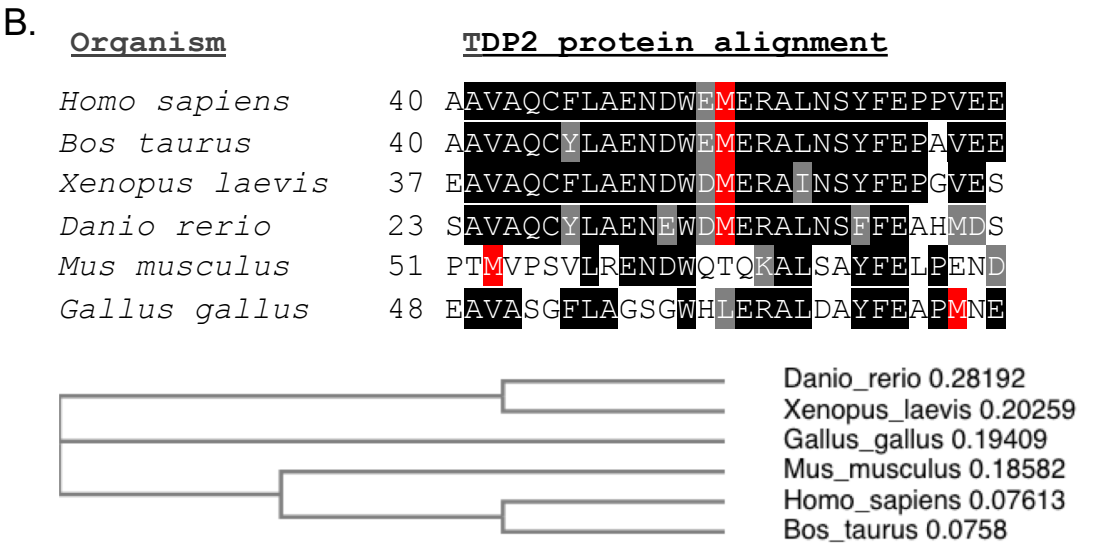
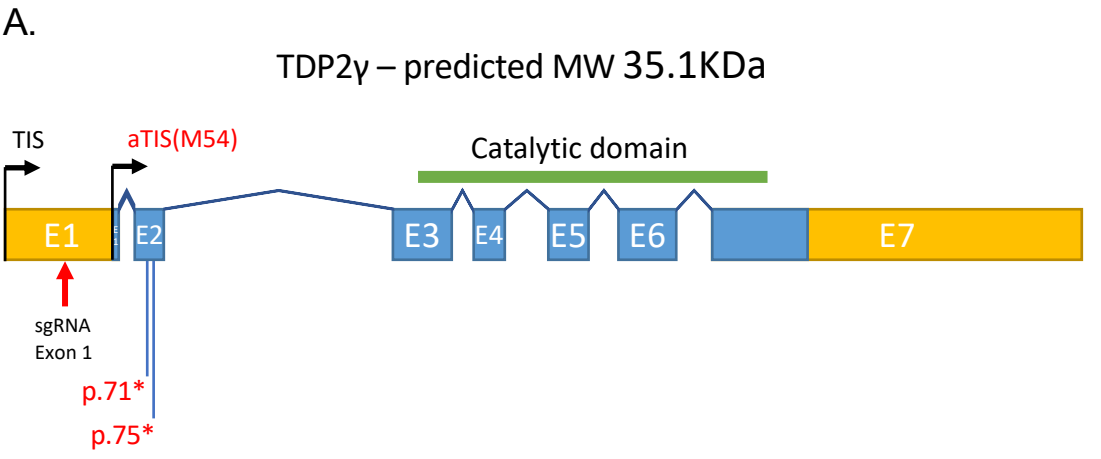
<u>Organism:</u>	<u>TDP2β start codon conservation</u>
<i>Homo sapiens</i>	CCTTCTATCGAAAA <b>ATGG</b> CAACTTGGGAAGTC
<i>Bos taurus</i>	TCCTCTACTAAAGA <b>ATG</b> GTAGCTTGGTGGTC
<i>Mus musculus</i>	TTTCTCCTGAAGA <b>ATG</b> GTAACCTGTAAGTC
<i>Gallus gallus</i>	CCCTGTGGTAGCACATCTTGCCTTTGACTGG

**Figure 4.6 – TDP2 isoform beta (TDP2β) – (A) (top)** Diagram of TDP2 gene structure. Transcription start site (TSS) and translation initiation site (TIS) are indicated. TDP2 accession number on Ensembl and UniProt databases and its molecular weight (MW) are given. **(bottom)** Diagram of a putative TDP2 isoform, namely TDP2 beta (TDP2β), as per sequence described in Ensembl and UniProt databases, with predicted molecular weight (MW) of 34.4 KDa. TDP2β alternative transcription start site (aTSS) (located at intron 2) and novel N-terminus sequence is highlighted. Exon coding sequence is shown in blue and untranslated regions are shown in yellow. **(B)** Genome alignment of multiple organisms shows that the putative start codon (ATG) position in TDP2β is conserved in mammals but absent in other vertebrates. Black highlight indicates conservation in all species and grey highlight indicates conservation in at least two species.

might arise from an alternative translation initiation site (aTIS) at Met54 (Figure 4.7.A). This isoform, named here as TDP2 gamma (TDP2 $\gamma$ ), would possess some of exon 1 but would no longer encode the premature stop codons at positions 71 and 75. Like TDP2 $\beta$ , TDP2 $\gamma$  was predicted to contain the TDP2 phosphodiesterase catalytic domain, and to have a molecular mass of 35.1 kDa (Figure 4.7.A). Interestingly, multiple alignments of different TDP2 protein sequences revealed that the region spanning Met54 was highly conserved, even in evolutionarily divergent organisms including *Homo sapiens*, *Bos Taurus*, *Xenopus laevis* and *Danio rerio* (Figure 4.7.B top). In addition, even in organisms such as *Mus musculus* and *Gallus gallus* in which Met54 was not conserved, a nearby in-frame Met was present that would allow translation of a closely related isoform of a similar molecular mass (Figure 4.7.B top and bottom). The Met residues identified here were in sequence contexts harbouring a Kozak consensus translational start sequence (A/G)XXATGG(Kozak, 1986) (Figure 4.7.C).

#### 4.2.4. TDP2 $\beta$ and TDP2 $\gamma$ are present in human cells

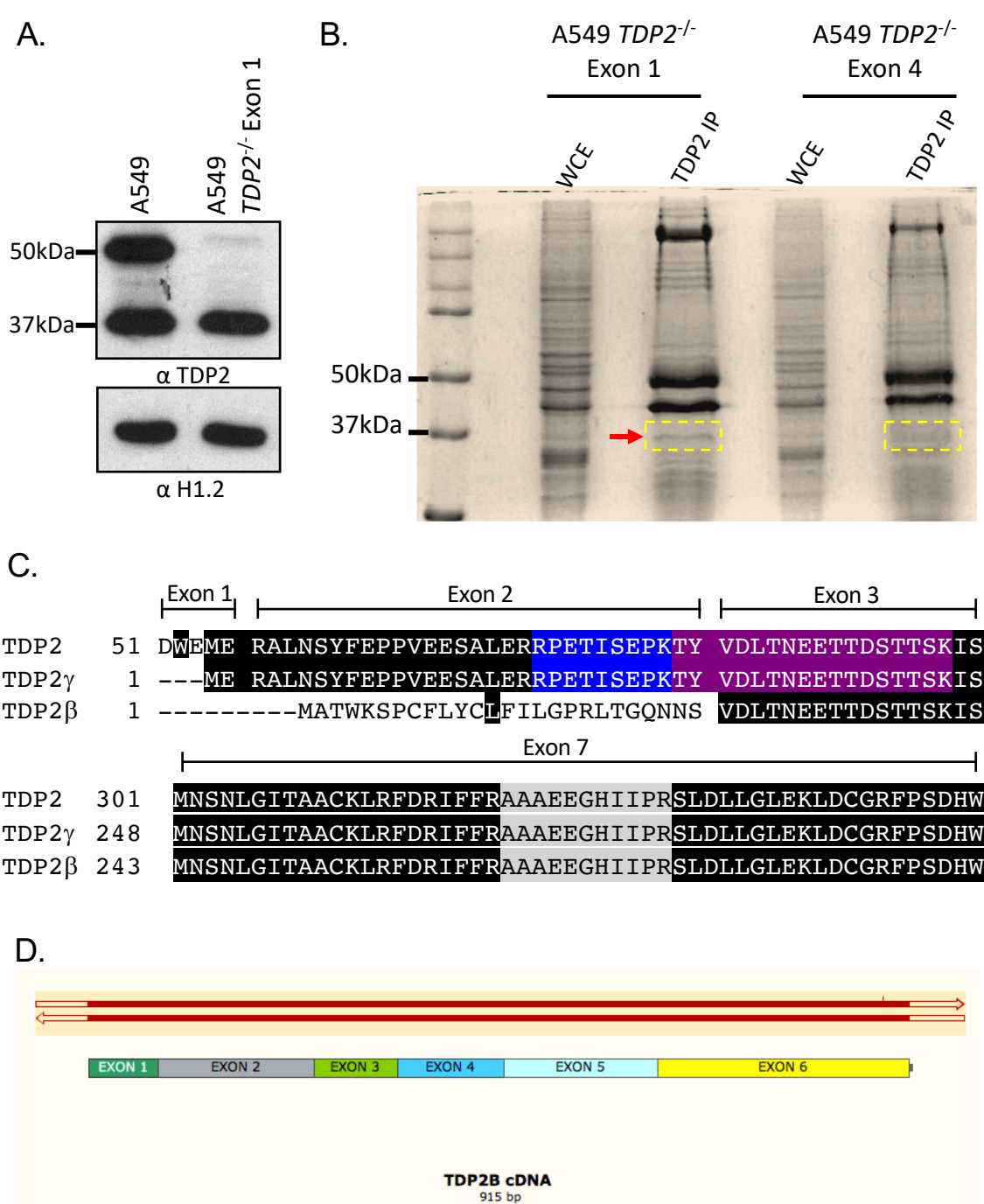
To examine whether the two putative shorter TDP2 isoforms exist in cells, I employed A549 cells in which TDP2 was gene edited using the sgRNAs that targeted either exon 1 or exon 4, so that the only TDP2 protein present in the experiment was TDP2 $\beta$  and/or TDP2 $\gamma$  (Figure 4.8.A). A549 cells were chosen for these experiments because these cells appeared to possess higher levels of the 37 kDa isoform than either CAL-51 or RPE-1 cells, accounting for approximately half of the total detected TDP2 (compare Figures 4.8.A, 3.7.A and 3.8.A). TDP2 was immunoprecipitated from cell extracts, resulting in the detection on silver stained gels of a ~37 kDa protein that was present in immunoprecipitates of the exon1, but not exon 4, gene edited cells (Figure 4.8.B). This protein was excised from the gel along with the material present at the analogous position in the immunoprecipitate from the exon 4 gene edited cells (yellow boxes in Figure 4.8.B), and examined by mass spectrometry (MS). Of note it is important to highlight that the prominent bands observed in the silver staining at 50kDa and ~42kDa are correspondent to antidoby heavy chain and not TDP2 protein, as these are also observed in A549 TDP2<sup>-/-</sup> cells that do not have detectable TDP2 protein. Only three peptides from TDP2 were detected by mass spectrometry; a C-terminal peptide (aa322-



**C.**

	<u>KOZAK Consensus</u> : [A/G] xx <b>ATGG</b>	<u>Methionine position</u>
<i>Homo sapiens</i>	CTGG <b>GAGATG</b> GAAAGGGC	M54
<i>Bos taurus</i>	CTGG <b>GAGATG</b> GAAAGAGC	M54
<i>Xenopus leavis</i>	CTGG <b>GATATG</b> GAAAGGGC	M53
<i>Danio rerio</i>	GTGG <b>GACATG</b> GAGAGAGC	M50
<i>Mus musculus</i>	CCCC <b>ACGATG</b> TCCCCAG	M53
<i>Gallus gallus</i>	GGCGCC <b>GTG</b> AACGAGCA	M73

**Figure 4.7 – TDP2 isoform gamma (TDP2γ).** **(A)** Diagram of an isoform of TDP2, namely TDP2 gamma (TDP2γ) with predicted molecular weight (MW) of 35.1 KDa. Coding sequence (blue rectangles) starts at an alternative translation initiation site (aTIS) at methionine 54 and should not be affected by CRISPR-Cas9 indels on exon 1. Exon coding sequence is shown in blue and untranslated regions are shown in yellow. **(B)** Alignment of TDP2 protein in multiple organisms indicate that methionine at N-terminus is conserved (top alignment) even in sequences that are evolutionarily divergent as indicated by phylogenetic tree (bottom). **(C)** Alignment of N-termini start codon (ATG) in multiple species, indicates that it is evolutionarily conserved and presents the Kozak consensus [A/G] xxATG.

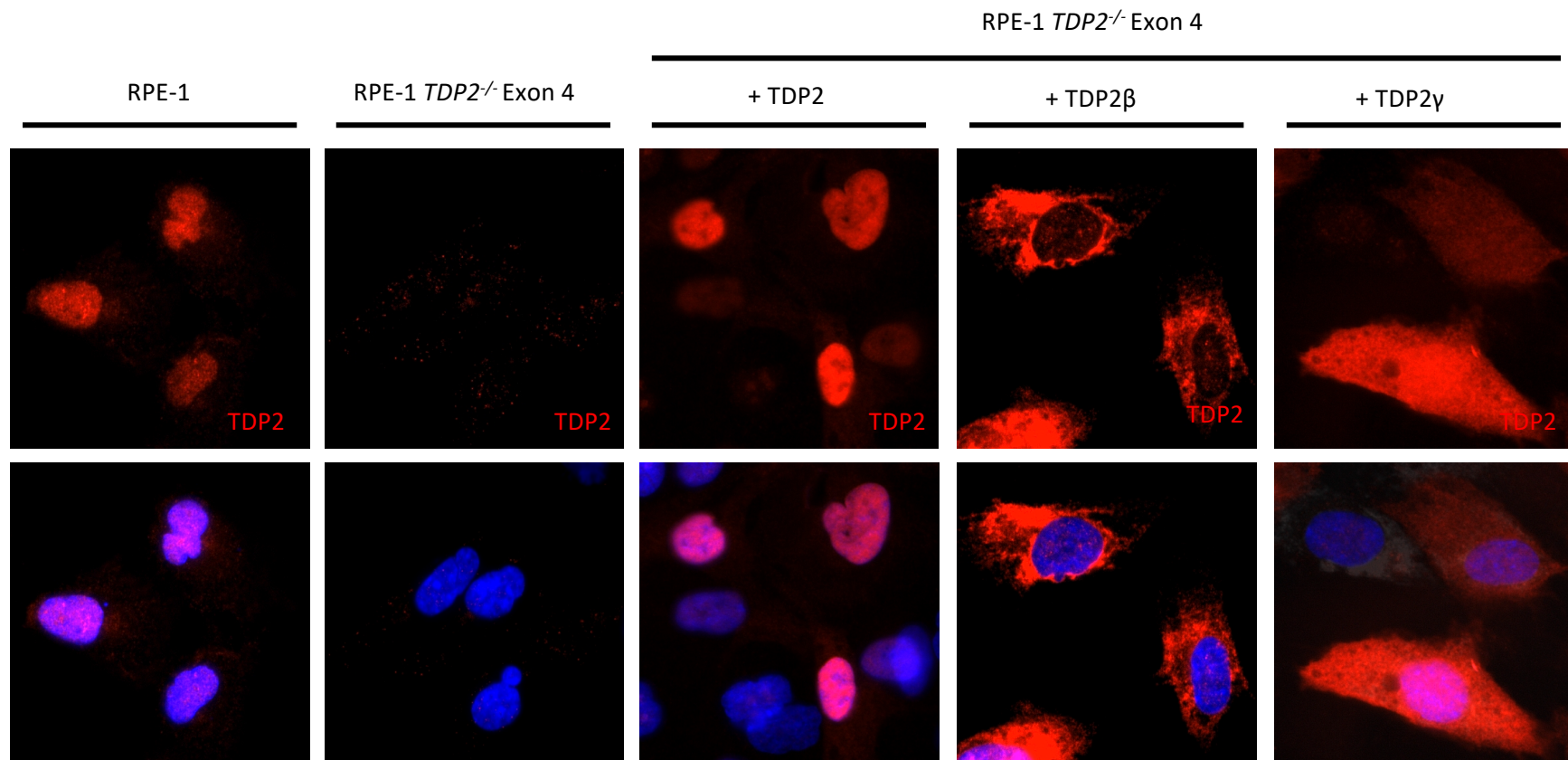


**Figure 4.8 – Mass spectrometry and sequencing analysis of TDP2 isoforms. (A)** Western blot against TDP2 in A549 and A549 *TDP2*<sup>-/-</sup> Exon 1. Anti-histone 1.2(H1.2) used as loading control **(B)** Silver staining of SDS-PAGE of whole cell extract (WCE) or TDP2 immunoprecipitation (IP) in A549 *TDP2*<sup>-/-</sup> Exon 1 versus A549 *TDP2*<sup>-/-</sup> Exon 4. The position of the TDP2 isoform at 37KDa and the band excised for mass spectrometry analysis are indicated by red arrow and yellow box, respectively. **(C)** Protein alignment of TDP2, TDP2 $\beta$  and TDP2 $\gamma$ . Peptide sequences identified by MS are highlighted in blue (aa 74-82), mangenta (aa. 83-99) and grey (aa.322-332). TDP2 exons boundaries are indicated above the sequences **(D)** TDP2 $\beta$  predicted cDNA structure (Exon 1 to Exon 6) with forward and reverse alignment of PCR fragment sequenced from cDNA library prepped in A549 *TDP2*<sup>-/-</sup> Exon 1 cells.

332) that would be expected to be shared by all of the isoforms, an N-terminal peptide located in exon 2 (aa 74-82), and a third spanning the sequence coded by TDP2 exon 2 and 3 (aa 83-89) (Figure 4.8.C). The presence of the latter supported the existence of the TDP2 $\gamma$  isoform, because a peptide spanning exons 2 and 3 would only be present in the 37 kDa protein if the alternative translation initiation at Met54 occurred as predicted (Figure 4.7.A). In contrast, unique peptides that demonstrate the existence of TDP2 $\beta$  were not detected. However, I examined whether the cDNA encoding this isoform was present in A549 cells in which TDP2 was gene edited in exon 1, by preparing a cDNA library from RNA extracted from this clone followed by PCR amplification using TDP2 $\beta$  specific primers and subsequent Sanger sequencing. This experiment recovered the predicted coding sequence for TDP2 $\beta$  (Figure 4.8.D), suggesting that both of the 37kDa isoforms are most likely expressed in cells.

#### **4.2.5. TDP2 isoforms do not localise exclusively to nucleus**

Since the nuclear localisation signal (NLS) of TDP2 is encoded by exon 1 (Rao *et al.*, 2016), I investigated the subcellular localisation of TDP2 $\beta$  and TDP2 $\gamma$ . First, I confirmed using indirect immunofluorescence that full length TDP2 localises to the nucleus in wild type RPE-1 cells (Figure 4.9). The anti-TDP2 signal was specific for TDP2 because it was absent from *TDP2*<sup>-/-</sup> RPE-1 cells in which exon 4 was gene edited (Figure 4.9). I then transfected these *TDP2*<sup>-/-</sup> RPE-1 cells with expression constructs encoding either full length TDP2, TDP2 $\beta$  or TDP2 $\gamma$  (Figure 4.9). Cells transfected with the full length TDP2 construct showed a strong nuclear localisation of TDP2, consistent with the signal detected in wild type RPE-1 cells. In contrast, cells transfected with expression construct encoding TDP2 $\beta$  or TDP2 $\gamma$  exhibited cytoplasmic or pan-cellular anti-TDP2 signal, respectively (Figure 4.9). The pan-cellular localisation of TDP2 $\gamma$  is consistent with absence of the NLS encoded by exon 1, because this protein is most likely small enough to diffuse into and out of the nucleus (Knockenbauer and Schwartz, 2016). In contrast, TDP2 $\beta$  localisation was almost entirely cytoplasmic, suggesting that this protein may harbour a cytoplasmic localisation signal. I therefore employed standard software programmes that predict the subcellular localisation of a protein, which predicted that



**Figure 4.9 – TDP2 isoforms cellular localisation.** – Immunostaining using anti-TDP2 in RPE-1 and RPE-1 *TDP2*<sup>-/-</sup> Exon 4 cells or RPE-1 *TDP2*<sup>-/-</sup> Exon 4 transfected with TDP2, TDP2 $\beta$  or TDP2 $\gamma$  isoform. DAPI is used as nucleus counterstain.

TDP2 $\beta$  will most likely present in the endoplasmic reticulum or, to a lesser extent, the mitochondria (Table 4.1). These software also predicted the presence of a Signal peptide sequence and cleavage site between amino acid 22 and 23 of TDP2 $\beta$  exon 1. Signal peptides are recognised by the signal recognition particle which targets translation to occur directly to the endoplasmic reticulum lumen. After translation the protein will either stay in endoplasmic reticulum lumen, be modified and exported to cytoplasm or secreted from cell (Hegde and Bernstein, 2006).

To validate these findings, I immunostained the *TDP2*<sup>-/-</sup> RPE-1 cells transfected with TDP2 $\beta$  for markers of the mitochondria and endoplasmic reticulum (Figure 4.10). Although, TDP2 $\beta$  was located close to the mitochondria, as defined by MitoTracker Deep Red stain, we did not observe co-localisation of TDP2 $\beta$  and mitochondria. In contrast, co-staining using anti-protein disulphide isomerase (PDI), an endoplasmic reticulum (ER) marker, indicated that TDP2 $\beta$  signal was mostly localised to the ER, as predicted (Table 4.1 and Figure 4.10).

#### **4.2.6. Re-expression of TDP2 $\gamma$ isoform leads to recovery of DNA repair and survival**

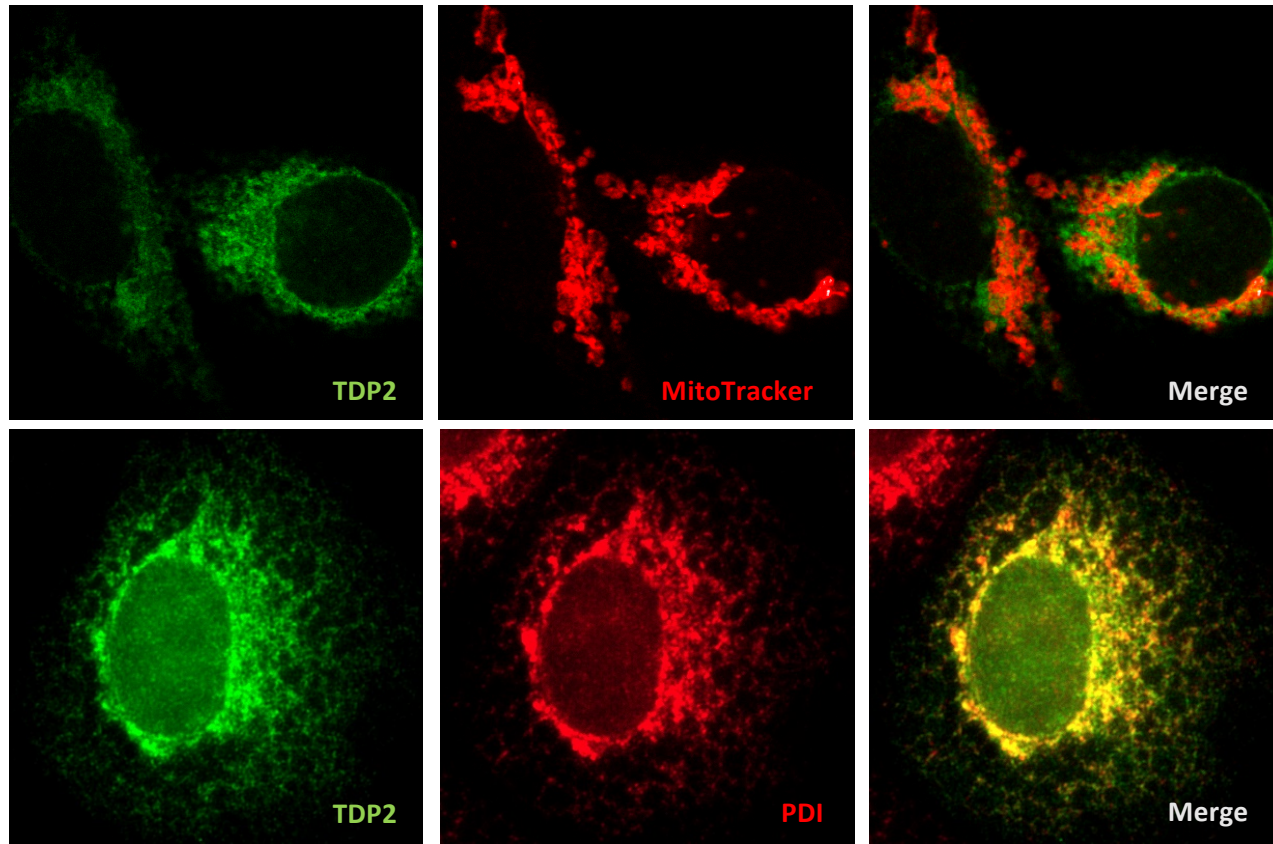
Since my experiments suggested that both TDP2 $\beta$  and TDP2 $\gamma$  are likely to exist, I decided to characterise their physiological relevance by transfecting expression constructs encoding these isoforms into the *TDP2*<sup>-/-</sup> RPE-1 cells in which exon 4 of TDP2 was gene edited. cDNA ORFs encoding wild type TDP2, TDP2 $\beta$  and TDP2 $\gamma$  were sub-cloned into pEGFP such that the encoded TDP2 proteins were tagged at the C-terminus with enhanced green fluorescent protein (eGFP). After transfection and antibiotic selection in medium containing G418, a stable expressing population of *TDP2*<sup>-/-</sup> RPE-1 was obtained for each construct. The expression of TDP2 on this population of cells stably expressing eGFP vector control, TDP2-GFP, TDP2 $\beta$ -GFP and TDP2 $\gamma$ -GFP was confirmed by western blotting (Figure 4.11.A). In addition, the western blotting of the tagged transgenes confirmed that TDP2 $\beta$  and TDP2 $\gamma$  have similar molecular weights (Figure 4.11.A).

Software	TDP2 $\beta$ localisation	Probability score
TargetP 1.1	Secretory pathway (endoplasmic reticulum)	p=0.859
MitoProt II v1.101	export to mitochondria	p=0.4338
MitoFates	export to mitochondria	P=0.01
SignalP 4.1	Signal peptide (endoplasmic reticulum)	p=0.642

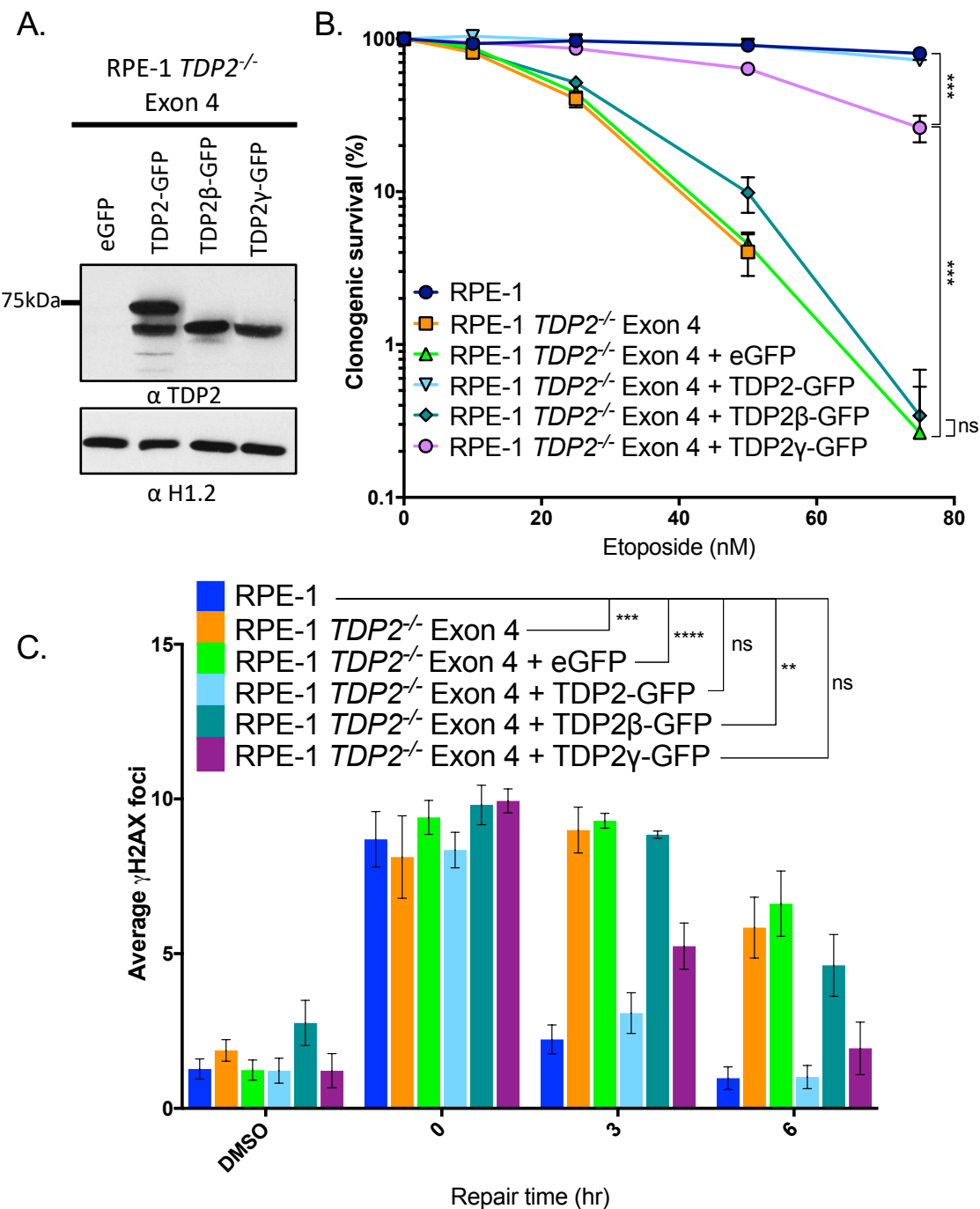
**Table 4.1 – Software prediction of TDP2 $\beta$  cellular localisation -**



RPE-1 *TDP2*<sup>-/-</sup> Exon 4 + TDP2 $\beta$



**Figure 4.10 – TDP2 $\beta$  localises mostly to endoplasmic reticulum.** – Immunostaining using anti-TDP2 in RPE-1 *TDP2*<sup>-/-</sup> Exon 4 cells transfected with TDP2 $\beta$ . MitoTracker Deep Red is used as mitochondria counterstain and anti-protein disulfide isomerase (PDI) is used as a marker for endoplasmic reticulum.



**Figure 4.11 – Exogenous expression of TDP2γ-GFP can complement *TDP2*<sup>-/-</sup> sensitivity and DSB repair delay . (A)** Western blot of against TDP2 in RPE-1 *TDP2*<sup>-/-</sup> Exon 4 complemented with eGFP vector control, TDP2-GFP, TDP2β-GFP and TDP2γ-GFP. Anti H1.2 used as loading control. **(B)** Clonogenic survival following continuous exposure to the indicated doses of etoposide in RPE-1, RPE-1 *TDP2*<sup>-/-</sup> Exon 4 and RPE-1 *TDP2*<sup>-/-</sup> Exon 4 complemented with eGFP vector control, TDP2-GFP, TDP2β-GFP and TDP2γ-GFP. **(C)** DSB repair kinetics was measured in same cells as (A), following 30 min exposure to 25 μM etoposide. Repair was measured by incubating cells for stated period in drug-free medium. Cells were gated to G1 according to DAPI profile. All statistical significance is given by two-way ANOVA. All data are mean ± SEM of three independent replicates. ns=p>0.05, \*\*p<0.01, \*\*\* p<0.001 and \*\*\*\*p<0.0001.

Following cell line complementation with TDP2 transgenes, I moved on to characterise the impact of the TDP2 isoforms in mediating cell survival and DSB repair in the presence of TOP2 DSBs. First, I conducted clonogenic survival assays employing continuous exposure to low concentrations of etoposide (Figure 4.11.B). As observed previously, *TDP2*<sup>-/-</sup> RPE-1 cells were hypersensitive to etoposide. As expected, hypersensitivity to etoposide was also observed in *TDP2*<sup>-/-</sup> RPE-1 expressing eGFP control vector, and was absent in *TDP2*<sup>-/-</sup> RPE-1 cells expressing wild type TDP2-GFP transgene. Intriguingly, whereas expression of TDP2 $\gamma$ -GFP restored etoposide resistance in *TDP2*<sup>-/-</sup> RPE-1 cells, albeit not completely, TDP2 $\beta$ -GFP failed to do so (Figure 4.11.B). To measure DSB repair I once again employed  $\gamma$ -H2AX immunostaining, following an acute treatment with etoposide in cells present in the G1 phase of the cell cycle. As expected, *TDP2*<sup>-/-</sup> RPE-1 cells and those transfected with empty vector displayed delayed DSB repair kinetics, as indicated by the persistence of elevated numbers of  $\gamma$ H2AX foci at 3 hr and 6 hr (Figure 4.11.C). Also, as expected, the number of  $\gamma$ -H2AX foci in cells transfected with TDP2-GFP were comparable to wild type RPE-1 cells at both time points. Interestingly, however, TDP2 $\gamma$ -GFP restored the rate of DSB repair to level between that of wild type and *TDP2*<sup>-/-</sup> RPE-1 cells at 3 hr, and restored repair to normal levels at 6 hr after drug removal. In contrast, consistent with the clonogenic survival assays, *TDP2*<sup>-/-</sup> RPE-1 cells expressing TDP2 $\beta$ -GFP exhibited rates of DSB repair similar to those harbouring empty vector (Figure 4.11.C).

### 4.3. CONCLUSIONS AND DISCUSSION

The results presented in this chapter demonstrate for the first time that in addition to TDP2, cells have two alternative TDP2 isoforms, TDP2 $\beta$  and TDP2 $\gamma$ . Our evidence for TDP2 $\beta$ , was based in the presence of its cDNA in A549 cells (Figure 4.8.C). In addition, TDP2 $\beta$  coding sequence seems to be conserved in multiple organisms (Figure 4.6.B). Whilst this work was in progress, an isoform of TDP2 with the same sequence as TDP2 $\beta$  was described in the mitochondria (Huang *et al.*, 2018). The published data on this isoform agrees with our observations, and previous publications, that multiple cell lines have a short TDP2 isoform that is detected at around 37 kDa on

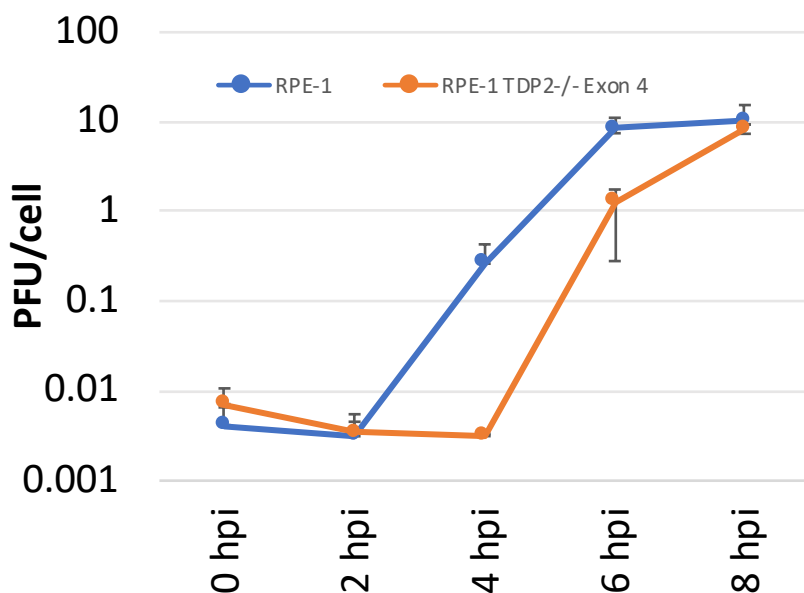
a western blot (Figure 4.1.A) (Li *et al.*, 2011; Huang *et al.*, 2018). Huang and colleagues data, confirms that TDP2 $\beta$  is present in murine tissues, as we predicted, and that this protein is mostly cytoplasmic (Figure 4.6.B and 4.10). Nevertheless, our prediction and experimental results indicate that this isoform is mostly localised to the endoplasmic reticulum contradictory to Huang et al. 2018 observation that this isoform is a mitochondrial protein (Table 4.1 and Figure 4.10). This discrepancy between our observations, could be due to the fact that the mitochondrion and endoplasmic reticulum have multiple point of contacts, so called mitochondrial-associated endoplasmic reticulum membranes (MAM), that could lead to contamination in the preparation of specific cellular fractions such as those used by Huang et al. 2018 (Rieusset, 2018; Eisenberg-Bord *et al.*, 2016). Furthermore, even as authors in Huang et. al claim a mitochondrial role for TDP2 $\beta$ , upon review of the data and personal communication with the authors it is clear that the most of this protein is not located in the mitochondria (Huang *et al.*, 2018).

The role of TDP2 $\beta$  in cells is not clear. The initial report of TTRAP from Pype et al. 2000 found that either in human or murine tissues, expression of multiple TDP2 mRNAs are present, including a short transcript detected specifically in human testis (Pype *et al.*, 2000). While TDP2 $\gamma$  would be coded by the same mRNA as TDP2, this shorter transcript could have been the first observation of TDP2 $\beta$ . It certainly warrants future analysis of tissue expression analysis to see if the expression of this isoform is tissue dependent. TDP2 $\beta$  localisation to the cytoplasm does not seem to support a direct role in the DNA damage response. Complementation of cells with TDP2 $\beta$  tagged with GFP failed to promote cell survival and mediate DSB repair after etoposide in RPE-1 *TDP2*<sup>-/-</sup> cells (Figure 4.11.B and 4.11.C). TDP2 was shown to be the protein previously denoted VPg unlinkase, which is required for the replication of some picornavirus genomes and thus cell infection (Virgen-Slane *et al.*, 2012; Maciejewski *et al.*, 2018). In order to verify that this was the case, our lab setup a collaboration with the lab of Dr. Bert Semler, who made the discovery that TDP2 was indeed the said VPg unlinkase. First, it was confirmed that the requirement of TDP2 to mediate viral infection in RPE-1 cells. Using poliovirus as a model of picornaviruses infection, it was shown that *TDP2*<sup>-/-</sup> RPE-1 cells targeted by gene editing at Exon 4 displayed a reduction of viral yield following infection

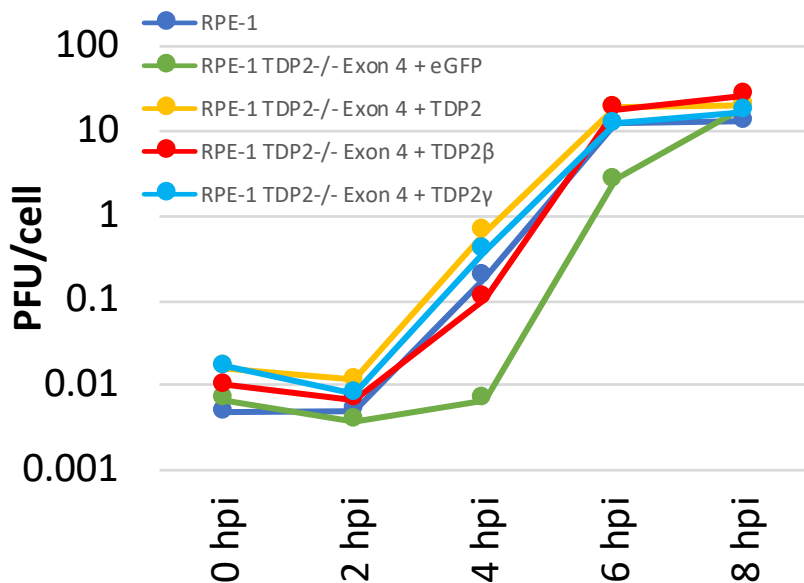
(Figure 4.12.A). Next, it was investigated whether the cytoplasmic localisation of TDP2 $\beta$  could support the viral unlinkase function of TDP2, even though it has not displayed a role in the DNA damage response. Using the same stable population of *TDP2*<sup>-/-</sup> RPE-1 Exon 4 cells as used for the DNA repair assays in Figure 4.11.A-C, it was shown that expression of all TDP2 isoforms was sufficient to increase viral yields (Figure 4.12B). This result confirmed that TDP2 $\beta$  was functional and could be exploited by the virus to mediate infection. Nevertheless, its canonical function in cells remains enigmatic. The endoplasmic reticulum, to our knowledge, does not contain RNA or DNA in which a phosphotyrosyl bond would require TDP2 activity. Thus, one could predict that TDP2 could be acting as a phosphatase at phosphorylated tyrosine residues instead. TDP2 protein does share homology with Mg<sup>2+</sup> dependent family of Inositol polyphosphate-5-phosphatases, indicating that this enzyme might mediate a similar function (Whisstock *et al.*, 2000; Rodrigues-Lima *et al.*, 2001). Unfortunately, when phosphotyrosine bovine serum albumin (BSA) was added to TDP2 in an *in vitro* activity assay, it could not compete with TDP2 activity on oligonucleotides (Zeng *et al.*, 2011). Indicating that at least that if TDP2 $\beta$  presents a phosphatase activity it might be target specific or dependent on protein interaction. Nevertheless, as the result of my own observations and the recent report by Huang *et al* 2018, I conclude that TDP2 $\beta$  is a bona fide isoform of TDP2 (Figure 4.3.A-B and 4.4.A-B)

The second identified isoform was TDP2 $\gamma$ . This isoform was first predicted to be encoded by alternative translation initiation on TDP2 coding sequence (Figure 4.7.A-C). TDP2 cytoplasmic localisation and the putative expression through alternative translation initiation was previously suggested but not investigated (Li *et al.*, 2011). The TDP2 $\gamma$  alternative start codon was found to be conserved in multiple organisms and present within the Kozak consensus supporting the expression of this isoform from TDP2 mRNA (Figure 4.7.B and 4.7.C). Using immunoprecipitation coupled with mass spectrometry, I was able to show that TDP2 $\gamma$  was indeed present in cells that lacked TDP2 following genome editing targeting exon 1 (Figure 4.8.A-C). I also observe that TDP2 $\gamma$  have a pan-cellular localisation, fitting with the loss of TDP2 nuclear localisation signal located in exon 1 and capability of proteins of up to 40 kDa to diffuse to the nucleus if a nuclear export signal is not present (Knockenbauer and Schwartz, 2016). The

A.



B.



**Figure 4.12 – TDP2 $\beta$  and TDP2 $\gamma$  isoforms can mediate viral infection** (A) Single cycle growth analyses carried out in RPE-1 and RPE-1 TDP2<sup>-/-</sup> Exon 4 cells following poliovirus infection. Cells and supernatant were collected every 2 h up to 8 hours post-infection (hpi). Virus yields were quantified by plaque assay on HeLa cells and represented as plaque forming units per cell (PFU/cell). Viral yields are plotted on a logarithmic scale and values represent the mean of triplicate experiments  $\pm$  SEM. (B) Same as (A) however RPE-1 and RPE-1 TDP2<sup>-/-</sup> Exon 4 complemented with eGFP, TDP2-GFP, TDP2 $\beta$ -GFP and TDP2 $\gamma$ -GFP were used instead. hpi= hours post infection [Contributing data from Autumn Candace Holmes from the lab of Dr. Bert Semler ]

availability of TDP2 $\gamma$  in the nucleus is also fitting with the observation that this isoform can mediate cell survival and DSB break repair in the absence of the canonical TDP2 (Figures 4.11.B and 4.11.C). In addition, I observed that expression of TDP2 cDNA (TDP2-GFP), even under an exogenous promoter, was sufficient to promote expression of a lower migrating protein band, which had similar molecular weight as TDP2 $\gamma$  (Figure 4.11.A). As a result of the all these observations, I believe that the TDP2 $\gamma$  is the isoform that is mediating repair and cell survival in cells *TDP2*<sup>-/-</sup> gene edited at exon 1 and that its deletion by targeting exon 4 accounts for the stronger phenotypes in exon 4 deleted cells, compared to exon 1 deleted cells (Figure 4.3.A-B and 4.4.A-B). As with TDP2 $\beta$ , I still do not know what the role of TDP2 $\gamma$  in cells would be. From the observations from Li et al. 2011, it seems that this isoform is overexpressed in multiple non-small-cell lung carcinoma (NSCLC) cell lines, such as A-549 which we used for its identification through MS (figure 4.8.B). The authors even suggest that stronger expression of the cytoplasmic signal, i.e. TDP2 $\gamma$  expression, in lung cancer tissue is correlative with disease progression (Li *et al.*, 2011). As a result, the mechanism by which cells control the expression of this isoform from TDP2 endogenous promoter seems of relevance to cancer treatment. As etoposide and other TOP2 targeting drugs are widely used to treat cancer, one could imagine that levels of TDP2 $\gamma$  in cells could correlate with increased resistance to such drugs. I believe that the initial observation by Li et al. 2011, in addition to results obtained here warrants the study of how this isoform might be regulated in tumorigenesis and its expression effects on patient outcome.

The presence of multiple TDP2 isoforms in cells also questions the importance of TDP2 N-terminal UBA domain for DNA repair response, since this domain is missing from the shorter TDP2 Isoforms. It was previously shown that complementation of *TDP2*<sup>-/-</sup> DT40 cells with TDP2 wild-type, truncated of its first 100 amino-acids, or the mutant F62R, resulted in only a partial rescue on etoposide clonogenic survival assay (Rao *et al.*, 2016). These results, were presented in support of TDP2 UBA domain role in binding di-ubiquitin chain linked via lysine (K) 48 or K63 role in the DNA damage response (Rao *et al.*, 2016). The four alpha-helices of TDP2 UBA are coded by its N-terminus, which means that both isoforms lack fully (TDP2 $\beta$ ) or partially (TDP2 $\gamma$ ) this domain altogether. Nonetheless, the complementation studies presented here suggests that when TDP2

catalytic activity is present, human cells do not require the ubiquitin binding domain for TDP2 functionality (Figure 4.11.B and 4.11.C). This discrepancy between results could be due to different requirement of TDP2 function in different organisms, as highlighted here, TDP2 sequence from *Gallus gallus* was the most evolutionarily divergent of the sequences analysed and showed little evidence of having more than one TDP2 isoform (Figure 4.6.B and 4.7.B).

Lastly, it is important to highlight that multiple isoforms of TDP2 were previously found in plants, also named TDP2 $\alpha$ , TDP2 $\beta$  and TDP2 $\gamma$  (Confalonieri *et al.*, 2013). There is little sequence conservation between human and plant TDP2, however all the C-terminal catalytic metal binding sites are conserved. Plant TDP2 $\alpha$  has two zinc finger Ras-related Nuclear (Ran) binding protein 2 (ZNF RanBP2) type domain at its N-terminus, plant TDP2 $\beta$  has one of the same motif, while they are altogether absent in plant TDP2 $\gamma$  (Confalonieri *et al.*, 2013). The function of these motifs are not clear in plant TDP2, but these domains were previously reported to be involved in binding RanGDP and mediate nuclear transport and splicing in humans (Steggerda and Paschal, 2002). Nevertheless, they do not seem essential as multiple species lack one or more of these isoforms (Confalonieri *et al.*, 2013). The observations that plants isoforms are absent in multiple organisms match the predictions for *Gallus gallus*, which seems to lack both TDP2 $\beta$  and TDP2 $\gamma$  isoforms identified in this chapter. It seems that different species have evolved different requirement for TDP2 isoforms, so only better understanding of their biological roles could clarify their apparent redundancy.

In conclusion, the results of this chapter reveal that human cells have multiple TDP2 isoforms. Here I demonstrated, using specific CRISPR-Cas9 deletion of TDP2 targeting either exon 1 or exon 4, that multiple cell lines are more sensitive to etoposide once exon 4 is targeted by genome editing. The difference in behaviour from TDP2<sup>-/-</sup> exon 1 and exon 4 cells was found to be the result of two putative TDP2 isoforms present in cells, namely TDP2 $\beta$  and TDP2 $\gamma$ . Although the function of these isoforms is not known their possible relevance for viral infection and tumorigenesis warrants further studies to better understand their roles in cells.



## 5. CHAPTER FIVE – ESTABLISHING AN ASSAY FOR TYROSYL DNA PHOSPHODIESTERASE ACTIVITY IN WHOLE CELL EXTRACTS

### 5.1. INTRODUCTION AND AIMS

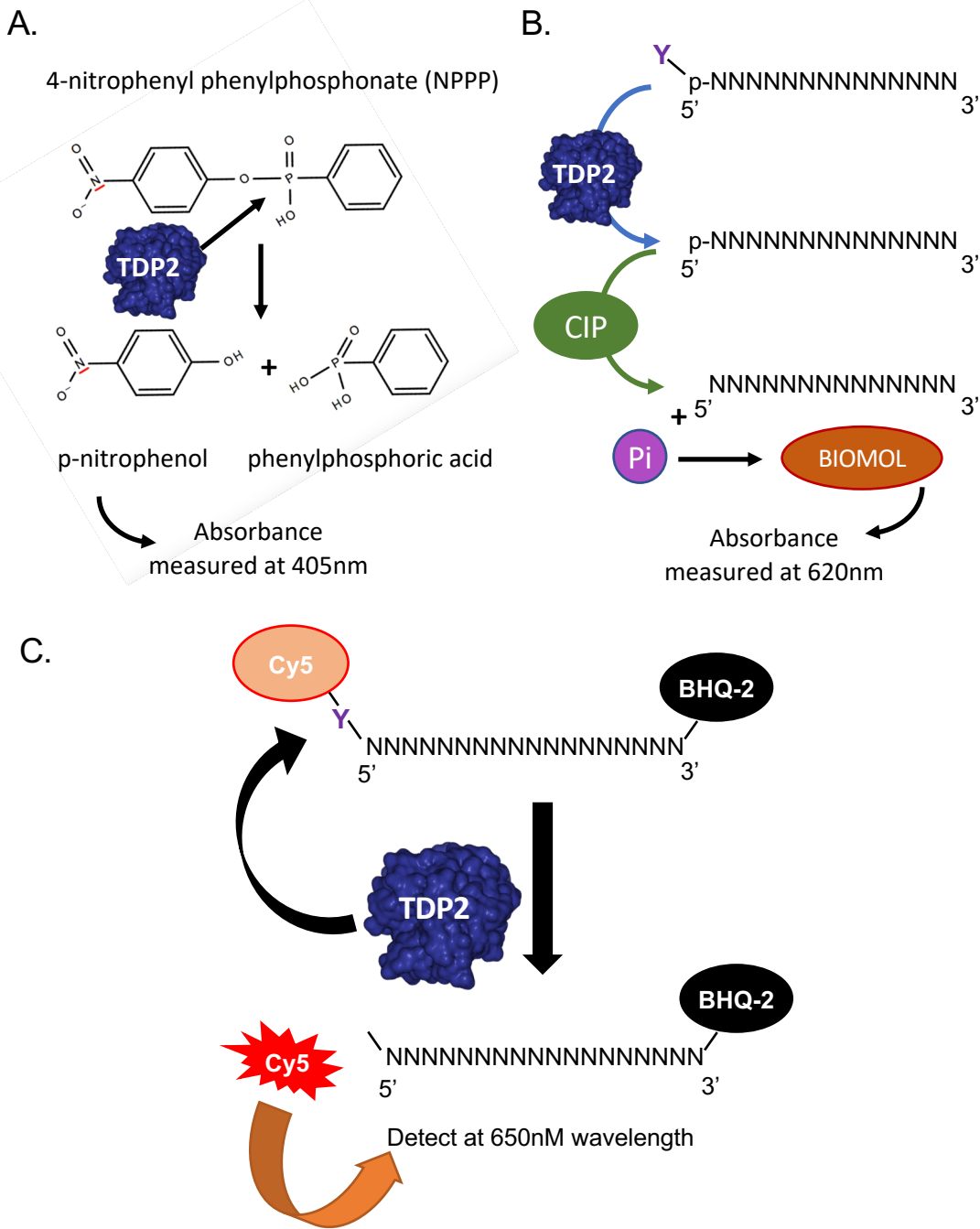
One of the emerging hallmarks of cancer is genomic instability (Negrini *et al.*, 2010; Hanahan and Weinberg, 2011). In addition, the DNA damage response is the most prevalent target of radiotherapy and chemotherapy (Jackson and Bartek, 2009). The ubiquitous requirement of TOP2 in replication and transcription makes this a target in multiple tumours (Pommier *et al.*, 2016). TOP2 ‘poisons’ such as etoposide and doxorubicin promote tumour cell death by inducing abortive TOP2 cleavage complexes that are comprised of protein-linked DSBs and are frequently used in chemotherapy (Nitiss, 2009b; Pommier *et al.*, 2010).

Nevertheless, these abortive TOP2 cleavage complexes require cellular pathways of DNA repair for their removal. As described earlier, it has been proposed that TDP2 is the enzyme primarily responsible for repairing such lesions, although it is clear that alternative repair pathways are present in cells (Zagnoli-Vieira and Caldecott, 2017). As a result, targeting TDP2 in cancer could provide an alternative treatment to current chemotherapeutic strategies that rely on DNA damage induction by targeting DNA damage repair instead. Nonetheless, the evidence for TDP2 role in cancer has so far been conflicting. On one hand, TDP2 expression was reportedly elevated in lung carcinoma tissue and non-small-cell lung carcinoma (NSCLC) cell lines (Li *et al.*, 2011). In the same study, exogenous expression of TDP2 was sufficient to accelerate cancer cell line growth and its depletion resulted in growth defects in cancer cell line and cancer xenograft studies (Li *et al.*, 2011). In addition, TDP2 expression was found to be controlled by the Tumour Protein 53 (TP53) and to be upregulated in TP53 mutant tumours possibly contributing to resistance to topoisomerase 2 targeting chemotherapy (Do *et al.*, 2012). TDP2 overexpression in lung carcinoma was also correlated with TP53 status and found to be present in TP53 positive tumours (Do *et al.*, 2012). On the other hand, exogenous TDP2 expression was reported to delay cell growth and promote G2/M cell cycle arrest in the osteosarcoma cell lines U-2-OS and SAOS-2 (Zhou *et al.*, 2013). I have not observed any alteration of cell cycle between our multiple CRISPR-Cas9 *TDP2*<sup>-/-</sup> clones, other than that observed due to clonal variation (unpublished observations).

Nevertheless, we have recently reported that if TDP2 is depleted in the breast cancer cell line MCF-7, these cells accumulate more transcription induced DSB in the presence of the hormone estradiol (Gómez-Herreros *et al.*, 2017). This is consistent with previous reports that TDP2 is required for repair of DSBs induced by TOP2 $\beta$  during transcription of androgen and estrogen receptor driven genes, highlighting the possible relevance of TDP2 to prostate and breast cancers (Ju and Rosenfeld, 2006; Haffner *et al.*, 2010; Gómez-Herreros *et al.*, 2017). Taking into account all these observations, we put forward the model that TDP2 could promote survival of tumour cells by dealing with the increased TOP2cc lesions resulting from increased hormonal driven transcription and replication rates, and that its inhibition could be exploited in such tumours.

The only reported and verified TDP2 inhibitors are the deazaflavin series of compounds (Raoof *et al.*, 2013). Deazaflavins bind to a hydrophobic shelf in the enzyme, essentially blocking access of DNA into the active site (Hornyak *et al.*, 2016). However, we have not been able to confirm the efficacy of these inhibitors in cells (unpublished observations). To enable rapid biochemical screening for TDP2 inhibitors that are efficacious in the context of a complex protein environment such as that present in cells, I attempted to establish a cell free assay in which TDP2 activity is measured in whole cell protein extract.

To date, biochemical TDP2 activity has so far only been described using purified enzyme and either radiolabelled or fluorophore labelled oligonucleotides harbouring a 5'-phosphotyrosyl (Ledesma *et al.*, 2009; Zeng *et al.*, 2011). Whilst the oligonucleotide substrate and products of TDP2 activity can be distinguished by gel electrophoresis, this is time consuming and not appropriate for high-throughput screening of many thousands of potential inhibitors. Two high-throughput fluorescence-based TDP2 assays have been reported in the literature. The first employs a direct measurement of fluorescence using the 4-nitrophenyl phenylphosphonate (NPPP) molecule (Figure 5.1.A) (Thomson *et al.*, 2013). Once TDP2 is incubated with NPPP, it hydrolyses the bond between p-nitrophenol and phenylphosphoric acid thereby releasing the p-nitrophenol moiety, the fluorescence of which can be detected at 405nm wavelength (Figure 5.1.A) (Thomson *et al.*, 2013). The second approach that has been employed was an indirect measurement of TDP2 activity, measuring the release of inorganic phosphate by alkaline phosphatase (CIP). This assay was based on fluorescence of BIOMOL green at 620nm



**Figure 5.1 – Overview of the 5'-Tyrosyl DNA phosphodiesterase (5'-TDP) assay.** (A) The first 5'-TDP assay developed by Thomson et al. 2013 employs fluorescence measurement of p-nitrophenol, a product of TDP2 activity on 4-nitrophenyl phenylphosphonate (NPPP). (B) The second 5'-TDP assay developed by Thomson *et al.* 2013, measures the release of inorganic phosphate (Pi) from a single-stranded oligodeoxyribonucleotide by alkaline phosphatase (CIP), using the BIOMOL green phosphate sensor reagent, to indirectly measure TDP2 activity. (C) A novel high throughput 5' TDP assay that employs a single-stranded oligodeoxyribonucleotide with a Cy5 moiety linked to the 5'-terminus by as 5'-phosphotyrosyl bond. The 3' terminus harbours a black hole quencher 2 (BHQ-2) that absorbs fluorescence of the Cy5 molecule, until the Cy5 moiety is released from the 5'-terminus (e.g. by TDP2) allowing detection of Cy5 fluorescence at 650nm.

wavelength as it reacted with the released phosphate, which was only available for removal by CIP if TDP2 had first hydrolysed the phosphotyrosyl bond linking a tyrosine moiety to the phosphate on the 5'-terminus (Figure 5.1.B) (Thomson *et al.*, 2013).

Although these fluorescence approaches are appropriate for high-throughput screening (HTS), neither combined a direct assay with a DNA based substrate. Consequently, as part of a project involving the Caldecott group and Sussex Drug Discovery, Dr. Sarah Walker in the Sussex Drug Discovery Group (SDDG) developed a system based on fluorescence resonance energy transfer (FRET) (Figure 5.1.C). FRET is the physical process by which energy is transferred from an excited chromophore, the donor, to a second chromophore, the acceptor (Clegg, 1995). It is essential for FRET that the emission spectra of the donor overlaps with the absorption spectra of the acceptor chromophore (Clegg, 1995; Didenko, 2001). Using this approach, Dr. Walker designed an assay employing a single-stranded oligonucleotide on which the 5' terminus possesses a cyanine 5 (Cy5) fluorophore-coupled tyrosine moiety (Y) and the 3' terminus is coupled to a black hole quencher 2 (BHQ-2) moiety (Figure 5.1.C). In the presence of tyrosyl DNA phosphodiesterase activity, the tyrosyl bond is hydrolysed and the Cy5 moiety is released from the black hole quencher resulting in fluorescence that can be detected in real time at 650nm wavelength.

#### **5.1.1. Aim**

In this chapter, I will further develop the new TDP high throughput screen to enable its use with whole cell extracts, thereby allowing analysis of the efficacy of putative new TDP2 inhibitors during drug screening and also the assessment of TDP2 activity in cell extracts of the CRISPR-Cas9 cell lines generated on Chapters 3 and 4. I will investigate if this assay is suitable to measure TDP activity, not only in purified TDP2 protein but also on potential patient samples. In addition, I will investigate the putative utility of TDP2 inhibitors in breast cancer using bioinformatic tools to interrogate existing data sets, as part of a TDP2 target validation.

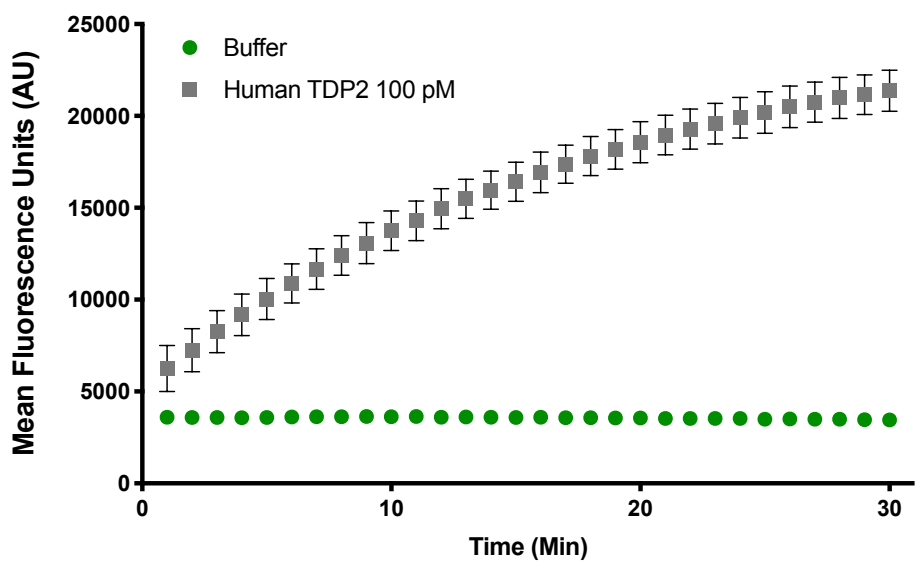
## **5.2. RESULTS**

### 5.2.1. 5'-Tyrosyl DNA phosphodiesterase (5'-TDP) assay

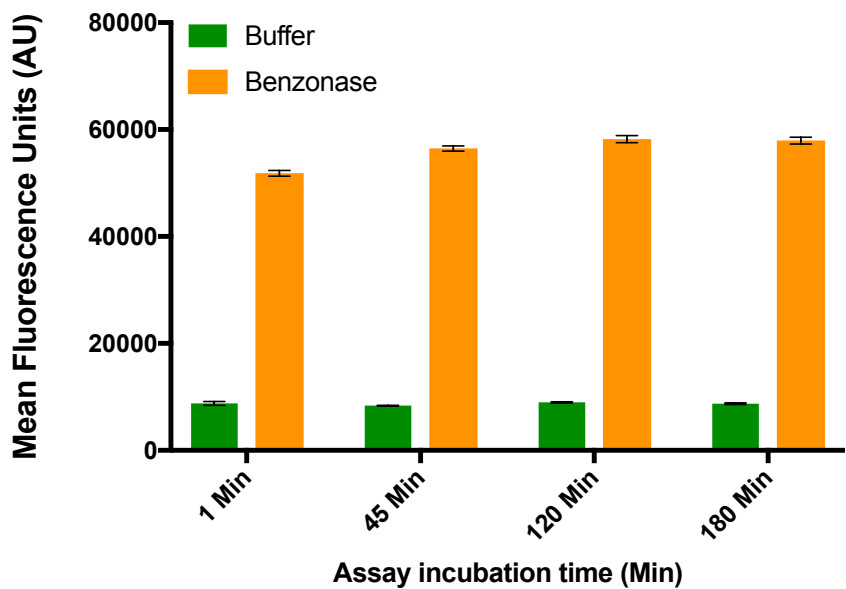
I first confirmed that the new fluorescence based assay was functional, using purified human TDP2 protein (Figure 5.2.A). Incubation of the substrate with 100 pM of purified TDP2 enzyme lead to a steady increase in the fluorescence signal detect at 650nm over 30 minutes, confirming the efficacy of the assay. A control reaction containing only sample buffer and conducted in parallel failed to result in fluorescent signal, indicating that the oligonucleotide substrate was stable during the time of the assay (Figure 5.2.A). As well as detecting TDP2 activity, I reasoned that this ought to be able to detect DNA endonuclease activities that cleave the substrate, because these too will separate the fluorophore and the black hole quencher. Although not a problem in experiments employing purified protein, this could become an issue when measuring TDP2 activity in whole cell extracts. In addition, however, it may also allow the detection of bona fide repair endonucleases that process TOP2 cleavage complexes. To address this issue, the TDP2 oligonucleotide substrate was incubated with Benzonase endonuclease (250 units). Notably, most of single-strand oligonucleotide substrate was cleaved within one minute, and was completed within 120 minutes (Figure 5.2.B). This result indicated that the 5'-TDP substrate is sensitive not only to TDP activity but also to cellular endonucleases.

Having confirmed that the 5'-TDP assay was working I next attempted to employ it to measure TDP activity in whole cell extracts (WCE). Given that endonucleases can also process this substrate I decided to first titrate the amount of cell extract required for optimal activity. A titration of WCE from wild type RPE-1 cells indicated that while 5'-TDP activity could be detected at multiple concentrations, only the data set obtained with 15 µg produced a steady increase in fluorescence that was clearly above background over a 30 minute incubation period (Figure 5.3.A). In similar experiments using WCE from A549 cells, I observed much higher levels of 5'-TDP activity, with 15 µg of total cellular protein resulting in ~2-fold more product over the same incubation period, and with 1.5 µg of A549 total cellular protein generating a similar amount of product over 30 min as that generated by 15 µg of whole cell protein from RPE-1 cells (Figure 5.3.B). This difference was made evident by calculating the rate of reaction (RoR), using the time of assay duration and mean fluorescence units (MFU) in the formula

A.

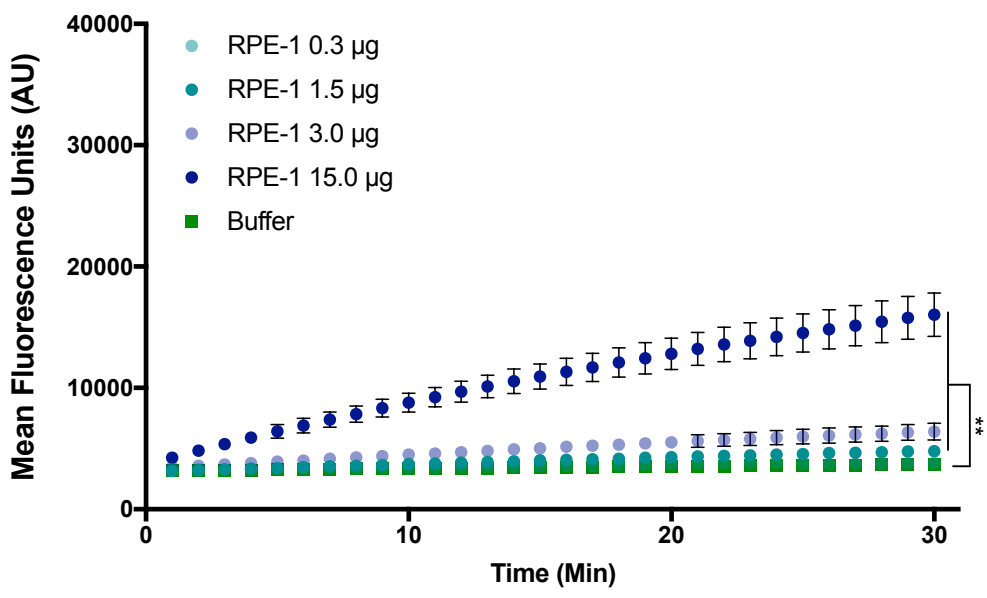


B.

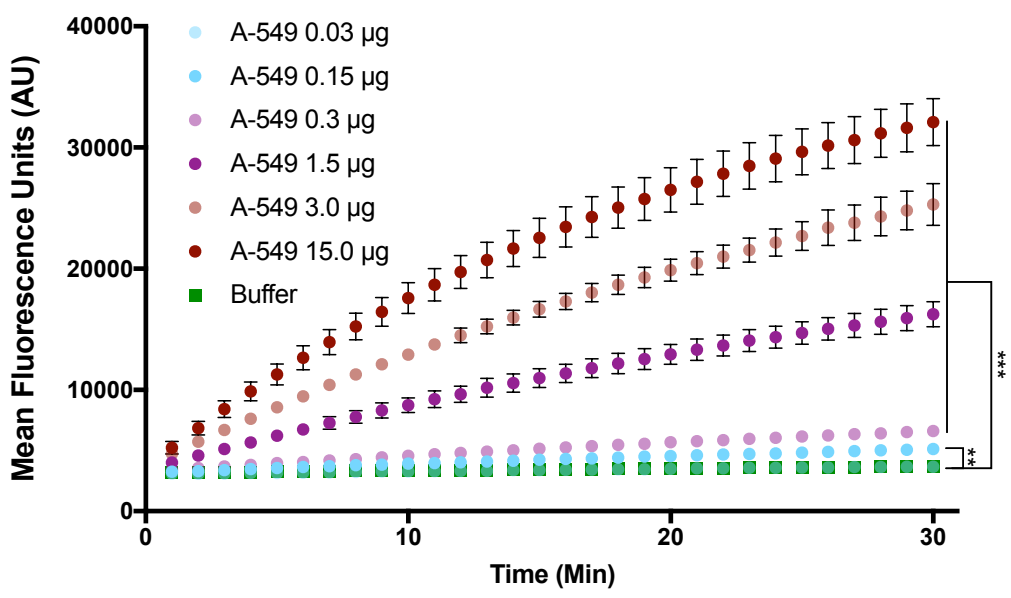


**Figure 5.2 – 5'-TDP activity of TDP2 protein.** (A) TDP2-dependent 5'-TDP activity measured using the oligo substrate described in Figure 5.1.C. Purified human TDP2 protein (100 pM) was incubated with 40 nM oligo substrate and Cy5 fluorescence measured in real time. Buffer control was employed to quantify the background fluorescence. (B) The 5'-TDP substrate is susceptible to endonuclease activity. 5'-TDP substrate (40nM) was incubated with 250U of Benzonase endonuclease, resulting in rapid increase in Cy5 fluorescence. Data are the mean (+/- SD) of three technical replicates.

A.



B.



**Figure 5.3 – 5'-TDP activity in whole cell extracts. (A)** Real-time measurement of 5'-TDP activity present in the stated amounts of whole cell extract from RPE-1 cells. Buffer control was used to quantify background fluorescence. **(B)** As (A) employing A549 cell extract. Data are mean (+/- SD) of three technical replicates. Statistical significance was established by t-test at the 30 minute time point. \*\*p<0.01 and \*\*\*p<0.001.

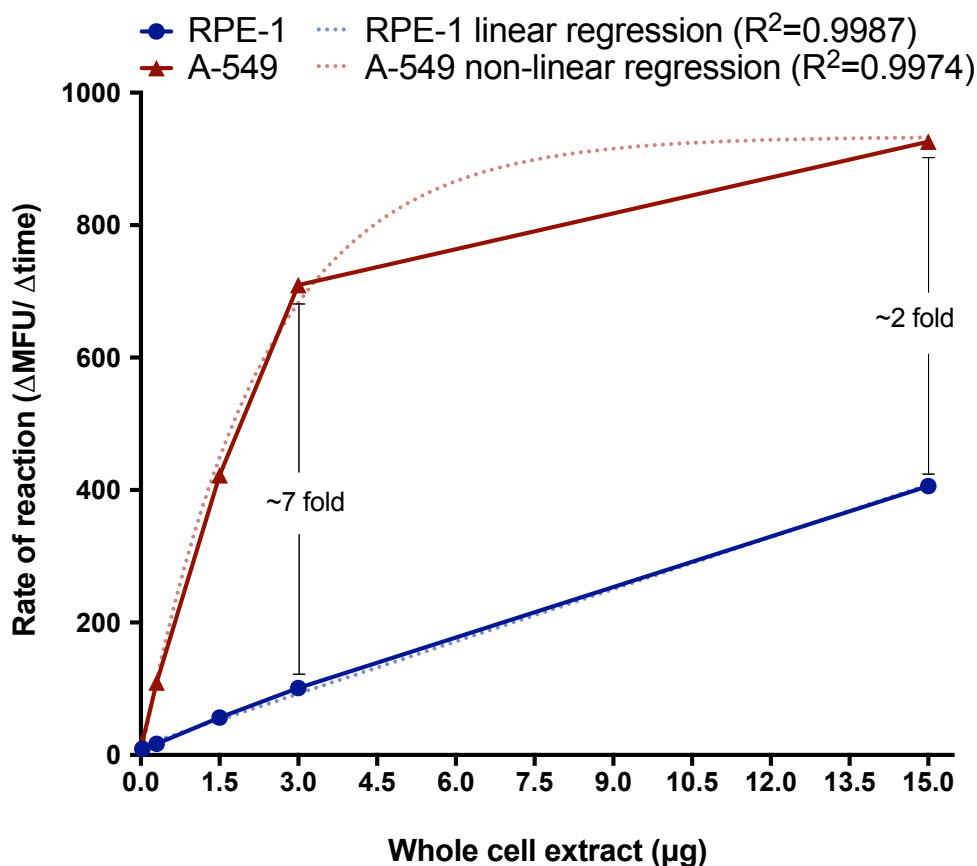
$RoR = \Delta MFU (MFU_{final} - MFU_{initial}) / \Delta time (time_{final} - time_{initial})$ . While RPE-1 cells presented a linear rate of reaction over the WCE concentrations used, A-549 cells presented an exponential rate of reaction that resulted in the final 2 fold difference (Figure 5.4). The rate of reaction also indicated that A-549 cells had a linear rate of reaction on concentrations of up to 3  $\mu$ g, indicating that the oligonucleotide concentration used was limiting to higher concentrations of WCE. Comparison of rates of reaction in this linear portion of the graph does indicated that A-549 cells have approximately 7 fold more 5'-TDP activity than RPE-1 cells.

To ensure that the signal that we were detecting in whole cell extracts in the 5'-TDP assay was TDP2, I measured 5'-TDP activity in WCE from *TDP2*<sup>-/-</sup> RPE-1 cells. Reassuringly, WCE from the *TDP2*<sup>-/-</sup> RPE-1 cells in which exon 1 was gene edited exhibited markedly less (1.5 fold) 5'-TDP activity, which was further reduced in WCE from *TDP2*<sup>-/-</sup> RPE-1 cells in which exon 4 was gene edited which possessed <20% of the activity present in wild type cells (Figure 5.5.A). This difference in activity between the two gene edited cell lines supported my identification by western blotting of the additional two TDP2 isoforms in the exon 1 gene edited clone (see Chapter 4). These data suggest that the TDP2 isoforms in the RPE-1 cells were responsible for approximately 50% of TDP2 activity (Figure 5.4.A). Additionally, I noted in these experiments that there was some residual 5'-TDP activity even in the *TDP2*<sup>-/-</sup> RPE-1 cells in which exon 4 was gene edited, presumably reflecting cellular nuclease activity, a hypothesis further discussed below.

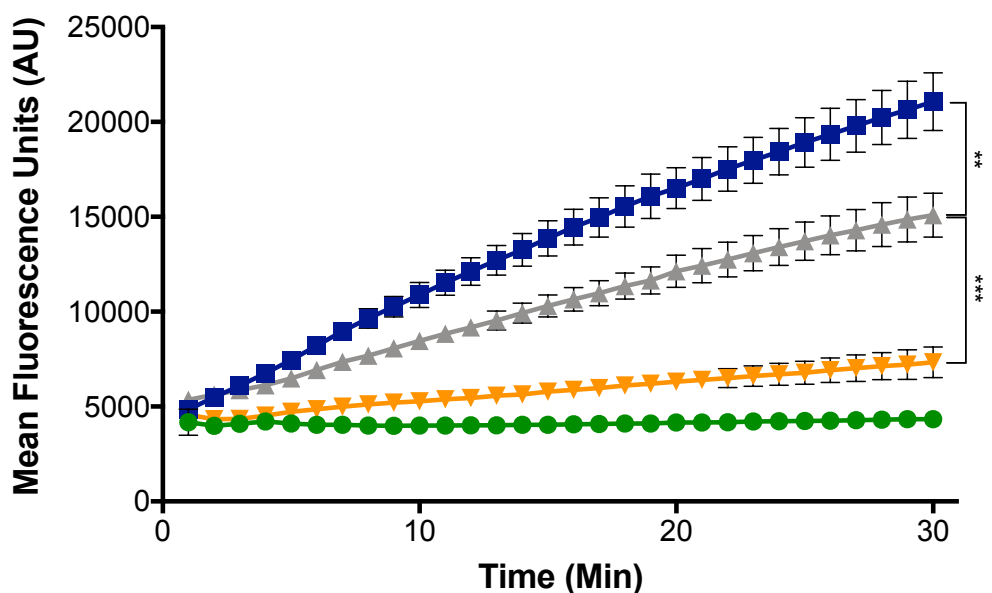
### 5.2.2. 3'-Tyrosyl DNA Phosphodiesterase (3'-TDP) Assay

I next determined whether the new fluorescence based assay was capable of measuring 3'-TDP activity, which is conducted by TDP1. The rationale was that this would provide an excellent control for my TDP2 assay, ensuring that the differences observed when assaying 5'-TDP activity were not the result of differences in extract amount or quality. To do this, I employed a similar oligonucleotide substrate in which the position of the fluorophore and black hole quencher were switched, generating a substrate with a fluorophore-linked 3'-phosphotyrosyl bond and a 5'-linked black hole quencher (Figure 5.6.A). Using both the 5'-TDP and 3'-TDP substrates described above





**Figure 5.4 – Rate of reaction of 5'-TDP activity in whole cell extracts.** Data from Figure 5.3 is presented as the rate of reaction for whole cell extract from RPE-1 and A-549 cells. RPE-1 cells present a linear rate of reaction, while the best-fit for the activity seen in A-549 cells was an exponential rate. Coefficient of determination ( $R^2$ ) for each regression are indicated.



■ RPE-1 RoR=562.1± 8.835;  $R^2=0.9931$

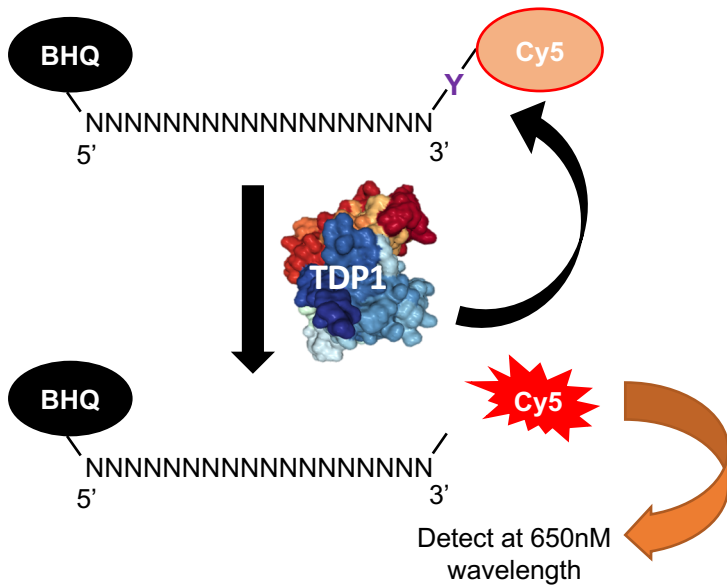
▲ RPE-1 *TDP2*<sup>-/-</sup> Exon 1 RoR=349.8±2.749;  $R^2=0.9983$

▼ RPE-1 *TDP2*<sup>-/-</sup> Exon 4 RoR=103.5±1.419;  $R^2=0.9948$

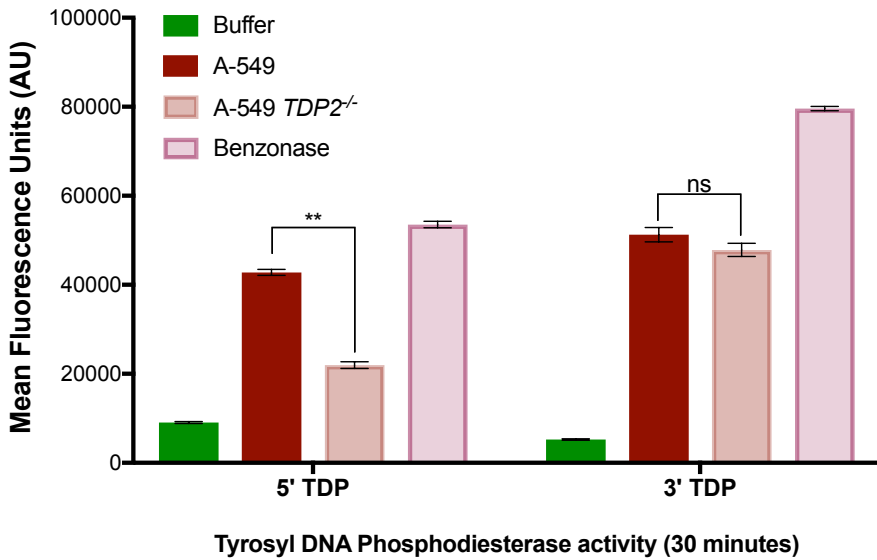
● Buffer

**Figure 5.5 – 5' TDP activity is TDP2 dependent. (A)** Real-time measurement of 5'-TDP activity present in 15  $\mu$ g of whole cell extract from wild type RPE-1, and from RPE-1 cells in which TDP2 exon 1 or exon 4 were targeted by CRISPR-Cas9 gene editing. The Rate of reaction (RoR) and coefficient of determination ( $R^2$ ) for linear regression are given. Buffer control was used as measurement of background fluorescence. All data are the mean (+/- SD) of three technical replicates. Statistical significance is given by t-test at 30 minute time point. \*\* $p < 0.01$  and \*\*\* $p < 0.001$ .

A.



B.



**Figure 5.6 – 3'-TDP assay.** (A) Overview of the 3'-TDP assay that employs a single-stranded oligodeoxynucleotide substrate in which Cy5 is linked to the 3'-terminus via a phosphotyrosyl bond. The 5' terminus harbours a black hole quencher 2 (BHQ-2) moiety that absorbs the fluorescence of Cy5 until the latter is released (e.g. by TDP1) from the 3' terminus, allowing fluorescence to be detected at 650nm. (B) 5'- or 3'-TDP activity was measured in 15 µg of whole cell extract from A549 or A549 *TDP2*<sup>-/-</sup> cells following 30 minutes incubation with stated substrate. Benzonase endonuclease (250U) was used as positive control and indicator of maximum Cy5 release. Buffer control was used as a measurement of background fluorescence. WCE data are the mean (+/- SE) of of three independent experiments. The buffer and benzonase data are the mean (+/- SE) of three technical replicates. Statistical significance was established by t-test. ns=p>0.05 and \*\*p<0.01.alone

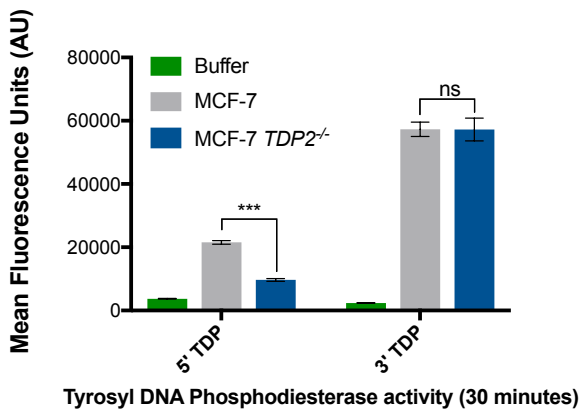
in reactions containing WCE from wild type or *TDP2*<sup>-/-</sup> A549 (Exon 4) cells, I was able to reproduce my finding that *TDP2*<sup>-/-</sup> A549 cells possess reduced 5'-TDP activity (Figure 5.6.B) and confirm that these cell lines possess similar levels of 3'-TDP activity. Once again, I employed Benzonase endonuclease to detect the level of fluorescence associated with complete conversion of substrate to the product (Figure 5.6.B).

### 5.2.3. Versatility and sensitivity of the TDP assays

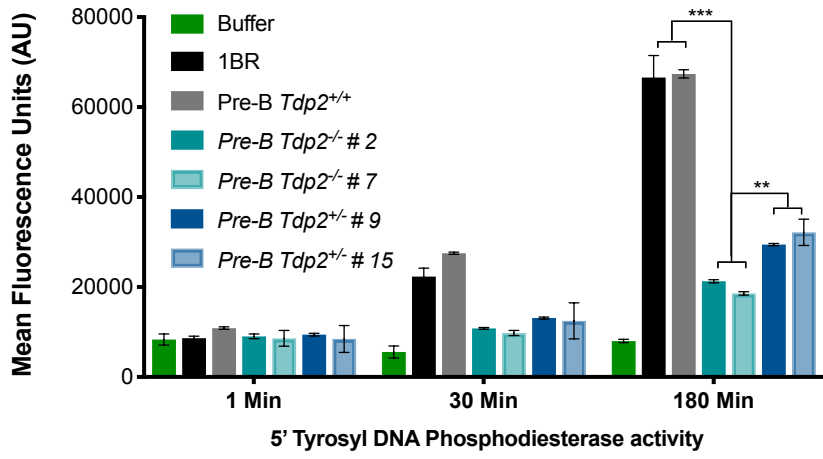
Since both the 5'- and 3'-TDP assays were shown to work well with A549 and RPE-1 whole cell extracts, I next decided to test whether these assays can be used to quantify TDP levels in other cell types. First, I employed the breast cancer cell line, MCF-7, and measured both 5'-TDP and 3'-TDP activities (Figure 5.7.A). With wild type MCF-7 WCE, I obtained similar levels of 5'-TDP activity as with RPE-1 cells. This activity was again reduced in the absence of TDP2 (MCF-7 *TDP2*<sup>-/-</sup>), whereas both wild type and *TDP2*<sup>-/-</sup> WCE possessed similar levels of 3'-TDP activity (Figure 5.7.A). Next, we measured 5'- and 3'-TDP activity in human fibroblasts and mouse pre-B cells (Figure 5.7.B-C). Multiple clones of murine Pre-B cells lacking Tdp2 were created using CRISPR-Cas9 by the group of Dr. Andre Nussenzweig and following genomic characterisation were found to have *Tdp2* at either homozygous or heterozygous status (personal communication). *Tdp2*<sup>+/-</sup> Pre-B cells WCE had similar levels of TDP2 activity as 1BR primary human fibroblasts cells (Figure 5.7.B). This activity was reduced in both *Tdp2*<sup>+/-</sup> and *Tdp2*<sup>-/-</sup> WCE (Figure 5.7.B). The assay was sufficiently sensitivity to distinguish between *Tdp2*<sup>+/-</sup> and *Tdp2*<sup>-/-</sup> WCE, though longer incubation was required. Importantly, residual 5'-TDP activity was again detected in the *Tdp2*<sup>-/-</sup> defective cells. Since these cells were generated by targeted deletion of Tdp2, this observation supports the idea that this residual activity reflects nucleases in the cell extract, rather than residual TDP2 activity. Once again, 3'-TDP activity was not significantly different between the different mouse cell extracts, confirming validity of observations with the 5'-TDP substrate (Figure 5.7.C).

Lastly, I examined whether I could detect the impact of the deazaflavin series of TDP2 inhibitor in this assay. As discussed earlier, this assay was initially developed to support current efforts at University of Sussex Drug Discovery Group to identify novel TDP2 inhibitors. To verify the efficacy of the only currently available inhibitor in WCE, I

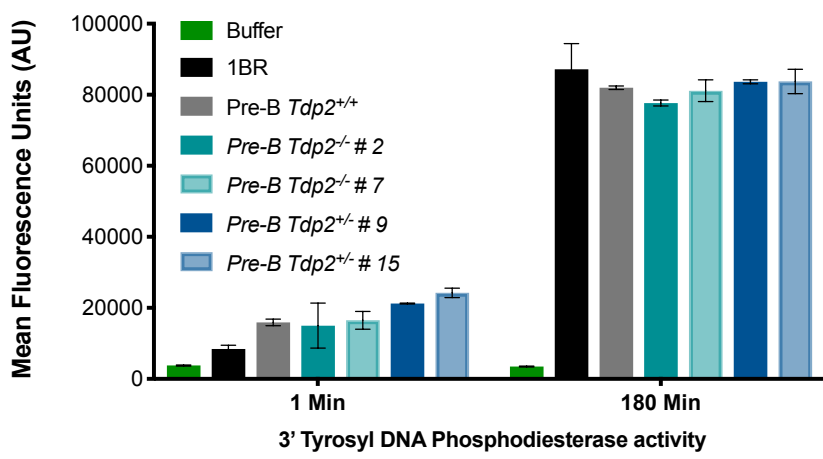
A.



B.



C.



**Figure 5.7 – The 5'-TDP assay is versatile and sensitive. (A)** 5' or 3' TDP activity was measured in 15 µg of whole cell extract (WCE) from the breast cancer cell line MCF-7 and from *TDP2*<sup>-/-</sup> MCF-7 cells following 30 minutes incubation with stated substrate. **(B)** 5'-TDP and **(C)** 3'-TDP activity was measured at the stated times in WCE from 1BR primary fibroblasts (5 µg) and mouse pre-B cells (5 µg). Pre-B cell clones (clone number stated) in which one (heterozygous) or both (homozygous) *TDP2* alleles were deleted by homologous recombination mediated gene targeting were employed to confirm the sensitivity of the 5'-TDP assay to gene dose. Buffer control was used as measurement of background fluorescence. All data are the mean (+/- SD) of three technical replicates. Statistical significance is given by t-test. ns=p>0.05, \*\*p<0.01 and \*\*\*p<0.001.

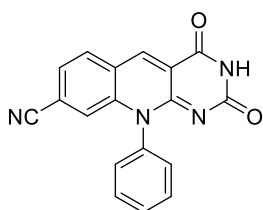
employed the compound UoS12248; a deazaflavin that has been characterised extensively for its inhibitory activity with purified TDP2, including at the structural level (Figure 5.8.A) (Hornyak *et al.*, 2016; Marchand *et al.*, 2016). Strikingly, pre-incubation of WCE from wild type A549 cells with 50  $\mu$ M of UoS12248 inhibited TDP2 activity to similar levels observed in A549 *TDP2*<sup>-/-</sup> cells, confirming that the compound is efficacious in a complex mixture of cellular proteins (Figure 5.8.B). Having confirmed its activity, I next measured its inhibitory profile at different concentrations. Using the extracts from A549 *TDP2*<sup>-/-</sup> cells as our baseline activity, I found that the inhibitory concentration (IC<sub>50</sub>) of this compound is 4.88  $\mu$ M in WCE when compared to the IC<sub>50</sub> of 0.32  $\mu$ M observed with purified protein (Figure 4.8.C) (IC<sub>50</sub> data for purified protein from SDDG). The ability of the deazaflavine to reduce activity in wild type extracts to only the same level as that in *TDP2*<sup>-/-</sup> cells is informative, because it supports the idea that this residual activity is the result of enzymes other than TDP2, a notion confirmed in additional experiments in Chapter 6.

#### 5.2.4. TDP2 in Cancer

Finding a working TDP2 inhibitor is not the only step necessary to target this enzyme in cancer, a suitable tumour target is also required. There are multiple databases containing tumour data that could help us to predict such suitable tumour candidates. For example, The Catalogue of Somatic Mutations In Cancer (COSMIC) latest release contains data from over one million samples, with mutation and copy number variation and gene expression data (Forbes *et al.*, 2017; COSMIC, 2018). Using COSMIC data, we found that TDP2, while not frequently mutated or amplified, is over-expressed in multiple malignancies, of which liver (18.23%), skin (12.26%) and breast (10.69%) tumours are the most frequently over-expressed (Table 5.1). Gene over-expression is common characteristic of oncogenes, and thought to be linked to oncogene addiction, which is the maintenance of a trait that is beneficial to the tumour growth (Sharma and Settleman, 2007). Using this hypothesis, we investigated if TDP2 expression could be behaving in a such way using patient survival data.

Although we could not find easily accessible data from liver or skin tumours we observed that TDP2 overexpression could be a potential marker in refractory breast

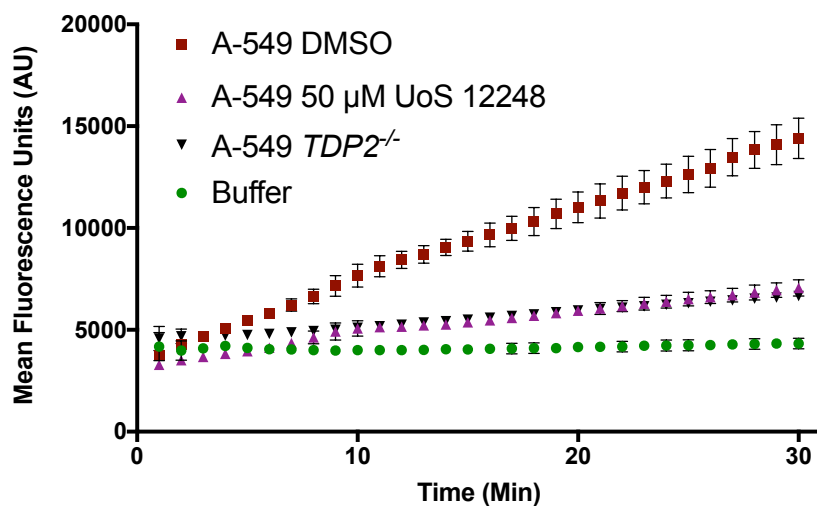
A.



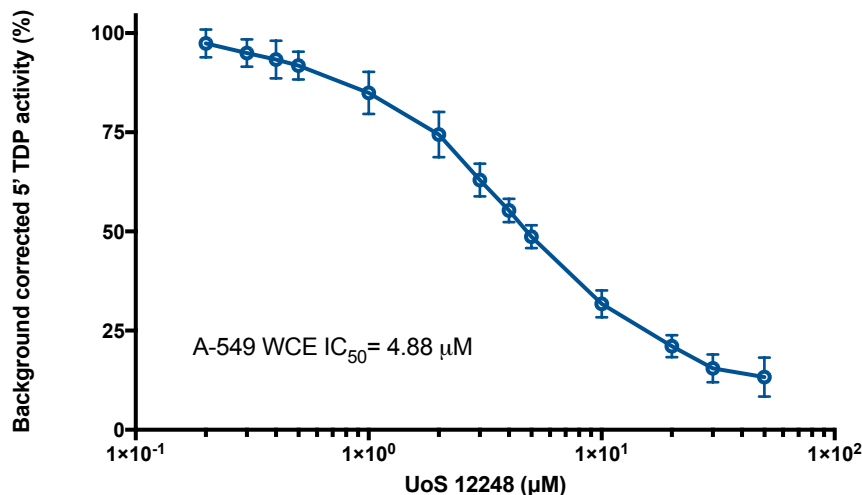
UoS 12248

EC<sub>50</sub> : 90 nM(Hornayak *et al.* 2016)

B.



C.



**Figure 5.8 – Measuring 5' TDP activity in the presence of a TDP2 inhibitor. (A)** Chemical structure of the TDP2 inhibitor UoS12248 described by Hornayak *et al.* 2016 **(B)** Real-time measurement of 5'-TDP activity in WCE from A549 cells for the stated time following pre-incubation of 30 minutes with either DMSO (vehicle) or 50 μM of UoS 12248. *TDP2*<sup>-/-</sup> A549 cell extract was used as a control as a measure of background (presumably endonuclease) activity. Buffer control was used as a measure of background fluorescence. Data are the mean (+/- SD) of three independent experiments. **(C)** Measurement of inhibitory concentration (IC<sub>50</sub>) of UoS12248. UoS12248 was pre-incubated at the stated concentrations with 15 μg WCE from A549 cells. 5'-TDP activity was measured after 30 minutes. All data were corrected for background by subtraction using A549 *TDP2*<sup>-/-</sup> cells measurement. The IC<sub>50</sub> was calculated using the formulae  $y = -0.238 \ln(x) + 0.8775$ . Data are the mean +/- SE of three independent experiments.

Cancer Primary Tissue	Copy Number Variation		Gene Expression	
	TDP2 Gain (%)	Samples tested	TDP2 Over-expression (%)	Samples tested
Liver	1.73	692	18.23	373
Skin	2.31	650	12.26	473
Breast	0.52	1544	10.69	1104
Lung	0.93	1185	8.83	1019
Endometrium	0.50	598	6.81	602
Stomach	0.20	501	5.96	285
Upper aerodigestive tract	0.18	563	5.36	522
Pancreas	0.11	929	5.03	179
Haematopoietic and lymphoid	0.12	835	4.98	221
Prostate	-	-	4.62	498
Large intestine	0.13	771	4.43	610
Oesophagus	0.55	546	4.00	125
Soft tissue	0	286	3.80	263
Adrenal gland	0	268	3.80	79
Central nervous system	0.18	1093	3.59	697
Cervix	0.96	313	3.26	307
Urinary tract	3.10	419	3.19	408
Ovary	1.65	729	3.01	266
Kidney	0	1027	3.00	600
Thyroid	-	-	2.34	513

**Table 5.1 – TDP2 gene variation according to the Catalogue of Somatic Mutations in Cancer (COSMIC) database.** (Data collected on 01/08/2018)

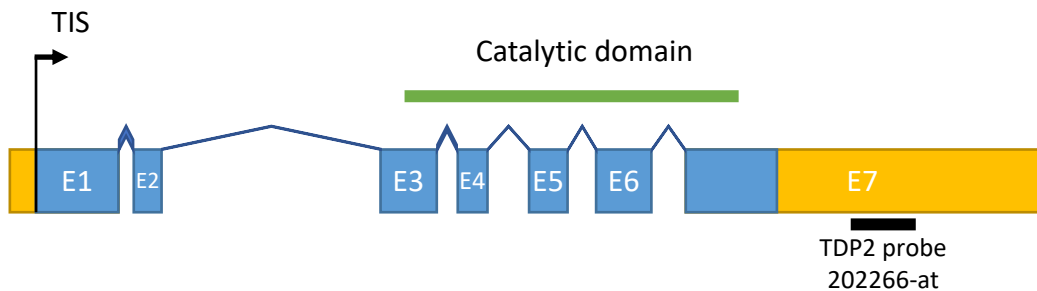


tumours. Using expression data of a probe localised at the 3' untranslated region of TDP2 (Figure 5.9.A), we found first, that changes in TDP2 expression did not lead to a decreased overall breast cancer patient survival (Figure 5.9.B). Nonetheless, we observed a significant increase in the probability of disease relapse in patients that have high TDP2 expression (Figure 5.9.C). Indicative that, using a simplistic explanation, TDP2 increased expression correlates with resistance and eventual disease relapse. To investigate if this apparent addiction to TDP2 over-expression was related to the role of topoisomerases in the transcription of oestrogen receptor target genes, we have stratified the survival data by tumour molecular phenotype. Breast cancer molecular phenotype is linked mainly to four markers, the proliferation marker Ki67, and the expression of the human epidermal growth factor receptor 2 (HER2/*ERBB2*), oestrogen (ESR) and progesterone (PR) receptors (Perou *et al.*, 2000; Sørbye *et al.*, 2001; Cancer Genome Atlas Network, 2012). From the presence of these four markers, tumours are categorised into HER2<sup>+</sup>, Luminal A (HER2<sup>-</sup>, ESR<sup>+</sup>, PR<sup>+</sup>, low Ki67), Luminal B (HER2<sup>-</sup>, ESR<sup>+</sup>, PR<sup>+</sup>, high Ki67) and triple negative breast cancers/basal like tumours (TNBC) (HER2<sup>-</sup>, ESR<sup>-</sup>, PR<sup>-</sup>) (Cancer Genome Atlas Network, 2012). While TDP2 over-expression was found to be present in all four molecular subtypes, it was only significantly elevated in relapsed Luminal A and B tumours, i.e. those that are oestrogen receptor positive (Figure 5.10.A-D). Interestingly, further stratification of the available patient data of luminal A tumours by cancer treatment also reveals that high TDP2 expression strongly correlates with worse disease prognosis only when primary tumours were systemically treated, i.e. patients have received chemotherapy including topoisomerase 'poisons' and adjuvant hormone therapy, and not in tumours that were surgically excised (Figure 5.11.A-B).

### 5.3. CONCLUSIONS AND DISCUSSION

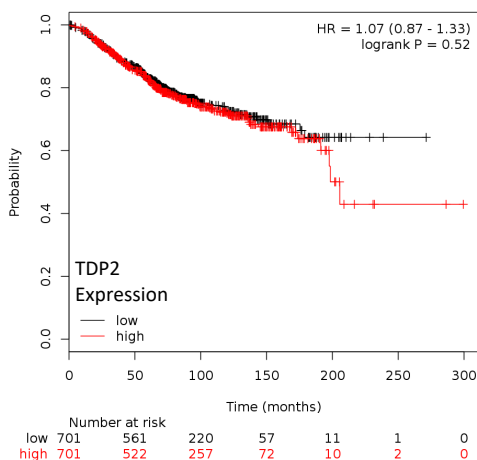
Here, I have demonstrated that TDP2 activity can be assayed using whole cell extract. Using an oligonucleotide substrate with a 5'-phosphotyrosyl bond, I found that 5'-TDP activity could be measured in whole cell extracts from multiple cell lines and that this activity was dependent on the presence and amount of TDP2 protein (Figure 5.3.A-B and 5.5). Furthermore, I was able to use this assay to confirm the TDP2 status in cell lines in which *TDP2* is present in heterozygous or homozygous state (Figure 5.7.A-B).

A.



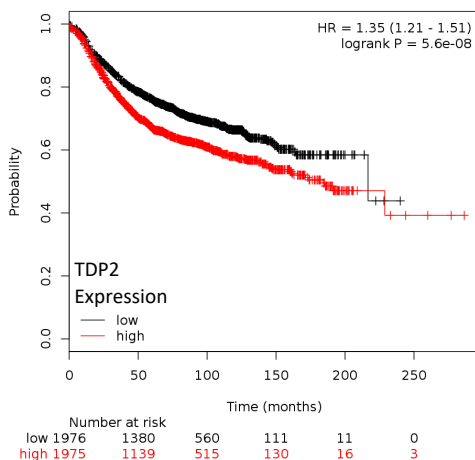
B.

Overall survival



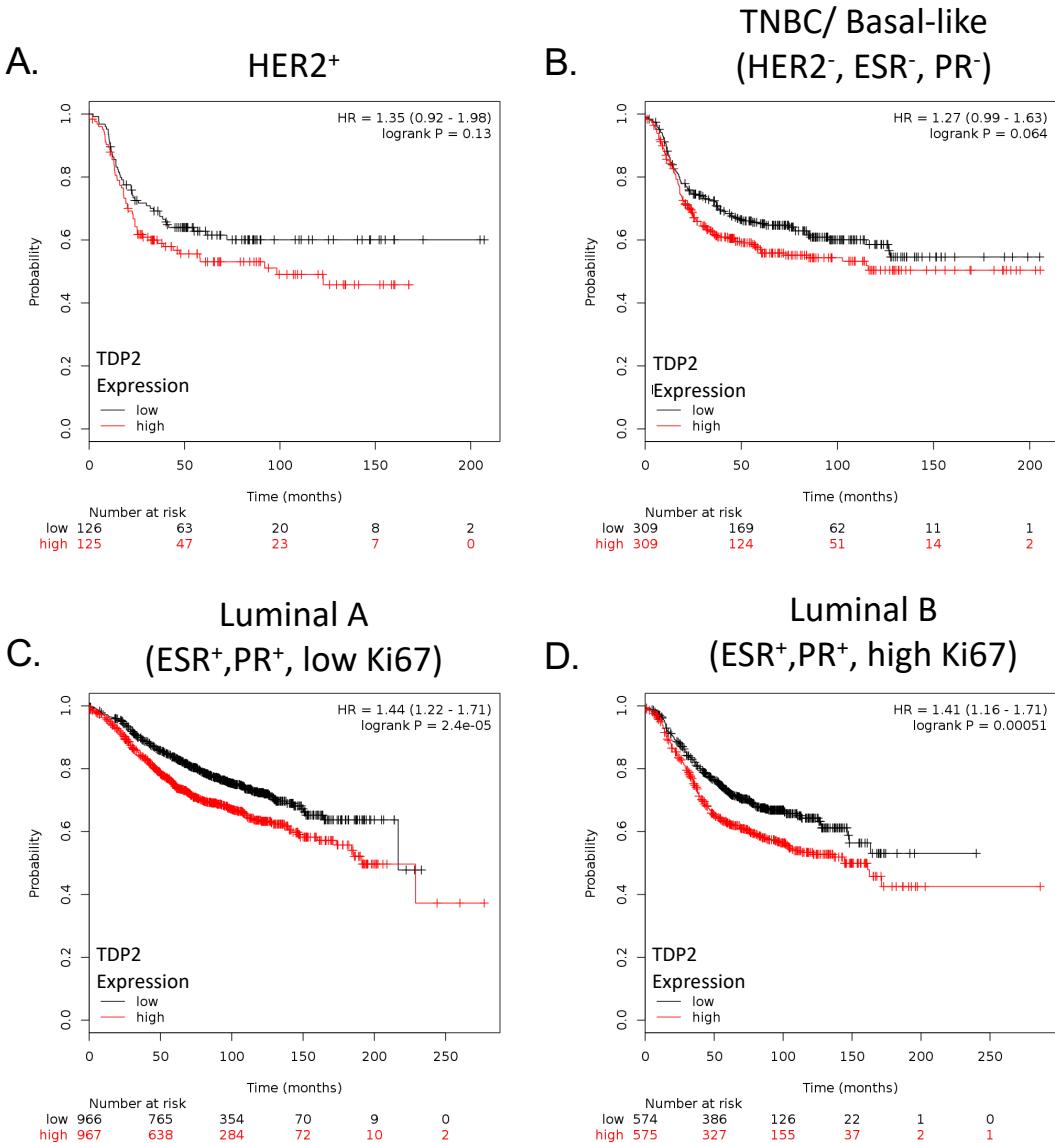
C.

Relapse free survival

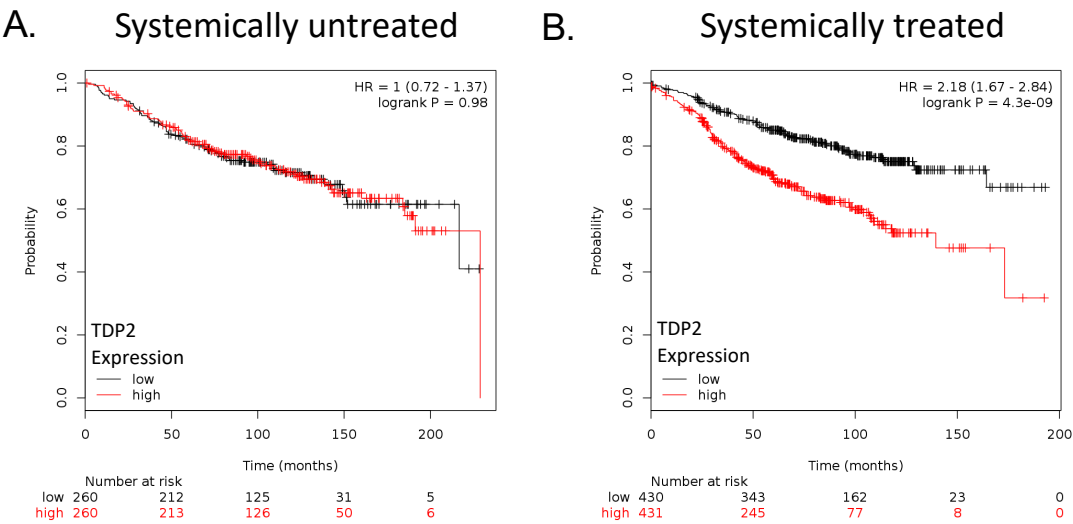


**Figure 5.9 - Breast cancer and TDP2 expression levels as possible prognosis marker.**

**(A)** Diagram of TDP2 showing location of Affymetrix probe 202266\_at for TDP2 3' untranslated region on Exon 7. **(B)** Overall survival of breast cancer patients does not correlated with TDP2 expression. **(C)** Breast cancer relapse free survival (probability of disease regression) strong correlated with low TDP2 expressing tumour ( $p=5.6 \times 10^{-8}$ ). These plots were created using kmplot tool by Györfy et al. 2010 accessible at <http://kmplot.com> and by filtering data for affymetrix probe 202266\_at (TTRAP/TDP2) (Györfy et al., 2010, Mihály and Györfy, 2013).



**Figure 5.10 - Breast cancer molecular phenotypes and TDP2 expression levels in relapse free survival.** Better Breast cancer prognosis correlates with lower expression of TDP2 in HER2<sup>+</sup> (A), triple negative breast cancers (TNBC)/basal-like (B), Luminal A (C) and Luminal B (D) molecular phenotypes. However the strongest correlation between lower level of TDP2 expression and disease relapse can be observed in luminal A ( $p=2.4 \times 10^{-5}$  and Luminal B ( $p=5.1 \times 10^{-4}$ ) tumours, which are positive for expression of oestrogen (ESR) and progesterone (PR) hormone receptors. These plots were created using Kmplot tool by Györfy et al. accessible at <http://kmplot.com> and by filtering data for affymetrix probe 202266\_at (TTRAP) and relevant molecular subtype (Györfy et al., 2010, Mihály and Györfy, 2013).



**Figure 5.11 - TDP2 addiction and cancer relapse in Luminal A Breast cancers –** Stratification of Luminal A molecular subtype patient data in systemically untreated, i.e. surgically removed tumours **(A)**, and systemically treated tumours **(B)**, i.e. patients have received chemotherapy and/or endocrine therapy. Systemically treated tumours show a strong correlation ( $p=4.3\times10^{-9}$ ) between high level of TDP2 expression and disease relapse can be observed (right-pane). These plots were created using Kmplot tool by Györffy et al. accessible at <http://kmplot.com> and by filtering data for affymetrix probe 202266\_at (TTRAP), above described molecular subtype and treatment conditions (Györffy et al., 2010, Mihály and Györffy, 2013).

Although these experiments confirmed that the assay was specific enough to match cell genotype, some discrepancies remain. For example, 5'-TDP activity in the *Tdp2*<sup>+/-</sup> cells was not half of that observed in wild type cells, contrary to expected (Figure 5.7.B). This may reflect chromosomal and gene expression alterations that can occur in cell lines following isolation of a single clone and that protein levels do not correlate with copy number variation (Veitia et al., 2013). I also confirmed that this substrate was sensitive to endonucleolytic cleavage using Benzonase endonuclease (Figure 5.2.B).

Residual removal of the 5'-phosphotyrosyl moiety was also observed in WCE of cells lacking TDP2, suggesting that in cells other enzymes such as endonucleases can process the oligonucleotide substrate used here (Figure 5.5 and 5.6.B). This notion is supported by (i) that *Tdp2*<sup>-/-</sup> cells also possess residual activity in the 5'-TDP assay, despite these cells having had Tdp2 deleted by homologous recombination mediated targeted gene disruption, rather than CRISPR/Cas9 mediated gene editing. (ii) that deazaflavin TDP2 inhibitor was unable to reduce activity in the 5'-TDP activity below that present in TDP2-disrupted cells, and does not further reduce this activity in TDP2-mutant patient cells (see next Chapter). Further optimisation of the assay might be required to remove this residual activity. For example, the use of a double-stranded oligonucleotide might reduce nuclease susceptibility, or the use of excess unlabelled oligonucleotide lacking a 5'-phosphotyrosine moiety might reduce this activity (Gómez-Herreros et al., 2014).

In addition, using a counter assay measuring 3'-phosphotyrosyl activity, I ensured that differences when measuring TDP2 activity were not due to sample loading (Figure 5.6.B). Furthermore, I found that the assay could be used in support of drug discovery efforts (figure 5.8.A-C). Using a compound previously shown to inhibit TDP2 protein purified in exogenous systems, I showed that the same compound does work in whole cell extracts, however a much higher concentration is required (Figure 5.8.C) (Raoof et al., 2013; Hornyak et al., 2016). These data highlight the utility of my assay for drug development approaches, although as mentioned earlier an appropriate tumour target for TDP2 inhibition still not known.

Better understanding of the molecular mechanisms underlying the DNA damage response and its signalling network, has so far allowed the development of many chemotherapeutic targets that exploit defects of repair pathways in cancer cells and

allows selective killing of tumours. However, despite the fact that much progress has been achieved, tumours heterogeneity and development of resistance to current treatments requires the continuous development of new targets to use as combination or monotherapy agents. One possible target that we propose is TDP2. Since its discovery it has been implicated in chemotherapy resistance to etoposide, chemotherapy induced chromosomal translocation and secondary malignancy (Do *et al.*, 2012; Gómez-Herreros *et al.*, 2017)

Furthermore, TOP2-induced breaks have been proposed to be required for activation of gene transcription in the oestrogen dependent MCF7 breast cancer cell line (Ju *et al.*, 2006). Recently, Gomez-Herreros and colleagues demonstrated that TDP2 is required for androgen driven transcription programmes that are dependent in TOP2 activity (Gómez-Herreros *et al.*, 2014). Given the hormonal dependence of some tumours, and high levels of TOP2 driven gene transcription and DNA breakage, we postulate that TDP2 expression might itself become a biomarker for such tumours. In agreement with this concept, data from patients with breast cancer indicates that TDP2 high expression, while it does not affect overall survival, it does correlate with worse disease prognosis (Figure 5.9.A-B). This correlation is present in all four molecular subtypes of Breast cancer (Figure 5.10.A-D). However, only in the luminal subtype, where there is overexpression of the oestrogen and progesterone receptors, this correlation becomes significant. Further stratification of the available patient data hormonal driven tumours by treatment received, further reveals that TDP2 high expression strongly correlates with worse disease prognosis only when primary tumours were systemically treated i.e. patients have received chemotherapy and endocrine therapy (Figure 5.11.B). We envisage that TDP2 addiction of such tumours could be avoided by adding a TDP2 inhibitor to complement current therapies resulting in inhibition of oestrogen driven transcription programmes and causing DSB-induced cell death selectively in such tumour cells. This approach is consistent with recent observations that in the MCF-7 cell line the absence of TDP2 leads to increased oestrogen transcriptional breaks and impaired transcription of oestrogen dependent genes (Gómez-Herreros *et al.*, 2017).

In summary, this Chapter highlights the potential of TDP2 as a target on cancer therapy. I characterise a new tyrosyl DNA phosphodiesterase activity assay that is

versatile and sensitive. This assay works in cell extracts and is able to detect TDP2 expression in multiple cells lines. While I have not measured TDP2 expression or activity in tumour samples, current data available from multiple studies indicate that its overexpression correlates with a worse disease prognosis, though whether this correlation reflects causality requires experimental approaches.

## 6. CHAPTER SIX – INVESTIGATING THE EFFECT OF TDP2 ABSENCE IN SCAR23 FIBROBLASTS

### 6.1. INTRODUCTION AND AIMS

If DNA repair is defective in a non-regenerative tissue, such as the brain, this can lead to neuronal dysfunction and/or loss and lead to disease (McKinnon, 2017). In contrast, if DNA repair proteins are faulty or absent during neural tissue development, that might lead to developmental problems and/or embryonic lethality (Rulten and Caldecott, 2013). The role of DNA repair in neurological diseases has been studied in multiple syndromes, with clinical presentations sometimes providing the basis of research to understand the DNA repair response itself. Such was the case for the autosomal recessive syndrome ataxia telangiectasia (AT), which is associated with cerebellar ataxia, telangiectasia, oculomotor apraxia, cancer, and immune-deficiency (Taylor *et al.*, 1975; Savitsky *et al.*, 1995). In addition to AT, many other neurological syndromes with mutations in DNA repair proteins have been identified (McKinnon, 2013; 2017). These syndromes not only arise from a general signalling defect of protein kinases such as ATM in AT and ATR in Seckel syndrome but also from mutations in proteins known to process specific DNA lesions such as Aprataxin (APTX), Polynucleotide Kinase 3'-Phosphatase (PNKP), TDP1 and TDP2 (O'Driscoll and Jeggo, 2003; Jazayeri *et al.*, 2006; Yoon and Caldecott, 2018; Zagnoli-Vieira *et al.*, 2018).

Aprataxin removes 5'-adenosine monophosphate (5'-AMP) moieties that arise following abortive activity of DNA ligases (Ahel *et al.*, 2006). Mutations in this gene lead to ataxia with oculomotor apraxia type 1 (AOA1), which is associated with cerebellar ataxia, oculomotor apraxia, and peripheral neuropathy (Date *et al.*, 2001; Moreira *et al.*, 2001). PNKP processes DNA termini with 3'-phosphate and/or 5'-hydroxyl moieties, to allow DNA ligation (Jilani *et al.*, 1999; Karimi-Busheri *et al.*, 1999). Mutations in PNKP result in two separate clinical presentations; microcephaly, seizures, and developmental delay (MCSZ) and ataxia-oculomotor apraxia-4 (AOA4) (Shen *et al.*, 2010; Bras *et al.*, 2015). TDP1 is mutated in multiple individuals of a family from Saudi Arabia, which presented with cerebellar ataxia, peripheral axonal motor and sensory neuropathy in a condition denoted spinocerebellar ataxia with axonal neuropathy (SCAN1) (Takashima *et al.*, 2002).



The involvement of TDP2 in neurological disease was first described in a family of Irish descendance in which three brothers presented with early onset seizures, intellectual disability, progressive ataxia and dysmorphic features, now denoted by the HUGO as Spinocerebellar Ataxia, autosomal recessive 23 (SCAR23) (Gómez-Herreros *et al.*, 2014; Zagnoli-Vieira *et al.*, 2018). All three patients harboured the homozygous mutation c.425+1G>A in a TDP2 intron splice site, resulting in a predicted premature stop codon upstream of the catalytic active site and in nonsense mediated decay of the TDP2 mRNA (Gómez-Herreros *et al.*, 2014). In addition, these brothers harboured a homozygous mutation in the gene ZNF193 (c.914A>G; p.His305Arg), which although predicted not to be pathogenic or related to the clinical presentation could not be unequivocally discarded as a contributing factor. Both TDP2 protein and 5'-TDP activity are absent in the Irish patients, and cells from these individuals present with delayed DSB repair and cellular hypersensitivity following treatment with the TOP2 poison, etoposide (Gómez-Herreros *et al.*, 2014). The molecular phenotype observed in patient-derived material was also observed in material from the TDP2<sup>Δ1-3</sup> mouse, although the mutant mouse did not share the clinical phenotypes presented by the patients (Gómez-Herreros *et al.*, 2013; 2014). Nonetheless, the Tdp2<sup>Δ1-3</sup> mouse did present with a mild loss of interneurons in the molecular layer of the cerebellum and a transcriptional defect of TOP2B dependent genes, consistent with the idea that TOP2B can induce transcriptional breaks that require TDP2-dependent repair (Gómez-Herreros *et al.*, 2014; Madabhushi *et al.*, 2015). The absence of an overt phenotype in the Tdp2<sup>Δ1-3</sup> mice highlights the difference in dependency on DNA repair proteins for normal neurological function, between mice and humans. Similarly, *Tdp1*<sup>-/-</sup> mice do not reproduce the neurological phenotypes observed in SCAN1 patients (Takashima *et al.*, 2002; Katyal *et al.*, 2007; Hirano *et al.*, 2007).

#### 6.1.1. Aim

In this chapter, I will focus on the characterisation of fibroblasts from a newly identified SCAR23 patient from the United States (US), again harbouring the homozygous TDP2 mutation, c.425+1G>A. Although the first SCAR23 patient was characterised by Gomez-Herreros and colleagues in 2014, up to now there has not been

a report of additional patients with TDP2 mutations. Furthermore, the probands characterised by Gomez-Herreros et al. 2014 also harboured a non-synonymous mutation on the gene ZSCAN9 (c.914A>G; reported as ZNF193). In addition, in collaboration with a group in Italy we have identified an additional two new patients with different mutations in TDP2 c.636+3\_636+6del. Thus, to better understand the role of TDP2 in disease in this chapter I will describe the characterisation of primary human fibroblasts from these novel SCAR23 individuals.

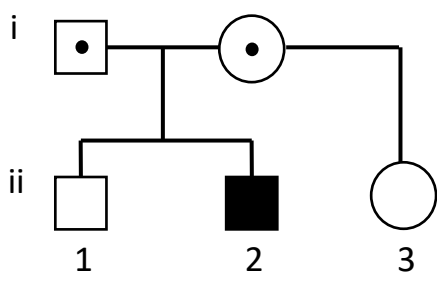
## 6.2. RESULTS

### 6.2.1. Spinocerebellar Ataxia, Autosomal Recessive 23 (SCAR23) patient of American origin

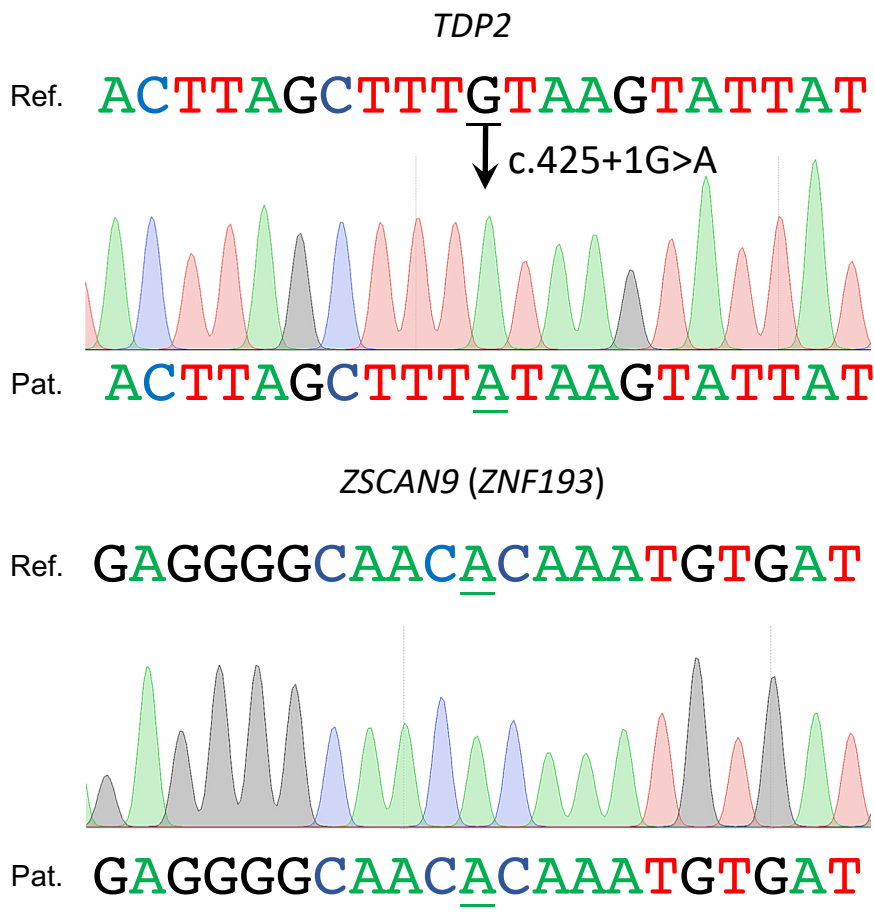
Through a collaboration with Dr. Dmitriy Niazov, a clinician in the United States of America, we were alerted to a boy with developmental delay, microcephaly, seizures and ataxia in which the TDP2 c.425+1G>A mutation was identified in homozygous state by whole exome sequencing (WES) (Niyazov, 2016). The proband was the second child of an unaffected American couple that carry of the TDP2 mutation in heterozygous state (Figure 6.1.A) (Niyazov, 2016; Zagnoli-Vieira *et al.*, 2018).

To allow better characterisation of the disease, the family kindly agreed to a skin biopsy which was used to generate primary fibroblasts, denoted 850BR. Using genomic DNA from these fibroblasts, I performed PCR followed by Sanger sequencing and was able to confirm the TDP2 mutations initially identified by whole exome sequencing (Figure 6.1.B, top). Importantly, using the same approach, we were able to show that the proband did not harbour the ZSCAN9 mutation that was additionally present in the Irish patients (Figure 6.1.B, bottom) reported by Gomez-Herreros et al. 2014, ruling out this mutation as a confounding factor. Despite this, however, haplotype analysis by our collaborator revealed that the US and Irish probands share the same haplotype spanning the TDP2 mutations, strongly suggesting a common founder mutation (Zagnoli-Vieira *et al.*, 2018).

A.



B.



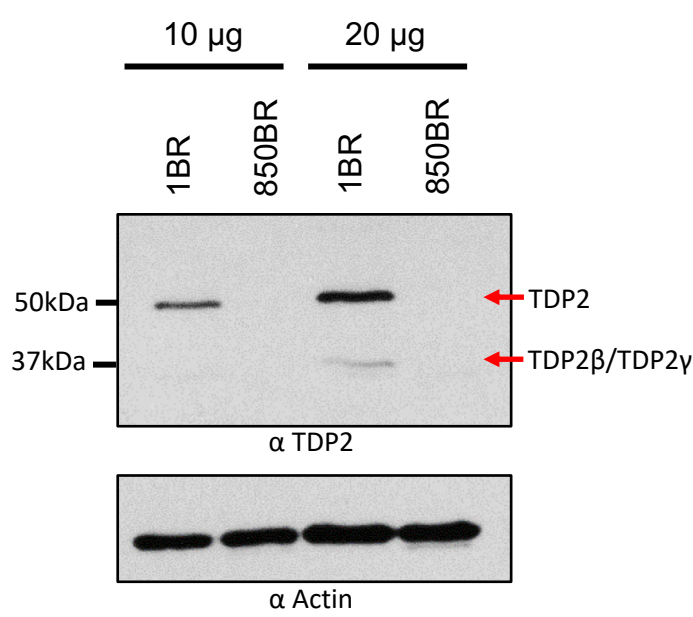
**Figure 6.1 –Mutation c.425+1G>A in TDP2 in a family from the USA. (A)** Pedigree analysis of SCAR23 patients of American origin. White symbols indicate wild type *TDP2* status. Black symbols; homozygosity for the *TDP2* mutation c.425+1G>A. Hybrid symbols indicate heterozygosity for the *TDP2* mutation. **(B)** The results of Sanger sequencing of PCR-amplified *TDP2* in genomic DNA isolated from proband II.2 demonstrating the *TDP2* mutation (top) and wild type status of *ZSCAN9* (bottom). Top sequence is the reference (ref.) genomic sequence and the bottom sequence is the result of Sanger sequencing of the patient (pat.) sample.

### 6.2.2. Tyrosyl phosphodiesterase activity in patient fibroblasts

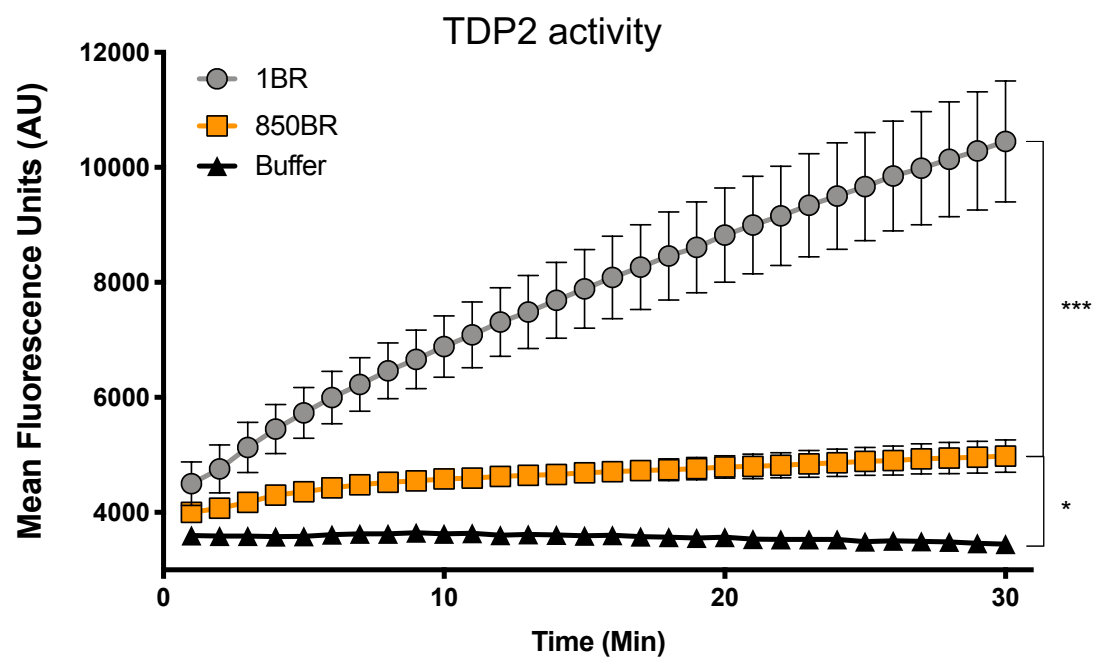
Having confirmed the mutant TDP2 status of the patient fibroblasts, the availability of the fibroblast cell line allowed further characterisation of protein levels and activity. Using an unrelated primary fibroblast as control (1BR), I discovered by western blotting that the patient cell fibroblasts (850BR) lacked detectable TDP2 protein (Figure 6.2.A). The apparent lack of TDP2 protein in the patient cells is consistent with previous results from analysis of lymphoblastoid cells from the Irish patients carrying the same mutation (Gómez-Herreros *et al.*, 2014).

To confirm that the absence of protein translated to lack of activity I employed the TDP assays I developed and described in Chapter 5. Using whole cell extract from patient-derived (850BR) and control primary fibroblasts (1BR), I was able to confirm that little 5'-TDP activity remained in patient WCE over 30 minutes of measurement compared to buffer control, whereas significant reduction was measured when comparing the wild type and patient WCE (Figure 6.2.B). This difference was not due to discrepancies in the quality or quantity of WCE employed, because the level of 3'-TDP activity was similar in wild type and patient WCE (Figure 6.3). The low level of substrate conversion by patient extracts, albeit significant, is most likely due to nucleases in the cell extract, and is consistent with my experiments in chapter 5, in which a low level of residual activity was detectable even in human cells in which TDP2 was deleted using CRISPR/Cas9 or homologous recombination mediated gene targeting (Figure 5.4.A and Figure 5.7.B). To confirm that this residual activity was not due to TDP2, I repeated the 5'-TDP activity assay in the presence of the TDP2 inhibitor described in Chapter 5 (UOS 12248) (Figure 6.4). I observed that pre-incubation of WCE from wild type human fibroblasts (1BR) with TDP2 inhibitor reduced 5'-TDP activity in wild-type cells (1BR) to a level of activity similar to that detected in WCE from 850BR patient fibroblasts. Moreover, addition of TDP2 inhibitor to patient WCE failed to further reduce the residual activity, confirming that the residual activity detected in these cells was not due to residual TDP2.

A.



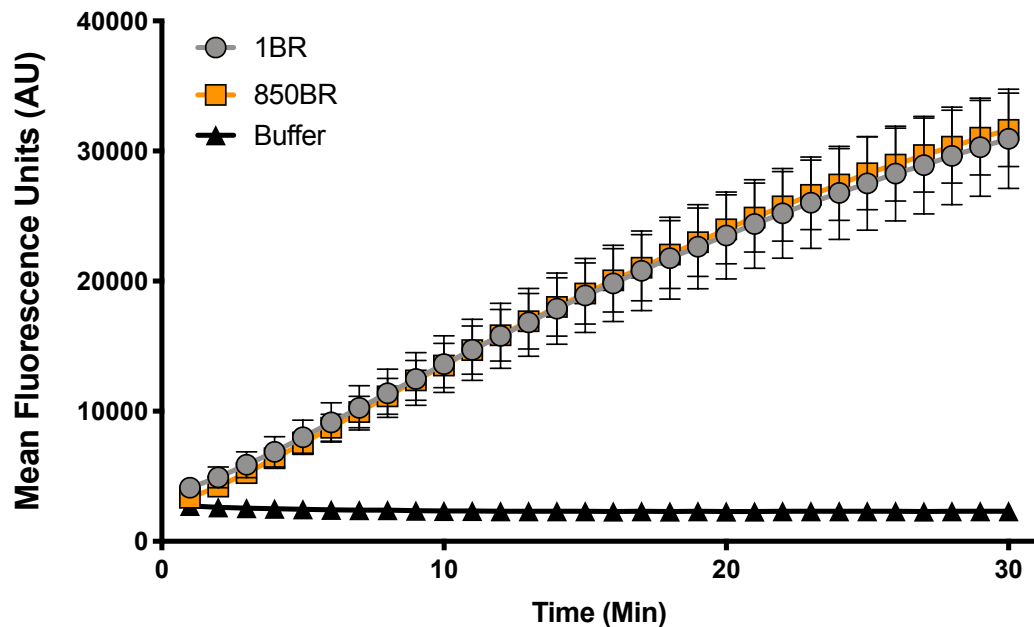
B.



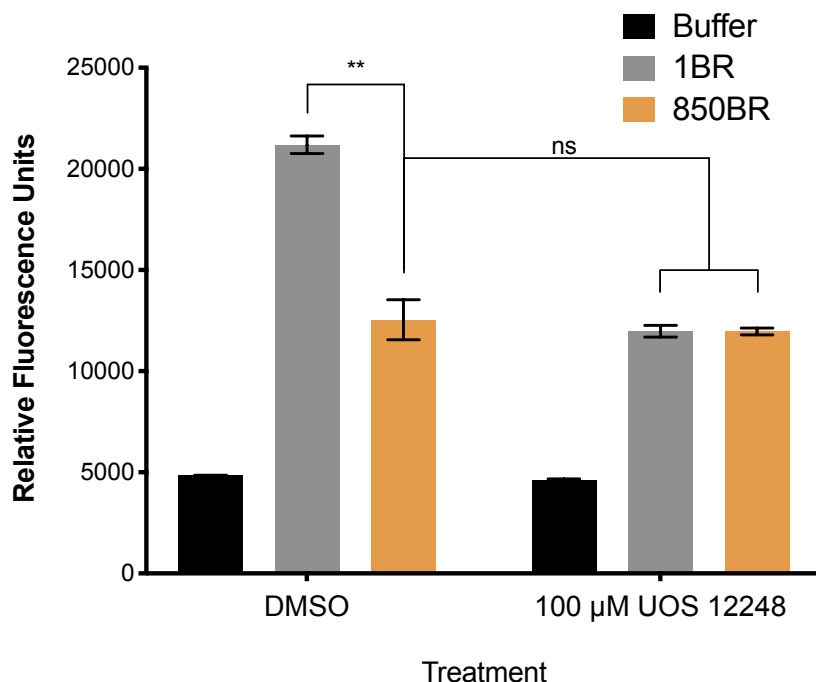
**Figure 6.2 – SCAR23 fibroblasts lacks detectable TDP2 protein and 5'-TDP activity – (A)** Western blot showing TDP2 protein levels in whole cell extract from 1BR normal and 850BR patient primary fibroblasts using either 10  $\mu$ g or 20  $\mu$ g total protein. Actin was employed as a loading control. Red arrows denote the full length and shorter isoforms (TDP2 $\beta$  and TDP2 $\gamma$ ) of TDP2. **(B)** Real-time measurements of 5'-TDP activity in WCE (15 $\mu$ g total protein) of 1BR normal or 850BR patient primary fibroblasts. Lysis buffer was used as a negative control. Data are the mean (+/- SEM) of three independent experiments. Statistical significance was determined by two-way ANOVA, \*p<0.05 and \*\*\*p<0.001

A.

TDP1 activity



**Figure 6.3 – SCAR23 fibroblasts displays normal 3'-TDP activity – (A)** Real time measurements of 3'-TDP activity in WCE (15µg) of 1BR normal or 850BR patient primary fibroblasts. Lysis buffer used as negative control. Data are the mean +/- SEM of three independent biological replicas.



**Figure 6.4 – Residual ‘5’-TDP’ activity in 850BR fibroblasts is not the result of residual TDP2 – (A)** 5’-TDP activity was measured in 1BR or 850BR WCE following preincubation for 30 min with either DMSO or 100 μM UOS 12248 (TDP2 inhibitor) and following a subsequent incubation for 30 min. Lysis buffer was employed as a negative control. Data are the mean (+/- SD) of two independent experiments. Statistical significance were calculated using a t-test. ns= $p > 0.05$ , \*\* $p < 0.01$ .

### 6.2.3. DNA repair and cellular sensitivity of SCAR23 fibroblasts following induction of TOP2-induced DSBs

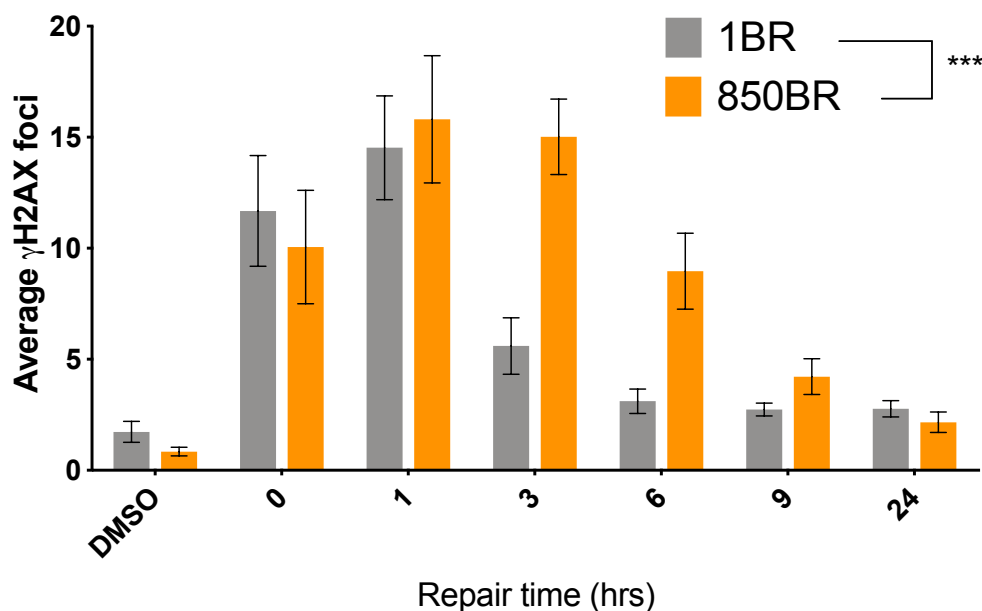
I next measured the efficiency of DSB repair in the patient derived fibroblasts. Using the topoisomerase ‘poison’, etoposide, I observed that confluency arrested primary patient cells (850BR) displayed delayed DSB repair kinetics when compared to unrelated control cells (1BR) (Figure 6.5.A). The repair defect was mostly significant at earlier repair time-points, 3 and 6 hours, and presented similar rates to control at later time-points. Interestingly, this repair defect seems to be specific to topoisomerase induced damage as the DSB repair rates after  $\gamma$ -irradiation were no different from control cells (Figure 6.5.B).

To ensure that the delayed DSB defect observed in patient cells was TDP2 dependent, patient cells were complemented with exogenous TDP2 expression. Patient and control fibroblasts were firstly immortalised by re-expression of the human telomerase reverse transcriptase (hTERT), generating 850BR hTERT and 1BR hTERT cell lines. The immortalised patient cell line was then transfected with expression construct harbouring a *TDP2* transgene tagged with green fluorescent protein (TDP2-GFP) or GFP control vector, and selected with G418 for a stable pool of transfectants (Figure 6.6.A). Once the stable expressing population was obtained, foci was measured after etoposide treatment once more (Figure 6.6.B). The delayed repair kinetics previously observed in primary patient fibroblasts by  $\gamma$ -H2AX immunostaining was also observed in immortalised patient fibroblasts. However, exogenous expression of TDP2-GFP was sufficient to correct this DSB repair defect, confirming TDP2 mutation as the cause of the DSB repair defect in the patient cells (Figure 6.6.B).

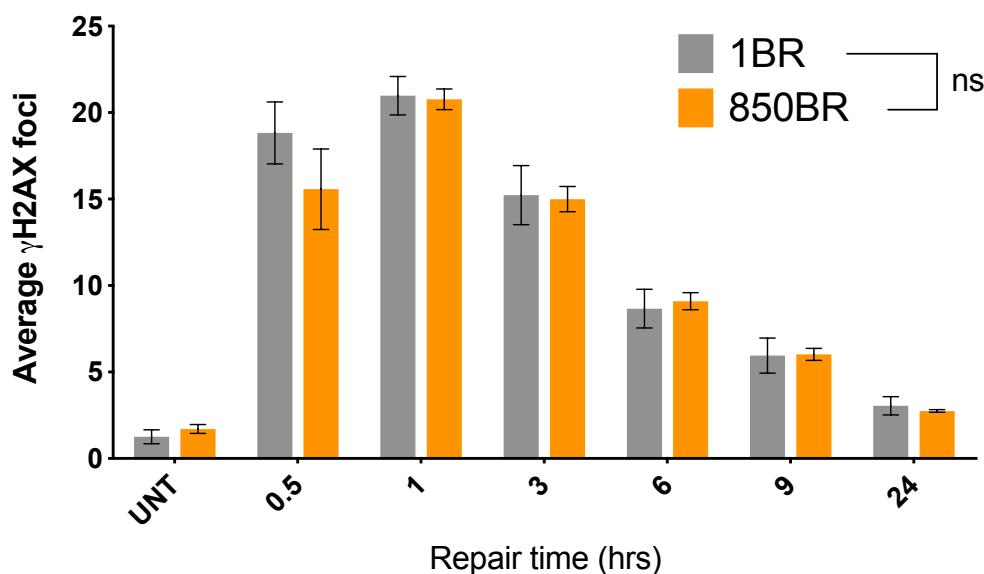
In addition to etoposide, I examined the impact of TDP2 on hypersensitivity to ionising radiation. Whereas 850BR patient primary fibroblasts were hypersensitive to etoposide (Figure 6.7.A), these cells displayed a level of sensitivity to x-rays that was similar to normal 1BR fibroblasts (Figure 6.7.B). Further suggesting that TDP2 function is specific to the topoisomerase 2 induced double-strand breaks.



A.

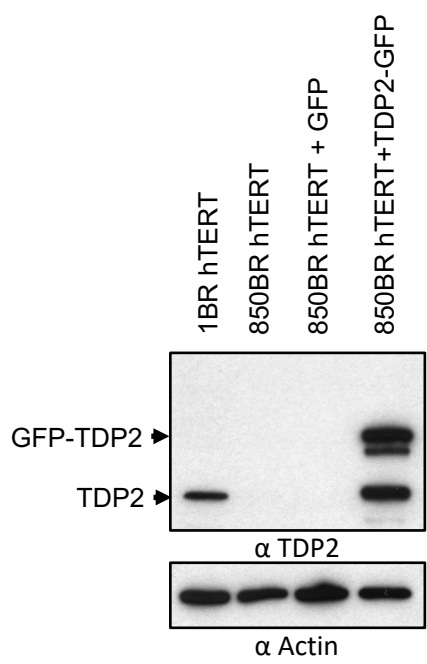


B.

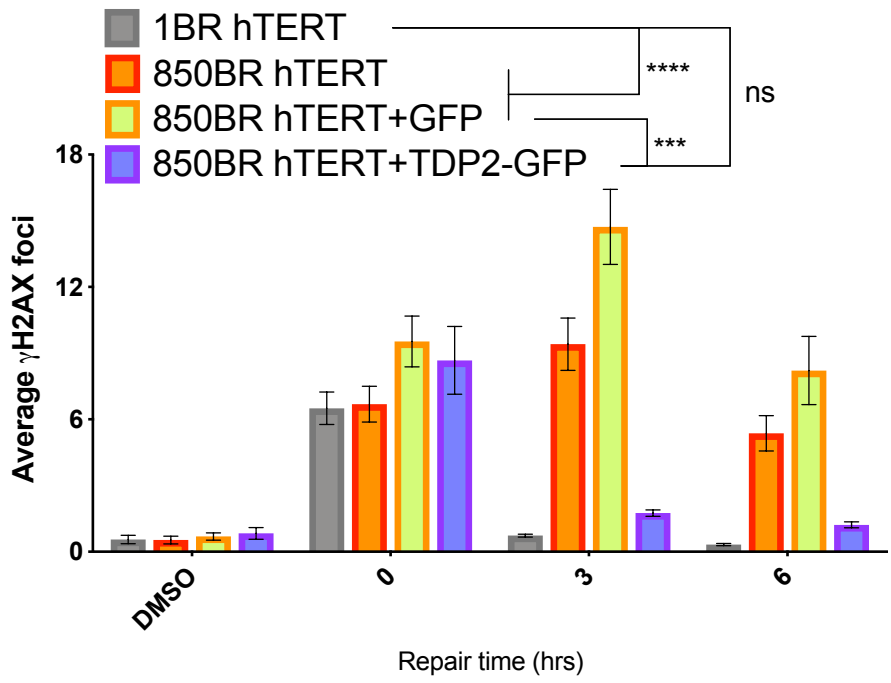


**Figure 6.5 – SCAR23 fibroblasts exhibit slower DSB repair following TOP2-induced DNA damage – (A)** DSBs were quantified indirectly as  $\gamma$ H2AX foci counts by ScanR high content imaging in >1000 1BR normal and 850BR patient primary fibroblasts, following 30min incubation with 25 $\mu$ M etoposide and the indicated repair time in drug-free medium. Cells were scored in G1 by gating using DAPI staining and CENP-F expression. **(B)** DSB repair rates measured as  $\gamma$ H2AX foci in 1BR and 850BR primary fibroblasts following mock treatment (untreated; UNT) or 2Gy  $\gamma$ -irradiation at the indicated times, quantified as above. Data are the mean (+/-SEM) of three independent experiments. Statistical significance was determined by two-way ANOVA (ns=not significant, \*\*\*p<0.001).

A.

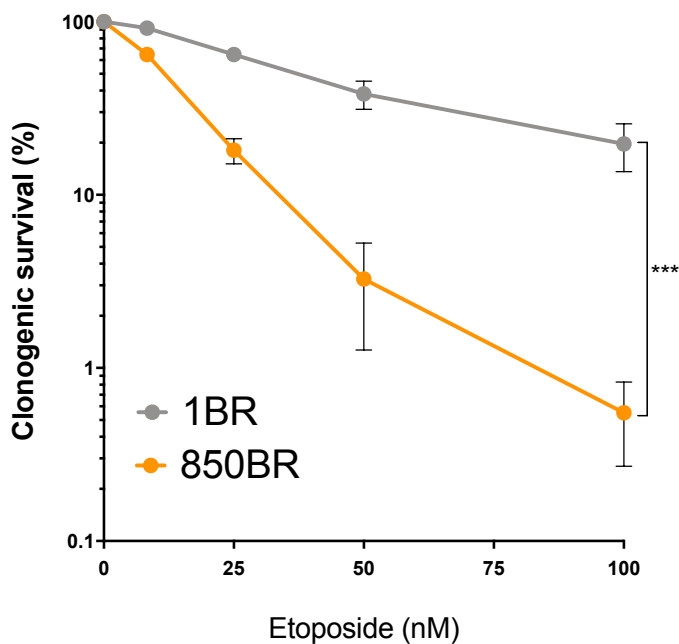


B.

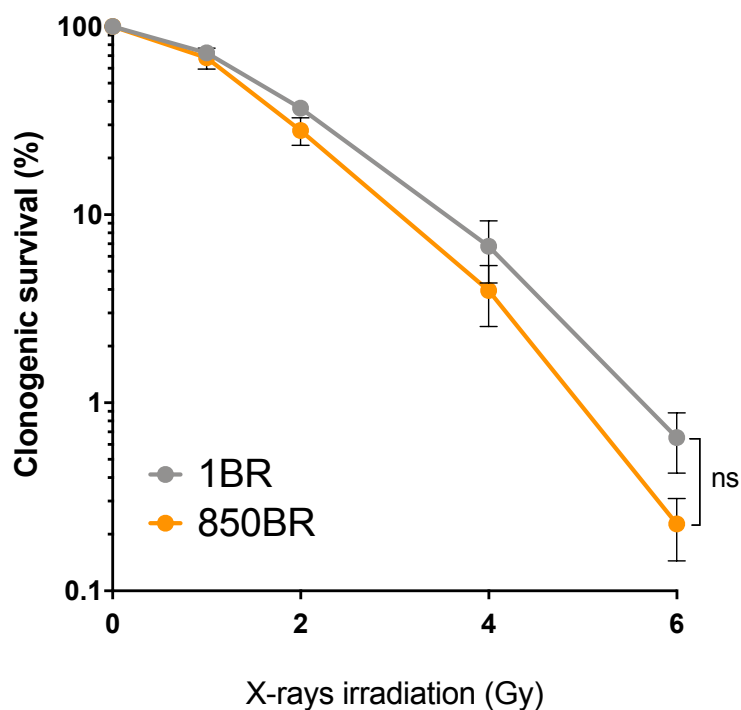


**Figure 6.6 – TDP2 exogenous expression complements the DSB repair defect in SCAR23 fibroblasts – (A)** Anti-TDP2 western blot of immortalized normal 1BR hTERT, patient 850BR hTERT, and patient 850BR hTERT fibroblasts stably expressing GFP or TDP2-GFP transgenes employed in the experiments in (B). Anti-Actin was used as a loading control. **(B)** DSBs were quantified as  $\gamma$ H2AX foci by ScanR high content imaging in >1000 cells as in (A), following 30min incubation with 25 $\mu$ M etoposide and the indicated repair time in drug-free medium. Cells were scored in G1 by gating according to DAPI. Data are the mean (+/-SEM) of three independent experiments. Statistical significance was determined by two-way ANOVA. ns=p>0.05, \*\*\*p<0.001 and \*\*\*\*p<0.0001.

A.



B.



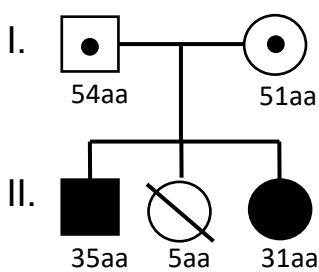
**Figure 6.7 – SCAR23 fibroblasts are hypersensitive to TOP2-induced DNA damage –**  
**(A)** Clonogenic survival of 1BR normal and 850BR patient primary fibroblasts following continuous incubation with the indicated doses of etoposide. **(B)** As in (A), following X-rays irradiation with the stated doses. Data are the mean (+/-SEM) of three independent experiments. Statistical significance was determined by two-way ANOVA. ns= $p>0.05$  and \*\*\* $p<0.001$ .

#### 6.2.4. SCAR23 patients of Italian origin

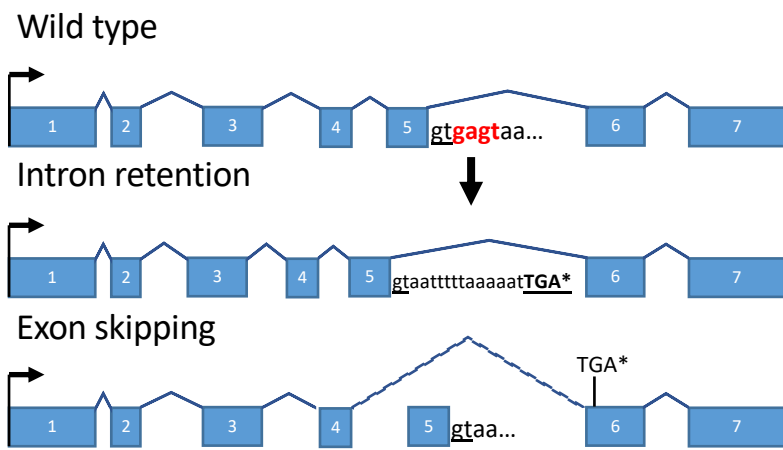
The identification and characterisation of *TDP2* mutation (c.425+1G>A) in SCAR23 in a newly identified patient from the USA described above is important, because it confirms the association of defective TDP2 in this disease. Nevertheless, it is even more informative from a molecular pathogenicity perspective if patients with different TDP2 mutations can be identified. With this in mind, we were enthused when presented with two Italian siblings (brother and sister) with clinical features reminiscent of SCAR23 including developmental delay, seizures, and dysmorphic facial features (Errichiello and Zuffardi, 2017). Exome sequencing revealed that both probands possessed the homozygous *TDP2* mutation c.636+6\_636+3del; a four base pair deletion close to the donor splice site at the 5'-end of intron 5 (highlighted on Figure 6.8.A-B). This intronic mutation was predicted to lead either to intron retention or to exon skipping, both of which are predicted outcomes leading to a premature stop codon and nonsense mediated decay of the mRNA (Figure 6.8.B). To identify the outcome of this mutation, Dr. Edoardo Errichiello, performed PCR and sanger sequencing on the cDNA from these proband and found that skipping of exon 5 was indeed occurring (Figure 6.8.C).

Once again, I characterised primary fibroblasts isolated from skin biopsies from these patients. The fibroblasts from neither the female (MG) or the male (MS) proband did not possess detectable TDP2 protein, reminiscent of the cells from the proband of American origin (850BR) (Figure 6.9.A). In addition, these fibroblasts also displayed delayed DSB repair kinetics following TOP2-induced DSBs, comparable to the defect I observed in 850BR fibroblasts (Figure 6.9.B). Finally, I examined the primary fibroblasts from the Italian patient for 5'-TDP activity using my cell free assay. On these experiments, because of cell availability constraints WCE were diluted 10-fold from normal 15 µg per reaction to 1.5 µg and incubation time was increased from 30 minutes to 180 minutes. These changes did not seem to affect the assay, as I observed that similarly to the 850BR cells both MG and MS primary fibroblasts also lacked TDP2 activity (Figure 6.10.A). Cell availability constraints also did not allow 3'TDP activity to be carried, however the WCE used in 5'TDP assay was found to contain similar levels of protein by western blotting (Figure 6.9.A).

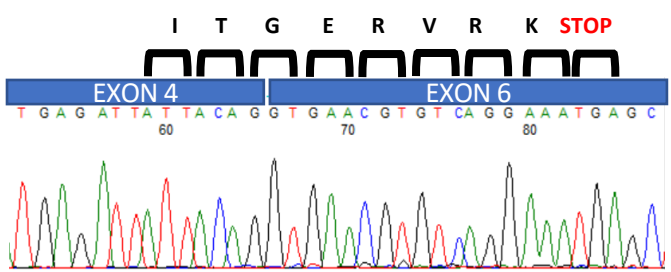
A.



B.

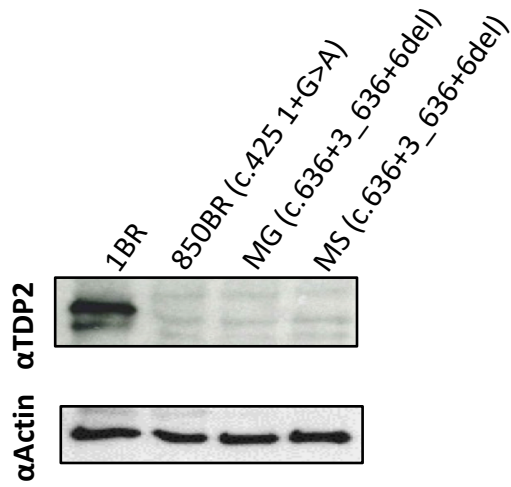


C.

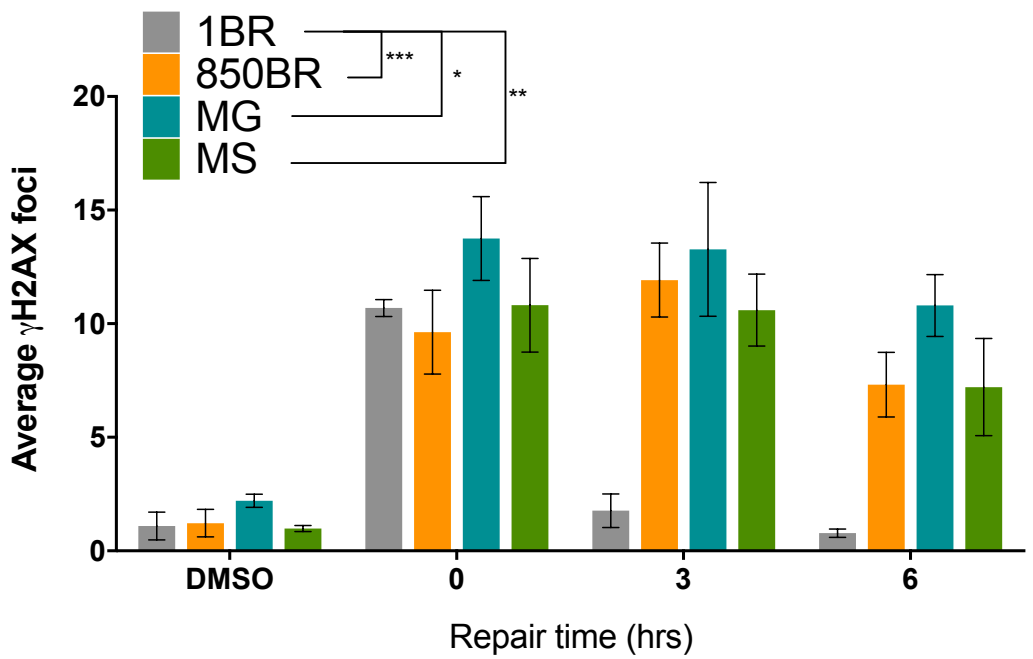


**Figure 6.8 - Intronic homozygous TDP2 c.636+3\_636+6del mutation in Italian SCAR23 patients. (A)** Pedigree analysis of SCAR23 patients of Italian origin. White symbols indicate wild type *TDP2* status. Black symbols indicate homozygosity for the *TDP2* mutation c.636+3\_636+6del. Hybrid symbols indicate carrier status. Crossed symbol indicate deceased individuals. **(B)** (Top) Diagram of *TDP2* gene structure depicting intronic mutation c.636+3\_636+6del (highlighted in red) and possible outcomes of intron retention (middle) and exon skipping (bottom) leading to premature stop codon. **(C)** Sanger sequencing of PCR amplified *TDP2* cDNA from proband II.1 and II.2 showing that *TDP2* Exon4 is spliced to exon 6. Corresponding amino acids are present on top with the premature stop codon indicated. (Data kindly provided by Dr. Edoardo Errichiello)

A.

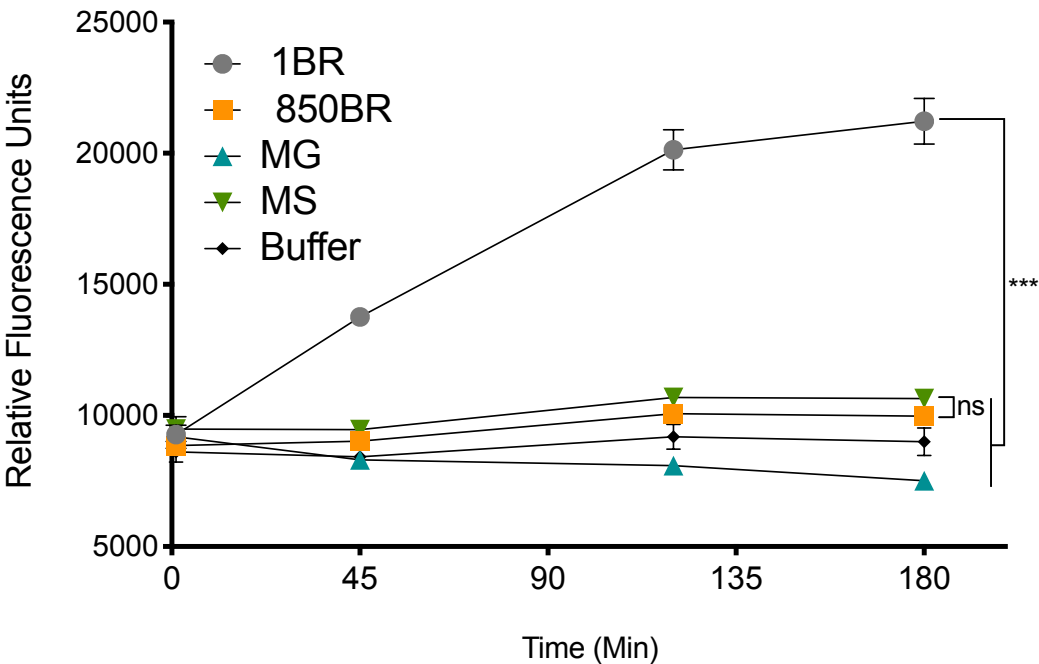


B.



**Figure 6.9 – The TDP2 mutation c.636+3\_636+6del leads to loss of TDP2 protein and reduced DSB repair - (A)** Western blot of TDP2 protein in 1BR normal, 850BR patient, and MG and MS TDP2 patient primary fibroblasts. The corresponding *TDP2* patient mutations are shown in brackets. An anti-actin blot is shown as loading control. **(B)** DSBs were measured by quantification of  $\gamma$ H2AX foci by ScanR high content imaging in 1BR normal, 850BR patient, and both MG and MS Italian patient primary fibroblasts, following 30min incubation with 25 $\mu$ M etoposide and the indicated repair time in drug-free medium. Cells were scored in G1 by gating according to DAPI staining. Data are the mean (+/-SEM) of three independent experiments. Statistical significance was determined by ANOVA. \* $p < 0.05$ , \*\* $p < 0.01$  and \*\*\* $p < 0.001$ .

A.



**Figure 6.10 – TDP2 mutation c.636+3\_636+6del leads to protein loss and defective repair - (A)** Real time measurements of 5'-TDP activity in WCE (1.5 µg) from 1BR normal and 850BR, MG and MS patient primary fibroblasts over the indicated reaction time. Lysis buffer was used as a negative control. WCE used in this experiment was also used for western-blotting analysis of TDP2 and actin in figure 6.9.A. Data are the mean (+/-SD) of three technical replicas. Statistical significance was determined by two-way ANOVA, ns= $p>0.05$  and \*\*\* $p<0.001$

### 6.3. CONCLUSIONS AND DISCUSSION

In this chapter, I provide the first independent report that the TDP2 mutation c.425+1G>A is causative of SCAR23 and initial evidence that the TDP2 mutation c.636+4\_636+6del could also contribute to the disease aetiology. I also provide the first analysis of primary fibroblasts from TDP2-mutated patients.

Using patient-derived fibroblasts, I confirmed that both of the TDP2 mutations c.425+1G>A and c.636+4\_636+6del lead to the loss of TDP2 protein (Figure 6.2.A and 6.9.A). In addition, I show that both mutations lead to greatly reduced 5'-TDP activity in whole cell extracts (Figure 6.2.B and 6.10). The residual conversion of substrate into product by the patient extracts is likely to reflect nuclease activity in the extracts rather than residual TDP2 activity for several reasons. First, the mutations are predicted to result in nonsense mediated decay of the mutated mRNA, and any mRNA that is translated is predicted to result in premature termination of translation upstream of the catalytic active site. Consistent with this, I failed to detect any residual TDP2 protein in the patient fibroblasts by western blotting (Figure 6.2.A and 6.9.A). Second, TDP2 inhibitor failed to reduce that residual activity in patient extracts, despite reducing the total activity observed in wild type extracts to the level observed in patient extracts (Figure 6.4). Finally, similar background 5'TDP activity was also detected in cells that TDP2 was deleted using gene editing (see Chapter 5).

In agreement with previous publications, and with my results using TDP2 in gene edited cell lines, I observed reduced DSB repair kinetics in all three TDP2 patient fibroblasts examined, following treatment with etoposide (Figure 6.5.A and 6.9.B). Furthermore, complementation with TDP2 transgene was sufficient to rescue normal rates of DSB repair in fibroblasts with the mutation c.425+1G>A (Figure 6.6.A-B). Additionally, we show that these patient fibroblasts are not sensitive to DSB introduced by ionising radiation, further implicating TOP2-induced DSBs as the lesions most likely responsible for SCAR23 (Figure 6.5.B). Indeed, consistent with this notion, I show that the TDP2 mutations in the patient fibroblasts result in a significant decrease of cell survival in response to TOP2-induced DSBs, but not to DSBs induced by ionising radiation (Figure 6.7.A-B).



Taken together the data on this chapter, establishes the TDP2 c.425+1G>A mutation as cause of SCAR23. Interestingly, the identified proband with the TDP2 c.425+1G>A mutation, presents more severe symptoms, while presenting similar molecular phenotype, than that of the three brothers previously reported in Ireland with same mutation (Table 6.1) (Gómez-Herreros *et al.*, 2014). Many of the additional symptoms, such as hyponatremia, hypotonia, neutropenia and fatigability with prolonged periods of sleep, agreed with mitochondrial dysfunction (Zagnoli-Vieira *et al.*, 2018). Consistently, analysis of the patient muscle tissue found deficiency of electron transport chain (ETC) complex I+III and II+III activity, indicating lack of co-enzyme Q10 (Niyazov, 2016; Zagnoli-Vieira *et al.*, 2018). Nevertheless, analysis of the patient-derived fibroblasts or TDP2<sup>-/-</sup> cells could not detect abnormal levels of ETC complexes I, II, III, IV and V protein levels or activity (Zagnoli-Vieira *et al.*, 2018). Secondary mitochondrial dysfunction is present sometimes in a setting of a non-mitochondrial disorder and has been shown to be the case in patients with mutations of DNA repair proteins such as aprataxin (Garcia-Diaz *et al.*, 2015; Niyazov *et al.*, 2016). AOA1 patients have also reported to have deficiency of co-enzyme Q10 (Quinzii *et al.*, 2005; Le Ber *et al.*, 2007). Nonetheless, this deficiency seems heterogenous within patients with the same mutations and also to variate according to which patient tissue is used for assessment (Castellotti *et al.*, 2011). Aprataxin was shown to localise to the mitochondria, however the deficiency in co-enzyme Q10 was linked to a transcription defect of the transcription factor nuclear respiratory factor 1 (NRF1) (Sykora *et al.*, 2011; Garcia-Diaz *et al.*, 2015).

Unlike Aprataxin, we could not find clear evidence of TDP2 localisation to the mitochondria, including its isoforms as discussed on Chapter 4. A small amount of TDP2 protein was identified in the mitochondria fraction of HeLa cells, however this small fraction was not present in mitoplasts, which is the fraction containing proteins that are localised to the mitochondrial matrix and inner membrane. In addition, we could not find the decreased levels of ETC complexes protein or activity using the 850BR fibroblasts from the TDP2 patient, indicating that a transcription defect, such as observed in AOA1 cells, is unlikely to explain the clinical presentation of this patient. In contrast, TDP2 was reported by Huang *et al.* 2018 to be located to the mitochondria and mediate transcription of mitochondrial genes (Huang *et al.*, 2018). Taking this in conjunction with our results we have to raise the possibility that, if present, TDP2

Proband (age at report)	Family origin	TDP2 Mutation	Seizures/ Epilepsy (age at onset)	Intellectual disability	Ataxia	Molecular features	Patient material analysed	Publication
32-years old	Irish	c.425+1G>A	Yes (<2 months old)	Yes	Yes	No 5'TDP activity; no TDP2 protein; etoposide sensitive; delayed DSB repair after etoposide	Blood extract; Lymphoblast derived cell line	Gomez-Herreros et. al 2014
26-years old	Irish	c.425+1G>A	Yes (<12 years old)	Yes	Yes	No 5'TDP activity; no TDP2 protein; etoposide sensitive	Blood extract; Lymphoblast derived cell line	
23-years old	Irish	c.425+1G>A	Yes (<6 months old)	Yes	No	No 5'TDP activity; no TDP2 protein; etoposide sensitive	Blood extract; Lymphoblast derived cell line	
ND	Egyptian	c.413_414delinsAA	Yes (ND)	Yes	Yes	No 5'TDP activity	Blood extract	
06-years old	American	c.425+1G>A	Yes (<5 months old)	Yes	Yes	No 5'TDP activity; no TDP2 protein; etoposide sensitive; delayed DSB repair after etoposide; DSB repair proficient	Primary fibroblasts	Zagnoli-Vieira et. al 2018
35-years old	Italian	c.636+3_636+6del	Yes (ND)	Yes	ND	No 5'TDP activity; no TDP2 protein; delayed DSB repair after etoposide	Primary fibroblasts	unpublished
31-years old	Italian	c.636+3_636+6del	Yes (ND)	Yes	ND	No 5'TDP activity; no TDP2 protein; delayed DSB repair after etoposide	Primary fibroblasts	

**Table 6.1 – Summary of all reported SCAR23 patients including major clinical and molecular features.** Data presented on this thesis is highlighted in blue. ND= Not disclosed

requirement in the mitochondria could be tissue dependent, even though the current data does not support this. So far, the only supporting information for this hypothesis is that, as with AOA1 patients that received ubiquinol therapy, the proband with the TDP2 c.425+1G>A mutation did report an improvement in energy levels, development progress and reduced number of seizures following liposomal ubiquinol, carnitine and leucovorin (Niyazov, 2016; Zagnoli-Vieira *et al.*, 2018).

Here, we also provide the first evidence that patients with the TDP2 c.636+6\_636+3del mutation share molecular and clinical phenotype with SCAR23 patients. While, this TDP2 mutation identified on the probands of Italian origin is the first to be independently linked to SCAR23, it was also reported in other patient studies. The TDP2 c.636+6\_636+3del mutation was found at the frequency of  $1.60 \times 10^{-3}$  of genotypes analysed by the National Heart, Lung and Blood Institute (NHLBI) Exome sequencing project (ESP, 2018). This frequency includes eight homozygous individuals and is approximately ten times more frequent than the mutation c.425+1G>A present at  $1.599 \times 10^{-4}$  of all genotypes sequenced (ESP, 2018). It is not clear, if these homozygous patients identified through whole exome sequencing targeting disorders from heart, lung and blood also presents the symptoms from SCAR23. Therefore, further investigation will be necessary to establish this link of SCAR23 with any secondary disorders.

In summary, the results presented on this chapter provides the first independent report of a patient with SCAR23 that does not have mutation on the ZSCAN9 gene, while adding evidence that the TDP2 c.636+6\_636+3del mutation is pathogenic. The data also provides evidence that TDP2 is not located or involved in mitochondria transcription. Finally, it highlights the importance of TDP2 to the repair of TOP2 lesions, raising the question on how significant and frequent such lesions could be on the human brain. Previously, DSBs were shown to be induced following neuronal activity (Suberbielle *et al.*, 2013). In addition, TOP2B mediated DSB breaks were shown to mediate transcription of neuronal genes (Madabhushi *et al.*, 2015; Madabhushi, 2018). How and when this breakage occurs still unclear. Nonetheless, it seems that TDP2 is the preferred way for cells to deal with such structures, therefore, we need a better understanding of the role of TDP2 in endogenous damage.

## 7. CHAPTER SEVEN – DISCUSSION AND CONCLUDING REMARKS

### 7.1. Role of TDP2 isoforms

Here I have identified two new isoforms of TDP2, namely TDP2 $\beta$  and TDP2 $\gamma$ . TDP2 $\beta$  was identified by capture from a cDNA library and TDP2 $\gamma$  through mass spectrometry (Figure 4.8.A-D). Although I was able to confirm that TDP2 $\gamma$  was a pan-cellular protein that could mediate repair and survival following etoposide treatment (Figure 4.11.B-C), the role of this protein in cells is not clear. Whereas for TDP2 $\beta$ , no involvement with nuclear DNA repair has been found even though it was found to be active in cytoplasm and mediate viral infection (Figure 4.12.B). TDP2 $\beta$  was found to be located in the endoplasmic reticulum contrary to a recent report which located this TDP2 isoform in the mitochondria (Figure 4.10). Like TDP2 $\gamma$ , the function of TDP2 $\beta$  in cells also remains elusive.

Nonetheless, TDP2 was initially identified as part of the TNFR signalling cascade which is active in the cytoplasm (Pype *et al.*, 2000). It may well be that the protein first identified was one of these isoforms. Therefore, it will be important to characterise the protein interactions of these isoforms to try to dissect out TDP2 functions.

### 7.2. TDP2 in cancer and drug discovery

TDP2 activity in whole cell extracts from multiple cell lines was also investigated here. It was found that an oligonucleotide with 5'-tyrosyl linked Cy5 moiety coupled with a black hole quencher at the 3'-termini was a sensitive assay that could be used for high throughput screening purposes (Figure 5.1.C). This assay performed consistently well and was able to support the absence of TDP2 in gene edited and patient cells lines (Figure 5.7.A-C and Figure 6.2.B). In addition, it was also shown to be a good method to test novel inhibitors of TDP2. Using a previously characterised TDP2 inhibitor, I was able to show its inhibitory coefficient in cells and show that any background fluorescence in cells lacking TDP2 was due to the presence of cellular endonucleases (Figure 5.8.A-C and Figure 6.4).

The assay characterised here, could have multiple applications pertinent to cell biology, drug discovery and clinical exploitation. This assay was devised initially to support current drug discovery efforts at the University of Sussex and was used in conjunction with purified TDP2 to screen multiple compound libraries. So far, initial hit validation is being carried out. The next proposed step is to use the TDP assay in whole cell extracts as described here to verify inhibitor activity in a complex subcellular environment, prior to compound use in cell cultures, making it an important tool in the drug discovery pipeline. Secondly, this assay could be used to aid the discovery of any additional enzyme that processes the 5'-phosphotyrosyl DNA bonds. *TDP2*<sup>-/-</sup> cell extracts possessed some residual activity in this assay that could be further investigated. For example, the TDP assay could be modified to allow capture of such enzyme. One of such enzymes could be TDP1 as it was previously reported that on *Tdp2*<sup>-/-</sup> DT40 cells loss of Tdp1 is additive to sensitivity to etoposide (Murai *et al.*, 2012). Even though additional reports do not support this observation (Zeng *et al.*, 2012). It would be interesting to investigate if TDP1 is indeed acting on such substrates in the absence of TDP2 in human cells. In addition, we could potentially test the activity not only of TDP1 but also of specific nucleases such as Artemis, MRE11, CtIP, XPG, FEN1 on this substrate in support of the discovery of alternative repair pathways that deal with such structures in the absence of TDP2.

Finally, the sensitivity of use of the TDP assay makes it suitable for analysis of TDP2 activity in patient biopsies. This assay could be used to identify patients lacking 5' phosphodiesterase activity, i.e. SCAR23 patients, even prior to sequencing of patient mutation. In addition, this approach could be used to confirm the observations that some tumours overexpress TDP2. In addition, as the activity detected on this assay correlates with the amount of TDP2 it could be used to screen tumours in which topoisomerase 'poisons' would be more effective.

### **7.3. TDP2 absence and DNA damage repair**

The effects of the absence of TDP2 in regards to DNA damage repair was investigated in this thesis in non-transformed, cancer and patient-derived cell lines. Upon treatment with the topoisomerase 2 'poison' etoposide, all these cells presented

hypersensitivity in clonogenic survival assays and delays in DSB repair (Figure 4.3.A-B, Figure 4.4.A-B, Figure 6.5.A and Figure 6.7.A). This observation was consistent with previous reports in the literature that TDP2 is a key enzyme that deals with covalently trapped topoisomerase 2 on to DNA (Zeng *et al.*, 2011; Gómez-Herreros *et al.*, 2013; Zagnoli-Vieira and Caldecott, 2017). The delay in DSB repair following etoposide was observed in G1 and S/G2 cells. In G1 cells, it was observed that DSBs persist longer in *TDP2*<sup>-/-</sup> cells, as measured by  $\gamma$ H2AX foci, but also that they do decline at later time points (Figure 6.5.A), indicative of a TDP2 independent repair process. TDP2 was shown to be epistatic to *Ku70*<sup>-/-</sup> in clonogenic survival assays using etoposide in mouse embryonic fibroblasts (Gómez-Herreros *et al.*, 2013), however it is still not known how repair takes place in the absence of TDP2 in G1 cells when HRR is not available.

The existence of an alternative pathway to TDP2 in G1 cells is supported in part by the presence of patients with no detectable TDP2 protein or 5'-TDP activity (Figure 6.9.A-B and Figure 6.10). If all repair was mediated only via TDP2, its loss would lead to embryonic lethality. Nonetheless, we can speculate that this alternative pathway is not as efficient as TDP2, since patient pathology is degenerative (Gómez-Herreros *et al.*, 2014). What repair pathway might be acting in the absence of TDP2 during G1 phase and in post-mitotic cells such as neurons? In RPE-1 cells lacking LIG4, DSB repair as measured by  $\gamma$ H2AX does not occur in G0/G1 over a 24h post-etoposide treatment period (unpublished observations). Thus, all damage measured in G1 is dependent in the NHEJ pathway for repair. NHEJ exhibits different kinetics and efficiency depending on DSB end complexity (Löbrich and Jeggo, 2017). The slow repair during NHEJ was shown to be dependent on ATM and Artemis (Riballo *et al.*, 2004). In addition, Artemis was recently shown to mediate a resection dependent NHEJ pathway in G1 cells (Biehs *et al.*, 2017). This pathway was dependent in the activities of CtIP, MRE11 exonuclease, BRCA1 and EXO1 to mediate resection. It is possible that DSB in *TDP2*<sup>-/-</sup> cells are repaired by this NHEJ mechanism. Interestingly, cells depleted of TDP2 in G1 have shown a requirement of ATM for DSB repair (Alvarez-Quilón *et al.*, 2014). Although, the role of Artemis on TOP2 repair is not clear, it was shown that after etoposide treatment these cells could repair DSBs with the same kinetics as wild type cells (Riballo *et al.*, 2004). However, there was also observations that Artemis defective cells are resistant to etoposide treatment (Hosono *et al.*, 2011). Furthermore, MRE11 was shown to be

involved in the repair of abortive TOP2cc, even in cells which TDP2 is present (Quennet *et al.*, 2011; Hoa *et al.*, 2016). Moreover, MRE11 endonuclease activity was shown to be promoted by NBN and cells from NBS patients were also defective to TOP2 induced breaks (Quennet *et al.*, 2011; Deshpande *et al.*, 2016). As a result, a pathway in which an endonucleolytic incision removes the TOP2cc from the breaks, followed by resection and processing by nucleases can be proposed. Presumably, TOP2cc can be repaired in cells by either nuclease or TDP2-dependent NHEJ pathways. TDP2 would likely be the preferred mechanism in non-replicating cells, because it facilitates potentially error-free DSB repair involving simple ligation of 4-bp cohesive termini. However, in cells lacking *TDP2*<sup>-/-</sup>, MRE11, together with ATM promote resection-dependent NHEJ repair. This latter process is potentially mutagenic, possibly resulting in deletions or insertions at the site of the break. Finally, in S/G2, the option remains for homologous recombination-mediated repair, in which the nuclease-resected DSBs can be also be repaired by an error-free mechanism.

#### 7.4. TDP2 and SCAR23

TDP2 absence was also investigated in relation to human genetic disease. Here we provided confirmation of the TDP2 mutation c.425+1G>A and first evidence of the mutation c.636+6\_636+3del involvement in SCAR23. Both mutations lead to defective splicing of *TDP2* pre-mRNA and nonsense-mediated decay of *TDP2* mRNA; resulting in the lack of TDP2 protein and activity (Figure 6.9.A-B and Figure 6.10). The molecular defect as well as the clinical symptoms of these patients matches previous patients' data (Table 6.1). Nonetheless the patient of American origin reported here presents additional symptoms that were linked to a possible secondary mitochondrial disorder (Zagnoli-Vieira *et al.*, 2018). Although the presentation of these new patient helps us to better understand SCAR23, it still does not make clear how prevalent this disease is. It has been highlighted in Chapter 6 that the c.636+6\_636+3del mutation has been identified in other clinical studies, therefore it will be essential for better understanding of this disease to try to identify these other patients, any comorbidities, in addition to any other mutation on TDP2 that could lead to the SCAR23 aetiology.

There are at least two aspects that require better understating in SCAR23. The first, is how repair of TOP2cc takes place in the absence of TDP2, which was discussed above (Section 7.2). The second is how endogenous TOP2cc arise in cells. From our own observations and others, we know that deletion of TDP2 in cells or mice, do not lead to any overt phenotype unless they are exposed to topoisomerase ‘poisons’ or inhibitors (Gómez-Herreros *et al.*, 2013; 2017; Zagnoli-Vieira *et al.*, 2018). Therefore, what is causing the SCAR23 aetiology? Recent evidence suggests that TOP2 $\beta$  could be at fault. First, physiological brain activity was shown to lead to  $\gamma$ H2AX accumulation (Suberbielle *et al.*, 2013). Later studies, linked transcription to DSB formation in TSS and long neural genes in neuronal progenitor cells (Schwer *et al.*, 2016; Wei *et al.*, 2016). Recent evidence also suggested that abortive TOP2 $\beta$  activity is required for the expression of early response genes, and it was shown that DSBs are associated with, and possibly required for, the initiation of transcriptional programs associated with learning and stress response (Madabhushi *et al.*, 2015). In addition to its role in transcription, there is mounting evidence that TOP2 $\beta$  activity is important for chromatin organisation (Canela *et al.*, 2017). TOP2 $\beta$  was shown to colocalise with the protein CCCTC-binding factor (CTCF) and cohesin, which was found to promote chromosome looping and chromatin organisation (Uusküla-Reimand *et al.*, 2016; Canela *et al.*, 2017). Therefore, it is feasible that increased abortive TOP2 $\beta$  activity is the cause of disease in SCAR23. Indeed, in the *Tdp2*<sup>41-3</sup> mouse a transcriptional defect was detected in the brain (Gómez-Herreros *et al.*, 2014). However, what is leading to neurodegeneration? Is it the transcriptional defect or the increased number of unrepaired DSBs? Also, could the absence of TDP2 lead to deleterious changes in chromosome conformation or structure due to increased persistence of TOP2 $\beta$  induced breaks at CTCF sites? Hopefully the cell lines generated here will help to answer some of these questions.



## 8. REFERENCES

- Ahel, I., Rass, U., El-Khamisy, S.F., et al. (2006) The neurodegenerative disease protein aprataxin resolves abortive DNA ligation intermediates. *Nature*, 443 (7112): 713–716
- Ahmad, M., Xu, D. and Wang, W. (2017) Type IA topoisomerases can be “magicians” for both DNA and RNA in all domains of life. *RNA biology*, 14 (7): 854–864
- Alt, F.W. and Schwer, B. (2018) DNA double-strand breaks as drivers of neural genomic change, function, and disease. *Dna Repair*
- Alt, F.W., Zhang, Y., Meng, F.-L., et al. (2013) Mechanisms of programmed DNA lesions and genomic instability in the immune system. *Cell*, 152 (3): 417–429
- Altschul, S.F., Gish, W., Miller, W., et al. (1990) Basic local alignment search tool. *Journal of molecular biology*, 215 (3): 403–410
- Alvarez-Quilón, A., Serrano-Benítez, A., Lieberman, J.A., et al. (2014) ATM specifically mediates repair of double-strand breaks with blocked DNA ends. *Nature communications*, 5 (1): 3347
- Austin, C.A., Sng, J.H., Patel, S., et al. (1993) Novel HeLa topoisomerase II is the II beta isoform: complete coding sequence and homology with other type II topoisomerases. *Biochimica et biophysica acta*, 1172 (3): 283–291
- Azarova, A.M., Lyu, Y.L., Lin, C.-P., et al. (2007) Roles of DNA topoisomerase II isozymes in chemotherapy and secondary malignancies. *Proceedings of the National Academy of Sciences of the United States of America*, 104 (26): 11014–11019
- Bagger, F.O., Sasivarevic, D., Sohi, S.H., et al. (2016) BloodSpot: a database of gene expression profiles and transcriptional programs for healthy and malignant haematopoiesis. *Nucleic acids research*, 44 (D1): D917–D924
- Baranello, L., Wojtowicz, D., Cui, K., et al. (2016) RNA Polymerase II Regulates Topoisomerase 1 Activity to Favor Efficient Transcription. *Cell*, 165 (2): 357–371
- Barrangou, R., Fremaux, C., Deveau, H., et al. (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science (New York, N.Y.)*, 315 (5819): 1709–1712
- Baxter, J. (2015) “Breaking up is hard to do”: the formation and resolution of sister chromatid intertwines. *Journal of molecular biology*, 427 (3): 590–607
- Ben-David, U., Siranosian, B., Ha, G., et al. (2018) Genetic and transcriptional evolution alters cancer cell line drug response. *Nature*, 560 (7718): 325–330
- Berger, J.M., Gamblin, S.J., Harrison, S.C., et al. (1996) Structure and mechanism of DNA topoisomerase II. *Nature*, 379 (6562): 225–232

- Bergerat, A., de Massy, B., Gadelle, D., et al. (1997) An atypical topoisomerase II from Archaea with implications for meiotic recombination. *Nature*, 386 (6623): 414–417
- Beucher, A., Birraux, J., Tchouandong, L., et al. (2009) ATM and Artemis promote homologous recombination of radiation-induced DNA double-strand breaks in G2. *Embo Journal*, 28 (21): 3413–3427
- Bhat, K.P. and Cortez, D. (2018) RPA and RAD51: fork reversal, fork protection, and genome stability. *Nature structural & molecular biology*, 25 (6): 446–453
- Biehs, R., Steinlage, M., Barton, O., et al. (2017) DNA Double-Strand Break Resection Occurs during Non-homologous End Joining in G1 but Is Distinct from Resection during Homologous Recombination. *Molecular Cell*, 65 (4): 671–684.e5
- Biesecker, L.G. and Green, R.C. (2014) Diagnostic clinical genome and exome sequencing. *The New England journal of medicine*, 371 (12): 1170–1170
- Blackford, A.N. and Jackson, S.P. (2017) ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. *Molecular Cell*, 66 (6): 801–817
- Bolotin, A., Quinquis, B., Sorokin, A., et al. (2005) Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology (Reading, England)*, 151 (Pt 8): 2551–2561
- Bras, J., Alonso, I., Barbot, C., et al. (2015) Mutations in PNKP cause recessive ataxia with oculomotor apraxia type 4. *American journal of human genetics*, 96 (3): 474–479
- Brouwer, I., Sitters, G., Candelli, A., et al. (2016) Sliding sleeves of XRCC4-XLF bridge DNA and connect fragments of broken DNA. *Nature*, 535 (7613): 566–569
- Buck, D., Malivert, L., de Chasseval, R., et al. (2006) Cernunnos, a novel nonhomologous end-joining factor, is mutated in human immunodeficiency with microcephaly. *Cell*, 124 (2): 287–299
- Caldecott, K.W. (2008) Single-strand break repair and genetic disease. *Nature reviews. Genetics*, 9 (8): 619–631
- Caldecott, K.W., Tucker, J.D. and Thompson, L.H. (1992) Construction of human XRCC1 minigenes that fully correct the CHO DNA repair mutant EM9. *Nucleic acids research*, 20 (17): 4575–4579
- Cancer Genome Atlas Network (2012) Comprehensive molecular portraits of human breast tumours. *Nature*, 490 (7418): 61–70
- Canela, A., Maman, Y., Jung, S., et al. (2017) Genome Organization Drives Chromosome Fragility. *Cell*, 170 (3): 507–521.e18
- Capranico, G., Tinelli, S., Austin, C.A., et al. (1992) Different patterns of gene expression of topoisomerase II isoforms in differentiated tissues during murine development. *Biochimica et biophysica acta*, 1132 (1): 43–48

- Carpenter, A.J. and Porter, A.C.G. (2004) Construction, characterization, and complementation of a conditional-lethal DNA topoisomerase II $\alpha$  mutant human cell line. *Molecular biology of the cell*, 15 (12): 5700–5711
- Castellotti, B., Mariotti, C., Rimoldi, M., et al. (2011) Ataxia with oculomotor apraxia type1 (AOA1): novel and recurrent aprataxin mutations, coenzyme Q10 analyses, and clinical findings in Italian patients. *Neurogenetics*, 12 (3): 193–201
- Champoux, J.J. (2001) DNA topoisomerases: structure, function, and mechanism. *Annual review of biochemistry*, 70 (1): 369–413
- Chan, K.-L., North, P.S. and Hickson, I.D. (2007) BLM is required for faithful chromosome segregation and its localization defines a class of ultrafine anaphase bridges. *Embo Journal*, 26 (14): 3397–3409
- Chang, H.H.Y., Pannunzio, N.R., Adachi, N., et al. (2017) Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nature reviews. Molecular cell biology*, 18 (8): 495–506
- Chanut, P., Britton, S., Coates, J., et al. (2016) Coordinated nuclease activities counteract Ku at single-ended DNA double-strand breaks. *Nature communications*, 7: 12889
- Chapman, J.R., Barral, P., Vannier, J.-B., et al. (2013) RIF1 is essential for 53BP1-dependent nonhomologous end joining and suppression of DNA double-strand break resection. *Molecular Cell*, 49 (5): 858–871
- Chen, B.P.C., Chan, D.W., Kobayashi, J., et al. (2005) Cell cycle dependence of DNA-dependent protein kinase phosphorylation in response to DNA double strand breaks. *Journal of Biological Chemistry*, 280 (15): 14709–14715
- Chen, S.H., Chan, N.-L. and Hsieh, T.-S. (2013) New mechanistic and functional insights into DNA topoisomerases. *Annual review of biochemistry*, 82 (1): 139–170
- Claros, M.G. and Vincens, P. (1996) Computational method to predict mitochondrially imported proteins and their targeting sequences. *European journal of biochemistry*, 241 (3): 779–786
- Clegg, R.M. (1995) Fluorescence resonance energy transfer. *Current opinion in biotechnology*, 6 (1): 103–110
- Confalonieri, M., Faè, M., Balestrazzi, A., et al. (2013) Enhanced osmotic stress tolerance in *Medicago truncatula* plants overexpressing the DNA repair gene *MtTdp2 $\alpha$*  (tyrosyl-DNA phosphodiesterase 2). *Plant Cell, Tissue and Organ Culture (PCTOC)*, 116 (2): 187–203
- Cong, L., Ran, F.A., Cox, D., et al. (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science (New York, N.Y.)*, 339 (6121): 819–823
- COSMIC (2018) Catalogue Of Somatic Mutations In Cancer (COSMIC) [online]. Available from: [cancer.sanger.ac.uk/](http://cancer.sanger.ac.uk/) [Accessed 31 August 2018]

- Costantini, S., Woodbine, L., Andreoli, L., et al. (2007) Interaction of the Ku heterodimer with the DNA ligase IV/Xrcc4 complex and its regulation by DNA-PK. *Dna Repair*, 6 (6): 712–722
- Costello, M., Fleharty, M., Abreu, J., et al. (2018) Characterization and remediation of sample index swaps by non-redundant dual indexing on massively parallel sequencing platforms. *BMC genomics*, 19 (1): 332
- Cowell, I.G., Sondka, Z., Smith, K., et al. (2012) Model for MLL translocations in therapy-related leukemia involving topoisomerase II $\beta$ -mediated DNA strand breaks and gene proximity. *Proceedings of the National Academy of Sciences of the United States of America*, 109 (23): 8989–8994
- Cruz-García, A., López-Saavedra, A. and Huertas, P. (2014) BRCA1 accelerates CtIP-mediated DNA-end resection. *Cell reports*, 9 (2): 451–459
- Curtin, N.J. (2012) DNA repair dysregulation from cancer driver to therapeutic target. *Nature reviews. Cancer*, 12 (12): 801–817
- D'Arpa, P., Beardmore, C. and Liu, L.F. (1990) Involvement of nucleic acid synthesis in cell killing mechanisms of topoisomerase poisons. *Cancer research*, 50 (21): 6919–6924
- Das, B.B., Huang, S.-Y.N., Murai, J., et al. (2014) PARP1-TDP1 coupling for the repair of topoisomerase I-induced DNA damage. *Nucleic acids research*, 42 (7): 4435–4449
- Date, H., Onodera, O., Tanaka, H., et al. (2001) Early-onset ataxia with ocular motor apraxia and hypoalbuminemia is caused by mutations in a new HIT superfamily gene. *Nature genetics*, 29 (2): 184–188
- Debéthune, L., Kohlhagen, G., Grandas, A., et al. (2002) Processing of nucleopeptides mimicking the topoisomerase I-DNA covalent complex by tyrosyl-DNA phosphodiesterase. *Nucleic acids research*, 30 (5): 1198–1204
- Delgado, J.L., Hsieh, C.-M., Chan, N.-L., et al. (2018) Topoisomerases as anticancer targets. *Biochemical Journal*, 475 (2): 373–398
- Demuth, I. and Digweed, M. (2007) The clinical manifestation of a defective response to DNA double-strand breaks as exemplified by Nijmegen breakage syndrome. *Oncogene*, 26 (56): 7792–7798
- DePristo, M.A., Banks, E., Poplin, R., et al. (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature genetics*, 43 (5): 491–498
- Deriano, L. and Roth, D.B. (2013) Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. *Annual review of genetics*, 47 (1): 433–455

- Desai, S.D., Li, T.K., Rodriguez-Bauman, A., et al. (2001) Ubiquitin/26S proteasome-mediated degradation of topoisomerase I as a resistance mechanism to camptothecin in tumor cells. *Cancer research*, 61 (15): 5926–5932
- Deshpande, R.A., Lee, J.-H., Arora, S., et al. (2016) Nbs1 Converts the Human Mre11/Rad50 Nuclease Complex into an Endo/Exonuclease Machine Specific for Protein-DNA Adducts. *Molecular Cell*, 64 (3): 593–606
- Dev, H., Chiang, T.-W.W., Lescale, C., et al. (2018) Shieldin complex promotes DNA end-joining and counters homologous recombination in BRCA1-null cells. *Nature cell biology*, 20 (8): 954–965
- Deweese, J.E. and Osheroff, N. (2009) The DNA cleavage reaction of topoisomerase II: wolf in sheep's clothing. *Nucleic acids research*, 37 (3): 738–748
- Didenko, V.V. (2001) DNA probes using fluorescence resonance energy transfer (FRET): designs and applications. *BioTechniques*, 31 (5): 1106–16– 1118– 1120–1
- Do, P.M., Varanasi, L., Fan, S., et al. (2012) Mutant p53 cooperates with ETS2 to promote etoposide resistance. *Genes & development*, 26 (8): 830–845
- Doudna, J.A. and Charpentier, E. (2014) Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science (New York, N.Y.)*, 346 (6213): 1258096–1258096
- Downs, J.A. and Jackson, S.P. (2004) A means to a DNA end: the many roles of Ku. *Nature reviews. Molecular cell biology*, 5 (5): 367–378
- Eisenberg-Bord, M., Shai, N., Schuldiner, M., et al. (2016) A Tether Is a Tether Is a Tether: Tethering at Membrane Contact Sites. *Developmental cell*, 39 (4): 395–409
- El-Khamisy, S.F., Saifi, G.M., Weinfeld, M., et al. (2005) Defective DNA single-strand break repair in spinocerebellar ataxia with axonal neuropathy-1. *Nature*, 434 (7029): 108–113
- Emanuelsson, O., Brunak, S., Heijne, von, G., et al. (2007) Locating proteins in the cell using TargetP, SignalP and related tools. *Nature protocols*, 2 (4): 953–971
- Errichiello, E. and Zuffardi, O. (2017) Personal Communication.
- Escribano-Díaz, C., Orthwein, A., Fradet-Turcotte, A., et al. (2013) A cell cycle-dependent regulatory circuit composed of 53BP1-RIF1 and BRCA1-CtIP controls DNA repair pathway choice. *Molecular Cell*, 49 (5): 872–883
- ESP, N.G.E.S.P. (2018) Exome Variant Server [online]. Available from: <http://evs.gs.washington.edu/EVS/> [Accessed 27 August 2017]
- Falck, J., Coates, J. and Jackson, S.P. (2005) Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage. *Nature*, 434 (7033): 605–611

- Flett, F.J., Ruksenaite, E., Armstrong, L.A., et al. (2018) Structural basis for DNA 3'-end processing by human tyrosyl-DNA phosphodiesterase 1. *Nature communications*, 9 (1): 24
- Forbes, S.A., Beare, D., Boutselakis, H., et al. (2017) COSMIC: somatic cancer genetics at high-resolution. *Nucleic acids research*, 45 (D1): D777–D783
- Fradet-Turcotte, A., Canny, M.D., Escribano-Díaz, C., et al. (2013) 53BP1 is a reader of the DNA-damage-induced H2A Lys 15 ubiquitin mark. *Nature*, 499 (7456): 50–54
- Fu, Y., Sander, J.D., Reyon, D., et al. (2014) Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nature Biotechnology*, 32 (3): 279–284
- Fukasawa, Y., Tsuji, J., Fu, S.-C., et al. (2015) MitoFates: improved prediction of mitochondrial targeting sequences and their cleavage sites. *Molecular & cellular proteomics : MCP*, 14 (4): 1113–1126
- Gao, R., Huang, S.-Y.N., Marchand, C., et al. (2012) Biochemical characterization of human tyrosyl-DNA phosphodiesterase 2 (TDP2/TTRAP): a Mg(2+)/Mn(2+)-dependent phosphodiesterase specific for the repair of topoisomerase cleavage complexes. *Journal of Biological Chemistry*, 287 (36): 30842–30852
- Gao, R., Schellenberg, M.J., Huang, S.-Y.N., et al. (2014) Proteolytic degradation of topoisomerase II (Top2) enables the processing of Top2·DNA and Top2·RNA covalent complexes by tyrosyl-DNA-phosphodiesterase 2 (TDP2). *Journal of Biological Chemistry*, 289 (26): 17960–17969
- Garcia-Diaz, B., Barca, E., Balreira, A., et al. (2015) Lack of aprataxin impairs mitochondrial functions via downregulation of the APE1/NRF1/NRF2 pathway. *Human molecular genetics*, 24 (16): 4516–4529
- Gasiunas, G., Barrangou, R., Horvath, P., et al. (2012) Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 109 (39): E2579–86
- Georgoulis, A., Vorgias, C.E., Chrousos, G.P., et al. (2017) Genome Instability and  $\gamma$ H2AX. *International journal of molecular sciences*, 18 (9): 1979
- Ghezraoui, H., Oliveira, C., Becker, J.R., et al. (2018) 53BP1 cooperation with the REV7-shieldin complex underpins DNA structure-specific NHEJ. *Nature*, 560 (7716): 122–127
- Gioanni, J., Le Francois, D., Zanghellini, E., et al. (1990) Establishment and characterisation of a new tumorigenic cell line with a normal karyotype derived from a human breast adenocarcinoma. *British Journal of Cancer*, 62 (1): 8–13
- Goodarzi, A.A. and Jeggo, P.A. (2013) The repair and signaling responses to DNA double-strand breaks. *Advances in genetics*, 82: 1–45

- Goodarzi, A.A., Yu, Y., Riballo, E., et al. (2006) DNA-PK autophosphorylation facilitates Artemis endonuclease activity. *Embo Journal*, 25 (16): 3880–3889
- Gómez-Herreros, F., Romero-Granados, R., Zeng, Z., et al. (2013) TDP2-dependent non-homologous end-joining protects against topoisomerase II-induced DNA breaks and genome instability in cells and in vivo. *Maizels, N. (ed.). PLoS genetics*, 9 (3): e1003226
- Gómez-Herreros, F., Schuurs-Hoeijmakers, J.H.M., McCormack, M., et al. (2014) TDP2 protects transcription from abortive topoisomerase activity and is required for normal neural function. *Nature genetics*, 46 (5): 516–521
- Gómez-Herreros, F., Zagnoli-Vieira, G., Ntai, I., et al. (2017) TDP2 suppresses chromosomal translocations induced by DNA topoisomerase II during gene transcription. *Nature communications*, 8 (1): 233
- Grawunder, U., Wilm, M., Wu, X., et al. (1997) Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. *Nature*, 388 (6641): 492–495
- Grue, P., Grässer, A., Sehested, M., et al. (1998) Essential mitotic functions of DNA topoisomerase II $\alpha$  are not adopted by topoisomerase II $\beta$  in human H69 cells. *Journal of Biological Chemistry*, 273 (50): 33660–33666
- Grundy, G.J., Rulten, S.L., Arribas-Bosacoma, R., et al. (2016) The Ku-binding motif is a conserved module for recruitment and stimulation of non-homologous end-joining proteins. *Nature communications*, 7: 11242
- Grundy, G.J., Rulten, S.L., Zeng, Z., et al. (2013) APLF promotes the assembly and activity of non-homologous end joining protein complexes. *Embo Journal*, 32 (1): 112–125
- Guo, C., Nakazawa, Y., Woodbine, L., et al. (2015) XRCC4 deficiency in human subjects causes a marked neurological phenotype but no overt immunodeficiency. *The Journal of allergy and clinical immunology*, 136 (4): 1007–1017
- Gupta, R., Somyajit, K., Narita, T., et al. (2018) DNA Repair Network Analysis Reveals Shieldin as a Key Regulator of NHEJ and PARP Inhibitor Sensitivity. *Cell*, 173 (4): 972–988.e23
- Györfy, B., Lanczky, A., Eklund, A.C., et al. (2010) An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast cancer research and treatment*, 123 (3): 725–731
- Haffner, M.C., Aryee, M.J., Toubaji, A., et al. (2010) Androgen-induced TOP2B-mediated double-strand breaks and prostate cancer gene rearrangements. *Nature genetics*, 42 (8): 668–675
- Hanahan, D. and Weinberg, R.A. (2011) Hallmarks of cancer: the next generation. *Cell*, 144 (5): 646–674

Hanzlikova, H., Gittens, W., Krejcikova, K., et al. (2017) Overlapping roles for PARP1 and PARP2 in the recruitment of endogenous XRCC1 and PNKP into oxidized chromatin. *Nucleic acids research*, 45 (5): 2546–2557

Hanzlikova, H., Kalasova, I., Demin, A.A., et al. (2018) The Importance of Poly(ADP-Ribose) Polymerase as a Sensor of Unligated Okazaki Fragments during DNA Replication. *Molecular Cell*, 71 (2): 319–331.e3

Hegde, M.L., Hazra, T.K. and Mitra, S. (2008) Early steps in the DNA base excision/single-strand interruption repair pathway in mammalian cells. *Cell research*, 18 (1): 27–47

Hegde, R.S. and Bernstein, H.D. (2006) The surprising complexity of signal sequences. *Trends in biochemical sciences*, 31 (10): 563–571

Heigwer, F., Kerr, G. and Boutros, M. (2014) E-CRISP: fast CRISPR target site identification. *Nature Methods*, 11 (2): 122–123

Her, J. and Bunting, S.F. (2018) How cells ensure correct repair of DNA double-strand breaks. *Journal of Biological Chemistry*, 293 (27): 10502–10511

Hirano, R., Interthal, H., Huang, C., et al. (2007) Spinocerebellar ataxia with axonal neuropathy: consequence of a Tdp1 recessive neomorphic mutation? *Embo Journal*, 26 (22): 4732–4743

Hoa, N.N., Shimizu, T., Zhou, Z.W., et al. (2016) Mre11 Is Essential for the Removal of Lethal Topoisomerase 2 Covalent Cleavage Complexes. *Molecular Cell*, 64 (3): 580–592

Hoch, N.C., Hanzlikova, H., Rulten, S.L., et al. (2017) XRCC1 mutation is associated with PARP1 hyperactivation and cerebellar ataxia. *Nature*, 541 (7635): 87–91

Hoeijmakers, J.H.J. (2009) DNA damage, aging, and cancer. *The New England journal of medicine*, 361 (15): 1475–1485

Hornyak, P., Askwith, T., Walker, S., et al. (2016) Mode of action of DNA-competitive small molecule inhibitors of tyrosyl DNA phosphodiesterase 2. *Biochemical Journal*, 473 (13): 1869–1879

Hosono, Y., Abe, T., Ishiai, M., et al. (2011) The role of SNM1 family nucleases in etoposide-induced apoptosis. *Biochemical and Biophysical Research Communications*, 410 (3): 568–573

Hsiang, Y.H., Hertzberg, R., Hecht, S., et al. (1985) Camptothecin induces protein-linked DNA breaks via mammalian DNA topoisomerase I. *Journal of Biological Chemistry*, 260 (27): 14873–14878

Hsu, P.D., Lander, E.S. and Zhang, F. (2014) Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell*, 157 (6): 1262–1278



- Hu, J., Selby, C.P., Adar, S., et al. (2017) Molecular mechanisms and genomic maps of DNA excision repair in *Escherichia coli* and humans. *Journal of Biological Chemistry*, 292 (38): 15588–15597
- Huang, S.-Y.N., Dalla Rosa, I., Michaels, S.A., et al. (2018) Mitochondrial tyrosyl-DNA phosphodiesterase 2 and its TDP2S short isoform. *EMBO reports*, 19 (3): e42139
- Hudson, J.J.R., Chiang, S.-C., Wells, O.S., et al. (2012) SUMO modification of the neuroprotective protein TDP1 facilitates chromosomal single-strand break repair. *Nature communications*, 3 (1): 733
- Huertas, P. and Jackson, S.P. (2009) Human CtIP mediates cell cycle control of DNA end resection and double strand break repair. *Journal of Biological Chemistry*, 284 (14): 9558–9565
- Hung, P.J., Johnson, B., Chen, B.-R., et al. (2018) MRI Is a DNA Damage Response Adaptor during Classical Non-homologous End Joining. *Molecular Cell*, 71 (2): 332–342.e8
- Interthal, H., Chen, H.J., Kehl-Fie, T.E., et al. (2005) SCAN1 mutant Tdp1 accumulates the enzyme--DNA intermediate and causes camptothecin hypersensitivity. *Embo Journal*, 24 (12): 2224–2233
- Interthal, H., Pouliot, J.J. and Champoux, J.J. (2001) The tyrosyl-DNA phosphodiesterase Tdp1 is a member of the phospholipase D superfamily. *Proceedings of the National Academy of Sciences of the United States of America*, 98 (21): 12009–12014
- Ishida, R., Miki, T., Narita, T., et al. (1991) Inhibition of intracellular topoisomerase II by antitumor bis(2,6-dioxopiperazine) derivatives: mode of cell growth inhibition distinct from that of cleavable complex-forming type inhibitors. *Cancer research*, 51 (18): 4909–4916
- Iwasaki, H., Takahagi, M., Shiba, T., et al. (1991) *Escherichia coli* RuvC protein is an endonuclease that resolves the Holliday structure. *Embo Journal*, 10 (13): 4381–4389
- Jackson, S.P. and Bartek, J. (2009) The DNA-damage response in human biology and disease. *Nature*, 461 (7267): 1071–1078
- Jazayeri, A., Falck, J., Lukas, C., et al. (2006) ATM- and cell cycle-dependent regulation of ATR in response to DNA double-strand breaks. *Nature cell biology*, 8 (1): 37–45
- Jiang, F. and Doudna, J.A. (2017) CRISPR-Cas9 Structures and Mechanisms. *Annual review of biophysics*, 46 (1): 505–529
- Jilani, A., Ramotar, D., Slack, C., et al. (1999) Molecular cloning of the human gene, PNKP, encoding a polynucleotide kinase 3'-phosphatase and evidence for its role in repair of DNA strand breaks caused by oxidative damage. *Journal of Biological Chemistry*, 274 (34): 24176–24186

- Jinek, M., Chylinski, K., Fonfara, I., et al. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (New York, N.Y.)*, 337 (6096): 816–821
- Jinek, M., East, A., Cheng, A., et al. (2013) RNA-programmed genome editing in human cells. *eLife*, 2: e00471
- Jiricny, J. (2013) Postreplicative mismatch repair. *Cold Spring Harbor perspectives in biology*, 5 (4): a012633–a012633
- Ju, B.-G. and Rosenfeld, M.G. (2006) A breaking strategy for topoisomerase II $\beta$ /PARP-1-dependent regulated transcription. *Cell cycle (Georgetown, Tex.)*, 5 (22): 2557–2560
- Ju, B.-G., Lunyak, V.V., Perissi, V., et al. (2006) A topoisomerase II $\beta$ -mediated dsDNA break required for regulated transcription. *Science (New York, N.Y.)*, 312 (5781): 1798–1802
- Karimi-Busheri, F., Daly, G., Robins, P., et al. (1999) Molecular characterization of a human DNA kinase. *Journal of Biological Chemistry*, 274 (34): 24187–24194
- Katyal, S., El-Khamisy, S.F., Russell, H.R., et al. (2007) TDP1 facilitates chromosomal single-strand break repair in neurons and is neuroprotective in vivo. *Embo Journal*, 26 (22): 4720–4731
- Katyal, S., Lee, Y., Nitiss, K.C., et al. (2014) Aberrant topoisomerase-1 DNA lesions are pathogenic in neurodegenerative genome instability syndromes. *Nature neuroscience*, 17 (6): 813–821
- Kawale, A.S. and Povirk, L.F. (2018) Tyrosyl-DNA phosphodiesterases: rescuing the genome from the risks of relaxation. *Nucleic acids research*, 46 (2): 520–537
- Kawale, A.S., Akopiants, K., Valerie, K., et al. (2018) TDP1 suppresses mis-joining of radiomimetic DNA double-strand breaks and cooperates with Artemis to promote optimal nonhomologous end joining. *Nucleic acids research*, 17: 703
- Keeney, S. (2008) Spo11 and the Formation of DNA Double-Strand Breaks in Meiosis. *Genome dynamics and stability*, 2 (Chapter 26): 81–123
- Keeney, S. and Kleckner, N. (1995) Covalent protein-DNA complexes at the 5' strand termini of meiosis-specific double-strand breaks in yeast. *Proceedings of the National Academy of Sciences of the United States of America*, 92 (24): 11274–11278
- Keeney, S., Giroux, C.N. and Kleckner, N. (1997) Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. *Cell*, 88 (3): 375–384
- Kerzendorfer, C. and O'Driscoll, M. (2009) Human DNA damage response and repair deficiency syndromes: linking genomic instability and cell cycle checkpoint proficiency. *Dna Repair*, 8 (9): 1139–1152

- Khanna, A. (2015) DNA damage in cancer therapeutics: a boon or a curse? *Cancer research*, 75 (11): 2133–2138
- Kim, N. and Jinks-Robertson, S. (2017) The Top1 paradox: Friend and foe of the eukaryotic genome. *Dna Repair*, 56: 33–41
- King, I.F., Yandava, C.N., Mabb, A.M., et al. (2013) Topoisomerases facilitate transcription of long genes linked to autism. *Nature*, 501 (7465): 58–62
- Knockenbauer, K.E. and Schwartz, T.U. (2016) The Nuclear Pore Complex as a Flexible and Dynamic Gate. *Cell*, 164 (6): 1162–1171
- Konstantinopoulos, P.A., Ceccaldi, R., Shapiro, G.I., et al. (2015) Homologous Recombination Deficiency: Exploiting the Fundamental Vulnerability of Ovarian Cancer. *Cancer discovery*, 5 (11): 1137–1154
- Kosicki, M., Tomberg, K. and Bradley, A. (2018) Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nature Biotechnology*, 36 (8): 765–771
- Koster, D.A., Croquette, V., Dekker, C., et al. (2005) Friction and torque govern the relaxation of DNA supercoils by eukaryotic topoisomerase IB. *Nature*, 434 (7033): 671–674
- Kozak, M. (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*, 44 (2): 283–292
- Kunkel, T.A. and Erie, D.A. (2015) Eukaryotic Mismatch Repair in Relation to DNA Replication. *Annual review of genetics*, 49 (1): 291–313
- Lam, I. and Keeney, S. (2014) Mechanism and regulation of meiotic recombination initiation. *Cold Spring Harbor perspectives in biology*, 7 (1): a016634
- Lander, E.S. (2016) The Heroes of CRISPR. *Cell*, 164 (1-2): 18–28
- Langereis, M.A., Feng, Q., Nelissen, F.H.T., et al. (2014) Modification of picornavirus genomic RNA using “click” chemistry shows that unlinking of the VPg peptide is dispensable for translation and replication of the incoming viral RNA. *Nucleic acids research*, 42 (4): 2473–2482
- Le Ber, I., Dubourg, O., Benoist, J.F., et al. (2007) Muscle coenzyme Q10 deficiencies in ataxia with oculomotor apraxia 1. *Neurology*, 68 (4): 295–297
- Ledesma, F.C., El-Khamisy, S.F., Zuma, M.C., et al. (2009) A human 5'-tyrosyl DNA phosphodiesterase that repairs topoisomerase-mediated DNA damage. *Nature*, 461 (7264): 674–678
- Li, C., Fan, S., Owonikoko, T.K., et al. (2011) Oncogenic role of EAPII in lung cancer development and its activation of the MAPK-ERK pathway. *Oncogene*, 30 (35): 3802–3812

- Li, H. and Durbin, R. (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* (Oxford, England), 26 (5): 589–595
- Li, J., Summerlin, M., Nitiss, K.C., et al. (2017) TDP1 is required for efficient non-homologous end joining in human cells. *Dna Repair*, 60: 40–49
- Li, W. and Wang, J.C. (1998) Mammalian DNA topoisomerase III $\alpha$  is essential in early embryogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, 95 (3): 1010–1013
- Li, W., Cowley, A., Uludag, M., et al. (2015) The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic acids research*, 43 (W1): W580–4
- Liao, H., Winkfein, R.J., Mack, G., et al. (1995) CENP-F is a protein of the nuclear matrix that assembles onto kinetochores at late G2 and is rapidly degraded after mitosis. *The Journal of cell biology*, 130 (3): 507–518
- Libura, J., Slater, D.J., Felix, C.A., et al. (2005) Therapy-related acute myeloid leukemia-like MLL rearrangements are induced by etoposide in primary human CD34+ cells and remain stable after clonal expansion. *Blood*, 105 (5): 2124–2131
- Lieber, M.R. (2010) NHEJ and its backup pathways in chromosomal translocations. *Nature structural & molecular biology*, 17 (4): 393–395
- Lim, D.S. and Hasty, P. (1996) A mutation in mouse rad51 results in an early embryonic lethal that is suppressed by a mutation in p53. *Molecular and Cellular Biology*, 16 (12): 7133–7143
- Lin, C.-P., Ban, Y., Lyu, Y.L., et al. (2009) Proteasome-dependent processing of topoisomerase I-DNA adducts into DNA double strand breaks at arrested replication forks. *Journal of Biological Chemistry*, 284 (41): 28084–28092
- Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature*, 362 (6422): 709–715
- Lindahl, T. (2016) The Intrinsic Fragility of DNA (Nobel Lecture). *Angewandte Chemie* (International ed. in English), 55 (30): 8528–8534
- Liu, L., Liu, C.-C. and Alberts, B.M. (1980) Type II DNA topoisomerases: Enzymes that can unknot a topologically knotted DNA molecule via a reversible double-strand break. *Cell*, 19 (3): 697–707
- Lord, C.J. and Ashworth, A. (2017) PARP inhibitors: Synthetic lethality in the clinic. *Science* (New York, N.Y.), 355 (6330): 1152–1158
- Lord, C.J., Tutt, A.N.J. and Ashworth, A. (2015) Synthetic lethality and cancer therapy: lessons learned from the development of PARP inhibitors. *Annual review of medicine*, 66 (1): 455–470

- Löbrich, M. and Jeggo, P. (2017) A Process of Resection-Dependent Nonhomologous End Joining Involving the Goddess Artemis. *Trends in biochemical sciences*, 42 (9): 690–701
- Lyu, Y.L., Lin, C.-P., Azarova, A.M., et al. (2006) Role of topoisomerase II $\beta$  in the expression of developmentally regulated genes. *Molecular and Cellular Biology*, 26 (21): 7929–7941
- Maciejewski, S., Ullmer, W. and Semler, B.L. (2018) VPg unlinkase/TDP2 in cardiovirus infected cells: Re-localization and proteolytic cleavage. *Virology*, 516: 139–146
- Madabhushi, R. (2018) The Roles of DNA Topoisomerase II $\beta$  in Transcription. *International journal of molecular sciences*, 19 (7): 1917
- Madabhushi, R., Gao, F., Pfenning, A.R., et al. (2015) Activity-Induced DNA Breaks Govern the Expression of Neuronal Early-Response Genes. *Cell*, 161 (7): 1592–1605
- Mahajan, K.N., Nick McElhinny, S.A., Mitchell, B.S., et al. (2002) Association of DNA polymerase mu (pol mu) with Ku and ligase IV: role for pol mu in end-joining double-strand break repair. *Molecular and Cellular Biology*, 22 (14): 5194–5202
- Mailand, N., Bekker-Jensen, S., Faustrup, H., et al. (2007) RNF8 ubiquitylates histones at DNA double-strand breaks and promotes assembly of repair proteins. *Cell*, 131 (5): 887–900
- Makarova, K.S., Grishin, N.V., Shabalina, S.A., et al. (2006) A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biology direct*, 1 (1): 7
- Mali, P., Esvelt, K.M. and Church, G.M. (2013a) Cas9 as a versatile tool for engineering biology. *Nature Methods*, 10 (10): 957–963
- Mali, P., Yang, L., Esvelt, K.M., et al. (2013b) RNA-guided human genome engineering via Cas9. *Science (New York, N.Y.)*, 339 (6121): 823–826
- Marchand, C., Abdelmalak, M., Kankanala, J., et al. (2016) Deazaflavin Inhibitors of Tyrosyl-DNA Phosphodiesterase 2 (TDP2) Specific for the Human Enzyme and Active against Cellular TDP2. *ACS chemical biology*, 11 (7): 1925–1933
- Marteijn, J.A., Lans, H., Vermeulen, W., et al. (2014) Understanding nucleotide excision repair and its roles in cancer and ageing. *Nature reviews. Molecular cell biology*, 15 (7): 465–481
- Matos, J. and West, S.C. (2014) Holliday junction resolution: regulation in space and time. *Dna Repair*, 19: 176–181
- Matsuoka, S., Ballif, B.A., Smogorzewska, A., et al. (2007) ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science (New York, N.Y.)*, 316 (5828): 1160–1166

- Mattioli, F., Vissers, J.H.A., van Dijk, W.J., et al. (2012) RNF168 ubiquitinates K13-15 on H2A/H2AX to drive DNA damage signaling. *Cell*, 150 (6): 1182–1195
- McKenna, A., Hanna, M., Banks, E., et al. (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research*, 20 (9): 1297–1303
- McKinnon, P.J. (2013) Maintaining genome stability in the nervous system. *Nature neuroscience*, 16 (11): 1523–1529
- McKinnon, P.J. (2017) Genome integrity and disease prevention in the nervous system. *Genes & development*, 31 (12): 1180–1194
- Minotti, G., Menna, P., Salvatorelli, E., et al. (2004) Anthracyclines: molecular advances and pharmacologic developments in antitumor activity and cardiotoxicity. *Pharmacological reviews*, 56 (2): 185–229
- Modrich, P. (2016) Mechanisms in E. coli and Human Mismatch Repair (Nobel Lecture). *Angewandte Chemie (International ed. in English)*, 55 (30): 8490–8501
- Moreira, M.C., Barbot, C., Tachi, N., et al. (2001) The gene mutated in ataxia-ocular apraxia 1 encodes the new HIT/Zn-finger protein aprataxin. *Nature genetics*, 29 (2): 189–193
- Morham, S.G., Kluckman, K.D., Voulomanos, N., et al. (1996) Targeted disruption of the mouse topoisomerase I gene by camptothecin selection. *Molecular and Cellular Biology*, 16 (12): 6804–6809
- Moynahan, M.E. (2002) The cancer connection: BRCA1 and BRCA2 tumor suppression in mice and humans. *Oncogene*, 21 (58): 8994–9007
- Murai, J., Huang, S.-Y.N., Das, B.B., et al. (2012) Tyrosyl-DNA phosphodiesterase 1 (TDP1) repairs DNA damage induced by topoisomerases I and II and base alkylation in vertebrate cells. *Journal of Biological Chemistry*, 287 (16): 12848–12857
- Murray, J.M. and Carr, A.M. (2018) Integrating DNA damage repair with the cell cycle. *Current opinion in cell biology*, 52: 120–125
- Negrini, S., Gorgoulis, V.G. and Halazonetis, T.D. (2010) Genomic instability--an evolving hallmark of cancer. *Nature reviews. Molecular cell biology*, 11 (3): 220–228
- Nicholls, T.J., Nadalutti, C.A., Motori, E., et al. (2018) Topoisomerase 3 $\alpha$  Is Required for Decatenation and Segregation of Human mtDNA. *Molecular Cell*, 69 (1): 9–23.e6
- Nielsen, H. (2017) Predicting Secretory Proteins with SignalP. *Methods in molecular biology (Clifton, N.J.)*, 1611 (Chapter 6): 59–73
- Nitiss, J.L. (2009a) DNA topoisomerase II and its growing repertoire of biological functions. *Nature reviews. Cancer*, 9 (5): 327–337

- Nitiss, J.L. (2009b) Targeting DNA topoisomerase II in cancer chemotherapy. *Nature reviews. Cancer*, 9 (5): 338–350
- Niyazov, D. (2016) Clinical report.
- Niyazov, D.M., Kahler, S.G. and Frye, R.E. (2016) Primary Mitochondrial Disease and Secondary Mitochondrial Dysfunction: Importance of Distinction for Diagnosis and Treatment. *Molecular syndromology*, 7 (3): 122–137
- Noordermeer, S.M., Adam, S., Setiaputra, D., et al. (2018) The shieldin complex mediates 53BP1-dependent DNA repair. *Nature*, 560 (7716): 117–121
- O'Driscoll, M. and Jeggo, P.A. (2003) Clinical impact of ATR checkpoint signalling failure in humans. *Cell cycle (Georgetown, Tex.)*, 2 (3): 194–195
- O'Driscoll, M., Cerosaletti, K.M., Girard, P.M., et al. (2001) DNA ligase IV mutations identified in patients exhibiting developmental delay and immunodeficiency. *Molecular Cell*, 8 (6): 1175–1185
- O'Driscoll, M., Ruiz-Perez, V.L., Woods, C.G., et al. (2003) A splicing mutation affecting expression of ataxia-telangiectasia and Rad3-related protein (ATR) results in Seckel syndrome. *Nature genetics*, 33 (4): 497–501
- Paull, T.T. (2018) 20 Years of Mre11 Biology: No End in Sight. *Molecular Cell*, 71 (3): 419–427
- Pei, H., Yordy, J.S., Leng, Q., et al. (2003) EAP1 interacts with ETS1 and modulates its transcriptional function. *Oncogene*, 22 (18): 2699–2709
- Pellegrino, S., Michelena, J., Teloni, F., et al. (2017) Replication-Coupled Dilution of H4K20me2 Guides 53BP1 to Pre-replicative Chromatin. *Cell reports*, 19 (9): 1819–1831
- Perlman, S.L., Boder Deceased, E., Sedgewick, R.P., et al. (2012) Ataxia-telangiectasia. *Handbook of clinical neurology*, 103: 307–332
- Perou, C.M., Sørlie, T., Eisen, M.B., et al. (2000) Molecular portraits of human breast tumours. *Nature*, 406 (6797): 747–752
- Pommier, Y. (2006) Topoisomerase I inhibitors: camptothecins and beyond. *Nature reviews. Cancer*, 6 (10): 789–802
- Pommier, Y. (2009) DNA topoisomerase I inhibitors: chemistry, biology, and interfacial inhibition. *Chemical reviews*, 109 (7): 2894–2902
- Pommier, Y., Barcelo, J.M., Rao, V.A., et al. (2006) Repair of topoisomerase I-mediated DNA damage. *Progress in nucleic acid research and molecular biology*, 81: 179–229
- Pommier, Y., Huang, S.-Y.N., Gao, R., et al. (2014) Tyrosyl-DNA-phosphodiesterases (TDP1 and TDP2). *Dna Repair*, 19: 114–129

Pommier, Y., Leo, E., Zhang, H., et al. (2010) DNA topoisomerases and their poisoning by anticancer and antibacterial drugs. *Chemistry & biology*, 17 (5): 421–433

Pommier, Y., Sun, Y., Huang, S.-Y.N., et al. (2016) Roles of eukaryotic topoisomerases in transcription, replication and genomic stability. *Nature reviews. Molecular cell biology*, 17 (11): 703–721

Pouliot, J.J., Yao, K.C., Robertson, C.A., et al. (1999) Yeast gene for a Tyr-DNA phosphodiesterase that repairs topoisomerase I complexes. *Science (New York, N.Y.)*, 286 (5439): 552–555

Prakash, R., Zhang, Y., Feng, W., et al. (2015) Homologous recombination and human health: the roles of BRCA1, BRCA2, and associated proteins. *Cold Spring Harbor perspectives in biology*, 7 (4): a016600

Pype, S., Declercq, W., Ibrahimi, A., et al. (2000) TTRAP, a novel protein that associates with CD40, tumor necrosis factor (TNF) receptor-75 and TNF receptor-associated factors (TRAFs), and that inhibits nuclear factor-kappa B activation. *Journal of Biological Chemistry*, 275 (24): 18586–18593

Quennet, V., Beucher, A., Barton, O., et al. (2011) CtIP and MRN promote non-homologous end-joining of etoposide-induced DNA double-strand breaks in G1. *Nucleic acids research*, 39 (6): 2144–2152

Quinzii, C.M., Kattah, A.G., Naini, A., et al. (2005) Coenzyme Q deficiency and cerebellar ataxia associated with an aprataxin mutation. *Neurology*, 64 (3): 539–541

Rao, T., Gao, R., Takada, S., et al. (2016) Novel TDP2-ubiquitin interactions and their importance for the repair of topoisomerase II-mediated DNA damage. *Nucleic acids research*, 44 (21): 10201–10215

Raoof, A., Depledge, P., Hamilton, N.M., et al. (2013) Toxoflavins and deazaflavins as the first reported selective small molecule inhibitors of tyrosyl-DNA phosphodiesterase II. *Journal of medicinal chemistry*, 56 (16): 6352–6370

Ray Chaudhuri, A. and Nussenzweig, A. (2017) The multifaceted roles of PARP1 in DNA repair and chromatin remodelling. *Nature reviews. Molecular cell biology*, 18 (10): 610–621

Reynolds, J.J. and Stewart, G.S. (2013) A nervous predisposition to unrepaired DNA double strand breaks. *Dna Repair*, 12 (8): 588–599

Riballo, E., Kühne, M., Rief, N., et al. (2004) A pathway of double-strand break rejoining dependent upon ATM, Artemis, and proteins locating to gamma-H2AX foci. *Molecular Cell*, 16 (5): 715–724

Rieusset, J. (2018) The role of endoplasmic reticulum-mitochondria contact sites in the control of glucose homeostasis: an update. *Cell death & disease*, 9 (3): 388



- Rodrigues-Lima, F., Josephs, M., Katan, M., et al. (2001) Sequence analysis identifies TTRAP, a protein that associates with CD40 and TNF receptor-associated factors, as a member of a superfamily of divalent cation-dependent phosphodiesterases. *Biochemical and Biophysical Research Communications*, 285 (5): 1274–1279
- Rogakou, E.P., Boon, C., Redon, C., et al. (1999) Megabase chromatin domains involved in DNA double-strand breaks in vivo. *The Journal of cell biology*, 146 (5): 905–916
- Rogakou, E.P., Pilch, D.R., Orr, A.H., et al. (1998) DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *Journal of Biological Chemistry*, 273 (10): 5858–5868
- Rulten, S.L. and Caldecott, K.W. (2013) DNA strand break repair and neurodegeneration. *Dna Repair*, 12 (8): 558–567
- Sallmyr, A. and Tomkinson, A.E. (2018) Repair of DNA double-strand breaks by mammalian alternative end-joining pathways. *Journal of Biological Chemistry*, 293 (27): 10536–10546
- Sambrook, J. and Russell, D. (2001) *Molecular Cloning: A Laboratory Manual*. Vol. 3. Cold Springs Laboratory Press
- Santos-Pereira, J.M. and Aguilera, A. (2015) R loops: new modulators of genome dynamics and function. *Nature reviews. Genetics*, 16 (10): 583–597
- Sarbajna, S. and West, S.C. (2014) Holliday junction processing enzymes as guardians of genome stability. *Trends in biochemical sciences*, 39 (9): 409–419
- Saredi, G., Huang, H., Hammond, C.M., et al. (2016) H4K20me0 marks post-replicative chromatin and recruits the TONSL–MMS22L DNA repair complex. *Nature*, 534 (7609): 714–718
- Savitsky, K., Bar-Shira, A., Gilad, S., et al. (1995) A Single Ataxia-Telangiectasia Gene with a Product Similar to Pi-3 Kinase. *Science (New York, N.Y.)*, 268 (5218): 1749–1753
- Schellenberg, M.J., Appel, C.D., Adhikari, S., et al. (2012) Mechanism of repair of 5'-topoisomerase II-DNA adducts by mammalian tyrosyl-DNA phosphodiesterase 2. *Nature structural & molecular biology*, 19 (12): 1363–1371
- Schellenberg, M.J., Lieberman, J.A., Herrero-Ruiz, A., et al. (2017) ZATT (ZNF451)-mediated resolution of topoisomerase 2 DNA-protein cross-links. *Science (New York, N.Y.)*, 357 (6358): 1412–1416
- Schellenberg, M.J., Perera, L., Strom, C.N., et al. (2016) Reversal of DNA damage induced Topoisomerase 2 DNA-protein crosslinks by Tdp2. *Nucleic acids research*, 44 (8): 3829–3844
- Schimmel, J., Kool, H., van Schendel, R., et al. (2017) Mutational signatures of non-homologous and polymerase theta-mediated end-joining in embryonic stem cells. *Embo Journal*, 36 (24): 3634–3649

Schwer, B., Wei, P.-C., Chang, A.N., et al. (2016) Transcription-associated processes cause DNA double-strand breaks and translocations in neural stem/progenitor cells. *Proceedings of the National Academy of Sciences of the United States of America*, 113 (8): 2258–2263

Sfeir, A. and Symington, L.S. (2015) Microhomology-Mediated End Joining: A Back-up Survival Mechanism or Dedicated Pathway? *Trends in biochemical sciences*, 40 (11): 701–714

Shamanna, R.A., Lu, H., de Freitas, J.K., et al. (2016) WRN regulates pathway choice between classical and alternative non-homologous end joining. *Nature communications*, 7: 13785

Sharma, S.V. and Settleman, J. (2007) Oncogene addiction: setting the stage for molecularly targeted cancer therapy. *Genes & development*, 21 (24): 3214–3231

Shen, J., Gilmore, E.C., Marshall, C.A., et al. (2010) Mutations in PNKP cause microcephaly, seizures and defects in DNA repair. *Nature genetics*, 42 (3): 245–249

Shibata, A., Conrad, S., Birraux, J., et al. (2011) Factors determining DNA double-strand break repair pathway choice in G2 phase. *Embo Journal*, 30 (6): 1079–1092

Shiloh, Y. and Ziv, Y. (2013) The ATM protein kinase: regulating the cellular response to genotoxic stress, and more. *Nature reviews. Molecular cell biology*, 14 (4): 197–210

Sibanda, B.L., Chirgadze, D.Y., Ascher, D.B., et al. (2017) DNA-PKcs structure suggests an allosteric mechanism modulating DNA double-strand break repair. *Science (New York, N.Y.)*, 355 (6324): 520–524

Sievers, F., Wilm, A., Dineen, D., et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7 (1): 539–539

Silverman, J., Takai, H., Buonomo, S.B.C., et al. (2004) Human Rif1, ortholog of a yeast telomeric protein, is regulated by ATM and 53BP1 and functions in the S-phase checkpoint. *Genes & development*, 18 (17): 2108–2119

Singleton, B.K., Torres-Arzayus, M.I., Rottinghaus, S.T., et al. (1999) The C terminus of Ku80 activates the DNA-dependent protein kinase catalytic subunit. *Molecular and Cellular Biology*, 19 (5): 3267–3277

Spivak, G. (2015) Nucleotide excision repair in humans. *Dna Repair*, 36: 13–18

Spycher, C., Miller, E.S., Townsend, K., et al. (2008) Constitutive phosphorylation of MDC1 physically links the MRE11-RAD50-NBS1 complex to damaged chromatin. *The Journal of cell biology*, 181 (2): 227–240

Steggerda, S.M. and Paschal, B.M. (2002) Regulation of nuclear import and export by the GTPase Ran. *International review of cytology*, 217: 41–91

- Stewart, G.S., Panier, S., Townsend, K., et al. (2009) The RIDDLE syndrome protein mediates a ubiquitin-dependent signaling cascade at sites of DNA damage. *Cell*, 136 (3): 420–434
- Stewart, G.S., Stankovic, T., Byrd, P.J., et al. (2007) RIDDLE immunodeficiency syndrome is linked to defects in 53BP1-mediated DNA damage signaling. *Proceedings of the National Academy of Sciences of the United States of America*, 104 (43): 16910–16915
- Stoll, G., Pietiläinen, O.P.H., Linder, B., et al. (2013) Deletion of TOP3 $\beta$ , a component of FMRP-containing mRNPs, contributes to neurodevelopmental disorders. *Nature neuroscience*, 16 (9): 1228–1237
- Stucki, M., Clapperton, J.A., Mohammad, D., et al. (2005) MDC1 directly binds phosphorylated histone H2AX to regulate cellular responses to DNA double-strand breaks. *Cell*, 123 (7): 1213–1226
- Suberbielle, E., Sanchez, P.E., Kravitz, A.V., et al. (2013) Physiologic brain activity causes DNA double-strand breaks in neurons, with exacerbation by amyloid- $\beta$ . *Nature neuroscience*, 16 (5): 613–621
- Super, H.J., McCabe, N.R., Thirman, M.J., et al. (1993) Rearrangements of the MLL gene in therapy-related acute myeloid leukemia in patients previously treated with agents targeting DNA-topoisomerase II. *Blood*, 82 (12): 3705–3711
- Syed, A. and Tainer, J.A. (2018) The MRE11-RAD50-NBS1 Complex Conducts the Orchestration of Damage Signaling and Outcomes to Stress in DNA Replication and Repair. *Annual review of biochemistry*
- Sykora, P., Croteau, D.L., Bohr, V.A., et al. (2011) Aprataxin localizes to mitochondria and preserves mitochondrial function. *Proceedings of the National Academy of Sciences of the United States of America*, 108 (18): 7437–7442
- Symington, L.S. and Gautier, J. (2011) Double-strand break end resection and repair pathway choice. *Annual review of genetics*, 45 (1): 247–271
- Sørli, T., Perou, C.M., Tibshirani, R., et al. (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proceedings of the National Academy of Sciences of the United States of America*, 98 (19): 10869–10874
- Takashima, H., Boerkoel, C.F., John, J., et al. (2002) Mutation of TDP1, encoding a topoisomerase I-dependent DNA damage repair enzyme, in spinocerebellar ataxia with axonal neuropathy. *Nature genetics*, 32 (2): 267–272
- Taylor, A.M.R., Groom, A. and Byrd, P.J. (2004) Ataxia-telangiectasia-like disorder (ATLD)-its clinical presentation and molecular basis. *Dna Repair*, 3 (8-9): 1219–1225
- Taylor, A.M.R., Harnden, D.G., Arlett, C.F., et al. (1975) Ataxia telangiectasia: A human mutation with abnormal radiation sensitivity. *Nature*, 258 (5534): 427–429

- Thierry-Mieg, D. and Thierry-Mieg, J. (2006) AceView: a comprehensive cDNA-supported gene and transcripts annotation. *Genome biology*, 7 Suppl 1 (Suppl 1): S12.1–14
- Thomson, G., Watson, A., Caldecott, K., et al. (2013) Generation of assays and antibodies to facilitate the study of human 5'-tyrosyl DNA phosphodiesterase. *Analytical Biochemistry*, 436 (2): 145–150
- Thorslund, T., Ripplinger, A., Hoffmann, S., et al. (2015) Histone H1 couples initiation and amplification of ubiquitin signalling after DNA damage. *Nature*, 527 (7578): 389–393
- Trotter, K.W., King, H.A. and Archer, T.K. (2015) Glucocorticoid Receptor Transcriptional Activation via the BRG1-Dependent Recruitment of TOP2 $\beta$  and Ku70/86. *Molecular and Cellular Biology*, 35 (16): 2799–2817
- Tubbs, A. and Nussenzweig, A. (2017) Endogenous DNA Damage as a Source of Genomic Instability in Cancer. *Cell*, 168 (4): 644–656
- Tuduri, S., Crabbé, L., Conti, C., et al. (2009) Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. *Nature cell biology*, 11 (11): 1315–1324
- UniProt Consortium, T. (2018) UniProt: the universal protein knowledgebase. *Nucleic acids research*, 46 (5): 2699–2699
- Uusküla-Reimand, L., Hou, H., Samavarchi-Tehrani, P., et al. (2016) Topoisomerase II beta interacts with cohesin and CTCF at topological domain borders. *Genome biology*, 17 (1): 182
- Vaz, B., Popovic, M. and Ramadan, K. (2017) DNA-Protein Crosslink Proteolysis Repair. *Trends in biochemical sciences*, 42 (6): 483–495
- Várady, G., Sarkadi, B. and Fátýol, K. (2011) TTRAP is a novel component of the non-canonical TRAF6-TAK1 TGF- $\beta$  signaling pathway. *Rishi, A. (ed.). PloS one*, 6 (9): e25548
- Veitia, R.A., Bottani, S. and Birchler, J.A. (2013) Gene dosage effects: nonlinearities, genetic interactions, and dosage compensation. *Trends in genetics : TIG*, 29 (7): 385–393
- Virgen-Slane, R., Rozovics, J.M., Fitzgerald, K.D., et al. (2012) An RNA virus hijacks an incognito function of a DNA repair enzyme. *Proceedings of the National Academy of Sciences of the United States of America*, 109 (36): 14634–14639
- Vodák, D., Lorenz, S., Nakken, S., et al. (2018) Sample-Index Misassignment Impacts Tumour Exome Sequencing. *Scientific reports*, 8 (1): 5307
- Vos, S.M., Tretter, E.M., Schmidt, B.H., et al. (2011) All tangled up: how cells direct, manage and exploit topoisomerase function. *Nature reviews. Molecular cell biology*, 12 (12): 827–841

- Wang, B.-Y., Xu, G.-L., Zhou, C.-H., et al. (2010) PhiC31 integrase interacts with TTRAP and inhibits NFkappaB activation. *Molecular biology reports*, 37 (6): 2809–2816
- Wang, J.C. (1971) Interaction between DNA and an Escherichia coli protein  $\omega$ . *Journal of molecular biology*, 55 (3): 523–IN16
- Wang, J.C. (1996) DNA topoisomerases. *Annual review of biochemistry*, 65 (1): 635–692
- Wang, J.C. (2002) Cellular roles of DNA topoisomerases: a molecular perspective. *Nature reviews. Molecular cell biology*, 3 (6): 430–440
- Wang, J.C. and Liu, L.F. (1979) “DNA Topoisomerases: Enzymes That Catalyze the Concerted Breaking and Rejoining of DNA Backbone Bonds.” *In* *Molecular Genetics*. Elsevier. pp. 65–88
- Wang, Y., Lyu, Y.L. and Wang, J.C. (2002) Dual localization of human DNA topoisomerase IIIalpha to mitochondria and nucleus. *Proceedings of the National Academy of Sciences of the United States of America*, 99 (19): 12114–12119
- Watson, J.D. and Crick, F.H. (1953) Genetical implications of the structure of deoxyribonucleic acid. *Nature*, 171 (4361): 964–967
- Wei, P.-C., Chang, A.N., Kao, J., et al. (2016) Long Neural Genes Harbor Recurrent DNA Break Clusters in Neural Stem/Progenitor Cells. *Cell*, 164 (4): 644–655
- Wei, Y., Diao, L.-X., Lu, S., et al. (2017) SUMO-Targeted DNA Translocase Rrp2 Protects the Genome from Top2-Induced DNA Damage. *Molecular Cell*, 66 (5): 581–596.e6
- Wendorff, T.J., Schmidt, B.H., Heslop, P., et al. (2012) The structure of DNA-bound human topoisomerase II alpha: conformational mechanisms for coordinating inter-subunit interactions with DNA cleavage. *Journal of molecular biology*, 424 (3-4): 109–124
- Whisstock, J.C., Romero, S., Gurung, R., et al. (2000) The inositol polyphosphate 5-phosphatases and the apurinic/apyrimidinic base excision repair endonucleases share a common mechanism for catalysis. *Journal of Biological Chemistry*, 275 (47): 37055–37061
- Woessner, R.D., Mattern, M.R., Mirabelli, C.K., et al. (1991) Proliferation- and cell cycle-dependent differences in expression of the 170 kilodalton and 180 kilodalton forms of topoisomerase II in NIH-3T3 cells. *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research*, 2 (4): 209–214
- Woodbine, L., Gennery, A.R. and Jeggo, P.A. (2014) The clinical impact of deficiency in DNA non-homologous end-joining. *Dna Repair*, 16: 84–96
- Woodbine, L., Neal, J.A., Sasi, N.-K., et al. (2013) PRKDC mutations in a SCID patient with profound neurological abnormalities. *The Journal of clinical investigation*, 123 (7): 2969–2980

- Wright, W.D., Shah, S.S. and Heyer, W.-D. (2018) Homologous recombination and the repair of DNA double-strand breaks. *Journal of Biological Chemistry*, 293 (27): 10524–10535
- Wyatt, H.D.M. and West, S.C. (2014) Holliday junction resolvases. *Cold Spring Harbor perspectives in biology*, 6 (9): a023192–a023192
- Wyatt, H.D.M., Laister, R.C., Martin, S.R., et al. (2017) The SMX DNA Repair Tri-nuclease. *Molecular Cell*, 65 (5): 848–860.e11
- Xu, D., Shen, W., Guo, R., et al. (2013) Top3 $\beta$  is an RNA topoisomerase that works with fragile X syndrome protein to promote synapse formation. *Nature neuroscience*, 16 (9): 1238–1247
- Yang, S.W., Burgin, A.B., Huizenga, B.N., et al. (1996) A eukaryotic enzyme that can disjoin dead-end covalent complexes between DNA and type I topoisomerases. *Proceedings of the National Academy of Sciences of the United States of America*, 93 (21): 11534–11539
- Yang, X., Li, W., Prescott, E.D., et al. (2000) DNA topoisomerase II $\beta$  and neural development. *Science (New York, N.Y.)*, 287 (5450): 131–134
- Yano, K.-I., Morotomi-Yano, K., Wang, S.-Y., et al. (2008) Ku recruits XLF to DNA double-strand breaks. *EMBO reports*, 9 (1): 91–96
- Yoon, G. and Caldecott, K.W. (2018) Nonsyndromic cerebellar ataxias associated with disorders of DNA single-strand break repair. *Handbook of clinical neurology*, 155: 105–115
- Yu, M. and Ren, B. (2017) The Three-Dimensional Organization of Mammalian Genomes. *Annual review of cell and developmental biology*, 33 (1): 265–289
- Zagnoli-Vieira, G. and Caldecott, K.W. (2017) TDP2, TOP2, and SUMO: what is ZATT about? *Cell research*, 27 (12): 1405–1406
- Zagnoli-Vieira, G., Bruni, F., Thompson, K., et al. (2018) Confirming TDP2 mutation in spinocerebellar ataxia autosomal recessive 23 (SCAR23). *Neurology Genetics*, 4 (4): e262
- Zeng, Z., Cortes Ledesma, F., El-Khamisy, S.F., et al. (2011) TDP2/TTRAP is the major 5'-tyrosyl DNA phosphodiesterase activity in vertebrate cells and is critical for cellular resistance to topoisomerase II-induced DNA damage. *Journal of Biological Chemistry*, 286 (1): 403–409
- Zeng, Z., Sharma, A., Ju, L., et al. (2012) TDP2 promotes repair of topoisomerase I-mediated DNA damage in the absence of TDP1. *Nucleic acids research*, 40 (17): 8371–8380
- Zerbino, D.R., Achuthan, P., Akanni, W., et al. (2018) Ensembl 2018. *Nucleic acids research*, 46 (D1): D754–D761

Zhang, A., Lyu, Y.L., Lin, C.-P., et al. (2006) A protease pathway for the repair of topoisomerase II-DNA covalent complexes. *Journal of Biological Chemistry*, 281 (47): 35997–36003

Zhang, J.-Q., Wang, J.-J., Li, W.-J., et al. (2009) Cellular protein TTRAP interacts with HIV-1 integrase to facilitate viral integration. *Biochemical and Biophysical Research Communications*, 387 (2): 256–260

Zhou, C., Shen, Q., Xue, J., et al. (2013) Overexpression of TTRAP inhibits cell growth and induces apoptosis in osteosarcoma cells. *BMB reports*, 46 (2): 113–118

Zou, L. and Elledge, S.J. (2003) Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science (New York, N.Y.)*, 300 (5625): 1542–1548

## 9. APPENDIX

### 9.1. Published work

Gómez-Herreros, F., **Zagnoli-Vieira, G.**, Ntai, I., Martínez-Macías, M. I., Anderson, R. M., Herrero-Ruíz, A. and Caldecott, K. W. (2017) 'TDP2 suppresses chromosomal translocations induced by DNA topoisomerase II during gene transcription.', *Nature communications*, 8(1), p. 233. doi: 10.1038/s41467-017-00307-y

**Zagnoli-Vieira, G.** and Caldecott, K. W. (2017) 'TDP2, TOP2, and SUMO: what is ZATT about?', *Cell research*, 27(12), pp. 1405–1406. doi: 10.1038/cr.2017.147.

**Zagnoli-Vieira, G.**, Bruni, F., Thompson, K., He, L., Walker, S., de Brouwer, A. P. M., Taylor, R., Niyazov, D. and Caldecott, K. W. (2018) 'Confirming TDP2 mutation in spinocerebellar ataxia autosomal recessive 23 (SCAR23)', *Neurology Genetics*. Wolters Kluwer Health, Inc. on behalf of the American Academy of Neurology, 4(4), p. e262. doi: 10.1212/NXG.0000000000000262.





## ARTICLE

DOI: 10.1038/s41467-017-00307-y

OPEN

# TDP2 suppresses chromosomal translocations induced by DNA topoisomerase II during gene transcription

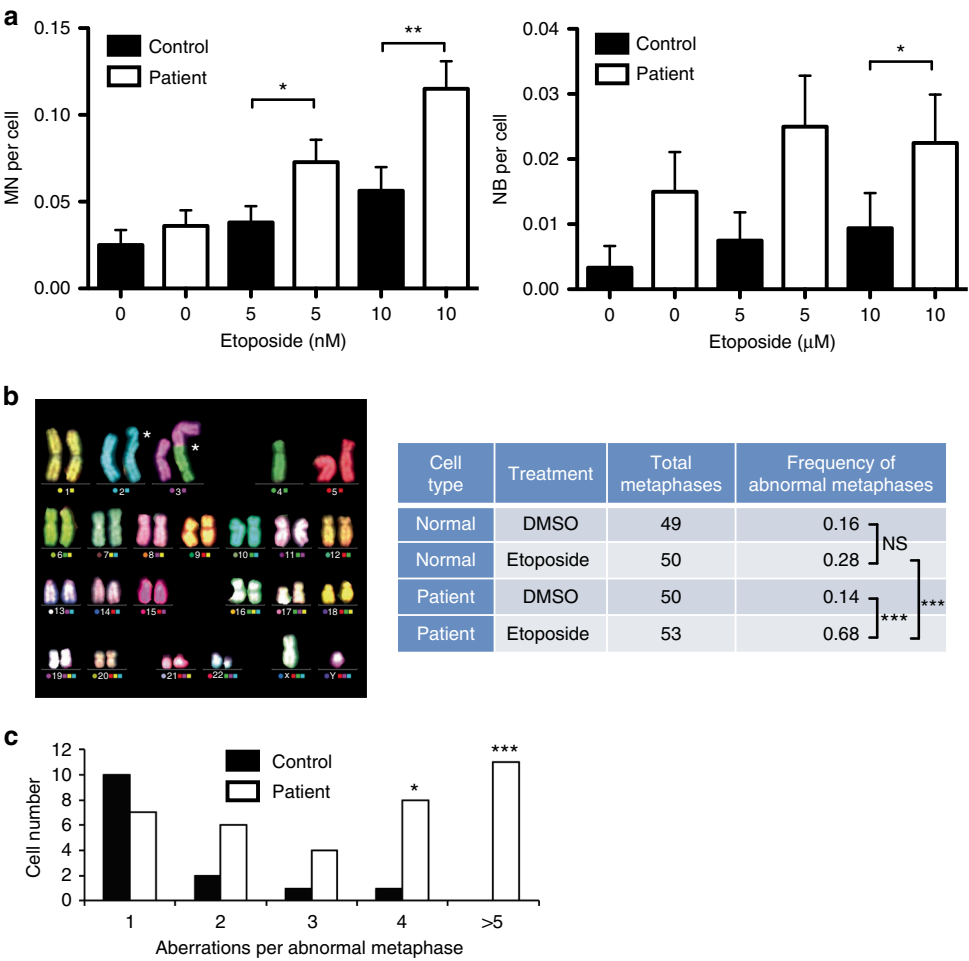
Fernando Gómez-Herreros<sup>1,2,3</sup>, Guido Zagnoli-Vieira<sup>1</sup> , Ioanna Ntai<sup>1</sup> , María Isabel Martínez-Macías<sup>1</sup>, Rhona M. Anderson<sup>4</sup> , Andrés Herrero-Ruiz<sup>1,5</sup> & Keith W. Caldecott<sup>1</sup>

DNA double-strand breaks (DSBs) induced by abortive topoisomerase II (TOP2) activity are a potential source of genome instability and chromosome translocation. TOP2-induced DNA double-strand breaks are rejoined in part by tyrosyl-DNA phosphodiesterase 2 (TDP2)-dependent non-homologous end-joining (NHEJ), but whether this process suppresses or promotes TOP2-induced translocations is unclear. Here, we show that TDP2 rejoins DSBs induced during transcription-dependent TOP2 activity in breast cancer cells and at the translocation 'hotspot', *MLL*. Moreover, we find that TDP2 suppresses chromosome rearrangements induced by TOP2 and reduces TOP2-induced chromosome translocations that arise during gene transcription. Interestingly, however, we implicate TDP2-dependent NHEJ in the formation of a rare subclass of translocations associated previously with therapy-related leukemia and characterized by junction sequences with 4-bp of perfect homology. Collectively, these data highlight the threat posed by TOP2-induced DSBs during transcription and demonstrate the importance of TDP2-dependent non-homologous end-joining in protecting both gene transcription and genome stability.

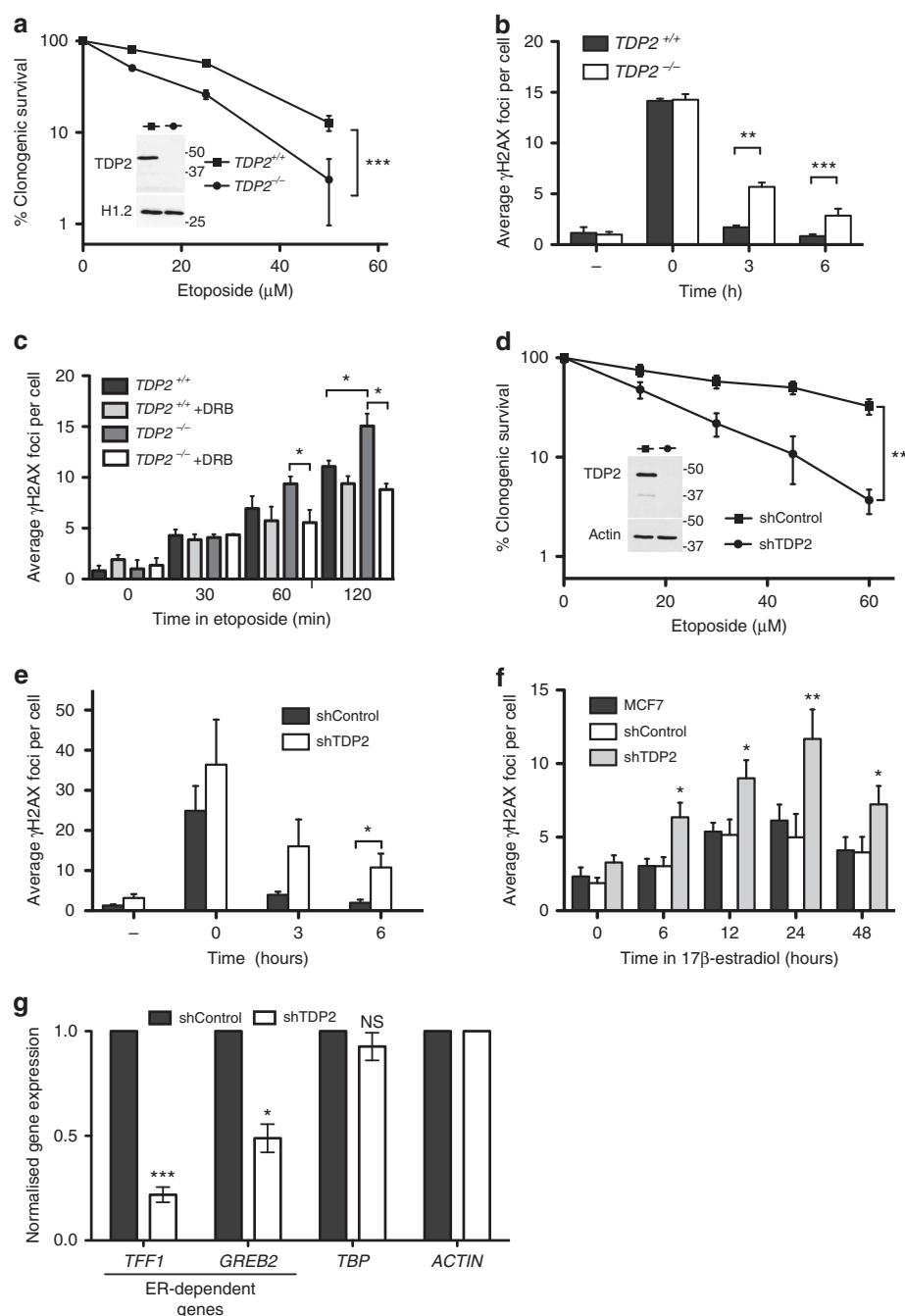
<sup>1</sup>Genome Damage and Stability Centre, School of Life Sciences, University of Sussex, Falmer, Brighton BN1 9RQ, UK. <sup>2</sup>Instituto de Biomedicina de Sevilla (IBiS), Hospital Virgen del Rocío-CSIC-Universidad de Sevilla, Seville 41013, Spain. <sup>3</sup>Departamento de Genética, Universidad de Sevilla, Seville 41080, Spain. <sup>4</sup>Institute of Environment, Health and Societies, Brunel University London, Uxbridge UB8 3PH, UK. <sup>5</sup>Present address: Centro Andaluz de Biología Molecular y Medicina Regenerativa (CABIMER), CSIC-Universidad de Sevilla, Seville 41092, Spain. Correspondence and requests for materials should be addressed to F.G.-H. (email: [fgomezhs@us.es](mailto:fgomezhs@us.es)) or to K.W.C. (email: [k.w.caldecott@sussex.ac.uk](mailto:k.w.caldecott@sussex.ac.uk))

**T**umorigenesis is a multistage process that involves the accumulation of genetic changes. Cytogenetic aberrations are common features of neoplasia and have a key impact in cancer development. Whereas gross chromosomal abnormalities such as duplications, deletions, and chromosome loss or gain are commonly involved in late stages of tumor development, reciprocal translocations represent single events that contribute to initial stages of cellular disarray. The molecular mechanisms that underlie the formation of reciprocal translocations are of great interest, because they can result chimeric proteins and/or deregulated transcription programmes that can drive oncogenesis in a variety of solid tumors and leukemias<sup>1–3</sup>. A number of reciprocal translocations involve partner genes that are co-located in a common transcription factory such as *IgH* and *myc* in Burkitt’s lymphoma<sup>4</sup>, *TPR22* and *ERG* in prostate cancer<sup>5</sup>, and mixed lineage leukemia (*MLL*), *AF4* and *AF9* in leukemia<sup>6,7</sup>. Consistent with the idea that spatial proximity is an important factor in reciprocal translocation formation<sup>8</sup>, translocations arise from DNA double-strand breaks (DSBs) arising in nearby genes<sup>2</sup>. Nonetheless, the molecular mechanisms that

underlie DSB induction and gene fusion during the translocation of actively transcribed genes are unclear and controversial. Indeed, some studies suggest that DSBs that drive translocations are obligate and/or are ‘programmed’ intermediates of some transcription events<sup>9–12</sup>. One class of enzymes implicated in DSB induction during gene expression are DNA topoisomerases. Topoisomerases, are enzymes that remove torsional stress in DNA by introducing transient DNA breaks and are involved in a wide variety of nuclear processes including transcription and DNA replication<sup>13</sup>. Type II topoisomerases (e.g. TOP2) pass one DNA duplex through another via a DSB intermediate known as the ‘cleavage complex’; within which the topoisomerase has cleaved both strands of DNA and is covalently linked to the 5’-terminus of the DSB via a phosphotyrosyl bond<sup>14</sup>. The cleavage complex is normally transient, because the topoisomerase reveals the break at the end of its catalytic cycle. However, cleavage complexes are the targets of a class of anti-tumor agents such as etoposide that act as topoisomerase ‘poisons’, resulting in the formation of ‘abortive’ cleavage complexes that require removal by DSB repair.



**Fig. 1** TDP2 suppresses TOP2-induced genome instability in human cells. **a** Micronuclei (MN, left) and nucleoplasmic bridges (NB, right) were quantified in control and TDP2 patient lymphoblastoid cells following continuous treatment (24 h) with the indicated concentration of etoposide. Data are the average ( $\pm$ s.e.m.) of a minimum of 300 cells from three independent experiments. Statistical significance was determined by *T*-test (\* $P$  < 0.05, \*\* $P$  < 0.01). **b** M-FISH karyotyping of control and TDP2 patient lymphoblastoid cells following continuous treatment with 50 nM etoposide for 20 h. Left, example of a patient metaphase depicting a chromatid break on Chr2 and a dicentric chromosome involving Chr3 and Chr4 (white asterisks). Right, quantification of abnormal metaphases by M-FISH in control and TDP2 patient lymphoblastoid cells. Statistical significance was determined by GLM ANOVA (\*\*\* $P$  < 0.005, NS, not significant). **c** Distribution of chromosome aberrations per abnormal metaphase. The total number of events (chromosome and chromatid exchanges/breaks) were counted in each abnormal metaphase (5 or more events per cell were pooled as > 5). Statistical significance was determined by ANOVA (\* $P$  < 0.05, \*\*\* $P$  < 0.005)



**Fig. 2** TDP2 protects gene transcription from TOP2-induced DSBs. **a** Clonogenic survival of wild-type and *TDP2*<sup>-/-</sup> RPE-1 cells following treatment (3 h) with the indicated concentrations of etoposide. Data are the mean (±s.e.m.) of three independent experiments. Statistical significance was determined by two-way ANOVA (\**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.005). Molecular weight markers are in kDa. **b** The number of γH2AX foci in wild-type and *TDP2*<sup>-/-</sup> RPE-1 cells before and 30 min after treatment with 20 μM etoposide, and after the indicated repair periods in drug-free medium. Data are the mean (±s.e.m.) of at least three independent experiments. Statistical significance was determined by *T*-tests (\**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.005). **c** The number of γH2AX foci in serum-starved wild-type and *TDP2*<sup>-/-</sup> RPE-1 cells was measured following treatment with 5 μM etoposide. Where indicated, cells were pre-incubated with the transcription inhibitor DRB (100 μM) for 3 h prior to etoposide treatment. Other details are as in **b**. **d** Clonogenic survival of mock-depleted (shControl) or TDP2-depleted (shTDP2) MCF7 cells following treatment (3 h) with the indicated concentrations of etoposide. Other details are as in **a**. **e** The number of γH2AX foci in mock-depleted (shControl) or TDP2-depleted (shTDP2) MCF7 cells before and 30 min after treatment with 20 μM etoposide, and after the indicated repair periods in drug-free medium. Other details as in **b**. **f** The number of γH2AX foci in untransfected MCF7 cells or in mock-depleted (shControl) or TDP2-depleted (shTDP2) MCF7 cells following incubation for the indicated times with 100 nM 17β-estradiol. Data are the mean (±s.e.m.) of five independent experiments. Statistical significance was determined by *T*-tests (\**P* < 0.05, \*\**P* < 0.01) and are comparisons with shControl cells. **g** Basal mRNA levels of two ER-regulated genes (*TFF1* & *GREB2*) and one ER-independent gene (*TBP*) in mock-depleted (shControl) or TDP2-depleted (shTDP2) MCF7 cells. mRNA levels were quantified by qRT-PCR and normalized relative to *ACTB* levels under the same experimental conditions. The normalized value from TDP2 patient cells was then expressed relative to the normalized value from the mock-depleted control cells. Data are the mean (±s.e.m.) of three independent experiments. Statistical significance was determined by *T*-tests (\**P* < 0.05, \*\*\**P* < 0.005, NS, not significant).

It is likely that cleavage complexes become abortive at some frequency even in the absence of TOP2 poisons. Indeed, TOP2-induced DSBs have been implicated in oncogenic translocation events not only in response to exposure to chemotherapeutic TOP2 poisons, but also in the formation of de novo translocations<sup>5, 10, 15</sup>. Consistent with this, cells possess DNA repair enzymes with activities targeted specifically at removing trapped topoisomerase activity from DNA 5' termini, such as tyrosyl-DNA phosphodiesterase 2 (TDP2)<sup>16–18</sup>. TDP2 cleaves the phosphotyrosyl bond between the topoisomerase and the 5' phosphate of the DNA generating ligatable ends that can be processed by the non-homologous end-joining (NHEJ) pathway<sup>18, 19</sup>. Here, we have addressed directly the impact of TDP2-dependent repair of TOP2-induced DSBs on genome stability and in chromosome translocations in human cells. We show that TDP2 rejoins DSBs induced during transcription-dependent TOP2 activity in breast cancer cells and at the translocation 'hotspot', *MLL*. Moreover, we find that TDP2 suppresses chromosome rearrangements induced by TOP2 and reduces TOP2-induced chromosome translocations that arise during gene transcription. Collectively, these data highlight the threat posed by TOP2 in genome instability during gene transcription and demonstrate the importance of TDP2-dependent NHEJ for suppressing this instability.

## Results

### Human TDP2 suppresses TOP2-induced genome instability.

We previously reported that TDP2-dependent NHEJ is required for genome stability in murine cells<sup>18</sup>. To determine whether TDP2 is required to maintain genomic stability in human cells we employed lymphoblastoid cells from a patient in which TDP2 is mutated<sup>11</sup>. These cells lack detectable TDP2 activity, are deficient in the repair of TOP2 breaks, and are hypersensitive to TOP2 poisons<sup>11</sup>. TDP2 patient cells exhibited a significantly greater increase in micronuclei and nucleoplasmic bridges than did normal control lymphoblastoid cell lines (LCLs) following treatment with low doses of etoposide, suggesting that genome instability is increased in the absence of TDP2 at sites of TOP2-induced DSBs (Fig. 1a). To measure the impact of TDP2 on chromosome stability directly, we conducted M-FISH of fixed chromosome preparations (Fig. 1b). Although similar numbers of abnormal metaphases were observed in control and TDP2 patient cells in the absence of etoposide treatment, treatment with an etoposide concentration that only slightly increased abnormal metaphases in control cells increased abnormal metaphases in TDP2 patient cells ~fivefold; from 0.14 per metaphase to 0.68 per metaphase. Moreover, TDP2 patient cells harbored on average ~10-fold more chromosome aberrations per cell than did control cells following etoposide treatment (Supplementary Fig. 1; 3.96 and 0.4 aberrations per cell, respectively). These were comprised of both chromosome and chromatid-type events but the latter were far more prevalent; increasing ~30-fold in etoposide-treated TDP2 patient cells from 0.04 to 3.53 aberrations per cell, compared with little or no increase in normal cells (Supplementary Fig. 1). Notably, patient cell populations exhibited a particularly high frequency of metaphases with multiple aberrations, including complex events involving multiple chromosomes (Fig. 1c). Collectively, these results demonstrate that TDP2 is required to maintain chromosome stability in human cells following the induction of DSBs by abortive TOP2 activity.

### TDP2 protects gene transcription from TOP2-induced DSBs.

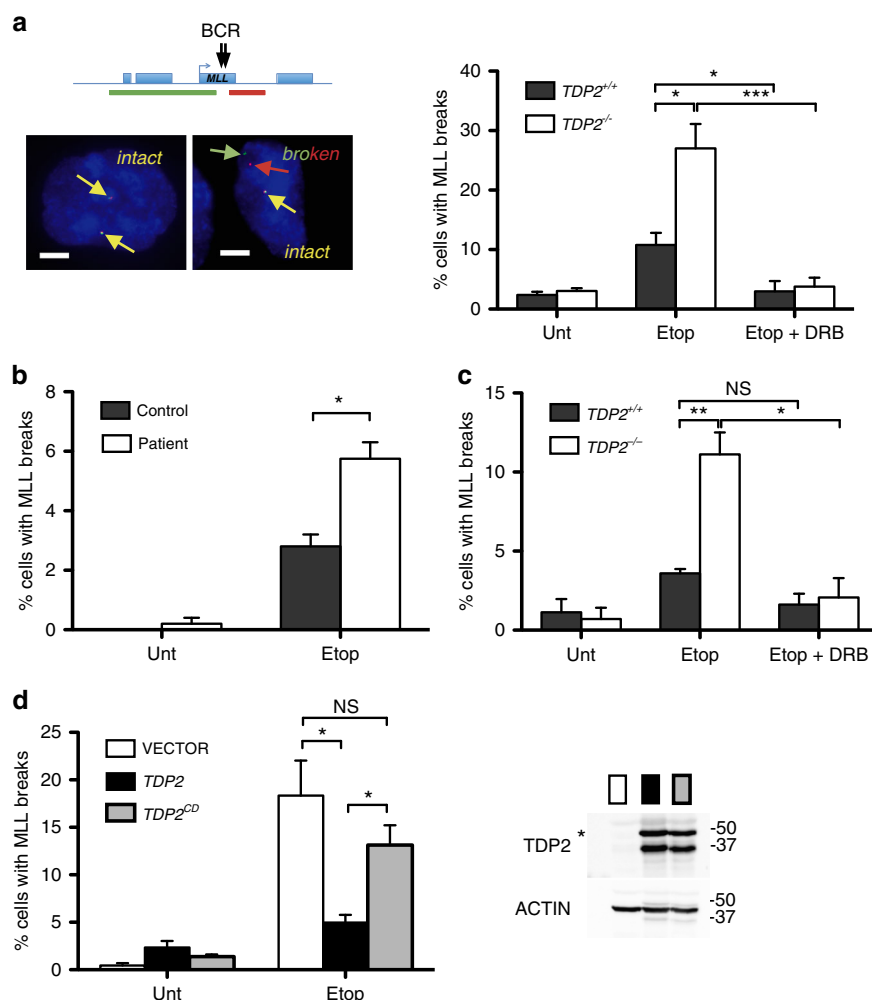
Although TDP2 can promote repair of TOP2-induced DSBs arising during S phase, we have proposed that the error-free

nature of this pathway may have particular utility in non-cycling cells that lack alternative error-free pathways such as homologous recombination repair<sup>18</sup>. Consequently, TDP2 may have particular utility in removing TOP2-induced DSBs arising during gene transcription, which can arise independently of cell cycle status. To address this, we employed diploid human RPE-1 hTERT cells in which we disrupted *TDP2* gene using CRISPR-Cas9. *TDP2*<sup>-/-</sup> RPE-1 cells exhibited hypersensitivity to etoposide and a reduced rate of repair of etoposide-induced DSBs, as measured using  $\gamma$ H2AX immunofoci as a marker of DSBs<sup>20</sup> (Fig. 2a, b). Importantly, in quiescent cells, the elevated accumulation of etoposide-induced DSBs in *TDP2*<sup>-/-</sup> cells was prevented by pre-incubation with the RNA polymerase II transcription inhibitor DRB, suggesting that TDP2 is important for the repair of TOP2-induced DSBs arising during transcription (Fig. 2c).

TOP2 activity has been implicated previously in DSB induction during hormone-regulated gene transcription<sup>9, 10</sup>. Consistent with this, we showed previously that TDP2 is required to maintain normal levels of TOP2-dependent gene transcription stimulated by androgen<sup>11</sup>. To further address the role of TDP2 in the repair of DSBs induced during transcription we depleted TDP2 in the ER-positive breast cancer cell line MCF7 using anti-TDP2 shRNA. TDP2-depleted MCF7 cells exhibited reduced DSB repair following treatment with etoposide and correspondingly an increase in hypersensitivity to this drug (Fig. 2d, e). Interestingly, the addition of estradiol alone resulted in the accumulation of  $\gamma$ H2AX foci in MCF7 cells, and did so to a greater extent in MCF7 cells depleted of TDP2 (Fig. 2f). More importantly, both the basal level and estradiol-stimulated level of transcription of *TFF1* and *GREB1*, two genes known to be regulated by estrogen, was reduced in TDP2-depleted MCF7 cells when compared to wild-type MCF7 cells (Fig. 2g and Supplementary Fig. 2). Collectively, these results implicate TDP2 in the removal of TOP2-induced DSBs arising during gene transcription and in the protection of transcription from inhibition by TOP2-induced DSBs.

**TDP2 protects *MLL* from TOP2-induced breakage.** Next, we investigated the impact of TDP2 on DNA breaks occurring at the *MLL* chromosome locus, which are also associated with TOP2 activity. Translocations involving the breakpoint cluster region (BCR) of *MLL* gene are amongst the most common chromosome rearrangements observed in leukaemia<sup>7</sup>. Rearrangements of the *MLL* gene can involve one of a large number of partners and are implicated in both therapy-related acute leukemia and infant acute leukaemia<sup>21</sup>. Numerous potential mechanisms for *MLL* translocation have been proposed, including nucleases involved in VDJ recombination, Alu-mediated recombination, apoptotic nucleases, and TOP2-induced breaks<sup>7, 15</sup>. To examine whether TDP2 is required to protect *MLL* from TOP2-induced breakage, we employed FISH probes flanking this locus so that we could detect etoposide-induced breaks via their impact on the spatial proximity of these probes (Fig. 3a).

The addition of etoposide increased to ~10% the fraction of wild-type RPE-1 cells harboring a visible break at one of the two *MLL* loci (Fig. 3a). More importantly, this fraction increased to ~30% in *TDP2*<sup>-/-</sup> RPE-1 cells, implicating TDP2 in the repair of TOP2-induced DSBs at *MLL*. Similar results were observed in lymphoblastoid cells, in which etoposide-induced more *MLL* breaks in TDP2 patient cells than in control cells (Fig. 3b). The breaks detected in these experiments were not apoptotic breaks<sup>22</sup>, because we did not detect any evidence for significant apoptosis under the conditions employed (Supplementary Fig. 3). Importantly, the presence of elevated etoposide-induced breaks at *MLL*



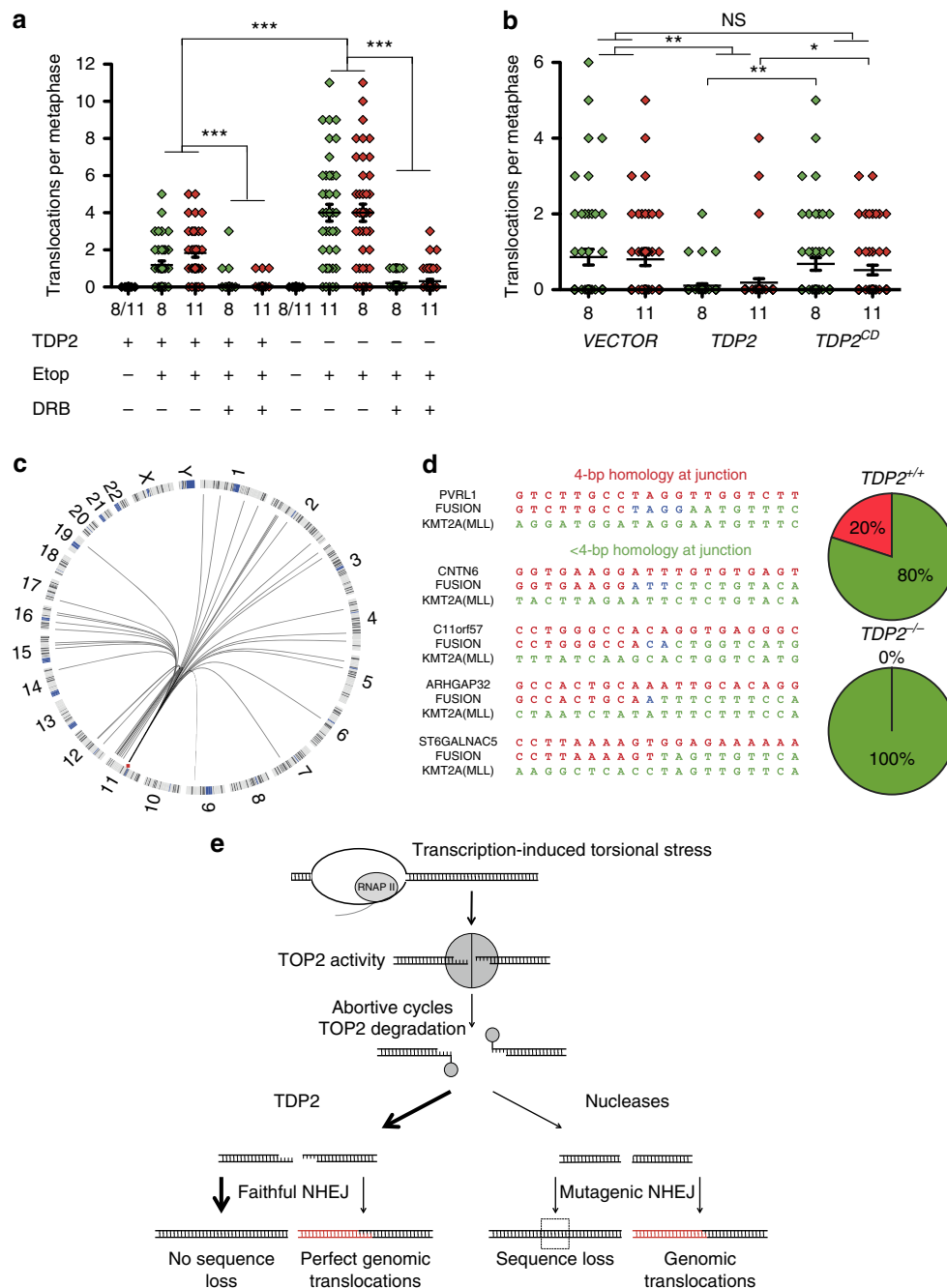
**Fig. 3** TDP2 is required for the repair of TOP2-induced DSBs at *MLL*. **a** Detection of broken *MLL* loci by FISH in wild-type and *TDP2*<sup>-/-</sup> RPE-1 cells following mock-treatment or treatment with 100  $\mu$ M etoposide for 6 h. Where indicated, cells were pre-incubated with DRB (100  $\mu$ M) for 3 h prior to etoposide treatment. *Left*, representative images of intact (yellow) and broken (red and green) *MLL* loci (Scale bar 5  $\mu$ m). The position of the FISH probes (red and green lines in top cartoon) flanking the *MLL* breakpoint cluster are shown. *Right*, data are the mean ( $\pm$  s.e.m.) of three independent experiments. Statistical significance was measured by *T*-tests (\**P* < 0.05, \*\**P* < 0.005, \*\*\**P* < 0.001). **b** Detection of broken *MLL* loci in normal and TDP2 patient lymphoblastoid cells, measured as above. Data are the mean ( $\pm$  s.e.m.) of two independent experiments. **c** Detection of broken *MLL* loci by FISH in serum-starved wild-type and *TDP2*<sup>-/-</sup> RPE-1 cells. Data as in **a**. **d** Detection of broken *MLL* loci by FISH in *TDP2*<sup>-/-</sup> RPE-1 cells infected with either empty lentiviral vector or vector expressing either wild-type TDP2 (*TDP2*) or catalytic mutant TDP2<sup>D262A</sup> (*TDP2*<sup>CD</sup>). Data as in **a**. See inset for levels of TDP2 protein (full-length TDP2 is indicated by an asterisk). Molecular weight markers in kDa

in the absence of TDP2 was also observed in quiescent cells, consistent with the source of these breaks being TOP2 activity during transcription rather than DNA replication (Fig. 3c). Indeed, consistent with this, pre-incubation with the transcription inhibitor DRB greatly reduced or ablated the *MLL* breaks detected by FISH in both wild-type and *TDP2*<sup>-/-</sup> RPE-1 cells (Fig. 3a, c). The impact of DRB did not reflect an influence on cell cycle distribution<sup>23</sup>, because we did not observe any impact of DRB in this respect (Supplementary Fig. 3). Finally, to confirm that the role of TDP2 in preventing *MLL* breakage required its catalytic activity we compared recombinant wild-type TDP2 and catalytic mutant TDP2<sup>D262A</sup><sup>16, 17</sup> for ability to complement this defect in *TDP2*<sup>-/-</sup> RPE-1 cells. Indeed, whereas wild-type TDP2 significantly reduced TOP2-induced *MLL* breakage, mutant TDP2 did not (Fig. 3d).

Collectively, these data indicate a role for TDP2 activity in the repair of TOP2-induced DSBs generated during gene transcription at the breakpoint cluster region of *MLL*; a chromosome locus that is a hotspot for oncogenic translocations.

### TDP2 suppresses TOP2-induced chromosome translocations.

To examine directly the impact of TDP2 on the frequency of TOP2-induced chromosome translocations we again employed FISH. Wild-type and *TDP2*<sup>-/-</sup> RPE-1 cells were maintained in G0/G1 during and following treatment with etoposide, to avoid the impact of TOP2-induced DSBs arising during DNA replication, and then released into cell cycle to enable the detection of translocations involving chromosome 8 or 11 by whole-chromosome FISH (Supplementary Fig. 4). Translocations were observed in both cell lines following etoposide treatment but were ~4 times more frequent in *TDP2*<sup>-/-</sup> RPE-1 cells, indicating that TDP2 suppresses the formation of chromosome translocations induced by TOP2 (Fig. 4a). Importantly, the translocations associated with loss of TDP2 were prevented by pre-incubation with DRB, implicating gene transcription by RNA polymerase II as the source of the TOP2-induced DSBs that resulted in translocation (Fig. 4a). Notably, expression of wild-type recombinant TDP2 but not the catalytic mutant TDP2 reduced the frequency of translocations in *TDP2*<sup>-/-</sup> cells (Fig. 4b). These results indicate



**Fig. 4** TDP2 suppresses TOP2-induced chromosome translocations. **a** Translocation frequencies (translocations/metaphase) in chromosomes 8 and 11 were quantified in serum-starved (G1/G0) cells in metaphase spreads prepared 48 hr after mock-treatment or etoposide treatment (1 h, 100  $\mu$ M). In the absence of etoposide, translocations for both chromosomes are plotted together (8/11). Where indicated, cells were pre-treated with DRB (100  $\mu$ M) for 3 h prior to, during, and 4 h after etoposide treatment to inhibit transcription. Data are from two independent experiments and statistical significance was measured by *T*-tests ( $***P < 0.001$ ). **b** Translocation frequencies in etoposide-treated TDP2<sup>-/-</sup> RPE-1 infected with empty lentiviral vector or with lentiviral vector encoding wild-type TDP2 or catalytic mutant TDP2<sup>D262A</sup> (TDP2<sup>CD</sup>). Data are from two independent experiments and statistical significance was determined by *T*-tests ( $*P < 0.05$ ,  $**P < 0.005$ , NS, not significant). **c** Circos plot showing the genomic distribution of MLL translocation partners in etoposide-treated wild-type RPE-1 cells. The location of MLL in chromosome 11 is indicated (red). Resolution; cytogenetic bands. **d** Left, representative junction sequences of etoposide-induced MLL translocations with 4-bp (top) or <4-bp regions of junction homology (bottom three). Right, distribution of translocation junctions with 4-bp or <4-bp regions of homology in WT and TDP2<sup>-/-</sup> RPE-1 cells. **e** Model depicting the influence of TDP2-dependent and independent NHEJ on translocations at TOP2-induced DSBs during transcription. Left, TDP2 primarily promotes rejoining of the correct (adjacent) termini (left, thick arrow) but occasionally may promote joining of incorrect (distal) termini with the same TOP2 cleavage sequence (left, thin arrow). We propose that the latter events are rare but result in a translocation with 4-bp of junction homology. Right, in the absence of TDP2, processing of 5'-phosphotyrosine termini requires nuclease activity and consequently error-prone NHEJ resulting in microdeletions (if the correct adjacent termini are rejoined) or translocations (if incorrect distal termini are joined) with  $\neq$  4-bp of junction homology



that TDP2 activity suppresses TOP2-induced translocations during gene transcription.

To analyze the nature and sequence of the translocations induced by TOP2 in the presence and absence of TDP2 we employed inverse PCR to amplify the breakpoint cluster region (BCR) of *MLL* (Supplementary Fig. 5). Etoposide treatment increased the appearance of PCR products that were different in size from that expected for wild-type product, suggesting that we had successfully isolated translocation breakpoints (Supplementary Fig. 5). Of a total of forty-three different *MLL* translocation breakpoints that we isolated, 84% involved partner loci that were located in actively transcribed regions, consistent with transcription stimulating the *MLL* translocations (Supplementary Table 2). Interestingly, 26% of the *MLL* translocation partners were present in the same or adjacent cytogenetic chromosome bands as *MLL* (Fig. 4c). This is in agreement with the notion that physical proximity greatly influences the choice of translocation partner<sup>3</sup>. We failed to detect a significant variation between wild-type and *TDP2*<sup>-/-</sup> cells in the distribution of *MLL* partners, suggesting that TDP2 does not influence the site of TOP2-mediated cleavage (Fig. 4c and Supplementary Table 2). Intriguingly, however, Sanger sequencing of the translocation junctions suggested that whereas some translocation events in wild-type cells involved 4 bp of identical sequence at the breakpoint in both *MLL* and the translocation partner, these events were absent in *TDP2*<sup>-/-</sup> cells (Fig. 4d). Thus, while TDP2 overall greatly suppresses chromosome translocations at sites of TOP2 cleavage, it might be responsible for a small subset of translocations that are characterized by 4-bp of perfect homology at the translocation junction.

## Discussion

In this study we have addressed the impact of TDP2-dependent DSB repair on chromosome rearrangements mediated by the abortive activity of TOP2. M-FISH revealed that lymphoblastoid cells from a TDP2-defective patient incur elevated chromosome aberrations following the induction of TOP2-induced DNA breaks by etoposide. These results are in agreement with previous cytogenetic analyses of lymphoid tissue in etoposide-treated *Tdp2*<sup>-/-</sup> mice<sup>18</sup>. Chromatid-type breaks were particularly prevalent in etoposide-treated TDP2 patient cells, consistent with the high level of TOP2 activity associated with chromosome duplication during S/G<sub>2</sub>-phase<sup>24</sup>.

Another major source of genome instability and tumorigenesis are DSBs that arise during gene transcription, which can result in chromosome translocations. However, the origins of these DSBs and the molecular processes by which they result in translocations are unclear. Here, we have shown that TDP2 suppresses transcription-associated DSBs induced by TOP2 and prevents the accumulation of DSBs arising during estrogen stimulation. It is currently unclear whether the estrogen-induced breaks in TDP2-defective cells reflected estrogen-stimulated transcription or estrogen-stimulated entry into S phase. However, the former possibility is consistent with the requirement for TDP2 for normal levels of androgen-driven transcription in prostate cancer cells<sup>11</sup>. Indeed, in the current study, TDP2-depleted cells exhibited reduced expression of estrogen-regulated genes in breast cancer cells, suggesting that TDP2 does indeed protect estrogen-regulated gene transcription from inhibition by TOP2-induced DSBs.

Strikingly, TDP2 also prevented the accumulation of etoposide-induced DSBs within the breakpoint cluster region of the mixed lineage leukemia (*MLL*) locus. These observations have important implications for carcinogenesis because TOP2-induced DSBs are strongly implicated in oncogenic translocations<sup>15</sup>, including those

arising at *MLL*<sup>10, 15</sup>. Moreover, based on whole-chromosome FISH of chromosome 8 and 11, our results suggest that TDP2 suppresses the frequency of etoposide-induced chromosome translocations genome-wide. Importantly, the translocation events that were suppressed by TDP2 were almost entirely transcription dependent. These data highlight both the threat posed by abortive TOP2 activity to genome stability during transcription and the importance of TDP2-dependent DSB repair in suppressing TOP2-induced genome instability. It should be noted, however, that we did not detect an impact of TDP2 on genome instability in the absence of etoposide. This is consistent with the similar lack of genome instability in *Tdp2*<sup>-/-</sup> mice, in the absence of etoposide treatment. The apparent absence of genome instability in *TDP2*<sup>-/-</sup> cells and mice at endogenous levels of TOP2-induced DNA breakage most likely reflects the availability of alternative mechanisms for removing TOP2 from DSB termini, such as Mre11-dependent NHEJ and/or HR<sup>25</sup>.

Sequence analysis of the translocation junctions at *MLL* in our experiments revealed that ~80% of the translocation partners were transcriptionally active regions. Moreover, there was strong bias towards translocation partners located on the same chromosome as *MLL*, consistent with the proximity of TOP2-induced DSBs influencing the likelihood of rejoining incorrect chromosome termini. Since TDP2 activity generates DSB termini that are ligated by the canonical pathway for cNHEJ<sup>18</sup>, our discovery that TDP2 suppresses the overall frequency of etoposide-induced chromosome aberrations and translocations implicates a protective role for NHEJ events involving TDP2 in genome stability. This is in contrast to DSBs induced by recombinant nucleases, which in human cells implicate NHEJ in the formation of chromosome translocations<sup>26</sup>. We propose that this difference reflects the nature of the termini at the DSBs. In contrast to most DSBs, those induced by TOP2 and processed by TDP2 result in 4-bp cohesive 5'-overhangs that enable the rapid ligation of adjacent termini without further processing, thereby reducing the chance of joining together incorrect termini<sup>19</sup>. This idea may also explain our observation that a small subset of translocation junctions involving 4-bp of perfect homology appeared to be absent in *TDP2*<sup>-/-</sup> cells. We suggest that while adjacent 4-bp cohesive DSB termini generated by TDP2 are normally correctly re-ligated, on rare occasions these termini re-ligate incorrectly with a distal TOP2-induced DSB harboring an identical 4-bp sequence overhang. This type of junction has been reported for a therapy-related *MLL* translocation in leukemia<sup>27, 28</sup>, and suggested to arise by TOP2 domain swapping<sup>21, 29, 30</sup>. While further work is required to confirm this observation, our data suggest that this class of translocation may result from TOP2-induced DSBs processed by TDP2 (Fig. 4e). It is important to note, however, that most of the TOP2-induced translocations arising in wild-type and *TDP2*<sup>-/-</sup> RPE-1 cells either lacked homology at the junction or harbored smaller regions of microhomology, consistent with previous observations<sup>31</sup> and with their generation by TDP2-independent inaccurate NHEJ.

In summary, transcriptionally active loci are regions of the genome commonly associated with translocations<sup>32</sup>, and often these loci are associated with high levels of TOP2 activity<sup>9, 10, 12</sup>. Our data demonstrate the threat posed by TOP2-induced DSBs to gene transcription and genome stability, and highlight the role of TDP2-dependent NHEJ in protecting transcription and genome stability from these breaks.

## Methods

**Cell culture.** Normal and TDP2 patient<sup>11</sup> LCLs were maintained in RPMI supplemented with 10% FCS and with penicillin and streptomycin. RPE-1 cells (originally purchased from ATCC) were propagated in DMEM/F12 medium supplemented with 10% FCS. *TDP2*<sup>-/-</sup> RPE-1 cells expressing wild-type human

TDP2 or catalytic mutant TDP2<sup>D262A</sup> (TDP2<sup>CD</sup>) were generated using the lentiviral vector *pSIN-DUAL-GFP*. Cells were sorted for GFP positivity before use, and all experiments were performed in cultures with at least 70% of GFP-positive cells. For serum starvation cells were grown until confluency and then cultured in 1% FCS for 5–8 days. All cell lines were grown at 37 °C, 5% CO<sub>2</sub>. For ER-dependent gene induction cells were washed with serum-free medium three times for 30 min and incubated in medium containing 5% charcoal-stripped FBS for at least 48 h before experiment. Etoposide, 17 $\beta$ -estradiol (estradiol) and 5,6-dichlorobenzimidazole 1- $\beta$ -D-ribofuranoside (DRB) were purchased from Sigma. All cell lines were tested for mycoplasma contamination.

**Western blotting.** Anti-TDP2 antibody was employed at 1:5000 in western blotting and has been described previously<sup>33</sup>. Anti-Actin (Sigma A4700) was employed at 1:2000 and anti-Histone 1.2 (Abcam ab17677) at 1:20,000. Full blots are included in Supplementary Fig. 6.

**Generation of TDP2-depleted MCF7 cells and TDP2<sup>-/-</sup> RPE-1 cells.** To stably deplete TDP2, MCF7 cells were co-transfected with pcD2E vector and either pSuper-TTRAP/TDP2 or pSuper as described previously<sup>16</sup> and selected in 0.5 mg ml<sup>-1</sup> G418 to isolate single clones. To delete TDP2, we employed CRISPR-Cas9<sup>34</sup>. A 17-bp (minus the PAM) Tru-guide<sup>35</sup> RNA sequence targeting TDP2 exon 1 (5'-GCGGCGACTTCTGTGTG-3') was selected using the E-CRISP tool E-CRISP (<http://www.e-crisp.org/E-CRISP/>) and cloned into the guide RNA vector #41824 (AddGene). Briefly, the Tru-guide sequence minus the PAM were annealed and extended into a 98-mer double-stranded fragment using Phusion polymerase (NEB) which was then subcloned into the guide RNA vector using Gibson Assembly (NEB)<sup>36</sup>. The TDP2 guide construct was co-transfected with hCas9 expressed from plasmid #41815 (AddGene) using a Neon electroporation system (Invitrogen). Transfected cells were enriched by selection in 0.5 mg/ml G418 (ThermoFisher) for 5 days prior to isolation of single clones and screening for loss of TDP2 expression by western blotting.

**Clonogenic survival assays.** Cells were plated in 10 mm plates and 3 h later treated with indicated concentrations of etoposide for 1 h. Cells were rinsed twice with PBS and incubated with fresh drug-free media for 10 days and then fixed in 70% ethanol/1% methylene blue. The surviving fraction at each dose was calculated by dividing the average number of colonies (>50 cells) in treated dishes by the average number in untreated dishes.

**Micronuclei and nucleoplasmic bridges.** To detect micronuclei (MN) and nucleoplasmic bridges (NB), cytochalasin B (Sigma) was added to cell cultures at 4 mg ml<sup>-1</sup>. After 24 h, cells were fixed and stained with DAPI.

**Metaphase spreads.** For metaphase spreads, cells were incubated with demecolcine (Sigma) at 0.2  $\mu$ g/ml for 4 h and then harvested. Cells were collected using standard cytogenetic techniques and fixed in 3:1 methanol:acetic acid. Fixed cells were dropped onto acetic acid-humidified slides.

**Immunofluorescence and FISH analysis.** For immunofluorescence (IF), cells were grown on coverslips for 2 days (for cycling cultures) or 8 days (for serum-starved and confluency-arrested cell cultures) and then treated as indicated. Cells were fixed (10 min in PBS-4% paraformaldehyde), permeabilized (2 min in PBS-0.2% Triton X-100), blocked (30 min in PBS-5% BSA), and incubated with the indicated primary antibodies for 1–3 h in PBS-1% BSA. Cells were then washed (3  $\times$  5 min in PBS-0.1% Tween 20), incubated for 30 min with the corresponding AlexaFluor-conjugated secondary antibody (1:1000 dilution in 1% BSA-PBS) and washed again as described above. Finally, cells were counterstained with DAPI (Sigma) and mounted in Vectashield (Vector Labs).  $\gamma$ H2AX foci were manually counted (double-blind) in 20 to 40 cells per data point per experiment. For the analysis of G0/G1 cells, only CENP-F negative cells were scored. Primary anti- $\gamma$ H2AX (Millipore, 05-636) were employed at 1:1000 and anti-CENP-F (Abcam, ab5) at 1:500. FISH was performed according to manufacturer's protocol (CytoCell). For M-FISH, a modified version of Metasystems protocol was used. Chromosome and chromatid aberrations were analyzed and categorized as exchanges or breaks. Exchange type aberrations were further categorized as 'simple' (involving maximum of two breaks in two chromosomes) or 'complex' (any exchange involving >3 breaks)<sup>37</sup>.

**Inverse PCR.** For the detection of single chromosomal translocations an adaptation of an inverse PCR protocol<sup>38</sup> was performed. After treatment, genomic DNA was extracted with the DNeasy Blood and Tissue Kit (Qiagen). 10  $\mu$ g of DNA were incubated for 1 h with Shrimp Alkaline Phosphatase (NEB). After inactivation and digestion with either Bgl2 or Sac1 (NEB) DNA was purified with QIAquick PCR Purification Kit (Qiagen). To ensure ligation in *cis* of DNA molecules DNA was diluted at 18  $\mu$ g/ml and ligated for 20 h at 4 °C. After purification and quantification 40 ng of DNA were used per PCR reaction. Two consecutive PCR reactions (LongAmp polymerase, NEB) were performed to ensure specific amplification. First-round primers: MLLBgl2\_1: (5'-

GAGGAAATCAGCACCAACTGGGGGAA-3') and MLLBgl2\_2: (5'-GATCCTGTGGACTCCATCTGCTGGAA-3') or MLLSac1\_1 (5'-ATATGGGTGCAAAGCACTGTAT-3') and MLLSac1\_2 (5'-ACCACTCCTTCAACTTCTGGG-3'). One microliter of a 1:10 dilution of the first-round PCR mix was amplified with second-round primers: MLLBgl2\_3: (5'-CATTAGCAGGTGGGTTTAGCGCTGGG-3') and MLLBgl2\_4: (5'-TTCTCCTGCTTATGACCGGAGGTGG-3') or MLLSac1\_3 (5'-CAGTA-GACCCCTGGCACTTG-3') and MLLSac1\_4 (5'-AGCAATCTCAGAGGGTTCCT-3'). Novel-size bands were purified from agarose gel and sequenced using MLLBgl2\_4 or MLLSac1\_5 (5'-AGTGCAGTGGCGTGATAATG-3') primers. Translocations were represented as by a Circos plot<sup>39</sup> using the OmicCircos package of R.

**Data availability.** All data are available from the authors on request.

Received: 10 January 2017 Accepted: 19 June 2017

Published online: 10 August 2017

## References

- Stratton, M. R., Campbell, P. J. & Futreal, P. A. The cancer genome. *Nature* **458**, 719–724 (2009).
- Mitelman, F., Johansson, B. & Mertens, F. The impact of translocations and gene fusions on cancer causation. *Nat. Rev. Cancer* **7**, 233–245 (2007).
- Mani, R.-S. & Chinnaiyan, A. M. Triggers for genomic rearrangements: insights into genomic, cellular and environmental influences. *Nat. Rev. Genet.* **11**, 819–829 (2010).
- Osborne, C. S. et al. Myc dynamically and preferentially relocates to a transcription factory occupied by Igh. *PLoS Biol.* **5**, e192 (2007).
- Lin, C. et al. Nuclear receptor-induced chromosomal proximity and DNA breaks underlie specific translocations in cancer. *Cell* **139**, 1069–1083 (2009).
- Cowell, I. G. & Austin, C. A. Mechanism of generation of therapy related leukemia in response to anti-topoisomerase II agents. *Int. J. Environ. Res. Public Health* **9**, 2075–2091 (2012).
- Wright, R. L. & Vaughan, A. T. M. A systematic description of MLL fusion gene formation. *Crit. Rev. Oncol. Hematol.* **91**, 283–291 (2014).
- Foster, H. A. et al. Relative proximity of chromosome territories influences chromosome exchange partners in radiation-induced chromosome rearrangements in primary human bronchial epithelial cells. *Mutat. Res.* **756**, 66–77 (2013).
- Ju, B.-G. et al. A topoisomerase II $\beta$ -mediated dsDNA break required for regulated transcription. *Science* **312**, 1798–1802 (2006).
- Haffner, M. C. et al. Androgen-induced TOP2B-mediated double-strand breaks and prostate cancer gene rearrangements. *Nat. Genet.* **42**, 668–675 (2010).
- Gómez-Herreros, F. et al. TDP2 protects transcription from abortive topoisomerase activity and is required for normal neural function. *Nat. Genet.* **46**, 516–521 (2014).
- Madabhushi, R. et al. Activity-induced DNA breaks govern the expression of neuronal early-response genes. *Cell* **161**, 1592–1605 (2015).
- Pommier, Y., Sun, Y., Huang, S.-Y. N. & Nitiss, J. L. Roles of eukaryotic topoisomerases in transcription, replication and genomic stability. *Nat. Rev. Mol. Cell Biol.* **17**, 703–721 (2016).
- Deweese, J. E. & Osheroff, N. The DNA cleavage reaction of topoisomerase II: wolf in sheep's clothing. *Nucleic Acids Res.* **37**, 738–748 (2009).
- Cowell, I. G. et al. Model for MLL translocations in therapy-related leukemia involving topoisomerase II $\beta$ -mediated DNA strand breaks and gene proximity. *Proc. Natl Acad. Sci. USA* **109**, 8989–8994 (2012).
- Cortes Ledesma, F., El-Khamisy, S. F., Zuma, M. C., Osborn, K. & Caldecott, K. W. A human 5'-tyrosyl DNA phosphodiesterase that repairs topoisomerase-mediated DNA damage. *Nature* **461**, 674–678 (2009).
- Zeng, Z., Cortes Ledesma, F., El-Khamisy, S. F. & Caldecott, K. W. TDP2/TTRAP is the major 5'-tyrosyl DNA phosphodiesterase activity in vertebrate cells and is critical for cellular resistance to topoisomerase II-induced DNA damage. *J. Biol. Chem.* **286**, 403–409 (2011).
- Gómez-Herreros, F. et al. TDP2-dependent non-homologous end-joining protects against topoisomerase II-induced DNA breaks and genome instability in cells and in vivo. *PLoS Genet.* **9**, e1003226 (2013).
- Caldecott, K. W. Tyrosyl DNA phosphodiesterase 2, an enzyme fit for purpose. *Nat. Struct. Mol. Biol.* **19**, 1212–1213 (2012).
- Rogakou, E. P., Pilch, D. R., Orr, A. H., Ivanova, V. S. & Bonner, W. M. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J. Biol. Chem.* **273**, 5858–5868 (1998).
- Meyer, C. et al. The MLL recombinome of acute leukemias in 2013. *Leukemia* **27**, 2165–2176 (2013).



22. Hars, E. S., Lyu, Y. L., Lin, C.-P. & Liu, L. F. Role of apoptotic nuclease caspase-activated DNase in etoposide-induced treatment-related acute myelogenous leukemia. *Cancer Res.* **66**, 8975–8979 (2006).
23. Wang, Z.-Q. et al. Inhibition of P-TEFb by DRB suppresses SIRT1/CK2 $\alpha$  pathway and enhances radiosensitivity of human cancer cells. *Anticancer. Res.* **34**, 6981–6989 (2014).
24. Nitiss, J. L. Targeting DNA topoisomerase II in cancer chemotherapy. *Nat. Rev. Cancer* **9**, 338–350 (2009).
25. Hoa, N. N. et al. Mre11 is essential for the removal of lethal topoisomerase 2 covalent cleavage complexes. *Mol. Cell* **64**, 580–592 (2016).
26. Ghezraoui, H. et al. Chromosomal translocations in human cells are generated by canonical nonhomologous end-joining. *Mol. Cell* **55**, 829–842 (2014).
27. Whitmarsh, R. J. et al. Reciprocal DNA topoisomerase II cleavage events at 5'-TATTA-3' sequences in MLL and AF-9 create homologous single-stranded overhangs that anneal to form der(11) and der(9) genomic breakpoint junctions in treatment-related AML without further processing. *Oncogene* **22**, 8448–8459 (2003).
28. Meyer, C. et al. Diagnostic tool for the identification of MLL rearrangements including unknown partner genes. *Proc. Natl Acad. Sci. USA* **102**, 449–454 (2005).
29. Ikeda, H., Aoki, K. & Naito, A. Illegitimate recombination mediated in vitro by DNA gyrase of *Escherichia coli*: structure of recombinant DNA molecules. *Proc. Natl Acad. Sci. USA* **79**, 3724–3728 (1982).
30. Pommier, Y., Zwelling, L. A., Kao-Shan, C. S., Whang-Peng, J. & Bradley, M. O. Correlations between intercalator-induced DNA strand breaks and sister chromatid exchanges, mutations, and cytotoxicity in Chinese hamster cells. *Cancer Res.* **45**, 3143–3149 (1985).
31. Han, Y. H., Austin, M. J., Pommier, Y. & Povirk, L. F. Small deletion and insertion mutations induced by the topoisomerase II inhibitor teniposide in CHO cells and comparison with sites of drug-stimulated DNA cleavage in vitro. *J. Mol. Biol.* **229**, 52–66 (1993).
32. Wei, P.-C. et al. Long Neural Genes Harbor Recurrent DNA Break Clusters in Neural Stem/Progenitor Cells. *Cell* **164**, 644–655 (2016).
33. Thomson, G. et al. Generation of assays and antibodies to facilitate the study of human 5'-tyrosyl DNA phosphodiesterase. *Anal. Biochem.* **436**, 145–150 (2013).
34. Mali, P. et al. RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).
35. Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M. & Joung, J. K. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* **32**, 279–284 (2014).
36. Hanzlikova, H., Gittens, W., Krejčíková, K., Zeng, Z. & Caldecott, K. W. Overlapping roles for PARP1 and PARP2 in the recruitment of endogenous XRCC1 and PNKP into oxidized chromatin. *Nucleic Acids Res.* **45**, 2546–2557 (2017).
37. Hoa, N. N. et al. Mre11 Is essential for the removal of lethal topoisomerase 2 covalent cleavage complexes. *Mol. Cell* **64**, 1010 (2016).
38. Blanco, J. G., Edick, M. J. & Relling, M. V. Etoposide induces chimeric MLL gene fusions. *FASEB J.* **18**, 173–175 (2004).
39. Krzywinski, M. et al. Circos: an information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645 (2009).

## Acknowledgements

We thank F. Cortés-Ledesma for providing TDP2 lentiviral vectors, A. Rodríguez-Gil for his help with Circos plots, and M.J. Castro-Pérez for FACS technical support. F.G.-H. was funded by a CR-UK Programme grant to KWC (C6563/A16771) and by the Spanish Ministry of Economy and Competitiveness (RYC-2014-16665 and BFU2016-76446-P). M.I. M.-M. and G.Z.V. were funded by a CR-UK Programme grant to KWC (C6563/A16771). A.H.-R. was recipient of a Leonardo da Vinci Vocational Education and Training program (MERCURIO 2013-1-ES1-LEO02-66587).

## Author contributions

All experiments were performed by F.G.-H. unless indicated. G.Z.-V. and I.N. generated TDP2<sup>-/-</sup> RPE-1 cells and TDP2-depleted MCF7 cells, respectively. G.Z.-V., I.N., and M. I.M.-M. conducted clonogenic survival assays,  $\gamma$ H2AX assays, and transcription assays. R. M.A. performed and scored M-FISH on LCLs. A.H.-R. performed MLL break-apart FISH on LCLs. K.W.C. and F.G.-H. conceived the study, designed the experiments, interpreted the results, and wrote the manuscript.

## Additional information

**Supplementary Information** accompanies this paper at doi:10.1038/s41467-017-00307-y.

**Competing interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017

## TDP2, TOP2, and SUMO: what is ZATT about?

*Cell Research* (2017) 27:1405–1406. doi:10.1038/cr.2017.147; published online 21 November 2017

**Recently, ZATT (also known as ZNF451 or Zpf451) was reported by Schellenberg *et al.* to aid the removal of Topoisomerase II cleavage complexes by stimulating the phosphodiesterase activity of Tyrosyl DNA Phosphodiesterase 2. Although the full implication of this discovery is unknown, it will help us understand how cells respond to topoisomerase-induced genome damage and chemotherapeutic topoisomerase ‘poisons’.**

Sources of endogenous DNA lesions are diverse, as are the biochemical mechanisms that comprise the DNA damage response (DDR) and remove them. DNA double strand breaks (DSBs) are potentially clastogenic and cytotoxic lesions, resulting not only from attack of DNA by electrophilic molecules such as reactive oxygen species but also from ‘programmed’ sources such as topoisomerase activity [1]. Topoisomerase II (TOP2) removes torsional stress from DNA during transcription and replication. To achieve this, it generates a transient DSB intermediate known as the ‘cleavage complex’ (TOP2cc), within which the topoisomerase is covalently linked to the 5′-termini of DSBs via 5′ phosphotyrosyl bonds. This enables intact DNA duplex to pass through the break and thereby decatenate DNA and remove DNA torsional stress [2]. While TOP2 normally reseals the DSB at the end of each catalytic cycle, under some conditions the cleavage complex becomes abortive, generating a potentially pathogenic DSB that requires DSB repair for removal.

The enzyme 5′-tyrosyl DNA phosphodiesterase 2 (TDP2) can initiate the re-

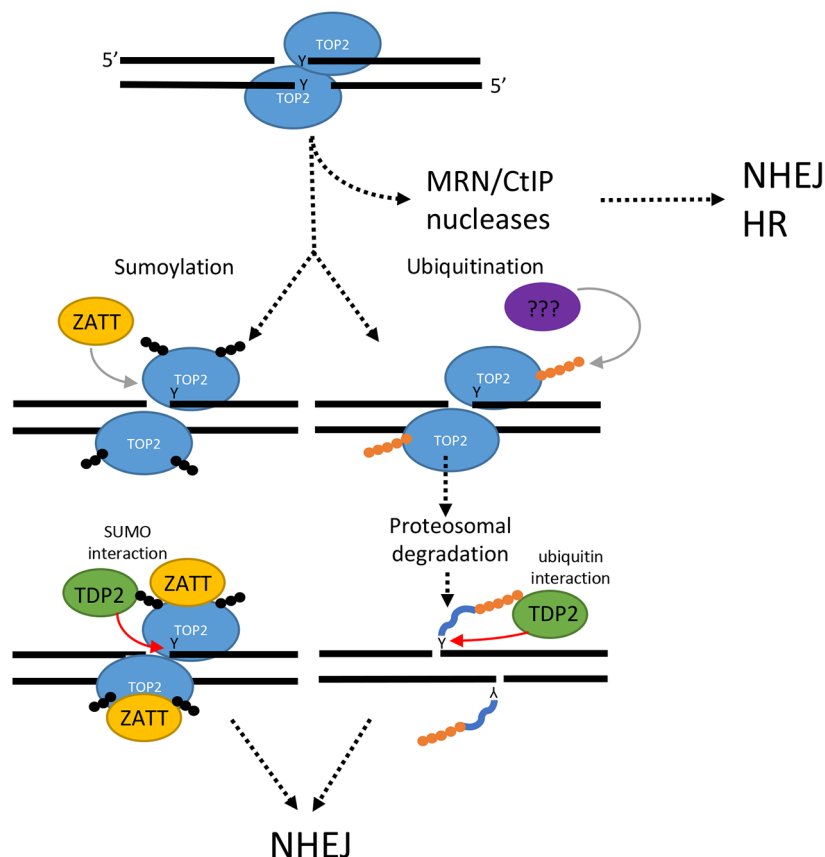
pair of TOP2cc by hydrolytic removal of trapped topoisomerase peptide from the 5′-termini of DSBs, thereby allowing the DSB to be religated [3]. The importance of this enzyme is illustrated by its proposed role in oncogenic translocations, resistance to chemotherapeutic topoisomerase ‘poisons’ and by causative mutations in the *TDP2* gene in Spinocerebellar Ataxia Autosomal Recessive 23 (SCAR23) [4,5]. The denaturation and/or degradation of TOP2cc is thought to be a prerequisite for TDP2 activity, to enable access of the 5′-phosphotyrosyl bond within the active site of topoisomerase [6]. In the current work, Schellenberg *et al.* [7] challenge this model and clarify how TDP2 processes TOP2cc.

In contrast to previous models, Schellenberg *et al.* [7] identify a new TDP2-dependent DNA repair pathway that involves TOP2 sumoylation independent of ubiquitination and proteasome activity. The authors show that TDP2 interacts with ZNF451 (aka Zpf451), a SUMO E3/E4 ligase/elongase that they denote ZATT (Zinc finger protein Associated with TDP2 and TOP2). ZNF451/ZATT employs tandem sumo interacting motifs (SIMs) and a C<sub>2</sub>H<sub>2</sub> zinc finger to promote SUMO2 conjugation and chain elongation *in vitro*, both in unperturbed cells and in perturbed cells treated with topoisomerase I poison, camptothecin [8]. ZNF451/ZATT was also identified as a TOP2-interacting protein by mass spectrometry, further implicating this protein in topoisomerase biology [9]. Schellenberg *et al.* identify a ‘split-SIM’ domain within TDP2 that confers SUMO2 binding specificity, and mutation of which ablates interaction with

TOP2 and ZNF451/ZATT.

Schellenberg and colleagues demonstrate that ZNF451/ZATT promotes both TOP2 sumoylation and TDP2 phosphodiesterase activity on intact TOP2cc. However, one of the most striking features of the work is that the latter stimulation occurs largely independently of SUMOylation, suggesting that direct interaction with ZNF451/ZATT alters the conformation of TOP2cc in a way that renders the DSB more accessible to TDP2 (Figure 1). The role of SUMO in this pathway is more complex than postulated, as suggested by the surprising finding that mutation of the SUMO binding interface in TDP2 did not affect cell survival or DSB repair following etoposide treatment, even in the absence of proteasome activity [7]. In addition, this complexity is most likely further increased by the presence of a ubiquitin associated (UBA) domain located at the N-terminus of TDP2, which has been shown to interact with mono- and polyubiquitin chains [10].

A second striking feature of the work is that TDP2 appears to be the major if not only mechanism for removal of TOP2cc in the absence of proteasome activity. Loss of ZNF451/ZATT also leads to cellular hypersensitivity and delayed DSB repair following etoposide treatment in the absence of proteasome, albeit not to the same extent as loss of TDP2. Whilst ZNF451/ZATT thus appears to play a role in the response of cells to TOP2 perturbation, this may not be its primary function. Nevertheless, the identification of ZNF451/ZATT as a new player in the cellular response to topoisomerase inhibition raises new



**Figure 1** A model for the repair of abortive TOP2 cleavage complexes. **Top**, TOP2 becomes transiently linked to the 5'-terminus of a DSB during its normal catalytic cycle via a phosphotyrosine bond, represented by the 'Y' on the DNA strand. Occasionally, these TOP2 cleavage complexes (TOP2cc) can become abortive, and thus require DSB repair for their removal. TOP2 can be removed from the 5'-termini of the DSBs by a potentially mutagenic mechanism that employs nucleases such as MRE11/RAD50/NBS1 (MRN) complex and CtIP. The DNA double strand break can then be repaired by homologous recombination (HR) or non-homologous end-joining (NHEJ). Alternatively, abortive TOP2cc can be repaired by TDP2-dependent mechanisms, providing an error-free route that does not require DNA degradation. **Middle**, the abortive TOP2cc is sumoylated by ZATT (left), and/or ubiquitinated (right). **Bottom**, TDP2 is recruited to the abortive TOP2cc by interaction with SUMO (left), and/or possibly ubiquitin (right), releasing TOP2 from the DSB by a hydrolytic mechanism. Note that the pathway on the left does not require proteasome activity, because ZATT and/or SUMO renders the phosphotyrosine bond accessible to TDP2. In contrast, the pathway on the right employs proteasome activity to render the 5'-phosphotyrosine bond visible to TDP2. Both TDP2-dependent pathways result in a DSB that can be ligated during NHEJ without nucleolytic processing, providing an error-free mechanism for DSB repair. Small black circles represent SUMO2 chains and orange circles represent ubiquitin chains.

provides a fascinating new insight into topoisomerase biology and the cellular processes that regulate it. Like all exciting studies, the work raises more questions than answers. For example, what is the relative contribution of SUMO and ubiquitin in the processing of TOP2 cc, and how do cells choose which mechanism to employ to recognize, process, and repair these structures? The cell cycle could provide us with some clues, since TOP2 is reportedly heavily sumoylated during mitosis [11]. In any case, further examination of TOP2 post-translation modifications and their regulation will allow a better understanding of how cells protect their genome from topoisomerase-induced DNA damage.

Guido Zagnoli-Vieira<sup>1</sup>,  
Keith W Caldecott<sup>1</sup>

<sup>1</sup>Genome Damage and Stability Centre, University of Sussex, Falmer, Brighton, BN1 9RQ, United Kingdom

Correspondence: Keith W Caldecott<sup>a</sup>,  
Guido Zagnoli-Vieira<sup>b</sup>

<sup>a</sup>E-mail: k.w.caldecott@sussex.ac.uk

<sup>b</sup>E-mail: G.Zagnoli-Vieira@sussex.ac.uk

## References

- 1 Tubbs A, Nussenzweig A. *Cell* 2017; **168**:644-656.
- 2 Pommier Y, Sun Y, Huang SN, *et al. Nat Rev Mol Cell Biol* 2016; **17**:703-721.
- 3 Ledesma FC, El Khamisy SF, Zuma MC, *et al. Nature* 2009; **461**:674-678.
- 4 Gómez-Herreros F, Schuurs-Hoeijmakers JH, McCormack M, *et al. Nat Genet* 2014; **46**:516-521.
- 5 Gómez-Herreros F, Zagnoli-Vieira G, Ntai I, *et al. Nat Commun* 2017; **8**:233.
- 6 Gao R, Schellenberg MJ, Huang SY, *et al. J Biol Chem* 2014; **289**:17960-17969.
- 7 Schellenberg MJ, Lieberman JA, Herrero-Ruiz A, *et al. Science* 2017; **357**:1412-1416.
- 8 Eisenhardt N, Chaugule VK, Koidl S, *et al. Nat Struct Mol Biol* 2015; **22**:959-967.
- 9 Uusküla-Reimand L, Hou H, Samavarchi-Tehrani P, *et al. Genome Biol* 2016; **17**: 182.
- 10 Rao T, Gao R, Takada S, *et al. Nucleic Acids Res* 2016; **44**:10201-10215.
- 11 Azuma Y, Arnautov A, Anan T, *et al. EMBO J* 2005; **24**:2172-2182.

questions concerning the role of SUMO with respect to cancer incidence, chemotherapy, and hereditary neurological

diseases that are related to topoisomerase biology.

The work by Schellenberg *et al.*

# Confirming TDP2 mutation in spinocerebellar ataxia autosomal recessive 23 (SCAR23)

Guido Zagnoli-Vieira, BSc, Francesco Bruni, PhD,\* Kyle Thompson, PhD, Langping He, MD, PhD, Sarah Walker, PhD, Arjan P.M. de Brouwer, PhD, Robert Taylor, PhD, FRCPATH, Dmitriy Niyazov, MD, and Keith W. Caldecott, PhD

**Correspondence**  
Dr. Caldecott  
k.w.caldecott@sussex.ac.uk

*Neurol Genet* 2018;4:e262. doi:10.1212/NXG.0000000000000262

## Abstract

### Objective

To address the relationship between mutations in the DNA strand break repair protein tyrosyl DNA phosphodiesterase 2 (TDP2) and spinocerebellar ataxia autosomal recessive 23 (SCAR23) and to characterize the cellular phenotype of primary fibroblasts from this disease.

### Methods

We have used exome sequencing, Sanger sequencing, gene editing and cell biology, biochemistry, and subcellular mitochondrial analyses for this study.

### Results

We have identified a patient in the United States with SCAR23 harboring the same homozygous *TDP2* mutation as previously reported in 3 Irish siblings (c.425+1G>A). The current and Irish patients share the same disease haplotype, but the current patient lacks a homozygous variant present in the Irish siblings in the closely linked gene *ZNF193*, eliminating this as a contributor to the disease. The current patient also displays symptoms consistent with mitochondrial dysfunction, although levels of mitochondrial function in patient primary skin fibroblasts are normal. However, we demonstrate an inability in patient primary fibroblasts to rapidly repair topoisomerase-induced DNA double-strand breaks (DSBs) in the nucleus and profound hypersensitivity to this type of DNA damage.

### Conclusions

These data confirm the *TDP2* mutation as causative for SCAR23 and highlight the link between defects in nuclear DNA DSB repair, developmental delay, epilepsy, and ataxia.

\*Current address: Department of Biosciences Biotechnologies and Biopharmaceutics, University of Bari Aldo Moro, Bari, Italy.

From the Genome Damage and Stability Centre (G.Z.-V., K.W.C.), University of Sussex, Falmer, Brighton, United Kingdom; Wellcome Centre for Mitochondrial Research (F.B., K.T., L.H., R.T.), Institute of Neuroscience, Newcastle University, Tyne, United Kingdom; Sussex Drug Discovery Centre (S.W.), University of Sussex, Falmer, Brighton, United Kingdom; Department of Human Genetics (A.P.M.d.B.), Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, The Netherlands; and Medical Genetics (A.P.M.d.B., D.N.), Ochsner Health Center for Children, New Orleans, LA.

Funding information and disclosures are provided at the end of the article. Full disclosure form information provided by the authors is available with the full text of this article at [Neurology.org/NG](http://Neurology.org/NG).

The Article Processing Charge was funded by the authors.

This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (CC BY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Glossary

**DSB** = double-strand break; **ETC** = electron transport chain; **FCS** = fetal calf serum; **NHEJ** = nonhomologous end joining; **SCAR23** = spinocerebellar ataxia, autosomal recessive 23; **TDP2** = tyrosyl DNA phosphodiesterase 2; **WCE** = whole-cell extract; **WES** = Whole-exome sequencing.

---

DNA is under constant threat from attack by endogenous and exogenous electrophilic molecules,<sup>1</sup> and DNA topoisomerase enzymes can introduce DNA breaks as abortive intermediates of their activity.<sup>2–4</sup> Topoisomerase “poisons” such as etoposide inhibit the ligation activity of topoisomerase 2 (TOP2), thereby promoting the formation of abortive DNA double-strand break (DSB) intermediates that require DSB repair. DSBs are repaired in cells by either homologous recombination-mediated repair or by nonhomologous end joining (NHEJ).<sup>2</sup> The repair of TOP2-induced DSBs by NHEJ involves the enzyme tyrosyl DNA phosphodiesterase 2 (TDP2), which removes trapped topoisomerase peptide from the 5′-termini at the DSB and thereby allows the DNA ends to be ligated.<sup>3–5</sup> The loss of TDP2 in mouse results in reduced expression of >100 genes in the brain,<sup>6</sup> and *TDP2* mutation in humans has been associated with intellectual disability, seizures, and ataxia,<sup>6</sup> a disease now denoted as spinocerebellar ataxia, autosomal recessive 23 (SCAR23). To date, our understanding of SCAR23 has been limited by the availability of only 3 Irish siblings with a mutation in *TDP2* and by the lack of availability of fibroblast cell lines from these patients for molecular and cellular characterization. Here, we have addressed these limitations and identified an SCAR23 patient in the United States with the same homozygous *TDP2* mutation as present in the Irish siblings, confirming the association of this disease with mutated *TDP2*. In addition, we have characterized at the molecular and cellular level primary patient fibroblasts from the current SCAR23 patient.

## Methods

### Patient case report

The patient is currently a 6-year-old caucasian boy from the United States presenting with developmental delay, microcephaly, and failure to thrive. His mother and father are aged 29 years and 32 years, respectively, and there is no previous family history of related disease, but there is possible consanguinity (3rd cousins). The patient started walking at the age of 14 months and talking at 3 years, his hearing was reportedly normal, but he did not have auditory brainstem response. The patient is easily fatigued, and his parents reported that he eats excessively, becomes irritable, and falls asleep sometimes for days or up to 2 weeks. He is now below the fifth percentile for weight and 5% for height, with a body mass index at <5%. He has a G-tube and is followed by a dietitian and receives nutritional liquid supplements (PediaSure; Abbott, Lake Forest, IL). He has a history of constipation and is followed up by a gastroenterologist. He has a 12-year-old brother who is

also easily fatigued and a 14-year-old maternal half-sister with attention deficit hyperactivity disorder (ADHD).

The patient has difficulty keeping balance and has an ataxic gait. He has been tested several times for abnormalities in blood, urine, and CSF, and by MRI and EEG, and only the latter is abnormal. Seizures began at the age of 5 months, and his EEGs show increased slowing and occasional spikes in the right posterior quadrant, and during drowsiness, he had 2 generalized bursts of polyspike and slow-wave activity at a frequency of 4–5 Hz. The patient’s 180K oligoarray, very long chain fatty acids, carbohydrate-deficient transferrin, CSF lactate, and neurotransmitters and plasma amino acids were normal.

The patient exhibited several phenotypes consistent with mitochondrial dysfunction such as hypotonia, low energy, fatigability, hypersomnia, failure to thrive, short stature, constipation, neutropenia, hyponatremia, cardiac arrhythmia, and gastrointestinal dysmotility. Consistent with this, electron transport chain (ETC) studies on muscle biopsy at age 1 year showed a severe reduction of complex I + III and II + III activity, which satisfied the major Walker criteria after correction for increased citrate synthase activity. CoQ10 deficiency was suggested based on the complex I + III and II + III deficiency, and his metabolic tests were abnormal with the high lactate and high lactate/pyruvate ratio. The patient was placed on ubiquinol, carnitine, and leucovorin, and anecdotally, he responded well (particularly to liposomal ubiquinol, the active or reduced form of CoQ10) in terms of energy and developmental progress. He is reportedly no longer sleepy and lethargic and has a stable gait, but is “still behind” in language and cognitive skills and is in special education. He still has idiopathic fevers but has not been admitted to the hospital for 15 months. He does not have as many infections but still has gastric dysmotility. His seizures have reduced in frequency.

### Standard protocol approvals, registrations, and patient consents

We confirm that we have received approval from an institutional ethics standards committee for this work, and we have written informed consent for research from the guardian of the patient for participation in this study.

### Exome sequencing and haplotype analysis

Whole-exome sequencing (WES) (trio study) of the patient was performed by GeneDx using the Agilent Clinical Research Exome kit to target the exonic regions and flanking splice junctions of the genome. These targeted regions were sequenced simultaneously by massively parallel (NextGen, Irvine, CA)



sequencing on an Illumina HiSeq sequencing system with 100bp paired-end reads. Bidirectional sequence was assembled, aligned to reference gene sequences based on human genome build GRCh37/UCSC hg19, and analyzed for sequence variants in the selected genes or regions of interest using a custom-developed analysis tool (Xome Analyzer). Capillary sequencing was used to confirm all potentially pathogenic variants identified in this individual. Sequence alterations were reported according to the nomenclature guidelines of the Human Genome Variation Society. The WES identified a homozygous splice site mutation (c.425+1G>A) in the *TDP2* gene. For comparison of the current patient with the Irish pedigree previously reported,<sup>6</sup> we conducted haplotype analysis. Variants were considered for homozygosity if they were (1) covered by at least 4 reads or more, (2) present in 80% of all reads or more, (3) designated as a substitution, (4) uniquely positioned in the human genome, and (5) present in the exome data of both individuals. Homozygous regions were determined using a sliding window, accepting 2 or less homozygous variants per 10 variants assessed.

### Mitochondrial preparation and subcellular fractionation

Mitochondria were prepared as described previously,<sup>7</sup> with few modifications. HeLa cells and fibroblasts (control and patient) were harvested, resuspended in homogenization buffer (HB [0.6 M mannitol, 10 mM Tris-HCl pH 7.4, 1 mM (ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid) (EGTA), 0.1% bovine serum albumin (BSA) (wt/vol)], and subjected to differential centrifugation. Mitochondria were pelleted at 11,000g for 10 minutes at 4°C and resuspended in HB; the postmitochondrial supernatant was retained after centrifugation ("post-mito spin"). For submitochondrial fraction preparation, HeLa cell mitochondria (300 μg) were treated with 1.6 μg of proteinase K on ice for 30 minutes, followed by the addition of 5 mM phenylmethanesulfonyl fluoride (PMSF). This fraction was pelleted at 11,000g for 10 minutes at 4°C and resuspended in HB. Mitoplasts were obtained by resuspending PK-treated mitochondria in 9 volumes of 10 mM Tris-HCl (pH 7.4) and treated with PK, as described earlier. Inner mitochondrial membrane proteins were extracted in the presence of 100 mM Na<sub>2</sub>CO<sub>3</sub>, followed by centrifugation at 100,000g for 15 minutes at 4°C. Proteins (30 μg) from each fraction were loaded onto 12% SDS-PAGE gel, transferred to the polyvinylidene difluoride (PVDF) membrane, and analyzed by immunoblotting using primary antibodies to apoptosis inducing factor (AIF) (NEB), eIF4E (Cell Signalling), EF-Tu (custom made), NDUFB8 (Mitosciences), and TDP2 (see Western Blotting, below).

### Cell culture and vectors

Human A549 cells were grown in Dulbecco Modified Eagle Medium (Gibco, ThermoFisher, Waltham, MA) containing 10% fetal calf serum (FCS), 2 mM glutamine, penicillin (100 units/mL), and streptomycin (100 μg/mL). Human fibroblasts were grown in Minimum Essential Media (Gibco) containing 15% FCS, 2 mM glutamine, penicillin (100 units/mL), and streptomycin (100 μg/mL). All cells were grown at 5% CO<sub>2</sub> at

37°C. *TDP2*-mutated patient primary human fibroblasts were established from a patient's skin biopsy and were denoted 850-BR. The control human fibroblast cell line 1-BR.<sup>3</sup> (denoted here for simplicity as 1-BR) was previously derived from an unrelated normal individual and has been described previously.<sup>8</sup> For complementation experiments, we used 1-BR cells that were immortalized previously with hTERT (denoted as 1-BR hTERT) and a derivative of 850-BR immortalized in the current study by retroviral-mediated hTERT expression and selected in a medium containing 1 μg/ml puromycin (denoted as 850-BR hTERT). For complementation with human *TDP2*, 850-BR hTERT cells were transfected with either empty eGFP-N1 vector or eGFP-N1 construct encoding GFP-tagged human *TDP2* (denoted as TDP2-GFP-N1) and stable transfectants selected for 21 days by growth in a medium containing 0.5 mg/mL G418 (Gibco, ThermoFisher, Waltham, MA). TDP2-GFP-N1 was generated by PCR amplification of the human *TDP2* ORF using the primers *TDP2\_FW* (5'-AAAGAATTCATGGAGTTGGGGAGTTGCCTG-3') and *TDP2\_RV* (5'-AAAGGATCCAATATTATATCTAAGTTGCACAGAAGACC-3') and subcloning the PCR product into the *EcoRI/BamHI* sites of eGFP-N1.

### CRISPR/Cas9 gene editing

A549 *TDP2*<sup>-/-</sup> cells were created as previously reported.<sup>9</sup> In brief, we used a 17-bp (minus the PAM) RNA sequence targeting *TDP2* exon 4 (5'-GTAGAAATATCACATCT-3'), which was selected using the tool E-CRISP (e-crisp.org/E-CRISP/) and cloned into the guide RNA vector #41824 (AddGene). The *TDP2* guide construct was cotransfected with hCas9 expressed from plasmid #41815 (AddGene) using Amaxa Nucleofector platform (Lonza, Basel, Switzerland) Kit T program X-001. Transfected cells were enriched by selection in 1 mg/mL G418 (ThermoFisher) for 5 days before isolation of single clones and screening for loss of *TDP2* expression by Western blotting.

### Sanger sequencing

DNA was extracted from 850-BR cells using the DNeasy Blood & Tissue kit (Qiagen, Manchester, UK). PCR reactions used Phusion HF-DNA Polymerase (NEB) and the following primers: *TDP2\_FW* GCCAGTGGT-GACCTAACCAATGAAGA; *TDP2\_RV* CTGTAGAAA-TATCACATCTGGGCTGTACC; *ZSCAN9\_FW* ATGAAGT AACCAAGACTGAGGACAGAGAG; and *ZSCAN9\_RV* AGACCAGCTCAGCCACTGTGTGGATCT. PCR products were purified before sequencing using a QIAquick PCR purification kit (Qiagen).

### Western blotting

Anti-TDP2 antibody was used at 1:5000 in Western blotting and has been described previously.<sup>10</sup> Anti-Actin (Sigma A4700) was used at 1:2000. Western blot assessment of OXPHOS components in patient primary fibroblasts was conducted as described previously,<sup>11</sup> using antibodies against NDUFB8 (Abcam cat# ab110242), SDHA (Abcam cat# ab14715), UQCRC2 (Abcam cat# ab14745), COXI (Abcam

cat# ab14705), ATP5A (Abcam cat# ab14748), SDHB (Abcam cat# ab14714), COX II (Abcam cat# 110258), and VDAC1 (Abcam cat# ab14734).

Clonogenic survival assays

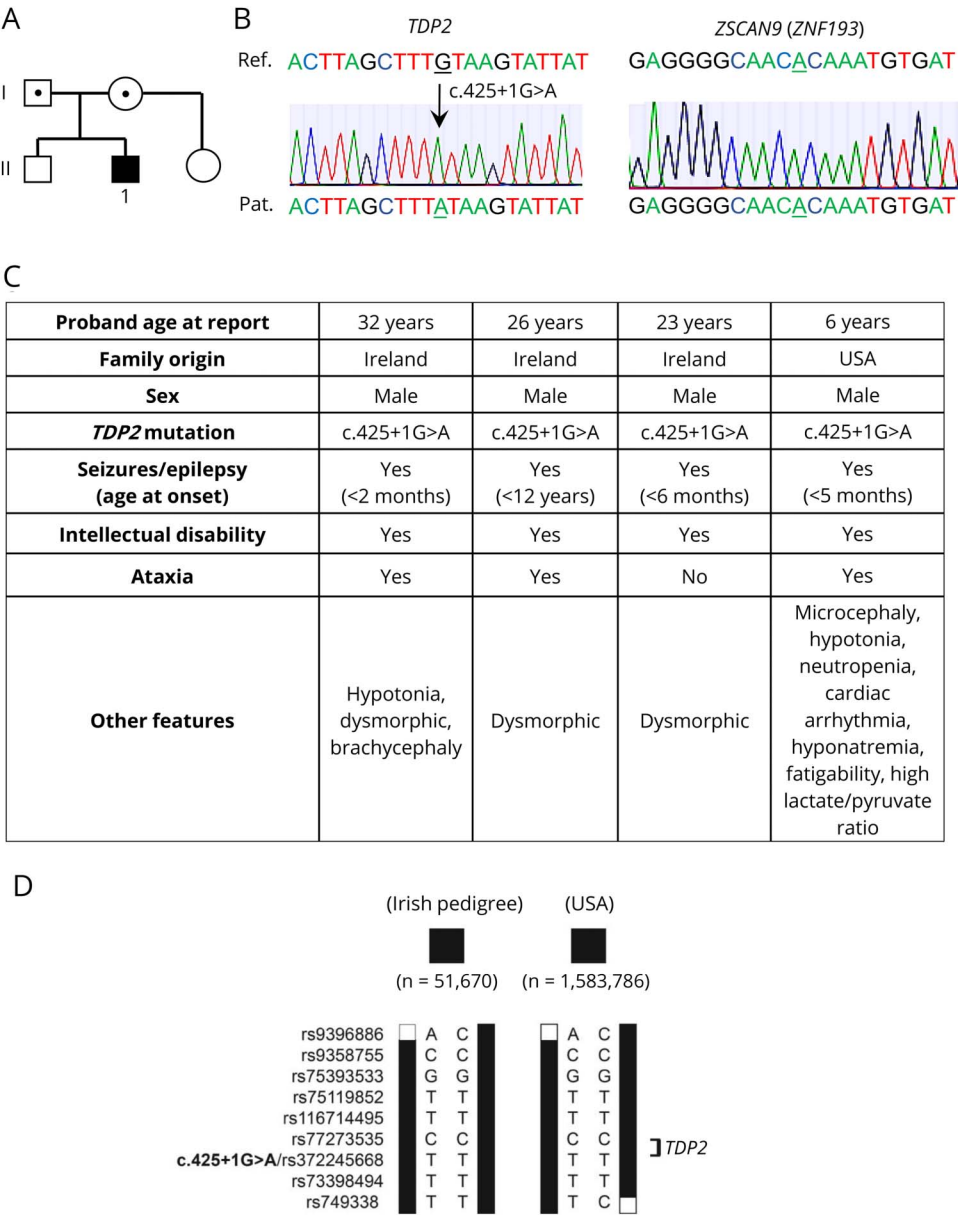
1-BR and 850-BR primary fibroblasts cells were plated onto feeder layers (see below) and 3 hours later treated with indicated concentrations of etoposide for 21 days to allow the formation of macroscopic colonies, which were rinsed in phosphate buffered saline (PBS) and fixed/stained in 70% ethanol/1% methylene blue. A549 and *TDP2*<sup>-/-</sup> A549 were plated 4 hours before treatment with the indicated concentrations of etoposide for 12 days and stained as described earlier. The surviving fraction at each dose was calculated by dividing the average number of

colonies (>50 cells) in treated dishes by the average number in untreated dishes. For feeder layers, 1-BR cells were irradiated (35 Gy) and plated 24 hours before use at 5 × 10<sup>4</sup> cells/10 cm dish.

Tyrosyl DNA phosphodiesterase assays

Whole-cell extract (WCE) was prepared by resuspension of 1-BR or 850-BR primary fibroblast cell pellets (1×10<sup>6</sup> cells) in 100μL lysis buffer (40 mM Tris/HCl pH 7.5, 100 mM NaCl, 0.1% Tween-20, 1 mM DTT, 1 mM PMSF, 1x EDTA free protease cocktail inhibitor), followed by 30 minutes of incubation on ice and mild sonication. The WCE was clarified by centrifugation for 10 minutes at 4°C at 16000g in a microfuge and the protein concentration quantified using the bicinchoninic acid (BCA) assay reagent (ThermoFisher).

Figure 1 *TDP2* splice site mutation in an individual from the United States with SCAR23



(A) Pedigree analysis. White symbols: wild-type *TDP2*. Black symbols: homozygosity for the *TDP2* splice site mutation c.425+1G>A (IVS3+1G>A) in the proband ("1"). Dotted symbols, heterozygosity for c.425+1G>A (IVS3+1G>A). (B) Sanger sequencing of the proband, demonstrating (left) the homozygous *TDP2* mutation c.425+1G>A (IVS3+1G>A) and (right) the absence of the *ZSCAN9* variant c914A>G. Patient sequences are shown at the bottom, and reference sequences are shown at the top. The 2 nucleotides relevant to the mutations are underlined. (C) Pathologic features of the Irish and US patients. (D) Haplotypes of Irish patients (51670; Irish pedigree) and the current patient from the United States (1583786) carrying the homozygous c.425+1G>A variant in *TDP2*. Only the variants bordering and directly within the homozygous regions and with a minor allele frequency of 5% are shown. Variants are indicated by their accession number in the dbSNP database, and the position of c.425+1G>A is in bold. The minimal overlapping region is delimited by rs9396886 and rs749338 and is 15.1 Mb in size. Black bars represent the homozygous haplotype as inferred from exome sequencing data. *TDP2* = tyrosyl DNA phosphodiesterase 2.

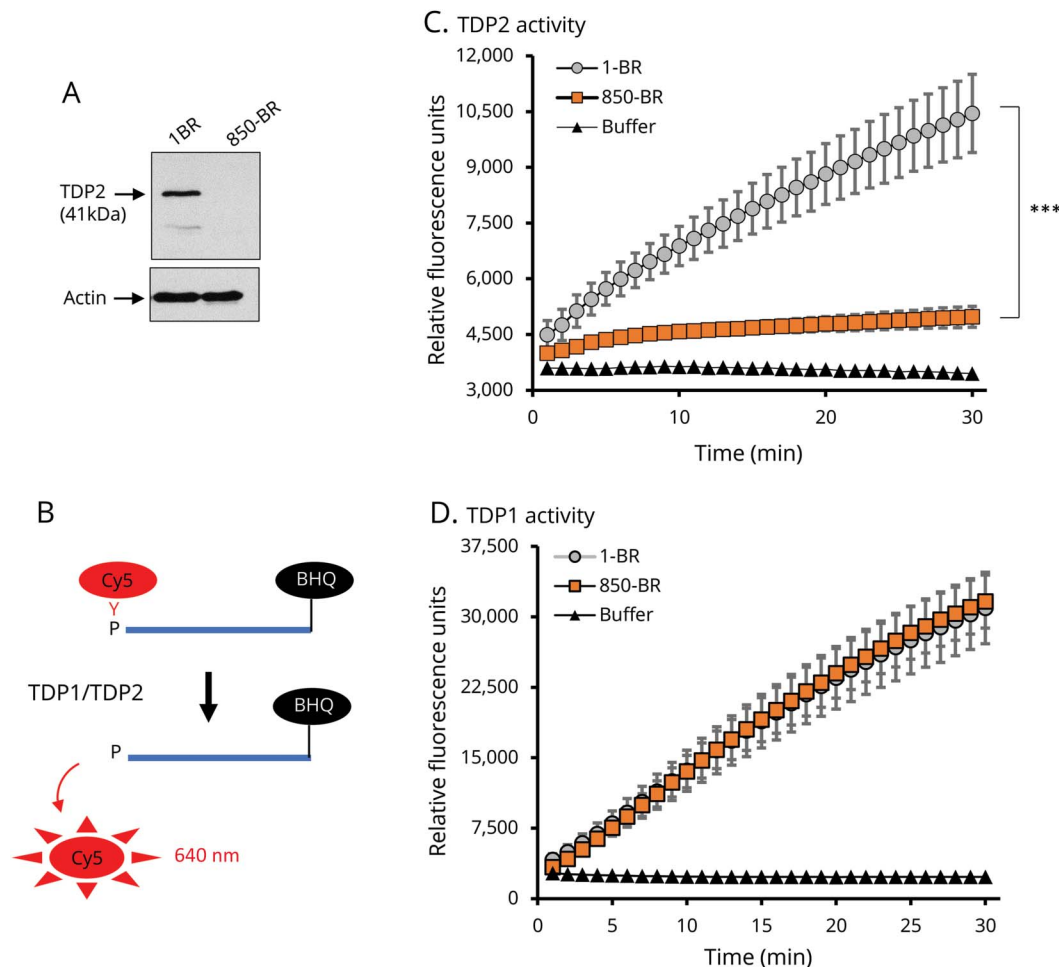
Clarified WCE (15µg total protein) was incubated with 40 nM TDP2 substrate (Cy5-5'Tyrosine-ssDNA<sub>19</sub>-BHQ) or TDP1 substrate (BHQ-ssDNA<sub>13</sub>-3'Tyrosine-Cy5) in reaction buffer (50 mM Tris/HCl pH8.0, 10 mM MgCl<sub>2</sub>, 80 mM KCl, and 1 mM DTT, 0.05% Tween-20) in a total volume of 6µL at room temperature, and Cy5 fluorescence was measured at 640 nm at the indicated time intervals on a BMG PHERAstar plate reader.

## DSB repair assays

Cells were grown on coverslips until confluent and then treated for 30 minutes with 25µM etoposide or irradiated with x-rays (2 Gy). After treatment, cells were rinsed and fixed for 10 minutes in PBS containing 4% paraformaldehyde at the indicated time points. Cells were permeabilized (20 minutes in PBS-0.2% Triton X-100),

blocked (1 hour in PBS-5% BSA), and incubated with anti-γH2AX (Millipore, 05-636, 1:2500) and anti-CENP-F (Abcam, ab5, 1:2500) antibodies for 3 hours in PBS containing 5% BSA. Cells were then washed (3 × 5 minutes in PBS containing 0.1% Tween-20), incubated for 1h with the corresponding Alexa Fluor conjugated secondary antibody (1:1000, 5% BSA), and washed again as described earlier. Finally, cells were counterstained with DAPI (Sigma, Gillingham, UK) and mounted in VECTASHIELD (Vector Labs, Peterborough, UK). Images were acquired on an automated wide-field microscopy Olympus ScanR system (motorized IX83 microscope) with ScanR Image Acquisition and Analysis Software, 20×/0.45 (LUCPLFLN 20×PH) dry objectives, and Hamamatsu ORCA-R2 digital CCD camera C10600. For the analysis of G0/G1 cells in patient primary fibroblasts, only CENP-F-negative cells

**Figure 2** Greatly reduced *TDP2* protein and activity in 850-BR patient fibroblasts



(A) TDP2 and actin protein levels in WCE (20 µg total protein) from 1-BR normal and 850-BR patient primary fibroblasts, as measured by immunoblotting. (B) TDP1 and TDP2 biochemical assay. Hydrolysis of a 3'-phosphotyrosyl (TDP1 assay) or 5'-phosphotyrosyl (TDP2 assay) bond releases the associated Cy5 fluorophore and alleviates quenching by the BHQ located on the opposite oligonucleotide terminus, resulting in elevated fluorescence at 640 nm. (C) Real-time (0–30 minutes) measurements of 5'-tyrosyl DNA phosphodiesterase activity (TDP2 assay) in reaction buffer (as a negative control) or WCE (15 µg protein) from 1-BR normal or 850-BR patient primary fibroblasts, using the single-stranded oligonucleotide (40 nM) with a 5'-phosphotyrosyl-linked Cy5 fluorophore and 3'-BHQ. (D) Real-time measurements (0–30 minutes) of 3'-tyrosyl DNA phosphodiesterase (TDP1 assay) conducted as above but using a 3'-phosphotyrosyl-linked Cy5 fluorophore and 5'-BHQ. Data are mean of 3 independent experiments ±SEM. Statistical significance was determined by two way analysis of variance (ANOVA). \*\*\**p* < 0.001. BHQ = black hole quencher; TDP2 = tyrosyl DNA phosphodiesterase 2; WCE = whole-cell extract.



were scored. For A549 and A549 *TDP2*<sup>-/-</sup>, cells were gated to the G1 population according to the DAPI profile of ScanR Image Analysis Software. For patient fibroblast complementation experiments, cells were gated accordingly to eGFP expression on ScanR Analysis Software, in addition to the G1 population according to the DAPI content.

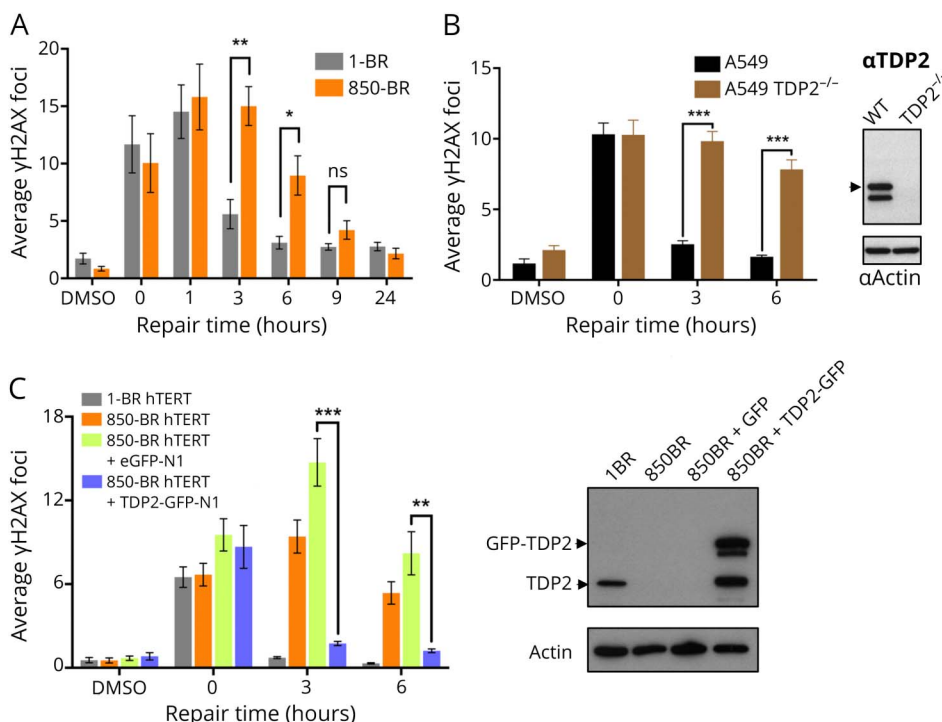
## Results

We recently described mutations in *TDP2* in three Irish patients from the same family with intellectual disability, seizures, and ataxia, a disease now denoted as spinocerebellar ataxia 23 (SCAR23)<sup>6</sup>. Here, we describe a 6-year-old patient in the United States with very similar pathology including developmental delay, epilepsy, and ataxia and in whom we identified by whole-exome and Sanger sequencing possesses the same homozygous splice site mutation in *TDP2* (c.425+1G>A) (figure 1, A–C). Whether there is a connection between the current patient and the Irish family is not clear, but a comparison of the WES data from the two families revealed that the two apparently unrelated affected individuals of whom their exome was sequenced share the same homozygous haplotype of ~15.1 Mb (figure 1D). It is important to note that in contrast to the Irish patients, however, the current patient lacks the mutation c.914A>G (p.His305Arg) in *ZNF193*, a zinc finger protein of unknown function (figure 1B, right). We previously concluded that this variant was not

disease causing in the Irish patients,<sup>6</sup> and the absence of this variant in the current patient confirms that this and also that *TDP2* mutation is the cause of SCAR23.

To enable additional molecular and cellular analyses of SCAR23, we generated primary fibroblasts (denoted 850-BR) from a skin biopsy kindly provided by the current patient. Western blotting failed to detect *TDP2* protein in the 850-BR patient fibroblasts (figure 2A), consistent with the lack of detectable *TDP2* protein previously reported in lymphoblastoid cells from the Irish patients.<sup>6</sup> The absence of detectable *TDP2* protein in the lymphoblastoid cell lines was explained by the impact of the *TDP2* mutation on splicing, which resulted in nonsense-mediated decay and greatly reduced (<20%) levels of *TDP2* mRNA.<sup>6</sup> To confirm the impact of the splice site mutation on *TDP2* activity, we used a highly sensitive tyrosyl DNA phosphodiesterase biochemical assay. This assay uses a single-stranded oligonucleotide substrate in which hydrolytic release of a Cy5-labeled tyrosine moiety present on one terminus of the oligonucleotide results in increased fluorescence due to evasion of a black hole quencher present on the opposite terminus (figure 2B). By situating the fluorescent-labeled tyrosine on either the 3' or 5' terminus, this assay can detect *TDP1* or *TDP2* activity, respectively. Although whole-cell extract from normal 1-BR human fibroblasts exhibited robust 5'-tyrosyl DNA phosphodiesterase activity, 850-BR patient fibroblasts did not (figure 2C). The small amount of activity observed in

**Figure 3** Reduced DSB repair in *TDP2*-mutated patient fibroblasts and A549 cells after topoisomerase 2-induced DNA damage



(A) DSBs were measured by γH2AX immunostaining in normal 1-BR and patient 850-BR primary fibroblasts before and after treatment for 30 minutes with DMSO vehicle or 25 μM etoposide ("0"), followed by subsequent incubation in a drug-free medium for the indicated repair periods. (B) DSBs were measured as above, in wild-type A549 cells and in A549 cells in which *TDP2* was deleted by CRISPR/Cas9 gene editing (*TDP2*<sup>-/-</sup> A549). The level of *TDP2* and actin (loading control) in the cell lines used for these experiments is shown by Western blotting on the right. (C) DSBs were measured as above in hTERT-immortalized 1-BR fibroblasts and 850-BR fibroblasts and in the latter cells after transfection with the empty GFP vector or vector encoding recombinant human *TDP2*-GFP. The level of *TDP2* and actin (loading control) in the cell lines used for these experiments is shown by Western blotting on the right. Data are mean ± SEM of 3 independent experiments, and statistically significant differences were determined by the *t* test (ns = not significant, \**p* < 0.05; \*\**p* < 0.01; \*\*\**p* < 0.001). DSB = double-strand break; GFP = green fluorescent protein; *TDP2* = tyrosyl DNA phosphodiesterase 2.

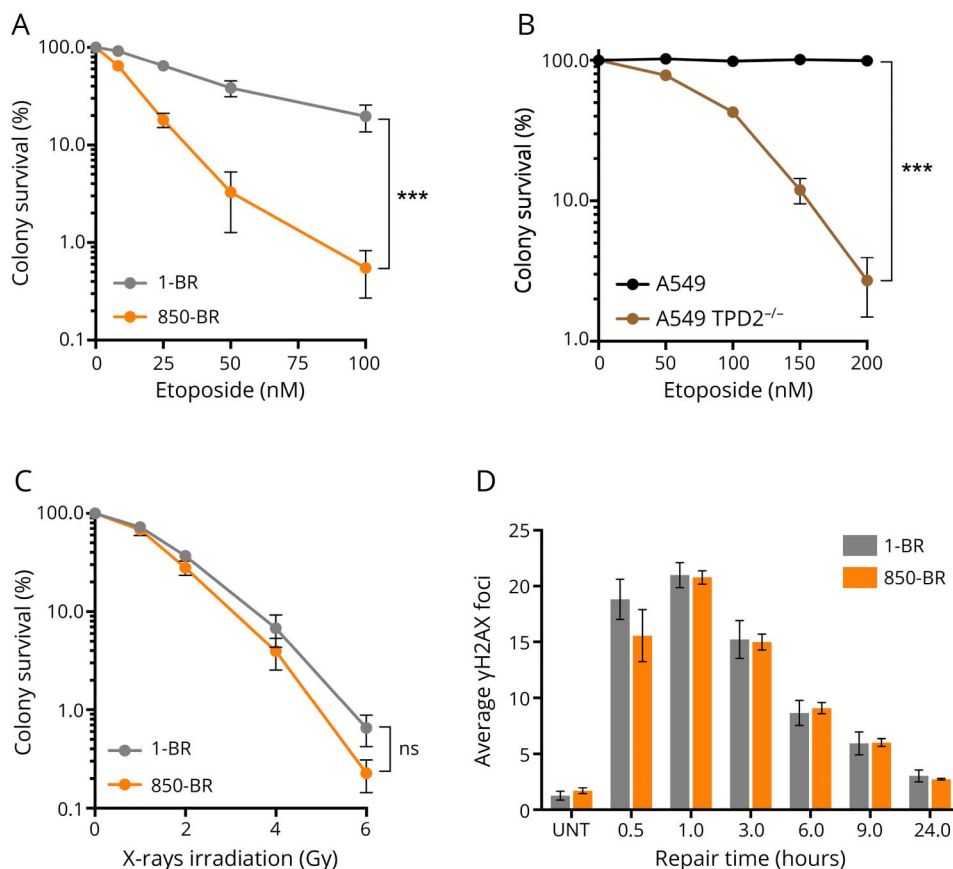
850-BR cell extract, compared with reactions supplemented with reaction buffer alone, reflects nucleases in the cell extract rather than residual TDP2 activity because similar results were observed when we used cell extract from human cells in which TDP2 was deleted by CRISPR/Cas9 (unpublished observations). The lack of TDP2 activity in patient cells did not reflect differences in the technical quality of cell extracts because the level of TDP1 activity in these extracts was similar to that in normal human 1-BR fibroblasts (figure 2D). We conclude from these data that the *TDP2* splice site mutation greatly reduces and most likely ablates both TDP2 protein and activity in 850-BR patient fibroblasts.

Next, we addressed the impact of the TDP2 splice site mutation on nuclear DSB repair rates in 850-BR patient fibroblasts using immunofluorescent detection of  $\gamma$ H2AX as an indirect measure of DSBs.<sup>12</sup> Although similar levels of nuclear  $\gamma$ H2AX foci were present in normal and patient fibroblasts immediately after treatment with etoposide for 30 minutes, the levels of these  $\gamma$ H2AX foci decreased far more slowly in the patient fibroblasts during subsequent incubation in a drug-free medium (figure 3A). However, DSB repair was completed in patient cells within 8–24 hours, consistent with the established existence of alternative, nuclease-dependent

mechanisms for repair of TOP2-induced DSBs in human cells.<sup>4,13,14</sup> Similar results were observed in human A549 cells in which *TDP2* was mutated by CRISPR/Cas9 gene editing, confirming the importance of TDP2 for repair of TOP2-induced DSBs (figure 3B). More importantly, complementation of 850-BR cells with expression construct encoding recombinant human TDP2 restored normal rates of nuclear DSB repair, confirming that the DNA repair defect in the patient fibroblasts was the result of the *TDP2* mutation (figure 3C). This defect in DSB repair was accompanied by cellular hypersensitivity to TOP2-induced DSBs because both 850-BR patient fibroblasts and *TDP2*<sup>-/-</sup> A549 cells were hypersensitive to etoposide in clonogenic survival assays (figure 4, A and B). In contrast to etoposide-induced DSBs, both DSB repair rates and levels of cell survival were normal in 850-BR patient fibroblasts after treatment with ionizing radiation (figure 4, C and D), confirming the specificity of the DNA repair defect in the patient's cells for DNA breaks induced by TOP2.

Finally, because analyses of muscle biopsy from the patient (see the case report) suggested the presence of possible defects in the ETC, we examined 850-BR patient fibroblasts for defects in mitochondrial function. However, the patient

**Figure 4** Hypersensitivity of TDP2-mutated patient fibroblasts and A549 cells to DNA damage induced by etoposide but not  $\gamma$ -rays

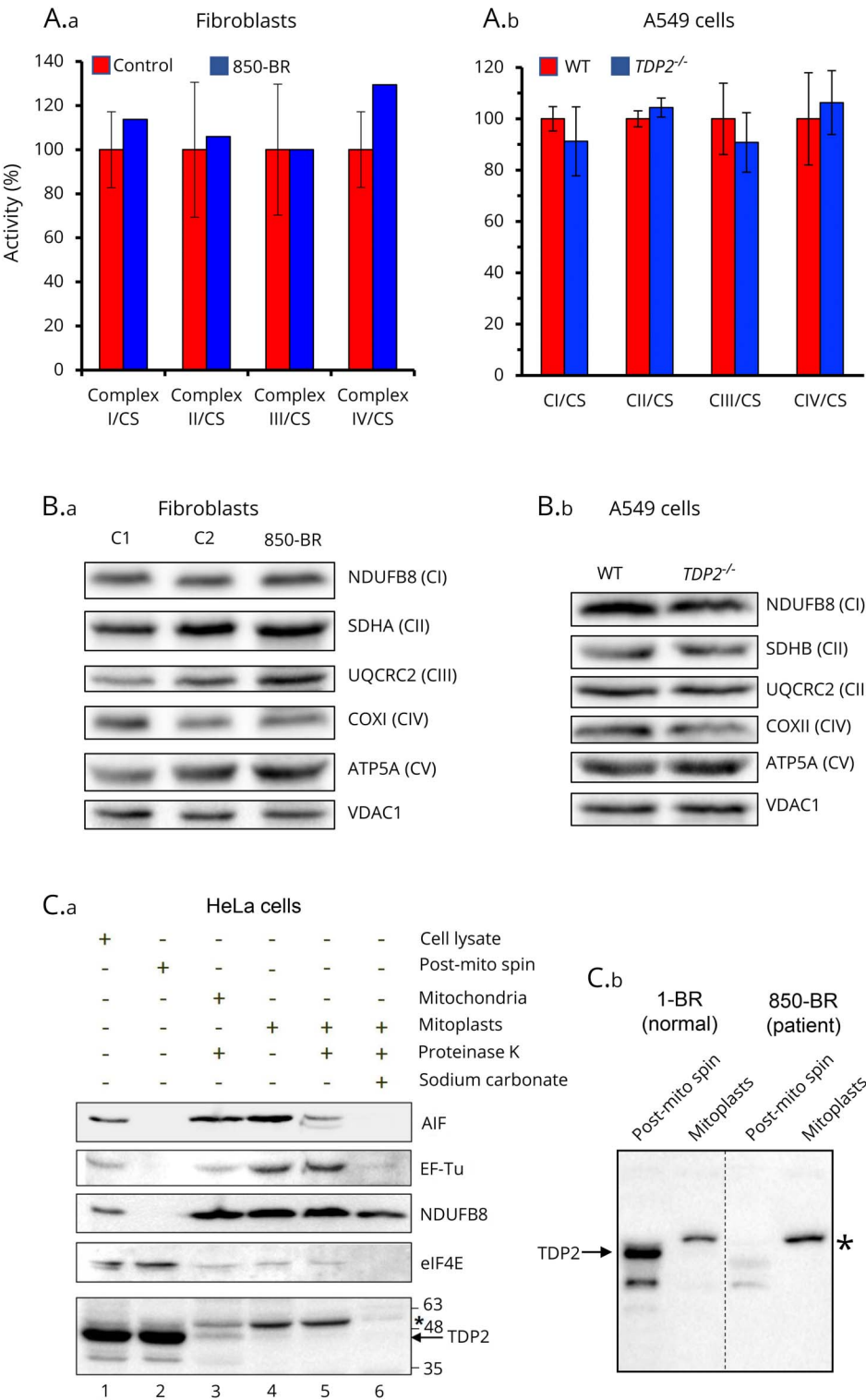


(A) Clonogenic survival of 1-BR normal and 850-BR patient primary fibroblasts in a medium containing the indicated concentrations of etoposide. (B) Clonogenic survival of wild-type A549 and *TDP2*<sup>-/-</sup> A549 cells in a medium containing the indicated concentrations of etoposide. (C) Clonogenic survival of 1-BR normal and 850-BR patient primary fibroblasts as above, following  $\gamma$ -irradiation. (D), DSBs were measured by  $\gamma$ H2AX immunostaining in normal 1-BR and patient 850-BR primary fibroblasts before and after the indicated periods after  $\gamma$ -irradiation (2 Gy). Data are mean  $\pm$  SEM of 3 independent experiments, and statistically significant differences were determined by two-way ANOVA (ns = not significant, \*\*\* $p$  < 0.001). TDP2 = tyrosyl DNA phosphodiesterase 2.

fibroblasts failed to exhibit significant defects in respiratory chain complexes in biochemical assays compared with a range of different normal (control) human fibroblasts (figure 5A.a). Similar results were observed when we compared wild-type A549 cells with A549 cells in which

*TDP2* was mutated by gene editing (figure 5A.b). We also failed to identify any impact of the *TDP2* mutation or deletion in 850-BR fibroblasts and A549 cells, respectively, in the level of respiratory chain proteins as measured by immunoblotting (figure 5B).

**Figure 5** Mitochondrial functionality in normal and *TDP2*-mutated cells



(A) Activity of mitochondrial respiratory complexes in normal (red bars) and patient 850-BR (blue bars) primary human fibroblasts (A.a) and in wild-type human A549 (red bars) and *TDP2*<sup>-/-</sup> A549 cells (blue bars) (A.b) was determined as previously described.<sup>15</sup> The mean enzyme activities in control cells (n = 8) are set to 100%, and error bars represent SD. (B) Levels of mitochondrial respiratory complex proteins in age-matched control fibroblasts (C1 and C2) and patient 850-BR fibroblasts (B.a) and in wild-type A549 cells and *TDP2*<sup>-/-</sup> A549 cells (B.b), as measured by immunoblotting for subunits of CI (NDUFB8), CII (SDHA), CIII (UQCRC2), CIV (COXI), and CV (ATP5A) and using VDAC1 as a mitochondrial loading control. (C) Subcellular localization of *TDP2* in HeLa cells (C.a) and in normal (1-BR) and patient (850-BR) primary human fibroblasts (C.b). Left, cell-equivalent amounts of HeLa total cell lysate (30 g total protein; lane 1), cell lysate depleted of mitochondria ("post-mito spin"; lane 2), mitochondria treated with proteinase K (to remove proteins associated with the outer membrane; lane 3), mitoplasts (mitochondrial matrix plus inner membrane proteins; lane 4), mitoplasts treated with proteinase K (lane 5), and inner membrane mitochondrial proteins (extracted with sodium carbonate; lane 6) were immunoblotted for *TDP2* and for protein markers of the cytosol (eIF4E) and each of the mitochondrial compartments, intermembrane space (AIF), mitochondrial matrix (EF-Tu), and inner mitochondrial membrane (NDUFB8). Right, cell-equivalent amounts of 1-BR or 850-BR fibroblast total cell lysates (30µg protein) depleted of mitochondria ("post-mito spin") and of mitoplasts were immunoblotted for *TDP2* as above. The position of full-length *TDP2* (arrow) and a nonspecific band detected by the antibody (asterisk; \*) are indicated. *TDP2* = tyrosyl DNA phosphodiesterase 2.

We next fractionated HeLa cells and both 1-BR and 850-BR fibroblasts into subcellular components, including intact mitochondria and mitoplasts, to examine whether TDP2 is located within the mitochondria and thus in the correct location to repair mitochondrial DNA (figure 5C). Most of the material detected by the anti-TDP2 antibody and which migrated to the position expected for TDP2 was present, as expected, in the cellular fractions containing cytosolic and nuclear proteins (figure 5C.a, lanes 1 and 2). However, a very small amount of this material was detected in intact mitochondria and mitoplasts (figure 5C.a, lanes 3–5). In 1-BR normal human fibroblasts, most of the material detected by anti-TDP2 antibody was again present in the cytosolic/nuclear fraction, and this material was TDP2 because it was absent from parallel preparations from patient 850-BR fibroblasts (figure 5C.b). However, if TDP2 was present in mitoplasts prepared from 1-BR cells, it was below the level of detection in our experiments (figure 5C.b, arrow). We did note the presence of a slower migrating band that was detected by the TDP2 antibody in both HeLa and 1-BR mitoplasts, but this band was also present in parallel preparations from 850-BR patient fibroblasts, indicating that it was not TDP2 (figure 5C.b, asterisk). Collectively, although we cannot rule out the presence of a small amount of endogenous TDP2 in mitochondria, we cannot detect defects in mitochondrial function in SCAR23 patient fibroblasts. Consequently, we suggest that the neurologic disease pathology in SCAR23 is most likely the result of a defect in nuclear DSB repair.

## Discussion

We recently identified *TDP2* mutations in a recessive hereditary genetic disorder associated with intellectual disability, seizures, and ataxia, now denoted as spinocerebellar ataxia autosomal recessive 23 (SCAR23)<sup>6</sup>. In the three Irish siblings originally described, we also identified the mutation c.914AG (p.His305Arg) in *ZNF193*, a zinc finger protein of unknown function, which we concluded was not disease causing but which we could not exclude as a contributor to the disease.<sup>6</sup> However, in contrast to the Irish patients, the new patient identified in the United States and reported here harbors the same *TDP2*-associated haplotype but lacks the mutation in *ZNF193*, ruling out a contribution of the latter and confirming the *TDP2* mutation as the cause of SCAR23.

The *TDP2*-mutated patient fibroblasts established here are the first such reported and have provided valuable information concerning the molecular and cellular defect in SCAR23. Consistent with *TDP2* defects in other cell types, we observed reduced rates of DSB repair and elevated cellular sensitivity after treatment with the TOP2 poison etoposide. These phenotypes were the result of the *TDP2* mutation in the patient fibroblasts because they were phenocopied in *TDP2*<sup>-/-</sup> A549 cells created by gene editing and were complemented by the reintroduction of the wild-type human *TDP2* transgene. Of interest, the magnitude of these phenotypes is greater in patient

fibroblasts and gene-edited A549 cells than what we have seen previously with our previous gene-edited cell lines.<sup>9</sup> We believe that this is because the cell lines reported here are effectively *TDP2* null, whereas our previously generated *TDP2* gene-edited human cells retained a minor isoform of *TDP2* that is expressed at low levels and is primarily cytoplasmic.

It is notable that TOP2 poisons are used widely in the clinic to treat a variety of cancers, and we point out that the use of these agents in patients with *TDP2* mutations should be avoided or at the very least treated with extreme caution. Although *TDP2*-defective cancer cells will be hypersensitive to this type of chemotherapy, normal cells will also be hypersensitive, as illustrated in the current work by the etoposide sensitivity observed in both *TDP2*<sup>-/-</sup> A549 lung cancer cells and SCAR23 primary fibroblasts.

The presence of mitochondrial dysfunction in the SCAR23 patient described here is suggested by ETC studies and the high lactate and lactate/pyruvate ratio in muscle (see the Patient Case Study in the Methods section). In addition, the current patient exhibits phenotypes consistent with mitochondrial dysfunction such as hypotonia, low energy, and fatigability. It is unclear whether phenotypes consistent with mitochondrial defects are also present in the Irish patients, although none were originally noted.<sup>6</sup> Unfortunately, neither fibroblasts nor muscle biopsy are available from these patients to address this possibility. The lack of detectable mitochondrial defect in primary skin fibroblasts from the current patient and in gene-edited A549 cells does not support a major role for *TDP2* in the repair of mitochondrial DNA. However, the absence of mitochondrial defects in skin fibroblasts from patients with mitochondrial defects in muscle is not uncommon.<sup>15</sup> Nevertheless, given the dramatic impact of the *TDP2* mutation on nuclear DNA repair, we suggest that any association of SCAR23 with mitochondrial disease is most likely an indirect or secondary dysfunction resulting from a nuclear disorder.<sup>16</sup>

## Author contributions

G. Zagnoli-Vieira, F. Bruni, K. Thompson, and L. He designed and conducted mitochondrial experiments. S. Walker designed TDP assay. R. Taylor supervised mitochondrial experiments. A.P.M. de Brouwer conducted haplotype analysis. D. Niyazov identified and assessed the patient and identified the *TDP2* mutation. K.W. Caldecott designed, directed, and coordinated the project and wrote the manuscript.

## Acknowledgment

The authors thank the patient and the patient's family, and Limei Ju/Elena Korneeva for generation of primary 850-BR patient fibroblasts and hTERT-immortalized derivatives.

## Study funding

This study was funded by Programme grants to KWC from the MRC (MR/P010121/1), CRUK (C6563/A16771), ERC (SIDSCA Advanced Grant, 694996), and by an MRC PhD Studentship for GZV (MRN/N50189X/1).



## Disclosure

G. Zagnoli-Vieira has no financial disclosures to report. F. Bruni serves on the editorial boards of *Biomarkers* journal and *Brain and Nervous System Current Research* and receives research support from ANVUR. K. Thompson, L. He, S. Walker, R. Taylor, and A.P.M de Brouwer report no disclosures. D. Niyazov has served on the scientific advisory boards of Alexion Pharmaceuticals and Malinckrodt ARD Inc.; has received funding for travel/speaker honoraria from Sanofi Genzyme Corporation; serves on the editorial board of *Molecular Syndromology*; has consulted with the Phase II clinical trial of Elamipretide in Mitochondrial Myopathy; is a current member of the Sanofi Genzyme Corporation Speakers' Bureau; and has served as an expert witness for a legal case regarding vaccine injury in a child with mitochondrial disease. K.W. Caldecott reports no disclosures. Full disclosure form information provided by the authors is available with the full text of this article at [Neurology.org/NG](http://Neurology.org/NG).

Received December 20, 2017. Accepted in final form May 24, 2018.

## References

1. Lindahl T. Instability and decay of the primary structure of DNA. *Nature* 1993;362:709–715.
2. Goodarzi AA, Jeggo PA. The repair and signaling responses to DNA double-strand breaks. *Adv Genet*. Elsevier 2013;82:1–45.
3. Cortes Ledesma F, El-Khamisy SF, Zuma MC, Osborn K, Caldecott KW. A human 5'-tyrosyl DNA phosphodiesterase that repairs topoisomerase-mediated DNA damage. *Nature* 2009;461:674–678.
4. Gómez-Herreros F, Romero-Granados R, Zeng Z, et al. TDP2-Dependent non-homologous end-joining protects against topoisomerase II-induced DNA breaks and genome instability in cells and. *Vivo Plos Genet* 2013;9:e1003226.
5. Zeng Z, Cortes Ledesma F, El-Khamisy SF, Caldecott KW. TDP2/TTRAP is the major 5'-tyrosyl DNA phosphodiesterase activity in vertebrate cells and is critical for cellular resistance to topoisomerase II-induced DNA damage. *J Biol Chem* 2011;286:403–409.
6. Gómez-Herreros F, Schuurs-Hoeijmakers JHM, McCormack M, et al. TDP2 protects transcription from abortive topoisomerase activity and is required for normal neural function. *Nat Genet* 2014;46:516–521.
7. Bruni F, Gramegna P, Oliveira JMA, Lightowlers RN, Chrzanowska-Lightowlers ZMA. REXO2 is an oligoribonuclease active in human mitochondria. *PLoS One* 2013;8:e64670.
8. Hoch NC, Hanzlikova H, Rulten SL, et al. XRCC1 mutation is associated with PARP1 hyperactivation and cerebellar ataxia. *Nature* 2016;541:87–91.
9. Gómez-Herreros F, Zagnoli-Vieira G, Ntai I, et al. TDP2 suppresses chromosomal translocations induced by DNA topoisomerase II during gene transcription. *Nat Commun* 2017;8:233.
10. Thomson G, Watson A, Caldecott K, et al. Generation of assays and antibodies to facilitate the study of human 5'-tyrosyl DNA phosphodiesterase. *Anal Biochem* 2013;436:145–150.
11. Thompson K, Majd H, Dallabona C, et al. Recurrent de novo dominant mutations in SLC25A4 cause severe early-onset mitochondrial disease and loss of mitochondrial DNA copy number. *Am J Hum Genet* 2016;99:860–876.
12. Rogakou EP, Pilch DR, Orr AH, Ivanova VS, Bonner WM. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J Biol Chem* 1998;273:5858–5868.
13. Hoa NN, Shimizu T, Zhou Z-W, et al. Mre11 Is Essential for the Removal of Lethal Topoisomerase 2 Covalent Cleavage Complexes. *Mol Cell* 2016;64:580–592.
14. Zagnoli-Vieira G, Caldecott KW. TDP2, TOP2, and SUMO: what is ZATT about? *Cell Res* 2017;17:182–1406.
15. Kirby DM, Thorburn DR, Turnbull DM, Taylor RW. Biochemical assays of respiratory chain complex activity. *Methods Cell Biol* 2007;80:93–119.
16. Niyazov DM, Kahler SG, Frye RE. Primary mitochondrial disease and secondary mitochondrial dysfunction: importance of distinction for diagnosis and treatment. *Mol Syndromol* 2016;7:122–137.