Abstract

Radiotherapy plays an important role in the curative management of cancer. Despite technical developments in treatment delivery, it is still associated with potentially severe side effects due to damage caused to normal tissues surrounding the tumour. Targeting tumour specific determinants of radiosensitivity has the potential to improve the therapeutic index. The identification of novel determinants of tumour intrinsic radiosensitivity is therefore a key-step to develop innovative therapies.

In this study, two high-throughput radiosensitivity screens using different functional genomics approaches are presented. The first one, an arrayed RNAi-based radiosensitivity screen, was used to investigate a subset of about 10,000 'druggable' genes in order to identify novel potential clinical targets. The second one, a pooled loss-of-function CRISPR radiosensitivity screen, made use of a genome-scale approach and therefore interrogated over 20,000 genes in the context of radiation response.

Interestingly, TRIM37 was identified from both screens as a novel gene whose depletion causes radiation sensitivity. TRIM37 (tripartite motif containing 37) is an E3 ubiquitin ligase belonging to the TRIM family proteins. TRIM37 germline mutations are associated with a rare autosomal recessive disorder known as MULIBREY nanism. Some of the clinical features are growth failure, infertility and higher incidence of cancer. TRIM37 has a known oncogenic function in somatic cells and it is overexpressed in a subset of tumours, where high expression of TRIM37 is correlated with lower survival rates. The depletion of TRIM37 as a modulator of radiosensitivity was validated in a wide panel of cancer cell lines. Also, the mechanism underlying TRIM37's role in the response to ionising radiation was investigated. As for the pleiotropic nature of E3 ubiquitin ligases, it was found that TRIM37 depletion impacts on different aspects of ionising radiation response, but mainly, on cell cycle regulation. The exact mechanism by which TRIM37 depletion modifies cellular response to radiation is yet to be identified. However, this study shows that cells lacking TRIM37 expression have increased genomic instability independently from radiation treatment, providing evidences in agreement with MULIBREY disease phenotype. Evidences of genomic instability were amplified in presence of IR-induced DNA damage.

Taken together with the observed overexpression of TRIM37 in cancer cells, our findings suggest that TRIM37 is a potential target to modulate tumour radiosensitivity.

Giovanna Granata, St Catherine's College

Submitted for the degree of DPhil, University of Oxford, Trinity Term, 2019

To Daniela Mignone,

Vittorio Granata,

Luigi Agnano.

Acknowledgements

First and foremost, I would like to thank my supervisors, Professor Geoff Higgins and Professor Gillies McKenna, for allowing me to carry out my doctoral research in their laboratory. It has been a great opportunity for me to work for them.

I am particularly thankful to Dr Gonzalo Rodriguez Berriguete who supported me during these years with several inputs and helpful discussions about my project, but also about science in general. A special thank you to Dr Remko Prevo who helped me to settle in the new lab and department and taught me how to be an organised scientist; without his help, this thesis would have not been possible. I would also like to thank the other lab members Dr Katharine Herbert, Ms Rathi Puliyadi for the never-ending patience they had with me, and for always been smiley and cheerful people. I would also like to wish the best of luck for their careers to my DPhil colleagues James Coates, Tzveta Pokrovska and Karen Sayal, you made my DPhil path less miserable and became much more than just course mates to me.

This thesis would not have been possible without the collaboration of many other scientists in and outside of the department. A special thank you to Dr Vincenzo D'Angiolella and Professor Nils Cordes, the examiners of my thesis, for their valuable feedback on my work, to Professor Francesca Buffa for teaching me the CRISPR screening data analysis, to Dr Daniel Ebner for the screening facilities as well as advice, to Dr Dylan Jones for helping me with the performance of the screening, and thank you to all those who challenged me scientifically and gave me the ideas that led to the results presented in this thesis.

I am very thankful to Marie Skłodowska-Curie actions (MSCA), for both the economic support to my work and my scholarship and for giving me the unique opportunity of being part of the prestigious ITN EU RADIATE network, I have learnt so much from every meeting with both senior and young members of this group. Thank you to Cancer Research UK and St Catherine's College for the economical contribution to my scholarship.

I would like to thank my nearest and dearest friends. The ones I made here in Oxford, who became like a family away from home. A massive thank you to

the ‘CS’ group composed by Angela, Andrea, Mattia, Martin and Virginia, because of you, I have never felt alone here in Oxford. In particular, I would to thank Ioanna for her limitless psychological and scientific support during the darkest times of my Oxford life. Thank you, Angela, for walking every step of this path with and next to me, for sharing with me joys and sorrows, for cheering me up, for being my buddy. And thank you, Andrea, because you were incredibly able to make the stressful time of the thesis writing enjoyable. Also, I would like to thank my dearest friends which were able to be close to me despite the physical distance separating us, Imma, Rita, Fiorella and Laura.

Finally, I would like to thank my family, because you have always believed in me even when I gave you many reasons to not do so. My achievements would not have been possible without your endless and untiring support.

Table of Contents

CHAPTER 1 GENERAL INTRODUCTION 7

1.1 Cancer and radiotherapy 8

1.2 Irradiation-induced DNA-damage response 9

1.2.1 DNA damage by ionising radiation 9

1.2.2 The DNA damage response in mammalian cells 11

1.2.3 Non-homologous end joining repair 13

1.2.4 Homologous recombination repair 14

1.2.5 Cell-cycle checkpoints 16

1.3 Intrinsic radiosensitivity 19

1.3.1 The linear-quadratic (LQ) model 20

1.4 High-throughput functional genomics screenings 21

1.4.1 siRNA high-throughput screenings to identify novel radiosensitisers 21

1.4.2 CRISPR/Cas9 genome mutagenesis 22

1.4.3 LOF CRISPR/Cas9 genome-wide pooled screenings to identify novel determinants of radiosensitivity 24

1.5 Aims of this thesis 26

CHAPTER 2 MATERIAL AND METHODS 28

2.1 Cell culture 29

2.1.1 Cell line maintenance 29

2.1.2 Cell line storage and resuscitation 29

2.1.3 Passaging cell lines 30

2.1.4 Cell counting 30

2.1.5 Cell synchronisation 30

2.2 Plasmid transfection 30

2.3 Reverse siRNA transfection 31

2.4 Colony formation assay 31

2.4.1 Statistical analysis 32

2.5 Western blotting 32

2.5.1 Preparation of protein lysates 32

2.5.2 SDS-PAGE and protein transfer 33

2.5.3 Blocking, antibody incubation and signal detection..... 33

2.6 Immunofluorescence analysis 34

2.7 Flow cytometry 35

2.7.1 Cell cycle analysis by propidium iodide staining 35

2.7.2 BrdU/Propidium iodide staining..... 35

2.8 I-SceI-based reporter assays 36

2.9 Genome-wide CRISPR/Cas9 Screen..... 36

2.9.1 Cell transduction 36

2.9.2 Timeline of transduction procedure..... 38

2.9.3 Cell treatments..... 39

2.9.4 Genomic DNA extraction..... 39

2.9.5 PCR..... 39

2.10 Statistical Analysis..... 41

2.10.1 Quantification of Sensitisation..... 41

CHAPTER 3 IDENTIFICATION AND VALIDATION OF NOVEL DETERMINANTS OF RADIOSENSITIVITY THROUGH HIGH-THROUGHPUT RNAI SCREEN 42

3.1 Introduction..... 43

3.2 Results 46

3.2.1 Gene selection and identification of novel determinants of radiosensitivity 46

3.2.2 Validation of TPR as a novel radiosensitiser..... 54

3.2.3 Validation of TRIM37 as a novel radiosensitiser..... 58

3.3 Discussion 63

CHAPTER 4 IDENTIFICATION OF NOVEL DETERMINANTS OF RADIOSENSITIVITY THROUGH A GENOME-WIDE CRISPR KNOCKOUT SCREEN 68

4.1 Introduction..... 69

4.2 Results 72

4.2.1 Data analysis..... 72

4.2.2 Identification of novel determinants of radiosensitivity and gene selection..... 75

4.3 Discussion 84

CHAPTER 5 IONISING RADIATION RESPONSE IN TRIM37-DEPLETED CELLS 88

5.1 Introduction 89

5.2 Results 94

5.2.1 Analysis of DNA damage response in TRIM37-depleted cells 94

5.2.2 Analysis of cell cycle progression after TRIM37 depletion 101

5.2.3 Genomic instability analysis in TRIM37 depleted cells 109

5.3 Discussion 113

CHAPTER 6 GENERAL DISCUSSION 117

6.1 Summary of findings 118

6.2 Significance of findings 120

6.3 Concluding remarks 123

List of abbreviations and acronyms

~	circa
53BP1	P53-binding protein 1
ATCC	American Type Culture Collection
ATM	Ataxia telangiectasia mutated
ATR	Ataxia telangiectasia and Rad3-related protein
BrdU	Bromodeoxyuridine
BSA	Bovine serum albumin
°C	Celsius degrees
Cdc2	Cyclin-dependent kinase 1 (CDK1)
CDK	Cyclin-dependent kinase
CFA	Colony formation assay
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
C_{score}	CRISPR Score
DAPI	4',6-diamidino-2-phenylindole
DBD	DNA-binding domain
DDR	DNA-damage response
DMEM	Dulbecco's modified Eagle's medium

DNA	Deoxyribonucleic acid
DNA-PK	Complex formed of DNA-PKcs and Ku70-Ku80
DNA-PKcs	DNA-dependent protein kinase catalytic subunit
DSB	Double-strand break
dsRNA	Double stranded RNA
DTT	Dithiothreitol
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
FBS	Foetal bovine serum
FDR	False discovery rate
FITC	Fluorescein isothiocyanate FU Follow-up
g	Earth's gravitational force (= 9.807 m/s2)
G1	Gap 1 phase of the cell cycle
G2	Gap 2 phase of the cell cycle
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
gRNA	Guide RNA
Gy	Gray (defined as the absorption of one joule of radiation energy per kilogram of matter)
h	Hour
H2AX	H2A histone family, member X

HDR	Homology-directed repair
HR	Homologous recombination
HRP	Horseradish peroxidase
HTS	High-throughput screen
IF	Immunofluorescence
IR	Ionising radiation
kb	Kilobase
KD	Knockdown
kDa	Kilodalton
KO	Knockout
Ku70	Lupus Ku autoantigen protein p70
Ku80	Lupus Ku autoantigen protein p80
Lig4	DNA Ligase IV
LOF	Loss-of-function
M-phase	Mitotic phase
mA	Milliampere
MDM2	Mouse double minute 2 homologue
min	Minute
MOI	Multiplicity of infection
MRN	MRE11/RAD50/NSB1

mRNA	Messenger ribonucleic acid
NBS1	Nibrin
NCI	National Cancer Institute
NDM	Nonsense-mediated decay
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells ng Nanograms
NGS	Next generation sequencing
NHEJ	Non-homologous end joining
NT	Non-targeting
nt	Nucleotide
p21	CDKN1A also known as p21waf1/cip1
PAM	Protospacer-adjacent motif
PARP-1	Poly [ADP-ribose] polymerase 1
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PE	Plating efficiency
Pen	Penicillin
PFA	Paraformaldehyde
PFP	Percentage of false positives
PI	Propidium iodide

PRKDC	DNA-PKcs
psi+	Psi packaging signal
Puro	puromycin selection marker
qPCR	Real-time polymerase chain reaction
Rb	Retinoblastoma protein
RIPA	radioimmunoprecipitation assay
RNAi	RNA interference
ROS	Reactive oxygen species
RPMI	Roswell Park Memorial Institute
RRE	Rev response element
RT	Room temperature
S-phase	S phase of the cell cycle
SD	Standard deviation
SDS-Page	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SER10	Survival enhancement ratio at 0.10 surviving fraction
SF	Surviving fraction
siNT	Non-targeting siRNA silencing
siRNA	Small-interfering RNA
siTPR	TPR siRNA silencing

siTRIM37	TRIM37 siRNA silencing
SSB	Single strand break
TPR	Translocated promoter region
TRIM37	Tripartite motif containing 37
Tris	Tris-hydroxy methylamino methane
TUBA1C	Tubulin alpha-1C chain
TUBA8	Tubulin alpha-8 chain
Tween20	Polyethylene glycol sorbitan monolaurate
WPRE	Post-transcriptional regulatory element
XRCC4	X-ray repair cross-complementing protein 4
γ-H2AX	Phosphorylated-H2AX histone
μ	Mean
σ	Standard deviation

Chapter 1 GENERAL INTRODUCTION

1.1 Cancer and radiotherapy

Cancer is a group of diseases characterised by uncontrolled growth of abnormal cells, which will eventually spread to surrounding body parts and/or other organs [1]. In United Kingdom, it is estimated that 1 every 2 people will develop cancer within their lifetime [2]. Cancer is the second leading cause of death globally, accounting 17 millions of new cases and 9.6 millions death in 2018 [1], [3].

Cancer disease management consists of different therapeutic strategies including surgery, chemotherapy and radiotherapy, which can be used as single or combined approaches. Treatment options may vary according to the tumour type, stage and available resources [1].

Radiotherapy is used in nearly 50% of all cancer patients [4]. It can be used either alone or in combination with other treatments such as surgery or chemotherapy [5]. Radiation therapy is usually prescribed as a cancer treatment to patients with tumours located at sites such as the central nervous system, cervix, bladder, breast, lung, prostate, oesophageal, stomach and head and neck [5]. Additionally, it can be used as a palliative treatment for symptoms control in patients with late stage disease [4]–[6].

Radiotherapy makes use of ionising radiation to provoke DNA damage and induce cell death of the targeted cells. Despite progress in dose delivery accuracy, radiotherapy is still associated with potentially severe side effects due to the damage of normal tissues surrounding the tumour [6], [7].

1.2 Irradiation-induced DNA-damage response

1.2.1 DNA damage by ionising radiation

When ionising radiation interact with biological systems, it triggers a series of events that can be categorised in three main phases; physical phase, chemical phase a biological phase [6]. During the physical phase, which is the first one occurring, the charged particle interacts with the atoms which compose the tissue causing their excitation and ionisation. The following phase, described as the chemical phase, consists of the rapid chemical reactions of the excited and ionised atoms with other cellular components, which form highly reactive molecules known as “free radicals” [8]. The free-radical propagation continues until the electronic charge equilibrium is restored and terminates within approximately 1 ms from radiation exposure [8], [6]. Scavenging factors are acceptors of electrons and therefore contribute to the electronic balancing process [6], [8]. Finally, the biological phase refers to all the subsequent cellular processes activated in response to the detection of the ionisation damage. It includes DNA-damage repair and, eventually, senescence or cell death which can be triggered few days after the damage formation [9].

Radiotherapy makes use of photons with energies typically from 100 keV to several MeV to ionise the biological matter. Ionizing radiation damages DNA by direct deposition of energy and also indirectly, by ionization of water molecules to produce hydroxyl radicals that attack the DNA [10], [11] (Figure 1.1).

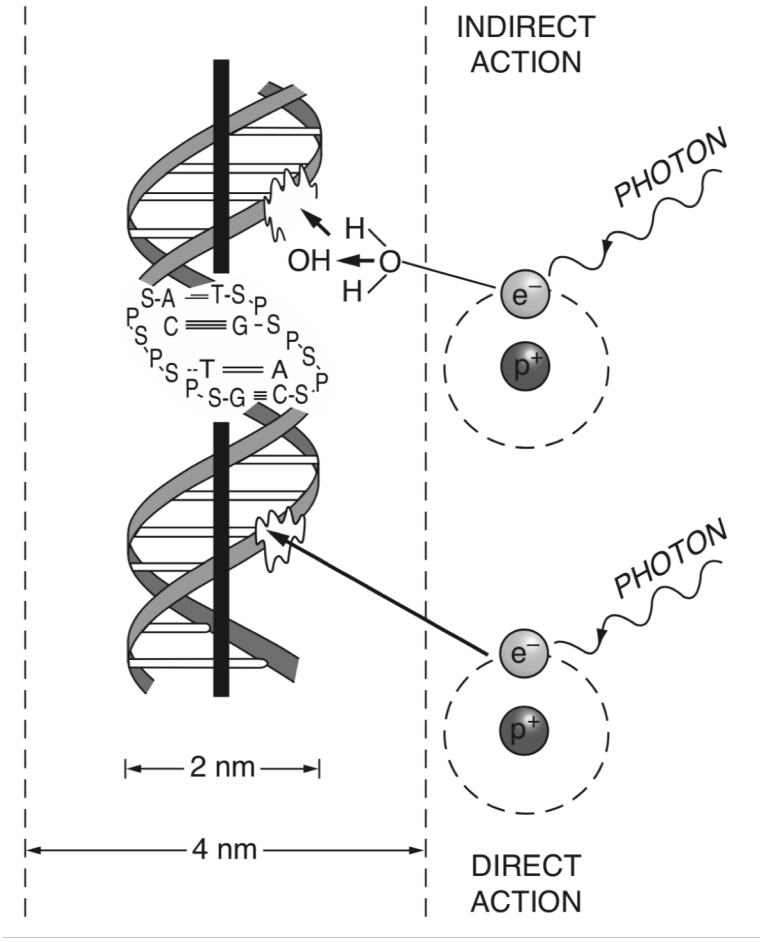


Figure 1.1. **IR induced DNA damage.** Graphical representation of the direct and indirect actions of radiation on DNA molecules (Figure from [11]). In the direct action (bottom), the electron resulting from the absorption of a photon induces the DNA damage. The electron originating from a photon absorption can also interacts with a water molecule to produce a hydroxyl radical (OH·), which in turn produces the damage to the DNA.

Ionisation damage disrupts the normal function of different biomolecules within the cellular compartment, however, the most important one for cellular survival is the DNA [9]. It is estimated that 1 Gy of radiation accounts nearly 1000 single-strand breaks (SSBs) and around 20-40 double-strand breaks (DSBs) per cell, together with base damage and DNA-protein cross-links [6]. Among them, DNA DSB is the most critical lesion, which when misrepaired or unrepaired, can

lead to genomic instability, cell cycle arrest and cell death [9]. For this reason, cells have developed more than one mechanism to detect and repair DNA DSBs.

1.2.2 The DNA damage response in mammalian cells

DNA damage response (DDR) is initiated by the recognition of the DNA lesions, which in turn promotes a signalling cascade to stimulate DNA repair and cell cycle checkpoint activation (Figure 1.2). Three related kinases are primarily involved in the control of the DDR signalling cascade: ATM, ATR and DNA-PK [12]. ATM kinase is recruited at the DSB site by the MRN complex (consisting of three proteins: MRE11, RAD50 and NBS1) and has multiple targets of phosphorylation, among which histone H2AX, CHK2 and p53 [6]. DNA-PK requires the dimer Ku70/Ku80 to sense the DNA DSBs and phosphorylates many substrates to activate and amplify DDR response, among which histone H2AX [12]. ATR represents the third kinase that, following DNA damage, phosphorylates histone H2AX. In contrast to ATM and DNA-PK, ATR is mainly activated by DNA SSBs and DNA damage raising from stalled or broken replication forks [6]. Therefore, ATM, ATR and DNA-PK represent the “DNA damage sensors”, and all of them have the serine 139 of histone H2AX as a target of phosphorylation [12]. The cascade of H2AX phosphorylation, forming γ-H2AX, is a critical step for DNA DSB repair. It generates thousands of γ-H2AX molecules surrounding one DSB which serve as platform for the accumulation of many DNA repair factors [13].

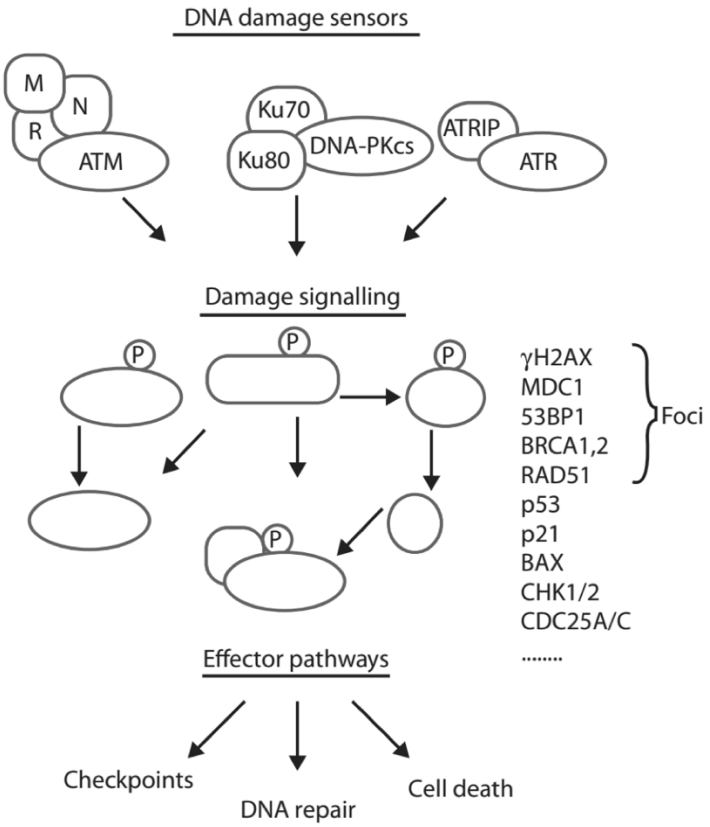


Figure 1.2. **The DNA damage response cascade.** Scheme outlying the DNA damage response (Figure from [6]). The DNA damage sensors (top) activate the DNA damage signalling cascade (middle), which in turn can promote cell cycle checkpoints activation, DNA repair and cell death (bottom).

Once DNA damage has been recognised, mammalian cells can activate two major pathways for the repair of DSBs that differ in their requirements for DNA homology: homologous recombination repair (HR) and non-homologous end joining (NHEJ) [9]. HR is an accurate form of repair, which requires an undamaged sister chromatid to act as a DNA template and functions only after DNA replication (S and G2 cell cycle phases) [14], [15]. In contrast, NHEJ is active throughout the cell cycle and is considered the major pathway for the repair of IR-induced DSBs in human cells [16]. In its simplest sense, NHEJ entails direct

ligation of DNA ends. This can lead to loss of nucleotides from either side of the break, making NHEJ potentially error-prone [16].

1.2.3 Non-homologous end joining repair

It is estimated that nearly 80% of the DNA DSBs are repaired by NHEJ [17]. The NHEJ pathway joins the two ends of a DSB through a process that is homology-independent and that therefore produces junctions that can vary in their sequence composition. NHEJ can be divided into three stages: end-detection and tethering, processing, and ligation [17]. End-detection and tethering is promoted by the Ku70/80 heterodimer which recognise the broken DNA ends and recruits the DNA-dependent protein kinase catalytic subunit (also known as DNA-PKcs) [18]. Together, DNA-PKcs and Ku tether the DNA ends in a synaptic complex. DNA-PKcs forms a physical bridge between the two broken ends and, at the same time, phosphorylates a large number of substrates involved in DNA metabolism. The second step is carried on by the nuclease Artemis and several DNA polymerases in order to process the loose DNA ends to make them suitable for ligation [19]. The final step of NHEJ consists of the ligation of processed DNA ends to repair the DNA break. The joining of the DNA ends is mediated by the DNA ligase LIG4, which exists in complex with XRCC4 and NHEJ1 [6], [20]. After ligation, the DNA break is repaired and the NHEJ machinery can dissociate.

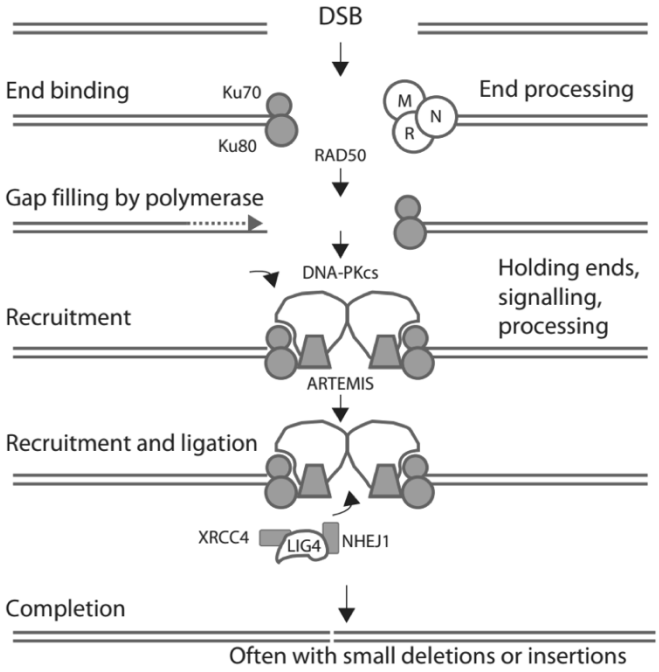


Figure 1.3. **NHEJ DNA repair pathway.** Schematic representation of NHEJ mediated DNA DSB repair (Figure from [6]). The heterodimers Ku70 and Ku80 and the MRN complex bind the loose DNA ends and start the recruitment of DNA-PKcs and Artemis. Then, the DNA ligation is mediated by proteins XRCC4, NHEJ1 and DNA Ligase IV. The NHEJ repair pathway can lead to small deletions or insertions in the DNA sequence during DNA ends processing.

1.2.4 Homologous recombination repair

DSB repair by HR results in the precise repair of the DNA lesion. HR requires the presence of homologous sequences elsewhere in the genome (e.g. a homologous chromosome or a sister chromatid), therefore its activity is restricted to late S and G2 phases of the cell cycle.

HR repair is initiated by DSB ends recognition, followed by DNA ends processing to single-stranded DNA (ssDNA). Recognition and processing of DNA DSBs require the MRE11-RAD50-NBS1 (MRN) complex. The MRN complex functions as a sensor that recruits ATM kinase to the DSB in order to enhance

its kinase activity which triggers the activation of proteins involved in the DNA repair. The exonuclease MRE11 component promotes the DNA filament processing by creating a 3' overhang ssDNA region. Then, the ssDNA tails are bound by RPA which protect them from nucleases activity. Finally, BRCA2 protein promotes the replacement of RPA by RAD51 [21]. Subsequently, RAD51-coated ssDNA can initiate the DNA strand invasion generating a D-loop intermediate. RAD51 paralogs such as RAD51B, RAD51C, RAD51D, XRCC2 and XRCC3 help with the search for homologous DNA and sister chromatid invasion processes. The last phase is characterised by DNA synthesis from the DNA template. DNA helicases such as BLM help to uncoil the DNA template allowing the DNA synthesis promoted by different polymerases. After the new DNA has been synthesised, the repair is completed. Specialized nucleases resolve the two Holliday junctions, which are then ligated to reset chromatin in its original configuration [6], [9], [17], [22], [23]. A schematic representation of the HR repair pathway is shown in Figure 1.4.

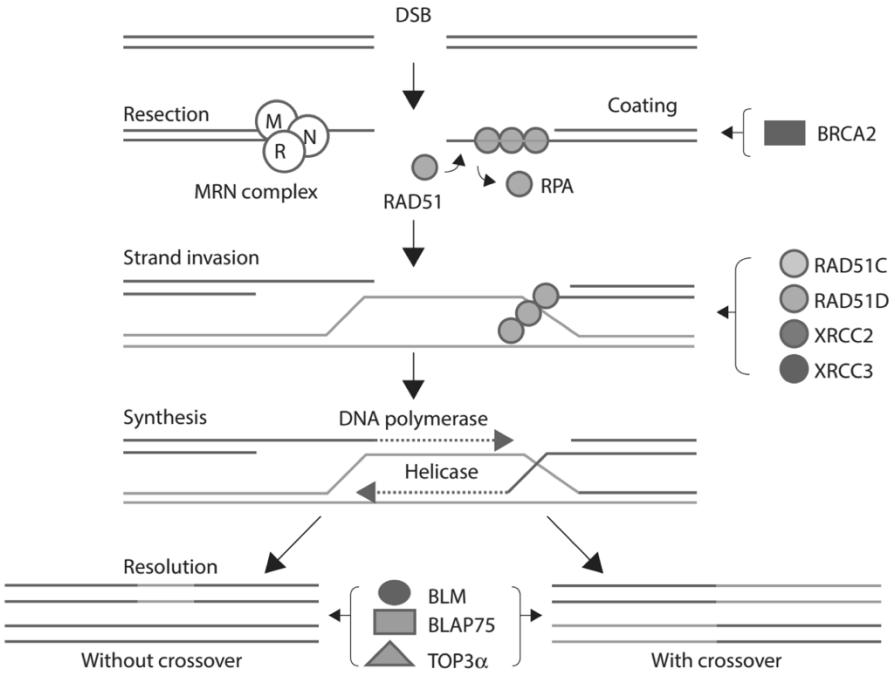


Figure 1.4. **Homologous recombination DNA repair.** Schematic representation of HR repair pathway (figure from [6]). After DSBs recognition, RAD51 is loaded on the ssDNA and promotes strand invasion to search for homologies in order to guarantee accurate DNA repair. The DNA synthesis mediated by DNA polymerases produces two Holliday junctions which can be resolved with or without sequence crossover.

1.2.5 Cell-cycle checkpoints

The cell cycle is the process that allows cells to replicate their content in order to generate two identical daughter cells. It is divided into four stages:

- The Gap phase 1 (G1), in which cells prepare for DNA replication
- DNA synthesis phase (S), during which the DNA is replicated
- Gap phase 2 (G2), in which proper DNA replication is verified
- Mitosis (M), during which cells divides into two identical daughter cells.

As cells stop proliferating, they enter a quiescent phase called G0 (Figure 1.5).

Cell cycle progression through the different phases is regulated by the expression and/or activation of a class of proteins called cyclin dependent kinases (CDKs). The kinase activity of these proteins initiates the signalling cascade accounting the progression through the cell cycle. In order to be able to phosphorylate their targets, CDKs need their cyclin partners (hence their name). The combination of different cyclin/CDK complexes varies within the cell cycle and promotes its progression through the different phases (Figure 1.5) [24].

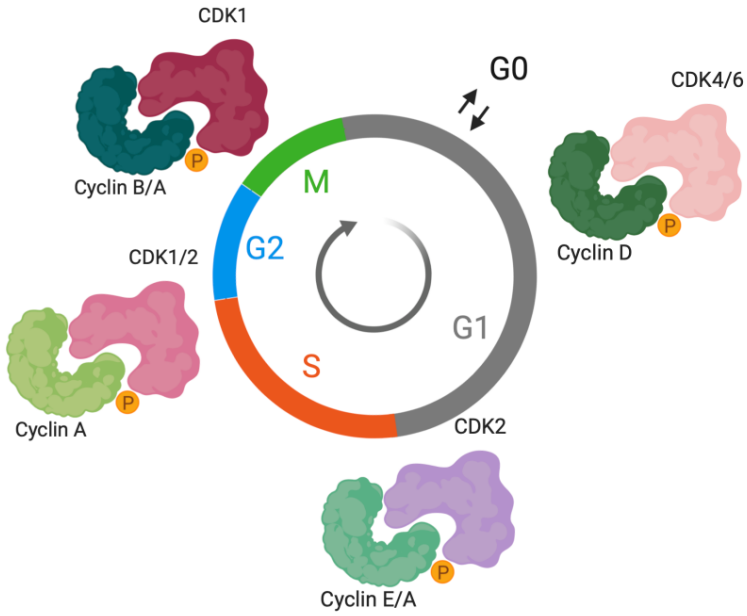


Figure 1.5. **Representation of the four cell cycle stages and the cyclin/CDK complexes regulating the transition between each phase.** Cyclin D-CDK4/6 complex regulates via phosphorylation the transcription of the factors required for DNA replication. Then, during late G1 phase, E2F promotes the transcription of both the cyclin E/A-CDK2 complex and its activator CDC25A, in order to promote the initiation of S phase. Mitotic Cyclin B/A-CDK1/2 complexes are synthesised during late S and G2 phases and get activated to start mitosis.

In presence of DNA damage, to prevent daughter cells inheriting mutated DNA, progenitor cells need to halt cell cycle progression to allow DNA repair. Cell cycle arrest is indeed finely regulated by DDR. According to the phase of the cell

cycle in which the DNA damage occur, four radiation-induced checkpoints can be activated.

The G1/S phase checkpoint prevent cells to start DNA replication in presence of DNA damage. The major responsible of its activation is ATM. In presence of DNA damage, ATM promotes nuclear accumulation of p53 through the phosphorylation of p53 itself and MDM2, an E3 ubiquitin ligase that regulates p53 degradation by proteasome [25]–[27]. p53 accumulation activates or repress a number of downstream target genes involved in various signalling pathways [28]. The activation of p53 induces the expression of its direct transcriptional target, p21 (encoded by CDKN1A gene), causing the G1 phase checkpoint activation and cell cycle arrest. The cyclin dependent kinase inhibitor p21 inhibits the G1 cyclin/CDK complexes (cyclinD/CDK4 and cyclinE/CDK2), which in turn are responsible of retinoblastoma protein (Rb) phosphorylation. Rb phosphorylation releases the E2F transcription factor, which regulates the transcription of several genes necessary to initiate DNA replication [6], [24].

The S phase checkpoint is activated in order to delay the DNA synthesis and allow its repair before continuing with chromosomes replication. The S checkpoint is regulated by CHK1 and CHK2 kinases which are directs phosphorylation targets of ATR and ATM respectively [29]. One of the phosphorylation targets of both CHK1 and CHK2 is CDC25A, which keeps CDK2 active by dephosphorylating it, in order to promote S/G2 transition. When CDC25A is phosphorylated, it becomes inactive or it is degraded by the proteasome [30]. Therefore, CHK1 and CHK2 kinase activity impedes the cells to progress through S phase [6], [29].

To arrest damaged cells in G2 phase there are two distinct checkpoints [31]. Both of them operate similarly to the S phase one, requiring the phosphatase activity of CDC25A/C to activate the cyclin/CDK complexes. The 'early' G2 checkpoint relies on the ATM-CHK2-CDC25A/C axis and it is activated by cells irradiated in G2 phase. The activation of this checkpoint prevents cells entry into mitosis. The 'late' G2 checkpoint applies to cells that reach G2 after being previously irradiated while in G1 or S phases. This checkpoint activation is dose-dependent and can cause a G2 delay of several hours when high radiation doses are delivered. This checkpoint activation is ATM independent and instead it is activated by ATR and CHK1 [6], [31].

1.3 Intrinsic radiosensitivity

According to their genetic background, cancer cells have different susceptibility to ionising radiation treatment compared to healthy cells [32]. Specific gene mutations are indeed associated to changes in cellular intrinsic radiosensitivity, which in turn can impact on radiation response at different clinical endpoints, including survival, local control and disease-free survival [33], [34]. This is often due to genetic mutations causing loss-of-function (LOF) of important proteins involved in the maintenance of genetic stability. Therefore, the identification of novel genes whose depletion accounts an increased genomic instability can reveal novel targets for innovative therapies.

Intrinsic radiosensitivity associated to the loss of a specific gene expression can be measured *in vitro* by using colony formation assay (CFA) in combination with gene depletion. This technique tests the ability of single cells to form colonies after radiation treatment upon a specific gene depletion.

1.3.1 The linear-quadratic (LQ) model

Radiosensitisation effects can be conveniently analysed in terms of mathematical dose-effect relationships based on the linear-quadratic (LQ) model of cell reproductive death as a function of the radiation dose [35]. The LQ formula for cell reproductive death, the surviving fraction (SD) of cells exposed to radiation dose (D), compared to the survival of unirradiated cells (S0) is described by an inverse exponential approximation: $SD/S0 = \exp-(\alpha D+\beta D^2)$ and contains two parameters, α (Gy-1) and β (Gy-2). The linear parameter, α, represents lethal damage from single particle events and describes the low dose area. In the other side, the quadratic parameter, β, indicating sub lethal damage (SLD) dominates the effectiveness in the high dose region. The linear-quadratic model is based on well-accepted biophysical concepts, involving the assumption that lethal damage can be induced by single-particle tracks and by interaction of damage from multiple particles. It has been found to describe the low-dose region of the survival curves up to 6 Gy rather accurately. When the additional treatment results in an increase of the value of the α-parameter, that indicates that this treatment radiosensitise at clinically relevant doses (1.5-2.5 Gy) [35]. When the additional treatment influences the value of the β-parameter, it indicates that the additional treatment has an effect on the repair of sublethal damage [35].

1.4 High-throughput functional genomics screenings

Forward genomic screens are potent tools for the identification and functional annotation of genes in clinically relevant contexts, such as radiation response. These approaches are based on the modulation of gene expression followed by the selection of the cells or organisms with a phenotype of interest.

In this thesis, two high-throughput screens (HTS) aimed at the identification of novel determinants of radiosensitivity are presented. The first one, shown in chapter 3, makes use of RNA interference (RNAi) to reduce the expression of targeted genes and identify novel modulator of radiosensitivity by the subsequent phenotype. This screen was performed by Dr Gaganpreet Tiwana, Dr Gonzalo Rodriguez-Berriguete and Dr Remko Prevo. The second one, presented in chapter 4, makes use of CRISPR/Cas9 system to achieve genome wide knockout of distinct genes in order to analyse clonal proliferation following radiation treatment. This screen was performed by me in collaboration with Dr Dylan Jones from the Target Discovery Institute (TDI), Oxford.

1.4.1 siRNA high-throughput screenings to identify novel radiosensitisers

Forward arrayed RNAi screens rapidly spread after the discovery of the system regulating gene expression by making use of small interfering RNAs (siRNAs) [36], [37]. Gene silencing by RNAi consists of transfecting cells with synthetic siRNAs, normally formed by 20-25 base pairs, engineered to target a specific RNA transcripts (mRNAs). The recognition of the mRNA by the complementary siRNA induce its degradation instead of translation, causing reduced expression of specific gene products.

The arrayed siRNA radiosensitivity screen presented in this thesis make use of RNAi to knockdown specific genes and identify them as modulator of radiosensitivity according to the loss of proliferative potential caused by radiation treatment. The proliferation assay used for the screen is the gold standard technique to assess cellular survival upon radiation treatment: the colony formation assay (CFA). Our laboratory developed a 96-well plate colony formation assay enabling the performance of several high-throughput radiosensitivity screens [38]–[40]. Using this approach, a high-throughput siRNA radiosensitivity screen was performed designed to target the 'druggable' genome by using a siRNA Library containing 9,962 siRNA pools [39]. The 'druggable' genome consists of a group of genes predicted to be suitable candidates for pharmacological targeting [41].

The major disadvantages of genetic screens using siRNA gene silencing are the incomplete and transient gene knockdown and the extensive off-target activity of the siRNAs. In contrast, sequence-specific programmable nucleases, such as the CRISPR/Cas9, have emerged as interesting new genetic perturbation systems, providing targeted modification of the DNA sequence itself. Recently, the CRISPR/Cas9 system has been extensively used for LOF genetic screenings, as it allows to achieve targeted and permanent genetic modifications which lead to gene knockout and subsequentially complete and stable depletion of the gene products.

1.4.2 CRISPR/Cas9 genome editing

Genome editing by the CRISPR/Cas9 system makes use of the Cas9 nuclease which is an element of the type II CRISPR bacterial adaptive immune

system [42], [43]. Targeted gene mutagenesis can be achieved by introducing a DNA DSB in specific portions of a gene and then exploit the eukaryotic DNA DSB repair pathways to create mutations at the required site [44], [45]. The nuclease Cas9 can induce DSBs to specific locations by using short single guide RNAs (sgRNAs or gRNAs) which are complementary to the target DNA. The gRNA usually consists of a 20 base pair guide sequence and the complementary target DNA is followed by a 3-base pair protospacer-adjacent motif (PAM). The CRISPR/Cas9 system relies both on the homology-directed repair (HDR) and NHEJ. In order to introduce exogenous DNA sequences in the genome (knock-in), the HDR pathway is exploited by providing a synthetic DNA sequence as a template for repair. In contrast, to generate knockout genetic mutations, NHEJ is exploited as an error-prone repair mechanism, which can introduce insertion or deletion (indel mutations) at the targeted site [44], [45]. Indel mutations cause a coding frameshift, resulting in a premature stop codon and the initiation of nonsense-mediated decay (NMD) of the transcript. However, an early frameshift mutation or large indels might be enough to generate a non-functional protein too. Because indel mutations are more likely to deactivate gene products when they occur at the beginning of a gene sequence, early exons are typically preferred for targeting [46].

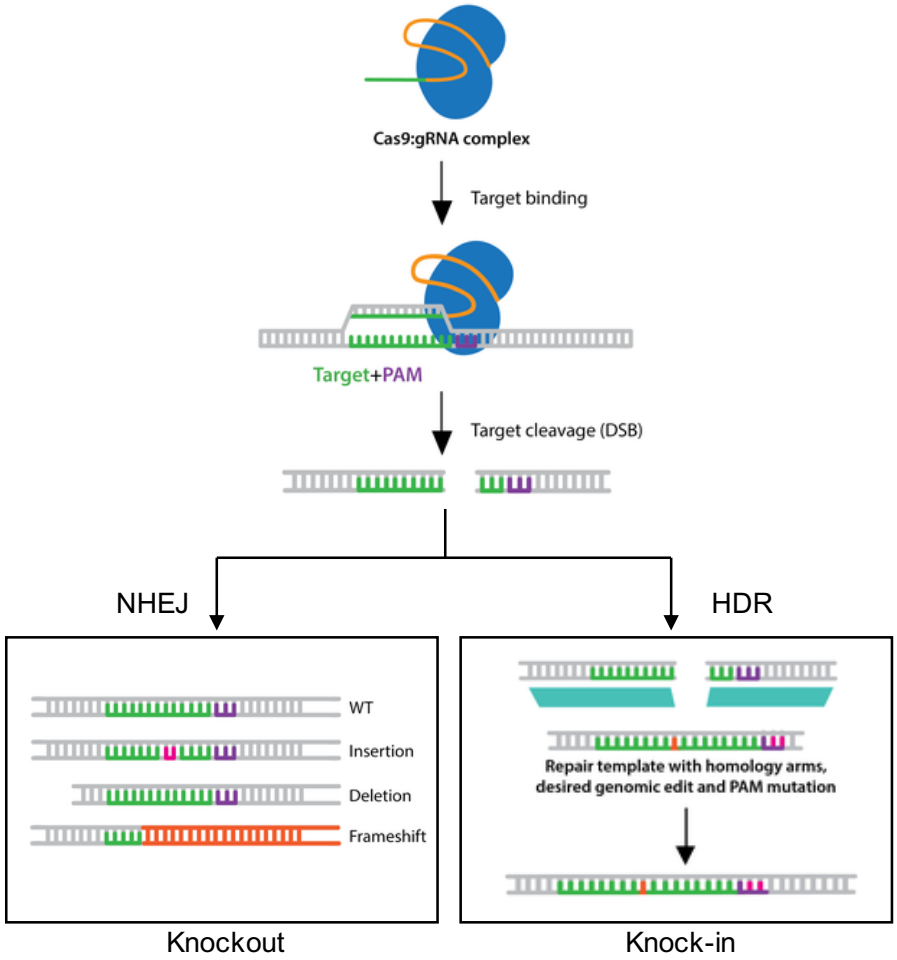


Figure 1.6. **CRISPR/Cas9 genome mutagenesis.** A graphical representation of the gene mutagenesis produced by the Cas9 nuclease in order to promote gene KO or knock-in. NHEJ repair induces indel or frameshift mutations which in turn result in NMD of the transcript. HDR repair uses a synthetic DNA template which provides the exogenous sequence to be inserted in the target gene. Figure adapted from [47].

1.4.3 LOF CRISPR/Cas9 genome-wide pooled screenings to identify novel determinants of radiosensitivity

LOF CRISPR/Cas9 genome-wide pooled screens are based on systematic induction of genetic perturbations in order to map them to a phenotype of interest. This involves knockout of individual genes, accomplished by the Cas9 activity, combined with genome-wide gRNA library, which mediates genome-scale

mutagenesis, followed by DNA sequencing, which allows the discovery of the genetic perturbations involved in the phenotypic alteration.

In this study a LOF whole genome CRISPR/Cas9 screen was performed in order to identify novel regulator of radiation response. To the best of my knowledge, this is the first study to investigate novel determinants of radiosensitivity by using a CRISPR genome-scale approach in MRC-5 cells. The genome-wide library used for the screen was the GeCKOv2 library, developed by Feng Zhang and Ophir Shalem [48]. The GeCKOv2 library was designed to target 19,052 genes, each targeted by 6 different gRNAs. The library was transduced using the lentiCRISPRv2 vector, a lentiviral vector able to transduce and stably maintain the constructs as genomic integrants during subsequent cell replications. The pooled library was transduced using a MOI such as single viral integration per cell was achieved. MRC-5 cells, a lung fibroblast cell line, were transduced with the pooled GeCKOv2 library and treated with IR in order to identify novel modulators of radiation response. Finally, the gRNA sequences were detected by DNA sequence analysis and the abundance of each gRNA was determined in each sample. By comparing the abundance of each gRNA in the irradiated and control samples, genes whose depletion caused an increased sensitivity to radiation were identified.

1.5 Aims of this thesis

The aim of this doctorate thesis it to identify and validate novel genes associated to radiation response. At first, this thesis implements and compares both well-established (RNAi) and novel approaches (CRISPR/Cas9) of functional genomics HTSs. Interestingly, both approaches identified a common target as a novel modulator of radiosensitivity, TRIM37. Therefore, the last part of this work focuses on the validation of TRIM37 as a novel determinant of radiosensitivity and investigates its role in radiation response. The thesis aims to address the following questions:

- Is CRISPR/Cas9 genome editing technology suitable for LOF genome-wide functional genomics screenings?
- Can LOF CRISPR/Cas9 genome-wide screenings be used to identify novel regulators of radiosensitivity?
- Is TRIM37 a modulator of radiosensitivity in a range of cancer cell lines?
- Is TRIM37 overexpressed in cancer?
- Does the lack of TRIM37 expression affect the DNA damage response?
- Does TRIM37 depletion affect the DNA damage repair efficiency?
- Can TRIM37 depletion alter the cell cycle progression in response to radiation?
- How does TRIM37 depletion alter the cell cycle progression of asynchronous and synchronous cells in response to radiation?
- Is TRIM37 important for genomic integrity?

The presented results show that LOF CRISPR/Cas9 genome-wide screenings allow the investigation of thousands of genes in the context of radiation, overcoming the limitations of the RNAi-based HTSs. Both approaches are valid tools for functional genomics studies in the context of radiation. Also, the results show that TRIM37 is a novel modulator of radiosensitivity and provides some evidences that link TRIM37 function to genome stability maintenance.

Chapter 2 MATERIAL AND METHODS

2.1 Cell culture

2.1.1 Cell line maintenance

HCT116 (colorectal), HeLa (cervical), MCF-7 (breast), T47D (breast), BT-549 (breast), MRC5 (lung fibroblast), DU145 (prostate), T24 (bladder) cell lines were purchased from American Type Culture Collection (ATCC). DR-GFP-HEK293 were provided by Prof. Jeremy Stark and DR-GFP-U2OS were gifted from Prof. Tim Humphrey.

Cells were grown in an incubator at 37°C with 5% CO2. The cell medium was Dulbecco's Modified Eagle Medium (DMEM, Sigma) for HeLa, DU145, MCF-7, T47D and HCT116. DR-GFP-HEK293 and DR-GFP-U2OS were grown in DMEM medium supplemented with 1 μg/mL puromycin. T24 and BT-549 cells were grown using Roswell Park Memorial Institute (RPMI-1640, Sigma) medium. Minimum Essential Medium (MEM, Sigma) was used for MRC-5 cells. All media were supplemented with 10% Fetal Bovine Serum (PAN Biotech). All cell lines were tested for mycoplasma using MycoAlert (Lonza) and authenticated by LGC standards (ATCC) by short tandem repeat (STR) profiling.

2.1.2 Cell line storage and resuscitation

Cells were kept in cryovials stored in vapour phase liquid nitrogen. Cells were thawed in a water bath at 37°C. Cells were immediately diluted in 10 ml of media in a 50 ml falcon tube and centrifuged at 240 x g for 5 minutes. The media was removed, and the cells were resuspended in 5 ml of media and seeded in a T75 flask.

2.1.3 Passaging cell lines

Media was removed from the flasks. Cells were washed in 5-10 mL of 1 x phosphate buffered saline (1xPBS) and 1-2 mL of trypsin-ethylenediaminetetraacetic acid (EDTA) was added to the flask. Cells were then returned to the incubator for 5 minutes. Medium was added after cells had lifted, and the cell suspension was diluted appropriately in medium. Typically, the cells were passaged at 1:10, or 1:20 according to the growth rate and returned to the incubator. Cell lines were passaged for a maximum of 3 weeks and were passaged when cells reached ~ 60-80% confluency.

2.1.4 Cell counting

Cells were trypsinised as described above. 10 μL of cell suspension was loaded into a cell counting slide from Luna™ and counted using the Luna-II™ automated cell counter. Otherwise, a 50 μL cell suspension was diluted 1:10 in DMEM in a microcentrifuge tube and counted using the Scepter 2.0 (Millipore) hand-held automated cell counter.

2.1.5 Cell synchronisation

Cell synchronisation in G2/M phase was achieved using a nocodazole block in a 6-well plate format. 100 ng/ml of nocodazole was added to adherent cells at about 50-60% confluency. Cells were left growing with nocodazole until sample collection. Cell cycle stage was confirmed by BrdU and propidium iodide FACS analysis.

2.2 Plasmid transfection

Plasmid transfection was performed by incubating growing cells with transfection complexes diluted in 1 mL of antibiotic-free medium for 5 to 6 hours.

Transfection complexes consisted of 2 ug of expression vector diluted into 250 μL of Opti-MEM (Invitrogen), mixed with 6 μL of Lipofectamine 2000 (Invitrogen) diluted in 250 μL of Opti-MEM (Invitrogen). Transfection complexes were left incubating for 30 minutes before being added to growing cells. Then, plates were shacked back and forth several times to mix the transfection medium and returned to the tissue culture incubator for 5 to 6 h. Then, transfection medium was replaced with fresh medium.

2.3 Reverse siRNA transfection

Reverse siRNA transfection was performed by incubating suspended cells with siRNAs at 20 nmol concentration and RNAiMAX transfection reagent at 1:1000 dilution in 6 well plates. First, siRNAs (Ambion) and the RNAiMAX transfection reagent (Sigma) were diluted in Opti-MEM (Invitrogen). Then, the diluted siRNA and RNAiMAX were mixed and incubated for 20-30 minutes at room temperature before addition to cell suspension. siRNA transfection mix was left incubating overnight and replaced with fresh medium.

2.4 Colony formation assay

Three days after transfection, siRNA transfected cells were plated as single cells in 6 well plates, allowed to settle for 4 to 6 hours and then given the appropriate dose of IR using an IBL637 caesium-137 irradiator. Experimental plating numbers vary depending on the cell line used and are shown in Table 2.1 below. Colonies were grown for two weeks, stained and fixed with crystal violet in methanol (MeOH) and counted using a GelCount colony counter (Oxford Optronix).

Cell line	CFA Plating Numbers			
	0 Gy	2 Gy	4 Gy	6 Gy
HeLa	200	400	800	1,800
MRC-5	1,000	2,000	4,000	10,000
T24	200	400	800	1,600
BT-549	400	800	1,600	4,000
DU145	200	400	800	1,600
HCT116	200	400	800	1,600
MCF-7	500	1,000	2,000	5,000
T47D	200	400	800	1,600

Table 2.1. **CFA plating numbers.** Plated cell numbers at each dose of radiation for colony formation assay in 6-well plates.

2.4.1 Statistical analysis

CFA survival data was curve fitted using GraphPad Prism by non-linear regression of the linear quadratic (LQ) model (Equation 2.1):

$$SF = e^{-\alpha D - \beta D^2}$$

Equation 2.1. **LQ model equation.** SF denotes surviving fraction, D is radiation dose expressed in Gray (Gy), and α (Gy^{-1}) and β (Gy^{-2}) are parameters to fit.

Each dose of radiation had N = 3 technical replicates and no fewer than three experimental repeats.

2.5 Western blotting

2.5.1 Preparation of protein lysates

Cells grown in 6 well plates were washed in 1x PBS and lysed by scraping in 100 µL RIPA buffer supplemented with protease inhibitor (Roche), phosphatase inhibitor cocktail 2 (Sigma) and cysteine protease inhibitor N-ethylmaleimide (NEM, Sigma). The cell lysates were transferred to 1.5 mL microcentrifuge tubes, incubated on ice for 5 minutes, and then sonicated for 10 seconds. Lysates were

centrifuged at 165,000 rpm at 4°C. The supernatants were transferred to fresh microcentrifuge tubes and stored in -80°C. Protein quantification was performed using a BCA assay kit (Thermo Scientific) according to manufacturer's instructions.

2.5.2 SDS-PAGE and protein transfer

Protein extracts (25 µg) were diluted 1:4 into 4x SDS loading buffer (Expedeon) supplemented with 20 mM of Dithiothreitol (DTT) reducing agent and then denatured for 5 minutes at 100°C prior to loading. The samples were loaded into RunBlue pre-cast gels (Expedeon) and subjected to SDS-PAGE alongside a Precision Plus Protein Dual Colour Standards (Bio Rad). RunBlue SDS Running Buffer (1x, Expedeon) was used to perform the gel electrophoresis. The proteins were transferred to a Hybond nitrocellulose transfer membrane (GE-healthcare) using the Xcell II™ Blot Module system (Invitrogen) with 1x transfer buffer (20 % Methanol, 20 mM Tris, 190 mM glycine).

2.5.3 Blocking, antibody incubation and signal detection

The membranes were blocked in 5% Milk made in 1x TBST (50 mM Tris-Cl pH 7.6, 150 mM NaCl, 0.1% Tween) for 1 hour at room temperature. The rabbit anti-TRIM37 antibody (Bethyl Laboratories, #A301-173A; 1:500), the rabbit anti-TPR antibody (Novus Bio, #NB100-2866; 1:1,000), and mouse anti-β-Actin (Sigma, #A1978; 1:100,000) were diluted in 5% Milk made in 1x TBST. The membranes were incubated overnight at 4°C with rotation, washed three times in 1x TBST and incubated for 2 hours with the anti-rabbit and anti-mouse HRP antibodies (Thermoscientific, #31460 and #31335, respectively) diluted 1:20,000 in 5% Milk made in 1x TBST. Finally, the blots were washed 3 times in 1x TBST,

treated with Immobilon Western Chemiluminescent HRP Substrate (Millipore) and developed in the dark room using BioMax photographic film (Kodak).

2.6 Immunofluorescence analysis

siRNA-transfected cells were plated in 24-well plates on glass coverslips. At various time-points following ionising radiation (IR) treatment, cells were washed with 1xPBS and fixed with 300 µL 4% paraformaldehyde in 1xPBS for 10 minutes. After washing with 1xPBS, cells were blocked and permeabilised in 500 µL 1xPBS with 0.1% Triton, 1% BSA and 1% Goat serum at room temperature for at least 1 hour. The blocking buffer was removed, and cells were probed with anti-γ-H2AX Ser139 (Upstate/Millipore, #05-636), anti-53BP1 (MERK, #BP13) and anti-RAD51 (Santa Cruz, #sc-8349) antibodies at a concentration of 1:4,000, 1:800 and 1:1,500, respectively diluted in 1% BSA, 1% Goat serum in 1xPBS overnight at 4℃. Plates were washed thrice with 1xPBS with the second wash lasting 10 minutes. Cells were then incubated with Alexa Flour 488 goat anti-mouse (Invitrogen #A11001, 1:2000) and Alexa 488 goat anti-rabbit (Invitrogen #A11070, 1:2000) conjugated secondary antibodies and 1 µg/mL 6-diamidino-2-phenylindole (DAPI) for at least 1 hour in dark. The coverslips were then washed, removed from wells and mounted on a microscope slide using VECTASHIELD antifade mounting media. Foci were detected using epifluorescence microscope (Nikon Ni-E) and the images analysed using an ImageJ macro. At least 300 nuclei per sample were counted for the analysis.

2.7 Flow cytometry

2.7.1 Cell cycle analysis by propidium iodide staining

Cells were lifted using Trypsin-EDTA, re-suspended in medium, transferred into FACS tubes and pelleted by centrifugation at 200 x g for 5 minutes. The medium was removed, and cells were fixed in ice-cold 70% ethanol. Samples were then centrifuged at 400 x g for 5 minutes and the ethanol was removed. The cells were washed once in 1xPBS to remove the ethanol, and then 300 µL propidium iodide (PI) staining solution (50 µg/ml PI, 200 µg/ml RNase) was added. Cells were vortexed briefly and incubated at room temperature in the dark for 10 minutes. Stained cells were analysed using a BD FACSCalibur flow cytometer. The cytometer settings were set at a flow rate of 150-300 events per second and the working voltage varied between 200 V and 400 V. Histogram plots were produced using open source Flowing Software 2.5.1.

2.7.2 BrdU/Propidium iodide staining

20 µM BrdU (Sigma) was added to adherent cells in a 6-well plate 30 minutes before being lifted. After samples had been fixed in 70% ethanol and subsequently washed with 1xPBS, 2 M HCl containing 0.1 mg/ml pepsin was added and samples were incubated at room temperature for 20 minutes. Samples were then washed three times with 2% FBS (Foetal Bovine Serum)/1xPBS and then resuspended in 100 µL of 2% FBS/1xPBS with 1 µL Mouse anti-BrdU monoclonal antibody (1:100, BD Bioscience). After 90 minutes incubation at room temperature, mixing every 30 minutes, 1xPBS was added, samples were centrifuged at 350 x g and the pellet was resuspended in 100 µL 2% FBS/1xPBS with 0.5 µL Goat anti-Mouse AF488 antibody (1:200 dilution) and

incubated for 60 minutes at room temperature in the dark, mixing every 30 minutes. Finally, samples were washed in 1 ml 1xPBS and resuspended in 0.5 ml 1xPBS containing 50 µg/ml propidium iodide, diluted as necessary to give a flow rate on the BD FACSCalibur cytometer between 150 and 300.

2.8 I-SceI-based reporter assays

The I-SceI-based assays were performed according to the method described by Gunn and Stark [49]. U2OS and HEK293 cells harbouring the DR-GFP (HR) reporters were reverse transfected with siRNAs using the siRNA reverse transfection protocol described above (2.3 Reverse siRNA transfection). The day after, cells were transfected with the I-SceI plasmid as described in section 2.2 Plasmid transfection. Seventy-two hours after siRNA transfection GFP-positive cells were scored by flow cytometry analysis using a BD FACSCalibur cytometer at a flow rate between 150 and 300.

2.9 Genome-wide CRISPR/Cas9 Screen

2.9.1 Cell transduction

MRC-5 cells were transduced with the human genome wide CRISPR library (GeCKOv2), a gift from Feng Zhang (Addgene). The library is divided into two parts (Library A and B) and consists of 122,411 guide RNAs (gRNA) that target and knockout 19,050 genes and 1,864 genes coding for miRNAs (detailed information of the GeCKOv2 libraries are shown in Table 2.2). Each library also contains a negative control consisting of 1,000 non-targeting gRNAs.

GeCKOv2 library specification	
Number of genes targeted	19,050
Targeting constructs per gene	6 per gene (3 in Library A, 3 in Library B)
Number of miRNAs targeted	1,864
Targeting constructs per miRNA	4 per miRNA
Non-targeting control guides	1,000
Total gRNA constructs	122,411 (65,383 in Library A, 58,028 in Library B)

Table 2.2. **GeCKOv2 library specifications.** The genome-wide library was divided into two libraries, GeCKOv2 library A and B.

MRC-5 cells were infected with the GeCKOv2 library A and B using 3rd generation Lentivirus. The lentiCRISPRv2 vector (Addgene) mediates the expression of the gRNAs in the transduced cells. Each construct (shown in Figure 2.1) encodes for the gRNA sequence, the Cas9 gene and the puromycin resistance gene. The lentiviral infection allows the integration of the vector into the genome of the cells to achieve a stable expression of the construct and subsequently a stable single-gene knockout.

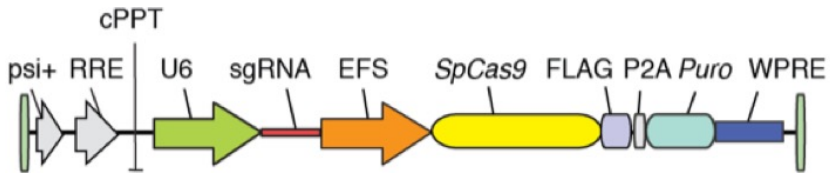


Figure 2.1. **LentiCRISPRv2 construct used to transfect the cells with the GeCKOv2 libraries.** The construct consists in lentiviral expression vector for S. pyogenes Cas9 (SpCas9) and guide RNA (sgRNA). psi+, Psi packaging signal; RRE, Rev response element; cPPT, central polypurine tract; EFS, elongation factor 1a short promoter; Flag, Flag octapeptide tag; P2A, 2A self-cleaving peptide; Puro, puromycin selection marker; WPRE, post-transcriptional regulatory element.

In order to obtain sufficient yet manageable library coverage within the cell population, the viral transduction procedure was aimed at obtaining 500

transduced cells per single gRNA CRISPR. Furthermore, in order to ensure that each cell integrated not more than a single gRNA construct, the transduction was set up with a virus concentration expected to yield only a 30% infection rate (MOI 0.3). Any non-infected cells were subsequently removed by puromycin selection. The number of cells needed for the transduction procedure was calculated by taking into account a 500x coverage and a MOI of 0.3 (Table 2.3).

GeCKO Library	Total gRNA constructs	Cells per gRNA	Total cells for 500X coverage	Cells needed for seeding
A	65,383	500	32,691,500	65,383,000
B	58,383	500	29,014,000	58,028,000

Table 2.3. **GeCKO library transduction details.** The number of cells needed to achieve a 500X coverage for both libraries and the cells needed for transduction when using a MOI of 0.3.

2.9.2 Timeline of transduction procedure

Cells were plated and left over-night to allow them to attach to the flasks. The day after, the Lentiviral infection was performed in all the flasks apart from one in order to have a no-transduction control. One day later, the medium containing the virus was aspirated and replaced with fresh media, which was left for another 48 hours. These extra 48 hours allowed the lentivirus DNA to be integrated in the host cells and therefore become puromycin resistant. All the cells were then incubated with 0.75 μg/mL puromycin and kept under selection until there were no more surviving cells in the non-transduced cells (approximately 7 days). Whereas non-transduced cells looked viable after 3 days, puromycin concentration was increased up to 1 μg/mL. The week after transduction is also meant to allow repair of the double strand breaks generated by Cas9 enzyme, followed by the depletion of the targeted transcript and protein. Finally,

transduced cells have been expanded to reach the required amount for the screen. A scheme of the transduction timeline is shown in Figure 2.2.

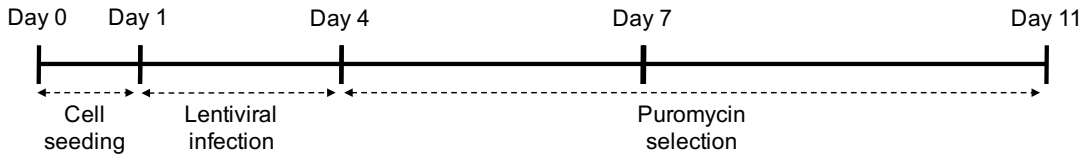


Figure 2.2. **Timeline of the transduction of the GeCKOv2 library in MRC-5 cells.** A scheme showing the steps of the library transduction over time.

2.9.3 Cell treatments

Cells were treated with γ-radiation at 2 Gy using an IBL637 caesium-137 irradiator.

2.9.4 Genomic DNA extraction

Genomic DNA extraction was performed using the Genomic DNA Maxi Kit (QIAGEN). The gDNA extraction was carried out following the user manual protocol.

2.9.5 PCR

The gRNA sequences were amplified by PCR in two steps. In order to achieve 500X coverage of the libraries, the amount of input genomic DNA was calculated to be 220 µg of DNA for library A and 200 µg of DNA for library B (assuming 6.6 µg of gDNA for 10^6 cells). For each sample, 22 (library A) and 20 (library B) separate 100 µL reactions with 10 µg of gDNA each were performed, using Herculase II Fusion DNA Polymerase (Agilent) in 18 cycles. Since all the

gRNA constructs have identical 5' and 3' terminus, they were amplified using the same Forward and Reverse primers. The primers sequences are:

F1 – AATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCG

R1 – CTTTAGTTTGTATGTCTGTTGCTATTATGTCTACTATTCTTTCC

The resulting amplicons were pooled and purified using agarose gel extraction. In order to render the pooled samples suitable for sequence analysis, the gRNAs amplicons were modified with Illumina adapters, staggers and barcodes in a second PCR, using the 12 different primers in order to increase the complexity of the library. Figure 2.3 shows a scheme of the second PCR primers.

Forward primer scheme

5' – Illumina adapter F : Stagger : Barcode F : Priming site – 3'

Reverse primer scheme

5' – Illumina adapter R : Barcode R : Illumina adapter R : Stagger : Priming site – 3'

Figure 2.3. **Forward and Reverse primers scheme.** The Illumina adapters are necessary to allow the annealing of the sequences to the patterned flow cells. The Barcode is an 8bp sequence used to mix different samples in one single run of sequencing. The Stagger consists of a variable length sequence to increase library complexity to achieve a more efficient read.

DNA amplification was carried out with 5 µL of the 1st PCR product in 100 µL volume. Each sample was amplified in 8 separate reactions. Resulting amplicons from the second PCR were gel extracted, quantified, mixed and sequenced using HiSeq 4000 (Illumina).

2.10 Statistical Analysis

The significance of differences between means was measured by two-tailed t-test using the GraphPad Prism program version 8.0 (GraphPad Software, USA). Significance was determined according to two-tail p-values where *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001 was considered statistically significant.

2.10.1 Quantification of Sensitisation

The sensitisation enhancement ratio (SER), was used to quantify the effectiveness of combination treatments (Equation 2.2):

$$SER_{10 = siNT\ D_{10}/\ siX\ D_{10}}$$

where SER_{10} represents the SER at 10% survival, siNT D_{10} is the dose required to achieve 10% survival with the non-targeting siRNA control, and siX D_{10} is the dose needed to achieve 10% survival with the siRNA targeting the investigated gene (X). In general, SER values greater than 1 indicate and increased sensitivity to radiation, values between 1 and 2 indicate an additive effect, as the SER factor grows larger, the effect become synergistic.

Chapter 3 IDENTIFICATION AND VALIDATION OF NOVEL DETERMINANTS OF RADIOSENSITIVITY THROUGH HIGH-THROUGHPUT RNAI SCREEN

3.1 Introduction

While radiotherapy has been widely and effectively used as a therapeutic agent for decades [6], the cascade of the events triggered by radiation damage has not yet been fully explored. Indeed, still nowadays there are numerous molecules and pathways discovered to have a role in radiation response. These novel regulators of radiation response represent potential targets of innovative treatments to improve cancer patients' therapies outcome. High-throughput functional genomics approaches, such as RNAi radiosensitivity screenings, represent powerful tools for novel therapeutic targets discovery. Moreover, functional genomics methods provide a better understanding of the events occurring after IR-induced damage and can therefore offer greater insights into the cellular response to radiation.

The gold standard technique to assess cellular survival upon radiation treatment is the colony formation assay. Our laboratory has developed a 96-well plate colony formation assay enabling the performance of high-throughput radiosensitivity screens and the discovery of novel targets of radiosensitivity, [38]–[40]. Using this approach, Dr Gaganpreet Tiwana, Dr Gonzalo Rodriguez-Berriguete and Dr Remko Prevo performed a high-throughput siRNA radiosensitivity screen in HeLa cells treated with 0 Gy and 7 Gy of IR [39]. The screen was designed to target the 'druggable' genome by using a siRNA Library containing 9,962 siRNA pools (four strands per gene). The 'druggable' genome consists of a subset of genes which are predicted to be suitable candidates for pharmacological targeting, according to several criteria, such as the presence of binding sites for bioavailable compounds [41]. The library also contained a

non-targeting siRNA pool and a siRNA pool targeting the PRKDC gene (DNA-PKcs) to be used as negative and positive controls, respectively. From the primary screen, the top 240 genes were selected for a secondary screen, using a human custom-made siRNA library, which was run in triplicate. For both primary and secondary screen, the radiosensitisation score (R_{score}) was calculated as described in the Equation 3.1 below, with lower values corresponding to greater radiosensitisation effect.

$$R_{score} = \frac{(normalised\ SF - 1)}{(meanabsolute\ deviation\ of\ normalised\ siNT\ SF)}$$

Equation 3.1. **Radiosensitisation score calculation.** Formula used to calculate the radiosensitisation score (Rscore) from the siRNA screening surviving fractions.

The genes in each replica were ranked according to their R_{score}. For the secondary screen, rank product analysis was carried out between replicates. Figure 3.1 shows the heatmap of the secondary screen, with a magnification of the top 100 genes. The quality of the screen could be also verified by the discovery of known determinants of radiosensitivity such as LIG4 and ATM [20], [50].

Genes that had not been previously associated with radiosensitivity but ranked within the top 30 genes were selected for further validation and investigation. This led to the identification of the nuclear pore complex protein NUP54 as novel player in the response to radiation [39]. This chapter describes the investigation of the genes ranked within the top 100 of the secondary screen.

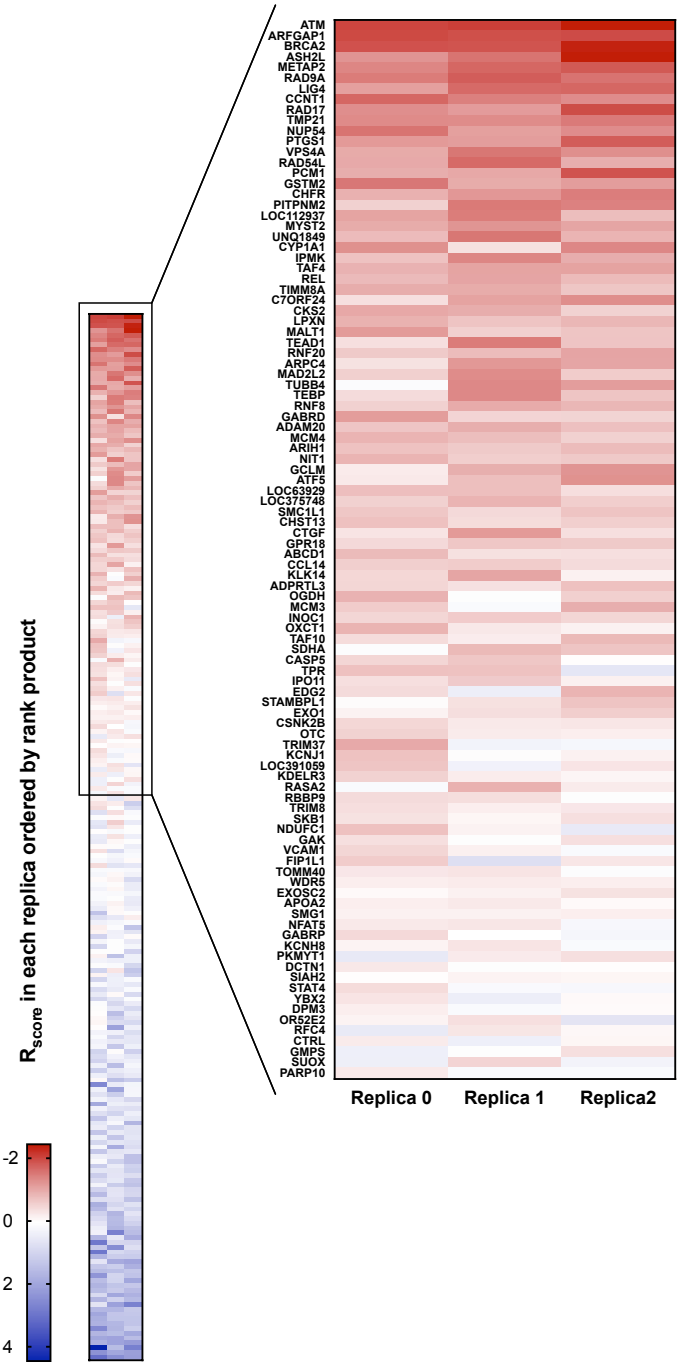


Figure 3.1. **Druggable genome siRNA radiosensitivity screen.** A high-throughput siRNA colony formation screen performed by Gaganpreet Tiwana, Gonzalo Rodriguez-Berriguete and Remko Prevo in our laboratory identified a range of candidate genes whose knockdown modulates radiosensitivity in HeLa cells. The figure shows heat maps of the secondary screen displaying the three replicas of each gene ordered by rank product. Low value R-score are labelled in red, high R-scores in blue. A magnification of the top 100 genes is showed on the right.

3.2 Results

3.2.1 Gene selection and identification of novel determinants of radiosensitivity

Among the top-100 hits, genes were selected for further analysis according to some of the following criteria:

- They should not have been previously associated with radiation response. A literature review to assess the novelty of the candidate genes was performed.
- Whether their expression was high or exclusive in tumours. Differential gene expression between tumour and normal tissue was assessed by using Oncomine 4.5 database [51].
- Whether there were good commercial reagents available. Availability of reagents such as antibodies is crucial for investigating the mechanism of action of radiosensitisation. Therefore, antibodies availability was determined by using the research tool BenchSci.
- Whether there was a known correlation with cancer prognosis. Through literature review, the clinical impact of each candidate target was investigated.

Giving the criteria above, eight genes (listed in Table 3.1 below and shown in Figure 3.2) were selected within the top 100 hits to confirm their radiosensitisation effect in HeLa cells (cervix carcinoma). The gene depletion was performed using an siRNA from a different vendor than the one used in the screen, thus reducing the chance of false positives due to siRNA off-target effects. At the same time, it was assessed whether their depletion radiosensitised also normal cell lines

(MRC-5, lung fibroblasts), as the main aim was to identify tumour specific radiosensitisation targets.

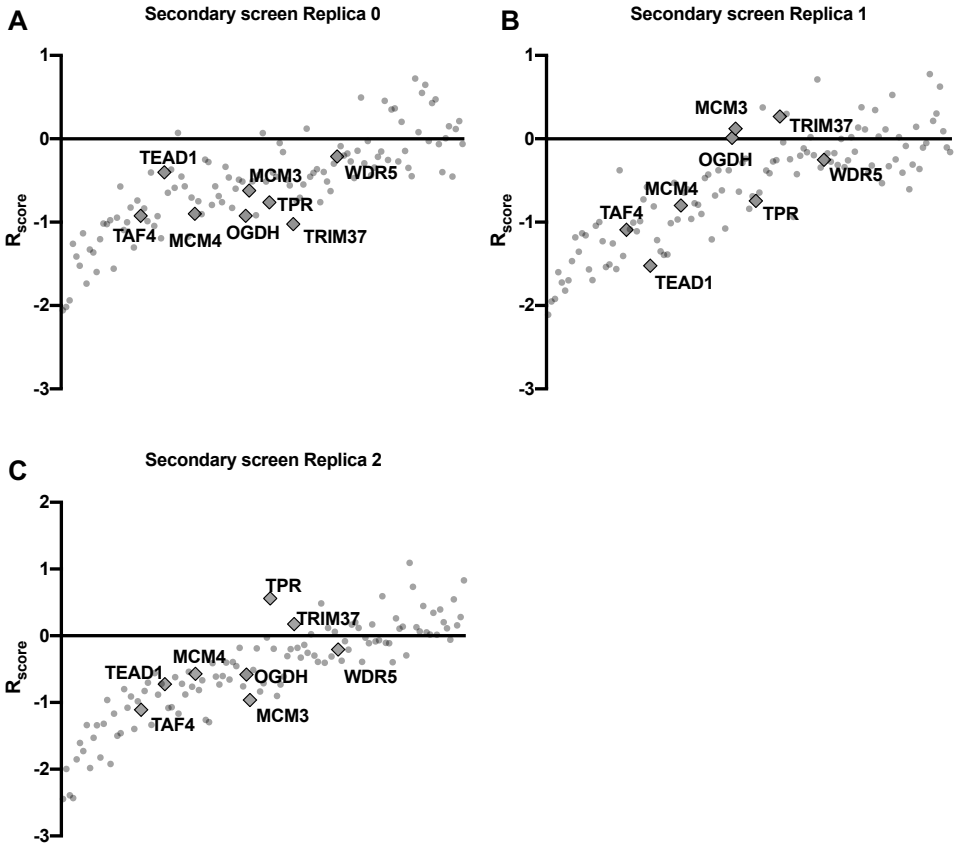


Figure 3.2. **Dot-plot chart showing the R-score for each gene screened in HeLa cells ordered by rank product.** Each first (A), second (B) and third (C) repeat is shown. The selected genes MCM3, MCM4, OGDH, TAF4, TEAD1, TPR, TRIM37 and WDR5 are highlighted and labelled. Negative values of Rscore denote radiosensitisation. Those genes whose knockdown reduced the plating efficiency below 10% were excluded from the analysis.

Gene symbol	Full name	Secondary screening ranking
MCM3	Minichromosome maintenance complex component 3	56th
MCM4	Minichromosome maintenance complex component 4	40th
OGDH	Oxoglutarate Dehydrogenase	55th
TAF4	TATA-box binding protein associated factor 4	24th
TEAD1	TEA domain transcription factor 1	31st
TPR	Translocated promoter region, Nuclear basket protein	62nd
TRIM37	Tripartite motif containing 37	69th
WDR5	WD repeat domain 5	82nd

Table 3.1. **Druggable targets chosen for further investigation.** The table shows the selected hits for validation and the secondary screen ranking value for each siRNA, where 1st corresponds to the most radiosensitising siRNA. The ranking is based on rank product calculation of the three screening replicates.

In order to assess whether the depletion of the selected genes (Table 3.1) modified the response to ionising radiation, HeLa and MRC-5 cells were transfected with the indicated siRNAs. After 72 hours, cells were lifted, counted and resuspended at the required concentration for a colony forming assay (CFA) plated in a 6-well plate and irradiated at 6 Gy once the cells had attached. About two weeks after radiation treatment, colonies were stained and counted. The surviving fraction was determined and plotted in a bar graph to compare each

siRNA to the non-targeting control in both HeLa and MRC-5 cells (Figure 3.3 and Figure 3.4, respectively).

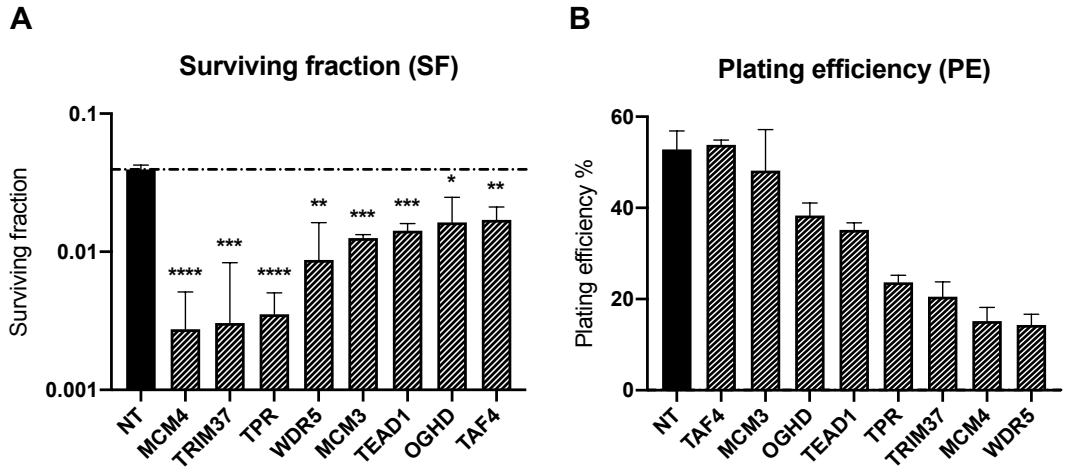


Figure 3.3. **Validation of the selected genes in HeLa cells.** Cells were transfected with the indicated siRNA. 72 hours later, cells were trypsinised, counted and seeded in fresh medium and irradiated at 6 Gy. Colonies were stained and counted after 14 days. Surviving fraction values are plotted in A, their relative plating efficiency values are shown in B. The dashed line in A shows the surviving fraction of the non-targeting siRNA (NT) (black column), in order to compare it to the surviving fraction of each indicated siRNA. Unpaired t-test was used to calculate the statistical significance of the difference in surviving fraction between NT and the siRNAs indicated. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

At first, HeLa cells were tested to confirm by colony formation assay in a 6-well plate format that MCM3, MCM4, OGDH, TAF4, TEAD1, TPR, TRIM37 and WDR5 depletion was causing radiosensitisation (Figure 3.3). All the selected targets from the siRNA druggable gene screen were confirmed in HeLa cells as radiosensitisers (Figure 3.3), as all of them showed a statistically significant radiosensitisation effect compared to the non-targeting control at 6 Gy dose, further confirming the quality of the original screen. The fold change decrease of

the surviving fraction of each siRNA compared to the non-targeting control was calculated and shown in Table 3.2.

siRNA	Surviving fraction at 6 Gy	Fold changes compare to NT
NT	3.9%	
MCM4	0.3%	-14.35
TRIM37	0.3%	-12.93
TPR	0.4%	-11.20
WDR5	0.9%	-4.52
MCM3	1.3%	-3.14
TEAD1	1.4%	-2.77
OGHD	1.6%	-2.42
TAF4	1.7%	-2.32

Table 3.2. **Surviving fraction at 6 Gy and relative fold change for each siRNA compared to the non-targeting control in HeLa cells.** Greater fold change reduction indicates greater radiosensitisation effect.

MCM4, TRIM37 and TPR showed a greater degree of radiosensitisation compare to the other's gene knockdown. Therefore, these genes were considered for further investigation.

Additionally, the candidate genes were tested by CFA in MRC-5 cells to assess whether their radiosensitisation effect was tumour-specific (Figure 3.4).

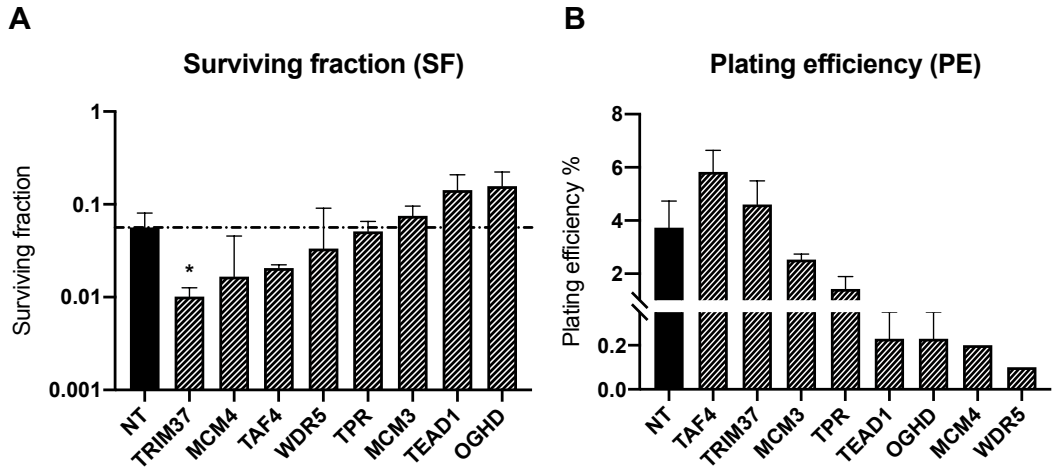


Figure 3.4. **Validation of the selected genes in MRC-5 cells.** Cells were transfected with the indicated siRNA. 72 hours later, cells were trypsinised, counted and seeded in fresh medium and irradiated at 6 Gy. Colonies were stained and counted after 14 days. Surviving fraction values are plotted in A, their relative plating efficiency values are shown in B. The dashed line in A shows the surviving fraction of the non-targeting siRNA (NT) (black column), in order to compare it to the surviving fraction of each indicated siRNA. Unpaired t-test was used to calculate the statistical significance of the difference in surviving fraction between NT and the siRNAs indicated. *$p < 0.05$.

Among the genes analysed, as shown in Figure 3.4, all but TRIM37 were shown not to radiosensitise MRC-5 cells. Although the siRNA targeting MCM4, TAF4 and WDR5 showed a lower surviving fraction at 6 Gy compared to the non-targeting control, this difference was not statistically significant. Finally, by analysing the plating efficiency of each siRNA, it was possible to identify genes whose depletion caused cell death independently from the radiation treatment (Figure 3.4B). Those genes whose depletion caused a decrease of the plating efficiency below 1% (TEAD1, OGHD, MCM4 and WDR5) were excluded from further investigation.

Finally, the two candidate genes resulting to be the most interesting ones to be validated were TPR and TRIM37. Indeed, TPR depletion radiosensitised HeLa but not MRC-5 cells, suggesting it could be a potential tumour-specific radiosensitisation target. TRIM37 was also considered an interesting target, as its knockdown strongly radiosensitised both HeLa and MRC-5 cells, suggesting it could play an important role in radiation response.

To assess whether TPR and TRIM37 could be attractive clinical targets, their differential expression in tumour and normal tissue was investigated by using Oncomine 4.5 database [51]. Differential expression between tumour and normal tissue makes a target more interesting for clinical approaches as increased or exclusive expression in tumours would allow a tumour-specific targeted therapy.

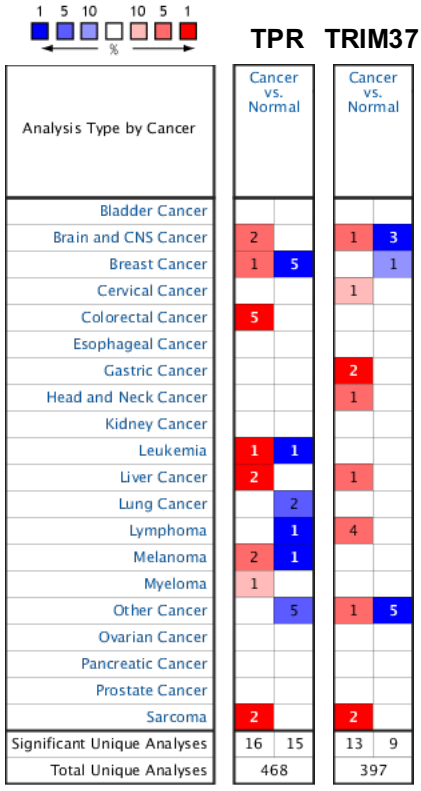


Figure 3.5. **Oncomine 4.5 Database disease summary for TPR and TRIM37.** Data were obtained from Oncomine web platform [51]. Data represent the up-to-date collection of all the published works containing microarray data about the enquired gene. Red colour represents high and blue low mRNA expression of cancer versus normal tissue. Each analysis also indicates the number of significant studies in which the listed gene had differential expression in cancer compared to normal tissue.

TPR was found overexpressed in 8, and underexpressed in 5 within the 20 tumour types analysed. TRIM37 was found to be overexpressed in 8, and underexpressed in 3 types of cancer. Both genes thus showed a certain degree of overexpression in cancer compared to normal tissue.

The remainder of this chapter will focus on further validation work performed for these two genes.

3.2.2 Validation of TPR as a novel radiosensitiser

TPR (translocated promoter region), also known as nuclear basket protein, is a 2094 amino-acids coiled-coil protein that makes part of the nuclear pore complex [52]. TPR function has not been fully defined yet, but it has been shown to be involved in the spindle assembly checkpoint [53], [54]. Besides, TPR mutations have been associated to increase radiation sensitivity in a genomic feature screen [55]. In light if these observations, TPR's role in radiation response was further investigated.

The first step to confirm TPR as a novel radiosensitiser was to verify that its depletion was able to radiosensitise additional cancer cells lines other than HeLa. Therefore, colony formation assays were performed in a panel of cancer cell lines. Among them, T24 (bladder carcinoma), BT-549 (breast ductal carcinoma) and DU145 (prostate cancer) were tested at three radiation doses, 2, 4, and 6 Gy and showed a significant radiosensitisation upon TPR depletion (Figure 3.6). TPR knockdown reduced the clonal growth of T24 and BT-549 cells (Figure 3.6B and E) also in the absence of radiation, whereas it increased the colonies formation of DU145 cells (Figure 3.6H). These data show that TPR depletion is able to increase the sensitivity to radiation in different tumour types.

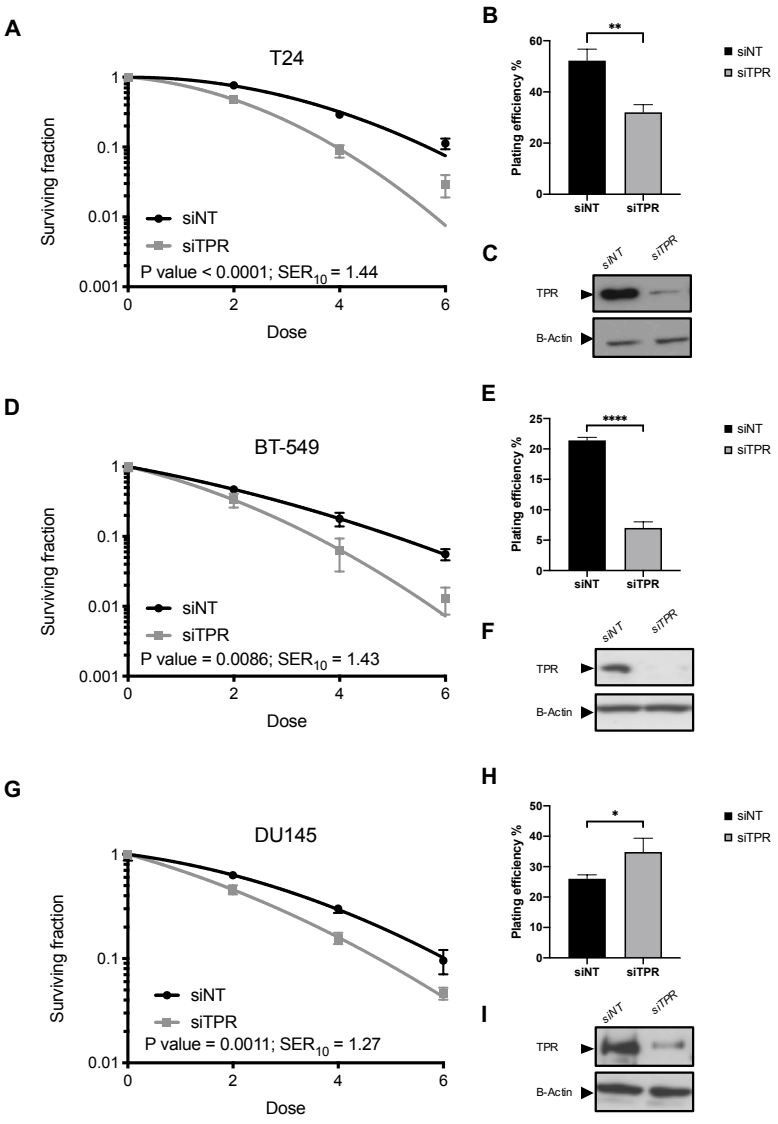


Figure 3.6. **TPR depletion radiosensitise different cancer cell lines.** Colony formation assays performed after TPR KD in different cancer cell lines: T24 (A), BT-549 (D), DU145 (G) cells. 72 hours after siRNA transfection, cells were lifted, plated and irradiated. Colonies were stained about two weeks after IR according to the growth rate of each cell line. The figure shows plating efficiency (B, E, H) and western blot analysis (C, F, I) to confirm TPR knock-down to the right of each CFA experiment's plot. β-Actin was used as loading control. All data are representative of three independent experiments and presented as mean +/- SD from triplicate wells. p-values were generated by factorial 2-way ANOVA. SER10 = Sensitisation Enhancement Ratio at 0.1 surviving fraction calculated from non-linear regression of linear quadratic model. Unpaired t-test was used to calculate the statistical significance of the difference in plating efficiency between NT and the siRNAs indicated. *p < 0.05, **p < 0.01, ***p < 0.001.

The following step was to confirm that the identification of TPR was not due to siRNA off-target effects. Therefore, a CFA experiment was performed in HeLa cells using three distinct strands of siRNAs (siTPR #1, siTPR #2, siTPR #3) targeting TPR gene transcript in different regions. All three siRNAs tested showed significant radiosensitisation of HeLa cells (Figure 3.7). Also, each siRNA produced a comparable and sufficient degree of TPR protein depletion, as shown in the immunoblot analysis in Figure 3.7 F. Overall, these results show that TPR knockdown increases the radiosensitivity of HeLa cells using three independent siRNAs suggesting that the radiosensitisation effect is due to TPR depletion itself and not because of an off-target effect.

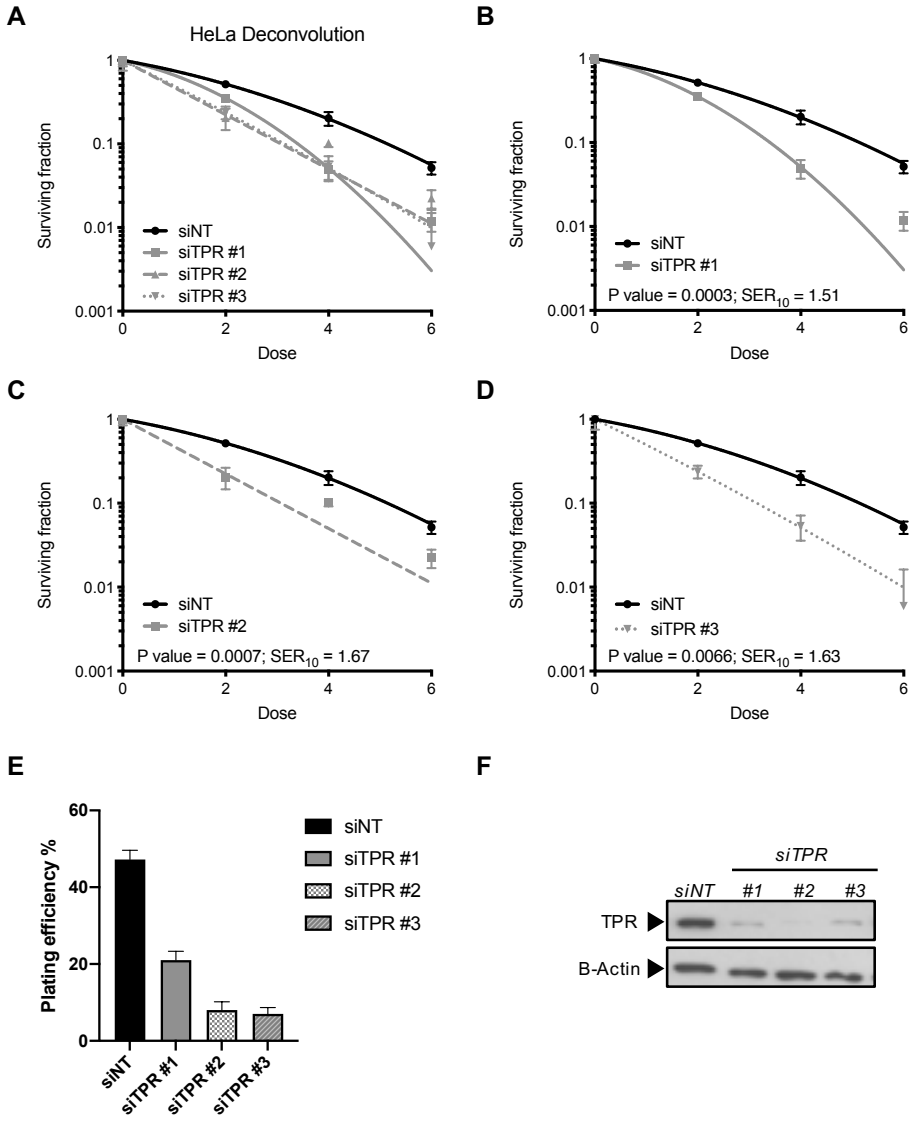


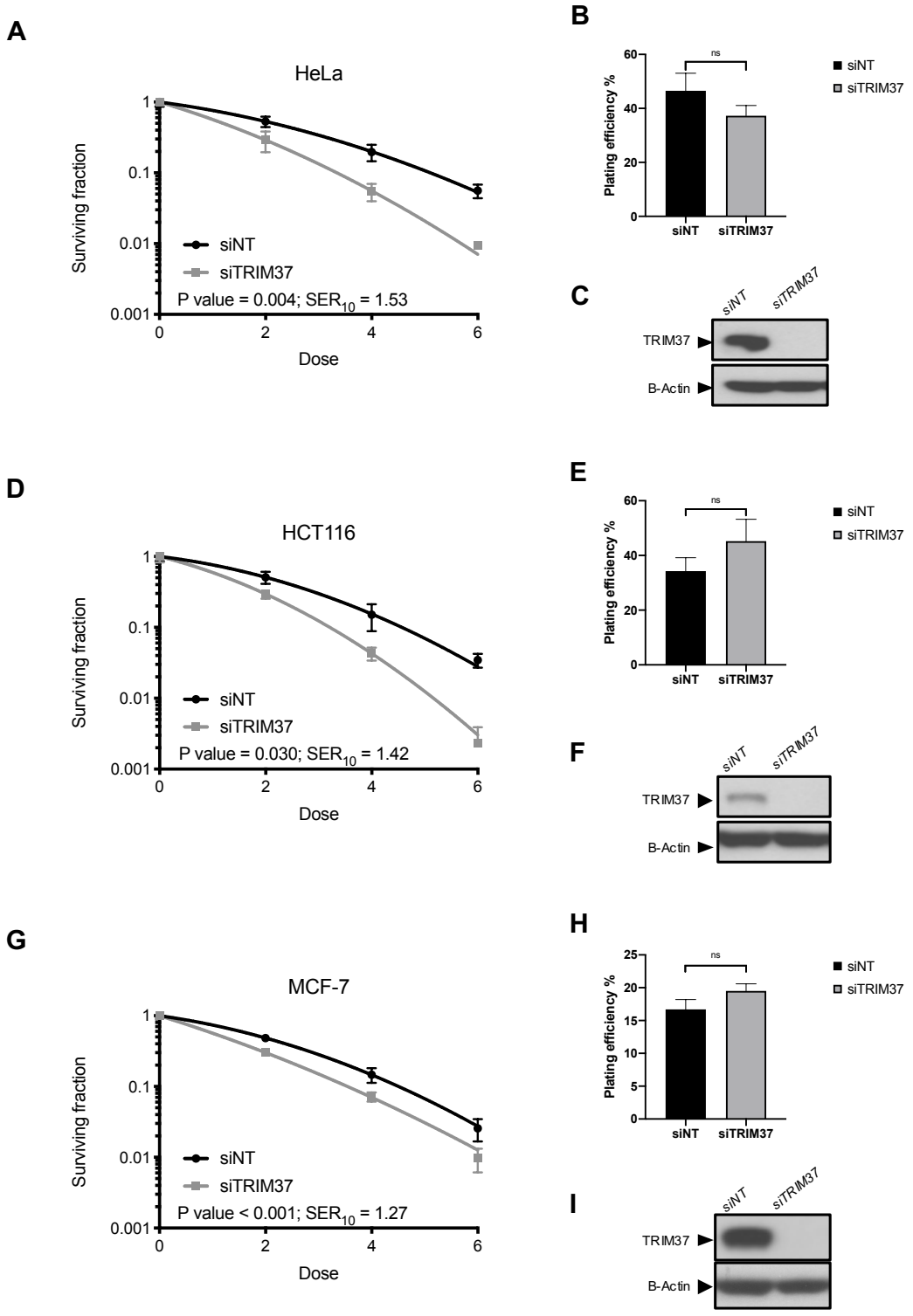
Figure 3.7. **Colony formation assays performed after TPR KD in HeLa cells using three different strands of siRNAs.** Panel A shows the deconvolution experiment. Panels B, C and D show the surviving fraction for each indicated siRNA strand. (E) Plating efficiency of each siRNA strand. (F) Western blot analysis to confirm TPR knockdown by all siRNA strands. All data are representative of three independent experiments and presented as mean +/- SD from triplicate wells. p-values were generated by factorial 2-way ANOVA. SER10 = Sensitisation Enhancement Ratio at 0.1 surviving fraction calculated from non-linear regression of linear quadratic model.

3.2.3 Validation of TRIM37 as a novel radiosensitiser

TRIM37 (tripartite motif containing 37), also known as MUL, is an E3 ubiquitin ligase belonging to the TRIM family proteins [56]. TRIM37 is constituted by 964 amino acids and its molecular mass is 108 kDa. TRIM37 protein consists of three domains: the RING-finger domain, the B2 Box and coiled-coil region, and the MATH domain (Meprin and TRAF-homology domain) [57]. There are several studies showing that TRIM37 is overexpressed in tumour tissues compared to the surrounding normal ones, and its overexpression has been correlated to poor prognosis and cancer progression [58]–[67], making TRIM37 an attractive target for cancer therapy. Moreover, it has been recently shown the contribution of TRIM37 in the activation of the NF-κB pathway after genotoxic stress [68], confirming the screening results suggesting TRIM37 as a regulator of IR cellular response.

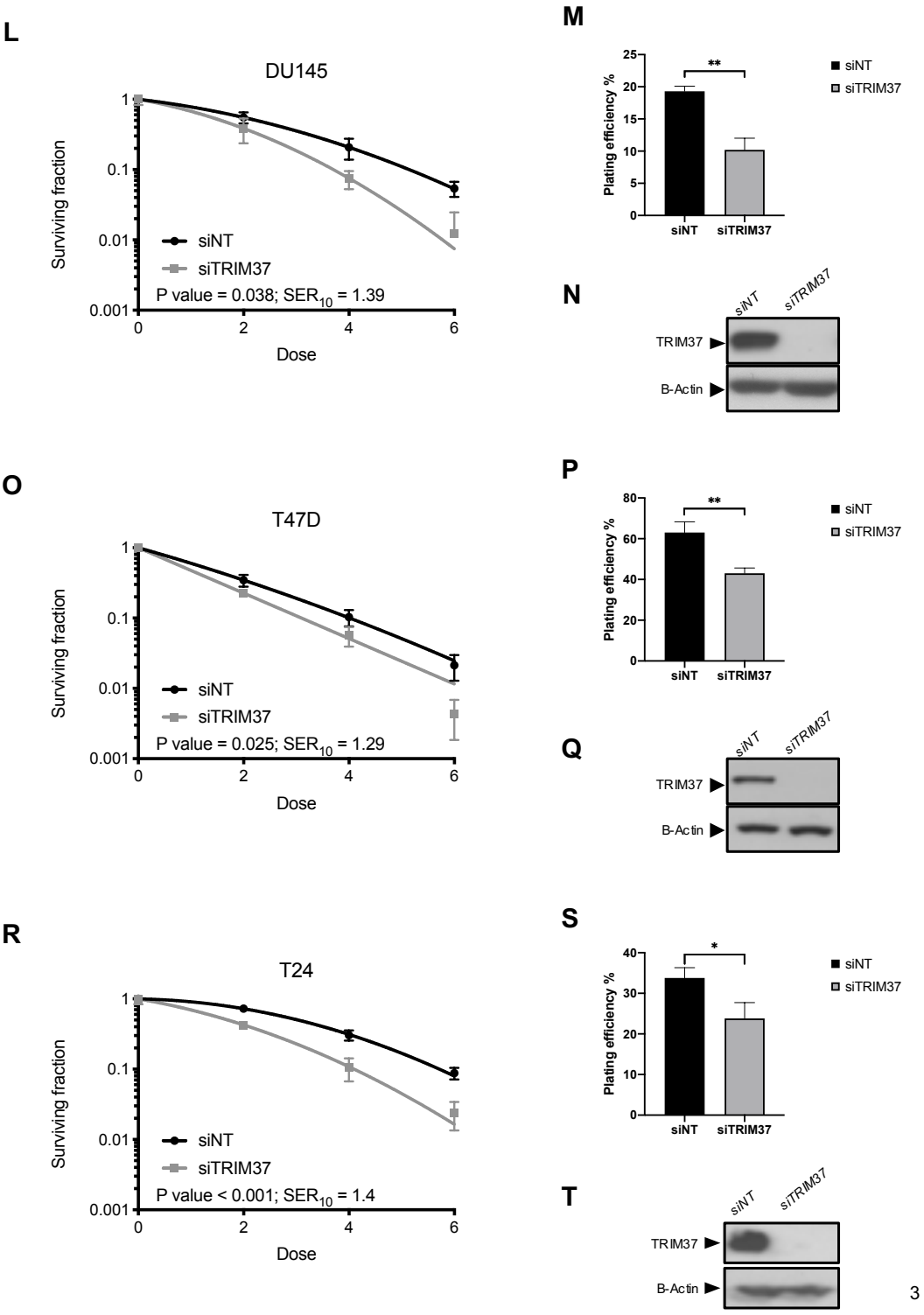
Therefore, the novel role of TRIM37 as a molecular determinant of radiosensitivity was validated. Again, the first step was to verify that TRIM37 depletion radiosensitised cell lines other than HeLa and MRC-5. CFA experiments were performed in a wide panel of cancer cell lines, including HCT116 (colorectal cancer), MCF-7 (breast carcinoma), DU145 (prostate cancer), T47D (ductal carcinoma), and T24 (bladder carcinoma) cells. The siRNA depletion of TRIM37 produced efficient protein knockdown and caused a significant radiosensitisation in the cell lines shown in Figure 3.8. For HeLa, HCT116 and MCF-7 cells, TRIM37 knockdown did not have a significant effect on the colony formation at 0 Gy (Figure 3.8 B, E and H). Whereas for DU145, T47D and T24 cells, TRIM37 depletion caused a decrease of clonal growth (Figure 3.8 M, P and S). These results showed that targeting TRIM37 increases

the radiosensitivity of a wide panel of cancer cell lines and several solid tumour types.



1,2

[1] Figure continues on next page
[2] See figure legend on page 61



[3]

[3] See figure legend on next page

Figure 3.8. **TRIM37 depletion radiosensitise a wide panel of cancer cell lines.** Colony formation assays performed after TRIM37 KD in different cancer cell lines: HeLa (A), HCT116 (D), MCF-7 (G), DU145 (L) and T47D (O), T24 (R) cells. 72 hours after siRNA transfection, cells were lifted, plated and irradiated. Colony formation was analysed about two weeks after IR according to grow rate of each cell line. The figure shows plating efficiency (B, E, H, M, P, S) and western blot analysis (C, F, I, N, Q, T) to confirm TRIM37 knock down to the right of each CFA experiment. β-Actin was used as loading control. All data are representative of three independent experiments and presented as mean +/- SD from triplicate wells. p-values were generated by factorial 2-way ANOVA. SER10 = Sensitisation Enhancement Ratio at 0.1 surviving fraction calculated from non-linear regression of linear quadratic model. Unpaired t-test was used to calculate the statistical significance of the difference in plating efficiency between NT and the siRNAs indicated. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

The final step to validate TRIM37 as a novel radiosensitivity modulator, was to confirm the absence of siRNA off-target effects. Therefore, three distinct siRNA strands, targeting TRIM37 transcript in different regions, were tested using CFA experiments in HeLa cells. The three different siRNA strands produced equally efficient protein knock-down and caused a similar radiosensitisation effect as shown in Figure 3.9, showing overlapping surviving fraction curves. Taken together, these results suggest that the radiosensitisation observed upon TRIM37 depletion is due to a gene-specific effect.

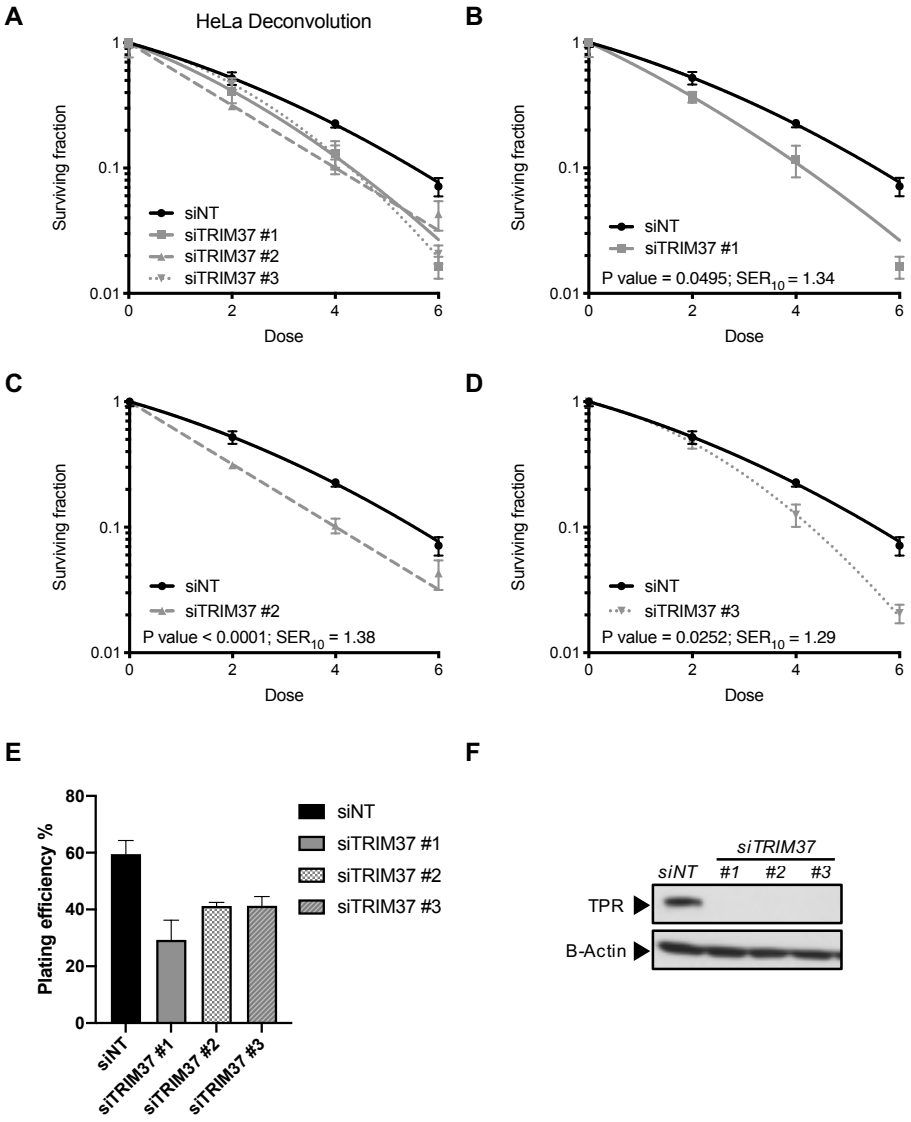


Figure 3.9. **Colony formation assays performed after TRIM37 KD in HeLa cells using three different strands of siRNAs.** Panel A shows the deconvolution experiment. Panels B, C and D show the surviving fraction for each siRNA strand. (E) Plating efficiency of each siRNA strand. (F) Western blot analysis to confirm TRIM37 knockdown by all siRNA strands. All data are representative of three independent experiments and presented as mean +/- SD from triplicate wells. p-values were generated by factorial 2-way ANOVA. SER10 = Sensitisation Enhancement Ratio at 0.1 surviving fraction calculated from non-linear regression of linear quadratic model.

3.3 Discussion

Our laboratory focuses on the identification of novel radiosensitisation targets by using high-throughput screen (HTS). In this chapter, an arrayed radiosensitivity RNAi screen in 96-well plate format has been presented. This screen's approach is based on the ability of single cells to form colonies after a specific gene knockdown combined with ionising radiation. A total of 9,962 genes was screened in a first stage, allowing the identification of 240 candidate genes that were subjected to a secondary screen, run in triplicate.

High-throughput screens are powerful tools, commonly used in medical research to identify novel therapeutic targets [46], [69]–[71]. Yet, the list of identified candidate genes requires validation because of the possibility of false-positive or false-negative target discovery. Regarding high-throughput siRNA-based screenings, their limitation is that only HeLa cells have been found to be suitable to form countable colonies in a 96-well plate. Therefore, only the genetic background of this specific cell line can be explored. Also, because cells need to be transfected with each siRNA pool individually, the more targets are investigated, the more labour-intensive and time-consuming the screen becomes. Finally, siRNA knockdown allows only transient and partial depletion of the targeted gene because it works by promoting the degradation of its mRNA transcript, and not of the gene or protein itself. Therefore, the protein depletion is subjected to transfection efficiency, mRNA and protein half-life, and experimental variability.

As stated above, validation work of the identified targets is a crucial step of high-throughput screenings. At first, our laboratory performed validation work of

those targets highly ranked and statistically significant by rank product analysis. Rank product analysis was performed using all three secondary screen's replicates. It was observed that many of the genes listed in the ranking just below the significance threshold (p-value = 0.05) had one outlier among the three replicates. Because of the high rate of false-negative target discovery proper of HTS and the limitation of siRNA-based screenings explained above, I decided that these genes still warranted further investigation.

MCM3, MCM4, TAF4, TEAD1, TPR, TRIM37 and WDR5 were selected among the top 100-hits for further investigation and validation based on their novelty in the context of radiation, their differential expression in normal tissues compared to cancerous ones, the availability of reagents, and their clinical impact on cancer disease.

Validation work of all the selected genes, confirmed that their depletion caused radiosensitisation of HeLa cells. Among the 7 hits investigated, MCM4, TRIM37 and TPR showed the greatest degree of radiosensitisation. Beside HeLa cells, a non-cancerous lung-derived fibroblast cell line (MRC-5), was tested for radiosensitivity upon depletion of the selected targets. In this case, the aim of the experiment was to find targets whose knockdown showed absence of radiosensitisation in MRC-5 cells. MCM4 knockdown had a variable effect on MRC-5 cells radiosensitivity, and its depletion caused a strong reduction of the plating efficiency, indicating a negative effect on cell survival independent from radiation. Consequently, it was excluded from any further investigation. TPR depletion had no effect on radiation sensitivity of MRC-5 cells, indicating its potential as a tumour-specific radiosensitiser. Also, TPR knockdown had only

minor effects on MRC-5 cells plating efficiency, making it an interesting target to further investigate in radiation response context. Finally, TRIM37 depletion did not modify MRC-5 cells plating efficiency. However, it increased their radiosensitivity to some extent, meaning that TRIM37 expression is not essential for cell survival, but critical for radiation response. Consequently, TRIM37 was also considered for additional analysis in the context of the response to radiation.

The discovery of novel radiosensitivity modulators is crucial to develop innovative therapeutic strategies and better understand the cellular response to radiation damage. Therefore, the clinical impact of a novel radiosensitisation target needs to be assessed by testing it on different type of tumours. The following step after the identification of TPR and TRIM37 as novel radiosensitisation targets, was to establish their clinical potential and importance in radiation response in a panel of cancer cell lines. TPR radiosensitisation effect was tested on a panel of cancer cell lines. Among them, TPR depletion caused an increase of radiosensitivity in T24 (bladder carcinoma), BT-549 (breast ductal carcinoma) and DU145 (prostate cancer) cells. However, other cell lines such as HCT116, H460 and A549, did not change their sensitivity to IR when TPR expression was abrogated (data not shown). TRIM37 depletion caused increased radiosensitivity of a larger group of cell lines – HCT116 (colorectal cancer), MCF-7 (breast carcinoma), DU145 (prostate cancer), T47D (ductal carcinoma), and T24 (bladder carcinoma) cells – confirming its theorised importance in radiation response.

The final stage to validate these two targets, was to assess the absence of off-target effects. Therefore, both TPR and TRIM37 were depleted using three

distinct siRNA strands targeting the mRNA transcript in different regions. For both genes, the three distinct siRNA strands produced a significant radiosensitisation in HeLa cells beside protein depletion, indicating that the identification of these target was specific to their depletion.

In light of these findings, both TRIM37 and TPR seemed to be interesting targets of radiosensitivity to be further explored in terms of identifying the mechanism by which their expression promotes radioresistance. However, I decided to further investigate TRIM37 rather than TPR because, as shown in the next chapter of this thesis, TRIM37 was identified by another independent screening I performed as part of my DPhil project.

TRIM37 is an E3 ubiquitin ligase belonging to the TRIpartite Motif protein family (TRIM). Mutations in TRIM37 gene result in a rare autosomal recessive disorder known as MULIBREY nanism (muscle-liver-brain-eye) [72]. Patients affected by MULIBREY nanism, together with the mouse model of the disease, have been previously described to have an higher risk of developing tumours, suggesting that TRIM37 is important to maintain genetic stability [58], [59], [61], [64]. TRIM37 has been reported to be overexpressed in many tumours type, such as non-small-cell lung cancer, colorectal cancer, breast cancer, hepatocellular carcinoma, glioblastoma and osteosarcoma, and its overexpression has been linked with tumour progression and aggressiveness [58]–[67]. High expression of TRIM37 is also associate with poor prognosis [60], [64]. TRIM37 has been identified as a breast cancer oncoprotein as for its overexpression in many breast cancer types [64]. The overexpression of TRIM37 in breast cancer causes an increased monoubiquitylation of histone H2A on lysine 119 associated with

transcriptional repression of a subset of tumour suppressor genes [64]. Moreover, it has been recently shown the contribution of TRIM37 in the activation of the NF-κB pathway after genotoxic stress [68], confirming our findings of TRIM37 as a regulator of IR cellular response.

Chapter 4 IDENTIFICATION OF NOVEL DETERMINANTS OF RADIOSENSITIVITY THROUGH A GENOME-WIDE CRISPR KNOCKOUT SCREEN

4.1 Introduction

The identification of novel determinants of radiosensitivity represents one of the main purposes of our laboratory. Our group has performed siRNA-based radiosensitivity screens in the past, leading to identification of genes previously unknown to be associated with radiosensitivity [38]–[40].

Arrayed siRNA-based screens are commonly used tools in functional genomics. They rely on reducing endogenous protein expression to investigate genetic LOF. However, RNAi can only partially suppress gene expression for a limited time and often causes off-target events [70]. In addition, this approach is often not suitable for slow-replicative cells and requires automation and specialised equipment for testing multiple conditions, as needed for HTS. To overcome these issues, novel tools that rely on precise genome editing rather than gene expression level reduction have been developed. The CRISPR-Cas9 genome editing system is currently one of the mostly used methods to precisely achieve systematic gene depletion [48], [70], [73], [74], [46].

Therefore, it was decided to design a radiosensitivity genome-wide pooled screen exploiting the CRISPR-Cas9 genome editing system. This system uses the enzyme Cas9 site-specific nuclease activity, guided by a guide RNA (gRNA), to achieve gene knockout (KO) [73]. In this screen, a library of gRNAs was introduced into a cell pool to inactivate single genes in individual cells. Then, as the library construct contained the gene conferring puromycin resistance, stable knockout cells were selected by using the selection marker for one week. Next, the knockout pooled cells were treated with ionising radiation and cultured for about 30 days to allow the depletion of sensitised cell populations, and the

enrichment of resistant cell populations. Finally, the different gene knockout clones were identified through DNA sequencing of the gRNA integrated in the genome of each cell. By comparing the abundance of gRNAs at the beginning and at the end of the screen, it is possible to associate a particular gene inactivation to a change in radiosensitivity. A schematic representation of the strategy used for this screen is shown below in Figure 4.1.

The aim of this screen was to identify and validate those genes whose depletion causes intrinsic radiosensitisation in non-cancerous cells, which are devoid of mutated genetic background and would be difficult to use in siRNA-based screen. In this regard, the human GeCKO v2 genome-wide library has been used. The human GeCKO v2 library was designed by Zhang et al. to target about 20,000 genes, using 6 different gRNAs for each gene, targeting as many regions within the gene [48], [75]. The screen was performed using MRC-5 cells, a diploid human fibroblast cell line with a 46, XY karyotype derived from lung tissue. This chapter describes the results of this screen.

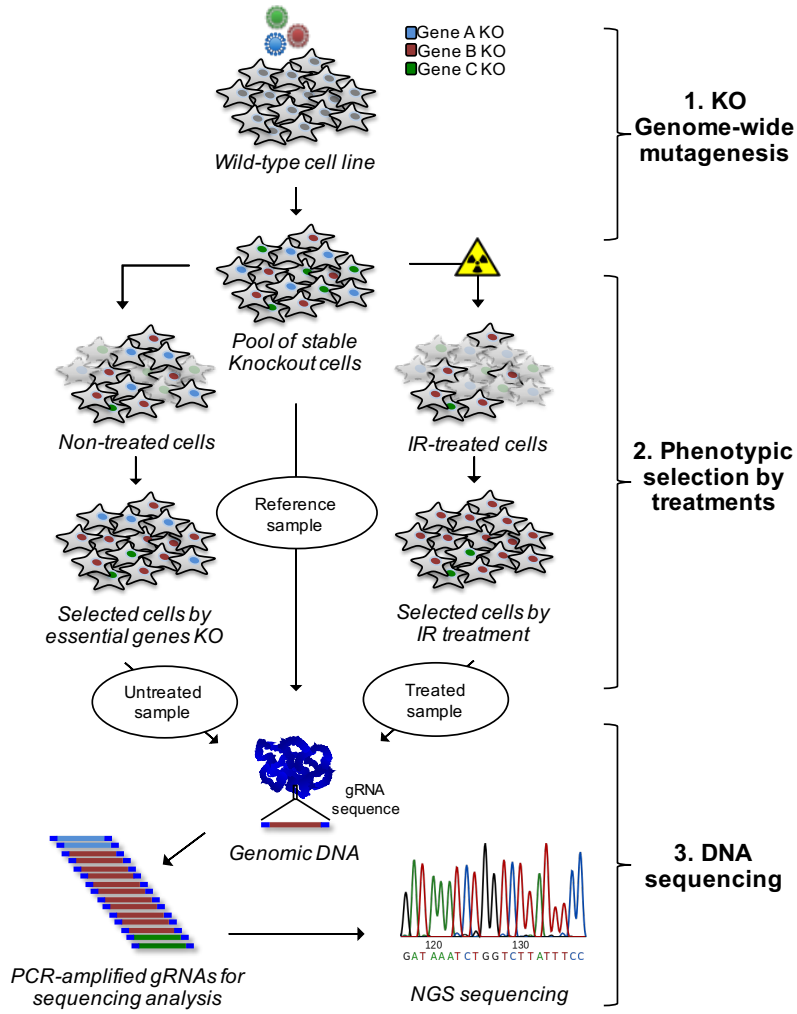


Figure 4.1. **Schematic representation of the CRISPR-Cas9 screen protocol.** At first, cells were transduced with the gRNA library. Then, transduced cells were divided into three samples: the reference one, to assess transduction efficiency at the beginning of the screen; the non-irradiated one, to interrogate the effect of gene depletion on its own; and the irradiated one, to analyse the effect of gene depletion combined with radiation. Treated and non-treated cells were cultured for 4 weeks to allow enrichment of selected gene knock-out clones. Finally, samples were collected and analysed by NGS sequencing. The screening protocol is represented using three genes (called A, B and C) as examples. Gene A knockout, shown in light blue, causes an increase of radiosensitivity, therefore the cells transduced with the correspondent gRNA disappear from the treated samples. Gene B knockout (in red) renders the cells resistant to radiation treatment, consequently the cells carrying the gRNA against gene B are enriched in the treated sample. Gene C knockout (in green) affects cells survival irrespective of radiation. Therefore, it is found poorly represented in both not treated and treated samples.

4.2 Results

4.2.1 Data analysis

In order to identify novel determinants of radiosensitivity, a pooled genome-wide screen in MRC-5 cells was performed using the CRISPR/Cas9 genome editing technology. This method allowed the interrogation of over 20,000 human genes in the contest of radiation response.

At first, MRC-5 cells were expanded to reach an adequate number in order to be able transduce each gRNA of the library in at least 500 cells. Also, to achieve single viral integration per cell a MOI of 0.3 was used, which resulted in about 70% of cell killing by the selection marker (puromycin). Considering the library consisted of 122,411 sgRNAs and the low MOI used, it was necessary to transduce over 220,000,000 cells. After one week of puromycin selection, the pool of stable KO cells was expanded again for the screening process. In order to obtain enough cells to reach a library coverage of 500 cells per sgRNA, the total amount of cells needed was about 200,000,000 cells. The KO cell pool was divided in three branches, one to be used as a non-treated control, one to be treated by IR and one to be used as reference sample (to verify the efficiency of the transduction). Before starting the phenotypic selection, the reference sample was collected and frozen for genomic DNA (gDNA) extraction, to be performed at the end of the screen simultaneously with the other samples. The IR treated arm was irradiated at 2 Gy. MRC-5 cells were successively allowed to grow for 4 weeks to enable phenotypic selection. Every 3 days they were lifted, mixed and reseeded at a confluency of 25%. Finally, cells were harvested and gDNA was extracted simultaneously from all the samples (reference, control and irradiated).

The genome-incorporated gRNAs were amplified by PCR and the amplicons were prepared for sequencing analysis by Illumina seq.

Negative selection screens rely on the depletion or reduction of certain gRNAs as a result of a perturbation, in this case the ionising radiation treatment and essential genes depletion. The final step of the screen consisted of DNA sequence analysis extracted from the three samples: reference, untreated-control and ionising radiation-treated. The gRNAs abundance was computed by using the web-based platform CRISPRAnalyzeR, which performs both NGS data extraction and mapping [76], [77]. The CRISPRAnalyzeR analysis generated a read count file for each sample that was successively analysed using R software.

At first, the read counts were transformed in log2 values. In order to dismiss those gRNAs not efficiently transduced before starting the screen, all the gRNAs represented less than 10 times in the reference sample were excluded from the analysis. Then, the gRNAs decreasing in the non-treated sample below 10 read counts were omitted from the analysis. This was done to avoid false-negative results, as the genes targeted by gRNAs depleted in the non-treated sample might be essential for cell survival independently from the effect of the radiation.

In order to identify those genes whose depletion causes radiosensitisation, the fold change between irradiated and untreated control samples was calculated as described in the Equation 4.1 below:

$$\boldsymbol{gRNA\ C_{score}} = [\log_2(gRNA\ abundance\ treated\ sample) - \log_2(gRNA\ abundance\ non\ treated\ sample)]$$

Equation 4.1. **Formula used to calculate the gRNA C_{score}.** The gRNA C_{score} is equal to the log2 fold change of the irradiated sample compared to the non-treated control one.

The gRNA CRISPR score (gRNA C_{score}) corresponds to the radiosensitisation effect of each gRNA determined by its targeted-gene depletion.

As each gene was targeted by 6 different gRNAs, the mean gRNA abundance for each gene was calculated using the Equation 4.2 below:

$$\boldsymbol{gene\ C_{score}} = \frac{\sum gRNA\ C_{score}}{n\ gRNA}$$

Equation 4.2. **Formula used to calculate the gene C_{score}.** The gene C_{score} is the mean log2 fold change of gRNAs targeting the same gene (n=6).

The gene CRISPR score (gene C_{score}) represents the radiosensitisation effect of each gene determined by its depletion.

To normalise these values, both *gene* and *gRNA C_{score}*, were transformed in z-scores using the following Equation 4.3:

$$\boldsymbol{Z\text{-}score} = \frac{C_{score} - \mu C_{score}}{\sigma C_{score}}$$

Equation 4.3. **Z-score calculation formula.** Data were normalised by z-score calculation. In the formula above, μ stands for mean and σ for standard deviation of the CRISPR scores.

4.2.2 Identification of novel determinants of radiosensitivity and gene selection

Once NGS data were extracted, mapped and normalised, it was possible to analyse the hits identified by the CRISPR screen.

At first, gRNAs where ranked according to their z-score. Due to the depletion of the targeted genes, lower values of z-scores are associated to an increased sensitivity to radiation, whereas higher values correspond to an increased resistance to radiation treatment. As the scope of this thesis is to identify novel targets of radiosensitivity, only gRNAs with negative values were considered for further analysis. Moreover, as the GeCKO library contained 2,000 non-targeting control guides, they were used to set thresholds in order to differentiate those gRNAs that had an effect on radiation response compared to the negative control. In Table 4.1, the top gRNAs associated with a radiosensitisation effect are shown. Notably, the second ranked gRNA targets TRIM37 gene, which had already been identified as a novel radiosensitivity target by the siRNA screen described in chapter 3.

Rank	gRNA	gRNA C_{score}	Z-score
1	**C11orf45**_HGLibB_05137	-7.09	-5.58
2	**TRIM37**_HGLibB_51414	-6.95	-5.46
3	**TLR7**_HGLibB_49602	-6.94	-5.45
4	**STK31**_HGLibB_47423	-6.77	-5.30
5	**FRY**_HGLibB_18030	-6.71	-5.25
6	**OR5H14**_HGLibB_34247	-6.66	-5.20
7	**DAPP1**_HGLibB_12320	-6.55	-5.11
8	**TEX30**_HGLibB_49034	-6.47	-5.03
9	**B3GNT2**_HGLibB_03982	-6.44	-5.01
10	**SLC34A2**_HGLibB_44887	-6.30	-4.89
11	**MSMP**_HGLibB_30188	-6.29	-4.87
12	**ANAPC4**_HGLibB_01880	-6.19	-4.79
13	**PEX6**_HGLibA_36123	-6.13	-4.73
14	**DAG1**_HGLibB_12279	-6.13	-4.73

Table 4.1. **The top depleted gRNAs ranked according to their z-score identified by the CRISPR screen.** In the gRNA column, the name of the gene targeted is written in bold, followed by the ID of the gRNA. The row corresponding to TRIM37 targeting gRNA is highlighted in light blue.

Then, the gRNA z-scores were plotted in ascending order in a dot plot chart shown in Figure 4.2. Two out of six gRNAs targeting TRIM37 were found with z-scores lower than three standard deviations (3σ) of the mean of the non-targeting control guides (μ(NT)). TRIM37 depletion promoted by these two gRNAs caused a drastic increase of MRC-5 cells sensitivity to radiation. The other four gRNAs targeting TRIM37 gene did not show substantial changes to radiation treatment, as their z-score values lie within two standard deviations (2σ) of the non-targeting control guides' mean. These results suggested that a

different gRNA sequence can have variable KO efficiency and therefore it can considerably change the cellular phenotype [78].

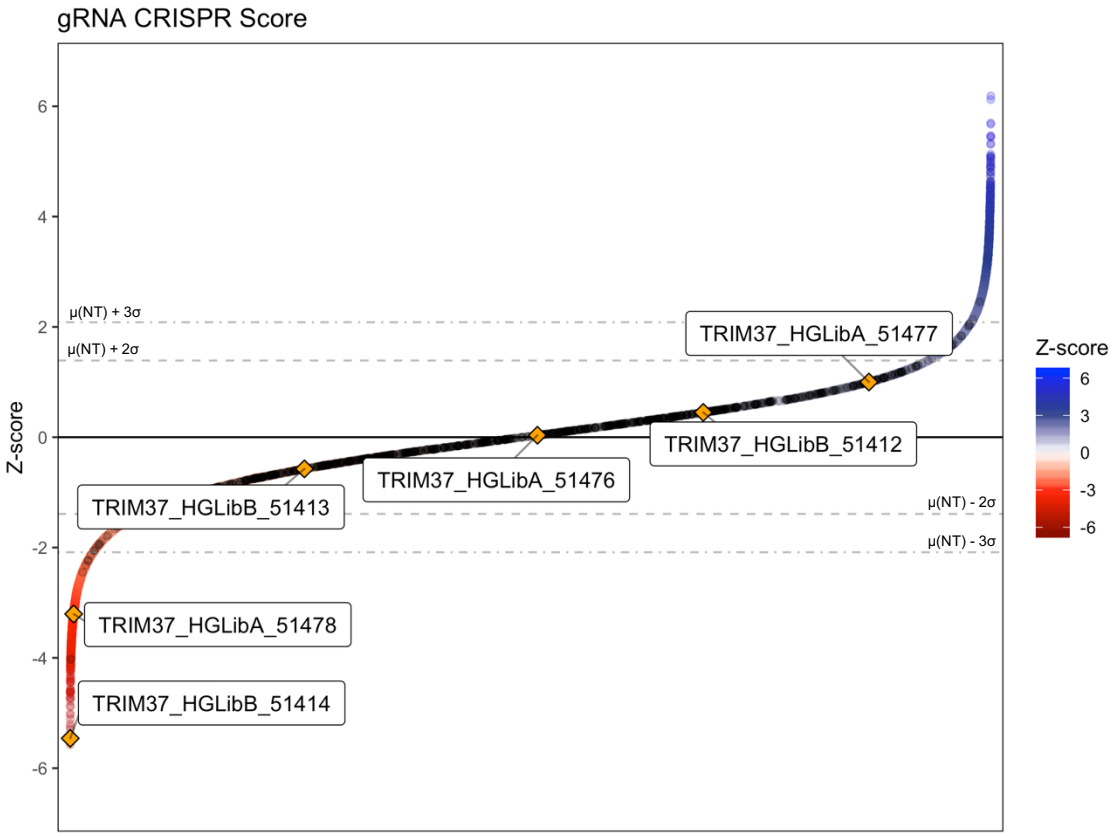


Figure 4.2. **Identification of genes involved into radiation response according to their gRNA C_{score}.** On the y axis each gRNA z-score is plotted in ascending order. The dot plot shows in red those gRNAs causing radiosensitisation and in blue the gRNAs associated with radioprotection effect. The non-targeting control RNA guides' Cscore, are shown in the plot coloured in black. The gRNAs targeting TRIM37 are highlighted in orange and labelled with the corresponding gRNA ID.

As stated above, the library adopted for the screen, has been designed to target each gene with 6 different gRNAs. The gene CRISPR score was calculated, and genes were ranked according to their z-score in ascending order. The top-30 hits discovered by analysing the gene CRISPR score are shown in Table 4.2. Among these genes, TRIM37 was found ranked 24th, indicating that

the 6 gRNAs targeting TRIM37, taken together, were likely to regulate radiosensitivity in MRC-5 cells. Gene z-score values and their standard deviation were plotted in a dot plot chart and are shown in Figure 4.3. Although TRIM37 showed a high standard deviation, reflecting the fact that only two out of six gRNAs had an effect on MRC-5 radiosensitivity, its z-score lie outside three standard deviations (3σ) from the mean of the non-targeting control guides. Notably, ATR gene, an already known target of radiosensitivity and crucial regulator of DNA repair, was also found with a significant low z-score.

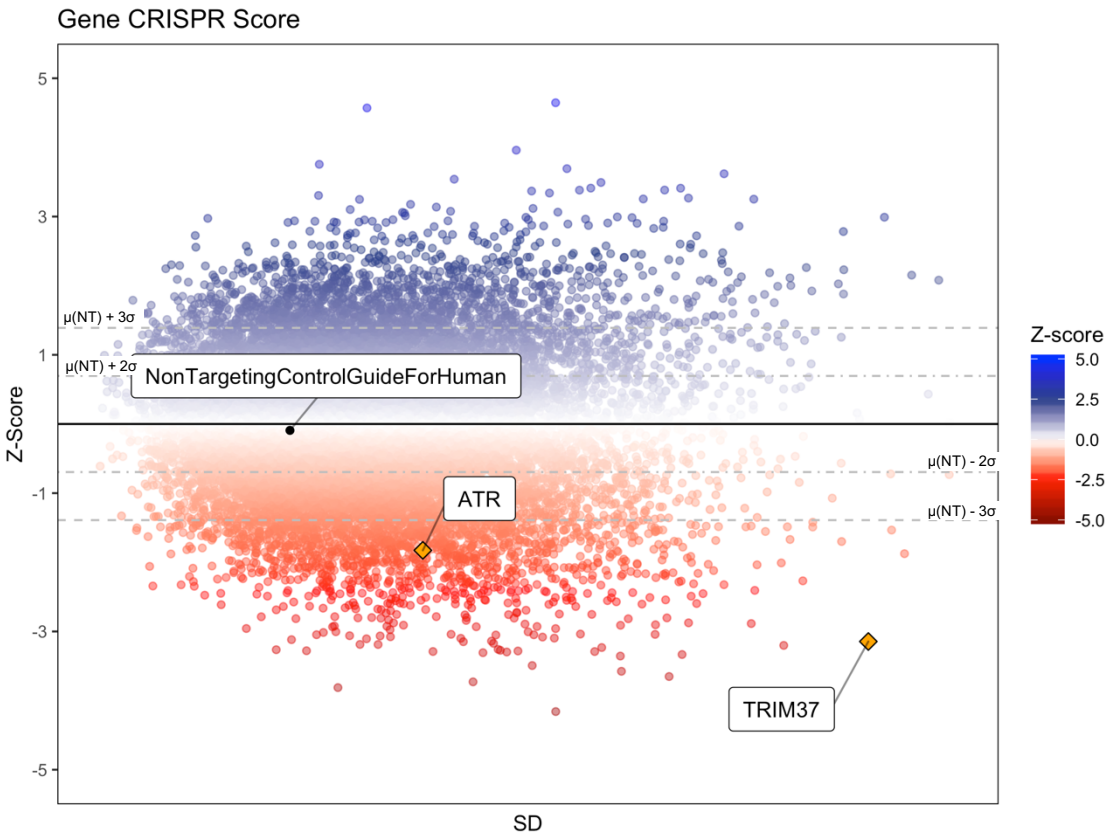


Figure 4.3. **Identification of genes involved into radiation response according to their gene C_{score}.** Dot plot showing the C_{score} of each gene included into the analysis. Genes whose knockout caused radioprotection are shown in blue, whereas those ones associated to an increase of radiosensitivity are shown in red. The increased intensity of the two colours denote a stronger effect. The non-targeting control guides' (negative control), ATR (positive control), and selected gene, TRIM37, are highlighted in the plot and labelled.

Rank	Gene name	gene C_{score}	SD	Z-score
1	KIF2C	-2.69	1.53	-4.16
2	PAXBP1	-2.53	0.85	-3.81
3	EIF3J	-2.49	1.27	-3.73
4	SRI	-2.45	1.89	-3.65
5	DMTF1	-2.42	1.74	-3.58
6	RSRC1	-2.38	1.46	-3.49
7	SEC61G	-2.32	1.73	-3.36
8	ANAPC4	-2.31	1.93	-3.33
9	HDDC2	-2.29	1.30	-3.31
10	ZNF552	-2.29	1.38	-3.30
11	SPANXN2	-2.29	1.65	-3.29
12	BOD1	-2.29	1.14	-3.29
13	NLGN3	-2.28	1.41	-3.28
14	FSD1L	-2.28	0.75	-3.28
15	FSIP2	-2.28	1.36	-3.27
16	PRAMEF7	-2.27	0.65	-3.26
17	MED27	-2.27	1.35	-3.26
18	MIPOL1	-2.26	1.68	-3.23
19	FBLN5	-2.24	2.25	-3.20
20	FOXI3	-2.24	0.97	-3.19
21	TMEM64	-2.24	0.92	-3.19
22	FHL2	-2.23	1.34	-3.17
23	NKX2-2	-2.22	1.24	-3.15
24	TRIM37	-2.22	2.51	-3.15
25	PWWP2A	-2.21	1.33	-3.14
638	ATR	-1.60	1.11	-1.83
8234	Non-targeting control	-0.79	0.70	-0.09

Table 4.2. **The top depleted genes ranked according to their z-score identified by the CRISPR screen.** The row corresponding to TRIM37 gene is highlighted in light blue. The non-targeting control and ATR are also shown with their corresponding ranking in the last rows of the table.

Ideally, the calculated CRISPR scores would be derived from biological replicates. However, we could not replicate the screen as it was beyond the scope of this study. Therefore, Rank product analysis was performed using each gRNA targeting the same gene as a replicate (n=6). Rank product analysis identifies differentially expressed genes based on statistical technique derived from biological reasoning [79]. This approach is commonly used for analysing large and noisy data sets, such as microarray experiments or high-throughput array screens.

The results of the rank product analysis are shown in Table 4.3 and the data are plotted in Figure 4.4. An extensive table including all the hits with a p-value < 0.05 can be found in the appendix section (Table A). Rank product analysis identified 762 targets of radiosensitivity using a p-value cut-off of 0.05. It confirmed the targets identified by analyses based on gRNA and gene C_{score} ranking where the non-targeting control distribution was used as a cut-off.

Rank product	Gene Name	p-value	Gene C_{score}	z-score
1	KIF2C	0.00005	-2.69	-4.16
2	DMTF1	0.00008	-2.42	-3.58
3	SRI	0.00021	-2.45	-3.65
4	SEC61G	0.00023	-2.32	-3.36
5	TRIM37	0.00024	-2.22	-3.15
6	ANAPC4	0.00030	-2.31	-3.33
7	PAXBP1	0.00038	-2.53	-3.81
8	B3GNT9	0.00045	-2.10	-2.89
9	NPBWR1	0.00057	-2.10	-2.90
10	MED27	0.00070	-2.27	-3.26
11	MIPOL1	0.00070	-2.26	-3.23
12	HABP4	0.00080	-1.62	-1.88
13	RSRC1	0.00084	-2.38	-3.49
14	HOXA4	0.00096	-2.06	-2.80
15	CLEC4E	0.00107	-2.19	-3.10
379	ATR	0.02490	-1.60	-1.83

Table 4.3. **The top 15 targets identified by the CRISPR radiosensitivity screen.** The table shows the top 15 genes identified ranked according to rank product analysis. Also, p-value, gene C_{score}, z-score and the ranking based on the z-score are shown for each gene. The row highlighted in light blue indicates TRIM37 gene, a gene previously identified as a radiosensitisation target by the siRNA screen presented in chapter 3.

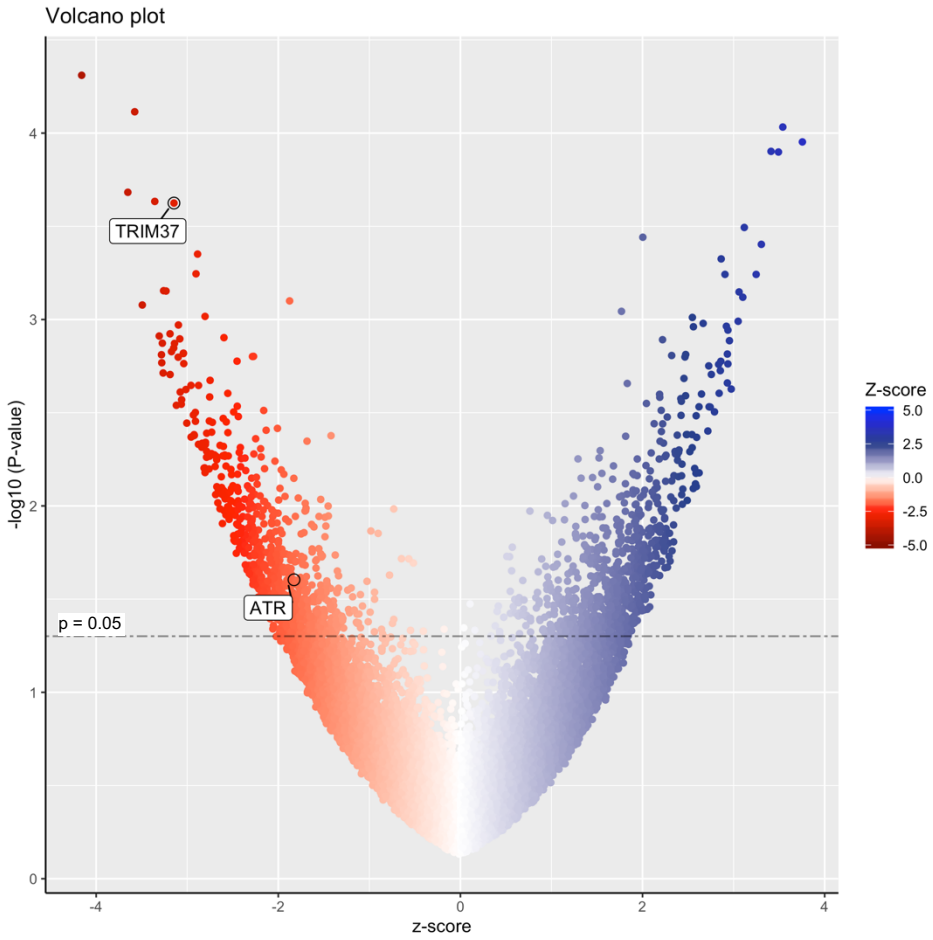


Figure 4.4. **Volcano plot showing the CRISPR radiosensitivity screen results.** The p-value and z-score of each gene analysed is plotted in the volcano plot above. The genes whose depletion causes an increase of cell viability are shown in blue, whereas the genes whose depletion causes radiosensitisation are shown in red. Genes with a p-value lower than 0.05 are found above the horizontal dot-dashed line. TRIM37 gene and ATR are labelled on the plot.

Once again, TRIM37 gene was identified as a determinant of radiosensitivity together with ATR gene. Therefore, TRIM37 gene, which has been already discovered as a novel target of radiosensitivity by an independent screen and validated in several cancer cell lines (data shown in chapter 3), was finally chosen for further investigation, aiming to identify the mechanism behind the radiosensitisation mediated by its depletion.

4.3 Discussion

High throughput screens like the one described in this chapter are potent tools for identifying novel determinants of radiosensitivity. To the best of my knowledge, the screen described here is the first genome-wide screen designed to identify novel radiosensitisers in MRC-5 cells and the first to use a CRISPR LOF pooled screening approach in combination with radiation.

Because of the risk of identifying cell line-specific or false positive targets, the cell line selection for the screening represented the first challenge of this work. In principle, a cancer cell line might appear as the best model for the purpose of identifying tumour-radiosensitisers. However, cancer cells are known to have abnormal numbers of chromosomes (aneuploidy), chromosomes structural abnormalities and genetic mutations [22], [32]. These alterations can interfere with the reliability of the CRISPR screen results because of several reasons. First, genetic mutations alter gene sequences which interfere with the efficiency of the Cas9 to recognise and target specific sequences within the genome. Then, genetic backgrounds characterised by multiple copies of the same gene represents a challenge in order to achieve complete KO of the desired gene. Additionally, mutated genetic backgrounds can lead to the consequent identification of cell line-specific hits.

Therefore, the choice of a standard system, such as a diploid cell line rather than a cancer one, allows to screen larger libraries with a lower risk to incur in false positive and cell line-specific hits. However, downstream validation of the identified targets in additional cell lines is always crucial, especially to translate the identified targets into clinically relevant context.

For these reasons, the non-cancerous MRC-5 cell line was selected for this screen. Their diploid 46, XY genome makes achieving complete KO easier compared to cancer cell lines, making it more likely to obtain reliable data from a CRISPR screen.

As there is no gold standard approach to analyse pooled CRISPR negative-selection screenings, the analysis was performed using three different approaches. At first, the logarithmic fold change abundance of each gRNA in the treated sample was compared to the non-treated one, denoted in this manuscript as the gRNA C_{score}, was calculated. Then, gRNAs were ranked according to these scores. This approach should ideally be confirmed by biological replicates, but additional replicates of the screen were beyond the scope of this work.

As the GeCKO library used for the screen was designed to target the same gene with 6 different gRNAs, we confirmed our findings by calculating the gene C_{score}, i.e. the mean value of the 6 gRNAs' C_{score} targeting the same gene. Although one expected the gRNAs targeting the same gene having similar phenotypes, the efficiency of the CRISPR KO can vary dramatically depending on the portion of the gene targeted or the secondary structure of the gRNA . [80], [81]. However, those genes with more than one gRNA found decreased or depleted after radiation, are more likely to be genuine radiosensitivity targets to be considered for further validation.

Finally, rank product analysis was performed, considering each individual gRNA as a replicate. Rank product analysis is a biologically motivated test commonly used for microarray experiment analysis, detecting gene differential expression between two experimental conditions by using ranks based on fold

changes [79]. It performs very well with large and noisy data sets and would therefore be a suitable test to be used for the CRISPR screen data analysis.

In conclusion, the three analysis approaches identified TRIM37 as a modulator of radiosensitivity. As TRIM37 had also been identified by the siRNA screen described chapter 3, it emerged as an even more interesting gene to further investigate.

The CRISPR genome-wide screening presented in this chapter was an extremely laborious and time-consuming work and, unavoidably, it had the limitation of being performed in a single biological replicate. However, the screen successfully identified TRIM37 among the top hits, which was extensively validated as a determinant of radiosensitivity. Conversely, many other known genes associated to radiation response did not appear in the top hits as expected. ATR gene itself, which depletion significantly radiosensitised MRC-5 cells (p-value = 0.025), was only ranked 379th in this screen.

The reason why known targets of radiosensitivity did not appear in the top hits of this study can be addressed to the combination of two highly variable methods: CRISPR/Cas9 genome mutagenesis and irradiation. The way CRISPR/Cas9 system generates gene KO depends on the randomness of the DNA DSBs repair mediated by NHEJ, which in turn induce DNA mutagenesis and the deletion of the targeted gene [82]. Then, cell death induced by irradiation has a known stochastic nature [83], meaning that there is a certain probability that clonal selection of irradiated cells can be random and not necessarily due to radioresistance. Finally, pooled screens only allow confirmation of genome gRNA

integration by sequence analysis, which does not necessarily demonstrate protein depletion.

The addition of multiple biological replicates is critical to achieve more reliable screening results, especially when the genetic perturbations are combined with radiation treatment. Indeed, when the non-treated control sample was compared to the reference one in order to identify genes whose depletion cause either cell death or increased proliferation, the expected hits showed highly ranked in the results (data shown in the appendix section, Table B and Table C). For example, TP53 was the first gene ranked among the ones whose depletion caused an increased proliferation, as well as GAPDH and RPL23A were found among the top ranked essential genes for MRC-5 cells survival. These data show that introducing an additional perturbation, such as ionising radiation, drastically increase the experimental variability. However, increasing the number of replicates can improve the screening outcome by reducing the probability of false-hits discovery.

Chapter 5 IONISING RADIATION RESPONSE IN TRIM37-DEPLETED CELLS

5.1 Introduction

In this study, TRIM37, an E3 ubiquitin ligase belonging to the TRIM family, has been identified for the first time as a modulator of radiosensitivity (Figure 3.1 and Figure 4.4). Interestingly, TRIM37 was identified by two independent screenings, performed in two different cell lines (HeLa and MRC-5 cells), and by using distinct approaches (siRNA array screening and CRISPR knockout pooled screening). Also, in this study, TRIM37 depletion has been shown to radiosensitise a wide panel of cancer cell lines (Figure 3.8). Furthermore, TRIM37 is overexpressed in tumour (Figure 3.5), and its overexpression has been correlated to poor prognosis and cancer progression [17–23], making TRIM37 an attractive target for cancer therapy. However, the mechanism by which TRIM37 depletion causes an increased radiosensitivity is still to be described.

In order to address this question, the effects of TRIM37 depletion on the cellular response to ionising radiation were investigated by using several different approaches.

Ionising radiation (IR) is a type of radiation commonly used for the treatment of patients with radiotherapy. IR rapidly induces cell damage by ionising different molecules within the cells, of which DNA is critical for cells survival. Indeed, DNA is present in only two copies, has a limited turnover and has a central role in all the cellular functions. IR generates different types of DNA damage, such as single-strand breaks (SSBs), damage to bases, double-strand breaks (DSBs) and DNA cross links [9]. As DNA DSBs represent the most lethal type of DNA damage, the cellular response to their occurrence is object of study of this

chapter. In normal conditions, DNA DSBs formation activate the DNA damage response (DDR), a highly complex and coordinated system that ensure the genome integrity by promoting DNA repair, cell cycle checkpoints activation and, eventually, cell death [6].

The first event to activate the DDR is the detection of the DNA damage, which in turn recruits at the damage site the proteins that are needed to repair the genetic insults. DNA DSBs can be repaired by two main pathways, non-homologous end joining (NHEJ) and homologous recombination (HR) [12]. NHEJ pathway joins two DNA ends together without using a template and therefore it is an error-prone repair that often results in permanent change to DNA [14]. HR repair pathway uses instead the homologous sister chromatid as template to repair DNA DSBs with high-fidelity. Because of this, it can only be used by cells when they are in S or G2 phase of the cell cycle [15], [84]. In contrast, NHEJ can be used by the cells in all the cell cycle phases. NHEJ is the predominant pathway in the G1 and G2 phase of the cell cycle, while HR repairs only 15–20% of the DSBs in late S and G2 phase [85]. The earliest proteins recruited at the damage site often serve as a platform to transduce the DDR signalling and assemble larger complexes that are the actual players of the DNA repair. Also, several post-translational modifications are involved at different levels of the DDR cascade, from DSB recognition to repair pathway choice. These include phosphorylation, ubiquitylation, sumoylation, acetylation or parylation [6].

Some of the DNA damage signalling proteins can be detected by immunofluorescence (IF) assays, and therefore can be used to monitor the DDR. The phosphorylation of histone H2AX at serine 139 residue (forming γ-H2AX)

and the localisation of 53BP1 at the damaged site are used to evaluate the levels of DSBs after IR treatment both in terms of early response (few minutes after IR) and in terms of persistent DNA damage (at later time points, when DNA DSBs are expected to have been repaired). 53BP1 localisation at the DSBs site also prevents DNA ends resection and therefore promotes DNA repair through the canonical NHEJ (c-NHEJ) over the alternative NHEJ (alt-NHEJ) and HR repair pathways [86], [87]. RAD51 represents another damage signalling protein which can be used to examine specifically HR DNA repair pathway. RAD51 is recruited on the single stranded DNA to promote the strand invasion and therefore HR. At this stage, RAD51 forms foci which can be microscopically visualised by IF analysis.

Another important process controlled by the DDR is the regulation of the cell cycle progression. A proliferation delay allows cells more time to repair DNA damage and prevent its propagation to the daughter cells, preventing in turn genetic instability [6]. The DDR activates cell cycle checkpoints in response to IR taking place at different stages within the cell cycle. The G1/S checkpoint prevents the cells irradiated in G1 to enter S phase with damaged DNA. Intra-S-phase checkpoint reduces the DNA synthesis rate and therefore slows down the progression over the S-phase. Both two G2/M checkpoints prevents the entry into mitosis and can be distinguished in “early” and “late” checkpoints. The “early” G2/M applies to the cells irradiated in G2 and results in a block of the cells at the end of G2 phase within short time after irradiation, preventing the entry into mitosis. In contrast, the “late” G2 checkpoint is observed at later time points after irradiation as a result of G1- or S-phase irradiated cells that escaped

the previous checkpoints. Finally, the spindle-assembly checkpoint (SAC) prevents mis-segregation of the chromosomes between two daughter cells.

Different flow cytometry assays have been developed and can be used to discriminate cell populations depending on their current cell cycle phase. For example, propidium iodide (PI) is commonly used to measure DNA content of each nucleus and therefore it detects whether a cell is in G1 or G2/M phase. Bromodeoxyuridine (BrdU) can be used to distinguish S phase cells based on the levels of incorporation in newly synthesised DNA. Phosphorylated histone H3 can be used to determine the number of cells in mitosis over a population.

When the DDR fails to activate cell cycle checkpoint and repair the DNA damage, programmed cell death is activated to prevent mitosis and the propagation of the damage to the daughter cells. However, the activation of apoptosis is not always effective, especially in cancer cells which often carry mutations in genes involved in the regulation of the apoptosis pathway [32]. Therefore, cells can undergo through mitosis with mis-repaired and unrepaired DNA, which in turn causes an impaired DNA segregation and aberrant mitosis. This process can lead to the loss of entire chromosomes or chromosomal portions and the physical inability to replicate and separate the genetic material correctly. The cell death that results from, or follows, this phenomenon is defined as mitotic catastrophe (MC). MC is morphologically associated with the accumulation of micronuclei, multinuclear cells and chromosomes aberration.

As MC is morphologically associated with the accumulation of micronuclei, multinuclear cells and chromosome aberration, it can be investigated by using IF assays. Indeed, by staining the DNA using DAPI and the cytoplasm by using

FITC, micronuclear and multinuclear cells can be scored and quantified in specific experimental conditions.

5.2 Results

5.2.1 Analysis of DNA damage response in TRIM37-depleted cells

The primary cellular response to DNA double-strand breaks (DSBs) consists of histone H2AX phosphorylation, which forms γ-H2AX foci at the damage site. H2AX phosphorylation is the first event triggering a cascade of events activating DNA damage response. Another early response to DSBs formation is the recruitment of 53BP1 at the damage site in the form of foci, promoted by monoubiquitylation of histone H2A. γ-H2AX foci are specific markers of DNA DSBs and their resolution are thought to correspond to DNA DSBs repair. As 53BP1 recruitment promotes the DNA DSBs repair by using NHEJ [86], [87], it is used as a sensor of this pathway activity. Beside NHEJ, DNA DSBs can be also repaired by HR repair pathway. HR makes use of a large variety of proteins to accomplish the precise repair of DSBs. Among them, RAD51 can be visualised microscopically in form of foci and therefore can be used to analyse HR activity in cells under specific conditions.

The first step to investigate TRIM37's role in ionising radiation response, was to assess whether its depletion affected DNA DSBs repair kinetic. γ-H2AX and 53BP1 foci formation and decay were analysed in TRIM37-depleted HeLa cells after 4 Gy of ionising radiation treatment. The experiment consisted of a time course starting from 30 minutes, when the γ-H2AX and 53BP1 foci detection is expected to reach its highest peak, to 24 hours after IR, when it is expected to return at the baseline values [88], [89]. Foci were quantified in HeLa cells nuclei by using immunofluorescence assays as described in material and methods section. Images were batch processed by using a Fiji software macro to count

foci in each nucleus. For each sample, at least 10 microscopy image frames were acquired, and at least 300 nuclei scored. Foci counts were processed by using R software to analyse the foci number distribution per sample and the mean count. Data were plotted by using Prism software. The results, shown in Figure 5.1A and Figure 5.1B, revealed no statistically significant difference between control and TRIM37 depleted cells, suggesting that TRIM37 might not be involved in DNA DSBs repair by NHEJ pathway.

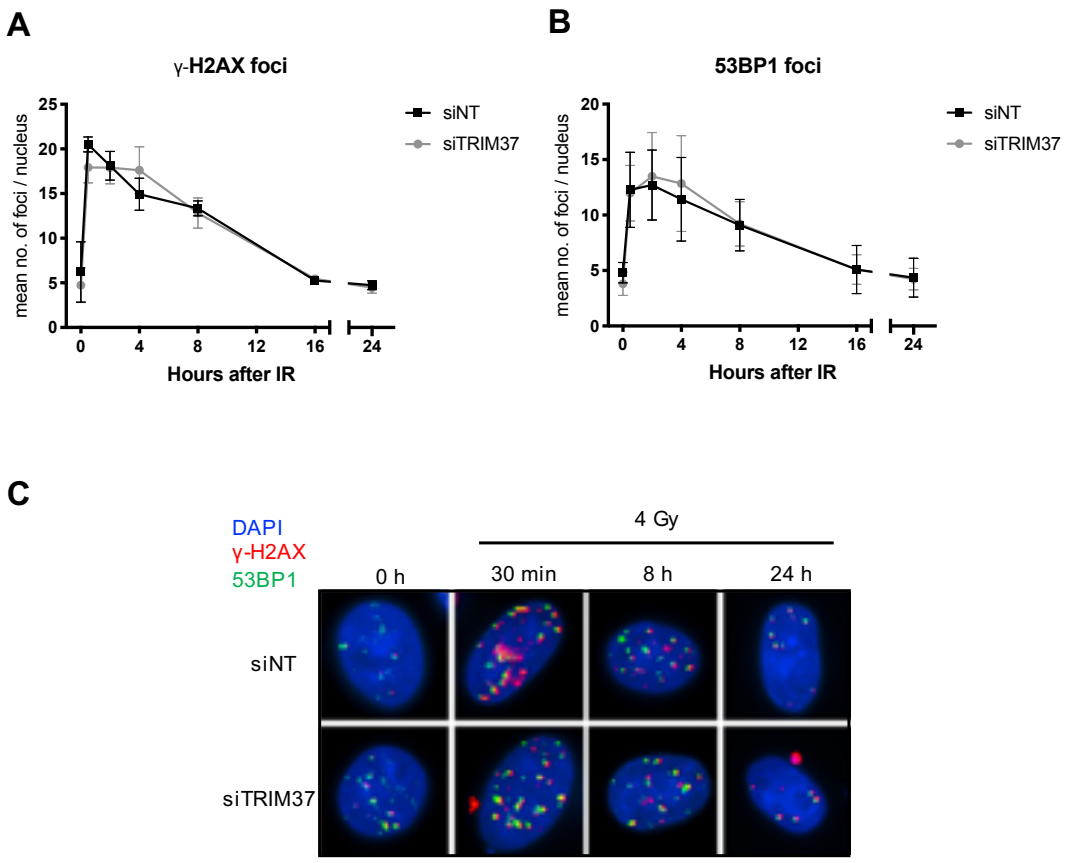


Figure 5.1. **γ-H2AX and 53BP1 foci scoring in TRIM37-depleted cells.** (A) The mean number of γ-H2AX foci per nucleus at 0 hours, 30 minutes, 2 hours, 4 hours, 8 hours, 16 hours and 24 hours following 4 Gy IR, 72 hours after transfection with NT or TRIM37 siRNAs. (B) The mean number of 53BP1 foci per nucleus at 0 hours, 30 minutes, 2 hours, 4 hours, 8 hours, 16 hours and 24 hours after 4 Gy IR. (C) Representative microscopy images of the assay above at 0 hours, 30 minutes, 8 hours and 24 hours. DNA, stained with DAPI, is shown in blue, γ-H2AX in red and 53BP1 in green. Values shown in graphs correspond to mean ± SD from three independent experiments.

Since TRIM37 did not seem to partake in NHEJ pathway, it was investigated whether TRIM37 played a role in HR pathway. At first, it was interrogated whether TRIM37 depletion affected the kinetic of RAD51 foci formation by scoring them through immunofluorescence assays. As for γ-H2AX and 53BP1 foci, a time course experiment was performed to score foci starting from 2 hours, when

RAD51 is expected to be recruited in the nucleus to promote HR, to 24 hours after IR, when foci are expected to be resolved. Foci were quantified in HeLa cells using the same approach described above. The results are shown in Figure 5.2A. TRIM37-depleted cells revealed a lower mean count of foci per nucleus at 2 and 8 hours after IR, however this difference was small and not statistically significant. Interestingly, When RAD51 foci formation was analysed by counting the proportion of positive cells to RAD51 foci, the small differences observed at 2 and 8 hours became more evident, showing a significant decrease of RAD51-foci-positive cells at 8 hours after IR treatment.

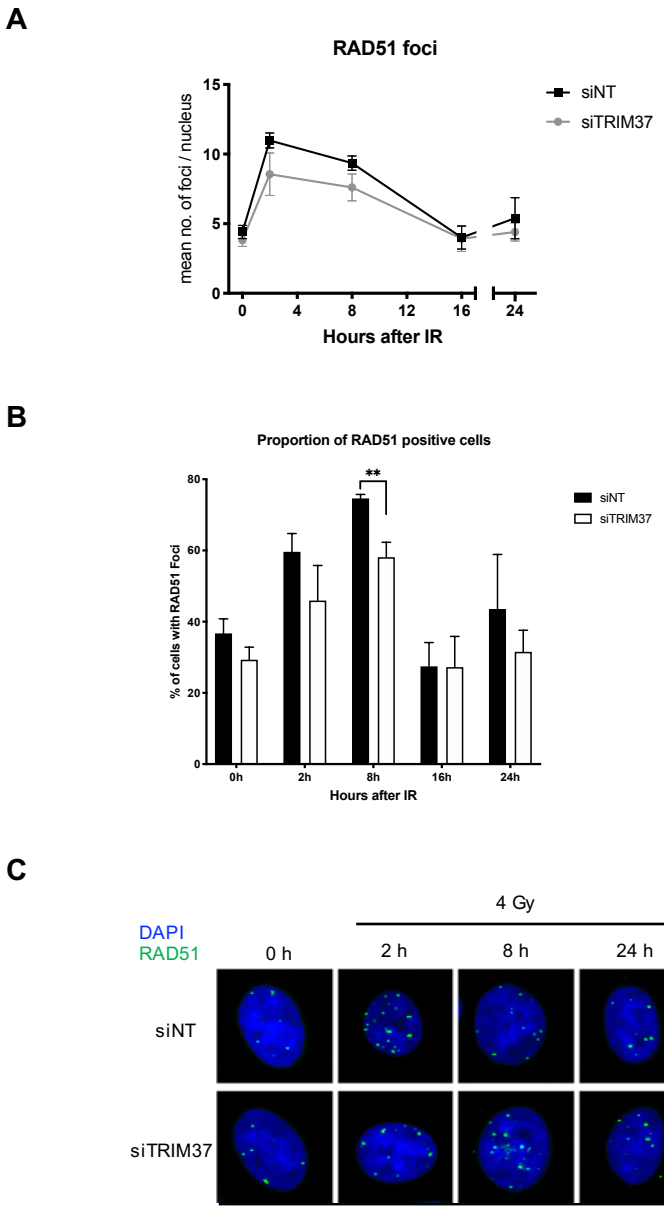


Figure 5.2. **RAD51 foci scoring in TRIM37-depleted cells.** (A) The mean number of RAD51 foci per nucleus at 0 hours, 2 hours, 8 hours, 16 hours and 24 hours following 4 Gy IR, 72 hours after transfection with NT or TRIM37 siRNAs. (B) Proportion of RAD51 positive cells (cells with more than 5 foci) in each sample. (C) Representative microscopy images of the assay above at 0 hours, 2 hours, 8 hours and 24 hours. DNA, stained with DAPI, is shown in blue and RAD51 in green. Values shown in graphs correspond to mean ± SD from three independent experiments. (**p < 0.001; two-sided Student's t-test).

As the results of RAD51 foci scoring were not conclusive to assess whether TRIM37 partakes in HR pathway, it was beneficial to use an additional assay to

investigate HR pathway efficiency. In this regard, it was used a I-SceI-based assay to examine DNA DSBs repair mediated by HR pathway [49]. The assay exploits genetically modified human cell lines (U2OS and HEK293 cells) that harbour a GFP-based reporter with recognition sites for the endonuclease I-SceI. The reporter contains an inactive expression cassette for green fluorescent protein (GFP) interrupted by a recognition sites for the restriction enzyme. The reporter is designed such that the repair of the DSB generated by the I-SceI enzyme leads to restoration of the GFP expression cassette. Accordingly, the efficiency of the repair was measured by knocking down TRIM37, transiently expressing I-SceI, culturing the cells to allow completion of repair, then determining the percentage of GFP-positive cells by flow cytometry analysis. The experiments were performed using RAD51 depletion as a positive control.

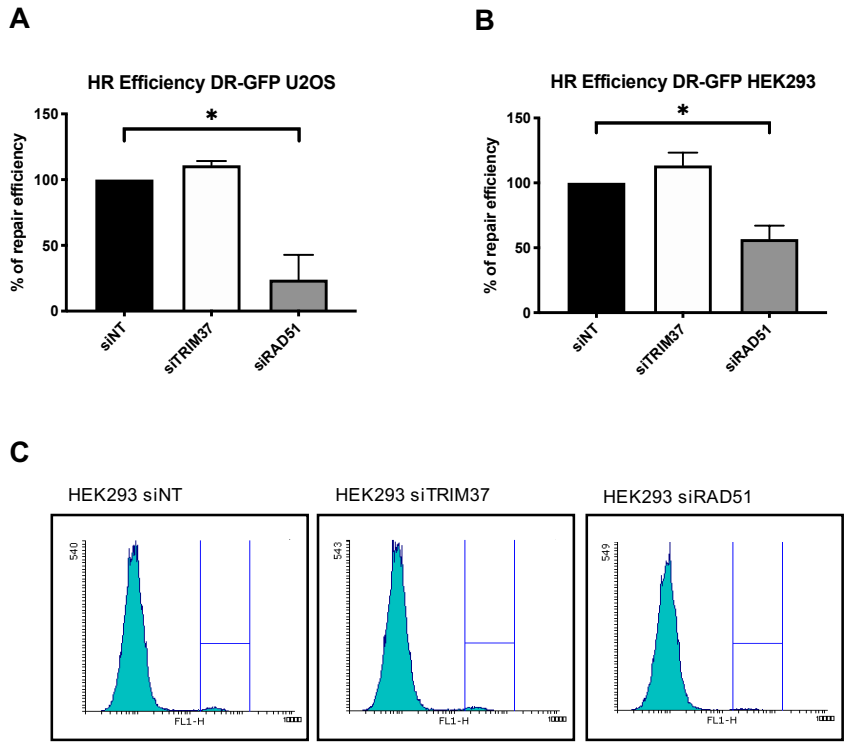


Figure 5.3. **HR pathway efficiency in TRIM37-depleted cells.** U2OS and HEK293 cells were first reverse transfected with non-targeting (siNT), siTRIM37 and siRAD51 siRNAs. Twenty-four hours later, cells were transiently transfected with a plasmid coding for the endonuclease I-SceI. In order to allow complete DNA DSBs repair, cells were left to grow for 72 hours. Then, GFP-positive cells were scored by using flow cytometry. The GFP-positive cells scored in the siNT sample were used to normalise the HR-mediated DSBs repair efficiency in the other samples. (A) A bar graph showing the measure of HR- mediated DNA DSBs repair efficiency in U2OS cells after depletion of TRIM37 and RAD51. (B) A bar graph showing the measure of HR- mediated DNA DSBs repair efficiency in U2OS cells after depletion of TRIM37 and RAD51. (C) Representative flowcytometry histograms of GFP-positive cells of the experiments showed in A and B. The gates, outlined in blue, were used to select the GFP-positive cells. Values shown in graphs correspond to mean ± SD from three independent experiments (*$p < 0.05$; two-sided Student's t-test).

The results, shown in Figure 5.3A and B, did not show any difference in HR efficiency upon TRIM37 depletion, suggesting that the impaired recruitment of RAD51 at the DNA damage site, shown in Figure 5.2A, is not due to a direct

involvement of TRIM37 in HR-mediated DNA DSBs repair. Besides, RAD51 depletion triggered a significant decrease of HR efficiency, confirming the reliability of this assay in measuring HR efficiency.

Provided the results in Figure 5.1, Figure 5.2, and Figure 5.3, TRIM37 does not seem to have a direct function in DNA DSBs repair. Therefore, the increased radiosensitivity of TRIM37 depleted cells (shown in Figure 3.8) was probably not due to an impairment of the DNA damage repair itself. However, given that RAD51 is mostly active in S and G2 cells, it was speculated that the small difference at 2 and 8 hours for RAD51 foci presented in Figure 5.2A might be caused by alterations in the cell cycle progression in TRIM37-depleted cells.

5.2.2 Analysis of cell cycle progression after TRIM37 depletion

The faithful repair of DSBs is possibly one of the most critical events for a cell in order to maintain its genomic integrity. Mis-repaired DNA damage can indeed have catastrophic consequences for the cells, such as chromosome breaks or rearrangements, mutations, cell death or cancer. DNA damage response (DDR) activates different pathways in order to delay the cell cycle progression and therefore allow completion of DNA damage repair. In this respect, cells can activate the G_1/S, intra S and G_2/M cell cycle checkpoints which are specific points in the cell cycle at which progression of the cell into the next phase can be blocked or slowed.

To investigate the role of TRIM37 in this process, the cell cycle profile was examined at different time points after radiation treatment in TRIM37-depleted HeLa cells.

In order to study cell cycle alterations, bromodeoxyuridine (BrdU) and propidium iodide (PI) cellular staining were analysed by flow cytometry assay. BrdU was added at a 10 μM concentration to growing cells 30 minutes before being trypsinised and fixed. BrdU is a thymidine analogue and gets incorporated into the newly synthesized DNA of replicating cells during S phase of the cell cycle. PI is an intercalating agent and a fluorescent molecule which binds to the DNA and makes it possible to quantify its cellular content. Since DNA content differs in each phase of the cell cycle, PI quantification makes possible to distinguish between cell cycle stages. These two molecules were used in combination in flow cytometry experiments to divide the cell population into four different groups depending on the cell cycle stage (G1, early S, late S, G2/M). G1 cells have low PI signal and low BrdU signal, early S phase cells have low PI signal and high BrdU signal, late S phase cells have high PI signal and high BrdU signal and G2/M cells have high PI signal and low BrdU signal.

The results, shown in Figure 5.4 and Table 5.1, showed that TRIM37 depletion did not affect cell cycle progression in asynchronous cells before IR treatment. The addition of IR treatment showed a progressive increase of cells in G2/M phases until 12 hours in both control and TRIM37-depleted cells, then, at 16 hours, the G2/M population amount returned to non-irradiated cells proportions, reflecting cells activating the G2/M checkpoints to allow DNA repair. As HeLa cells are p53 defective, cells in G1 phase do not increase in response to radiation treatment until 16 hours, reflecting, rather than G1/S checkpoint activation, the G2/M irradiated cells returning to proliferate and re-entering into the cell cycle. However, when comparing the irradiated TRIM37-depleted cells to the non-targeting control sample, a slight accumulation of TRIM37-depleted cells

in G1 was observed at 6 and 8 hours, with a simultaneous reduction of cells in both late S and G2/M phases. As this experiment was performed by using asynchronous cell populations, it was no possible to establish whether these differences were due to an increased cell cycle arrest at the G1/S checkpoint or to a compromised arrest of the cells at the G2/M one. To address this question, a different approach to analyse cell cycle profile was proposed.

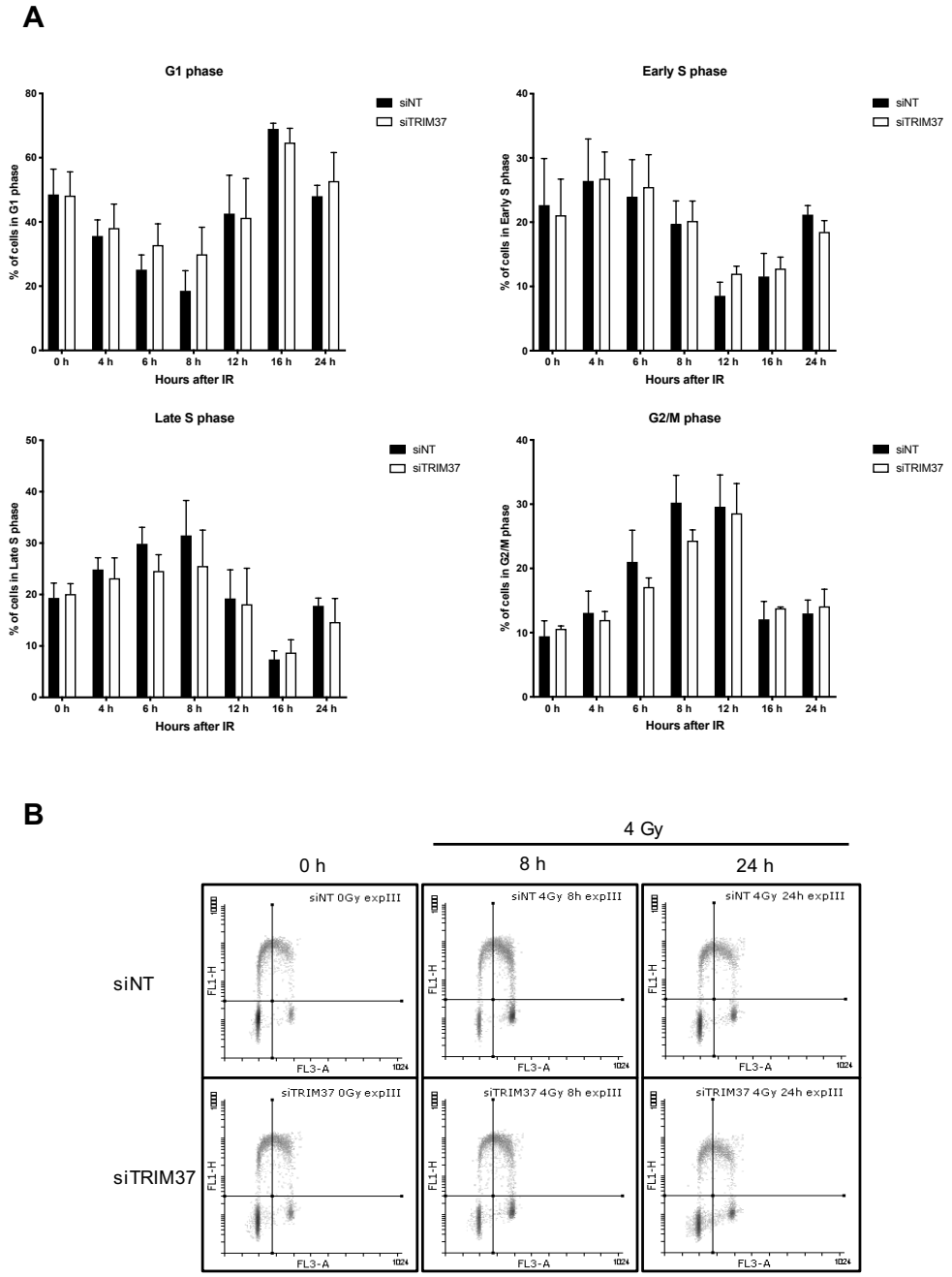


Figure 5.4. **Cell cycle progression of HeLa cells post TRIM37 siRNA depletion and IR.** (A) Proportion of cells in each cell cycle phase scored by using PI vs BrdU staining and monitored by flow cytometry. Cells were irradiated with 4 Gy IR at 72 hours post siRNA transfection and then fixed at different time points. BrdU was added 30 minutes before lifting each sample to allow incorporation into the DNA. After fixation, PI was added to intercalate with the DNA. (B) Representative Bromodeoxyuridine (BrdU) vs PI dot plots of the flow cytometry experiment above. G1: bottom-left quadrant. Early S: top-left quadrant. Late S: top- right quadrant. G2/M bottom-right quadrant.

		Hours after IR						
Cell Cycle phase		0	4	6	8	12	16	24
G1	siNT	48.6	35.6	25.2	18.6	42.6	69.0	48.0
	siTRIM37	48.2	38.1	32.8	29.9	41.3	64.7	52.7
	Δ	-0.4	2.5	7.7	11.4	-1.3	-4.3	4.7
Early S	siNT	22.7	26.4	24.0	19.7	8.5	11.6	21.2
	siTRIM37	21.1	26.8	25.5	20.2	12.0	12.8	18.5
	Δ	-1.6	0.3	1.5	0.5	3.4	1.2	-2.7
Late S	siNT	19.3	24.9	29.9	31.5	19.2	7.4	17.8
	siTRIM37	20.1	23.2	24.6	25.6	18.1	8.7	14.7
	Δ	0.8	-1.7	-5.3	-5.9	-1.1	1.3	-3.1
G2/M	siNT	9.4	13.1	21.0	30.2	29.6	12.1	13.0
	siTRIM37	10.6	12.0	17.1	24.3	28.6	13.8	14.1
	Δ	1.2	-1.1	-3.9	-5.9	-1.0	1.7	1.1

Table 5.1. **Quantification of the flow cytometry dot plots for asynchronous HeLa cells presented in Figure 5.4 above.** The difference between siTRIM37 and siNT is shown for each cell cycle phase and timepoint in the corresponding Δ rows. Values are coloured in gradient red when Δ<-3 and blue when Δ>3. Proportions were calculated with Flowing Software.

Cells have different responses and sensitivity to radiation according to the cell cycle phase they collocate at the moment of the treatment. For example, cells in G1 and early S phases can only access to NHEJ to repair DSBs, with a consequent possible formation of imprecise DNA repair, chromosomes rearrangements and genome mutagenesis. Besides, cells irradiated in G2 and M phases have less time to repair DNA before mitosis and can rely only on G2 checkpoints activation. This can cause a higher risk of undergoing through mitosis with unrepaired DNA and, consequently, mitotic catastrophe.

In order to characterise the fate of cells irradiated in G1, cell cycle changes were followed by PI staining over 24 hours after IR with the addition of Nocodazole, a chemical agent which arrests the cells at the spindle mitotic checkpoint by inhibiting microtubules polymerisation. Early and late G2 checkpoints activation was examined by staining for phospho-histone H3 (pH3), a molecular marker of mitosis.

The results are shown in Figure 5.5 and Figure 5.6. Cells positive to pH3 staining (shown in Figure 5.5A) augmented over time, reflecting the accumulation of cells in mitosis caused by the addition of Nocodazole. Non-irradiated TRIM37-depleted cells showed a greater proportion of cells in mitosis compared to the non-targeting control, which persisted over time after IR treatment. As the scope of the experiment was to investigate the effect of TRIM37 depletion on irradiated cells, the IR mitotic index was calculated by normalising the mitotic fraction of each sample after IR to the one at 0 Gy. The mitotic index, shown in Figure 5.5B, was decreased in TRIM37-depleted cells at later timepoints (12, 16 and 24 hours) but not at early ones (4, 6 and 8 hours), suggesting that cells irradiated in late S and G2/M phase were not affected by TRIM37 knockdown. However, siTRIM37-treated cells irradiated in G1 and early S phase seemed to accumulate later in mitosis, which might be explained by a prolonged arrest in either G1 or S phase.

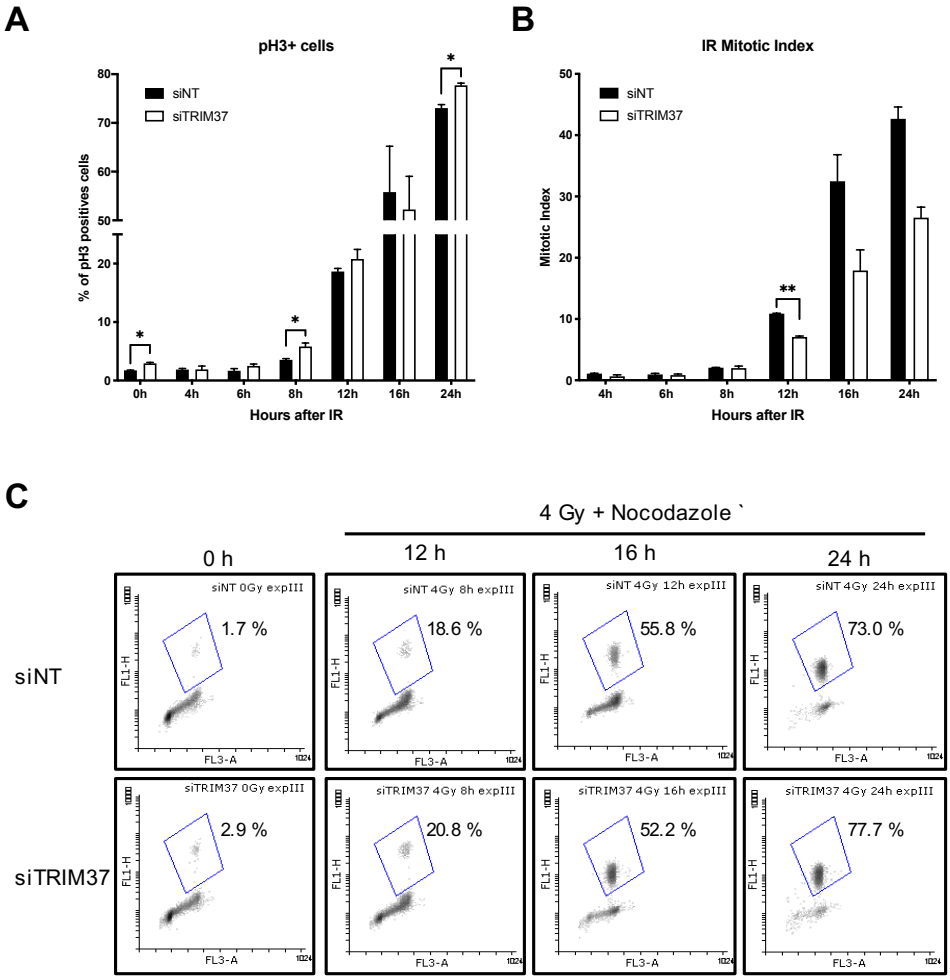


Figure 5.5. **Mitotic index of HeLa cells post TRIM37 siRNA depletion, IR and Nocodazole block.** (A) Proportion of phospho-histone H3 (pH3) positive cells scored by using PI vs pH3 staining and monitored by flow cytometry. Cells were irradiated with 4 Gy IR at 72 hours post siRNA transfection and immediately after incubated with Nocodazole, then, fixed at the time points indicated. After fixation, PI was added to intercalate with the DNA. (B) Quantification of the mitotic index calculated by normalising the mitotic fraction of each sample after IR to the one at 0 Gy. (C) Representative pH3 vs PI dot plots of the flow cytometry experiment above. Values shown in graphs correspond to mean ± SD from three independent experiments (*$p < 0.05$, **$p < 0.01$; two-sided Student's t-test).

In regard to the PI staining analysis (Figure 5.6), over time, it was observed a decrease of cells in G1 phase with a corresponding augmentation of cells in G2/M phase, due to the Nocodazole block. Besides, TRIM37-depleted cells

showed an increased proportion of cells in G1 compared to the non-targeting control at all the time points after radiation, suggesting a prolonged arrest in G1 phase. A consequent decrease of TRIM37 knockdown cells in G2 phase at later time points (12, 16 and 24 hours), was observed as well, suggesting that cells irradiated in G1 phase have more sustained G1/S checkpoint upon TRIM37 depletion and consequently a delay to progress to G2/M phases.

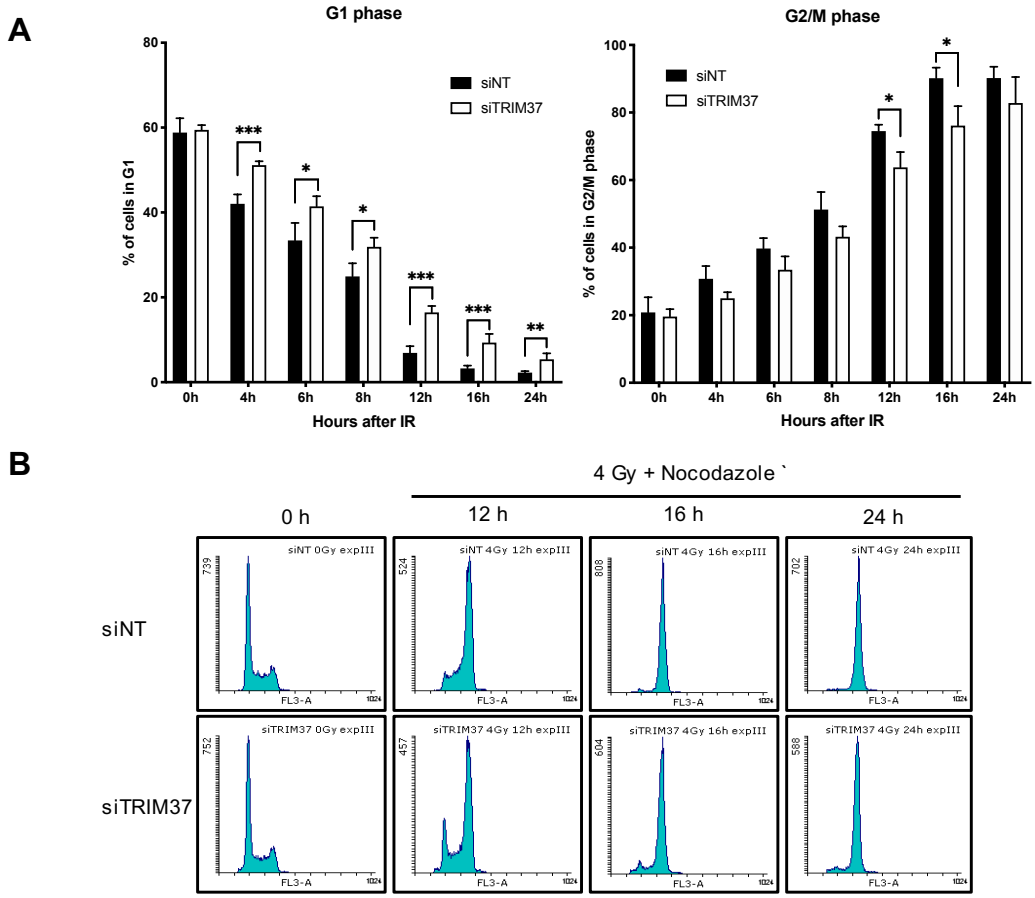


Figure 5.6**. Cell cycle progression of HeLa cells post TRIM37 siRNA depletion, IR and Nocodazole block.** (A) Proportion of cells in G1 and G2 phases scored by using PI staining and monitored by flow cytometry. (B) Representative PI histograms of the flow cytometry experiment above. Values shown in graphs correspond to mean ± SD from three independent experiments (*p < 0.05, **p < 0.01, ***p < 0.001; two-sided Student's t-test).

These results, taken together, seem to associate TRIM37 with G1/S phase transition. Interestingly, TRIM37 was already described to be involved in p53 signalling [90] and it has been described as a target of ATM kinase activity to activate the NF-kB pathway upon genotoxic stress [68].

5.2.3 Genomic instability analysis in TRIM37 depleted cells

IR treatment causes DNA DSBs formation in the nucleus, which cells attempt to repair in order to preserve their genome integrity. However, defects in DNA damage response can lead to mis-repaired or unrepaired DNA, which in turn cause chromosomal rearrangements and genome mutagenesis. When cells attempt to divide with unrepaired DNA, changes in nuclear morphology, failure to complete cytokinesis, and cell death may occur. Similarly, genome mutagenesis can lead to cancer transformation when tumour suppressor or oncogenes are mutated. Patients carrying TRIM37 mutation and TRIM37 KO mice have been described in literature to have an higher incidence of tumours, suggesting that TRIM37 may be involved in genome integrity surveillance [58], [59], [61], [64].

Therefore, the effects of TRIM37 depletion were investigated in terms of nuclear morphology. Alterations of nuclear morphology arise when incorrect chromosomes separation occur and can be distinguished in micronucleation and multinucleation. Micronucleation is characterised by a cell containing a nucleus and a smaller portion of a second one in the cytoplasm. Multinucleation is categorised by a single cytoplasm containing two or more nuclei. Examples of nuclear morphology alterations are shown in Figure 5.7C. Nuclear morphology was examined in TRIM37-depleted cells with and without radiation treatment by immunofluorescence assays. The immunostaining was performed using DAPI, a

fluorescent DNA-binding agent used to stain the nucleus, and FITC, a protein binding agent used to stain the cytoplasm. These experiments and their analysis were performed under my supervision by Katya Marks, a former FHS student who joined our lab to perform a research project. The results of the assay are shown in Figure 5.7.

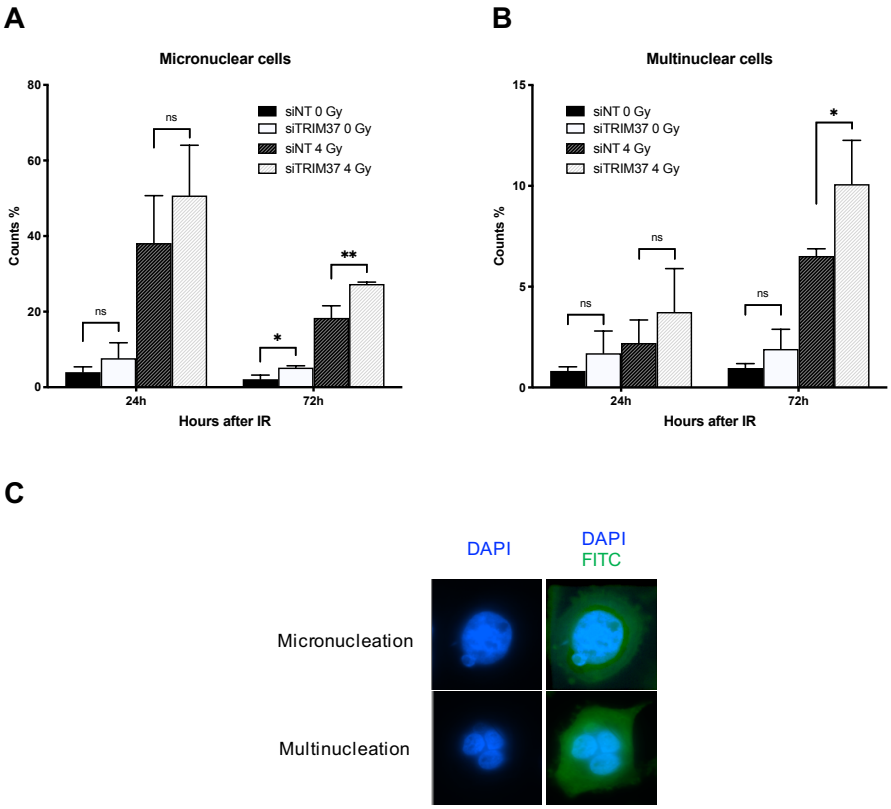


Figure 5.7. **Cellular morphology analysis in TRIM37 depleted cells.** (A) Quantification of the percentage counts of micronuclear cells in TRIM37 depleted HeLa cells (showed at 0 Gy in white and at 4 Gy in grey-banded white) and control cells (showed at 0 Gy in black and at 4 Gy in grey-banded black). At least 300 cells were counted per condition. (B) Quantification of the percentage counts of multinuclear cells in TRIM37 depleted HeLa cells (showed at 0 Gy in white and at 4 Gy in grey-banded white) and control cells (showed at 0 Gy in black and at 4 Gy in grey-banded black). At least 300 cells were counted per condition. (C) Representative microscopy images from the immunofluorescence analysis of HeLa cells 72 hours after 4 Gy irradiation. DAPI-stained nuclei are shown in blue. FITC-stained cytoplasm is shown in green. Values shown in graphs correspond to mean ± SD from three independent experiments (*p < 0.05, **p < 0.01; two-sided Student's t-test).

Micronuclei formation was found to be higher in TRIM37-depleted cells in both irradiated and unirradiated samples (Figure 5.7A). At 72 hours after IR the depletion of TRIM37 accounted a statistically significant increase of micronuclei formation in both irradiated and unirradiated samples. As in 24-hours samples

there was considerable variation across replicates, the micronuclei formation changes were not found to be statistically significant. However, there was a consistent trend across three repeats of increased micronuclei formation after both 0 Gy and 4 Gy IR. Multinuclear cells were also scored to investigate TRIM37's role in genome's integrity maintenance. The proportion of multinuclear cells was found to be higher in TRIM37-depleted HeLa cells at both 24 and 72 hours in irradiated and unirradiated samples (Figure 5.7B). The most striking effect of TRIM37 depletion was observed at 72 hours after 4 Gy of IR.

Overall, there was a greater increase of both micronucleation and multinucleation when TRIM37 depletion was combined with IR, indicating relevance of these results with respect to TRIM37's role in modifying cellular radiosensitivity. Also, these results suggested that TRIM37 may play a role in promoting accurate cellular division, independently from the presence of genotoxic stress.

Taken together, the results presented in this chapter suggest that upon IR treatment TRIM37 may be involved in regulating the cell cycle progression. It is possible that the altered progression over the cell cycle causes an unbalanced regulation of the DNA repair pathway choice in favour of error-prone ones, which in turn increases the chances of generating mis-repaired DNA and chromosomal rearrangements. After a few cell divisions, the accumulation of these alterations causes the occurrence of micronuclear and multinuclear cells, which represents a marker of mitotic catastrophe.

5.3 Discussion

Following its identification as a novel modulator of radiosensitivity, TRIM37 was selected to be further investigated in order to examine its role in ionising radiation response. In this chapter, several approaches have been undertaken to identify the mechanism by which the lack of TRIM37 induces an increase of radiosensitivity.

TRIM37 is an E3 ubiquitin ligase of the TRIpartite Motif protein family (TRIM). Mutations in TRIM37 gene are associated to a rare autosomal recessive disorder known as MULIBREY nanism (muscle-liver-brain-eye) [72]. These patients, as well as a TRIM37$^{-/-}$ mouse model of disease, have been previously described to have an higher incidence of tumours, suggesting that TRIM37 is important to maintain genetic stability [58], [59], [61], [64]. In light of these observations, together with the evidences provided by this study that TRIM37 depletion increases sensitivity to IR, it was hypothesised that TRIM37 could partake in DNA repair. Interestingly, TRIM37 depletion did not cause an overall change of DNA DSBs repair kinetics when γ-H2AX and 53BP1 foci were examined. Nevertheless, a small decrease in RAD51 foci detection was observed at early time points after IR. This decrease was not statistically significant because of the variability among replicates; however, RAD51 foci were consistently found decreased across replicates. Also, when RAD51-foci-positive cells were scored, the difference detected at 2 and 8 hours became more marked, with a decrease of positive cells at 8 hours statistically significant ($p < 0.001$). Considering the altered RAD51 foci formation upon radiation treatment in TRIM37-depleted cells, together with the importance of RAD51 in regulating HR [91], it was speculated

that an HR defect could occur because of TRIM37 absence. Conversely, when the efficiency of HR-mediated DNA DSBs repair was investigated, no difference between TRIM37 knockdown and control cells was observed. These results suggested that TRIM37 did not directly participate to the DNA damage repair, however, a more comprehensive panel of DNA damage repair markers, such as RPA, BRCA1, or BRCA2, need to be examined to fully exclude TRIM37 direct involvement in DNA repair. Also, NHEJ and Alt-NHEJ efficiency could be tested by using GFP-reporter cell lines to address whether any of these pathways are altered in TRIM37 depleted cells [49]. Yet, the reason why TRIM37 appeared to be important for cellular survival upon IR could be investigated by exploring other aspects of DDR functioning.

DNA DSBs repair can occur by using different pathways according to the cellular context at the moment of the damage. The choice of the adequate pathway to be used to repair DNA depends on many factors, including cell cycle phase and the local chromatin environment [92]. In a well-regulated cell, the decision process is controlled to preserve the DNA sequence integrity, minimising error-prone outcomes. However, alterations in genes regulating DSBs repair, combined with overwhelming DNA damage conditions, can perturb this balance and reduce the accuracy of the DNA lesion repair. For example, catastrophic chromosomes rearrangements can be induced by an unbalanced activation of the alternative NHEJ pathway over the canonical one [16].

TRIM37 has been shown to play a role in cell cycle regulation. In particular, TRIM37 appears to be involved in centrioles duplication, an essential stage to form functional centrosomes [93]. Functional centrosomes guarantee a correct

DNA segregation during mitosis, and their activity is required for cell cycle progression [94], [95]. Taking into consideration the role already described of TRIM37 in cell cycle regulation, together with the influence that cell cycle phase context has on the DDR regulation, it was investigated whether the observed radiosensitivity of TRIM37 depleted cells was due to an altered cell cycle progression. Remarkably, irradiated TRIM37-depleted cells showed an increased arrest in G1 phase. The G1 arrest became more accentuated when mitotic arrest was chemically induced by the addition of Nocodazole, suggesting that cells irradiated in G1 had a prolonged arrest at the G1/S checkpoint. These results, considering the role of IR in centrosomes overduplication [96]–[98], and the requirement of functional centrosomes to progress through the cell cycle at the G1/S checkpoint [94], suggest that TRIM37 may be important in regulating the cell cycle progression upon radiation by ensuring a correct centrosome assembly. Furthermore, the fact that both patients affected by MULIBREY nanism and TRIM37 KO mice have a higher incidence of cancer [58], [59], [61], [64], suggests that the lack of TRIM37 function may affect the cellular genomic stability.

The exact mechanism by which TRIM37 modify cellular radiosensitivity is yet to be investigated. A supplementary experiment to clarify at which stage of the cell cycle TRIM37 is functioning could be to analyse its expression over cell cycle phases. Additionally, RAD51 foci scoring together with cell cycle phase identification markers, could verify whether the reduction of RAD51-positive cells in TRIM37-depleted samples is due to cell cycle modifications. Also, identifying molecular interactors of TRIM37 could be crucial to find ubiquitination targets involved in cell cycle regulation pathways. Post-translational modification, such

as ubiquitination, are known to be crucial to regulate several cellular function, including cell cycle progression [99]. For example, Cyclin F has been described to regulate the cell cycle progression by ubiquitinating CP110, an essential protein for centrosome duplication [100], [101].

TRIM37 has been described in literature as a caretaker of genomic stability by regulating centriole replication and consequentially ensuring correct DNA segregation during mitosis. In particular, increased lagging chromosome were observed in mitotic cells when TRIM37 was depleted [93], suggesting that the radioprotection effect provided by TRIM37 could be also due to its role in preventing aberrant mitosis and therefore mitotic catastrophe (MC). Besides, it is known that radiation induces centrosomes overduplication, which in turn causes the assembly of multipolar spindles and mitotic failure [96]–[98]. Interestingly, in this study, an increased occurrence of markers of MC, such as micronuclear and multinuclear cells, was observed upon TRIM37 depletion. This increase was more prominent when cells were treated with IR, suggesting that an increase in MC could explain the augmented radiosensitivity of TRIM37 depleted cells. To support the role of TRIM37 in genomic integrity surveillance, additional experiments such as quantification of chromosomes aberrations and chromatid exchanges in cells lacking TRIM37 expression could be performed. TRIM37 depletion has been linked to an increased apoptosis upon genotoxic stress through altered activation of the NF-kB pathway [68]. Therefore, an assay to examine apoptosis activation, such Annexin V/PI staining, could be performed to assess whether the increased radiosensitivity observed in cells lacking TRIM37 expression is due to an improved apoptosis activation.

Chapter 6 GENERAL DISCUSSION

6.1 Summary of findings

This study had the scope of identifying novel regulators of radiosensitivity by using different functional genomics approaches.

Starting from a RNAi-based screening, performed in our laboratory by Dr. Gaganpreet Tiwana, Dr. Gonzalo Rodriguez-Berriguete and Dr. Remko Prevo, I selected among the top-100 identified hits MCM3, MCM4, TAF4, TEAD1, TPR, TRIM37 and WDR5 to assess their radiosensitisation effect in both HeLa and MRC-5 cells. All the analysed hits showed a significant radiosensitisation when depleted in HeLa cells, confirming the original siRNA screen findings and quality. Therefore, the method developed in our laboratory, combining 96-well plate colony formation assay with high-throughput radiosensitivity RNAi-based screens resulted to be a successful approach for the identification of novel genes involved in radiation response. Among the investigated targets, two novel molecular determinants of radiosensitivity, TPR and TRIM37 appeared to be the most interesting targets, therefore, validation work was carried on. Both genes were found overexpressed in cancer, their depletion sensitised a panel of diverse cancer cell lines to IR treatment, and siRNA off-target effect was excluded by comparing the radiosensitisation effect given by different siRNA strands targeting the same gene transcript.

The second approach to identify novel regulators of radiation response presented in this thesis is a pooled LOF CRISPR genome-wide screen. The screening combined systematic perturbation of each gene of the genome with radiation treatment in MRC-5 cells. Interestingly, among the top radiosensitivity hits identified by the screen, TRIM37, a gene already discovered by the

siRNA-based HTS previously presented, was associated to radiation response modulation. The second top-ranked gRNA identified by the screen as a modulator of radiosensitivity (among over 120,000 gRNAs) was TRIM37_HGLibB_51414 with a gRNA C_{score} = -6.95. According to the gene C_{score} ranking, TRIM37 was found ranked 24th (among over 20,000 genes) with a gene C_{score} = -2.22. Finally, rank product analysis ranked TRIM37 as the 5th gene whose depletion caused radiosensitisation with a p-value = 0.00024. Therefore, the different approaches used to analyse the CRISPR screen consistently found TRIM37 among the top targets of radiosensitivity. Also ATR, a gene whose deficiency is already known to increase cellular radiosensitivity [102], [103], was identified by this screen to significantly sensitise MRC-5 to radiation when depleted (p-value = 0.025). Considering the previous identification of TRIM37 as a novel determinant of radiosensitivity by an independent screening, its validation in a panel of cancer cell lines, and the identification of an already known radiosensitiser, such as ATR, this study showed that LOF CRISPR radiosensitivity screens are suitable approaches in order to identify novel genes whose function is associated to radiation response.

Finally, because TRIM37 was not described in literature as a modulator of radiosensitivity before, its molecular function in radiation response context was investigated. Different endpoints were analysed in order to assess TRIM37 function in radiation response, yet, it was no possible to establish the exact mechanism by which TRIM37 modulates the response to radiation. A direct role of TRIM37 in DNA repair activation cascade was excluded. When investigating cell cycle changes, it was found that TRIM37 depletion modified the cellular progression of irradiated cells through cell cycle checkpoints. In particular,

TRIM37-depleted cells appeared to accumulate in G1 phase of the cell cycle upon radiation treatment. Lastly, the genomic stability of cells lacking TRIM37 expression was investigated. The absence of TRIM37 expression caused an increased formation of nuclear alterations in both unirradiated and irradiated cells. Nuclear alterations, such as micronuclei formation and multinuclear cells, represent cellular markers of genomic instability and are considered the precursors of mitotic catastrophe.

Taken together, these results identify TRIM37 as a novel molecular determinant of radiosensitivity and associate its function to genome stability surveillance.

6.2 Significance of findings

The aim of this study was the identification of novel genes involved in radiation response in order to discover new clinical targets of radiosensitivity. For this scope, an innovative method, such as genome-wide CRISPR mutagenesis technology was adopted. Since its discovery in 2013 [45], [44], the use of CRISPR technology as a tool for genome mutagenesis has increased exponentially. Because of its versatility and the possibility of targeting specific sequences within the genome, CRISPR genome editing quickly became a valid tool for high-throughput functional genomics approaches [48], [104]–[106].

To the best of my knowledge, the screen described in this study is the first genome-wide screen designed to discover novel clinical targets of radiosensitivity and the first to use a LOF CRISPR pooled screening approach in combination with radiation. This method represents an extremely powerful approach to investigate gene function at specific condition in a genome-scale

setup. For example, the presented results only showed the negative selection screen data, meaning that only the subset of genes whose depletion caused decreased survival were considered for further analysis. However, because the experimental readout was the measure of clonal survival, positive selection screen data were also available (data not shown), meaning that in a single experimental setup it is possible to concurrently identify both genes responsible of radiosensitivity and genes responsible of radioprotection. Moreover, a secondary screen to identify additional targets of radiosensitivity could be performed, as only one of the identified hits was investigated and validated.

Although the CRISPR radiosensitivity screen allowed the identification of an extensively validated target of radiosensitivity among the first hits (TRIM37) and identified an already known determinant of radiosensitivity (ATR), there are still some elements of the screen design that could have been improved. For example, other genome-wide gRNA libraries (such as the Brunello and Sabatini libraries) have been developed since the human GeCKOv2 one, and have been showed to be more effective at distinguishing essential and non-essential genes when compared to the GeCKOv2 library [104], [107], [108]. Another limitation of this screen, as previously stated in chapter 4, was the lack of additional biological replicates, as they are essential to reduce the noisiness of the data, the risk of identifying false-positive hits, and increase the statistical power of the analysis. Finally, recent findings, which are dated later than the performance of this screen, identify that in p53-proficient cells, Cas9 generated DSBs activate the DDR and therefore trigger a transient, p53-dependent cell cycle arrest [109]. This p53-dependent cell cycle arrest in G1 phase can cover guide-specific effects by decreasing the efficiency of the genome editing. Therefore, in dropout screens

the p53-proficiency can cause the occurrence of false-negative hits. This can explain why some of expected hits, such as DNA-PK, ATM and LIG4, did not appear among the significant identified targets. In light of these findings, the cell line choice represents a crucial step for the performance of screenings that are aimed at identifying genes whose loss leads to cell death or decreased cell proliferation. Also, for radiosensitivity screens it might be interesting investigating possible interactions between the DDR response activated by both Cas9 nuclease activity and the ionising radiation induced DNA damage.

CRISPR/Cas9 system can be used not only for achieving gene KO and knock-in, but also as a modulator of gene expression at targeted loci [110], [111]. The Cas9-based transcriptional inhibition and activation, referred as CRISPRi and CRISPRa respectively, represent valuable tools for more recently developed functional genomics screens in which they are combined with single-cell transcriptomics approaches such as RNAseq [112]. CRISPRi allows indeed direct verification of its efficiency by transcriptomic analysis as it reduces the transcript of the targeted genes instead of perturbing the gene sequence itself. CRISPRi is also beneficial when genome-scale permutations are used, because it is not affected by the genetic variability associated to the indel mutations formation characteristic of CRISPR KO. However, the same group showed that the combination of CRISPR KO and single-cell RNA profiling (also called Perturb-Seq) represents an equally valuable approach for pooled genetic screens [113].

In conclusion, the Perturb-Seq, a novel methods developed by Aviv Regev group for the performance of pooled genome-wide genetic screens [112], [113],

represents an interesting subsequent step to implement the findings provided by the LOF CRISPR KO screen presented in this study.

6.3 Concluding remarks

The two methods described in this study have been shown to be valuable approaches for the identification of novel determinants of radiosensitivity. Moreover, both screenings identified TRIM37 as a novel modulator of radiosensitivity, making it an interesting target to be further investigated and eventually translated into clinical approaches.

Whether TRIM37 can be exploited as a target of radiosensitivity or as a biomarker to predict tumour response to radiation therapy, is yet to be defined. However, its identification opens to several investigation paths and will be considered for further investigation in the future.

Moreover, this thesis added further evidences that can provide a mechanistic explanation of some of the clinical features of the MULIBREY syndrome, such as growth failure, infertility and higher incidence of cancer.

References

[1] World Health Organisation, “WHO | Cancer.” [Online]. Available: https://www.who.int/health-topics/cancer. [Accessed: 29-Sep-2019].

[2] Cancer Research UK, “CRUK | Cancer Risk Statistics 2018.” [Online]. Available: https://www.cancerresearchuk.org/health-professional/cancer-statistics/risk. [Accessed: 29-Sep-2019].

[3] Cancer Research UK, “Worldwide cancer statistics.” [Online]. Available: https://www.cancerresearchuk.org/health-professional/cancer-statistics/worldwide-cancer#heading-One. [Accessed: 29-Sep-2019].

[4] Cancer Research UK, “CRUK | Radiotherapy,” 2018. [Online]. Available: https://www.cancerresearchuk.org/about-cancer/cancer-in-general/treatment/radiotherapy/about. [Accessed: 29-Sep-2019].

[5] G. Delaney, S. Jacob, C. Featherstone, and M. Barton, “The role of radiotherapy in cancer treatment: Estimating optimal utilization from a review of evidence-based clinical guidelines,” *Cancer*. 2005.

[6] M. Joiner and A. van der Kogel, *Basic Clinical Radiobiology, Fifth edition*. 2019.

[7] A. C. Begg, F. A. Stewart, and C. Vens, “Strategies to improve radiotherapy with targeted drugs,” *Nat. Publ. Gr.*, vol. 11, 2011.

[8] B. Halliwell and O. I. Aruoma, “DNA damage by oxygen-derived species,” *Febs*, 1991.

[9] A. Ciccia and S. J. Elledge, “The DNA Damage Response: Making It Safe to Play with Knives,” *Molecular Cell*. 2010.

[10] J. A. Reisz, N. Bansal, J. Qian, W. Zhao, and C. M. Furdui, “Effects of ionizing radiation on biological molecules--mechanisms of damage and emerging methods of detection.,” *Antioxid. Redox Signal.*, vol. 21, no. 2, pp. 260–92, Jul. 2014.

[11] E. J. Hall and A. J. Giaccia, *Radiobiology for the radiologist: Seventh edition*. 2012.

[12] A. N. Blackford and S. P. Jackson, “ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response,” *Mol. Cell*, vol. 66, no. 6, pp. 801–817, Jun. 2017.

[13] O. A. Martin *et al.*, “Statistical analysis of kinetics, distribution and co-localisation of DNA repair foci in irradiated cells: Cell cycle effect and implications for prediction of radiosensitivity,” *DNA Repair (Amst).*, 2013.

[14] K. M. Kramer, J. A. Brock, K. Bloom, J. K. Moore, and J. E. Haber, “Two different types of double-strand breaks in Saccharomyces cerevisiae are repaired by similar RAD52-independent, nonhomologous recombination events.,” *Mol. Cell. Biol.*, 1994.

[15] R. D. Johnson and M. Jasin, “Double-strand-break-induced homologous recombination in mammalian cells,” in *Biochemical Society Transactions*, 2001.

[16] D. Simsek and M. Jasin, “Alternative end-joining is suppressed by the canonical NHEJ component Xrcc4-ligase IV during chromosomal translocation formation,” *Nat. Struct. Mol. Biol.*, 2010.

[17] A. Kakarougkas and P. A. Jeggo, “DNA DSB repair pathway choice: An

orchestrated handover mechanism," *British Journal of Radiology*. 2014.

[18] T. M. Gottlieb and S. P. Jackson, "The DNA-dependent protein kinase: Requirement for DNA ends and association with Ku antigen," *Cell*, 1993.

[19] P. A. Jeggo and M. Löbrich, "Artemis links ATM to double strand break rejoining," *Cell Cycle*. 2005.

[20] E. Riballo *et al.*, "Cellular and Biochemical Impact of a Mutation in DNA Ligase IV Conferring Clinical Radiosensitivity," *J. Biol. Chem.*, 2001.

[21] J. Liu, T. Doty, B. Gibson, and W. D. Heyer, "Human BRCA2 protein promotes RAD51 filament formation on RPA-covered single-stranded DNA," *Nat. Struct. Mol. Biol.*, 2010.

[22] B. Alberts *et al.*, *Molecular Biology of the Cell - Sixth Edition*. 2015.

[23] R. Ceccaldi, B. Rondinelli, A. D. D'Andrea, and A. D. D'Andrea, *Repair Pathway Choices and Consequences at the Double-Strand Break*, vol. 26, no. 1. 2016, pp. 52–64.

[24] J. H. Wilson and T. Hunt, "Molecular biology of the cell, sixth edition : the problems book." Garland Science, New York ; Abingdon, 2015.

[25] M. B. Kastan *et al.*, "A mammalian cell cycle checkpoint pathway utilizing p53 and GADD45 is defective in ataxia-telangiectasia," *Cell*, 1992.

[26] C. E. Canman *et al.*, "Activation of the ATM kinase by ionizing radiation and phosphorylation of p53," *Science (80-.).*, 1998.

[27] R. Khosravi, R. Maya, T. Gottlieb, M. Oren, Y. Shiloh, and D. Shkedy, "Rapid ATM-dependent phosphorylation of MDM2 precedes p53 accumulation in response to DNA damage," *Proc. Natl. Acad. Sci. U. S. A.*,

1999.

[28] T. Riley, E. Sontag, P. Chen, and A. Levine, “Transcriptional control of human p53-regulated genes,” *Nat. Rev. Mol. Cell Biol.*, 2008.

[29] J. Bartek, C. Lukas, and J. Lukas, “Checking on DNA damage in S phase,” *Nature Reviews Molecular Cell Biology*. 2004.

[30] L. Busino *et al.*, “Degradation of Cdc25A by β-TrCP during S phase and in response to DNA damage,” *Nature*, 2003.

[31] B. Xu, S.-T. Kim, D.-S. Lim, and M. B. Kastan, “Two Molecularly Distinct G2/M Checkpoints Are Induced by Ionizing Irradiation,” *Mol. Cell. Biol.*, vol. 22, no. 4, pp. 1049 LP – 1059, Feb. 2002.

[32] D. Hanahan and R. A. Weinberg, “Hallmarks of cancer: The next generation,” *Cell*. 2011.

[33] C. M. L. West, S. E. Davidson, S. A. Roberts, and R. D. Hunter, “The independence of intrinsic radiosensitivity as a prognostic factor for patient response to radiotherapy of carcinoma of the cervix,” *Br. J. Cancer*, 1997.

[34] A. Fourquet, D. Stoppa-Lyonnet, Y. M. Kirova, B. Sigal-Zafrani, and B. Asselain, “Familial breast cancer: Clinical response to induction chemotherapy or radiotherapy related to brca1/2 mutations status,” *Am. J. Clin. Oncol. Cancer Clin. Trials*, 2009.

[35] N. Franken *et al.*, “Cell survival and radiosensitisation: modulation of the linear and quadratic parameters of the LQ model,” *Int. J. Oncol.*, 2013.

[36] A. Fire, S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver, and C. C. Mello, “Potent and specific genetic interference by double-stranded RNA

in caenorhabditis elegans," *Nature*, 1998.

[37] M. T. McManus and P. A. Sharp, "Gene silencing in mammals by small interfering RNAs," *Nature Reviews Genetics*. 2002.

[38] G. S. Higgins *et al.*, "A Small Interfering RNA Screen of Genes Involved in DNA Repair Identifies Tumor-Specific Radiosensitization by POLQ Knockdown," *Cancer Res.*, vol. 70, no. 7, pp. 2984–2993, Apr. 2010.

[39] G. Rodriguez-Berriguete *et al.*, "Nucleoporin 54 contributes to homologous recombination repair and post-replicative DNA integrity," *Nucleic Acids Res.*, vol. 46, no. 15, 2018.

[40] G. S. Tiwana *et al.*, "Identification of vitamin B1 metabolism as a tumor-specific radiosensitizing pathway using a high-throughput colony formation screen," *Oncotarget*, vol. 6, no. 8, pp. 5978–5989, Mar. 2015.

[41] A. L. Hopkins and C. R. Groom, "The druggable genome," *Nat. Rev. Drug Discov.*, 2002.

[42] G. Gasiunas, R. Barrangou, P. Horvath, and V. Siksnys, "Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria," *Proc. Natl. Acad. Sci. U. S. A.*, 2012.

[43] M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, and E. Charpentier, "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity," *Science (80-.).*, 2012.

[44] L. Cong *et al.*, "Multiplex genome engineering using CRISPR/Cas systems," *Science (80-.).*, 2013.

[45] P. Mali *et al.*, "RNA-guided human genome engineering via Cas9," *Science*

(80-.)., 2013.

[46] O. Shalem, N. E. Sanjana, and F. Zhang, “High-throughput functional genomics using CRISPR–Cas9,” *Nat. Rev. Genet.*, vol. 16, no. 5, pp. 299–311, May 2015.

[47] Addgene, “Addgene - CRISPR guide.” [Online]. Available: https://www.addgene.org/crispr/guide/.

[48] O. Shalem *et al.*, “Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells,” *Science (80-.)*., vol. 343, no. 6166, pp. 84–87, Jan. 2014.

[49] A. Gunn and J. M. Stark, “I-SceI-based assays to examine distinct repair outcomes of mammalian chromosomal double strand breaks,” *Methods Mol. Biol.*, 2012.

[50] S. E. Morgan, C. Lovly, T. K. Pandita, Y. Shiloh, and M. B. Kastan, “Fragments of ATM which have dominant-negative or complementing activity.,” *Mol. Cell. Biol.*, 1997.

[51] D. R. Rhodes *et al.*, “ONCOMINE: A Cancer Microarray Database and Integrated Data-Mining Platform,” *Neoplasia*. 2004.

[52] P. Bangs, B. Burke, C. Powers, R. Craig, A. Purohit, and S. Doxsey, “Functional Analysis of Tpr: Identification of Nuclear Pore Complex Association and Nuclear Localization Domains and a Role in mRNA Export,” *J. Cell Biol.*, vol. 143, no. 7, pp. 1801–1812, 1998.

[53] N. Schweizer, C. Ferrás, D. M. Kern, E. Logarinho, I. M. Cheeseman, and H. Maiato, “Spindle assembly checkpoint robustness requires Tpr-mediated regulation of Mad1/Mad2 proteostasis.,” *J. Cell Biol.*, vol. 203,

no. 6, pp. 883–93, Dec. 2013.

[54] K. Rajanala *et al.*, “Phosphorylation of nucleoporin Tpr governs its differential localization and is required for its mitotic function,” *J. Cell Sci.*

[55] B. D. Yard *et al.*, “A genetic basis for the variation in the vulnerability of cancer to DNA damage,” *Nat. Commun.*, vol. 7, p. 11428, Apr. 2016.

[56] J. Kallijärvi, K. Avela, M. Lipsanen-Nyman, I. Ulmanen, and A.-E. Lehesjoki, “The TRIM37 Gene Encodes a Peroxisomal RING-B-Box-Coiled-Coil Protein: Classification of Mulibrey Nanism as a New Peroxisomal Disorder,” *Am. J. Hum. Genet.*, 2002.

[57] S. Hatakeyama, “TRIM Family Proteins: Roles in Autophagy, Immunity, and Carcinogenesis,” *Trends in Biochemical Sciences*. 2017.

[58] K. M. Kettunen *et al.*, “Trim37-deficient mice recapitulate several features of the multiorgan disorder Mulibrey nanism,” *Biol. Open*, 2016.

[59] S. Karlberg, M. Lipsanen-Nyman, H. Lassus, J. Kallijärvi, A. E. Lehesjoki, and R. Butzow, “Gynecological tumors in Mulibrey nanism and role for RING finger protein TRIM37 in the pathogenesis of ovarian fibrothecomas,” *Mod. Pathol.*, 2009.

[60] J. Jiang, C. Yu, M. Chen, S. Tian, and C. Sun, “Over-expression of TRIM37 promotes cell migration and metastasis in hepatocellular carcinoma by activating Wnt/β-catenin signaling,” *Biochem. Biophys. Res. Commun.*, 2015.

[61] N. Karlberg, S. Karlberg, R. Karikoski, S. Mikkola, M. Lipsanen-Nyman, and H. Jalanko, “High frequency of tumours in Mulibrey nanism,” *J. Pathol.*,

2009.

[62] J. Jiang, S. Tian, C. Yu, M. Chen, and C. Sun, “TRIM37 promoted the growth and migration of the pancreatic cancer cells,” *Tumor Biol.*, 2016.

[63] C.-E. E. Hu and J. Gan, “TRIM37 promotes epithelial-mesenchymal transition in colorectal cancer,” *Mol. Med. Rep.*, vol. 15, no. 3, pp. 1057–1062, Mar. 2017.

[64] S. Bhatnagar *et al.*, “TRIM37 is a new histone H2A ubiquitin ligase and breast cancer oncoprotein,” *Nature*, vol. 516, no. 7529, pp. 116–20, Nov. 2014.

[65] P. Zhao, H.-T. Guan, Z.-J. Dai, Y.-G. Ma, X.-X. Liu, and X.-J. Wang, “Knockdown of Tripartite Motif-Containing Protein 37 (TRIM37) Inhibits the Proliferation and Tumorigenesis in Colorectal Cancer Cells,” *Oncol. Res. Featur. Preclin. Clin. Cancer Ther.*, vol. 25, no. 1, pp. 115–122, Jan. 2017.

[66] S. Tang, Y. Gao, and H. Wen-zhong, “Knockdown of TRIM37 suppresses the proliferation, migration and invasion of glioma cells through the inactivation of PI3K/Akt signaling pathway,” *Biomed. Pharmacother.*, vol. 99, pp. 59–64, Mar. 2018.

[67] Y. Tao *et al.*, “TRIM37 promotes tumor cell proliferation and drug resistance in pediatric osteosarcoma.,” *Oncol. Lett.*, vol. 14, no. 6, pp. 6365–6372, Dec. 2017.

[68] G. Wu *et al.*, “An ATM/TRIM37/NEMO axis counteracts genotoxicity by activating nuclear-to-cytoplasmic NF-κB signaling,” *Cancer Res.*, 2018.

[69] T. S. Klann *et al.*, “CRISPR–Cas9 epigenome editing enables high-

throughput screening for functional regulatory elements in the human genome," *Nat. Publ. Gr.*, 2017.

[70] H.-Y. Xue, L.-J. Ji, A.-M. Gao, P. Liu, J.-D. He, and X.-J. Lu, "CRISPR-Cas9 for medical genetic screens: applications and future perspectives.," *J. Med. Genet.*, vol. 53, no. 2, pp. 91–7, Dec. 2015.

[71] J. Pan *et al.*, "A kinome-wide siRNA screen identifies multiple roles for protein kinases in hypoxic stress adaptation, including roles for IRAK4 and GAK in protection against apoptosis in VHL-/- renal carcinoma cells, despite activation of the NF-κB pathway.," *J. Biomol. Screen.*, vol. 18, no. 7, pp. 782–96, Aug. 2013.

[72] J. Perheentupa, S. Autio, S. Leisti, and C. Raitta, "Mulibrey-nanism: dwarfism with muscle, liver, brain and eye involvement.," *Acta Paediatr. Scand. Suppl.*, 1970.

[73] J. A. Doudna and E. Charpentier, "Genome editing. The new frontier of genome engineering with CRISPR-Cas9.," *Science*, vol. 346, no. 6213, p. 1258096, Nov. 2014.

[74] J. L. Schmid-Burgk *et al.*, "A genome-wide CRISPR screen identifies NEK7 as an essential component of NLRP3 inflammasome activation.," *J. Biol. Chem.*, vol. 291, no. 1, pp. 103–9, Nov. 2015.

[75] N. E. Sanjana, O. Shalem, and F. Zhang, "Improved vectors and genome-wide libraries for CRISPR screening," 2014.

[76] J. Winter *et al.*, "CRISPRAnalyzeR: Interactive analysis, annotation and documentation of pooled CRISPR screens," *bioRxiv*, p. 109967, Jan. 2017.

[77] J. Winter *et al.*, “CRISPRAnalyzeR.” .

[78] J. G. Doench *et al.*, “Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation,” *Nat. Biotechnol.*, 2014.

[79] R. Breitling, P. Armengaud, A. Amtmann, and P. Herzyk, “Rank products: A simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments,” *FEBS Lett.*, 2004.

[80] S. B. Thyme, L. Akhmetova, T. G. Montague, E. Valen, and A. F. Schier, “Internal guide RNA interactions interfere with Cas9-mediated cleavage,” *Nat. Commun.*, 2016.

[81] A. Wroblewska *et al.*, “Protein Barcodes Enable High-Dimensional Single-Cell CRISPR Screens,” *Cell*, 2018.

[82] F. A. Ran *et al.*, “Double nicking by RNA-guided CRISPR cas9 for enhanced genome editing specificity,” *Cell*, 2013.

[83] S. Nelson, “A stochastic model of the effects of ionizing radiation on mammalian cells in vitro,” *Bull. Math. Biol.*, vol. 46, no. 3, pp. 423–466, 1984.

[84] R. D. Johnson, “Sister chromatid gene conversion is a prominent double-strand break repair pathway in mammalian cells,” *EMBO J.*, 2000.

[85] A. J. Nakamura, V. A. Rao, Y. Pommier, and W. M. Bonner, “The complexity of phosphorylated H2AX foci formation and DNA repair assembly at DNA double-strand breaks,” *Cell Cycle*, vol. 9, no. 2, pp. 389–397, Jan. 2010.

[86] A. Bothmer, D. F. Robbiani, N. Feldhahn, A. Gazumyan, A. Nussenzweig,

and M. C. Nussenzweig, “53BP1 regulates DNA resection and the choice between classical and alternative end joining during class switch recombination,” *J. Exp. Med.*, vol. 207, no. 4, pp. 855–865, Apr. 2010.

[87] A. Bothmer *et al.*, “Regulation of DNA End Joining, Resection, and Immunoglobulin Class Switch Recombination by 53BP1,” *Mol. Cell*, 2011.

[88] E. L. Leatherbarrow, J. V. Harper, F. A. Cucinotta, and P. O’Neill, “Induction and quantification of γ-H2AX foci following low and high LET-irradiation,” *Int. J. Radiat. Biol.*, 2006.

[89] L. G. Mariotti *et al.*, “Use of the γ-H2AX assay to investigate DNA repair dynamics following multiple radiation exposures,” *PLoS One*, 2013.

[90] F. Meitinger *et al.*, “53BP1 and USP28 mediate p53 activation and G1 arrest after centrosome loss or extended mitotic duration,” *J. Cell Biol.*, vol. 214, no. 2, pp. 155–166, Jul. 2016.

[91] A. Shinohara, H. Ogawa, and T. Ogawa, “Rad51 protein involved in repair and recombination in S. cerevisiae is a RecA-like protein,” *Cell*, 1992.

[92] R. Scully, A. Panday, R. Elango, and N. A. Willis, “DNA double-strand break repair-pathway choice in somatic mammalian cells,” *Nature Reviews Molecular Cell Biology*. 2019.

[93] F. Balestra, P. Strnad, I. Flückiger, and P. Gönczy, “Discovering regulators of centriole biogenesis through siRNA-based functional genomics in human cells,” *Dev. Cell*, 2013.

[94] E. H. Hinchcliffe, F. J. Miller, M. Cham, A. Khodjakov, and G. Sluder, “Requirement of a centrosomal activity for cell cycle progression through

G1 into S phase," *Science (80-.).*, 2001.

[95] E. H. Hinchcliffe, C. Li, E. A. Thompson, J. L. Maller, and G. Sluder, "Requirement of Cdk2-cyclin E activity for repeated centrosome reproduction in xenopus egg extracts," *Science (80-.).*, 1999.

[96] N. Sato *et al.*, "A possible role for centrosome overduplication in radiation-induced cell death," *Oncogene*, 2000.

[97] E. Bourke *et al.*, "DNA damage induces Chk1-dependent centrosome amplification," *EMBO Rep.*, 2007.

[98] N. Sato, K. Mizumoto, M. Nakamura, and M. Tanaka, "Radiation-induced centrosome overduplication and multiple mitotic spindles in human tumor cells," *Exp. Cell Res.*, 2000.

[99] S. Gilberto and M. Peter, "Dynamic ubiquitin signaling in cell cycle regulation," *Journal of Cell Biology*. 2017.

[100] V. Dangiolella *et al.*, "SCF Cyclin F controls centrosome homeostasis and mitotic fidelity through CP110 degradation," *Nature*, 2010.

[101] J. Li *et al.*, "USP33 regulates centrosome biogenesis via deubiquitination of the centriolar protein CP110," *Nature*, 2013.

[102] H. Wang, H. Wang, S. N. Powell, G. Iliakis, and Y. Wang, "ATR affecting cell radiosensitivity is dependent on homologous recombination repair but independent of nonhomologous end joining," *Cancer Res.*, 2004.

[103] R. Prevo *et al.*, "The novel ATR inhibitor VE-821 increases sensitivity of pancreatic cancer cells to radiation and chemotherapy," *Cancer Biol. Ther.*, 2012.

[104] T. Wang *et al.*, “Identification and characterization of essential genes in the human genome.,” *Science*, vol. 350, no. 6264, pp. 1096–101, Nov. 2015.

[105] H. Koike-Yusa, Y. Li, E. P. Tan, M. D. C. Velasco-Herrera, and K. Yusa, “Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library,” *Nat. Biotechnol.*, 2014.

[106] Y. Zhou *et al.*, “High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells,” *Nature*, 2014.

[107] R. J. Park *et al.*, “A genome-wide CRISPR screen identifies a restricted set of HIV host dependency factors,” *Nat. Genet.*, 2017.

[108] K. R. Sanson *et al.*, “Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities,” *Nat. Commun.*, 2018.

[109] E. Haapaniemi, S. Botla, J. Persson, B. Schmierer, and J. Taipale, “CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response,” *Nat. Med.*, 2018.

[110] L. A. Gilbert *et al.*, “Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation,” *Cell*, 2014.

[111] S. Konermann *et al.*, “Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex,” *Nature*, 2015.

[112] B. Adamson *et al.*, “A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response,” *Cell*, 2016.

[113] A. Dixit *et al.*, “Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens,” *Cell*, 2016.

Appendix

Supplementary tables.

Table A. Top-depleted genes identified as modulators of radiosensitivity by the CRISPR genome-wide radiosensitivity screen in MRC-5 cells presented in Chapter 4.

Gene Name	Rank	False Positive Ratio (FPF)	p value
KIF2C	1	0.710	0.00005
DMTF1	2	0.558	0.00008
SRI	3	1.005	0.00021
SEC61G	4	0.844	0.00023
TRIM37	5	0.689	0.00024
ANAPC4	6	0.731	0.00030
PAXBP1	7	0.778	0.00038
B3GNT9	8	0.809	0.00045
NPBWR1	9	0.917	0.00057
MED27	10	1.017	0.00070
MIPOL1	11	0.928	0.00070
HABP4	12	0.961	0.00079
RSRC1	13	0.933	0.00084
HOXA4	14	0.998	0.00096
CLEC4E	15	1.035	0.00107
FOXI3	16	1.082	0.00119
HDDC2	17	1.047	0.00123
PON3	18	1.009	0.00125
JAKMIP2	19	0.970	0.00127
FSIP2	20	0.972	0.00134
PWWP2A	21	0.929	0.00134
NKX2-2	22	0.938	0.00142
FHL2	23	0.939	0.00149
ZBTB17	24	0.916	0.00152
TTC30A	25	0.884	0.00152
NLGN3	26	0.863	0.00155
ADAMTS6	27	0.848	0.00158
GPAM	28	0.818	0.00158
DNM1	29	0.797	0.00159
BIRC7	30	0.810	0.00167
FSD1L	31	0.799	0.00171
SEC61A2	32	0.782	0.00172
PRAMEF7	33	0.851	0.00194
TMEM64	34	0.842	0.00197
SIPA1L2	35	0.879	0.00212

Gene Name	Rank	False Positive Ratio (FPF)	p value
ZBTB45	36	0.910	0.00226
MSMP	37	0.885	0.00226
LGR6	38	0.907	0.00238
DPF2	39	0.912	0.00245
SH2B3	40	0.903	0.00249
PRKAG3	41	0.922	0.00260
ZBBX	42	0.931	0.00269
FAM98C	43	0.961	0.00285
SESN2	44	0.952	0.00289
OSBPL6	45	0.941	0.00292
TRAF3IP2	46	0.971	0.00308
OSTM1	47	0.968	0.00313
BLID	48	0.953	0.00315
ALG9	49	0.963	0.00325
HAGHL	50	0.963	0.00332
CARD16	51	0.966	0.00339
MGAT1	52	0.978	0.00350
SIX2	53	0.963	0.00352
LOC100505549	54	0.952	0.00354
TEX36	55	0.940	0.00356
PTPRJ	56	0.932	0.00360
TMX2	57	0.977	0.00384
RNFT2	58	0.983	0.00393
TRIM43B	59	0.989	0.00402
NAALAD2	60	0.978	0.00404
DTX4	61	0.969	0.00407
RIMBP2	62	0.971	0.00415
ABCF1	63	0.964	0.00418
TFAM	64	0.952	0.00420
MRPS2	65	0.940	0.00421
C7orf43	66	0.940	0.00427
SLC27A5	67	0.945	0.00436
TEX101	68	0.959	0.00449
NAV2	69	0.959	0.00456
ZBED3	70	0.956	0.00461
MS4A3	71	0.955	0.00467
HACE1	72	0.942	0.00468
PPP1R35	73	0.941	0.00474
SWSAP1	74	0.941	0.00480
COQ6	75	0.938	0.00485
NPAS4	76	0.928	0.00486
DISP2	77	0.930	0.00493
SMC1B	78	0.947	0.00509

Gene Name	Rank	False Positive Ratio (FPF)	p value
ZNF70	79	0.949	0.00516
TSNARE1	80	0.951	0.00524
TLR7	81	0.943	0.00527
TNIP3	82	0.942	0.00532
SEC22A	83	0.933	0.00534
KLHDC7B	84	0.931	0.00539
APLNR	85	0.921	0.00539
THAP6	86	0.913	0.00541
ANKRD36	87	0.903	0.00541
MFRP	88	0.894	0.00542
ARL4A	89	0.890	0.00546
PIGW	90	0.884	0.00548
ZNF322	91	0.883	0.00554
DDX20	92	0.879	0.00557
SGTA	93	0.882	0.00565
DAG1	94	0.888	0.00575
GORASP1	95	0.934	0.00612
ZNF778	96	0.929	0.00615
SLCO1B7	97	0.922	0.00616
CCDC38	98	0.926	0.00625
RNF125	99	0.919	0.00627
SPIRE2	100	0.916	0.00631
NEK5	101	0.911	0.00634
ALG8	102	0.906	0.00637
TMEM147	103	0.903	0.00641
LRRC37A2	104	0.902	0.00646
CSK	105	0.916	0.00663
PDGFRB	106	0.925	0.00676
TLR1	107	0.917	0.00676
ZCCHC4	108	0.913	0.00680
PDX1	109	0.929	0.00697
ZFP69	110	0.927	0.00702
C1orf95	111	0.928	0.00710
GPM6B	112	0.921	0.00711
FGF7	113	0.982	0.00765
LYPD2	114	0.992	0.00780
SUN5	115	0.985	0.00780
RAB3GAP2	116	0.987	0.00789
ZWINT	117	0.988	0.00796
EIF4EBP1	118	0.980	0.00797
PAK4	119	0.981	0.00805
U2SURP	120	0.975	0.00806
GPR19	121	0.967	0.00807

Gene Name	Rank	False Positive Ratio (FPF)	p value
C12orf74	122	0.962	0.00808
VPS54	123	0.971	0.00823
CDCA5	124	0.968	0.00827
KIAA1456	125	0.965	0.00831
ANKRD9	126	0.962	0.00835
RIF1	127	0.962	0.00842
ERCC5	128	0.958	0.00845
VPS37C	129	0.974	0.00866
LY6G6C	130	0.967	0.00867
CC2D2A	131	0.963	0.00870
HRSP12	132	0.957	0.00870
RXRB	133	0.962	0.00882
NCK1	134	0.957	0.00883
WNT10A	135	0.956	0.00889
RASL12	136	0.954	0.00894
NADK	137	0.961	0.00907
HIC1	138	0.964	0.00917
PDS5B	139	0.961	0.00920
C11orf45	140	0.957	0.00923
HYI	141	0.957	0.00930
LCE5A	142	0.957	0.00936
C16orf93	143	0.953	0.00939
MESP1	144	0.952	0.00944
PARP4	145	0.947	0.00946
WDR43	146	0.941	0.00947
IMPA2	147	0.950	0.00962
HCN1	148	0.947	0.00966
KHDC1	149	0.941	0.00966
ZMYM4	150	0.937	0.00968
INTS2	151	0.932	0.00969
ETFB	152	0.950	0.00995
ZNF629	153	0.945	0.00996
C19orf52	154	0.945	0.01002
CDH16	155	0.943	0.01007
BCDIN3D	156	0.937	0.01007
DAXX	157	0.934	0.01010
PCSK1N	158	0.932	0.01015
DMD	159	0.929	0.01017
ABR	160	0.924	0.01019
BLNK	161	0.927	0.01028
IFI16	162	0.928	0.01036
OR5H14	163	0.924	0.01038
C7orf57	164	0.921	0.01040

Gene Name	Rank	False Positive Ratio (FPF)	p value
TRPV2	165	0.920	0.01046
MYPOP	166	0.917	0.01049
KPNA6	167	0.914	0.01052
SPINK5	168	0.913	0.01057
FRY	169	0.910	0.01059
PPIP5K2	170	0.905	0.01060
NOL3	171	0.900	0.01061
FAM20C	172	0.895	0.01061
EID2	173	0.893	0.01065
FGF21	174	0.895	0.01073
ZNF599	175	0.899	0.01084
OFD1	176	0.902	0.01094
TAF5	177	0.905	0.01104
PRDM15	178	0.910	0.01116
COMMD10	179	0.907	0.01118
SLBP	180	0.906	0.01124
C15orf53	181	0.907	0.01132
APLN	182	0.908	0.01138
TMEM39A	183	0.903	0.01139
PSMD12	184	0.899	0.01139
HHEX	185	0.895	0.01140
TCFL5	186	0.890	0.01140
YPEL5	187	0.889	0.01145
C5orf52	188	0.889	0.01151
INA	189	0.901	0.01173
UVRAG	190	0.899	0.01176
TSR3	191	0.896	0.01179
ZAR1	192	0.892	0.01180
SHQ1	193	0.902	0.01199
EPYC	194	0.909	0.01214
TPST1	195	0.910	0.01222
TEKT1	196	0.908	0.01226
CYBA	197	0.909	0.01234
PRKD3	198	0.911	0.01243
HP1BP3	199	0.911	0.01248
CRCT1	200	0.906	0.01249
HUWE1	201	0.903	0.01251
BST2	202	0.904	0.01257
NUTM1	203	0.908	0.01270
PDK1	204	0.911	0.01280
TMEM17	205	0.907	0.01282
DDX60	206	0.912	0.01294
FECH	207	0.908	0.01296

Gene Name	Rank	False Positive Ratio (FPF)	p value
VPS18	208	0.918	0.01316
CENPP	209	0.917	0.01320
PDIA4	210	0.919	0.01329
FAM229B	211	0.926	0.01346
MAP3K3	212	0.923	0.01348
GJB2	213	0.929	0.01363
UBR5	214	0.928	0.01369
LRRC42	215	0.926	0.01372
CLECL1	216	0.931	0.01385
DCTN1	217	0.930	0.01391
STRN4	218	0.930	0.01396
T	219	0.926	0.01397
SOX12	220	0.926	0.01404
ZDHHC6	221	0.923	0.01406
GID8	222	0.920	0.01407
NKD2	223	0.918	0.01410
KCNN1	224	0.916	0.01413
DERL3	225	0.914	0.01417
C1orf210	226	0.913	0.01422
SLC9A8	227	0.911	0.01425
NPPB	228	0.919	0.01443
MTMR3	229	0.915	0.01444
SLC22A3	230	0.913	0.01447
SLC45A2	231	0.911	0.01450
REEP3	232	0.915	0.01463
CD46	233	0.922	0.01480
NUDT21	234	0.932	0.01502
SSTR2	235	0.928	0.01503
RASL10A	236	0.934	0.01519
FAM110C	237	0.935	0.01526
RPA3	238	0.931	0.01527
FAM83F	239	0.931	0.01534
GHRHR	240	0.928	0.01534
PRR23A	241	0.932	0.01548
GPT2	242	0.929	0.01549
ZIC5	243	0.927	0.01553
PORCN	244	0.927	0.01559
PRPF4	245	0.928	0.01566
CST6	246	0.933	0.01582
KATNAL2	247	0.933	0.01588
CTSF	248	0.947	0.01619
NPPA	249	0.947	0.01625
RPS26	250	0.962	0.01657

Gene Name	Rank	False Positive Ratio (FPF)	p value
TSPAN11	251	0.969	0.01675
HSPA4	252	0.971	0.01686
C14orf183	253	0.968	0.01688
TBC1D9B	254	0.966	0.01691
B3GNT2	255	0.965	0.01696
LMNB2	256	0.965	0.01702
MPP5	257	0.961	0.01702
MMP23B	258	0.958	0.01703
BZRAP1	259	0.956	0.01706
AMDHD1	260	0.956	0.01712
OR2AK2	261	0.952	0.01713
OR2A25	262	0.978	0.01766
IFNA16	263	0.981	0.01777
SEMA7A	264	0.983	0.01789
SLC6A19	265	0.980	0.01789
TBXA2R	266	0.976	0.01790
DNAJB3	267	0.974	0.01792
PPCS	268	0.971	0.01793
FST	269	0.968	0.01794
ARHGAP9	270	0.967	0.01798
ATAD1	271	0.964	0.01800
RRP15	272	0.965	0.01808
FCGR2A	273	0.962	0.01809
RPP40	274	0.964	0.01820
C1QTNF4	275	0.961	0.01821
ZNF732	276	0.958	0.01822
OR52B6	277	0.955	0.01822
ELMSAN1	278	0.954	0.01828
USF2	279	0.952	0.01830
LIN9	280	0.950	0.01833
ZSWIM1	281	0.948	0.01835
ZFP36	282	0.950	0.01846
HIST1H1D	283	0.949	0.01850
C20orf166	284	0.946	0.01850
MSTO1	285	0.943	0.01853
OR52I2	286	0.941	0.01854
DEFA6	287	0.944	0.01866
PFDN6	288	0.943	0.01872
UBA5	289	0.941	0.01874
SLC2A3	290	0.940	0.01877
SMIM4	291	0.941	0.01886
ARHGAP21	292	0.941	0.01893
KANK3	293	0.940	0.01897

Gene Name	Rank	False Positive Ratio (FPF)	p value
COQ2	294	0.943	0.01909
RAB11A	295	0.944	0.01919
ERC2	296	0.942	0.01920
PEX6	297	0.940	0.01923
FAM76B	298	0.947	0.01944
BTBD8	299	0.948	0.01953
ACSL1	300	0.946	0.01955
IGHMBP2	301	0.943	0.01955
MBP	302	0.943	0.01962
C14orf119	303	0.943	0.01968
SH3PXD2B	304	0.949	0.01987
BRPF1	305	0.950	0.01996
MKL1	306	0.949	0.02001
SCN4B	307	0.950	0.02009
CYB561A3	308	0.950	0.02017
RASSF8	309	0.950	0.02023
FAM129B	310	0.948	0.02024
TUBA1A	311	0.945	0.02026
SPC25	312	0.957	0.02058
HSPB2	313	0.955	0.02059
LGR4	314	0.954	0.02064
KIF4A	315	0.952	0.02066
CBLN4	316	0.949	0.02066
ZCCHC3	317	0.948	0.02071
ZCCHC2	318	0.947	0.02075
PHF21A	319	0.945	0.02078
C11orf87	320	0.942	0.02078
GUCY2F	321	0.942	0.02083
LRRIQ4	322	0.942	0.02090
MRGPRD	323	0.940	0.02093
ATP8B4	324	0.945	0.02109
OR2V1	325	0.944	0.02114
DIEXF	326	0.942	0.02116
CTSC	327	0.942	0.02123
AKAP3	328	0.940	0.02125
PDCD1	329	0.942	0.02136
ZBTB22	330	0.942	0.02141
PNLIPRP3	331	0.940	0.02143
LYSMD4	332	0.942	0.02156
HMX3	333	0.945	0.02168
CLSTN2	334	0.953	0.02192
IL1RL1	335	0.956	0.02206
MICU1	336	0.955	0.02210

Gene Name	Rank	False Positive Ratio (FPF)	p value
SYNE1	337	0.959	0.02226
EMC7	338	0.958	0.02231
DSEL	339	0.956	0.02233
PPID	340	0.954	0.02235
SFTPA1	341	0.955	0.02243
PJA1	342	0.952	0.02244
SHROOM4	343	0.955	0.02257
CRYGC	344	0.953	0.02260
LATS2	345	0.951	0.02260
FUT4	346	0.951	0.02266
MYO10	347	0.948	0.02267
XYLT1	348	0.947	0.02271
DEFB1	349	0.948	0.02280
SENP1	350	0.956	0.02306
CLMN	351	0.959	0.02319
DHCR24	352	0.961	0.02330
GOLGA7	353	0.959	0.02333
RUFY1	354	0.958	0.02337
C10orf55	355	0.958	0.02343
CELF3	356	0.959	0.02351
AOX1	357	0.961	0.02363
OR4M2	358	0.960	0.02368
ARMC4	359	0.958	0.02370
CWF19L1	360	0.956	0.02371
NDUFA2	361	0.955	0.02375
TWF1	362	0.952	0.02376
SLC28A1	363	0.954	0.02387
HFE2	364	0.953	0.02391
RRAGB	365	0.952	0.02393
UQCRC2	366	0.951	0.02399
KCNS3	367	0.954	0.02412
EML3	368	0.954	0.02418
0302	369	0.956	0.02431
UPRT	370	0.958	0.02441
C9orf163	371	0.958	0.02448
SEMA4D	372	0.958	0.02456
PTN	373	0.959	0.02464
VPS53	374	0.960	0.02475
CDK5RAP3	375	0.960	0.02479
CASP8	376	0.959	0.02485
SLC2A14	377	0.958	0.02488
OGDHL	378	0.956	0.02489
ATR	379	0.954	0.02490

Gene Name	Rank	False Positive Ratio (FPF)	p value
STAP1	380	0.954	0.02497
CYP26A1	381	0.952	0.02499
MLST8	382	0.952	0.02505
STX10	383	0.952	0.02511
EMC10	384	0.952	0.02519
ATP5H	385	0.952	0.02525
PRKCD	386	0.949	0.02525
ZC3HAV1	387	0.948	0.02526
WIPI2	388	0.947	0.02532
DUSP1	389	0.946	0.02534
CTBP1	390	0.944	0.02538
CD6	391	0.943	0.02539
GAB2	392	0.951	0.02567
INPP5B	393	0.951	0.02576
OR56A4	394	0.950	0.02578
RNF152	395	0.948	0.02581
TAF4	396	0.946	0.02581
UGP2	397	0.944	0.02583
SLC26A9	398	0.943	0.02587
ATAD2B	399	0.944	0.02594
ASNA1	400	0.945	0.02604
MAN2C1	401	0.943	0.02604
OSBPL3	402	0.948	0.02625
TCTEX1D4	403	0.947	0.02629
RFNG	404	0.945	0.02632
CYB5A	405	0.944	0.02635
METTL17	406	0.942	0.02635
SLC2A9	407	0.940	0.02637
ZCCHC7	408	0.939	0.02640
0903	409	0.938	0.02643
LACTB	410	0.941	0.02657
POU1F1	411	0.939	0.02658
XRCC2	412	0.941	0.02671
UBXN7	413	0.942	0.02680
CYP20A1	414	0.941	0.02683
TOMM6	415	0.939	0.02684
RAD23A	416	0.937	0.02686
LCE3B	417	0.936	0.02688
NTF4	418	0.935	0.02692
UGT2B4	419	0.939	0.02710
PDRG1	420	0.946	0.02738
IQSEC1	421	0.951	0.02758
MTNR1B	422	0.950	0.02761

Gene Name	Rank	False Positive Ratio (FPF)	p value
CCDC157	423	0.948	0.02764
ZBTB3	424	0.947	0.02767
FAM115C	425	0.945	0.02767
C19orf57	426	0.947	0.02779
RUSC2	427	0.947	0.02786
FHOD3	428	0.948	0.02794
CASKIN2	429	0.946	0.02797
DNAJC14	430	0.944	0.02797
BAI2	431	0.943	0.02801
LTB4R	432	0.941	0.02802
PSG4	433	0.939	0.02802
VAPA	434	0.937	0.02803
EXO5	435	0.936	0.02805
TUBGCP6	436	0.936	0.02813
PIP4K2B	437	0.936	0.02818
CDC37	438	0.936	0.02823
FAM207A	439	0.935	0.02828
FMO2	440	0.934	0.02830
RASGRP2	441	0.936	0.02843
ADO	442	0.934	0.02844
ZSCAN5A	443	0.934	0.02849
HOXA13	444	0.932	0.02851
ZXDA	445	0.933	0.02860
ZNF713	446	0.931	0.02862
RARA	447	0.936	0.02881
LOC100129520	448	0.935	0.02886
QSOX2	449	0.936	0.02897
IRAK1	450	0.935	0.02898
CLDN22	451	0.934	0.02901
NSF	452	0.933	0.02905
PCSK6	453	0.932	0.02910
ZBTB34	454	0.931	0.02911
ACTL8	455	0.931	0.02918
LYSMD1	456	0.930	0.02923
KIAA1755	457	0.929	0.02926
BAZ2A	458	0.929	0.02930
CHRND	459	0.930	0.02940
TTBK1	460	0.932	0.02952
C3orf72	461	0.931	0.02956
WAS	462	0.935	0.02977
PUF60	463	0.937	0.02989
ATP6V0A4	464	0.936	0.02991
ARHGEF2	465	0.935	0.02994

Gene Name	Rank	False Positive Ratio (FPF)	p value
USP24	466	0.938	0.03010
SLC39A11	467	0.936	0.03012
CLYBL	468	0.935	0.03014
NSMCE1	469	0.934	0.03017
ITGB2	470	0.936	0.03031
PKD2	471	0.941	0.03055
SCN1A	472	0.940	0.03058
DPPA2	473	0.940	0.03063
MED4	474	0.940	0.03068
USP49	475	0.938	0.03070
POP5	476	0.937	0.03074
UCKL1	477	0.936	0.03075
P2RY10	478	0.936	0.03082
KRTAP19-3	479	0.934	0.03084
SERPINB1	480	0.933	0.03086
PPP2R3A	481	0.932	0.03087
ZNF433	482	0.936	0.03107
ANKFN1	483	0.935	0.03110
ATP6V1G2	484	0.933	0.03111
ALDH6A1	485	0.933	0.03118
RIOK1	486	0.931	0.03118
MYCBPAP	487	0.932	0.03126
LHCGR	488	0.930	0.03127
MUCL1	489	0.929	0.03130
TMEM234	490	0.930	0.03141
DDX5	491	0.930	0.03146
CDC6	492	0.928	0.03146
LHX5	493	0.927	0.03149
KLHDC7A	494	0.929	0.03161
TMEM185B	495	0.927	0.03162
RELB	496	0.926	0.03163
CPLX1	497	0.925	0.03166
CYP8B1	498	0.924	0.03170
CDC34	499	0.922	0.03170
IRF9	500	0.923	0.03180
HMGXB3	501	0.923	0.03184
DKK3	502	0.921	0.03187
UBE2B	503	0.920	0.03187
WNT2	504	0.920	0.03195
MDM1	505	0.919	0.03198
DHRS13	506	0.920	0.03208
TSPYL6	507	0.919	0.03209
PDCD10	508	0.919	0.03218

Gene Name	Rank	False Positive Ratio (FPF)	p value
TMEM151A	509	0.921	0.03230
EXO1	510	0.921	0.03236
OR8J3	511	0.920	0.03237
LENEP	512	0.922	0.03252
TIPIN	513	0.922	0.03258
CNBD2	514	0.924	0.03272
THRA	515	0.929	0.03295
PGP	516	0.928	0.03299
PITHD1	517	0.928	0.03306
DMRT3	518	0.927	0.03307
CYB5R2	519	0.926	0.03310
CIRBP	520	0.928	0.03326
TSPAN13	521	0.931	0.03341
CDX4	522	0.929	0.03343
C3orf80	523	0.933	0.03361
OTUD1	524	0.934	0.03374
SLC5A3	525	0.934	0.03377
LCNL1	526	0.933	0.03380
TMEM57	527	0.931	0.03380
ASAH2B	528	0.931	0.03388
MTSS1	529	0.930	0.03391
TTC21B	530	0.930	0.03395
CCT6A	531	0.929	0.03397
LHX2	532	0.930	0.03409
DDX51	533	0.930	0.03416
NUBP1	534	0.929	0.03419
SRF	535	0.928	0.03422
IFT43	536	0.927	0.03423
VAT1	537	0.927	0.03428
SLC19A3	538	0.926	0.03433
ATPIF1	539	0.925	0.03435
KLF1	540	0.924	0.03438
TCF4	541	0.924	0.03445
EVX1	542	0.924	0.03451
HIST1H2AK	543	0.924	0.03456
FBXW7	544	0.929	0.03481
ANTXR2	545	0.929	0.03487
TMEM82	546	0.927	0.03488
DEGS1	547	0.927	0.03495
AGBL3	548	0.926	0.03496
HEPHL1	549	0.925	0.03498
SLC25A41	550	0.924	0.03501
ATAD3B	551	0.925	0.03513

Gene Name	Rank	False Positive Ratio (FPF)	p value
CUEDC2	552	0.924	0.03514
LOC728392	553	0.925	0.03523
GH1	554	0.923	0.03523
TUFT1	555	0.924	0.03535
RASGRF2	556	0.924	0.03540
SNX17	557	0.930	0.03567
PRR14	558	0.929	0.03572
FAM124B	559	0.928	0.03574
ITPKA	560	0.929	0.03584
PPOX	561	0.927	0.03585
TAF9	562	0.927	0.03588
SLAMF9	563	0.925	0.03589
GPR119	564	0.925	0.03594
ITGB1BP1	565	0.924	0.03596
TRPA1	566	0.922	0.03596
PCSK7	567	0.921	0.03597
ANKRD30B	568	0.921	0.03602
TAGLN2	569	0.919	0.03604
TMEM54	570	0.920	0.03613
OR3A1	571	0.919	0.03614
DPPA5	572	0.918	0.03619
TCEB2	573	0.917	0.03622
TCEAL5	574	0.921	0.03641
RPS6KA1	575	0.921	0.03650
SMC5	576	0.922	0.03659
SEMA6B	577	0.921	0.03660
PLXNB2	578	0.923	0.03675
ZMYND19	579	0.922	0.03677
GAST	580	0.922	0.03684
PROSER1	581	0.924	0.03701
C14orf142	582	0.925	0.03711
HTT	583	0.925	0.03716
ZNF695	584	0.924	0.03718
MTA3	585	0.926	0.03733
FGFR1OP2	586	0.925	0.03736
YRDC	587	0.925	0.03741
AUH	588	0.924	0.03743
OR5V1	589	0.923	0.03745
FAM110D	590	0.923	0.03751
QKI	591	0.924	0.03762
ASB2	592	0.923	0.03765
GPR158	593	0.929	0.03795
GAL3ST1	594	0.928	0.03799

Gene Name	Rank	False Positive Ratio (FPF)	p value
EFNA3	595	0.927	0.03799
BTBD16	596	0.926	0.03801
NAP1L3	597	0.925	0.03803
CCDC84	598	0.924	0.03808
SLC39A6	599	0.926	0.03821
C10orf62	600	0.924	0.03822
GBP6	601	0.926	0.03834
ATP5B	602	0.928	0.03847
PNCK	603	0.928	0.03855
GRIK4	604	0.929	0.03868
CHRNB3	605	0.930	0.03876
MAGI2	606	0.928	0.03876
KRTAP12-3	607	0.929	0.03886
S100A7A	608	0.932	0.03904
WDR63	609	0.931	0.03904
TAS2R46	610	0.931	0.03912
ACAP1	611	0.929	0.03912
C3orf30	612	0.928	0.03914
TECRL	613	0.927	0.03917
WDR83	614	0.926	0.03918
SUPT7L	615	0.929	0.03937
WNT3A	616	0.929	0.03944
ELMO3	617	0.928	0.03944
CMA1	618	0.926	0.03944
IFNGR2	619	0.932	0.03976
C14orf169	620	0.932	0.03979
C11orf88	621	0.931	0.03984
FEM1B	622	0.931	0.03991
CAMSAP3	623	0.930	0.03992
HMX1	624	0.929	0.03993
SLC35A2	625	0.928	0.03996
ZNF98	626	0.931	0.04016
ALDH5A1	627	0.930	0.04017
SAMD9L	628	0.929	0.04021
TMEM161B	629	0.930	0.04029
C2	630	0.931	0.04039
SHFM1	631	0.931	0.04047
NFE2	632	0.931	0.04053
THRSP	633	0.930	0.04054
KCNIP2	634	0.929	0.04056
LCP1	635	0.929	0.04062
FOXN4	636	0.927	0.04064
RBM47	637	0.926	0.04064

Gene Name	Rank	False Positive Ratio (FPF)	p value
RPTOR	638	0.926	0.04070
CXCL16	639	0.925	0.04071
KIAA0895L	640	0.924	0.04076
OR1K1	641	0.932	0.04115
TRAK1	642	0.933	0.04127
COX5B	643	0.934	0.04139
TTC16	644	0.934	0.04142
PPP1R14A	645	0.933	0.04147
NOP9	646	0.934	0.04156
PFKFB1	647	0.933	0.04158
F10	648	0.935	0.04175
APOL5	649	0.934	0.04176
CCNB1	650	0.933	0.04180
LITAF	651	0.933	0.04184
AQPEP	652	0.934	0.04195
ZNF362	653	0.933	0.04196
HMMR	654	0.932	0.04200
ABHD1	655	0.932	0.04208
ITM2A	656	0.934	0.04222
GDAP1	657	0.933	0.04224
PLA2G4F	658	0.934	0.04236
ST7L	659	0.934	0.04239
KCNJ8	660	0.933	0.04242
TNFSF10	661	0.932	0.04245
MCIDAS	662	0.931	0.04247
PNRC1	663	0.930	0.04249
EXOC1	664	0.931	0.04257
FAM78A	665	0.930	0.04259
MYOM3	666	0.929	0.04262
LCORL	667	0.928	0.04265
NCAPG2	668	0.929	0.04274
ARMC7	669	0.928	0.04279
AATK	670	0.931	0.04297
RYR1	671	0.931	0.04305
DIS3L	672	0.930	0.04307
FGL2	673	0.930	0.04311
TLX1NB	674	0.929	0.04315
CUL1	675	0.929	0.04320
CIRH1A	676	0.929	0.04326
HIATL1	677	0.929	0.04332
TEX22	678	0.928	0.04334
BCL9	679	0.927	0.04335
CLVS2	680	0.928	0.04347

Gene Name	Rank	False Positive Ratio (FPF)	p value
VMA21	681	0.927	0.04349
CBX3	682	0.928	0.04359
SH3BGR	683	0.928	0.04368
POU5F1	684	0.929	0.04376
AURKAIP1	685	0.928	0.04379
BSND	686	0.928	0.04386
S100A8	687	0.928	0.04390
FOXN3	688	0.928	0.04397
ZNF835	689	0.928	0.04406
PSMC1	690	0.927	0.04409
TNFRSF13B	691	0.929	0.04424
ZNF415	692	0.929	0.04428
DRP2	693	0.929	0.04434
SLC2A11	694	0.928	0.04436
GLYATL2	695	0.930	0.04454
CBLN3	696	0.929	0.04455
DLX4	697	0.930	0.04465
C14orf37	698	0.931	0.04477
ADAM11	699	0.932	0.04486
COX5A	700	0.931	0.04488
RTEL1	701	0.930	0.04490
LOC650293	702	0.931	0.04501
KXD1	703	0.929	0.04502
ZCCHC12	704	0.931	0.04515
IFI30	705	0.930	0.04518
MASTL	706	0.930	0.04522
ZMIZ1	707	0.929	0.04524
MED31	708	0.929	0.04530
GNAI2	709	0.928	0.04531
CCK	710	0.929	0.04542
ELP2	711	0.927	0.04542
DDHD2	712	0.928	0.04551
ZNF426	713	0.930	0.04569
MUT	714	0.930	0.04575
PAICS	715	0.931	0.04588
ATP6V1G1	716	0.933	0.04601
DNAJC1	717	0.932	0.04603
CHRNA5	718	0.932	0.04609
CNTN2	719	0.931	0.04614
CPSF6	720	0.937	0.04650
THAP4	721	0.936	0.04651
PNISR	722	0.935	0.04651
KLHL13	723	0.934	0.04654

Gene Name	Rank	False Positive Ratio (FPF)	p value
SLX4	724	0.934	0.04657
DGKQ	725	0.934	0.04664
PPIE	726	0.939	0.04698
LRRC57	727	0.938	0.04700
FAM53C	728	0.938	0.04706
GSG1	729	0.937	0.04708
SH3GL1	730	0.938	0.04717
SELV	731	0.937	0.04721
ZNF655	732	0.936	0.04722
BCL11B	733	0.935	0.04722
SOCS4	734	0.935	0.04727
AMMECR1L	735	0.935	0.04735
GPR110	736	0.935	0.04739
CELSR3	737	0.934	0.04740
TPST2	738	0.933	0.04744
CXXC11	739	0.938	0.04775
ATXN2	740	0.938	0.04780
GAP43	741	0.937	0.04785
PDS5A	742	0.936	0.04786
SEPSECS	743	0.935	0.04787
RB1	744	0.938	0.04807
DENND1C	745	0.938	0.04812
WDR88	746	0.938	0.04822
PAFAH2	747	0.939	0.04834
MMADHC	748	0.938	0.04835
GHRH	749	0.937	0.04838
MSH3	750	0.937	0.04839
LAMA1	751	0.936	0.04842
LURAP1L	752	0.936	0.04851
GSE1	753	0.935	0.04853
MCM3AP	754	0.941	0.04889
STK11	755	0.942	0.04900
ZNF484	756	0.943	0.04910
SLC25A31	757	0.943	0.04921
PPP5C	758	0.943	0.04927
GABRR1	759	0.945	0.04944
VPS72	760	0.945	0.04948
OVCH2	761	0.946	0.04959
CCL25	762	0.950	0.04986

Table B. Top-50-depleted genes identified as essential genes for MRC-5 cells survival.

Gene Name	Rank	False Positive Ratio (FPF)	p value
RPL23A	1	0.00010290	0.00000001
SRP9	2	0.00022510	0.00000003
GAPDH	3	0.00017200	0.00000004
SOD1	4	0.00060900	0.00000017
CCT4	5	0.00059660	0.00000021
PGAM1	6	0.00259000	0.00000107
SNRPG	7	0.00278200	0.00000134
UBL5	8	0.00875300	0.00000482
EIF2B2	9	0.00996700	0.00000618
WASF2	10	0.00902000	0.00000622
TCP1	11	0.00906800	0.00000687
RPL18A	12	0.01389000	0.00001149
HAUS3	13	0.01412000	0.00001265
RPL10	14	0.01812000	0.00001748
TEAD1	15	0.01849000	0.00001911
SLC25A28	16	0.01773000	0.00001955
IMP4	17	0.01747000	0.00002047
RPL35	18	0.01680000	0.00002083
TRIP13	19	0.01700000	0.00002225
DHX33	20	0.01901000	0.00002620
RPS3A	21	0.02090000	0.00003024
RPS24	22	0.02255000	0.00003418
SNRPE	23	0.02191000	0.00003472
PSMB3	24	0.02181000	0.00003606
GINS3	25	0.02183000	0.00003760
PSMA3	26	0.03147000	0.00005637
KTI12	27	0.03712000	0.00006905
POLR2D	28	0.03880000	0.00007485
LSM11	29	0.03832000	0.00007656
PTPMT1	30	0.03742000	0.00007735
NCAPG	31	0.04242000	0.00009060
RPL3	32	0.04344000	0.00009578

Gene Name	Rank	False Positive Ratio (FPF)	p value
SPATA5	33	0.04705000	0.00010700
GFER	34	0.04636000	0.00010860
TBCC	35	0.05025000	0.00012120
TRAPPC1	36	0.05629000	0.00013960
NOP16	37	0.05532000	0.00014100
TAF8	38	0.06049000	0.00015840
FERMT2	39	0.06071000	0.00016310
RPL9	40	0.06682000	0.00018420
RHOG	41	0.06923000	0.00019560
ARFRP1	42	0.07225000	0.00020910
MSTO1	43	0.07141000	0.00021160
TPX2	44	0.08030000	0.00024340
CDC16	45	0.08609000	0.00026690
SUMO2	46	0.08941000	0.00028340
YEATS2	47	0.08891000	0.00028790
NAA10	48	0.08749000	0.00028930
SAP130	49	0.09276000	0.00031320
RPS15A	50	0.09324000	0.00032120

Table C. Top-50-enriched genes identified as genes whose depletion increased proliferation of MRC-5 cells.

Gene Name	Rank	False Positive Ratio (FPF)	p value
TP53	1	0.0000015120	0.0000000001
PDCD10	2	0.0000015520	0.0000000002
LATS2	3	0.0000457900	0.0000000095
PRDM6	4	0.0489800000	0.0000135000
TAOK1	5	0.0749800000	0.0000258300
UBE2L3	6	0.1100000000	0.0000454900
CAB39	7	0.2949000000	0.0001423000
SH3GL1	8	0.3359000000	0.0001851000
CC2D2A	9	0.3115000000	0.0001932000
RNF146	10	0.4239000000	0.0002921000
ZAR1	11	0.4225000000	0.0003202000
HBM	12	0.4422000000	0.0003656000
MAGEE2	13	0.4096000000	0.0003668000
RAP2C	14	0.4033000000	0.0003890000
ZBBX	15	0.3923000000	0.0004054000
EIF4EBP1	16	0.4547000000	0.0005012000
ZBTB34	17	0.5041000000	0.0005904000
AUTS2	18	0.5525000000	0.0006853000
FUT4	19	0.5710000000	0.0007475000
CCDC107	20	0.5558000000	0.0007658000
TMEM57	21	0.5421000000	0.0007844000
ISX	22	0.5202000000	0.0007886000
PI3	23	0.5460000000	0.0008653000
AHDC1	24	0.5620000000	0.0009293000
HIC1	25	0.5640000000	0.0009715000
NUCKS1	26	0.5531000000	0.0009907000
SOX12	27	0.5365000000	0.0009981000
GPAM	28	0.5249000000	0.0010130000
GALNT15	29	0.5209000000	0.0010410000
ZNF212	30	0.5211000000	0.0010770000
TMEM179	31	0.5053000000	0.0010790000
DCTN1	32	0.4996000000	0.0011020000

Gene Name	Rank	False Positive Ratio (FPF)	p value
TRAF3	33	0.4919000000	0.0011180000
KRTAP10-9	34	0.5263000000	0.0012330000
RIMBP2	35	0.5381000000	0.0012980000
MAGEB6	36	0.5835000000	0.0014470000
ECH1	37	0.5795000000	0.0014770000
KRTAP3-3	38	0.5861000000	0.0015340000
NLGN3	39	0.6084000000	0.0016350000
TUBB6	40	0.6358000000	0.0017520000
OR51S1	41	0.6229000000	0.0017600000
MEGF6	42	0.6380000000	0.0018460000
TMEM181	43	0.6792000000	0.0020120000
STAP1	44	0.6656000000	0.0020180000
HSPB8	45	0.6725000000	0.0020850000
SPATA21	46	0.6622000000	0.0020990000
CRYBB1	47	0.6869000000	0.0022240000
WBP2	48	0.6828000000	0.0022580000
LATS1	49	0.6894000000	0.0023270000
SLC8A1	50	0.6757000000	0.0023280000