

Molecular characterisation of putative WNT signalling protein,
Leucine Zipper and ICAT domain containing (LZIC)

Thesis submitted for degree of

Doctor of Philosophy

at University of Leicester

by

George Ludwig Skalka BSc MRes

MRC Toxicology Unit

September 2019

Molecular characterisation of putative WNT signalling protein Leucine zipper and ICAT domain containing (LZIC) – George Ludwig Skalka

The regulatory events which control cell division are referred to as the cell cycle. The cell cycle is arrested in non-permissive conditions by cell cycle checkpoints. The functioning of these checkpoints in response to stimuli prevents incorporation of mutations and acquisition of aneuploidy, while failure is a promoter of oncogenesis and hyper proliferative disorders. Interestingly, both the development and the treatment of cancer are dependent on the modulation of these checkpoints. Therefore, identification of proteins which regulate these cell cycle checkpoints can both provide mechanistic insight and treatment targets for cancer.

Ionising radiation (IR) is a well-established cancer therapy and a potent activator of the cell cycle checkpoints. However, the mechanisms governing cellular response to IR are under investigated compared to many small molecular inhibitors and chemotherapeutic agents. The Leucine zipper and ICAT containing (LZIC) protein is poorly characterised but has been implicated in the development of IR induced tumorigenesis. Interactome analysis of LZIC highlights an enrichment for spliceosome components. In parallel, I show that the transcriptional response of LZIC knock-out cells to IR is altered, with emphasis on MYC signalling and G2/M checkpoint. Analysis of the cell cycle checkpoint activation by flow cytometry and western blot indicate an early release phenotype from the late G2/M checkpoint, with partial recovery of the phenotype being observed following treatment with protein phosphatase inhibitor. In addition, quantification of chromosome number in LZIC KO cell lines shows an increased aneuploidic state. Survival analysis for multiple human cancers shows decreased prognosis of patients with reduced LZIC expression.

My findings suggest that LZIC is a new component of the cell cycle regulatory machinery with potential usage as a biomarker for IR cancer therapy sensitivity.

Acknowledgements

Firstly, I would like to thank my supervisor, Dr Michal Malewicz, for the opportunity to undertake a PhD in his lab and all the guidance he gave during the project. I would also like to thank my co-supervisor, Professor Martin Bushell, for all the advice and the opportunity to continue my learning in his lab for the last portion of my PhD.

I would like to extend my gratitude to the other members of the Malewicz lab: Dr Andrew Craxton, Dr Deeksha Munnar, and Dr Joanna Somers for all their help with establishing experimental procedures and analysing data. I would like to thank the MRC Toxicology unit proteomics department: Professor Kelvin Cain, Dr Claudia Langlais, and Rebekah Jukes-Jones for helping me to generate an interactome for LZIC. A special thank you also goes to Rebekah for always being a friend to talk to during my time and for filling most days with obscure quotes. I am also very grateful to Dr Joanna Fox, who provided advice and aid during a brief foray into caspase cleavage.

I would like to express my external gratitude to my partner, Holly (Dr Holly Hall), for being my scientific partner, my copy editor, and my keeper. In particular, thank you for constantly supporting me when I needed it and for listening to me endlessly talk about my project (and nothing else). I must also thank my friends, especially, Dr Julian Wang, Dr Callum Rakhit, and Dr Ben Hawley who always made life interesting and without whom I would never have visited the Arnold Schwarzenegger museum.

Finally, I would like to extend my utmost gratitude to my family: Sara, Marek, William, and Edward, without their support I am not sure I would have managed to write this thesis. I would also like to thank them for all the support they have given me over the years and think that they can take some of the credit for the work that has been done.

Table of Contents

Molecular characterisation of putative WNT signalling protein Leucine zipper and ICAT domain containing (LZIC) – George Ludwig Skalka	2
Acknowledgements	3
List of Figures.....	9
List of Tables.....	12
List of Abbreviations	13
<i>Chapter 1.....</i>	<i>20</i>
<i>Introduction.....</i>	<i>20</i>
1.1 Background.....	21
1.2 Organisation of DNA by Histones	21
1.3 Chromosome structure and model for replication.....	23
1.4 Cell cycle regulation.....	23
1.4.1 G1/ S-phase progression	27
1.4.2 S-phase to G2 phase progression	29
1.4.3 G2/M phase	32
1.4.4 Mitosis.....	34
1.4.5 Mitotic exit	37
1.5 DNA Damage	38
1.5.1 Ionising radiation and radiomimetic drugs – cause or cure?	39
1.5.2 Double-stranded DNA damage and recognition	40
1.5.3 Non-homologous end-joining.....	44
1.5.4 Homologous Recombination Repair	45
1.6 Cell cycle and the DDR	48
1.6.1 G1/ S-phase checkpoint	49
1.6.2 S-phase checkpoint.....	51
1.6.3 G2/M checkpoint.....	51

1.6.4 Spindle Assembly checkpoint	53
1.7 Checkpoint activation recovery	53
1.8 WNT Signalling pathway	54
1.9 Clustered regularly interspaced short palindromic repeats (CRISPR)	58
1.10 Hypothesis and Objectives.....	59
Chapter 2.....	60
Methods.....	60
2.1 Conservation analysis of LZIC sequences and phylogenetic tree generation	61
2.2 Predictive structure generation	61
2.3 Biogrid and String network analysis	61
2.4 Bacterial culture and transformation	61
2.5 Expression and Purification of GST tagged LZIC proteins	62
2.6 Cell culture.....	63
2.7 LZIC Knock-out Line Generation.....	63
2.8 Cloning of Wild-type, N-terminal-Flag, and C-terminal-Flag tagged LZIC pBABE- Puromycin expression and viral production plasmid	63
2.9 Retroviral reintroduction of LZIC-Flag into CRISPR line.....	64
2.10 Plasmid and DNA sequencing.....	65
2.11 Hypotonic fraction isolation and high salt treatment.....	65
2.12 Mass spectrometry analysis.....	66
2.13 Microarray analysis of LZIC KO cells	66
2.14 Microarray bioinformatics analysis.....	67
2.15 Gene set enrichment analysis.....	67
2.16 Quantitative polymerase chain reaction	67
2.17 Cell Cycle profile analysis	68
2.18 Immunofluorescence	69
2.19 Western Blot	69

2.20 WST-1 Assay	70
2.21 Metaphase spread analysis	70
2.22 Kaplan Meier plot generation	71
Chapter 3.....	72
<i>Investigation of Leucine zipper and ICAT containing protein (LZIC) interactome ..</i>	72
3.1 Introduction	73
3.1.1 Leucine zipper and ICAT containing (LZIC)	73
1.9 Armadillo repeat proteins	73
3.1.2 Protein interactome analysis.....	74
3.1.3 Hypothesis and chapter aims	75
3.2 Results.....	76
3.2.1 Evolutionary conservation of LZIC sequence.....	76
3.2.2 LZIC spliceforms comparison and domain characterisation	78
3.2.3 Evolutionary conservation of LZIC functional domains.....	81
3.2.4 Meta-analysis of LZIC Biogrid interactors.....	83
3.2.5 Assessment of recombinant LZIC-GST protein quality.....	85
3.2.6 Affinity purification of LZIC-GST	88
3.2.7 Determination of LZIC interactors by yeast-2-hybrid.....	90
3.2.8 Generation of LZIC knock-out HEK293 cell lines by CRISPR.....	93
3.2.9 CRISPR cut site sequencing	95
3.2.10 Optimisation of mammalian expression system for recombinant LZIC.....	97
3.2.11 Generation of LZIC KO HEK293 with stably expressed recombinant LZIC-c- flag.....	99
3.2.12 Data normalisation and technical quality control of LZIC-c-flag MS/MS spectral counts.	101
3.2.13 Repeat normalisation and quality control of LZIC-c-flag MS/MS spectral counts following removal of aberrant sample.....	103

3.2.14 Identification of significant LZIC-c-flag interactors	105
3.2.15 Gene ontology term enrichment of LZIC-c-flag interactors.....	108
3.2.16 Protein domain term enrichment of LZIC-c-flag interactors.....	110
3.3 Discussion	112
Chapter 4.....	115
<i>Impact of LZIC loss on basal transcriptome and transcriptomic response to ionising radiation.....</i>	<i>115</i>
4.1 Introduction	116
4.1.1 WNT signalling is linked to vital transcriptional regulatory cascades	116
4.1.2 Transcriptional response to gamma radiation exposure	116
4.1.3 Methods of transcriptome analysis	117
4.1.4 Hypothesis and chapter aims.....	118
4.2.1 Analysis of IR responsive gene expression profile over 24hr time course.....	119
4.2.2 Comparison of CRISPR control cells to LZIC KO clones for quality control....	121
4.2.3 Variable analysis	124
4.2.4 Differential gene expression analysis of normalised gene probe intensity from all experimental conditions.....	126
4.2.5 Validation of differential gene expression analysis by qPCR	129
4.2.6 Gene ontology analysis of IR dependent and IR independent LZIC regulated transcripts	131
4.2.7 Z-score analysis of most variable genes following loss of LZIC expression	133
4.2.8 Gene set enrichment analysis of CRISPR control and LZIC KO gene expression profiles.....	136
4.3 Discussion	138
Chapter 5.....	141
<i>The Role of LZIC in Regulation of Cell Cycle Following Genotoxic Stress</i>	<i>141</i>
5.1 Introduction	142

5.1.1 Activation and reversion mechanisms for the G2/M checkpoint.....	142
5.1.2 Checkpoint adaptation and genome instability.....	142
5.1.3 Hypothesis and aims.....	143
5.2 Results.....	144
5.2.1 Cell cycle analysis of G2/M checkpoint in response to IR exposure	144
5.2.2 Cell Cycle analysis of LZIC KO cells following treatment with cellular stressors	146
5.2.3 Cell cycle analysis of LZIC KO cells following reintroduction of LZIC-c-flag...	148
5.2.4 Quantification of mitotic marker histone 3 serine 10 phosphorylation following exposure to IR	150
5.2.5 Analysis of Early G2/M checkpoint activation in response to IR exposure	152
5.2.6 p53 phosphorylation status following IR exposure	154
5.2.7 Phosphorylation status and total protein abundance of master checkpoint regulators CHK1 and CHK2 following exposure to IR.	156
5.2.8 Analysis of PIKK protein activation in LZIC KO cells following IR exposure ..	158
5.2.9 Analysis of CDC2 and Cyclin B1 phosphorylation and total protein abundance changes in response to IR exposure.....	160
5.2.10 PP1 and PP2A total protein abundance following IR exposure.....	162
5.2.11 Cell cycle profile of cells following IR exposure and PP2A inhibition.	164
5.2.12 Hypothesised LZIC position in G2/M control pathway.....	166
5.2.13 Development of defective G2/M checkpoint induced aneuploidy	168
5.2.14 Metaphase spread analysis of LZIC KO cells karyotype basally and following exposure to IR	170
5.2.15 Viability impact of exposure to IR following loss of LZIC protein expression..	172
5.2.16 Impact of LZIC expression on prognosis of renal clear cell carcinoma and neuroendocrine tumours.....	174
5.3 Discussion	177
<i>Chapter 6</i>	<i>181</i>

<i>Implications and Future Work</i>	<i>181</i>
6.1 Implications and future work.....	182
<i>Bibliography</i>	<i>185</i>
<i>Appendix</i>	<i>234</i>

List of Figures

Figure 1.1 Histone structure.....	22
Figure 1.2. CDC25 family member specificity.....	26
Figure 1.3. CDKi families and substrate specificity	27
Figure 1.4. G1 to S-phase transition cascade	29
Figure 1.5. S-phase transition	31
Figure 1.6. S-phase to G2 transition	33
Figure 1.7. G2/M transition.....	36
Figure 1.8. DNA damage sources.....	38
Figure 1.9. Double-strand break response cascade.....	41
Figure 1.10. NHEJ Complex	45
Figure 1.11. HR pathway.....	47
Figure 1.12. Cell cycle progression	48
Figure 1.13. G1 Checkpoint	50
Figure 1.14. G2/M Checkpoint.....	52
Figure 1.15. Canonical WNT signalling	56

Figure 3.1 Evolutionary conservation of LZIC sequence.....	77
Figure 3.2 LZIC spliceforms comparison	80
Figure 3.3 Evolutionary conservation of LZIC protein domains.	82
Figure 3.4 Network analysis of previously identified LZIC interactors.	84
Figure 3.5 Assessment of recombinant LZIC-GST protein quality	87
Figure 3.6 LZIC-GST protein interactor visualization	89
Figure 3.7 LZIC interactome determination by Yeast-2-hybrid	92
Figure 3.8 Establishment of LZIC knock-out line by CRISPR.	94
Figure 3.9 Impact of CRISPR mediated cut on LZIC mRNA and predicted protein expression.....	96
Figure 3.10 Expression of recombinant LZIC in LZIC KO HEK293 cells.	98
Figure 3.11 Comparison of recombinant LZIC expression to WT expression in parental and CRISPR control.....	100
Figure 3.12 Data normalization and quality control checks of LZIC-c-flag MS/MS spectral counts	102
Figure 3.13 Quality control checks for LZIC-c-flag MS/MS analysis post removal of aberrant experiment	104
Figure 3.14 significantly identified LZIC-c-flag interactors.	107
Figure 3.15 Gene ontology enrichment for all significantly identified LZIC interactors.	109
Figure 3.16 Domain enrichment analysis for all significantly identified LZIC-c-flag interactors.....	111
Figure 4.1 Meta-analysis of temporal gene expression changes in response to IR.....	120
Figure 4.2 RNA quality analysis and microarray background normalisation.	123
Figure 4.3. principle component analysis of microarray probe intensities from all experimental conditions.....	125

Figure 4.4 Differential expression analysis of LZIC KO Clone 1 vs CRISPR control.	128
Figure 4.5 qPCR validation of significantly identified genes by differential expression analysis.	130
Figure 4.6 Gene ontology analysis of IR independent and IR dependent gene groups. ...	132
Figure 4.7 Z-score analysis of most variable genes due to loss of LZIC.	135
Figure 4.8 Gene set enrichment analysis of CRISPR control and LZIC KO gene expression profiles.	137
Figure 5.1 Analysis of cell cycle profile in LZIC KO cell lines exposure to IR.	145
Figure 5.2 Analysis of cell cycle in LZIC KO cell lines following cellular stress and genotoxic stress.	147
Figure 5.3 Cell cycle analysis following reintroduction of LZIC-flag in KO cell lines.	149
Figure 5.4 Analysis of pS10 H3 immunofluorescence in LZIC KO clones following 2Gy ionising radiation.	151
Figure 5.5 Early G2/M checkpoint response to IR exposure.	153
Figure 5.6 Analysis of p53 serine 15 phosphorylation status and p53 steady-state levels.	155
Figure 5.7 Phosphorylation status and total protein abundance of master checkpoint regulators CHK1 and CHK2 following exposure to IR.	157
Figure 5.8 Analysis of ATM and ATR activation in LZIC KO cells following IR treatment.	159
Figure 5.9 Analysis of phosphorylation and steady state levels of Cyclin B1 and CDC2 following exposure to IR.	161
Figure 5.10 Expression levels of protein phosphatase 2A subunits and protein phosphatase 1 following IR exposure.	163
Figure 5.11 Cell cycle distribution analysis following treatment with OA and IR.	165
Figure 5.12 Schematic of LZIC function within the G2/M checkpoint signalling cascade.	167

Figure 5.13 Schematic for the development of aneuploidy following the disruption of G2/M checkpoint.....	169
Figure 5.14 Karyotyping of CRISPR control and LZIC KO cell lines basally and following exposure to IR.....	171
Figure 5.15 Viability impact of exposure to IR following loss of LZIC protein expression.	173
Figure 5.16 Overall survival correlates with LZIC expression for renal clear cell carcinoma and neuroendocrine tumours.	176

List of Tables

Table 2.1. LZIC cloning primers for pBABE insertion	64
Table 2.2 qPCR primers	68

List of Abbreviations

DNA – Deoxyribonucleic acid

RNA – ribonucleic acid

H2A – Histone 2A

H2B – Histone 2B

H3 – Histone 3

H4 – Histone 4

H1 – Histone 1

G0 – Quiescence

G1 – Growth phase 1

S-phase – Synthesis phase

G2 – Growth phase 2

M-phase – Mitosis

CDK – Cyclin-dependent kinase

CDKi – CDK inhibitor

CAK – CDK activating kinase

CDC25 – Cell division cycle 25

GSK3 β – Glycogen Synthase kinase 3 beta

RB – retinoblastoma

PCNA – Proliferating cell nuclear antigen

MCM7 – Minichromosome complex component 7

MPF - Mitosis promoting factor

NEB - Nuclear envelope breakdown

NF-Y – Nuclear transcription factor Y

Thr14 – Threonine 14

Tyr15 – Tyrosine 15

PLK1 – Polo-like kinase 1

SMC – Structural maintenance of chromosomes

MTOC – Microtubule organisation centre

CENP-A – Centromere protein A

CENP-B – Centromere protein B

PP2A – Protein phosphatase 2A

APC/C – Anaphase-promoting complex/cyclosome (APC/C)

CDH1 – CDC20 homolog 1

PP1 – protein phosphatase 1

DSB – DNA double-strand break

INDELs – insertions and deletions

NHEJ – non-homologous end joining

HR – Homologous recombination

DNA-PKcs – DNA-protein kinase catalytic subunit

MRE11 – Meiotic recombination 11

NBS1 – Nibrin

ATM – Ataxia telangiectasia mutated

PIKK – Phosphatidylinositol 3-kinase related

MDC1 – Mediator of DNA damage checkpoint protein 1

RNF8 – Ring finger protein 8

RNF168 – Ring finger protein 168

HUWE1 - HECT, UBA And WWE Domain Containing E3 Ubiquitin Protein Ligase 1

53BP1 – p53 binding protein 1

BRCA1 – Breast and ovarian cancer susceptibility protein 1

BRCT – breast cancer susceptibility

RIF1 – Rap1 interacting factor

CtIP – C-terminal binding protein 1 interacting protein

XRCC4 – X-ray cross complementing protein 4

PAXX – Paralog of XLF and XRCC4

XLF – XRCC4-like factor

RPA – Replication protein A

BRCA2 – Breast and ovarian cancer susceptibility protein 2

ATR – Ataxia telangiectasia and Rad3 related

ATRIP – ATR-interacting protein

G2/M – G2 phase/mitosis

SAC – Spindle assembly checkpoint

CHK1 – Checkpoint protein 1

CHK2 – Checkpoint protein 2

Thr68 – Threonine 68

MDM2 – Mouse double minute 2 homolog

HU – hydroxyurea

S345 – Serine 345

S317 – Serine 317

IR – Ionising radiation

MCC- Mitotic checkpoint complex

WIP1 – Wild-type p53 induced phosphatase 1

WNT – Wingless and INT-1

Fz – Frizzled

LRP5/6 – Lipoprotein receptor-related protein 5/6

Dvl – Dishevelled

PKC – Protein kinase C

NFAT – Nuclear factor associated with T-cells

KAP3 – Kinesin-associated protein 3

ROS – Reactive oxygen species

siRNA – silencing RNA

CRISPR – clustered regularly interspaced short palindromic repeats

KO – knock-out

PAM – Protospacer adjacent motif

HPV – Human papillomavirus

NCBI – national centre for biotechnology information

GO – gene ontology

LB – Luria bertani

IPTG – Isopropyl B-D-1 thiogalatopyranoside

DMEM – Dulbecco’s modified Eagle’s medium

HEK293 – Human embryonic kidney

HDMS_e – Data-independent acquisition and ion mobility

MMC – Mitomycin C

MMS – Methyl Methanosulfonate

CC – Cobalt chloride

UV – Ultra-violet

LZIC – Leucine zipper and ICAT domain containing

MS – Mass spectrometry

HPLC – High pressure liquid chromatography

SILAC – Stable isotope labelling with amino acids in cell culture

LFQ – Label free quantification

MS/MS – tandem MS

PLGS – Protein lynx global server

LAH – Long alpha hairpin

ZNF521 – Zinc finger 521

ZNF558 – Zinc finger 558

ZNF3 – Zinc finger 3

CHIP – Chromatin immunoprecipitation

BACH1 – BTB and CNC homology

CARM1 – Histone arginine methyl transferase

HSPA5 – Heat shock protein family A member 5

ZNF197 – Zinc finger 197

HIF1 α – Hypoxia inducible factor 1 α

pVHL – von-Hippel-Lindau tumour suppressor

gRNA – guide RNA

LZIC-c-flag – C-terminus tagged LZIC sequence

LZIC-n-flag – N-terminus tagged LZIC sequence

LZIC-WT – Wild-type LZIC sequence

PCA – Principal component analysis

FACT – Facilitates chromatin transcription

DDX47 – Dead box helicase 47

SRSF9 – Serine and arginine rich splicing factor

MCM4 – Maintenance complex component 4

IQSEC2 – IQ motif and Sec7 domain 2

mRNA – Messenger RNA

Pol II – RNA polymerase II

RIN – RNA integrity number

SOX11 – SRY box 11

Lnc-CTNNA2 – long-non-coding-catenin-alpha 2

PCDH7 – Protocadherin 7

HOXB6 – Homeobox B6

HOXC4 – Homeobox C4

EPCAM – Epithelial cell adhesion molecules

FHL2 – four-and-a-half LIM domains protein 2

SFN – Stratifin

CCNB1 – Cyclin B1

CALCA – Calcitonin related polypeptide alpha

JRK – JRK helix-turn-helix protein

lncRNA – long-non-coding RNA

MEG3 – maternally expressed 3

DHRS2 – Dehydrogenase/reductase 2

ATF3 – Activating transcription factor 3

CREB – cAMP responsive element binding family

MASTL - Microtubule Associated Serine/Threonine Kinase Like

LOH – Loss of heterozygosity

PI – Propidium Iodide

pS10 H3 – Phosphorylation serine 10 histone 3

OA – Okadaic acid

Chapter 1

Introduction

1.1 Background

Deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are two molecules that store information for protein production. Both consist of nucleic acids, referred to as bases, bound together in string structures. The monomeric bases adenine (A), thymine (T), guanine (G), and cytosine (C) are segregated into two major families, purines and pyrimidines. The purine group comprises adenine and guanine, and the pyrimidine group includes thymine and cytosine. The bases can interact through hydrogen bonds to form base pairs. The biochemical restraints of the interaction between bases of two sequence strands forces DNA to predominantly form a double stranded antisense helix, which produces the classical structure of DNA suggested by Watson and Crick in 1953 (Watson and Crick, 1953). Typically, the bases preferentially interact intragroup, however, base pairing irrespective of group is possible under certain conditions (Kyogoku, Lord and Rich, 1966). RNA is also constituted from sequences of these nucleic acids, however, T is exchanged for uracil (U) and is typically single stranded.

Eukaryotes, prokaryotes, and a sub-set of viruses utilise DNA for genetic storage, however, the size of a genome is significantly different between these groups. For example, the human genome project initially identified 2.9 billion base pairs, which has now been revised to 3.2 billion base pairs (Venter *et al.*, 2001). In comparison, the *Escherichia coli* genome constitutes 4.6 million base pairs (Blattner *et al.*, 1997). The significant disparity in genome size has yielded divergent organisational mechanisms to facilitate processes such as, replication and gene expression. In eukaryotes a specialised family of proteins, termed histones, are utilised to organise DNA into more compact structures.

1.2 Organisation of DNA by Histones

The core eukaryotic histone is comprised of 4 subunits, present in duplicate, referred to as histone 2A (H2A), histone 2B (H2B), histone 3 (H3), and histone 4 (H4). Each histone core is surrounded by 146 bp of DNA, with a distribution of a histone complex every 200 bp throughout the genome (Luger *et al.*, 1997). This structure of histone core distribution prevents the direct interaction between individual histone cores, which is required to form the higher-order structure of DNA, referred to as the 30 nm fibre (Robinson *et al.*, 2006).

Therefore, the linker protein histone 1 (H1) is required to mediate interaction between histone clusters. It was originally postulated that bacterial genomes were structured by histone-like proteins (Pettijohn, 1988) (Figure 1.1). However, it was later discovered that bacterial genomes are compacted by the process of supercoiling instead of through histone proteins. The genome is predominantly negatively supercoiled with the specific deviation from this state can significantly impact transcription and replication (Lal *et al.*, 2016).

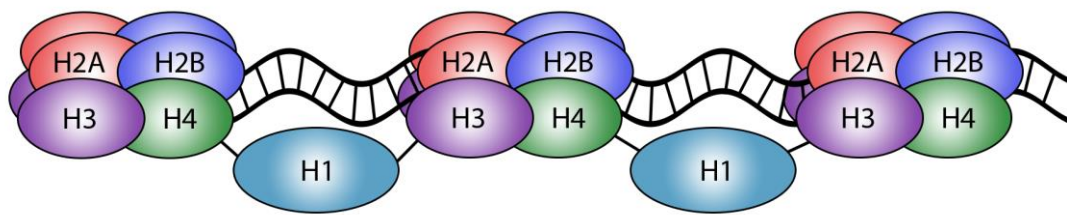


Figure 1.1 Histone structure.

Core histone structure with two copies of each histone subunit: H2A, H2B, H3, and H4 forming an overall octameric structure. Each histone core is connected through the linker histone, H1.

The DNA in a eukaryotic cell is contained within a specialised lipid membrane-bound organelle – the nucleus. Within this subcellular compartment, the entirety of DNA is not persistent in a completely condensed form. Instead, the DNA is split into two major states: heterochromatin or euchromatin (Mello, 1983). Heterochromatin defines DNA containing tightly compacted histones which can restrict gene expression from these regions (Allshire and Madhani, 2018). In comparison, euchromatin represents less compacted DNA regions thus facilitating transcription. The process of cellular differentiation is a vital step during embryonic development, with each lineage being defined by specific gene expression programs. The transition of chromatin surrounding regulatory gene regions between heterochromatin and euchromatin is vital to assure correct gene expression profiles in differentiating cells (Ueda *et al.*, 2014). The transition of DNA between heterochromatin and euchromatin is controlled by the post-translational modification of c-terminal regions, referred to as “histone tails”. An example of a “histone tail” modification which is particularly associated with heterochromatin is the tri-methylation of H3 (Litt *et al.*, 2001; Gessaman and Selker, 2017; Janssen, Sidoli and Garcia, 2017). In contrast, a characteristic mark of euchromatin is the hyper-acetylation

of histone tails which alters the structure increasing co-factor accessibility of nucleosomes (Görisch *et al.*, 2005). In addition to regulation of histone state by post-translational modification of histone tails, an additional layer of control is conferred by incorporation of alternative histone variants into the core complex (Henikoff and Smith, 2015). These subunits can have vital roles for fundamental processes such as replication and response to DNA damage. For example, the H2A variant H2A.z is incorporated into histone cores and acetylated to facilitate DNA replication and without it genome instability increases (Hardy *et al.*, 2009; Kim *et al.*, 2009).

1.3 Chromosome structure and model for replication

While on a microscale sub-domains of DNA are organised by histones, on a macroscale the 3.2 billion base pairs constituting the human genome are not stored in one continuous stretch of DNA. Instead, the genetic information is distributed across 23 asymmetrical units, referred to as chromosomes (Tjio and Levan, 1956). Each time a human cell divides both daughter cells must subsequently contain 23 chromosomes, therefore the number of chromosomes must first double to 46. The original model of semi-conservative chromosomal replication was proposed by Watson and Crick in their seminal paper on DNA structure (Watson and Crick, 1953). This hypothesis was corroborated in 1958 by Meselson and Stahl who demonstrated this mechanism of DNA replication through the labelling of DNA by different isotopes of nitrogen (Meselson and Stahl, 1958). The experiments demonstrated that each strand of starting DNA forms the template for new strand synthesis, thereby semi-conservative as half of the genetic information from the parent cell is maintained.

The process of DNA regulation requires intricate cellular control to prevent the incorporation of errors and avoid division of cells before this process is complete. Therefore, the progression of cells through this event and those required for division occurs in a specific order termed the cell cycle.

1.4 Cell cycle regulation

The cell cycle is split into 5 major stages: quiescence (G0), growth phase 1 (G1), DNA synthesis phase (S-phase), growth phase 2 (G2) and mitosis (M-phase). The transitions

between each of these cell cycle phases are carefully regulated and dependent upon the conserved activity of multiple signalling pathways. The convergence point for these signalling pathways are proteins of the cyclin-dependent kinase family (CDK) which are master regulators of cell cycle phase transitions.

The proteins of the CDK family are highly conserved, however, the number of CDK family members varies dependent upon species. The number of CDK proteins has increased throughout evolution, with 8 being identified in *Saccharomyces cerevisiae* and 22 being found in *Homo sapiens* (Liu and Kipreos, 2000). All of the identified CDK proteins share a similar structure, which is constituted from 2 lobes, referred to as the N and C lobes, with a substrate-binding cleft between. In addition to the main structural domains, there are two regulatory regions consisting of a PSTAIRE domain and a T-loop. The PSTAIRE domain acts as a cyclin binding site, and the T-loop structure can be phosphorylated to facilitate complete activation of the CDK (Pavletich, 1999).

The quintessential CDK binding partners are members of the cyclin family, with specific CDK/cyclin complexes controlling each of the five cell cycle phase transitions (Sherr, 1993; Malumbres *et al.*, 2009). Upon binding of cyclin protein to the respective CDK through the PSTAIRE domain, the substrate-binding cleft opens and the T-loop twists exposing the substrate binding cleft (Andzelm, Lew and Taylor, 1995; Jeffrey, Tong and Pavletich, 2000). This structural alteration increases substrate binding capability of the complex, however, without further regulatory steps full activation of the CDK/cyclin complex is not achieved (De Bondt *et al.*, 1993). A further three steps are required for activation of CDK protein kinase activity: phosphorylation of a conserved threonine residue within the T-loop, removal of inhibitory phosphorylation, and proteolytic degradation of CDK inhibitor (CDKi) proteins.

The CDK T-loop domain is a site of regulatory phosphorylation which controls activation of overall CDK kinase activity. For example, the phosphorylation of the CDK2 T-loop induces domain shift exposing the substrate cleft and increasing contact points between CDK2 and cyclin A (Russo, Jeffrey and Pavletich, 1996). A CDK activating Kinase (CAK) is responsible for phosphorylating the T-loop of CDK proteins. In humans a single CAK has been identified, CDK7, and is responsible for activation of all cell cycle CDK proteins (Mäkelä *et al.*, 1994). As with other CDK proteins, CDK7 requires binding of

its partner, cyclin H for full activation (Patel and Simon, 2010). However, in contrast to other members of the CDK family the stability and kinase activity of the complex requires additional interaction with the protein, MAT1 (Devault *et al.*, 1995). Additionally, in contrast to other CDK/cyclin complexes, the expression of CDK7/cyclin H is not linked to specific cell cycle phases (Fisher and Morgan, 1994).

In conjunction with CAK dependent phosphorylation of T-loop residues, the removal of additional inhibitory phosphorylation is also required for activation of CDK/cyclin complexes. This process is reliant on the activity of the cell division cycle 25 (CDC25) phosphatase family. The CDC25 family is constituted of three members: CDC25A, CDC25B, and CDC25C (Aressy and Ducommun, 2012; Sur and Agrawal, 2016). The role of these proteins is to remove specific inhibitory phosphate groups from tyrosine and threonine residues of CDK proteins increasing kinase activity (Rudolph, 2007). CDC25A regulates both the progression of cells through G1/S-phase and G2/M by specifically removing the inhibitory phosphorylation from tyrosine 17 and tyrosine 24 of CDK4 and CDK6, respectively, and tyrosine 15 and threonine 14 of CDK2 and CDK1 (Shen and Huang, 2012). In contrast, CDC25B and CDC25C are only active on CDK/cyclin complexes regulating the G2/M transition and remove phosphorylation from tyrosine 15 and threonine 14 (Mitra and Enders, 2004) (Figure 1.2). Each member of the CDC25 family has a specific set of regulatory sites which restricts activity to specific CDK proteins. For example, hypo-phosphorylated CDC25C has low substrate specificity for CDK1/cyclin B, however, after entering a hyperphosphorylated state the substrate specificity increases dramatically (Gabrielli *et al.*, 1997; Boutros, Dozier and Ducommun, 2006).

The CDKi proteins are associated with CDK/cyclin complexes and prevent activation of kinase activity. Therefore, to promote cell cycle progression CDKi proteins must first be degraded. In general, the CDKi proteins are segregated into two major families based on their evolutionary conservation: the INK4 family and the Cip/Kip family. The INK4 family is composed of 4 proteins: p15, p16, p17, and p18, which are specific for G1 to S-phase transition associated CDK proteins, CDK4 and CDK6 (Cánepa *et al.*, 2007) (Figure 1.3A). Inhibition of CDK activity occurs due to distortion of both the ATPase domain and the cyclin binding domain following binding of the INK4 family member (Jeffrey *et al.*, 1995). The degradation of these proteins is induced by phosphorylation induced

proteasome recruitment (Thullberg, Bartek and Lukas, 2000). The dynamic nature of the INK inhibitor function can be demonstrated by the shuttling of CDK4 between the HSP90-cDc37 complex and the INK bound state, which prevents cyclin association until conditions are optimal (Hallett *et al.*, 2017).

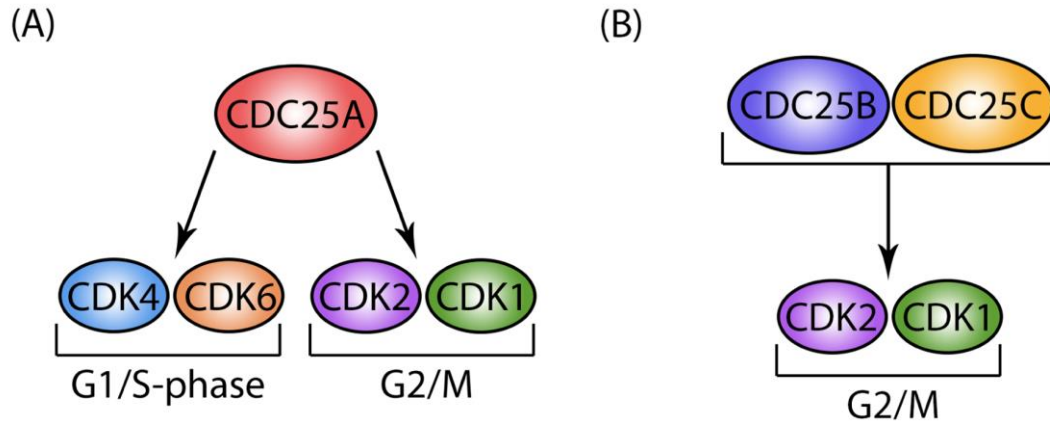


Figure 1.2. CDC25 family member specificity

(A) The phosphatase activity of CDC25A can remove inhibitory phosphorylation moieties from both G1 and G2/M transition CDK proteins. (B) The phosphatase activity of CDC25B and CDC25C is restricted to CDK proteins of the G2/M transition.

The Cip/Kip family is composed of p21, p27, and p57 proteins, and bind to both CDK and cyclin proteins through specialised amino-terminal motifs (Nakanishi *et al.*, 1995) (Figure 1.3B). As with the interaction of INK4 family members and CDK, this interaction can prevent CDK/cyclin complex activity, although the substrate specificity is wider with the potential to bind CDK/cyclin complexes from multiple cell cycle boundaries. Additionally, Cip/Kip family members have a dual role during the CDK/cyclin complex formation. For example, the interaction of p21 with CDK4/cyclin D complexes is required for complex kinase activity (Cheng, 1999; Cerqueira *et al.*, 2014). The degradation of this family of proteins is mediated through the same mechanism as INK4 (Starostina and Kipreos, 2012).

The convergence of signalling pathways on the activity of CDK/cyclin complex provides a multi-layered regulatory system which prevents cell progression until all conditions are optimal. While these events have been discussed in broad terms the following sections will investigate the individual requirements for progression through each cell cycle phase.

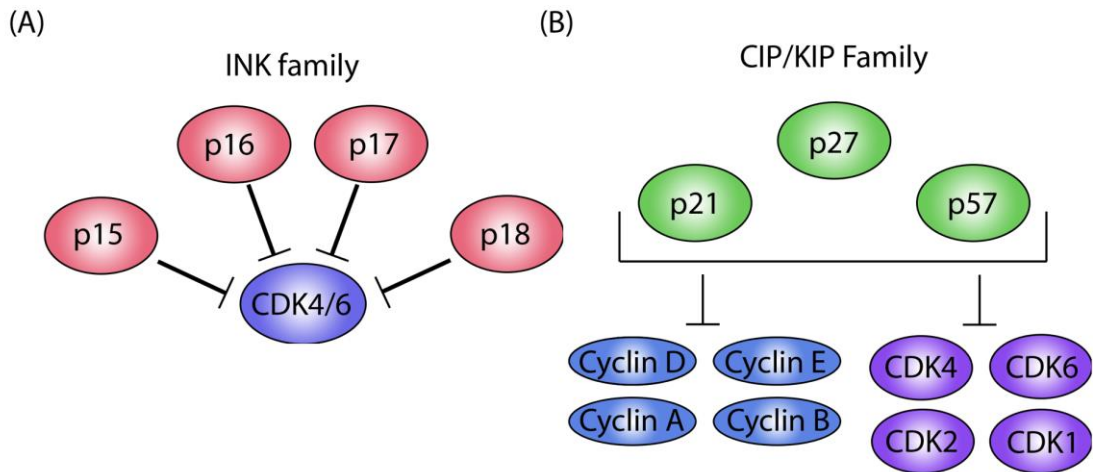


Figure 1.3. CDKi families and substrate specificity

(A) The INK family has specific substrate specificity for CDK4 and CDK6 which regulate the G1 to S-phase transition. (B) The CIP/KIP family can inhibit CDK and cyclin proteins from all cell cycle transitions.

1.4.1 G1/ S-phase progression

Each of the cell cycle stage transitions is controlled by a specific CDK/cyclin complex or a combination of these complexes. The transition from quiescence through G1 into S-phase is controlled by cyclin D/CDK4, cyclin D/CDK6, and cyclin E/CDK2 complexes (Dong *et al.*, 2018). In addition to the canonical cyclin E/CDK2 complex, cyclin D will also associate with CDK2 in multiple tissues, however, the role of this complex and the activity is widely debated with many pieces of conflicting evidence (Jirawatnotai *et al.*, 2011). Jahn *et al.* attempted to address this controversy and found that a complex involving p21/CDK2/Cyclin D can phosphorylate components of the centrosome, which have been linked to increased genome instability in cancer rather than the canonical role at the G1 phase to S-phase transition (Jahn *et al.*, 2013).

The transition phase between G1 and S-phase was coined the “restriction point” by Arthur Pardee, which refers to the mitogen signalling cascades required for progression (Pardee, 1974). In growth factor deprived conditions cyclin D1 is phosphorylated by the GSK3 β complex which targets it for ubiquitin-mediated proteasome degradation. Upon cellular detection of mitogens the transcription of cyclin D1 is upregulated (Won *et al.*, 1992; Diehl, Zindy and Sherr, 1997). In addition, activation of growth factor pathways

such as WNT signalling reduces the activity of GSK3 β , which in turn increases the stability of cyclin D1 (Diehl *et al.*, 1998). However, the increase in cyclin D mRNA and protein stability does not induce progression through the restriction point. As previously described the phosphorylation of the CDK4/CDK6 T-loop residues by CDK7/cyclin H must occur (Schachter *et al.*, 2013). In parallel, the INK family proteins undergo proteasomal degradation and the remaining inhibitory phosphorylation moieties are removed by CDC25A, facilitating release and activation of CDK4/cyclin D complex kinase activity (Blomberg and Hoffmann, 1999) (Figure 1.4).

Once the restriction point has been passed the major gatekeepers of progression into S-phase are members of the E2F transcription factor family. This family consists of 5 members, with E2F1-3 being transcriptional activators and E2F4-5 having an inhibitory activity on transcription (Attwooll, Denchi and Helin, 2004). Prior to pathway activation inhibitory E2F factors are associated with promoter regions of genes specific for S-phase transition and activator E2F factors are constitutively bound by members of the retinoblastoma protein family. The canonical family member is retinoblastoma (RB), with other family members including, p107 and p110 (Friend *et al.*, 1986; Ewen *et al.*, 1991; Hannon, Demetrick and Beach, 1993). The RB protein, in particular, binds to E2F family members through the large A/B pocket (Knudsen and Wang, 1997). This interaction can inhibit E2F mediated transcription by two methods. The first of these is to directly block the interaction of E2F and promoter regions. While the second is the recruitment of chromatin-modifying enzymes, such as histone deacetylases (HDAC), which prevent transcriptional activity at these sites (Brehm *et al.*, 1998; Trimarchi and Lees, 2002). The association of E2F family members and RB protein is controlled by the phosphorylation state of RB at 16 CDK dependent phosphorylation sites. While none of the 16 phosphorylation sites is individually responsible for the loss of interaction with E2F proteins the phosphorylation of multiple sites leads to disassociation of RB (Adams *et al.*, 1999; Brown, Phillips and Gallie, 1999; Knudsen *et al.*, 1999). Due to this the RB protein has three broad states of phosphorylation, these being, un-phosphorylated RB, partially phosphorylated RB and hyper-phosphorylated RB. The importance for RB mediated regulation of the G1/S-phase checkpoint is demonstrated by the highly pathogenic consequences of RB loss, with the development of retinoblastoma in patients with none functional protein (Lee *et al.*, 1987).

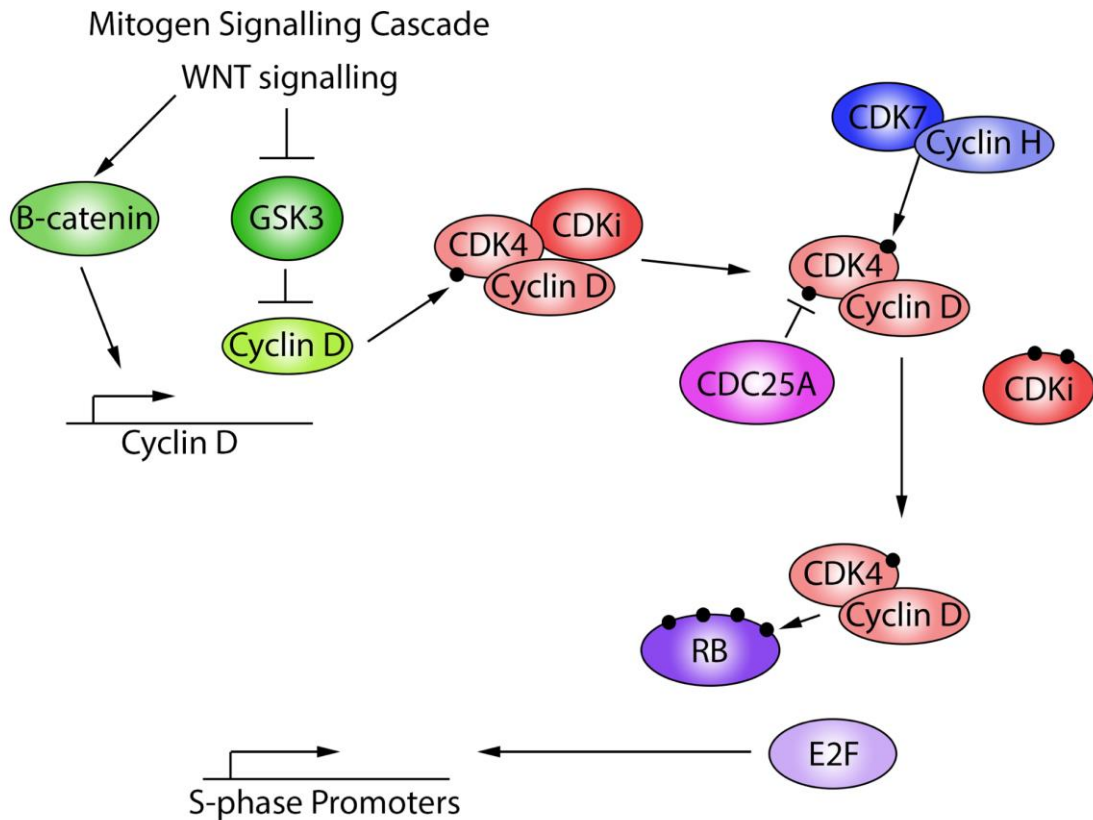


Figure 1.4. G1 to S-phase transition cascade

Detection of mitogens can activate multiple growth promoting cell surface receptor pathways, such as, WNT signalling. The WNT signalling family member β -catenin, promotes the transcription of cyclin D, which increases cyclin D protein levels. The WNT signalling pathway also inhibits activity of GSK3 β which typically induces proteolytic degradation of cyclin D. The collaboration of these mechanisms increases cyclin D protein levels which promotes formation of the cyclin D/CDK4 complex. Initially this complex is bound by CDKi proteins. The phosphorylation of CDKi proteins alters structure and induces dissociation from the cyclin D/CDK4 complex. In addition, the cyclin H/CDK7 complex phosphorylates activator residues in the T-loop structure of CDK4 and the CDC25A phosphatase removes inhibitory phosphorylation moieties. These three events increase the cyclin D/CDK4 complex activity, which can subsequently hyper-phosphorylate retinoblastoma protein causing dissociation from E2F transcription factors. The uninhibited E2F transcription factor family promote transcription of S-phase progression controlling proteins.

1.4.2 S-phase to G2 phase progression

Following correct deactivation of RB and the activation of E2F transcriptional cascades, cell cycle progresses into S-phase. During this phase of the cell cycle the genome is replicated. The initial steps of this process in humans are widely debated as the initiation of replication, referred to as origin firing, is not convincingly shown to be either random

or region-specific (Kaykov and Nurse, 2015). Despite this debate, it is well established that the protein complex associated with replication origins is referred to as the replisome. The replisome is constituted of proteins which facilitate replication of DNA, which includes DNA polymerase III, RNase H, primase, PCNA, topoisomerase and multiple addition co-factors and complexes which regulate the activity of these factors (Yao and O'Donnell, 2010; Gao *et al.*, 2019). The activity of this complex is dependent upon cyclin A/CDK2 providing a link directly to the master regulators of cell cycle progression.

To drive expression of cyclin A, following the entry into S-phase, E2F1 specifically associates with E2F binding domains within the cyclin E promoter increasing transcription of cyclin E mRNA and as a result protein levels (Ohtani, Degregori and Nevins, 1995; Takahashi, Rayman and Dynlacht, 2000). Cyclin E forms the CDK2/cyclin E complex and induces further activation of the E2F signalling cascade which subsequently increases expression of cyclin A (Zerfass-Thome *et al.*, 1997). The increased expression of cyclin A drives formation of the Cyclin A/CDK2 complex which interacts with components of the replisome, such as minichromosome maintenance complex component 7 (MCM7), to promote S-phase progression (Chibazakura *et al.*, 2011) (Figure 1.5). The importance of this complex for correct genome replication has been demonstrated by RNAi studies which show a reduction in replication following knockdown of cyclin A (Girard *et al.*, 1991). However, a comprehensive spectrum of cyclin A/CDK2 targets has not been identified and therefore more work is required in this area of research. In addition to a cyclin A/CDK2 complex, a cyclin A/CDK1 complex also forms which is required for origin firing during DNA replication, a loss of this protein complex resulted in prolonged S-phase and incomplete replication firing (Katsuno *et al.*, 2009). The major role of cyclin A/CDK1 complex is to drive progression through G2 phase and is required for progression from S-phase to mitosis as loss leads to slowing of G2 phase progression (Mitra and Enders, 2004).

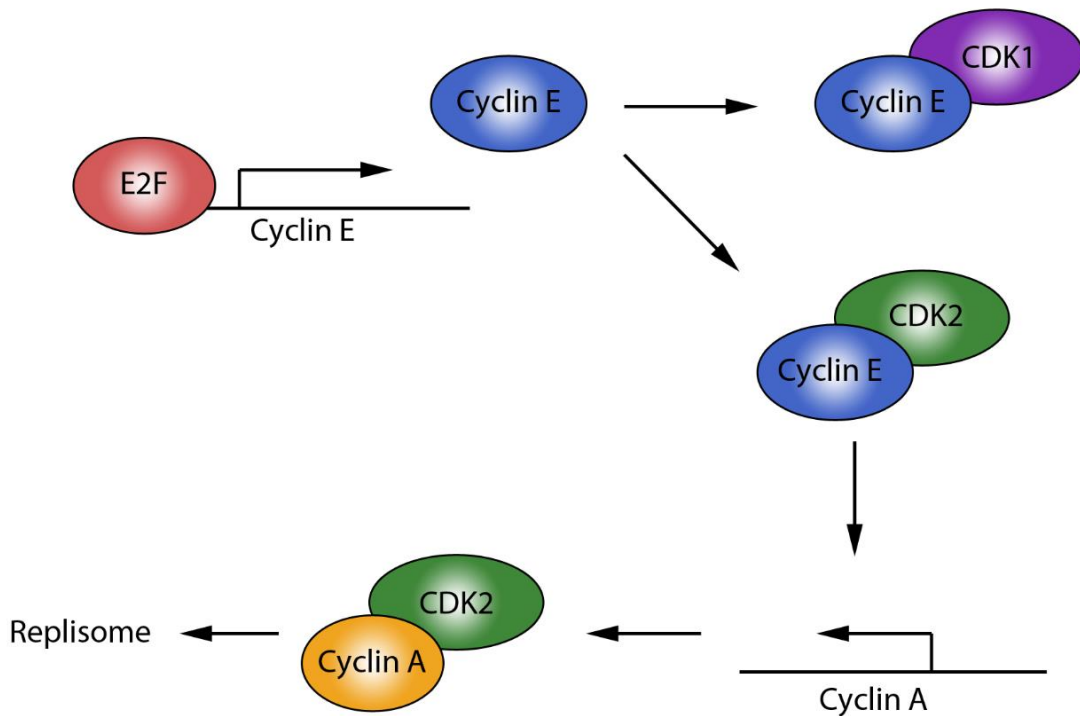


Figure 1.5. S-phase transition

The activation of E2F transcription factor family members induces transcription of multiple genes which promote the progression from G1 to S-phase, including cyclin E. The increased transcription of cyclin E also increases the protein levels which facilitates formation of two complexes: cyclin E/CDK1 and cyclin E/CDK2. The cyclin E/CDK2 complex promotes the transcription and protein stabilisation of cyclin A. The increase of cyclin A protein yields formation of cyclin A/CDK2 complex, which phosphorylates components of the replisome and promotes genome replication.

While the processes which have been outlined so far are centred on kinase activation cascades, in parallel large-scale structural changes occur within the cell to permit mitosis. The progression of cells through S-phase also alters the dynamics of cellular cytoskeleton in preparation for chromosomal segregation during mitosis. The centrosome, which is formed from two centrioles, form the nucleation points for cytoskeletal proteins, such as actin (Farina *et al.*, 2016). The centrioles are formed from nine triplet microtubules arranged in a cylindrical structure (Winey and O'Toole, 2014). The space surrounding the centrioles is then dominated by proteins which constitute the pericentriolar material (Woodruff, Wueseke and Hyman, 2014). During S-phase, the centrioles begin to duplicate forming two centromeres (Holland, Lan and Cleveland, 2010). These structures continue to elongate and mature until entry into mitosis (Tsou and Stearns, 2006).

1.4.3 G2/M phase

Following replication of the genome a cascade of events which culminates in cell division is initiated. This is carefully regulated during two cell cycle phases referred to as G2 and mitosis. The G2 phase of the cell cycle occurs immediately prior to the induction of mitosis. Mitosis is split into 5 stages: prophase, metaphase, anaphase, telophase, and cytokinesis. These two stages of the cell cycle share a significant overlap of signalling cascade activation which overall facilitates the progression of five major events which are required for final cell division: increase of mitosis promoting factor (MPF) expression, activation of MPF, chromatin condensation, nuclear envelope breakdown (NEB), and migration of centromeres.

One of the major events of G2 and prophase is increasing cyclin B protein abundance. Cyclin B is the predominant cyclin family member that regulates the cellular progression through G2 and mitosis and through association with CDK1 forms the MPF (Gavet and Pines, 2010). This increase of cyclin B protein is driven by two cyclin A/CDK complexes. As G2 phase progresses the kinase activity of both the cyclin A/CDK2 and cyclin A/CDK1 complexes increases. A major phosphorylation target of both cyclin A/CDK complexes is the transcription factor, nuclear transcription factor Y (NF-Y). The phosphorylation of NF-Y increases the binding to the cyclin B1 promoter site and thereby increases cyclin B1 expression (Yun *et al.*, 2003). In parallel to activation of NF-Y the increased activity of the acetyl-transferase E1A associated protein p300 (p300) also increases transcriptional activity at the cyclin B1 gene loci (Bolognese *et al.*, 1999; Wasner *et al.*, 2003).

As with the previous cell cycle phase transition points, in addition to increasing cyclin expression, the removal of inhibitory phosphorylation must also be achieved. The inactivity of MPF is initially maintained by phosphorylation of CDK1 residues threonine 14 (Thr14) and tyrosine 15 (Tyr15) by the Wee1 kinase family (Ruiz, Vilar and Nebreda, 2010). The CDC25 family member, CDC25C, is the major phosphatase required for removal of the inhibitory phosphorylation moieties from MPF (Hoffmann *et al.*, 1993). The activation of CDC25C acts as a convergence point for multiple regulatory pathways of G2/M progression. While the predominant mechanism of CDC25C activation is mediated by a positive feedback loop between the MPF and polo-like kinase 1 (PLK1)

(Hoffmann *et al.*, 1993; Toyoshima-Morimoto, Taniguchi and Nishida, 2002). Multiple CDC25C phosphorylation sites have also been identified which are not dependent on these kinases and instead require the activity of ERK-MAP and JNK signalling pathways (Wang *et al.*, 2007; Gutierrez *et al.*, 2010). The continued activity of CDC25C gradually changes the balance between MPF and WEE1 mediated signalling, thereby targeting the WEE1 kinase with inhibitory phosphorylation and increasing proteasome-mediated degradation (Watanabe *et al.*, 2004) (Figure 1.6). While CDC25C is the predominant CDC25 family member that regulates G2/M transition, significant redundancy can be observed between CDC25A and CDC25C. This is demonstrated by overexpression of CDC25A at G2/M which induces high kinase activity of the MPF, without removal of inhibitory phosphorylation at Tyr15 and forces aberrant mitosis (Timofeev *et al.*, 2010).

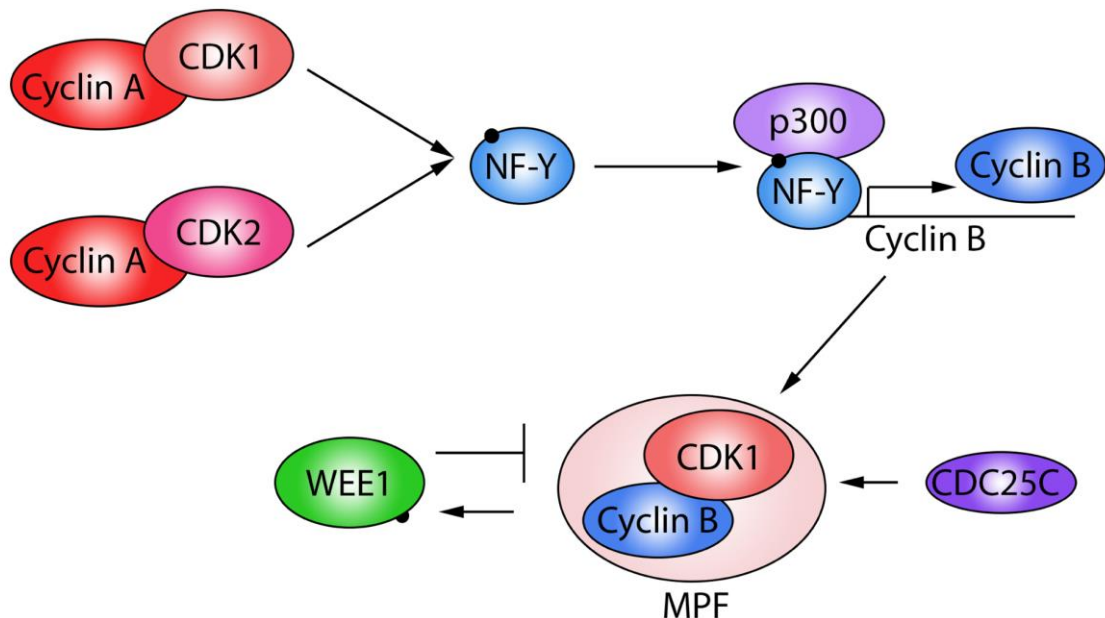


Figure 1.6. S-phase to G2 transition

Cyclin A forms two complexes with CDK protein family members: cyclin A/CDK1 and cyclin A/CDK2. Through collaboration of these complexes the transcription factor NF-Y is phosphorylated increasing affinity for DNA. In collaboration with the acetyltransferase, p300, the phosphorylated NF-Y promotes transcription of the G2/M cyclin family member, cyclin B. Increased protein abundance of cyclin B promotes formation of the cyclin B/CDK1 complex, commonly referred to as the MPF. Typically, the WEE1 kinase phosphorylates the MPF at inhibitory residues to prevent kinase activity. The increase of MPF abundance and the increasing activity of CDC25C will increase the reciprocal inhibitory phosphorylation on WEE1 reducing activity and forming a positive feedback loop, eventually yielding full MPF kinase activation.

The activation of MPF is a prerequisite for two further major preparatory events for mitosis initiation: The condensation of chromosomes and nuclear envelope alterations. Fully functional MPF activates the structural maintenance of chromosomes (SMC) family proteins which form the condensin complex and actively increase the condensation of DNA (Hirano, 2016). The primary hypothesis for the compaction method is the formation of loops which condense the DNA (Goloborodko *et al.*, 2016). This was supported by the demonstration that the condensin complex acts as a mechanochemical motor capable of ATP dependent movement along the DNA, in turn generating force to produce loops of DNA (Terakawa *et al.*, 2017).

The nuclear membrane is constituted of a lipid bilayer, with a section of the outer membrane forming the endoplasmic reticulum and the inner membrane containing specific nuclear pore proteins (Hetzer, 2016). While a continuous membrane, the regions forming an outer nuclear membrane and those forming ER are selectively enriched for non-overlapping protein groups (Watson, 1955). Nuclear membrane breakdown (NEB) during mitosis is initiated by dispersal of membrane-associated proteins into the cytosol (Ellenberg *et al.*, 1997). The phosphorylation of nuclear lamins by CDK1/ Cyclin B1 induces rapid depolymerisation inducing break down of the nuclear membrane (Heald and McKeon, 1990; Güttinger, Laurell and Kutay, 2009).

The activation of MPF complex is highly regulated and occurs before NEB (Gavet and Pines, 2010). Upon entry of dividing human cells into prophase, the nuclear membrane will start to break down and duplicated centrioles will begin to migrate to spindle poles. This process is referred to as open mitosis and is common between all higher eukaryotes. In organisms such as budding yeast, mitosis occurs through closed mitosis which does not require a breakdown of the nuclear envelope. In this case, the centrioles, also known as the microtubule organisation centre (MTOC), is associated with the nuclear envelope throughout mitosis (Ding *et al.*, 1997).

1.4.4 Mitosis

Following the G2 phase cells will enter mitosis through the action of MPF, however at this stage the process becomes less reliant on cyclin/CDK complex kinase activity. Instead, mitosis is regulated by alternative kinases, such as PLK1 and aurora kinase

family. PLK1 is a protein kinase which is distinguished by the presence of two polo-boxes at the c-terminus (García-Álvarez *et al.*, 2007; Kothe *et al.*, 2007). The CDK1/cyclin A complex activates PLK1 through both direct interaction and also stimulation of Aurora A kinase, which increases phosphorylation of the PLK1 T-loop residue tyrosine 210. The expression of a PLK1 tyrosine 210 phosphorylation mimic overcomes the requirement for aurora A activation to promote cell cycle progression (Macůrek *et al.*, 2008; Gheghiani *et al.*, 2017; Vigneron *et al.*, 2018). One of the events mediated by activated PLK1 during the end of prophase and early metaphase is the deposition of CENP-A onto centromeres (McKinley and Cheeseman, 2014). The canonical centromere is a region of highly specialised proteins, predominantly nucleated by CENPA and CENPB which act as a site for protein binding to form the kinetochore (Earnshaw and Rothfield, 1985). CENPA shares distinct similarities to histone core subunit, H3. During mitosis, CENPA displaces H3 within the centromere. However, this displacement is not complete and instead interlocking subdomains of CENPA and H3 are formed (Yoda *et al.*, 2000; Ross, Woodlief and Sullivan, 2016). The CENPA and H3 within this region have a specific pattern of chromatin tail modification, including hypoacetylation of H3 (Sullivan and Karpen, 2004). CENPB, in contrast to CENPA, recognises stretches of DNA. The DNA surrounding a centromere is highly repetitive due to the abundance of alpha satellite DNA (Grady *et al.*, 1992). The CENPB DNA binding subdomain recognises the repetitive DNA of alpha satellite sites referred to as a CENPB-box. The binding of CENPB to DNA can be prevented with the removal of a DNA binding protein subdomain (Pluta *et al.*, 1992) (Figure 1.7).

Upon completion of prophase in higher eukaryotes, the nuclear membrane will have been removed, the chromatin will have condensed and the centrioles will have migrated to the cell poles. This stage of mitosis is referred to as metaphase. The chromosomes subsequently align along the cellular midline, or metaphase plate. Two mechanisms referred to as direct congression and peripheral congression act to align chromosomes (Maiato *et al.*, 2017). Direct congression occurs when chromosomes are already positioned close the metaphase plate upon NEB and does not require the activity of centriole motor proteins, such as CENP-E for realignment (Barisic *et al.*, 2014). In contrast, peripheral congression occurs when chromosomes are not bi-oriented upon NEB and initially become polarised before trafficking, in a centriole motor dependent manner, to the metaphase plate (Bancroft *et al.*, 2015).

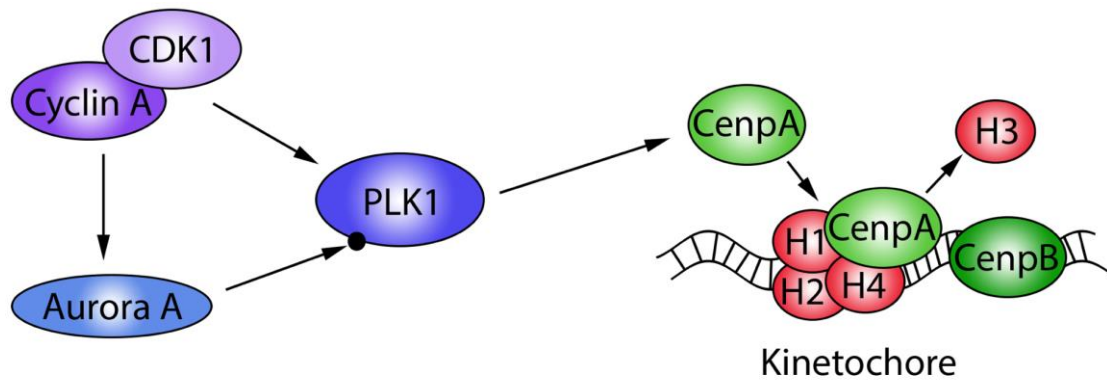


Figure 1.7. G2/M transition

The cyclin A/CDK1 complex promotes activation of aurora A kinase activity. The collaboration of cyclin A/CDK1 and activated aurora A, increases the PLK1 kinase activity. The phosphorylation of CENPA by PLK1 promotes replacement of H3 in the regions surrounding the kinetochore. The CENPA/H3 replacement is vital for definition of the kinetochore region.

Both the process of chromosome congression and the eventual separation of chromosomes requires spindle fibre and centromere. The docking sites between spindle fibres and the centromere forms the kinetochore. This site is an active site of protein signalling, which is responsible for regulating chromatid separation. To promote exit from metaphase the phosphorylation of CDC25C is reduced by the activity of PP2A, increasing the phosphorylation of CDK1 Threonine 15 and Tyrosine 15, inactivating the MPF prior to cyclin B degradation (Forester *et al.*, 2007). Cyclin B is rapidly degraded by the ubiquitin-mediated protein degradation pathway, specifically mediated by the anaphase-promoting complex/cyclosome (APC/C) (Hershko, 1999). The importance of cyclin B degradation can be demonstrated by ectopic expression of non-degradable cyclin B1, which prevents the progression of cells into anaphase, with only 30% of endogenous protein required to prevent progression (Chang, Xu and Luo, 2003). Activated PLK1 increases activity of the APC/C, which degrades the MPF and securin promoting entry into anaphase (Descombes and Nigg, 1998; Kotani *et al.*, 1998; Qiao *et al.*, 2016). Securin is an inhibitor of the mitotic protease separase, which is released upon degradation and cleaves kleisin, a component of the cohesin complex and the major regulator of chromosome maintenance at the metaphase plate (Lin, Luo and Yu, 2016).

The serine-threonine kinase, Greatwall (gwl) was initially identified in *Drosophila* as a factor required for correct chromosomal condensation (Yu *et al.*, 2004; Voets and Wolthuis, 2010). MastL inhibits the activity of PP2A-B55 and regulates the phosphorylation status of CDK1, with loss of MastL expression increasing dephosphorylation of CDK1 and promoting premature entry into mitosis from G2 (Lorca and Castro, 2013; Diril *et al.*, 2016).

1.4.5 Mitotic exit

The segregation of chromosomes to spindle poles signals the end of anaphase following which mitotic signalling pathways are deactivated to prepare cellular structures for re-entry into G1 or quiescence. For re-entry of the cell into G1 phase – the nuclear envelope must be reformed and the DNA decondensed from mitotic chromosomes.

Upon progression into anaphase the APC/C targets cyclin B1 for degradation reducing overall levels. The APC/C complex will subsequently associate with CDC20 homolog 1 (CDH1), which increases the size of substrate repertoire and facilitates degradation of mitosis related proteins, such as, aurora A kinase (Floyd, Pines and Lindon, 2008; van Leuken *et al.*, 2009). In parallel with the degradation of key substrates by APC/C, the activity of protein phosphatase family members: PP1 and PP2A, increases. The substrate specificity of PP1 is regulated by sub-unit proteins which associate with the complex. For example, the sub-unit Repo-Man is associated with sequential dephosphorylation of multiple H3 phosphorylation sites which are required for mitotic exit (Qian *et al.*, 2011). In addition, the PP2A-B55 complex can reduce reformation of the nuclear envelope when expression levels are reduced (Mehsen *et al.*, 2018).

The progression through the cell cycle will proceed unabated unless DNA damage is detected either endogenously or through toxic insult. One potential impact of toxin exposure is DNA damage due to either internal or external sources of DNA altering agents. The formation of a DNA double-strand break (DSB) is the most mutagenic form of DNA damage and also the break type which is most significantly associated with defects in late-stage cell cycle, such as, mitotic separation of chromosomes. The generation of double-strand breaks and the canonical repair pathways will now be explored.

1.5 DNA Damage

Alterations to the genetic code are the major basis for evolution and without this process adaptation of organisms over generations would cease. However, there are thousands of genomic damage events each day and if all of these induce genetic code changes then the organism may not survive. Therefore, several sophisticated mechanisms have evolved to protect the genome from these events and repair damage, substantially reducing the number of incorporated mutations.

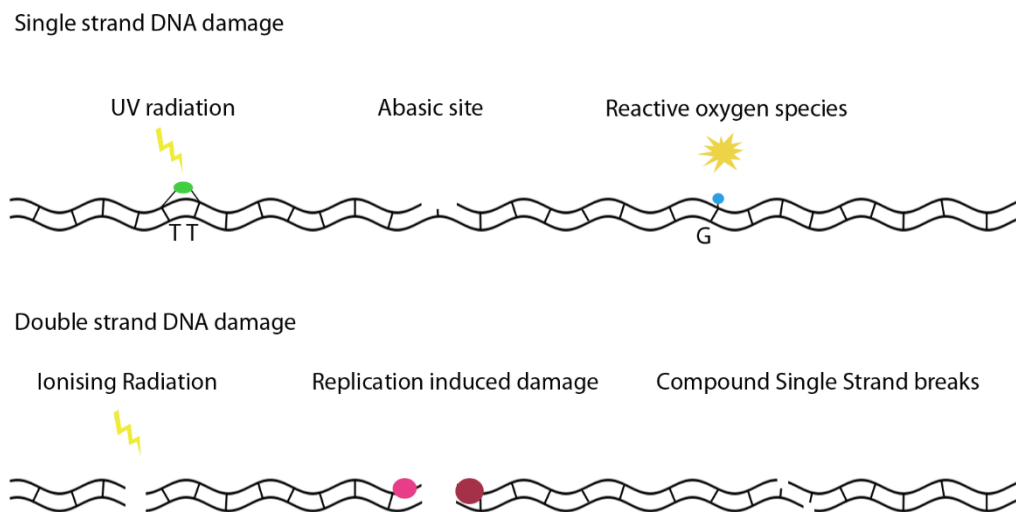


Figure 1.8. DNA damage sources

Single strand DNA damage induction can occur following exposure to multiple processes. Three examples of these processes are double-thymidine dimers formed through UV radiation, development of abasic sites following removal of mismatched bases, and oxidation of bases, most commonly guanine bases, by reactive oxygen species. The development of double strand breaks in comparison can occur directly through ionisation of bases separating the DNA backbones. The collapse of DNA replication machinery can induce formation of double strand DNA breaks. Finally the development of single strand DNA breaks in close proximity can form double strand breaks.

DNA can be damaged by a number of sources both exogenous and endogenous to the cell. Endogenously, many processes critical to cellular viability are responsible for generating DNA damage events, such as the stalling of DNA polymerases during replication of DNA. In addition, exogenous events can cause damage, including, exposure to high energy radiation and DNA binding chemicals. This damage of DNA can be in the form of a single strand break or a DSB which is the more mutagenic form of

damage (Figure 1.8). This event in the most extreme cases can lead to the incorrect fusing of chromosomes, referred to as chromosomal translocations. A prime example of the severe pathology associated with this outcome from DNA damage is translocation of chromosome 21. This translocation increases the copy number of all genes within the long arm of chromosome 21 and yields a Down's syndrome like phenotype (Petersen *et al.*, 1991). Each source of genomic damage will have a particular mechanism of action which requires specific response pathways to be activated. Pathway activation orchestrates the repair of damage, prevents cell cycle progression, and initiates-controlled cell death pathways.

1.5.1 Ionising radiation and radiomimetic drugs – cause or cure?

Background levels of high-energy radiation are always present, however, exposure of organisms to high dose emission sources can be extremely harmful and can even be fatal if the dose is high enough. In addition, several drugs have been synthesised which mimic the effect of high dose radiation (Umezawa, 1976). High energy radiation such as, γ -radiation and X-rays, and radiomimetic drugs such as, bleomycin, cause ionisation of molecules within the cell. In particular, the exposure to IR can ionise water molecules producing ROS and either through energy transfer from ionised water or direct ionisation impact biomolecules such as proteins (Kaplan, 1960; Reisz *et al.*, 2014; Franco *et al.*, 2016). These events have two major cellular consequences, altered mitochondrial function and direct damage to DNA.

The mitochondria are critical organelles responsible for multiple homeostatic processes, but most importantly the production of ATP by oxidative phosphorylation. This process utilises a group of proteins referred to collectively as the electron transport chain. As a by-product of oxidative phosphorylation ROS are produced which can modify both proteins and DNA. Proteins can be oxidised at cysteine residues by ROS which has significant impact upon structure and function, and multiple studies have identified the importance of this modification for cell proliferation and conversely apoptosis induction (Ray, Huang and Tsuji, 2012; Redza-Dutordoir and Averill-Bates, 2016). The disturbance of ROS homeostasis by exposure of cells to IR can have a significant impact upon mitochondrial dynamics. One alteration to dynamics is a change in activity of the electron transport chain increasing cellular ROS levels (Kam and Banati, 2013). A

proposed mechanism is a transient calcium dependent mitochondrial membrane permeability transition which releases ROS into the cell (Leach *et al.*, 2001). In addition, the number of mitochondria increases following exposure to IR, amplifying the ROS production (Yamamori *et al.*, 2012). In response to this increased cellular ROS level a cell cycle checkpoint arrest is induced, which prevents the further progression of cell cycle without reduction of ROS levels. The cell cycle arrest induced by ROS is compounded by direct and indirect induction of DNA damage. The exposure of DNA to ROS can lead to both single strand DNA and double stranded DNA damage (Cadet and Richard Wagner, 2013). Given both these cellular impacts the exposure of cells can be highly mutagenic, potentially causing cancer. However, these phenotypes can also be utilised to achieve death of cancer cells.

Medical physics has utilised high energy radiation as a cancer therapy and in diagnostics through the use of X-rays for many years. For example, a subset of breast cancer patients possesses a mutation in the protein BRCA1 protein, which as discussed earlier orchestrates the homologous recombination response (Section 1.5.4). In these tumours it was expected that IR would be highly efficient at inducing death of cancer cells due to perturbed response to DSB induction, however, the data from multiple studies are inconclusive (Kan and Zhang, 2015). This is a problem widely experienced with potential biomarkers of IR therapy and as such the need for increased effort to identify new markers is paramount to improving treatment (Zeegers *et al.*, 2017). A recent break-through identified the FDXR transcript as being tightly linked to IR treatment *in vivo* and can be used to give an effective readout of dose, however, no prognostic data for cancers (O'Brien *et al.*, 2018).

1.5.2 Double-stranded DNA damage and recognition

DNA DSBs occur when both pentose-phosphate backbones of DNA are broken. A prime example of a toxic insult which yields this break type is exposure to high energy radiation (Cannan and Pederson, 2016). Unlike ssDNA repair pathways, which can utilise the intact DNA strand as a template, a DSB does not have an inherent template strand by which to repair. Which in addition to the chance for chromosomal translocation also increases the likelihood of small insertions and deletions (INDELs) of sequence at the break sites.

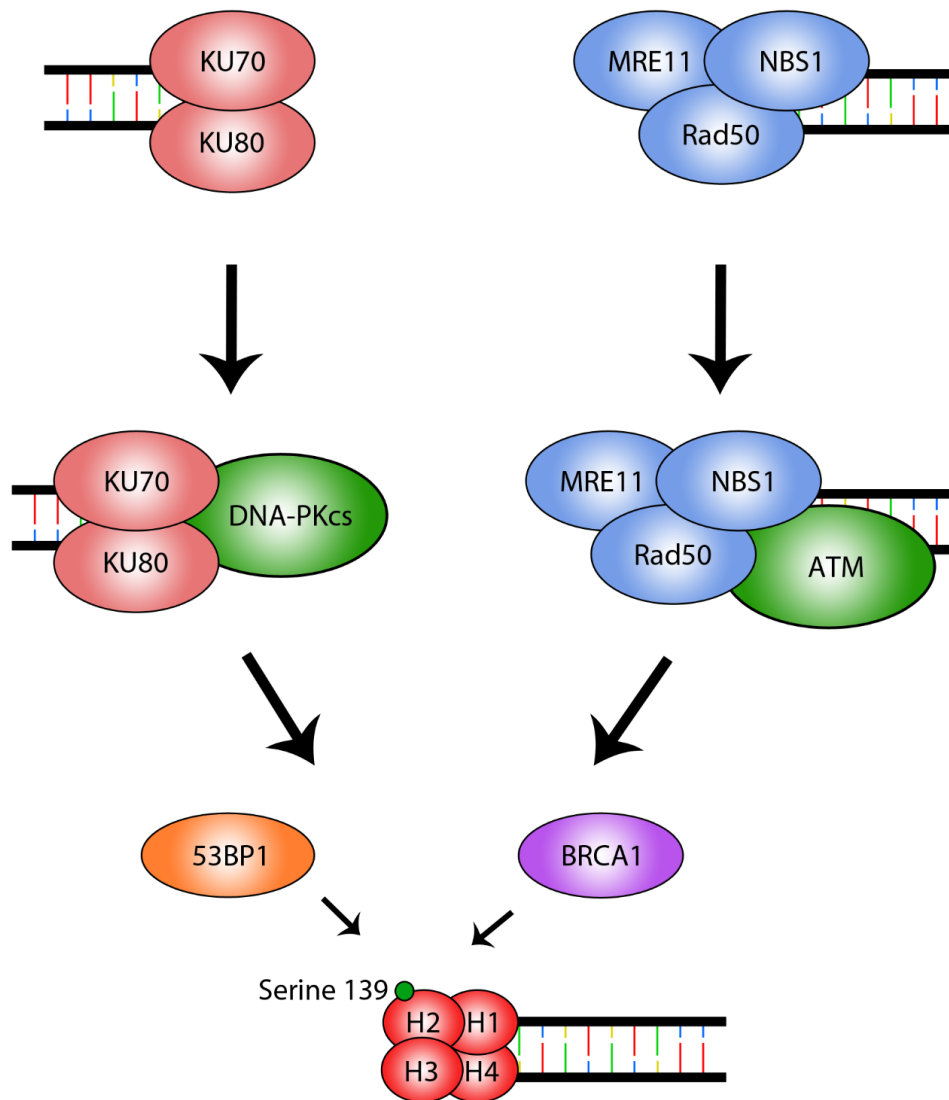


Figure 1.9. Double-strand break response cascade

Recognition of a free double strand DNA break end will be through either the Ku70/Ku80 complex or the MRN complex. If the Ku70/Ku80 complex recognises the break end then DNA-PKcs is recruited and the full DNA-PK complex is formed. If the MRN complex is recruited then ATM will become associated. At this point of the cascade multiple processes control the recruitment of further factors, such as, 53BP1 and BRCA1, including the phosphorylation of H2 at serine 139.

The incorporation of INDEL sequences and translocations can alter the specific arrangement of genetic elements on a chromosome, which are paramount for regulatory sequences both proximal and distal from genes, and can significantly alter transcriptional activity (Riethoven, 2010). The pathological consequences of chromosomal translocations can be demonstrated by two examples, the first example is the

intrachromosomal translocation of chromosome 3 which increases transcription levels of BCL6 in B-cells non-Hodgkin lymphoma, driving tumorigenesis (Keller *et al.*, 2006). The second example is the interchromosomal translocation of sequence during the development of Burkitt's lymphoma in which the DH region of chromosome 14 becomes associated with the MYC oncogene, resulting in constant activation of c-MYC, and inducing aberrant cell growth (Haluska, Tsujimoto and Croce, 1987). To prevent these pathological consequences repair pathways have evolved to quickly repair DSBs before translocations can occur. The repair mechanisms which have evolved are generally split into 2 major pathways – non-homologous end-joining (NHEJ) and homologous recombination (HR). Two protein complexes, the KU heterodimer (Ku70 and Ku86) and the MRN complex (MRE11, NBS1, and Rad50) both recognise the double strand break site. While the KU heterodimer and MRN complex share a high affinity for DSB ends there are major structural and signalling distinctions between the two complexes (Blier *et al.*, 1993). The KU heterodimer binds to the DSB end and recruits the DNA-PK catalytic subunit (DNA-PKcs) (Hammel *et al.*, 2010). This complex conducts a process referred to as synapsis which stabilises the two DSB ends and prevents drift (Graham, Walter and Loparo, 2016). The MRN complex, composed of meiotic recombination 11 (MRE11), Rad50, and nibrin (Nbs1), binds to the DSB end through the DNA recognition site of MRE11 (Haber, 1998). Following end recognition dimerization of Rad50 proteins from MRN complexes on opposite sides of the break facilitate break end tethering. The final complex component, Nbs1 is considered to be an interaction platform by which the MRN complex associates with other proteins of the DNA damage response cascade (Williams, Williams and Tainer, 2007). Following DSB binding MRN acts to activate the ataxia telangiectasia mutated (ATM) kinase (Uziel *et al.*, 2003). While hypotheses have been suggested for the mechanism of pathway selection, there are multiple steps of this process which require more investigation. However, the initial competition of these two DSB end recognition mechanisms and cell cycle position is suggested play a role in the selection of repair pathway (Chapman *et al.*, 2013) (Figure 1.9).

The proteins ATM and DNA-PK are both members of the Phosphatidylinositol 3-kinase related (PIKK) family (Smith and Jackson, 2010). The activation of these kinases induces multiple signalling pathways, one of which is the chromatin response cascade. This step of DSB response is critical for further progression of DSB repair and recruitment of factors which expand the response to a cellular level. Initially, DNA-PK or ATM

phosphorylates histone subunit H2A.X (H2A.X) forming γ H2AX foci which extends for multiple kilobases surrounding the break-site (Stiff *et al.*, 2004). The mediator of DNA damage checkpoint protein 1 (MDC1) binds to this histone mark and directly interacts with ring finger protein 8 (RNF8) (Mailand *et al.*, 2007). RNF8 is an E3 ubiquitin-protein ligase that ubiquitinates H2A.X at lysines 119 and 120 of the histone tail (Huen *et al.*, 2007; Ma, Keller and Yu, 2011). The sustained ubiquitination of these residues also requires the activity of a second ring finger protein 168 (RNF168) (Gatti *et al.*, 2012). The interplay of RNF8-RNF168 produces the characteristic ubiquitin marks on H2A.X. However, this complex requires ubiquitination of H1 subunit by HUWE1 for stabilisation on chromatin (Mandemaker *et al.*, 2017). The regulation of histone ubiquitin marks is an example of the nuanced temporal and spatial regulation that occurs to achieve the correct progression of DSB repair. At this point pathway selection between NHEJ and HR occurs by the interplay of two major protein complexes centred around 53BP1 and breast and ovarian cancer susceptibility protein 1 (BRCA1).

The p53 binding protein 1 (53BP1) is the central component of a protein complex that binds to chromatin and promotes the progression of NHEJ. 53BP1 is recruited by interaction with the breast cancer susceptibility (BRCT) domain of MDC1 and binding to histone ubiquitination through a C-terminal ubiquitin interacting domain (Stewart *et al.*, 2003; Eliezer *et al.*, 2009; Fradet-Turcotte *et al.*, 2013). Following stabilisation on chromatin, 53BP1 acts as a nucleation factor with multiple domains which act as sites for recruitment of factors critical for NHEJ pathway selection (Panier and Boulton, 2014). Rap1 interacting factor (RIF1) is a major interacting partner of 53BP1 which binds the N-terminal phosphor-SQ-TQ domain (Chapman *et al.*, 2013). RIF1 acts to inhibit both resection of DNA and accumulation of BRCA1 at damage sites, promoting the progression of NHEJ (Zimmermann *et al.*, 2013). To promote the progression of break repair towards HR 53BP1 must be displaced by the antagonist, BRCA1, which is recruited to chromatin through interaction with the Rap80 complex (BRCC36, Abraxas, Merit40, and BRCC45) (Wu *et al.*, 2012). The RAP80 complex is recruited through interaction with poly-ubiquitinated H2AX, with the loss of any factor reducing the efficiency of BRCA1 recruitment (Sobhian *et al.*, 2007; Coleman and Greenberg, 2011; Hu *et al.*, 2011). Subsequently the KU complex is removed from the DSB by the BRCA1 interacting partner, C-terminal binding protein 1 interacting protein (CtIP) (Coates *et al.*, 2016). This promotes the progression of HR to repair the DSB. The competition between

BRCA1 and 53BP1 is thought to be the last selection point before pathway commitment occurs.

1.5.3 Non-homologous end-joining

The defining characteristic of NHEJ is the direct ligation of DSBs with only minor processing and no long-range resection or homology searching (Lieber, 2010). This is particularly true for blunt end DSB or compatible ends, which *in vitro* can immediately be re-ligated by the minimal complex of Ku70/Ku80, XRCC4, and ligase IV (H. H. Y. Chang *et al.*, 2016). However, the DSB events induced *in vivo* are rarely simplistic blunt end or compatible ends, therefore a large number of additional factors are required (Figure 1.10). For example the process of synapsis is reliant on either DNA-PK or MRN complexes, but this process is enhanced by XRCC4 and its closely related paralogs - paralog of XLF and XRCC4 (PAXX) and XRCC4-like factor (XLF), which stabilise the DSB by forming filamentous structures along DNA surrounding the break site (Andres *et al.*, 2007; Ropars *et al.*, 2011). Interestingly, a major role of XRCC4 *in vivo* is to regulate the stability of ligase IV, with a direct link between the protein stability of ligase IV and XRCC4 abundance being observed (Nick McElhinny *et al.*, 2000). In addition to synapsis, *in vivo* DSB events will commonly be modified with adducts which require removal by resection. Artemis is a dual function nuclease which can cleave a variety of damage overhangs (Ma, Schwarz and Lieber, 2005; Chang, Watanabe and Lieber, 2015). This process can result in the formation of abasic sites which must be filled by specific polymerases associated with NHEJ, predominantly polymerase λ and polymerase μ (Ramsden, 2011). Interestingly, the activity of polymerase λ is augmented by association with PAXX, demonstrating the complicated interplay between factors which modulate activity (Craxton *et al.*, 2018). While DNA repair utilising NHEJ occurs rapidly following damage induction, due to the lack of template the NHEJ pathway is error prone due to the lack of template and is, therefore, the preferred pathway only when template DNA is not available.

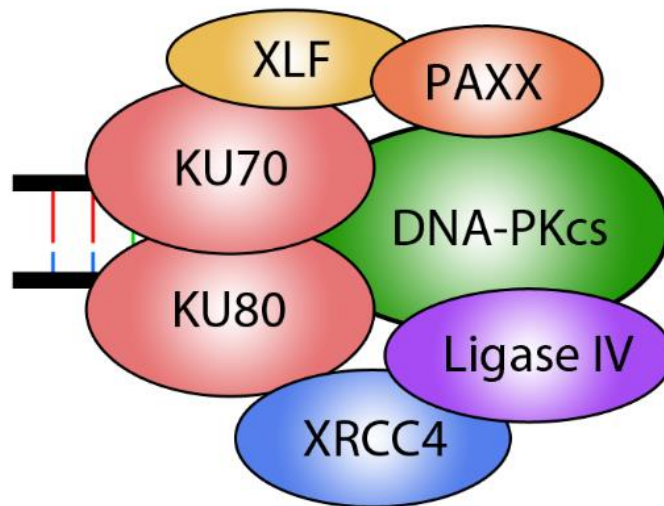


Figure 1.10 NHEJ Complex

The complex which mediates repair of the double strand DNA break by non-homologous end joining (NHEJ). In addition to the factors represented many alternative factors are required depending on the specific structure of each break site.

1.5.4 Homologous Recombination Repair

The HR pathway requires two major conditions to successfully activate, the first of these is the presence of a separate undamaged template DNA molecule and the second is long-range resection surrounding the break site (Rothkamm *et al.*, 2003). Initially, MRN and CtIP are recruited to break sites and resect DNA surrounding the break, producing 15-20 nucleotides ssDNA (Cannavo and Cejka, 2014; Makharashvili and Paull, 2015). During this process, MRN removes KU from the break end, fully committing to break repair through HR (Langerak *et al.*, 2011; Coates *et al.*, 2016). Further long-range resection complexes are recruited to these sites, known as resection complex 1 (BLM, DNA2, RPA, and MRN) or complex 2 (EXO1, BLM, RPA, and MRN) (Nimonkar *et al.*, 2011). The interaction of CtIP and resection complex 1 has been shown to enhance the exonuclease activity of the complex, demonstrating the link between short- and long-range resection in HR (Daley *et al.*, 2017). As resection progresses the single-stranded DNA is coated with replication protein A (RPA). The RPA complex is heterotrimeric and formed from 3 components: RPA70, RPA32, and RPA14 (Zou *et al.*, 2006). The major roles of this complex are to prevent further nuclease activity and support the correct polarity of exonucleases (Fanning, Klimovich and Nager, 2006; Nimonkar *et al.*, 2011).

Upon pathway progression the RPA complex is replaced by the RAD51 protein, this is mediated by a combination of Rad52 and BRCA2 in humans, with filament formation initiating from binding loci of these proteins (Jensen, Carreira and Kowalczykowski, 2010; Liu *et al.*, 2010). Once coated in RAD51 the filament will conduct a homology search of the sister chromosome, the detection of homology is dependent upon an 8bp region of microhomology (Qi *et al.*, 2015). The necessity for a template DNA strand restricts the use of HR to S-phase and G2 phase of the cell cycle. Many of the mechanisms that restrict DNA repair pathway of choice to a specific cell cycle phase require more intensive research. However, one mechanism identified involves the inhibitory phosphorylation of BRCA2 by CDK proteins blocks its interaction with Rad51, preventing replacement of RPA and homology searching (Esashi *et al.*, 2005). Following identification of homology, the Rad51 coated strand undergoes a process of strand invasion which is resolved by a Holliday junction (Bizard and Hickson, 2014). The utilisation of a template strand increases the fidelity of HR repair substantially reducing the chance of incorporating mutations (Figure 1.11).

In addition to the two PIKK proteins already mentioned a third member of this family, Ataxia telangiectasia and Rad3 related (ATR), is activated by the presence of single-stranded DNA coated in RPA. ATR is specifically recruited to these sites through the activity of its binding partner ATR interacting protein (ATRIP) (Zou and Elledge, 2003). Following recruitment, ATR is responsible for single-stranded dependent phosphorylation of the RPA32 which stabilises the RPA complex (Shiotani and Zou, 2009). The increase in activity of ATR following resection leads to a switch between ATM activity and ATR activation in HR response (Shiotani and Zou, 2009). Additionally, the pathway involvement of ATR is much broader than either DNA-PK or ATM due to the activation following the detection of single-stranded DNA. Transcriptional stress is an important physiological event that is heavily reliant on ATR signalling. For example, the phosphorylation of p53 on Serine 15, an important activation residue, is ATR and RPA dependent following transcriptional stress induction (Derheimer *et al.*, 2007).

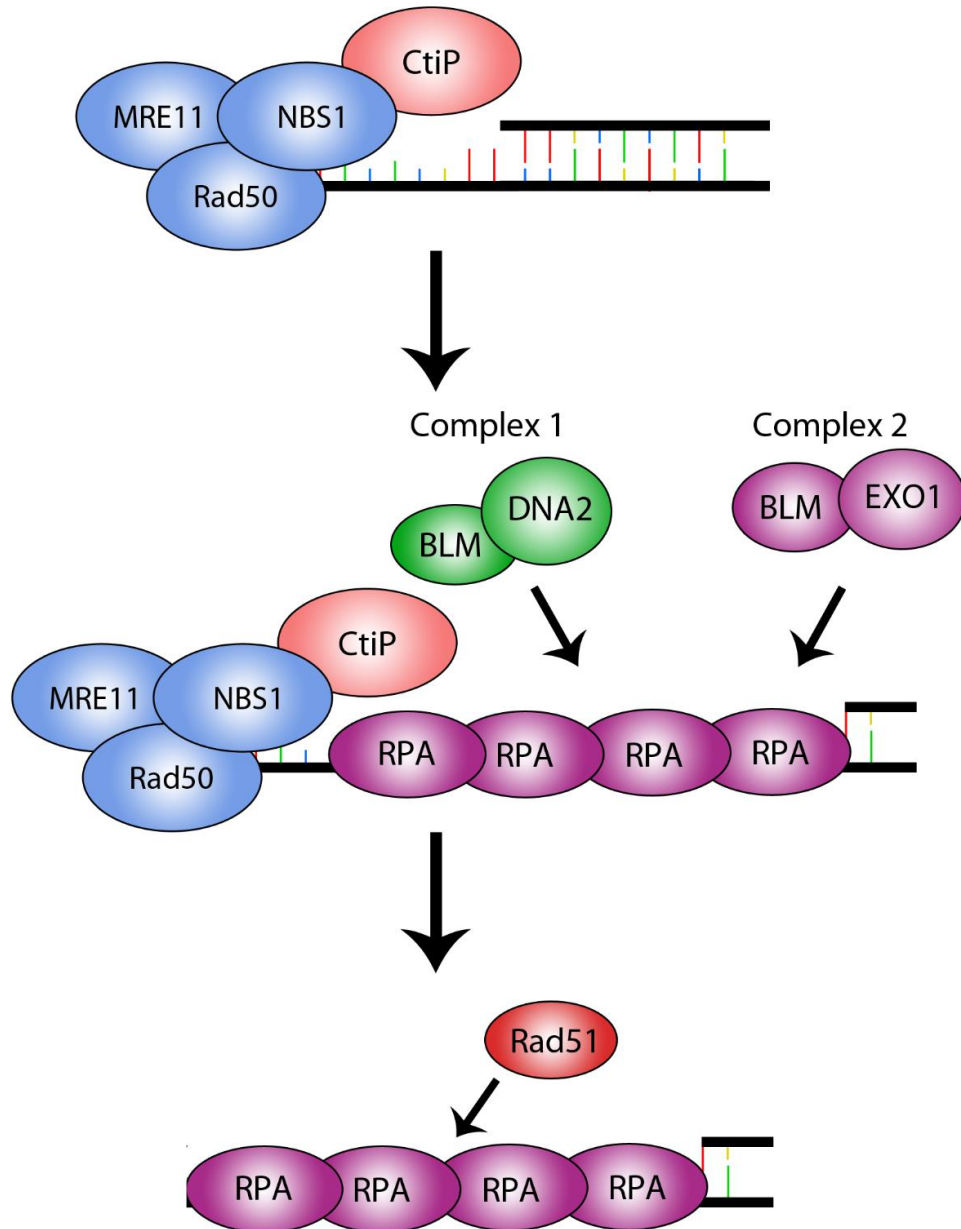


Figure 1.11. HR pathway

Recognition of a DNA double strand break end by MRN promotes the recruitment of CtIP, with this complex promoting short range resection of double stranded DNA away from the break site. The resulting single stranded DNA is coated in RPA complex prior to recruitment of either resection complex 1 or complex 2, which resect and form single stranded DNA hundreds of bases from the break site. The RPA complex coated strand will undergo a process of replacement by the RAD51 protein before strand invasion by homologous searching occurs.

1.6 Cell cycle and the DDR

The replication of DNA during S-phase requires the translocation of the replisome along DNA (Gao *et al.*, 2019). The collision of the replisome with sites of DNA damage can increase the severity of the break and therefore mechanisms have evolved to prevent this eventuality. These blockades are referred to as cell cycle checkpoints. The canonical cell cycle checkpoint is the G1 checkpoint or “restriction point” which arrests the cell in response to lack of mitogen signalling and following DNA damage occurring within G1 phase (Pardee, 1974). In addition to the “Restriction point” three further checkpoints exist in mammalian cells: the S-phase checkpoint, the G2 phase/Mitosis (G2/M) checkpoint, and the spindle assembly checkpoint (SAC) (Figure 1.12). The latter two checkpoints occur post DNA replication and are instead utilised to prevent incorrect chromosome segregation due to structural abnormalities or loss of mitosis regulation.

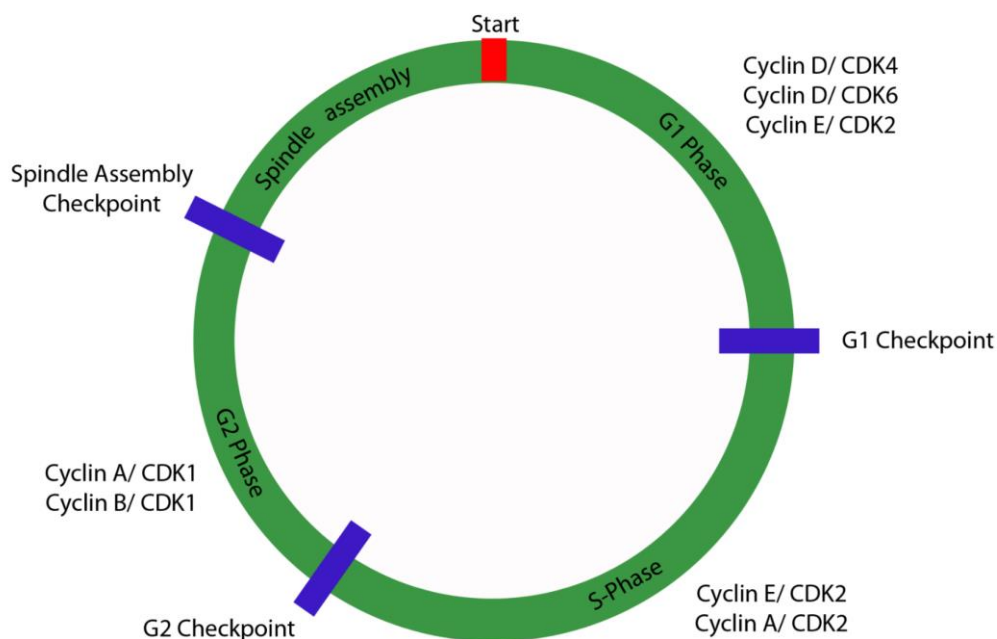


Figure 1.12. Cell cycle progression

The cell cycle is separated into 4 stages: G1, S-phase, G2, and mitosis. The progression through each of cell cycle phase is controlled by the formation of specific cyclin/CDK complexes. In addition to these cell cycle phases, four cell cycle checkpoints have been found: G1 checkpoint, S-phase checkpoint, G2/M checkpoint, and spindle assembly checkpoint.

Currently, there is significant debate as to the definition between G2/M Checkpoint and SAC. In this review, the G2/M checkpoint will be defined as any event that regulates the activation of kinases within the MPF and other CDK/cyclin complexes (Rieder and Maiato, 2004). In contrast, the spindle assembly checkpoint will include any event regulating the separation of chromosomes at anaphase through kinetochore or centromeres. A multitude of events can lead to DNA damage and activation of PIKK proteins. Three master checkpoint regulating proteins are activated by the PIKK proteins, these being: checkpoint protein 1 (CHK1), checkpoint protein 2 (CHK2), and PLK1

1.6.1 G1/ S-phase checkpoint

The G1/S-phase checkpoint prevents cell cycle entry of cells which are undergoing active DNA break repair. The master regulator, CHK2 is predominantly required for activation of the G1 checkpoint in response to DNA damage. Following DNA damage, CHK2 is activated by ATM-mediated phosphorylation, specifically at threonine 68 (Thr68) within the T-loop (Ahn *et al.*, 2000). This phosphorylation promotes dimerisation of two CHK2 proteins. The dimerisation of CHK2 causes auto-phosphorylation of the T-loop further increasing kinase activity in a positive feedback loop (Ahn *et al.*, 2002). In addition to phosphorylation by ATM, the alternative kinases DNA-PK and MSH1 have also been suggested to regulate activation of CHK2, however, this is in more limited circumstances (Adamson *et al.*, 2005; Shang *et al.*, 2010).

The transcription factor p53 is referred to as the guardian of the genome and is constantly degraded by mouse double minute 2 homolog (MDM2) mediated ubiquitination (Moll and Petrenko, 2003). Upon detection of DNA damage CHK2 phosphorylates serine 20 of p53 which reduces binding of MDM2, with MDM2 being ubiquitinated and degraded (Banin *et al.*, 1998; Brooks and Gu, 2010). The non-degraded p53 subsequently forms an oligomeric complex and following nuclear translocation, activates a host of genes associated with the cellular stress response. For example, the Cip/Kip family member p21 gene promoter contains a p53 binding site and mediates cell cycle arrest following DNA damage through interaction with multiple CDK family members (Harper *et al.*, 1993; Gartel and Tyner, 1999). In addition, CHK2 can phosphorylate CDC25A at serine 123, which induces CDC25A degradation and prevents removal of an inhibitory

phosphorylation on CDK4/CDK6, preventing cell cycle progression (Falck *et al.*, 2001) (Figure 1.13).

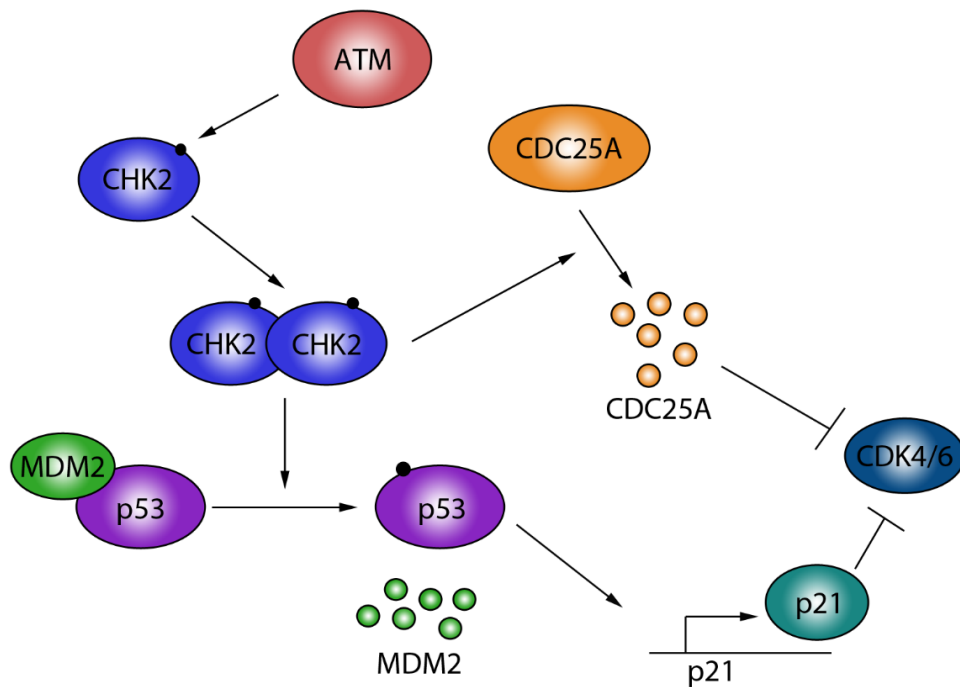


Figure 1.13. G1 Checkpoint

The detection of a DNA break will induce activation of the kinase ATM. The activated ATM can phosphorylate CHK2, inducing the dimerisation and full activation of CHK2 kinase activity. The activated CHK2 phosphorylates p53 promoting release and degradation of the inhibitor MDM2 and increases protein abundance of p21, a G1 checkpoint activator, that inhibits CDK4/CDK6 kinase activity. In addition, activated CHK2 promotes degradation of CDC25A, preventing removal of inhibitory phosphorylation moieties from CDK4/CDK6.

To prevent unwanted stalling of the cell cycle CHK2 is maintained in an inactive form by the activity of Protein phosphatase 2 (PP2A) which removes T-loop phosphate groups (Freeman and Morrison, 2011). Due to the importance of CHK2 for transferring the signal from the DNA damage site to further cascades, mutation of CHK2 can lead to loss of the G1/S-phase checkpoint. Wu. *et al.* analysed four CHK2 mutants, two of which contained frameshift mutations which lead to loss of kinase activity, these mutants in particular prevented the activation following treatment with IR (Wu, Webster and Chen, 2001).

1.6.2 S-phase checkpoint

The intra S-phase checkpoint predominantly responds to stalled replication forks and damage events which impede genome replication. Upon stalling of the replicative polymerase, the DNA unwinding complex will continue to unwind ahead of the lesion, generating a stretch of single-stranded DNA which must be coated with RPA complex to prevent degradation (Longhese, 1996). This initial step is vital for response to the stalling of replication forks. The coating of DNA with RPA recruits the 9-1-1 complex (Rad1-Rad9-Hus1), containing TopBP1, claspin, and ATRIP, which induces activation of ATR (Bermudez *et al.*, 2003; Kumagai *et al.*, 2006). The activity of CHK1/ATR, is required to limit replication origin firing following damage by inhibitory phosphorylation of CDK1, with the loss of this process inducing uncontrollable origin firing (Moiseeva *et al.*, 2019). Upon damage detection surrounding a replication fork, secondary “licensed” origin sites surrounding the break are activated, while in contrast, overall genome duplication is inhibited (Yekezare, Gómez-González and Diffley, 2013). In addition to activation of ATR, the importance for ATM within this process can be demonstrated by the induction of radioresistant DNA synthesis, in response to treatment with IR following a loss of ATM activity (Painter, 1981; Gatei *et al.*, 2003). This phenotype is due to the failure of an ATM/CHK2/CDC25A complex which is responsible for the degradation of CDC25A and prevention of cell cycle progression (Falck *et al.*, 2001).

1.6.3 G2/M checkpoint

The G2/M checkpoint arrests cells prior to engagement of the spindle complex. CHK1 is the master signal transducer for the G2/M and SAC checkpoints. Following treatment with IR, UV or hydroxyurea (HU) CHK1 is phosphorylated at the canonical activation site, Serine 345 (S345), predominantly by ATR (Liu *et al.*, 2000). Multiple phosphorylation sites have been identified within CHK1. Full activation in response to IR required the phosphorylation of serine 317 (S317) (Zhao and Piwnicka-Worms, 2001). This phosphorylation event is mainly mediated by ATR, however, ATM can also phosphorylate S317 following treatment with IR (Gatei *et al.*, 2003). Following activation, CHK1 will interact with WEE1, a second vital kinase responsible for inhibition of G2/M progression. The active CHK1/WEE1 complex has multiple targets which prevent the progression of the cell cycle (Blasius *et al.*, 2011). For example,

CHK1/WEE1 phosphorylated the inhibitory tyrosine 15 (tyr15) site of CDK1 preventing cell cycle progression (O'Connell *et al.*, 1997). CHK1 also has substrates which are independent of WEE1 including phosphorylation of threonine 507 within CDC25a which mediates binding of 14-3-3γ relocating the complex to the cytoplasm for proteolytic degradation (Kasahara *et al.*, 2010). Mutation of threonine 507 prevents interaction with 14-3-3γ and prematurely activates the MPF inducing premature progression into mitosis following DNA damage (Chen, Ryan and Piwnica-Worms, 2003). The 14-3-3 family are vital regulators of cell cycle progression and has multiple members which can recognise all three CDC25 family members and induce cytoplasmic translocation preventing function (Gardino and Yaffe, 2011) (Figure 1.14).

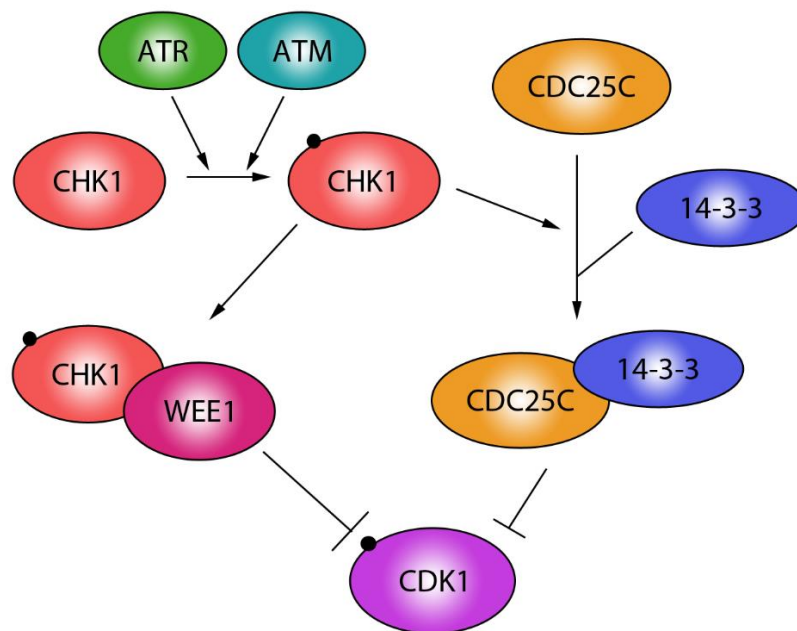


Figure 1.14. G2/M Checkpoint

The detection of DNA damage will activate the damage kinases, ATR and ATM. This activation event promotes the phosphorylation of CHK1, the master regulator of G2/M checkpoint. Activated CHK1 can form a complex with WEE1 which increases the inhibitory phosphorylation of CDK1, preventing kinase activity and cell cycle progression. In addition, CHK1 also phosphorylates CDC25C, promoting association with the 14-3-3 family. The CDC25C/14-3-3 complex is nuclear translocated and removal of inhibitory phosphorylation from CDK1 is inhibited.

This checkpoint is still relatively under characterised in comparison to G1 checkpoint. However, the importance is demonstrated by the threshold nature of activation, with less

than 20 DSBs not inducing activation, therefore any factor which alters this fidelity can have dire consequences for the cell (Deckbar *et al.*, 2007).

1.6.4 Spindle Assembly checkpoint

The spindle assembly checkpoint (SAC) arrests cells during mitosis, predominantly due to incorrect attachment of spindle proteins to the securin complex. Following the entry into mitosis cell cycle, arrest is dependent upon the spindle-assembly checkpoint (SAC). This is predominantly triggered if the spindle proteins fail to attach correctly to securin on the chromosomes (Lara-Gonzalez, Westhorpe and Taylor, 2012). Following incorrect spindle attachment, a complex containing BUBR1, BUB3, CDC20, and MAD2, which is referred to as the mitotic checkpoint complex (MCC), forms and inhibits the activity of APC (Sudakin, Chan and Yen, 2001). This complex also mediates the binding of PLK1 to CDC20, which induces an inhibitory phosphorylation preventing progression through mitosis (Jia, Li and Yu, 2016). The failure to form the MCC can result in multipolar separation events which lead to random chromosome distribution across all daughter cells (Gisselsson *et al.*, 2008).

Activation of these cell cycle checkpoints is paramount to resolving the cellular damage caused by toxic insult or preventing proliferation in sub-optimal conditions. Without the function of checkpoint major pathogenic consequences such as DNA mutation, chromosomal instability, and organelle damage can be observed. However, the failure to release cells from checkpoints can be equally pathogenic with unwanted apoptosis being a common phenotype.

1.7 Checkpoint activation recovery

The activation of cell cycle checkpoints is driven by the activity of kinases, therefore in order to reverse this activity phosphatases play a major role. The reversal of phosphorylation at serine 15 of p53 is vital for restoring correct cell cycling following G1 checkpoint arrest. It has been postulated that three candidate phosphatases, in particular, are involved in the removal of this phosphorylation, these being: PP4, PTPRN2, and PTPN6 (Lee *et al.*, 2012). In particular, PP4 is required, as it

dephosphorylates the p53 inhibitory protein kruppel-associated box domain-containing protein 1, which in turn inhibits the transcription of p21 (Shaltiel *et al.*, 2014).

The claspin protein is required during S-phase to mediate DNA replication. The degradation of claspin is required upon entry into mitosis for cell cycle progression. Following DNA damage, the claspin is stabilised by CHK1 phosphorylation preventing cell cycle progression. The removal of CHK1 dependent phosphorylation from claspin by PLK1 triggered processes is required for checkpoint recovery (Mamely *et al.*, 2006). This activity of PLK1 is promoted by Aurora A/bora complex activation during checkpoint reactivation (Macûrek *et al.*, 2008). In addition, the activated pATM (S1981) is recovered to resting state by the activity of PP1 and WIP1 (Shreeram *et al.*, 2006; Peng *et al.*, 2010; Shaltiel *et al.*, 2015).

The activity of wild-type p53 induced phosphatase 1 (WIP1) works in opposition to p53 induced signalling on MPF components and releases cells from the G2/M checkpoint (Lindqvist *et al.*, 2009; Shaltiel *et al.*, 2014). The importance of CDC25 family proteins for correct progression through the G2/M checkpoint has been highlighted in previous sections. While CDC25A predominantly regulates G1/S-phase transition and CDC25C regulates the G2/M transition, CDC25B has a particular impact upon the number of cells which re-enter cell cycle following damage (Bansal and Lazo, 2007). This function of CDC25B is corroborated by the finding that it is an initiator of the positive feedback loop between CDC25C and CDK1/cyclin B1 complex (Cans, Ducommun and Baldin, 1999).

1.8 WNT Signalling pathway

The detection of mitogenic signalling molecules is a key step during the induction of the cell cycle. Multiple families of extracellular signalling molecules can induce cell division predominantly by controlling the activity of cyclin D1/CDK4 (Lukas, Bartkova and Bartek, 1996). However, while partially redundant the nuances of cellular induction by each mitogen signalling cascade are distinct. One such pathway is wingless and INT-1 (WNT) signalling which is separated into three major sub-pathways: canonical WNT, Calcium-dependent WNT, and Planar WNT. WNT signalling was originally coined as a combination of the Wingless gene in *Drosophila* and the mouse gene INT-1, which were identified as homologs (Nüsslein-volhard and Wieschaus, 1980; Nusse and Varmus,

1982). The pathway is highly conserved in humans, mice, *Drosophila*, and *Xenopus* (Cadigan and Nusse, 1997; Clevers, 2006). However, while similar in function the pathway has become evolutionarily divergent in nematodes (Mizumoto and Sawa, 2007).

The canonical WNT signalling pathway is an autocrine/paracrine signalling pathway, which initially requires the secretion of WNT ligands into the extracellular matrix. The WLS protein, acts to transport WNT ligands to the cell surface (Yu *et al.*, 2014). There are currently 19 known mammalian WNT ligands which are secreted from cells (MacDonald, Tamai and He, 2009). Once secreted, these ligands will predominantly interact with the GTPase coupled receptors of the Frizzled (Fz) and lipoprotein receptor-related protein 5/6 (LRP5/6) family (MacDonald and He, 2012). In the absence of extra-cellular ligand receptor binding the WNT pathway is maintained in an inactive state.

The primary gene targets of canonical WNT signalling are members of the TCF/LEF gene family. The transcriptional activity of TCF/LEF genes is predominantly regulated by the binding of β -catenin to promoter regions, therefore, regulation of β -catenin activity is critical to prevent incorrect transcription (Cadigan and Waterman, 2012). The activity of β -catenin is controlled by three regulatory steps, protein degradation, nuclear/cytoplasmic shuttling, and association with inhibitory proteins. The first of these steps is the targeting of β -catenin for degradation which is primarily controlled by two kinases, glycogen synthase kinase 3 β (GSK3 β) and Casein kinase 1 (CK1) (Peifer, Pai and Casey, 1994; Liu *et al.*, 2002). Following phosphorylation, β -catenin is ubiquitinated and degraded by the proteasome, primarily through interaction with the ubiquitin ligase β -TrCP (Hart *et al.*, 1999; Kitagawa *et al.*, 1999; Latres, Chiaur and Pagano, 1999). The combination of these proteins in complex with Axin and protein phosphatase 2A (PP2A) is referred to as the “Destruction complex” (Stamos and Weis, 2013). Upon the binding of WNT signalling molecules to the frizzled receptor an intracellular Dishevelled (Dvl) binding site is exposed, comprising of three discontinuous motifs (Tauriello *et al.*, 2012). The frizzled receptor bound Dvl stabilises axin and β -catenin at these sites, preventing phosphorylation and subsequent degradation by the destruction complex (Cliffe, Hamada and Bienz, 2003). The non-degraded β -catenin can undergo nuclear translocation to initiate transcription of TCF/LEF genes. The nuclear translocation is predominantly promoted by TCF4 and BCL9, however, APC may also

have a role in this process (Cadigan and Waterman, 2012). While in the majority of conditions the increased retention of β -catenin in the nucleus increases transcription of TCF/LEF mRNAs the association with inhibitory proteins can further inhibit pathway activation. The ICAT protein is a binding partner of β -catenin and can directly inhibit the DNA interaction of β -catenin/TCF complex, preventing transcriptional activation (Daniels and Weis, 2002). The interaction between the two proteins is mediated through the armadillo repeat domain in β -catenin (Tago *et al.*, 2000; Graham *et al.*, 2002) (Figure 1.15).

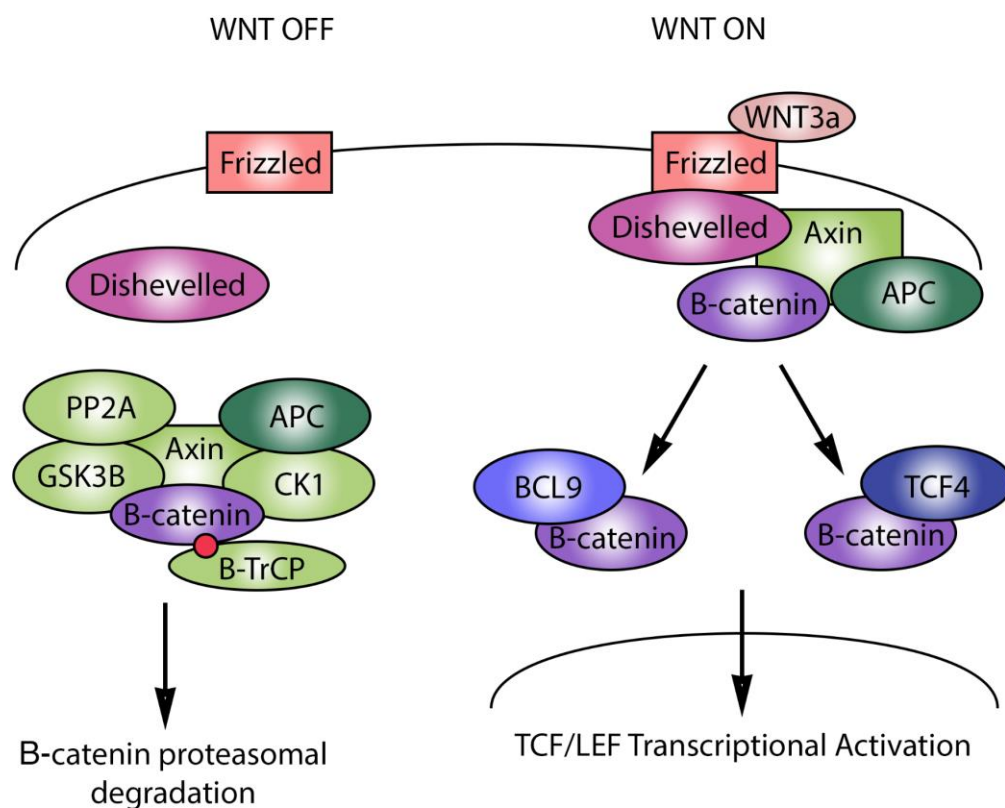


Figure 1.15. Canonical WNT signalling

Recognition of extracellular WNT signalling molecules by the frizzled receptor induces intracellular structural changes, which permit the association of dishevelled. The activated dishevelled cascade will prevent the degradation of β -catenin. The stabilised β -catenin undergoes nuclear translocation and activates transcription of genes from the TCF/LEF family.

The calcium-dependent WNT signalling pathway shares a common induction step with the canonical pathway. Upon binding of WNT ligands to a Fz receptor at the cell surface, Dvl is recruited preventing destruction complex mediated degradation of β -catenin.

However, during calcium-dependent WNT signalling there is an increase in intracellular calcium levels in a Fz receptor-dependent manner, yet independent of β -catenin, which induces protein kinase C (PKC) activity (Kühl *et al.*, 2000). The increased levels of intracellular calcium can activate proteins such as calcinurin and nuclear factor associated with T-cells (NFAT) (Hogan *et al.*, 2003).

The Planar WNT pathway has a role in regulating cellular polarity and is most active during the definition of mesenchymal to epithelial transition (Yang and Mlodzik, 2015). This pathway requires the association of Fz receptors and intracellular Dvl interaction with binding partners Vang and Flamingo (Gray, Roszko and Solnica-Krezel, 2011). The interaction of these receptors and proteins act to transfer signalling across a cell layer and directing planar development of polarity. The involvement of WNT signalling during induction of cell growth and differentiation leads to common dysregulation of this pathway during tumorigenesis. An example of which is APC, which is a potent tumour suppressor and expression loss has been linked to the development of cancer (Fodde, 2002).

Given the importance of all three WNT signalling pathways for regulation of cell proliferation and polarity, the characterisation of novel pathway factors could yield important insight into these complicated processes. A proteomic screen conducted by Craxton *et al.* identified leucine zipper and ICAT domain containing (LZIC) as an interacting partner of DNA-PKcs, PAXX, and XLF (Craxton *et al.*, 2018). LZIC is a putative member of the canonical WNT signalling family that was initially identified in a genetic screen for sequences sharing homology to the β -catenin protein inhibitor, ICAT (Katoh, 2001). The cellular function of LZIC is unknown, however, a morpholino induced knockdown of LZIC expression lead to failure of midline brain development in Zebrafish (Clements and Kimelman, 2005). In addition, the expression of LZIC has been linked to IR induced osteosarcoma development and gastric cancer outcome (Katoh, 2001; Daino *et al.*, 2009). Further investigation of LZIC interaction with NHEJ components by immunoprecipitation and western blot was unsuccessful (Unpublished data). Therefore, the role of LZIC within the cellular response to IR is seemingly not reliant upon interaction with the NHEJ complex. One method by which to conduct analysis of protein function is to produce a gene knock-out (KO) cell line. This can be

achieved by multiple techniques, such as, clustered regularly interspaced short palindromic repeats (CRISPR).

1.9 Clustered regularly interspaced short palindromic repeats (CRISPR)

The study of protein function regularly reduces protein expression by the use of synthetic silencing RNA (siRNA). This method, while highly effective, does not achieve a total loss of protein expression levels. In contrast, clustered regularly interspaced short palindromic repeats (CRISPR) is a technique that allows specific editing of the genome through either the removal or addition of sequence and can yield total protein knock-out (KO) or exclusively mutant protein expression. The system utilises the CAS9 protein isolated from bacteria, to cut DNA either as a double-stranded endonuclease or as a single-stranded nickase. Originally this pathway was utilised by multiple species of bacteria as a mechanism to degrade bacteriophages (Jinek *et al.*, 2012). The system utilises a 20 nucleotide guide RNA which is specific to a region of DNA, this then acts as a scaffold for tracerRNA and the CAS9 to be recruited. The site must be located adjacent to a protospacer adjacent motif (PAM), with the canonical sequence being 5'-NGG'-3'. Once the DNA has been cut the break can be repaired by the DSB repair pathways. If repaired by NHEJ the repair of these breaks can lead to the insertion of small indels within the gene which can frameshift the coding sequence and cause loss of protein expression. A second method which has been used to remove genomic regions is the incorporation of a construct which is inserted by homologous regions around the break site. A major limiting factor for the utilisation of the CRISPR technology was the requirement for a specific PAM sequence. Through mutation, many different CAS9 variants have now been established each with specific PAM configurations which have majorly improved the number of genomic regions that can be targeted. The xCas9 has a wide PAM range, able to recognise NG, GAA, and GAT, increasing the applicability of this enzyme (Hu *et al.*, 2018).

In addition to the benefits provided by this technology for the fields of cellular and molecular biology, CRISPR technology is also being utilised to combat human viruses and cancer. An example of the use of CRISPR to combat viruses that causes persistent infection is the treatment of human papillomavirus (HPV). Certain subgroups of the HPV family are directly linked to the formation of cervical cancer (White, Pagano and Khalili,

2014). CRISPR was used in an attempt to mutate the HPV genome in transformed cell lines and showed high efficacy and preventing HPV tumorigenesis in mice models (Zhen *et al.*, 2014).

1.10 Hypothesis and Objectives

The role of LZIC in development of IR induced osteosarcoma suggests a function within cellular response to IR. However, as the validation of interaction with the NHEJ complex was unsuccessful, LZIC is hypothesised to have alternative activities to direct regulation of DNA repair. With these findings in mind this thesis will be centred on characterising the function of LZIC both within normal cellular processes and during IR response. This will be conducted with three broad objectives. The first objective will be to characterise the LZIC protein interactome in an attempt to generate hypotheses about function. Secondly, the importance for transcriptome regulation of WNT signalling suggests a potential impact of LZIC loss upon this process. Therefore, the transcriptional changes in response to LZIC loss in both untreated and IR treated conditions has been analysed. Finally, the deregulated pathways identified within the initial proteomics and transcriptome profiling have been characterised by utilisation of cell biology and molecular biology techniques.

Chapter 2

Methods

2.1 Conservation analysis of LZIC sequences and phylogenetic tree generation

All available sequences for LZIC were retrieved from the National centre for biotechnology information (NCBI) database. The sequences were then aligned using ClustalW within MacVector software. The conservation score was given dependent upon the number of bases which matched between each sequence. The phylogenetic tree was generated in MacVector.

2.2 Predictive structure generation

The online software PHYRE2 was used in intensive modelling mode to generate the predicted structure of LZIC domains in all isoforms (Kelley *et al.*, 2015).

2.3 Biogrid and String network analysis

Biogrid database was interrogated to determine previous LZIC interactors (Oughtred *et al.*, 2019). The String network analysis tool was used to determine interaction and clustering to specific gene ontology terms of both Biogrid determined LZIC interactors and significant gene lists from both mass spectrometry analysis and microarray analysis (Szklarczyk *et al.*, 2019). The Gene ontology (GO) terms were visualised with REVIGO, utilising the SIMREL to compress GO terms (Supek *et al.*, 2011).

2.4 Bacterial culture and transformation

Transformation of bacteria was conducted as per the manufacturer's instructions (NEB-5 α , New England Biosciences). In brief, the backbone and cloned amplicon were mixed at a ratio of 1:3 molar concentrations with Quick ligase (New England Bioscience) for 5 minutes at room temperature. Ligation reaction was mixed with NEB-5 α and incubated on ice for 30 minutes. The bacterial suspension was heated at 42°C for 45 seconds, before cooling on ice for a further 5 minutes. SOC media was added and bacteria were incubated at 37°C for 60 minutes. Bacteria were spread on Ampicillin (100 μ g/ml) or Kanamycin (50 μ g/ml) supplemented agar plate, depending on resistance and incubated overnight at 37°C.

For plasmid expansion and cloning, the bacterial strain, Neb-5 α (New England Bioscience), was cultured in the desired volume of Luria Bertani (LB) media containing Ampicillin (100 μ g/ml) at 37°C in a shaking incubator.

For expression of GST-Tagged proteins the bacterial strain, Rosetta2 (New England Bioscience) was used. This strain was transformed using the protocol shown above and grown in LB media containing Ampicillin (50 μ g/ml) and Chloramphenicol (17 μ g/ml).

2.5 Expression and Purification of GST tagged LZIC proteins

Previous cloning of LZIC tagged with GST was conducted by Joanna Somers. I then conducted all following steps noted here. An antibiotic resistant colony was selected and expanded into 100ml of LB media containing ampicillin (50 μ g/ml) and chloramphenicol (17 μ g/ml). Following overnight incubation at 37 degrees the culture was expanded into 500ml LB media. The cultures OD600 was monitored and upon reaching OD600 = 0.5 the culture was induced with IPTG (1 mM). For expression of GST-LZIC were incubated overnight at 18 degrees.

To extract expressed protein, bacterial cell pellets were collected by centrifugation (6000g, 30 minutes, 4°C). Cell pellets were re-suspended in lysis buffer (50mM Tris-HCl pH 7.5, 150mM NaCl, 1% (v/v) Triton X-100, 0.1% (v/v) β -mercaptoethanol, 1mM PMSF, 1mM Benzamidine, 1x (w/v) protein inhibitor cocktail, 1mM EDTA, 1mM EGTA) and sonicated for 10 x 10 seconds. The resulting cell lysate was centrifuged (20,000g, 30 minutes, 4°C). Glutathione Sepharose beads (GE healthcare) were pre-washed with lysis buffer and the cell lysate added. Following a 1 hour incubation at room temperature on a rocker, the beads were collected by centrifugation (300g, 5 minutes, 4°C) and the supernatant removed. Beads were washed 3x with wash buffer I (50mM Tris-HCL pH 7.5, 500mM NaCl, 1% (v/v) Triton X-100, 0.1% (v/v) β -Mercaptoethanol, 1mM Benzamidine, 0.5x (w/v) PIC, 1mM EDTA, 1mM EGTA) at 4 degrees.

To elute proteins from the beads with GST tag intact, beads were further washed with wash buffer II (50mM Tris-HCl pH 8.0, 150mM NaCl, 1mM DTT, 1mM EDTA, 1mM EGTA) and 10mM reduced glutathione was used to elute. To cleave the GST tag, beads were re-suspended in cleavage buffer (50mM Tris-HCl pH 7.0, 150mM NaCl, 1mM

DTT) and Prescission protease was added to a concentration of 15U before incubating overnight at 4°C. Gel filtration (Superdex 75) was used to further purify both the GST-tagged and the cleaved version of WT LZIC.

2.6 Cell culture

HEK293H were cultured in Dulbecco's-modified Eagle's medium (DMEM), supplemented with 4.5g/l D-glucose, GlutaMAX (Life Technologies, Carlsbad, CA, USA) and 10% fetal bovine serum. LZIC constructs were transfected into HEK293H cells by lipofectamine LTX (Invitrogen) in Optimem (Gibco). Cells were incubated for 48 hours before analysis was conducted. CRISPR-derived HEK293H clones were cultured in the same medium as wild type cells. For conversion of HEK293H cells to freestyle growth media, cells were harvested by trypsinisation and following three PBS washes, incubated in serum free freestyle medium (Gibco) (37°C, 12% CO₂, 120 rpm agitation).

2.7 LZIC Knock-out Line Generation

To generate a CRISPR mediated LZIC knock-out cell line the manufacturer's instructions from the Origene kit were followed. Puromycin selection was added to the cells at a concentration of 0.5 µg/ml and 12 clones for both guide RNA 1 and 2 were isolated by colony selection and expanded for western blot analysis.

2.8 Cloning of Wild-type, N-terminal-Flag, and C-terminal-Flag tagged LZIC pBABE-Puromycin expression and viral production plasmid

HEK293 cells were cultured and harvested before total RNA was extracted with RNeasy kit (Qiagen). The total RNA was converted to cDNA by the use of reverse transcriptase, super script II (Thermo Fisher Scientific). Primers were used to amplify LZIC sequence, the flag sequence was contained within primer sequences, in addition to a restriction enzyme nuclease site (Table 2.1). The amplicons were treated with BamHI and SalI restriction enzymes (New England Bioscience) and separated by size using agarose gel electrophoresis. The region corresponding to 600-700 bp was excised and DNA extracted by gel purification kit (Qiagen) as per manufacturer's instructions. In parallel, pBABE-

puromycin (Addgene – 21836) was digested by BamHI and SalI and phosphatase treated as per the manufacturer's instructions (New England Bioscience). The resulting amplicons: LZIC WT, LZIC N-terminal flag, and LZIC C-terminal flag, and plasmid were mixed at a ratio of 3:1 with Quick ligase (New England Bioscience) for 5 minutes at room temperature. . The transformation protocol for NEB-5 α (New England Bioscience) using ligation mixes was conducted as stated in the manufacturer's instructions. The resulting colonies were selected and cultures prepared for extraction of plasmid by mini-prep. The restriction enzymes used for cloning of amplicons was used to screen amplified colonies for presence of insert. The positive colonies were amplified by maxiprep and submitted for sequencing. The flag-tagged versions of WT LZIC are expressed from pBABE-puro (Addgene – 21836) plasmid.

Table 2.1. LZIC cloning primers for pBABE insertion		
Construct variant	Forward Primer (5' – 3')	Reverse Primer (5' – 3')
LZIC WT	AAAAGGATCCGCCACCATG GCTTCC	AAAAGTCGACTCATTTTTTT GTTT
LZIC N-terminal Flag sequence	AAAAGGATCCGCCACCATG GACTACAAGGACGACGAC GACAAGGCAATGGCTTCCA GA	AAAAGTCGACTCATTTTTTT GTTT
LZIC C-terminal Flag sequence	AAAAGGATCCGCCACCATG GCTTCC	AAAAGTCGACTCATGCCTT GTCGTCGTCGTCCTTGTA TCTTTTTTTGT

2.9 Retroviral reintroduction of LZIC-Flag into CRISPR line

The transduction of HEK293 cells by live retroviral vector was conducted by Michal Malewicz. All subsequent steps noted here were conducted by myself. Removal of the puromycin resistance cassette from LZIC KO Clone 2 genome was achieved by transfection with constitutively expressed Cre-recombinase. After 48hrs the cells were plated at low density to generate single cell colonies. After 9 days of growth the single colonies were selected by cloning disk and expanded. Screening with puromycin sensitivity was subsequently conducted to confirm of cassette. Sensitive colonies were transduced with pBABE LZIC-Flag expressed from retroviral vector with puromycin

resistance cassette. The cells were incubated for 24hrs before exposure to 1ug/ml puromycin and clonally selected. The resistant clones were expanded and screened for LZIC-Flag expression.

2.10 Plasmid and DNA sequencing

LZIC KO clone 1 and 2 were cultured for 24hrs prior to harvesting. The cells were washed with PBS and trypsinised before quenching with DMEM. Cells were centrifuged (300g, 3 minutes, room temperature). Genomic DNA was extracted from cell pellet by use of DNeasy Blood & Tissue Kits (Qiagen). Amplification of region surrounding the CRISPR cut site was conducted by using specifically designed primers: Forward Primer – AGAATCATCACACATGGGCCT and Reverse Primer – AGAGATTAAAATGGCTTCCAGAGG. Amplicons were submitted for sanger sequencing. The resulting sequencing results were analysed with FinchTV and aligned with clustalW in Macvector against the NCBI sequence for LZIC isoform A.

2.11 Hypotonic fraction isolation and high salt treatment

Between 180 million and 200 million HEK293 freestyle cells stably expressing LZIC-c-flag were harvested by centrifugation (200g, 5 min, 4 degrees). The cell pellets were washed twice with ice-cold PBS-MC (PBS, 1 mM MgCl₂, 1 mM CaCl₂). The cells were gently resuspended in hypotonic buffer (10 mM Hepes pH 7.9, 10 mM KCl, 0.1 mM EDTA, 0.1 mM EGTA, 1x (w/v) protease inhibitor cocktail, 10 mM NaF, 1mM Na₃VO₄, 10 µM MG132, 1 mM DTT, 1 mM PMSF). Following 15 minutes on ice the cells were vortexed at full speed for 10 seconds and centrifuged (2300g, 5 mins). The supernatant was removed and the pellet washed with hypotonic buffer before being centrifuged (2300g, 5 mins). The resulting nuclear pellet was over layed with high salt buffer (20 mM Hepes pH 7.9, 420 mM NaCl, 1.5 mM MgCl₂, 20% (v/v) glycerol, 1x (w/v) protease inhibitor cocktail, 10 mM NaF, 1 mM Na₃VO₄, 10 µM MG132, 1 mM DTT, 1 mM PMSF) and left on ice for 15 minutes with gentle mixing to resuspend the pellet. Tubes were centrifuged (15,000g, 30 minutes, 4°C) and diluted with 2 volumes of Dilution buffer (20 mM Hepes pH 7.9, 20% (v/v) glycerol, 0.5% (v/v) Igepal CA630). Following a 30 minute incubation on ice, tubes were recentrifuged (15,000g, 30 minutes, 4°C) and all supernatant pooled. Equal volumes of nuclear extract were incubated with prewashed

anti-flag M2 agarose (Sigma Aldrich) and incubated overnight at 4°C. The supernatant was removed and beads were washed x5 with Wash Buffer (20mM Hepes pH 7.9, 140 mM NaCl, 0.5 mM MgCl₂, 20% (v/v) glycerol, 10 mM NaF, 1 mM Na₃VO₄, 10 µM MG132, 1 mM DTT, 1 mM PMSF, 0.5% (v/v) Igepal CA630). The proteins were eluted with 3X Flag peptide for 30 minutes at 4°C. Resulting elutant was used for Mass spectrometry.

2.12 Mass spectrometry analysis

The mass spectrometry analysis of LZIC-flag isolated proteins was conducted as in Craxton *et al.* 2015 (Craxton *et al.*, 2015). In brief, protein samples were separated by gel electrophoresis and sectioned before application of trypsin. All further sample processing steps were conducted in collaboration with the MRC toxicology proteomics department, however, data analysis was conducted independently. Peptides were extracted from the gel sections and lyophilised and subsequently resuspended in 10% acetonitrile (9:1) and 5% formic acid, with spiked massPREP standards (Waters corporation, Manchester, UK). UPLC separation of sample was achieved by sample injection into nanoAcquity UPLC system (Waters Corporation) with 25cm x 75µm I.D, 1.7 µm BEH130 C18 analytical reverse phase column for 90 minutes across a 3-40% acetonitrile gradient. Analysis of resulting peptide fractions was performed with Waters Synapt G2-S HDMS mass spectrometer. Data-independent acquisition and ion mobility (HDMSE) mode was used with peptide fragmentation achieved by cycling between 4eV and 20-50eV. Resulting mass spectra were processed with Waters ProteinLynx Global server version 3.0 (PLGS, Waters). The resulting peptides assignment were compared to uniprot data base before further analysis with Perseus software (Max Plank).

2.13 Microarray analysis of LZIC KO cells

The CRISPR control and LZIC KO clone 1 were seeded for untreated and 5Gy IR treatment. All clones were plated in duplicate for both conditions. After 24 hours, cells were exposed to 5Gy IR and incubated for a further 24hrs before harvesting. Untreated cells were harvested 48hrs post seeding. Cells were harvested using trypsin and EDTA before RNA extraction using RNeasy kit (Qiagen). The RNA quality was assessed by bioanalyzer, as per the manufacturer's instructions (Agilent). Samples were subsequently

labelled as per manufacturer's instructions in One-colour microarray-based gene expression analysis – low input quick amp labelling (Agilent Technologies). The chipset reference was G4858A, GE 8 x 60K with design 039494 V3. The Microarray chips were imaged by SureScan high-resolution imaging in DNA microarray scanner (Agilent technologies). Adapted from (Skalka *et al.*, 2019).

2.14 Microarray bioinformatics analysis

Analysis of resulting data from single colour microarray was performed by utilised the Limma pipeline in R (Ritchie *et al.*, 2015). In brief, the probe intensity values were log₂ transformed. The samples were background corrected by using the inbuilt Limma function – backgroundcorrect(X, method = normexp, offset = 16), where X is dataset. Normalisation between the arrays was then conducted by the function - normaliseBetweenArrays(X, method = quantile), where X is dataset. Average intensities were generated by collapsing each set of conditional replicates. A linear model was generated for each probe set within the data set by function lmFit. Determination of differential expression was conducted by comparison of each condition to linear models defined in the previous step by function contrasts.fit. Adapted from (Skalka *et al.*, 2019).

2.15 Gene set enrichment analysis

Gene set enrichment for detected mRNA changes was conducted by comparison of genes sets to the molecular signatures database by the Broad Institute (Mootha *et al.*, 2003; Subramanian *et al.*, 2005). In brief, the analysis was conducted by ranking all genes within a condition by the T-statistic. The ranked gene sets were subsequently compared to gene sets for hall mark pathways, as released with the GSEA tool using default settings. Adapted from (Skalka *et al.*, 2019).

2.16 Quantitative polymerase chain reaction

HEK293 cells and LZIC KO clones were exposed to 5 Gy IR and following 24 hr, RNA was extracted using RNeasy kit (Qiagen kit). 1000 ng of extracted total RNA was reverse transcribed to cDNA by Superscript II (Thermo Fisher Scientific) as per the manufacturer's instructions. The qPCR was conducted using SYBR green reagent (Applied Biosystems, ThermoFisher) and plates were analysed on Quantstudio 6 flex

(Applied Biotechnologies). Delta-delta CT calculation was conducted using GAPDH as a reference gene. Primer sequences are shown in table 2.1. Adapted from (Skalka *et al.*, 2019).

Table 2.2 qPCR primers		
Gene	Forward Primer (5' – 3')	Reverse Primer (5' – 3')
GAPDH	GGAGTCAACGGATTTGGTCGTA	GAATTTGCCATGGGTGGAAT
LZIC	AGTCTCTACAGACCTTGGCTC	ACAAGCTTCTGCACCATGTC
CCNB1	AACTTTGCCTGAGCCTATTTT	TTGGTCTGACTGCTTGCTCTT
SOX11	CGGTCAAGTGCGTGTTTCTG	CACTTTGGCGACGTTGTAGC
NREP	CTGTCTTTCTAGCATGTTGCCC	CCAGGGAGACCAACAGACAA
FLNA	GTCACAGTGTCAATCGGAGGT	TGCACGTCACTTTGCCTTTG
POU3F2	TTGTGTTGCCCCCTTCTTCGT	TTGCCTTCGATAAAGCGGGT
CPNE7	CACCCTGGGGCAGATTGTG	TCACCGTGATGGTGGACTTG
SFN	CGCTGTTCTTGCTCCAAAGG	ATGACCAGTGGTTAGGTGCG
LGALS3	GGGCCACTGATTGTGCCTTA	TCACCGTGCCCAGAATTGTT
IFI30	TACGGAAACGCACAGGAACA	CAGGCCTCCACCTTGTTGAA

2.17 Cell Cycle profile analysis

HEK293 cells and LZIC KO clones were treated with 5Gy IR, camptothecin (20 μ M), cobalt chloride (200 μ M), or UV (20 mJ) and incubated for 24 hours before being harvested. After washing with PBS, ice cold 70% ethanol was slowly added under slight agitation. Cells were left at 4°C for 24hrs to fix, PBS washed and Propidium Iodide and RNase A were added to final concentration of 10 μ g/ml and 100 μ g/ml, respectively. Samples were heated to 37°C for 30 minutes and then incubated at 4°C for at least 4 hours before reading. Flow cytometry analysis of cells was conducted on a BD biosciences FACS canto. In experiments ultising okadaic acid treatment all conditions were kept the

same as shown above, however, 1hr prior to IR exposure cells were treated with 1nM okadaic acid (Tocris biotechnique). Adapted from (Skalka *et al.*, 2019).

Cell cycle analysis of early G2/M checkpoint was conducted as in Xu *et al.* (Xu *et al.*, 2002). An additional sample was included which was treated with ATM inhibitor (10 μ M final concentration, Sigma Aldrich) 1 hr prior to IR exposure. Cells were stained with phospho-serine 10 H3 (Cell signalling, 1/100) and secondary antibody goat-anti-rabbit 488 (Abcam, 1/500). Attune NXT (Life technologies) was used to analyse stained cells. Adapted from (Skalka *et al.*, 2019).

2.18 Immunofluorescence

Parental HEK293 and LZIC KO clones were seeded and treated with 2Gy IR. The cells were incubated for 24hrs before supernatant was removed and washing with PBS. 4% Paraformaldehyde was used to fix cells for 10 minutes at room temperature before treatment with blocking buffer (0.3% triton X-100 in PBS supplemented with 5% goat serum). Fixed cells were treated with primary antibody overnight at 4°C. Cells were washed 3x with PBS before addition of secondary antibody and incubated at room temperature for 1hr. Cells were mounted with hard set mounting medium (vector hard set mounting medium, Vector labs). Images were acquired using a Zeiss LSM 510. Primary antibody - Phospho-serine 10 Histone 3 antibody (cell signalling technology, 1/1000). Secondary antibody – (Thermo Fisher scientific). Adapted from (Skalka *et al.*, 2019).

2.19 Western Blot

Cells were extracted for western blot analysis by addition of RIPA buffer (150mM NaCl, 50mM Tris pH 7.5, 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS, 1mM DTT, 0.4mM PMSF, Protease inhibitor cocktail) and sonicated for 2 x 10 seconds. Samples were diluted with lamelli buffer to a final concentration of 1x (v/v) before being heated to 90°C for 10 minutes. Proteins were separated by SDS-Page gel electrophoresis and transferred to nitrocellulose membrane by use of Bio-Rad Transblot Turbo. The antibodies used were: LZIC (Bethyl, 1/1000), Tubulin (Sigma Aldrich, 1/5000), pCHK2 T68 (Cell signalling, 1/2000), CHK2 total (Bethyl, 1/2000), pCHK1 S345 (Bethyl, 1/2000), pCHK1 S317 (Bethyl, 1/2000), CHK1 total (Bethyl, 1/2000), pATR Tyr1981

(Cell signalling, 1/1000), ATR total (Cell signalling, 1/1000), Cyclin B1 (cell signalling, 1/2000), pCyclin B1 Ser147 (Cell signalling, 1/2000), CDC2 total (Cell signalling, 1/2000), pCDC2 Tyr15 (Cell signalling, 1/2000). Horseradish peroxidase conjugated secondary antibody (Thermo Fisher), specific to the species of each primary antibody was visualised by ECL. PP2A – subunit A (Cell signalling, 1/1000), PP2A – subunit B (Cell signalling, 1/1000), PP2A – subunit C (Cell signalling, 1/1000), PP1 (Santa Cruz, 1/1000), pATM Serine 1981 (Cell signalling, 1/1000), ATM total (Cell signalling, 1/1000), Vinculin (Abcam, 1/5000), p53 total (Santa Cruz, 1/1000), and p-p53 Serine 15 (Cell signalling, 1/1000), were instead visualised by LICOR compatible secondary antibodies with conjugated fluorophore. Adapted from (Skalka *et al.*, 2019).

2.20 WST-1 Assay

Parental HEK293 and LZIC KO clones were seeded and treated with either: Mitomycin C (MMC), Camptothecin, Methyl Methanesulfonate (MMS), Cobalt Chloride (CC), and IR. Following a 24hr exposure to these toxins, WST-1 (Sigma-Aldrich) was added to wells and incubated for 2hrs. Plates were shaken to mix and read on a (Biotek Powerwave XS2) at an absorbance of 450nm with a reference absorbance of 600nm. The cell viability was standardised to vehicle control treated cells. Adapted from (Skalka *et al.*, 2019).

2.21 Metaphase spread analysis

Parental HEK293 and LZIC KO clones were seeded and treated with 2Gy IR and incubated for 48hrs. Cells were harvested by trypsinisation and centrifuged at 300g 5 minutes before swelling buffer was added (75mM KCl). Cells were incubated for 10 minutes at room temperature before addition of fixative solution (methanol and acetic acid 3:1 ratio). Cells were centrifuged at 200g for 5 minutes and supernatant was removed. This step was repeated twice. Pellet was resuspended in fixative to give cell suspension and dropped from a height of 30cm onto slides (Superfrost plus, Thermo scientific). Slides were dried at room temperature for 2 minutes before steaming for 10 seconds. Slides were left in a humidity box overnight to dry. Cells were stained with DAPI (1/5000) diluted in PBS and then mounted. Images were acquired using a Zeiss LSM 510. Adapted from (Skalka *et al.*, 2019).

2.22 Kaplan Meier plot generation

The PROGgene V2 database was used to generate Kaplan Meier plots for LZIC expression in cancers (Goswami and Nakshatri, 2014). The overall survival of patients was analysed with no stratification apart from LZIC expression. Adapted from (Skalka *et al.*, 2019).

Chapter 3

Investigation of Leucine zipper and ICAT containing protein (LZIC) interactome

3.1 Introduction

3.1.1 Leucine zipper and ICAT containing (LZIC)

The WNT signalling cascade has multiple vital roles during the control of differentiation and cell growth. As discussed in section 1.8, the canonical WNT signalling cascade has been extensively characterised, however, the alternative pathways and many canonical factors still require further investigation to fully understand the breadth of regulation. Leucine zipper and ICAT containing (LZIC) protein was identified in a genetic screen for sequences sharing homology to the ICAT protein and as such was characterised as a putative member of the WNT signalling cascade (Katoh, 2001). The function of LZIC has not been established, however, morpholino induced knockdown of LZIC in Zebrafish leads to failure of midline brain development (Clements and Kimelman, 2005). In addition, loss of LZIC expression has been linked to IR induced osteosarcoma development and gastric cancer outcome (Katoh, 2001; Daino *et al.*, 2009).

The main splice-form of LZIC is 190 amino acids in length and contains two domains – an N-terminal coiled-coil and a C-terminal ICAT domain. Previous conservation analysis identified over 90% homology between human, mouse, and zebrafish LZIC (Cadigan and Nusse, 1997; Clevers, 2006). The long-alpha hairpin or coiled-coil domain has multiple functions but may facilitate DNA binding (Shao, 2000). The ICAT domain was initially postulated to bind β -catenin through armadillo repeats in a similar mechanism to ICAT protein (Section 1.8). However, the interaction between LZIC and β -catenin could not be demonstrated (Clements and Kimelman, 2005). While, β -catenin possesses the canonical armadillo repeat this domain structure is not unique and a family of armadillo repeat containing proteins has been identified (Graham *et al.*, 2002).

1.9 Armadillo repeat proteins

Proteins of the armadillo repeat-containing family are highly evolutionarily conserved and have important roles within embryogenesis and cellular differentiation. The protein, β -catenin possesses an armadillo repeat and is the quintessential member of this family, with the domain being named after the *Drosophila* homolog *Armadillo* (Peifer, Berg and Reynolds, 1994). This family includes multiple proteins with a variety of cellular

functions including, Kinesin-associated protein 3 (KAP3), which predominantly regulates trafficking of proteins along the actin cytoskeleton, however, it has also been shown to have roles in mitosis (Carpenter *et al.*, 2015). The domain itself is constituted of 42 residues forming three alpha-helices subunits, however, the number of these domains can vary, for example β -catenin contains seven (Huber, Nelson and Weis, 1997). In addition to the number of domains the sequence of the domain is also highly divergent and so identification of these domains in whole-genome studies has not been possible. A typical method of assessing protein function is to identify interacting partners which can be conducted by multiple “omic” methods, such as mass spectrometry.

3.1.2 Protein interactome analysis

Mass spectrometry (MS) measures the mass/charge (m/z) ratio of ions and can be used for both chemical and protein identification. Two major approaches are available for identification of protein, referred to as “top-down” and “bottom-up”. “Top-down” proteomics involves the MS analysis of protein samples without any prior protein fragmentation steps. In comparison, “bottom-up” approaches involve protease-mediated digestion of proteins and MS analysis of the resulting peptides which when compared to an *in silico* generated digestion can be used to assign protein identity (Zhang *et al.*, 2013). A major problem with using top-down proteomics is the reduced dynamic range of protein detection, which can pose multiple problems during discovery MS analysis. In comparison, bottom-up proteomics has a significantly increased protein dynamic range due to the fragmentation of protein with protease and the separation of the resulting peptides by high-pressure liquid chromatography (HPLC)(Churchwell *et al.*, 2005). The most commonly used protease for peptide fragmentation is trypsin which cleaves proteins at the carboxyl-group of lysine and arginine residues generating a library of peptides (Simpson, 2006). Following HPLC, in brief, the separated peptides are charged by ionisation and accelerated through a charged environment altering the path of flight relative to peptide charge. The resulting spectra are defined as m/z ratio, which can be compared to an *in silico* generated data set to assign detected peptides to proteins.

Quantification of proteins within a sample can be conducted by two predominant methods: stable isotope labelling with amino acids in cell culture (SILAC) and label-free quantification (LFQ). SILAC, utilise different molecular weight isotopes of amino acid

to investigate quantitatively the changes in protein levels or interaction characteristics within a cell across a time course or changing conditions (Lanucara and Eyers, 2011). In contrast, LFQ does not require the use of amino acid labelling and has been made possible by the development of tandem MS (MS/MS). MS/MS requires two rounds of m/z analysis with a further fragmentation step separating the two, this method allows very specific allocation of sequence for each peptide and quantification of peptide levels (Shalit *et al.*, 2015).

Following acquisition of peptide spectral counts from MS/MS the assignment of peptide signals to specific protein sequences must be conducted. This requires comparison of spectral counts to an *in silico* generated peptide database, one such database being the protein lynx global server (PLGS). Finally, the comparison and quality control of samples can be performed. Multiple software packages have been developed for this purpose, examples of which are MaxQuant and Perseus (Cox and Mann, 2008; Tyanova *et al.*, 2016).

3.1.3 Hypothesis and chapter aims

Sequence conservation analysis has been conducted on LZIC protein sequences from human, mouse, and zebrafish. In general, the sequence of WNT signalling protein family members is particularly high. Therefore, to assess the alignment of LZIC with this particular characteristic of the WNT signalling family, conservation analysis of LZIC will be extended to further species. In addition, the ICAT domain within the LZIC c-terminal region does not interact with β -catenin, this could suggest a non-canonical role for this domain. Therefore, interactome analysis utilising MS will be conducted to establish LZIC interacting partners.

3.2 Results

3.2.1 Evolutionary conservation of LZIC sequence

WNT signalling proteins are generally highly conserved through multiple taxonomic groups (Cadigan and Nusse, 1997; Clevers, 2006). The identification of an ICAT domain within LZIC sequence suggests it is a new member of the WNT signalling pathway. LZIC protein sequence is well conserved when comparing humans, mice, and zebrafish (Clements and Kimelman, 2005). During the intervening period since Clements *et al.* published their seminal work on LZIC in 2005, the number and quality of sequences available have increased exponentially. To extend the previous analysis sequences for human, mouse, zebrafish, *Xenopus*, nematode, and slime mould LZIC were compared. Nematode sequence was used as a negative control for conservation as the WNT signalling pathway significantly diverged in this species (Mizumoto and Sawa, 2007). A phylogenetic tree created from each of the species sequences shows slime mould (*Dictyostelium discoideum*) as the most distant point of the tree (Figure 3.1A).

The conservation of the total protein across the various species was then assessed. The overall conservation of LZIC sequence across all species analysed is over 50%, except for nematode sequence (Figure 3.1B). The highest level of sequence conservation observed was between human (*Homo sapiens*) and mouse (*Mus musculus*) with 94% homology. The conservation of LZIC from *Homo sapiens* to zebrafish (*Danio reiro*) is 85% in line with previous work (Clements and Kimelman, 2005). The most distant species from *Homo sapiens* used within this analysis was *Dictyostelium discoideum* in this species was 58%, despite a significant separation in evolutionary terms. In contrast, Nematodes sequence (*Caenorhabditis remanei*) shows conservation of only 38%.

Overall, LZIC is highly conserved although with reduced sequence identity when comparing nematode LZIC sequence to human as has been previously been demonstrated for other WNT signalling proteins (Mizumoto and Sawa, 2007). Conservation analysis was conducted using the canonical LZIC sequence referred to as isoform A, however, three further splice forms were previously identified (Katoh, 2001). The multiple isoforms of LZIC are not conserved in the species analysed and so the difference of these isoforms from the canonical sequence was analysed.

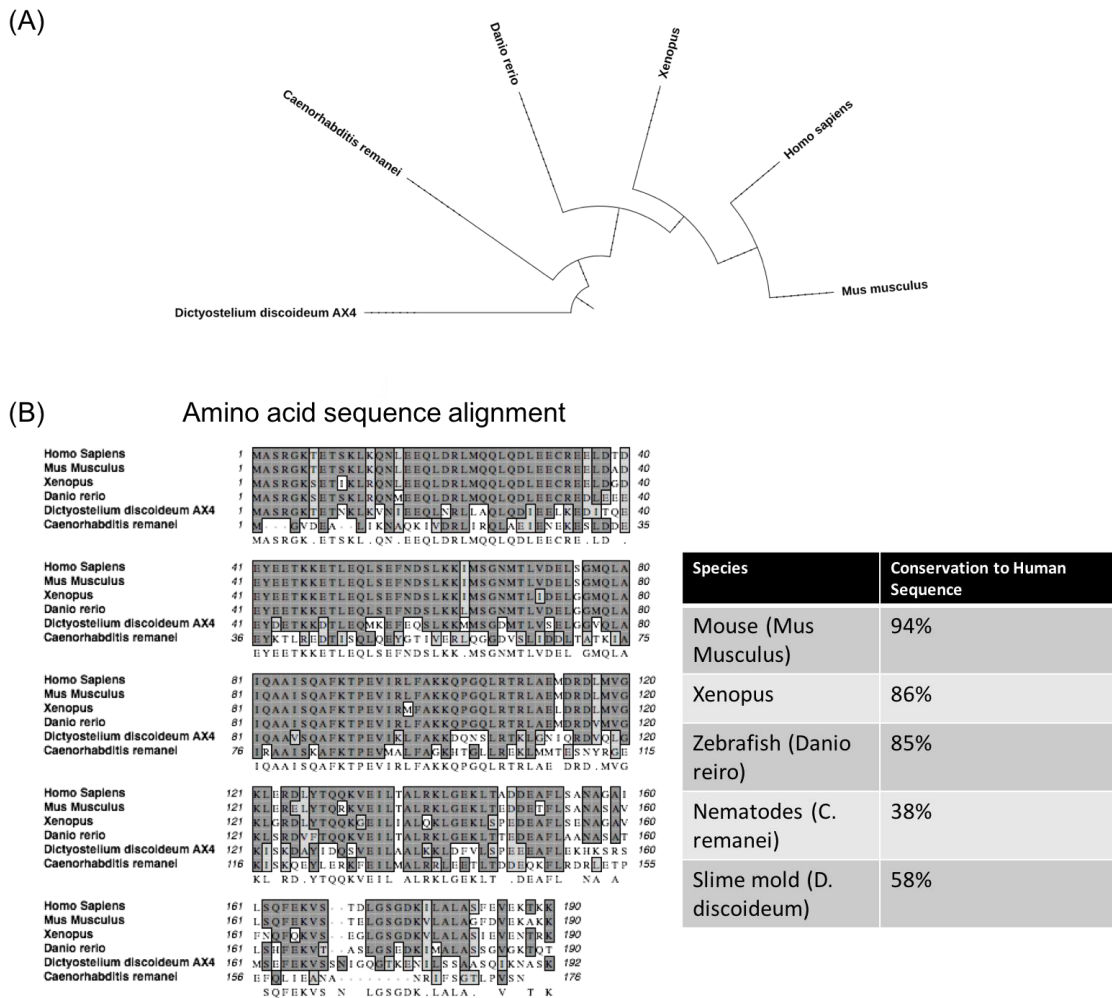


Figure 3.1 Evolutionary conservation of LZIC sequence.

(A) Phylogenetic tree showing the relationship between the LZIC amino acid sequences of Human, Mouse, Zebrafish, Xenopus, Nematode, and Slime mould. (B) LZIC amino sequences from multiple species was aligned by using Clustal W in Macvector, with the human sequence as the reference. The overall conservation value for each species was then determined by quantifying the number of conserved residues between human LZIC and the species of interest.

3.2.2 LZIC spliceforms comparison and domain characterisation

The alternative splicing of genes can have a significant impact on the functional characteristics of the eventual protein. Therefore, the impact on functional domains of LZIC following alternative splicing was investigated. LZIC isoform A has previously been shown to be the predominant form of LZIC (Katoh, 2001). Therefore, LZIC isoform A was used as the subject sequence for alignment. The alignment of other isoforms indicates multiple differences to isoform A (Figure 3.2A). Isoform B is longer by 21 amino acids due to an extended N-terminus. However, despite this extension, the rest of the sequence aligns with isoform A completely. Isoform C contains the same N-terminal sequence as LZIC isoform A, however, the C-terminus of the protein is altered. Isoform D is the most divergent from isoform A. Isoform D contains a 7 amino insertion between position 112 and 113 of isoform A. Additionally the C-terminus is substantially altered with the loss of 50 amino acids, significantly shortening the protein.

Functional allocation of domains to LZIC identified a leucine zipper domain at the N-terminus and an ICAT domain at the C-terminus (Katoh, 2001). Phyre2 is a method by which to analyse *in silico* the domain structure of an amino acid sequence (Kelley *et al.*, 2015). The certainty of structural modelling produced by Phyre2 is represented by two values, confidence and percentage identity. Confidence is the probability of true homology between subject and query sequence. Percentage identity indicates the actual sequence homology to the domain. Guideline values are given with confidence of >90% and %Id. value of >30% being an accurate predicted structure and allocation. The domain structure of each isoform shows some striking differences. As expected by previous data Isoform A contains a C-terminal ICAT domain and an N-terminal long alpha hairpin domain. The scores for the ICAT domain are 99.3 and 48% for confidence and %Id. respectively. However, the long-alpha hairpin domain has a high level of %Id. at 48% with a much lower confidence score of 35.5. This indicates that the structure is less stringent but the homology to that domain sequence is still appropriate to allocate that domain (Figure 3.2B). Interestingly the C-terminal alteration in isoform C has a direct impact upon the N-terminal when total protein folding is modelled the position of the long-alpha hairpin domain is shifted leading to a reduction of identity and confidence. This could indicate the loss of this domain in isoform C. As previously shown isoform D loses a substantial portion of its C-terminus through alternative splicing. This has a direct

impact upon the ICAT domain with a reduction to 88% and 16%, confidence and %Id. respectively. These values are reduced below the threshold for accurate modelling and so could represent a loss of this domain (Figure 3.2C). The structure for each domain represented by Phyre2 is built using the specific sequence input and so, therefore, gives a representation of the folding (Figure 3.2D).

The comparison of LZIC domain structure indicates significant differences between the splice forms. However, the isoforms have so far only been shown in humans and could indicate divergence of function between the isoforms. While the overall conservation of LZIC is high, the actual conservation of each domain independently could indicate particular importance for the protein functionality.

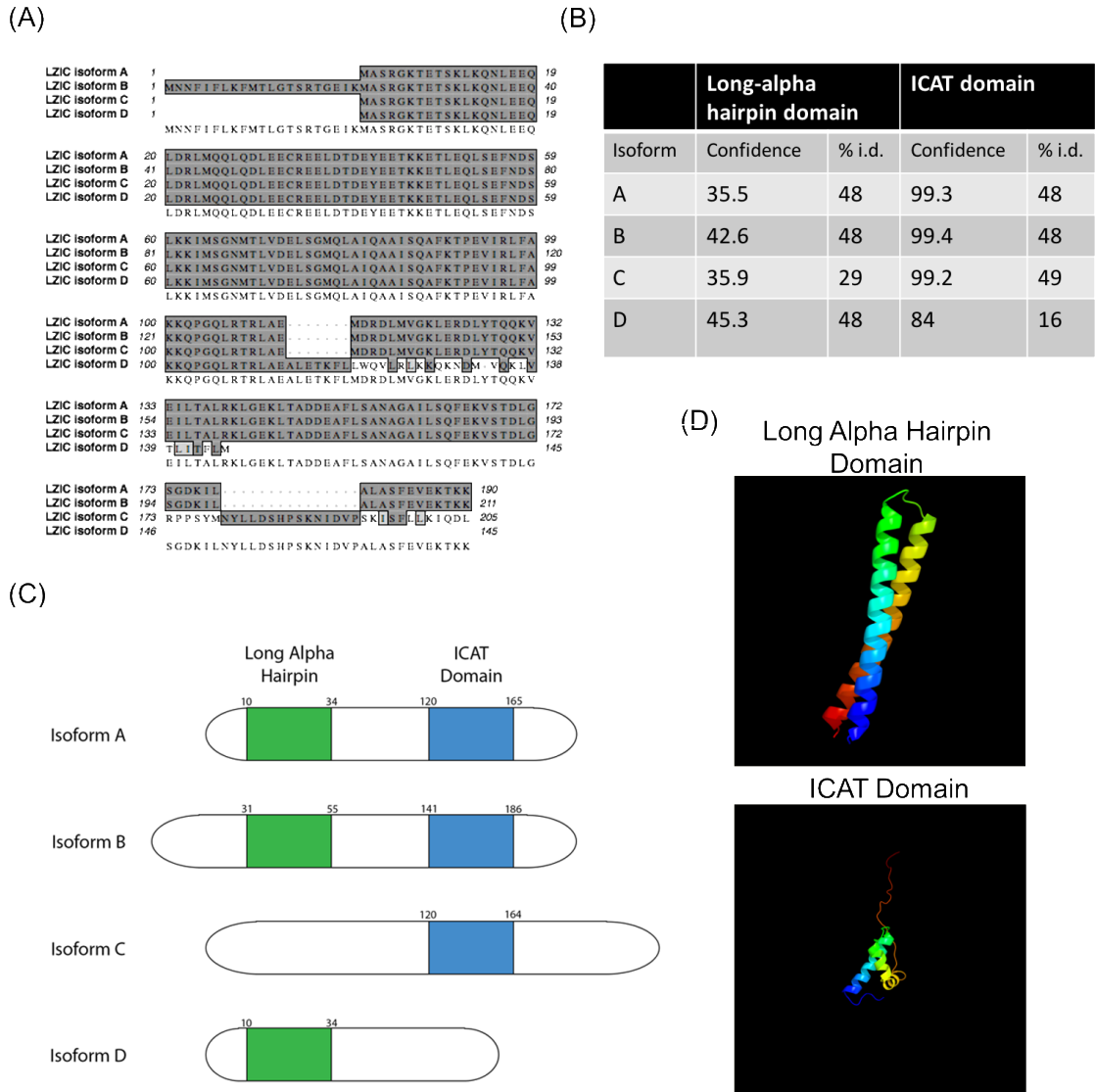


Figure 3.2 LZIC spliceforms comparison

(A) Clustal W alignment of LZIC amino acid sequence for each of the four spliceforms by use of MACvector, with the LZIC isoform A sequence as the reference. (B) Schematics of domain presence on LZIC isoforms. (C) Phyre2 analysis of LZIC isoform showing domain allocation. The confidence of domain represents a score of between 0-100 and is the probability of a true structural homology between the subject and query sequence. The % i.d. indicates the absolute number of residues conserved between the subject and query domain sequences. (D) Structural folding analysis of ICAT domain and long alpha hairpin produced by Phyre2 using *in silico* inferred folding.

3.2.3 Evolutionary conservation of LZIC functional domains

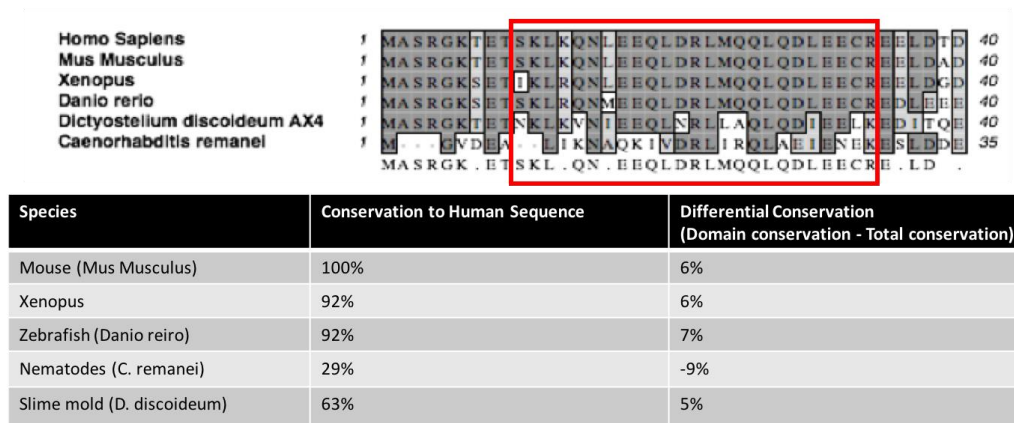
The specific conservation of each domain in isolation was determined by comparison of the long-alpha hairpin and ICAT domain region sequence from the species utilised previously (Figure 3.1A). To assess the conservation of each region in comparison to total protein, the total sequence conservation percentage of LZIC in each species was subtracted from the regional conservation percentage within each species. A positive number represents increased conservation and a negative value represents decreased conservation compared to total protein. Initially, the long-alpha hairpin domain was analysed, in all sequences, this was found between residue 10 and 34. The differential conservation values show that conservation is higher in the long-alpha hairpin domain compared to an average for the whole protein in all species apart from *Caenorhabditis remanei* which shows reduced conservation compared to whole protein (Figure 3.3A).

The conservation of the ICAT domain was next analysed. In comparison to the conservation seen in the long-alpha hairpin domain all species show a negative conservation score (Figure 3.3B). This represents that this domain is less well conserved when compared to total protein. For example, the conservation of this domain in mouse sequence shows a differential drop of 10% compared to an increase of 6% for the LAH domain.

Overall, this result may indicate that the N-terminal hairpin is more important for the conserved function of LZIC than the ICAT domain. In addition, while the ICAT domain is a characterised protein-protein interaction domain, there are indications that in addition to being a DNA interaction domain, the long-alpha hairpin can be utilised to mediate protein-protein interaction (Shao, 2000).

The conservation of LZIC sequence matches the profile of a canonical WNT signalling factor. However, the specific conservation of the sequence related to the ICAT domain indicates reducing conservation. In addition, LZIC does not bind to β -catenin (Katoh, 2001). This may indicate that the function of the LZIC ICAT domain is diverging which would be represented in an altered binding protein partner profile. This provided rationale to pursue characterisation of the LZIC interactome.

(A) Long Alpha hairpin conservation



(B) ICAT Domain

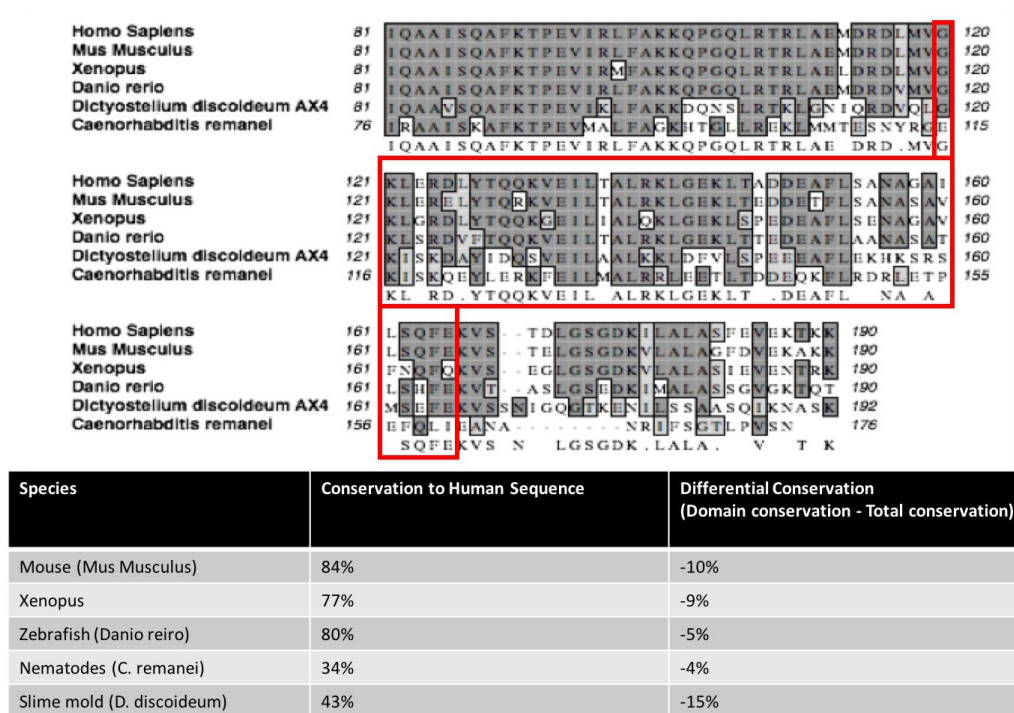


Figure 3.3 Evolutionary conservation of LZIC protein domains.

(A) Clustal W alignment of LZIC Long Alpha hairpin domain by use of MACvector.
 (B) Clustal W alignment of LZIC ICAT domain by use of MACvector. The differential conservation compares the total conservation of whole protein to the domain in order to determine specific conservation.

3.2.4 Meta-analysis of LZIC Biogrid interactors

While interactome analysis has not specifically been conducted for LZIC, it has been identified in a small number of high throughput proteomic screens (Fenner, Scannell and Prehn, 2010; Gupta *et al.*, 2015; Hein *et al.*, 2015; Wan *et al.*, 2015; Arumughan *et al.*, 2016; Huttlin *et al.*, 2017). To determine any functional link between the proteins identified across the studies a network analysis was conducted. The edges between nodes represent protein-protein interaction. In total 15 previous interacting partners were identified (Figure 3.4A). A small network is identified between LZIC, CDK8, CDK19, and cyclin C with these proteins being components of the mediator complex, which is an important regulator of RNA polymerase II activity at transcriptional loci (Soutourina, 2018). In addition, both GINS2 and GINS4 were identified, which in humans are referred to as PSF4 and SLD5 respectively, and are components of the replisome (MacNeill, 2010). However, the majority of proteins show no direct interaction. Therefore, the enrichment of specific functional groups using gene ontology (GO) terms was conducted. Four GO terms were identified which are enriched with the top 2 being related to cell cycle regulation (Figure 3.4B).

WNT signalling proteins as regulators of the cell cycle is well established (Section 1.8). To obtain a more comprehensive analysis of LZIC interactome specific analysis using LZIC as a bait protein is required.

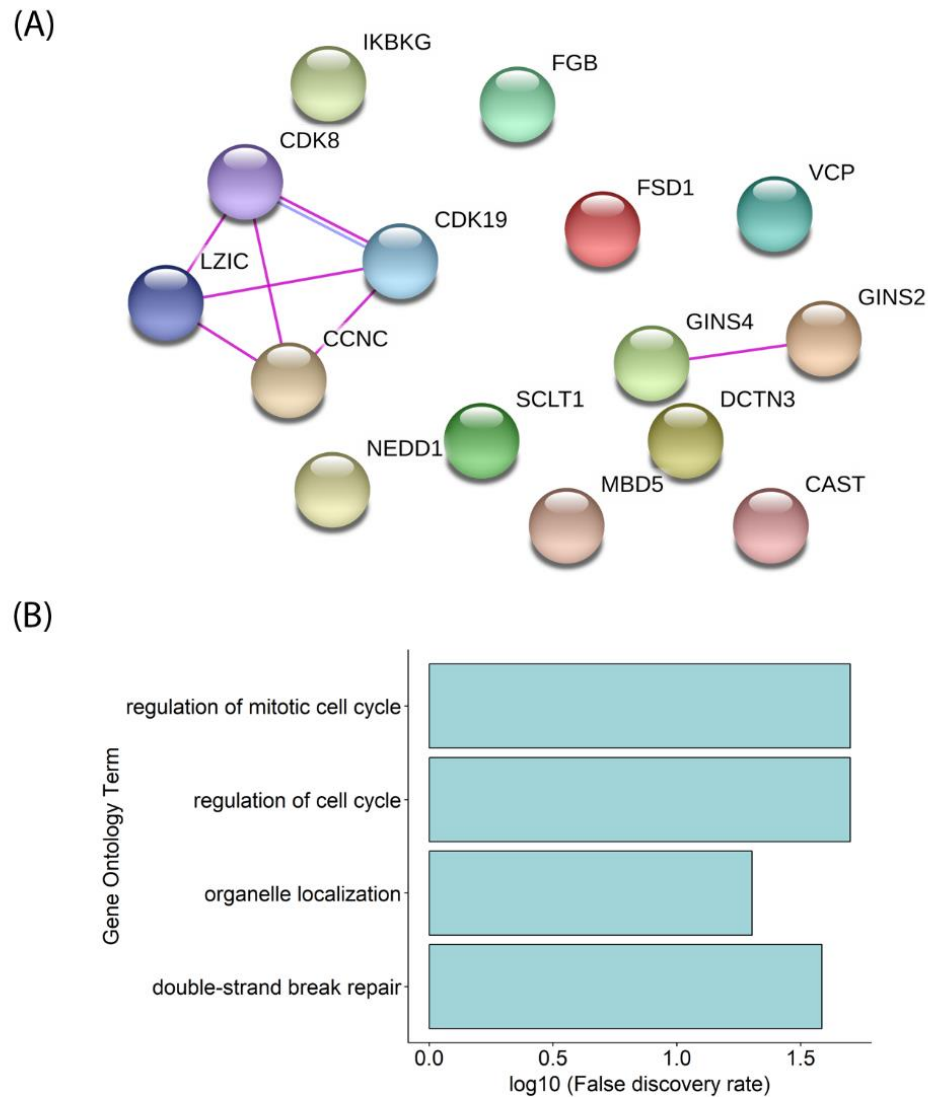


Figure 3.4 Network analysis of previously identified LZIC interactors.

(A) Network analysis conducted using STRING. Pink lines are representative of previous evidence of interaction between proteins and lilac are representative of protein homology. All proteins shown are currently available interacting partners of LZIC on the BioGrid interactome database. (B) gene ontology (GO) term analysis of all proteins represented in (A), all terms with fewer than 100 proteins were removed.

3.2.5 Assessment of recombinant LZIC-GST protein quality

Previous attempts to isolate endogenous LZIC protein have been unsuccessful. Affinity purification of endogenous LZIC was not possible due to non-specific binding of the commercially available antibodies. To address this expression of exogenous flag-tagged LZIC in HEK293 cells was attempted, however, this yields truncation of products and low expression of the full-length protein. Flag-tagging of protein is only one method by which to achieve isolation of proteins and interacting partners. The overexpression of protein within a mammalian system can lead to degradation due to nutrient depletion or lack of co-factor availability (Moriya, 2015). Expression of the protein in bacteria can avoid these issues. A common protein tag used to isolate proteins in bacteria is GST, the sequence for which can be added to the CDS for the protein of interest. It has also been indicated that GST tags can aid the folding of proteins through a chaperone-like function (Harper and Speicher, 2011). The LZIC CDS was recloned into a GST containing vector, pGEX and the plasmids sequenced to determine that no mutations had been incorporated during this process.

Recombinant LZIC-GST was produced in Rosetta2 bacteria and purified by size exclusion chromatography. The protein was expressed from a pGEX plasmid construct, with a lac operon promoter, allowing induction by the addition of IPTG. The bacteria were lysed and then a GST-mediated affinity purification conducted to isolate LZIC-GST protein with high stringency washes to remove bacterial protein binding partners. The purified protein was subsequently eluted from beads with reduced GST and further isolated by superdex 200 size exclusion chromatography column, with flow controlled by an AKTA gel filtration system. In tandem with the purification of LZIC-GST, pure GST protein was also purified to use as a binding control in all pull-down experiments. The resulting fractions were analysed by coomassie gel (Figure 3.5A). The purification yielded full-length protein. A small number of non-specific bands can be seen. These bands may represent a low-level of degradation products of LZIC that formed during the process of isolation.

To quantify the LZIC-GST and GST produced a BSA standard was analysed in tandem. This shows that the concentration of LZIC is between 0.25mg and 0.5mg/ml, with the

GST concentration being closer to 1 mg/ml (Figure 3.5B). This LZIC-GST was pure enough that it could then be used for the optimisation of the GST-Pull down protocol.

Overall, the optimisation of this process indicated that LZIC-GST can be produced in high concentration with minimal bacterial protein contamination. This provided an appropriate construct to conduct LZIC-GST pull-down experiments.

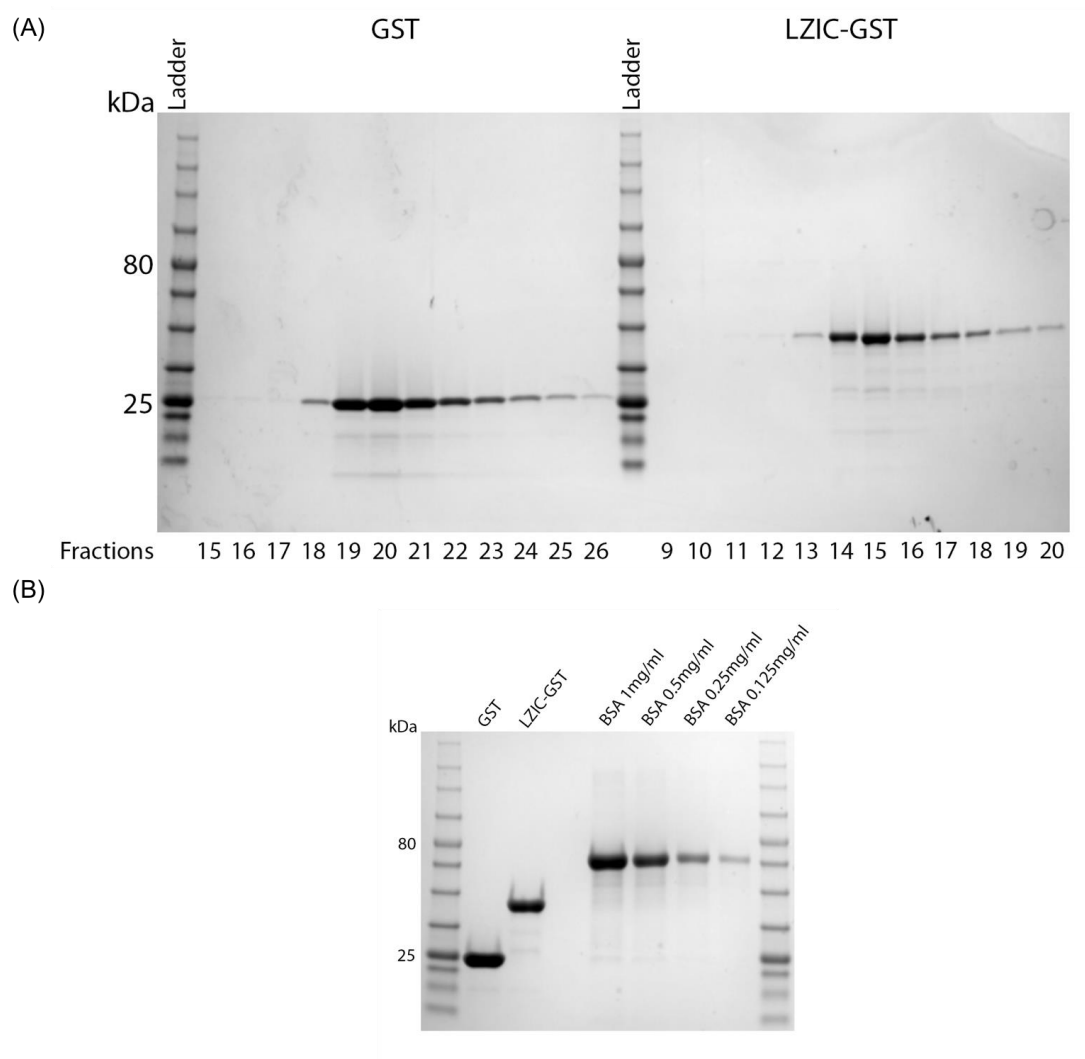


Figure 3.5 Assessment of recombinant LZIC-GST protein quality

(A) Rosetta2 protein expression E.coli previously transformed with pGEX LZIC-GST or pGEX GST were induced overnight at 15°C by IPTG. Rosetta2 were lysed and homogenized by sonication before isolation of expressed LZIC-GST and GST proteins by GST immunoprecipitation. Further purification of LZIC-GST and GST was achieved by gel filtration using Superdex 75 and AKTA system. Those fractions corresponding to peaks of protein concentration on UV trace were separated by gel electrophoresis and visualized with Coomassie blue. This indicates that the LZIC is expressed with no heat shock protein present and with few degradation products. GST was produced in tandem to act as a control in all experiments. (B) The fractions containing highest concentrations of LZIC-GST or GST were pooled and the resulting mix was compared to a Bovine Serum Albumin control standard.

3.2.6 Affinity purification of LZIC-GST

Following confirmation of full-length LZIC-GST expression in bacteria without high levels of truncation products, this construct was used to affinity purify interacting partners from mammalian cell lysate. The fractionation of cells was conducted as the predominant localisation of LZIC is in the nucleus (Uhlen *et al.*, 2015). HEK293 cells were grown in freestyle cultures and then fractionated to yield both the cytoplasmic fraction and the enriched nuclear fraction (Section 2.12). Fractionation will improve the background binding profile of MS analysis by removal of multiple non-specific interactors which can be found in the cytoplasm. The LZIC-GST was rebound to sepharose beads and incubated overnight with cell lysate. The elution of LZIC was conducted by using Precision protease, this cleaves the GST leaving it attached to the beads. This allows removal of the GST molecules before MS analysis. The success of the elution was checked by silver stain analysis. It can be seen that all GST has been removed from the control lane and the LZIC band has returned to 21kDa in size representing successful removal of the GST. However, the banding profile of each pull-down lane does not indicate any strong unique interactors (Figure 3.6). Due to the lack of interactor isolation, this construct was not used to conduct mass spectrometry analysis.

Due to the failure of interactor isolation utilising recombinant GST-tagged LZIC two further techniques were pursued to establish interactome. The first of these is a yeast-2-hybrid analysis of LZIC.

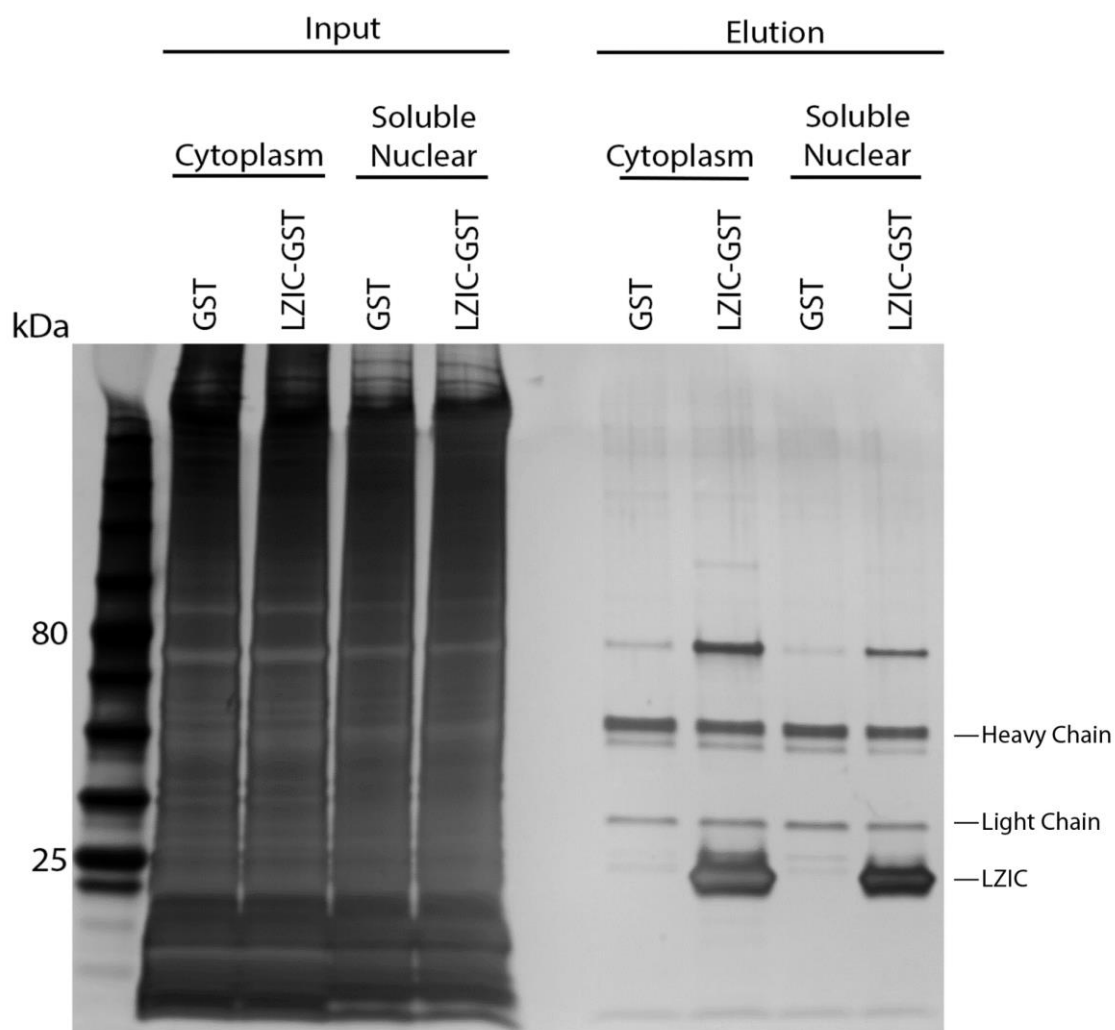


Figure 3.6 LZIC-GST protein interactor visualization

Purified GST and LZIC-GST were bound to glutathione sepharose beads by rotation at 4°C. In parallel HEK293H cells were lysed and fractionated into the enriched cytoplasmic and soluble nuclear compartments. The fractionated lysate for either cytoplasm or soluble nuclear were incubated at 4°C with LZIC-GST or GST bound glutathione beads. GST was used as a control pull-down to show proteins which non-specifically bind to the beads. Beads were washed and LZIC eluted by precision protease cleavage of the GST tag. The resulting eluate was diluted in lamelli buffer and boiled. The protein samples were separated by gel electrophoresis and visualized by silver stain. The LZIC-GST lanes show good presence of LZIC at ~21 kDa.

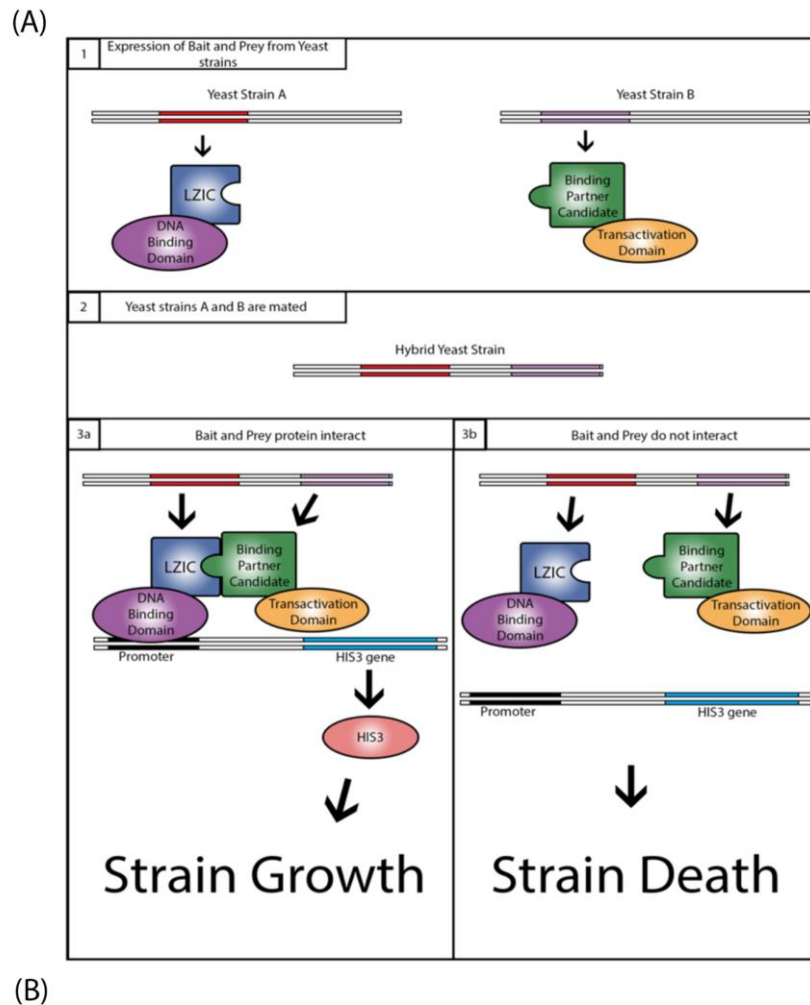
3.2.7 Determination of LZIC interactors by yeast-2-hybrid

The previous attempts to conduct interactome analysis utilising LZIC-GST were unsuccessful. Therefore, LZIC sequence was utilised for yeast-2-hybrid screening against the human placenta library (All yeast-two-hybrid data was generated by Hybrigenics). Yeast-2-hybrid analysis requires the ligation of transcription factor elements to two proteins of interest, one is ligated to the DNA binding domain and the other ligated to the transactivation domain. Upon interaction of proteins, the two domains collaborate to activate transcription of a selection gene, which can either confer a colony colour change or improve survival on amino-acid negative medium (Brückner *et al.*, 2009) (Figure 3.7A).

In total 6 proteins were identified which interact with LZIC (Table 3.1). The confidence of an interaction is given by a rating of A-F: A-C indicates a high confidence interaction, D is assigned to interactions which are either barely detectable or are possible false positives, E is a non-specific interaction, and F is an experimentally proven technical artefact. Therefore, Zinc finger 521 (ZNF521) was not investigated further, as it is an artefact. The highest confidence interactor, zinc finger 558 (ZNF558), is a zinc finger protein with no characterised biological function assigned to the protein, however, as a zinc finger protein, it is putatively assigned the capability to bind DNA and regulate transcription. Zinc finger (ZNF3) much like ZNF558 has no specifically investigated role, however, it has been detected in Chromatin-immunoprecipitation (CHIP) analysis of BTB and CNC homology (BACH1) which is a response element to low levels of cellular heme (Warnatz *et al.*, 2011). The three further interactors: Histone-arginine methyltransferase CARM1, Heat shock protein family A member 5 (HSPA5), and zinc finger 197 (ZNF197) were all weakly detected. CARM1 is assigned a D and therefore was a weak interactor. This protein is a methyl-transferase and has roles in the regulation of gene expression following β -catenin signalling and p53 signalling (An, Kim and Roeder, 2004; Ou *et al.*, 2011). HSPA5 also referred to as GRP78 or BiP, is associated with the endoplasmic reticulum membrane and regulates the unfolded protein response. Importantly, BiP has been associated with oncogenesis and response to a toxic insult (Y. W. Chang *et al.*, 2016; Wang *et al.*, 2017). Finally, ZNF197 has no previously prescribed function, however, a gene product that is produced from this locus, which is lacking the

22 zinc finger repeats, regulates the hypoxia-induced factor 1 alpha (HIF1 α) cascade through binding to von-Hippel-Lindau tumour suppressor (pVHL) (Li *et al.*, 2003).

Overall, the proteins which were detected within the yeast-2-hybrid screen have a role within the regulation of various transcriptional pathways, with zinc-finger proteins being canonical transcription factors. However, the yeast-2-hybrid identified a small number of proteins. The difficulty to achieve interaction of LZIC with binding partners suggested problems with folding or stability. Therefore, a more nuanced expression system was sought to address these concerns.



(B)

Gene Identified	Confidence (A – F)
Carm1	D
HSPA5	D
ZNF197	D
ZNF3	C
ZNF521	F
ZNF558	A

Figure 3.7 LZIC interactome determination by Yeast-2-hybrid.

Experiments conducted by Hybrigenics. To assess interaction of the protein of interest with multiple candidate proteins. A yeast line was produced which expresses LZIC linked to the DNA binding domain of LexA. In addition, a panel of yeast lines was created with potential interacting partners linked to the transactivation domain of Gal4. The mating of two yeast lines leads to co-expression of the two protein constructs. If interaction occurs then the transcription factor domains interact and drive expression of Gal4, leading to strain survival on selection media. If interaction does not occur, no Gal4 is expressed leading to death on selection media. The lettering system A-F is representative of confidence the identified protein is not a false positive. A-C are proteins which are identified with confidence, D is assigned to weak interactors which are possible false positives. E is a non-specific interactor, and F is a proven technical artefact.

3.2.8 Generation of LZIC knock-out HEK293 cell lines by CRISPR

The truncation of LZIC protein following overexpression of both wild type and flag-tagged protein suggests incorrect folding or localisation of the protein (Chapter 3.2.6). Specialised factors exist in all cells that can stabilise protein expression and prevent degradation, for example, XRCC4 increases the stability of Ligase IV protein through a chaperone-type activity (Bryans, Valenzano and Stamato, 1999). Therefore, the truncation of LZIC-flag protein following plasmid expression could be a result of insufficient availability of stabilisation complex. CRISPR provides a platform by which to produce total KO cells for specific proteins and through re-introduction of flag-tagged peptide allows protein replacement.

The cell line selected was HEK293, this is due to the ease with which this cell type can be genetically manipulated, additionally, the utilisation of the freestyle HEK293 system provides a system by which to produce material for affinity purification experiments (Portolano *et al.*, 2014). To generate KO cell lines, two gRNAs directed to ~80 base pairs upstream of the LZIC ATG were used in conjunction with a CAS9 and an insertion cassette, containing a GFP and Puromycin resistance gene, both of which were coded for on the same plasmid (Figure 3.8A). After initial transfection the cells underwent 8 passages, increasing the chance for loss of the unincorporated plasmid before antibiotic selection. Cells were serially diluted to generate clones. These clones were expanded and screened by western blot to determine whether loss of protein expression had been achieved (Figure 3.8B). Multiple clones had to be analysed due to the variability of CRISPR efficiency that can occur between individual cells. The efficiency of KO cell line generation varied significantly between the two gRNAs utilised with 20% and 75% efficiency for gRNA1 and gRNA2, respectively (Figure 3.8C). This can be due to the chromatin status of the region being cut but also the nucleotide sequence (Liu *et al.*, 2016; Kallimasioti-Pazi *et al.*, 2018). It is possible that the change in cut position between gRNA1 and gRNA2 leads to this reduction of efficiency.

Two knock-out clones were selected for further molecular and cellular analysis (LZIC-/- clone 1 and clone 2). Additionally, gRNA 2 clone 1 (CRISPR Control) was selected due to the continued expression of LZIC, as a control for the CRISPR procedure. To determine the genetic nature of the knock-out it was necessary to sequence the break-site.

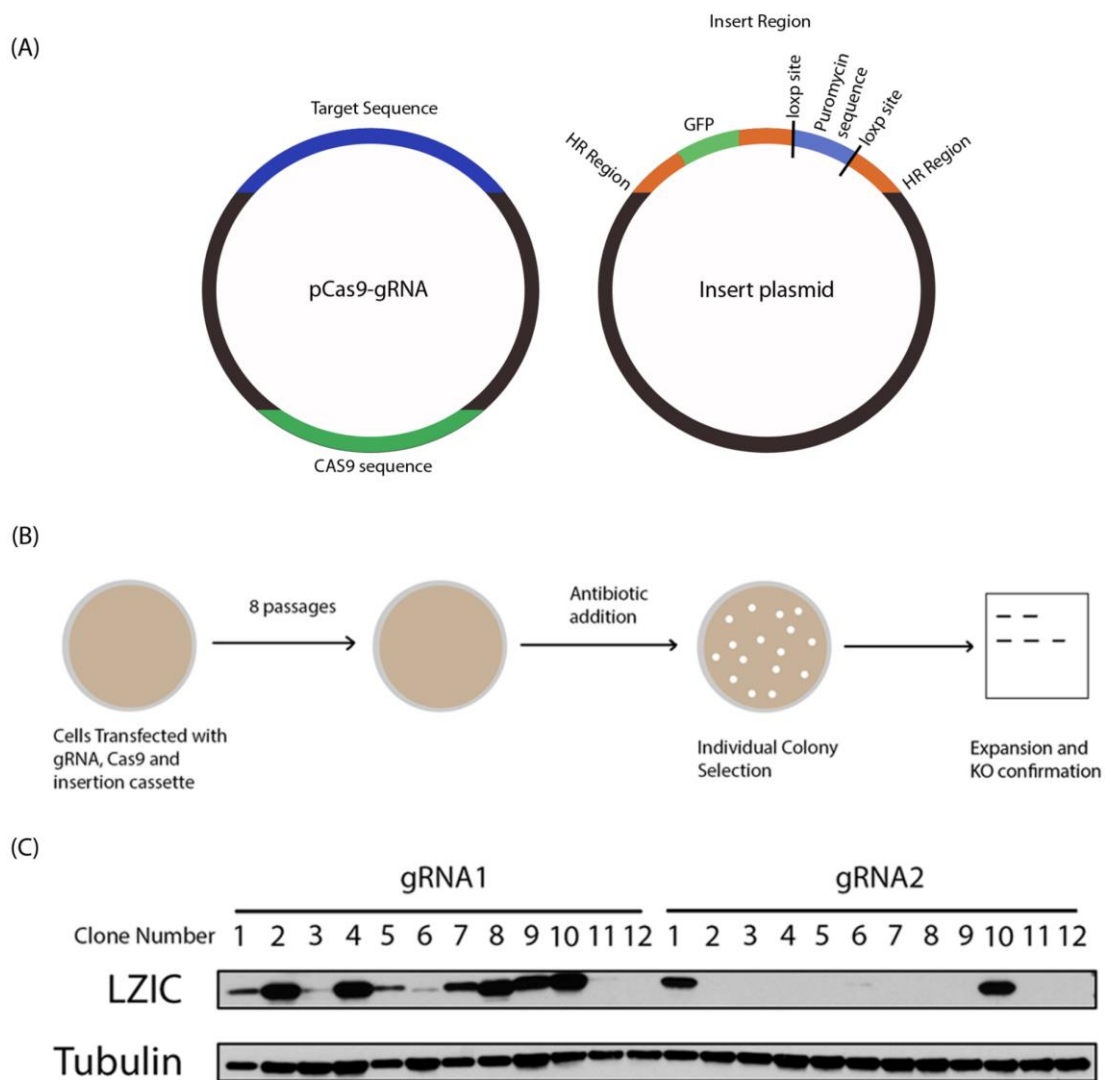


Figure 3.8 Establishment of LZIC knock-out line by CRISPR.

(A) Structure of plasmids used for CRISPR protocol. The Cas9-gRNA plasmid expresses both the gRNA sequence specific to LZIC and the Cas9 transcript. The second plasmid contains the cassette to be inserted by homologous recombination. In addition to the cut region sequence homology arms the sequence for GFP and puromycin is contained in the cassette to facilitate selection of clones. The puromycin resistance sequence is flanked by loxP sites. (B) Schematic showing process of CRISPR line generation in HEK293 cells. HEK293H cells were cotransfected with both plasmids outlined in (A) and underwent 8 passages. Puromycin selection was added to cells and colony selection conducted. The resulting colonies were screened for loss of LZIC expression by western blot (C) Screening of HEK293 single cell clones for expression of LZIC, indicates that successful knockout of LZIC in multiple clones from gRNA2 treated selection.

3.2.9 CRISPR cut site sequencing

The sequence alterations following CRISPR complex activity are important to determine, with the potential for multiple types of mutation to cause protein loss. Firstly, primers were designed to either side of the Cas9 cut site and PCR used to amplify this region. The separation of amplicons by gel electrophoresis was expected to yield sizes of between 2500-3500bps in length. However, this was not the case with amplicon lengths being much shorter than expected. This result indicated that a possible short length deletion or insertion was present in this region instead. The amplified regions were submitted for sequencing and the resulting profiles analysed by alignment to the genomic sequence for that region by use of NCBI blast align. Three different mutations were detected in the LZIC KO clone 1 and two mutations detected in LZIC KO clone 2. These constitute an 8bp deletion in all cases which leads to a frameshift (Figure 3.9A). The WT isoform 1 of the LZIC protein is 191 Amino acids in length, with 4 out of the 5 deletions generating a truncated translational product of LZIC equating to 35 Amino acids in length and the final deletion mutant detected in LZIC KO clone 1 generating a 50 amino acid product (Figure 3.9B). The number of mutations detected in the LZIC KO clone 1 genotype was three compared to just two separate sequences detected in LZIC KO clone 2. The resistance cassette could not be detected in the cut region, however, the cells are resistant to puromycin. It was concluded that the cassette has randomly incorporated into the genome.

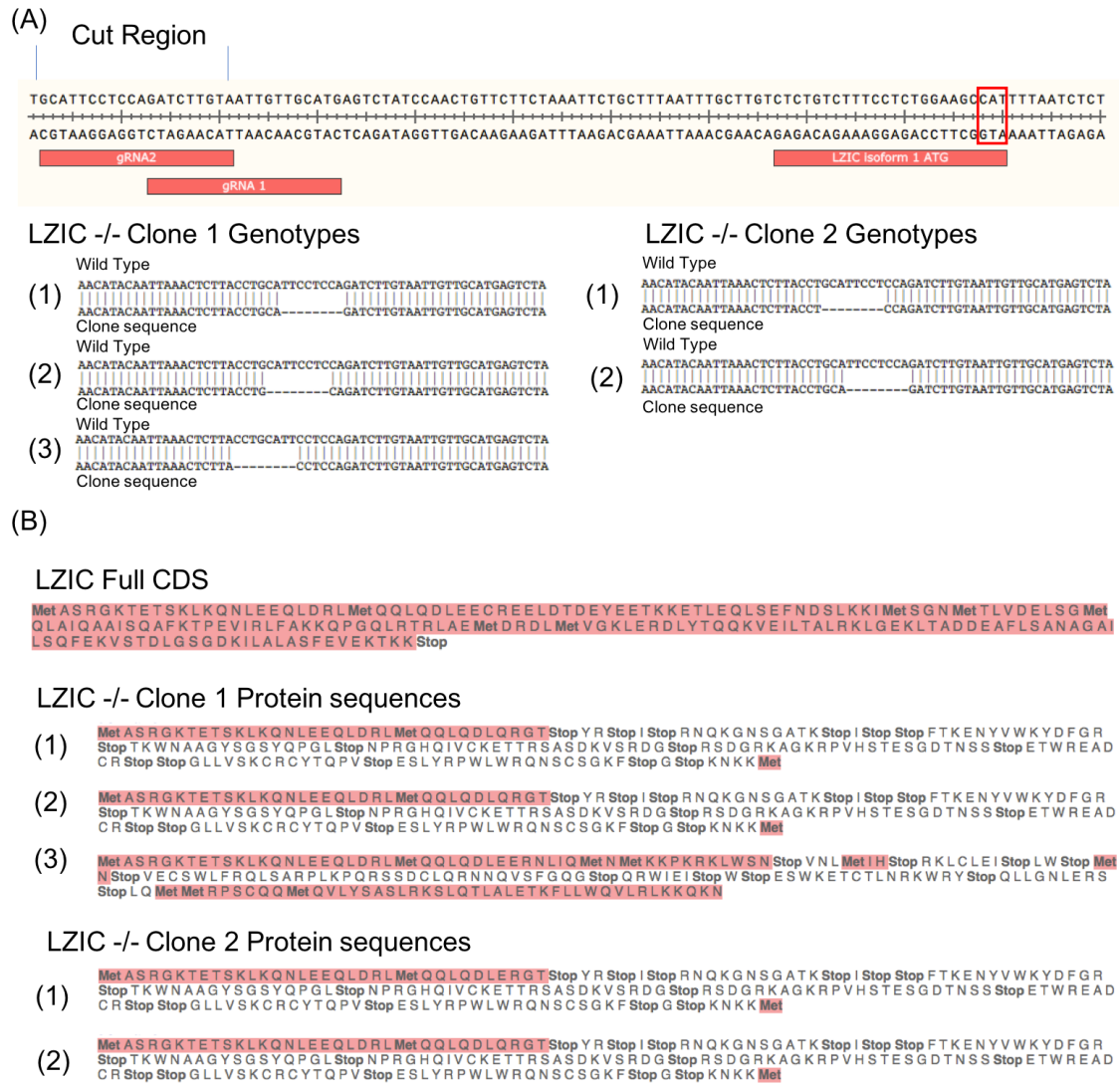


Figure 3.9 Impact of CRISPR mediated cut on LZIC mRNA and predicted protein expression.

(A) Sanger sequencing of break site in CRISPR generated knockout lines. The LZIC sequence is present on the reverse strand and the ATG is highlighted. The sequence results indicate that an 8bp deletion has been incorporated into the LZIC protein knock-out clones. This is repeated in all events with slight variation in position in sequence. (B) Full LZIC protein sequence and resulting protein changes due to the deletion in CDS. The amino acid sequence for all mutations detected indicates a frame shift, which introduces a stop codon and prematurely truncates the protein sequence.

3.2.10 Optimisation of mammalian expression system for recombinant LZIC.

The successful generation of LZIC KO HEK293 cells provided a system to re-express recombinant LZIC, without the potential competition with endogenous LZIC for co-factors and complex association. Therefore, this method may improve upon the previous systems tested. LZIC KO cell lines were used to re-introduce several exogenous LZIC pBABE expression constructs: C-terminally tagged LZIC (LZIC-c-flag), N-terminally tagged LZIC (LZIC-n-flag), and wild-type LZIC (LZIC-wt) sequence. LZIC KO Clone 2 cells were transfected with a range of plasmid concentrations between 1 µg/ml and 0.067µg/ml (Figure 3.10). The resulting transfected cells were harvested after 48hrs of transfection and analysed by western blot. All the constructs produce an alternative band at ~15kDa. The ratio between full-length and alternative LZIC increases in all cases with higher concentrations of the plasmid. The LZIC-c-flag expression construct showed the lowest level of degradation with only full-length protein detected at 0.5 µg/ml plasmid transfection concentration. The LZIC-c-Flag construct also expresses LZIC with flag-tag, in the case of LZIC-n-flag the tag is not detectable and so possibly cleaved.

Overall, this analysis indicated that expression of LZIC-c-flag can yield non-degraded and flag-tagged LZIC. However, to increase the consistency of expression for the construct a stable reintroduction line utilising LZIC KO clone 2 would be beneficial.

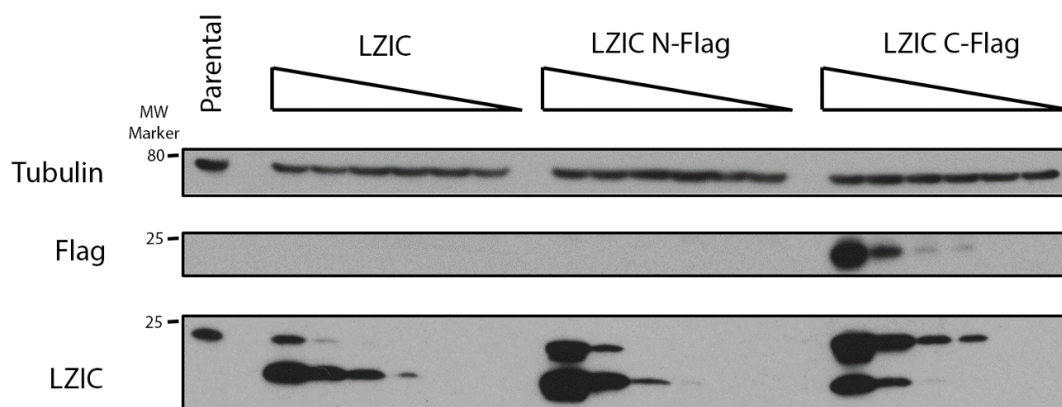


Figure 3.10 Expression of recombinant LZIC in LZIC KO HEK293 cells.

Transfection titration using concentrations of pBABE-LZIC ranging from 1 $\mu\text{g/ml}$ to 0.067 $\mu\text{g/ml}$. Included in this analysis is pBABE coded WT LZIC, N-terminally flag tagged LZIC, and C-terminally flag tagged LZIC. The parental is included to indicate true size of LZIC and flag blot is indicates whether the construct successfully expresses the tag. High levels of cleavage can be seen for constructs resulting in reduced levels of expression for full length protein.

3.2.11 Generation of LZIC KO HEK293 with stably expressed recombinant LZIC-c-flag

The transfection of plasmid for expression of recombinant LZIC-c-flag yields a significant level of truncated protein at concentrations above 250µg/ml (figure 3.10). However, both 250µg/ml and 125µg/ml of plasmid do not show any truncation product, while expression of full-length flag-tagged LZIC can be detected. This suggests that a strict protein expression limit exists per cell which when exceeded leads to a truncated protein. The successful generation of LZIC KO HEK293 cells provided a platform to reintroduce LZIC-c-flag expression at sub-physiological levels. The incorporated construct used to select LZIC KO lines during the CRISPR process contained a puromycin resistance cassette flanked by CRE recombinase cut sites. To select cells stably expressing exogenous LZIC with puromycin the cassette was removed by CRE recombinase treatment and clonal selection conducted by the recovery of sensitivity to puromycin. Following successful isolation of a puromycin sensitive variant of LZIC KO Clone 2 the LZIC-c-flag plasmid was introduced by plasmid transfection, however, this technique did not generate any stable clones. Therefore, the LZIC-c-flag sequence was recloned into a retroviral expression plasmid to be used for transduction. Following viral transduction of LZIC KO clone 2, the clonal selection was conducted as previously and a stably expressing line selected. The expression level of this clone was then compared to parental and CRISPR control lines. While expression of the Flag-tagged LZIC can be detected the level of protein detected is lower than endogenous, however, no degradation products are seen (Figure 3.11).

Following successful expression of LZIC-flag without truncation of protein, it is possible to conduct flag mediated affinity purification.

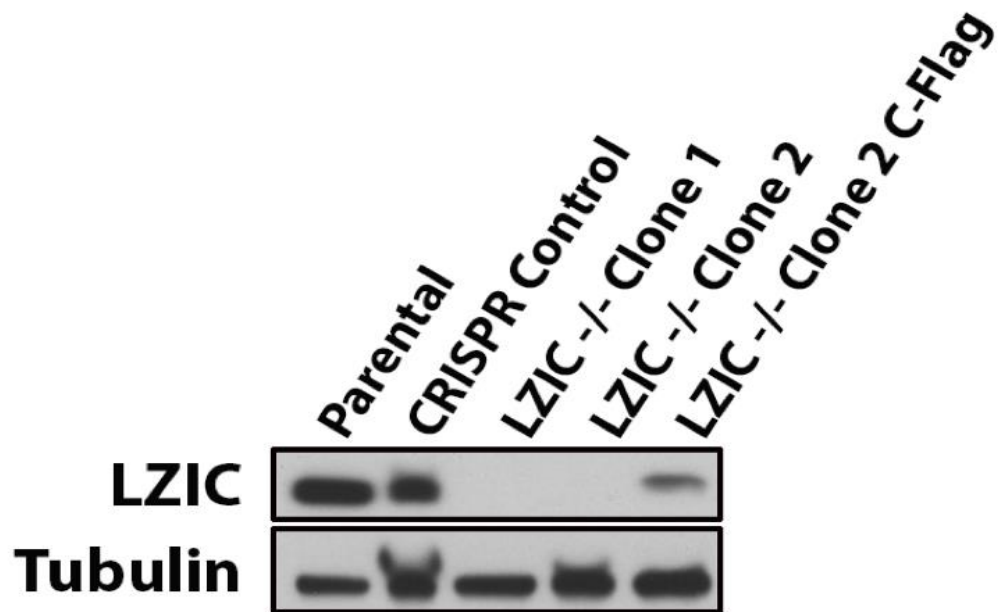


Figure 3.11 Comparison of recombinant LZIC expression to WT expression in parental and CRISPR control.

LZIC KO clone 2 cells were initially transfected with plasmid expressing CRE recombinase to remove the puromycin resistance cassette. These cells were then isolated into single cell clones and screened for sensitivity to puromycin. The puromycin sensitive LZIC KO clone 2 Cells were transduced with retrovirus, produced using pBABE vector containing C-Flag LZIC, and treated with puromycin to select cells which were successfully transduced. The cell line was further clonally selected and a clone was selected which expresses LZIC-flag at sub-physiological levels.

3.2.12 Data normalisation and technical quality control of LZIC-c-flag MS/MS spectral counts.

Samples were generated for affinity purification, utilising LZIC KO clone 2 as a control for non-specific interactors. As previously mentioned the cells were fractionated by high salt extraction of the nucleus to enrich for nuclear interacting partners (Craxton *et al.*, 2015). With the resulting proteins purified by LZIC-flag affinity purification being analysed by MS/MS and LFQ. The resulting peptide spectra were initially processed using the Perseus software (Tyanova *et al.*, 2016).

The detection of peptides is inherently susceptible to systemic bias against low abundance peptides. This results in non-identification and can cause significant problems for downstream statistical analysis. Imputation is a method by which to artificially add values which represent these “missing values” and facilitate downstream analysis processes (Taylor *et al.*, 2017). Following imputation, the normal distribution of each sample is improved, with a reduction of bias against low values. The native normal distribution is shown in blue with the computationally added values being shown in red (Figure 3.12A). Principal component analysis (PCA) is a method of comparing the proportion of variance in a data set which can be accounted for by specific components (Lever, Krzywinski and Altman, 2017). Conducting this analysis on the imputed data set indicates that the 1st component only accounts for 32.4% of the variance, which represents the difference between control and LZIC IP, with the 2nd component representing 21.2% of the variance, which is split by individual IP and may represent technical variation (Figure 3.12B). This shows that these two known components only account for a total of 53.6% of the total variance. This analysis indicates that the control from repeat 1 is aberrant and does not cluster with the other control samples.

To investigate the relatedness of samples, hierarchical clustering can be performed. This analysis indicated that the control from experiment 1 clustered incorrectly and precluded further processing (Figure 3.12C).

Overall, the analysis of data by PCA and hierarchical clustering indicated an incorrectly clustering experimental control. To address this the experimental repeat was removed and the normalisation steps repeated.

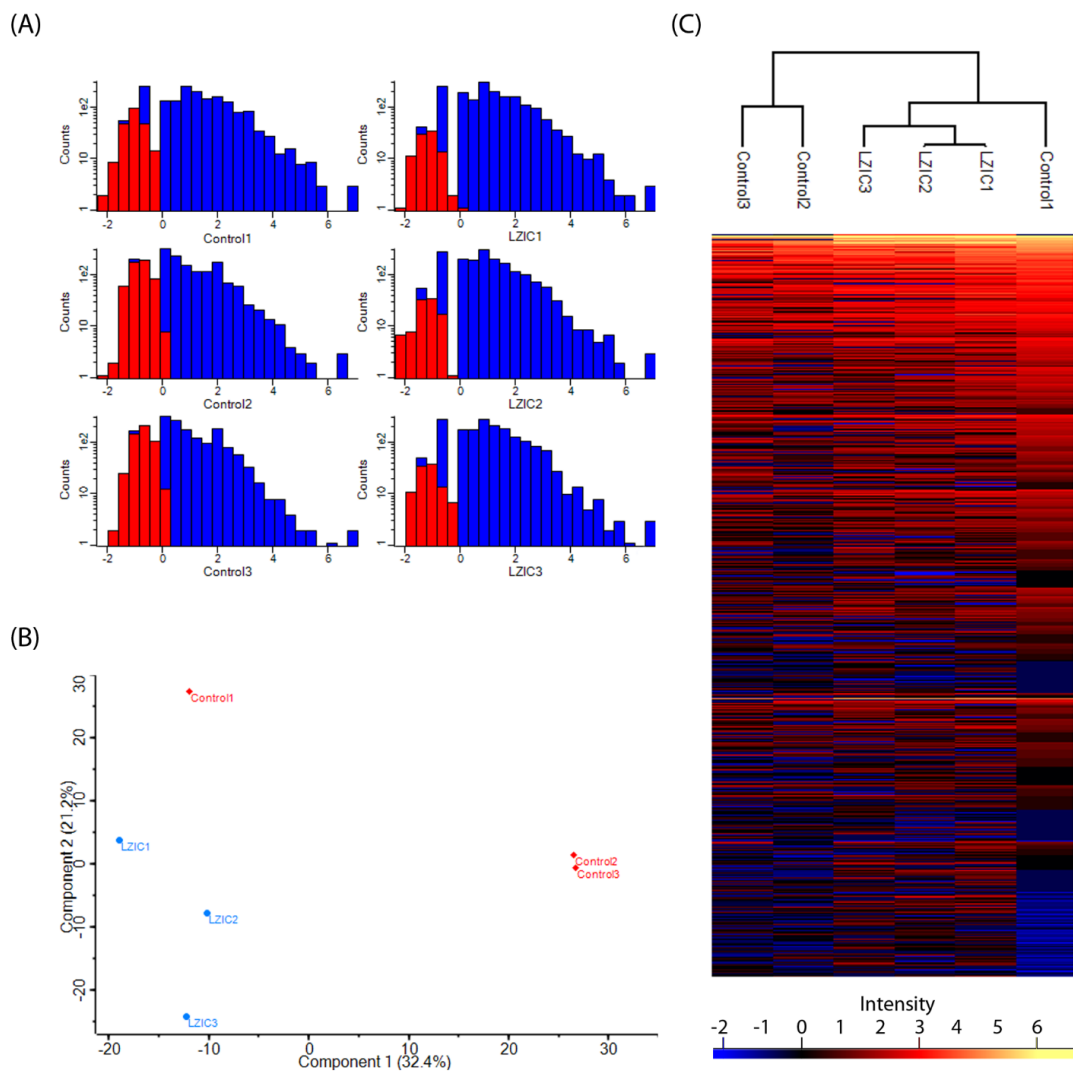


Figure 3.12 Data normalization and quality control checks of LZIC-c-flag MS/MS spectral counts.

(A) Histogram representing the distribution of peptide counts. Shown in blue is the actual counts, with red representing the artificially imputed values. (B) Principal component analysis (PCA) for each affinity purification, representing relationship between the conditions in component 1 and component 2. The names given show condition followed by experiment number. This separation indicates that the IgG control in experiment 1 is aberrant (C) Hierarchical clustering of all conditions by column, therefore representing condition. The intensity represents the peptide counts compared between columns.

3.2.13 Repeat normalisation and quality control of LZIC-c-flag MS/MS spectral counts following removal of aberrant sample

Due to the identification of a biased control sample, the entire experimental repeat was removed and data reanalysed. As samples were removed prior to analysis pipeline instigation, the imputation step was repeated using the same settings as previously (Figure 3.13A).

The PCA analysis of samples following removal of the divergent experimental repeat significantly alters variance accounted for by the first 2 components. The 1st and 2nd component accounting for 46.8% and 28.2% of the variance, respectively (Figure 3.13B). This represents a total of 75% of experimental variance which is a substantial increase and will remove variance that could detrimentally alter the downstream statistical analysis. The separation of sample, Control 2 and Control 3 by PCA may be representative of different background non-specific binding proteins which are found in absence of bait, this can be confirmed by the enrichment profile of hierarchical clustering which identifies substantial differences between the two samples, in contrast, the LZIC-flag samples show very similar clustering (Figure 3.13C).

To determine the proteins which were specifically enriched in LZIC-Flag affinity purification samples over those detected in the control IP, unpaired student T-Test was used.

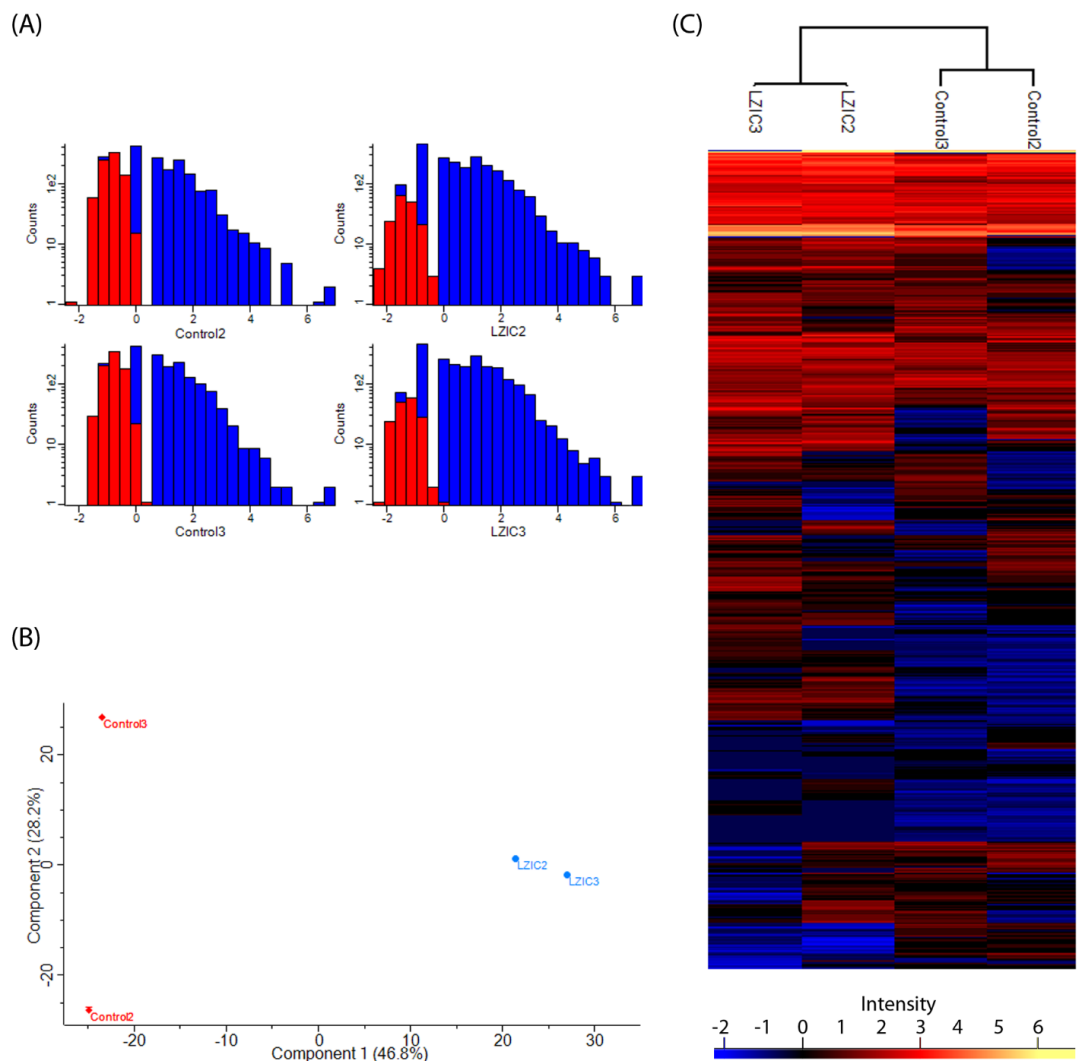


Figure 3.13 Quality control checks for LZIC-c-flag MS/MS analysis post removal of aberrant experiment.

(A) Histogram representing the distribution of peptide counts. Shown in blue is the actual counts, with red representing the artificially imputed values. (B) Principal component analysis (PCA) for each affinity purification, representing relationship between the conditions in component 1 and component 2. The names given show condition following by experiment number. Removal of the aberrant repeat improves separation between the two conditions. (C) Hierarchical clustering of all conditions by column, therefore representing condition. The intensity represents the peptide counts compared between columns.

3.2.14 Identification of significant LZIC-c-flag interactors

An unpaired student T-test was implemented to determine the significance of identified proteins with a p-value cut-off of 0.1 which was applied to the data and an additional requirement for a difference score of over 1. The difference score indicates the log-fold enrichment over detection in the IgG following imputation. LZIC is detected as one of the 5 genes with the highest intensity relative to the control IP, supporting the successful affinity purification of c-flag LZIC. The three further genes showing high difference were ACTA, ACTC, and ACTS.

The top 20 proteins identified were analysed by literature search for function with specific functional clusters being identified during this process (Figure 3.14A). The first of these are proteins which control transcription but also have associated roles during DSB repair. SP16H is the most confidently detected protein. This is a central component of the Facilitates chromatin transcription (FACT) complex and is vital for correct chromatin remodelling surrounding regions of transcriptional activation (Belotserkovskaya *et al.*, 2003). In addition to this role, the FACT complex regulates chromatin dynamics following DSB induction, specifically homologous recombination (Kari *et al.*, 2011; Oliveira *et al.*, 2014). A second identified protein is ATM is the master regulator of response to DSB induction by HR (Section 1.5.4). Dead-box helicase 47 (DDX47) and Serine and arginine Rich splicing factor (SRSF9) were detected and both have a role in splicing and processing of RNA following transcription. ILF2 is a regulator of both transcription and differentiation of various cell types, in addition, to having a role regulating DNA repair response (Marchesini *et al.*, 2017; Ye *et al.*, 2017).

The second group identified suggest a functional impact on transcription and cell cycle progression. Periphilin 1 was detected and is a regulatory member of the HUSH complex which regulates epigenetic silencing of transcription (Robbez-Masson *et al.*, 2018). The loss of periphilin 1 also alters the regulation of cell cycle progression particularly through S-phase (Kurita *et al.*, 2007). Compounding this finding the protein minichromosome maintenance complex component 4 (MCM4) was identified and has a role in the initiation of replication during S-phase and is involved in origin firing (Sheu and Stillman, 2010).

The final functional role identified within the top interactors is control of neuronal development. IQ motif and Sec7 Domain 2 (IQSEC2) protein is required for brain development and loss of this protein has been associated with significant intellectual deficiencies (Shoubbridge *et al.*, 2013).

The identification of significant proteins is conducted following imputation to replace conditions in which no peptide was detected. Another classical method of classifying significant interactors within mass spectrometry analysis is to identify interactors which are not detected in the control immunoprecipitation. The peptide counts of the top 20 most significant proteins were ranked by abundance in the LZIC-c-flag IP (Figure 3.14B). The most abundant proteins detected are the actin subunits ACTA, ACTC, and ACTS which may be due to the abundance of cytoskeletal protein. LZIC is the fourth most abundant protein identified and no background binding to the IgG control can be detected. In the general, the remaining identified proteins align with the determination of statistical significance and no binding to the IgG can be detected.

Previous analysis within this chapter has attempted to investigate LZIC interactome by yeast-two hybrid and interrogation of the interactome database Biogrid. The comparison of the interaction partners of LZIC identified by these methods compared to the LZIC-c-flag interactome demonstrates no overlap (Figure 3.15C).

The proteins which have been highlighted are the most significant interactors of LZIC detected in this study. However, in total 199 proteins were identified as significantly enriched over the control. In order to determine the gene function enrichment of all the genes identified a GO term analysis was conducted.

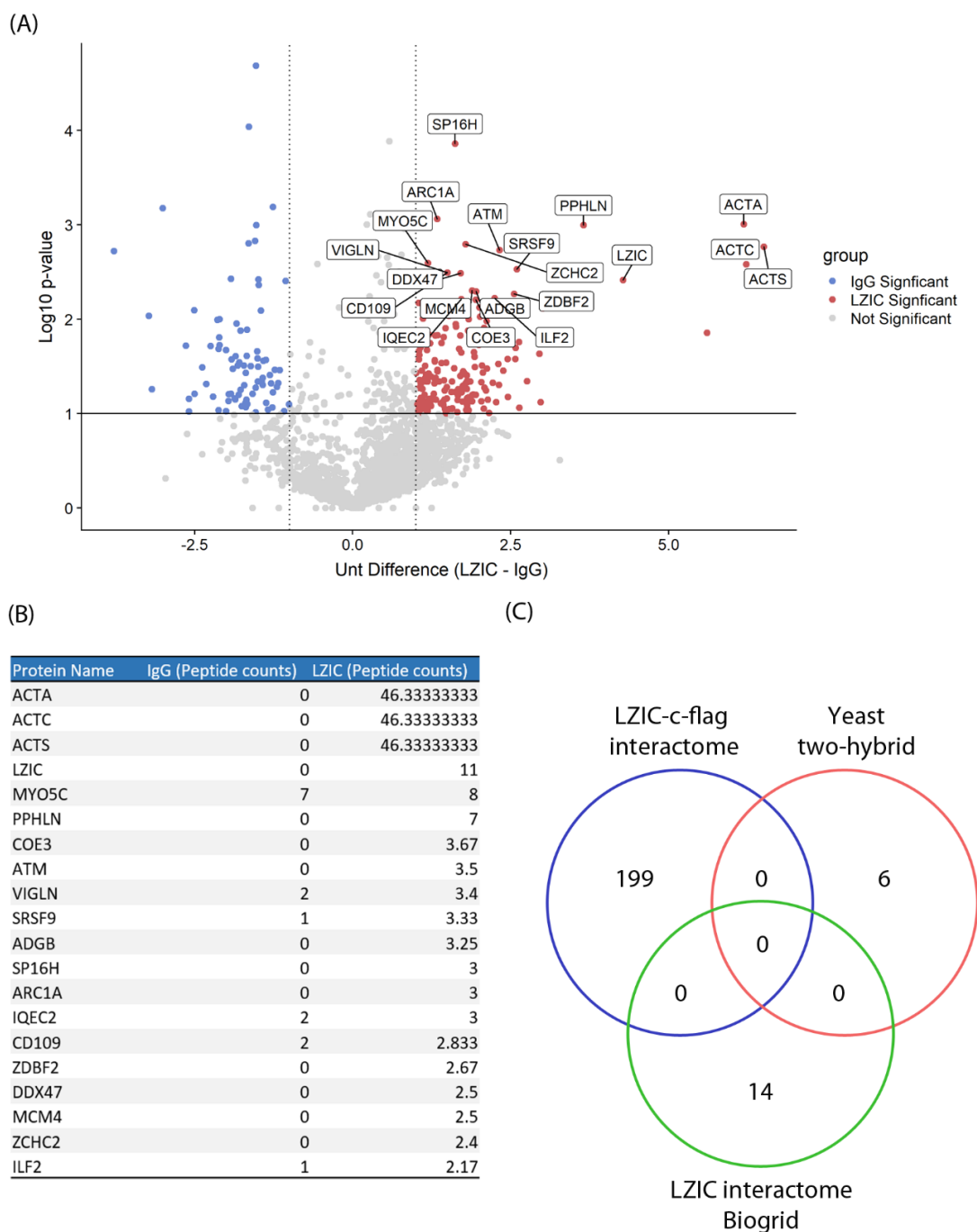


Figure 3.14 significantly identified LZIC-c-flag interactors.

(A) The log₁₀ p-value is determined by unpaired Student-T-test between the IgG and LZIC conditions. While the difference represents the differential of log₂ peptide counts observed between the two conditions, with positive representing enriched in LZIC interactome and negative indicating enrichment in IgG interactome. The red and blue coloured dots represent interactors which have either over a difference of 1 or -1, respectively, and a log₁₀ p-value of >1. (B) The peptide counts for each of the top 20 most statistically significant proteins as detected in the IgG and LZIC conditions. (C) Comparison of the proteins detected as LZIC interacting partners in this study to previously identified LZIC interactors on Biogrid and detected during yeast-2 hybrid analysis.

3.2.15 Gene ontology term enrichment of LZIC-c-flag interactors

The enrichment of large groups of genes for a related function can be established by conducting gene ontology analysis. This process compares the query gene lists against reference gene sets which have been assigned specific functional linkage. A caveat of this modality of analysis is the repeat of multiple genes for small overlapping groups. Therefore, methods of compacting these terms have been developed. This analysis was conducted by Revigo using semantic similarity index (SimRel)(Supek *et al.*, 2011).

Five major clusters of gene ontology terms can be identified by this analysis: mRNA splicing via spliceosome, ribonucleoprotein biogenesis, actin filament-based process, anatomical structure development, and establishment of RNA localisation (Figure 3.15). The most enriched cluster is that of mRNA splicing via spliceosome, alterations to this GO term can have a significant impact upon the transcriptome of the cell in response to extracellular stimuli and during differentiation. The enrichment for ribonucleoprotein complex biogenesis includes all processes which control assembly of the ribosome, including modification and processing of ribosomal RNA. Anatomical structure development and tissue migration contain genes which are involved during the development and differentiation of tissues. The uniqueness score for both of these GO terms is low indicating that multiple of the terms from the collapsed lists fall under these categories. Interestingly these processes are also regulated by planar WNT signalling (Wallingford and Mitchell, 2011).

In addition to these major clusters, some smaller groups can also be identified. One such group is the regulation of response to gamma radiation. As previously discussed exposure of cells to IR induces the formation of DSBs with the repair of these events being paramount to cell survival. In addition, developmental processes GO term is identified which coupled to anatomical structure development could be an explanation for previous phenotypes of LZIC loss (Clements and Kimelman, 2005).

Overall, the GO term analysis of LZIC interactome suggests a variety of functions, however, predominantly, these results suggest a function for LZIC as a spliceosome component.

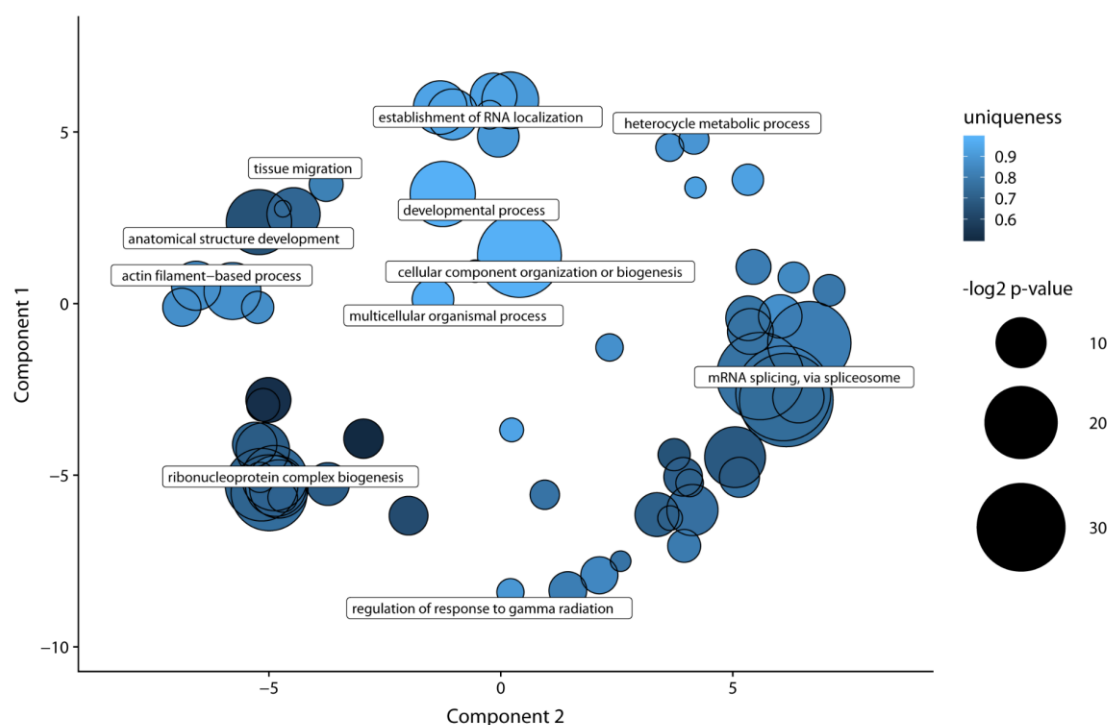


Figure 3.15 Gene ontology enrichment for all significantly identified LZIC interactors.

All genes with a p-value below 0.1 and a log-fold change of >1 above the control IgG were used for determining enriched gene ontology terms. Gene ontology analysis was conducted using STRING, with visualisation and compressing of GO terms performed with REVIGO through SimRel. REVIGO measures similarity between GO terms to determine no redundant terms for visualization. The $-\log_2 p\text{-value}$ represents the significance of each term identified. Uniqueness value represents the number of similar GO terms that were identified prior to SimRel. The most enriched groups of term clusters are those for mRNA splicing, via spliceosome and ribonucleoprotein biogenesis.

3.2.16 Protein domain term enrichment of LZIC-c-flag interactors

As previously discussed WNT signalling proteins are characteristically enriched for domains such as armadillo repeats (Tewari *et al.*, 2010). Therefore, domain enrichment analysis of LZIC interactors was conducted. This analysis identified 10 enriched domains. The least significantly enriched group is that of helicase domains (figure 3.16). Helicase domains are required for unwinding secondary structure that can be generated by complementary bases within single-stranded nucleotide stretches (Gorbalenya and Koonin, 1993). Helicase domain-containing proteins have vital roles within processes, such as translational control and DNA repair (Jaramillo *et al.*, 1991; Brosh, 2013). In addition to the helicase domains identified, RNA recognition and binding motifs are also identified, which may suggest that the helicases domains are predominately active on RNA secondary structure.

The most highly enriched groups identified in this analysis are three groups containing the ankyrin repeat and a group representing armadillo type fold. The ankyrin repeat is found in proteins which are related cell surface-associated, however, incidences of this fold type have been identified in WNT signalling cascade proteins (Schwarz-Romond *et al.*, 2002). Interestingly, a significant enrichment can be observed for proteins which contain an armadillo type fold, which while not identical to canonical armadillo folds are found within many proteins from both WNT signalling and other pathways.

Overall, these results indicate that LZIC interactome is enriched for domains which are involved in RNA processing and signal transduction from cell-surface through ankyrins and armadillo-like repeat-containing proteins.

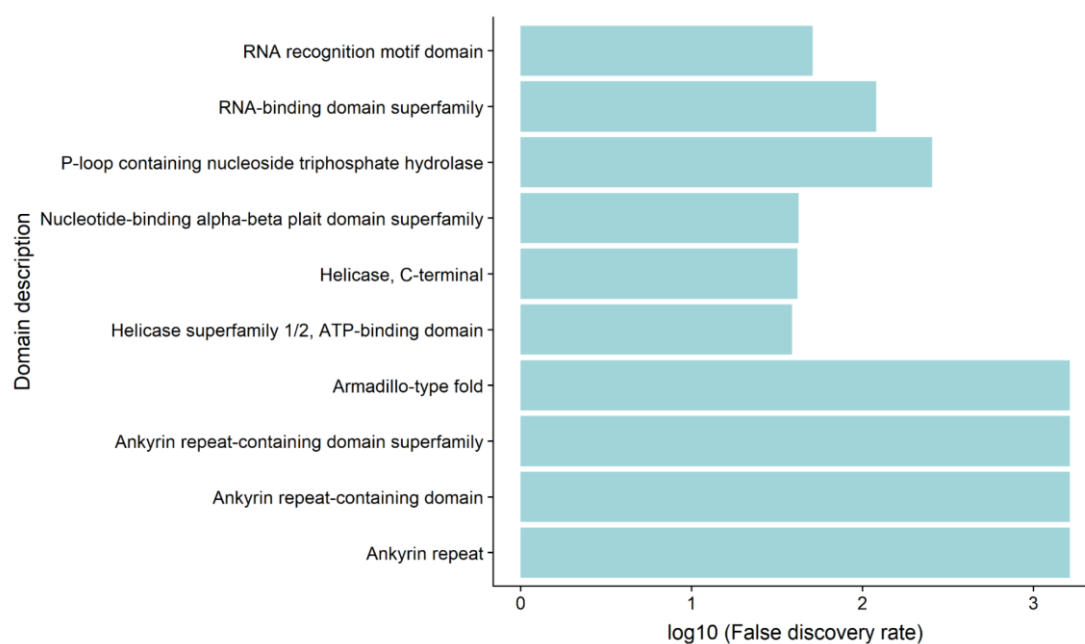


Figure 3.16 Domain enrichment analysis for all significantly identified LZIC-c-flag interactors.

All genes with a p-value below 0.1 and a log-fold change of >1 above the control IgG were used for determining enriched gene ontology terms. Domain enrichment analysis was conducted using STRING database and a cut off of >100 genes per domain term was applied to the resulting lists.

3.3 Discussion

LZIC is well conserved yet poorly characterised putative member of the WNT signalling cascade. Previous investigations have suggested a role for both development and tumorigenesis (Kato, 2001; Clements and Kimelman, 2005; Daino *et al.*, 2009). However, the molecular investigation is lacking which could provide insight into the contribution of LZIC to these phenotypes. This chapter aimed to extend the previous evolutionary conservation analysis of LZIC and elucidate the interaction partners of LZIC by utilising mass spectrometry.

The overall conservation analysis of LZIC supports previous findings that LZIC sequence is maintained in multiple species (Clements and Kimelman, 2005). Protein size is directly correlated with evolutionary conservation. In *Homo sapiens* conservation increases with length of protein where it was found that the majority of conserved proteins are over 400 residues in length (Lipman *et al.*, 2002). This shows the high level of protein conservation for LZIC is atypical due to its small size of 190 residues. Interestingly, the divergence of LZIC sequence in nematodes matches that of multiple WNT signalling proteins (Mizumoto and Sawa, 2007). Therefore, this finding may provide further evidence for the classification of LZIC as a WNT signalling protein. However, the regional analysis of LZIC conservation suggests the loss of conservation within the ICAT domain, which could suggest that the function performed by this domain and so the protein is diverging from the WNT signalling cascade. Future work should aim to characterise the nucleotide-binding capabilities of LZIC *in vitro* and relate this to cellular function.

The interacting partners of LZIC are unknown and during this chapter, three sets of interacting partners have been identified by meta-analysis, yeast-2 hybrid and flag-LZIC isolation from LZIC KO cell lines. There is no overlap between the proteins identified by these three methods, however, multiple reasons are plausible to explain this phenomenon. Firstly, the different studies have been conducted in multiple cell types with widely variable proteomes. In addition, the method of bait purification varies, with some of the studies using peptide tagged baits and others utilising endogenous affinity purification. The MS analysis conducted in this chapter is the first to utilise LZIC as a bait protein. In addition, the utilisation of an LZIC KO line for stable supraphysiologic expression of the bait protein may have benefits over typical over-expression systems as the bait protein

will not be competing with endogenous untagged variants for incorporation into complex and the lack of over-expression can reduce cellular stress. Finally, the methods of protein isolation vary between each of the previous studies that identified LZIC. The protein isolation method utilised during this chapter involved fractionation of cells to enrich for the nuclear compartment, in comparison to the whole-cell lysis used by many of the studies. This difference can substantially alter enrichment of potential prey proteins and so may account for the poor overlap between the interactome generated in this chapter and previous work. An important caveat of the investigation conducted in this chapter is that the targets have not been validated by alternative methods. Therefore, the interaction partners of LZIC identified are only suggestive and further work should focus on confirming these interactions utilising methods such as immunoprecipitation.

While the identity of the proteins detected in the various studies mentioned above varies, the overall consensus of functional groups is consistent. Predominantly, transcriptional regulators are identified, either direct transcription factors or spliceosome components. This may provide some functional evidence that LZIC is a component of the WNT signalling pathway, as the major outcome of pathway activation is transcriptional regulation. However, this study did not identify β -catenin interaction with LZIC regardless of the analysis performed, supporting previous data that the ICAT homology domain is unable to bind the canonical armadillo repeat within β -catenin (Clements and Kimelman, 2005). Interestingly, the identification of armadillo-like repeats and ankyrin domains during the domain enrichment analysis, suggests that while LZIC can't bind the canonical armadillo repeat, it may have the capacity to bind the armadillo-like repeat family (Figure 3.2.16). To address this point recombinant armadillo-like repeat could be generated *in vitro* and utilising affinity purification, the capacity of LZIC to bind such substrate assessed.

Enrichment for factors which have a role for regulation of cell cycle and DSB repair responses was observed in both the meta-analysis and LZIC-C-flag MS. The WNT signalling pathway has multiple roles regulating the G1 checkpoint progression cascade and the regulation of mitosis through microtubule dynamic (Shtutman *et al.*, 1999; Niehrs and Acebron, 2012). Therefore, the identification of these GO terms is not unexpected for a putative WNT cascade member. However, while some link between WNT signalling and IR has been suggested, such as the interaction between β -catenin and ligase

IV, the identification of gamma radiation response proteins may provide a further link between these two pathways, which will be investigated in further chapters (Jun *et al.*, 2016). These interactors are also further supported by the finding of LZIC loss promoting the formation of IR induced osteosarcoma (Daino *et al.*, 2009). However, due to the nature of MS/MS analysis further work would include validation of the most abundant identified factors to demonstrate these were not false-positive events.

The GO analysis of MS identified proteins revealed multiple functional groups in addition to spliceosome regulators. There is significant enrichment for proteins involved in ribonucleoprotein biogenesis. However, ribosomal subunits are typical contaminants of many mass spectrometry experiments, and so would require careful confirmation (Mellacheruvu *et al.*, 2013). Another significantly identified group is anatomical structure development, which encompasses proteins regulate the developmental process, such as neuronal growth. With previous findings that LZIC loss induces apoptosis of the midline brain region of zebrafish, proteins identified in this group could provide candidates for further analysis to address this phenotype (Clements and Kimelman, 2005)(Figure 3.15).

Overexpression of LZIC leads to significant truncation of the protein with truncation of the wild-type untagged version also being observed. This may indicate the requirement for a chaperone which either aids with the folding of LZIC or modulates the subcellular localisation. This is supported by the lack of truncation observed when expressing LZIC in a KO cell line at supra-physiological levels and it could explain the lack of interactions observed for LZIC-GST (Figure 3.6 & 3.10). In addition, it could suggest a post-translational modification of LZIC is required for stability, which in further work could be detected by specialised mass spectrometry analysis pipelines.

Overall, in this chapter further LZIC conservation analysis has shown a similar profile to proteins of the canonical WNT pathway. An LZIC KO cellular model has been successfully generated which was utilised to reveal a protein interactome for LZIC. The identified proteins predominantly congregate on the regulation of transcriptome and alterations to cell cycle and DNA repair response. Therefore, the next chapter will attempt to determine the changes to transcriptome induced by loss of LZIC and whether an altered transcriptional response to gamma radiation can be detected.

Chapter 4

Impact of LZIC loss on basal transcriptome and transcriptomic response to ionising radiation

4.1 Introduction

4.1.1 WNT signalling is linked to vital transcriptional regulatory cascades

Transcription is a process by which the genetic information stored in DNA is replicated into RNA. Transcription is performed by members of the RNA polymerase family, with the major polymerase responsible for transcription of messenger RNA (mRNA) being RNA polymerase II (Pol II) (Dignani, Lebovitz and Roeder, 1983). The initiation of Pol II driven transcription is dependent upon the permissive chromatin state of gene promoter regions (Fuda, Ardehali and Lis, 2009). Transcription factors promote the activation of gene regions by binding to specific DNA sequences and recruiting factors, such as chromatin re-modellers (Lambert *et al.*, 2018). To prevent the constitutive transcription of gene regions, the abundance and post-translational modification state of transcription factors are modulated. This process is the last layer of signal transduction from both extracellular and intracellular signalling cascades and therefore the breakdown of this process can lead to severely compromised gene expression profiles.

Canonical WNT signalling regulates the transcription of the TCF/LEF family genes through binding of the transcription factor β -catenin (Section 1.8). However, canonical WNT signalling also controls growth-promoting transcription factors, such as the MYC family (He *et al.*, 1998). The MYC transcription factor family: c-MYC, N-MYC, and L-MYC are classical oncogenes, with the overexpression of c-MYC, in particular, being linked to increased DNA mutations and genome instability (Kuzyk and Mai, 2014). An example of the oncogenic potential for the dysregulation of the c-MYC-WNT cascade is the formation of BRAF positive lung cancer (Juan *et al.*, 2014). The pathological amplification of c-MYC activity drives the expression of its target genes, cyclin D1 and cyclin B1, which forces progression through cell cycle checkpoints at G1 and S-phase (Felsher and Bishop, 1999).

4.1.2 Transcriptional response to gamma radiation exposure

The cellular response to IR exposure is a temporal process, with early events, such as phosphorylation of γ H2AX occurring within minutes of exposure and induction of the early G2/M checkpoint within 4 hrs (Xu *et al.*, 2002; Cucinotta *et al.*, 2008). However,

increased γ H2AX levels can be observed up to 48 hrs after exposure and the late G2/M checkpoint activates after 8 hrs (Xu *et al.*, 2002; Redon *et al.*, 2009). A meta-analysis of the transcriptional response to DNA damage in five cell types demonstrates the tight control of responses by p53 target genes (Rashi-Elkeles *et al.*, 2011). However, this study analyses transcriptional responses which occur between 3 and 6 hours post-exposure to IR. Further studies have attempted to analyse the gene expression changes at late time points in response to IR treatment and identified response pathway shift away from p53 control (Tsai *et al.*, 2006; Sokolov and Neumann, 2015).

The signalling pathways activated following exposure to IR will have a significant impact on whether cells are radiosensitive or radioresistant. The treatment of glioblastoma cells with IR induces the nuclear translocation of β -catenin and expression of TCF/LEF family transcripts increasing tumour invasiveness (Dong *et al.*, 2015). Additionally, the alternating complex formation between TCF-4 and either PARP1 and Ku80, in a DNA damage dependent manner is required to prevent TCF-4 transcriptional activity following treatment of cells with the radiomimetic, bleomycin (Idogawa *et al.*, 2007). These changes demonstrate the role of intracellular WNT signalling for the response to IR, however, the WNT signalling cascade also controls multiple paracrine molecules which can alter tissue responses to a toxic insult (Section 1.8). For example, exposure of fibroblasts to IR induces increased expression of WNT16B, which in turn initiates dendritic cell-driven recruitment of regulatory T-cells. The latter effect is believed to aid in immune evasion of cancer (Shen *et al.*, 2014).

4.1.3 Methods of transcriptome analysis

The predominant techniques of transcriptome analysis stem from variations of either RNA sequencing or microarray. RNA sequencing became more available with the reduction in the cost of next-generation sequencing technology and the development of more accessible bioinformatics analysis tools and pipelines (Kukurba and Montgomery, 2015). RNA sequencing can be used to detect mutations incorporated into mRNA or detected differential expression of splice variants, however, unless these data are required the nature of the procedure it is still a significant cost investment and other techniques are more applicable. Microarray technology was originally developed in the late 1990s and can assess relative RNA expression changes with pre-established data analysis

pipelines (Johnston, 1998). The benefit of this method is that each chip can specifically analyse the abundance changes of individual mRNA species, without the requirement to assess all RNA species in the cell. The basic process of transcriptome analysis by microarray involves reverse-transcription of an RNA to produce a fluorescence tagged cDNA, before exposure to anti-sense gene probes on the chips surface. The fluorescence intensity of each probe is representative of gene abundance.

4.1.4 Hypothesis and chapter aims

As a putative component of the WNT signalling pathway, LZIC is hypothesised to be involved in the regulation of cellular transcription. Additionally, the LZIC interacting partners identified in the previous chapter suggests interaction with components of the RNA polymerase II and spliceosome complex (chapter 3). Changes to these complexes can potentially alter cellular transcriptional response at both the basal level and in response to stress. In addition, previous analysis of LZIC suggests involvement in tumour development following exposure to IR, however, whether this phenotype has a transcriptional basis is unknown. In order to address these hypotheses, the impact of LZIC upon both basal transcription and transcriptional response to IR treatment will be investigated.

4.2.1 Analysis of IR responsive gene expression profile over 24hr time course

The gene expression changes in response to IR is a dynamic process requiring orchestration of early, predominantly p53 driven transcriptional changes and late transcriptional events (Rashi-Elkeles *et al.*, 2011). Jen *et al.* assessed transcriptome changes at different time points following treatment with two doses of IR, 3Gy or 10Gy (Jen and Cheung, 2003). The study characterised gene expression profiles of lymphoblastoid cells in response to IR, over a time course of, 1, 2, 6, 12, and 24 hours. In total 319 and 816 IR responsive genes were identified for the 3Gy and 10Gy doses, respectively, over the time course.

Analysis of all significantly altered genes identified at each IR dose indicates two distinct temporal profiles. The treatment of cells with 3Gy yields an overall decrease in the expression of IR responsive genes at 24hrs, in comparison to 10Gy which shows an overall increase (Figure 4.1). While there are detectable alterations in the log-fold change of IR responsive genes at 1hr post-IR treatment, the altered expression of genes reaches maximal deviation from 0hr at 24hrs post IR.

In conclusion re-analysis of this data set suggests that the largest number of IR response-related genes can be analysed by performing transcriptomic analysis at the 24hr time point.

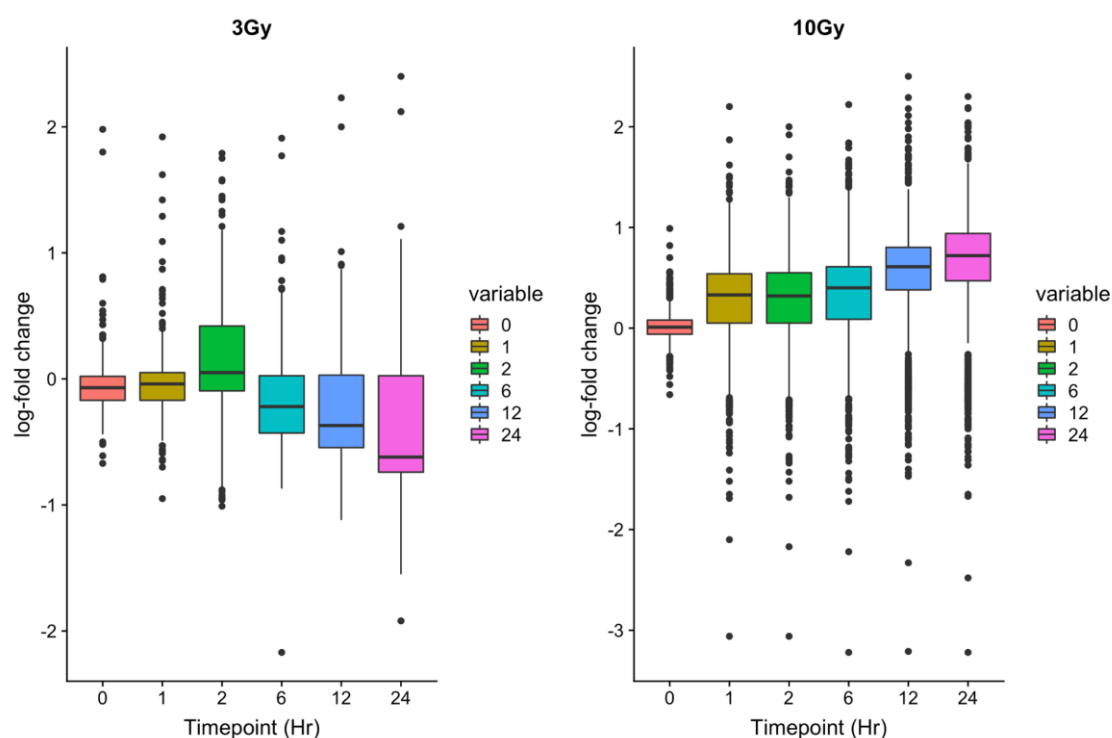


Figure 4.1 Meta-analysis of temporal gene expression changes in response to IR.

Data used for meta-analysis obtained from publication by Jen and Cheung *et al.* (Jen and Cheung, 2003). The study involved exposure of lymphoblastoid cell lines to IR and measurement of mRNA abundance across a post exposure time course. Log fold values for all significant genes identified in the study following treatment with either 3Gy or 10Gy IR were compared at each of the 6 time points. Overall, while the gene expression changes differ between the two doses overall the largest divergence of expression can be observed at 24hrs.

4.2.2 Comparison of CRISPR control cells to LZIC KO clones for quality control

The classification of LZIC as a putative member of the WNT signalling family suggests a function as a transcriptional regulator, however, no assessment of transcriptional changes in response to LZIC loss have been previously characterised. Therefore, the LZIC KO cell lines provide the ideal platform by which to analyse this phenotype. The correct transcriptional alterations following IR exposure are paramount to cellular adaptation and with the identification of LZIC loss as a priming event for the development of IR induced osteosarcoma, the LZIC KO cell lines also provide a system in which to assess whether this phenotype has any transcriptional basis (Daino *et al.*, 2009). To assess the impact of LZIC loss on basal transcription CRISPR control and LZIC KO clone 1 were incubated for 48hrs post-seeding, before RNA extraction. In addition, given the peak of the transcriptional response to IR at 24hrs post-exposure, half the seeded cells were exposed to 5Gy IR at 24hrs post-seeding and then harvested in parallel with untreated conditions.

Given that in the process of RNA extraction substantial degradation can occur. The RNA integrity was assessed by detection of the ribosomal RNA 18s and 23s, with the increased intensity of the 23s compared to 18s yielding a high RNA integrity number (RIN). All RNA samples assessed had a RIN score of 10 indicating successful extraction of the RNA with no degradation (Figure 4.2A). Single colour, Cy5 labelled probe, microarray was used to determine the overall abundance of genes within all conditions analysed.

Comparison of overall intensity profile between samples can identify outlying samples, due to either technical issues with microarray chip or due to biological variation between samples. This analysis of samples showed the general agreement of the intensity profile (Figure 4.2B Top). To perform differential expression analysis between conditions the data must first be normalised. Background correction was performed by mathematical fitting of convoluted normal and exponential models for background intensity vs foreground intensity, with subsequent quantile normalisation to standardise the $\log_2$ expression values between the two array chips used, both of these processes were conducted by Limma Voom (Figure 4.2B bottom)(Bolstad *et al.*, 2003; Silver, Ritchie and Smyth, 2009).

Overall, no degradation of samples was found prior to analysis and normalisation of intensities from each sample shows a strong correlation. However, further insight into data variables can be achieved by analysis with principal component analysis (PCA).

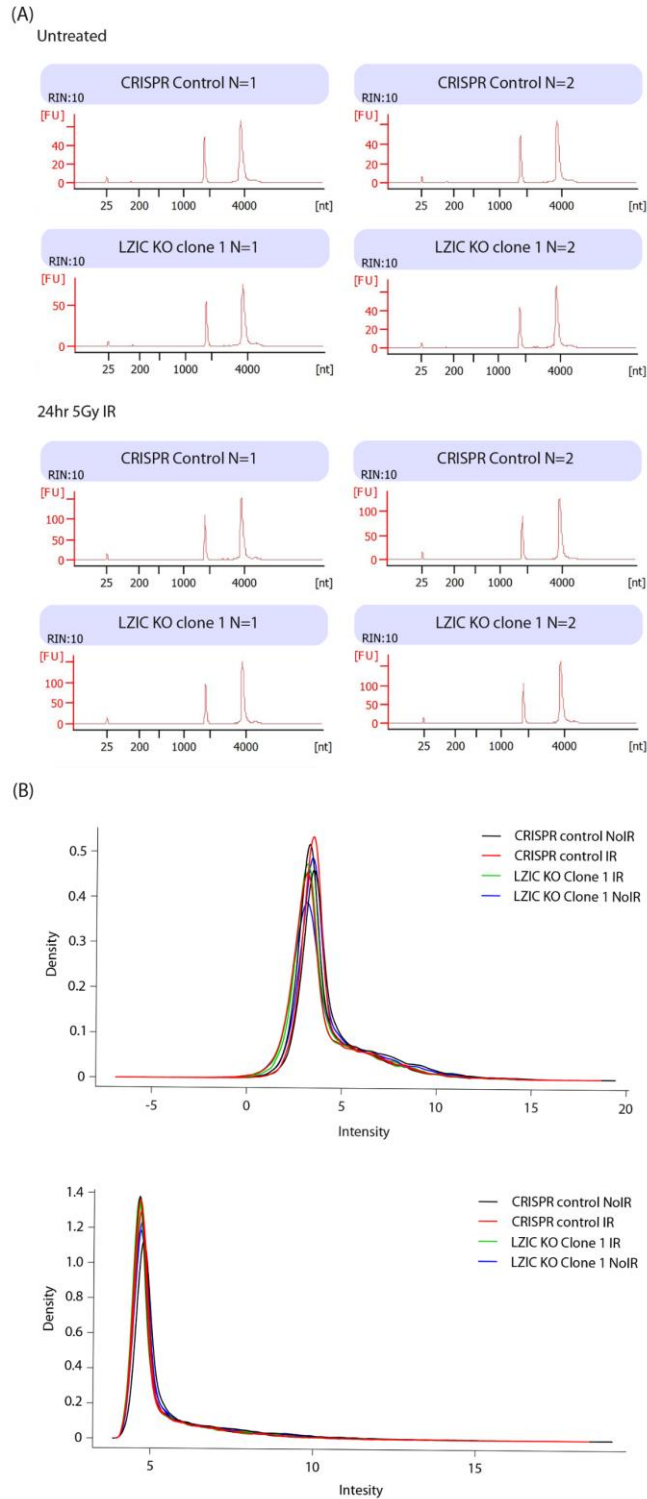


Figure 4.2 RNA quality analysis and microarray background normalisation.

(A) RNA integrity was determined by the ratio of 18s to 28s ribosomal RNA and this used to generate an RNA integrity score (RIN score). A RIN of 10 was generated for all conditions which indicates no degradation. (B) Following microarray chip imaging the raw intensity values for each probe were background corrected and normalised between arrays by Limma microarray analysis pipeline. The normalization methods used for each step Normexp and quantile, for background correction and between array normalization, respectively.

4.2.3 Variable analysis

The application of differential expression analysis to microarray samples is most powerful when the major source of variance within the data set is related to the biological question of interest. The application of PCA to a data set can identify the sources of variation (Lever, Krzywinski and Altman, 2017). This data manipulation splits the variance for each sample set into multiple dimensions. The component with the most variance will constitute dimension 1 with the ranked decrease of variance within the further dimensions.

The comparison of all dimensions indicates that 98.2% of the variance is accounted for by the 1st dimension (Figure 4.3A). The analysis of these dimensions can be performed in a 2D space to determine sample clustering. Firstly, the comparison of dimension 1 and dimension 2 indicates that 98.2% of the variance is due to the different arrays (Figure 4.3B). Scatter of dimension 2 and 3 split the samples by experimental condition with the 2nd dimension representing variation introduced by IR treatment and the 3rd dimension representing variation related to loss of LZIC expression (Figure 4.3C).

Overall, the major source of variation is due to the technical element of probe intensity between arrays and so this may reduce the number of statistically significant genes identified during differential expression analysis.

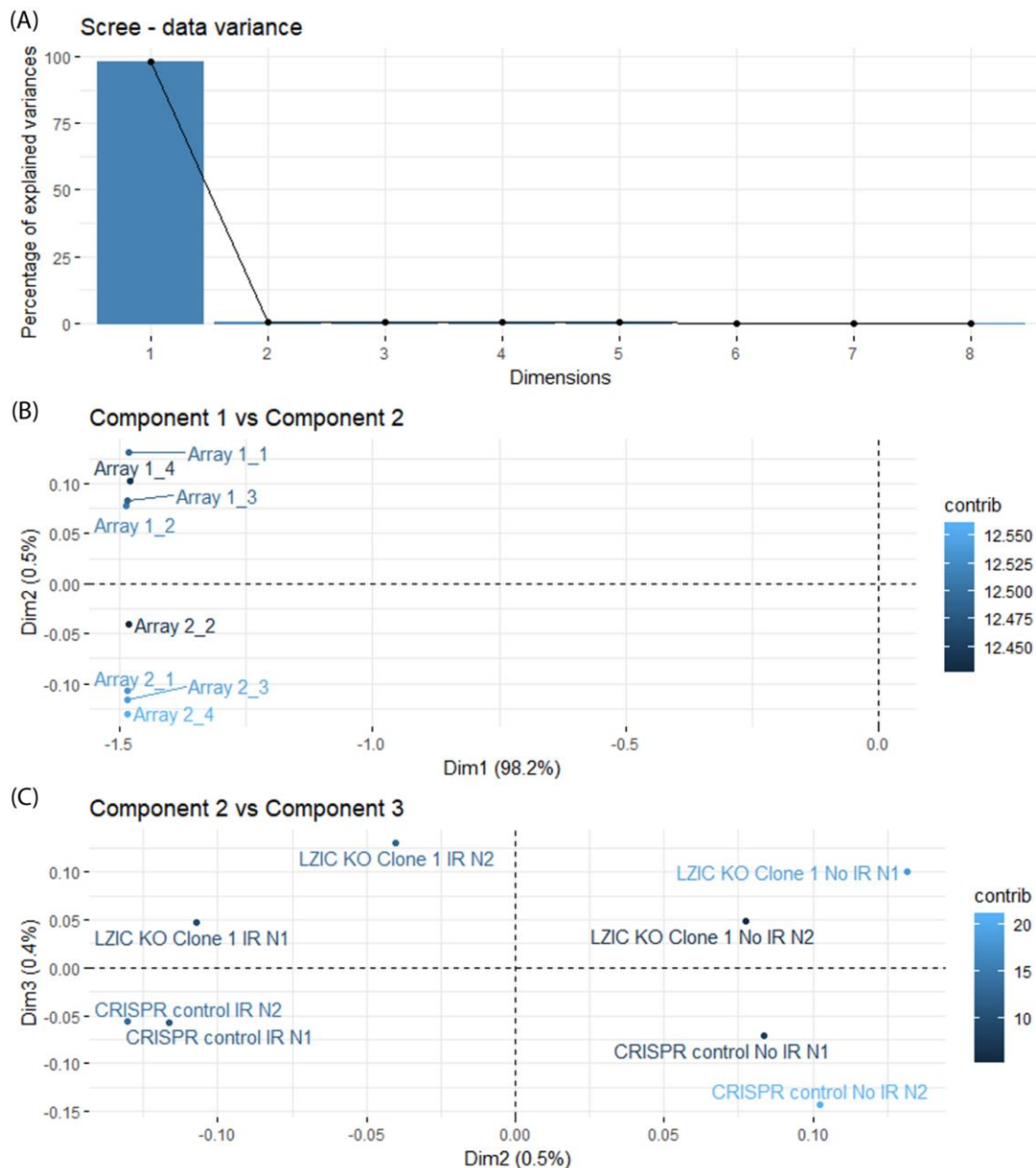


Figure 4.3. principle component analysis of microarray probe intensities from all experimental conditions.

(A) The Scree plot represents the contribution of each source of variance to total experimental variance. Overall, this analysis demonstrates that the majority of variance within the data set is generated by component 1. (B) Principal component analysis plot for variance component 1 and component 2. This comparison indicates that the majority of the variance within the data set is due to array. (C) Principal component analysis plot for component 2 and component 3. This comparison demonstrates that samples are separated by experimental condition, once variance associated with the array is removed.

4.2.4 Differential gene expression analysis of normalised gene probe intensity from all experimental conditions.

The identification of differentially expressed genes was achieved by comparison of CRISPR control to LZIC KO clone 1 both basally and following treatment with IR. The most down-regulated gene in response to LZIC loss is SRY-Box 11 (SOX11) with a log fold change of -1.8 and -1.4 in untreated and treated conditions, respectively (Figure 4.4A). The altered abundance of SOX11 occurs irrespective of IR treatment, suggesting that this is a constitutively LZIC dependent transcript. Correct SOX11 expression is required for the development of the nervous system and the altered regulation can promote tumorigenesis (Huang *et al.*, 2019; Liu *et al.*, 2019). In contrast, long-non-coding-catenin-Alpha 2 (lnc-CTNNA2) has the highest abundance change, with a coefficient of 1.992, this is only observed in basal conditions as this change is lost following exposure of LZIC KO cells to IR. While a function for lnc-CTNNA2 has not been established, in general, long-non-coding RNA can have a significant impact upon gene expression by regulation of the chromatin state (Rinn and Chang, 2012).

The identification of SOX11 as the most differentially regulated gene may provide a link to the neurological phenotypes associated with LZIC loss. In addition, further developmental regulators were identified, the first of these is protocadherin 7 (PCDH7) which is upregulated in LZIC KO cells both basally and following treatment with IR. Overexpression of PCDH7 has been linked with decreased neuronal survival through induction of apoptosis and interestingly the protocadherin family has significant links to the regulation of the WNT signalling pathway (Mah and Weiner, 2017; Xiao *et al.*, 2018). The homeobox gene family are canonical regulators of development and the identification of family members: homeobox B6 (HOXB6) and homeobox C4 (HOXC4), may suggest priming of LZIC KO cells for developmental defects (Luo, Rhie and Farnham, 2019).

In addition to the identification of factors associated with development, altered expression of genes that regulate cell cycle and DNA damage responses were also identified. Epithelial cell adhesion molecules (EPCAM) primary role is to mediate calcium-dependent cell adhesion, however, through association with four-and-a-half LIM domains protein 2 (FHL2) EPCAM directly upregulates expression of cyclin D1

(Chaves-Pérez *et al.*, 2013). Stratifin (SFN), which is also referred to as 14-3-3 σ , is a negative regulator of cell cycle progression with links to increased disease aggressiveness in lung cancer (Shiba-Ishii *et al.*, 2015). Cyclin B1 (CCNB1) abundance is reduced in both basal and IR treated conditions, and when associated with CDK1 forms the MPF and promotes entry of cells into mitosis (Section 1.4.3). Therefore, the reduced expression could have an impact on LZIC KO cell competency to pass through the G2/M phase of the cell cycle.

Interestingly, LZIC mRNA abundance is identified as upregulated in both conditions, but more significantly following treatment with IR. This may suggest that LZIC is specifically up-regulated in response to IR or the presence of a negative feedback loop from the protein which is lost with protein expression.

In general, the altered genes can be segregated into 2 major categories of IR dependence or independence with a total of 65 genes identified as IR independent and 96 genes identified as IR dependent (Figure 4.4B).

With the identification of multiple genes which support previously identified functions of LZIC of gene ontology identified was analysed between IR dependent and IR independent gene sets. However, prior to this analysis, validation of the most significantly identified genes was conducted to confirm the statistical analysis did not identify false positives and negatives.

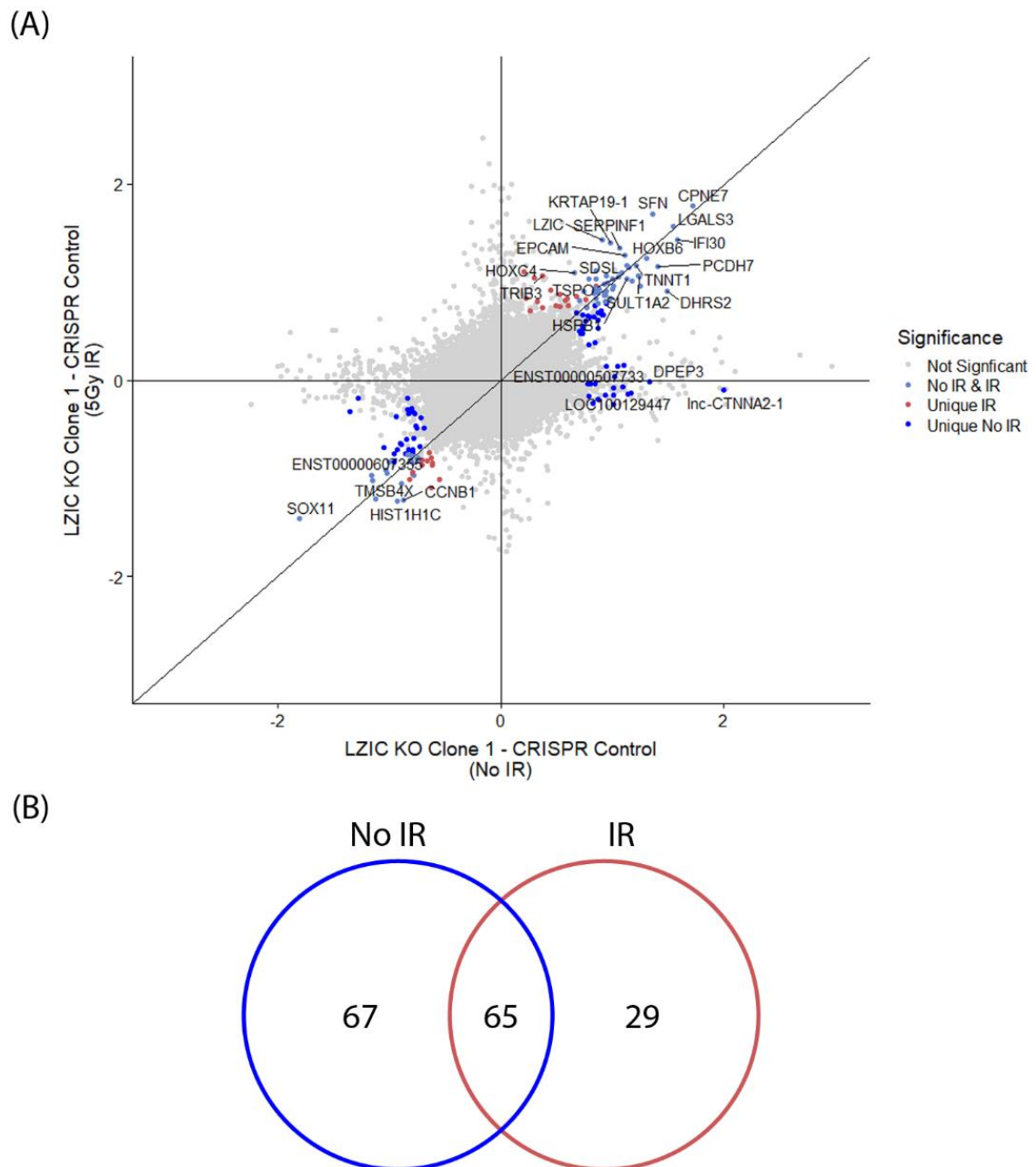


Figure 4.4 Differential expression analysis of LZIC KO Clone 1 vs CRISPR control.

(A) Differential $\log_2$ -fold changes between the two cell lines in both basal and IR treated conditions. Negative value for differential $\log_2$ -fold indicates increased expression in CRISPR Control line, with a positive value indicating increased expression in LZIC KO Clone 1. Highlighted points represent genes with a p -adj-value < 0.1 and are coloured as shown in legend. (B) Quantification of statistically significant gene numbers which are unique or shared between IR treated and untreated conditions.

4.2.5 Validation of differential gene expression analysis by qPCR

The identification of significantly altered genes both in the basal state and following treatment with IR requires validation of changes by a second method. Quantitative PCR is a typical method by which to confirm that changes observed in high throughput methods are not experimental artefacts. Ten genes were selected in total. Five of which showed increased expression: CPNE7, IFI30, LGAL, SFN and LZIC. Five genes which showed decreased expression: FLNA, NREP, POU3, SOX11, CCNB1.

The expression of all genes was consistent with the identified expression changes in the differential expression analysis (Figure 4.5). This analysis demonstrates that the result from the microarray can be validated by qPCR.

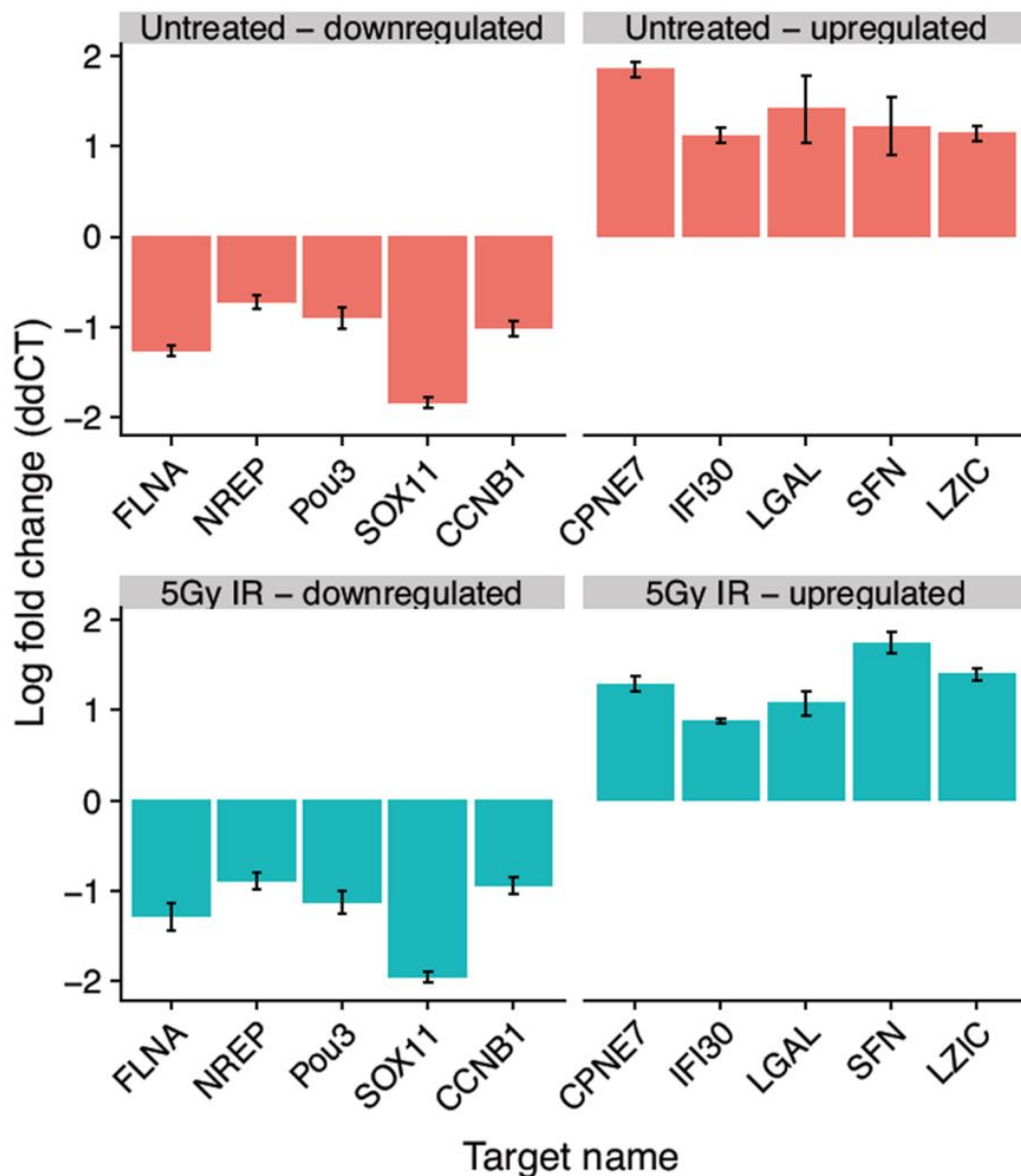


Figure 4.5 qPCR validation of significantly identified genes by differential expression analysis.

The mRNA abundance of 10 gene identified as significantly altered following loss of LZIC expression both under basal conditions and following treatment with IR were quantified by qPCR. Five genes down regulated genes and five up-regulated genes following LZIC loss were selected. GapDH was used as the loading reference and values shown are the ddCT of LZIC KO clone 1 vs CRISPR Control. Error bars represent Standard error mean. Overall, the results of this mRNA quantification match the results obtained by microarray quantification. Adapted from Skalka *et al.* 2019.

4.2.6 Gene ontology analysis of IR dependent and IR independent LZIC regulated transcripts

The analysis of enriched gene ontology groups within those LZIC dependent genes can provide invaluable information about the functional role of LZIC. The significantly altered genes identified in differential expression analysis were segregated into two groups for this analysis, IR independent and IR dependent genes.

The GO term enrichment for IR independent gene sets indicates 5 predominant clusters of terms. The largest of these clusters is the “regulation of cell proliferation” and “regulation of cell communication”. In addition to these larger groups, multiple smaller more unique groups of GO terms are identified, with these groups being enriched for processes which control organismal development (Figure 4.6A).

The resulting GO term analysis for LZIC following IR treatment shows clustering of genes involved in “cellular response to chemical stimulus” and “response to UV-A” which is a result of the treatment with IR as these groups are not present within the IR independent groups (Figure 4.6B). In addition, terms for “negative regulation of cellular processes” and “negative regulation of DNA binding” are identified which are characteristic of cellular response to IR exposure.

In general, both gene sets suggest that LZIC KO cells are pre-primed to respond differently to stimuli, whether that be treatment with IR or other chemical insults to the cell.

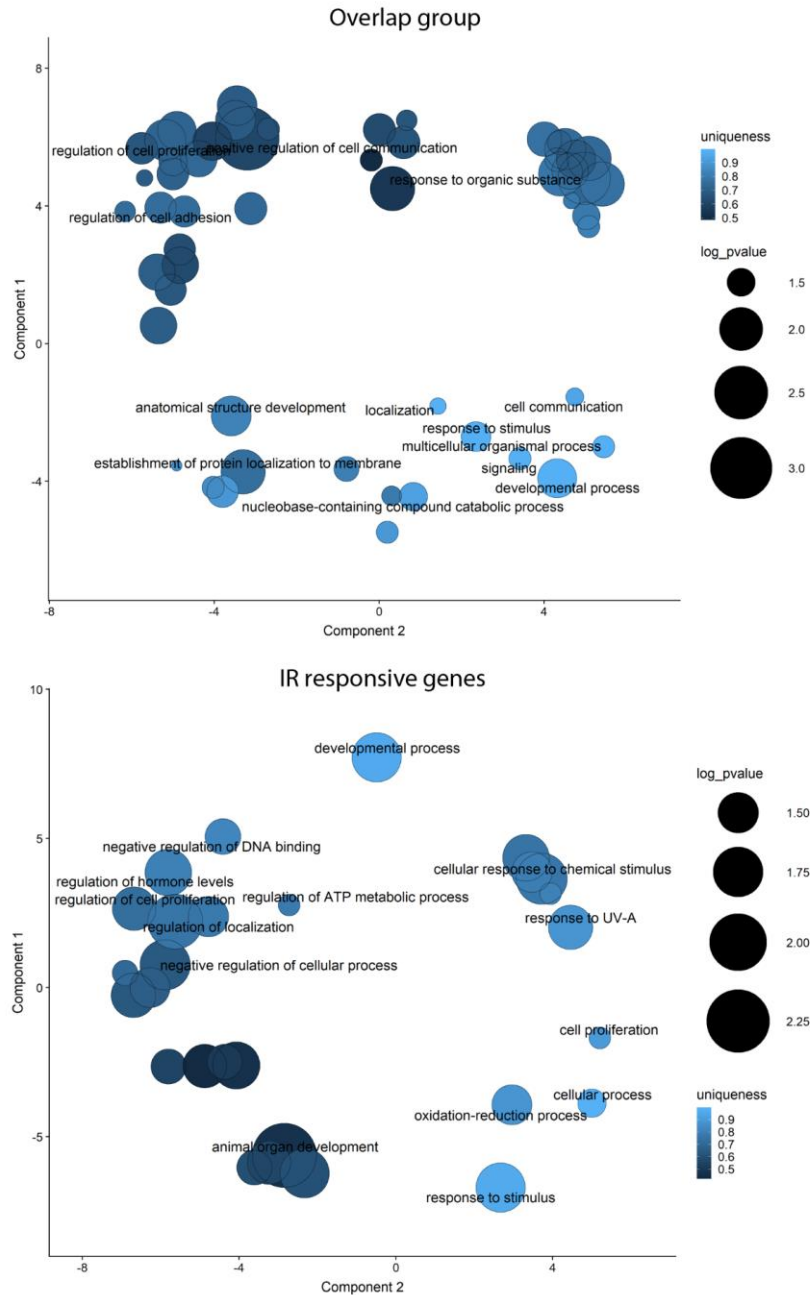


Figure 4.6 Gene ontology analysis of IR independent and IR dependent gene groups.

(A) Gene ontology analysis of all genes with significantly altered expression in both untreated and IR treated samples was conducted using STRING database, with visualisation and compressing of GO terms performed with REVIGO through SimRel. The log_{pvalue} represents the -log₁₀ p-value and defines the significance of each term identified. Uniqueness value represents the number of similar GO terms that were identified prior to SimRel. This analysis identifies the three largest GO term clusters as: regulation of cell proliferation, positive regulation of cell communication and response to organic substances. (B) All genes with significantly altered expression in both untreated and IR treated samples were analysed as in (A). The largest identified groups are: animal organ development and cellular response to chemical stimuli.

4.2.7 Z-score analysis of most variable genes following loss of LZIC expression

The differential of log-fold gene expression between conditions indicates those genes which are most regulated. However, identification of trends for increased or decreased expression levels within each of the conditions can also be established without overall increased log-fold increase or decrease through z-score analysis. This method has the benefit of overcoming initial large deviations in raw intensity values and can in some cases identify more nuanced regulatory pathways (Cheadle *et al.*, 2003).

The z-score normalisation was performed by condition, with hierarchical clustering representing gene clusters. The top 50 genes identified are shown with multiple gene clusters being identified (Figure 4.7). The first group is those genes which are down-regulated both basally and in response to treatment with IR. This identifies SOX11 and CCNB1 which were also found by log-fold change based differential expression analysis. The remaining genes identified in this group include calcitonin related polypeptide alpha (CALCA) and The JRK helix-turn-helix protein (JRK). CALCA acts as a signalling molecule in the central nervous system and is required for calcium regulation (Menon *et al.*, 2011). Whereas, the JRK protein is a positive regulator of β -catenin (Pangon *et al.*, 2016). The identification of changed expression for JRK provides an additional link between LZIC and the WNT signalling cascade.

The second group of genes is up-regulated in basal conditions and down-regulated following exposure to IR and the loss of LZIC expression. The majority of these transcripts are members of the long-non-coding RNA (lncRNA) family and do not have identified cellular functions. However, two lncRNA species were identified with previously characterised roles. Maternally expressed 3 (MEG3) is a lncRNA which inhibits tumour cell proliferation and can interact with p53 (Zhang *et al.*, 2019). RAD51-AS1 is an antisense transcript which is negatively regulated by E2F1 and has links to cancer prognosis (Zhang *et al.*, 2017).

The third group of genes are those upregulated following exposure to IR and are unresponsive to loss of LZIC under basal conditions. The exception to this trend is Dehydrogenase/Reductase 2 (DHRS2) which is upregulated in both conditions. The

DHRS2 protein catalyses the removal of dicarbonyl groups and through this activity can stabilise p53 and increase the expression of p21, thereby inhibiting the progression of the cell cycle (Deisenroth *et al.*, 2010). Activating transcription factor 3 (ATF3) shows an upregulation following treatment with IR, with no change observed in basal conditions. ATF3 is a member of the cAMP-responsive element binding family (CREB) and as such regulates the activity of transcriptional responses to extracellular stimuli (Rohini, Haritha Menon and Selvamurugan, 2018).

Overall, the z-score analysis of microarray data indicates further potential LZIC targets which could contribute to potential phenotype observed in LZIC KO cells. Of note, analysis of differentially expressed genes suggests to a possible role for LZIC in cell cycle checkpoint progression either through G1 growth factor signalling cascades, such as WNT signalling or checkpoint induction through p53 and p21.

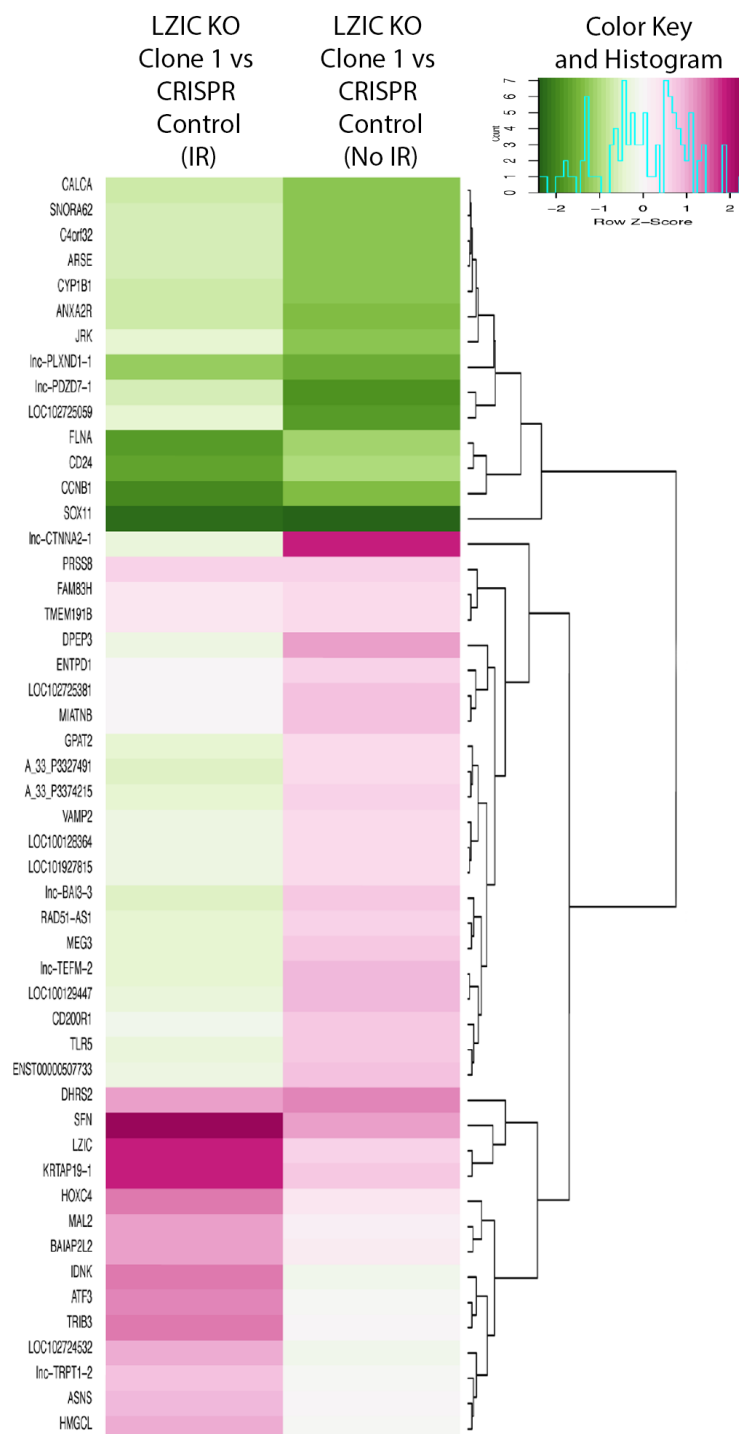


Figure 4.7 Z-score analysis of most variable genes due to loss of LZIC. Following identification of significantly altered genes between the CRISPR control and LZIC KO clone 1 under both IR treated and untreated conditions. The 50 genes with the highest variance between IR treated and untreated were compared. A z-score normalisation was conducted by column to facilitate comparison, with hierarchal clustering of grouped genes sets. The most variable genes can be clustered into 4 cohorts depending on the expression profile. Adapted from Skalka *et al.* 2019.

4.2.8 Gene set enrichment analysis of CRISPR control and LZIC KO gene expression profiles.

Gene set enrichment analysis utilises the full breadth of gene expression values without weighting for significance to determine overall pathway function alteration (Subramanian *et al.*, 2005). The determination of significance for pathway enrichment is then represented by FDR q-value.

In general, multiple critical pathways for regulation of growth and differentiation were identified prior to exposure to IR (Figure 4.8A). MYC targets are the most confidently identified pathway with the dysregulation of MYC targets being linked to the induction of multiple oncogenic pathways (Dang, 2012). In addition, altered regulation of the WNT and NOTCH signalling cascade are identified which further suggests that LZIC has a role in the regulation of not only transcriptional responses but also a specific impact on WNT signalling functionality. The identification of E2F targets and G2/M checkpoint could suggest that the regulation of these processes is dysregulated in basal conditions following the loss of LZIC.

The GSEA analysis was extended to gene expression following treatment with 5Gy IR. This indicates that as with untreated samples MYC signalling is the most significantly affected pathway (Figure 4.8B & 4.8C). However, in contrast to basal conditions in which E2F targets and TGF- β signalling are the second and third most significantly altered pathways, respectively, after IR instead G2/M checkpoint and WNT/ β -catenin signalling are the second and third most significantly altered pathways. The increased confidence of WNT signalling and mitotic spindle dysregulation may represent an increase in the number of genes within these pathways perturbed following treatment with IR, this could indicate a more severe cell cycle deficiency during the late stages of the cell cycle following IR.

Overall, these results indicate a possible alteration of MYC signalling and G2/M checkpoint signalling in LZIC KO cells in response to IR, with alterations to multiple mitogenic signalling pathways under basal conditions.

(A)

Hallmark Gene Group	FDR q-value
MYC targets	0.027
E2F targets	0.067
TGF- β signaling	0.092
Notch signaling	0.111
WNT/ β -catenin signaling	0.112
G2/M checkpoint	0.153
IL2/STAT5 signaling	0.195
DNA repair	0.195
Unfolded srotein response	0.201
Mitotic spindle	0.227

(B)

Hallmark Gene Group	FDR q-value
MYC targets	0.005
G2/M checkpoint	0.009
WNT/ β -catenin signaling	0.058
Androgen response	0.118
DNA repair	0.164
Notch signaling	0.204
Hedgehog signalling	0.241
TNF α signaling via NFkB	0.212
Mitotic spindle	0.208
P53 pathway	0.390

(C)

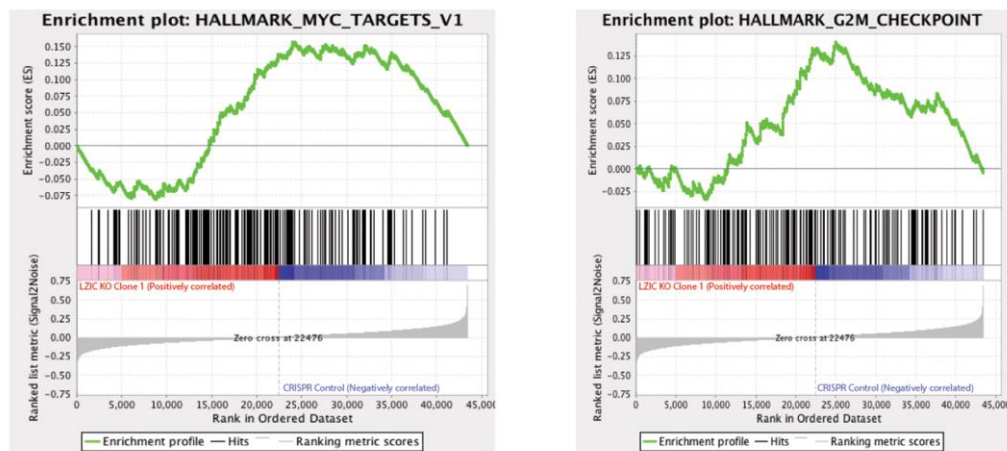


Figure 4.8 Gene set enrichment analysis of CRISPR control and LZIC KO gene expression profiles.

(A) Gene set enrichment analysis was conducted using T-statistic. Represented are the top 10 altered GSEA pathways for basal transcriptional alterations following LZIC loss. This shows that MYC targets is the most enriched pathway within the analysis. (B) The analysis was conducted as in part (A) and the top 10 altered GSEA pathways for IR induced transcriptional alterations following LZIC loss. The most enriched pathways are MYC targets and G2/M checkpoint following IR exposure. (C) Barcode plots for MYC targets and G2/M checkpoint demonstrating the distribution of enrichment changes across the entire dataset. Adapted from Skalka *et al.* 2019.

4.3 Discussion

The canonical WNT signalling pathway is critical for transduction of mitogenic signalling from cell surface receptors to the nucleus to affect gene expression alterations (Section 1.8). LZIC was originally categorised as a putative WNT signalling factor due to the presence of a well-conserved ICAT-like domain. However, whether LZIC affects the WNT pathway is not clear. Analysis of the LZIC interactome in the previous chapter identified enrichment for components of the RNA Polymerase II/Spliceosome complex (Section 3.2.15). Therefore, this chapter aimed to identify the impact of LZIC loss upon the basal cellular transcription. The loss of LZIC protein expression is associated with the development of IR induced osteosarcoma (Daino *et al.*, 2009). Therefore, the impact of LZIC on the late transcriptional response to IR exposure was also assessed.

The previous chapter identified enriched interaction with proteins that contain ankyrin domain and armadillo-like repeats which are canonical domains within WNT signalling proteins (Chapter 3). However, no functional impact upon the WNT signalling cascade was assessed. Loss of LZIC expression leads to significant alteration of WNT signalling and MYC signalling, which suggests a direct or indirect involvement of LZIC with these pathways. This is supported by the identification of individual factors with known roles in WNT signalling. Specifically, PCDH7 and JRK expression are altered following LZIC deletion. Direct regulation of β -catenin by JRK provides a possible mechanism of LZIC connection to canonical WNT signalling (Benchabane *et al.*, 2011). Further work should assess the activity of β -catenin in LZIC KO cell lines to determine whether a regulatory loop exists between the two proteins.

The exposure of cells to IR has multiple pathological consequences. One of the initial cellular impacts of IR exposure is increased reactive oxygen species (ROS) production by the mitochondria (Yamamori *et al.*, 2012). This elevated ROS production coupled to the direct damage to DNA can cause the formation of single-stranded and double-stranded DNA breaks. In order to prevent the pathological consequence of progression through the cell cycle with DNA damage, cell cycle checkpoints are activated. As previously discussed, alterations to c-MYC signalling cascades are positive drivers of cell proliferation with the potential to force cells through the checkpoints and so are typically repressed following treatment with IR both *in vitro* and *in vivo* (Huang, Traugh

and Bishop, 2004; Manning *et al.*, 2017). The identification of altered c-MYC signalling following LZIC loss suggests a role regulating this pathway. This could be mediated through the canonical function of WNT signalling or through altering the degradation of MYC in response to IR, and further work should address this question (Kim *et al.*, 2011).

The G2/M checkpoint is strongly associated with IR treatment of cells, especially in those which are lacking a functional G1 checkpoint (Fernet *et al.*, 2010). The alteration of G2/M checkpoint could indicate that this checkpoint is perturbed in LZIC KO cell following the treatment with IR. Interestingly, the G2/M checkpoint also requires p53 dependent signalling for some facets of activation. The identification of altered DHRS2 in LZIC KO cells could be a cause of significant p53 signalling divergence, which may impact upon checkpoint fidelity. The predominant function of p53 is G1 checkpoint activation and regulation, however, due to the transformation with adenovirus E1a and E1b, HEK293 cells have an altered G1 checkpoint response. Therefore, further work utilising cell lines with functional G1 checkpoint to assess the impact of LZIC loss upon transcriptome could yield critical information about the breadth of LZIC functions.

WNT signalling is important for neuronal proliferation and differentiation, predominantly through canonical extracellular WNT signalling factors (Mulligan and Cheyette, 2012). In particular, the development of the midbrain is reliant upon the concerted signalling activity of WNT3a, WNT1 and WNT5a (Mattes *et al.*, 2012). The LZIC KD induces apoptosis of midbrain neurones during development (Daino *et al.*, 2009). This study has identified alteration of protocadherin 7, Inc-CTNNA2, and SOX11, which are all regulators of the WNT signalling cascade and their loss is commonly associated with neuronal developmental defects when the expression is lost (Uemura and Takeichi, 2006; Kormish, Sinner and Zorn, 2010). The identification of differential regulation of these genes could provide functional insight into the mechanism by which LZIC causes defects in midbrain formation.

Cyclin B1 is a critical component of the MPF (Section 1.4.3). The finding that RNA levels of cyclin B1 are altered prior to treatment of cells with IR may suggest that the MPF is partially functional in basal conditions. The cyclin B1 gene has a homolog, cyclin B2, which has a redundant function and can recover the phenotype of reduced cyclin B1 in actively replicating cells (Bellanger, de Gramont and Sobczak-Thépot, 2007).

A striking finding during the differential expression analysis was the identification of LZIC as upregulated following IR treatment. The CRISPR process successfully perturbed the protein expression of LZIC, however, this process did not prevent transcription of the LZIC mRNA (section 3.2.8 & section 4.2.4). The increased expression of LZIC following treatment of cells with IR could indicate an incomplete negative feedback loop in LZIC KO cells which is perturbed when LZIC is lost. The regulatory regions surrounding LZIC could also be altered by the mutation of LZIC coding sequence.

The PCA analysis indicates that the largest source of variation between the samples was due to the different array chips used within the experiment. The quality of extracted RNA was high and so the variation may have been introduced when samples were loaded onto arrays. The issue of small replicate numbers which can be significantly impacted by technical variance elements is well documented and multiple techniques have been developed to attempt to resolve these issues. Some of these have been employed during the normalisation steps in this chapter (Bolstad *et al.*, 2003; Ritchie *et al.*, 2015). In addition, to the technical caveats this study has only utilised a single LZIC CRISPR knock-out clone. Therefore, in further work the transcriptome changes which have been identified in this study should be validated in a second knock-out clone to assure that the changes are not clone specific.

This chapter has successfully identified an LZIC dependent transcriptome for both basal conditions and those genes which are linked to LZIC following IR treatment. However, to improve the analysis further work would utilise next-generation sequencing technologies to assess the changes to splice variants of individual genes. Given the implications of LZIC for neuronal development and tumorigenesis, the transcriptomes within these contexts would be analysed to determine contextual dependent LZIC transcription.

Chapter 5

The Role of LZIC in Regulation of Cell Cycle Following Genotoxic Stress

5.1 Introduction

5.1.1 Activation and reversion mechanisms for the G2/M checkpoint

Cell cycle checkpoints exist to arrest cell division when conditions are not optimal for division, either due to reduced mitogen signalling or active DNA damage repair. Currently, four cell cycle checkpoints have been identified: G1 checkpoint, S-phase checkpoint, G2/M checkpoint, and the spindle assembly checkpoint. The G2/M and spindle assembly cell cycle checkpoints are critical for the preservation of genome integrity. DNA damage induction activates the PIKK family of kinases, which phosphorylate downstream checkpoint proteins: CHK1, CHK2, and PLK1 leading to cell cycle progression stall (Section 1.8). Resolution of DNA damage reverses checkpoint activation partly due to the activity of protein phosphatases that remove activating phosphorylation moieties on checkpoint proteins restoring the basal state. The two major protein phosphatase family members that control the reversion of cell cycle checkpoints are PP1 and PP2A. The overactivation of PP1 and PP2A can incorrectly reverse activation of checkpoint proteins and promote the early release. For example, the phosphorylation of PP2A, by MASTL, increases phosphatase activity and promotes entry into mitosis (Burgess *et al.*, 2010). Like many tightly regulated biological processes, the G2/M checkpoint is dependent upon thresholds of signalling. An example of this phenomenon is the link between MASTL expression and activity. Complete loss of MASTL arrests cells in G2, in comparison, when levels are partially decreased the cells proceed through the G2/M checkpoint harbouring multiple defects which then prevent them from activating the SAC (Lorca and Castro, 2012). In general, the bypass of cell cycle checkpoints is through specific signalling cascade deviations, one of which is checkpoint adaptation.

5.1.2 Checkpoint adaptation and genome instability

Checkpoint adaptation is a process whereby cells progress through the G2/M checkpoint following treatment with IR, irrespective of DNA damage resolution. This mechanism is dependent upon the loss of CHK1 and PLK1 activation, which in turn reduces inhibitory phosphorylation of the MPF complex and induces the early release of claspin complex (Yoo *et al.*, 2004; Syljuåsen *et al.*, 2006). The resulting progression of cells through the

G2/M checkpoint with DNA damage following checkpoint adaptation can lead to significant genome instability and micro-nuclei formation (Krempler *et al.*, 2007; Kalsbeek and Golsteyn, 2017). The increased genome instability induced by checkpoint adaptation has been attributed to the development of cancer (Hanahan and Weinberg, 2011). Therefore, the identification of factors which regulate this process could provide mechanistic insight into the process by which genome instability increase in cancer cells.

5.1.3 Hypothesis and aims

The altered expression of LZIC is detected in multiple cancers originating in different tissues. The previous chapters have both identified enrichment for cell cycle regulating pathways, in particular in response to treatment with IR (Chapter 3 and Chapter 4). This chapter will aim to characterise the impact of LZIC loss upon the activation of the cell cycle in response to IR and to investigate the molecular underpinning for any phenotype identified.

5.2 Results

5.2.1 Cell cycle analysis of G2/M checkpoint in response to IR exposure

Propidium iodide (PI) staining was used to assess the cell cycle distribution of LZIC WT vs LZIC KO cell lines. This compound binds to DNA and fluoresces which can be quantified by flow cytometry. The intensity of fluorescence is directly proportional to DNA content. Cells present in G1 have the lowest DNA content, with a steadily increasing DNA content through S-phase until a peak in G2. In addition to the assessment of basal cell cycle, PI staining can be used to investigate cell cycle arrest following stimuli (Kim and Sederstrom, 2015). The cell lines used in this experiment included, the parental line from which all CRISPR lines are derived, a CRISPR control line, and two separate LZIC KO clones.

The cell cycle profiles indicate that in basal condition there is no difference in cell cycle distribution between the cell lines analysed. To assess whether the treatment of cells with IR induced the canonical G2/M checkpoint arrest, all cell lines were exposed to 5Gy IR and cell cycle distribution analysed at 8hrs and 24hrs (Figure 5.1)(Gogineni *et al.*, 2011). The induction of G2/M checkpoint at 8hrs following treatment with IR in all cell lines suggests activation of the G2/M checkpoint. However, 24hrs post-exposure to IR the number of cells in the G2/M phase in LZIC KO cell lines was significantly reduced in comparison to control lines. This is observed with a concomitant increase in the number of cells progressing through the cell cycle into G1 and S-phase.

Overall these data demonstrate that the G2/M checkpoint is successfully activated in all cell lines but this is lost at the 24hr with loss of LZIC leads to an increased percentage of cells with a G1 DNA content at 24hrs post IR.

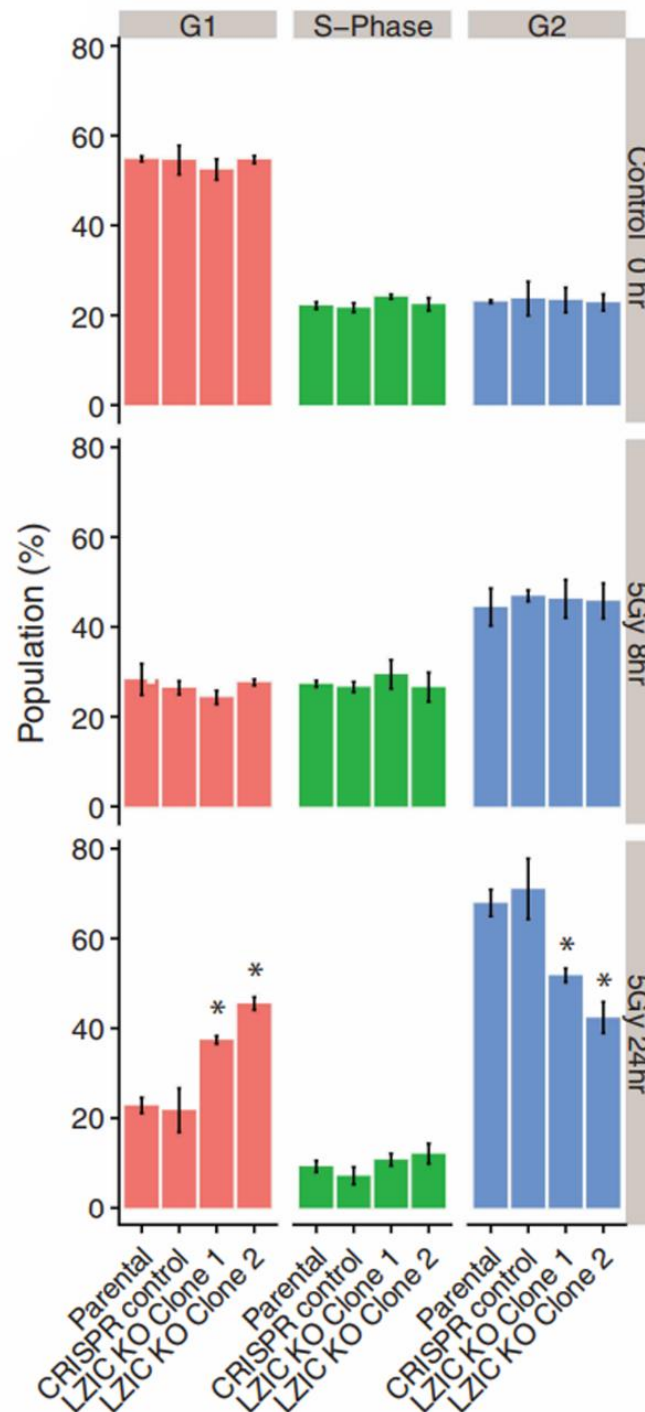


Figure 5.1 Analysis of cell cycle profile in LZIC KO cell lines exposure to IR.

All cell types were treated with 5Gy IR and incubated for 24hrs. Cells were harvested and fixed with 70% ethanol before RNase A treatment to remove RNA and staining with propidium iodide. Fixed cells were then analysed on a FACS canto and G1/S-phase/G2 populations quantified. No difference can be seen in cell cycle distribution for 0 and 8hr timepoints. At 24hrs the LZIC KO clones show a reduced G2/M population. Significance was tested by unpaired student T-Test, * = p-value < 0.05. Adapted from Skalka *et al.* 2019.

5.2.2 Cell Cycle analysis of LZIC KO cells following treatment with cellular stressors

To assess the IR specificity of the G2/M checkpoint defect observed, LZIC KO cells were exposed to a range of cellular stress and DNA damage agent. Initially, all cell types were treated with Cobalt Chloride (CC) which induces hypoxic conditions in the cell. The detection of hypoxia activates the G1 cell cycle checkpoint by the increase of p27 (Goda *et al.*, 2003; Ortmann, Druker and Rocha, 2014). The treatment with CC does not induce any changes to cell cycle profile at either 8hrs or 24hrs (Figure 5.2). However, HEK293 cells have a perturbed G1 checkpoint and so this may impact upon the ability of these cells to induce a G1 checkpoint arrest inducing toxin. To assess the impact of S-phase related damage, the topoisomerase inhibitor, camptothecin was used. This toxin induces replication stress during the S-phase of cell cycle and can induce both the intra-S-phase checkpoint and G2/M arrest (Goldwasser *et al.*, 1996). The cell cycle profile indicates a reduction in the G2/M population at the 8hr timepoint for all cell lines, with a concomitant increase in the G1 population (Figure 5.2). However, no deviation between control lines and CRISPR KO lines can be observed. Finally, the treatment of cells with ultra-violet (UV) radiation induces the formation of thymidine-dimers which lead to DSB formation, recapitulating this specific phenotype of treatment with IR (Rastogi *et al.*, 2010). An increase in the G1 population can be observed in all cell types at the 24hr time point, however, no deviation can be seen between the cell types (Figure 5.2).

Overall, this result suggests that LZIC is required specifically for response to IR as a response to other toxins is not perturbed.

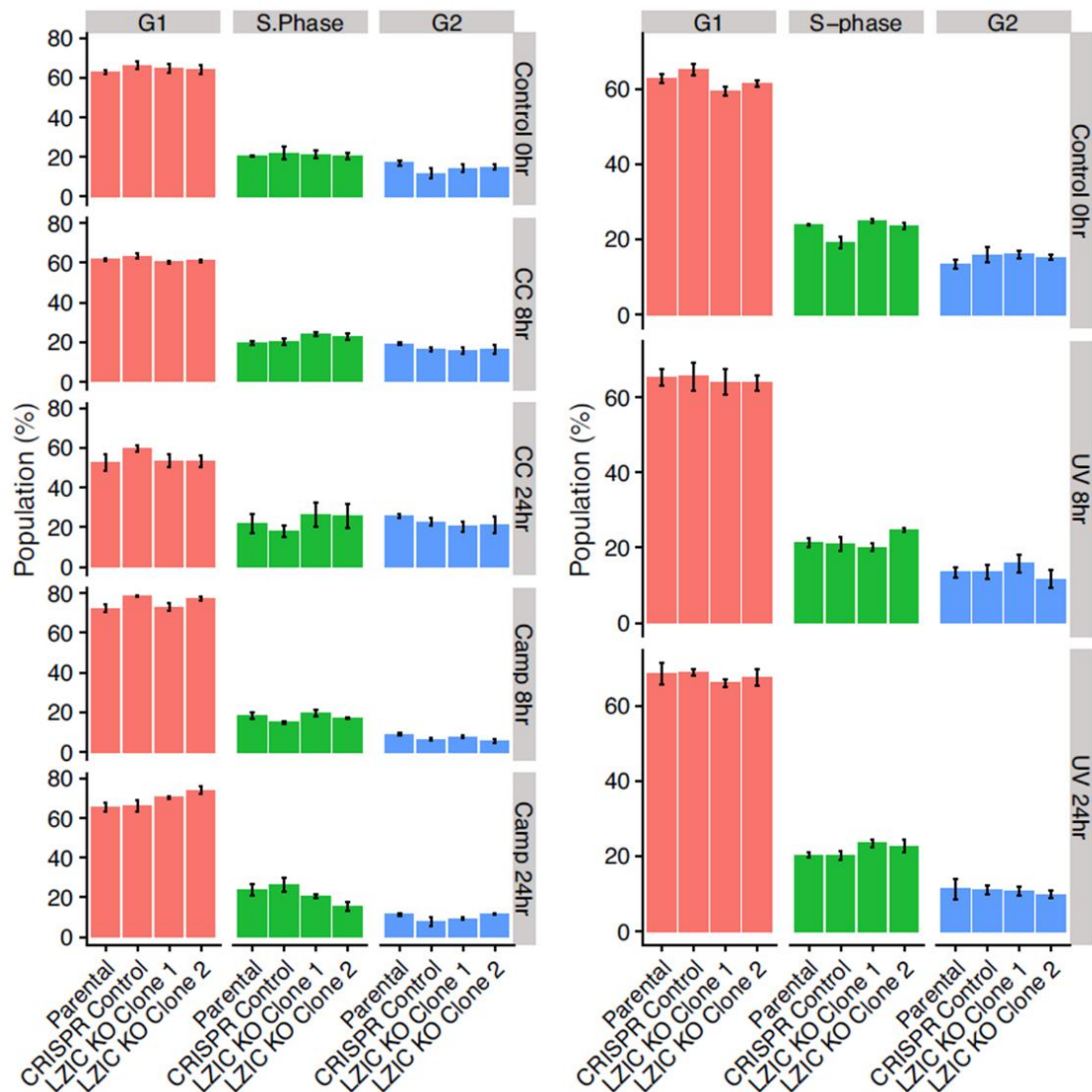


Figure 5.2 Analysis of cell cycle in LZIC KO cell lines following cellular stress and genotoxic stress.

Cells were treated with toxin compounds: Cobalt Chloride (CC), camptothecin (Camp), and Ultra-violet light (UV), 24hrs prior to harvesting and fixing with 70% ethanol. The cells were treated with RNase A and stained with propidium iodide overnight. Cell cycle distribution following treatment with the different toxins showed no difference between the different cell types. Adapted from Skalka *et al.* 2019.

5.2.3 Cell cycle analysis of LZIC KO cells following reintroduction of LZIC-c-flag

The CRISPR cut site sequencing of LZIC KO cell lines and qPCR analysis of LZIC mRNA expression indicate that the protein expression is lost but dysfunctional mRNA is still transcribed (Figure 3.9). The observed phenotype could be due to altered expression of LZIC mRNA. In addition, the CRISPR process can lead to KO of unintended genes due to non-specific cuts. Therefore, to confirm neither of the phenomena was responsible for the phenotype observed, the flag tagged LZIC protein was stably expressed in LZIC KO clone 2. The partial restoration of the G2/M checkpoint arrest in LZIC KO cells expressing LZIC-flag supports that the phenotype observed is not due to off-target effects (Figure 5.3). However, the reversion using this method seems to be expression level-dependent, as the overall expression of ectopic LZIC is lower than physiological levels.

Overall, this experiment has shown that the altered function of the G2/M checkpoint in response to IR is directly due to the loss of LZIC protein expression.

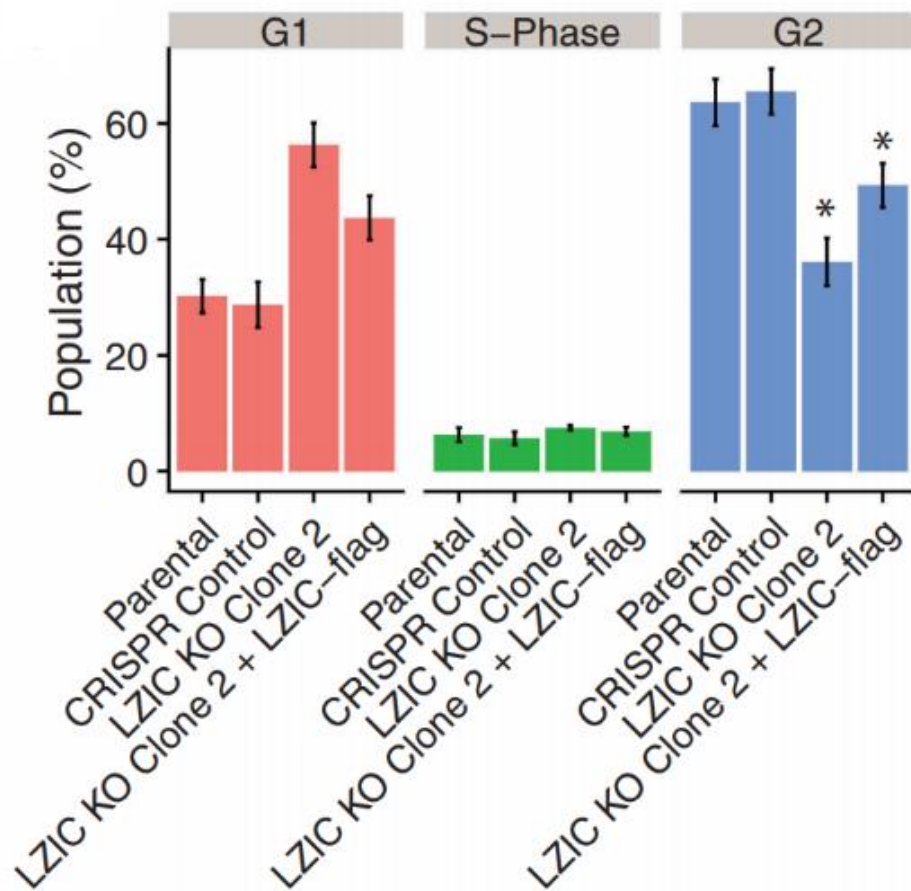


Figure 5.3 Cell cycle analysis following reintroduction of LZIC-flag in KO cell lines.

A stable sub-clone of LZIC KO Clone 2 expressing LZIC CDS with an additional C-terminal flag tag was generated for comparison to LZIC KO Clone 2. Cell lines were exposed to 5Gy IR and incubated for 24hrs prior to harvesting. The cells were treated with RNase A and stained with PI overnight before analysis for cell cycle distribution by flow cytometry. The reintroduction of C-flag LZIC restores partial function of the G2/M checkpoint. * - Unpaired student T-test with a p-value <0.05, tests used the parental line as the null hypothesis. Adapted from Skalka *et al.* 2019.

5.2.4 Quantification of mitotic marker histone 3 serine 10 phosphorylation following exposure to IR

IR induced activation of the G2/M checkpoint occurred in LZIC KO cells, however, maintenance of this arrest was perturbed. To further confirm this phenotype a direct marker of mitotic cells was analysed. Phosphorylation of serine 10 within histone 3 (pS10H3) is a characteristic marker of condensing chromosomes, which is a vital stage within mitosis (Kim *et al.*, 2017). Importantly, while chromosomes undergoing mitosis are modified with this marker it is then lost upon entry into G1. Staining of pS10 H3 is possible for analysis of cells either by FACS or immunofluorescence to determine the size of the mitotic population. Immunofluorescence followed by manual quantification of positive cells was utilised to determine differences in the size of the mitotic population between the cell lines. At 24hrs post-IR the mitotic population is reduced in LZIC KO cells when compared to control cells (Figure 5.4). This data combined with the cell cycle indicates an increased population of cells have left mitosis and entered G1 phase.

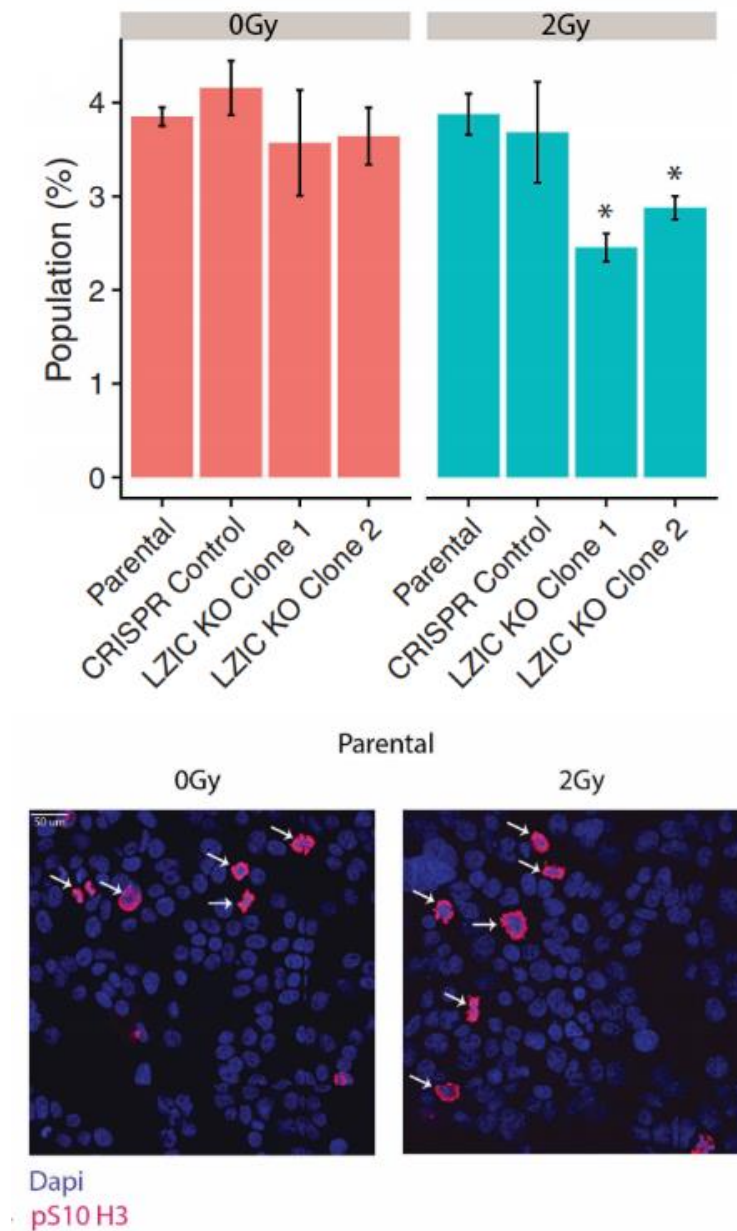


Figure 5.4 Analysis of pS10 H3 immunofluorescence in LZIC KO clones following 2Gy ionising radiation.

Cell lines were treated with 2Gy IR and incubated for 24hrs before harvesting. Cells were subsequently fixed and permeabilised for immunofluorescence and stained for pS10 H3 (Red) and nucleus with DAPI (Blue). Result of this indicates that the population of pS10 positive cells decreases in LZIC KO cells following IR. Significance determined by unpaired student T-test. * = p-value < 0.05. Adapted from Skalka *et al.* 2019.

5.2.5 Analysis of Early G2/M checkpoint activation in response to IR exposure

The treatment of cell with IR induces activation of two distinct G2/M checkpoints, referred to as the early and late checkpoint (Xu *et al.*, 2002). The early checkpoint is ATM-dependent and predominantly affects those cells already in G2 when damage is induced. This arrest of G2 cells causes a rapid reduction in the number of cells passing through mitosis, however, the impact is minimal when compared to the late G2/M checkpoint. Detection of early checkpoint activation can be conducted by flow cytometry analysis of pS10 H3 levels up to 4 hours following exposure of cells to IR.

The treatment with IR leads to a reduction of mitotic ratio for all cell lines, apart from the parental line additionally treated with ATM inhibitor which shows reduced response across the time course but most strikingly at the 1hr time point (Figure 5.5). The ATM inhibitor sample was included to indicate the maximal extent of deviation that can occur at this checkpoint following the loss of ATM function. The comparison of control lines to LZIC KO lines indicates consistent induction of the early G2/M checkpoint, which is supported by the unaltered cell cycle profile at the 8hr timepoint as observed by PI staining.

Overall, this result indicates no early G2/M checkpoint deficiency in LZIC KO cell lines and suggests that regulation of the late time point is selectively perturbed.

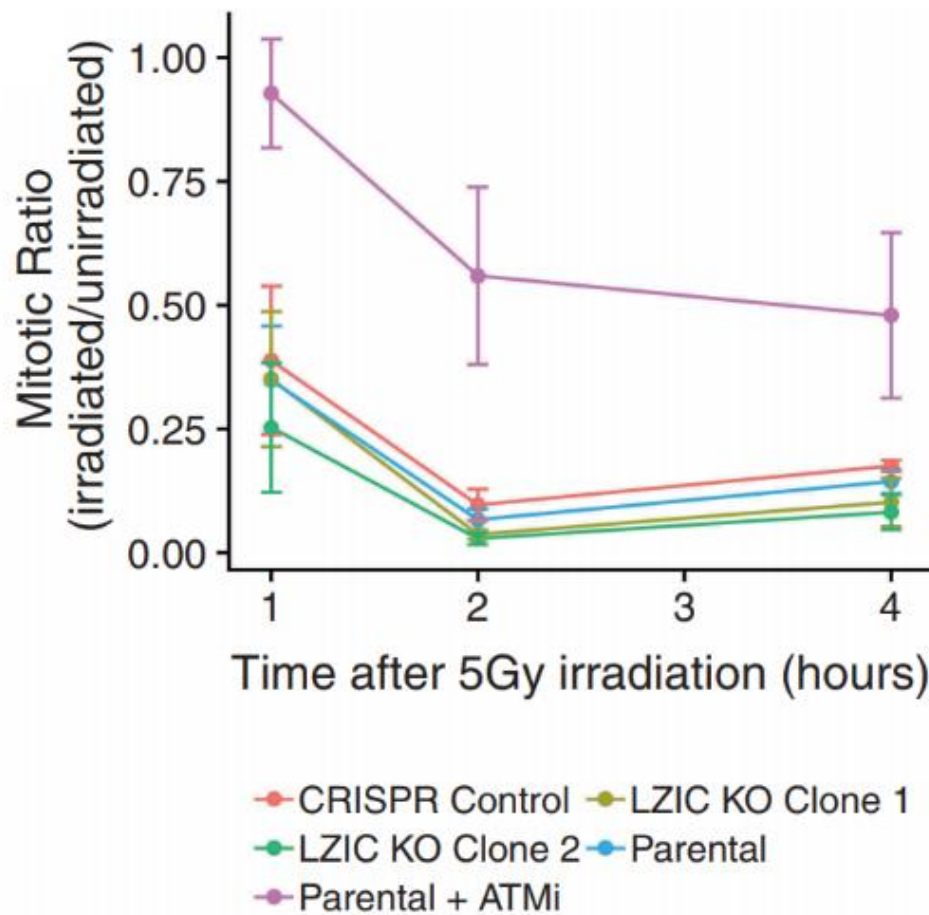


Figure 5.5 Early G2/M checkpoint response to IR exposure.

All cell lines were treated with 5Gy IR and at time points indicated harvested. The cell pellets were subsequently fixed and blocked with BSA. Fixed cell pellets were dual stained with antibody specific to phosphorylated serine10 of histone 3 and DNA content by propidium iodide. Mitotic ratio represents the ratio between pS10 H3 positive cells in untreated vs treated cells specifically in G2/M at the time points shown. Overall, there is no difference of early G2/M checkpoint activation in response to IR exposure following LZIC protein expression loss, however, the inhibition of ATM successfully perturbs the checkpoint. Adapted from Skalka *et al.* 2019.

5.2.6 p53 phosphorylation status following IR exposure

The critical role of p53 for the regulation of cellular response to toxic insult has been examined in depth (Section 1.6.1). The predominant modification site is serine 15 which is phosphorylated by ATR and ATM in response to DNA damage (Loughery *et al.*, 2014). Activation of p53 by phosphorylation of serine 15 occurs rapidly following DNA damage detection and so up to 4hrs following IR exposure was analysed. The treatment of cells with IR induces phosphorylation of serine 15 which incrementally increases up to the final time point of 4 hours in all cell lines (Figure 5.6). In addition, the exposure of cells to IR increases the total p53 present in the cell.

Overall, this result suggests that the altered checkpoint response of LZIC KO cells is not due to deficiency of p53 activation by members of the PIKK family immediately following exposure to IR.

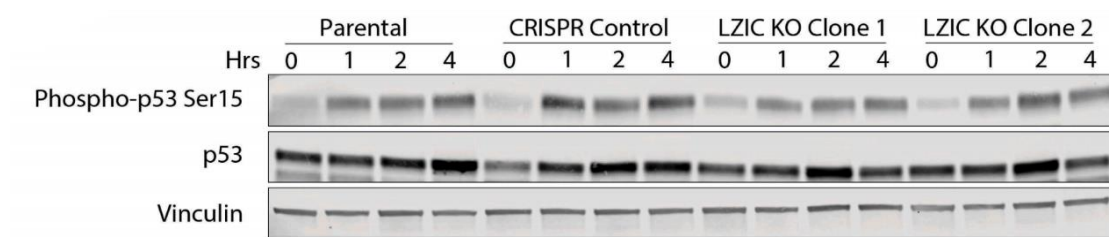


Figure 5.6 Analysis of p53 serine 15 phosphorylation status and p53 steady-state levels.

Total protein was harvested from all cell lines, utilizing RIPA buffer and disruption of chromatin with sonication, at the time points indicated following 5Gy IR exposure. Overall, there is no difference in either the total expression of p53 or the phosphorylation of p53 serine 15 in response to IR exposure. Adapted from Skalka *et al.* 2019.

5.2.7 Phosphorylation status and total protein abundance of master checkpoint regulators CHK1 and CHK2 following exposure to IR.

The premature release of cells from the G2/M checkpoint can indicate a form of checkpoint adaptation. To progress through checkpoint adaptation CHK1/CHK2 must be degraded or dephosphorylated (Syljuåsen *et al.*, 2006; Kalsbeek and Golsteyn, 2017). Therefore, the CHK1/CHK2 activation profile was determined in LZIC KO cells following damage.

Initially, CHK2 expression levels and activation was investigated. The phosphorylation of CHK2 in response to IR shows no defect in activation between control lines and LZIC KO lines. However, the G2/M checkpoint is predominantly regulated by the master kinase CHK1 in response to genotoxic stress. Thus, phosphorylation of CHK1 at the two major activation sites serine 345 (S345) and serine 317 (S317) was investigated. Modification of Chk1 at S317 is required to fully activate CHK1 functionality through a combinatorial role with the S345 site (Wilsker and Bunz, 2009). The activation of CHK1 by phosphorylation on S345 can be seen at 8hrs post IR and is consistent across all cell types. However, phosphorylation at S317 in LZIC negative cells is almost completely lost (Figure 5.7).

Defects in phosphorylation of CHK1 can prevent activation of the G2/M checkpoint and permit cells to proceed through mitosis.

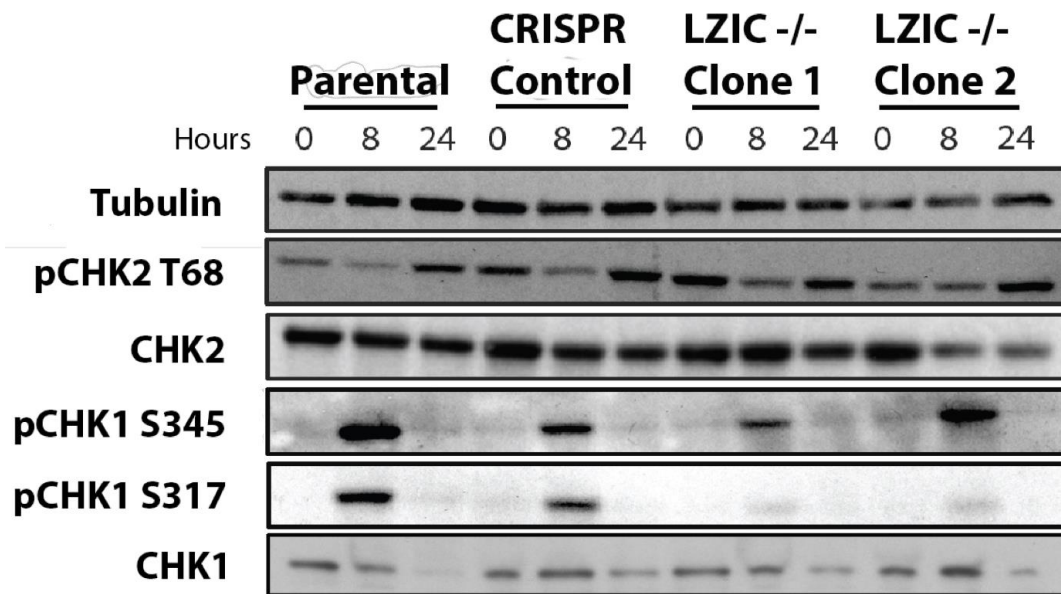


Figure 5.7 Phosphorylation status and total protein abundance of master checkpoint regulators CHK1 and CHK2 following exposure to IR.

Cell lines were treated with 5Gy IR and harvested, utilizing RIPA buffer and disruption of chromatin with sonication, at 8 and 24hrs post treatment. The resulting protein quantification demonstrates that activation of CHK2 is consistent between all cell lines but the activation of CHK1, specifically at serine 317, is perturbed following exposure to IR in LZIC null cells. Adapted from Skalka *et al.* 2019.

5.2.8 Analysis of PIKK protein activation in LZIC KO cells following IR exposure

One explanation for the severe reduction in CHK1 phosphorylation at the S317 is defective signalling from the DSB site. Three kinases are responsible for transferring signals to checkpoint proteins, DNA-PK, ATR and ATM. At later time points, ATR is the predominant kinase that controls the G2/M checkpoint and is activated by phosphorylation on tyrosine 1981 (Nam *et al.*, 2011). Analysis of ATR Thr1981 phosphorylation status following IR at 8 and 24hr timepoints indicates that the activation of this protein is maintained across all cell types (Figure 5.8). In addition, the kinase ATM functions in concert with ATR at this checkpoint and as with ATR, the activation following DNA damage is conserved across all cell types.

Overall, this data shows that the activation of PIKK proteins is correct and so indicating that LZIC loss causes pathway disruption downstream of this point.

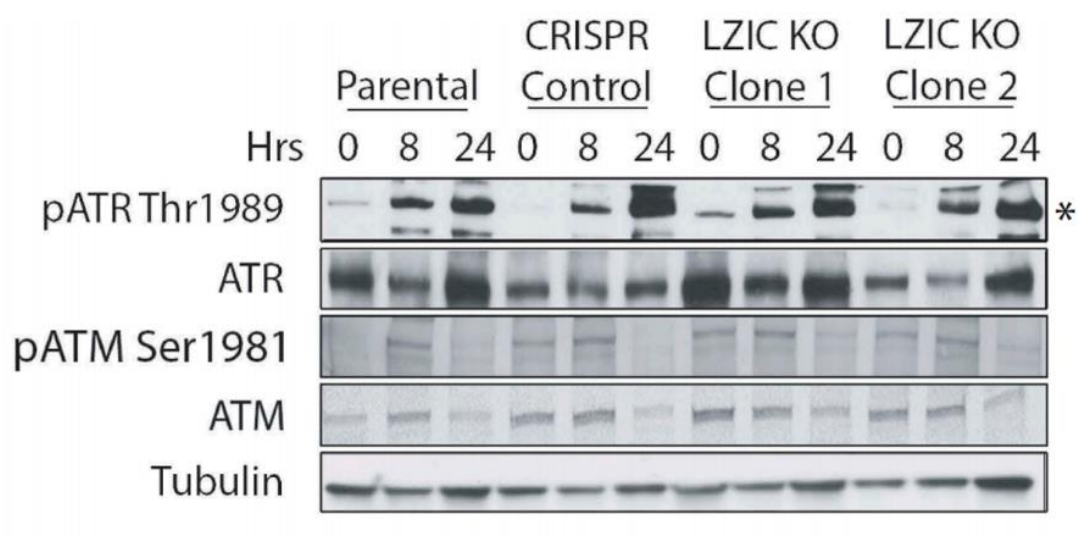


Figure 5.8 Analysis of ATM and ATR activation in LZIC KO cells following IR treatment.

The cells were treated with 5Gy IR and then harvested, utilizing RIPA buffer and disruption of chromatin with sonication, at 8 and 24hrs following exposure. Overall, the result indicates correct activation of ATM and ATR in response to IR treatment in all cell lines, irrespective of LZIC expression status. * - indicates the band for full length ATR. Adapted from Skalka *et al.* 2019.

5.2.9 Analysis of CDC2 and Cyclin B1 phosphorylation and total protein abundance changes in response to IR exposure.

The major regulators of G2/M progression are Cyclin B1 and CDC2, through the formation of the MPF. Activation of the MPF kinase activity is tightly linked to phosphorylation of specific activator sites, such as the CDC2 T-loop threonine 170, and removal of inhibitory phosphorylation moieties. Activation of the DNA damage cascade will increase the phosphorylation of inhibitory sites predominantly through CHK1/WEE1 and the inhibition of CDC25C activity. The decreased phosphorylation of CHK1 S317 could indicate a reduced kinase activity, which in turn would be represented by decreased inhibitory phosphorylation of the MPF.

The cyclin B1 serine 147 site is phosphorylated in response to DNA damage and induces differential sub-cellular localisation (Hagting *et al.*, 1998). A comparison of this phosphorylation site between all the cell lines indicates a reduced and shifted phosphorylation response of cyclin B1 at serine 147 to IR exposure following LZIC depletion (Figure 5.9). In addition, the total expression of cyclin B1 changes cyclically with cell cycle phase and expression peaks in the G2/M phase. Upon DNA damage if a G2/M checkpoint arrest is induced the levels of cyclin B1 will increase until the release from the checkpoint is completed. Interestingly, cyclin B1 total protein levels are decreased in LZIC KO clones compared to control lines.

A second DNA damage responsive site within the MPF is found on CDC2 at tyrosine 15. This is the site of inhibitory phosphorylation mediated by WEE1 and prevents the activity of the MPF complex (Welburn *et al.*, 2007). Following exposure of control lines to DNA damage the phosphorylation of this site increases. However, the phosphorylation of this site shows little response to treatment with IR in LZIC KO cells (Figure 5.9). The overall, expression of CDC2 across all the cell types is unaltered and therefore differential activity is likely due to phosphorylation state regulation.

These data indicate that the major regulators of the G2/M progression are not inhibited by appropriate phosphorylation sites responsible for the cessation of cell cycle allowing aberrant progression through mitosis to G1, which supports the data collected during cell cycle analysis and phospho-serine10 H3 staining.

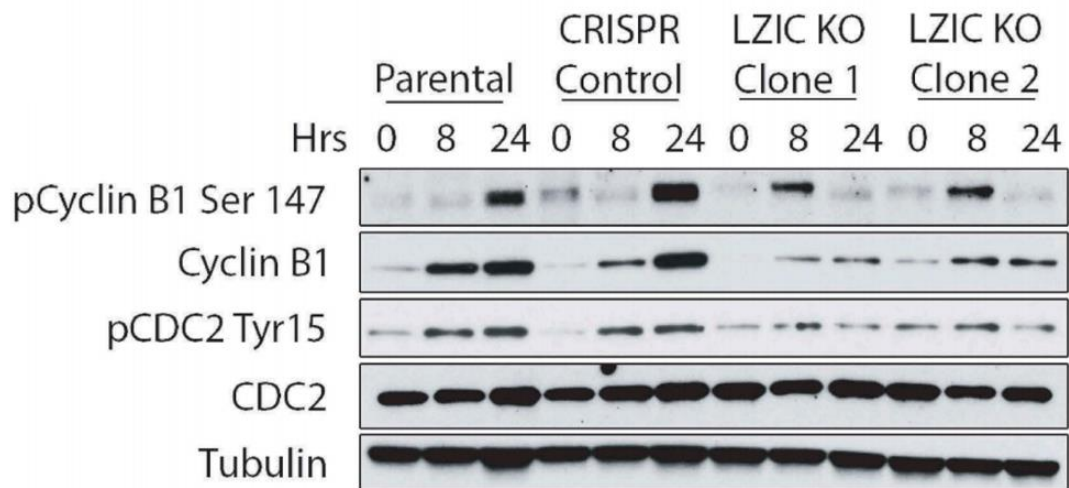


Figure 5.9 Analysis of phosphorylation and steady state levels of Cyclin B1 and CDC2 following exposure to IR.

Cells were treated with 5Gy IR and harvested, utilizing RIPA buffer and disruption of chromatin with sonication, at 8 and 24hrs following exposure. Overall, the total protein levels of cyclin B are reduced in LZIC KO conditions compared to control cell lines. In addition, the phosphorylation of inhibitory site serine 147 is temporally shifted in comparison to control lines following treatment with IR. The protein expression of CDC2 is consistent between conditions and cell lines, however, phosphorylation of the inhibitory site, tyrosine 15, is reduced in LZIC KO cell lines following IR exposure. Adapted from Skalka *et al.* 2019.

5.2.10 PP1 and PP2A total protein abundance following IR exposure

The western blot analysis of CHK1 indicated that while phosphorylation on Serine 345 is maintained, the phosphorylation of S317 is largely lost. The modification of this site is regulated by the kinase ATR and the phosphatase PP2A and PP1 (Freeman and Monteiro, 2010). Therefore, the incorrect phosphorylation status of this site could be predominantly due to two mechanisms. Either lack of ATR activation in response to DNA damage events or the altered activity of PP2A/PP1 leading to the release from checkpoint arrest too quickly. Analysis of the PIKK protein family members indicates correct activation in response to IR exposure (Figure 5.8). Therefore, altered activity or abundance of the protein phosphatase family members may be responsible for the reduced phosphorylation of MPF components. PP1 core enzyme constitutes a single main unit with the association of different subunits altering activity on substrates. PP2A a heterotrimeric complex which is comprised of three subunits: A, B, and C (Nilsson, 2019). The expression levels of PP1 and PP2A subunits were analysed to determine if over-expression of these proteins was observed. This analysis demonstrates that total expression levels of both PP1 and PP2A are the same between all cell lines analysed across a time course of 24hrs post exposure to IR (Figure 5.10).

These data suggest that the loss of phosphorylation at DNA damage related target sites of the MPF is not due to overexpression of the PP family, however, next the activity of the phosphatase family was analysed.

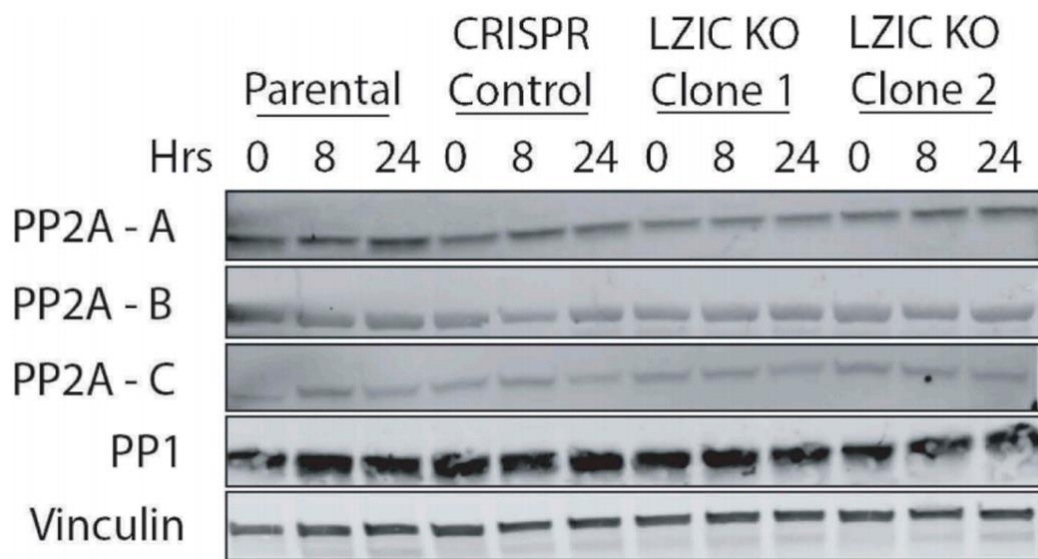


Figure 5.10 Expression levels of protein phosphatase 2A subunits and protein phosphatase 1 following IR exposure.

Cells were treated with 5Gy IR and harvested, utilizing RIPA buffer and disruption of chromatin with sonication, at 8 and 24hrs. Overall, the quantification of PP1 and the subunits of PP2A indicated no alteration to expression levels following exposure to IR or loss of LZIC protein expression. Adapted from Skalka *et al.* 2019.

5.2.11 Cell cycle profile of cells following IR exposure and PP2A inhibition.

The expression of phosphatases is not altered in LZIC KO cell lines, basally or following exposure to IR. However, the activity of the protein phosphatase family members can vary significantly without changes in total expression. To assess whether the reduction of phosphatase activity could impact the phenotype, okadaic acid (OA), a PP2A inhibitor was utilised.

Cells lines, were treated with 5Gy IR as previously either with or without additional exposure to okadaic acid (Figure 5.11). The addition of OA leads to partial restoration of the G2/M checkpoint function in IR treated cells. LZIC KO clone 2 cells show an increase of G2/M population from 45% to 53.3% and a drop in the G1 population. This is also consistent with LZIC KO clone 3 cells which show an increase in G2 population from 33.9% to 43.9%. The treatment of control lines with OA in addition to IR has no impact cell cycle profile supporting the competitive effect with LZIC KO. As can be seen there is an increase in G2/M arrest in LZIC KO clones, however, this is not to the level seen following treatment with 5Gy IR.

Overall, these data indicate that the inhibition of PP2A function partially restores the functionality of the G2/M checkpoint in LZIC knock-down cells. This suggests that loss of LZIC leads to overactivation of PP2A.

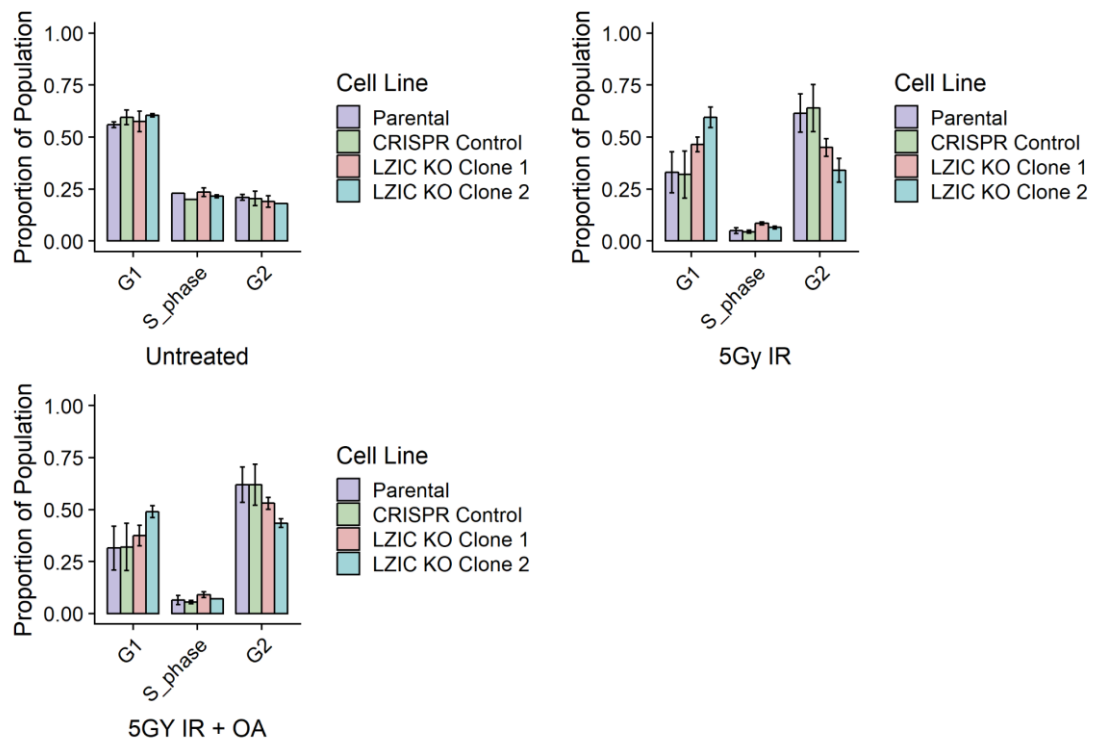


Figure 5.11 Cell cycle distribution analysis following treatment with OA and IR.

Cells were treated with 10nM okadaic acid 1hr prior to exposure to 5Gy IR and harvested at 8 and 24hrs following. The cells were fixed/permeabilised before treatment with RNase A and staining with propidium iodide. Overall, these results suggested that pre-treatment with okadaic acid could revert the loss of G2/M checkpoint activation following IR exposure in LZIC KO cell lines. Error bars represent standard error mean (SEM)(N=2).

5.2.12 Hypothesised LZIC position in G2/M control pathway

The analysis of LZIC KO cells suggests a function for LZIC within the control of G2/M checkpoint activation (Figure 5.12). Following the induction of DNA damage and detection of DNA damage in both WT and LZIC KO cells, ATR and ATM are activated. At this stage of the signalling cascade, the WT and LZIC KO cells response diverges. While in WT cells activated PIKK phosphorylates CHK1 at positions S345 and S317 and the phosphorylation of Cyclin B1 and CDC2 is increased stabilising protein levels which prevent progression into mitosis. In LZIC KO cells phosphorylation of CHK1 is lost and thereby the stabilisation of MPF factors. The data collected so far indicate that ATR signalling is not lost in LZIC KO cells, however, the phosphorylation of CHK1 is prematurely lost. Therefore, this suggests PP2A is activated which maintains CHK1 in a hypo-phosphorylated state and promotes progression into mitosis. This is further confirmed by initial data with okadaic acid treatment.

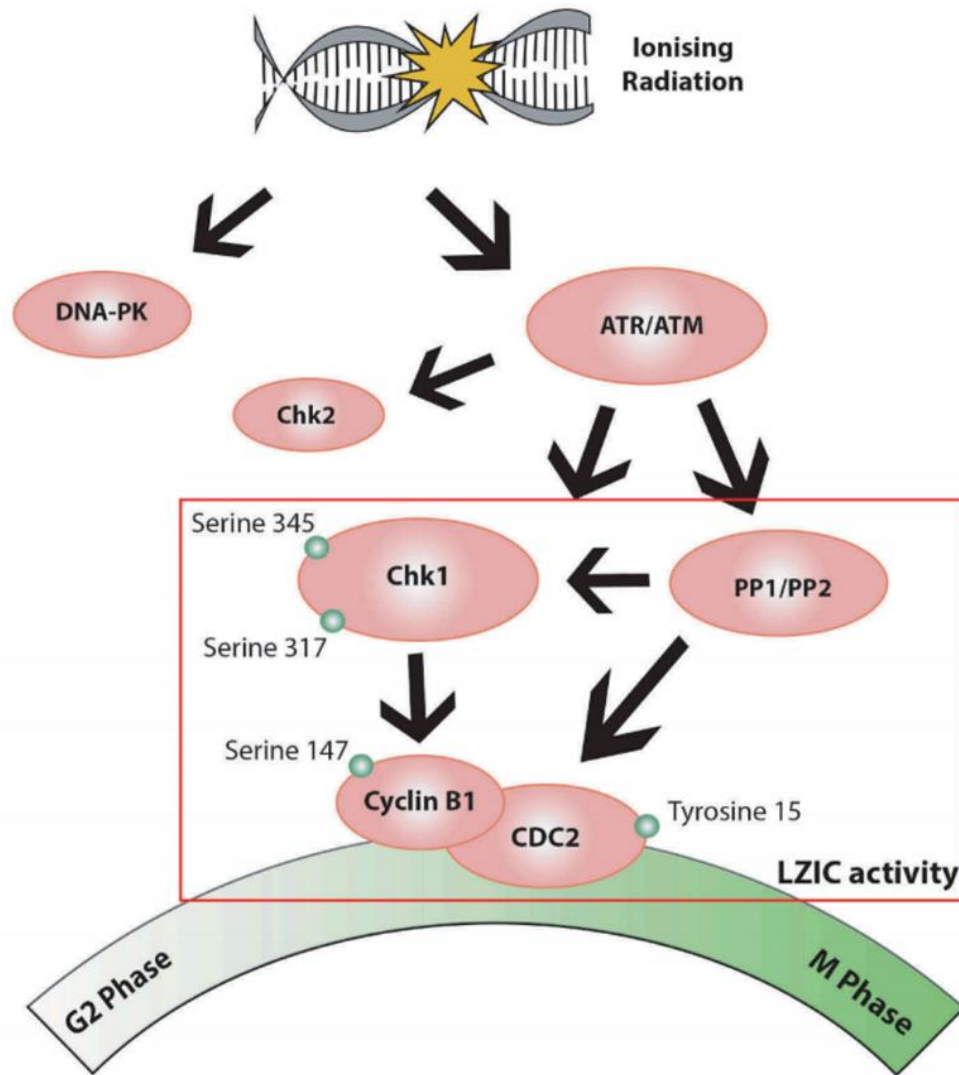


Figure 5.12 Schematic of LZIC function within the G2/M checkpoint signalling cascade.

Upon exposure of a cell to IR DNA damage can occur, this will activate the PIKK family members – DNA-PK, ATM, and ATR. The activation of G2/M checkpoint cascade is predominantly controlled by ATR and ATM, which activate CHK1 through phosphorylation of Serine 345 and 317. The activated CHK1 in turn phosphorylates Cyclin B1 at serine 147 and CDC2 at tyrosine 15, these are damage activated sites which prevent progression of cell cycle. The overall phosphorylation levels is controlled by the interplay between PIKK activity and the PP2A Phosphatase activity. Current experimentation places LZIC within the cascade at the level of phosphorylation regulation on CHK1. Adapted from Skalka *et al.* 2019.

5.2.13 Development of defective G2/M checkpoint induced aneuploidy

The loss or gain of chromosomes is a state known as aneuploidy and is a common phenotype associated with the process of tumorigenesis (Figure 5.13). Many mechanisms contribute to the development of aneuploidy including defects in the induction and maintenance of the G2/M checkpoint. Upon treatment with IR, a typical cell will arrest at the G2/M transition until DNA repair has been completed. The outcome of this will be to maintain the correct number of chromosomes. However, following dysfunction of the G2/M checkpoint the cells will progress through the G2/M with incompletely repaired DNA increasing the chances of chromosomes mis-segregating.

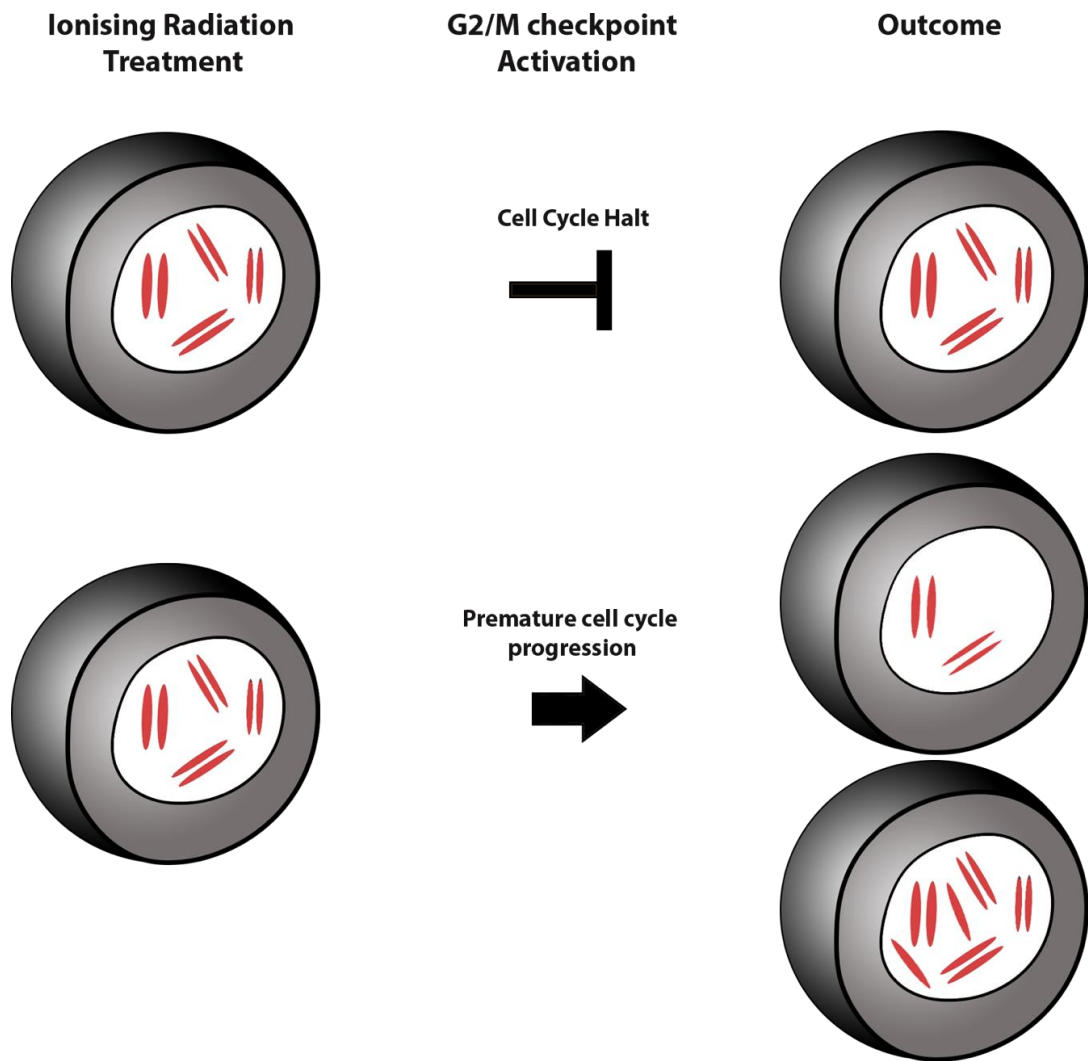


Figure 5.13 Schematic for the development of aneuploidy following the disruption of G2/M checkpoint.

Upon treatment with IR cellular response will arrest a population of cells at the G2/M checkpoint. Arrest will be sustained until DNA repair has occurred. This prevents the cells from miss segregating chromosomes during mitosis and maintains the correct number of chromosomes. However, disruption of the G2/M checkpoint prematurely releases cells into mitosis prior to break repair. The chromosomes can miss segregate generating aneuploidy.

5.2.14 Metaphase spread analysis of LZIC KO cells karyotype basally and following exposure to IR

An altered number of chromosomes can be quantified by the technique called metaphase spread analysis. This allows visualisation of chromosomes by immunofluorescence or light microscopy and quantification.

The analysis of chromosomes numbers in all cell lines both before treatment with IR and additionally 48hrs after treatment. The resulting chromosome numbers indicate that the loss of LZIC causes a reduction of chromosome numbers from 61 to 51 on average (Figure 5.14). Indicating an increase in chromosomal instability following LZIC loss. Analysis of cells 48hrs after treatment with 2Gy IR shows that in all cell types the number of chromosomes decreases, however, there is no difference between the cell types at this point. This may indicate a threshold of aneuploidy which is not exceeded in LZIC cells with the addition of IR.

Overall, these data indicate that LZIC KO induces severe genome instability in basal conditions. However, the genome instability does not increase following IR compared to parental and CRISPR control cell lines.

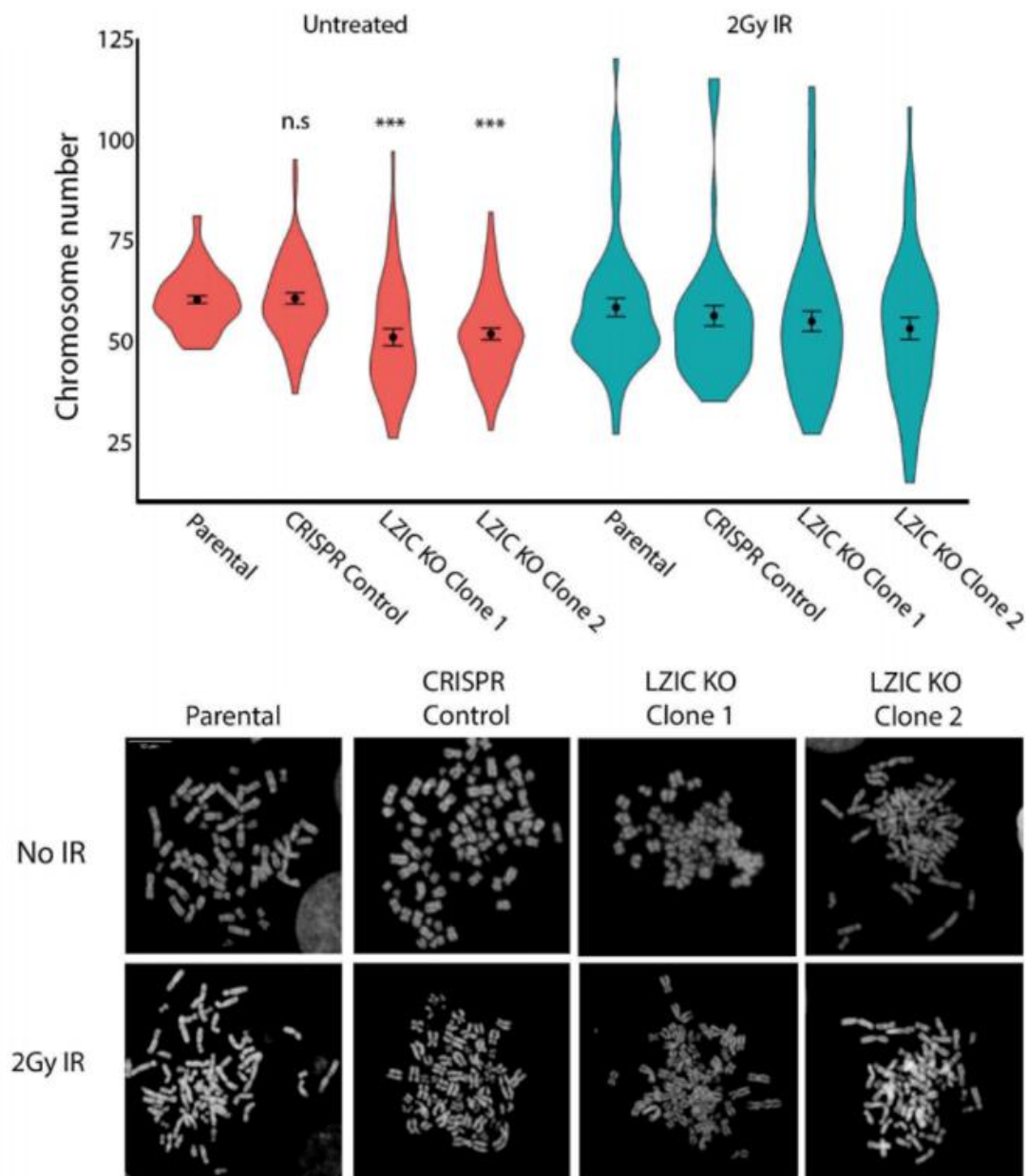


Figure 5.14 Karyotyping of CRISPR control and LZIC KO cell lines basally and following exposure to IR.

Cells were treated with 2Gy IR and following 48hrs harvested and processed for analysis of metaphase spreads. The spreads were quantified by manual counting following staining with DAPI. Total number of chromosomes for each spread was quantified. Overall, the quantification of chromosome numbers in basal conditions indicated a significant decrease of chromosome numbers in LZIC KO cell lines. Following exposure to IR the range of chromosome numbers increased dramatically for all lines, however, instability is similar between all lines. Error bars represent standard error mean (SEM)(N=3) and significance was determined by Student T-test *** = pvalue < 0.01. Adapted from Skalka *et al.* 2019.

5.2.15 Viability impact of exposure to IR following loss of LZIC protein expression

The chromosomal instability of LZIC does not increase following treatment with IR. Therefore, it is possible to hypothesise that cells which overcome this threshold undergo apoptosis. The WST-1 assay was used to measure cell viability at 24hrs post exposure to IR. The viability of all cells decreases across the doses of IR. The CRISPR control lines show the most resistance to IR induced cell death and have significantly increased viability when compared to the parental lines. In comparison, the LZIC KO lines are most sensitive to IR induced cell death with the highest sensitivity observed at 40Gy (Figure 5.15).

Overall, this experiment demonstrates that LZIC KO cells are more sensitive to IR exposure but this only a minor sensitivity.

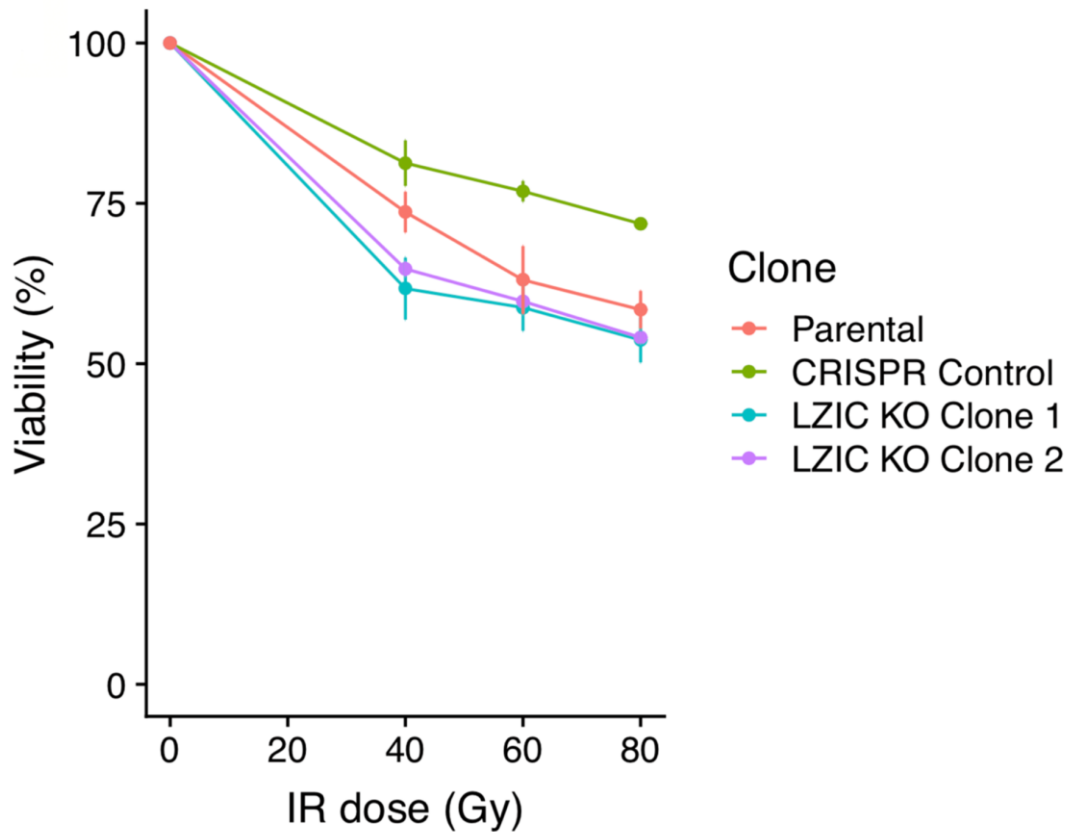


Figure 5.15 Viability impact of exposure to IR following loss of LZIC protein expression.

Cell lines were treated with IR and incubated for 24hrs before addition of the WST-1 reagent. Cells were incubated with WST-1 for 2hrs and then absorbance readings taken to show overall cell viability. Overall, the viability of LZIC KO cell lines is reduced compared to parental following IR exposure, in comparison, the CRISPR control line is more resistant to reduced cell viability. Error bars represent standard error mean (SEM). Three independent biological repeats were conducted. Adapted from Skalka *et al.* 2019.

5.2.16 Impact of LZIC expression on prognosis of renal clear cell carcinoma and neuroendocrine tumours.

The significant impact of LZIC loss upon genome stability has strong implications for diseases, such as, cancer. Interestingly, the altered expression of LZIC protein has been associated with the poor prognosis of multiple cancers. Neuroblastoma develops from nervous tissue, predominantly the adrenal glands (Shohet and Foster, 2017). Chromosome section 1p is frequently deleted in many aggressive neuroblastomas (Attiyeh *et al.*, 2005). The overlap between all section 1p deletions yields a region of 55 genes which contains LZIC, with the overall expression of LZIC found to be significantly reduced in tumours with a poor prognosis (Fransson, Martinsson and Ejeskär, 2007).

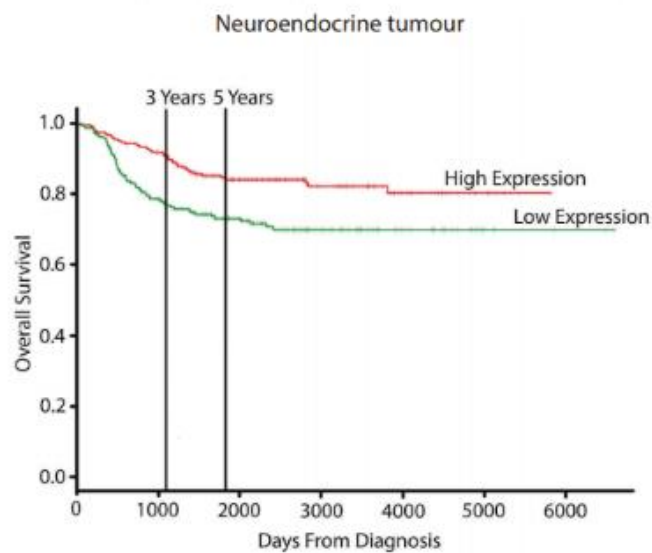
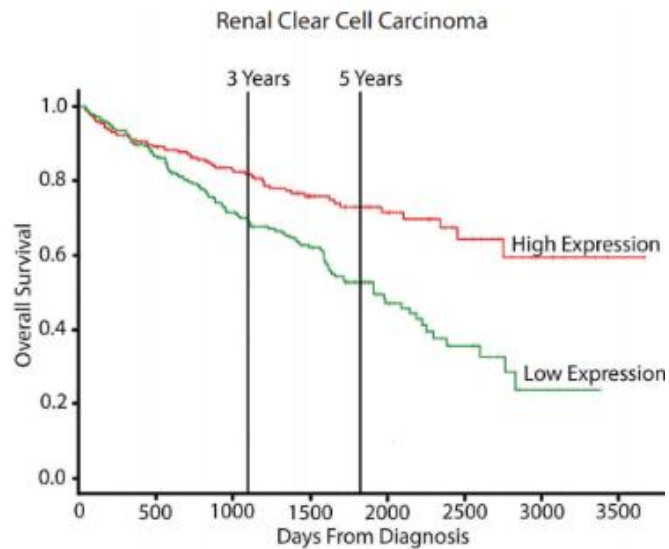
The loss of heterozygosity (LOH) occurs by a crossover of chromosomes and the duplication of one allele. This phenomenon was originally demonstrated by analysis of the RB locus, which indicated that a mitotic recombination event was the driver for establishing homozygosity at the locus (Cavenee *et al.*, 1983). The LOH can be highly pathogenic in situations whereby a non-mutated allele sequence is replaced by the mutated sequence by homologous recombination. Wilms tumour is a member of the nephroblastoma family and exhibits the common presence of LOH regions. LZIC was identified as a gene present in a common LOH region in Wilms tumour which is considered a driver region of tumour progression (Tamimi *et al.*, 2007).

Osteosarcoma is a cancer of the bone and originates from osteoblasts (Gianferante, Mirabello and Savage, 2017). The use of a rat model to investigate gene expression profiles of alpha radiation-induced osteosarcoma identified LZIC expression as significantly reduced in developed tumours (Daino *et al.*, 2009).

The most recent study to analyse LZIC expression in cancers specifically investigated pancreatic cancer. This cancer is highly aggressive and has a poor 5-year survival prognosis upon diagnosis (Ilic and Ilic, 2016). In contrast to the previous studies which identified reduced LZIC expression, LZIC expression is increased in pancreatic cancer with lymph node involvement (Yang *et al.*, 2013). To assess whether any additional tissue-specific cancer show LZIC expression linked to survival, the Progene V2 database was mined. In addition to those cancers previously identified LZIC expression was found

to strongly correlate with prognosis of renal clear cell carcinoma and neuroendocrine tumours.

This analysis demonstrates that LZIC expression can serve as a surrogate marker for prognosis of multiple cancers.



Cancer	Hazard Ratio	P-Value
Renal Clear cell carcinoma	0.42	6.8×10^{-9}
Neuroendocrine Tumour	0.36	7×10^{-7}

Figure 5.16 Overall survival correlates with LZIC expression for renal clear cell carcinoma and neuroendocrine tumours.

The PROGgene V2 database was used to investigate prognosis of cancer types that correlate with LZIC expression levels. Adapted from Skalka *et al.* 2019.

5.3 Discussion

Previous interactome and transcriptome analysis highlighted an enrichment for proteins and pathways which regulate the G2/M checkpoint. Therefore, this chapter aimed to assess the cell cycle phenotypes of LZIC KO cells following exposure to IR. The findings in this chapter indicate that loss of LZIC expression alters the maintenance of the G2/M checkpoint cascade, predominantly with altered phosphorylation of CHK1 and MPF factors observed, which culminates in the generation of genome instability. In addition, preliminary data suggest inhibition of PP2A can restore checkpoint functionality.

LZIC expression loss leads to a loss of IR induced late G2/M checkpoint arrest at the 24hr timepoint with no alteration to induction of early G2/M checkpoint or initial late G2/M checkpoint arrest at 8hrs (Figure 5.1 & 5.5). The reduced phosphorylation of CHK1 and components of the MPF in response to IR exposure show a characteristic of checkpoint adaptation (Syljuåsen *et al.*, 2006). The process of checkpoint adaptation increases genome instability and promotes the formation of micronuclei (Kalsbeek and Golsteyn, 2017). LZIC KO cells have high levels of genome instability, however, micronuclei formation was not specifically investigated and so this would be an interesting avenue for further work. In addition, the specific molecular impact of LZIC upon the known components of checkpoint adaptation should be investigated. Candidates for further explorations would be, predominantly, WEE1 and claspin. The activity of both CHK1 and PLK1 are required for checkpoint adaptation and in further work, the activation and localisation of PLK1 would also need to be addressed. In addition, the process of checkpoint adaptation facilitates the progression of cells through cell cycle checkpoints with active damage sites and so staining for DNA damage markers, such as γ H2AX, should be conducted.

CHK1 serine 317 is the primary site phosphorylated in response to DNA damage and is required to fully activate kinase function (Wilsker and Bunz, 2009). This site is associated with loss of the G2/M checkpoint. Reduced phosphorylation at S317 typically reduces phosphorylation of CHK1 S345, which in this study is not observed. However, the experiments conducted by Wilsker *et al.* had two fundamental differences, the first of these is the treatment of cells with hydroxyurea to assess the linkage between S317 and S345 instead of IR. Therefore, this lack of linkage may be an IR specific response as the

loss of G2/M checkpoint is corroborated. The cell types utilised to conduct experiments were also different between the two studies, therefore the background mutation profile could have a significant impact upon pathway activation.

The inhibition or knockdown of protein phosphatase function stalls proteins at the G2/M checkpoint by constitutive phosphorylation of the inhibitory Y15 site on CDC2 (Wei *et al.*, 2013). The phosphorylation of this site is determined by an intricate balance between WEE1 and CDC25C. The reduced phosphorylation of this site at the 24hr timepoint suggests altered regulation of these proteins. The loss of WEE1 activity is associated with early release from the G2/M checkpoint, which LZIC KO cells exhibiting the same phenotype (Fernet *et al.*, 2010). Further work will aim to characterise the activation state of these two proteins and assess whether differential regulation of these targets is involved in the phenotype observed in LZIC KO cells.

The MASTL kinase regulates the G2/M checkpoint transition in a protein level dependent manner. Reduced activity of MASTL signalling can permit progress through the G2/M and SAC checkpoint with structural defects that would typically arrest progression (Diril *et al.*, 2016). This phenotype is similar to that observed in LZIC KO cells, in addition, the MASTL phenotype can be reversed by treatment with PP2A inhibitor Okadaic acid, as observed in LZIC KO cells. Therefore, it is possible to suggest that LZIC plays a role in the MASTL pathway and leads to altered PP2A activity. For further work the activity of MASTL in LZIC KO cells should be analysed, both the total protein level and the phosphorylation status.

The activity of the 14-3-3 family members shuttles G2/M regulating proteins between the cytoplasm and nucleus as a way of modulating activity (Gardino and Yaffe, 2011). Interestingly, previous investigations into LZIC localisation have shown that the protein can be both cytoplasmic and nuclear, which seems to be dependent upon the cell type investigated (Uhlen *et al.*, 2015). Cell type-specific localisation may provide insight into the wide variety of interactors detected for LZIC. Whether specific shuttling of LZIC is observed in response to IR may also provide insight into the molecular mechanism.

The cyclin proteins have multiple isoforms with subtly differing functions. One example of this is the difference between cyclin B1 and cyclin B2 (Gong and Ferrell, 2010). The

loss of both these proteins leads to break down of the G2/M checkpoint transition and causes cell death. However, the loss of either allows progression through the G2/M checkpoint. The presence of only cyclin B2, however, increases the level of genome instability. In LZIC KO cells the previous microarray study shows that there is a reduced expression level of cyclin B1, this could impact upon the balance between cyclin B1 and B2 increasing the genome instability (Gong and Ferrell, 2010). For further work, qPCR should be conducted to analyse the expression levels of both of these proteins both under endogenous conditions and following treatment with IR. In addition, phenotype reversion studies could be conducted by overexpression of cyclin B1.

Cell viability analysis was conducted utilising WST-1 and found that the CRISPR control is more resistant to IR induced cell death than parental and LZIC KO lines. The phenotypes observed were corroborated by the parental line and CRISPR control, this may indicate that this phenomenon is due to off-target effects of the CRISPR. The quantification of cell death following treatment with toxin was conducted by using the WST-1 assay. This readout will give an indication when the mitochondrial functionality of a cell drops, which is concurrent with reduced cellular viability. However, it is not a direct read-out of apoptosis. To address this future work should utilise either caspase cleavage or Annexin/Propidium iodide staining to specifically determine the levels of apoptosis induction.

The analysis of various cancers indicates that linkage between LZIC expression and prognosis is highest for clear cell renal carcinoma. Large scale expression analysis of LZIC across multiple tissues has indicated that expression of the protein is highest in cells of the kidney (Canela-Xandri, Rawlik and Tenesa, 2018). Interestingly the cell model utilised within this study was HEK293 cells which originate from the kidney and are phenotypically related to renal cell lines. It would be of interest to separate clear cell renal carcinoma prognosis by a combination of treatment modality and LZIC status to assess whether any prognostic use for LZIC can be established. Typical cancer therapy will utilise drugs such as bleomycin which are classed as radiomimetic and induce similar DNA damage profiles to IR. To follow up from the finding that LZIC is involved with cell cycle checkpoint response to IR, the correlation between LZIC expression level and genome instability should be analysed in primary tumour lines, with the addition of segregation by treatment type. Finally, the development of either a total body KO or

tissue-specific KO mouse models of LZIC expression could yield vital data about the impact of LZIC expression on tumorigenesis initiation and response to IR exposure on an organismal level.

The treatment of LZIC KO cells with multiple stress-inducing compounds confirmed that the G2/M arrest and subsequent loss following IR treatment are specific. However, a previous investigation has shown that treatment of cells with camptothecin should also induce a G2/M arrest, however, the study used different cells (Jayasooriya *et al.*, 2018). To further address this phenomenon the cell cycle profiles would be conducted at earlier time points to determine if a G2/M arrest can be observed.

Overall this chapter has confirmed the involvement of LZIC in the regulation cascade of the G2/M checkpoint cascade and identified a range of further experimental avenues which to address in further work.

Chapter 6

Implications and Future Work

6.1 Implications and future work

The activation of cell cycle checkpoints is paramount to avoid damage to the genome following genotoxic stress. While the regulators of the G1 checkpoint have been intensely characterised, less work has been conducted to establish the full breadth of G2/M checkpoint factors. Of particular interest for this project was the putative WNT signalling factor, LZIC. While the role of WNT signalling pathway as a G1/S-phase activator is well-established, the functions of the various WNT signalling cascades during G2/M progression is less understood. This study has utilised proteomics, transcriptome analysis, and cell biology techniques to assess the impact of LZIC expression loss both basally and in response to IR. The combination of these techniques and the development of LZIC KO HEK293 enabled this study to successfully address the initial aims.

The identification of LZIC protein interacting partners showed an enrichment for components of the spliceosome. Previous links have been found between the spliceosome and WNT signalling pathway, with over expression of β -catenin in colorectal cells altering global gene splicing (Gonçalves, Matos and Jordan, 2008). Further work utilising RNA-sequencing will be required to assess the impact on alternative splicing following the loss of LZIC, as microarray analysis utilised in this study did not generate data to address this question. Interestingly, during the optimisation of the LZIC-c-flag reintroduction HEK293 cells, for mass spectrometry, the instability of exogenously over expressed LZIC was identified. This instability of over-expressed protein is characteristic of the requirement for specific co-factor binding or post-translational modification mediated stability. The interactome of LZIC included kinases which could post-translationally modify LZIC with phosphorylation moieties increasing stability of the protein.

The analysis of transcriptomic data by GSEA identified MYC targets as the most altered in response to IR. While not investigated within this study, the altered activity of MYC has been linked to aberrant G2/M checkpoint (Sheen, Woo and Dickson, 2003; Yang *et al.*, 2018). Interestingly, altered WNT signalling has been directly linked to changes of MYC transcription (Zhang *et al.*, 2012). While LZIC does not interact with β -catenin, the enrichment for armadillo-like repeats could link LZIC to the WNT signalling pathway. In addition, the established MYC interactor, NEMO, was also identified as an

LZIC interacting partner providing a second link between the two pathways (Kim *et al.*, 2011)(Figure 3.4). A combination of these results suggests that further investigation of the impact of LZIC loss upon WNT signalling is required and how this specifically impacts upon the MYC signalling cascade.

Previous work investigating LZIC function has identified a neuronal development defect when LZIC expression is reduced in zebrafish (Clements and Kimelman, 2005). Analysis of both LZIC interactome and basal transcriptional changes dependent on LZIC expression identified factors required for anatomical development and neuronal development (Figure 3.14 & Figure 4.4). The factors identified within these groups such as, SOX11, should be investigated further to identify the mechanism of gene expression regulation by LZIC. In addition, to assess the neuronal development phenotype of LZIC KO in a mammalian system, a mouse model of LZIC KO would be beneficial. This model would also facilitate the identification of LZIC expression related changes across different tissues, such as the kidney.

The final chapter utilised molecular and cellular biology techniques to assess the cell cycle checkpoint activation of LZIC KO cell lines following IR exposure. As predicted by the transcriptome analysis and the proteomics the response of the LZIC KO cell lines was altered, with early release from G2/M checkpoint arrest being observed. The early release from the G2/M checkpoint is referred to as “checkpoint adaptation” and while a vital phenotype of cancer cells the molecular mechanism by which cells can undergo this process is debated. A prime candidate for further analysis in LZIC KO cell lines is gwl. The gwl cascade regulates the activity of PP2A and is required for restoration of the cell cycle following damage (Peng *et al.*, 2010; Wong *et al.*, 2016). To address this question, mRNA and protein expression abundance of gwl will be assayed and the phosphorylation of known gwl dependent sites would be investigated.

Tumour biomarkers for chemotherapeutic agents and small molecular inhibitors is yielding increased treatment efficacy and improved patient morbidity (Ong *et al.*, 2012). However, despite IR being a widely utilised treatment modality, the identification of sensitivity and response biomarkers are lacking. This study has shown that LZIC both interacts with proteins that regulate the response to IR and that loss of LZIC protein expression alters G2/M checkpoint response, specifically to IR exposure. Therefore,

these results suggest that stratification of clinical data sets by LZIC expression and treatment modality could determine whether LZIC is a clinical biomarker of IR responsiveness.

Overall, LZIC interactome analysis suggests an interaction with the spliceosome. While, transcriptomic analysis demonstrates that basally LZIC loss particular causes alterations to MYC signalling and following IR the MYC signalling pathway deviation G2/M cell cycle checkpoint and WNT signalling are also identified. Finally, the study of cell cycle checkpoint activation following IR treatment identified a loss of G2/M checkpoint maintenance which was supported by loss of canonical inhibitory signals in complexes, such as, the MPF.

.

Bibliography

Adams, P. D. *et al.* (1999) 'Retinoblastoma Protein Contains a C-terminal Motif That Targets It for Phosphorylation by Cyclin-cdk Complexes', *Molecular and Cellular Biology*. American Society for Microbiology, 19(2), pp. 1068–1080. doi: 10.1128/mcb.19.2.1068.

Adamson, A. W. *et al.* (2005) 'Methylator-induced, mismatch repair-dependent G2 arrest is activated through Chk1 and Chk2', *Molecular Biology of the Cell*, 16(3), pp. 1513–1526. doi: 10.1091/mbc.E04-02-0089.

Ahn, J. Y. *et al.* (2000) 'Threonine 68 phosphorylation by ataxia telangiectasia mutated is required for efficient activation of Chk2 in response to ionizing radiation.', *Cancer research*, 60(21), pp. 5934–6. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11085506> (Accessed: 13 September 2019).

Ahn, J. Y. *et al.* (2002) 'Phosphorylation of threonine 68 promotes oligomerization and autophosphorylation of the Chk2 protein kinase via the forkhead-associated domain', *Journal of Biological Chemistry*, 277(22), pp. 19389–19395. doi: 10.1074/jbc.M200822200.

Allshire, R. C. and Madhani, H. D. (2018) 'Ten principles of heterochromatin formation and function.', *Nature reviews. Molecular cell biology*, 19(4), pp. 229–244. doi: 10.1038/nrm.2017.119.

An, W., Kim, J. and Roeder, R. G. (2004) 'Ordered cooperative functions of PRMT1, p300, and CARM1 in transcriptional activation by p53.', *Cell*, 117(6), pp. 735–48. doi: 10.1016/j.cell.2004.05.009.

Andres, S. N. *et al.* (2007) 'Crystal structure of human XLF: a twist in nonhomologous DNA end-joining.', *Molecular cell*, 28(6), pp. 1093–101. doi: 10.1016/j.molcel.2007.10.024.

Andzelm, E. R., Lew, J. and Taylor, S. (1995) 'Bound to activate: conformational consequences of cyclin binding to CDK2', *Structure*, 3(11), pp. 1135–1141. doi: 10.1016/S0969-2126(01)00249-0.

Aressy, B. and Ducommun, B. (2012) 'Cell Cycle Control by the CDC25 Phosphatases', *Anti-Cancer Agents in Medicinal Chemistry*. Bentham Science Publishers Ltd., 8(8), pp. 818–824. doi: 10.2174/187152008786847756.

Arumughan, A. *et al.* (2016) 'Quantitative interaction mapping reveals an extended UBX domain in ASPL that disrupts functional p97 hexamers', *Nature Communications*. Nature Publishing Group, 7. doi: 10.1038/ncomms13047.

Attiyeh, E. F. *et al.* (2005) 'Chromosome 1p and 11q deletions and outcome in neuroblastoma', *New England Journal of Medicine*, 353(21), pp. 2243–2253. doi: 10.1056/NEJMoa052399.

Attwooll, C., Denchi, E. L. and Helin, K. (2004) 'The E2F family: Specific functions and overlapping interests', *EMBO Journal*, pp. 4709–4716. doi: 10.1038/sj.emboj.7600481.

Bancroft, J. *et al.* (2015) 'Chromosome congression is promoted by CENP-Q- and CENP-E-dependent pathways', *Journal of Cell Science*. Company of Biologists Ltd, 128(1), pp. 171–184. doi: 10.1242/jcs.163659.

Banin, S. *et al.* (1998) 'Enhanced phosphorylation of p53 by ATM in response to DNA damage.', *Science (New York, N.Y.)*, 281(5383), pp. 1674–7. doi: 10.1126/science.281.5383.1674.

Bansal, P. and Lazo, J. S. (2007) 'Induction of Cdc25B regulates cell cycle resumption after genotoxic stress', *Cancer Research*, 67(7), pp. 3356–3363. doi: 10.1158/0008-5472.CAN-06-3685.

Barisic, M. *et al.* (2014) 'Kinetochore motors drive congression of peripheral polar chromosomes by overcoming random arm-ejection forces', *Nature Cell Biology*. Nature Publishing Group, 16(12), pp. 1249–1256. doi: 10.1038/ncb3060.

Bellanger, S., de Gramont, A. and Sobczak-Thépot, J. (2007) 'Cyclin B2 suppresses mitotic failure and DNA re-replication in human somatic cells knocked down for both cyclins B1 and B2.', *Oncogene*, 26(51), pp. 7175–84. doi: 10.1038/sj.onc.1210539.

Belotserkovskaya, R. *et al.* (2003) 'FACT facilitates transcription-dependent nucleosome alteration.', *Science (New York, N.Y.)*, 301(5636), pp. 1090–3. doi: 10.1126/science.1085703.

Benchabane, H. *et al.* (2011) 'Jerky/Earthbound facilitates cell-specific Wnt/Wingless signalling by modulating β -catenin-TCF activity.', *The EMBO journal*, 30(8), pp. 1444–58. doi: 10.1038/emboj.2011.67.

Bermudez, V. P. *et al.* (2003) 'Loading of the human 9-1-1 checkpoint complex onto DNA by the checkpoint clamp loader hRad17-replication factor C complex in vitro', *Proceedings of the National Academy of Sciences of the United States of America*, 100(4), pp. 1633–1638. doi: 10.1073/pnas.0437927100.

Bizard, A. H. and Hickson, I. D. (2014) 'The dissolution of double Holliday junctions', *Cold Spring Harbor Perspectives in Biology*. Cold Spring Harbor Laboratory Press, 6(7). doi: 10.1101/cshperspect.a016477.

Blasius, M. *et al.* (2011) 'A phospho-proteomic screen identifies substrates of the checkpoint kinase Chk1', *Genome Biology*, 12(8). doi: 10.1186/gb-2011-12-8-r78.

Blattner, F. R. *et al.* (1997) 'The complete genome sequence of Escherichia coli K-12', *Science*, pp. 1453–1462. doi: 10.1126/science.277.5331.1453.

Blier, P. R. *et al.* (1993) 'Binding of Ku Protein to DNA', *Journal of Biological Chemistry*, 268(10), pp. 7594–7601.

Blomberg, I. and Hoffmann, I. (1999) 'Ectopic Expression of Cdc25A Accelerates the G 1 /S Transition and Leads to Premature Activation of Cyclin E- and Cyclin A-Dependent Kinases ', *Molecular and Cellular Biology*. American Society for Microbiology, 19(9), pp. 6183–6194. doi: 10.1128/mcb.19.9.6183.

Bolognese, F. *et al.* (1999) 'The cyclin B2 promoter depends on NF-Y, a trimer whose CCAAT-binding activity is cell-cycle regulated', *Oncogene*. Nature Publishing Group, 18(10), pp. 1845–1853. doi: 10.1038/sj.onc.1202494.

Bolstad, B. M. *et al.* (2003) 'A comparison of normalization methods for high density oligonucleotide array data based on variance and bias', *Bioinformatics*, 19(2), pp. 185–193. doi: 10.1093/bioinformatics/19.2.185.

De Bondt, H. L. *et al.* (1993) 'Crystal structure of cyclin-dependent kinase 2', *Nature*, 363(6430), pp. 595–602. doi: 10.1038/363595a0.

Boutros, R., Dozier, C. and Ducommun, B. (2006) 'The when and wheres of CDC25 phosphatases', *Current Opinion in Cell Biology*. Elsevier Ltd, 18(2), pp. 185–191. doi: 10.1016/j.ceb.2006.02.003.

Brehm, A. *et al.* (1998) 'Retinoblastoma protein recruits histone deacetylase to repress transcription.', *Nature*, 391(6667), pp. 597–601. doi: 10.1038/35404.

Brooks, C. L. and Gu, W. (2010) 'New insights into p53 activation', *Nature Publishing Group*, 20(6), pp. 614–621. doi: 10.1038/cr.2010.53.

Brosh, R. M. (2013) 'DNA helicases involved in DNA repair and their roles in cancer', *Nature Reviews Cancer*, pp. 542–558. doi: 10.1038/nrc3560.

Brown, V. D., Phillips, R. A. and Gallie, B. L. (1999) 'Cumulative effect of phosphorylation of pRB on regulation of E2F activity.', *Molecular and cellular biology*, 19(5), pp. 3246–56. doi: 10.1128/mcb.19.5.3246.

Brückner, A. *et al.* (2009) 'Yeast two-hybrid, a powerful tool for systems biology', *International Journal of Molecular Sciences*, pp. 2763–2788. doi: 10.3390/ijms10062763.

Bryans, M., Valenzano, M. C. and Stamato, T. D. (1999) 'Absence of DNA ligase IV protein in XR-1 cells: Evidence for stabilization by XRCC4', *Mutation Research - DNA Repair*, 433(1), pp. 53–58. doi: 10.1016/S0921-8777(98)00063-9.

Burgess, A. *et al.* (2010) 'Loss of human Greatwall results in G2 arrest and multiple mitotic defects due to deregulation of the cyclin B-Cdc2/PP2A balance', *Proceedings of the National Academy of Sciences of the United States of America*, 107(28), pp. 12564–

12569. doi: 10.1073/pnas.0914191107.

Cadet, J. and Richard Wagner, J. (2013) ‘DNA base damage by reactive oxygen species, oxidizing agents, and UV radiation’, *Cold Spring Harbor Perspectives in Biology*, 5(2). doi: 10.1101/cshperspect.a012559.

Cadigan, K. M. and Nusse, R. (1997) ‘Wnt signaling: A common theme in animal development’, *Genes and Development*. Cold Spring Harbor Laboratory Press, pp. 3286–3305. doi: 10.1101/gad.11.24.3286.

Cadigan, K. M. and Waterman, M. L. (2012) ‘TCF/LEFs and Wnt signaling in the nucleus.’, *Cold Spring Harbor perspectives in biology*, 4(11). doi: 10.1101/cshperspect.a007906.

Canela-Xandri, O., Rawlik, K. and Tenesa, A. (2018) ‘An atlas of genetic associations in UK Biobank’, *Nature Genetics*. Nature Publishing Group, 50(11), pp. 1593–1599. doi: 10.1038/s41588-018-0248-z.

Cánepa, E. T. *et al.* (2007) ‘INK4 proteins, a family of mammalian CDK inhibitors with novel biological functions’, *IUBMB Life*, pp. 419–426. doi: 10.1080/15216540701488358.

Cannan, W. J. and Pederson, D. S. (2016) ‘Mechanisms and Consequences of Double-Strand DNA Break Formation in Chromatin’, *Journal of Cellular Physiology*. Wiley-Liss Inc., pp. 3–14. doi: 10.1002/jcp.25048.

Cannavo, E. and Cejka, P. (2014) ‘Sae2 promotes dsDNA endonuclease activity within Mre11-Rad50-Xrs2 to resect DNA breaks.’, *Nature*, 514(7520), pp. 122–5. doi: 10.1038/nature13771.

Cans, C., Ducommun, B. and Baldin, V. (1999) ‘Proteasome-dependent degradation of human CDC25B phosphatase.’, *Molecular biology reports*, 26(1–2), pp. 53–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10363647> (Accessed: 14 September 2019).

Carpenter, B. S. *et al.* (2015) ‘The heterotrimeric kinesin-2 complex interacts with and

regulates GLI protein function', *Journal of Cell Science*. Company of Biologists Ltd, 128(5), pp. 1034–1050. doi: 10.1242/jcs.162552.

Cavenee, W. K. *et al.* (1983) 'Expression of recessive alleles by chromosomal mechanisms in retinoblastoma', *Nature*, 305(5937), pp. 779–784. doi: 10.1038/305779a0.

Cerqueira, A. *et al.* (2014) 'Genetic Characterization of the Role of the Cip/Kip Family of Proteins as Cyclin-Dependent Kinase Inhibitors and Assembly Factors', *Molecular and Cellular Biology*. American Society for Microbiology, 34(8), pp. 1452–1459. doi: 10.1128/mcb.01163-13.

Chang, D. C., Xu, N. and Luo, K. Q. (2003) 'Degradation of cyclin B is required for the onset of anaphase in mammalian cells', *Journal of Biological Chemistry*, 278(39), pp. 37865–37873. doi: 10.1074/jbc.M306376200.

Chang, H. H. Y. *et al.* (2016) 'Different DNA end configurations dictate which NHEJ components are most important for joining efficiency', *Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology Inc., 291(47), pp. 24377–24389. doi: 10.1074/jbc.M116.752329.

Chang, H. H. Y., Watanabe, G. and Lieber, M. R. (2015) 'Unifying the DNA end-processing roles of the artemis nuclease: Ku-dependent artemis resection at blunt DNA ends', *Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology Inc., 290(40), pp. 24036–24050. doi: 10.1074/jbc.M115.680900.

Chang, Y. W. *et al.* (2016) 'Deacetylation of HSPA5 by HDAC6 leads to GP78-mediated HSPA5 ubiquitination at K447 and suppresses metastasis of breast cancer', *Oncogene*. Nature Publishing Group, 35(12), pp. 1517–1528. doi: 10.1038/onc.2015.214.

Chapman, J. R. *et al.* (2013) 'RIF1 Is Essential for 53BP1-Dependent Nonhomologous End Joining and Suppression of DNA Double-Strand Break Resection', *Molecular Cell*, 49(5), pp. 858–871. doi: 10.1016/j.molcel.2013.01.002.

Chaves-Pérez, A. *et al.* (2013) 'EpCAM regulates cell cycle progression via control of

cyclin D1 expression', *Oncogene*, 32(5), pp. 641–650. doi: 10.1038/onc.2012.75.

Cheadle, C. *et al.* (2003) 'Analysis of microarray data using Z score transformation.', *The Journal of molecular diagnostics: JMD*, 5(2), pp. 73–81. doi: 10.1016/S1525-1578(10)60455-2.

Chen, M.-S., Ryan, C. E. and Piwnica-Worms, H. (2003) 'Chk1 Kinase Negatively Regulates Mitotic Function of Cdc25A Phosphatase through 14-3-3 Binding', *Molecular and Cellular Biology*. American Society for Microbiology, 23(21), pp. 7488–7497. doi: 10.1128/mcb.23.21.7488-7497.2003.

Cheng, M. (1999) 'The p21Cip1 and p27Kip1 CDK `inhibitors' are essential activators of cyclin D-dependent kinases in murine fibroblasts', *The EMBO Journal*. Wiley, 18(6), pp. 1571–1583. doi: 10.1093/emboj/18.6.1571.

Chibazakura, T. *et al.* (2011) 'Cyclin A Promotes S-Phase Entry via Interaction with the Replication Licensing Factor Mcm7', *Molecular and Cellular Biology*. American Society for Microbiology, 31(2), pp. 248–255. doi: 10.1128/mcb.00630-10.

Churchwell, M. I. *et al.* (2005) 'Improving LC-MS sensitivity through increases in chromatographic performance: Comparisons of UPLC-ES/MS/MS to HPLC-ES/MS/MS', *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*, 825(2), pp. 134–143. doi: 10.1016/j.jchromb.2005.05.037.

Clements, W. K. and Kimelman, D. (2005) 'LZIC regulates neuronal survival during zebrafish development.', *Developmental biology*, 283(2), pp. 322–34. doi: 10.1016/j.ydbio.2005.04.026.

Clevers, H. (2006) 'Wnt/beta-catenin signaling in development and disease.', *Cell*, 127(3), pp. 469–80. doi: 10.1016/j.cell.2006.10.018.

Cliffe, A., Hamada, F. and Bienz, M. (2003) 'A role of dishevelled in relocating axin to the plasma membrane during wingless signaling', *Current Biology*. Cell Press, 13(11), pp. 960–966. doi: 10.1016/S0960-9822(03)00370-1.

Coates, J. *et al.* (2016) ‘Coordinated nuclease activities counteract Ku at’. doi: 10.1038/ncomms12889.

Coleman, K. A. and Greenberg, R. A. (2011) ‘The BRCA1-RAP80 complex regulates DNA repair mechanism utilization by restricting end resection Running Title: BRCA1-RAP80 regulates DNA repair pathway choice Correspondence to’. doi: 10.1074/jbc.M110.213728.

Cox, J. and Mann, M. (2008) *MaxQuant summer school 2010 must read*, *Nature biotechnology*. doi: 10.1038/nbt.1511.

Craxton, A. *et al.* (2015) ‘XLS (c9orf142) is a new component of mammalian DNA double-stranded break repair.’, *Cell death and differentiation*, 22(6), pp. 890–7. doi: 10.1038/cdd.2015.22.

Craxton, A. *et al.* (2018) ‘PAXX and its paralogs synergistically direct DNA polymerase λ activity in DNA repair’, *Nature Communications*. Nature Publishing Group, 9(1). doi: 10.1038/s41467-018-06127-y.

Cucinotta, F. A. *et al.* (2008) ‘Biochemical Kinetics Model of DSB Repair and Induction of γ -H2AX Foci by Non-homologous End Joining’, *Radiation Research*. Radiation Research Society, 169(2), pp. 214–222. doi: 10.1667/rr1035.1.

Daino, K. *et al.* (2009) ‘Gene expression profiling of alpha-radiation-induced rat osteosarcomas: identification of dysregulated genes involved in radiation-induced tumorigenesis of bone.’, *International journal of cancer*, 125(3), pp. 612–20. doi: 10.1002/ijc.24392.

Daley, J. M. *et al.* (2017) ‘Enhancement of BLM-DNA2-Mediated Long-Range DNA End Resection by CtIP’, *CellReports*, 21, pp. 324–332. doi: 10.1016/j.celrep.2017.09.048.

Dang, C. V (2012) ‘MYC on the path to cancer.’, *Cell*, 149(1), pp. 22–35. doi: 10.1016/j.cell.2012.03.003.

Daniels, D. L. and Weis, W. I. (2002) 'ICAT inhibits beta-catenin binding to Tcf/Lef-family transcription factors and the general coactivator p300 using independent structural modules.', *Molecular cell*, 10(3), pp. 573–84. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12408825> (Accessed: 13 September 2019).

Deckbar, D. *et al.* (2007) 'Chromosome breakage after G2 checkpoint release', *Journal of Cell Biology*, 176(6), pp. 749–755. doi: 10.1083/jcb.200612047.

Deisenroth, C. *et al.* (2010) 'Mitochondrial HEP27 Is a c-Myb Target Gene That Inhibits Mdm2 and Stabilizes p53', *Molecular and Cellular Biology*. American Society for Microbiology, 30(16), pp. 3981–3993. doi: 10.1128/mcb.01284-09.

Derheimer, F. A. *et al.* (2007) 'RPA and ATR link transcriptional stress to p53', *Proceedings of the National Academy of Sciences of the United States of America*, 104(31), pp. 12778–12783. doi: 10.1073/pnas.0705317104.

Descombes, P. and Nigg, E. A. (1998) 'The polo-like kinase Plx1 is required for M phase exit and destruction of mitotic regulators in *Xenopus* egg extracts', *EMBO Journal*, 17(5), pp. 1328–1335. doi: 10.1093/emboj/17.5.1328.

Devault, A. *et al.* (1995) 'MAT1 ('menage à trois') a new RING finger protein subunit stabilizing cyclin H-cdk7 complexes in starfish and *Xenopus* CAK.', *The EMBO journal*, 14(20), pp. 5027–36. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/7588631> (Accessed: 12 September 2019).

Diehl, J. A. *et al.* (1998) 'Glycogen synthase kinase-3 β regulates cyclin D1 proteolysis and subcellular localization', *Genes and Development*. Cold Spring Harbor Laboratory Press, 12(22), pp. 3499–3511. doi: 10.1101/gad.12.22.3499.

Diehl, J. A., Zindy, F. and Sherr, C. J. (1997) 'Inhibition of cyclin D1 phosphorylation on threonine-286 prevents its rapid degradation via the ubiquitin-proteasome pathway.', *Genes & development*, 11(8), pp. 957–72. doi: 10.1101/gad.11.8.957.

Dignani, J. D., Lebovitz, R. M. and Roeder, R. G. (1983) 'Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei',

Nucleic Acids Research, 11(5), pp. 1475–1489. doi: 10.1093/nar/11.5.1475.

Ding, R. *et al.* (1997) ‘The spindle pole body of *Schizosaccharomyces pombe* enters and leaves the nuclear envelope as the cell cycle proceeds’, *Molecular Biology of the Cell*. American Society for Cell Biology, 8(8), pp. 1461–1479. doi: 10.1091/mbc.8.8.1461.

Diril, M. K. *et al.* (2016) ‘Loss of the Greatwall Kinase Weakens the Spindle Assembly Checkpoint’, *PLoS Genetics*. Public Library of Science, 12(9). doi: 10.1371/journal.pgen.1006310.

Dong, P. *et al.* (2018) ‘Cyclin D/CDK4/6 activity controls G1 length in mammalian cells’, *PLoS ONE*. Public Library of Science, 13(1). doi: 10.1371/journal.pone.0185637.

Dong, Z. *et al.* (2015) ‘Wnt/ β -catenin pathway involvement in ionizing radiation-induced invasion of U87 glioblastoma cells.’, *Strahlentherapie und Onkologie: Organ der Deutschen Rontgengesellschaft ... [et al]*, 191(8), pp. 672–80. doi: 10.1007/s00066-015-0858-7.

Earnshaw, W. C. and Rothfield, N. (1985) ‘Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma.’, *Chromosoma*, 91(3–4), pp. 313–21. doi: 10.1007/bf00328227.

Eliezer, Y. *et al.* (2009) ‘The direct interaction between 53BP1 and MDC1 Is required for the recruitment of 53BP1 to sites of damage’, *Journal of Biological Chemistry*, 284(1), pp. 426–435. doi: 10.1074/jbc.M807375200.

Ellenberg, J. *et al.* (1997) ‘Nuclear membrane dynamics and reassembly in living cells: Targeting of an inner nuclear membrane protein in interphase and mitosis’, *Journal of Cell Biology*, 138(6), pp. 1193–1206. doi: 10.1083/jcb.138.6.1193.

Esashi, F. *et al.* (2005) ‘CDK-dependent phosphorylation of BRCA2 as a regulatory mechanism for recombinational repair’, *Nature*, 434(7033), pp. 598–604. doi: 10.1038/nature03404.

Ewen, M. E. *et al.* (1991) ‘Molecular cloning, chromosomal mapping, and expression of

the cDNA for p107, a retinoblastoma gene product-related protein', *Cell*, 66(6), pp. 1155–1164. doi: 10.1016/0092-8674(91)90038-Z.

Falck, J. *et al.* (2001) 'The ATM-Chk2-Cdc25A checkpoint pathway guards against radioresistant DNA synthesis.', *Nature*, 410(6830), pp. 842–7. doi: 10.1038/35071124.

Fanning, E., Klimovich, V. and Nager, A. R. (2006) 'A dynamic model for replication protein A (RPA) function in DNA processing pathways', *Nucleic Acids Research*, pp. 4126–4137. doi: 10.1093/nar/gkl550.

Farina, F. *et al.* (2016) 'The centrosome is an actin-organizing centre', *Nature Cell Biology*. Nature Publishing Group, 18(1), pp. 65–75. doi: 10.1038/ncb3285.

Felsher, D. W. and Bishop, J. M. (1999) 'Transient excess of MYC activity can elicit genomic instability and tumorigenesis', *Proceedings of the National Academy of Sciences of the United States of America*, 96(7), pp. 3940–3944. doi: 10.1073/pnas.96.7.3940.

Fenner, B. J., Scannell, M. and Prehn, J. H. M. (2010) 'Expanding the substantial interactome of NEMO using protein microarrays', *PLoS ONE*, 5(1). doi: 10.1371/journal.pone.0008799.

Fernet, M. *et al.* (2010) 'Control of the G2/M checkpoints after exposure to low doses of ionising radiation: Implications for hyper-radiosensitivity', *DNA Repair*, 9(1), pp. 48–57. doi: 10.1016/j.dnarep.2009.10.006.

Fisher, R. P. and Morgan, D. O. (1994) 'A novel cyclin associates with MO15/CDK7 to form the CDK-activating kinase.', *Cell*, 78(4), pp. 713–24. doi: 10.1016/0092-8674(94)90535-5.

Floyd, S., Pines, J. and Lindon, C. (2008) 'APC/CCdh1 Targets Aurora Kinase to Control Reorganization of the Mitotic Spindle at Anaphase', *Current Biology*, 18(21), pp. 1649–1658. doi: 10.1016/j.cub.2008.09.058.

Fodde, R. (2002) 'The APC gene in colorectal cancer', *European Journal of Cancer*,

38(7), pp. 867–871. doi: 10.1016/S0959-8049(02)00040-0.

Forester, C. M. *et al.* (2007) ‘Control of mitotic exit by PP2A regulation of Cdc25C and Cdk1’, *Proceedings of the National Academy of Sciences of the United States of America*, 104(50), pp. 19867–19872. doi: 10.1073/pnas.0709879104.

Fradet-Turcotte, A. *et al.* (2013) ‘53BP1 is a reader of the DNA-damage-induced H2A Lys 15 ubiquitin mark.’, *Nature*, 499(7456), pp. 50–4. doi: 10.1038/nature12318.

Franco, A. *et al.* (2016) ‘Rays Sting: The Acute Cellular Effects of Ionizing Radiation Exposure.’, *Translational medicine @ UniSa*, 14, pp. 42–53. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/27326395> (Accessed: 13 September 2019).

Fransson, S., Martinsson, T. and Ejeskär, K. (2007) ‘Neuroblastoma tumors with favorable and unfavorable outcomes: Significant differences in mRNA expression of genes mapped at 1p36.2’, *Genes Chromosomes and Cancer*, 46(1), pp. 45–52. doi: 10.1002/gcc.20387.

Freeman, A. K. and Monteiro, A. N. A. (2010) ‘Phosphatases in the cellular response to DNA damage’, *Cell Communication and Signaling*. doi: 10.1186/1478-811X-8-27.

Freeman, A. K. and Morrison, D. K. (2011) ‘14-3-3 Proteins: Diverse functions in cell proliferation and cancer progression’, *Seminars in Cell and Developmental Biology*. Elsevier Ltd, pp. 681–687. doi: 10.1016/j.semcdb.2011.08.009.

Friend, S. H. *et al.* (1986) ‘A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma’, *Nature*, 323(6089), pp. 643–646. doi: 10.1038/323643a0.

Fuda, N. J., Ardehali, M. B. and Lis, J. T. (2009) ‘Defining mechanisms that regulate RNA polymerase II transcription in vivo’, *Nature*, pp. 186–192. doi: 10.1038/nature08449.

Gabrielli, B. G. *et al.* (1997) ‘Hyperphosphorylation of the N-terminal domain of cdc25 regulates activity toward cyclin B1/cdc2 but not cyclin A/cdk2’, *Journal of Biological*

Chemistry, 272(45), pp. 28607–28614. doi: 10.1074/jbc.272.45.28607.

Gao, Y. *et al.* (2019) ‘Structures and operating principles of the replisome’, *Science*. American Association for the Advancement of Science, 363(6429). doi: 10.1126/science.aav7003.

García-Álvarez, B. *et al.* (2007) ‘Molecular and structural basis of polo-like kinase 1 substrate recognition: Implications in centrosomal localization’, *Proceedings of the National Academy of Sciences of the United States of America*, 104(9), pp. 3107–3112. doi: 10.1073/pnas.0609131104.

Gardino, A. K. and Yaffe, M. B. (2011) ‘14-3-3 proteins as signaling integration points for cell cycle control and apoptosis’, *Seminars in Cell and Developmental Biology*. Elsevier Ltd, pp. 688–695. doi: 10.1016/j.semcdb.2011.09.008.

Gartel, A. L. and Tyner, A. L. (1999) ‘Transcriptional regulation of the p21((WAF1/CIP1)) gene’, *Experimental Cell Research*. Academic Press Inc., 246(2), pp. 280–289. doi: 10.1006/excr.1998.4319.

Gatei, M. *et al.* (2003) ‘Ataxia-telangiectasia-mutated (ATM) and NBS1-dependent phosphorylation of Chk1 on Ser-317 in response to ionizing radiation’, *Journal of Biological Chemistry*, 278(17), pp. 14806–14811. doi: 10.1074/jbc.M210862200.

Gatti, M. *et al.* (2012) ‘A novel ubiquitin mark at the N-terminal tail of histone H2As targeted by RNF168 ubiquitin ligase’, *Cell Cycle*. Taylor and Francis Inc., 11(13), pp. 2538–2544. doi: 10.4161/cc.20919.

Gavet, O. and Pines, J. (2010) ‘Progressive Activation of CyclinB1-Cdk1 Coordinates Entry to Mitosis’, *Developmental Cell*, 18(4), pp. 533–543. doi: 10.1016/j.devcel.2010.02.013.

Gessaman, J. D. and Selker, E. U. (2017) ‘Induction of H3K9me3 and DNA methylation by tethered heterochromatin factors in *Neurospora crassa*’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 114(45), pp. E9598–E9607. doi: 10.1073/pnas.1715049114.

Gheghiani, L. *et al.* (2017) 'PLK1 Activation in Late G2 Sets Up Commitment to Mitosis', *Cell Reports*. Elsevier B.V., 19(10), pp. 2060–2073. doi: 10.1016/j.celrep.2017.05.031.

Gianferante, D. M., Mirabello, L. and Savage, S. A. (2017) 'Germline and somatic genetics of osteosarcoma - connecting aetiology, biology and therapy.', *Nature reviews. Endocrinology*, 13(8), pp. 480–491. doi: 10.1038/nrendo.2017.16.

Girard, F. *et al.* (1991) 'Cyclin a is required for the onset of DNA replication in mammalian fibroblasts', *Cell*, 67(6), pp. 1169–1179. doi: 10.1016/0092-8674(91)90293-8.

Gisselsson, D. *et al.* (2008) 'When the genome plays dice: Circumvention of the spindle assembly checkpoint and near-random chromosome segregation in multipolar cancer cell mitoses', *PLoS ONE*, 3(4). doi: 10.1371/journal.pone.0001871.

Goda, N. *et al.* (2003) 'Hypoxia-Inducible Factor 1 Is Essential for Cell Cycle Arrest during Hypoxia', *Molecular and Cellular Biology*. American Society for Microbiology, 23(1), pp. 359–369. doi: 10.1128/mcb.23.1.359-369.2003.

Gogineni, V. R. *et al.* (2011) 'Chk2-mediated G2/M cell cycle arrest maintains radiation resistance in malignant meningioma cells', *Cancer Letters*, 313(1), pp. 64–75. doi: 10.1016/j.canlet.2011.08.022.

Goldwasser, F. *et al.* (1996) 'Correlations between S and G2 arrest and the cytotoxicity of camptothecin in human colon carcinoma cells.', *Cancer research*, 56(19), pp. 4430–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8813137> (Accessed: 13 September 2019).

Goloborodko, A. *et al.* (2016) 'Compaction and segregation of sister chromatids via active loop extrusion', *eLife*. eLife Sciences Publications Ltd, 5(MAY2016). doi: 10.7554/eLife.14864.

Gonçalves, V., Matos, P. and Jordan, P. (2008) 'The β -catenin/TCF4 pathway modifies alternative splicing through modulation of SRp20 expression', *RNA*, 14(12), pp. 2538–

2549. doi: 10.1261/rna.1253408.

Gong, D. and Ferrell, J. E. (2010) 'The roles of cyclin A2, B1, and B2 in early and late mitotic events', *Molecular Biology of the Cell*, 21(18), pp. 3149–3161. doi: 10.1091/mbc.E10-05-0393.

Gorbalenya, A. E. and Koonin, E. V. (1993) 'Helicases: amino acid sequence comparisons and structure-function relationships', *Current Opinion in Structural Biology*, 3(3), pp. 419–429. doi: 10.1016/S0959-440X(05)80116-2.

Görisch, S. M. *et al.* (2005) 'Histone acetylation increases chromatin accessibility', *Journal of Cell Science*, 118(24), pp. 5825–5834. doi: 10.1242/jcs.02689.

Goswami, C. P. and Nakshatri, H. (2014) 'PROGgeneV2: Enhancements on the existing database', *BMC Cancer*. BioMed Central Ltd., 14(1). doi: 10.1186/1471-2407-14-970.

Grady, D. L. *et al.* (1992) *Highly conserved repetitive DNA sequences are present at human centromeres*, *Biochemistry*.

Graham, T. A. *et al.* (2002) 'The crystal structure of the β -catenin/ICAT complex reveals the inhibitory mechanism of ICAT', *Molecular Cell*. Cell Press, 10(3), pp. 563–571. doi: 10.1016/S1097-2765(02)00637-8.

Graham, T. G. W., Walter, J. C. and Loparo, J. J. (2016) 'Two-Stage Synapsis of DNA Ends during Non-homologous End Joining.', *Molecular cell*, 61(6), pp. 850–8. doi: 10.1016/j.molcel.2016.02.010.

Gray, R. S., Roszko, I. and Solnica-Krezel, L. (2011) 'Planar cell polarity: coordinating morphogenetic cell behaviors with embryonic polarity.', *Developmental cell*, 21(1), pp. 120–33. doi: 10.1016/j.devcel.2011.06.011.

Gupta, G. D. *et al.* (2015) 'A Dynamic Protein Interaction Landscape of the Human Centrosome-Cilium Interface.', *Cell*, 163(6), pp. 1484–99. doi: 10.1016/j.cell.2015.10.065.

- Gutierrez, G. J. *et al.* (2010) 'JNK-mediated phosphorylation of Cdc25C regulates cell cycle entry and G₂/M DNA damage checkpoint', *Journal of Biological Chemistry*, 285(19), pp. 14217–14228. doi: 10.1074/jbc.M110.121848.
- Güttinger, S., Laurell, E. and Kutay, U. (2009) 'Orchestrating nuclear envelope disassembly and reassembly during mitosis', *Nature Reviews Molecular Cell Biology*, pp. 178–191. doi: 10.1038/nrm2641.
- Haber, J. E. (1998) 'The many interfaces of Mre11', *Cell*. Cell Press, pp. 583–586. doi: 10.1016/S0092-8674(00)81626-8.
- Hagting, A. *et al.* (1998) 'MPF localization is controlled by nuclear export', *EMBO Journal*, 17(14), pp. 4127–4138. doi: 10.1093/emboj/17.14.4127.
- Hallett, S. T. *et al.* (2017) 'Differential Regulation of G1 CDK Complexes by the Hsp90-Cdc37 Chaperone System', *Cell Reports*, 21. doi: 10.1016/j.celrep.2017.10.042.
- Haluska, F. G., Tsujimoto, Y. and Croce, C. M. (1987) 'The t(8;14) chromosome translocation of the Burkitt lymphoma cell line Daudi occurred during immunoglobulin gene rearrangement and involved the heavy chain diversity region.', *Proceedings of the National Academy of Sciences of the United States of America*, 84(19), pp. 6835–6839. doi: 10.1073/pnas.84.19.6835.
- Hammel, M. *et al.* (2010) 'Ku and DNA-dependent protein kinase dynamic conformations and assembly regulate DNA binding and the initial non-homologous end joining complex', *Journal of Biological Chemistry*, 285(2), pp. 1414–1423. doi: 10.1074/jbc.M109.065615.
- Hanahan, D. and Weinberg, R. A. (2011) 'Hallmarks of cancer: The next generation', *Cell*, pp. 646–674. doi: 10.1016/j.cell.2011.02.013.
- Hannon, G. J., Demetrick, D. and Beach, D. (1993) 'Isolation of the Rb-related p130 through its interaction with CDK2 and cyclins', *Genes and Development*, 7(12 A), pp. 2378–2391. doi: 10.1101/gad.7.12a.2378.

Hardy, S. *et al.* (2009) ‘The euchromatic and heterochromatic landscapes are shaped by antagonizing effects of transcription on H2A.Z deposition’, *PLoS Genetics*, 5(10). doi: 10.1371/journal.pgen.1000687.

Harper, J. W. *et al.* (1993) ‘The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases.’, *Cell*, 75(4), pp. 805–16. doi: 10.1016/0092-8674(93)90499-g.

Harper, S. and Speicher, D. W. (2011) ‘Purification of Proteins Fused to Glutathione S-Transferase’, in, pp. 259–280. doi: 10.1007/978-1-60761-913-0_14.

Hart, M. *et al.* (1999) ‘The F-box protein β -TrCP associates with phosphorylated β -catenin and regulates its activity in the cell’, *Current Biology*. Current Biology Ltd, 9(4), pp. 207–211. doi: 10.1016/S0960-9822(99)80091-8.

He, T. C. *et al.* (1998) ‘Identification of c-MYC as a target of the APC pathway.’, *Science (New York, N.Y.)*, 281(5382), pp. 1509–12. doi: 10.1126/science.281.5382.1509.

Heald, R. and McKeon, F. (1990) ‘Mutations of phosphorylation sites in lamin A that prevent nuclear lamina disassembly in mitosis’, *Cell*, 61(4), pp. 579–589. doi: 10.1016/0092-8674(90)90470-Y.

Hein, M. Y. *et al.* (2015) ‘A Human Interactome in Three Quantitative Dimensions Organized by Stoichiometries and Abundances’, *Cell*. Cell Press, 163(3), pp. 712–723. doi: 10.1016/j.cell.2015.09.053.

Henikoff, S. and Smith, M. M. (2015) ‘Histone variants and epigenetics’, *Cold Spring Harbor Perspectives in Biology*. Cold Spring Harbor Laboratory Press, 7(1). doi: 10.1101/cshperspect.a019364.

Hershko, A. (1999) ‘Mechanisms and regulation of the degradation of cyclin B’, *Philosophical Transactions of the Royal Society B: Biological Sciences*. Royal Society, 354(1389), pp. 1571–1576. doi: 10.1098/rstb.1999.0500.

Hetzer, M. W. (2016) ‘The Nuclear Envelope: Methods and Protocols, Methods in

Molecular Biology', *Online*, 1411, pp. 255–267. doi: 10.1101/cshperspect.a000539.

Hirano, T. (2016) 'Leading Edge Review Condensin-Based Chromosome Organization from Bacteria to Vertebrates', *Cell*, 164, pp. 847–857. doi: 10.1016/j.cell.2016.01.033.

Hoffmann, I. *et al.* (1993) 'Phosphorylation and activation of human cdc25-C by cdc2--cyclin B and its involvement in the self-amplification of MPF at mitosis.', *The EMBO journal*, 12(1), pp. 53–63. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8428594> (Accessed: 12 September 2019).

Hogan, P. G. *et al.* (2003) 'Transcriptional regulation by calcium, calcineurin, and NFAT', *Genes and Development*, pp. 2205–2232. doi: 10.1101/gad.1102703.

Holland, A. J., Lan, W. and Cleveland, D. W. (2010) 'Centriole duplication: A lesson in self-control', *Cell Cycle*. Taylor and Francis Inc., pp. 2731–2736. doi: 10.4161/cc.9.14.12184.

Hu, J. H. *et al.* (2018) 'Evolved Cas9 variants with broad PAM compatibility and high DNA specificity', *Nature*. Nature Publishing Group, 556(7699), pp. 57–63. doi: 10.1038/nature26155.

Hu, Y. *et al.* (2011) 'RAP80-directed tuning of BRCA1 homologous recombination function at ionizing radiation-induced nuclear foci.', *Genes & development*, 25(7), pp. 685–700. doi: 10.1101/gad.2011011.

Huang, J. *et al.* (2019) 'Sox11 promotes head and neck cancer progression via the regulation of SDCCAG8', *Journal of Experimental and Clinical Cancer Research*. BioMed Central Ltd., 38(1). doi: 10.1186/s13046-019-1146-7.

Huang, Z., Traugh, J. A. and Bishop, J. M. (2004) 'Negative Control of the Myc Protein by the Stress-Responsive Kinase Pak2', *Molecular and Cellular Biology*. American Society for Microbiology, 24(4), pp. 1582–1594. doi: 10.1128/mcb.24.4.1582-1594.2004.

Huber, A. H., Nelson, W. J. and Weis, W. I. (1997) 'Three-dimensional structure of the

armadillo repeat region of beta-catenin.’, *Cell*, 90(5), pp. 871–82. doi: 10.1016/s0092-8674(00)80352-9.

Huen, M. S. Y. *et al.* (2007) ‘RNF8 Transduces the DNA-Damage Signal via Histone Ubiquitylation and Checkpoint Protein Assembly’, *Cell*, 131(5), pp. 901–914. doi: 10.1016/j.cell.2007.09.041.

Huttlin, E. L. *et al.* (2017) ‘Architecture of the human interactome defines protein communities and disease networks.’, *Nature*, 545(7655), pp. 505–509. doi: 10.1038/nature22366.

Idogawa, M. *et al.* (2007) ‘Ku70 and poly(ADP-ribose) polymerase-1 competitively regulate beta-catenin and T-cell factor-4-mediated gene transactivation: possible linkage of DNA damage recognition and Wnt signaling.’, *Cancer research*, 67(3), pp. 911–8. doi: 10.1158/0008-5472.CAN-06-2360.

Ilic, M. and Ilic, I. (2016) ‘Epidemiology of pancreatic cancer’, *World Journal of Gastroenterology*. Baishideng Publishing Group Co., Limited, pp. 9694–9705. doi: 10.3748/wjg.v22.i44.9694.

Jahn, S. C. *et al.* (2013) ‘Assembly, activation, and substrate specificity of cyclin D1/Cdk2 complexes’, *Biochemistry*, 52(20), pp. 3489–3501. doi: 10.1021/bi400047u.

Janssen, K. A., Sidoli, S. and Garcia, B. A. (2017) ‘Recent Achievements in Characterizing the Histone Code and Approaches to Integrating Epigenomics and Systems Biology’, in *Methods in Enzymology*. Academic Press Inc., pp. 359–378. doi: 10.1016/bs.mie.2016.10.021.

Jaramillo, M. *et al.* (1991) ‘RNA unwinding in translation: assembly of helicase complex intermediates comprising eukaryotic initiation factors eIF-4F and eIF-4B.’, *Molecular and Cellular Biology*. American Society for Microbiology, 11(12), pp. 5992–5997. doi: 10.1128/mcb.11.12.5992.

Jayasooriya, R. G. P. T. *et al.* (2018) ‘Camptothecin induces G2/M phase arrest through the ATM-Chk2- Cdc25C axis as a result of autophagy-induced cytoprotection:

Implications of reactive oxygen species', *Oncotarget*. Impact Journals LLC, 9(31), pp. 21744–21757. doi: 10.18632/oncotarget.24934.

Jeffrey, P. D. *et al.* (1995) 'Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex', *Nature*, 376(6538), pp. 313–320. doi: 10.1038/376313a0.

Jeffrey, P. D., Tong, L. and Pavletich, N. P. (2000) 'Structural basis of inhibition of CDK-cyclin complexes by INK4 inhibitors'. doi: 10.1101/gad.851100.

Jen, K. Y. and Cheung, V. G. (2003) 'Transcriptional response of lymphoblastoid cells to ionizing radiation', *Genome Research*, 13(9), pp. 2092–2100. doi: 10.1101/gr.1240103.

Jensen, R. B., Carreira, A. and Kowalczykowski, S. C. (2010) 'Purified human BRCA2 stimulates RAD51-mediated recombination.', *Nature*, 467(7316), pp. 678–83. doi: 10.1038/nature09399.

Jia, L., Li, B. and Yu, H. (2016) 'The Bub1-Plk1 kinase complex promotes spindle checkpoint signalling through Cdc20 phosphorylation', *Nature Communications*. Nature Publishing Group, 7. doi: 10.1038/ncomms10818.

Jinek, M. *et al.* (2012) 'A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity', *Science*. American Association for the Advancement of Science, 337(6096), pp. 816–821. doi: 10.1126/science.1225829.

Jirawatnotai, S. *et al.* (2011) 'A function for cyclin D1 in DNA repair uncovered by protein interactome analyses in human cancers', *Nature*, 474(7350), pp. 230–234. doi: 10.1038/nature10155.

Johnston, M. (1998) 'Gene chips: Array of hope for understanding gene regulation', *Current Biology*, 8, pp. 171–174. Available at: <http://biomednet.com/elecref/09609822008R0171> (Accessed: 14 September 2019).

Juan, J. *et al.* (2014) 'Diminished WNT → β -catenin → c-MYC signaling is a barrier for malignant progression of BRAFV600E-induced lung tumors', *Genes and Development*.

Cold Spring Harbor Laboratory Press, 28(6), pp. 561–575. doi: 10.1101/gad.233627.113.

Jun, S. *et al.* (2016) ‘LIG4 mediates Wnt signalling-induced radioresistance’, *Nature Communications*. Nature Publishing Group, 7. doi: 10.1038/ncomms10994.

Kallimasioti-Pazi, E. M. *et al.* (2018) ‘Heterochromatin delays CRISPR-Cas9 mutagenesis but does not influence the outcome of mutagenic DNA repair’, *PLoS Biology*. Public Library of Science, 16(12). doi: 10.1371/journal.pbio.2005595.

Kalsbeek, D. and Golsteyn, R. M. (2017) ‘G2/M-phase checkpoint adaptation and micronuclei formation as mechanisms that contribute to genomic instability in human cells’, *International Journal of Molecular Sciences*. MDPI AG. doi: 10.3390/ijms18112344.

Kam, W. W. Y. and Banati, R. B. (2013) ‘Effects of ionizing radiation on mitochondria’, *Free Radical Biology and Medicine*, pp. 607–619. doi: 10.1016/j.freeradbiomed.2013.07.024.

Kan, C. and Zhang, J. (2015) ‘BRCA1 mutation: A predictive marker for radiation therapy?’, *International Journal of Radiation Oncology Biology Physics*. Elsevier Inc., pp. 281–293. doi: 10.1016/j.ijrobp.2015.05.037.

Kaplan, H. S. (1960) ‘Cellular effects of ionizing radiation.’, *Bulletin of the New York Academy of Medicine*, 36, pp. 649–61. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/13751287> (Accessed: 13 September 2019).

Kari, V. *et al.* (2011) ‘The H2B ubiquitin ligase RNF40 cooperates with SUPT16H to induce dynamic changes in chromatin structure during DNA double-strand break repair’, *Cell Cycle*. Taylor and Francis Inc., 10(20), pp. 3495–3504. doi: 10.4161/cc.10.20.17769.

Kasahara, K. *et al.* (2010) ‘14-3-3 γ 3 mediates Cdc25A proteolysis to block premature mitotic entry after DNA damage’, *EMBO Journal*, 29(16), pp. 2802–2812. doi: 10.1038/emboj.2010.157.

Katoh, M. (2001) ‘Molecular cloning and characterization of LZIC, a novel gene

encoding ICAT homologous protein with leucine zipper domain.’, *International journal of molecular medicine*, 8(6), pp. 611–5. doi: 10.3892/ijmm.8.6.611.

Katsuno, Y. *et al.* (2009) ‘Cyclin A-Cdk1 regulates the origin firing program in mammalian cells’, *Proceedings of the National Academy of Sciences of the United States of America*, 106(9), pp. 3184–3189. doi: 10.1073/pnas.0809350106.

Kaykov, A. and Nurse, P. (2015) ‘The spatial and temporal organization of origin firing during the S-phase of fission yeast’, *Genome Research*. Cold Spring Harbor Laboratory Press, 25(3), pp. 391–401. doi: 10.1101/gr.180372.114.

Keller, C. E. *et al.* (2006) ‘Intrachromosomal rearrangement of chromosome 3q27: an under recognized mechanism of BCL6 translocation in B-cell non-Hodgkin lymphoma’, *Human Pathology*, 37(8), pp. 1093–1099. doi: 10.1016/j.humpath.2006.03.016.

Kelley, L. A. *et al.* (2015) ‘The Phyre2 web portal for protein modeling, prediction and analysis1. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc 10:845–858.’, *Nature Protocols*, 10(6), pp. 845–858. doi: 10.1038/nprot.2015.053.

Kim, B. Y. *et al.* (2011) ‘Phosphorylation and stabilization of c-Myc by NEMO renders cells resistant to ionizing radiation through up-regulation of γ -GCS’, *Oncology Reports*, 26(6), pp. 1587–1593. doi: 10.3892/or.2011.1432.

Kim, H. S. *et al.* (2009) ‘An acetylated form of histone H2A.Z regulates chromosome architecture in *Schizosaccharomyces pombe*’, *Nature Structural and Molecular Biology*, 16(12), pp. 1286–1293. doi: 10.1038/nsmb.1688.

Kim, J. Y. *et al.* (2017) ‘The value of phosphohistone H3 as a proliferation marker for evaluating invasive breast cancers: A comparative study with Ki67’, *Oncotarget*. Impact Journals LLC, 8(39), pp. 65064–65076. doi: 10.18632/oncotarget.17775.

Kim, K. H. and Sederstrom, J. M. (2015) ‘Assaying cell cycle status using flow cytometry’, *Current Protocols in Molecular Biology*. Blackwell Publishing Inc., 2015, pp. 28.6.1–28.6.11. doi: 10.1002/0471142727.mb2806s111.

Kitagawa, M. *et al.* (1999) 'An F-box protein, FWD1, mediates ubiquitin-dependent proteolysis of β -catenin', *EMBO Journal*, 18(9), pp. 2401–2410. doi: 10.1093/emboj/18.9.2401.

Knudsen, E. S. and Wang, J. Y. (1997) 'Dual mechanisms for the inhibition of E2F binding to RB by cyclin-dependent kinase-mediated RB phosphorylation.', *Molecular and cellular biology*, 17(10), pp. 5771–83. doi: 10.1128/mcb.17.10.5771.

Knudsen, K. E. *et al.* (1999) 'Cyclin A is a functional target of retinoblastoma tumor suppressor protein-mediated cell cycle arrest', *Journal of Biological Chemistry*, 274(39), pp. 27632–27641. doi: 10.1074/jbc.274.39.27632.

Kormish, J. D., Sinner, D. and Zorn, A. M. (2010) 'Interactions between SOX factors and Wnt/ β -catenin signaling in development and disease', *Developmental Dynamics*, pp. 56–68. doi: 10.1002/dvdy.22046.

Kotani, S. *et al.* (1998) 'PKA and MPF-activated polo-like kinase regulate anaphase-promoting complex activity and mitosis progression', *Molecular Cell*. Cell Press, 1(3), pp. 371–380. doi: 10.1016/S1097-2765(00)80037-4.

Kothe, M. *et al.* (2007) 'Structure of the catalytic domain of human polo-like kinase 1.', *Biochemistry*, 46(20), pp. 5960–71. doi: 10.1021/bi602474j.

Krempler, A. *et al.* (2007) 'An imperfect G₂/M checkpoint contributes to chromosome instability following irradiation of S and G₂ phase cells', *Cell Cycle*. Taylor and Francis Inc., pp. 1682–1686. doi: 10.4161/cc.6.14.4480.

Kühl, M. *et al.* (2000) 'Ca²⁺/calmodulin-dependent protein kinase II is stimulated by Wnt and Frizzled homologs and promotes ventral cell fates in *Xenopus*', *Journal of Biological Chemistry*, 275(17), pp. 12701–12711. doi: 10.1074/jbc.275.17.12701.

Kukurba, K. R. and Montgomery, S. B. (2015) 'RNA Sequencing and Analysis Kimberly', *Cold Spring Harbor protocols*, 2015(11), pp. 951–69. doi: 10.1101/pdb.top084970.RNA.

Kumagai, A. *et al.* (2006) 'TopBP1 activates the ATR-ATRIP complex.', *Cell*, 124(5), pp. 943–55. doi: 10.1016/j.cell.2005.12.041.

Kurita, M. *et al.* (2007) 'CR/periphilin is a transcriptional co-repressor involved in cell cycle progression.', *Biochemical and biophysical research communications*, 364(4), pp. 930–6. doi: 10.1016/j.bbrc.2007.10.090.

Kuzyk, A. and Mai, S. (2014) 'c-MYC-Induced Genomic Instability', *Cold Spring Harbor Perspectives in Medicine*, 4(4), pp. a014373–a014373. doi: 10.1101/cshperspect.a014373.

Kyogoku, Y., Lord, R. C. and Rich, A. (1966) 'Hydrogen bonding specificity of nucleic acid purines and pyrimidines in solution.', *Science*, 154(748), pp. 518–520. doi: 10.1126/science.154.3748.518.

Lal, A. *et al.* (2016) 'Genome scale patterns of supercoiling in a bacterial chromosome', *Nature Communications*. Nature Publishing Group, 7. doi: 10.1038/ncomms11055.

Lambert, S. A. *et al.* (2018) 'The Human Transcription Factors.', *Cell*, 172(4), pp. 650–665. doi: 10.1016/j.cell.2018.01.029.

Langerak, P. *et al.* (2011) 'Release of Ku and MRN from DNA ends by Mre11 nuclease activity and Ctp1 is required for homologous recombination repair of double-strand breaks', *PLoS Genetics*, 7(9). doi: 10.1371/journal.pgen.1002271.

Lanucara, F. and Eyers, C. E. (2011) 'Mass spectrometric-based quantitative proteomics using SILAC.', *Methods in enzymology*, 500, pp. 133–50. doi: 10.1016/B978-0-12-385118-5.00008-6.

Lara-Gonzalez, P., Westhorpe, F. G. and Taylor, S. S. (2012) 'Review- The Spindle Assembly Checkpoint (Higher Eukaryotes)', *Current Biology*, 22, pp. R966–R980. doi: 10.1016/j.cub.2012.10.006.

Latres, E., Chiaur, D. S. and Pagano, M. (1999) 'The human F box protein β -Trcp associates with the Cul1/Skp1 complex and regulates the stability of β -catenin',

Oncogene, 18(4), pp. 849–854. doi: 10.1038/sj.onc.1202653.

Leach, J. K. *et al.* (2001) *Ionizing Radiation-induced, Mitochondria-dependent Generation of Reactive Oxygen/Nitrogen 1, CANCER RESEARCH.*

Lee, D. H. *et al.* (2012) ‘Phosphoproteomic analysis reveals that PP4 dephosphorylates KAP-1 impacting the DNA damage response’, *EMBO Journal*, 31(10), pp. 2403–2415. doi: 10.1038/emboj.2012.86.

Lee, W. H. *et al.* (1987) ‘Human retinoblastoma susceptibility gene: cloning, identification, and sequence.’, *Science (New York, N.Y.)*, 235(4794), pp. 1394–9. doi: 10.1126/science.3823889.

van Leuken, R. *et al.* (2009) ‘Polo-like kinase-1 controls Aurora A destruction by activating APC/C-Cdh1’, *PLoS ONE*, 4(4). doi: 10.1371/journal.pone.0005282.

Lever, J., Krzywinski, M. and Altman, N. (2017) ‘Points of Significance: Principal component analysis’, *Nature Methods*. Nature Publishing Group, pp. 641–642. doi: 10.1038/nmeth.4346.

Li, Z. *et al.* (2003) ‘The VHL protein recruits a novel KRAB-A domain protein to repress HIF-1 α transcriptional activity.’, *The EMBO journal*, 22(8), pp. 1857–67. doi: 10.1093/emboj/cdg173.

Lieber, M. R. (2010) ‘The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway.’, *Annual review of biochemistry*, 79, pp. 181–211. doi: 10.1146/annurev.biochem.052308.093131.

Lin, Z., Luo, X. and Yu, H. (2016) ‘Structural basis of cohesin cleavage by separase’, *Nature*. Nature Publishing Group, 532(7597), pp. 131–134. doi: 10.1038/nature17402.

Lindqvist, A. *et al.* (2009) ‘Wip1 confers G2 checkpoint recovery competence by counteracting p53-dependent transcriptional repression’, *EMBO Journal*, 28(20), pp. 3196–3206. doi: 10.1038/emboj.2009.246.

Lipman, D. J. *et al.* (2002) 'The relationship of protein conservation and sequence length', *BMC Evolutionary Biology*, 2. doi: 10.1186/1471-2148-2-20.

Litt, M. *et al.* (2001) 'Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci.', *The EMBO journal*, 20(9), pp. 2224–35. doi: 10.1093/emboj/20.9.2224.

Liu, C. *et al.* (2002) 'Control of β -catenin phosphorylation/degradation by a dual-kinase mechanism', *Cell*. Cell Press, 108(6), pp. 837–847. doi: 10.1016/S0092-8674(02)00685-2.

Liu, J. *et al.* (2010) 'Human BRCA2 protein promotes RAD51 filament formation on RPA-covered single-stranded DNA', *Nature Structural and Molecular Biology*, 17(10), pp. 1260–1262. doi: 10.1038/nsmb.1904.

Liu, J. and Kipreos, E. T. (2000) 'Evolution of cyclin-dependent kinases (CDKs) and CDK-activating kinases (CAKs): Differential conservation of CAKs in yeast and metazoa', *Molecular Biology and Evolution*. Society for Molecular Biology and Evolution, 17(7), pp. 1061–1074. doi: 10.1093/oxfordjournals.molbev.a026387.

Liu, P.-P. *et al.* (2019) 'Polycomb Protein EED Regulates Neuronal Differentiation through Targeting SOX11 in Hippocampal Dentate Gyrus.', *Stem cell reports*, 13(1), pp. 115–131. doi: 10.1016/j.stemcr.2019.05.010.

Liu, Q. *et al.* (2000) 'Chk1 is an essential kinase that is regulated by Atr and required for the G2/M DNA damage checkpoint', *Genes & Development*. Cold Spring Harbor Laboratory Press, 14(12), pp. 1448–1459. doi: 10.1101/gad.14.12.1448.

Liu, X. *et al.* (2016) 'Sequence features associated with the cleavage efficiency of CRISPR/Cas9 system', *Scientific Reports*. Nature Publishing Group, 6. doi: 10.1038/srep19675.

Longhese, M. (1996) 'The 70 kDa subunit of replication protein A is required for the G1/S and intra-S DNA damage checkpoints in budding yeast', *Nucleic Acids Research*. Oxford University Press (OUP), 24(18), pp. 3533–3537. doi: 10.1093/nar/24.18.3533.

Lorca, T. and Castro, A. (2012) 'Deciphering the new role of the greatwall/ PP2A pathway in cell cycle control', *Genes and Cancer*, 3(11–12), pp. 712–720. doi: 10.1177/1947601912473478.

Lorca, T. and Castro, A. (2013) 'The Greatwall kinase: a new pathway in the control of the cell cycle.', *Oncogene*, 32(5), pp. 537–43. doi: 10.1038/onc.2012.79.

Loughery, J. *et al.* (2014) 'Critical role for p53-serine 15 phosphorylation in stimulating transactivation at p53-responsive promoters', *Nucleic Acids Research*. Oxford University Press, 42(12), pp. 7666–7680. doi: 10.1093/nar/gku501.

Luger, K. *et al.* (1997) 'Crystal structure of the nucleosome core particle at 2.8 Å resolution', *Nature*, 389(6648), pp. 251–260. doi: 10.1038/38444.

Lukas, J., Bartkova, J. and Bartek, J. (1996) 'Convergence of mitogenic signalling cascades from diverse classes of receptors at the cyclin D-cyclin-dependent kinase-pRb-controlled G1 checkpoint.', *Molecular and Cellular Biology*. American Society for Microbiology, 16(12), pp. 6917–6925. doi: 10.1128/mcb.16.12.6917.

Luo, Z., Rhie, S. K. and Farnham, P. J. (2019) 'The enigmatic hox genes: Can we crack their code?', *Cancers*. MDPI AG, 11(3). doi: 10.3390/cancers11030323.

Ma, T., Keller, J. A. and Yu, X. (2011) 'RNF8-dependent histone ubiquitination during DNA damage response and spermatogenesis', *Acta Biochimica et Biophysica Sinica*, pp. 339–345. doi: 10.1093/abbs/gmr016.

Ma, Y., Schwarz, K. and Lieber, M. R. (2005) 'The Artemis:DNA-PKcs endonuclease cleaves DNA loops, flaps, and gaps', *DNA Repair*, 4(7), pp. 845–851. doi: 10.1016/j.dnarep.2005.04.013.

MacDonald, B. T. and He, X. (2012) 'Frizzled and LRP5/6 receptors for wnt/ β -catenin signaling', *Cold Spring Harbor Perspectives in Biology*, 4(12). doi: 10.1101/cshperspect.a007880.

MacDonald, B. T., Tamai, K. and He, X. (2009) 'Wnt/ β -Catenin Signaling: Components,

Mechanisms, and Diseases’, *Developmental Cell*, pp. 9–26. doi: 10.1016/j.devcel.2009.06.016.

MacNeill, S. A. (2010) ‘Structure and function of the GINS complex, a key component of the eukaryotic replisome’, *Biochemical Journal*, pp. 489–500. doi: 10.1042/BJ20091531.

Macůrek, L. *et al.* (2008) ‘Polo-like kinase-1 is activated by aurora A to promote checkpoint recovery.’, *Nature*, 455(7209), pp. 119–23. doi: 10.1038/nature07185.

Mah, K. M. and Weiner, J. A. (2017) ‘Regulation of Wnt signaling by protocadherins.’, *Seminars in cell & developmental biology*, 69, pp. 158–171. doi: 10.1016/j.semcdb.2017.07.043.

Maiato, H. *et al.* (2017) ‘Mechanisms of chromosome congression during mitosis’, *Biology*. MDPI AG. doi: 10.3390/biology6010013.

Mailand, N. *et al.* (2007) ‘RNF8 ubiquitylates histones at DNA double-strand breaks and promotes assembly of repair proteins.’, *Cell*, 131(5), pp. 887–900. doi: 10.1016/j.cell.2007.09.040.

Mäkelä, T. P. *et al.* (1994) ‘A cyclin associated with the CDK-activating kinase MO15’, *Nature*, 371(6494), pp. 254–257. doi: 10.1038/371254a0.

Makharashvili, N. and Paull, T. T. (2015) ‘CtIP: A DNA damage response protein at the intersection of DNA metabolism.’, *DNA repair*, 32, pp. 75–81. doi: 10.1016/j.dnarep.2015.04.016.

Malumbres, M. *et al.* (2009) ‘Cyclin-dependent kinases: A family portrait’, *Nature Cell Biology*, pp. 1275–1276. doi: 10.1038/ncb1109-1275.

Mamely, I. *et al.* (2006) ‘Polo-like Kinase-1 Controls Proteasome-Dependent Degradation of Claspin during Checkpoint Recovery’, *Current Biology*, 16(19), pp. 1950–1955. doi: 10.1016/j.cub.2006.08.026.

Mandemaker, I. K. *et al.* (2017) 'DNA damage-induced histone H1 ubiquitylation is mediated by HUWE1 and stimulates the RNF8-RNF168 pathway', *Scientific Reports*. Nature Publishing Group, 7(1). doi: 10.1038/s41598-017-15194-y.

Manning, G. *et al.* (2017) 'Radiotherapy-associated long-term modification of expression of the inflammatory biomarker genes ARG1, BCL2L1, and MYC', *Frontiers in Immunology*. Frontiers Media S.A., 8(APR). doi: 10.3389/fimmu.2017.00412.

Marchesini, M. *et al.* (2017) 'ILF2 Is a Regulator of RNA Splicing and DNA Damage Response in 1q21-Amplified Multiple Myeloma.', *Cancer cell*, 32(1), pp. 88-100.e6. doi: 10.1016/j.ccell.2017.05.011.

Mattes, B. *et al.* (2012) 'Wnt3 and Wnt3a are required for induction of the mid-diencephalic organizer in the caudal forebrain', *Neural Development*, 7(1). doi: 10.1186/1749-8104-7-12.

McKinley, K. L. and Cheeseman, I. M. (2014) 'Polo-like kinase 1 licenses CENP-a deposition at centromeres', *Cell*. Cell Press, 158(2), pp. 397–411. doi: 10.1016/j.cell.2014.06.016.

Mehsen, H. *et al.* (2018) 'PP2A-B55 promotes nuclear envelope reformation after mitosis in *Drosophila*', *Journal of Cell Biology*. Rockefeller University Press, 217(12), pp. 4106–4123. doi: 10.1083/jcb.201804018.

Mellacheruvu, D. *et al.* (2013) 'The CRAPome: A contaminant repository for affinity purification-mass spectrometry data', *Nature Methods*, 10(8), pp. 730–736. doi: 10.1038/nmeth.2557.

Mello, M. L. S. (1983) 'Cytochemical properties of euchromatin and heterochromatin', *The Histochemical Journal*. Kluwer Academic Publishers, 15(8), pp. 739–751. doi: 10.1007/BF01003338.

Menon, S. *et al.* (2011) 'Association study of calcitonin gene-related polypeptide-alpha (CALCA) gene polymorphism with migraine', *Brain Research*, 1378, pp. 119–124. doi: 10.1016/j.brainres.2010.12.072.

Meselson, M. and Stahl, F. W. (1958) 'The replication of DNA in *Escherichia coli*', *Proceedings of the National Academy of Sciences*. Proceedings of the National Academy of Sciences, 44(7), pp. 671–682. doi: 10.1073/pnas.44.7.671.

Mitra, J. and Enders, G. H. (2004) 'Cyclin A/Cdk2 complexes regulate activation of Cdk1 and Cdc25 phosphatases in human cells', *Oncogene*, 23(19), pp. 3361–3367. doi: 10.1038/sj.onc.1207446.

Mizumoto, K. and Sawa, H. (2007) 'Two β s or not two β s: regulation of asymmetric division by β -catenin', *Trends in Cell Biology*, pp. 465–473. doi: 10.1016/j.tcb.2007.08.004.

Moiseeva, T. N. *et al.* (2019) 'An ATR and CHK1 kinase signaling mechanism that limits origin firing during unperturbed DNA replication', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 116(27), pp. 13374–13383. doi: 10.1073/pnas.1903418116.

Moll, U. M. and Petrenko, O. (2003) 'The MDM2-p53 interaction.', *Molecular cancer research*: MCR, 1(14), pp. 1001–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/14707283> (Accessed: 13 September 2019).

Mootha, V. K. *et al.* (2003) 'PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes', *Nature Genetics*, 34(3), pp. 267–273. doi: 10.1038/ng1180.

Moriya, H. (2015) 'Quantitative nature of overexpression experiments', *Molecular Biology of the Cell*. American Society for Cell Biology, pp. 3932–3939. doi: 10.1091/mbc.E15-07-0512.

Mulligan, K. A. and Cheyette, B. N. R. (2012) 'Wnt signaling in vertebrate neural development and function', *Journal of Neuroimmune Pharmacology*, pp. 774–787. doi: 10.1007/s11481-012-9404-x.

Nakanishi, M. *et al.* (1995) 'Identification of the active region of the DNA synthesis inhibitory gene p21Sdi1/CIP1/WAF1.', *The EMBO Journal*. European Molecular

Biology Organization, 14(3), p. 555.

Nam, E. A. *et al.* (2011) 'Thr-1989 phosphorylation is a marker of active ataxia telangiectasia- mutated and Rad3-related (ATR) kinase', *Journal of Biological Chemistry*, 286(33), pp. 28707–28714. doi: 10.1074/jbc.M111.248914.

Nick McElhinny, S. A. *et al.* (2000) 'Ku Recruits the XRCC4-Ligase IV Complex to DNA Ends', *Molecular and Cellular Biology*. American Society for Microbiology, 20(9), pp. 2996–3003. doi: 10.1128/mcb.20.9.2996-3003.2000.

Niehrs, C. and Acebron, S. P. (2012) 'Mitotic and mitogenic Wnt signalling', *EMBO Journal*, pp. 2705–2713. doi: 10.1038/emboj.2012.124.

Nilsson, J. (2019) 'Protein phosphatases in the regulation of mitosis', *Journal of Cell Biology*. Rockefeller University Press, pp. 395–409. doi: 10.1083/jcb.201809138.

Nimonkar, A. V. *et al.* (2011) 'BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair', *Genes and Development*, 25(4), pp. 350–362. doi: 10.1101/gad.2003811.

Nusse, R. and Varmus, H. E. (1982) 'Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome.', *Cell*, 31(1), pp. 99–109. doi: 10.1016/0092-8674(82)90409-3.

Nüsslein-volhard, C. and Wieschaus, E. (1980) 'Mutations affecting segment number and polarity in drosophila', *Nature*, 287(5785), pp. 795–801. doi: 10.1038/287795a0.

O'Brien, G. *et al.* (2018) 'FDXR is a biomarker of radiation exposure in vivo', *Scientific Reports*. Nature Publishing Group, 8(1). doi: 10.1038/s41598-017-19043-w.

O'Connell, M. J. *et al.* (1997) 'Chk1 is a wee1 kinase in the G2 DNA damage checkpoint inhibiting cdc2 by Y15 phosphorylation.', *The EMBO journal*, 16(3), pp. 545–54. doi: 10.1093/emboj/16.3.545.

Ohtani, K., Degregori, J. and Nevins, J. R. (1995) 'Regulation of the cyclin E gene by

transcription factor E2F1', *Proceedings of the National Academy of Sciences of the United States of America*, 92(26), pp. 12146–12150. doi: 10.1073/pnas.92.26.12146.

Oliveira, D. V. *et al.* (2014) 'Histone chaperone FACT regulates homologous recombination by chromatin remodeling through interaction with RNF20', *Journal of Cell Science*, 127(4), pp. 763–772. doi: 10.1242/jcs.135855.

Ong, F. S. *et al.* (2012) 'Personalized medicine and pharmacogenetic biomarkers: Progress in molecular oncology testing', *Expert Review of Molecular Diagnostics*, pp. 593–602. doi: 10.1586/erm.12.59.

Ortmann, B., Druker, J. and Rocha, S. (2014) 'Cell cycle progression in response to oxygen levels', *Cellular and Molecular Life Sciences*. Birkhauser Verlag AG, pp. 3569–3582. doi: 10.1007/s00018-014-1645-9.

Ou, C. Y. *et al.* (2011) 'A coactivator role of CARM1 in the dysregulation of β -catenin activity in colorectal cancer cell growth and gene expression', *Molecular Cancer Research*, 9(5), pp. 660–670. doi: 10.1158/1541-7786.MCR-10-0223.

Oughtred, R. *et al.* (2019) 'The BioGRID interaction database: 2019 update.', *Nucleic acids research*, 47(D1), pp. D529–D541. doi: 10.1093/nar/gky1079.

Painter, R. B. (1981) 'Radioresistant DNA synthesis: an intrinsic feature of ataxia telangiectasia', *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*, 84(1), pp. 183–190. doi: 10.1016/0027-5107(81)90061-0.

Pangon, L. *et al.* (2016) 'JRK is a positive regulator of β -catenin transcriptional activity commonly overexpressed in colon, breast and ovarian cancer.', *Oncogene*, 35(22), pp. 2834–41. doi: 10.1038/onc.2015.347.

Panier, S. and Boulton, S. J. (2014) 'Double-strand break repair: 53BP1 comes into focus', *Nature Reviews Molecular Cell Biology*, pp. 7–18. doi: 10.1038/nrm3719.

Pardee, A. B. (1974) 'A restriction point for control of normal animal cell proliferation', *Proceedings of the National Academy of Sciences of the United States of America*, 71(4),

pp. 1286–1290. doi: 10.1073/pnas.71.4.1286.

Patel, S. A. and Simon, M. C. (2010) ‘Functional analysis of the Cdk7·cyclin H·Mat1 complex in mouse embryonic stem cells and embryos’, *Journal of Biological Chemistry*, 285(20), pp. 15587–15598. doi: 10.1074/jbc.M109.081687.

Pavletich, N. P. (1999) ‘Mechanisms of cyclin-dependent kinase regulation: Structures of Cdks, their cyclin activators, and Cip and INK4 inhibitors’, *Journal of Molecular Biology*. Academic Press, 287(5), pp. 821–828. doi: 10.1006/jmbi.1999.2640.

Peifer, M., Berg, S. and Reynolds, A. B. (1994) ‘A repeating amino acid motif shared by proteins with diverse cellular roles.’, *Cell*, 76(5), pp. 789–91. doi: 10.1016/0092-8674(94)90353-0.

Peifer, M., Pai, L. M. and Casey, M. (1994) ‘Phosphorylation of the drosophila adherens junction protein armadillo: Roles for wingless signal and zeste-white 3 kinase’, *Developmental Biology*, 166(2), pp. 543–556. doi: 10.1006/dbio.1994.1336.

Peng, A. *et al.* (2010) ‘A novel role for greatwall kinase in recovery from DNA damage.’, *Cell cycle (Georgetown, Tex.)*, 9(21), pp. 4364–9. doi: 10.4161/cc.9.21.13632.

Petersen, M. B. *et al.* (1991) ‘Down syndrome due to de novo Robertsonian translocation t(14q;21q): DNA polymorphism analysis suggests that the origin of the extra 21q is maternal.’, *American journal of human genetics*, 49(3), pp. 529–36. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/1831959> (Accessed: 13 September 2019).

Pettijohn, D. E. (1988) *Histone-like Proteins and Bacterial Chromosome Structure**.

Pluta, A. F. *et al.* (1992) ‘Identification of a subdomain of CENP-B that is necessary and sufficient for localization to the human centromere’, *Journal of Cell Biology*. Rockefeller University Press, 116(5), pp. 1081–1093. doi: 10.1083/jcb.116.5.1081.

Portolano, N. *et al.* (2014) ‘Recombinant protein expression for structural biology in HEK 293F suspension cells: A novel and accessible approach’, *Journal of Visualized Experiments*. Journal of Visualized Experiments, (92). doi: 10.3791/51897.

Qi, Z. *et al.* (2015) 'DNA sequence alignment by microhomology sampling during homologous recombination.', *Cell*, 160(5), pp. 856–869. doi: 10.1016/j.cell.2015.01.029.

Qian, J. *et al.* (2011) 'PP1/Repo-man dephosphorylates mitotic histone H3 at T3 and regulates chromosomal aurora B targeting.', *Current biology : CB*, 21(9), pp. 766–73. doi: 10.1016/j.cub.2011.03.047.

Qiao, R. *et al.* (2016) 'Mechanism of APC/CCDC20 activation by mitotic phosphorylation', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 113(19), pp. E2570–E2578. doi: 10.1073/pnas.1604929113.

Ramsden, D. A. (2011) 'Polymerases in nonhomologous end joining: building a bridge over broken chromosomes.', *Antioxidants & redox signaling*, 14(12), pp. 2509–19. doi: 10.1089/ars.2010.3429.

Rashi-Elkeles, S. *et al.* (2011) 'Transcriptional modulation induced by ionizing radiation: P53 remains a central player', *Molecular Oncology*. John Wiley and Sons Ltd, pp. 336–348. doi: 10.1016/j.molonc.2011.06.004.

Rastogi, R. P. *et al.* (2010) 'Molecular mechanisms of ultraviolet radiation-induced DNA damage and repair', *Journal of Nucleic Acids*. doi: 10.4061/2010/592980.

Ray, P. D., Huang, B. W. and Tsuji, Y. (2012) 'Reactive oxygen species (ROS) homeostasis and redox regulation in cellular signaling', *Cellular Signalling*, pp. 981–990. doi: 10.1016/j.cellsig.2012.01.008.

Redon, C. E. *et al.* (2009) 'γ-H2AX as a biomarker of DNA damage induced by ionizing radiation in human peripheral blood lymphocytes and artificial skin', *Advances in Space Research*. Elsevier Ltd, 43(8), pp. 1171–1178. doi: 10.1016/j.asr.2008.10.011.

Redza-Dutordoir, M. and Averill-Bates, D. A. (2016) 'Activation of apoptosis signalling pathways by reactive oxygen species', *Biochimica et Biophysica Acta - Molecular Cell Research*. Elsevier B.V., pp. 2977–2992. doi: 10.1016/j.bbamcr.2016.09.012.

Reisz, J. A. *et al.* (2014) 'Effects of ionizing radiation on biological molecules - mechanisms of damage and emerging methods of detection', *Antioxidants and Redox Signaling*. Mary Ann Liebert Inc., pp. 260–292. doi: 10.1089/ars.2013.5489.

Rieder, C. L. and Maiato, H. (2004) 'Stuck in division or passing through: What happens when cells cannot satisfy the spindle assembly checkpoint', *Developmental Cell*, pp. 637–651. doi: 10.1016/j.devcel.2004.09.002.

Riethoven, J.-J. M. (2010) 'Regulatory Regions in DNA: Promoters, Enhancers, Silencers, and Insulators BT - Computational Biology of Transcription Factor Binding', in Ladunga, I. (ed.). Humana Press, pp. 33–42. doi: 10.1007/978-1-60761-854-6_3.

Rinn, J. L. and Chang, H. Y. (2012) 'Genome regulation by long noncoding RNAs.', *Annual review of biochemistry*, 81, pp. 145–66. doi: 10.1146/annurev-biochem-051410-092902.

Ritchie, M. E. *et al.* (2015) 'limma powers differential expression analyses for RNA-sequencing and microarray studies', *Nucleic Acids Research*. Oxford University Press, 43(7), pp. e47–e47. doi: 10.1093/nar/gkv007.

Robbez-Masson, L. *et al.* (2018) 'The hush complex cooperates with trim28 to repress young retrotransposons and new genes', *Genome Research*. Cold Spring Harbor Laboratory Press, 28(6), pp. 836–845. doi: 10.1101/gr.228171.117.

Robinson, P. J. J. *et al.* (2006) 'EM measurements define the dimensions of the "30-nm" chromatin fiber: Evidence for a compact, interdigitated structure', *Proceedings of the National Academy of Sciences of the United States of America*, 103(17), pp. 6506–6511. doi: 10.1073/pnas.0601212103.

Rohini, M., Haritha Menon, A. and Selvamurugan, N. (2018) 'Role of activating transcription factor 3 and its interacting proteins under physiological and pathological conditions', *International Journal of Biological Macromolecules*. Elsevier B.V., pp. 310–317. doi: 10.1016/j.ijbiomac.2018.08.107.

Ropars, V. *et al.* (2011) 'Structural characterization of filaments formed by human

Xrcc4-Cernunnos/XLF complex involved in nonhomologous DNA end-joining', *Proceedings of the National Academy of Sciences of the United States of America*, 108(31), pp. 12663–12668. doi: 10.1073/pnas.1100758108.

Ross, J. E., Woodlief, K. S. and Sullivan, B. A. (2016) 'Inheritance of the CENP-A chromatin domain is spatially and temporally constrained at human centromeres', *Epigenetics and Chromatin*. BioMed Central Ltd., 9(1). doi: 10.1186/s13072-016-0071-7.

Rothkamm, K. *et al.* (2003) 'Pathways of DNA Double-Strand Break Repair during the Mammalian Cell Cycle', *Molecular and Cellular Biology*. American Society for Microbiology, 23(16), pp. 5706–5715. doi: 10.1128/mcb.23.16.5706-5715.2003.

Rudolph, J. (2007) 'Cdc25 phosphatases: Structure, specificity, and mechanism', *Biochemistry*, pp. 3595–3604. doi: 10.1021/bi700026j.

Ruiz, E. J., Vilar, M. and Nebreda, A. R. (2010) 'A Two-Step Inactivation Mechanism of Myt1 Ensures CDK1/Cyclin B Activation and Meiosis I Entry', *Current Biology*, 20(8), pp. 717–723. doi: 10.1016/j.cub.2010.02.050.

Russo, A. A., Jeffrey, P. D. and Pavletich, N. P. (1996) 'Structural basis of cyclin-dependent kinase activation by phosphorylation', *Nature Structural Biology*, 3(8), pp. 696–700. doi: 10.1038/nsb0896-696.

Schachter, M. M. *et al.* (2013) 'A Cdk7-Cdk4 T-loop phosphorylation cascade promotes G1 progression.', *Molecular cell*, 50(2), pp. 250–60. doi: 10.1016/j.molcel.2013.04.003.

Schwarz-Romond, T. *et al.* (2002) 'The ankyrin repeat protein diversin recruits casein kinase I ϵ to the β -catenin degradation complex and acts in both canonical Wnt and Wnt/JNK signaling', *Genes and Development*, 16(16), pp. 2073–2084. doi: 10.1101/gad.230402.

Shalit, T. *et al.* (2015) 'MS1-based label-free proteomics using a quadrupole orbitrap mass spectrometer.', *Journal of proteome research*, 14(4), pp. 1979–86. doi: 10.1021/pr501045t.

Shaltiel, I. A. *et al.* (2014) 'Distinct phosphatases antagonize the p53 response in different phases of the cell cycle', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 111(20), pp. 7313–7318. doi: 10.1073/pnas.1322021111.

Shaltiel, I. A. *et al.* (2015) 'The same, only different - DNA damage checkpoints and their reversal throughout the cell cycle', *Journal of Cell Science*. Company of Biologists Ltd, pp. 607–620. doi: 10.1242/jcs.163766.

Shang, Z. F. *et al.* (2010) 'Inactivation of DNA-dependent protein kinase leads to spindle disruption and mitotic catastrophe with attenuated checkpoint protein 2 phosphorylation in response to DNA damage', *Cancer Research*, 70(9), pp. 3657–3666. doi: 10.1158/0008-5472.CAN-09-3362.

Shao, X. (2000) 'Common fold in helix-hairpin-helix proteins', *Nucleic Acids Research*. Oxford University Press (OUP), 28(14), pp. 2643–2650. doi: 10.1093/nar/28.14.2643.

Sheen, J. H., Woo, J. K. and Dickson, R. B. (2003) 'c-Myc alters the DNA damage-induced G2/M arrest in human mammary epithelial cells', *British Journal of Cancer*, 89(8), pp. 1479–1485. doi: 10.1038/sj.bjc.6601307.

Shen, C. C. *et al.* (2014) 'WNT16B from ovarian fibroblasts induces differentiation of regulatory T cells through β -catenin signal in dendritic cells', *International Journal of Molecular Sciences*. MDPI AG, 15(7), pp. 12928–12939. doi: 10.3390/ijms150712928.

Shen, T. and Huang, S. (2012) 'The role of Cdc25A in the regulation of cell proliferation and apoptosis', *Anti-cancer agents in medicinal chemistry*. NIH Public Access, 12(6), p. 631.

Sherr, C. J. (1993) 'Mammalian G1 cyclins', *Cell*, 73(6), pp. 1059–1065. doi: 10.1016/0092-8674(93)90636-5.

Sheu, Y.-J. and Stillman, B. (2010) 'The Dbf4-Cdc7 kinase promotes S phase by alleviating an inhibitory activity in Mcm4.', *Nature*, 463(7277), pp. 113–7. doi: 10.1038/nature08647.

Shiba-Ishii, A. *et al.* (2015) ‘Stratifin accelerates progression of lung adenocarcinoma at an early stage’, *Molecular Cancer*. BioMed Central Ltd., 14(1). doi: 10.1186/s12943-015-0414-1.

Shiotani, B. and Zou, L. (2009) ‘Single-Stranded DNA Orchestrates an ATM-to-ATR Switch at DNA Breaks’, *Molecular Cell*, 33(5), pp. 547–558. doi: 10.1016/j.molcel.2009.01.024.

Shohet, J. and Foster, J. (2017) ‘Neuroblastoma’, *BMJ*, p. j1863. doi: 10.1136/bmj.j1863.

Shoubridge, C. *et al.* (2013) ‘Europe PMC Funders Group Mutations in the guanine nucleotide exchange factor gene IQSEC2 cause nonsyndromic intellectual disability’, *Nature Genetics*, 42(6), pp. 486–488. doi: 10.1038/ng.588.Mutations.

Shreeram, S. *et al.* (2006) ‘Wip1 phosphatase modulates ATM-dependent signaling pathways.’, *Molecular cell*, 23(5), pp. 757–64. doi: 10.1016/j.molcel.2006.07.010.

Shtutman, M. *et al.* (1999) ‘The cyclin D1 gene is a target of the β -catenin/LEF-1 pathway’, *Proceedings of the National Academy of Sciences of the United States of America*, 96(10), pp. 5522–5527. doi: 10.1073/pnas.96.10.5522.

Silver, J. D., Ritchie, M. E. and Smyth, G. K. (2009) ‘Microarray background correction: maximum likelihood estimation for the normal-exponential convolution.’, *Biostatistics (Oxford, England)*, 10(2), pp. 352–63. doi: 10.1093/biostatistics/kxn042.

Simpson, R. J. (2006) ‘Fragmentation of Protein Using Trypsin’, *Cold Spring Harbor Protocols*. Cold Spring Harbor Laboratory, 2006(28), p. pdb.prot4550-pdb.prot4550. doi: 10.1101/pdb.prot4550.

Skalka, G. *et al.* (2019) ‘Leucine zipper and ICAT domain containing (LZIC) protein regulates cell cycle transitions in response to ionising radiation.’, *Cell Cycle*, p. 15384101.2019.1601476. doi: 10.1080/15384101.2019.1601476.

Smith, G. C. M. and Jackson, S. P. (2010) ‘The PIKK family of protein kinases’, in *Handbook of Cell Signaling*, 2/e. Elsevier Inc., pp. 575–580. doi: 10.1016/B978-0-12-

374145-5.00077-2.

Sobhian, B. *et al.* (2007) 'RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites.', *Science (New York, N.Y.)*, 316(5828), pp. 1198–202. doi: 10.1126/science.1139516.

Sokolov, M. and Neumann, R. (2015) 'Global gene expression alterations as a crucial constituent of human cell response to low doses of ionizing radiation exposure', *International Journal of Molecular Sciences*. MDPI AG. doi: 10.3390/ijms17010055.

Soutourina, J. (2018) 'Transcription regulation by the Mediator complex.', *Nature reviews. Molecular cell biology*, 19(4), pp. 262–274. doi: 10.1038/nrm.2017.115.

Stamos, J. L. and Weis, W. I. (2013) 'The β -catenin destruction complex', *Cold Spring Harbor Perspectives in Biology*. doi: 10.1101/cshperspect.a007898.

Starostina, N. G. and Kipreos, E. T. (2012) 'Multiple degradation pathways regulate versatile CIP/KIP CDK inhibitors', *Trends in Cell Biology*, pp. 33–41. doi: 10.1016/j.tcb.2011.10.004.

Stewart, G. S. *et al.* (2003) 'MDC1 is a mediator of the mammalian DNA damage checkpoint', *Nature*, 421(6926), pp. 961–966. doi: 10.1038/nature01446.

Stiff, T. *et al.* (2004) 'ATM and DNA-PK function redundantly to phosphorylate H2AX after exposure to ionizing radiation.', *Cancer research*, 64(7), pp. 2390–6. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15059890> (Accessed: 13 September 2019).

Subramanian, A. *et al.* (2005) 'Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles', *Proceedings of the National Academy of Sciences of the United States of America*, 102(43), pp. 15545–15550. doi: 10.1073/pnas.0506580102.

Sudakin, V., Chan, G. K. T. and Yen, T. J. (2001) 'Checkpoint inhibition of the APC/C in HeLa cells is mediated by a complex of BUBR1, BUB3, CDC20, and MAD2', *Journal of Cell Biology*, 154(5), pp. 925–936. doi: 10.1083/jcb.200102093.

Sullivan, B. A. and Karpen, G. H. (2004) 'Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin', *Nature Structural and Molecular Biology*, 11(11), pp. 1076–1083. doi: 10.1038/nsmb845.

Supek, F. *et al.* (2011) 'Revigo summarizes and visualizes long lists of gene ontology terms', *PLoS ONE*, 6(7). doi: 10.1371/journal.pone.0021800.

Sur, S. and Agrawal, D. K. (2016) 'Phosphatases and kinases regulating CDC25 activity in the cell cycle: clinical implications of CDC25 overexpression and potential treatment strategies', *Molecular and Cellular Biochemistry*. Springer New York LLC, pp. 33–46. doi: 10.1007/s11010-016-2693-2.

Syljuåsen, R. G. *et al.* (2006) 'Adaptation to the ionizing radiation-induced G2 checkpoint occurs in human cells and depends on checkpoint kinase 1 and polo-like kinase 1 kinases', *Cancer Research*. American Association for Cancer Research Inc., 66(21), pp. 10253–10257. doi: 10.1158/0008-5472.CAN-06-2144.

Szklarczyk, D. *et al.* (2019) 'STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets.', *Nucleic acids research*, 47(D1), pp. D607–D613. doi: 10.1093/nar/gky1131.

Tago, K. *et al.* (2000) 'Inhibition of Wnt signaling by ICAT, a novel β -catenin-interacting protein', *Genes & Development*. Cold Spring Harbor Laboratory Press, 14(14), pp. 1741–1749. doi: 10.1101/gad.14.14.1741.

Takahashi, Y., Rayman, J. B. and Dynlacht, B. D. (2000) 'Analysis of promoter binding by the E2F and pRB families in vivo: distinct E2F proteins mediate activation and repression', *Genes & Development*. Cold Spring Harbor Laboratory Press, 14(7), pp. 804–816. doi: 10.1101/gad.14.7.804.

Tamimi, Y. *et al.* (2007) 'Identification of a minimal region of loss on the short arm of chromosome 1 in Wilms tumor.', *Genes, chromosomes & cancer*, 46(4), pp. 327–35. doi: 10.1002/gcc.20413.

Tauriello, D. V. F. *et al.* (2012) 'Wnt/ β -catenin signaling requires interaction of the

Dishevelled DEP domain and C terminus with a discontinuous motif in Frizzled', *Proceedings of the National Academy of Sciences of the United States of America*, 109(14). doi: 10.1073/pnas.1114802109.

Taylor, S. L. *et al.* (2017) 'Effects of imputation on correlation: Implications for analysis of mass spectrometry data from multiple biological matrices', *Briefings in Bioinformatics*. Oxford University Press, 18(2), pp. 312–320. doi: 10.1093/bib/bbw010.

Terakawa, T. *et al.* (2017) 'The condensin complex is a mechanochemical motor that translocates along DNA', *Science*. American Association for the Advancement of Science, 358(6363), pp. 672–676. doi: 10.1126/science.aan6516.

Tewari, R. *et al.* (2010) 'Armadillo-repeat protein functions: questions for little creatures.', *Trends in cell biology*, 20(8), pp. 470–81. doi: 10.1016/j.tcb.2010.05.003.

Thullberg, M., Bartek, J. and Lukas, J. (2000) 'Ubiquitin/proteasome-mediated degradation of p19INK4d determines its periodic expression during the cell cycle.', *Oncogene*, 19(24), pp. 2870–6. doi: 10.1038/sj.onc.1203579.

Timofeev, O. *et al.* (2010) 'Cdc25 phosphatases are required for timely assembly of CDK1-cyclin B at the G2/M transition', *Journal of Biological Chemistry*, 285(22), pp. 16978–16990. doi: 10.1074/jbc.M109.096552.

Tjio, J. H. and Levan, A. (1956) 'THE CHROMOSOME NUMBER OF MAN', *Hereditas*, 42(1–2), pp. 1–6. doi: 10.1111/j.1601-5223.1956.tb03010.x.

Toyoshima-Morimoto, F., Taniguchi, E. and Nishida, E. (2002) 'Plk1 promotes nuclear translocation of human Cdc25C during prophase.', *EMBO reports*, 3(4), pp. 341–8. doi: 10.1093/embo-reports/kvf069.

Trimarchi, J. M. and Lees, J. A. (2002) 'Sibling rivalry in the E2F family', *Nature Reviews Molecular Cell Biology*, pp. 11–20. doi: 10.1038/nrm714.

Tsai, M.-H. *et al.* (2006) 'Transcriptional responses to ionizing radiation reveal that p53R2 protects against radiation-induced mutagenesis in human lymphoblastoid cells.',

Oncogene, 25(4), pp. 622–32. doi: 10.1038/sj.onc.1209082.

Tsou, M. F. B. and Stearns, T. (2006) ‘Controlling centrosome number: Licenses and blocks’, *Current Opinion in Cell Biology*, pp. 74–78. doi: 10.1016/j.ceb.2005.12.008.

Tyanova, S. *et al.* (2016) ‘The Perseus computational platform for comprehensive analysis of (prote)omics data.’, *Nature methods*, 13(9), pp. 731–40. doi: 10.1038/nmeth.3901.

Ueda, J. *et al.* (2014) ‘Heterochromatin dynamics during the differentiation process revealed by the DNA methylation reporter mouse, methylRO’, *Stem Cell Reports*. Cell Press, 2(6), pp. 910–924. doi: 10.1016/j.stemcr.2014.05.008.

Uemura, M. and Takeichi, M. (2006) ‘ α N-catenin deficiency causes defects in axon migration and nuclear organization in restricted regions of the mouse brain’, *Developmental Dynamics*, 235(9), pp. 2559–2566. doi: 10.1002/dvdy.20841.

Uhlen, M. *et al.* (2015) ‘Tissue-based map of the human proteome, Human Protein Atlas available from www.proteinatlas.org’, *Science*, 347(6220), pp. 1260419–1260419. doi: 10.1126/science.1260419.

Umezawa, H. (1976) ‘Structure and action of bleomycin.’, *Progress in biochemical pharmacology*, 11, pp. 18–27. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/63962> (Accessed: 13 September 2019).

Uziel, T. *et al.* (2003) ‘Requirement of the MRN complex for ATM activation by DNA damage’, *EMBO Journal*, 22(20), pp. 5612–5621. doi: 10.1093/emboj/cdg541.

Venter, C. J. *et al.* (2001) ‘The sequence of the human genome’, *Science*, 291(5507), pp. 1304–1351. doi: 10.1126/science.1058040.

Vigneron, S. *et al.* (2018) ‘Cyclin A-cdk1-Dependent Phosphorylation of Bora Is the Triggering Factor Promoting Mitotic Entry’, *Developmental Cell*. Cell Press, 45(5), pp. 637–650.e7. doi: 10.1016/j.devcel.2018.05.005.

Voets, E. and Wolthuis, R. M. F. (2010) 'MASTL is the human orthologue of Greatwall kinase that facilitates mitotic entry, anaphase and cytokinesis', *Cell Cycle*. Taylor and Francis Inc., 9(17), pp. 3591–3601. doi: 10.4161/cc.9.17.12832.

Wallingford, J. B. and Mitchell, B. (2011) 'Strange as it may seem: The many links between Wnt signaling, planar cell polarity, and cilia', *Genes and Development*, pp. 201–213. doi: 10.1101/gad.2008011.

Wan, C. *et al.* (2015) 'Panorama of ancient metazoan macromolecular complexes', *Nature*. Nature Publishing Group, 525(7569), pp. 339–344. doi: 10.1038/nature14877.

Wang, J. *et al.* (2017) 'HSPA5 Gene encoding Hsp70 chaperone BiP in the endoplasmic reticulum.', *Gene*, 618, pp. 14–23. doi: 10.1016/j.gene.2017.03.005.

Wang, R. *et al.* (2007) 'Regulation of Cdc25C by ERK-MAP kinases during the G2/M transition.', *Cell*, 128(6), pp. 1119–32. doi: 10.1016/j.cell.2006.11.053.

Warnatz, H. J. *et al.* (2011) 'The BTB and CNC homology 1 (BACH1) target genes are involved in the oxidative stress response and in control of the cell cycle', *Journal of Biological Chemistry*, 286(26), pp. 23521–23532. doi: 10.1074/jbc.M111.220178.

Wasner, M. *et al.* (2003) 'Cyclin B1 transcription is enhanced by the p300 coactivator and regulated during the cell cycle by a CHR-dependent repression mechanism', *FEBS Letters*. Elsevier, 536(1–3), pp. 66–70. doi: 10.1016/S0014-5793(03)00028-0.

Watanabe, Nobumoto *et al.* (2004) 'M-phase kinases induce phospho-dependent ubiquitination of somatic Wee1 by SCF β -TrCP', *Proceedings of the National Academy of Sciences of the United States of America*, 101(13), pp. 4419–4424. doi: 10.1073/pnas.0307700101.

Watson, J. D. and Crick, F. H. C. (1953) 'Molecular Structures of Nucleic Acids', *Nature*, pp. 737–738. doi: 10.1038/171737a0.

Watson, M. L. (1955) 'THE NUCLEAR ENVELOPE: ITS STRUCTURE AND RELATION TO CYTOPLASMIC MEMBRANES', *The Journal of Cell Biology*, 1(3),

pp. 257–270. doi: 10.1083/jcb.1.3.257.

Wei, D. *et al.* (2013) ‘Inhibition of protein phosphatase 2A radiosensitizes pancreatic cancers by modulating CDC25C/CDK1 and homologous recombination repair’, *Clinical Cancer Research*, 19(16), pp. 4422–4432. doi: 10.1158/1078-0432.CCR-13-0788.

Welburn, J. P. I. *et al.* (2007) ‘How tyrosine 15 phosphorylation inhibits the activity of cyclin-dependent kinase 2-cyclin A’, *Journal of Biological Chemistry*, 282(5), pp. 3173–3181. doi: 10.1074/jbc.M609151200.

White, M. K., Pagano, J. S. and Khalili, K. (2014) ‘Viruses and human cancers: A long road of discovery of molecular paradigms’, *Clinical Microbiology Reviews*. American Society for Microbiology, 27(3), pp. 463–481. doi: 10.1128/CMR.00124-13.

Williams, R., Williams, J. and Tainer, J. (2007) ‘Mre11-Rad50-Nbs1 is a keystone complex connecting DNA repair machinery, double-strand break signaling, and the chromatin template 1’, *Biochem. Cell Biol*, 85, pp. 509–520. doi: 10.1139/O07-069.

Wilsker, D. and Bunz, F. (2009) ‘Chk1 phosphorylation during mitosis: A new role for a master regulator’, *Cell Cycle*. Taylor and Francis Inc., pp. 1161–1163. doi: 10.4161/cc.8.8.8148.

Winey, M. and O’Toole, E. (2014) ‘Centriole structure.’, *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 369(1650). doi: 10.1098/rstb.2013.0457.

Won, K. A. *et al.* (1992) ‘Growth-regulated expression of D-type cyclin genes in human diploid fibroblasts’, *Proceedings of the National Academy of Sciences of the United States of America*, 89(20), pp. 9910–9914. doi: 10.1073/pnas.89.20.9910.

Wong, P. Y. *et al.* (2016) ‘MASTL(Greatwall) regulates DNA damage responses by coordinating mitotic entry after checkpoint recovery and APC/C activation’, *Scientific Reports*. Nature Publishing Group, 6(September 2015), pp. 1–12. doi: 10.1038/srep22230.

Woodruff, J. B., Wueseke, O. and Hyman, A. A. (2014) 'Pericentriolar material structure and dynamics', *Philosophical Transactions of the Royal Society B: Biological Sciences*. Royal Society of London. doi: 10.1098/rstb.2013.0459.

Wu, J. *et al.* (2012) 'RAP80 protein is important for genomic stability and is required for stabilizing BRCA1-A complex at DNA damage sites in vivo', *Journal of Biological Chemistry*, 287(27), pp. 22919–22926. doi: 10.1074/jbc.M112.351007.

Wu, X., Webster, S. R. and Chen, J. (2001) 'Characterization of Tumor-associated Chk2 Mutations', *Journal of Biological Chemistry*, 276(4), pp. 2971–2974. doi: 10.1074/jbc.M009727200.

Xiao, H. *et al.* (2018) 'Overexpression of protocadherin 7 inhibits neuronal survival by downregulating BIRC5 in vitro', *Experimental Cell Research*. Elsevier Inc., 366(1), pp. 71–80. doi: 10.1016/j.yexcr.2018.03.016.

Xu, B. *et al.* (2002) 'Two Molecularly Distinct G2/M Checkpoints Are Induced by Ionizing Irradiation', *Molecular and Cellular Biology*. American Society for Microbiology, 22(4), pp. 1049–1059. doi: 10.1128/mcb.22.4.1049-1059.2002.

Yamamori, T. *et al.* (2012) 'Ionizing radiation induces mitochondrial reactive oxygen species production accompanied by upregulation of mitochondrial electron transport chain function and mitochondrial content under control of the cell cycle checkpoint.', *Free radical biology & medicine*, 53(2), pp. 260–70. doi: 10.1016/j.freeradbiomed.2012.04.033.

Yang, S. *et al.* (2013) 'Inhibition of SCAMP1 suppresses cell migration and invasion in human pancreatic and gallbladder cancer cells.', *Tumour biology: the journal of the International Society for Oncodevelopmental Biology and Medicine*, 34(5), pp. 2731–9. doi: 10.1007/s13277-013-0825-9.

Yang, Y. *et al.* (2018) 'c-Myc regulates the CDK1/cyclin B1 dependent-G2/M cell cycle progression by histone H4 acetylation in Raji cells', *International Journal of Molecular Medicine*. Spandidos Publications, 41(6), pp. 3366–3378. doi: 10.3892/ijmm.2018.3519.

Yang, Y. and Mlodzik, M. (2015) 'Wnt-Frizzled/Planar Cell Polarity Signaling: Cellular Orientation by Facing the Wind (Wnt)', *Annual Review of Cell and Developmental Biology*. Annual Reviews, 31(1), pp. 623–646. doi: 10.1146/annurev-cellbio-100814-125315.

Yao, N. Y. and O'Donnell, M. (2010) 'SnapShot: The Replisome', *Cell*. Cell Press, pp. 1088-1088.e1. doi: 10.1016/j.cell.2010.05.042.

Ye, J. *et al.* (2017) 'NF45 and NF90/NF110 coordinately regulate ESC pluripotency and differentiation.', *RNA (New York, N.Y.)*, 23(8), pp. 1270–1284. doi: 10.1261/rna.061499.117.

Yekezare, M., Gómez-González, B. and Diffley, J. F. X. (2013) 'Controlling DNA replication origins in response to DNA damage - inhibit globally, activate locally', *Journal of Cell Science*, pp. 1297–1306. doi: 10.1242/jcs.096701.

Yoda, K. *et al.* (2000) 'Human centromere protein A (CENP-A) can replace histone H3 in nucleosome reconstitution in vitro', *Proceedings of the National Academy of Sciences of the United States of America*, 97(13), pp. 7266–7271. doi: 10.1073/pnas.130189697.

Yoo, H. Y. *et al.* (2004) 'Adaptation of a DNA replication checkpoint response depends upon inactivation of Claspin by the Polo-like kinase.', *Cell*, 117(5), pp. 575–88. doi: 10.1016/s0092-8674(04)00417-9.

Yu, J. *et al.* (2004) 'Greatwall kinase: A nuclear protein required for proper chromosome condensation and mitotic progression in *Drosophila*', *Journal of Cell Biology*, 164(4), pp. 487–492. doi: 10.1083/jcb.200310059.

Yu, J. *et al.* (2014) 'WLS Retrograde transport to the endoplasmic reticulum during Wnt secretion', *Developmental Cell*. Cell Press, 29(3), pp. 277–291. doi: 10.1016/j.devcel.2014.03.016.

Yun, J. *et al.* (2003) 'Cdk2-dependent phosphorylation of the NF-Y transcription factor and its involvement in the p53-p21 signaling pathway', *Journal of Biological Chemistry*, 278(38), pp. 36966–36972. doi: 10.1074/jbc.M305178200.

Zeegers, D. *et al.* (2017) 'Biomarkers of ionizing radiation exposure: A multiparametric approach', *Genome Integrity*. BioMed Central Ltd., 8(1). doi: 10.4103/2041-9414.198911.

Zerfass-Thome, K. *et al.* (1997) 'p27KIP1 blocks cyclin E-dependent transactivation of cyclin A gene expression.', *Molecular and Cellular Biology*. American Society for Microbiology, 17(1), pp. 407–415. doi: 10.1128/mcb.17.1.407.

Zhang, S. *et al.* (2012) 'Wnt/ β -Catenin Signaling Pathway Upregulates c-Myc Expression to Promote Cell Proliferation of P19 Teratocarcinoma Cells', *Anatomical Record*, 295(12), pp. 2104–2113. doi: 10.1002/ar.22592.

Zhang, X. *et al.* (2017) 'E2F1-regulated long non-coding RNA RAD51-AS1 promotes cell cycle progression, inhibits apoptosis and predicts poor prognosis in epithelial ovarian cancer', *Scientific Reports*. Nature Publishing Group, 7(1). doi: 10.1038/s41598-017-04736-z.

Zhang, Y. *et al.* (2013) 'Protein analysis by shotgun/bottom-up proteomics', *Chemical Reviews*, pp. 2343–2394. doi: 10.1021/cr3003533.

Zhang, Y. *et al.* (2019) 'Long noncoding RNA MEG3 inhibits breast cancer growth via upregulating endoplasmic reticulum stress and activating NF- κ B and p53', *Journal of Cellular Biochemistry*. Wiley-Liss Inc., 120(4), pp. 6789–6797. doi: 10.1002/jcb.27982.

Zhao, H. and Piwnica-Worms, H. (2001) 'ATR-Mediated Checkpoint Pathways Regulate Phosphorylation and Activation of Human Chk1', *Molecular and Cellular Biology*. American Society for Microbiology, 21(13), pp. 4129–4139. doi: 10.1128/mcb.21.13.4129-4139.2001.

Zhen, S. *et al.* (2014) 'In vitro and in vivo growth suppression of human papillomavirus 16-positive cervical cancer cells by CRISPR/Cas9.', *Biochemical and biophysical research communications*, 450(4), pp. 1422–6. doi: 10.1016/j.bbrc.2014.07.014.

Zimmermann, M. *et al.* (2013) '53BP1 regulates DSB repair using Rif1 to control 5' end resection', *Science*. American Association for the Advancement of Science, 339(6120),

pp. 700–704. doi: 10.1126/science.1231573.

Zou, L. and Elledge, S. J. (2003) ‘Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes.’, *Science (New York, N.Y.)*, 300(5625), pp. 1542–8. doi: 10.1126/science.1083430.

Zou, Y. *et al.* (2006) ‘Functions of human replication protein A (RPA): From DNA replication to DNA damage and stress responses’, *Journal of Cellular Physiology*, 208(2), pp. 267–273. doi: 10.1002/jcp.20622.

Appendix

RESEARCH PAPER



Leucine zipper and ICAT domain containing (LZIC) protein regulates cell cycle transitions in response to ionizing radiation

George Skalka ^{a*}, Holly Hall ^{a,b*}, Joanna Somers^{a*}, Martin Bushell^{a,b}, Anne Willis^a, and Michal Malewicz ^a

^aMRC Toxicology Unit, University of Cambridge, Leicester, UK; ^bBeatson Institute for Cancer Research, Glasgow, UK

ABSTRACT

Common hallmarks of cancer include the dysregulation of cell cycle progression and the acquisition of genome instability. In tumors, G1 cell cycle checkpoint induction is often lost. This increases the reliance on a functional G2/M checkpoint to prevent progression through mitosis with damaged DNA, avoiding the introduction of potentially aberrant genetic alterations. Treatment of tumors with ionizing radiation (IR) utilizes this dependence on the G2/M checkpoint. Therefore, identification of factors which regulate this process could yield important biomarkers for refining this widely used cancer therapy. Leucine zipper and ICAT domain containing (LZIC) downregulation has been associated with the development of IR-induced tumors. However, despite LZIC being highly conserved, it has no known molecular function. We demonstrate that LZIC knockout (KO) cell lines show a dysregulated G2/M cell cycle checkpoint following IR treatment. In addition, we show that LZIC deficient cells competently activate the G1 and early G2/M checkpoint but fail to maintain the late G2/M checkpoint after IR exposure. Specifically, this defect was found to occur downstream of PIKK signaling. The LZIC KO cells demonstrated severe aneuploidy indicative of genomic instability. In addition, analysis of data from cancer patient databases uncovered a strong correlation between LZIC expression and poor prognosis in several cancers. Our findings suggest that LZIC is functionally involved in cellular response to IR, and its expression level could serve as a biomarker for patient stratification in clinical cancer practice.

ARTICLE HISTORY

Received 19 August 2018
Revised 11 March 2019
Accepted 25 March 2019

KEYWORDS

Ionising radiation; DNA damage; cell cycle; checkpoint; G2/M; LZIC

Introduction

DNA damage can be induced by numerous internal and external sources, such as the collapse of DNA replication forks and exposure to exogenous high-energy radiation [1]. Upon recognition of DNA damage, cells mount a coordinated response of adaptive signaling pathways collectively termed the DNA damage response (DDR) [2]. In addition to DNA break repair pathways, the DDR includes a series of specialized DNA damage sensing and signaling proteins which arrest the cell at specific checkpoints during the cell cycle [3]. These checkpoints allow for the completion of DNA repair prior to DNA replication and cell division [4]. Importantly, checkpoints will activate depending on the specific modalities of damage, for example, activation of the G2/Mitosis (G2/M) checkpoint is associated with the exposure of cells to high-energy radiation [5,6]. The breakdown of cell cycle checkpoint control can be a precursor to multiple pathological conditions, such

as tumorigenesis. Most widely studied is the loss of p53 and p21 proteins resulting in failure to activate G1 checkpoint [7,8]. In these situations, the G2/M checkpoint becomes critically important for the maintenance of cell genome stability [9].

Activation and maintenance of the G2/M checkpoint is controlled by protein kinases. The phosphatidylinositol 3-kinase-related kinase (PIKK) family is activated following identification of DNA damage. Ataxia-telangiectasia mutated (ATM) and Ataxia-telangiectasia mutated and Rad3 related (ATR) are members of this family. One function of these proteins following damage is to activate the G2/M checkpoint signaling cascade [10]. To maintain the signal transduction cascade the master regulator of the G2/M signaling cascade, checkpoint protein 1 kinase (Chk1), is activated [11]. This requires phosphorylation of two serine residues at positions 345 (S345) and 317 (S317), which is mediated by ATR and ATM. Importantly, phosphorylated Chk1 is essential for

CONTACT Michal Malewicz  mzm23@mrc-tox.cam.ac.uk

*These authors contributed equally to this work

 Supplementary data for this article can be accessed [here](#).

© 2019 MRC Toxicology Unit, University of Cambridge. Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

the activation of the G2/M checkpoint in response to treatment with ionizing radiation (IR) [12]. Chk1 functions by phosphorylating specific inhibitory sites within cell cycle control proteins. An example of this is the phosphorylation of WEE1 by Chk1 in response to damage, which in turn induces an inhibitory phosphorylation event on Tyrosine 15 (Tyr15) of CDC2, inhibiting entry into mitosis [13]. The G2/M checkpoint is maintained until DNA repair has been completed at which point the checkpoint is deactivated and cells resume normal cell cycle. Release from cell cycle arrest is conducted by various protein phosphatase family members, such as PP2 and PP1. This activity is through the removal of phosphorylation from inhibitory sites on cell cycle controllers [14,15]. Incorrect functioning of any step within this process can lead to a dysfunctional G2/M checkpoint, which can result in chromosomal abnormalities, e.g., aneuploidy [16]. Cellular reaction to IR encompasses both direct repair response and induction of checkpoint signaling cascade. While many proteins which mediate these responses have been identified, further investigation into these response pathways is required to understand the nuances of control.

One protein, which was linked to cellular IR response, is the Leucine zipper and ICAT domain containing (LZIC) protein [17]. LZIC is a putative member of the WNT signaling family [17]. The LZIC protein is composed of 190 amino acids (21 kDa) and contains two domains, an N-terminal coiled-coil and a C-terminal ICAT-like domain (Supplemental Figure 1A). Unlike ICAT protein, which antagonizes WNT signaling by binding and inhibiting β -catenin, LZIC protein does not interact with β -catenin [18]. Furthermore, in a rat model of IR-induced osteosarcoma reduced LZIC expression was associated with the onset of oncogenesis [19–21]. To investigate the function of LZIC protein we have employed CRISPR technology to derive LZIC knockout (KO) cell lines. Our data show that LZIC is a component of the cellular response to IR. LZIC deficient cells show dysregulated transcription after IR treatment and fail to efficiently maintain the G2/M checkpoint, with the generation of severe genomic instability. Finally, analysis of patient databases identified a positive correlation between LZIC expression and average patient survival time in a number of cancers, suggesting that LZIC expression could serve as a biomarker for patient stratification.

Results

LZIC deletion leads to gene expression changes following treatment with ionizing radiation

LZIC is a putative member of the WNT signaling pathway, which typically regulates the activity of TCF/LEF family transcription factors and has been implicated in response to IR [22]. As such, we sought to determine the impact of LZIC loss on late transcriptional regulation following IR [23]. To address this question, CRISPR was used to generate an HEK293 cell line with a deletion of LZIC (LZIC KO Clone 1) and a control line, which has undergone the CRISPR process, but with no LZIC deletion (Supplementary Figure 1B).

Differential expression was determined by comparing whole genome expression profiling 24 h following 5 Gy of IR with cells which were left untreated for both LZIC KO and the CRISPR control (Figure 1A). Genes involved in the response to DNA damage were found to be differentially regulated following treatment with IR in the CRISPR control (Figure 1, Group A). Following the loss of LZIC expression, we detected 42 genes which are uniquely regulated (Figure 1, Group B). To further investigate the relationship between these groups a z-score analysis was conducted. This indicates strongly related clusters of genes between the cell lines, with the reduced expression of histone subunits being most conserved following treatment with IR (Figure 1B). However, differences between the two cell lines can be observed, with a dysregulation of several long-non-coding (lnc) RNAs and a downregulation DHRS2, which is involved in the p53 regulatory cascade, in LZIC KO conditions.

To directly investigate the loss of LZIC on the transcriptome, the differential expression between LZIC KO cells and CRISPR control was determined. In untreated conditions, we identified a total of 62 unique genes which are differentially regulated following LZIC loss (Figure 1A, Group C). Genes involved in neuronal development, such as FOXQ1 and Peripherin, are present in line with previous reports of LZIC function. In comparison, we found 24 unique genes which are differentially regulated in response to IR following LZIC loss (Figure 1A, Group D). This group includes genes such as PLK2, which has a role in cell division. Among the genes identified following LZIC KO, regardless of

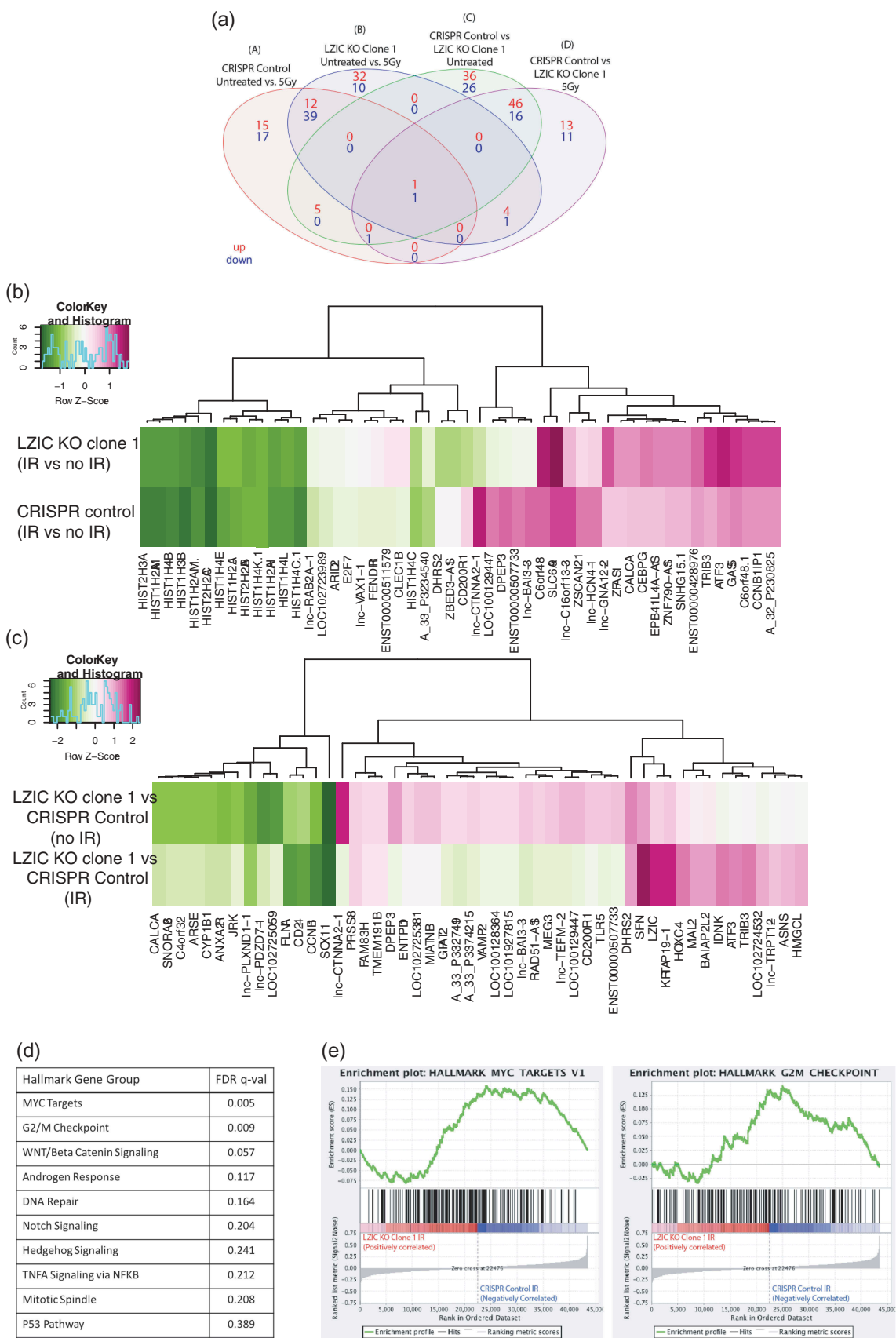


Figure 1. The loss of LZIC expression alters transcriptome under basal conditions and in response to ionizing radiation. (A) Venn diagram comparing the numbers of genes with significantly altered log-fold changes between LZIC KO clone vs CRISPR control cells in control and IR conditions. (B) Heatmaps representing Z-scores for genes comparing transcriptomic profiles in CRISPR control to LZIC KO Clone 1 cell lines in response to 5 Gy IR. (C) Heatmaps representing Z-scores for comparison of transcriptional profiles between CRISPR control and LZIC KO Clone 1 either under basal condition or following IR exposure. (D) Hallmark gene groups analyzed in GSEA and their associated FDR q-value (E) Barcode plots for significant GSEA hallmark gene groups. Microarray was repeated on two separate biological repeats with two technical repeats of each condition.

treatment with IR, were SFN and CCBN1, which are critical regulators of the G2/M checkpoint (Figure 1C). The 10 most significantly altered transcripts from each unique group are highlighted in table form (Supplementary Figure 1C). Differential expression of 10 mRNAs was validated by qPCR (Supplementary Figure 2A & 2B).

To examine specific pathways which were dysregulated in LZIC KO we utilized GSEA (Gene set enrichment analysis) using MSigDB (Molecular Signatures Database) hallmark gene sets. This revealed that LZIC KO causes alteration of MYC signaling and G2/M checkpoint pathways following treatment with IR (Figure 1D & 1E) [24,25]. This analysis was also performed on the basal conditions, identifying MYC targets and E2F targets, suggesting that MYC alterations are LZIC dependent, with changes to G2/M and E2F targets being treatment dependent (Supplementary Figure 2C).

Overall our transcriptome analysis found that LZIC KO cells had an altered transcriptional profile under both basal conditions and after treatment with IR, with a particular focus on cell cycle regulation.

LZIC loss leads to increased release from G2/M phase in response to IR

Our transcriptomic analysis found dysregulation of mRNA for critical G2/M checkpoint regulatory genes following treatment with IR in LZIC KO. Altered abundance of cyclin B1 and SFN, in particular, have been linked to progression through the G2/M transition with damaged DNA [26]. Therefore, we used flow cytometry to assess changes in cell cycle distribution in LZIC KO cells following IR treatment. The parental line and an additional LZIC knockout line (LZIC KO Clone 2) were included in this analysis to increase the robustness of derived conclusions (Supplementary Figure 3A & 3B). We observed G2/M checkpoint induction in all cell lines at 8-h post-IR (Figure 2A middle panel). Interestingly, when measured at 24hr post-IR LZIC KO cell lines showed a significantly reduced G2/M population, with a concurrent increase of cells present in the G1 phase (Figure 2A bottom panel). This effect was specific to exposure to IR, since cells treated with camptothecin (CPT), cobalt chloride (CoCl₂), or Ultraviolet light (UV) showed no phenotype

(Supplementary Figure 4A and 4B). To confirm that altered cell cycle distribution was LZIC KO specific and not due to off-target effects, a FLAG-tagged LZIC cDNA was stably introduced into the LZIC KO Clone 2 line. The expression of exogenous LZIC protein was lower than endogenous levels (Supplementary Figure 2A), despite this, exogenous LZIC partially reversed the KO phenotype confirming its specificity (Figure 2B). While the data suggests a defective G2/M checkpoint, the activation of G1 checkpoint following IR treatment was assessed. We used phosphorylation of p53 serine 15 (Ser 15) as a marker of G1 checkpoint signaling induction as it occurs in response to DNA damage and promotes association with p53-responsive promoters [27]. The phosphorylation of p53 Ser 15 is consistent across all the cell lines indicating correct induction of G1 checkpoint signaling irrespective of LZIC loss (Figure 2C).

Two G2/M checkpoints have been characterized: a minor immediate (within 1-h post-IR) ATM-dependent checkpoint and a major ATM-independent G2 accumulation checkpoint [5]. To determine whether induction of early G2/M checkpoint was perturbed, techniques demonstrated by Xu, et.al, were utilized [5]. ATM inhibitor-treated cells were utilized as an experimental control and show an increase in the mitotic ratio relative to the WT cells, indicating loss of the ATM-dependent early checkpoint (Figure 2D). In contrast, the LZIC KO cell lines show no deviation from the WT at the time points measured indicating a correct activation of the early G2/M checkpoint. Finally, the phosphorylation of histone 3 Serine 10 (pS10 H3) occurs upon entry into late G2 and persists until rapid dephosphorylation occurs in early G1 [28]. Quantification of the pS10 H3 population gives a further measure of those cells present in late G2 and mitosis. We found that at 24-h post-IR LZIC KO cells had a reduced positively stained population (Figure 2E). We conclude that LZIC KO cells successfully induce activation of cell cycle checkpoints but fail to sustain the late G2/M checkpoint and proceed to mitosis prematurely.

Defective signaling downstream of Chk1 in LZIC KO cells

The ATR and ATM kinases are essential for the establishment of the G2/M checkpoint following

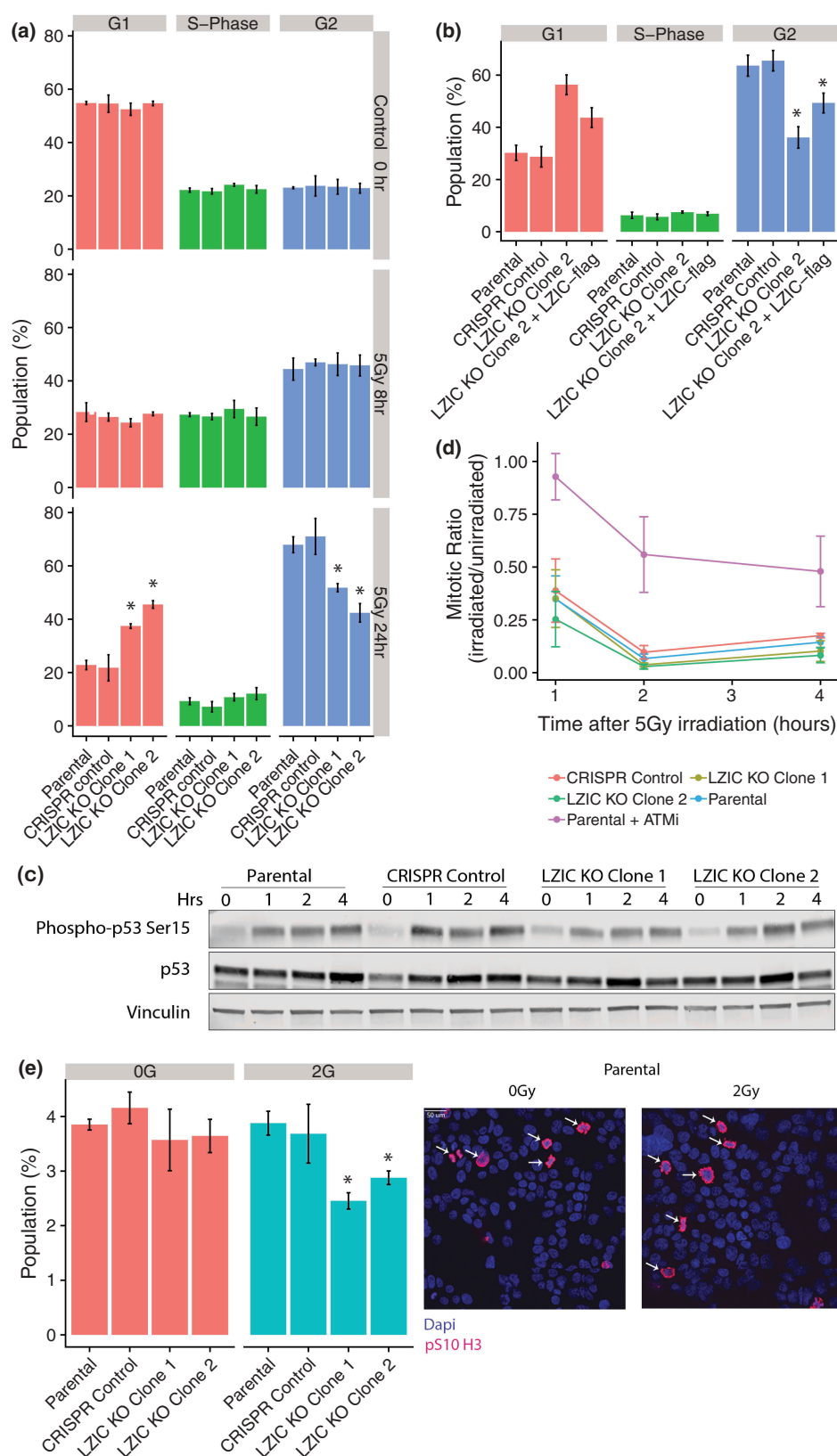


Figure 2. Late G2/M checkpoint arrest is perturbed following LZIC loss. (A) Cell cycle analysis of cell lines by propidium iodide staining. Cell lines were treated with 5 Gy IR and following a 24hr incubation harvested for analysis. Graphs are based on four separate biological repeats (B) Cell cycle analysis of LZIC KO Clone 2 and stable re-expression of LZIC-flag. Graphs based on three separate biological repeats. (C) p53 phosphorylation status in all cell lines, over a 4-h time course, following treatment with IR. A representative blot is shown from three separate biological repeats. (D) Activation of early G2/M checkpoint induction, following treatment with IR. The inclusion of Parental + ATMi provides a positive control for loss of early G2/M checkpoint activation. Graphs based on three separate biological repeats. (E) Quantification of phosphorylated serine 10 on Histone 3. Cells were treated with 2 Gy IR and harvested at 24 h. Images indicate staining profile with arrows to denote the quantified cells. Graphs based on three separate biological repeats. All statistical significance was determined with unpaired student T-Test, * = p-value < 0.05, n.s = non-significant. CRISPR control and LZIC KO Clones were compared to the parental line.

damage induction (Figure 3D) [11]. The activity of ATR and ATM following treatment with IR is regulated by phosphorylation on specific activation residues [29,30]. Analysis of canonical ATR and ATM activation sites show no impact of LZIC loss upon phosphorylation following IR treatment (Figure 3A). The major cell cycle targets of these kinases are both Chk1 and checkpoint protein 2 (Chk2). While Chk1 is the master regulator of G2/M, checkpoint interplay with Chk2 has been observed [31]. Analysis of Chk2 expression levels and activation showed no deviation between LZIC KO cells and control lines. In contrast, the phosphorylation of Chk1 serine 317 was reduced in the absence of LZIC (Figure 3B). The phosphorylation status of Chk1 has a direct impact upon its function, particularly, serine 317 which can reduce the activity of the other two major activation sites

serine 296 and serine 345 [12]. Therefore, the phosphorylation status of downstream components reliant on Chk1 activation was analyzed.

The mitosis promoting factor (MPF) is a complex containing cyclin B1 and CDC2 [32]. Phosphorylation of CDC2 at tyrosine 15, a DNA damage-induced phosphorylation site, occurs through a Chk1 mediated pathway and was reduced in LZIC KO clones [33]. Given interdependence between the MPF components, we further investigated the status of cyclin B1 in this condition, as the expression levels of cyclin B1 and phosphorylation of cytoplasmic to nuclear import sites directly effect the progression of cells through mitosis. LZIC KO cells showed reduced expression levels of cyclin B1 at 8-h and 24-h post-IR. Furthermore, phosphorylation of cyclin B1 at serine 147, a site involved in nuclear shuttling, was

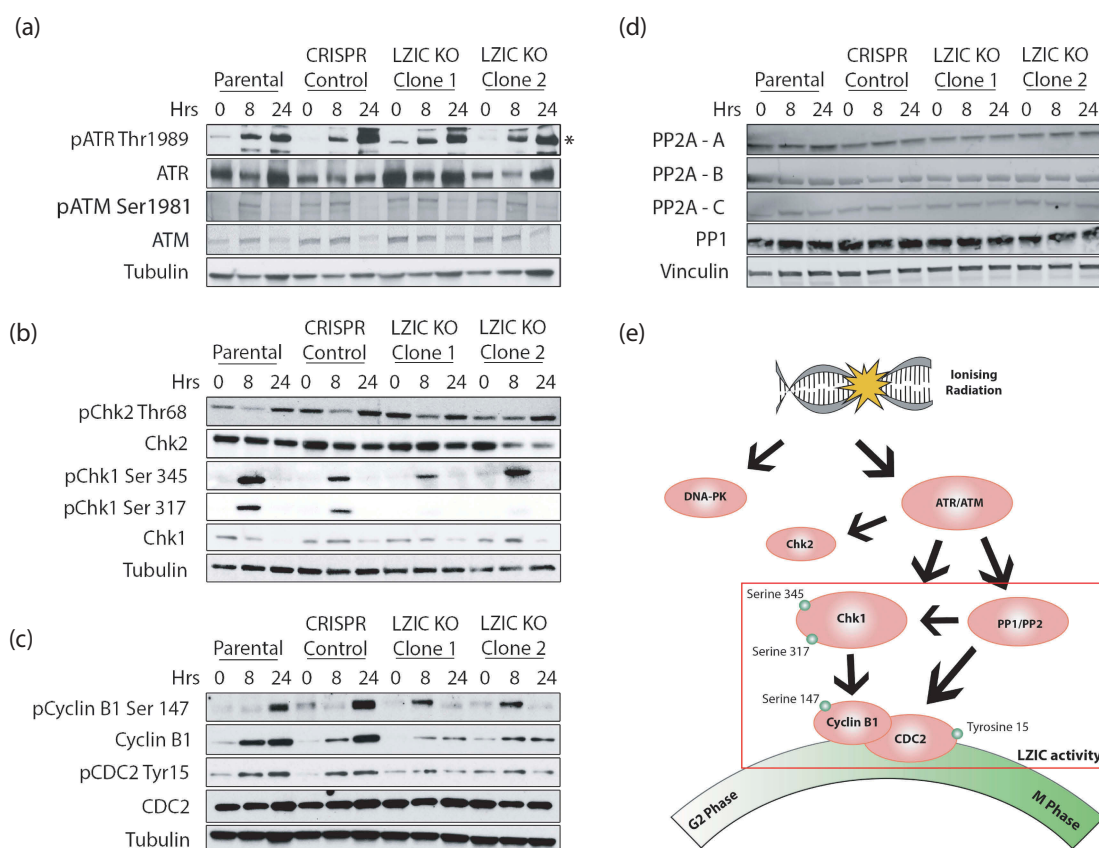


Figure 3. Cellular signaling for activation of G2/M checkpoint in response to IR is perturbed in LZIC KO cells. (A) Analysis of ATR and ATM phosphorylation status at 8 and 24-h post-treatment with 5 Gy IR. (*) indicates protein band corresponding to ATR protein. (B) Western blot analysis of checkpoint proteins at 8 h and 24 h in all cell types following treatment with 5 Gy IR. (C) Western blot analysis of Mitosis promoting factors at 8 and 24 h following treatment with 5 Gy IR. (D) Western blot analysis of major G2/M phosphatases, PP1 and PP2A, at 8 and 24 h following treatment with IR. (E) Schematic diagram of regulatory cascade showing key proteins involved in the G2/M cell cycle progression and their DNA damage-induced phosphorylation sites. All western blots shown are representative image of three separate biological repeats.

aberrant in LZIC KO cells with peak phosphorylation occurring at 8-h post-IR relative to 24 h seen in control cell lines (Figure 3C) [34]. This data suggests that the MPF complex regulation is altered in response to IR treatment, following LZIC loss, facilitating progression through the G2/M checkpoint into mitosis.

The phosphorylation status of Chk1 is controlled by the interplay between the PIKK proteins and removal of phosphorylation by the protein phosphatase family. To assure that LZIC loss did not lead to loss of phosphatase expression, overall expression of protein phosphatase 1 (PP1) and protein phosphatase 2 A (PP2A) was conducted. We show that the overall expression levels of the phosphatases are unchanged (Figure 3D). We conclude that LZIC operates downstream of PIKK signaling and that LZIC KO cells show a selective defect in the execution of IR-induced signaling which converges on the MPF (Figure 3E).

Loss of LZIC leads to genome instability and poor prognosis for clear renal cell carcinoma

A premature release of cells from the G2/M checkpoint increases the chance of chromosome loss and the development of aneuploidy (Figure 4A) [35]. Cells were either left untreated or exposed to IR and metaphase spreads were used to determine chromosome numbers. Under basal conditions, LZIC KO cell lines showed a reduced number of chromosomes when compared to controls. Similar chromosome loss was observed in control cells following IR exposure (Figure 4B). Notably, the genome instability observed in LZIC cells following IR does not increase beyond observed levels in the untreated population. Analysis of cell viability following treatment with IR indicates an increased sensitivity for LZIC KO, which could indicate that the population with increased genome instability are lost (Supplementary Figure 4C). These data suggest that LZIC KO cell lines had spontaneously undergone chromosome loss before IR treatment and that IR-induced instability can generate an equivalent outcome in control cells.

Next, we analyzed available cancer patient databases of RNA-seq data to determine whether there was a correlation of LZIC RNA expression levels with patient prognosis [36]. Although LZIC RNA expression correlated with poor patient survival for a range of cancers, the most striking effect was

observed for the clear renal cell carcinoma and neuroendocrine tumors, in which reduction of LZIC expression correlated with a severe decrease in average patient survival times (Figure 4C).

Discussion

A major treatment modality for cancer is IR, which is used in isolation or in combination with small molecular inhibitors and chemical chemotherapy. The identification of biomarkers for sensitivity to IR is important for improving response rates to this treatment. LZIC expression was shown to be specifically downregulated during the development of IR-initiated oncogenesis [21]. However, the cellular function of LZIC is currently unknown. This investigation aimed to identify the role of LZIC within the cell, and more specifically, the IR response cascade.

This study has generated human LZIC KO cell lines to investigate the effect of LZIC loss on the transcriptomic response to IR. From these data, we can conclude that following IR treatment, LZIC acts to regulate the cell cycle checkpoint cascade, more specifically at the G2/M checkpoint. To our knowledge, this is the first study to suggest such a function for LZIC. In general, the increased activity of WNT signaling proteins at the G2/M checkpoint, and during mitotic spindle assembly, has been widely characterized [37]. One example of which is the interplay between β -catenin and DNA ligase IV being an important radioresistance determinant [38]. Therefore, this finding agrees with current roles for WNT signaling proteins. In addition, WNT signaling is an important pathway during oncogenesis, with the identification of altered LZIC regulation having been established in multiple cancers [19–21]. The hypothesis presented here suggests reduced LZIC expression is linked to induced oncogenesis by decreased checkpoint control.

The transcriptomic analysis of LZIC KO cells identified altered MYC signaling in untreated and treated conditions. This suggested that the altered regulation of this pathway is not IR specific and is, instead, a direct response to the loss of LZIC. The regulation of MYC signaling by WNT pathway proteins, for example, the upregulation of c-myc by β -catenin, can promote cell proliferation and differentiation [39,40]. Further investigation would be required to

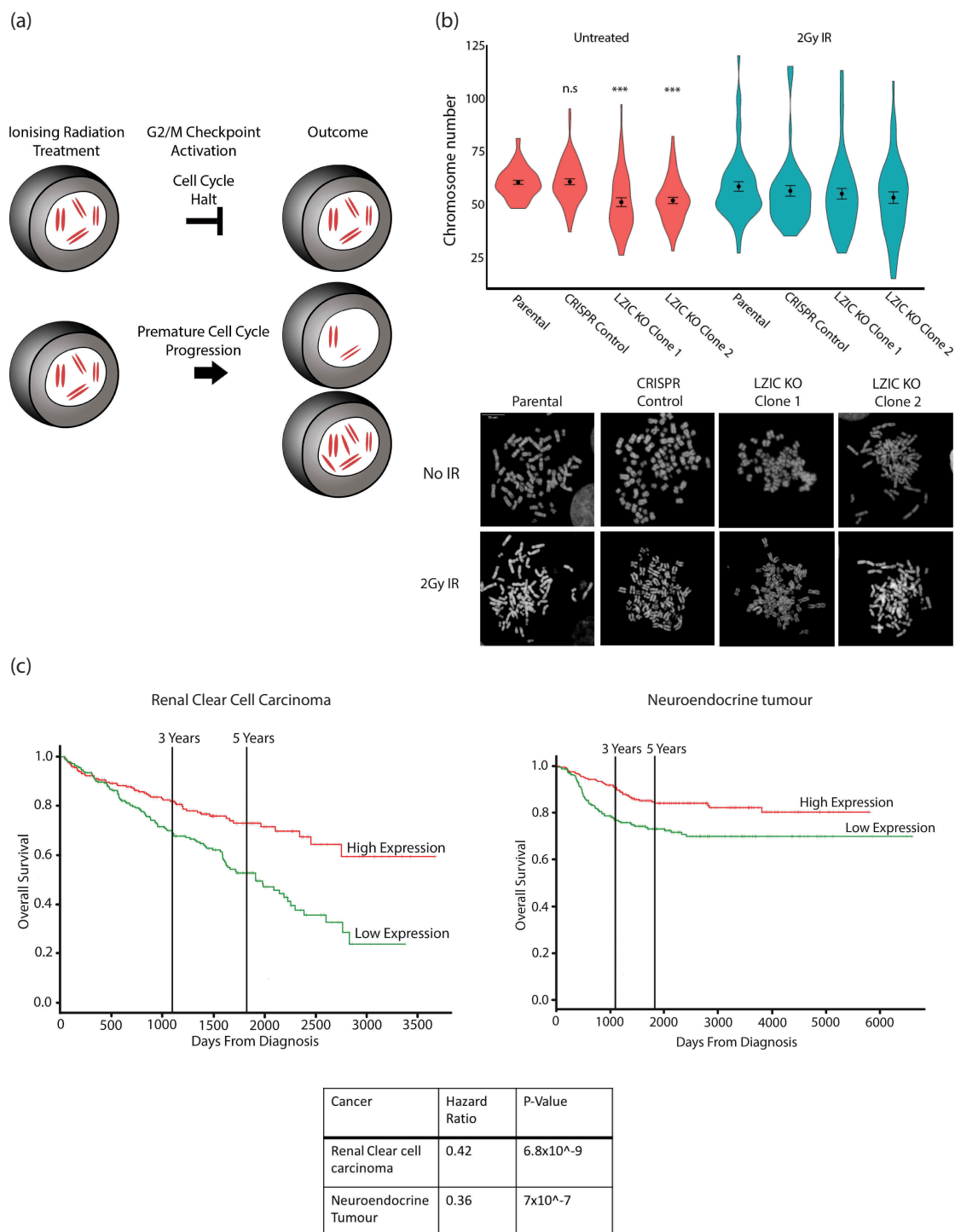


Figure 4. Loss of LZIC leads to genome instability and poor prognosis in clear cell renal carcinoma. (A) Schematic outlining the development of aneuploidy following loss of G2/M checkpoint control. (B) Metaphase spread quantification of chromosome numbers from Parental, CRISPR control line, and LZIC KO Clone 1 and 2. Data from 3 biological repeats counting at least 17 spreads per replicate. Statistical significance was determined using unpaired student T-Test, *** = p-value < 0.001, n.s = non-significant. CRISPR control and LZIC KO clones were compared to the parental line in the untreated condition. (C) Kaplan Meier plot showing overall survival of patients stratified by LZIC expression. The calculated hazard ratios and significance is also included.

determine a role for LZIC in the regulation of the MYC pathway. However, these data suggest a similar role to canonical WNT signaling components.

The analysis of pS10 H3 levels in LZIC KO cells following IR identified a reduced number of cells in late G2 and mitosis compared to control lines.

Previous literature shows that following release from the G2/M checkpoint, the mitotic population significantly increases [5]. However, the release from the G2/M checkpoint can begin as early as 12-h post-IR [5]. In this case, we hypothesize that LZIC KO cells undergo early G2/M checkpoint release prior to the 24-h time point, which causes the majority of the population to have passed through mitosis into G1. In G1 the mitotic pS10 H3 is rapidly lost, decreasing the observable population. These data are supported by the cell cycle analysis at 24 h indicating an increased G1 population (Figure 2A).

The phosphorylation of Chk1 S317 is mediated by ATR. The loss of this phosphorylation event has been shown to perturb the function of surrounding phosphorylation sites S345 and S296 [12]. Therefore, the reduced phosphorylation of this site in LZIC KO cells could have a detrimental impact on Chk1 activity. Interestingly, the altered phosphorylation status of the MPF components is downstream of both Chk1 and the protein phosphatase family [41]. We hypothesize that while expression levels of PP1 and PP2 are not altered, it is the interplay between these proteins and Chk1 which leads to the defect of checkpoint control.

The genome instability observed in LZIC KO cells is significant under basal conditions. The link between a dysfunctional G2/M checkpoint and increased genome instability has been previously shown [35]. In addition, a damage threshold must be overcome to successfully activate the spindle assembly checkpoint [42]. LZIC KO cells do not show changes to cell cycle prior to damage with IR; however, it is possible that LZIC cells possess a defect which increases the number of cells progressing through cell cycle with damage. We stipulate that LZIC cells reduce the fidelity of the G2/M checkpoint, which over time will yield the phenotypes of genome instability.

Following LZIC KO the transcriptional signatures identified showed altered regulation of multiple genes with known functions in neuronal differentiation and development. A previous study has identified LZIC as a factor required for the correct development of the zebrafish brain midline [18]. The high conservation of LZIC in zebrafish and the interaction with genes associated with the development may provide a basis for

further investigation into the regulation of these pathways, leading to a mechanism by which this process is controlled.

Collectively our data classifies LZIC as functionally involved in the IR response cascade. Clearly more mechanistic data on LZIC protein and its interacting factors are necessary to fully comprehend the contribution of this protein to mammalian DDR. However, even at this early stage, our data are suggestive of the usefulness of LZIC as a biomarker for patient stratification, given that its expression is strongly correlated with survival of patients suffering from clear cell renal carcinoma.

Materials and methods

LZIC protein evolutionary conservation analysis

National center for bioinformatics information (NCBI) nucleotide sequence database was interrogated manually and the nucleotide sequences for Human, Mouse, Xenopus, Zebrafish, Nematodes, and Slimemold were acquired. The previously identified domains were aligned, by ClustalW [43], and a percentage conservation score calculated by assessing the number of nucleotides conserved between sequences by the equation – total number of conserved nucleotides (Analyzed species)/total number of nucleotides (Humans).

Cell culture

HEK293 were cultured at 5% CO₂ in Dulbecco's-modified Eagle's medium (DMEM), supplemented with 4.5 g/l D-glucose, GlutaMAX (Life Technologies, Carlsbad, CA, USA) and 10% fetal bovine serum.

LZIC knock-out line generation

LZIC-targeting CRISPR-based knockout plasmid kit was purchased from Origene. HEK293 cells were transfected with plasmids provided in Origene kit using Lipofectamine LTX. Cells were cultured for eight passages before addition of antibiotic selection, as per manufacturer's instructions. Cells were reseeded and treated with Puromycin (0.5µg/ml) and individual colonies were selected, by the use of cloning discs. Individual colonies were expanded and

screened for LZIC expression by western blot. LZIC-Flag CDS was reintroduced into LZIC knockout (KO) Clone 2 by Lentiviral transduction. Prior to transduction LZIC KO clone 2 was transfected with Cre recombinase plasmid to remove puromycin resistance cassette from a cell line.

Microarray analysis of LZIC KO cells

All clones were plated in duplicate for both untreated and IR treated conditions. After 24 h, cells were exposed to 5 Gy IR and incubated for a further 24 h before harvesting. Untreated cells were harvested 48-h post seeding. Cells were harvested using trypsin and EDTA before RNA extraction using RNeasy kit (Qiagen) as per manufacturers instructions. Samples were subsequently labeled by low input quick amp labeling (Agilent Technologies) as per manufacturer's instructions in one-colour microarray-based gene expression analysis. The chipset reference was G4858A, GE 8 x 60K with design 039494 V3. 100 ng of RNA was used for analysis. Microarray was imaged on DNA microarray scanner with SureScan high-resolution imaging (Agilent technologies). The resulting raw data were analyzed using the R package Limma as conducted in previous studies [44,45]. Gene set enrichment analysis (GSEA) was conducted by comparing gene sets to the Molecular Signatures Database (MSigDB) [24].

qPCR analysis

HEK293 cells and LZIC KO clones were grown for 24 h prior to treatment with 5 Gy IR. Cells were harvested by trypsinization prior to extraction of RNA

using RNeasy kit (Qiagen kit). 1000 ng of extracted RNA was reverse transcribed to cDNA by Superscript II (Thermo Fisher Scientific) as per the manufacturer's instructions. The qPCR was conducted using SYBR green reagent (Applied Biosystems, Thermofisher) and plates were analyzed on Quantstudio 6 flex (Applied Biotechnologies). Delta-delta ct calculation was conducted using GAPDH as a reference gene. Primers sequences used are shown in Table 1.

Western blotting

HEK293 and CRISPR lines were seeded and 24 h following either left untreated or exposed to 5Gy IR. The cells were then harvested at the stated time points following IR. Cells extracts were generated by the addition of RIPA buffer (150mM NaCl, 50nM Tris pH 7.5, 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS, 1mM DTT, 0.4mM PMSF, Protease inhibitor cocktail) and sonication for 2 × 10 s. Samples were loaded with 1x Laemmli buffer before being heated to 90°C for 10 min. Samples were run on SDS-Page gels and transferred to nitrocellulose membrane by use of Bio-Rad Transblot Turbo. The antibodies used were: LZIC (Bethyl, 1/1000), Tubulin (Sigma Aldrich, 1/5000), pChk2 Thr68 (Cell signalling, 1/2000), Chk2 total (Bethyl, 1/2000), pChk1 S345 (Bethyl, 1/2000), pChk1 S317 (Bethyl, 1/2000), Chk1 total (Bethyl, 1/2000), pATR Tyr1981 (Cell signalling, 1/1000), ATR total (Cell signalling, 1/1000), pCyclin B1 Ser147 (Cell signalling, 1/2000), Cyclin B1 (Cell signalling, 1/2000), pCDC2 Tyr15 (Cell signalling, 1/2000), and CDC2 total (Cell signalling, 1/2000). Secondary goat antibody was horseradish peroxidase

Table 1. Primer sequences for qPCR.

Gene Name	Forward Primer (5'-3')	Reverse Primer (5'-3')
GapDH	GGAGTCAACGGATTGGTCGTA	GAATTTGCCATGGGTGGAAT
LZIC	AGTCTCTACAGACCTTGGCTC	ACAAGCTTCTGCACCATGTC
CCBN1	AACTTTCGCCTGAGCCTATTTT	TTGGTCTGACTGCTTGCTCTT
SOX11	CGGTCAAGTGCCTGTTTCTG	CACCTTGGCGACGTTGTAGC
NREP	CTGTCTTTCTAGCATGTTGCC	CCAGGGAGACCAACAGACAA
FLNA	GTCACAGTGTCATCGGAGGT	TGCACGTCACTTGCCTTTG
POU3F2	TTGTGTTGCCCTTCTTCGT	TTGCCTTCGATAAAGCGGGT
CPNE7	CACCTGGGGCAGATTGTG	TCACCGTGATGGTGGACTTG
SFN	CGCTGTTCTTGCTCAAAGG	ATGACCAGTGGTTAGGTGCG
LGALS3	GGGCCACTGATTGTCCTTA	TCACCGTGCCAGAAATTGTT
IFI30	TACGGAAACGCACAGGAACA	CAGGCCTCCACCTTGTGAA

conjugated with reactivity against mouse or rabbit (Thermo Fischer Scientific). All further antibodies were analyzed using LICOR system using goat-secondary with conjugated fluorescence. PP2A – subunit A (Cell signalling, 1/1000), PP2A – subunit B (Cell signalling, 1/1000), PP2A – subunit C (Cell signalling, 1/1000), PP1 (Santa Cruz, 1/1000), pATM Serine 1981 (Cell signalling, 1/1000), ATM total (Cell signalling, 1/1000), Vinculin (Abcam, 1/5000), p53 total (Santa Cruz, 1/1000), p-p53 Serine 15 (Cell signalling, 1/1000).

Cell cycle analysis

Twenty-four hours after seeding HEK293 cells and CRISPR clones were treated with IR (5 Gy), camptothecin (20 μ M), cobalt chloride (200 μ M), or UV (20 mJ) and incubated 24 h. Cells were incubated for further 8 and 24 h and then harvested. After washing with PBS, ice-cold 70% ethanol was slowly added under slight agitation. Cells were left at 4°C for 24 h to fix, PBS washed, and Propidium Iodide and RNase A were added to final concentration of 10 μ g/ml and 100 μ g/ml, respectively. Samples were heated to 37°C for 30 min and then incubated at 4°C for at least 4 h before reading. Flow cytometry analysis of cells was conducted on a BD biosciences FACS canto.

Early G2/M checkpoint activation

This method was conducted as shown in Xu, et.al. 2002 [5]. One set of control cells were additionally treated with ATMi (10 μ M final concentration, Sigma Aldrich) 1 h prior to exposure to 5 Gy IR. Cells were stained with pS10 H3 antibody (Cell Signalling, 1/100) and incubated with Goat-anti-rabbit 488 (Abcam, 1/500). The cells were then analyzed on Attune NXT (Life Technologies).

Immunofluorescence

Parental HEK293 and CRISPR clones were seeded and treated with 2 Gy IR. The cells were incubated for 24 h before supernatant was removed and cells washed with PBS. Four percent Paraformaldehyde was used to fix cells for 10 min at room temperature before treatment with blocking buffer (0.3% triton X-100 in PBS supplemented with 5% goat serum). Fixed cells were treated with primary antibody overnight at 4°C. Cells were washed 3x with

PBS before the addition of secondary antibody and incubated at room temperature for 1 h. Cells were mounted with hard set mounting medium (vector hard set mounting medium, Vector labs). Images were acquired using a Zeiss LSM 510 and images were processed with ZEN 2009 software. Primary antibody – Phospho-serine 10 Histone 3 antibody (Cell Signalling Technology, 1/1000). Secondary antibody was goat anti-rabbit with conjugated Cy5 (Thermo Fischer Scientific).

Metaphase spread analysis

Parental HEK293 and CRISPR clones were seeded and treated with 2 Gy IR before incubating for 48 h. Cells were harvested by trypsinization and centrifuged at 300 g 5 min before swelling buffer (75 mM KCl) was added. The cell pellet was incubated for 10 min at room temperature before addition of a fixative solution (Methanol and acetic acid 3:1 ratio). The Cells were centrifuged at 200 g for 5 min and the supernatant was removed. This step was repeated twice. Pellet was suspended in fixative to give cell suspension and dropped from a height of 30 cm onto slides (Superfrost plus, Thermo scientific). Slides were dried at room temperature for 2 min before steaming for 10 s. Slides were left in a humidity box overnight to dry. Cells were stained with Dapi (1/5000) diluted in PBS and then mounted. Images were acquired using a Zeiss LSM 510. Manual counts were conducted of spreads to determine chromosomal numbers.

Cell viability assay

Cells were treated with 0, 40, 60, or 80 Gy IR before a 24-h incubation. The WST-1 reagent (Sigma Aldrich) was used and data analyzed as per manufacturers instructions. With the following deviations, the WST-1 reagent was added 2 h prior to absorbance quantification. With the absorbance being read by Powerwave XS2 plate reader (BioTek).

Kaplan-meier plot generation

The PROGene V2 database was used to generate Kaplan-Meier plots for LZIC expression in cancers [36]. The overall survival of patients was analyzed with no stratification apart from LZIC expression.

Acknowledgments

This work was supported by the Medical Research Council (MRC) UK.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

George Skalka  <http://orcid.org/0000-0001-5117-2599>
 Holly Hall  <http://orcid.org/0000-0003-2779-2565>
 Michal Malewicz  <http://orcid.org/0000-0002-3872-3402>

References

- [1] Raschella G, Melino G, Malewicz M. New factors in mammalian DNA repair-the chromatin connection. *Oncogene*. 2017;36(33):4673–4681.
- [2] Polo SE, Jackson SP. Dynamics of DNA damage response proteins at DNA breaks: a focus on protein modifications. *Genes Dev*. 2011;25(5):409–433.
- [3] Lovejoy CA, Cortez D. Common mechanisms of PIKK regulation. *DNA Repair (Amst)*. 2009;8(9):1004–1008.
- [4] Dasika GK, Lin SC, Zhao S, et al. DNA damage-induced cell cycle checkpoints and DNA strand break repair in development and tumorigenesis. *Oncogene*. 1999;18(55):7883–7899.
- [5] Xu B, Kim S-T, Lim D-S, et al. Two molecularly distinct G(2)/M checkpoints are induced by ionizing irradiation. *Mol Cell Biol*. 2002;22(4):1049–1059.
- [6] Kastan MB, Bartek J. Cell-cycle checkpoints and cancer. *Nature*. 2004;432(7015):316–323.
- [7] Abbas T, Dutta A. p21 in cancer: intricate networks and multiple activities. *Nat Rev Cancer*. 2009;9(6):400–414.
- [8] Muller PA, Vousden KH. p53 mutations in cancer. *Nat Cell Biol*. 2013;15(1):2–8.
- [9] van Harn T, Foijer F, van Vugt M, et al. Loss of Rb proteins causes genomic instability in the absence of mitogenic signaling. *Genes Dev*. 2010;24(13):1377–1388.
- [10] Abraham RT. Cell cycle checkpoint signaling through the ATM and ATR kinases. *Genes Dev*. 2001;15(17):2177–2196.
- [11] Liu Q, Guntuku S, Cui XS, et al. Chk1 is an essential kinase that is regulated by Atr and required for the G (2)/M DNA damage checkpoint. *Genes Dev*. 2000;14(12):1448–1459.
- [12] Wilsker D, Petermann E, Helleday T, et al. Essential function of Chk1 can be uncoupled from DNA damage checkpoint and replication control. *Proc Natl Acad Sci U S A*. 2008;105(52):20752–20757.
- [13] O'Connell MJ, Raleigh JM, Verkade HM, et al. Chk1 is a wee1 kinase in the G2 DNA damage checkpoint inhibiting cdc2 by Y15 phosphorylation. *Embo J*. 1997;16(3):545–554.
- [14] Leung-Pineda V, Ryan CE, Piwnica-Worms H. Phosphorylation of Chk1 by ATR is antagonized by a Chk1-regulated protein phosphatase 2A circuit. *Mol Cell Biol*. 2006;26(20):7529–7538.
- [15] Den Elzen NR, O'Connell MJ. Recovery from DNA damage checkpoint arrest by PP1-mediated inhibition of Chk1. *Embo J*. 2004;23(4):908–918.
- [16] Sanchez-Vega F, Mina M, Armenia J, et al. Oncogenic signaling pathways in the cancer genome atlas. *Cell*. 2018;173(2):321–337 e10.
- [17] Katoh M. Molecular cloning and characterization of LZIC, a novel gene encoding ICAT homologous protein with leucine zipper domain. *Int J Mol Med*. 2001;8(6):611–615.
- [18] Clements WK, Kimelman D. LZIC regulates neuronal survival during zebrafish development. *Dev Biol*. 2005;283(2):322–334.
- [19] Tamimi Y, Ziebart K, Desaulniers N, et al. Identification of a minimal region of loss on the short arm of chromosome 1 in Wilms tumor. *Genes Chromosomes Cancer*. 2007;46(4):327–335.
- [20] Fransson S, Martinsson T, Ejeskär K. Neuroblastoma tumors with favorable and unfavorable outcomes: significant differences in mRNA expression of genes mapped at 1p36.2. *Genes Chromosomes Cancer*. 2007;46(1):45–52.
- [21] Daino K, Ugolin N, Altmeyer-Morel S, et al. Gene expression profiling of alpha-radiation-induced rat osteosarcomas: identification of dysregulated genes involved in radiation-induced tumorigenesis of bone. *Int J Cancer*. 2009;125(3):612–620.
- [22] Wang J-F, Liu C, Zhang Q, et al. Research progress in the radioprotective effect of the canonical Wnt pathway. *Cancer Biol Med*. 2013;10(2):61–71.
- [23] Cadigan KM, Waterman ML. TCF/LEFs and Wnt signaling in the nucleus. *Cold Spring Harb Perspect Biol*. 2012;4:11.
- [24] Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545–15550.
- [25] Liberzon A, Birger C, Thorvaldsdóttir H, et al. The molecular signatures database (MSigDB) hallmark gene set collection. *Cell Syst*. 2015;1(6):417–425.
- [26] Cui F, Hou J, Huang C, et al. C-Myc regulates radiation-induced G2/M cell cycle arrest and cell death in human cervical cancer cells. *J Obstet Gynaecol Res*. 2017;43(4):729–735.
- [27] Siliciano JD, Canman CE, Taya Y, et al. DNA damage induces phosphorylation of the amino terminus of p53. *Genes Dev*. 1997;11(24):3471–3481.
- [28] Pérez-Cadahía B, Drohic B, Davie JR. H3 phosphorylation: dual role in mitosis and interphase. *Biochem Cell Biol*. 2009;87(5):695–709.

- [29] Nam EA, Zhao R, Glick GG, et al. Thr-1989 phosphorylation is a marker of active ataxia telangiectasia-mutated and Rad3-related (ATR) kinase. *J Biol Chem.* **2011**;286(33):28707–28714.
- [30] So S, Davis AJ, Chen DJ. Autophosphorylation at serine 1981 stabilizes ATM at DNA damage sites. *J Cell Biol.* **2009**;187(7):977–990.
- [31] Bartek J, Lukas J. Chk1 and Chk2 kinases in checkpoint control and cancer. *Cancer Cell.* **2003**;3(5):421–429.
- [32] Dorée M, Hunt T. From Cdc2 to Cdk1: when did the cell cycle kinase join its cyclin partner? *J Cell Sci.* **2002**;115(Pt 12):2461–2464.
- [33] Niida H, Tsuge S, Katsuno Y, et al. Depletion of Chk1 leads to premature activation of Cdc2-cyclin B and mitotic catastrophe. *J Biol Chem.* **2005**;280(47):39246–39252.
- [34] Gavet O, Pines J. Activation of cyclin B1-Cdk1 synchronizes events in the nucleus and the cytoplasm at mitosis. *J Cell Biol.* **2010**;189(2):247–259.
- [35] Lobrich M, Jeggo PA. The impact of a negligent G2/M checkpoint on genomic instability and cancer induction. *Nat Rev Cancer.* **2007**;7(11):861–869.
- [36] Goswami CP, Nakshatri H. PROGeneV2: enhancements on the existing database. *BMC Cancer.* **2014**;14:970.
- [37] Niehrs C, Acebron SP. Mitotic and mitogenic Wnt signalling. *Embo J.* **2012**;31(12):2705–2713.
- [38] Jun S, Jung Y-S, Suh HN, et al. LIG4 mediates Wnt signalling-induced radioresistance. *Nat Commun.* **2016**;7:10994.
- [39] Zhang S, Li Y, Wu Y, et al. Wnt/beta-catenin signaling pathway upregulates c-Myc expression to promote cell proliferation of P19 teratocarcinoma cells. *Anat Rec (Hoboken).* **2012**;295(12):2104–2113.
- [40] Rennoll S, Yochum G. Regulation of MYC gene expression by aberrant Wnt/beta-catenin signaling in colorectal cancer. *World J Biol Chem.* **2015**;6(4):290–300.
- [41] Calonge TM, O'Connell MJ. Turning off the G2 DNA damage checkpoint. *DNA Repair (Amst).* **2008**;7(2):136–140.
- [42] Rieder CL, Maiato H. Stuck in division or passing through: what happens when cells cannot satisfy the spindle assembly checkpoint. *Dev Cell.* **2004**;7(5):637–651.
- [43] Larkin MA, Blackshields G, Brown NP, et al. Clustal W and Clustal X version 2.0. *Bioinformatics.* **2007**;23(21):2947–2948.
- [44] Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **2015**;43(7):e47.
- [45] Phipson B, Lee S, Majewski IJ, et al. Robust hyperparameter estimation protects against hypervariable genes and improves power to detect differential expression. *Ann Appl Stat.* **2016**;10(2):946–963.