

Characterisation of the RNA processing factor
CIP29 (HCC-1/SARNP) and its involvement in cell
proliferation and genome stability

GISELLE EDGE (BSC HONS) (MSC)

**Supervisors: Dr Elaine Taylor and Dr Howard Lindsay (Lancaster
University)**

LANCASTER MEDICAL SCHOOL

Doctor of Philosophy

Thesis

Acknowledgements

This body of work and the time spent to complete it is dedicated to my family, especially my mother who has always believed in me and supported my endeavours.

"Happiness can be found,
even in the darkest of times,
if one only remembers to turn on the light"

- Prof. Albus Dumbledore,
Harry Potter and the Prisoner of Azkaban

With thanks,

To Dr Elaine Taylor and Dr Howard Lindsay

Abbreviations

AA	Animo acid
ALY	ALYREF - THO complex subunit 4
APC	Anaphase-promoting complex
ATM	Ataxia Telangiectasia Mutated
ATR	Ataxia Telangiectasia and Rad 3 related protein
BAX	BCL-2-associated X protein
BRCA1	Breast cancer type 1 susceptibility protein
BRCA2	breast cancer type 2 susceptibility protein
Cas9	CRISPR associated protein 9
CBC	Cap binding complex
CDK	Cyclin-dependent kinase
CHK1	Checkpoint kinase 1
CHK2	Checkpoint kinase 2
CHTOP	Chromatin target of PRMT1 protein
CIN	Chromosome instability
CIP29	Cytokine Induced Protein 29
CLL	Chronic myelogenous leukemia
CP	Cell cycle checkpoints
CPC	Chromosome passenger complex
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DDK	Dbf4-dependent kinase
DDR	DNA damage response
DDR	DNA damage responses
DNA	Deoxyribonucleic acid
DNA-PK	DNA-dependent protein kinase
DNA-PKcs	DNA dependent protein kinase catalytic subunit
DSB	Double Strand Break
DSB	double strand breaks
dsDNA	Double-stranded DNA
EJC	Exon junction complex
EPO	Erythropoeitin
ESCRT-III	Endosomal sorting complex required for transport-III
FACS	Fluorescence-activated cell sorting
FUS	Fused in Sarcoma

G1	Gap phase 1
G2	Gap phase 2
H2AX	H2A (Histone Family Member X)
HCC-1	Hepatocellular carcinoma-1
HNPCC	Hereditary non-polyposis colorectal cancer
hnRNPs	Nuclear ribonucleoproteins
HR	Homologous recombination
Hsp	Heat shock proteins
LS	Lynch syndrome
LUZP4	Leucine Zipper Protein 4
M	Mitosis
MCC	Mitotic checkpoint complex
MDM2	Mouse double minute 2 homolog
MDS	Myelodysplastic syndromes
MMR	Mismatch Repair
MRN	Mre11–Rad50–Nbs1
mRNA	Messenger Ribonucleic acid
mRNP	Messenger ribonucleoprotein
MT	Microtubules
NER	Nucleotide excision repair
NHEJ	Non-Homologous End Joining
NMD	Nonsense-mediated mRNA decay
NPC	Nuclear pore complex
ORC	Origin recognition complex
PARP1	Poly (ADP-Ribose) Polymerase I
PI3KK	Phosphatidylinositol-3-OH kinase-related kinase
PIKKs	phosphatidylinositol-3-kinase related kinases
POLDIP3	Polymerase delta-interacting protein 3
pre-RC	pre-replication complexes
PTCs	Premature termination codons
RhoA	Ras homolog gene family, member A
RNA	Ribonucleic acid
RNAPII	RNA Polymerase II
S	DNA synthesis phase
SAC	Spindle assembly checkpoint
SAP	Scaffold attachment factor A/B ACINUS and PIAS
SARNP	SAP domain-containing ribonucleoprotein

Ser	Serine
SiRNA	Small Interfering Ribonucleoprotein
SSB	Single strand break
ssDNA	Single stranded DNA
TP53 /p53	Tumour suppressor P53
TREX	TRanscription and EXport
UAP56	Also known as DDX39B, BAT1 DEAD-box Helicase 39B
UBM	UAP56-binding motif
UIF	UAP56-interacting factor
URH49	Also known as DDX39A, DEAD-box Helicase 39A
UTRs	Untranslated regions
VEGF	Hyperactivating vascular endothelial growth factor
XCIP29	<i>X.laevis</i> Cytokine Induced Protein 29
XP	Xeroderma Pigmentosum
ZC3H11A	Zinc finger CCCH domain-containing protein 11A

Abstract	13
1 Introduction	14
1.1 Hallmarks of Cancer	14
1.2 Genome Instability and Cancer.....	17
1.2.1 DNA Damage and its repair.....	17
1.2.1.1 Base excision repair (BER)	17
1.2.1.2 Nucleotide excision repair (NER)	17
1.2.1.3 Mismatch repair (MMR).....	18
1.2.1.4 Recombination repair (HR) and non-homologous end joining (NHEJ)	18
1.3 The cell cycle and its key events	19
1.3.1 DNA replication	19
1.3.2 Mitosis.....	21
1.3.3 The Centrosome cycle.....	23
1.4 Cell Cycle Regulation.....	24
1.4.1 Cyclins and cyclin-dependent kinases (CDKs)	24
1.4.2 Cell cycle checkpoints and the DNA damage response	25
1.4.3 Mitotic checkpoints.....	28
1.4.4 Checkpoint dysregulation and chromosome instability (CIN)	31
1.5 Regulation of gene expression: mRNA processing and export.....	31
1.5.1 mRNA processing	32
1.5.2 mRNA export and The TRanscription and EXport (TREX) complex.....	33
1.6 The Influence of the TREX complex on Genome Instability.....	34
1.6.1 CIP29	37
1.7 Work leading up to this study.....	39
1.8 Aims and Objectives.....	41
1.8.1 Project aims.....	41
1.8.2 Project objectives.....	41
2 Materials and methods.....	43
2.1 Materials	43
2.2 Molecular biology methods	45
2.2.1 Restriction Endonuclease Digestion.....	45
2.2.2 Gel Electrophoresis	45
2.2.3 QIAquick Gel Extraction of DNA fragment	45
2.2.4 Purification of PCR products using the QIAquick® PCR Purification Kit.....	45

2.2.5	Ligation into Pepex, or PCI-NEO expression vectors.....	45
2.2.6	Polymerase chain reaction.....	46
2.2.7	A-tailing of PCR products	49
2.2.8	Ligation into pGEM®-T Easy vector	49
2.2.9	Transformation of plasmids into ultra-competent E. coli cells.....	49
2.2.10	Alkaline lysis miniprep method.....	49
2.2.11	QIAGEN® Midiprep method	49
2.2.12	DNA sequencing.....	49
2.3	Protein Methods:.....	50
2.3.1	Preparation of whole cell extracts	50
2.3.2	Preparation of HeLa nuclear cell extracts	50
2.3.3	Direct cell lysis for monitoring DNA damage response (DDR) activation	50
2.3.4	Cell fractionation assays	51
2.3.5	Bradford assay.....	51
2.3.6	SDS-PAGE gel production.....	51
2.3.7	Western blot semi-dry transfer	53
2.3.8	Antibody Staining of nitrocellulose membrane	53
2.3.9	Developing Western Blots by Enhanced Chemiluminescence (ECL).....	53
2.3.10	Stripping Western Blots	53
2.3.11	In vitro translation (IVT)	54
2.3.12	Affinity column purification of antibodies	54
2.3.13	Crosslinking antibodies to protein A Sepharose beads.....	55
2.3.14	Crosslinking antibodies to magnetic protein A/G Dynabeads	55
2.4	Immunoprecipitation	55
2.4.1	Standard Immunoprecipitation (IP) protocol.....	55
2.4.2	Immunoprecipitation protocol adapted from Dufu et al (2010)	56
2.4.3	Immunoprecipitation protocol adapted from Fujita et al., (2020)	56
2.5	Cell Culture Techniques	57
2.5.1	Cell culture conditions	57
2.5.2	Freezing and thawing cell stocks.....	57
2.5.3	Media supplements and additions.....	57
2.5.4	Transient transfection using FuGENE® HD Transfection.....	57
2.5.5	Stable transfection using FuGENE® HD Transfection.....	58
2.5.6	Isolation of stable clones	58

2.5.7	Cell proliferation assays	58
2.6	Flow cytometry methods	59
2.6.1	Ethanol fixation	59
2.6.2	Paraformaldehyde (PFA) fixation	59
2.6.3	Antibody staining	59
2.6.4	PI labelling incubation	59
2.7	Immunofluorescence analysis.....	61
2.7.1	Coverslip preparation and cell seeding.....	61
2.7.2	Immunofluorescence fixation and blocking.....	61
2.7.3	Immunofluorescence primary and secondary antibody incubation.....	61
2.7.4	Immunofluorescence secondary antibody incubation	61
2.7.5	Immunofluorescence mounting and imaging.....	62
2.8	Computational and Statistical analysis methods	63
2.8.1	Numerical counts of nuclear size	63
2.8.2	Numerical counts of Centrosome number	63
2.8.3	Statistical analysis and graph development.....	63
2.8.4	Structural analysis through Alpha fold.....	63
3	Chapter 3: Analysis of CIP29 Protein-Protein Interactions	64
3.1	Introduction	64
3.1.1	The TRanscription and EXport (TREX) Complex	64
3.1.2	CIP29 and novel interacting proteins.....	65
3.1.3	Project objective: Validation of CIP29 protein-protein interactions	66
3.2	Results.....	66
3.2.1	Development of CIP29 antibody reagents for detection and immunoprecipitation of CIP29	66
3.2.2	CIP29 does not associate with TREX complex components using the standard immunoprecipitation protocol.....	69
3.2.3	Does CIP29 associate with TREX/AREX components in the presence of ATP?	71
3.2.4	Validation of potential CIP29 interactors identified through mass spectrometry.	77
3.2.5	Analysis of CIP29-anillin interaction	78
3.2.6	Analysis of CIP29-KIF2C interaction	80
3.2.7	Discussion.....	86

3.2.8	The results suggest that CIP29 is indeed a TREX complex component whereby the association relies on ATP, however the α -CIP29 antibody used in this study may disrupt the association.....	86
3.2.9	Attempted validation of novel CIP29 interacting proteins	87
4	Chapter 4: Exploring the phenotype caused by loss of CIP29.....	89
4.1	Introduction	89
4.1.1	Project objective: Exploring the phenotype caused by loss of CIP29.	89
4.2	Results.....	90
4.2.1	Validation of CIP29 knockout and CIP29 rescue cell lines.	90
4.2.2	Investigating the role of CIP29 in cell proliferation	92
4.2.3	Investigating the role of CIP29 in cell cycle progression.....	94
4.2.4	Does the absence of CIP29 induce G2/M checkpoint activation?	97
4.2.5	Does the absence of CIP29 induce a mitotic delay?	99
4.2.6	Does the absence of CIP29 induce cytokinesis defect?	101
4.2.7	Do the CIP29-KO cells exhibit larger nuclei?	104
4.2.8	CDT1 staining and Flow Cytometry.....	111
4.2.9	Immunofluorescence analysis to investigate CDT1 expression within the CIP29-KO population.....	114
4.3	Discussion.....	116
5	Chapter 5: exploring the Structure-function analysis of CIP29.....	119
5.1	Introduction	119
5.1.1	Project objective: Obtain a greater understanding of the molecular structure of the CIP29 protein and its functional domains.	119
5.2	Results.....	120
5.2.1	Generation of CIP29 truncation mutants.....	120
5.2.2	In vitro expression analysis of CIP29 truncation mutants	125
5.2.3	In vivo expression analysis of CIP29 truncation mutants	126
5.2.4	Investigating the cell cycle profiles of the CIP29-truncations using flow cytometry analysis	128
5.2.5	Do the CIP29- truncations exhibit morphological defects like the CIP29-KOs?.....	130
5.2.5.1	Morphological defects: centrosome supernumerary	130
5.2.5.2	Do the CIP29-Truncations exhibit nuclei malformations?	133
5.2.6	Investigating the localisation of the CIP29 truncations	139
5.3	Discussion.....	143

6	Chapter 6: Discussion.....	145
6.1	Discussion.....	145
6.1.1	Chapter 1: Investigating CIP29 protein-protein interactions.....	145
6.1.2	Chapter 2: Explore the phenotype caused by loss of CIP29.	148
6.1.3	Chapter 3: Exploring the molecular structure of the CIP29 protein and its functional domains.152	
6.1.4	Summary and future perspectives.....	155
7	References.....	158

List of Figures

Figure 1: The Hallmarks of cancer.....	15
Figure 2: Common causes of DNA damage and their corresponding repair pathways.....	18
Figure 3: DNA Replication: Leading and Lagging Strand Synthesis.....	20
Figure 4: Phases of cell division (Zitouni et al., 2014).....	21
Figure 5: The Cell Cycle and Its Checkpoints.....	24
Figure 6: The DNA Damage response pathway	26
Figure 7: Mitotic Checkpoints: the spindle assembly checkpoint (SAC).....	29
Figure 8: Mitosis Checkpoints: Abcission Checkpoint	30
Figure 9: RNA Processing and the TREX Complex.....	34
Figure 10: Previous data: CIP29 is required for successful cell cycle progression and is phosphorylated during mitosis and in response to DNA damage.	40
Figure 11: Affinity purified CIP29 antibody specifically recognises and immunoprecipitates the CIP29 protein.....	68
Figure 12: TREX/AREX components UAP56/URH49 and ALY do not co-immunoprecipitate with CIP29	70
Figure 13: TREX/AREX components UAP56/URH49 and ALY do not co-immunoprecipitate with CIP29 regardless of ATP addition	72
Figure 14: Using Fujita et al., (2020) methodology CIP29 does Co-immunoprecipitate with TREX components UAP56/URH49 or ALY, however reproducibility is poor.....	74
Figure 15: CIP29 does co-immuoprecipitate with TREX/AREX components.	76

Figure 16: Various proteins that were identified as potential interactors through mass-spectrometry do not co-immunoprecipitate with CIP29 within MRC5-V1 cell extracts.	78
Figure 17: anillin, a potential interactor identified through mass-spectrometry does not reliably Co-immunoprecipitate with CIP29.	79
Figure 18: KIF2C, a potential interactor identified through mass-spectrometry does reliably Co-immunoprecipitate with CIP29.	80
Figure 19: KIF2C did not immunoprecipitate with CIP29 under <i>invitro</i> -translation conditions	82
Figure 20: Binding between CIP29 and KIF2C is absent within α -FLAG immunoprecipitation assays. 83	
Figure 21: KIF2C associations with CIP29 may be disingenuous.....	85
Figure 22: Validation of CRISPR-Cas9-induced Functional Knockout and Rescue of the CIP29 protein through western blot and protein localisation studies.	91
Figure 23: Measurement of the cell proliferation for MRC5-VI, CIP29-KNOCKOUT (CIP29-KO), CIP29-RESCUE (W6) and CIP29-RESCUE-FLAG (F17) using WST-1 assay.....	93
Figure 24: Flow Cytometric analysis through propidium iodide, (PI) staining of the wildtype (MRC5-VI), CIP29 knockout cells (CIP29-KO) and CIP29-RESCUES (W6 and F17).	96
Figure 25: Analysis of the DNA damage response of each cell type through the activation of DNA damage markers	98
Figure 26: Flow Cytometric analysis of each cell type investigating the expression of a key mitotic marker.....	100
Figure 27A: Immunofluorescence analysis of MRC5, CIP29-knockout (CIP29-KO), CIP29-Rescue (W6) CIP29-Rescue-flag (F17) cell lines.	102
Figure 28 CIP29-KO cells maintain a distinct population distribution which is not observed within their counterparts.....	105
Figure 29: Box plots detailing the nuclei area of each cell type.	107
Figure 30: CIP29-KO cells may exhibit centrosome amplification.	109
Figure 31: CIP29-KNOCKOUT cells do exhibit an increase in CDT1 Suggesting they are arrested or are slowly progressing through G1.	113
Figure 32: Immunofluorescence analysis suggests that CIP29-KO cells exhibit abnormal CDT1 expression.	115
Figure 33: Schematic representation of the CIP29 truncations.....	121
Figure 34: Development of CIP29 truncations using alpha fold protein structure data base (3 of 3).	124
Figure 35 Investigating the protein expression of each CIP29 truncation.....	127

Figure 36 Some stably transfected CIP29- truncations exhibit abnormal cell cycle profiles indicative of cell cycle defects.	129
Figure 37: Investigating if the CIP29 truncations exhibit centrosome amplification.	131
Figure 38: Examples of centrosome amplification through immunofluorescence staining 2/2. Immunofluorescence staining using DAPI (grey) γ -tubulin (red) and merge to identify the occurrence of centrosome amplification (yellow arrows) within each cell type (W6 (CIP29 rescue), CIP29-FL (CIP29- FLAG rescue), KO (CIP29-KNOCKOUT), and Truncations (B, C, D and Δ SAP). Images are scaled to 10 μ m. 2/2 (R=3).....	133
Figure 39: Figure 39: Box plots show the nuclei area (μ m ²) for each cell type with individual readings overlaid.	135
Figure 40: Density plots showing the distribution of nuclei area for each cell type.	137
Figure 41: Assessing the functionality of each CIP29-truncation through localisation studies.....	140
Figure 42: CIP29 schematic representation of the CIP29 truncations with regions of interest and potential regulation sites highlighted.	142

List of tables

Table 1: Primary antibodies used in this work.....	43
Table 2: Secondary antibodies used in this work.	44
Table 3: Primers designed for this work.	44
Table 4: PCR parameters and program details for generating HA-KIF2C, FLAG-KIF2C and HA-CIP29 DNA fragments.....	46
Table 5: Generating CIP29 truncation mutants through PCR: 3' truncations.....	47
Table 6: Generating CIP29 truncation mutants through PCR: SAP domain deletion (Δ SAP).	48
Table 7: Composition of resolving and stacking gels used for SDS-PAGE Polyacrylamide gel electrophoresis.	52
Table 8: DNA constructs investigated for protein synthesis using the In vitro translation (IVT).....	54
Table 9: Antibody staining conditions and corresponding fixative.....	60
Table 10: Antibody staining conditions and corresponding fixative.....	62
Table 11: CIP29- Truncations B and C exhibit cell populations that maintain a wide variation of nucleus sizes (μ m).....	134
Table 12: Testing the significance of the Nuclear Areas between the cell types using the Wilcox non-parametric test.....	138

Abstract

Cancer has become a leading cause of death worldwide, originating when cells incur genetic damage and altered gene expression that leads to dysregulated proliferation and other cancer hallmarks. Mechanisms that help to maintain genome integrity and the correct expression of this genetic information, such as effective DNA repair, accurate chromosome segregation and regulated gene expression are therefore fundamentally important in curbing the development and progression of cancer. Correct gene expression relies on both transcriptional regulation and correct processing, export and translation of mRNA transcripts, the mediators between the genetic instructions and protein synthesis. The journey an mRNA transcript undergoes during RNA processing can affect its stability and behaviour which in turn, can profoundly influence gene expression and cellular behaviour.

An RNA processing factor of interest CIP29, is often upregulated in cancer cells and may aid cancer cell development and growth. Although implicated in RNA processing and export, the precise function of CIP29 has not been fully characterized. To gain a greater understanding of the activities of CIP29 we attempted to identify and validate CIP29-interacting factors and to phenotypically characterise cells in which CIP29 had been abolished through CRISPR-Cas9 gene editing. Upon examination, only proteins that had been previously implemented within mRNA export were shown to reliably associate with the CIP29 protein, suggesting that the influence of CIP29 within the cell is achieved through its role as an mRNA exporter. CIP29 knockout cells showed no evidence of genetic damage but displayed a profound defect in cell cycle progression. We have characterised this cell proliferation problem as relating to defective chromosome segregation and cell division. These defects were not observed once the CIP29 protein was reintroduced through CIP29-rescue cell lines, suggesting that the mitotic defects observed were because of the loss of CIP29. To further obtain an understanding of the functional domains and regions of interest within the CIP29 protein, its molecular structure was examined through the development of truncated mutants which were then subjected to phenotypic analysis. The results obtained suggested that the C-terminal end of the protein was significant within the mitotic defects observed, as truncations without this region displayed a phenotype that was like or exceeded the CIP29-knockout. Overall, the results obtained suggest that the CIP29 protein is fundamental in the successful completion of the cell cycle, which may be due to its role in the processing and export of mRNA transcripts for key mitotic proteins.

1 [Introduction](#)

1.1 **Hallmarks of Cancer**

Cancer has become a global health burden which, despite decades of research into its detection and treatment, is associated with a persistently high level of mortality, with approximately 1 in 6 deaths worldwide attributed to cancer (Weir et al., 2016). In the UK alone, 1 in 2 people are diagnosed with cancer, with 165,000 deaths per year, or 450 deaths per day (Smittenaar et al., 2016). The term 'cancer' does not describe a single disease but a collection of over 200 diseases. Each arise from different cell types with unique characteristics and genetics, that no longer adhere to normal regulatory control systems. This diversity further complicates drug development, making it a slow and challenging process (Benz, 2017).

Despite the diversity, some fundamental features or hallmarks of cancer are commonly observed. In an attempt to rationalise the highly complex nature of tumour pathogenesis, these hallmarks were detailed as progressive characteristics that upon acquisition, cells would have the potential to become tumorigenic (Hanahan and Weinberg, 2000). Originally, six biological hallmarks of cancer were described: sustained cell proliferation, evasion of growth suppressors, replicative immortality, resistance to cell death, angiogenesis, tissue invasion and metastasis (Figure 1) (Hanahan and Weinberg, 2000).

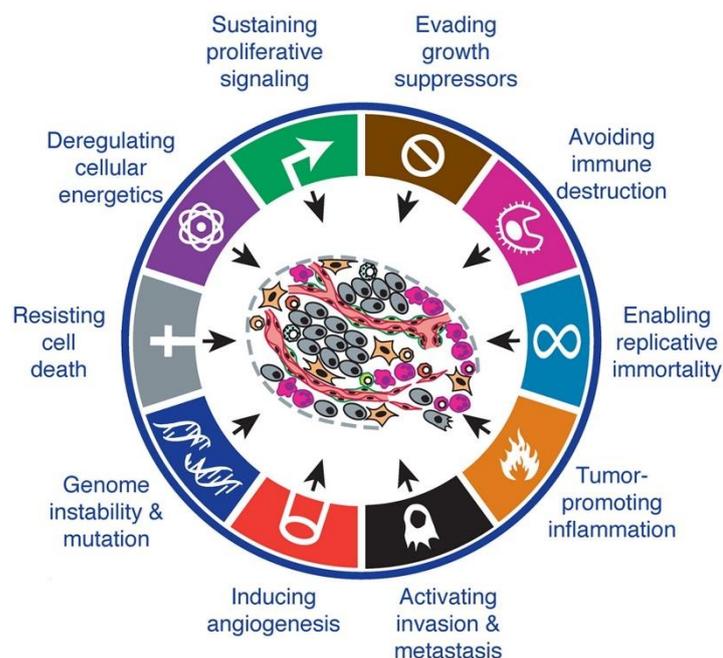


Figure 1: The Hallmarks of cancer

The collection of cancer hallmarks proposed by Hanahan and Weinberg in their 2000 and 2011 publications (adapted with permission from (Hanahan and Weinberg, 2011)).

Sustained proliferation is one of the most well recognised traits of cancer cells and manifests through the manipulation of a cells' own autocrine signalling pathways, or through the paracrine and endocrine pathways of nearby healthy cells. Such manipulation can induce the hyper-production and release of growth ligands or cytosolic signalling molecules, which enable a continual source of proliferation stimuli (Feitelson et al., 2015). To continue malignant development and achieve an **unlimited replicative potential**, cancer cells must overcome the functional decline and limited lifespan that is normally restrained by the progressive shortening of a chromosome's telomeres over repeated cell divisions. To achieve this, cancer cells often prolong the enzymatic activity of telomerase, an enzyme that synthesizes the telomeric DNA, by removing the transcriptional restrains associated with the telomerase gene (Bell et al., 2015). By extending telomeric DNA, telomere erosion can be counteracted to prolong a cells' replicative lifetime (Harley, 2008). For cellular growth to persist, malignant cells must **resist the action of growth suppressors**. It is estimated that 70% of all genetic changes within solid tumours are linked to tumour suppression evasion (Amin et al., 2015). For example, the common loss of the cyclin-dependent kinase inhibitor (*CDKN2A*) gene and its alternative open reading frame *p14ARF* can impact two distinct tumour

suppressors: p16INK4A and p14ARF. Both are key cell cycle regulators. p16INK4A binds to CDK4 and CDK6, to prevent the phosphorylation of Rb family members and promote cell cycle arrest in G1, while p14ARF binds to MDM2 to prevent MDM2-mediated degradation of p53 and causes cell cycle arrest within G1 and G2 (Brown et al., 2004). The loss of these tumour suppressors can override normal cell cycle regulation, which may result in unrestrained cell cycle progression and uncontrolled cellular growth, two common features of the carcinogenic process (Worsham et al., 2006). The detrimental effects of such characteristics are often further compounded by mutations within TP53 itself, which can lead to a dysfunctional cell death response and the acquisition of an additional hallmark; the **evasion of cell death** (Ichimura et al., 2000).

Unsurprisingly, excessive tumour cell proliferation can rapidly exceed the diffusion capacity of the surrounding vasculature, causing a 2–9% reduction in oxygen levels and leaving some regions hypoxic (Muz et al., 2015). To counteract this, tumour masses exploit a process normally limited to embryogenesis and wound healing known as **angiogenesis**, whereby the tumour microenvironment releases proangiogenic factors, such as the vascular endothelial growth factor (VEGF) and fibroblast growth factor (FGF). Both factors promote the vascularization of the tumour, thus restoring the oxygen supply (Niu and Chen, 2010). As this is such a vital development in cancer progression, other systems that also utilise angiogenesis have been widely studied with many parallels to cancer progression identified. For example, during wound healing cell proliferation is also enhanced, while specialised cells migrate through the wound bed to repair the damaged tissue (Sundaram et al., 2018). Such cellular migration is akin to **cancer metastasis**, whereby cells leave the primary tumour and colonize distant organs. However, unlike cancer the hyper-proliferation and cell migration of specialised wound healing cells is highly controlled and subsides once the repair is completed.

As research has progressed, this original list of hallmarks has been extended to include two additional hallmarks: **evasion of the immune response** and **deregulation of cellular energetics**. Two characteristics that are crucial to the development of such cancer hallmarks have also been recognised: **tumour-promoting inflammation** and **genome instability** (Hanahan and Weinberg, 2011). These additions redirect the focus of cancer research from a reductionist view to one of ever increasing complexity, with the majority of altered cellular behaviours and cancer development rooted in highly variable genetic paths, which gradually lead to acquisition of genome mutations until little uniformity remains and genome destabilization is widespread (Weinberg, 2014).

1.2 Genome Instability and Cancer

Progressive loss of genetic integrity can be advantageous to cancer development, as certain mutant genotypes can confer a selective advantage over their healthy counterparts. This can encourage the outgrowth and eventual dominance of mutant cell lineages over those that lack such genetic changes. Since multiple genetic alterations are necessary to achieve dominance during cancer development, any increased propensity for genetic change will support this process. The molecular mechanisms underlying genomic instability are rooted in defects in processes that normally help preserve genetic information during the cell cycle and cell division. Namely, DNA repair systems, accurate DNA replication, proper chromosome segregation and correctly orchestrated cell cycle checkpoints. A thorough understanding of these systems is therefore essential to provide an improved understanding of the genome instability found within cancers.

1.2.1 DNA Damage and its repair

Within a single cell genetic changes can occur between $1 \cdot 10^4$ – $1 \cdot 10^6$ times per day (De Bont and van Larebeke, 2004) and can be either inherited or more commonly arise from certain environmental exposures, including UV radiation, tobacco smoke and some viral infections. DNA repair pathways exist to identify and repair the damaged DNA. These networks can be subdivided into several distinct mechanisms based on the type of DNA lesion including, base excision repair (BER), nucleotide excision repair (NER), mismatch repair (MMR), homologous recombination (HR) and non-homologous end joining (NHEJ) (Figure 2).

1.2.1.1 Base excision repair (BER)

Base excision repair (BER) is mainly activated during G1 and repairs basic single base or short fragments of damage that are inflicted through endogenous sources such as oxidative stress. The BER pathway relies on DNA glycosylases to recognise and excise the damaged base from the DNA sequence, to leave an abasic site (AP-Site) which is then repaired by short-patch repair or long-patch repair depending on the length of the error (Krokan and Bjørås, 2013).

1.2.1.2 Nucleotide excision repair (NER)

Nucleotide excision repair (NER) is responsible for addressing complicated lesions including bulky adducts and cross-linking lesions that often develop through external sources such as exposure to UV light or DNA-alkylating carcinogens. When damage is detected, NER excises the damaged site along with its flanking base pairs ensuring that damage-containing nucleotides are removed. After excision the resulting gap is repaired using the intact complementary strand as a template (Schärer, 2013).

1.2.1.3 Mismatch repair (MMR)

Mismatch repair combats replication errors to maintain the fidelity of DNA replication and mainly resolves single nucleotide mismatches, small insertion loops and short repeated sequences known as microsatellite regions generated by DNA polymerase (Vilar and Gruber, 2010).

1.2.1.4 Recombination repair (HR) and non-homologous end joining (NHEJ)

Recombinational repair can be further divided into Homologous recombination repair (HR) and non-homologous end joining (NHEJ) and the choice between the two is dependent on the cell cycle stage. NHEJ is active throughout the cell cycle, but the repair pathway functions to modify the broken DNA ends, and ligate them together with little or no homology, which often generates additional deletions or insertions. While HR is nearly absent in G1, before peaking in S phase, and declining in G2/M. This is because HR requires an undamaged DNA template on the sister chromatid or homologous chromosome to repair the break. Although limiting its usage, it results in a complete reconstitution of the original DNA sequence.

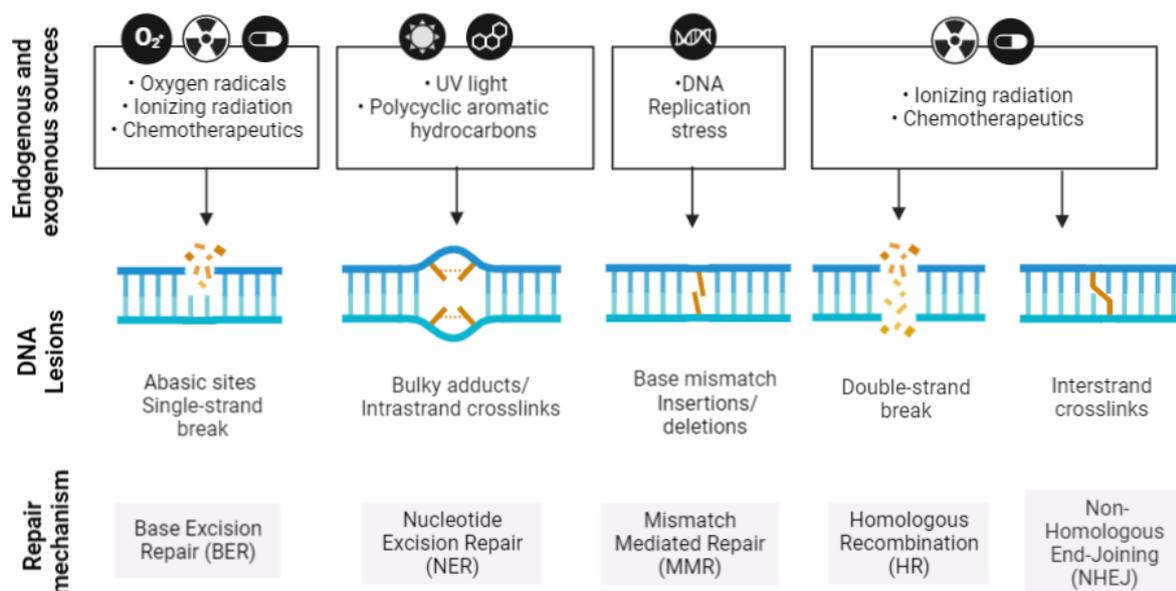


Figure 2: Common causes of DNA damage and their corresponding repair pathways.

Endogenous and exogenous agents are major sources of a range of DNA damage, from single base mismatches to significant double stand breaks. Each then requires repair via specific DNA repair pathways (adapted from Biorender database).

If DNA damage is not identified and removed by DNA repair systems, they can promote more severe lesions such as frameshift mutations, chromosome translocations and aneuploidy leading to chromosomal instability (Stading et al., 2021). The activation of repair pathways can also significantly affect the cell cycle, as repair pathways often promote the activation of DNA damage checkpoints to arrest the cell cycle when DNA damage is detected. Although this is a preventive measure, to enable repair, it can cause the cell additional stress and further reduce damage tolerance and genome integrity, especially during mitosis (Mao et al., 2008). This suggests that mutation and genome instability are common initiating events that have the potential to impact critical genes such as DNA repair or disrupt the cell cycle to result in a compounding effect which can promote cancer development (Negrini et al., 2010).

1.3 The cell cycle and its key events

The cell cycle is separated into four distinct stages G1, S, G2 and M phase. The phases can be grouped into mitosis (M phase) and the phases in-between (G1, S, G2 phase) which are collectively known as interphase. G1 is primarily a period of cellular growth whereby macromolecular synthesis and metabolic activity takes place in preparation for the Synthesis (S) phase. During S phase, the genome is copied, and the chromosomes are duplicated. Following this is a second period of cellular growth and protein synthesis, G2 which occurs in preparation for mitosis. In mitosis, the duplicated chromosomes situated within a single cell are segregated and the single cell subsequently divides into two identical daughter cells (Vermeulen et al., 2003).

1.3.1 *DNA replication*

An essential process within the cell cycle is genome replication, whereby chromosomal DNA is faithfully duplicated ready for transmission to a daughter cell. DNA must be precisely replicated. However, aberrant DNA structures, protruding DNA damage and replication obstacles such as slow-moving RNA polymerases can hinder this process, leading to replication fork arrest which stall DNA replication and increases the replication burden. Replication stress is a primary cause of genome instability and a common enabling characteristic of cancer (Leonard and Méchali, 2013). To ensure that the genome can be replicated efficiently, DNA replication is initiated simultaneously from multiple AT-rich sites known as origins. Replication origins are numerous within higher eukaryotes, with 30,000–50,000 origin sites situated throughout the human genome (Méchali, 2010). Despite this, replication is only facilitated from a subset of these origins which require licencing and activation before replication can commence. By separating these two mechanisms, the potential for

unauthorised origin firing is limited and the possibility of genome reduplication is reduced, thus, reducing replication stress. In G1 of the cell cycle, dormant DNA helicases (MCM2-7) are loaded on to sections of the DNA sequence by the origin recognition complex (ORC). Once loaded, replication initiator proteins Cdc6 and Cdt1 bind to the helicases to form a licensed, but inactive pre-replicative complex (pre-RC) (Kang et al., 2018). As the cell progresses to S phase, active Cdc45 and GINS proteins are incorporated into the pre-RC to form the active Cdc45-MCM2-7-GINS (CMG) helicases. In response, Dbf4-dependent and cyclin-dependent kinases (DDK and CDKs) promote origin firing and the activation of the now fully formed (CMG) helicases. Following this, origin DNA is melted and replication factors are recruited to establish bidirectional replisomes for DNA replication to progress (Burgers and Kunkel, 2017) (Bleichert, 2019) (Fenwick et al., 2016). Each DNA strand is replicated in a semi-conservative manner, as one strand is synthesized continuously by DNA polymerase ϵ (Pol ϵ) (leading strand), while the opposite lagging strand is copied in short ~ 200 base pair Okazaki fragments by Pol δ . Both strands are then later joined together by DNA ligase to produce a complete DNA strand (Stillman, 2008) (Figure 3).

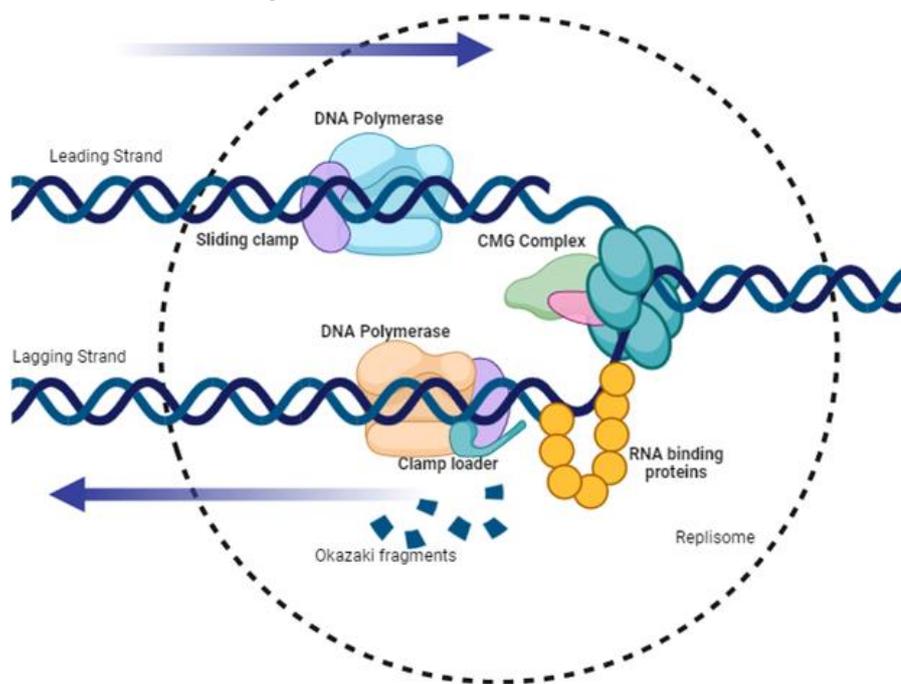


Figure 3: DNA Replication: Leading and Lagging Strand Synthesis

DNA replication begins with the unwinding of the DNA helix via the DNA helicase. Each unwound ssDNA is then used as a template for the synthesis of a new complementary strand. One strand is continuously replicated (leading strand), while the other is produced via a series of short Okazaki DNA fragments (lagging strand) which are later sealed together by DNA ligase. Consequently, the lagging-strand polymerase remains at the fork after synthesising each Okazaki fragment ensuring a single polymerase can be used to synthesize the entire lagging DNA strand. While the replication

fork remains open, the strands are protected via DNA-binding proteins which cover each of the single DNA stands throughout the replication process (created using biorender).

1.3.2 Mitosis

Mitosis is a highly complex event that can be subdivided into prophase, prometaphase, metaphase, anaphase, and telophase. During mitosis the duplicated chromosomes condense and are separated by mitotic spindles. The segregated chromosomes are then encased when the parental cellular membrane constricts to form two daughter cells, thus completing the cell cycle (Scholey et al., 2003) (Figure 4).

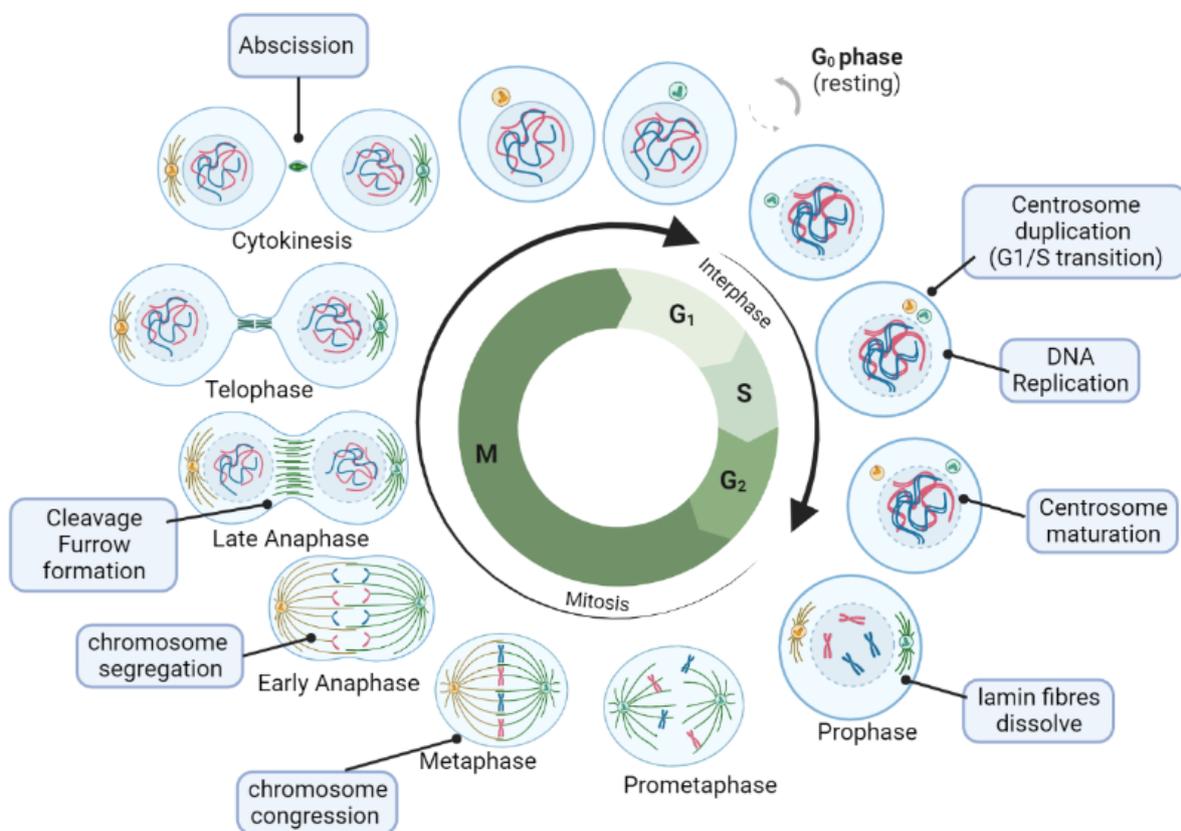


Figure 4: Phases of cell division (Zitouni et al., 2014)

Cells transition through several stages of growth (G₁, S and G₂ Phase) and division or Mitosis (M phase). Mitosis is divided into prophase, prometaphase, metaphase, anaphase, and telophase. Until cell division is completed during Cytokinesis (Zitouni et al., 2014). Successful chromosome segregation also relies on centrosomes which are major microtubule-organizing centres (MTOC) and are responsible for the formation of the mitotic spindles at each polar end of the cell. Mitotic spindles then act as anchor points for the microtubules to pull apart the duplicated chromosomes (created using biorender).

During **prophase**, the nucleolus disperses and the chromatin condenses with the aid of Condensin II, which wraps around histone octamers to displace transcription factors and prevent further transcription (Hirota et al., 2004). Meanwhile, CDK1 and Aurora A kinases are upregulated to promote the construction of γ -tubulin ring complexes. These complexes provide the foundations for centrosome-associated microtubules (MTs). As mitosis progresses, additional kinases gather at the centromere to assist in mitotic quality control (Blagden and Glover, 2003). **Prometaphase** begins with the weakening of the nuclear envelope, as the lamin fibres that line its inner surface dissolve. This merges the cytoplasm and nucleoplasm together and allows the MTs to access the kinetochores for their subsequent attachment. During **metaphase**, MTs align the chromosomes at the metaphase plate in a process known as chromosome congression (Maiato et al., 2017). During this process, each sister chromatid is assessed for its proper attachment to the mitotic spindle via the spindle assembly checkpoint (SAC). Once satisfied, the cohesins that bind the sister chromatids together are cleaved and the now independent sister chromatids can be separated by polarised MTs. The polarised movement of MTs relies on a positive-negative charged polar gradient. Negatively charged MTs that are situated at the spindle pole region begin to depolymerise and traverse towards the positively charged kinetochore region. This creates a constant poleward flow of MT and forces the chromosomes apart enabling mitosis to progress to **anaphase** (McIntosh et al., 2012). While the chromosomes are still condensed, **telophase** initiates to re-establish the previously dispersed nuclear envelope and produce a nuclear boundary around each of the daughter's nuclei (Güttinger et al., 2009).

Finally, cell division is completed with **cytokinesis**; whereby the cytoskeleton that separates the cytoplasm of each daughter cell is re-established. It is normally a highly ordered process, requiring an intricate interplay between cytoskeletal, chromosomal and cell cycle regulatory pathways (Miller, 2011). The process begins during chromosome segregation, whereby the contractile ring forms under the surface of the plasma membrane, which upon contraction, forces an ingression upon the plasma membrane and the development of the cleavage furrow. Constriction of the contractile ring is driven by bipolar filaments of myosin II which use their motor activity to move along two antiparallel actin filaments, this causes them to slide past each other and constrict the cell through a "purse string" action (Spira et al., 2017). Contraction continues until the cell is almost completely separated into two nascent daughter cells. However, the contractile ring alone is not sufficient to separate the cell into two and an intercellular bridge remains between the two daughter cells. It is only until abscission, whereby ESCRT-III filaments form cylindrical spirals that further constrict the intercellular bridge, forces the two tethered cells to separate (Figure 4) (Mierzwa and Gerlich, 2014).

1.3.3 *The Centrosome cycle.*

Cell division is in part, facilitated by centrosomes, cytoplasmic organelles that function as microtubule-organising centres (MTOCs) during the cell cycle. For a cell to divide successfully it requires two mature centrosomes which facilitate bipolar spindle formation at metaphase and are themselves segregated, one into each daughter cell, during cytokinesis. Centrosomes are self-replicative, and their duplication is strictly co-ordinated with cell cycle progression to ensure two mature centrosomes are available for chromosome segregation during mitosis. Centrosome replication begins with duplication of the pre-existing mother centriole which duplicates itself to form a daughter centriole at the G1/S transition and is completed during S-phase. Once duplicated, the daughter centriole elongates during S-phase and G2 to form procentrioles. However, procentrioles are still immature and do not possess the accessory appendages and centriolar satellite structures that a mature centrosome does. For a procentriole to mature, pericentriolar material (PCM) components are recruited causing it to increase in size and enhance its MT-nucleating capacity to ensure correct spindle pole formation during mitosis. Finally, the flexible proteinaceous linker that bridges the two centrioles is severed enabling complete separation of mother and daughter centriole and their migration to the cell's polar regions where the foundations for each mitotic spindle will be situated (Azimzadeh and Marshall, 2010) (Figure 4).

1.4 Cell Cycle Regulation

Progression through each phase of the cell cycle is controlled via cyclins, cyclin-dependent kinases (CDKs) and cell cycle checkpoints (CP). Cyclins and CDKs form active subunits that promote cell cycle progression while, cell cycle checkpoints can halt the cell cycle by triggering cell cycle arrest or cell death (Figure 5).

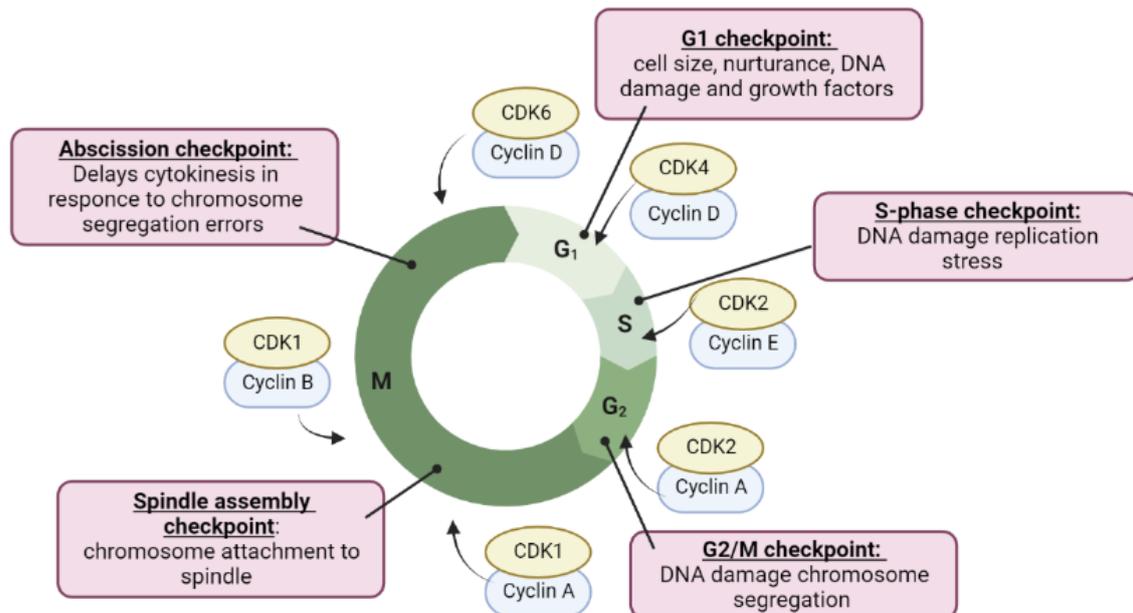


Figure 5: The Cell Cycle and Its Checkpoints

The cell cycle is promoted by cyclin dependent kinases and their activation partners, cyclins. In humans, the cell cycle involves three interphase CDKs (CDK2, CDK4 and CDK6), a mitotic CDK (CDK1) as well as cyclins of four different classes (the A-, B-, D- and E-type cyclins) (Malumbres and Barbacid, 2009). Checkpoint mechanisms exist to restrain inappropriate cell cycle progression (created using biorender).

1.4.1 Cyclins and cyclin-dependent kinases (CDKs)

There are three interphase CDKs (CDK2, CDK4 and CDK6) and a mitotic CDK (CDK1), the activity of each CDK is controlled via the availability of their cyclin regulatory subunits. There are a total of ten cyclin subunits from four different classes (A-, B-, D- and E-type cyclins), with each CDK forming a subcomplex with a specific class of cyclin. Specific cyclins are synthesised and destroyed at various stages of the cell cycle. Therefore, their availability is limited and the activity of their CDK counterparts is restrained ensuring that progression through the cell cycle is controlled (Malumbres and Barbacid, 2009).

For instance, during G1, D-type cyclins (D1, D2 and D3) bind and activate CDK4 and CDK6 to initiate DNA synthesis. CDK4 and CDK6 activation also promotes the expression of E-type cyclins (E1 and E2), which bind and partially activate CDK2. This drives the G1/S transition. During the final stages of DNA replication, CDK2 is further activated by an additional cyclin, A2 which promotes the progression from S phase to mitosis. At the close of interphase, CDK1 is activated by A-type cyclins to facilitate mitosis. As the cell cycle progresses through mitosis, A-type cyclins are then degraded and CDK1–cyclin B complexes are formed to promote the final stages of mitosis (Malumbres and Barbacid, 2009).

The activity of CDKs can be inhibited by CDK inhibitors (CDKi), of which there are two main families the inhibitor of kinase family (INK4s) and the CDK-interacting protein/kinase inhibitory protein family (CIP/KIP). The INK4 family regulates cell progression through G1 and the initiation of DNA synthesis during S phase. The family has four main members: p16INK4a, p15INK4b, p18INK4c, p19INK4d each only bind to CDK4 and CDK6 to regulate the activity of cyclin D. The INK4 proteins can be contrasted with more broadly acting inhibitors, the Cip/Kip family, which is composed of p21, p27 and p57 and can inhibit cyclin D-, E-, and A-dependent kinases (Besson et al., 2008) (Sherr and Roberts, 1999).

1.4.2 Cell cycle checkpoints and the DNA damage response

While the cyclical activation of cyclin-CDKs drives the cell cycle forward, checkpoint mechanisms also exist to restrain the cell cycle when errors are encountered to enable repair. If the repair is unsuccessful, a cell can transition into a permanent state of arrest, or the cell death pathway can be activated (Matson and Cook, 2017). The first of these checkpoints, the G1/S checkpoint and restriction point occurs before the cell has fully committed to the cell cycle. This checkpoint serves two functions, the first like other cell cycle checkpoints is to respond to DNA damage and halt the progression to S-phase. Its second role is to assess the cell for its fitness to continue through the cell cycle and divide. To achieve this, a cell must pass through the restriction point, but to do so a cell must be capable of relinquishing its reliance on external growth factors and begin to synthesize its own cell cycle promoting factors, cyclins. If this is not achieved, a cell can enter quiescence at G0 (Blagosklonny and Pardee, 2002). Several other checkpoints operate throughout the cell cycle in response to DNA damage including the intra S phase checkpoint; which safeguards DNA replication, the decatenation checkpoint; which limits entangled (catenated) sister DNA helices, the G2 checkpoint; which is situated at the G2/M boundary and offers last opportunity to repair damaged DNA before mitosis, and two additional checkpoints that are also found within M-phase; the spindle assembly checkpoint (SAC) and the abscission checkpoint (Paulovich et al., 1997) (Soliman et al., 2023).

Mitotic and interphase checkpoints are commonly activated via the DNA damage response (DDR), a multifaceted signalling network which relies on a family of phosphatidylinositol-3-kinase related kinases (PIKKs) to stimulate downstream signalling and coordinate the cell cycle. In mammals, ATM (ataxia telangiectasia mutated) and ATR (ATM- and Rad3-related) are key to this signalling network and feature throughout the network. ATM is primarily involved in the response to DNA double strand breaks (DSB) and activates the G1/S and G2/M cell cycle checkpoints. ATR primarily responds to single strand breaks (SSB) and replication stresses during S-phase. However, there is a degree of overlap with the activity of ATM and ATR, as some replication errors like fork collapse can produce DSBs, therefore ATM can be recruited during S phase, while ATR can respond to ssDNA intermediates formed during failed HR repair in G2/M (Figure 6) (Awasthi et al., 2015).

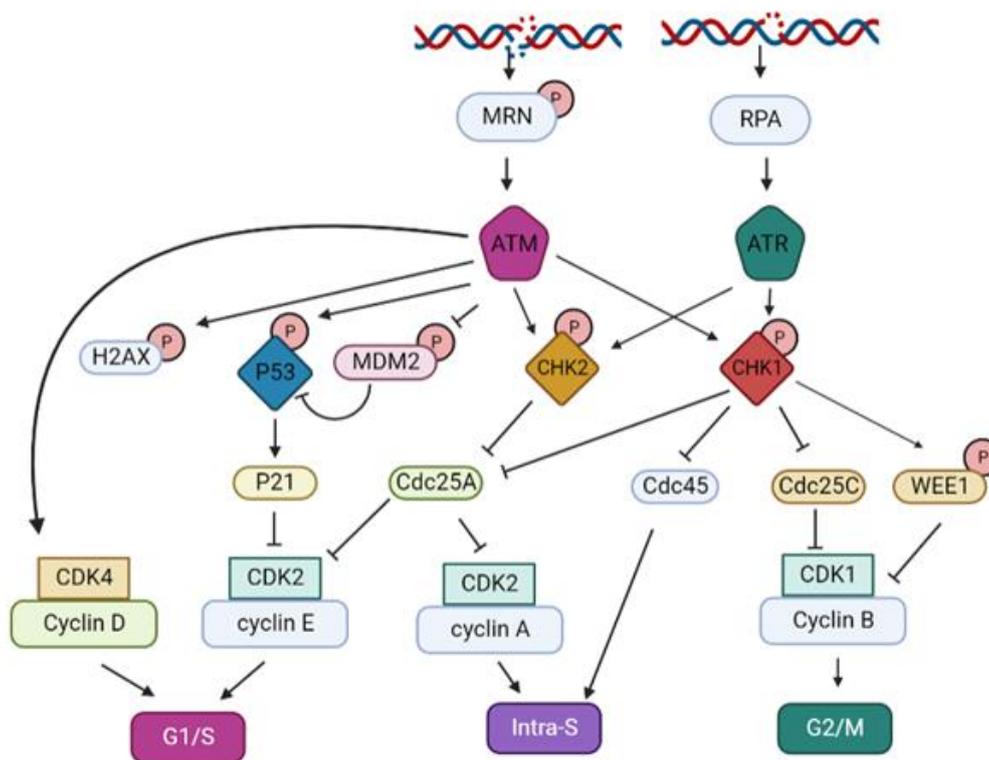


Figure 6: The DNA Damage response pathway

DNA damage lesions such as DSB and ssDNA are detected via sensors which activate PI3KK kinases ATM and ATR. Both initiate a cascade of effector proteins through phosphorylation events which coordinate cell-cycle checkpoints, DNA repair and apoptotic responses to DNA damage.

ATM is recruited by MRE11–RAD50–NBS (MRN) or the Mre11 complex, which detects DSBs and temporarily spans the DNA break. Upon recruitment, ATM becomes auto-phosphorylated and disassembles from its homodimer structure to become kinase active monomers. Once activated, ATM contributes to the phosphorylation of the histone variant H2AX, which initiates a signalling cascade to assemble additional DDR proteins to the damaged site and promote its repair (Dupré et al., 2006). ATM plays a crucial role in the activation of the G1/S cell cycle checkpoint and inhibiting the entry to S-phase in response to DNA damage. To achieve this, ATM directly phosphorylates p53, MDM2 and checkpoint kinase 2 (CHK2). Phosphorylation of p53 and MDM2 removes the autoregulatory feedback loop between the two proteins and allows for the accumulation p53. This consequently triggers the upregulation of p21 which inhibits the cyclin E/CDK2 complex to enforce the G1 checkpoint and prevent the transition from G1 to S-phase (Dasika et al., 1999). ATM activation of CHK2 induces ubiquitination and degradation of the S-phase, promoting phosphatase Cdc25A which prevents the activation of CDK2. Additionally, ATM has been shown to directly phosphorylate cyclin D1, again preventing cyclin D/CDK4 coupling (Deckbar et al., 2010).

Problems encountered during DNA replication can cause replication fork stalling, which gives rise to exposed single stranded DNA (ssDNA) as the replicative helicase moves ahead of the DNA polymerases. The exposed ssDNA is coated with ssDNA binding protein RPA, and ATR is recruited to initiate the intra-S phase checkpoint response (Ma et al., 2020). ATR inhibits origin firing via CHK1 which prevents the binding of Cdc45 to the pre-replicative complex (pre-RC), thus blocking further origin firing and providing time to resolve the source of replication stress (Truong and Wu, 2011). In addition, CHK1 inhibition of Cdc25A phosphatase prevents the dephosphorylation and subsequent activation of CDK2, thus suppressing the formation of cyclinE/CDK2 and cyclinA/CDK2 complexes, again preventing cell cycle progression. ATR can also directly prevent DNA damage by stabilising replication forks and phosphorylating MCM2 to prevent firing. Mitosis entry can also be blocked via ATR and ATM regulating the activity of cyclinB/CDK1. ATR and ATM target CHK1 and CHK2 to phosphorylate and limit Cdc25C phosphatase activity, preventing cyclinB/CDK1 activation (van Vugt et al., 2010). CHK1 also phosphorylates and increases the kinase activity of WEE1 which negatively regulates entry into mitosis by facilitating inhibitory phosphorylation of the cyclinB/CDK1 kinase (De Schutter et al., 2007).

The decatenation checkpoint operates in parallel with the G2 DDR to regulate the transition from G2-to-M as well as during mitosis, as the cells progress through Metaphase-to-Anaphase. This

checkpoint is triggered via the ATR signalling pathway to limit the formation of cyclin B1/Cdk1 complexes (Deming et al., 2001). The pathway relies on the activity of Topoisomerase II which disentangles chromosomes by passing one double helix through a transient double-strand break (DSB) in another double helix before resealing the break (Damelin and Bestor, 2007).

1.4.3 Mitotic checkpoints

Once a cell progresses to mitosis, a further checkpoint, the spindle assembly checkpoint (SAC), monitors chromosome segregation. During metaphase, components of the SAC, MAD1 and BUB1 monitor proper spindle microtubule-kinetochore attachment and tension. If problems arise, MAD2, BUBR1, BUB3, assemble at the kinetochore to form the mitotic checkpoint complex (MCC) which binds to CDC20 and inhibits the activation of the anaphase-promoting complex (APC). This ensures the stabilization of APC/CCDC20 substrates such as cyclin B which maintain the cell in a mitotic state of arrest by preventing further activation of its co-factor and cell cycle regulator CDK1. Therefore, preventing the transition to anaphase. Once each kinetochore is attached to the mitotic spindle, the SAC proteins are displaced and CDC20 is released, allowing the APC to target Cyclin B for degradation and promote mitotic progression through the accumulation of CDK1 (Lawrence and Engebrecht, 2015) (Figure 7). However, if mitotic arrest is prolonged this can cause the SAC to weaken, whereby MAD2 activation at unattached kinetochores is progressively diminished, which can result in the gradual disassembly of MCC. This may enable APC/CCDC20 activation and gradually degrade Cyclin B, eventually reducing it to a level insufficient for maintaining cell cycle arrest, finally, resulting in mitotic slippage whereby cells can exit mitosis without proper chromosome segregation and cytokinesis (Lok et al., 2020). This can lead to an increase in chromosome instability and abnormal chromosome ploidy.

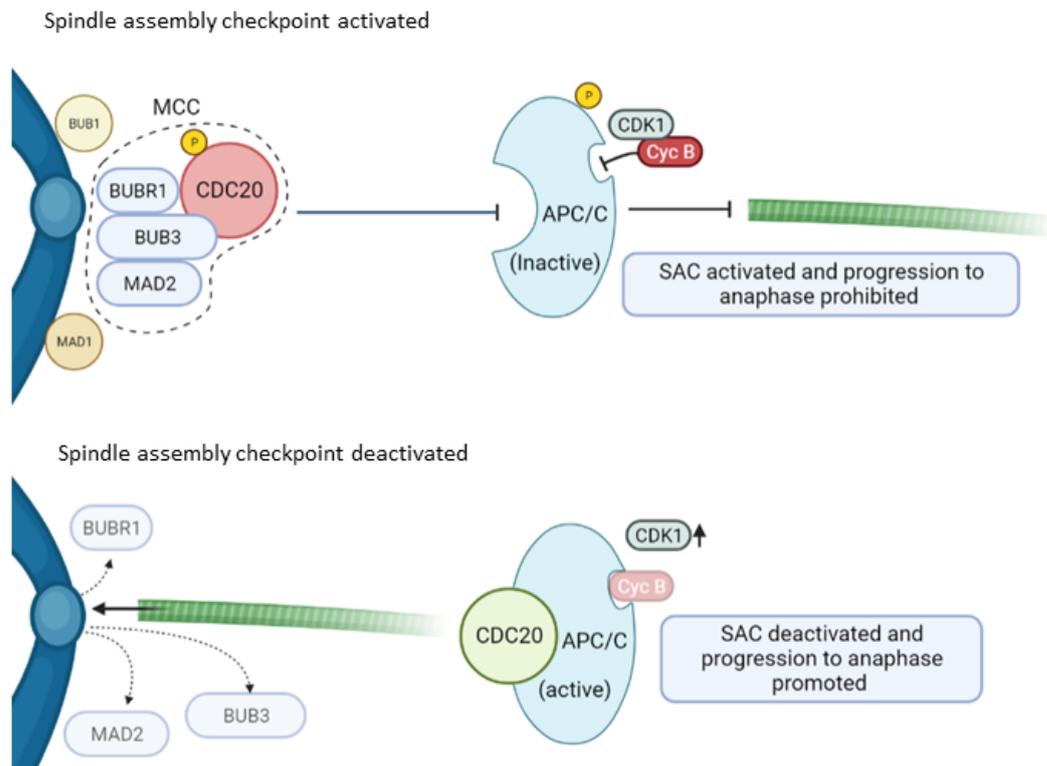


Figure 7: Mitotic Checkpoints: the spindle assembly checkpoint (SAC)

The spindle assembly checkpoint (SAC) and mitotic checkpoint complex (MCC) act as a surveillance mechanism delaying the onset of anaphase until each kinetochore is paired with a microtubule, ensuring accurate chromosome segregation during mitosis. When SAC is activated, MCC withholds CDC20 and limiting its association with the APC Complex preventing the accumulation of CDK1, prolonging mitotic arrest. Once proper kinetochore attachment is achieved, SAC is deactivated, CDC20 is released and associates with APC. This enables Cyclin B to be degraded and CDK1 to accumulate (Created in biorender).

The final cell cycle checkpoint, the abscission checkpoint, is activated during cytokinesis in response to chromatin bridges or lagging chromosomes trapped within the intercellular bridge of the separating daughter cells. Activated via CHK1 and CHK2 phosphorylation, Aurora B along with INCENP, Survivin and Borealin form a macromolecular complex known as the chromosome passenger complex (CPC) (Carmena et al., 2012). CPC is a key orchestrator of chromosomal and cytoskeletal events, which localises to the midpoint and imposes the abscission checkpoint by phosphorylating the endosomal sorting complex required for transport-III (ESCRT-III). This phosphorylation prevents the downstream activation of ESCRT-III by limiting ATPase association and its Adenosine triphosphate (ATP) energy source, therefore preventing its relocation to the abscission site and the delivery the final cut (Steigemann et al., 2009).

Once the chromatin bridges have been resolved, progression to the final cut can continue, whereby Tsg101 (ESCRT-I) and ALIX accumulate at the midbody and recruit ESCRT-III, the complex responsible for the scission activity. During this time, Spastin, is recruited to the midbody by ESCRT-III subunits and ensures the region is clear of any remaining microtubules. After which, ESCRT-III driven by the ATPase VPS4, assembles in filaments around the abscission site and makes the membrane curve, eventually driving the final break between daughter cells (Figure 8).

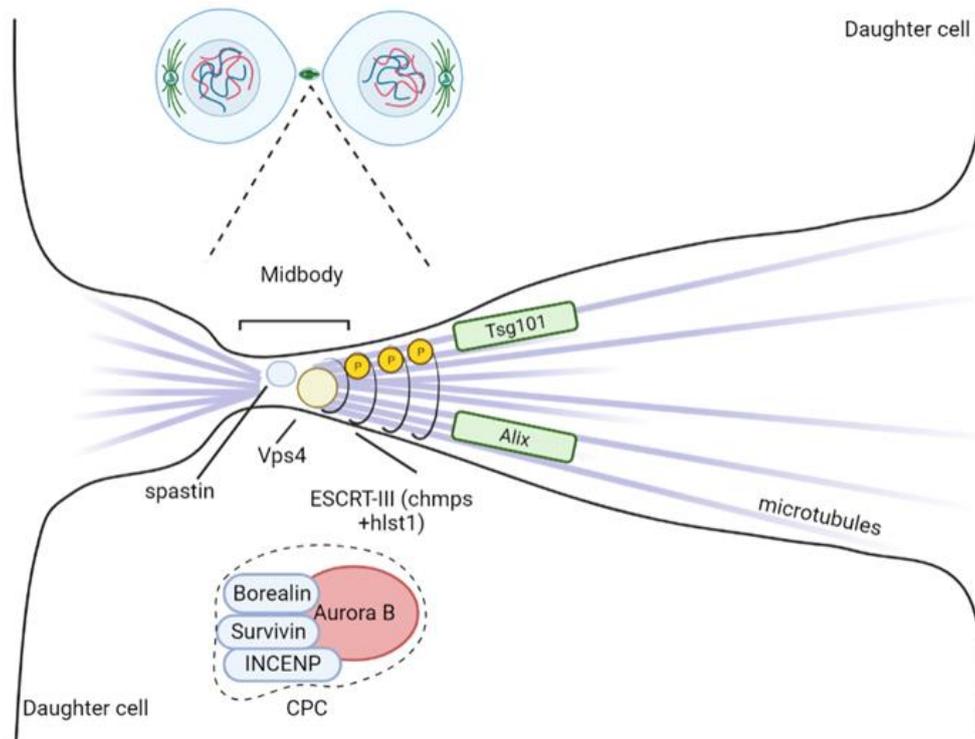


Figure 8: Mitosis Checkpoints: Abscission Checkpoint

During cytokinesis, complete cell division is prevented via the intercellular bridge. If chromosomal DNA spans this intracellular space (chromosome bridge), the abscission checkpoint is activated to prevent severe DNA damage such as aneuploidy. During this checkpoint, Aurora B and the CPC, phosphorylate ESCRT-III subunits to stabilise the midbody and delay the completion of cell division (created in biorender).

1.4.4 Checkpoint dysregulation and chromosome instability (CIN)

Cell cycle checkpoints are critical to the success and stability of the cell cycle, with their dysregulation profoundly affecting a cells' fate and continued prosperity. If a DNA damage checkpoint fails, the identification and repair of DNA damage will be compromised, and the cell will experience the consequences of replicating the damaged template. Although, these consequences will vary depending on the type of damage being replicated, even a small error can have compounding effects. For example, if a G1 cell acquires a single-stranded break within its DNA, but it can progress through to S phase the single strand lesion can potentially become a double strand break during replication. Which in turn will disrupt the replicative machinery and increase replication stress (Abbotts and Wilson, 2017). As well as acquiring additional lesions, some repair pathways may be lost or unavailable if the damage is not repaired and the cell continues to progress, again potentially introducing more severe lesions. For example, segregation of broken chromosomes can lead to their complete loss or rearrangement and result in chromosome instability (CIN) and abnormal chromosome ploidy (Vargas-Rondón et al., 2017). CIN indicates a persistent defect in the fidelity of chromosome segregation during mitosis and has been recognized as a major source of genetic variation that contributes to tumour development due to the potential for tumour suppressor loss and oncogene activation (Thompson et al., 2010). In fact, CIN is a characteristic of most solid tumours and is associated with poor prognosis due to the increased heterogeneity of the tumour through highly aneuploid karyotypes. Such variation can favour survival adaptations within stressful environments, which in turn can frequently cause cytotoxic drug resistance and treatment failure (Bakhoun and Cantley, 2018).

Cell cycle checkpoints can fail for a multitude of reasons, Firstly, as with all cellular processes, checkpoints maintain an intrinsic error rate (Paulovich et al., 1997). Secondly, like signalling pathways there is a degree of adaptation whereby, after a period of prolonged arrest, the cell may resume its progression in favour of cell death or senescence, even though the damage remains unrepaired. Additionally, as the checkpoints are a collection of proteins with different roles and responsibilities, they are affected by the relative gene expression and activity of those individual components (Bertoli et al., 2013).

1.5 Regulation of gene expression: mRNA processing and export

Gene expression is fundamental to cell function and cancer cells invariably exhibit dysregulation of gene expression as a contributing factor to development. Nearly all facets of malignancy such as survival, proliferation, metastasis and invasion are impacted by the dysregulation of selected messenger (m)RNAs. Multiple steps are involved in the process of gene expression including transcription, RNA processing, RNA export, translation and post-translational modifications.

As the transcription of mRNAs occurs in the nucleus, it is physically separated from their subsequent cytoplasmic translation into functional proteins. This separation offers a significant opportunity for regulation, which ultimately means that the transcriptome does not always predict the proteome with post-transcriptional regulation often decoupling transcription and translation.

1.5.1 mRNA processing

Before newly transcribed mRNA molecules can serve as a template for protein synthesis, each must transition to a mature mRNA form through a series of processing steps. If mRNA is not properly processed, it will not be exported and instead will be targeted for degradation, therefore reducing gene expression (Carmody and Wentz, 2009).

The first processing step involves the addition of a 7-methylated guanosine to the first nucleotide of the RNA by the cap binding complex (CBC) (Figure 9). This cap protects the 5' end of the pre-mRNA from degradation and provides a unique identifier during subsequent processing steps (Ramanathan et al., 2016). This is followed by the precise cleavage of introns, non-coding sequences dispersed within each mRNA, by the spliceosome. Once removed, only coding exon sequences remain. The variable retention of certain exons through alternative splicing ensures vast proteomic diversity despite relatively small gene pools (Shi, 2017). The third processing step protects the mRNA from premature degradation through the addition of a poly-A tail, whereby the end of the pre-mature mRNA is cleaved to reveal a 3' hydroxyl which is used as an attachment point for a chain of 100 to 250 adenosine nucleotides (Goodall and Wickramasinghe, 2021). Then, the mRNA must be packaged and exported from the nucleus through association with multiple mRNA binding proteins and export factors. Each of these processing steps are aided by a family of RNA-binding proteins known as nuclear ribonucleoproteins (hnRNPs), (Dreyfuss et al., 2002). HnRNPs associate with nascent transcripts as they are transcribed, surrounding them with six "core" hnRNP proteins (A1, A2, B1, B2, C1 and C2) in a bead on a string structure (Jean-Philippe et al., 2013). The resulting complex is highly dynamic and multifunctional, remodelling throughout RNA processing to accompany the mRNA to its maturity. Therefore, hnRNPs participate in all crucial aspects of RNA processing, including pre-mRNA splicing, mRNA export, localization, translation and stability (Dreyfuss et al., 2002). Consequently, specific processes such as splicing can be dramatically affected by the presence of hnRNPs within the splice sites of a pre-mRNA, as accessibility of the splice site can be variable and thus producing different mRNA isoforms to contribute to proteome diversity (Martinez-Contreras et al., 2007).

1.5.2 mRNA export and The TRanscription and EXport (TREX) complex

A fully mature mRNA transcript is licenced and delivered to the nuclear pore complex (NPC) for its exportation to the cytoplasm by export factors and the TRanscription and EXport (TREX) complex (Carmody and Wentz, 2009). This complex is highly conserved, with key elements observed in yeast, *X. laevis* and humans. Its core components include a multi-subunit THO complex; (THOC1, -2, -3, -5, -6, and -7), A DEAD-box RNA helicase; UAP56/DDX39B and an RNA export adapter ALY (Figure 9) (Pühringer et al., 2020).

In addition to this core structure, multiple subsidiary proteins have been identified as TREX subunits including CIP29 (SARNP), ZC3H11A and ALY-like proteins; CHTOP, UIF (also known as FYTTD1), LUZP4, and POLDIP3 (Strässer et al., 2002) (Dufu et al., 2010) (Heath et al., 2016). The full implications of these subunits within the TREX complex are yet to be fully identified, although many share redundancies with the core TREX proteins. For example, CHTOP functions in a similar manner to ALY, with two UAP56-binding motifs (UBM) which are also present in UIF and LUZP4. CHTOP can stimulate the ATPase and helicase activities of UAP56 and like ALY, it is dependent on UAP56 for mRNA loading (Chang et al., 2013). UIF is upregulated with the loss of ALY and freely associates with other TREX components (Hautbergue et al., 2009). LUZP4 also compensates for the loss of ALY expression, and partially compensates for the loss of both ALY and UIF (Viphakone et al., 2015). Both POLDIP3 and ZC3H11A associate with UAP56 in an ATP-dependent manner, with knockdown of ZC3H11A leading to an accumulation of mRNA in the nucleus (Folco et al., 2012). Due to these characteristics, it has been suggested that the TREX subunits may form alternative TREX complexes by replacing core components. Finally, the cytokine-induced protein 29 (CIP29) (SARNP/HCC-1) may also feature as a TREX complex component as CIP29 associates with the RNA export proteins Sub2/UAP56 and Yra1/Aly in *Saccharomyces cerevisiae* and has been linked to cell cycle progression (Leaw et al., 2004), (Fukuda et al., 2002).

Although the precise constitution of the TREX complex is still debated, recent structural investigations have revealed further details of its molecular structure. The TREX structure features two 14-subunit THO–UAP56 dimers that assemble into a flexible 28-subunit tetramer, accessible to multiple RNA processing complexes (Pühringer et al., 2020). The nascent transcript is bound by UAP56 and secured by ALY which bridges the two UAP56 dimers through its two UMB sites, before promoting licencing for export. As the TREX subunits (LUZP4, POLDIP3, UIF, CHTOP) hold a similar functionality to ALY through their own UBMs they may function to increase the binding affinity of UAP56 within the complex, and aid licencing. The assembled TREX–mRNP complex, can then be loaded to the export factor NXF1–NXT1 by ALY or alternative export adaptors. Once loaded, NXF1

interacts directly with the NPC transport channel to promote permeability and transport its cargo mRNA into the cytoplasm (Wickramasinghe and Laskey, 2015).

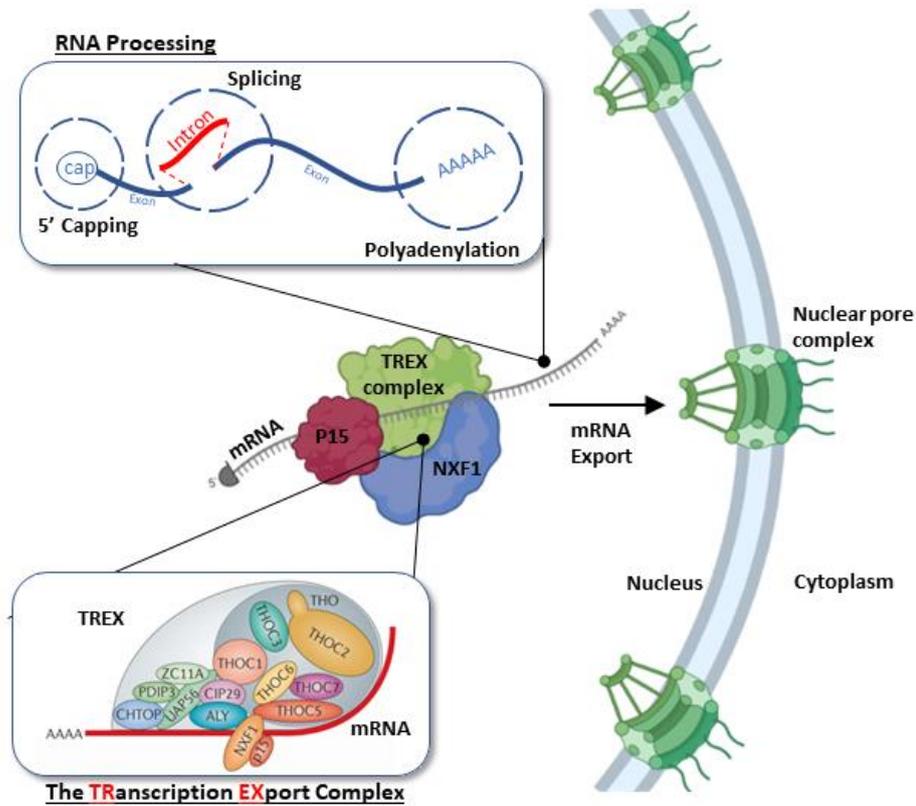


Figure 9: RNA Processing and the TREX Complex

Before export, a nascent mRNA undergoes multiple interconnected processing steps to become mature (5' capping, splicing and polyadenylation). Once successfully processed, exportation is directed via the export complex TREX (Transport Export complex). This complex along with transport factors NXF1 and p15 transport and chaperone mRNAs from the nucleus to the cytoplasm through the nuclear pore complex (NPC) (Pühringer et al., 2020) (Wickramasinghe and Laskey, 2015) (adapted using biorender).

1.6 The Influence of the TREX complex on Genome Instability

Increasingly, RNA processing factors have been found to be upregulated within various cancers suggesting RNA processing and mRNA export play a role in cancer development or progression (Obeng et al., 2019). Although the exact mechanisms of this are yet to be fully explored, every

aspect of RNA processing is potentially mutable, which could significantly impact the nascent transcript and thus gene expression.

A dysfunctional TREX complex may significantly advance the alteration of these processes, as it has been shown to assist the nascent transcript from its very transcription and interact with various RNA processing complexes.

The TREX complex has been shown to bind to CBC and the nuclear exosome targeting complex (NEXT) preventing mRNA degradation (Andersen et al., 2013) (Heath et al., 2016). This pathway is part of a key quality control mechanism whereby unfit transcripts are often degraded during RNA processing via nonsense-mediated mRNA decay (NMD), which inspects mRNAs for premature termination codons (PTCs). Interestingly, alterations to mRNA transcript decay through dysregulation of the TREX complex may offer a mechanism for cancer cells to exploit; as gene expression can be influenced independent of transcription and established transcriptional networks. This can be observed in BRCA1 in breast and ovarian cancers (Perrin-Vidoz et al., 2002), BRCA2 in ovarian cancers (Ware et al., 2006) and p53 in breast cancers (Anczuków et al., 2008). Associations between CBC and TREX also suggest an export route for intron-less mRNAs which could provide highly selective and rapid exportation (Nojima et al., 2007). As cancer cells often express mRNA isoforms with shorter 3' untranslated regions (UTRs) due to altered polyadenylation and alternative cleavage, the shorter mRNA isoforms exhibit increased stability during processing with some displaying a 100-fold difference in translational efficiency. Therefore, a rapid exportation route through upregulated TREX complexes could greatly influencing global gene expression (Mayr and Bartel, 2009), (Floor and Doudna, 2016).

TREX can also associating with the spliceosome, the exon junction complex (EJC), and key factors of poly(A) site choice (Dias et al., 2010) (Katahira et al., 2013). These associations functionally link the complex to each step of RNA processing in a co-transcriptional manner and may greatly influence gene expression (Reed and Cheng, 2005) (Nojima et al., 2007).

Defective RNA processing factors can also lead to increased replication stress and DNA damage, through the formation of RNA/DNA hybrid loops (R-loops). R-loops can be catastrophic to DNA replication and hinder gene expression as they prevent further DNA replication and often cause replication collisions. R-loops are formed when the nascent RNA treads back into the DNA duplex and displaces the non-template strand, leaving it as a ssDNA and exposed to degradation or mutation (Aguilera and García-Muse, 2012). mRNA processing and transcription factors can limit the formation of R-loops by chaperoning the nascent mRNA and preventing the detrimental hybridisation with exposed template DNA (Brickner et al., 2022). Fully formed R-loops can also be

corrected by RNA helicases such as DHX9 which dismantle the secondary structures that bind the RNA-DNA hybrids together (Chan et al., 2014). However, rectifying R-loops often results in an excess of DSB and SSB causing genetic information to be altered or lost, ultimately destabilising the genome. Such an increase in DNA damage can activate the DDR pathway, halting cell cycle progression and reducing gene expression. Therefore, the frequency and severity of R-loop formation can lead to genome instability and potentially, the development of cancer (Wickramasinghe and Venkitaraman, 2016). The ability of RNA processing factors to prevent and correct R-loops, again illustrates the importance of RNA processing factors in preventing DNA damage and limiting a cascade of errors which promote genome instability.

As well as indirectly assisting the DDR through R-loop reduction, RNA processing factors can selectively regulate the expression of DDR proteins, while DDR proteins also interact and regulate RNA processing factors. This suggests the possibility of feedback loops that regulate the transcription of these DDR proteins (Wickramasinghe and Venkitaraman, 2016). This may be highly selective, only affecting a small pool of genes involved in processes such as DNA replication, repair or the cell cycle. For example, DNA DSB repair is predominantly achieved via non-homologous end joining (NHEJ) and homologous recombination (HR) repair mechanisms. The TREX component ALY selectively exports transcripts encoding essential HR factors, while transcripts relating to NHEJ are unaffected (Wickramasinghe and Laskey, 2015). The ability to select one form of repair over another, highlights the regulatory control of selective export and its influence on genome stability. Selective export has also been shown to rapidly respond to cellular stress within yeast, which has proven critical for its stress recovery and survival. During heat stress, specific transcripts encoding heat shock proteins that aid thermotolerance and stress recovery, were selected for and exported via RNA export factors, while most other selective mRNAs were retained within the nucleus. This further illustrates the importance of RNA processing and export factors when responding to stress to ensure a sufficient number of DNA repair proteins are available if DNA damage is detected and consequently gene expression is reduced therefore preserving DNA integrity (Carmody et al., 2010) (Jackson and Bartek, 2009).

A rapidly growing body of evidence suggests that RNA processing factors also have a direct role in DNA repair, with RNA processing factors directly recruited and binding to DSB sites, while also interacting with DSB repair proteins and may aid in the re-unification of DNA DSBs (Zhang et al., 2009) (Ciccia and Elledge, 2010, Adamson et al., 2012) (Kim et al., 2013). This suggests that RNA processing factors may influence the DDR at multiple levels or maybe responsible for a cascade of DDR events which play a significant role in maintaining genome integrity.

1.6.1 CIP29

Originally identified due to its upregulation in hepatocellular carcinoma cells (Choong et al., 2001). The protein was subsequently isolated from the lysates of erythropoietin-stimulated UT-7/EPO cells as a 29kDa protein, which is often upregulated in highly proliferating cells such as adult bone marrow, fetal and malignant cells such as leukemic, liver and pancreatic cancer. CIP29 was also found to influence RNA maturation as an RNA exporter and respond to cytokines which may influence cellular growth regulation (Fukuda et al., 2002).

It is assumed that the proteins' possession of a bi-helical DNA binding motif, the SAP domain allows CIP29 to bind dsDNA and ssDNA, with a high affinity for ssDNA, through its RNA recognition motifs and therefore may influence these structures (Aravind and Koonin, 2000). The SAP domain is thought to be of great importance as SAP domains often flank functional genes, binding to matrix attachment regions of DNA and AT-rich residues located within intronic DNA. These proteins have also been found to influence transcription, DNA repair and replication stress (Hashii et al., 2004).

CIP29 has been linked to hnRNPs and binds to other RNA processing factors such as UAP56/BAT1/DDX39B and URH49/DDX39A, two very closely related DEAD-box helicases that facilitate mRNA biogenesis (Yamazaki et al., 2010) (Sugiura et al., 2007), (Leaw et al., 2004).

Genetic analysis in yeast suggests a functional involvement for the CIP29 orthologue, Tho1, in RNA processing, since Tho1 overexpression can partially suppress the loss of function of known RNA processing factors. In this system, RNA processing and export is controlled via a TREX complex containing a THO complex (Tho2, Hpr1, Mft1, and Thp2) which recruits Sub2/UAP56 and Yra1/ALY. In addition to nuclear export, the yeast TREX complex separates the nascent transcript from the DNA template and reduces R-loop formation. Mutations within THO components result in transcriptional impairment and mRNA export defects, however Tho1/CIP29 overexpression can partially suppress this. This recovery is lost with Sub2 mutants suggesting that Tho1/Sub2/Yra1 can provide partial THO function or offer alternative ways of RNA processing and export in the absence of a functional THO complex (Jimeno et al., 2006b).

Subsequently, CIP29 was identified as a human TREX complex component. Like its yeast counterpart, human TREX complex consists of a multi-subunit THO complex and additional components, however the exact composition is still debated (Dufu et al., 2010, Yamazaki et al., 2010, Fujita et al., 2020).

Initially, two contradictory studies reported differing interactions of CIP29 with the UAP56 and URH49 DEAD-box helicases. For example, one study suggested that CIP29 interacts with UAP56 and ALY, in an ATP-dependent manner to form a trimeric complex which regulated UAP56 activity, with CIP29 stimulating the helicase activity, while ALY stimulated the ATPase activity of UAP56 (Sugiura et

al., 2007), (Taniguchi and Ohno, 2008) (Dufu et al., 2010). To complete TREX assembly, the complex undergoes ATP-dependent remodelling with ALY and CIP29 relying on UAP56 to bind the THO complex. In a second study, CIP29 was shown to assemble with a second helicase URH49 to form an alternate mRNA export complex (AREX) rather than associating UAP56 or ALY to form the TREX complex (Yamazaki et al., 2010). However, both assumptions contradicted previous observations to some degree, for example, THOC2 was shown to associate to URH49 (Leaw et al., 2004) while UAP56, URH49, ALY and CIP29 have all been shown to associate with one another to some degree (Golovanov et al., 2006). Finally, it was suggested that UAP56 and URH49 have two complex formations. In the absence of ATP, UAP56 and the THO subcomplex form a precursor TREX complex, while URH49 associates with CIP29, to form a precursor AREX complex. With the addition of ATP, both complexes undergo restructuring to become highly similar complexes both featuring ALY, CIP29 and the THO subcomplex with either UAP56 or URH49. However, it is yet unclear if the two complexes share the exact same set of auxiliary components (Fujita et al., 2020). Overall, these findings suggest that CIP29 is linked to RNA export, possibly regulating gene expression in a bulk or selective manner depending on complex composition (Fujita et al., 2020, Yamazaki et al., 2010, Dufu et al., 2010). This is further supported as CIP29 knockdown through siRNA targeting, can reduce bulk mRNA export, with transcripts accumulating in the nucleus and resulting in mitotic defects, which although this effect was not significant, it suggested that CIP29 knockdown led to similar defects observed in URH49 knockdown, such as chromosome arm resolution and attachment defects which resulted in the failure of cytokinesis (Yamazaki et al., 2010).

In addition to its interactions with UAP56, URH49 and ALY, CIP29 has also been reported to associate with several other RNA processing factors, identified through mass spectrometry approaches, including FUS, DDX1, Matrin 3 and hnRNPU (Dufu et al., 2010). Interestingly, many of these processing factors have been identified as targets of DNA damage signalling (Gardiner et al., 2008) (Britton et al., 2009). Similarly, previous research using *X. laevis* egg extracts found that CIP29 is rapidly phosphorylated at a conserved serine residue, Ser95, in response to DNA DSBs by the DNA damage response kinase ATM (Holden et al., 2017). Phosphorylation of RNA processing factors may alter their activity, localisation or the interactions with other proteins or target mRNAs, thus impacting biological processes such as RNA biogenesis, gene expression and the maintenance of genome integrity.

1.7 Work leading up to this study.

Previous characterisation of CIP29 in human cells has relied on siRNA knockdown to ablate CIP29 function and examine the consequences for cell behaviour. However, siRNA knockdown can be highly variable only reducing rather than completely abolishing the expression of the gene of interest. Therefore, to overcome this limitation, a CIP29 knockout cell line was generated by CRISPR-Cas 9 gene editing, using a guide RNA targeting exon 2 of CIP29 to disrupt CIP29 expression within a transformed lung fibroblast cell line, MRC5-VI. Gene targeting was confirmed to induce frameshift mutations within CIP29 exon 2 of this knockout cell line, producing premature stop codons and abolishing CIP29 expression (Figure 10A). CIP29 gene disruption led to a significant reduction in cell proliferation. Upon flow cytometric analysis, a large number of cells accumulated in G2/M and a proportion of cells exhibited a greater than 4N DNA content (Armstrong, 2016) (Taylor, pers. comm.) (Figure 10B). These preliminary results were suggestive of a significant difficulty in cell cycle processes such as chromosome segregation and cytokinesis following CIP29 ablation (Armstrong, 2016). In addition, CIP29 was found to become phosphorylated in a cell cycle dependent manner, with CIP29 phosphorylation observed in response to late cell cycle arrest (G2/M), and in response to treatment with DNA DSB-causing agents. This was confirmed in both *X.laevis* egg and human cell extracts, thus suggesting that CIP29 function is likely to be regulated by phosphorylation, through the cell cycle and in response to DNA damage (Figure 10 C/D), (Taylor and Lindsay, pers.comm.). To identify proteins that interact with *X.laevis* CIP29, associated proteins were isolated by co-immunoprecipitation with XCIP29 and identified through mass spectrometry. The results indicated that XCIP29 interacts with some RNA processing and translation factors, although co-transcriptionally assembled protein complexes are commonly under-represented in *X.laevis* egg extracts, that are largely transcriptionally inactive due to the abundance of maternal mRNA stores. Interestingly, several proteins involved in mitosis were also identified as possible CIP29 interacting factors, including KIF2C/MCAK, a microtubule depolymerase responsible for chromosome-kinetochore fibre attachment which plays a significant role in chromosome segregation (Holden and Lindsay, Unpublished data).

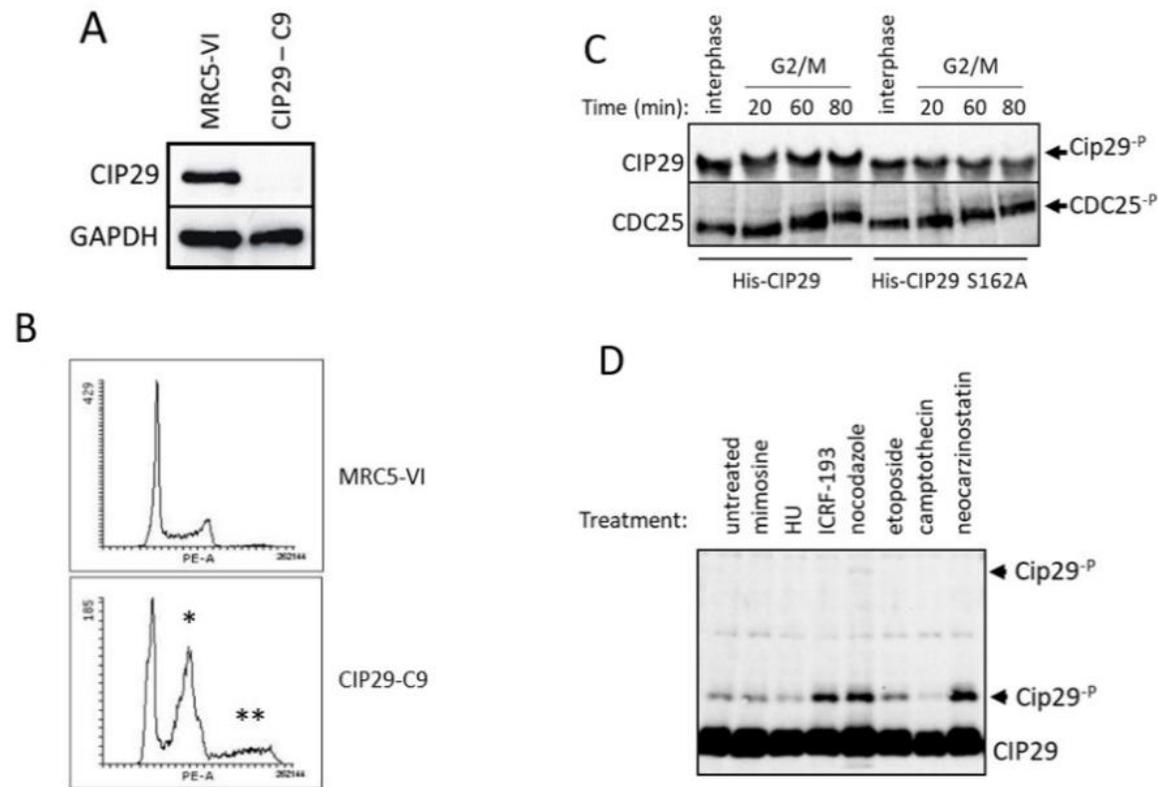


Figure 10: Previous data: CIP29 is required for successful cell cycle progression and is phosphorylated during mitosis and in response to DNA damage.

Western blot displaying CIP29 expression levels in parental MRC5-VI cells and CIP29 knockout (CIP29-C9) cells. GAPDH is used as a loading control (A).

DNA content of MRC5-VI and CIP29-C9 cells, determined by PI staining. When compared to the wildtype MRC5VI cells, CIP29-C9 cells accumulate in G2/M (*), with a proportion of cells displaying a greater than 4N DNA content (**). Recombinant wildtype His-CIP29 undergoes phosphorylation in mitotic egg extracts but mutation of Ser162 abolishes this modification. High CDK1 activity, associated with mitosis, can be inferred via CDC25 phosphorylation (C). Phos-tag gel analysis reveals increased phosphorylation when cells are blocked in G2 (ICRF-193) or mitosis (nocodazole) or after DNA damage (etoposide, neocarzinostatin). Phosphorylation is reduced with transcription blocking (camptothecin) (D).

1.8 Aims and Objectives

The purpose of this project is to build on the foundation of previous published and unpublished work to further characterise CIP29. The project aims to explore roles for CIP29 in cell division control and genome integrity maintenance. Such additional insight will provide mechanistic detail as to how this RNA processing and mRNA export factor influences cell proliferation and how it may influence cancer development through a range of strategies.

1.8.1 *Project aims.*

- Investigating CIP29 protein-protein interactions
- Explore the phenotype caused by the loss of CIP29.
- Obtain a greater understanding of the molecular structure of the CIP29 protein and its functional domains.

1.8.2 *Project objectives*

Chapter 1: To validate novel CIP29 protein-protein interactions.

At the time of initiating this research there was some ambiguity about the association between CIP29 with various core TREX complex components which required further clarification. While previous data from this laboratory suggested that the removal of CIP29 profoundly impacted the cell cycle suggesting a more direct role, rather than solely influencing gene expression indirectly through the TREX complex. Furthermore, interactions between XCIP29 and proteins known to aid cell cycle progression during mitosis and cytokinesis remained to be validated.

- The true association between CIP29 and UAP56 or ALY will be investigated through co-immunoprecipitation studies.
- It will also be determined if CIP29 associations are ATP dependant, thus distinguishing its involvement within the various TREX complex variations such as the Apo-TREX complex, the Apo-AREX complex or the fully assembled forms of these complexes.
- Investigate if the associations highlighted through mass spectrometry of XCIP29 are conserved in human CIP29 through co-immunoprecipitation assays.

Chapter 2: Explore the phenotype caused by the loss of CIP29.

The effects of CIP29 ablation and its reintroduction will be examined, clarifying if the loss of CIP29 is responsible for the previously observed cell cycle defects. To achieve this, various modified cell

types, MRC5.VI, CIP29 knockout (CIP29-KO) and CIP29-rescue (CIP29-KO with CIP29 genes reintroduced through stable transfection) will be utilised.

- Each cell line will be further characterised, with the removal and reintroduction of CIP29 confirmed through western blotting and functional localisation studies.
- The cell cycle profile of each cell line will also be examined through flow cytometry and Propidium Iodide (PI) staining to confirm that the cell cycle defects seen in the CIP29-KO cell line are as a consequence of CIP29 aberration.
- Proliferation assays will be conducted on the modified cell types described above, in order to examine the effect of reintroduced CIP29 on cell's ability to successfully complete the cell cycle and proliferate.
- The presence of DNA damage within the previously described CIP29 modified cell lines will be investigated through the presence and activation of key DNA damage markers via western blot.
- As the previously used Flow cytometry and Propidium Iodide (PI) staining fails to completely distinguish between S/G2/M phases of the cell cycle, cell populations will be further separated into these phases of interest through additional staining techniques.
- Immunofluorescence staining will be utilised to investigate the presence of morphological defects that are indicative of mitotic defects.

Chapter 3: Obtain a greater understanding of the molecular structure of the CIP29 protein and its functional domains.

To further identify which regions and domains within the CIP29 protein hold functional significance, truncations of the protein will be developed, and their cell cycle profiles, morphology and localisation will be assessed and compared to the previously described cell lines of interest.

- Stable truncations of the CIP29 protein will be developed.
- The cell cycle profile of each CIP29-truncation will be examined through flow cytometry and Propidium Iodide (PI) staining to assess if certain truncations share a similar cell defect to the CIP29- KNOCKOUT.
- Immunofluorescence staining will be utilised to investigate the presence of morphological defects that are indicative of mitotic defects.
- The truncated proteins' functional ability will be assessed through localisation studies.

2 [Materials and methods](#)

2.1 [Materials](#)

Table 1: Primary antibodies used in this work.

Primary Antibodies				
Antibody	Origin	Supplier and cat. No.	Working Conc.	Target Protein Mol. Wt. (kDa)
α -ALY	Rabbit	Cell Signalling Technology (12655s)	1:1000	27
α -anillin	Mouse	Insight Biotechnology (ab211872)	1:1000	123
α -CIP29	Rabbit	Raised In House	1:1000	24
α -GAPDH	Mouse	Upstate Biotechnology, Inc (NM-002046)	1:1000	36
α -KIF2C	Mouse	Insight Biotechnology (JB181-1G2)	1:1000	81
α -MAD1	Rabbit	Insight Biotechnology (NP-001013858)	1:500	83
α -UAP56	Rabbit	Cell Signalling Technology (47258s)	1:1000	49
α -NPL4	Mouse	Santa Cruz Biotechnology (SC-365796)	1:1000	69
α -ARP3	Mouse	Santa Cruz Biotechnology (SC-48344)	1:1000	47
α -DNM2	Mouse	Santa Cruz Biotechnology (SC-166669)	1:1000	98
α -NONO/p54	Mouse	Santa Cruz Biotechnology (SC-376804)	1:1000	53
α -ARS2	Mouse	Santa Cruz Biotechnology (SC-376716)	1:1000	98
α - γ -tubulin	Mouse	Abcam (AB133)	1:1000	55
α -phospho Histone H3 (pS10)	Rabbit	Sigma-Aldrich (AB422)	1:100	17
α -CDT1/DUP antibody	Rabbit	Abcam (EPR17891)	1:100	60

Table 2: Secondary antibodies used in this work.

Secondary antibodies			
Antibody	Origin	Supplier and cat. No.	working Concentration
Rabbit anti-mouse immunoglobulins /HRP	Rabbit	Dako (P0260)	1:1000
Goat anti-rabbit immunoglobulins /HRP	Goat	Dako (P0448)	1:1000
Goat anti rabbit Alexa Fluor 647 secondary Antibody	Goat	Thermo Fisher Scientific (A21245)	1:100/ 1:200/1:500
Goat anti-Mouse IgG (H+L) Secondary Antibody, Alexa Fluor 647	Goat	Thermo Fisher Scientific (A21235)	1:500
Goat anti-Mouse IgG (H+L) Secondary Antibody, Alexa Fluor 568	Goat	Thermo Fisher Scientific (A11031)	1:500

Table 3: Primers designed for this work.

Primer	Primer sequence (5'-3')
EP82	ATC TAG AAT GGA CTA CAA GGA CGA CGA TGA CAA GC
EP83	ATC TAG AAT GGG CTA CCC TTA CGA CGT TCC AG
EP84	AAC CAG TTG AAT CGA CAT CTG
EP85	GTT TAT TTT CCA AGT TAG AAA CTT C
K20/910	TAG TCG ACT CAC TGG GGC CGT TTC TTG C
622	TTG GAT CCT CAG GCA ATC CCA AAG CGC TCT GC
EP135 (CIP29-truncation (R) 1-56)	ATT AAT TAA CTA CAG TAC ATC TTC TTC ATT TGC CTC
EP136 (CIP29-truncation (R) 1-93)	ATT AAT TAA CTA CAG TAC ATC TTC TTC ATT TGC CTC
EP137 CIP29- truncation (R) 1-129)	ATT AAT TAA CTA AAT CCC AAA CCT AGC TGC CCG
EP138 CIP29- truncation (R) 1-182)	ATT AAT TAA CTA AAT CCC AAA CCT AGC TGC CCG
EP139 (CIP29- (R) FL)	ATT AAT TAA CTA CCC AAA TCG CTC CTT CCT CTT TTTC
EP40 (CIP29- truncation (F) DS)	ACT CGA GGT ACT GGG AGA TGA AAC AGA GG
HDL#623 (hc29F Xhol)	AAC TCG AGA TGG CGA CCG AGA CGG TGG AGC TCC

2.2 Molecular biology methods

2.2.1 *Restriction Endonuclease Digestion*

To generate DNA fragments for cloning, preparative-scale restriction digests were prepared in a total volume of 40 μ l, containing 4 μ g of vector DNA or 7 μ g of insert DNA alongside appropriate restriction enzymes (15 units of enzyme) and restriction enzyme buffer (NEB). Digests were incubated overnight (37°C), and successful digestion was confirmed through gel electrophoresis. To confirm successful cloning, 20 μ l diagnostic digests were conducted containing 2 μ l of plasmid DNA, 2 μ l of appropriate restriction buffer, 0.3 μ l RNAase (200 μ g/ml) and 6 units of appropriate restriction enzymes. Digests were incubated overnight (37°C).

2.2.2 *Gel Electrophoresis*

Each digest was diluted with H₂O and 6x loading buffer (30% glycerol, 0.25% bromophenol blue, 0.25% xylene cyanol FF) before loading on to an agarose gel (0.8% routinely, or 1.5% for smaller DNA fragments) prepared with 1 x TBE (0.1M Tris-base, 0.1M Boric acid, 2mM EDTA), and run at 80V for 1 hour. Gels were stained for 1 hour with Sybr safe[®] stain (50ml 1xTBE and 2 μ l Sybr safe[®]) and visualized on a UV transilluminator (wavelength VU-C).

2.2.3 *QIAquick Gel Extraction of DNA fragment*

DNA fragments were excised under UV light and processed in accordance with the recommended QIAquick[®] Gel Extraction Kit protocol (QIAGEN[®]).

2.2.4 *Purification of PCR products using the QIAquick[®] PCR Purification Kit*

PCR amplified DNA fragments were purified in accordance with QIAquick (QIAGEN[®]) PCR purification kit instructions. PCR product purity was confirmed through agarose gel electrophoresis.

2.2.5 *Ligation into Pepex, or PCI-NEO expression vectors*

DNA inserts were ligated into Pepex or pCI-NEO expression vectors in accordance with NEB Quick Ligation Protocol recommendations. 10 μ l ligation mixture (5 μ l 2x Quick ligase buffer (NEB[®]), 0.5 μ l of T4 DNA Ligase (Invitrogen™) combined with 0.5 μ l diluted vector and 4 μ l insert DNA) were incubated at room temperature for 5 minutes.

2.2.6 Polymerase chain reaction

DNA inserts were amplified through the polymerase chain reaction (PCR) using specific primers. Reaction mixes and PCR parameters are as follows in table 4 and 5. Successful amplification was confirmed through agarose gel electrophoresis.

Table 4: PCR parameters and program details for generating HA-KIF2C, FLAG-KIF2C and HA-CIP29 DNA fragments.

PCR reagents					
Reagent		Volume (μ l)			
Q5x buffer		5			
10mM dNTPs		1			
Q5 DNA polymerase		0.5			
dH2O		32.5			
Insert DNA (0.5ng/ μ l)		0.3			
PCR Primers for specific DNA inserts (μ l) (50 μ M)/ (100 μ M)*					
KIF2C-HA		KIF2C-FLAG		CIP29-HA	
Primer EP83	0.5	Primer EP82	0.5	Primer EP83	0.5
Primer 910	0.25*	Primer 910	0.25*	Primer 622	0.25*
PCR program for specific DNA inserts					
Profile	Temperature ($^{\circ}$ C)	Time			
Initialization	98 $^{\circ}$ C	30s			
Denaturation	98 $^{\circ}$ C	10s	30 cycles		
Annealing	55 $^{\circ}$ C	30s			
Elongation	72 $^{\circ}$ C	1min			
Final elongation	72 $^{\circ}$ C	2mins			

Table 5: Generating CIP29 truncation mutants through PCR: 3' truncations.

PCR reagents									
Reagent					Volume (μl)				
Q5x buffer					5				
10mM dNTPs					1				
Q5 PC1					0.5				
dH2O					32.5				
Insert DNA					1ng/μl				
PCR Primers for specific DNA inserts (μl) (50μM)/ (100μM)*									
CIP29 truncation 1-56		CIP29 truncation 1-93		CIP29 truncation 1-129		CIP29 truncation 1-182		CIP29 Full length	
Primer EP135	0.5	Primer EP136	0.5	Primer EP137	0.5	Primer EP138	0.5	EP139	0.5
Primer HDL#623	0.25*	Primer HDL#623	0.25*	Primer HDL#623	0.25*	Primer HDL#623	0.25*	Primer HDL#623	0.25*
PCR program for specific DNA inserts (please note the two annealing temperature alternatives depending on the primers used).									
Profile		Temperature (°C)				Time			
initialization		98°C				30s			
Denaturation		98°C				10s (30 cycles)			
Annealing		52°C				30s (30 cycles)			
Elongation		72°C				1min (30 cycles)			
Final elongation		72°C				2mins			

Table 6: Generating CIP29 truncation mutants through PCR: SAP domain deletion (Δ SAP).

PCR Primers for specific DNA inserts (μl) (50μM)		
CIP29 truncation DS		
Primer EP140		0.5
Primer EP139		0.5
PCR program for specific DNA inserts		
Profile	Temperature ($^{\circ}$C)	Time
initialization	98 $^{\circ}$ C	30s
Denaturation	98 $^{\circ}$ C	10s (30 cycles)
Annealing	56$^{\circ}$C	30s (30 cycles)
Elongation	72 $^{\circ}$ C	1min (30 cycles)
Final elongation	72 $^{\circ}$ C	2mins

2.2.7 *A-tailing of PCR products*

10 µl total reactions (1 µg purified PCR product, 1 µl 10x TAQ buffer, 1 µl TAQ polymerase (NEB®) and 1 µl 2 mM dATP) were incubated for 2 min (70°C).

2.2.8 *Ligation into pGEM®-T Easy vector*

10 µl ligation reactions (5µl 2x rapid ligation buffer, 4µl A-tailed PCR product, 0.5µl of pGEM®-T Easy vector (Promega®) and 0.5µl T4 DNA ligase (Promega®) were incubated at room temperature overnight.

2.2.9 *Transformation of plasmids into ultra-competent E. coli cells*

5 µl of ligation reaction was added to 50 µl ultra-competent *E. coli* cells, gently mixed and incubated on ice for 30 min. Cells were heat-shocked at 42°C for 45s and incubated on ice for a further 2 min. Cells were plated on agar plates containing 100 µg/ml ampicillin selection and incubated overnight at 37°C. When required (e.g., pGEM-T Easy cloning) colonies were selected using ampicillin and X-gal/IPTG for blue/white selection (X-gal (50 µg/ml) and IPTG (100 mM)).

2.2.10 *Alkaline lysis miniprep method*

Individual colonies were grown overnight in 2.5 ml LB medium + Carbenicillin (100 mg/ml), with shaking at 37°C. Cells were harvested and pelleted (13,000rpm, 30s (Heraeus Fresco 17) before resuspension in 100 µl of MP1 (50 mM glucose, 25 mM Tris pH8, 10 mM EDTA), followed by 200 µl of MP2 (200 mM NaOH, 1% SDS), and finally 150 µl of MP3 (3 M KOAc, 2 M HOAc). After mixing, samples were centrifuged (13,000 rpm, 30s) and the supernatant collected. 1 ml of 100% ethanol was added, mixed, and centrifuged (13,000 rpm, 10min). The resulting DNA pellet was resuspended in 30-50 µl of water.

2.2.11 *QIAGEN® Midiprep method*

Cell pellets from 100 ml LB medium + Carbenicillin were harvested (5000 xg for 20 min at 4°C) and DNA purified in accordance with QIAGEN® Midiprep method, following the Qiagen low copy number protocol.

2.2.12 *DNA sequencing*

Plasmid DNA (100 ng/µl) was sequenced by Source BioScience, utilising their in-house primers or providing our own primers (3.2pmol/µl).

2.3 Protein Methods:

2.3.1 *Preparation of whole cell extracts*

For SDS-PAGE analysis, frozen cell pellets or *in situ* adherent cells were routinely resuspended in 50 mM Tris pH 7.5, 150 mM NaCl, and then lysed by the addition of an equal volume of lysis buffer (50 mM Tris pH 7.5, 150 mM NaCl, 0.1% SDS, 1 µl/ml Base muncher (Expedeon), Protease Inhibitor Cocktail (Sigma-Aldrich®) at 1/100).

When preserving protein interactions for co-immunoprecipitation assays an alternative lysis buffer was used: (50 mM Tris pH 7.5, 150 mM NaCl, 1% NP40, 1µl/ml Base muncher (Expedeon)). To preserve protein phosphorylation, phosphatase inhibitors (10 mM NaF, 10 mM β-glycerophosphate and 2 mM Na₂VO₃) were included in the resuspension buffer as required. Cells were lysed on ice (10 mins), before extracts were clarified through centrifugation (13,000rpm, 3 min, 4°C (Heraeus Fresco 17)).

2.3.2 *Preparation of HeLa nuclear cell extracts*

Frozen HeLa nuclear extracts were resuspended in 20 mM HEPES, 100 mM KCl, 50ng/µl RNase, Protease Inhibitor Cocktail (Sigma-Aldrich®) at 1/100, and lysed by addition of an equal volume of lysis buffer(50 mM Tris pH 7.5, 150 mM NaCl, 1% NP40, 1µl/ml Base muncher (Expedeon)) (10 min, ice), before clarification through centrifugation (13,000rpm, 3 min, 4°C, Heraeus Fresco 17). When required, preparation of HeLa nuclear extracts was modified in accordance with Dufu et al (2010) or Fujita et al, (2020), as described in section methods 2.4.

2.3.3 *Direct cell lysis for monitoring DNA damage response (DDR) activation*

5x10⁵ MRC5-VI control cells were grown in the presence of etoposide (0.1µg/ml, 24 h) alongside cells of interest without treatment. Cells were harvested and cell pellets were washed once with PBS and resuspended in 100 µl of ice-cold resuspension buffer (50 mM Tris pH 7.5, 150 mM NaCl, 10 mM NaF, 10 mM β-glycerophosphate, 2 mM Na₂VO₃ and Protease Inhibitor Cocktail at 1/100 (Sigma-Aldrich®).

A 20 µl sample was lysed at room temperature (0.1% SDS and Basemuncher (0.1 µl/ml, 10 mins)) and used for protein concentration analysis via a Bradford assay (refer to methods 2.3.4), while 0.25 volumes of 5x SDS -PAGE loading buffer was added to the remaining cell suspension, which was mixed then boiled (3 mins, 95°C). Samples were incubated with Basemuncher (Expedeon), (0.2

µl/ml, 10 mins, room temperature) before further boiling (3mins, 95°C) before analysing through SDS-PAGE.

2.3.4 Cell fractionation assays

Cell fractionation assays were conducted using the Subcellular Protein Fractionation kit for Cultured Cells (Cat. 78840, Thermo Scientific™) in accordance with the manufacturer's protocol.

2.3.5 Bradford assay

Protein solution/cell extract (2 µl) was mixed with 600 µl Protein Assay Dye Reagent (Expedeon) in a cuvette and absorbance measured using a spectrophotometer at OD600. Protein concentrations were estimated relative to a buffer-only control sample.

2.3.6 SDS-PAGE gel production

Polyacrylamide gels were constructed with an 10-14% resolving gel and overlaid with a 4% stacking gel (see table 6). Samples were combined with 5x loading buffer (100mM b-mercaptoethanol, 0.25% bromophenol blue 0.5 M dithiothreitol, 50% glycerol, 10% SDS) and boiled (3 min) before loading. Proteins were separated by electrophoresis (voltage 120v, 120 mins) in a 1x SDS-PAGE running buffer (24.8mM Tris-base, 192mM glycine, 0.1% SDS).

Table 7: Composition of resolving and stacking gels used for SDS-PAGE Polyacrylamide gel electrophoresis.

Resolving gel		
10%	10 ml (Volumes (ml))	15 ml (Volumes (ml))
ddH ₂ O	4.0	5.9
30% acrylamide mix	3.3	5
1.5 M Tris (pH 8.8)	2.5	3.8
10% SDS	0.1	0.15
10% APS	0.1	0.15
TEMED	0.004	0.006
11%	10 ml (Volumes (ml))	15 ml (Volumes (ml))
ddH ₂ O	3.7	5.4
30% acrylamide mix	3.6	5.5
1.5 M Tris (pH 8.8)	2.5	3.8
10% SDS	0.1	0.15
10% APS	0.1	0.15
TEMED	0.004	0.006
14%	10 ml (Volumes (ml))	15 ml (Volumes (ml))
ddH ₂ O	2.7	3.9
30% acrylamide mix	4.6	7
1.5 M Tris (pH 8.8)	2.5	3.8
10% SDS	0.1	0.15
10% APS	0.1	0.15
TEMED	0.004	0.006
Stacking gel (3mls)		
Reagent	Volumes (ml)	
ddH ₂ O	2.1	
30% acrylamide mix	0.5	
1.5 M Tris (pH 6.8)	0.38	
10% SDS	0.03	
10% APS	0.03	
TEMED	0.003	

2.3.7 Western blot semi-dry transfer

Separated proteins were transferred to nitrocellulose by semi-dry transfer using transfer buffer (48 mM Tris-base, 39 mM Glycine, 20% methanol) at 20V for 55 mins.

2.3.8 Antibody Staining of nitrocellulose membrane

Nitrocellulose membranes were blocked in blocking buffer (5% powdered milk in PBS plus 0.1% Tween (PBS-T)) for 1 h at room temperature. Membranes were incubated in the chosen primary antibody overnight, 4°C with gentle rotation. Membranes were washed (3x 10 min in PBS-T), before incubation with secondary antibody in blocking buffer (1 h, room temperature) with gentle rotation, followed by further washing (3x 10 min in PBS-T).

2.3.9 Developing Western Blots by Enhanced Chemiluminescence (ECL)

Nitrocellulose membrane was incubated with ECL reagent (20 ml of 100 mM Tris pH 8.0 with 6 µl of hydrogen peroxide (30% solution), 50 µl of 90mM p-coumaric acid, and 100 µl of 250 mM luminol) before immediate imaging using a Bio-Rad Chemidoc XRS+ with an ultraviolet light (UV-B) luminescence (wavelength: 280-315 nm).

2.3.10 Stripping Western Blots

Under Mild Stripping Conditions: The nitrocellulose membrane was incubated twice in mild stripping buffer (200 mM glycine pH2, 0.1% SDS) for 5 mins, at room temperature. After stripping, the membrane was rinsed twice with PBS-T, and washed (3x 10 min in PBS-T).

2.3.11 *In vitro translation (IVT)*

For *in vitro* expression using a coupled transcription/translation reaction (Promega), plasmid DNA (0.5µg/µl) encoding CIP29 or KIF2C was incubated with 9µl TNT master mix (40µl TnT® T7 PCR Quick Master Mix, 1mM methionine and 7µl nuclease-free water) for 120 mins, 30°C. Translation products were diluted to the equivalent to 1 µl of IVT reaction (8 µl 5x SDS loading buffer, 24µl ddH₂O and loading 5µl per lane) and visualised through SDS-PAGE and western blotting (refer to methods 2.3.6-9). For detection of protein synthesis, a Transcend™ Non-Radioactive Translation Detection System was also utilised within the IVT reaction. *Please see table 7 for further details.*

For co-translation, plasmids were incubated in a 24µl of the IVT system. (40µl TnT® T7 PCR Quick Master Mix, 1mM methionine and 7µl nuclease-free water) for 90 mins, 30°C. A sample was taken for SDS-PAGE and western blotting analysis, while the remaining was diluted with PBS and utilised for an immunoprecipitation assay utilising CIP29 bound A- Sepharose beads (~1 in 7).

Immunoprecipitation assays were conducted as described (Methods 2.4). All SDS-PAGE samples were loaded equivalent to 0.5 µl of IVT reaction alongside the equivalent amount for the immunoprecipitation samples.

Table 8: DNA constructs investigated for protein synthesis using the *In vitro* translation (IVT)

DNA constructs investigated for protein synthesis using the <i>In vitro</i> translation (IVT)			
DNA constructs	Volume of DNA per IVT reaction	incubation period	Transcend™ Non-Radioactive Translation Detection System
CIP29-NT-PCI-NEO	0.2 µg/µl	90 mins, 30°C	NO
KIF2C-NT-PCI-NEO	0.2 µg/µl	90 mins, 30°C	NO
FLAG-CIP29 truncation 1-56-PCI- NEO	0.5 µg/µl	120 mins, 30°C	YES
FLAG -CIP29 truncation 1-93-PCI- NEO	0.5 µg/µl	120 mins, 30°C	YES
FLAG -CIP29 truncation 1-129-PCI- NEO	0.5 µg/µl	120 mins, 30°C	YES
FLAG-CIP29 truncation1-182-PCI- NEO	0.5 µg/µl	120 mins, 30°C	YES
FLAG-CIP29 Full length-PCI- NEO	0.5 µg/µl	120 mins, 30°C	YES
FLAG-CIP29 truncation DS-PCI- NEO	0.5 µg/µl	120 mins, 30°C	YES

2.3.12 *Affinity column purification of antibodies*

To remove residual azide, the CIP29 antigen column (previously prepared in this laboratory) was drained and washed with 10 column volumes of PBS. Serum was applied to the column and mixed by

inverting (O/N, 4°C). Serum flow-through was discarded and the column was washed four times with PBS, once with PBS + 500mM NaCl, before a final wash in PBS. Bound antibodies were eluted in 200mM glycine pH 2, with 1ml fractions collected into 1.5 M Tris pH 8.8 at approximately 1/10th the elution volume. Finally, the Protein (IgG) concentration in each eluate fraction was estimated using Bradford assay.

2.3.13 Crosslinking antibodies to protein A Sepharose beads

Protein A Sepharose beads were washed in PBS (3x) and incubated with the antibody overnight with rotary motion (4°C). Antibody-bound beads were washed (3x) in 0.2 M sodium borate, pH9) and a pre-crosslinking sample taken for SDS-PAGE analysis. Solid dimethylpimelimidate (DMP) was added to the remaining beads to a final concentration of 40 mM (103 mg/10 ml) and mixed for 1 hour with inversion (room temperature). A post-crosslink sample was taken for SDS-PAGE analysis. The remaining antibody-beads were washed (2x) then mixed for 2 h in 1.5 M Tris pH8.8 (room temperature) before washing in PBS (3x), (200 mM glycine pH2 (2x) and PBS (4x). Antibody beads were stored at 4°C in PBS containing sodium azide (0.05%).

2.3.14 Crosslinking antibodies to magnetic protein A/G Dynabeads

Antibody was diluted in PBS/0.02% Tween-20 (10 µg/400 µl), combined with 1.5 mg of protein A or protein G Dynabeads (~50 µl bead slurry) and incubated with rotation for 20mins, room temperature. Waste supernatant was removed and the antibody-bound Dynabeads was resuspended and washed in PBS/0.02% Tween-20 followed by 0.2 M sodium borate pH9 (2x). To crosslink, antibody bound Dynabeads were incubated in 0.2 M sodium borate pH9 containing DMP (10 mg/ml) for 1 hour with rotation, room temperature. Finally, the antibody-crosslinked Dynabeads were washed with PBS/0.02% Tween-20 (2x) and resuspended in PBS/0.02% Tween-20 + sodium azide (0.05%) and stored at 4°C.

2.4 Immunoprecipitation

2.4.1 Standard Immunoprecipitation (IP) protocol

Protein extracts were precleared by incubation with PBS-washed protein A- Sepharose beads, for 20mins with rotation, 4°C. Approximately 5µg of pre-cleared protein extract were incubated with washed and equilibrated antibody-crosslinked protein A- Sepharose beads (~25µl/IP, in a reaction volume of 200µl, for 90mins with rotation, 4°C. Following centrifugation (30s, 4000rpm, centrifuge), unbound protein samples were collected for SDS-PAGE analysis and antibody beads were washed in PBS + 0.1% NP40 (x3). Immunoprecipitated proteins were eluted by a 30sec incubation in 50µl 200

mM glycine, pH2 followed by centrifugation, as before. The eluted supernatant was neutralised by addition of 6µl 1.5 M Tris 8.8pH and any residual beads were removed by passing through a spin column (sigma, SC1000) (30secs/13000rpm) before being subjected to SDS-PAGE analysis.

2.4.2 Immunoprecipitation protocol adapted from Dufu et al (2010)

Equal measures of HeLa nuclear extract (300µl) and splicing reaction buffer ((20 mM HEPES at pH 7.9, 100 mM KCl), 500 µM ATP, 3.2 mM MgCl₂, 20 mM creatine phosphate, and 50 ng/µl RNase A) were incubated for 30 min at 30°C. 150µl of IP buffer (1× PBS, 0.1% Triton, 0.2 mM, Protease Inhibitor Cocktail (Sigma-Aldrich®) at 1/100) was added and the resulting sample was firstly pre-cleared with PBS-washed protein A- Sepharose beads, for 20mins with rotation, 4°C.

Once pre-cleared the samples were combined with 15µl of antibody-cross-linked A- Sepharose beads for 90 mins. Unbound protein samples were collected for SDS PAGE analysis.

Beads were washed (4x PBS, 0.1% Triton, 0.2 mM, Protease Inhibitor Cocktail (Sigma-Aldrich®)) and centrifuged (30secs/4000rpm). To elute the sample beads were incubated in 50µl 200 mM glycine, pH2 for 30s followed by centrifugation, as before. The eluted supernatant was neutralised by addition of 6µl 1.5 M Tris 8.8pH and any residual beads were removed by passing through a spin column (sigma, SC1000) (30secs/13000rpm) before being subjected to SDS-PAGE analysis.

2.4.3 Immunoprecipitation protocol adapted from Fujita et al., (2020)

HeLa nuclear extract was diluted with 1/4th the volume dilution buffer (1× PBS, 0.1% Triton, 0.2 mM, Protease Inhibitor Cocktail (Sigma-Aldrich®) at 1/100 and 0.5mM DTT) and incubated for 20mins, 30°C. Additional supplements were then added to both ATP+ and ATP- samples. ATP+ samples were supplemented with 500 µM ATP, 3.2 mM MgCl₂, 20 mM creatine phosphate, and 100 ng/µl RNase A while, ATP- samples contained 100 ng/µl RNase A. Samples were incubated for 30mins, 30°C, centrifuged (30s/ 13,000 rpm (Heraeus Fresco 17)) and the supernatant collected. Supernatant then pre-cleared by exposure to PBS-washed protein A- Sepharose beads, for 20mins with rotation, 4°C. before incubation with 15µl of antibody-cross-linked beads for 90 mins with rotation, 4°C).

Flow through from each extract (5 µl) was collected, for SDS PAGE analysis. Beads were washed 4x (PBS, 0.1% Triton, 0.2 mM, Protease Inhibitor Cocktail (Sigma-Aldrich®) at 1/100 and 0.5mM DTT) and centrifuged (30 s/4000 rpm (Heraeus Fresco 17)). To elute, samples were dissolved 50µl of elution buffer (250mM Tris-HCL Ph6.8, 1% SDS, 0.002% bromophenol blue, 40% glycerol) and incubated for 10mins at 37°C. Samples were boiled for 2 mins and beads were removed using a spin column (30secs/13000rpm). DTT was then added to a final concentration of 10mM.

2.5 Cell Culture Techniques

2.5.1 *Cell culture conditions*

MRC5-V1 cells and their derivatives, were routinely grown in DMEM medium (Lonza) containing 10% foetal bovine serum (Hyclone™) at 37°C in 5% CO₂.

For routine subculture, spent medium was removed and cells washed in PBS. Adherent cells were incubated in 0.5% trypsin solution (5% trypsin (Gibco™) at 1 in 10 in PBS). Detached cells were centrifuged (300 rpm for 5 mins) through an equal volume of medium. The cell pellet was resuspended in growth medium, and the resulting cell suspension used as required.

2.5.2 *Freezing and thawing cell stocks*

Cells (~1-2 x 10⁶/ml) were frozen in cryovials using medium containing 10% DMSO. Cryovials were frozen through a freezing vessel overnight, -80°C and transferred to liquid nitrogen for long-term storage. When required, cryovials were removed from liquid nitrogen storage and cells were quickly thawed via a water bath at 37°C. Once thawed, the cells were transferred to an appropriate flask and the media changed the following day.

2.5.3 *Media supplements and additions*

Etoposide cytotoxic media: to stimulate DNA damage, cell lines were grown in medium containing 0.1µg/ml of etoposide for 24 hours.

Colcemid treatment: to promote a mitotic block, cell lines were grown in a medium containing 1µg/ml Colcemid for 24 hours.

G418 selection media: to select for cells harbouring a neomycin-resistance gene, cells were grown in G418 Selection media; initial selection for stable transfectants and the development of novel cell lines was in medium containing 600 µg/ml G418, while 250 µg/ml G418 was used for selection maintenance thereafter.

2.5.4 *Transient transfection using FuGENE® HD Transfection*

Transient transfections were conducted in accordance with FuGENE® HD Transfection protocols using ~4x 10⁵ cells in a 6-well plate at approximately 80% confluency. Routinely, 180µl of serum-free antibiotic-free medium and 3.6µg of DNA of interest were combined with Fugene transfection reagent at a ratio of ~1:5 DNA/ transfection reagent. Once mixed the reaction was incubated at room temperature (10 mins). The transfection mixture was added dropwise to cells and the

transfection was left to incubate with the cells for 24 or 48 hours before collection and further analysis.

2.5.5 Stable transfection using FuGENE® HD Transfection

Transfections were conducted as previously described (Methods 2.5.4) except that cells were grown on 24 well plates with an approximate seeding density of 1.5×10^4 cells/well. Reagent volume was reduced to accommodate this (30 μ l of serum-free medium, 0.6 μ g DNA and 1.8 μ l Fugene transfection reagent (ratio of ~1:3 DNA/ transfection reagent). Cells remained in culture until confluent (~48/96 hours). Once confluent, the cells were resuspended and diluted either 1/10 or 1/5 and transferred to 10cm dishes. Cells were then incubated with G418 selection (600 μ g/ml) for 2 weeks.

2.5.6 Isolation of stable clones

Medium was removed, and the cells were gently washed in PBS. Single colonies were isolated using sterile cloning rings dipped in sterile silicone grease and incubated with 0.5% trypsin (10 μ l) to encourage detachment. Finally, each colony was transferred to a well of a 24 well plate containing cell culture medium with appropriate selection, for further expansion.

2.5.7 Cell proliferation assays

Cell proliferation assays were conducted over an 8-day period, sampling every 48 hours in accordance with WST-1 cell proliferation protocol (Abcam). Absorbency readings were obtained using a microplate (ELISA) reader at A450/A680.

2.6 Flow cytometry methods

2.6.1 *Ethanol fixation*

Cells were counted to ensure 1×10^6 cells/ml before cells were washed in PBS (x2) and the remaining pbs removed. The pelleted cells were resuspended in 50 μ l PBS and fixed on ice with the dropwise addition of 1 ml ethanol (70%), while continuously agitating the sample. Fixed cells were stored overnight at 4 °C or stored for 2 hours at -20 for same day usage.

Fixed samples were washed (2x) in 3ml 1% BSA in TBS, with centrifugation (500 xg for 5mins)) before a final wash in 1% IF grade TBS(Sigma). Ethanol fixed samples then proceeded to antibody incubation or propidium iodide (PI) staining.

2.6.2 *Paraformaldehyde (PFA) fixation*

Cell pellets were washed (2x PBS (500 μ l)). Drained cell pellets were fixed on ice by adding 100 μ l, 4% PFA (pierce)) which was added dropwise and with continuous agitation. Cells were incubated in darkness for 15 min, room temperature and washed with centrifugation (2x 3 ml, 1% BSA in TBS (500 xg for 5mins)) before a final wash (1% IF grade TBS, Sigma) and if required left overnight at 4°C. Samples were permeabilised (100 μ l of 1 X TBS containing 0.25% Triton X100) for 15 minutes in darkness, and pelleted (500 xg, 5 min). Samples were washed twice with centrifugation (3 mL, 1% BSA in TBS (5 min, 500 xg)) before a final wash (1% IF grade TBS, Sigma). Samples then proceeded to antibody or propidium iodide (PI) staining.

2.6.3 *Antibody staining*

The pelleted cells were resuspended and incubated with the primary antibody before washing (3x 1ml of 1% BSA in TBS) and recovery (500 xg for 5 mins). Followed by the secondary antibody incubation (Table 8), After incubation, the samples were washed (2x 1 ml of 1% BSA in TBS) (500 xg for 5mins). After which the remaining supernatant was removed.

2.6.4 *PI labelling incubation*

Pelleted cells were washed (1 ml 1 x TBS IF grade) and recovered with centrifugation (500 xg for 5mins) three times. Before the cells were gently resuspended (500 μ l PI buffer) for 5-10 mins in darkness, 4°C. Flow cytometry analysis was performed using the BD FACS Canto™ II flow cytometer with FACSDiva software.

Table 9: Antibody staining conditions and corresponding fixative.

Primary Antibody Staining			
Primary Antibody Staining	Incubation	Fixative	Additional Requirements
anti-phospho Histone H3 pS10, (AB422) (1:100) in 1% BSA in TBS	2 hours, room temperature in darkness	PFA	
Anti-CDT1/DUP antibody (EPR17891) (1:100) in 1% BSA, 0.1% NP40 in TBS,	1 hour, room temperature in darkness	ethanol	Wash steps included NP40 (1ml of 1% BSA in TBS+ 0.1%NP40)
Secondary antibody staining/ counterstain			
8R-goat anti rabbit Alexa Fluor 647 (molecular probes A21245) (1:100) in 1% BSA in TBS	1 hour, room temperature in darkness	PFA	
8R-goat anti rabbit Alexa Fluor 647 (molecular probes A21245) (1:100 in 1% BSA in TBS	1 hour, room temperature in darkness	Ethanol	wash steps included NP40 (1ml of 1% BSA in TBS+ 0.1%NP40)
FxCycle™ PI/RNase Staining Solution Invitrogen™ 500 µl /5*10 ⁵ cells	5-10 mins, room temperature in darkness	Ethanol or PFA	

2.7 Immunofluorescence analysis

2.7.1 *Coverslip preparation and cell seeding*

Glass coverslips (18 mm diameter, 1.5 mm thickness) were washed with ddh²O and sterilised by baking for 2 hours (200 °C) and stored in a sterile container. When required coverslips were transferred to a 24 well plate and cells of interest were seeded upon them at a density of 4×10^4 and incubated overnight (37°C, 5% CO₂).

2.7.2 *Immunofluorescence fixation and blocking*

Cells were washed TBS (x3) and fixed (100% methanol containing 5 mM EGTA (4 mins, -20°C)). Once fixed, cells were washed TBS (x3) and blocked (3% BSA in 1 X TBS) for 30 mins, room temperature with rotation.

2.7.3 *Immunofluorescence primary and secondary antibody incubation*

Primary antibodies were diluted (3% BSA in 1 x TBS (1:100)) (Table 9). Each coverslip was exposed to the antibody within a humid chamber overnight, (3°C). antibody solution was removed, and coverslips were washed TBS (x3).

2.7.4 *Immunofluorescence secondary antibody incubation*

Secondary antibodies were diluted (3% BSA in 1 x TBS (1:500)) (Table 9). Coverslips were incubated with antibody in darkness for 1 hour, room temperature and protected from light until imaging. The secondary antibody was removed, and the coverslips were washed TBS (x3).

Table 10: Antibody staining conditions and corresponding fixative.

Primary Antibody staining			
Primary Antibody staining	Incubation	Fixative	Additional requirements
anti- γ -tubulin (AB133) +1% BSA in TBS, 1:100	2 hours, room temperature in darkness	PFA or Ethanol	If using ethanol fixation Wash steps included NP40 (1ml of 1% BSA in TBS+ 0.1%NP40)
anti-phospho Histone H3 pS10, (AB422) +1% BSA in TBS, 1:100	2 hours, room temperature in darkness	PFA	
Anti-CDT1/DUP antibody (EPR17891) +1% BSA in TBS, +0.1% NP40, 1:100	1 hour, room temperature in darkness	Ethanol	Wash steps included NP40 (1ml of 1% BSA in TBS+ 0.1%NP40)
Secondary antibody staining			
Goat anti-Mouse IgG (H+L) Secondary Antibody, Alexa Fluora 568	1 hour, room temperature in darkness	ethanol or PFA	when fixing with ethanol wash steps included NP40 (1ml of 1% BSA in TBS+ 0.1%NP40)
8R-goat anti rabbit Alexa Fluor 647 (molecular probes A21245) (1:100) in 1% BSA in TBS, 1:100	1 hour, room temperature in darkness	PFA	
8R-goat anti rabbit Alexa Fluor 647 (molecular probes A21245) (1:100) in 1% BSA in TBS	1 hour, room temperature in darkness	ethanol	wash steps included NP40 (1ml of 1% BSA in TBS+ 0.1%NP40)

2.7.5 Immunofluorescence mounting and imaging.

Coverslips were mounted on to glass slides using Vectashield containing DAPI and secured with clear nail polish. Immunofluorescent images were obtained using a Zeiss LSM80 with Airyscan SR module confocal microscope and a 40x objective oil-based lens wavelengths included 568-712nm, 493-598nm and 410-507nm.

2.8 Computational and Statistical analysis methods

2.8.1 Numerical counts of nuclear size

Image analysis software, Image J was utilised to analysis immunofluorescent images stained with anti-Gamma tubulin and DAPI. Nuclei stained with DAPI enabled binary selection to identify each individual nuclei and the area (μm^2) to be determined computationally.

2.8.2 Numerical counts of Centrosome number

Immunofluorescent images stained with anti-Gamma tubulin to highlight the number of centrosomes within each cell were subject to individual counting using Image J. Counts were conducted blind with a minimum of 50 cells analysed per replicate.

2.8.3 Statistical analysis and graph development

The graphs including density plots, box plots, bar charts and line graphs were developed in R studio and Excel. Statistical analysis such as standard deviations, Wilcox non-parametric test and Cramér's V statistical analysis was produced using R studio.

2.8.4 Structural analysis through Alpha fold

The structural features of the CIP29 protein were illustrated through the AI system AlphaFold which utilises a protein's amino acid sequence and information gathered by the European Bioinformatics Institute (EMBL) to predict a protein's 3D structure (Jumper et al., 2021, Varadi et al., 2022).

3 Chapter 3: Analysis of CIP29 Protein-Protein Interactions

3.1 Introduction

Proteins are responsible for all aspects of a cell from its cellular structure to its functionality. The functional role of a protein is often alluded to through its interaction repertoire with other proteins of known function, as many proteins have a functional relationship with one another. However, compiling an accurate picture of protein interactions or an interactome, is a significant challenge due to the sheer number of potential proteins. For example, if the hypothesis of “one gene = one protein,” is followed there should be at least ~20,000 nonmodified human proteins (Ponomarenko et al., 2016). However, when accounting for alternative splicing, polymorphisms, RNA-editing and post translational modifications a single gene has the potential to produce as many as 100 different protein forms. Therefore, the true number of proteins within the human proteome could be closer to several million (Aebersold et al., 2018).

Adding to the complexity of protein-protein interactions, proteins often form large multi-protein complexes which can achieve a molecular weight of several megadaltons. As a result, it can be difficult to distinguish between direct or indirect associations within these structures, due to their interconnected nature. Furthermore, proteins within these structures may only exhibit transient or weak protein associations and could be easily disrupted if specific cellular conditions are not achieved. Thus, the analysis of protein-protein interactions can be illuminating but challenging.

3.1.1 *The TRanscription and EXport (TRES) Complex*

One mega complex of increasing interest is the TRanscription and EXport (TRES) Complex, this complex was discussed in detail within the introduction therefore only a short overview will be provided here. The TRES complex plays a major role in the processing, licencing and export of nascent mRNA transcript through its direct association with mRNA transcripts and through its interactions with other large multi-protein complexes as it accompanies the transcript through various aspects of mRNA maturation (Carmody and Wentz, 2009) (Dias et al., 2010) (Katahira et al., 2013) (Andersen et al., 2013) (Heath et al., 2016). The TRES complex's vast array of protein interactions and influence over mRNA processing is made possible through its robust core structure (the THO complex, UAP56 and ALY) and the multiple subunits that associate to this core including (CIP29, ZC3H11A and ALY-like proteins; CHTOP, UIF (also known as FYTDD1), LUZP4, and POLDIP) (Pühringer et al., 2020), (Strässer et al., 2002) (Dufu et al., 2010) (Heath et al., 2016). The core

proteins provide a base for the additional subunits to associate and disassociate in a highly dynamic nature which is thought to be driven through ATP, although this mechanism is also yet to be fully understood.

An ATP-dependent association between CIP29 and TREX components UAP56 and ALY was first reported by Dufu et al in 2010 yet, these findings were not substantiated in a second study. Additional proteins PDIP3 and ZC11A were also found to associate with the human TREX complex in an ATP-dependent manner although this again was not further validated and the associations between CIP29, UAP56 and ALY were not reproduced (Folco et al., 2012). In contrast, CIP29 was found to associate with UAP56 in the absence of ATP, while the association between CIP29 and ALY was not mentioned (Chi et al., 2013). Furthermore, it has been reported that a second CIP29-containing complex, known as the AREX complex, exists and functions in a similar manner to the TREX complex. To form the AREX complex, URH49 associates with CIP29 but not ALY or UAP56. However, it is yet unclear if other auxiliary TREX complex components are also assembled within this complex (Yamazaki et al., 2010, Fujita et al., 2020). As there is ambiguity within the literature regarding the TREX complex composition further clarification is required to determine if CIP29 is indeed a member of the complex.

3.1.2 *CIP29 and novel interacting proteins*

CIP29 was originally identified as a nuclear protein with the ability to bind nucleic acids that may participate in the control of cellular growth. CIP29 up-regulation has also been linked to an increase in cell proliferation which is likely to be rooted in the ability of CIP29 to regulate gene expression through its role in RNA processing/export. However, it is also possible that CIP29 could influence cell proliferation more directly through interaction with other proteins. To gain some insight into this, previous work from this laboratory took advantage of the fact that *Xenopus* egg extracts are largely transcriptionally inactive (relying on maternal mRNA stores not *de novo* transcription) to look for CIP29-interacting factors in the absence of global transcription and RNA processing. CIP29 was immunoprecipitated from *X. laevis* egg extracts and CIP29-associated proteins were identified through mass spectrometry. This analysis identified several RNA processing and translation initiation factors as CIP29-interactors, but also revealed several proteins involved in mitosis and cell division. The top two hits in duplicate experiments were KIF2C/MCAK, a microtubule depolymerase responsible for chromosome-kinetochore fibre attachment which plays a significant role in chromosome segregation and anillin (ANLN), a scaffolding protein which has been implicated in cytoskeletal dynamics during cellularization and cytokinesis. (Holden and Lindsay, Unpublished data). If these interactions could be confirmed, it may suggest that CIP29 plays a more direct role in

cell division, that is not necessarily mediated through mRNA processing and export. Furthermore, it is unclear if these potential interactions are independent of the TREX complex, and its auxiliary components and so further validation and investigation is required.

3.1.3 *Project objective: Validation of CIP29 protein-protein interactions*

Due to these contrasting results, the association between CIP29 and UAP56 or ALY was investigated through co-immunoprecipitation studies. This technique identifies physiologically relevant protein-protein interactions by using target protein-specific antibodies to isolate the target protein and indirectly capture any associated proteins. Similarly, since previous mass spectrometric analysis from this laboratory had suggested that *X. laevis* CIP29 may associate with several proteins that are known to aid cell cycle progression during mitosis and cytokinesis, we sought to validate these protein-protein interactions in human cells through co-immunoprecipitation assays.

3.2 **Results**

3.2.1 *Development of CIP29 antibody reagents for detection and immunoprecipitation of CIP29*

To investigate the protein interactions of CIP29, affinity purified antibodies that are specific to the CIP29 protein were required for both the immunoprecipitation of CIP29 and its subsequent detection through western blot. Therefore, a rabbit polyclonal antibody previously raised against recombinant full length XCIP29 (and known to cross-react with the human CIP29 protein) was affinity purified from serum using a XCIP29 antigen column (refer to M&M 2.3.12). The resulting affinity purified antibody detected a single band at the expected size of the CIP29 protein (~29kDa) in MRC5-VI whole cell extracts and HeLa nuclear extracts through western blot (Figure 11 A/B). Upon confirmation that the purified antibody specifically detected the human CIP29 protein, the antibody was crosslinked to protein A Sepharose (0.5µg antibody/ 1ml Sepharose beads) (refer to M&M 2.3.12). To determine whether the α -CIP29 antibody beads could successfully immunoprecipitate CIP29, an immunoprecipitation experiment was performed using either the α -CIP29 antibody beads (α -CIP29) or non-specific rabbit IgG-crosslinked beads as a control (MOCK), to pull down associated proteins from MRC5-VI whole cell extracts or HeLa nuclear extracts (refer to M&M standard IP 2.4.1). Western blot analysis of the immunoprecipitated samples indicated that the CIP29 protein was successfully immunoprecipitated from the MRC5-VI whole cell extract and the HeLa nuclear extract by the α -CIP29 beads, but not by the control IgG-beads (Figure 11C/D). Similarly, an equivalent immunoprecipitation result was observed when utilizing CIP29 protein that has been synthesized using an *in vitro* transcription/translation system, which synthetically

transcribes and translates protein from a provided DNA template (Figure 11E). Together, these results indicate that the CIP29 antibody is specific to the human CIP29 protein, and that the antibody is suitable for use in western blotting and immunoprecipitation analysis of the CIP29 protein derived from human cells. Therefore, this α -CIP29 antibody was utilised to investigate protein binding between CIP29 and other proteins of interest.

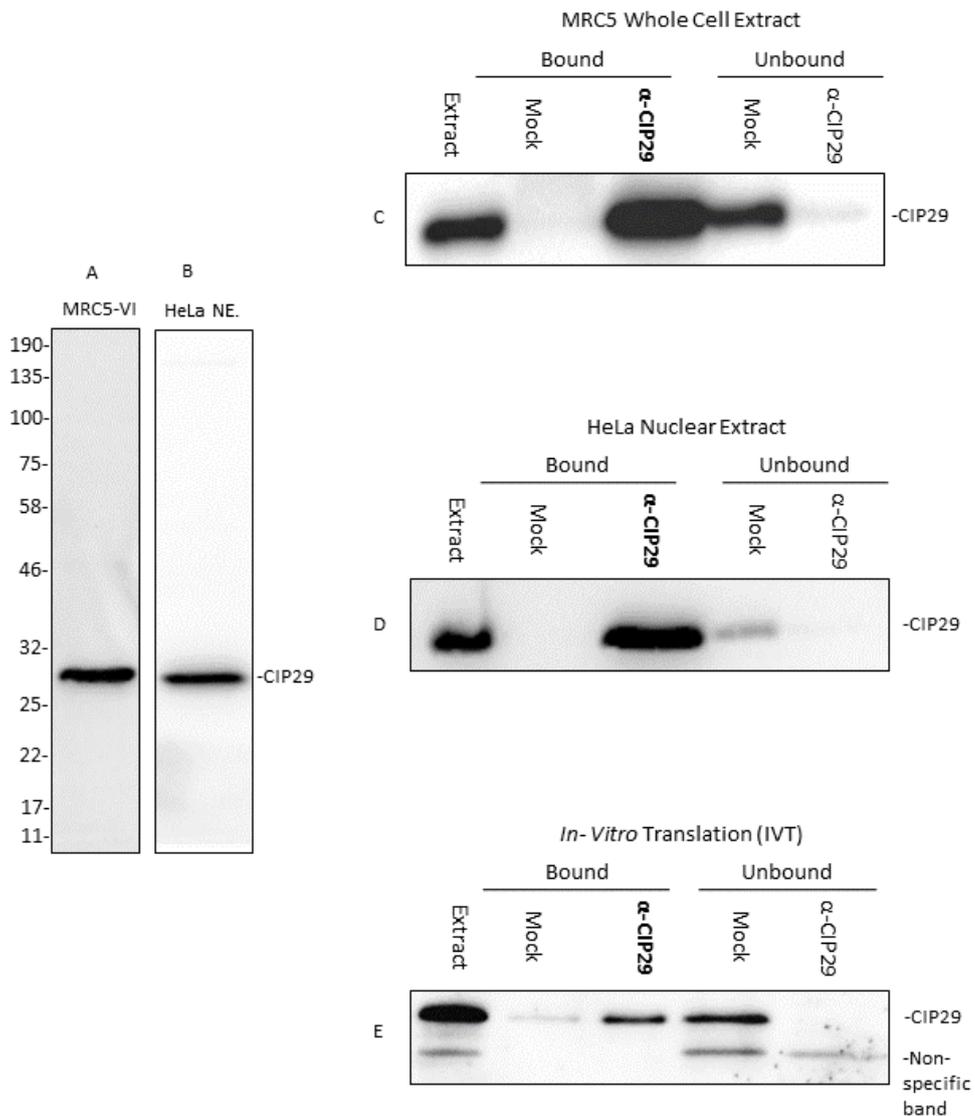


Figure 11: Affinity purified CIP29 antibody specifically recognises and immunoprecipitates the CIP29 protein.

Western blot of CIP29 from (A) MRC5-VI whole cell extracts and (B) HeLa nuclear extracts, confirms detection of CIP29 protein at \sim 29kd. Molecular weight markers are shown.

Western blot of immunoprecipitation samples, pulled down with mock beads or α -CIP29 antibody beads, from (C) MRC5-VI whole cell extracts, (D) HeLa nuclear extracts or (E) CIP29 in vitro transcription and translation reaction. 25% of input, 100% of bound fraction and 25% of unbound fractions were used for western blot. (A, B, C and D: R=3) (E: R=2)

3.2.2 *CIP29 does not associate with TREX complex components using the standard immunoprecipitation protocol.*

To confirm CIP29 as a component of the TREX complex, co-immunoprecipitation was used to investigate CIP29 binding to known TREX complex components; ALY and UAP56. Immunoprecipitations were performed, as before, from MRC5-VI whole cell extracts and HeLa nuclear extracts. (refer to M&M 2.4.1) The resulting immunoprecipitation samples were then analysed by Western blotting for the presence of CIP29, UAP56 or ALY protein. The results confirmed that the CIP29 protein was successfully immunoprecipitated from MRC5-VI cell extracts by the CIP29-specific antibody beads (α -CIP29). In contrast, none of the CIP29 protein was immunoprecipitated with the non-specific IgG mock beads (MOCK). The UAP56 protein was readily detectable within the input MRC5-VI whole extract sample, but only a very small proportion could be detected in the CIP29 immunoprecipitation sample. This was not reproducible in subsequent experiments and is therefore likely to represent a small amount of non-specific binding to the beads. Although the ALY protein was also detected within the input extract, ALY was not detected in the CIP29 immunoprecipitation (Figure 12A). Therefore, there was no appreciable co-immunoprecipitation of UAP56 or ALY with CIP29 when analysing MRC5-VI whole cell extracts. A similar co-immunoprecipitation experiment was also conducted using commercially prepared HeLa cell nuclear extracts. Once again, CIP29 was successfully immunoprecipitated using the α -CIP29 antibody, but neither UAP56 nor ALY were co-immunoprecipitated with CIP29 from the HeLa nuclear extract (Figure B).

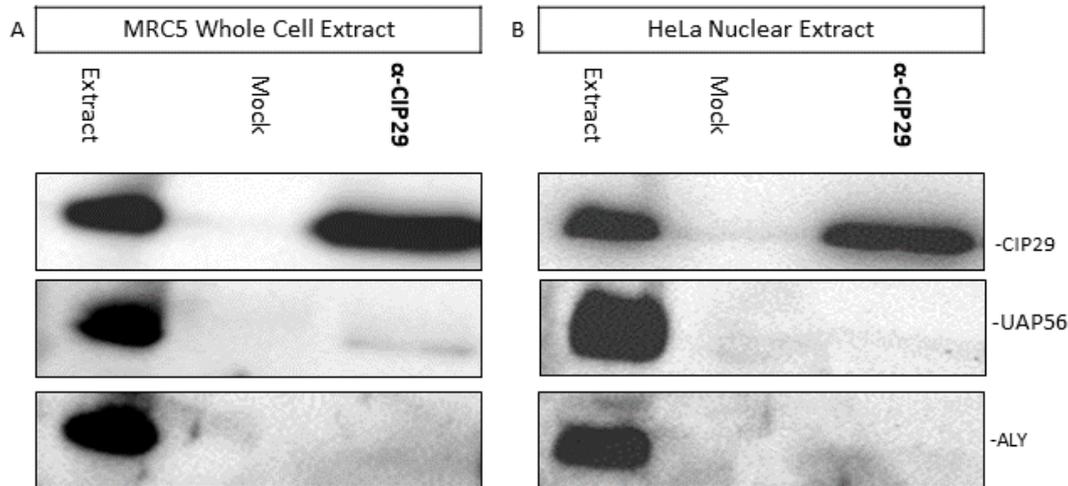


Figure 12: TREX/AREX components UAP56/URH49 and ALY do not co-immunoprecipitate with CIP29

Western blot of immunoprecipitation samples, pulled down with mock beads or α -CIP29 antibody beads, from (A) MRC5-VI whole cell extracts or (B) HeLa nuclear extracts. Proteins of interest are indicated alongside. 25% of input, 100% of bound fraction and 25% of unbound fractions were used for western blot. (R=3)

Although CIP29 did not co-immunoprecipitate with ALY or UAP56 in this experiment, the association of CIP29 with the TREX complex cannot be invalidated, as this has previously been reported to only be achieved under conditions that support RNA processing and is dependent on the addition of ATP (Dufu et al., 2010). In the absence of ATP, it has been reported that UAP56 forms an Apo-TREX complex containing the THO subcomplex, but not ALY or CIP29, while CIP29 forms part of an Apo-AREX complex with the UAP56 paralog URH49, but does not associate with UAP56, ALY or the THO subcomplex (Yamazaki et al., 2010). With the addition of ATP, both the Apo-TREX complex and the Apo-AREX complex are remodeled to highly similar ATP-TREX complexes containing the THO subcomplex, ALY, CIP29 and UAP56 (Yamazaki et al., 2010) (Fujita et al., 2020). This would be consistent with the lack of binding between UAP56 or ALY and CIP29 seen in Figure 12 A/B as this was in the absence of added ATP. It is possible that the small amount of UAP56 observed in the CIP29 pulldown from MRC5-VI cell extracts could potentially represent URH49 associating with CIP29 in the Apo-AREX complex, since the UAP56 antibody used here is likely to detect both UAP56 and URH49 given the sequence identity shared by these two proteins. However, this result was not reliably observed in subsequent immunoprecipitations from MRC5-VI cell extracts and was not

apparent in the immunoprecipitations conducted from HeLa nuclear extracts, therefore the weak association observed is most likely to represent some non-specific binding of UAP56/URH49.

3.2.3 Does CIP29 associate with TREX/AREX components in the presence of ATP?

Since the interaction between CIP29 and UAP56 or ALY had previously been reported in splicing-competent extracts, we aimed to reconstitute these conditions, through the addition of ATP and prior incubation at 30°C. Using such conditions, two groups had already successfully described CIP29 interactions with UAP56 and ALY (Dufu et al., 2010) and (Fujita et al., 2020). Dufu et al. utilised HeLa nuclear extracts that were supplemented with ATP and incubated at 30°C. Therefore, the following experiments recapitulated these conditions using both MRC5-V1 whole cell extract (Figure 13 A/B) and commercial HeLa nuclear extract (Figure 13 C/D). Each extract was supplemented with or without ATP, for 30-minutes at 30°C prior to performing α -CIP29 and MOCK immunoprecipitations (refer to M&M IP 2.4.2). Panels A and C show the immunoprecipitation results from both extracts in the absence of ATP while, panels B and D shows the results after supplementation of each extract with ATP. While CIP29 was effectively immunoprecipitated by the α -CIP29 antibody in each case, there was no detectable co-immunoprecipitation of UAP56 or ALY, regardless of ATP supplementation.

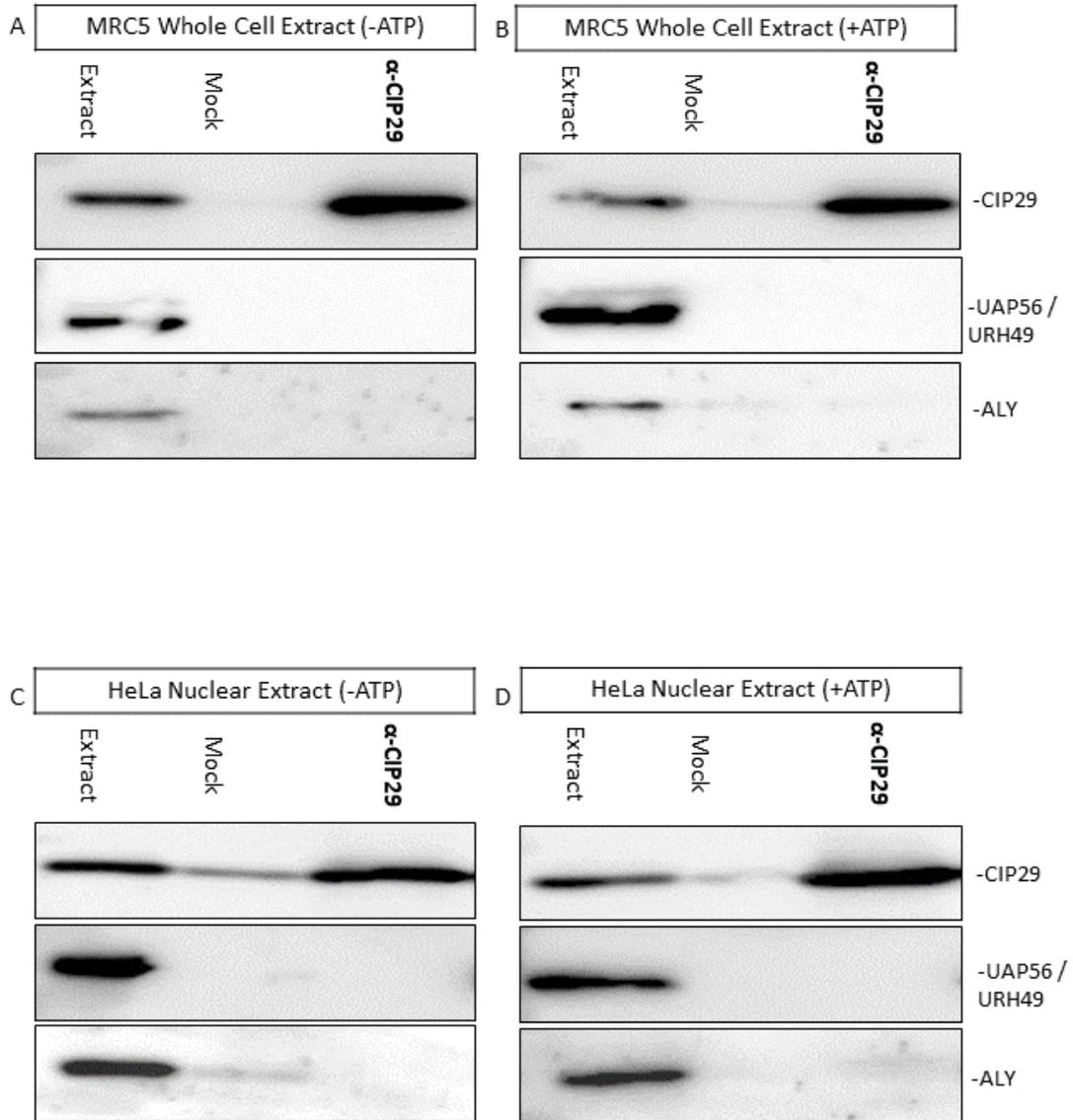


Figure 13: TREX/AREX components UAP56/URH49 and ALY do not co-immunoprecipitate with CIP29 regardless of ATP addition

Western blot of Immunoprecipitation samples, pulled down with mock beads or α-CIP29 antibody beads from MRC5-VI whole cell extract (A/B) and HeLa nuclear extract (C/D), with (B/D) or without (A/C) addition of ATP, in accordance with Dufu et al., (2010). The proteins detected are indicated alongside. 25% of input, 100% of bound fraction and 25% of unbound fractions were used for western blot (R=3).

The second, modified immunoprecipitation method through which the association of CIP29 with UAP56 and ALY was detected following ATP supplementation was reported by Fujita et al., (2020). This method included an additional incubation, in the presence of 0.1% Triton X-100, 0.2mM protease inhibitors and 0.5mM dithiothreitol, prior to ATP addition and incubation at 30°C (refer to M&M 2.4.3). The impact of these alterations can be seen in three independent repeats of the CIP29 immunoprecipitation reaction from HeLa nuclear extracts in Figure 14 (panels A, C and E show the immunoprecipitation results from HeLa nuclear extract in the absence of ATP, while panels B, D and F show the results after ATP supplementation). CIP29 was effectively immunoprecipitated by the α -CIP29 antibody in each case. While the initial experiment suggested that ALY and UAP56 co-immunoprecipitated with CIP29 in an ATP-dependent manner, this result was not reproducible, despite apparently identical experimental conditions in subsequent repeat experiments.

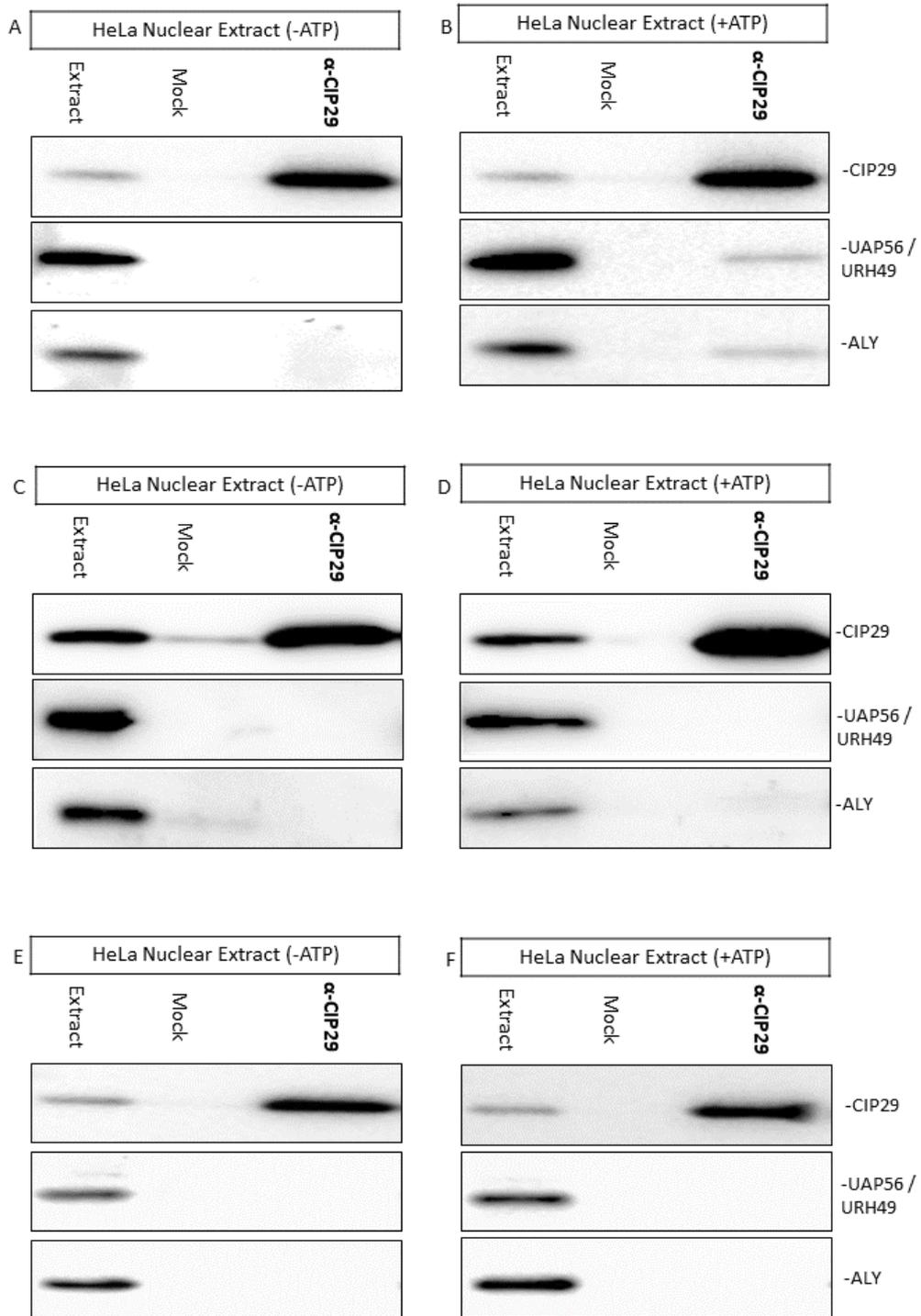


Figure 14: Using Fujita et al., (2020) methodology CIP29 does Co-immunoprecipitate with TREX components UAP56/URH49 or ALY, however reproducibility is poor.

Western blot of Immunoprecipitation samples, pulled down with mock beads or α-CIP29 antibody beads were conducted on HeLa nuclear extracts with (B/D/F) or without (A/C/E) addition of ATP, in accordance with Fujita et al., (2020). The proteins detected are indicated alongside. 25% of input, 100% of bound fraction and 25% of unbound fractions were used for western blot. (R=3, all repeats shown)

Thus far, it has not been possible to reproducibly co-immunoprecipitate either UAP56/URH49 or ALY with CIP29. It is not clear if this is because the association with CIP29 depends on conditions that have not yet been recapitulated or if the CIP29 antibody being used here for these immunoprecipitations is disrupting any interactions. To try to differentiate between these possibilities, immunoprecipitations were once again performed using HeLa nuclear extracts supplemented with ATP (refer to M&M 2.4.2) but using antibodies against α -UAP56 or α -ALY instead of α -CIP29 to pulldown the proteins of interest. Immunoprecipitations using α -UAP56 antibodies successfully pulled down UAP56/URH49 from HeLa nuclear extract, in both the absence and presence of ATP. In both cases, CIP29 was also pulled down by the α -UAP56 antibodies but not by the mock beads (Figure 15 A/B). Similarly, immunoprecipitations using α -ALY antibodies also immunoprecipitated ALY regardless of ATP supplementation. In this case however, CIP29 co-immunoprecipitation with ALY was only observed from HeLa nuclear extract that had been supplemented with ATP (Figure 15 C/D). Together, this confirms that CIP29 does associate with ALY in an ATP-dependent manner. However, the association of CIP29 with UAP56/URH49, is observed irrespective of ATP addition. This initially appears to contradict published data regarding the ATP-dependence of CIP29 association with TREX components. However, given that the UAP56 antibody used here is likely to cross react with URH49 due to their shared sequence identity, it is likely that CIP29 is associated with URH49 in the absence of ATP, forming the AREX complex while in the presence of ATP CIP29 can form the TREX complex with UAP56 and ALY. This would be consistent with previous published reports. These results also suggest that the α -CIP29 antibody may be disrupting the interaction between CIP29 and the TREX complex components during the immunoprecipitation studies, as ALY and UAP56 could not be pulled down by CIP29 using the α -CIP29 antibody but CIP29 was successfully pulled down by both UAP56 and ALY when using their specific antibodies. As these immunoprecipitation studies shared the same experimental conditions aside from the specific antibodies, it indicates that the conditions required for CIP29-containing complexes were recapitulated, therefore the pull down of ALY and UAP56 by CIP29 should have been possible.

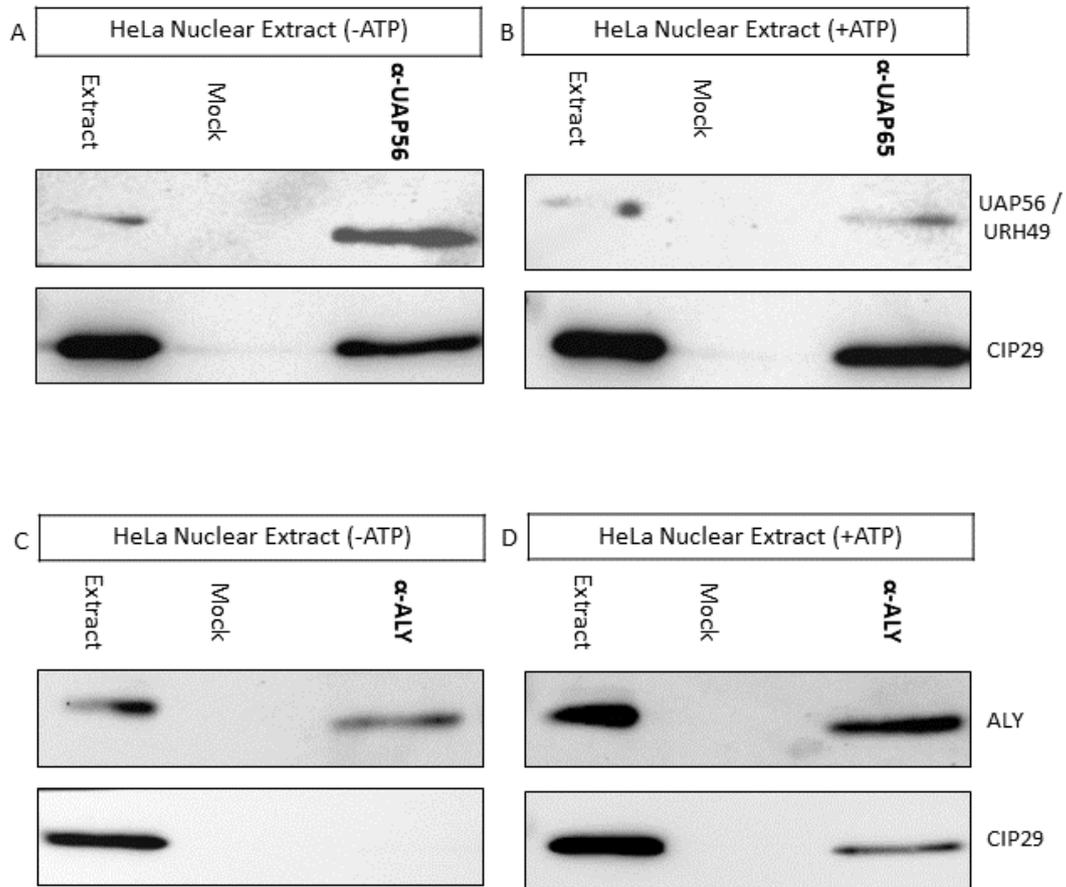


Figure 15: CIP29 does co-immunoprecipitate with TREX/ALEX components.

Western blot of Immunoprecipitation samples, pulled down with mock beads or α -UAP56 antibody beads or α -ALY were conducted on HeLa nuclear extracts with (B/D) or without (A/C) addition of ATP in accordance with DUFU et al., (2010). The proteins detected are indicated alongside. 25% of input, 100% of bound fraction and 25% of unbound fractions were used for western blot (R=2).

3.2.4 Validation of potential CIP29 interactors identified through mass spectrometry.

In addition to the reported protein interactions between CIP29 and TREX/AREX components, previous experiments from this laboratory have identified other proteins that may potentially associate with CIP29 -through immunoprecipitation of *X. laevis* CIP29 and mass spectrometric analysis of co-immunoprecipitated proteins. These potential CIP29 interactors had yet to be further validated but, since several of these interactions held phenotypic relevance, the project began to focus on this task.

The two strongest hits amongst the XCIP29 co-immunoprecipitated proteins were the kinesin-like protein, KIF2C and anillin which are involved in chromosome segregation and cytokinesis respectively. A variety of other proteins with functions relating to chromosome segregation, cell division and RNA processing were also identified, including Mitotic arrest deficiency 1 (MAD1), which functions as part of the spindle assembly checkpoint; Nuclear protein localization protein 4 (NPL4), which is linked to G2/M checkpoint signaling and may aid the suppression of chromosome instability; Actin-related protein 3 (ARP3), a protein linked to actin polymerisation; Dynamin 2 (DNM2), a fission protein that participates in endocytic vesicle formation and in the dynamics of microtubules; Nuclear scaffold protein p54 (NONO/p54), an RNA binding protein that facilitates transcriptional regulation and splicing and Arsenite resistance protein 2 (ARS2), an adapter protein that facilitates the association between various RNA maturation complexes (Schulze et al., 2018) (Luo et al., 2018) (Sato et al., 2019) (Liu et al., 2013) (Landrieu et al., 2013) (Shen et al., 2021). Confirmed binding between CIP29 and any of these additional proteins of interest could help to further elucidate the influence and functional roles of CIP29.

To investigate whether any of these protein interactions could be validated in human cells, we assessed whether they co-immunoprecipitated with CIP29 from MRC5-VI cell extracts (refer to M&M 2.4.1). As previously observed, CIP29 was successfully immunoprecipitated by the CIP29-specific antibody beads (Figure 16). However, there was no apparent association between CIP29 and MAD1, NPL4, ARP3, DNM2, NONO/p54 or ARS2 in these immunoprecipitations from human cell extract.

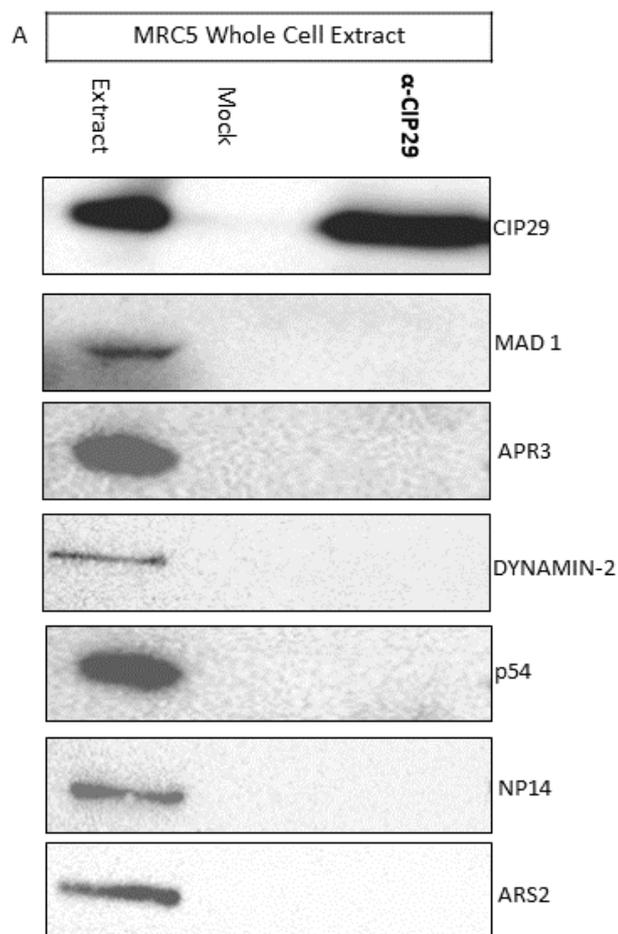


Figure 16: Various proteins that were identified as potential interactors through mass-spectrometry do not co-immunoprecipitate with CIP29 within MRC5-V1 cell extracts.

Western blot of Immunoprecipitation samples, pulled down with mock beads or α -CIP29 antibody beads were conducted on MRC5-VI whole cell extracts. The proteins detected are indicated alongside. 25% of input, 100% of bound fraction and 25% of unbound fractions were used for western blot Proteins of interest are shown alongside. (R=3).

3.2.5 Analysis of CIP29-anillin interaction

When assessing the interaction of CIP29 and anillin, initial experiments revealed that anillin was present in the α -CIP29 immunoprecipitation samples derived from MRC5-VI whole cell extract (Figure 17A). anillin was also readily detected following CIP29 pulldown from HeLa nuclear extract (Figure 17B). This suggested that the interaction between CIP29 and anillin identified in *X. laevis* egg extracts, was conserved in human cells. However, in subsequent experiments, these associations were found to be variable and somewhat inconsistent (Figure 17 C/D).

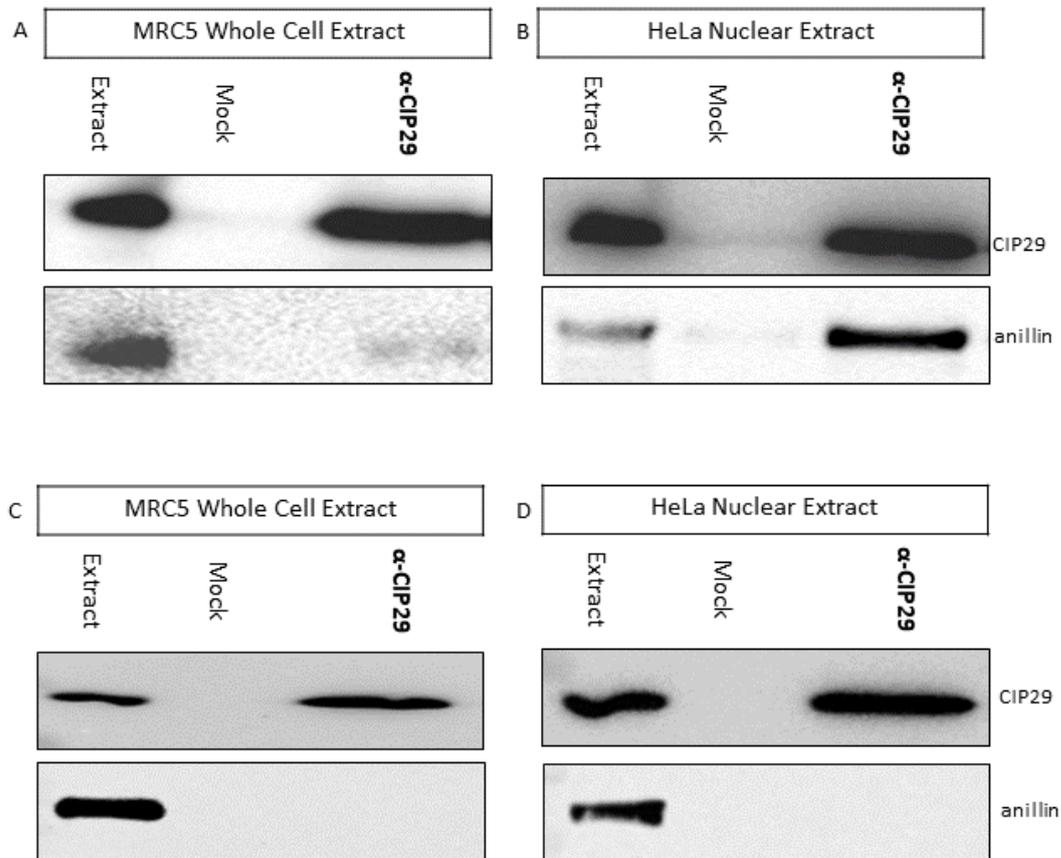


Figure 17: anillin, a potential interactor identified through mass-spectrometry does not reliably Co-immunoprecipitate with CIP29.

Western blot of Immunoprecipitation samples, pulled down with mock beads or α -CIP29 antibody beads were conducted on MRC5-VI whole cell extract (A/C) or HeLa nuclear extract samples (B/D). The proteins detected are indicated alongside. 25% of input, 100% of bound fraction and 25% of unbound fractions were used for western blot Proteins of interest are shown alongside (R=4, contrasting data shown).

3.2.6 Analysis of CIP29-KIF2C interaction

In contrast, KIF2C co-immunoprecipitated reproducibly with CIP29 from MRC5-VI whole cell extracts, HeLa nuclear extracts and MRC5-VI nuclear extracts (Figure 18).

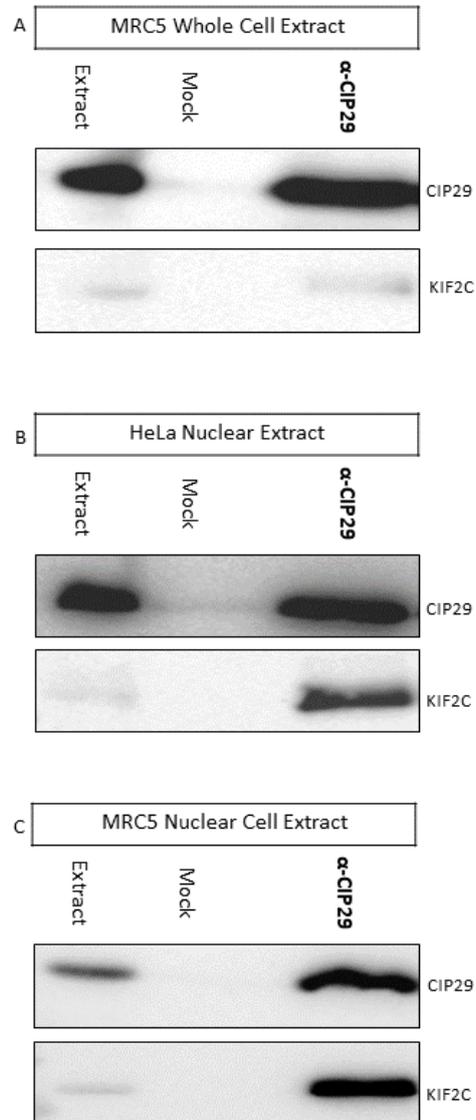


Figure 18: KIF2C, a potential interactor identified through mass-spectrometry does reliably Co-immunoprecipitate with CIP29.

Western blot of Immunoprecipitation samples, pulled down with mock beads or α -CIP29 antibody beads were conducted on MRC5-VI whole cell extract (A) HeLa nuclear extract samples (B) or MRC5-VI nuclear extract (C). The proteins detected are indicated alongside. 25% of input, 100% of bound fraction and 25% of unbound fractions were used for western blot Proteins of interest are shown alongside (R=3).

As the association between CIP29 and KIF2C was so reproducible, steps were taken to further characterize the interaction. Firstly, we wanted to determine if the association between CIP29 and KIF2C was through a direct interaction between the two proteins. To investigate this, the CIP29 and KIF2C genes were expressed using an *In Vitro* Translation System (IVT) (M&M 2.3.11), which synthetically transcribes and translates the DNA to form only the proteins of interest, thus ensuring no outside influences can be exerted on the proteins to stimulate or maintain protein binding (Steinle et al., 2017). *In vitro* transcription and translation of CIP29 gave rise to a product of ~29kDa, which is equivalent to the full length CIP29, as well as a minor product at ~25kDa, which was also recognised by the α -CIP29 antibodies on western blot. This smaller product most likely represents a truncated form of CIP29 or a non-specific cross-reacting band within the IVT reaction mix. Despite this, only the full length CIP29 was immunoprecipitated by the α -CIP29 antibodies. *In vitro* transcription and translation of KIF2C gave rise to a single band at the expected height for the full length KIF2C. The CIP29 and KIF2C *in vitro* transcription/translation products were mixed and incubated as described before being immunoprecipitated with either mock or α -CIP29 antibody beads (M&M 2.3.11, 2.4.1). While CIP29 was successfully immunoprecipitated under these conditions, KIF2C was not pulled down with CIP29 (Figure 19). This experiment ruled out a direct interaction between KIF2C and CIP29 or indeed, between KIF2C and the CIP29 antibody beads, under the conditions tested. It remains possible that CIP29 and KIF2C interact directly, but that optimal conditions for this binding have not yet been identified. For example, an interaction between CIP29 and KIF2C may be dependent on appropriate: posttranslational modifications, ATP or cofactor stimulation; or the presence of intermediates such as RNA, DNA or additional proteins.

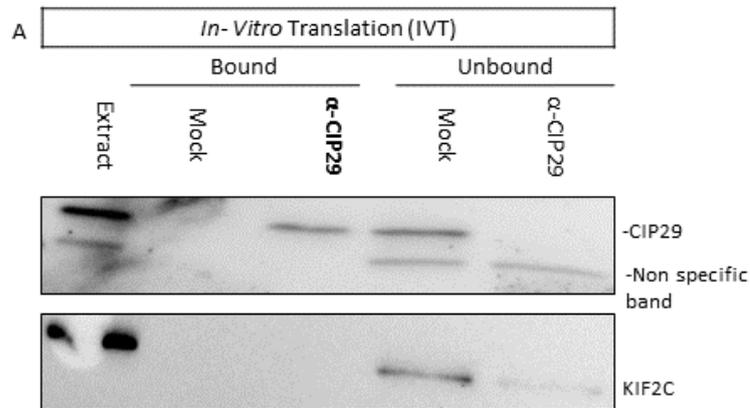


Figure 19: KIF2C did not immunoprecipitate with CIP29 under *in vitro*-translation conditions

Western blot of Immunoprecipitation samples, pulled down with mock beads or α -CIP29 antibody beads were conducted on IVT synthetically developed CIP29 and KIF2C proteins. The proteins detected are indicated alongside. 25% of input, 100% of bound fraction and 25% of unbound fractions were used for western blot Proteins of interest are shown alongside (R=3).

Since the interaction between CIP29 and KIF2C that was observed in cell extracts was not detected when the *in vitro* translated proteins were mixed, it is possible that their interaction is dependent on an intermediary or specific conditions that are present within cells. Therefore, to further validate our initial co-immunoprecipitation findings, that were based on using α -CIP29 antibodies to isolate associated proteins, we pursued two different approaches. The first was to perform the reciprocal immunoprecipitations of the endogenous CIP29 and KIF2C proteins using α -KIF2C antibodies. The second approach involved immunoprecipitations utilising epitope-tagged FLAG-CIP29 and HA-KIF2C following transfection and ectopic expression within MRC5-VI cells (refer to M&M 2.2.6, 2.4.1). The KIF2C immunoprecipitations were somewhat ambiguous, as CIP29 was initially pulled down with KIF2C although, some non-specific binding was seen in the related mock samples (Figure 20A). Upon repetition, very little KIF2C was pulled down by the α -KIF2C immunoprecipitation beads and CIP29 was not detected in the immunoprecipitated sample, therefore it was difficult to draw firm conclusions about the association of CIP29 and KIF2C from these experiments (Figure 20B). The second approach utilised mammalian expression vectors (pCI-neo) that expressed FLAG epitope-tagged CIP29 and HA epitope-tagged KIF2C following transient transfection into MRC5-VI cells. 24h after transfection, cell extracts were prepared, and FLAG-tagged proteins were isolated by binding to α -FLAG antibody beads as described (M&M 2.5.4, 2.4.1). Expression of FLAG-CIP29 was confirmed by western blot with α -CIP29 and α -FLAG antibodies while expression of HA-KIF2C was confirmed by western blot with α -KIF2C and α -HA antibodies. FLAG-CIP29 was readily pulled down using α -FLAG

antibody beads but no HA-KIF2C associated, and this protein was detected in the FLAG pulldown samples (Figure 20C).

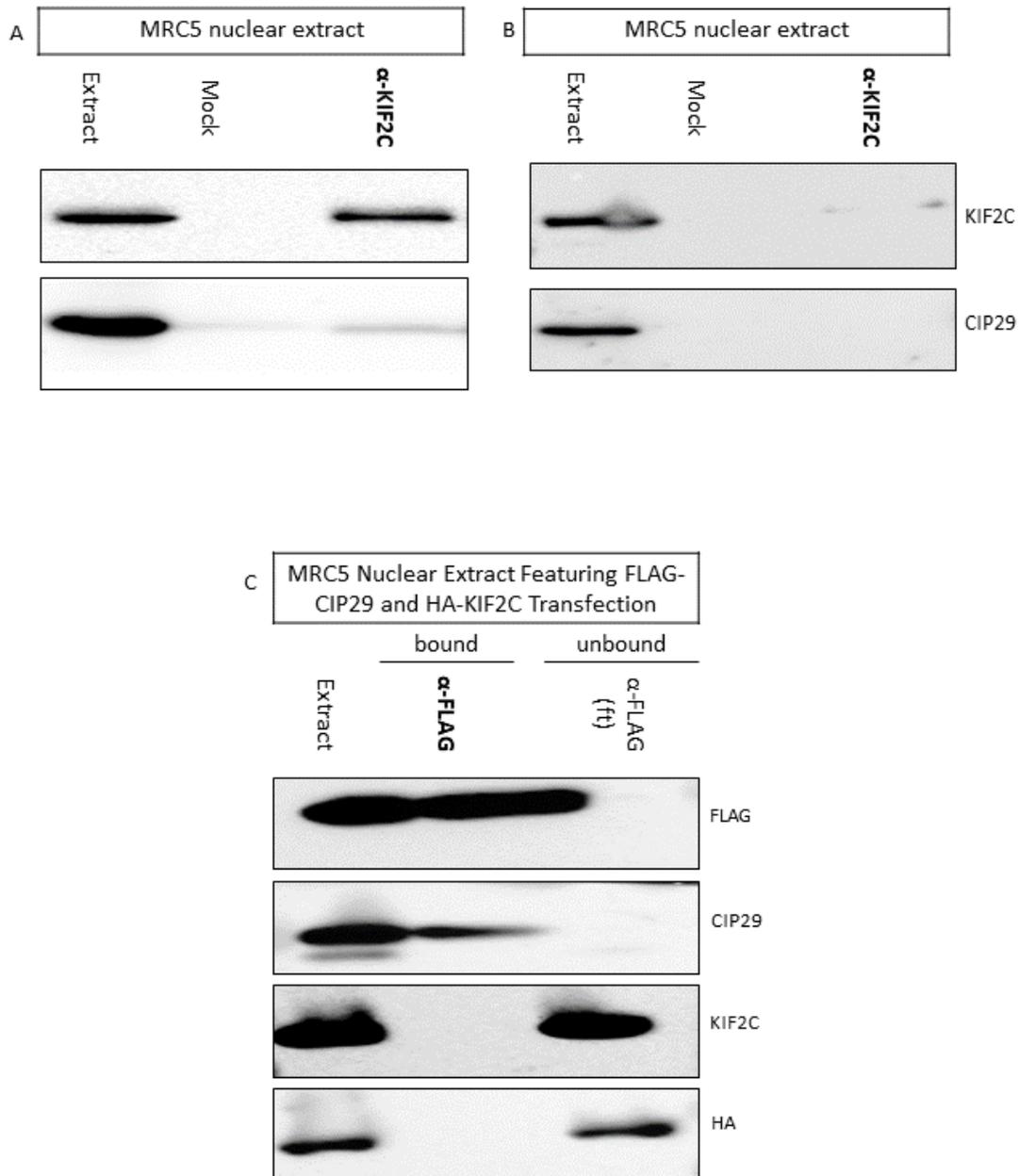


Figure 20: Binding between CIP29 and KIF2C is absent within α -FLAG immunoprecipitation assays.

Western blot of Immunoprecipitation samples, pulled down with mock beads or α -KIF2C antibody beads were conducted on MRC5-VI nuclear extracts (A/B) or MRC5-VI nuclear extracts supplemented with transfected FLAG-CIP29 and HA-KIF2C epitope tags (C). The proteins detected are indicated alongside. 25% of input, 100% of bound fraction and 25% of unbound fractions were used for western blot. (R=3, contrasting result is shown).

The results obtained suggested that the exogenously expressed epitope-tagged KIF2C and CIP29 do not interact with each other or with the endogenous CIP29 or KIF2C proteins. This contradicts the previous data which analysed endogenous CIP29 and KIF2C within whole cell, HeLa nuclear and MRC5 nuclear extracts which suggested a highly stable and reproducible interaction. This contradictory data may not discount the interaction between CIP29 and KIF2C as the endogenous proteins may already be fully associated with other endogenous proteins and therefore the epitope-tagged KIF2C and CIP29 remained unbound to give the appearance of a lack of association. Or the proteins may be available, but the association is hindered by the epitope tags.

Alternatively, the result may show that the previous data using endogenous CIP29 and KIF2C is an artefactual result. However, this is not supported when observing the mock controls or the IVT immunoprecipitations which discounts any direct interaction between KIF2C and the immunoprecipitation beads or the α -CIP29 antibody. Alternatively, a second protein that can be recognised and immunoprecipitated by the α -CIP29 antibody may be facilitating the interaction, while circumventing CIP29. This second protein would not be identified within the mock samples as the α -CIP29 antibody is not present. The protein itself would also not be present in the IVT derived immunoprecipitations as only the CIP29 protein could be produced, which would explain the lack of pull down in this experiment, as the interaction between KIF2C and the α -CIP29 antibody would not have been facilitated by the secondary protein. It would also not be isolated through the α -FLAG pulled down as the secondary protein would lack the FLAG epitope.

To unambiguously determine whether the KIF2C pulldown is dependent on an association with CIP29, we took advantage of the fact that the laboratory has developed a cell line in which the expression of CIP29 has been abrogated (this cell line is further discussed in Chapter 4). As the CIP29-knockout (CIP29-KO) cell lines are derived from the MRC5-VI cell lines they maintain all the biological attributes that are not dependent on CIP29. However, as the endogenous CIP29 gene has been disrupted through CRISPR-Cas9 gene editing no CIP29 protein is present within these cells. If the KIF2C pull down is dependent on the presence of CIP29, no association should be observed within these cells. For control purposes the immunoprecipitations from CIP29-KO nuclear extracts were conducted alongside MRC5-VI nuclear extracts, within which KIF2C is readily pulled down along with CIP29 as previously observed (Figure 21A). As there is no CIP29 within the CIP29-KO cells, both the extract and MOCK samples are clean of CIP29. While KIF2C is present within the extract sample, it is absent in the mock which suggests no non-specific binding to the immunoprecipitation beads. However, there is pulldown of KIF2C within the immunoprecipitation sample despite the absence of the CIP29 protein (Figure 21B).

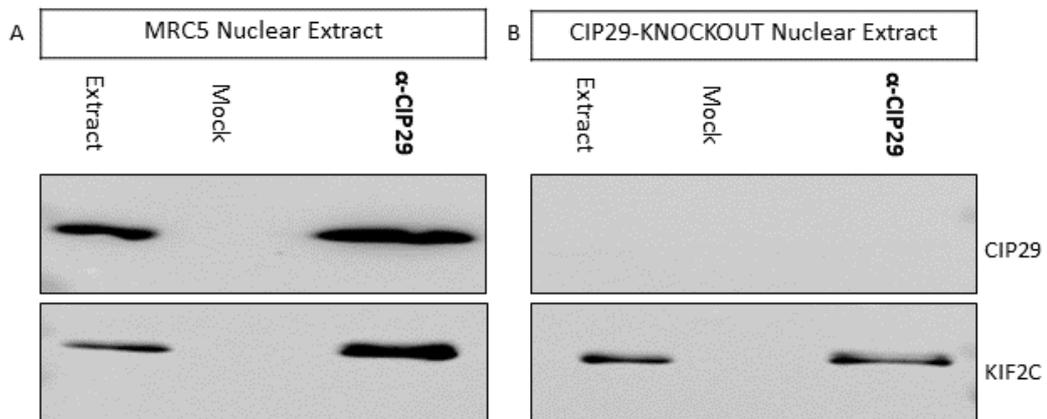


Figure 21: KIF2C associations with CIP29 may be disingenuous.

Western blot of Immunoprecipitation samples, pulled down with mock beads or α -CIP29 antibody beads were conducted on MRC5-VI nuclear cell extracts (A) or CIP29-KNOCKOUT nuclear extracts (B).

The proteins detected are indicated alongside. 25% of input, 100% of bound fraction and 25% of unbound fractions were used for western blots (R:3).

This suggests that KIF2C can be immunoprecipitated without the presence of CIP29 and therefore the previous binding observed between CIP29 and KIF2C may be disingenuous. Unfortunately, it was beyond the scope of this project to investigate this issue further.

3.2.7 Discussion

3.2.8 *The results suggest that CIP29 is indeed a TREX complex component whereby the association relies on ATP, however the α -CIP29 antibody used in this study may disrupt the association.*

When initiating this study CIP29 and its protein interactors were somewhat ambiguous, with conflicting data regarding the involvement of CIP29 within the TREX complex. However, as the research has progressed and subsequent papers have been published, the role of CIP29 in relation to the TREX complex is beginning to show further clarity. For example, a study using cryo-electron microscopy was able to confirm that the overall structure of the core TREX complex and the mechanics that facilitate these interactions relies heavily on ATP (Pühringer et al., 2020). These observations further expanded and supported previous reports such as Dufu et al (2010) and Folco et al (2012). While our own data suggested that no reliable interaction between CIP29 and the TREX complex components was observed in the absence of ATP (Figure 12) again this supports the suggestion that ATP is indeed necessary for the complex to form and function correctly.

Despite the requirement for ATP, when assessing the immunoprecipitations of α -UAP56 or α -ALY to pull down CIP29, binding between CIP29 and UAP56 was seen regardless of ATP, while the interaction with ALY required ATP (Figure 15). This indicates that CIP29 can associate with UAP56 independent of ATP, while its association with ALY is ATP dependent. However, as the α -UAP56 antibody potentially cross reacts with URH49 due to their shared sequence identity, this result may show the interaction between CIP29 and URH49 when forming the AREX (alternative mRNA export) complex, which has been reported to be independent of ATP. Once ATP is supplemented, CIP29 can form the TREX complex with UAP56 and ALY. It is important to note that an interaction between CIP29 and UAP56 in the absence of ATP is not unique, with previous reports documenting a similar result, which was ultimately accredited to the high protein concentrations within those specific samples, enabling the requirement for ATP to be bypassed (Crute et al., 1983, Snyder et al., 2004). Such reports and our own data highlight importance of sample handling and its impact on the reproducibility of such experiments therefore additional replicates and alternative experiments or methodologies is always beneficial to further clarify a result or assumption.

Even with the addition of ATP, the binding between CIP29 and UAP56 or ALY was still somewhat unstable with published methodologies consistently failing to produce stable interactions within our laboratory (Figure 13 and 14). This suggests that it is possible that other conditions may not have been achieved during some of the experiments. For example, mRNA availability, competition from other TREX complex components may have hindered the formation of a stable and consistent TREX complex or the α -CIP29 antibody may have disrupted the interactions.

As the interaction between CIP29 and the TREX complex components was more reliable when pulling down with α -ALY or α -UAP56 antibodies, it suggests that although the experimental conditions may not have always been absolutely optimal, they were sufficient for the interaction to exist. This suggests that the α -CIP29 antibodies may have contributed to the disruption of the final complex's formation. This again highlights the requirement for additional procedures such as reciprocal immunoprecipitation assays to highlight any possibility of false positive or negative results.

3.2.9 Attempted validation of novel CIP29 interacting proteins

Our laboratory's previous mass spectrometric analysis of XCIP29 co-immunoprecipitations suggested that XCIP29 may associate with multiple additional proteins, in addition to TREX complex components. Many of these suggested interactors held phenotypic relevance in which an association may begin to explain the significant impact CIP29 has on cell proliferation and its continued recognition as an oncogenic promotor. XCIP29 has significant sequence homology to the human CIP29 protein, sharing 72% sequence identity. XCIP29 also contains conserved features such as the N-terminal SAP motif and two C-terminal nuclear localisation sequences. Despite this high degree of conservation, the results obtained from previous analysis of XCIP29 interactors could not be validated in human cells, with human CIP29 failing to associate with several of the potential interactors including MAD1, NPL4, ARP3, DNM2, NONO/p54 and ARS2 through immunoprecipitation studies (Figure 16). Furthermore, these potential interactions do not appear in published protein interaction databases such as IntAct or String.

Some of the mass-spectrometry predictions were shown to exhibit some degree of interaction such as the interaction between CIP29 and anillin. This interaction held biological plausibility due to its role in cytoskeletal dynamics during cytokinesis. Furthermore, the concomitant work presented in subsequent chapters also suggests a possible cytokinesis influence for CIP29 and therefore it was pertinent to the project to further understand this potential interaction. Upon further analysis, the association was detected in MRC5-VI whole cell extracts and HeLa nuclear extracts, but this result was somewhat variable (Figure 17). As no interactions between CIP29 and anillin have been published within interaction databases to date, it suggests that the associations may require further validation before this interaction can be confirmed. It is not uncommon for certain antibodies to not perform well under immunoprecipitation, and therefore the failure of α -anillin antibody to immunoprecipitate during the reciprocal pull downs is not unsurprising, but without this experimental prospect, sound conclusions are difficult to obtain and therefore further exploration and optimisation of potential antibodies which may be better suited for the conditions of immunoprecipitation is required.

KIF2C also held biological plausibility, as a microtubule depolymerization protein with the association holding significant biological ramifications if disrupted. The initial investigation suggested the interaction between CIP29 and KIF2C was recapitulated in human cells and maintained significant reproducibility (Figure 18) although, upon further investigation, the IP analysis of *in vitro* translated proteins did not provide support for a direct association (Figure 19). Furthermore, while using epitope tagged CIP29 and KIF2C proteins, the interaction was not observed (Figure 20). This suggested that the initial interaction may have been disingenuous or that the epitope tags disrupted the interaction between CIP29 and KIF2C itself. Again, early certainty regarding the interaction was limited by the ambiguous results during the reciprocal immunoprecipitations as upon repetition the α -KIF2C antibodies failed to perform well under the required conditions (Figure 20). Further doubt was cast upon the interaction when utilising the CIP29-KO cell lines, which revealed that KIF2C can be immunoprecipitated in the absence of CIP29 (Figure 21). Although this suggests that KIF2C binds to the immunoprecipitation beads or the α -CIP29 antibody in a non-specific manner, this was not the case in the Mock immunoprecipitations or in the α -CIP29 immunoprecipitation of *in vitro* translated CIP29. Therefore, the interaction within the CIP29-KO cells may be facilitated by a nominal amount of CIP29, which is undetectable through western blot analysis, or a secondary protein that is similar enough to CIP29 for cross reaction with the α -CIP29 antibody beads, has facilitated the interaction between the α -CIP29 antibody and KIF2C. This study highlights that the choice of controls, antibodies, or epitopes and the availability of advanced technologies, can significantly influence the scientific investigation.

4 Chapter 4: Exploring the phenotype caused by loss of CIP29.

4.1 Introduction

As immunoprecipitation studies suggested that CIP29 did not associate with any of the additional proteins highlighted during the mass spectrometry analysis, any influence that CIP29 has on cellular functions is perhaps most likely through its association with the TREX or AREX complexes and the exportation of key genes that drive those related cellular functions.

CIP29 has been linked with cell proliferation and cell cycle regulation, with upregulation of CIP29 seen in highly proliferating cells. (Pryor et al., 2004) (Fukuda et al., 2002). Although there is contrasting data which suggests that over expression of CIP29 can result in G2/M phase arrest and growth inhibition (Leaw et al., 2004). These conflicting data may be due to differing experimental procedures, or CIP29 may form part of a negative feedback loop, which regulates the cell cycle (Fukuda and Pelus, 2005). Further to this, our own laboratory's data suggested that the loss of CIP29 resulted in significant cell cycle defects, whereby cells failed to progress normally through cell division. There is also biological plausibility that CIP29 may be involved in the DDR which may again indirectly impact the cell cycle through the activation of cell cycle checkpoints.

4.1.1 *Project objective: Exploring the phenotype caused by loss of CIP29.*

At present, the exact role of CIP29 remains unclear therefore, to begin to clarify its functionality, mode of action, and the biological systems it influences, mutant cell lines of CIP29 were developed within the laboratory. This ensured that the cellular responses in the absence of CIP29 and following its reintroduction could be studied in greater detail. These cell lines were originally derived from the parental MRC5-VI transformed lung fibroblast cell line and included a CIP29-Knockout (CIP29-KO) cell line in which the CIP29 gene was disrupted, and its expression abolished through CRISPR-Cas-9 gene targeting (refer to Armstrong thesis). Rescue cell lines, CIP29-RESCUE (W6) and FLAG-CIP29-RESCUE (F17), had also been developed immediately prior to this project, whereby CIP29 or CIP29 coupled to an N-terminal FLAG epitope were re-expressed in the CIP29-KO cell line following stable transfection (Taylor, personal comm.). Using these cell lines, the role of CIP29 within the cell cycle, cell proliferation and the DNA damage response will be assessed using cell proliferation assays, flow cytometry and immunofluorescence studies.

4.2 Results

4.2.1 Validation of CIP29 knockout and CIP29 rescue cell lines.

Before any further analysis could be conducted, previously produced CIP29-KO, W6 and F17 cell lines required validation to confirm the absence or presence of CIP29. Therefore, CIP29 protein expression levels of each cell type were determined by western blotting and compared against the parental MRC5-VI cell line (refer to methods 2.3.6-9) (FIGURE 22A). Western blot analysis revealed that the CIP29 protein was observed as expected in the control MRC5-VI cell lines, at a molecular weight of ~29 kilodaltons (kDa). In contrast, the CIP29-KO cells did not display any expression of the CIP29 protein although, expression was detected in the rescue cell lines W6 and F17. Furthermore, the CIP29 protein levels in both rescue cell lines were comparable to the endogenous CIP29 observed within the parental control. FLAG-CIP29 in F17 cells can be detected as a slightly higher band because of its increased molecular weight due to the additional FLAG peptide at the N-terminus of CIP29.

To further demonstrate that the reintroduced CIP29 within the rescue cell lines could function comparably to the endogenous CIP29 protein of MRC5-VI cells, the subcellular localization of CIP29 was examined within each of the cell lines. Subcellular fractionation was performed as described (refer to methods 2.3.4). Following this the cytoplasmic, membrane, soluble nuclear, chromatin-bound and cytoskeletal fractions were analysed through western blotting. To ensure that each subcellular fraction was successfully distinct from carryover from the previous, each fraction was probed for proteins only found within that subcellular fraction. These included GAPDH for the cytoplasmic and histone H3 for the chromatin-bound fractions. The results obtained suggested that the endogenous CIP29 within the MRC5-VI controls localised to the cytoplasmic, membrane, soluble nuclear and chromatin-bound subcellular fractions, but was not observed within the cytoskeletal fraction. As expected, CIP29-KO cell line lacked CIP29 protein as it was not significantly observed in most of the sub-cellular fractions, although a very faint detection of residual CIP29 is observed within the whole cell extract and the cytoplasmic fraction. When investigating the expression pattern of the re-introduced CIP29 protein in the rescue cell lines W6 and F17, the expression profile and localisation was consistent with the endogenous CIP29 of the MRC5-VI control. This suggests that the subcellular localisation of the re-expressed CIP29 behaved normally and was unaffected by the addition of the N-terminal FLAG peptide (Figure 22B).

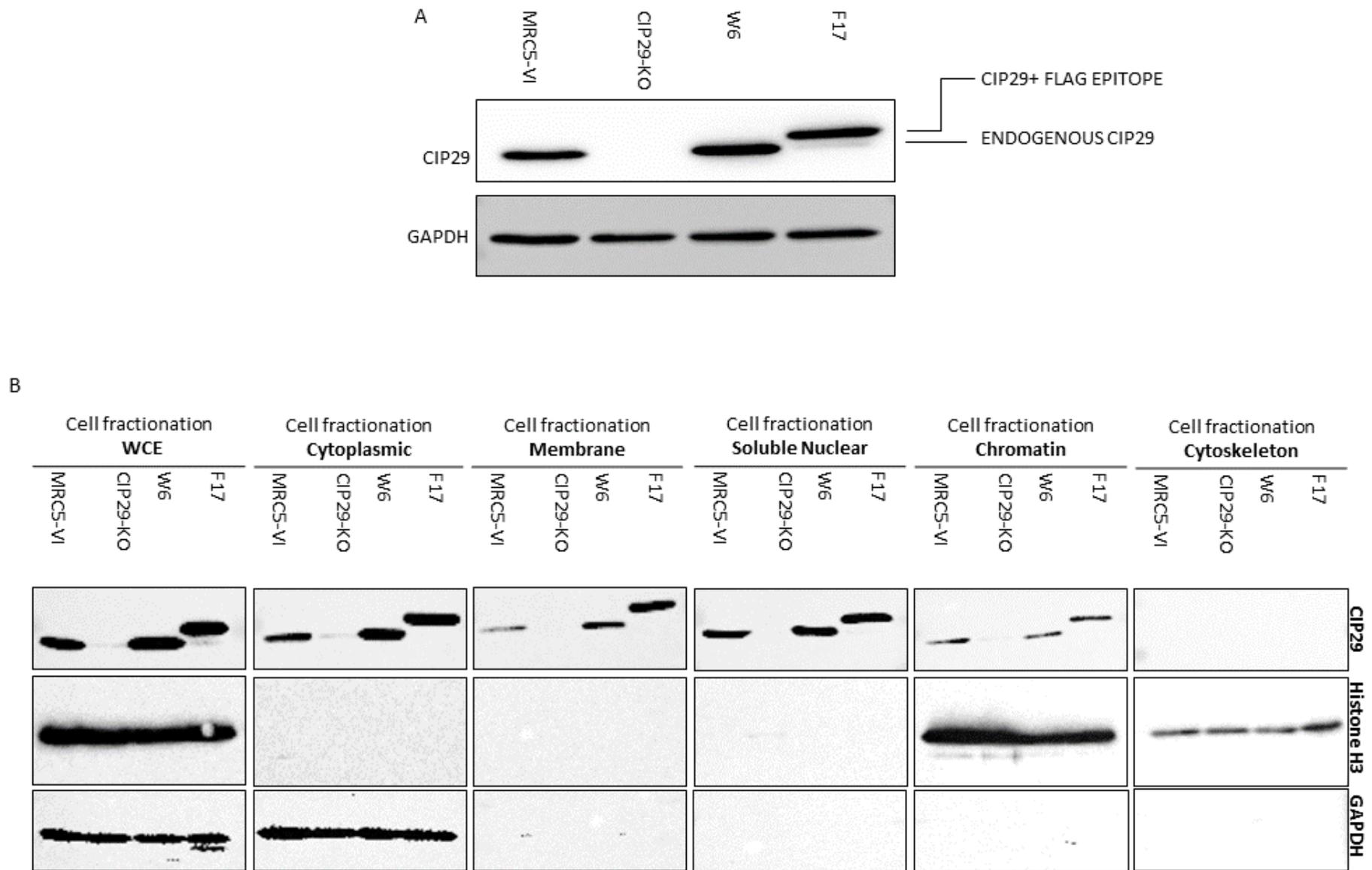


Figure 22: Validation of CRISPR-Cas9-induced Functional Knockout and Rescue of the CIP29 protein through western blot and protein localisation studies.

Western Blot shows the protein expression of CIP29 within each cell type; MRC5-VI control, CIP29-KO (CIP29-KNOCKOUT), W6 (CIP29-RESCUE), F17 (FLAG-CIP29-RESCUE). Specific proteins of interest that were detected are indicated alongside (R=3) (A). Western blot of Cell fractionation for each cell type; MRC5-VI control, CIP29-KO (CIP29-KNOCKOUT), W6 (CIP29-RESCUE), F17 (FLAG-CIP29-RESCUE). Cell fractionations include cytoplasmic, membrane, soluble nuclear, chromatin-bound and cytoskeletal. This is alongside whole cell extracts (WCE) of each cell type. Each fraction has been probed for specific proteins that should only be present in that individual fraction and are indicated alongside (R=2) (B).

4.2.2 Investigating the role of CIP29 in cell proliferation

As CIP29 had previously been implicated in cell proliferation and its increased expression was correlated to cancer progression, the effect of its absence and subsequent reintroduction was analysed by investigating the proliferation rate of each of the four cell lines using the WST-1 commercial proliferation assay (refer to methods 2.5.7). The WST-1 assay relies on the cleavage of the tetrazolium salt to formazan by cellular mitochondrial dehydrogenases. The WST-1 components within the assay stain the resulting formazan, enabling its detection as a measure of metabolic cellular activity. As the stained formazan directly corresponds to the number of viable cells, it can then be used to quantify the populations' proliferation rate. The results obtained suggested that the CIP29-KO cells proliferated more slowly than the MRC5-VI control, while the two rescue cell lines, either matched (F17) or slightly exceeded (W6) the proliferation rate of the MRC5-VI cell line (Figure 23). This suggests that the lack of CIP29 does impact cell proliferation in the CIP29-KO cells and that this negative effect can be overcome when CIP29 expression is restored. It is possible that the greater proliferation rate of W6 maybe a consequence of a slightly elevated expression levels of CIP29, compared to the endogenous CIP29 levels found within MRC5-VI. Unfortunately, the assay exhibited quite a large amount of variation between each replicate, so it is difficult to determine if the proliferation rate of the CIP29-KO is significantly less than the MRC5-VI and CIP29-Rescue controls. Ultimately, the assay may require further optimisation or additional replicates to reduce this variation.

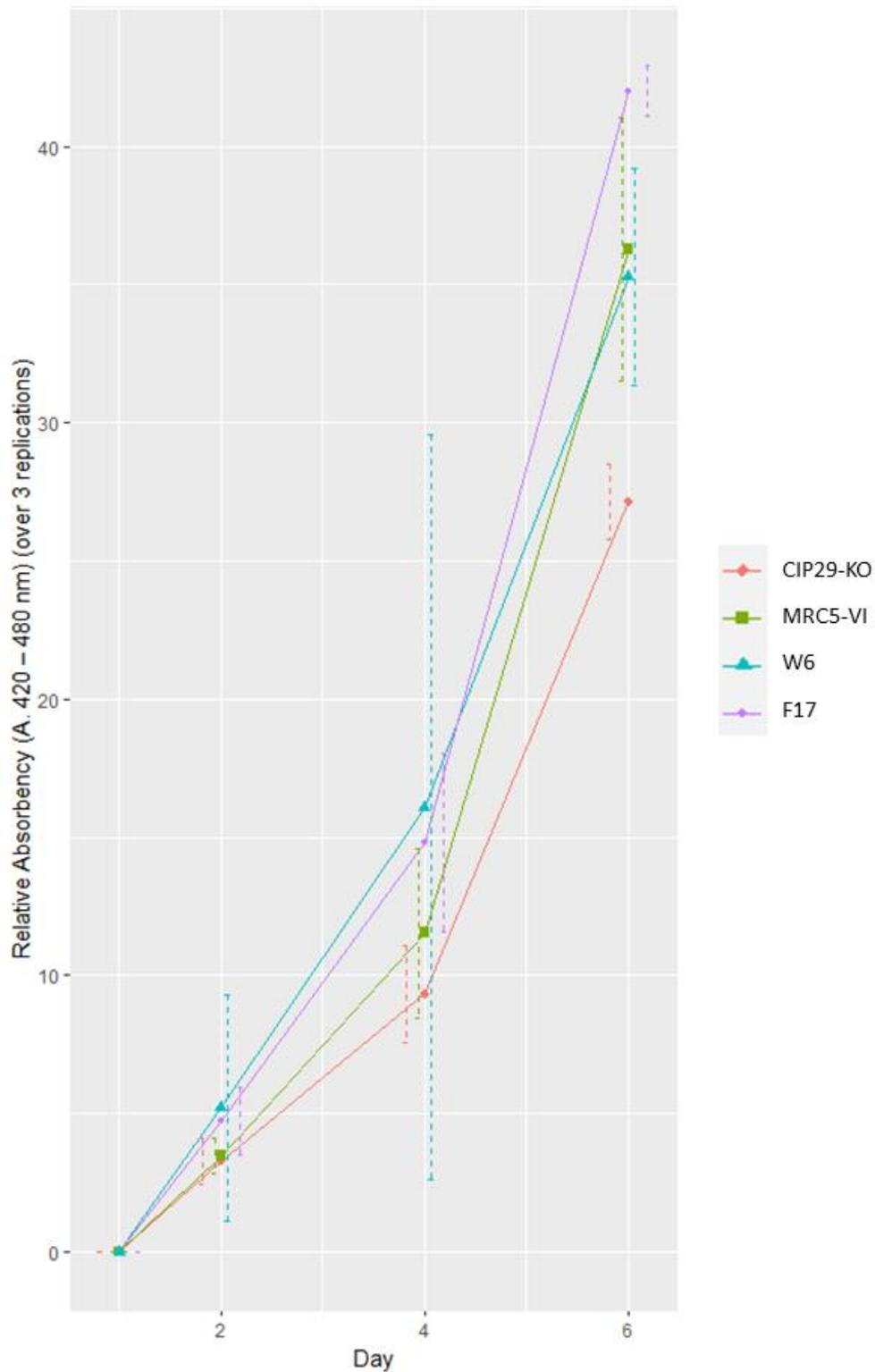


Figure 23: Measurement of the cell proliferation for MRC5-VI, CIP29-KNOCKOUT (CIP29-KO), CIP29-RESCUE (W6) and CIP29-RESCUE-FLAG (F17) using WST-1 assay.

Line graphs showing the relative absorbency of the formazan+WST-1 product produced over a period of 6 days. Each cell type is highlighted in colour CIP29-KO (CIP29-KNOCKOUT) (red), MRC5-VI (green), F17 (FLAG-CIP29-RESCUE) (blue) and W6 (CIP29-RESCUE) (purple). Absorbency readings were taken every 48 hours for each type. Each data point represents an average for each biological replicate, as well as a standard deviation range (dashed) with a confidence interval of 95% (R=3).

4.2.3 Investigating the role of CIP29 in cell cycle progression

Previous flow cytometric analysis of CIP29-KO cells within our own laboratory strongly supported a role for CIP29 within cell cycle progression. CIP29-KO cells were shown to accumulate within G2/M phases of the cell cycle and a proportion of the cells displayed a greater than 4N DNA content (see Introduction Figure 10B) (Armstrong, 2016). However, it was unclear if these cell cycle defects could be rectified following re-introduction of CIP29. Therefore, cell cycle analysis through flow cytometry was performed on MRC5-VI, CIP29-KO, W6 and F17 cell lines. Propidium iodide (PI) staining was utilized to highlight the major phases of the cell cycle through the difference in DNA ploidy (refer to methods 2.6.2 and 2.6.4). Cell populations can be separated into pre-replicative (G0/G1), whereby the cells possess a normal diploid chromosomal number and hence hold DNA content of 2n; the replicative (S) population, in which the population has a DNA content ranging between 2n and 4n as DNA is synthesized, and a post-replicative + mitotic population (G2/M) which contains a DNA content of exactly twice the 2n amount at 4n. Anything greater than 4n is considered abnormal as once the cells have completed mitosis and divided into subsequent daughter cells, their DNA content should return to 2n. As late S-phase and the G2/M phases of the cell cycle have a similar DNA content it can be difficult to discriminate between these phases based on PI staining alone, although it can quickly and effectively identify >4n populations and can be combined with additional cell cycle-specific markers to further delineate different cell cycle stages.

Upon analysis, the flow cytometric profiles of MRC5-VI and CIP29-KO cells were consistent with previous data, whereby the MRC5-VI cell cycle profiles indicated that most cells (approximately 56%), were in G1 phase with a DNA content of 2n, just under 20% were found to be in S phase and had DNA content between 2n and 4n, while almost all the remaining cells, (26%) had a 4n DNA content and were situated in the G2/M phase of the cell cycle. A very small proportion of cells (1%) were found to exhibit an abnormal increased ploidy which is highlighted in pink. In contrast, the CIP29-KO cells exhibited a cell cycle profile that was profoundly different. 29% of cells were found within G1 and 7% of cells in S phase. However, most cells (46%) were situated in the G2/M phases of the cell cycle, while a substantial proportion of cells, just under 20%, displayed abnormal ploidy. The cell cycle profiles of W6 cells strongly mirrored the MRC5-VI control whereby 56% of cells were found within G1, 15% within S phase and almost all the remaining cells 27%, were found within G2/M. Again, a very small 1% of cells were found to exhibit an abnormal >4n DNA content. F17 cells were similar, with 50% of cells situated in G1, 8% in S phase and 38% within G2/M. There is a slightly greater proportion of F17 cells in G2/M than for the MRC5-VI or W6 cells, although there are still only very few cells (3%) that exhibit abnormally increased ploidy. This may suggest that the W6

rescue is slightly more effective than its F17 counterpart. As very few cells within the rescue cell lines exhibited abnormal ploidy and the cell cycle profiles resembled the MRC5-VI control, it indicates that the re-introduction of CIP29 was able to recover the abnormalities of the CIP29-KO confirming that the loss of CIP29 is responsible for the cell cycle defects seen in CIP29-KO cell line (Figure 24).

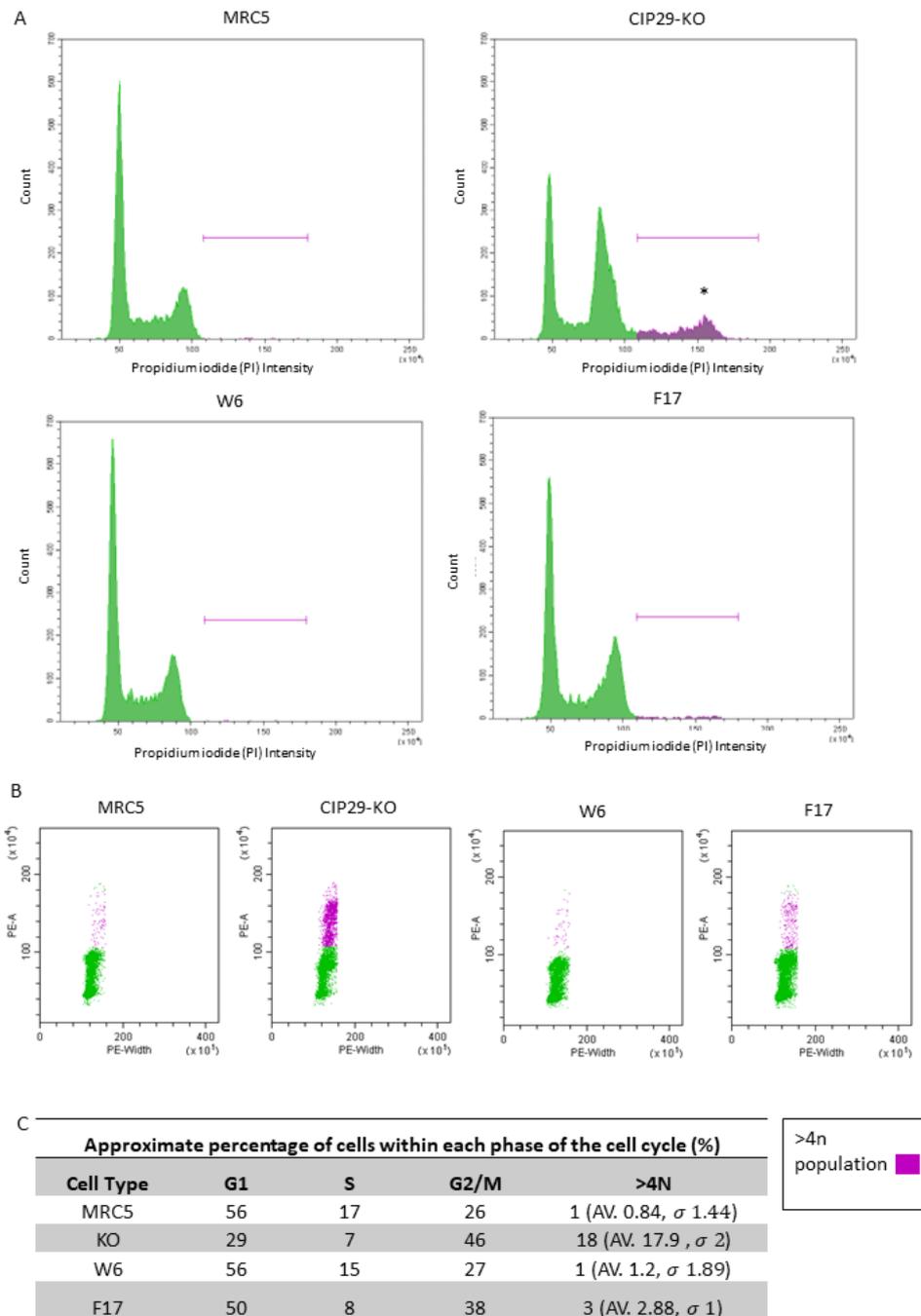


Figure 24: Flow Cytometric analysis through propidium iodide, (PI) staining of the wildtype (MRC5-VI), CIP29 knockout cells (CIP29-KO) and CIP29-RESCUES (W6 and F17).

Cells cycle profiles of wild type (MRC5-V1), CIP29 knockout (CIP20-KO), CIP29-RESCUE (W6) and FLAG-CIP29-RESCUE (F17) are illustrated as population histograms, DNA content (propidium iodide, PI intensity) versus number of cell counts (counts) (A). Dot plots of single cells, a cells' area (PE-A) in relation to its width (PE-width) (B). >4n populations are highlighted in pink with a sub-population of note marked (*). Table of the approximate percentage of cells within each cell cycle phase (G1, S, G2/M and >4n) (C) (R= 3, N: 10,000 cells).

4.2.4 *Does the absence of CIP29 induce G2/M checkpoint activation?*

As the flow cytometry data suggested that the CIP29-KO population exhibited a large proportion of cells within the G2/M phases of the cell cycle when compared to the MRC5-VI controls, the potential cause of this accumulation required further investigation. One possibility for this increase is that the CIP29-KO cells experienced some level of cell cycle arrest at the G2/M transition, due to activation of the DNA damage checkpoint following loss of CIP29. This is plausible due to the multiple reports that suggest that defective RNA processing factors often promote genome instability. Therefore, the loss of CIP29 could potentially give rise to DNA damage, which in turn may lead to DNA damage-dependent cell cycle arrest. Further to this, previous data from our laboratory has demonstrated that CIP29 is phosphorylated in response to DSBs, which may be an indication that it is regulated in response to DNA damage and, perhaps, that CIP29 functions to ameliorate DNA damage, although this is still an assumption (Holden et al, 2017).

To clarify this, any evidence of DDR activation in the absence of CIP29 was investigated by examining each cell line for the activation of key DNA damage markers (p53, histone H2AX) through western blot analysis. The Ser15 motif of p53 is a primary target of the DNA damage response and is phosphorylated by both the ATM and ATR protein kinases to regulate cell cycle arrest and apoptosis (Allison and Milner, 2003), while histone H2AX phosphorylation stimulates the assembly of DNA repair proteins at damaged chromatin sites and promotes the activation of cell cycle checkpoints to arrest the cell cycle (Podhorecka et al., 2010). As a positive control, MRC5-VI cells were treated with etoposide, which activates the cellular response to DNA damage through the inhibition of topoisomerase II (TopII) enzymes. TopII enzymes manipulate the DNA to reduce torque-tension during DNA replication and transcription, by cutting the DNA strands to enable one strand to pass over the other before religating the cut site. Etoposide can readily insert into the DNA cleavage site formed by the TopII enzyme, where it prevents TopII from religating the cleavage sites resulting in DSBs and the activation of the ATM, H2AX and p53 DNA damage response pathways (Li et al., 2013). This provides a reliable positive control for the presence of DSBs and the activation of these pathways. Whole cell extracts were prepared from etoposide-treated MRC5-VI cells and from untreated MRC5-VI, CIP29-KO, W6 and F17 before analysis through Western blotting (refer to methods 2.3.3, 2.3.6-9).

Although the etoposide-treated MRC5-VI controls exhibited both p53-Ser15 and histone H2AX phosphorylation, none of the untreated cells showed activation of the DNA damage response by these readouts. This absence of DNA damage-dependent phosphorylation of p53 and histone H2AX in CIP19-KO cells indicates that loss of CIP29 does not activate the DNA damage response pathway

through these avenues and there is no evidence that the CIP29-KO cells are arrested at the G2/M DNA damage checkpoint (Figure 25).

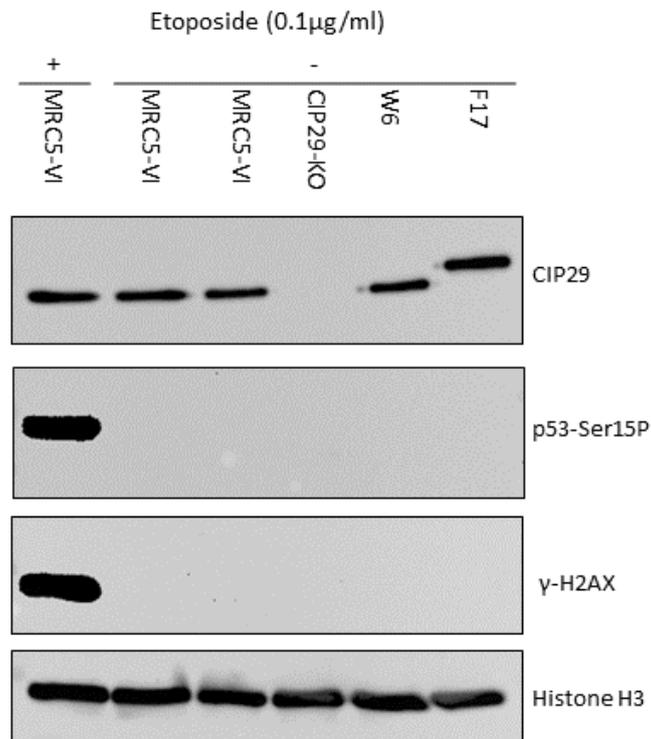


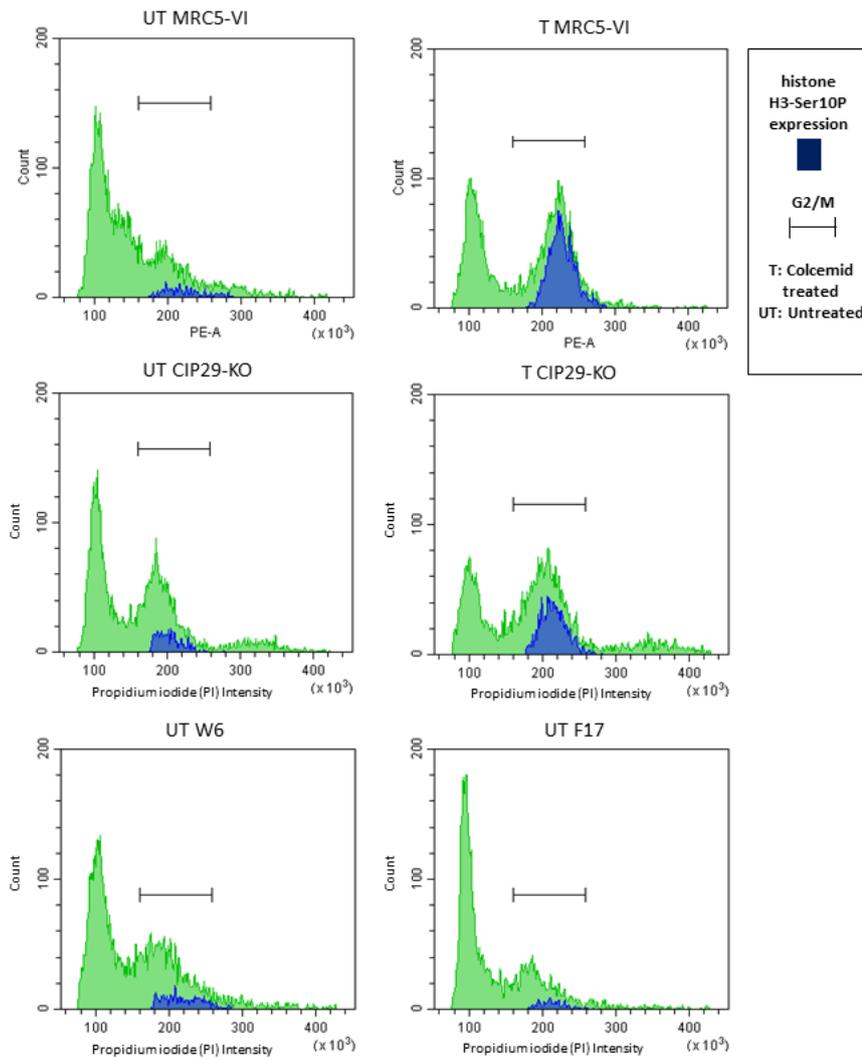
Figure 25: Analysis of the DNA damage response of each cell type through the activation of DNA damage markers

Western blot shows the DNA damage response of each cell type, wild type (MRC5-VI), CIP29-KNOCKOUT (CIP20-KO), CIP29-RESCUE (W6) and FLAG-CIP29-RESCUE (F17), alongside a 0.1 ug/ml Etoposide treated MRC5-VI as a control. Specific antibody probes are illustrated alongside (R=2).

4.2.5 Does the absence of CIP29 induce a mitotic delay?

The DNA damage response analysis suggested that CIP29-KO cells were not arrested at the G2/M checkpoint, which suggests the cells may instead be accumulating within mitosis. To test this assumption, flow cytometric analysis was performed using PI staining in combination with the detection of histone H3, phosphorylated on Ser10 (M&M 2.6.2-4). This phosphorylation is seen as a crucial event for the onset of mitosis and is linked to chromosome condensation and segregation. Histone H3-Ser10 phosphorylation continues until anaphase, when it begins to be rapidly dephosphorylated and remains so throughout the remainder of interphase (Crosio et al., 2002). Therefore, an increase in histone H3-Ser10 phosphorylation should give an indication as to whether the cells are blocked in mitosis. As a further control, MRC5-VI cells were blocked in mitosis, at metaphase, using colcemid treatment. Colcemid does not affect the rate at which cells enter mitosis, but when applied before nuclear envelope breakdown the drug completely inhibits the formation of microtubules (Rieder and Palazzo, 1992). Consequently, the chromosomes are released into the cytoplasm where they remain randomly dispersed inhibiting further progression through mitosis. Because of these properties, colcemid treatment offers a true representation of M phase arrest.

Like previous observations, the PI staining indicated that CIP29-KO cells maintained a large proportion of cells within the G2/M phases of the cell cycle and a sub-population of cells that exhibited a >4N DNA content (Figure 26 A/B). Colcemid treatment of both MRC5-VI and CIP29-KO cell lines resulted in an increase in the proportion of cells with a G2/M DNA content. The colcemid-treated MRC5-VI and CIP29-KO cells also showed an increase in the proportion of histone H3-Ser10^P-positive cells from 7% to 33% and from 10% to 23% respectively. This suggests that the colcemid treatment was successful and that a significant proportion of these cells are arrested within mitosis. This increase in histone H3-Ser10^P levels was not observed in the untreated MRC5, W6 or F17 cell lines with histone H3-Ser10^P positive cells remaining at around 10% of the total population, accounting for around 20-25% of the G2/M population. Although the untreated CIP29-KO cells exhibited a DNA content distribution profile that was similar their colcemid treated counterpart, they did not show a significant increase in histone H3-Ser10^P levels, in comparison with the other untreated cell types, with 10% of the total population and 26% of the G2/M cell populations respectively being histone H3-Ser10^P positive (figure 26 (Table)). Furthermore, the histone H3-Ser10^P staining did not encompass the >4N population. This suggests that despite the greater number of CIP29-KO cells accumulated in the G2/M phases of the cell cycle, based on through their DNA content, most of these cells do not contain detectable histone H3-Ser10^P. Therefore, these cells cannot be in the early phases of mitosis, despite their 4N DNA content.



Percentage of histoneH3-Ser10P for Each Cell Type		
Cell type	histoneH3-Ser10P as a % of total population	histoneH3-Ser10P as percentage of G2/M population (%)
T MRC5-VI	33 (AV. 35.4)	63 (AV. 74)
T KO	23 (AV. 20.6)	43 (AV. 53)
UT MRC5	7 (AV. 9.7)	24 (AV. 50)
UT KO	10 (AV. 8.7)	26 (AV. 38.5)
UT W6	6 (AV. 5.5)	21 (AV. 33.5)
UT F17	10 (AV. 6.5)	28 (AV. 35)

Figure 26: Flow Cytometric analysis of each cell type investigating the expression of a key mitotic marker.

Flow cytometry population histograms show the cell cycle profiles of each cell type wild type (MRC5-V1), CIP29- KNOCKOUT (CIP20-KO), CIP29-RESCUE (W6) and FLAG-CIP29-RESCUE (F17) are illustrated as population histograms, DNA content (propidium iodide, PI intensity) versus number of cell counts (counts). Histograms also highlight the histone H3 expression level (blue), and the GM/2 cell cycle phase. For comparison, Colcemid treated MRC5, and CIP29-KO population histograms are also presented (A). Table detailing the approximate percentage of cells for each cell type that expresses histone H3 within the total cell population and the G2/M only phase of the cell cycle (B) (R=2) N: 10,000 cells).

As no relative increase in histone H3-Ser10P was observed within the CIP29-KO cells it suggests that the increased proportion of cells in the 4n DNA population in this cell line may be situated in one of two phases of the cell cycle. For example, the cells may have duplicated their DNA in S-phase but and are yet to express phosphorylated histone H3 as they are slowed or held within the early stage of G2. As a complete arrest is not seen in the cell cycle profiles, this scenario suggests a very slow progression from the end of S-phase to late G2 when histone H3 would begin to be phosphorylated. A second possibility is that the CIP29-KO cells with a 4n DNA content may have already passed metaphase, and instead the cells are struggling to divide at the end of mitosis. In this scenario, they would maintain their higher DNA content within a single undivided cell, but the majority of histone H3 would have already been dephosphorylated by this point. If these cells were not lost from the population having failed to divide, they may be prevented from progressing further through to another cell cycle and be held in G1. The cells would then be mis-allocated to the G2/M phases of the cell cycle based on their 4n DNA content, but these cells would not express histone H3-Ser10P. Some of these cells may even exit G1 and continue through the cell cycle to replicate their DNA again. This would account for the small population of cells which display a >4N DNA content. To investigate the assumption that the CIP29-KO cells may fail to complete cytokinesis, immunofluorescence staining, and confocal microscopy was utilised to study the cells' morphology for discrepancies that are characteristic of this defect.

4.2.6 Does the absence of CIP29 induce cytokinesis defect?

If the CIP29-KO cells have failed to complete cytokinesis, the cells may exhibit morphological defects such as bi-nucleates, larger nuclei or multiple centrosomes. They may also exhibit micronuclei or nuclei buds, although these characteristics often coincide with DNA damage which we have not detected. To investigate the characteristics of each cell type and the potential occurrence of cytokinesis failure, immunofluorescence and confocal microscopic studies were conducted using DAPI and γ -tubulin staining. At first glance, some striking features were seen regarding the morphology of the CIP29-KO cells. For example, the overall size of the cells seemed much greater than the parental MRC5-VI cells and the W6 and F17 rescue cells (Figure 27A). Some cells also exhibited other morphological abnormalities such as multi-lobed nuclei and micronuclei, while a large proportion of cells exhibited an excessive number of centrosomes. However, very few fully formed binucleated cells were observed. These abnormalities were not observed in the MRC5-VI cells and were only very rarely seen in the W6 and F17 rescue cell lines (Figure 27B).

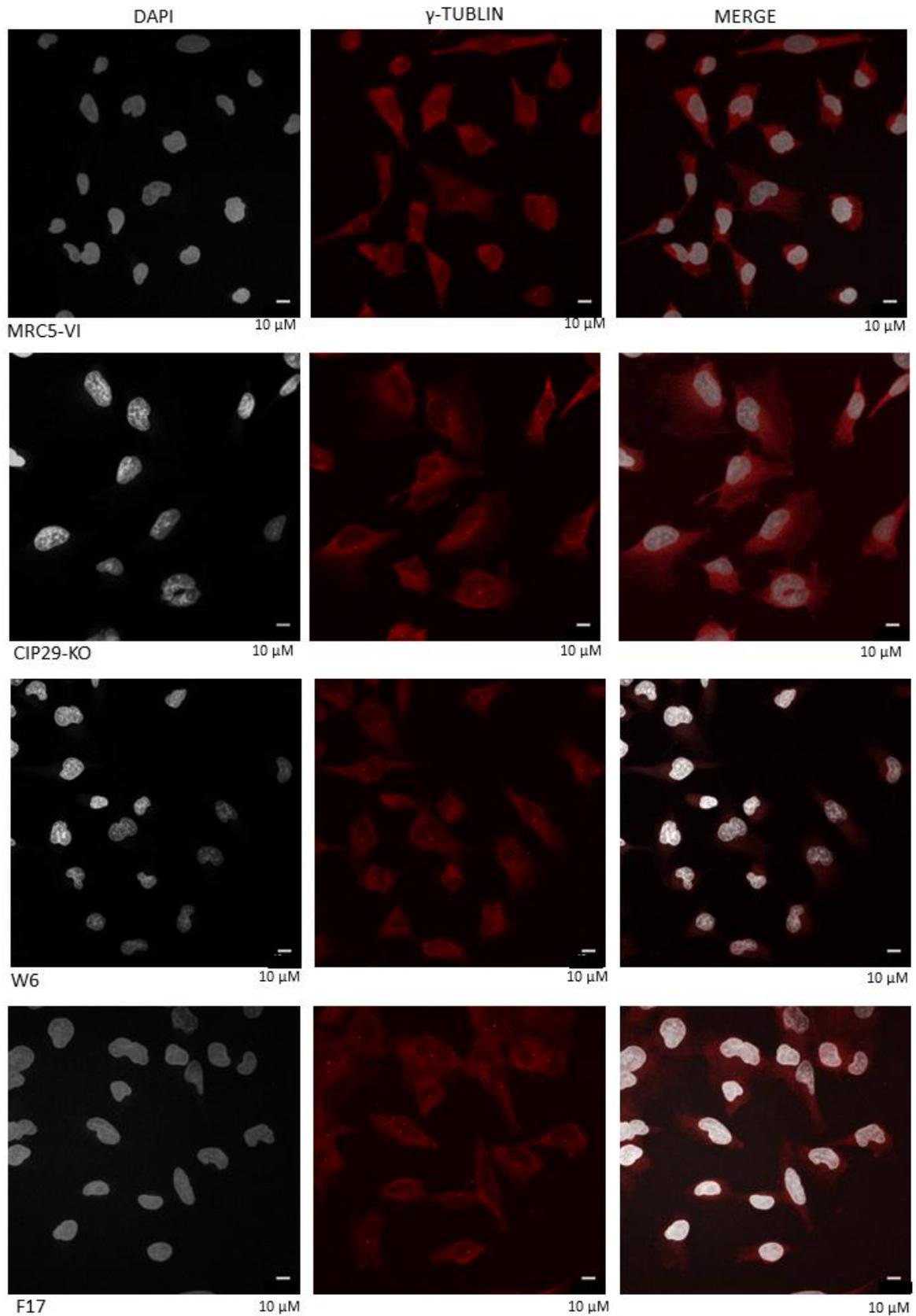
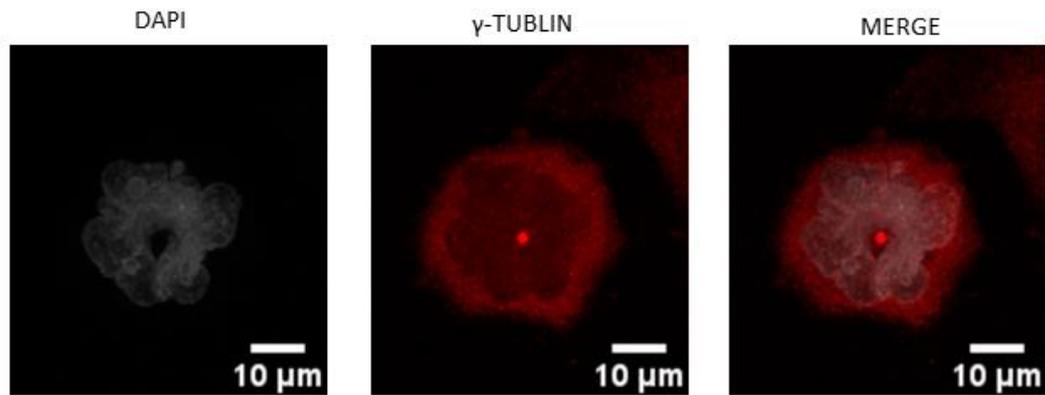
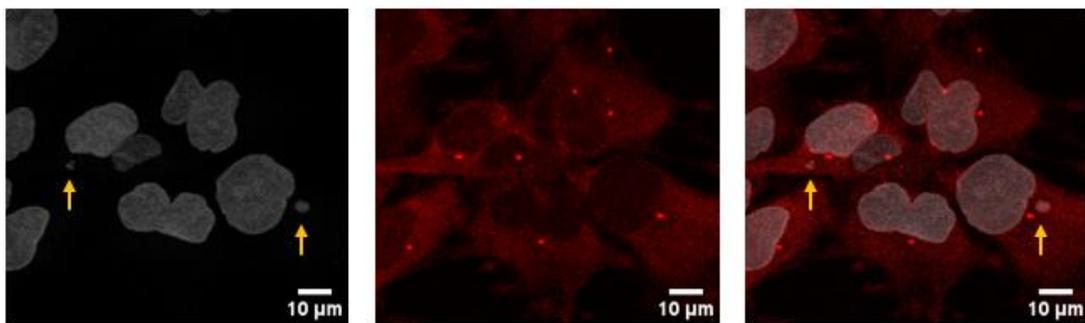


Figure 27A: Immunofluorescence analysis of MRC5, CIP29-knockout (CIP29-KO), CIP29-Rescue (W6) CIP29-Rescue-flag (F17) cell lines.

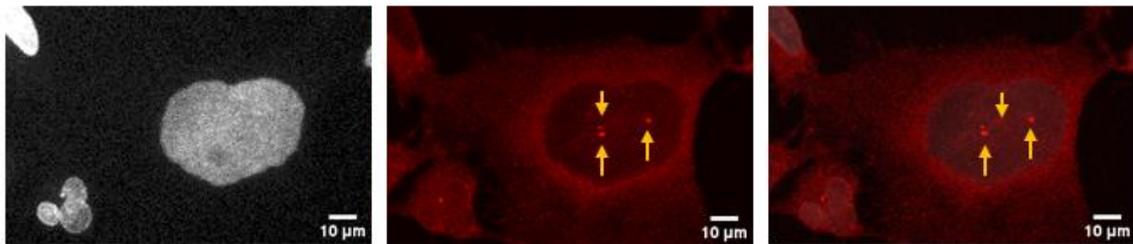
Each cell type, wild type (MRC5-VI), CIP29- KNOCKOUT (CIP20-KO), CIP29-RESCUE (W6) and FLAG-CIP29-RESCUE (F17) are stained with either DAPI (grey) or Gamma (γ -)tubulin (red), alongside a merge of both stains. Magnification: 40x, Scale: 10 μ m (R=3)



Example of a multi-lobbed nucleus within a CIP29-KO cell



Example of micro-nuclei within CIP29-KO cells



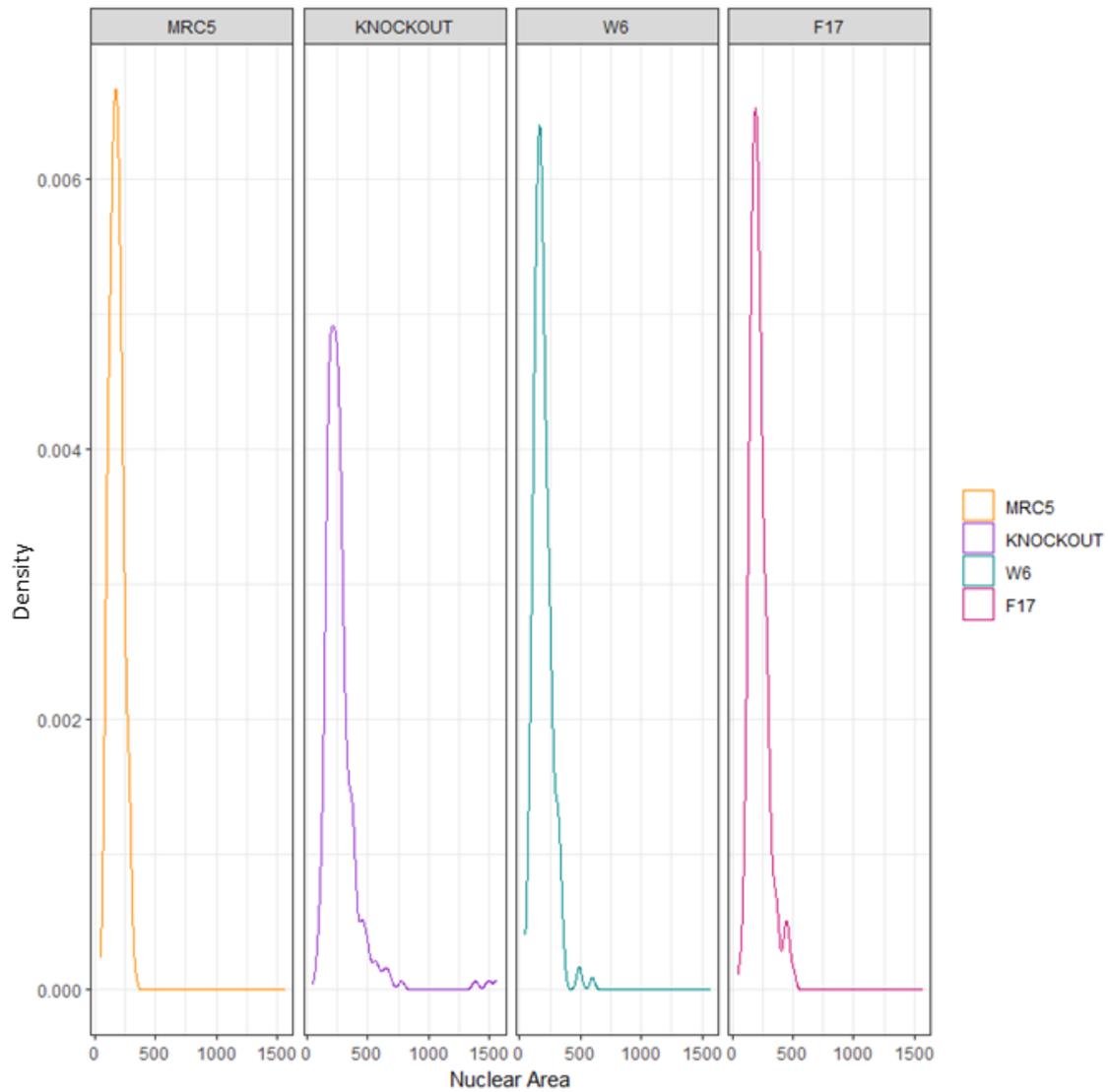
Example of multiple centrosomes a CIP29-KO cell

Figure 27B: Immunofluorescence analysis of abnormal CIP29-knockout (CIP29-KO) cells.

CIP29- KNOCKOUT (CIP20-KO) are stained with either DAPI (grey) or Gamma (γ -)tubulin (red), alongside a merge of both stains. Abnormalities such as multi-lobbed nuclei, micro-nuclei and multiple centrosomes are highlighted (arrow). Magnification: 40x, Scale: 10 μ m (R=3)

4.2.7 *Do the CIP29-KO cells exhibit larger nuclei?*

As initial analysis suggested that CIP29-KO cells contained larger nuclei than their counterparts, this required further investigation. Therefore, the nuclear area of each cell was examined and compared between cell types. The results obtained confirmed the previous observations, as even the minimum average nuclei size of the CIP29-KO cells outpaced the other groups, while the maximum values vastly exceed the other groups. Furthermore, the CIP29-KO cells potentially exhibited the beginnings of a bimodal distribution, due to the number of cells that possessed a nuclei area that exceeded $800\mu\text{m}^2$. This secondary population is not seen in the other cell populations, all of which maintained a uniform distribution (Figure 28A). The excessively large nuclei of this sub-population within the CIP29-KO cells suggest that these cells have undergone multiple rounds of DNA replication without fully completing the cell cycle and may exhibit the $>4N$ DNA content seen within the flow cytometry data. The potential for a bimodal distribution has implications for the statistical tests available when comparing each of the populations and limits the statistical analysis to non-parametric tests. Such tests take the different distribution types and the consequent standard deviation into account and will raise the threshold for statistical significance to accommodate this (Figure 28B (Table)).



B: CIP29-KNOCKOUT cells exhibit a large variation of nuclei sizes within their population

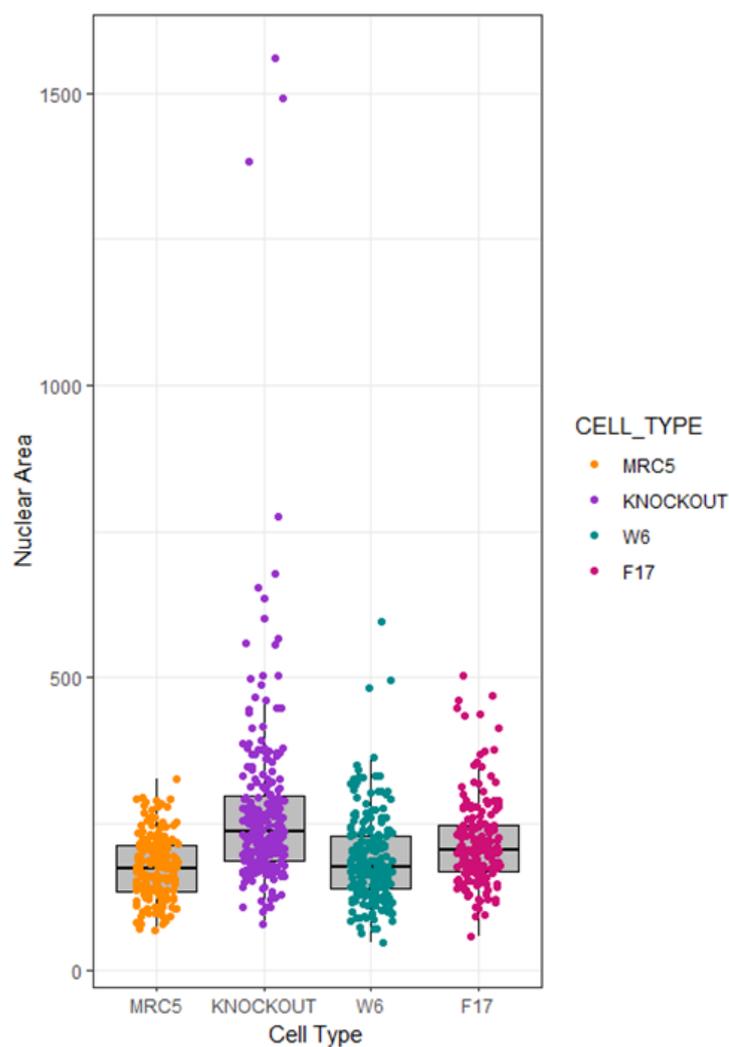
CELL TYPE	Media n	Mean	MIN	MAX	Standard Deviation
MRC5-VI	175	176	67	328	54.71
CIP29-KO	238	275	78	1559	169.47
W6	176	189	47	596	75.73
F17	205	218	57	505	76.76

Figure 28 CIP29-KO cells maintain a distinct population distribution which is not observed within their counterparts.

Density plots showing the distribution of nuclei area for each cell type. Each cell type is colour coded (A). Table displays the median, mean, minimum, maximum and standard deviation values for the nuclei area for each cell type: MRC5-VI, CIP29-KNOCKOUT, CIP29-RESCUE (W6) AND FLAG-CIP29-RESCUE (F17) (B) R= 3 (collated samples, N= ~250 cells).

Box plots and Wilcoxon non-parametric test of each of the cell types reveal the true extent of the differences between the CIP29-KO and the other cell types. Box plots of the MRC5-VI control display a tight collection of values regarding the nuclear areas which is mirrored within the W6 cell population with the overall averages comparable, despite a few larger values present within the population. In contrast, the CIP29-KO population displays a wide spread of values whereby the true difference in the potentially bimodal distribution can be observed. Although F17 maintains a smaller spread of values, which is comparable to the range of its W6 counterpart, it is the number of slightly larger values that force its characterisation towards that of the CIP29-KO population. This suggests that some restoration of function is achieved within the F17 cell line, although not to the same extent of the W6 rescue cell line (Figure 29A).

Statistical analysis reveals that despite accounting for the bimodal distribution of the CIP29-KO population, the large nuclear areas within this population are significantly unique when compared to the other cell types. In contrast, the MRC5-VI control and the rescue line W6 are not significantly different which further suggests that the re-introduction of CIP29 has been successful within this cell line and the cells within this population are able to regain functionality. When comparing the MRC5-VI and F17, the two are significantly different. In fact, F17 is significantly different to both the W6 and the CIP29-KO. This suggests that F17 may have recovered enough functionality for these cells to be distinct from the CIP29-KO, however this recovery is not enough to match the functionality of the MRC5-VI controls or the other W6 rescue cell line (Figure 29B Table).



B: Testing the significance of the Nuclear Areas between the cell types using the Wilcox non-parametric test

GROUP 1	GROUP 2	P.ADJUSTED	P. Values	Significance
KNOCKOUT	MRC5	4.1 e-21	>0.001	****
KNOCKOUT	W6	1.56e-17	>0.001	****
KNOCKOUT	F17	8.73e- 6	>0.001	****
MRC5	W6	2.57e- 1	0.257	NS
MRC5	F17	1.15e- 6	>0.001	****
W6	F17	7.36e- 5	>0.001	****

Figure 29: Box plots detailing the nuclei area of each cell type.

Box plots show the nuclei area (μm^2) for each cell type with individual readings overlaid. Each cell type is colour coded. (A) Table displays the results from a Wilcox non-parametric test which compares the nuclei area of each cell type and their significance. Significance is indicated as '****' 0.0001 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 'NS' not significant (B) (collated samples, N= ~250 cells).

Another striking morphological feature of the CIP29-KO cell line is the accumulation of centrosomes within the cells that exhibit larger nuclei. Centrosomes are major microtubule organizing centres that arrange the number, stability, polarity and localisation of microtubules to facilitate chromosome segregation and other cellular processes. Therefore, centrosomes play a significant role in the success of mitosis. Although a cells' centrosome cycle is synchronised with the cell cycle, it is a distinct process which relies on a semi-conservative duplication much like that of DNA replication. Like DNA, every centriole must duplicate once and only once during the cell cycle to maintain cell cycle control and remain error free.

In cycling cells, a G1 cell possesses a single centrosome that contains two centrioles. It is not until early S-phase when a new procentriole forms adjacent to each pre-existing centriole and begins to elongate. By the G2-M transition, the two centrosomes containing both a parent and a daughter centriole, separate to facilitate the formation of the two spindle poles. At anaphase, the microtubule spindle divides the centrosomes to ensure that each emerging daughter cell will inherit one copy, ready for the next round of the cell cycle (Vitre and Cleveland, 2012). Deregulation of centrosome duplication and maturation can impair the accurate inheritance of genetic material, in fact the link between numerical abnormalities of centrosomes and tumorigenesis was noted almost a century ago (Boveri, 1895) (Scheer, 2014). This was later rediscovered and observed in multiple cancer types such as breast (Lingle and Salisbury, 1999), lung (Jung et al., 2007) and pancreatic (Sato et al., 1999).

To assess each cell line for centrosome accumulation, centrosomes were stained with the γ -tubulin and cells were categorised into; cells that possessed a single centrosome (C1), cells that possessed 2 centrosomes (C2) and cells that maintained more than 2 centrosomes (>2C) (refer to methods 2.7.1-5 and 2.8.2). The results obtained suggested that the MRC5-VI controls experienced no centrosome amplification, with cells either displaying a single centrosome (71%) or 2 centrosomes (29%). In contrast CIP29-KO cells exhibited a high level of centrosome amplification with 46% of cells displaying a single centrosome, 24% exhibiting 2 centrosomes, while almost a third of cells exhibited more than 2 centrosomes at 30%. When examining the two rescue cell lines centrosome amplification was reduced, but it was not completely abolished as in the MRC5-VI controls. W6 cells exhibited 47% of cells with a single centrosome, 32% of cells with 2 centrosomes and 21% with more than 2 centrosomes. The F17 cell line displayed cell populations with 34% of cells containing a single centrosome, 39% with 2 centrosomes while, 27% displayed more than 2 centrosomes (Figure 30).

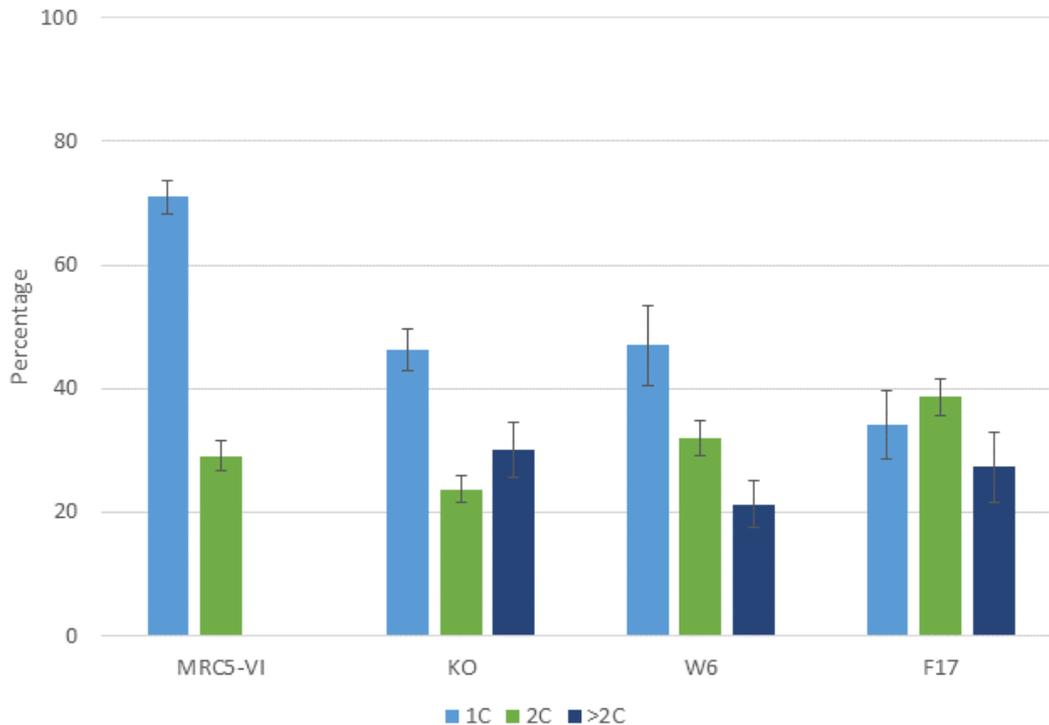


Figure 30: CIP29-KO cells may exhibit centrosome amplification.

Bar charts show the number of centrosomes per cell for each cell type (1C: 1 centrosome, 2C: 2 centrosomes or >2C: more than 2 centrosomes). Standard deviation for each data set is also calculated (A). (N: ~150 cells over 3 replicates).

Centrosome deregulation and the acquisition of extra centrosomes can occur through multiple mechanisms including the deregulation of centrosome biogenesis, cell fusion, mitotic slippage and cytokinesis failure. Fusion events, are unlikely to drive chromosome missegregation as these events are normally limited to development and the terminally differentiated cells that do not continue to divide (Fant et al., 2009). However, mitotic slippage and failure to complete cytokinesis have both been shown to induce centrosome amplification in cancer cells. The multiple centrosomes that are seen within the CIP29-KO cells also seem to populate together. This type of centrosome clustering is often used by cancer cells with multiple centrosomes to avoid cell death and enable bipolar division (Krämer et al., 2011). Furthermore, the location of centrosomes is not random, with a centrosome normally occupying either the centroid or edge of the cell, ready to organize microtubules for the progression of mitosis (Burakov and Nadezhdina, 2020). As many CIP29-KO cells exhibit centrosomes

that are clustered at the centroid, it suggests that these cells may be situated within interphase. But it is unclear if these cells have full control over the additional centrosomes.

As both the larger nuclei and the additional centrosomes seemed to be significant phenotypes for the CIP29-KO cells, the relationship between the two characteristics was examined to establish if the phenotypes influence each other. To do this, a comparative study using Cramér's V measurement of association between the two nominal variables, nuclei size and centrosome number was conducted. The results obtained suggested that the association was of a moderate significance and that the correlation between the two variables would not have occurred by chance (Table 10).

Table 10: Cramer V measure of association indicates that the two variables larger nuclei and centrosome amplification are statistically correlated within the CIP29-KO cells.

Cramer V	Lower confidence intervals	Upper confidence intervals
0.5 (medium association)	0.4	0.6

Pearson's Chi-squared test	
X-squared	64.05
Degrees of freedom	2
p-value	1.235e-14

4.2.8 CDT1 staining and Flow Cytometry

Following the finding that CIP29-KO cells do not have an increased proportion of cells with detectable histone H3-Ser10^P and are therefore not blocked in late G2 through to anaphase, along with the immunofluorescence data which showed that the CIP29-KO cells contain enlarged nuclei and multiple centrosomes, we speculated that a proportion of the cells placed in G2/M by flow cytometric analysis, were in fact G1 phase cells with duplicated DNA content. To test this assumption, all four cell lines MRC5-VI, CIP29-KO, W6 and F17, were stained with an anti-CDT1 antibody alongside PI for further flow cytometry analysis (Refer to methods 2.6.1-4).

CDT1 is a key cell cycle regulator and is required for the delivery of key licensing factors for DNA replication. The recruitment of licensing factors begins during the previous round of the cell cycle, during the latter phase of mitosis known as telophase once the nuclear envelopes have formed around the newly segregated mitotic chromosomes. CDT1 recruitment continues throughout the following G1 phase before finally ceasing at the G1/S-phase transition (Pozo and Cook, 2016). During this process, the Origin Recognition Complex (ORC) recruits the Cdc6 protein to the chromatin and awaits the arrival of CDT1 which is bound to the Mini chromosome Maintenance (MCM) to form a pre-replicative complex (pre-RC). The arrival of CDT1 and the MCM complex is crucial for DNA licensing to continue and subsequent DNA replication to initiate. Aside from delivering the MCM complex to OCR and CDC6, CDT1 may also maintain proper orientation or conformation of MCM when loading on the DNA, as CDT1 remains bound to MCM until DNA loading is completed successfully (Deegan and Diffley, 2016). As the recruitment of CDT1 only occurs in the following G1 phase, the MCM complexes' access to chromatin is limited and ensures that the DNA can only be replicated once per cell cycle. In fact, in the absence of CDT1, MCM complexes are never recruited to the DNA (Maiorano et al., 2000).

Due to the importance of CDT1, its activity is highly regulated, and it is proteolytically degraded through ubiquitin-dependent proteolysis by two E3 ubiquitin ligases, SCF-Skp2 and DDB1-Cul4-Cdt2. The remaining amounts of CDT1 are inactivated by geminin, which tightly binds to CDT1 and prevents its activity, ensuring that the re-initiation of replication is prevented (Wohlschlegel et al., 2000) (McGarry and Kirschner, 1998). Geminin accumulates in S and G2, but its degradation is initiated once the cell reaches anaphase, therefore enabling the assembly of pre-RC for the upcoming round of the cell cycle. As CDT1 is inactive when bound to geminin, it is only able to form part of the pre-RC at the end of mitosis when geminin levels decrease, due to its significant degradation.

CDT1 staining of cells for flow cytometric analysis was not compatible with the PFA fixation used in the previous flow cytometry experiments, staining cells indiscriminately with this fixative procedure. This may be because PFA fixes the cells through covalently crosslinking molecules, which can alter a protein's folding capabilities or preserve artifacts, which can increase background staining. However, after optimisation, specific CDT1 staining was achieved using 70% ethanol fixation (refer to methods 2.6.1,3 and 4). Unlike PFA, ethanol fixes the cells through precipitation and therefore cross-linking is reduced. Furthermore, as ethanol fixation also simultaneously permeabilizes the cells, this additional step is not required, and the overall procedure time was reduced.

As previously observed using PI staining, the CIP29-KO cells exhibited a larger proportion of cells with a 4n DNA content and a sub-population of cells that exhibit a >4n DNA content, that was not observed in the control cell lines. As expected, when analysing CDT1 levels, almost all CDT1-positive MRC5-VI cells coincided with the 2n DNA content (G1) population of cells, with very little CDT1 expression in the subsequent phases of the cell cycle. Contrary to this, although the majority of CDT1-positive CIP29-KO cells were also in G1 phase of the cell cycle, an increased proportion of CDT1-positive cells was also identified with a 4n DNA content, normally considered to represent G2/M phase cells. This increased population of 4n CDT1-positive cells was not observed in W6, whereas F17 displayed a small increase in CDT1-positive cells within the 4n DNA population. This is consistent with the rescue of this phenotype by the re-expression of CIP29 and partial rescue by the re-expression of FLAG-CIP29, like what was seen for other phenotypes described in this chapter, such as nuclear size and centrosome duplication.

Interestingly, the cells that displayed a >4N DNA content within the CIP29-KO do not present as expressing CDT1. As there was also very little histone H3-Ser10^P expression within this population, it is possible that these cells may be within the G0 phase of the cell cycle, whereby the cells could be in a senescent state and no longer express CDT1 or histone H3-Ser10^P.

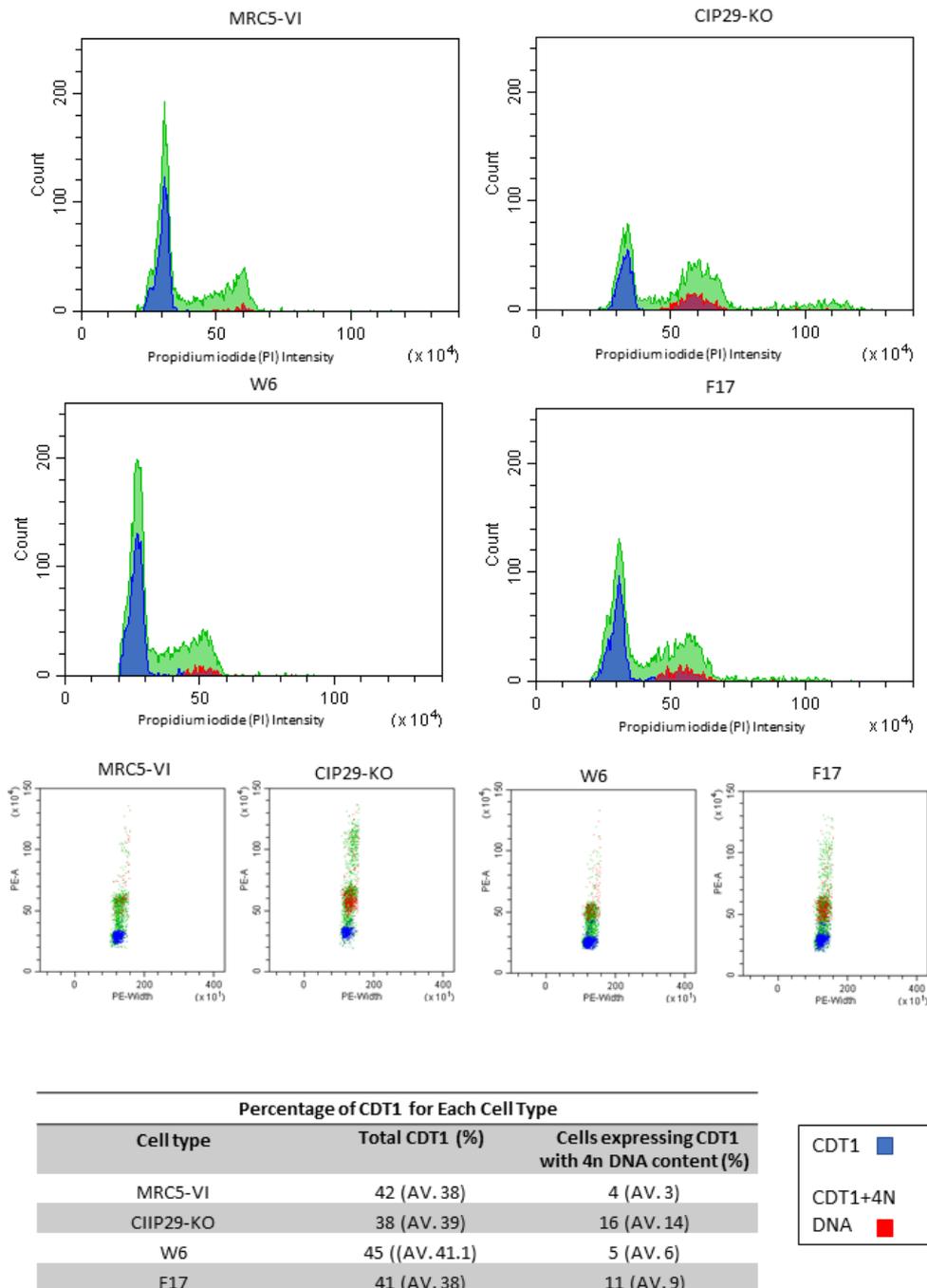


Figure 31: CIP29-KNOCKOUT cells do exhibit an increase in CDT1 Suggesting they are arrested or are slowly progressing through G1.

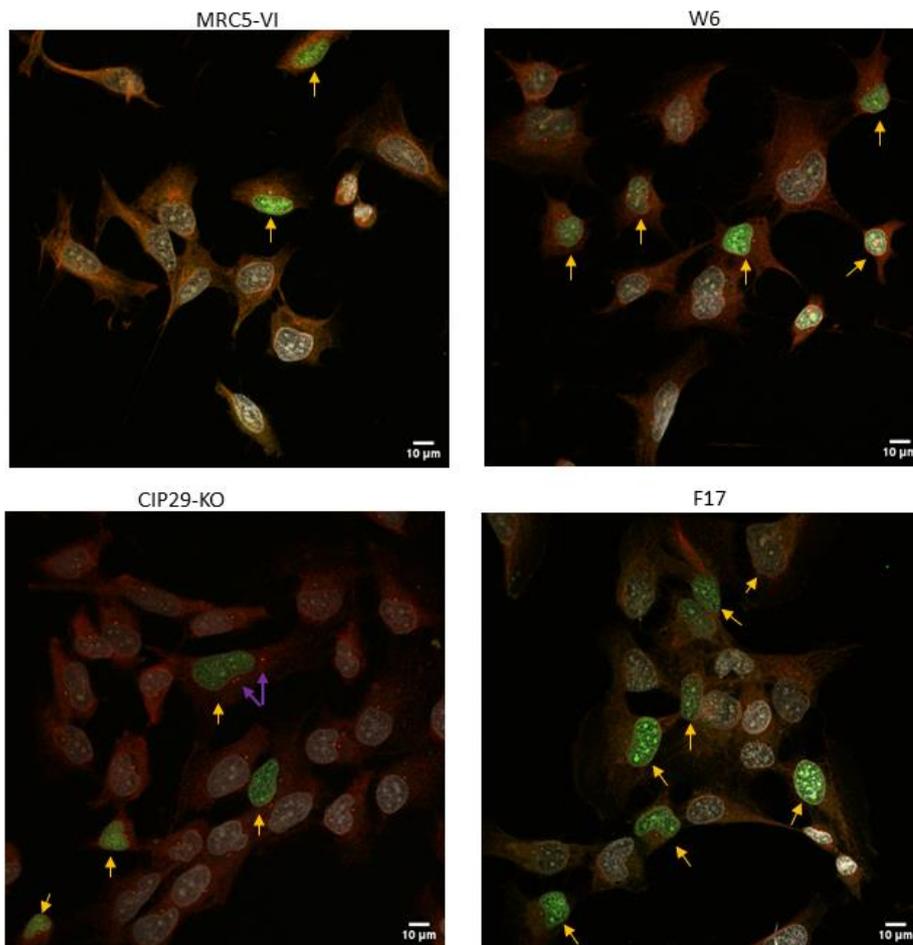
Flow cytometry population histograms and dot plots show the cell cycle profiles of each cell type wild type (MRC5-V1), CIP29- KNOCKOUT (CIP20-KO), CIP29-RESCUE (W6) and FLAG-CIP29-RESCUE (F17). Population histograms are illustrated as DNA content (propidium iodide, PI intensity) versus number of cell counts (counts). Histograms also highlight the CDT1 expression level (blue), and the CDT1 expressed within cells that exhibit a 4n DNA content. While dot plots indicate the cell area vs cell width highlighting the range of cellular size (A). Table detailing the approximate percentage of cells for each cell type that expresses CDT1 within the total cell population and the approximate

percentage of cells that express CDT1 while also maintaining a 4n DNA content (B) (R:2) N: 10,000 cells).

4.2.9 Immunofluorescence analysis to investigate CDT1 expression within the CIP29-KO population.

As the flow cytometry data suggested that CIP29-KO cells may, abnormally, express CDT1 in cells that have already duplicated their DNA, we wanted to investigate this further, at the cell rather than population level. To achieve this, immunofluorescence studies of each cell type were conducted as before, but the cells were stained with CDT1, DAPI and γ -tubulin to correlate CDT1 positive cells and centrosome number (refer to methods 2.7.1-5).

The results obtained suggested that of the MRC5-VI cells that expressed CDT1 89% presented with one centrosome typical of G1 cells, while only 11% expressed CDT1 above a background threshold when 2 centrosomes were present. When analysing the CIP29-KO cells that expressed CDT1, only 54% of CDT1-positive cells possessed a single centrosome, while 28% displayed 2 centrosomes and 18% exhibited more than 2 centrosomes. Matching the flow cytometry data, the CIP29-KO cells that maintained the largest nuclei and potentially a DNA content of $>4N$ did not express CDT1. The W6 rescue cell line was largely similar to the MRC5-VI control, in that, 79% of CDT1-positive cells displayed a single centrosome, with only 17% having 2 centrosomes and just 3% showed multiple centrosomes while expressing CDT1. Similarly, in the F17 cells, 72% of CDT1-positive cells had a single centrosome, 18% had 2 centrosomes, while 11% showed multiple centrosomes. These data suggests that there is a shift in the CDT1 expression within the CIP29-KO cells as only 54% of the cells that expressed CDT1 maintained a single centrosome, while this percentage is much higher in the other cell types. This suggest that only half of the CDT1 expressing CIP29-KO cells are behaving as expected, whereby these cells are within G1 and are preparing to progress to S-phase to replicate their DNA. As CDT1 expression initiates in the latter phases of the previous cell cycle it is not surprising that a small proportion of the MRC5-VI cells with 2 centrosomes also express CDT1. However, in CIP29-KO cells there are almost double the amount of CDT1-positive cells with 2 centrosomes as well as some cells with multiple centrosomes that also express CDT1, while the rescue cell lines, especially W6, display similar patterns to that of the control and only display very limited occurrences of multiple centrosomes and CDY1 expression. This suggests that the CIP29-KO cells that have multiple centrosomes may be situated within G1 either preparing to progress through the cell cycle or may be held within this phase. Further supporting this, the largest CIP29-KO cells that do display centrosome supernumerary and centrosome clustering do not express CDT1. This is also observed in the flow cytometry data and therefore suggests that after a few division cycles CDT1 expression is reduced and the cells may enter G0 to become senescent.



CDT1 expressing cells and their centrosome count expressed as percentages

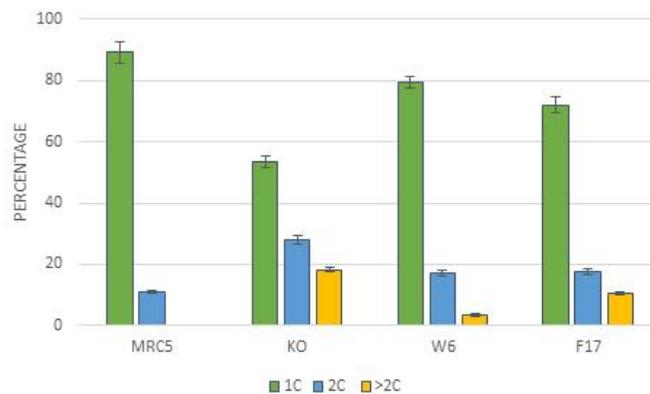


Figure 32: Immunofluorescence analysis suggests that CIP29-KO cells exhibit abnormal CDT1 expression.

Cells are stained with gamma tubulin (RED) and CDT1 (Green). A representation of Cells expressing CDT1 are highlighted (yellow arrows) while abnormal centrosome numbers within the CIP29-KO population are marked (purple arrows) (A). Bar chart showing the number of centrosomes (1C: 1 centrosome, 2C: 2 centrosomes or >2C: more than 2 centrosomes) for each cell that also expresses CDT1 for each cell type. Error bars are defined by standard deviation, (B) (N: ~60 cells, R=3).

4.3 Discussion

From the results obtained the removal of CIP29 has a profound effect on cell functionality that significantly affects the cell cycle, and possibly the final fate of a cell.

From the data presented, it is apparent that the removal of CIP29 results in an overall reduction in the proliferation rate (Figure 23). Unfortunately, as there are only a few replicates it is difficult to determine if the reduction in cell proliferation is significant when compared to the MRC5-VI controls. Furthermore, the level of significance may also be diluted by the assay used to measure the cells' proliferation. The WST-1 cell proliferation assay determines cell viability and the number of cells within the population through the metabolic activity of mitochondrial dehydrogenase. Mitochondrial activity is very dynamic and is greatly influenced by a cells' size (Miettinen and Björklund, 2017). For example, the exchange of nutrients and oxygen within a smaller cell is much more efficient than in a larger one, as the compounds do not have to travel vast intracellular distances to complete cellular processes. Therefore, a smaller cell can maintain a relatively high metabolic rate with only a few mitochondria. However, as a cell increases in size, so will the intracellular distances which can lead to metabolic limitations as the transport of metabolites and oxygen becomes less efficient. To counteract this, mitochondria can undergo mitochondrial fission and develop hyperfused mitochondrial networks which restore efficient nutrient exchange and maintain a higher metabolic rate. As the CIP29-KO cells are significantly larger than the MRC5-VI controls (Figure 28 and 29), the metabolic rate of one CIP29-KO cell may equate to more than one control cell. This could explain why, although the proliferation rate of the CIP29-KO cells is reduced using this assay, it is not significantly different to the MRC5-VI cells, and it is only in the latter stages of the experiment where differences between the two groups become more apparent. This may be because the MRC5-VI cells must achieve multiple divisions before the combined metabolic rate of multiple smaller MRC5-VI cells begins to match and overtake the metabolic rate of the significantly larger CIP29-KO cells. Although mitochondrial activity is scalable to cell size, a cells' metabolic rate will begin to decline once it has reached its ultimate size and the dynamic nature of mitochondria can no longer compensate for the loss in efficiency, due to the increased intracellular distances. Therefore, if the size of the CIP29-KO cells inadvertently influences the cell proliferation readings, these should begin to plateau as the CIP29-KO cells reach their final size. To capture these events the duration of the cell proliferation assay would need to be extended or perhaps an alternative method of measurement (e.g. cell counting) used, now that it is understood that the CIP29-KO cells are significantly larger than the other cell types.

Although the difference in cell proliferation between the CIP29-KO cells and the MRC5-VI controls was moderate, the cell cycle profiles obtained through flow cytometry show significant differences in cell cycle progression, with the CIP29-KO cells apparently accumulating in the later stages of the cell cycle, after genome duplication, while some cells contained a much higher DNA content than expected (Figure 24).

Despite these cell cycle defects, the cells do not seem to be arrested in G2/M as there is no significant increase in histone H3-Ser10^P expression that would indicate a mitotic block and no evidence of DNA damage checkpoint activation was detected (p53-Ser15^P or γ -H2AX) (Figure 25 and 26). This is consistent with the suggestion that the activation of p53 is much weaker following spindle damage when compared to its activation in response to DNA damaging agents (De Santis Puzzon et al., 2016). However, p53 activation cannot be ruled out, as many other studies that have investigated similar cell cycle defects and the development of aneuploidy cells through a cytokinesis failure also reported no phosphorylation of p53 on Ser 15. Instead, p53 was activated on Ser 33 (p53-Ser 33^P) to implement the DDR (Mikule et al., 2007). This residue is known to be phosphorylated by p38, a member of the DDR pathway that responds to cellular stress within the cell cycle and promotes cell senescence by inducing the G1/S checkpoint (Yee et al., 2004) (Kishi et al., 2001). As a significant loss of cells is not seen in the CIP29-KO population some of the cell may present as a secondary senescent population. A secondary population within the CIP29-KO cells was also highlighted when assessing the population density of each cell type (Figure 28). This second population was also found to have significantly larger nuclei which upon further investigation through immunofluorescence, also indicated that some of this population exhibited multi-lobed nuclei and supernumerary centrosomes (Figure 27) (Figure 30). Our data also indicated that a proportion of cells with a 4n DNA content expressed high levels of the G1 phase cell cycle marker, CDT1. (Figure 31 and 32). This suggests that the increase in cells with a 4n DNA content observed in CIP29-KO cells is indicative of a cell division defect that may, in part, delay mitotic completion (given the very slight increase in histone H3-Ser10^P) but, in many instances, cells appear to exit mitosis, perhaps through mitotic slippage, and re-enter G1 phase but without undergoing nuclear and cell separation. In some cases, these tetraploid G1 cells may attempt a further round of DNA replication, acquiring a greater than 4n DNA content, which may explain the abnormal phenotypes observed.

It is possible that the centrosome anomalies arise simply as a by-product of cytokinesis failure, whereby multiple centrosomes follow the lack of separation within the next round of the cell cycle. However, our Cramer V analysis suggests that the centrosome anomalies and larger nuclei size are intrinsically linked, and the two phenotypes may have a compounding effect upon one another (Table 10). As aberrant expression of proteins involved in centrosome biogenesis can drive

centrosome dysfunction or abnormal centrosome number, this could lead to improper mitotic spindle formation and chromosome missegregation which could cause cytokinesis failure and force the cells to increase its nuclei area to accommodate this lack of separation, although further research is required and at present this remains an assumption. Furthermore, it is important to note, that the observations of centrosome clustering and amplification have been made using immunofluorescent staining on fixed cellular specimens with antibodies that recognize pericentriolar material (PCM) components. Unfortunately, immunofluorescent staining has several limitations when used to detect centrosomes. For example, antibody quality, fixation and penetration can all affect the interpretation of the staining, but its most important shortcoming is that not all objects stained by γ -tubulin will represent true centrosomes. As a result of this, it is difficult to conclude with certainty that all observations of centrosome clustering within the CIP29-KO cells are genuine occurrences of this phenomenon. However, the frequency and its isolation to the CIP29-KO cells, suggests that it is an effect due to the loss of CIP29 and suggest a cascade of errors occur during the cell cycle as a result. Although the exact mechanisms that cause such defects requires further exploration.

5 [Chapter 5: exploring the Structure-function analysis of CIP29.](#)

5.1 **Introduction**

Historically, cancer treatment has relied on radiotherapy and chemotherapy which target cells that are more predisposed to cell death. However, both techniques fail to distinguish between cancer cells and healthy cells, resulting in significant toxicity during treatment. As a result, novel drug development has shifted to targeted cancer therapies such as small molecules inhibitors (SMIs). These have already shown much success and are already used to target dysfunctional DNA repair, apoptosis and RNA processing factors (Attwood et al., 2021, Cohen et al., 2021, Ferguson and Gray, 2018, Hogg et al., 2020, Greene et al., 2020, Bedard et al., 2020, Merino et al., 2018). Much of this success depends on the definition of the drug target, whereby its protein structure, its regions of interest and its specific conformation are fully understood (Wu, 2020, Derakhshan et al., 2017, Hengel et al., 2017). This knowledge enables the functional biology of the protein and its downstream activities to be inferred which in turn provides a greater justification for therapeutic intervention at a certain domain, while also improving the overall clinical safety of a suggested drug.

Although CIP29 is a relatively small protein of only 29 KDa very little is known about its protein structure, with only its evolutionarily conserved N terminal SAP domain previously investigated. This domain contains a conserved 35 residue motif that is composed of a collection of highly hydrophobic, bulky and polar amino acids, which are separated via a region of glycine residues (Leaw et al., 2004) (Choong et al., 2001). Many proteins that contain SAP domains have been found to localise to the nucleus and often contribute to nuclear processes such as the regulation of transcription, DNA repair, RNA processing or apoptotic chromatin degradation. In contrast to this, the C-terminal domains of the yeast homolog Tho1 have been suggested to be responsible for the RNA binding and the mRNA processing ability of the protein in yeast (Jimeno et al., 2006a). As the double helix motifs within the C-terminal end of the CIP29 protein is highly conserved there is a possibility that the SAP domain is not the only region of importance.

5.1.1 Project objective: Obtain a greater understanding of the molecular structure of the CIP29 protein and its functional domains.

Aside from the limited evaluations the CIP29 protein has yet to be fully structurally or functionally characterised, so it is unclear which regions hold the most importance. To begin to explore this, the CIP29 protein will be dissected so various regions of interest can be further examined and subjected

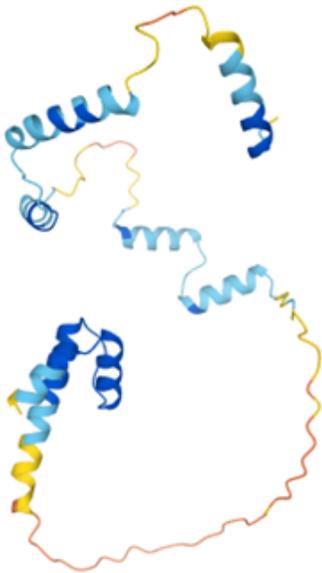
to the phenotypic studies that have been previously presented. The truncations will also be compared to the CIP29-KO and rescue cell lines.

5.2 Results

5.2.1 Generation of CIP29 truncation mutants

To further investigate each region of the CIP29 protein, its structural features were illustrated through the AI system AlphaFold (refer to methods 2.8.4). This AI system utilises a protein's amino acid sequence and information gathered by the European Bioinformatics Institute (EMBL) to predict a protein's 3D structure which consisted of an SAP domain, two double α -helical regions and a single helix. Each region is also separated via unstructured regions (Figure 33) (Jumper et al., 2021, Varadi et al., 2022).

A



B

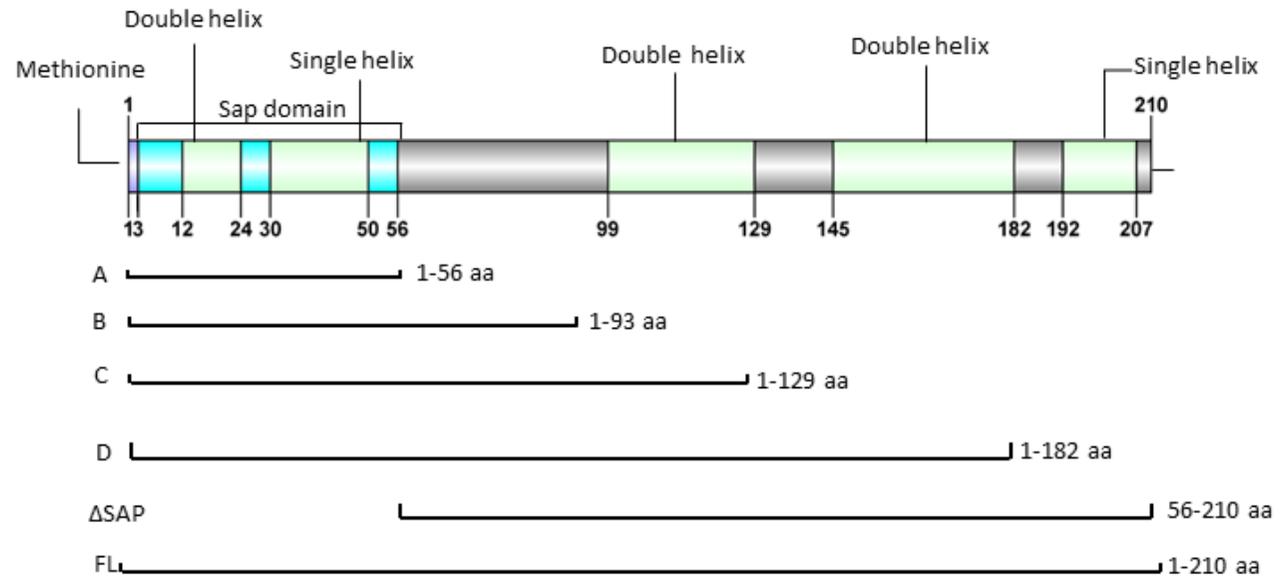


Figure 33: Schematic representation of the CIP29 truncations.

Molecular structure of the full length CIP29 protein generated through Alpha Fold AI assisted prediction software (A). Regions of interest are highlighted, from left to right: the Methionine start site (purple), the SAP domain (light blue: 1-56 aa), the helix domains (green: 12-24 aa, 30-50 aa, 99-129 aa, 145-182 aa and 192-207 aa) and unstructured regions (grey: 56-99 aa, 129-145 aa, 182-192 aa and 207-210 aa). Each spanning truncation is detailed below along with the full-length protein for reference (B).

As the functional significance of the SAP domain region and other structural motifs (α -helices) was somewhat unclear, CIP29 truncations of various lengths were designed to encompass the different regions of interest (Figure 34). The first and largest truncation (mutant A) gave rise to truncated protein comprising only the N-terminal SAP domain with a length of 56 amino acids (aa), the second truncation mutant, mutant B included the SAP domain and an adjacent unstructured region to reach a final length of 93 aa. The third truncation mutant C included the SAP domain, the unstructured region, and a double α -helical region immediately downstream to a length of 129 aa. The fourth truncation, mutant D included a further double α -helical region (1-182 aa). Finally, a fifth truncation mutant, Δ SAP (56-210 aa) was developed to include all regions of the CIP29 protein, except for the N-terminal SAP domain.

SAP domain + unstructured region + helical pair + helical pair + single helix

FL
(1-210aa)

10	20	30	40	50
MATETVELHK LKLAELKQEC LARGLETGKI KQDLIHRLQA YLEEHAEAAA				
60	70	80	90	100
NEEDVLGDET EEEETKPIEL PVKEEEPPEK TVDVAAEKKV VKITSEIPQT				
110	120	130	140	150
ERMQKRAERF NVPVSLESKK AARAARFGIS SVPTKGLSSD NKPMVNLDKL				
160	170	180	190	200
KERAQRFGLN VSSISRKSED DEKLKRRKER FGIVTSSAGT GTTEDTEAKK				
210				
RKRAERFGIA				

SAP domain only

A
(1-56aa)

10	20	30	40	50
MATETVELHK LKLAELKQEC LARGLETGKI KQDLIHRLQA YLEEHAEAAA				
60	70	80	90	100
NEEDVIGDET EEEETKPIEL PVKEEEPPEK TVDVAAEKKV VKITSEIPQT				
110	120	130	140	150
ERMQKRAERF NVPVSLESKK AARAARFGIS SVPTKGLSSD NKPMVNLDKL				
160	170	180	190	200
KERAQRFGLN VSSISRKSED DEKLKRRKER FGIVTSSAGT GTTEDTEAKK				
210				
RKRAERFGIA				

SAP domain + unstructured region

B
(1-93aa)

10	20	30	40	50
MATETVELHK LKLAELKQEC LARGLETGKI KQDLIHRLQA YLEEHAEAAA				
60	70	80	90	100
NEEDVLGDET EEEETKPIEL PVKEEEPPEK TVDVAAEKKV VKITSEIPQT				
110	120	130	140	150
ERMQKRAERF NVPVSLESKK AARAARFGIS SVPTKGLSSD NKPMVNLDKL				
160	170	180	190	200
KERAQRFGLN VSSISRKSED DEKLKRRKER FGIVTSSAGT GTTEDTEAKK				
210				
RKRAERFGIA				

Sap domain + unstructured region + helical pair

C
(1-129aa)

10	20	30	40	50
MATETVELHK LKLAELKQEC LARGLETGKI KQDLIHRLQA YLEEHAEAAA				
60	70	80	90	100
NEEDVLGDET EEEETKPIEL PVKEEEPPEK TVDVAAEKKV VKITSEIPQT				
110	120	130	140	150
ERMQKRAERF NVPVSLESKK AARAARFGIS SVPTKGLSSD NKPMVNLDKL				
160	170	180	190	200
KERAQRFGLN VSSISRKSED DEKLKRRKER FGIVTSSAGT GTTEDTEAKK				
210				
RKRAERFGIA				

Sap domain + unstructured region + helical pair + helical pair

D
(1-182aa)

10	20	30	40	50
MATETVELHK LKLAELKQEC LARGLETGKI KQDLIHRLQA YLEEHAEAAA				
60	70	80	90	100
NEEDVLGDET EEEETKPIEL PVKEEEPPEK TVDVAAEKKV VKITSEIPQT				
110	120	130	140	150
ERMQKRAERF NVPVSLESKK AARAARFGIS SVPTKGLSSD NKPMVNLDKL				
160	170	180	190	200
KERAQRFGLN VSSISRKSED DEKLKRRKER FGIVTSSAGT GTTEDTEAKK				
210				
RKRAERFGIA				

Unstructured region + helical pair + helical pair + single helix

**ΔSAP
(56-210aa)**

10	20	30	40	50
MATETVEIHK	LKLAELKQEC	LARGLETKGI	KQDLIHRLQA	YLEEHAE EEEE A
60	70	80	90	100
NEEDVL	<u>GDET</u> <u>EEEE</u> TKPIEL	<u>PVKEE</u> PPER	<u>TVDVAA</u> EKKV	<u>VKITSEI</u> <u>PQT</u>
110	120	130	140	150
<u>ERMQKRA</u> ERF	<u>NVPVS</u> LESKK	<u>AARAAR</u> FGIS	<u>SVPTKGL</u> SSD	<u>NKPMVN</u> LDKL
160	170	180	190	200
<u>KERAQR</u> FGLN	<u>VSSISR</u> KSED	<u>DEKLKR</u> KER	<u>FGIVT</u> SSAGT	<u>GTTE</u> DTEAKK
210				
<u>RKRAE</u> RFGIA				

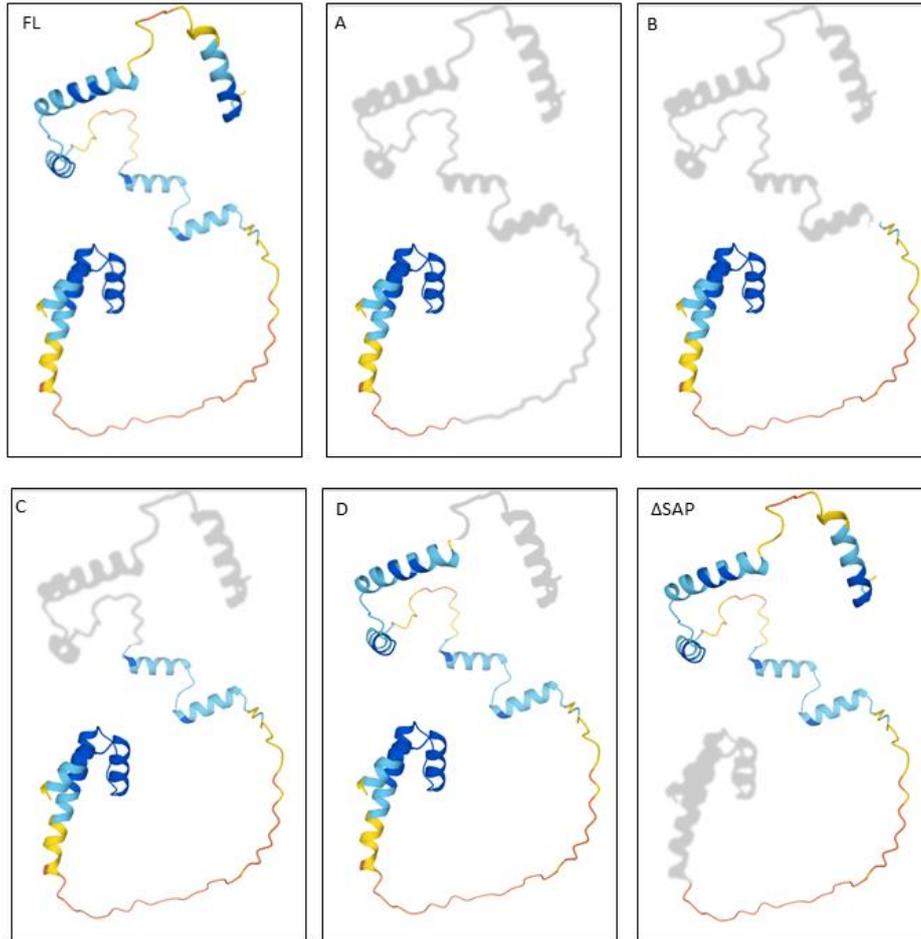


Figure 34: Development of CIP29 truncations using alpha fold protein structure data base (3 of 3). The amino acid sequence of the CIP29 protein is provided with each truncation highlighted (yellow).

Known regions of interest are also marked, the SAP domain (Bold) and the helix structures (Underlined). A visual representation of each truncation is also provided with the absent structure marked in grey. Truncations are denoted as A (1-56 aa), B (1-93 aa), C (1-129 aa), D (1-182 aa) ΔSAP (56-210 aa) alongside the full-length CIP29 protein (FL) for reference (Jumper et al., 2021, Varadi et al., 2022). Figure 3 of 3.

To generate these various CIP29 mutants, each relevant coding sequence was PCR amplified using primers to incorporate a *XhoI* restriction enzyme site at the 5' end (refer to methods 2.2.6). PCR products were initially cloned into pGEMT-Easy and sequence verified (refer to methods 2.2.8), before cloning (*XhoI*-*NotI*) into a mammalian expression vector, PCI-NEO-FLAG, in frame with an N-terminal FLAG epitope (refer to methods 2.2.5).

5.2.2 *In vitro* expression analysis of CIP29 truncation mutants

To confirm expression of each FLAG-tagged CIP29 truncation mutant, each expression plasmid was individually transcribed and translated using an *In Vitro* transcription/translation system (IVT) which featured a Transcend™ Non-Radioactive Translation Detection System. The Transcend™ detection system can validate protein synthesis by the detection of biotinylated lysine residues that are incorporated into the nascent protein during its translation. As the biotinylated lysines within each protein bind to Streptavidin-Alkaline Phosphatase, western blot analysis followed by Streptavidin-Alkaline Phosphatase detection was then used to confirming protein synthesis (refer to methods 2.3.11). The results obtained suggested that the full-length protein could be readily detected due to its 26 lysine residues at the expected molecular weight of just over 29KDa. This slightly higher molecular weight is due to the incorporation of the FLAG epitope. A second slightly lower band can also be observed, which represents the synthesis of the CIP29 protein but through the internal methionine at the start of the CIP29 protein rather than at the upstream methionine situated at the start of the FLAG epitope. Similarly, truncation mutants C (1-129), D (1-182) and Δ SAP exhibit a strong signal through colorimetric detection with each having an abundance of lysine residues to facilitate detection, (mutant C has 14 Lys residues, mutant D has 19 Lys residues, Δ SAP has 21 Lys residues). This detection confirms the predicted molecular weights of ~19KDa for C, 25KDa for D and 22KDa for Δ SAP (Stothard, 2000).

Again, a second band can be observed for C and D, but not the Δ SAP. This is because this truncation lacks the internal methionine of the endogenous CIP29 as this truncation is produced from the 56th aa to the final 210th aa to exclude the SAP region of the endogenous CIP29 protein. Therefore, only a single protein variant can be produced through the methionine situated within the FLAG sequence. In contrast to the other truncations, mutant A only contains 5 Lys residues and consequently exhibited no detection even after expression, while truncation mutant B (1-93) only displays a faint level of detection at approximately ~15KDa molecular weight, as this truncation only contains 8 lysine residues. When assessing the detection of the truncations through the α -CIP29 antibody, detection of the full length CIP29 protein could be readily confirmed along with truncation mutants: C, D and the Δ SAP. However, mutants A and B were not detected. This suggests that the α -CIP29

antibody identifies the CIP29 protein using regions that are predominately situated after the 93th aa, which suggests that the double helixes or the single helix within the C-terminal region maybe the source of binding, rather than the SAP domain (Figure 35A/B).

5.2.3 *In vivo expression analysis of CIP29 truncation mutants*

To initially confirm the expression of each truncation mutant *in vivo*, each construct was transiently transfected into mammalian MRC5-VI cells for 24 hours before expression of each truncation mutant was assessed through western blot of whole cell lysates (refer to methods 2.3.11). Western blot analysis confirmed that both the full length CIP29 protein and the truncation mutants B, C, D and Δ SAP could be successfully transiently transfected and expressed to produce a truncated CIP29 protein within MRC5-VI cells. However, mutant A the smallest truncation which featured only the SAP domain was not observed, despite attempting detection with both the α -CIP29 and α -FLAG antibodies (Figure 35B). It is possible that this truncation mutant was particularly unstable and may be readily targeted for degradation, therefore further investigation using this construct was not pursued. Although the transient transfections resulted in protein expression for each of the remaining truncation mutants, transfection efficiency can vary between each experiment, which can introduce unnecessary uncertainty within the results. Therefore, stably transfected cell lines expressing the various FLAG-tagged CIP29 truncation mutants were developed for further investigations. To develop stable lines of each of the CIP29 truncation mutants, CIP29-KO cells were transfected with expression constructs of each FLAG-tagged truncation mutants B (1-93 aa), C (1-129 aa), D (1-182 aa) and Δ SAP. 48 hours after transfection, selection (600 μ g/ml G418) was applied, and G418-resistant colonies were isolated after approximately 10-14 days (refer to methods 2.5.5-6). Expression levels of the truncated CIP29 proteins in individual clones were examined through western blot analysis using whole cell extracts and α -FLAG antibodies (Figure 35C). Through this screening approach, a cloned cell line displaying the expression of FLAG-CIP29 at the expected protein size and at comparable levels to endogenous CIP29 protein, was identified for each of the four CIP29 truncation mutants and full-length FLAG-CIP29, therefore analysis of phenotypic rescue in relation to each truncated protein could progress.

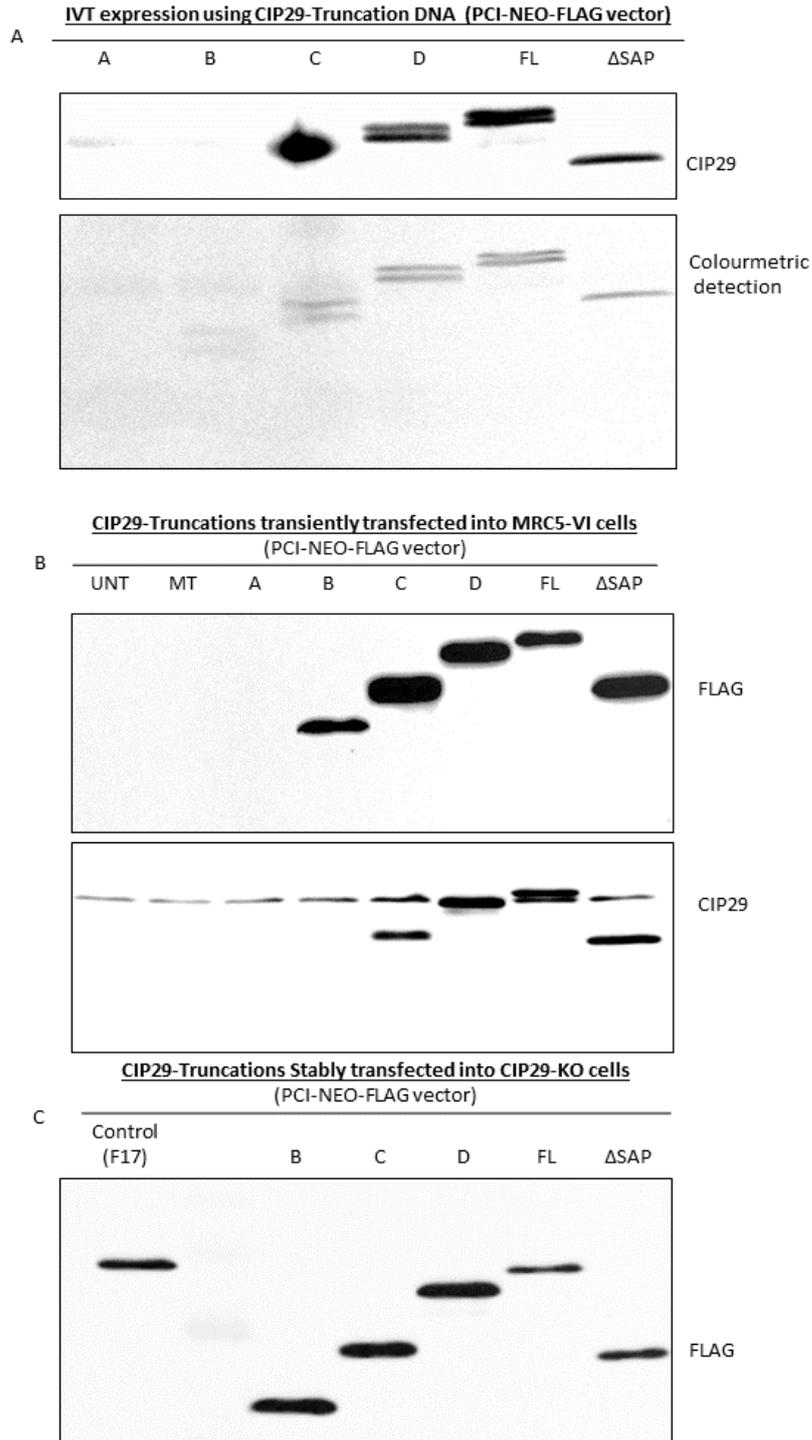


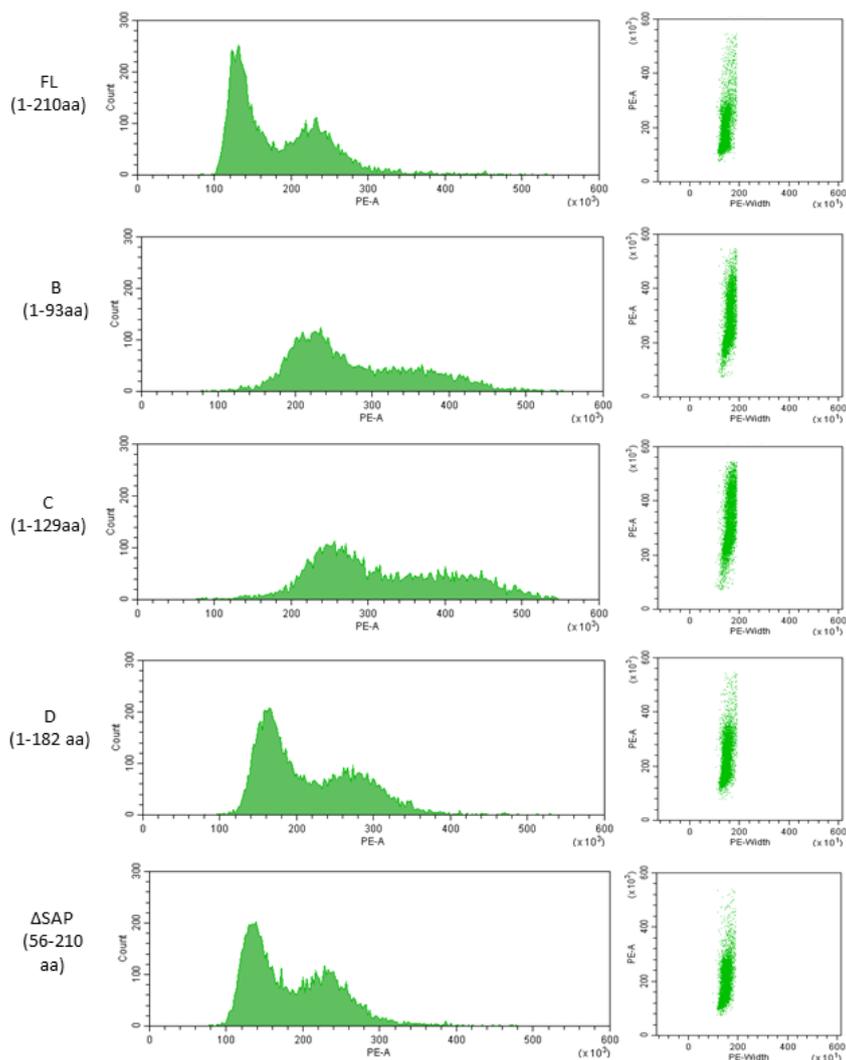
Figure 35 Investigating the protein expression of each CIP29 truncation.

CIP29-truncations (A, B, C, D, and Δ SAP) and the full-length protein (FL) was expressed within an in-vitro transcription translation system. Each expression construct was detected through enhanced colorimetric detection and a α -CIP29 specific antibody (A). FLAG-CIP29 protein expression 24h after transient transfection into MRC5-VI cells was detected by α -CIP29 and α -FLAG antibodies (B). Stably

transfected CIP29 truncated proteins within CIP29-KO cells with CIP29 expression assessed through α -FLAG antibodies (C) (A: R=2, B and C R=3).

5.2.4 Investigating the cell cycle profiles of the CIP29-truncations using flow cytometry analysis

As the CIP29-KO cells exhibited such extensive cell cycle defects it was pertinent to investigate if any of the CIP29-truncations shared similar effects, as this would highlight the molecular structures that influence such abnormal development. To achieve this, the cell cycle profiles of each cell line were investigated through flow cytometry with PI staining as previously described (refer to methods 2.6.2-4) (Figure 36). The results obtained confirmed results from the previous chapter, that CIP29-KO cells stably expressing full length CIP29 displayed a cell cycle profile with no obvious cell cycle defects. This normal cell cycle profile was closely matched in CIP29-KO cells expressing CIP29 truncations D and Δ SAP. However, the cell cycle profiles of cells expressing CIP29 truncations B and C displayed much wider and shallower peaks which had shifted significantly along the X-Axis. This suggested that many of the cells within these populations had already surpassed the diploid DNA content of $2n$ ($100-200 \times 10^3$) with many displaying a DNA content of greater than $4N$. Some even reached a DNA content that suggested pentaploidy which indicates that these cells may experience multiple rounds of the cell cycle and DNA replication without successfully completing cell division. As each peak is very poorly defined, it suggests that the DNA content within these cell populations is highly variable, indicating that the normal cycle from G1 diploid to G2/M tetraploid, and back to diploid is not routinely achieved. Furthermore, as very few cells displayed a diploid DNA content it suggests that these cells are not dividing and instead are accumulating more and more DNA therefore cell cycle regulation within these cells may not be functioning as required. Ultimately, these data suggests that a functioning C-terminal end of CIP29 is required for correct cell cycle progression, however more research is required to confirm this assumption.



B: Percentage of cells in each phase of the cell cycle for each cell type

Approximate percentage of cells within each phase of the cell cycle (%)			
Cell Type	G1	S	G2/M
FL	50	6	43
B	4	6	88
C	3	2	95
D	51	10	38
ΔSAP	46	8	44

Figure 36 Some stably transfected CIP29- truncations exhibit abnormal cell cycle profiles indicative of cell cycle defects.

Flow cytometry population histograms and dot plots offer an indication of the cell cycle profiles for each CIP29- truncation (B, C, D and ΔSAP) and the full length CIP29 protein (FL) (A). Table shows an estimation of the percentage of cells within each phase of the cell cycle (G1, S, G2/M) for each CIP29- truncations (B) (R=2).

5.2.5 Do the CIP29- truncations exhibit morphological defects like the CIP29-KOs?

As the cell cycle profiles suggested that some of the CIP29 truncations showed similar defects to the CIP29-KO, these attributes were analysed in more detail. As previously described the nuclear area and centrosome number of each cell mutant were analysed (refer to methods 2.7.1-5, 2.8.1 & 2.8.3). The immunofluorescence data obtained suggested that the smaller CIP29 truncation mutants, B and C, showed similar morphological defects to the CIP29-KO, while the larger truncation mutants, including the Δ SAP, displayed a more typical morphology.

5.2.5.1 Morphological defects: centrosome supernumerary

To investigate centrosome number, each of the CIP29 truncation mutants were subjected to immunofluorescence analysis and were stained with γ -tubulin and DAPI (refer to methods 2.7.1 - 2.7.5) before centrosomes were identified, counted and computationally analysed (refer to methods 2.8.2). Consistent with previous findings (Chapter 4), CIP29-KO cells in which either full length CIP29 (W6) or full-length FLAG-tagged CIP29 (FL) was re-expressed, the majority of cells contained one (~60%) or two (~30%) centrosomes, with only 4-5% of cells containing more than two centrosomes. In contrast, the CIP29-KO cells with no re-expression of CIP29, the proportion of cells with supernumerary centrosomes increased to 15%, with fewer cells containing a single centrosome. Suggesting that centrosome supernumerary may hold phenotypic significance. As with the full length CIP29, when CIP29 mutant D or the Δ SAP were expressed in CIP29-KO cells, these mutants displayed restored centrosome numbers to levels approximating the W6 cells, with few (4-5%) cells harbouring supernumerary centrosomes. This suggest that neither of these truncational mutants exhibit centrosome supernumerary nor are they hindered by centrosome accumulation. In contrast, CIP29-KO cells that housed CIP29 truncation mutants B and C exhibited centrosome characteristics that seemed to exceed even what had previously been observed within the CIP29-KO displayed multiple centrosome counts of 32% (B) and 40% (C) of the overall cell populations, vastly exceeding the CIP29-KO observation of 15%. While the duplicate centrosome numbers were again higher than the other cell lines and truncations with 42% (B) and 36%(C) for each population. Furthermore, as so many of the cells that maintained B and C truncations displayed multiple centrosomes, the number of cells with only a single centrosome was drastically reduced to 25% (B) and 24% (C). Whereas the CIP29-KO cells maintained a single centrosome population of double this, at 50% and the full-length cell line was greater still, at 62% (Figure 37). From the results obtained, it suggests that the smallest CIP29 truncations (B and C) exhibit centrosome amplification or centrosome breakdown. However, from the staining used it is difficult to determine which. The results also suggest that these

truncations are unable to reverse the centrosome defects once the centrosomes have been duplicated. As this phenotype does not seem to be shared by the remaining truncations (D and Δ SAP) it suggests that the C-terminal end of the CIP29 protein is required for the correct functionality of centrosome control.

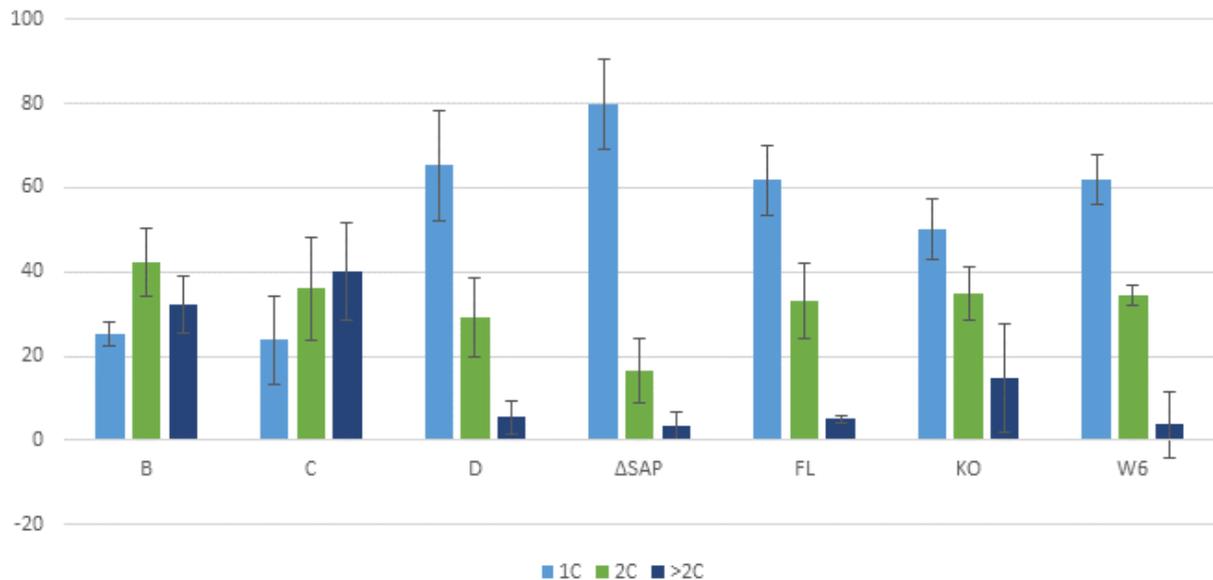
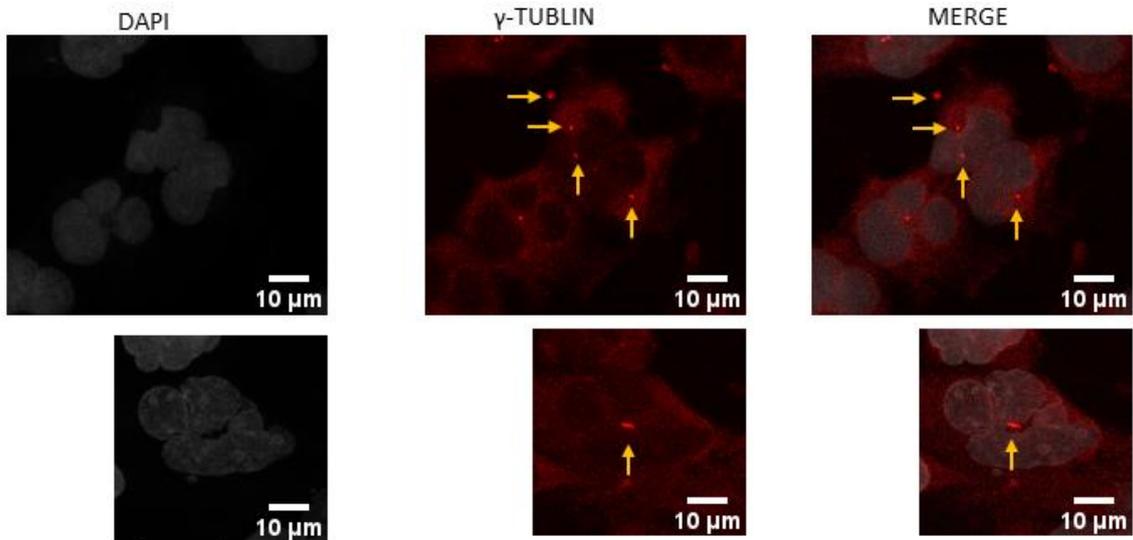
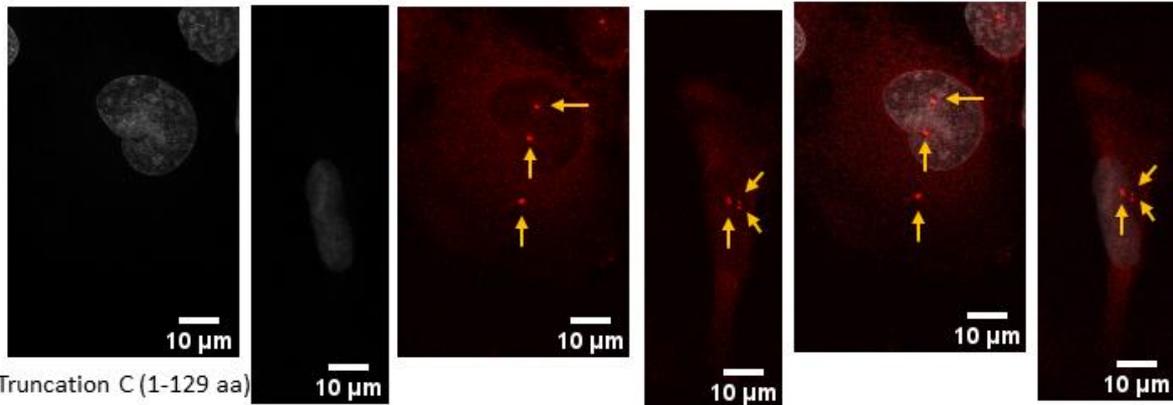


Figure 37: Investigating if the CIP29 truncations exhibit centrosome amplification.

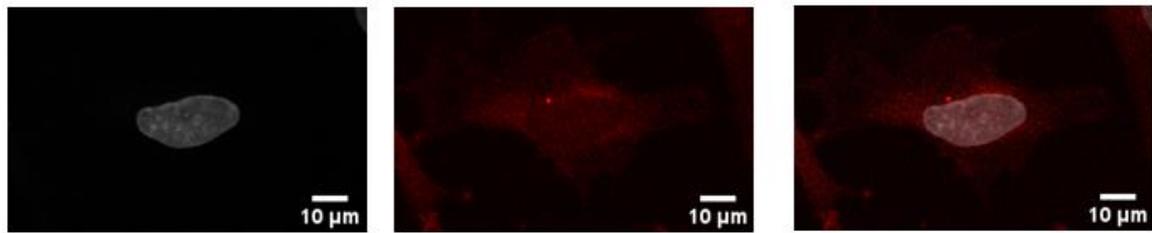
Bar charts show the percentage of centrosomes per cell for each cell type W6 (CIP29 rescue), CIP29-FL (CIP29-FLAG-rescue), KO (CIP29-KNOCKOUT) and each CIP29-truncation (B, C, D and Δ SAP). (1C: 1 centrosome, 2C: 2 centrosomes or >2C: more than 2 centrosomes) (A) (R= 3, N:~180 cells). Standard deviation shown as error bars.



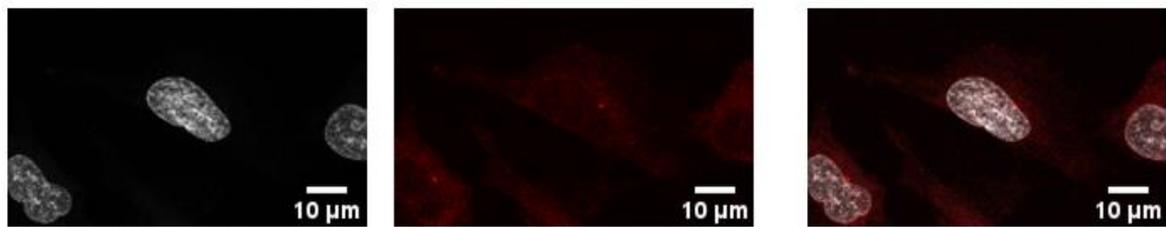
Truncation B (1-93 aa)



Truncation C (1-129 aa)



Truncation D (1-182 aa)



Truncation Δ SAP (56-210 aa)

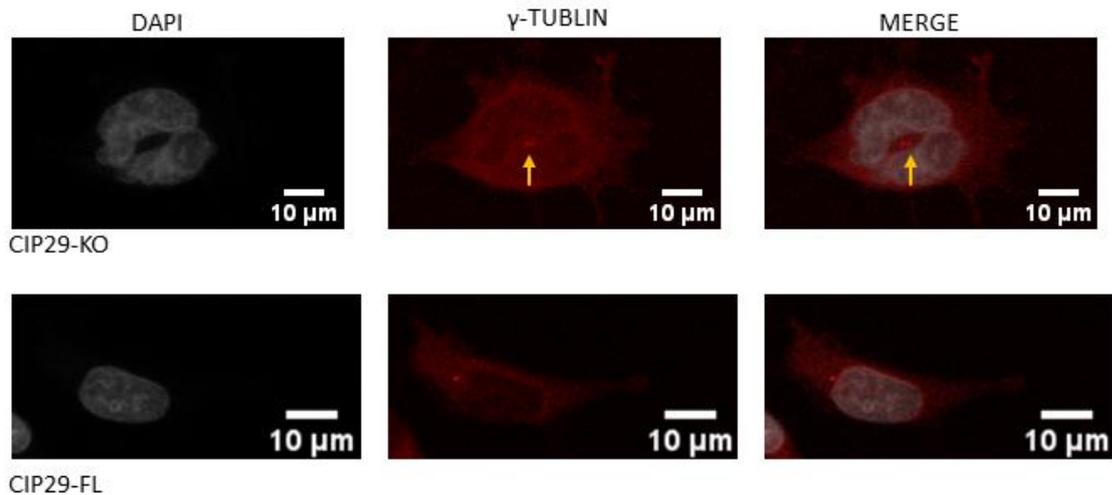


Figure 38: Examples of centrosome amplification through immunofluorescence staining 2/2. Immunofluorescence staining using DAPI (grey) γ -tubulin (red) and merge to identify the occurrence of centrosome amplification (yellow arrows) within each cell type (W6 (CIP29 rescue), CIP29-FL (CIP29- FLAG rescue), KO (CIP29-KNOCKOUT), and Truncations (B, C, D and Δ SAP). Images are scaled to 10 μ m. 2/2 (R=3)

5.2.5.2 Do the CIP29-Truncations exhibit nuclei malformations?

To analyse the nuclei sizes of each truncation mutant in more detail, each mutant cell line was subjected to immunostaining (γ -tubulin and DAPI) and computation analysis (refer to methods 2.7.1-5, 2.8.1). The results obtained suggested that the CIP29-KO cells in which either full length CIP29 (W6) or full-length FLAG-tagged CIP29 (FL) was re-expressed, displayed an average nuclei area that did not exceed 200 μ m. As previously observed (chapter 4) the CIP29-KO cells exceeded this. When comparing the nuclei area for the CIP29 truncated mutants the Δ SAP exhibited an average nuclei area that was most comparable to the W6 and FL, while the CIP29-truncation D cell line only slightly exceeded this. Suggesting that these two truncated mutants did not exhibit excessively large nuclei and therefore the cells may not have experienced cytokinesis failure. This is unlike the smaller truncation mutants (B and C) which exhibited nuclei that vastly exceeded even the CIP29-KO cell line. This suggests that these cell lines may have experienced a mitotic defect. When analysing the standard deviations of each of the cell lines the CIP29-KO cells and the smaller CIP29 truncations B and C maintain a high level of variation within these populations compared to their counterparts. Therefore, further statistical analysis was limited to non-parametric tests as these types have a higher threshold for statistical significance (Table 11 and Figure 39).

Table 11: CIP29- Truncations B and C exhibit cell populations that maintain a wide variation of nucleus sizes (μm)

Cell Type	Median	Mean	Min	Max	Standard deviation
W6	166	174	63.9	322	47.6
FL	182	195	64.8	465	69.0
KO	231	253	76.5	1155	108
B	297	316	123	926	105
C	286	309	118	877	111
D	222	229	78.2	394	60.2
ΔSAP	186	193	61.7	577	63.5

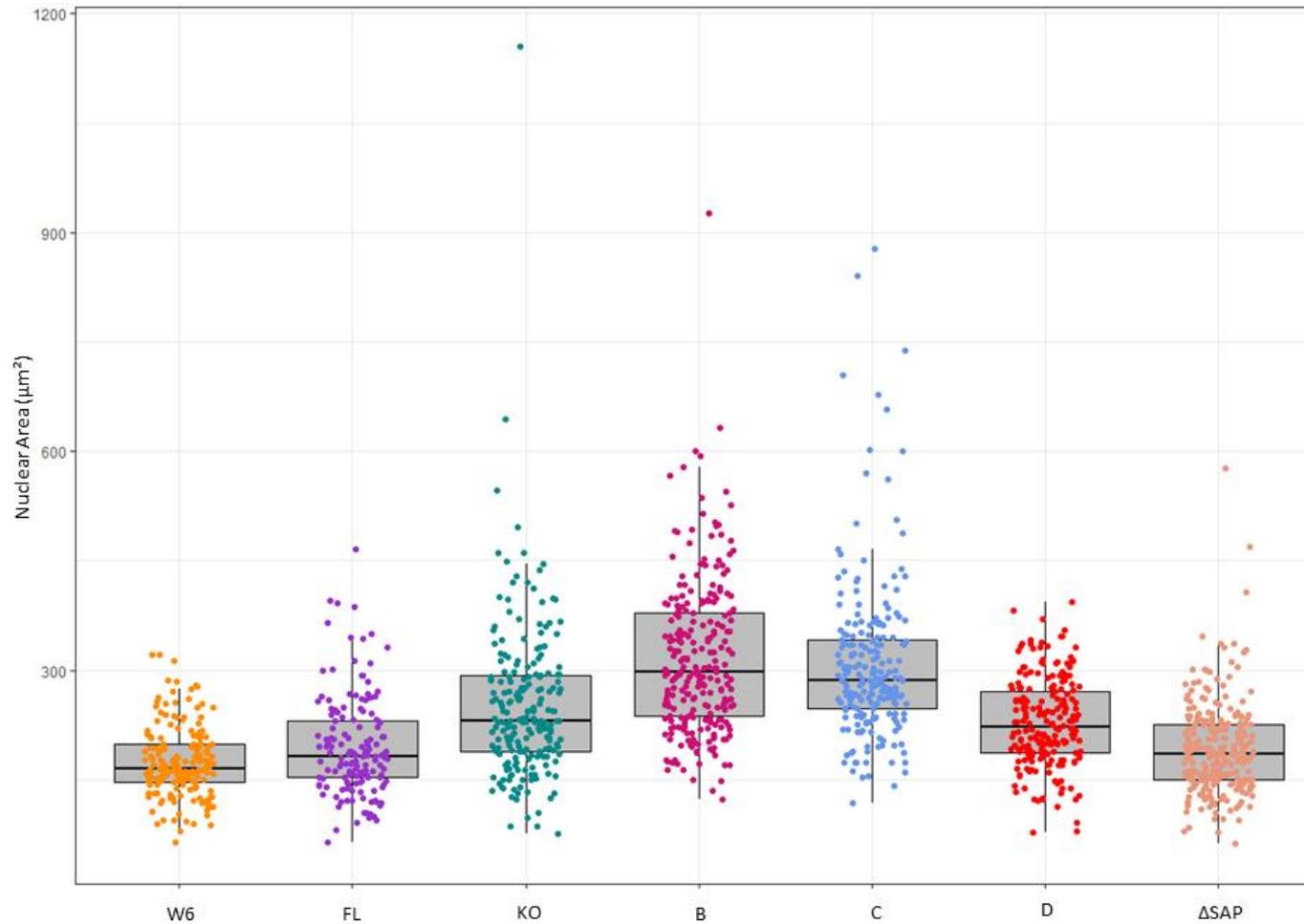


Figure 39: Figure 39: Box plots show the nuclei area (μm^2) for each cell type with individual readings overlaid.

Box plots display the nuclei area (μm^2) for each cell type W6 (CIP29 RESCUE), FL (FLAG-CIP29-RESCUE), Knockout (CIP29-KNOCKOUT) and each truncation (B, C, D, and ΔSAP). Each cell type is colour coded. (A) (R=3) (collated samples, N= ~ 250 cells)

As the standard deviation of the nuclear sizes suggested that some of the cell populations exhibited a broad range of nuclear sizes within each of the populations this was analysed in more detail through density plots (Figure 40). Although the justification for a bimodal distribution within these dataset is not as strong as previously observed within the CIP29-KO cells, significantly larger cells can still be identified, suggesting that this attribute may still be significant when characterising this cell population. Like previous observations of the CIP29-KO, the cell truncations B and C exhibited secondary density peaks, which suggest these populations may maintain subpopulations that exhibit larger than average nuclei sizes. Upon analysis, Truncation B displayed a secondary density peak relating to nuclei size of approximately $900 \mu\text{m}^2$ indicating a bimodal distribution, while truncation C maintained numerous additional peaks between $\sim 600 \mu\text{m}^2$ and $\sim 900 \mu\text{m}^2$ suggesting a multimodal distribution. These characteristics were not observed within truncation D which mirrored the distribution profile of the W6 rescue cell line. Interestingly, the ΔSAP truncation did display some additional sub-populations, although the nuclei sizes of these populations remained below $\sim 600 \mu\text{m}^2$.

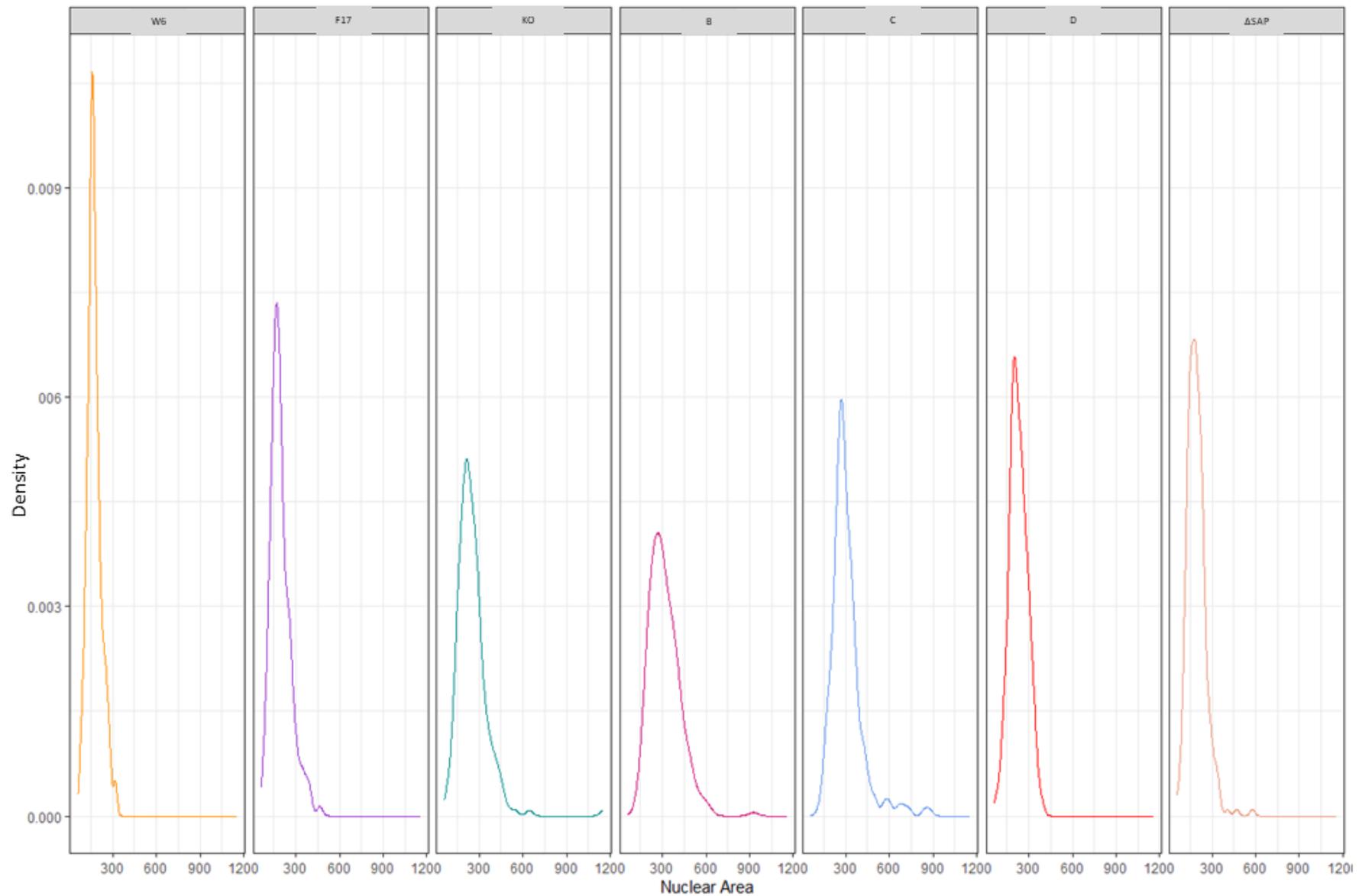


Figure 40: Density plots showing the distribution of nuclei area for each cell type.

Each cell type is colour coded. W6 (CIP29 rescue), CIP29-FL (CIP29- FLAG rescue), KO (CIP29-KNOCKOUT) and each truncation (B, C, D, and Δ SAP) (R=3) (collated samples, N= ~ 250 cells).

The Wilcoxon non-parametric test of each of the cell types reveals the true extent of the differences or not, between the cell populations (Table 12). Firstly, the W6 population only showed a slight difference (0.002 points) to the CIP29-FL suggesting these populations were highly similar in this aspect. Following this, W6 was also only slightly differently to the Δ SAP population. Furthermore, the Δ SAP population displayed no significant difference to the CIP29-FL, thus further indicating that the SAP domain may not influence the cell cycle phenotypes observed. Interestingly, the nuclear area size of truncation C was not significantly different to the CIP29-KO cell line. However, Truncations B and C were significantly different to all remaining cell lines, although these truncations were not significantly different from each other. This indicates truncations B and C may exhibit a phenotype which may exceed that of even the CIP29-KO cell lines.

Table 12: Testing the significance of the Nuclear Areas between the cell types using the Wilcoxon non-parametric test.

GROUP 1	GROUP 2	P.ADJUSTED	P. Values	Significance
W6	FL	4.8 E-2	0.048	*
W6	KO	5.5E-21	>0.001	****
W6	B	4.91 E-50	>0.001	****
W6	C	4.24 E-47	>0.001	****
W6	D	7.81 E-20	>0.001	****
W6	Δ SAP	6.00 E-3	0.006	**
KO	B	2.79 E-12	>0.001	****
KO	C	5.50 E-10	>0.001	****
KO	D	2.11 E-1	21.1	ns
KO	Δ SAP	2.43 E-13	>0.001	****
FL	KO	6.61 E-10	>0.001	****
FL	Δ SAP	8.51 E-1	85.1	ns
B	C	4.46 E-1	44.6	ns

5.2.6 *Investigating the localisation of the CIP29 truncations*

As the results so far had suggested that some of the CIP29 truncations exhibited a phenotype that mirrored that of the CIP29-KO cells, this was further investigated to provide a greater understanding of how the functionality of the truncations may have been lost. Therefore, the capability of each truncation to correctly localise was further investigated. To achieve this, subcellular fractionation was performed as described (refer to methods 2.3.4). After which the cytoplasmic, membrane, soluble nuclear, chromatin-bound and cytoskeletal fractions were analysed through western blotting. To ensure that each subcellular fraction was successfully distinct from carryover from the previous, each fraction was probed for proteins only found within that subcellular fraction. These included GAPDH for the cytoplasmic and histone H3 for the chromatin-bound fractions.

As CIP29 or its mutant derivatives were identified using either the α -FLAG or α -CIP29 antibody within the whole cell extracts, any absence of the CIP29- truncation mutants within a sub-cellular fraction would indicate a lack of localisation within this area. The results obtained suggested that the endogenous CIP29 within the MRC5-VI cells and the stably reintroduced FLAG-CIP29 within F17 cells, localised to the cytoplasmic, membrane, soluble nuclear and chromatin-bound subcellular fractions, but were not observed within the cytoskeletal fraction. When assessing the localisation of each of the truncations the results suggested that there was no significant hindrance in localisation within the cytoplasmic fraction for any of the truncations. Although, truncation B may struggle to localise within the membrane fraction, as this truncation was barely visible when detecting through either the α -FLAG or α -CIP29 antibodies. While the remaining truncations localised normally to this subcellular fraction. When analysing the soluble nuclear fraction, truncation B was again undetectable while, truncation C was only just detectible with the α -FLAG antibody. When investigating the chromatin bound fraction neither truncations B nor C were detectible through either antibody. In contrast, the localisation of truncations D and Δ SAP did not seem to be particularly affected and localised somewhat normally, as both truncations were detected alongside both controls in all the fractions, aside from the cytoskeletal (Figure 41).

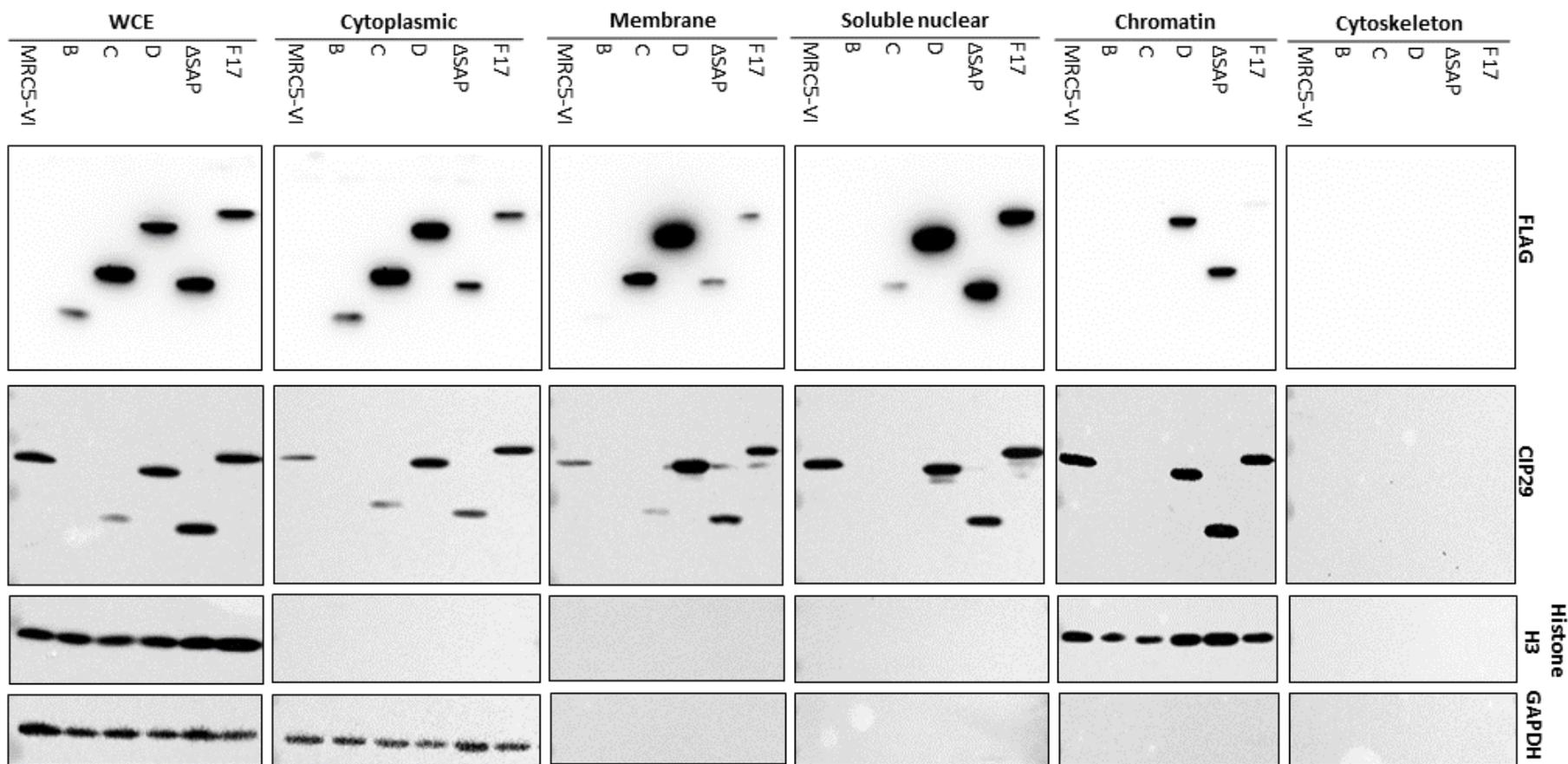


Figure 41: Assessing the functionality of each CIP29-truncation through localisation studies.

Cell fractionation of each cell type; MRC5-VI control, CIP29-truncations (B, C, D and Δ SAP) and F17 (CIP29- FLAG rescue). Cell fractionations include cytoplasmic, membrane, soluble nuclear, chromatin-bound and cytoskeletal. This is alongside whole cell extracts (WCE) of each cell type. Each fraction has been probed for specific proteins that should only be present in that individual fraction and are indicated alongside. The experiment was conducted in duplicate, and a representation is presented (R=2).

From the data collected it suggests that truncations B and C struggle to localise to the soluble nuclear and are completely absent from the chromatin bound subcellular fraction, this observation can be further supported by previous literature that suggests that CIP29 contains two nuclear localisation regions (NLR) within the C-terminal α -helical regions, which would both be absent from truncations B and C. As there is a small detection of truncation C within the soluble nuclear a third yet unknown NLR may be present, or perhaps a small amount of protein can be transported into the nucleus in association with binding partners. As well as housing the NLRs multiple phosphorylation and glycosylation sites can also be found within these helical regions. As protein modifications can also profoundly affect the activity and localisation of a protein, these sites may also play a role in stimulating the activity of CIP29, which in their absence may leave the protein defunct (Holden et al., 2017) (Fukuda et al., 2002) (Figure 42). This could explain the phenotypic defects that are observed within these cells as the failure to localise to the nucleus and the chromatin would likely be detrimental to the role of CIP29 within the TREX and AREX complexes.

Additionally, when investigating the highly conserved protein residues of CIP29, its C-terminal helices contain two very similar motifs. These motifs may interact with two copies of the same RNA sequence which is separated by a particular number of bases, or only allow a specifically orientated RNA to reside within the structural motifs (Jacobsen et al., 2016). This could potentially provide a high level of specificity for certain RNAs. Therefore, without this identification feature certain mRNAs may not be selected for and exported for the nucleus to become functional proteins.

Interestingly, the Δ SAP variant does not seem to be hindered when binding to chromatin as it is clearly detected within both the soluble nuclear and the chromatin bound fractions this is despite suggestions that the SAP domain is involved in chromosomal organization and DNA binding (Aravind and Koonin, 2000). This result may begin to explain why the Δ SAP truncation did not exhibit a significantly defective phenotype when compared to the rescue cell lines contrasting some the other CIP29-truncations such as B and C which suggests that the C-terminal end of the CIP29 holds a greater importance for the development of the cell cycle defects observed than the SAP domain. Furthermore, if the C-terminal region plays a fundamental role in specific mRNA selection, this could afford significant specificity for therapeutics as if CIP29 only handles a small subset, yet vitally important genes. Its disruption may not affect a cells' entire gene expression profile, which may reduce the potential for unwanted adverse effects such as the accumulation of DNA damage. This would require further analysis and additional clarification regarding which genes fail to be exported in the absence of CIP29, but a greater understanding of a proteins' molecular structure is key to deciphering its functionality and cellular impact which may then provide valuable information for drug discovery and therapeutic advancement.

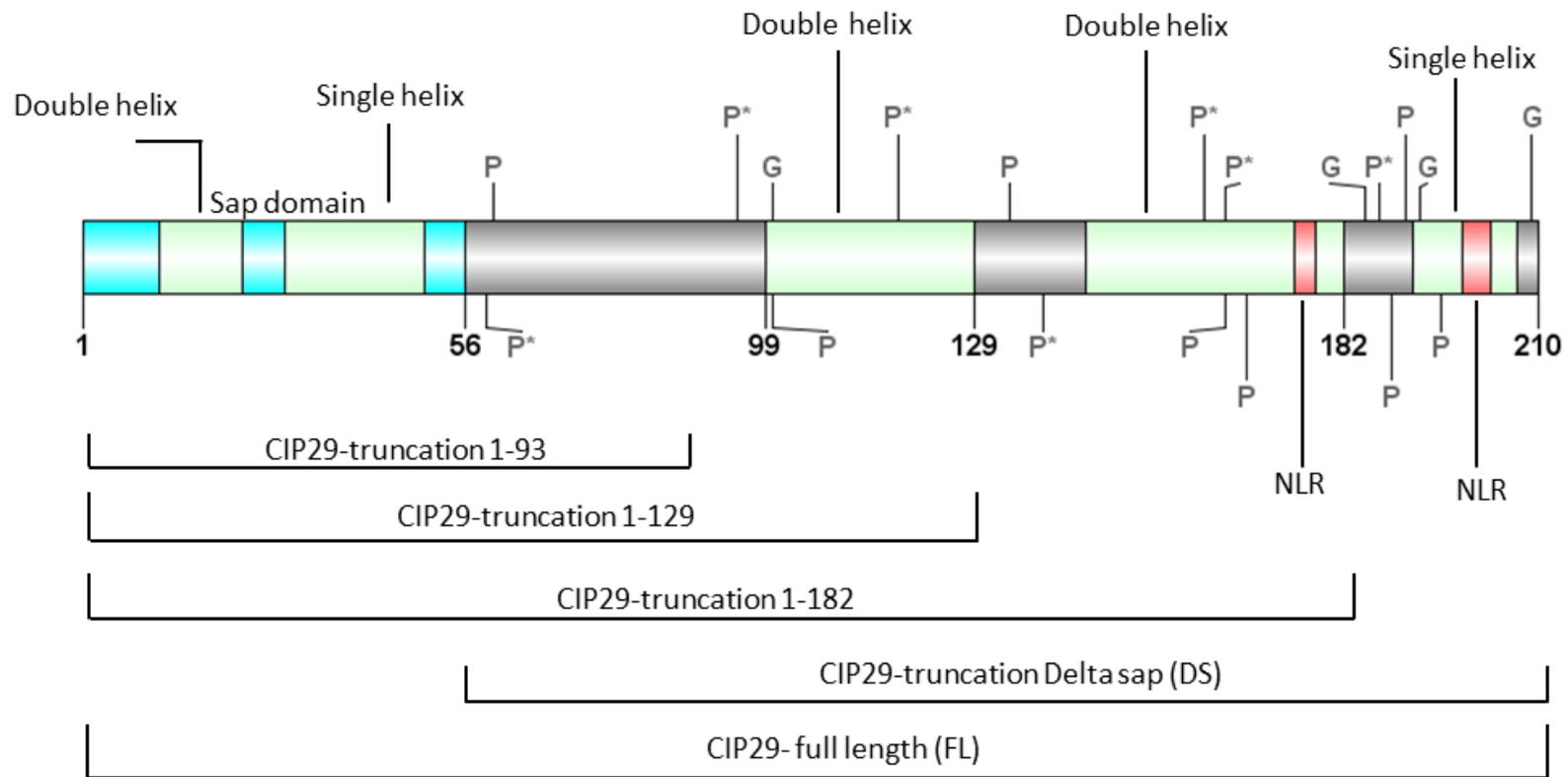


Figure 42: CIP29 schematic representation of the CIP29 truncations with regions of interest and potential regulation sites highlighted.

Regions of interest are highlighted, from left to right: the CIP29 protein structure featuring the SAP domain (light blue: 1-56aa), the helix domains (green: 12-24aa, 30-50aa, 99-129aa, 145-182aa and 192-207aa) and sites of interest (nuclear localisation regions (NLR), glycosylation sites (G) and phosphorylation sites that have been identified within our laboratory (P*) and other literature sources (P) (blue). Each spanning truncation is detailed below (Holden et al., 2017, Fukuda et al., 2002).

5.3 Discussion

To further investigate the molecular structure of the CIP29 protein, truncations of the protein were developed to various amino acid lengths of the full CIP29 protein including Truncation A (1-56 aa), Truncation B (1-93 aa), Truncation C (1-129 aa), Truncation D (1-182aa) and Δ SAP (55-210 aa). Each truncation featured various regions of interest such as the SAP domain or the C-terminal helix domains, while others did not, which enabled the different regions to be compared and evaluated for their importance and functionality (Figure 33 and 34). When developing the CIP29 truncations it was clear from the IVT and the transiently transfected western blot analysis that truncation A failed to be stability expressed and consequently, this construct was not further developed to a stable line and therefore the smallest CIP29 truncation used for further analysis was 1-93 aa in length, truncation B which featured only the SAP domain and a short unstructured region of 37 aa.

When analysing each of the truncations, it was clear that the smallest truncations B and C were significantly different to their counterparts. For example, these truncations displayed severe morphological defects such as multi lobed nuclei, large nuclei and supernumerary centrosomes. These morphological defects suggested that these cells struggled to separate their DNA properly and failed cytokinesis. This phenotype was also observed when analysing their cell cycle profiles through flow cytometry where again these truncations displayed a PI-stained profile of wide and shallow peaks that had significantly shifted along the X-Axis compared to the controls and other truncations (Figure 36). Additionally, these truncations displayed centrosome accumulation that even exceeded what was previously observed within the CIP29-KO cells, with only around 20% exhibiting a single centrosome, suggesting that once centrosome duplication had begun it was irreversible (Figures 37 and 38). However, from the γ -tubulin staining used it was difficult to determine if the centrosome defects are due to centrosome amplification or if a single centrosome had broken down into multiple fragments to only give the appearance of supernumerary centrosomes.

Although the failed cell division phenotype observed is quite severe, at present the mechanistic detail of how such a phenotype develops is still unclear. For example, do the smaller truncations and the CIP29-KO cells suffer from a reduction in the export of specific mRNAs due to a lack of the C-terminal region? Such a defect may cause a reduction in specific cell cycle proteins such as proteins that control the regulation and maturation of centrosomes. This disruption in turn may result in centrosome defects, leading to cytokinesis failure through a lack of spindle pole management and an inability to segregate the sister chromosomes to each daughter cell. This would leave the cells as

large mononucleated cells with a tetraploid DNA content, some of which may be capable of progressing through the cell cycle a second time and acquiring an even greater DNA content. However, the additional centrosomes may also be just a by-product of other cell cycle regulation or cytokinesis failures with multiple mature centrosomes existing within a single cell due to the lack of cell body division during cytokinesis and the ability for a cell to progress through the cell cycle multiple times without recombination.

The suggestion that the smallest CIP29 truncations also struggled to complete cytokinesis and become large mononucleated cells with a tetraploid DNA content was further supported by the fact that the nuclei of truncations B and C were indeed significantly larger than the MRC5-VI controls with an average nuclei size that exceeded even the CIP29-KO cells (table 12, figure 39). As B and C truncations displayed phenotypes that were often worse than what had previously been observed in the CIP29-KO cells it suggests that these truncations profoundly altered the molecular structure of the CIP29 protein and potentially the complexes that CIP29 associates with, such as the TREX or AREX complexes which may lead to other TREX or AREX components being displaced or blocked. Although further analysis is required to validate this assumption.

In contrast, the Δ SAP truncation displayed few morphological defects and exhibited a cell cycle profile like the CIP29-FL controls (Figure 36). This was somewhat surprising, as it has been suggested that the SAP domain may function in a similar manner to other such motifs often observed within RNA binding proteins, facilitating RNA or DNA binding and potentially playing a significant role in RNA processing and export. As the Δ SAP truncation did not share the same cell cycle or mitotic defects observed in truncations B and C or the CIP29-KO, it can be assumed that the SAP domain does not influence the cell division phenotype identified here, although other defects that have yet to be identified may reside within the Δ SAP truncation.

Our initial analysis of the localisation capabilities of each truncation further supports the assumption that the C-terminal helices are more influential to the cell division phenotype than the SAP region, as the truncations B and C struggled to localise to the nucleus and the chromatin regions when compared to their counterparts (Figure 41). This suggested that the nuclear localisation regions that reside within the C-terminal helices play a significant role in the proper function of the CIP29 protein. In addition, protein modifications such as phosphorylation or glycosylation events may also profoundly affect the activity and localisation of a protein as a significant number of these sites also reside within the C-terminal end of the CIP29 protein (Fukuda et al., 2002) (Holden et al., 2017). At this present time, it is unclear if CIP29 relies on glycosylation and the addition of glycans for its regulation however, CIP29 has been shown to be phosphorylated in a DNA damage dependent and

cell cycle dependent manner, therefore different regions of the protein could have alternative functions.

6 [Chapter 6: Discussion](#)

6.1 **Discussion**

6.1.1 *Chapter 1: Investigating CIP29 protein-protein interactions.*

At the initiation of this study very little was known about CIP29. The protein seemed to function as an RNA export factor which was often upregulated in cancer development and influenced cell proliferation (Fukuda et al., 2002). It had been suggested to function as an RNA export factor due to its SAP domain and localisation within the nucleus (Aravind and Koonin, 2000). CIP29 was often linked to the TREX complex, however its exact role within the complex and the proteins that it associated with was somewhat disputed (Sugiura et al., 2007) (Yamazaki et al., 2010) (Dufu et al., 2010). Over the course of this study, the evidence surrounding CIP29's involvement within the TREX complex became more substantial in both the literature and our own data, while its knockdown was shown to suppress mRNA export (Kang et al., 2020). To date no published literature has explored the phenotypic effect of CIP29-knockout, although it has been suggested that the export of key cytokinesis and cell proliferation genes may be effected when CIP29 is knocked down (Fujita et al., 2020).

To begin to expand on the previous characterisation of CIP29, our laboratory first aimed to substantiate the involvement of CIP29 within the TREX complex and confirm its' association with core components of the complex. The conformation that CIP29 plays a significant role in the TREX complex would suggest that the disruption of CIP29 maybe a significant driver of cancer progression due to the TREX complex' influence over gene expression. For example, when ALY is depleted, the activity of polymerase II is reduced, poly(A) RNAs accumulate within the nucleus and bulk gene expression is diminished (Stubbs and Conrad 2015).

Upon examination, the data strongly suggested that CIP29 does indeed interact with the core TREX complex components as a reliable interaction between CIP29 and UAP56 or ALY was achieved when utilising α -UAP56 and α -ALY antibodies during the immunoprecipitation studies. This association was also well documented in published literature through the course of the study. Unfortunately, the interaction was only captured on the rare occasion when using the α -CIP29 antibody to pull out UAP56 or ALY through CIP29. This discrepancy is likely due to the mechanics of the α -CIP29 antibody

which may masked the detection of the associated proteins or prevented the binding between CIP29 and the TREX complex components.

The study also suggested that ATP plays a significant role in these associations which follows previous reports whereby UAP56 utilizes ATP to fulfill its role as a mediator within the trimeric complex between UAP56, ALY and CIP29 (Dufu et al., 2010). This characteristic also explains the lack of association between CIP29 and ALY in the absence of ATP, which was observed in both the initial studies and the reciprocal immunoprecipitations. Without ATP, UAP56 would be unable to mediate the interaction between ALY and CIP29. Although some studies suggested that CIP29 could associate with UAP56 independent of ATP this result may show the formation of the AREX complex, which is independent of ATP, which may have been detected through the highly similar sequence homology of the two helicases (Yamazaki et al., 2010). Phosphorylation and the availability of ATP is strongly influenced by environmental stresses, such as hypoxia or heat shock (Bonora et al., 2012). Therefore, the potential for a level of selectivity between two helicases, depending on the cellular stimulus could be highly beneficial for restricting cancer development. For example, in a period of Hypoxia and ATP restriction, mRNA export would only be achieved via URH49 and CIP29 through the AREX complex. However, the upregulation of CIP29 could profoundly alter this dynamic, and may provide the opportunity to promote cancer progression as the AREX complex often exports mRNAs that encode cell cycle regulators and DNA repair proteins, which would enhance a cells' potential for survival (Fujita et al., 2020). Therefore, CIP29 may be vital for mRNA export within cancer cells, including solid tumors which often experience periods of hypoxia (Shi et al., 2021). Ultimately the targeting of CIP29 may offer significant avenues for cancer repression.

At the initiation of this study, there was the suggestion that CIP29 may play a more direct role in the cellular processes than through mRNA export alone. This hypothesis was suggested due to the detrimental cell cycle phenotype observed once CIP29 was removed, its phosphorylation in response to DNA damage and that mass-spectrometry utilizing *X. laevis* extracts, highlighted a collection of proteins may interact with the CIP29 protein that may not involve in RNA processing. These interacting proteins also held phenotypic relevance as many were implemented as cell cycle regulators and assistors. Unfortunately, human CIP29 failed to associate with several of the potentially novel proteins, including KIF2C which displayed some promise, but was found to be disingenuous. Despite this, some of the mass-spectrometry predictions were shown to exhibit a degree of interaction, such as CIP29 and anillin. This interaction was one of the top mass spectrometry analysis recommendations and held biological plausibility due to the protein's role in cytoskeletal dynamics and centrosome regulation during cytokinesis. However, upon further analysis

the associations detected became somewhat variable, while the α -anillin antibody failed during the reciprocal immunoprecipitations. Despite this, an association between the two proteins may still be plausible, although further clarification with alternative α -anillin antibodies that may respond better to immunoprecipitation conditions is still required. Additionally, the two proteins could be assessed for their localisation and affinity to one another through live cell imaging or cell fractionation studies. This may also highlight if anillin activity is hindered in the absence of CIP29.

Although at present, it is not completely clear how the disruption of CIP29 leads to the cellular defects observed. The results suggest that the cellular defects result from a defective mRNA export pathway, which may alter the gene expression of specific cell cycle proteins rather than CIP29 playing a direct role with specific cell cycle regulators like KIF2C or anillin, although some interactions cannot be completely disregarded at this time. Although anillin and proteins like it may not be directly associated with CIP29, they may still be subsequently influenced by the absence of CIP29 and a lack of functional mRNA export. This is supported when analysing the knockdown of UAP56 and URH49 (Yamazaki et al., 2010). Which upon investigation, either knockdown led to the accumulation of bulk poly(A) RNA within the nucleus, many of which were implemented in the cell cycle, DNA damage or centrosome regulators. For example, UAP56 knockdown caused a down regulation of genes such as: cyclin-dependent kinase 2 (CDK2), aurora kinase B, (AURKB) and centromere protein E, (CENPE), with an upregulation of cyclin-dependent kinase 3 (CDK3) and Checkpoint kinase 1 (CHEK1). While URH49 knockdown resulted in the down regulation of cyclin-dependent kinase 6 (CDK6), Aurora kinase B, (AURKB) and Centromere Protein A (CENPA) while the expression of cyclin E2 and Checkpoint kinase 1 (CHEK1) was stimulated. Furthermore, the study highlighted that the remaining helicase (UAP56 or URH49) was not able to compensate for the lost expression of specific genes. As CIP29 is implemented in both pathways, its absence may have a significant impact and may leave a myriad of un-exported genes without means of recovery. Therefore, the role of CIP29 within mRNA export may be enough alone to cause the cell cycle defects observed. To gain a greater understanding of this, the gene expression profiles of the CIP29-KO cells would need to be examined through RNA sequencing.

The results obtained suggest there are multiple opportunities for CIP29 to influence a cells' functionality and cellular fate, therefore CIP29 may hold promise as a therapeutic target, however there still significant characterisation to be achieved. The difficulties faced when conducting the immunoprecipitation assays and the potential that the epitopes or the antibody design may have hindered the studies' reliability, and potentially induced false results highlights the importance of understanding the molecular structure of a protein and ensuring that any additions do not interfere

with the protein of interest or its' associated proteins' functionality. The difficulties faced also highlight how little about CIP29 was known when the project first began, and therefore further phenotypic characterisation of the protein became a requirement for the project as if CIP29 did in fact have such a wide-reaching influence like the mass-spectrometry data suggested even if the additional research was unable to prove this at this present time, the effect of removing CIP29 would have a profound impact on the cell.

6.1.2 Chapter 2: Explore the phenotype caused by loss of CIP29.

To begin to elude the functional importance of CIP29, cellular functionality in the absence of CIP29 and its reintroduction was assessed through the study of CIP29-KO (MRC5-VI cells devoid of the protein through CRISPR-cas-9) and rescue cell lines W6 and F17 (in which the protein had been reintroduced). Before phenotypic analysis could begin the newly developed cell lines were validated through western blot and functional assessments were conducted. Localisation studies suggested that both the rescue cells lines localised normally, residing in the cytoplasmic, membrane, soluble nuclear and chromatin-bound subcellular fractions, but not within the cytoskeletal fraction. As this mirrored the expression of the endogenous CIP29, it suggested that the re-expressed CIP29 found within the rescue cells lines was expressed successfully and functioned as normal. The cell cycle profiles of the CIP29-KO cells mirrored previous data whereby cells accumulated within G2/M and exhibited a small tetraploid population. This was not observed within both the rescue cell lines, which exhibited cell cycle profiles like the MRC5-VI wildtype. This suggested that the defects seen within the CIP29-KO resulted from a loss of CIP29. Furthermore, when assessing the cell proliferation of each of the cell types the CIP29-KO cells proliferated more slowly compared to the MRC5-VI cells, while the two rescue cell lines, either matched (F17) or slightly exceeded (W6) the proliferation rate of the wildtype. This again suggested that the loss of CIP29 effected the cell's ability to proliferate normally and that this defect could be rectified with the re-introduction of CIP29. The accumulation of cells within the G2/M phases of the cell cycle and the reduced proliferation suggested that some level of cell cycle control remained within the CIP29-KO cells, and that the cells may experience some level of arrested or slowed progression within the G2/M phases of the cell cycle. It was plausible that this regulation may have resulted from a build-up of DNA damage from an absence of CIP29 and the lack of R-loop suppression (Wickramasinghe and Venkitaraman, 2016). Furthermore, previous laboratory studies had reported that within *X. laevis* cell-free extracts CIP29 was rapidly phosphorylated at Ser95 in response to DSBs by ATM. This phosphorylation was found to be cell cycle dependent, with CIP29 phosphorylated in response to late cell cycle arrest and DSBs within mitosis. This was observed in both *X. laevis* egg extracts and humans. However, when tested CIP29 was not found to have a direct role in DSB repair via DNA end joining. This suggests that DNA

damage dependent phosphorylation of CIP29 was linked to the cell cycle, but the exact role of CIP29 within the DNA damage response was still unclear (Holden et al., 2017) (Taylor, pers. comm.). To elude if this role is linked to the activation of the G2/M checkpoint CIP29-KO cells were examined for the activation of DNA damage response, but no activation of key DNA damage markers was observed. This suggested that despite their accumulation, the cells were not arrested within G2/M checkpoint.

As the subsequent DNA damage assessment concluded that the cells are not arrested within the G2/M transition despite the PI-stained flow cytometry data, it suggested that the CIP29-KO cells may have arrested within mitosis. However, upon further investigation the CIP29-KO cells including the >4N population, were not found to exhibit an increase in the key mitotic marker phosphorylated histone H3 with levels between prophase and anaphase of mitosis akin to its counterparts. This suggests that despite the greater number of CIP29-KO cells allocated to late G2/M phases of the cell cycle through their DNA content, the cells did not seem to reside within the late G2 or early M phases of the cells cycle. However, as histone H3 begins to become dephosphorylated after anaphase the CIP29-KO cells may still be accumulating in the lateral phases of mitosis between telophase and cytokinesis.

Upon further investigation through immunofluorescent staining, a significant number of CIP29-KO cells exhibited morphological defects that were indicative of cytokinesis failure, which is also a common mechanism for the development of tetraploid cells.

Cytokinesis is highly complex process with a multitude of proteins requiring direction and correct functionality for its completion. Therefore, cytokinesis failure can be linked to defective gene expression and a lack of activators. For example, if the correct initiator proteins are not available during the initiation of the cleavage furrow, ingression will fail. Or if it can be initiated, cytokinesis may not reach completion and result in furrow regression. This can lead to physical obstructions within the divisional plane and further hinder completion (Normand and King, 2010). The presence of either a single lagging chromosome or a chromatin bridge can cause cells to pause cleavage furrow ingression and attempt to stabilize the intercellular bridge through the activation of the abscission checkpoint (Petsalaki and Zachos, 2021). If the cleavage furrow is rapidly regressed, large amounts of chromatin or DNA can remain within the cleavage plane of the intercellular bridge, further blocking completion of cytokinesis and promoting chromosome instability (Fraschini, 2020). When assessing the morphological defects in more detail, the nuclei of the CIP29-KO cells were significantly larger than their counterparts, suggesting that these cells did indeed fail to complete cytokinesis and did not separate into two daughter cells. Furthermore, as the nuclei were large

mononucleates rather than separated binucleates it suggests that cytokinesis failure occurred early on, before the nuclear membrane has been established. Due to the range in nuclear area within the CIP29-KO cells the population displayed a bimodal distribution, indicating that the population that maintained a DNA content that was $>4n$ was substantial. This population also exhibited centrosome amplification which in some cases may have progressed to centrosome supernumerary and centrosome clustering. Although the exact number of incidents of these centrosome defects is unclear due to the staining procedures used. Unfortunately, staining PCMs can have several limitations. This is because: (i) PCM proteins can disassociate from the centrioles due to a breakdown in centrosome structure; (ii) PCM components can accumulate at microtubule minus ends independently of centrioles; and (iii) PCM proteins cluster in an asymmetric manner around nascent and parental centrioles, thus making the angle of visualization influence the perceived dimension of the centrosomes. Therefore, not all observations will represent true centrosomes (Nigg and Raff, 2009). To overcome these drawbacks, additional staining to identify the age of the centrosomes or a greater magnification of the centrosomes such as using super-resolution fluorescence microscopic techniques like SIM and STORM sub-diffraction, would aid the identification of true centrosomes. However, these procedures are both costly and time consuming (Mennella et al., 2012). Furthermore, the staining technique used does not distinguish between parental and juvenile centrosomes. Such a distinction could be achieved through staining with centrosomal protein 170 (CEP170) which would aid the understanding of how the additional centrosomes are made; either through the dysregulation of centrosome biogenesis or mitotic defects such as cytokinesis, as the additional centrosomes that are formed through cytokinesis are more likely to be fully mature (Kalkan et al., 2022).

As the CIP29-KO cells didn't not exhibit an increase in mitotic markers that would be indicative of G2/M arrest and that the cells were often larger with an increased DNA content, it was possible that the flow cytometry analysis was miss-allocating the cells into the G2/M phases of their cell cycle based on the cells' size and DNA content. Furthermore, as the cells showed signs of experiencing cytokinesis failure it was possible that the cells were in fact held within the G1 phase of the second cell cycle. To clarify this, the cells were examined for their CDT1 expression levels. Upon analysis, the CIP29-KO cell line was found to exhibit an abnormal CDT1 expression. CDT1 fluctuates depending on the phase of the cell cycle, its expression peaks during G1 to promote the formation of pre-replicative complexes and replication, before being degraded during S-phase to prevent re-replication. Finally, it begins to increase again during G2 and mitosis in preparation for the next cell cycle. Although, this is heavily regulated by geminin (Ballabeni et al., 2013).

Unfortunately, as CDT1 staining was conducted without S/G2-phase marker proteins, such as Cyclin A it is difficult to confirm that the CDT1 expression seen is truly within G1, but as the MRC5-VI controls only exhibited CDT1 expression when they only maintained a single centrosome, it suggests that the CDT1 staining is exclusively within G1 as such an expression and centrosome number is characteristic of a cell within G1 (Gergely and Basto, 2008).

As some of the CIP29-KO cells seem to undergo multiple rounds of replication these cells may undergo mitotic slippage. Mitotic slippage is thought to be a major contributing factor to drug resistance and is common within tumor cells that have chromosomal instability and are often exhibit abnormal ploidy (Sinha et al., 2019). As cells that undergo mitotic slippage do not enter mitosis, they do not exhibit the features of mitosis such as condensed chromosomes and nuclear envelope breakdown or histone H3 expression. Furthermore, a key feature of a cell that has experienced mitotic slippage is that its' DNA content increases by clearly delineated genome doublings which can be seen within the CIP29-KO cell line. This further suggests that the CIP29-KO cells that exhibit a >4N DNA content have obtained this through mitotic slippage. This is an important distinction from the aberrant process such as re-replication, which is characterized by uncontrolled and continuous re-initiation of DNA synthesis within a given S phase. This results in an increase in DNA content without a clearly recognizable genome doubling events (Lee et al., 2009).

The tetraploid sub-population did not seem to progress through multiple rounds of replication and instead, their DNA content remained at >4N but did not progress to further extremes such as 8n, it suggests further mitotic slippage was prevented. Interestingly, these cells do not seem to be lost from the population in significant numbers, as the cell proliferation of the CIP29-KO cells was not significantly reduced. Therefore, the tetraploid cells may exit the cell cycle and become senescent to avoid cell death. As the potential occurrence of mitotic slippage and cell senescence were only exhibited in a proportion of the CIP29-KO population, it suggests that some redundancy within the role of CIP29 remains and can compensate for the lack of CIP29 to some extent, leading to a more gradual and accumulative effect. As redundancy has already been shown within the TREX complex this again suggest that the role of CIP29 within the TREX complex and its absence within this complex is potentially the source of the defects observed.

To promote cell senescence, the potent cyclin-dependent kinase inhibitor (CKI) p21 needs to be activated, which can be achieved through the activation of p53. p21 is vital for cell cycle arrest in G1 as it controls inhibits cyclin D/CDK4 and cyclin E/CDK2 complexes to block the phosphorylation of protein substrates essential for the onset of S phase (Al Bitar and Gali-Muhtasib, 2019), (Abbas and Dutta, 2009).

Although our data suggests that the CIP29-KO cells do not exhibit the phosphorylation of p53 at Serine 15, it does not rule out the activation of the DNA damage response through different avenues. For example, the cellular stresses produced during the development of aneuploidy and tetraploid cells often trigger the MAP kinase p38. P38 phosphorylates p53 through serine 33, which then activates p21 to inhibit Cdk2–cyclin complexes and delay G1 (Harper et al., 1993) (el-Deiry et al., 1993) (Sherr and Roberts, 1999) (Mikule et al., 2007). It has also been suggested that abscission, the final stage of cytokinesis maybe an G1 event rather than occurring within mitosis. This theory is based on the time taken to progress through mitosis which for the majority (metaphase to telophase) is quite rapid, at under 20 minutes. However, this progression slows significantly during abscission, with the separation of the intercellular bridge taking over an hour to complete (Gershony et al., 2014). Considering this, it suggests that the CIP29-KO cells maybe under the control of G1 regulators when their journey through the cell cycle is disrupted. Furthermore, p38 and phosphorylated p53-Ser 33 have been shown to localise to centrosomes, to promote centrosome-associated cell-cycle arrest in response to centrosome protein depletion and a lack of centrosome integrity (Pihan et al., 1998). As the CIP29-KO cells also display a lack of centrosome integrity and regulation, such a DNA damage response may be activated. Clearly a more extensive investigation to the DNA damage response is required as our results indicated that this is a very nuanced response, without large amounts of DNA damage, despite the presence of aneuploidy.

As it is still unclear as to how a lack of CIP29 can have such an influential impact on the cell cycle, despite the high level of redundancy within its identified role as a TREX complex component, the molecular structure of CIP29 requires further investigation to identify regions of specific interest which may be vital to the functionality of the TREX complex and gene export. A better understanding of the molecular structure of CIP29 may also provide the foundation for translational medicine and drug development avenues.

6.1.3 Chapter 3: Exploring the molecular structure of the CIP29 protein and its functional domains.

As the results suggested that CIP29 may play an important role in the proper function of the cell cycle, a better understanding of the molecular structure of CIP29 may provide an indication as to why and may aid future investigations. For example, a greater understanding of the molecular structure would be beneficial when implementing secondary structures such as epitopes, antibodies and other modifications. Further analysis may also provide some level of explanation regarding the issues faced during the immunoprecipitation assays, which could be mitigated in future.

Furthermore, a greater knowledge may also provide an inference into its' most important functional domains, which may provide a suggestion to its' functionality and an indication of which regions

could be linked to the abhorrent phenotypes observed. Such information would be highly beneficial if CIP29 was to become a target for therapeutics as a knowledge of its valuable domains and regions may offer a starting point for targeted therapies.

To further investigate each region of the CIP29 protein, its currently known structural features were illustrated through the AI system AlphaFold. The use of an AI system was highly beneficial with each region of potential interest well portrayed in a 3-dimensional structure. This enabled the truncations to be easily and efficiently designed. The accuracy of the AI system is also highlighted in the stepwise decline in function as the truncations became more severe.

This difficulty can be observed within truncations B and C, which despite the presence of a SAP domain the truncations suffered significant morphological and cell cycle defects including larger nuclei, centrosome amplification and an increased in DNA content. Interestingly, these defects were indicative of what was observed with the CIP29-KO populations, suggesting that the C-terminal end of the CIP29 protein may be linked to cell cycle activity and potentially its regulation. The remaining truncations displayed a more typical morphology, including the Δ SAP which maintained a phenotype similar to the full-length protein in both the cell cycle profiles and nuclei sizes, as it was only found to be slightly different from the W6 rescue cell line. This suggests that despite the proposal that SAP domains are important for RNA and DNA binding of other RNA processing factors, the domain may not be as functionally important to CIP29. Or perhaps the two helices that reside within the SAP domain provide a level of mRNA selectivity for mRNAs that do not play a role in the biological processes identified the conducted phenotypic studies. While the double helices within the C-terminal end may offer a greater selectivity for genes that are needed for cell cycle regulation although further analysis is still required.

Furthermore, the C-terminal end houses multiple sites of interest, including phosphorylation sites, glycosylation sites and nuclear localisation regions (NLRs). These regions may control the proteins' activity, such as its potential for degradation or localisation. As CIP29 localized mainly to the nucleus with small amounts detected in the cytosol, its NLRs may play a significant role in its functionality. As a TREX complex component, CIP29 is required to exit and re-enter the nucleus to deliver mRNA cargo to the cytoplasm for its further translation and protein synthesis. While its potential role within chromosomal re-arrangement and organisation would also rely on functional NLS. This can be illustrated by the significant aberrant phenotypes of the B and C truncations and their inability or severe handicap to localise to the nucleus and chromatin regions, a defect that is not seen within the other truncations, including the Δ SAP. Furthermore, as the Δ SAP displayed few mitotic defects it suggests that once within the nucleus, the CIP29 protein can provide some level of function despite

the lack of an SAP domain. As the ability of the Δ SAP truncation to associate with core TREX complex components remains unexplored further clarification is required, however immunoprecipitation studies between each of the truncations and the core TREX or AREX complex components such as UAP56, URH49 and ALY would shed light on this.

Interestingly, the B and C truncations often exhibited defects that exceeded the CIP29-KO phenotype which suggested that not only did the truncations exhibit a reduction in functionality, they may have also been a hinderance to the cell. For example, as the truncations were not readily degraded, but lacked important regions such as glycosylation sites, these proteins may not be properly folded by the cell and therefore pose an additional cellular stress. Normally, glycosylation is completed within the Endoplasmic Reticulum (ER) whereby N-linked oligosaccharides are attached to the nascent polypeptide chain upon entry into the ER. These alterations aid protein folding and act as an important quality control measure (Schoberer et al., 2018). To prevent the accumulation or secretion of misfolded proteins, incompletely folded glycoproteins are removed from the ER by a highly conserved degradation process. During periods of cellular stress, this mechanism can become overwhelmed and improperly folded proteins can accumulate resulting in ER stress. If resolved, the unfolded protein response (UPR) can be activated. The upregulation of UPR-targeted genes not only increases the cells' capacity for protein folding, but also protein degradation and transport pathways, which help to alleviate the burden of misfolded protein within the ER (Read and Schröder, 2021). Although designed to promote cell survival in healthy cells, this pathway can also be beneficial for the progression of cancer (Yadav et al., 2014). At present it is unclear if the B and C truncations cause ER stress and UPR, but their ability to progress to a more severe phenotype without succumbing to apoptosis, suggest some protective mechanisms may be activated. Although this is another avenue which requires further exploration.

As well as glycosylation sites, truncations B and C also lack multiple phosphorylation sites. Phosphorylation is a key mechanism to regulate a proteins' activities and signalling potential (Ardito et al., 2017). Interestingly, previous data in *X. laevis* suggested that the phosphorylation in response to DNA damage was situated at Ser95 while the C-terminal region was found to be vastly phosphorylated, but there was no indication that this phosphorylation was because of the DNA damage response (Holden et al., 2017). As both truncations B and C would still house the Ser95 motif, it further suggests that the DNA damage response pathway may not be linked to the phenotypes observed. Regarding the phosphorylation within the C-terminal, the study did not rule out that other kinases may be responsible, including kinases that are implemented in cell cycle regulation such as the mitogen-activated protein kinase (MAPKs) or Polo-like kinases (PLKs). PLKs are

associated with tumour development and are over-expressed in many tumour cells which results in mitotic checkpoints being overridden to simulate immature cell division and aneuploidy two phenotypes that have been frequently observed within this study.

6.1.4 Summary and future perspectives

Although questions around the exact function of CIP29 still remain, the study has progressed the characterisation of CIP29 in a positive way and has provided novel avenues for further research. It can be concluded that despite issues surrounding the α -CIP29 antibody, that CIP29 is indeed a TREX complex component and that its incorporation within the final complex is ATP dependent. This conclusion is also supported in many published reports. The data presented here and a growing body of other published works also suggest that CIP29 is a vital component of the AREX complex, however specific URH49 antibodies would be required to confirm this within our own laboratory. Therefore, with the data collected it can be assumed that the mitotic defects observed within the CIP29-KO cell lines is a result of a mis-regulation of gene expression due to the role of CIP29 as an RNA export factor either through chromatin organisation and manipulation or as its role within mRNA export complexes such as the TREX and AREX complex, rather than CIP29 directly associating with and manipulating cell cycle regulators. To further clarify this, RNA sequencing could be used to give an indication of altered gene expression profiles due to the absence of CIP29 within the TREX or AREX complexes.

The loss of CIP29 produces a profound phenotype, which suggests that the cell cycle is severely hindered and mitotic defects can be readily observed. Including an increased DNA content which can progress to Tetraploidy, a reduction in cell proliferation, the presence of significantly larger, multi-lobed nuclei and centrosome supernumerary. These defects are indicative of cytokinesis failure, although further research such as advanced microscopy and live cell imaging is required to confirm this assumption. Furthermore, the lack of response within the G2/M cell cycle checkpoint suggests that the cell cycle is intact at the transition to mitosis, while no increase in histone H3 expression indicates that the cells are capable of progressing through the majority of M phase and may only encounter issues after anaphase, further indicating that the cells do indeed suffer a late mitotic defect such as cytokinesis failure. At present, the root cause of the mitotic failure is still unknown however, a reduction in specific cytokinesis proteins that coordinate the process could lead to its failure. This again suggests that the absence of CIP29 results in the deregulation of specific proteins due to a lack of mRNA processing and export that other RNA processing factors can not compensate for. This again could be validated through western blotting for protein expression and RNA sequencing. RNA sequencing could also be enhanced through immunoprecipitations and pull down

of specific transcripts that interact with CIP29 prior to sequencing such as e CLIP-seq (crosslinking with immunoprecipitation followed by sequencing) and RIP-seq (RNA immunoprecipitation followed by sequencing) (Hafner et al., 2021). Mitotic failure also be exasperated through a dysregulation of centrosome maturity and stability, which may lead to a disruption in chromosome segregation and a failure of cell division. The root cause of this may remain the same, with protein dysregulation and deficiency due to mRNA export defects. Further analysis into the maturity of the centrosomes such as staining with CEP170 would clarify this and may indicate if CIP29 plays a direct role in these defects or if the mitotic defects observed are a consequence of miss-regulated gene expression ((Rodríguez-Real et al., 2023).

The increase in CDT1 suggests that once the cells have failed to complete cytokinesis, they experience a slow progression through the subsequent G1 phase of the second cell cycle. At this point, the cells become miss-allocated by flow cytometry to reside within G2/M due to their increased size and DNA content following a lack of cell division. It is unclear at present what happens to these cells thereafter, although a combination of possible scenarios is most likely given the evidence gathered surrounding the CIP29-KO population. For example, the presence of a small but constant tetraploid population, a reduction in cell proliferation but not a significant lack of proliferation, and a lack of an extensive DNA damage response. These features suggest that despite failing to separate, some of the CIP29-KO cells can continue to progress through to a second round of DNA replication and become tetraploid cells but do not progress to a third DNA synthesis phase. This suggests that these cells are then either lost from the population through apoptosis or become permanently senescent. If the cells experience a level of cell senescence, they may remain metabolically active and therefore preventing the cell proliferation analysis of the population from dropping to significantly reduced levels. Again, further research into the accumulation of senescence makers such as p21 and b-galactosidase are required to provide an indication of the extent of senescence (Valieva et al., 2022). While alternative cell proliferation methodologies may be required to provide a more accurate interpretation of the population.

When analysing the molecular structure of CIP29, it became clear that the C-terminal helix domains play a significant role in the protein's functionality, while its SAP domain did not seem to be influential to the development of the mitotic defects observed within the CIP29-KO cells. The importance of the C-terminal end may be due to the specific sites that are situated within this section of the protein. It is clear its NLRs are significant to the protein as truncations B and C failed to localise to the nucleus and the chromatin regions, therefore it would be impossible for these truncations to function as mRNA export or processing factors. This again suggests that the mitotic

defects are because of a disruption within these processes, as these truncations also displayed the most severe mitotic defects. It is unclear whether the absence of the other sites such as glycosylation or phosphorylation sites further hinder the cell through mis-protein folding and the accumulation of cellular stress, but as the defects of the B and C truncations often exceeded previous observations within the CIP29-KO population it is likely, although further research is required. As the double helix domain within the C-terminal is also found near a NLR it is unclear if the lack of this domain also hinders the functionality of CIP29, as this region may provide a level of specificity to certain mRNAs that cannot be replicated by other TREX complex components. The targeting of this region or other sites, while leaving the NLR intact may offer a less severe disruption, as some CIP29 functionality within the nucleus may be retained. Therefore, further analysis of the C-terminal end from 145 aa – 210 aa may be required to fully distinguish the significance of each domain and site. Once clarified, significant regions of interest or combinations of regions could be targeted to disrupt CIP29's interaction within the TREX, AREX complexes or mRNA transcripts to reduce cell proliferation for therapeutic advancement whereby the synthesis of specific proteins could be controlled through the downregulation of CIP29. If this is possible, the regulation of CIP29 alone rather than the complete export pathway may reduce the possibility of adverse effects from a reduction of global transcript expression.

7 References

- ABBAS, T. & DUTTA, A. 2009. p21 in cancer: intricate networks and multiple activities. *Nat Rev Cancer*, 9, 400-14.
- ABBOTTS, R. & WILSON, D. M., 3RD 2017. Coordination of DNA single strand break repair. *Free Radic Biol Med*, 107, 228-244.
- ADAMSON, B., SMOGORZEWSKA, A., SIGOILLOT, F. D., KING, R. W. & ELLEDGE, S. J. 2012. A genome-wide homologous recombination screen identifies the RNA-binding protein RBMX as a component of the DNA-damage response. *Nature cell biology*, 14, 318-328.
- AEBERSOLD, R., AGAR, J. N., AMSTER, I. J., BAKER, M. S., BERTOZZI, C. R., BOJA, E. S., COSTELLO, C. E., CRAVATT, B. F., FENSELAU, C., GARCIA, B. A., GE, Y., GUNAWARDENA, J., HENDRICKSON, R. C., HERGENROTHER, P. J., HUBER, C. G., IVANOV, A. R., JENSEN, O. N., JEWETT, M. C., KELLEHER, N. L., KIESSLING, L. L., KROGAN, N. J., LARSEN, M. R., LOO, J. A., OGORZALEK LOO, R. R., LUNDBERG, E., MACCOSS, M. J., MALLICK, P., MOOTHA, V. K., MRKSICH, M., MUIR, T. W., PATRIE, S. M., PESAVENTO, J. J., PITTEI, S. J., RODRIGUEZ, H., SAGHATELIAN, A., SANDOVAL, W., SCHLÜTER, H., SECHI, S., SLAVOFF, S. A., SMITH, L. M., SNYDER, M. P., THOMAS, P. M., UHLÉN, M., VAN EYK, J. E., VIDAL, M., WALT, D. R., WHITE, F. M., WILLIAMS, E. R., WOHLSCHLAGER, T., WYSOCKI, V. H., YATES, N. A., YOUNG, N. L. & ZHANG, B. 2018. How many human proteoforms are there? *Nat Chem Biol*, 14, 206-214.
- AGUILERA, A. & GARCÍA-MUSE, T. 2012. R Loops: From Transcription Byproducts to Threats to Genome Stability. *Molecular Cell*, 46, 115-124.
- AL BITAR, S. & GALI-MUHTASIB, H. 2019. The Role of the Cyclin Dependent Kinase Inhibitor p21(cip1/waf1) in Targeting Cancer: Molecular Mechanisms and Novel Therapeutics. *Cancers (Basel)*, 11.
- ALLISON, S. J. & MILNER, J. 2003. Loss of p53 Has Site-Specific Effects on Histone H3 Modification, Including Serine 10 Phosphorylation Important for Maintenance of Ploidy. *Cancer Research*, 63, 6674.
- ANCZUKÓW, O., WARE, M. D., BUISSON, M., ZETOUNE, A. B., STOPPA-LYONNET, D., SINILNIKOVA, O. M. & MAZOYER, S. 2008. Does the nonsense-mediated mRNA decay mechanism prevent the synthesis of truncated BRCA1, CHK2, and p53 proteins? *Hum Mutat*, 29, 65-73.
- ANDERSEN, P. R., DOMANSKI, M., KRISTIANSEN, M. S., STORVALL, H., NTINI, E., VERHEGGEN, C., SCHEIN, A., BUNKENBORG, J., POSER, I., HALLAIS, M., SANDBERG, R., HYMAN, A., LACAVA, J., ROUT, M. P., ANDERSEN, J. S., BERTRAND, E. & JENSEN, T. H. 2013. The human cap-binding complex is functionally connected to the nuclear RNA exosome. *Nat Struct Mol Biol*, 20, 1367-76.
- ARAVIND, L. & KOONIN, E. V. 2000. SAP – a putative DNA-binding motif involved in chromosomal organization. *Trends in Biochemical Sciences*, 25, 112-114.
- ARDITO, F., GIULIANI, M., PERRONE, D., TROIANO, G. & LO MUZIO, L. 2017. The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy (Review). *Int J Mol Med*, 40, 271-280.
- ARMSTRONG, A. 2016. *Characterisation of CIP29 : a potential target of the eukaryotic DNA damage response*. Masters, Lancaster University.
- ATTWOOD, M. M., FABBRO, D., SOKOLOV, A. V., KNAPP, S. & SCHIÖTH, H. B. 2021. Trends in kinase drug discovery: targets, indications and inhibitor design. *Nat Rev Drug Discov*, 20, 839-861.
- AWASTHI, P., FOIANI, M. & KUMAR, A. 2015. ATM and ATR signaling at a glance. *Journal of Cell Science*, 128, 4255.
- AZIMZADEH, J. & MARSHALL, W. F. 2010. Building the centriole. *Curr Biol*, 20, R816-25.
- BAKHOUM, S. F. & CANTLEY, L. C. 2018. The Multifaceted Role of Chromosomal Instability in Cancer and Its Microenvironment. *Cell*, 174, 1347-1360.

- BALLABENI, A., ZAMPONI, R., MOORE, J. K., HELIN, K. & KIRSCHNER, M. W. 2013. Geminin deploys multiple mechanisms to regulate Cdt1 before cell division thus ensuring the proper execution of DNA replication. *Proc Natl Acad Sci U S A*, 110, E2848-53.
- BEDARD, P. L., HYMAN, D. M., DAVIDS, M. S. & SIU, L. L. 2020. Small molecules, big impact: 20 years of targeted therapy in oncology. *Lancet*, 395, 1078-1088.
- BELL, R. J., RUBE, H. T., KREIG, A., MANCINI, A., FOUSE, S. D., NAGARAJAN, R. P., CHOI, S., HONG, C., HE, D., PEKMEZCI, M., WIENCKE, J. K., WRENSCH, M. R., CHANG, S. M., WALSH, K. M., MYONG, S., SONG, J. S. & COSTELLO, J. F. 2015. Cancer. The transcription factor GABP selectively binds and activates the mutant TERT promoter in cancer. *Science*, 348, 1036-9.
- BENZ, E. J., JR. 2017. The Jeremiah Metzger Lecture Cancer in the Twenty-First Century: An Inside View from an Outsider. *Trans Am Clin Climatol Assoc*, 128, 275-297.
- BERTOLI, C., SKOTHEIM, J. M. & DE BRUIN, R. A. 2013. Control of cell cycle transcription during G1 and S phases. *Nat Rev Mol Cell Biol*, 14, 518-28.
- BESSON, A., DOWDY, S. F. & ROBERTS, J. M. 2008. CDK Inhibitors: Cell Cycle Regulators and Beyond. *Developmental Cell*, 14, 159-169.
- BLAGDEN, S. P. & GLOVER, D. M. 2003. Polar expeditions — provisioning the centrosome for mitosis. *Nature Cell Biology*, 5, 505-511.
- BLAGOSKLONNY, M. V. & PARDEE, A. B. 2002. The Restriction Point of the Cell Cycle. *Cell Cycle*, 1, 102-109.
- BONORA, M., PATERGNANI, S., RIMESSI, A., DE MARCHI, E., SUSKI, J. M., BONONI, A., GIORGI, C., MARCHI, S., MISSIROLI, S., POLETTI, F., WIECKOWSKI, M. R. & PINTON, P. 2012. ATP synthesis and storage. *Purinergic Signal*, 8, 343-57.
- BOVERI, T. 1895. Über die Befruchtungs- und Entwicklungsfähigkeit kernloser Seeigel-Eier und über die Möglichkeit ihrer Bastardirung. *Archiv für Entwicklungsmechanik der Organismen*, 2, 394-443.
- BRICKNER, J. R., GARZON, J. L. & CIMPRICH, K. A. 2022. Walking a tightrope: The complex balancing act of R-loops in genome stability. *Mol Cell*, 82, 2267-2297.
- BRITTON, S., FROMENT, C., FRIT, P., MONSARRAT, B., SALLES, B. & CALSOU, P. 2009. Cell nonhomologous end joining capacity controls SAF-A phosphorylation by DNA-PK in response to DNA double-strand breaks inducers. *Cell Cycle*, 8, 3717-3722.
- BROWN, V. L., HARWOOD, C. A., CROOK, T., CRONIN, J. G., KELSELL, D. P. & PROBY, C. M. 2004. p16INK4a and p14ARF Tumor Suppressor Genes Are Commonly Inactivated in Cutaneous Squamous Cell Carcinoma. *Journal of Investigative Dermatology*, 122, 1284-1292.
- BURAKOV, A. V. & NADEZHINA, E. S. 2020. Centering and Shifting of Centrosomes in Cells. *Cells*, 9.
- CARMENA, M., WHELOCK, M., FUNABIKI, H. & EARNSHAW, W. C. 2012. The chromosomal passenger complex (CPC): from easy rider to the godfather of mitosis. *Nature reviews. Molecular cell biology*, 13, 789-803.
- CARMODY, S. R. & WENTE, S. R. 2009. mRNA nuclear export at a glance. *Journal of Cell Science*, 122, 1933.
- CHAN, Y. A., HIETER, P. & STIRLING, P. C. 2014. Mechanisms of genome instability induced by RNA-processing defects. *Trends in Genetics*, 30, 245-253.
- CHANG, C. T., HAUTBERGUE, G. M., WALSH, M. J., VIPHAKONE, N., VAN DIJK, T. B., PHILIPSEN, S. & WILSON, S. A. 2013. Chtop is a component of the dynamic TREX mRNA export complex. *EMBO J*, 32, 473-86.
- CHOONG, M. L., TAN, L. K., LO, S. L., REN, E.-C., OU, K., ONG, S.-E., LIANG, R. C. M. Y., SEOW, T. K. & CHUNG, M. C. M. 2001. An integrated approach in the discovery and characterization of a novel nuclear protein over-expressed in liver and pancreatic tumors. *FEBS Letters*, 496, 109-116.
- CICCIA, A. & ELLEDGE, S. J. 2010. The DNA Damage Response: Making It Safe to Play with Knives. *Molecular Cell*, 40, 179-204.

- COHEN, P., CROSS, D. & JÄNNE, P. A. 2021. Kinase drug discovery 20 years after imatinib: progress and future directions. *Nat Rev Drug Discov*, 20, 551-569.
- CROSIO, C., FIMIA, G. M., LOURY, R., KIMURA, M., OKANO, Y., ZHOU, H., SEN, S., ALLIS, C. D. & SASSONE-CORSI, P. 2002. Mitotic phosphorylation of histone H3: spatio-temporal regulation by mammalian Aurora kinases. *Mol Cell Biol*, 22, 874-85.
- CRUTE, J. J., LADUCA, R. J., JOHANSON, K. O., MCHENRY, C. S. & BAMBARA, R. A. 1983. Excess beta subunit can bypass the ATP requirement for highly processive synthesis by the Escherichia coli DNA polymerase III holoenzyme. *Journal of Biological Chemistry*, 258, 11344-11349.
- DAMELIN, M. & BESTOR, T. H. 2007. The decatenation checkpoint. *Br J Cancer*, 96, 201-5.
- DASIKA, G. K., LIN, S.-C. J., ZHAO, S., SUNG, P., TOMKINSON, A. & LEE, E. Y. H. P. 1999. DNA damage-induced cell cycle checkpoints and DNA strand break repair in development and tumorigenesis. *Oncogene*, 18, 7883-7899.
- DE BONT, R. & VAN LAREBEKE, N. 2004. Endogenous DNA damage in humans: a review of quantitative data. *Mutagenesis*, 19, 169-85.
- DE SANTIS PUZZONIA, M., GONZALEZ, L., ASCENZI, S., CUNDARI, E. & DEGRASSI, F. 2016. Tetraploid cells produced by absence of substrate adhesion during cytokinesis are limited in their proliferation and enter senescence after DNA replication. *Cell Cycle*, 15, 274-82.
- DE SCHUTTER, K., JOUBÈS, J., COOLS, T., VERKEST, A., CORELLOU, F., BABIYCHUK, E., VAN DER SCHUEREN, E., BEECKMAN, T., KUSHNIR, S., INZÉ, D. & DE VEYLDER, L. 2007. WEE1 Kinase Controls Cell Cycle Arrest in Response to Activation of the DNA Integrity Checkpoint. *The Plant Cell*, 19, 211.
- DECKBAR, D., STIFF, T., KOCH, B., REIS, C., LÖBRICH, M. & JEGGO, P. A. 2010. The limitations of the G1-S checkpoint. *Cancer Res*, 70, 4412-21.
- DEMING, P. B., CISTULLI, C. A., ZHAO, H., GRAVES, P. R., PIWNICA-WORMS, H., PAULES, R. S., DOWNES, C. S. & KAUFMANN, W. K. 2001. The human decatenation checkpoint. *Proc Natl Acad Sci U S A*, 98, 12044-9.
- DERAKHSHAN, A., CHEN, Z. & VAN WAES, C. 2017. Therapeutic Small Molecules Target Inhibitor of Apoptosis Proteins in Cancers with Deregulation of Extrinsic and Intrinsic Cell Death Pathways. *Clin Cancer Res*, 23, 1379-1387.
- DIAS, A. P., DUFU, K., LEI, H. & REED, R. 2010. A role for TREX components in the release of spliced mRNA from nuclear speckle domains. *Nature Communications*, 1, 97.
- DREYFUSS, G., KIM, V. N. & KATAOKA, N. 2002. Messenger-RNA-binding proteins and the messages they carry. *Nat Rev Mol Cell Biol*, 3, 195-205.
- DUFU, K., LIVINGSTONE, M. J., SEEBACHER, J., GYGI, S. P., WILSON, S. A. & REED, R. 2010. ATP is required for interactions between UAP56 and two conserved mRNA export proteins, Aly and CIP29, to assemble the TREX complex. *Genes Dev*, 24, 2043-53.
- DUPRÉ, A., BOYER-CHATENET, L. & GAUTIER, J. 2006. Two-step activation of ATM by DNA and the Mre11–Rad50–Nbs1 complex. *Nature Structural & Molecular Biology*, 13, 451-457.
- EL-DEIRY, W. S., TOKINO, T., VELCULESCU, V. E., LEVY, D. B., PARSONS, R., TRENT, J. M., LIN, D., MERCER, W. E., KINZLER, K. W. & VOGELSTEIN, B. 1993. WAF1, a potential mediator of p53 tumor suppression. *Cell*, 75, 817-25.
- FANT, X., SRSEN, V., ESPIGAT-GEORGER, A. & MERDES, A. 2009. Nuclei of non-muscle cells bind centrosome proteins upon fusion with differentiating myoblasts. *PLoS One*, 4, e3303.
- FERGUSON, F. M. & GRAY, N. S. 2018. Kinase inhibitors: the road ahead. *Nat Rev Drug Discov*, 17, 353-377.
- FLOOR, S. N. & DOUDNA, J. A. 2016. Tunable protein synthesis by transcript isoforms in human cells. *eLife*, 5, e10921.
- FOLCO, E. G., LEE, C. S., DUFU, K., YAMAZAKI, T. & REED, R. 2012. The proteins PDIP3 and ZC11A associate with the human TREX complex in an ATP-dependent manner and function in mRNA export. *PLoS One*, 7, e43804.
- FRASCHINI, R. 2020. Cytokinesis in Eukaryotic Cells: The Furrow Complexity at a Glance. *Cells*, 9.

- FUJITA, K.-I., YAMAZAKI, T., HARADA, K., SENO, S., MATSUDA, H. & MASUDA, S. 2020. URH49 exports mRNA by remodeling complex formation and mediating the NXF1-dependent pathway. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 1863, 194480.
- FUKUDA, S. & PELUS, L. M. 2005. Growth inhibitory effect of Hcc-1/CIP29 is associated with induction of apoptosis, not just with G2/M arrest. *Cell Mol Life Sci*, 62, 1526-7.
- FUKUDA, S., WU, D. W., STARK, K. & PELUS, L. M. 2002. Cloning and characterization of a proliferation-associated cytokine-inducible protein, CIP29. *Biochem Biophys Res Commun*, 292, 593-600.
- GARDINER, M., TOTH, R., VANDERMOERE, F., MORRICE, NICHOLAS A. & ROUSE, J. 2008. Identification and characterization of FUS/TLS as a new target of ATM. *Biochemical Journal*, 415, 297-307.
- GERGELY, F. & BASTO, R. 2008. Multiple centrosomes: together they stand, divided they fall. *Genes Dev*, 22, 2291-6.
- GERSHONY, O., PE'ER, T., NOACH-HIRSH, M., ELIA, N. & TZUR, A. 2014. Cytokinetic abscission is an acute G1 event. *Cell Cycle*, 13, 3436-41.
- GOLOVANOV, A. P., HAUTBERGUE, G. M., TINTARU, A. M., LIAN, L.-Y. & WILSON, S. A. 2006. The solution structure of REF2-I reveals interdomain interactions and regions involved in binding mRNA export factors and RNA. *RNA (New York, N.Y.)*, 12, 1933-1948.
- GOODALL, G. J. & WICKRAMASINGHE, V. O. 2021. RNA in cancer. *Nature Reviews Cancer*, 21, 22-36.
- GREENE, B. L., KANG, G., CUI, C., BENNATI, M., NOCERA, D. G., DRENNAN, C. L. & STUBBE, J. 2020. Ribonucleotide Reductases: Structure, Chemistry, and Metabolism Suggest New Therapeutic Targets. *Annu Rev Biochem*, 89, 45-75.
- GÜTTINGER, S., LAURELL, E. & KUTAY, U. 2009. Orchestrating nuclear envelope disassembly and reassembly during mitosis. *Nat Rev Mol Cell Biol*, 10, 178-91.
- HAFNER, M., KATSANTONI, M., KÖSTER, T., MARKS, J., MUKHERJEE, J., STAIGER, D., ULE, J. & ZAVOLAN, M. 2021. CLIP and complementary methods. *Nature Reviews Methods Primers*, 1, 20.
- HANAHAAN, D. & WEINBERG, R. A. 2000. The Hallmarks of Cancer. *Cell*, 100, 57-70.
- HANAHAAN, D. & WEINBERG, R. A. 2011. Hallmarks of cancer: the next generation. *Cell*, 144, 646-74.
- HARPER, J. W., ADAMI, G. R., WEI, N., KEYOMARSI, K. & ELLEDGE, S. J. 1993. The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. *Cell*, 75, 805-16.
- HASHII, Y., KIM, J. Y., SAWADA, A., TOKIMASA, S., HIROYUKI, F., OHTA, H., MAKIKO, K., TAKIHARA, Y., OZONO, K. & HARA, J. 2004. A novel partner gene CIP29 containing a SAP domain with MLL identified in infantile myelomonocytic leukemia. *Leukemia*, 18, 1546-1548.
- HAUTBERGUE, G. M., HUNG, M.-L., WALSH, M. J., SNIJDERS, A. P. L., CHANG, C.-T., JONES, R., PONTING, C. P., DICKMAN, M. J. & WILSON, S. A. 2009. UIF, a New mRNA export adaptor that works together with REF/ALY, requires FACT for recruitment to mRNA. *Current biology : CB*, 19, 1918-1924.
- HEATH, C. G., VIPHAKONE, N. & WILSON, S. A. 2016. The role of TREX in gene expression and disease. *The Biochemical journal*, 473, 2911-2935.
- HENGEL, S. R., SPIES, M. A. & SPIES, M. 2017. Small-Molecule Inhibitors Targeting DNA Repair and DNA Repair Deficiency in Research and Cancer Therapy. *Cell Chem Biol*, 24, 1101-1119.
- HIROTA, T., GERLICH, D., KOCH, B., ELLENBERG, J. & PETERS, J.-M. 2004. Distinct functions of condensin I and II in mitotic chromosome assembly. *Journal of Cell Science*, 117, 6435.
- HOGG, S. J., BEAVIS, P. A., DAWSON, M. A. & JOHNSTONE, R. W. 2020. Targeting the epigenetic regulation of antitumour immunity. *Nat Rev Drug Discov*, 19, 776-800.
- HOLDEN, J., TAYLOR, E. M. & LINDSAY, H. D. 2017. Cip29 is phosphorylated following activation of the DNA damage response in *Xenopus* egg extracts. *PLoS One*, 12, e0181131.
- JACKSON, S. P. & BARTEK, J. 2009. The DNA-damage response in human biology and disease. *Nature*, 461, 1071-1078.

- JACOBSEN, J. O., ALLEN, M. D., FREUND, S. M. & BYCROFT, M. 2016. High-resolution NMR structures of the domains of *Saccharomyces cerevisiae* Tho1. *Acta Crystallogr F Struct Biol Commun*, 72, 500-6.
- JEAN-PHILIPPE, J., PAZ, S. & CAPUTI, M. 2013. hnRNP A1: the Swiss army knife of gene expression. *International journal of molecular sciences*, 14, 18999-19024.
- JIMENO, S., LUNA, R., GARCIA-RUBIO, M. & AGUILERA, A. 2006a. Tho1, a novel hnRNP, and Sub2 provide alternative pathways for mRNP biogenesis in yeast THO mutants. *Mol Cell Biol*, 26, 4387-98.
- JIMENO, S., LUNA, R., GARCÍA-RUBIO, M. & AGUILERA, A. 2006b. Tho1, a novel hnRNP, and Sub2 provide alternative pathways for mRNP biogenesis in yeast THO mutants. *Molecular and cellular biology*, 26, 4387-4398.
- JUMPER, J., EVANS, R., PRITZEL, A., GREEN, T., FIGURNOV, M., RONNEBERGER, O., TUNYASUVUNAKOOL, K., BATES, R., ŽÍDEK, A., POTAPENKO, A., BRIDGLAND, A., MEYER, C., KOHL, S. A. A., BALLARD, A. J., COWIE, A., ROMERA-PAREDES, B., NIKOLOV, S., JAIN, R., ADLER, J., BACK, T., PETERSEN, S., REIMAN, D., CLANCY, E., ZIELINSKI, M., STEINEGGER, M., PACHOLSKA, M., BERGHAMMER, T., BODENSTEIN, S., SILVER, D., VINYALS, O., SENIOR, A. W., KAVUKCUOGLU, K., KOHLI, P. & HASSABIS, D. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596, 583-589.
- JUNG, C. K., JUNG, J. H., LEE, K. Y., KANG, C. S., KIM, M., KO, Y. H. & OH, C. S. 2007. Centrosome abnormalities in non-small cell lung cancer: correlations with DNA aneuploidy and expression of cell cycle regulatory proteins. *Pathol Res Pract*, 203, 839-47.
- KALKAN, B. M., OZCAN, S. C., QUINTYNE, N. J., REED, S. L. & ACILAN, C. 2022. Keep Calm and Carry on with Extra Centrosomes. *Cancers (Basel)*, 14.
- KANG, G. J., PARK, M. K., BYUN, H. J., KIM, H. J., KIM, E. J., YU, L., KIM, B., SHIM, J. G., LEE, H. & LEE, C. H. 2020. SARNP, a participant in mRNA splicing and export, negatively regulates E-cadherin expression via interaction with pinin. *J Cell Physiol*, 235, 1543-1555.
- KATAHIRA, J., OKUZAKI, D., INOUE, H., YONEDA, Y., MAEHARA, K. & OHKAWA, Y. 2013. Human TREX component Thoc5 affects alternative polyadenylation site choice by recruiting mammalian cleavage factor I. *Nucleic Acids Research*, 41, 7060-7072.
- KIM, E. R., SELYUTINA, A. A., BULDAKOV, I. A., EVDOKIMOVA, V., OVCHINNIKOV, L. P. & SOROKIN, A. V. 2013. The proteolytic YB-1 fragment interacts with DNA repair machinery and enhances survival during DNA damaging stress. *Cell cycle (Georgetown, Tex.)*, 12, 3791-3803.
- KISHI, H., NAKAGAWA, K., MATSUMOTO, M., SUGA, M., ANDO, M., TAYA, Y. & YAMAIZUMI, M. 2001. Osmotic shock induces G1 arrest through p53 phosphorylation at Ser33 by activated p38MAPK without phosphorylation at Ser15 and Ser20. *J Biol Chem*, 276, 39115-22.
- KRÄMER, A., MAIER, B. & BARTEK, J. 2011. Centrosome clustering and chromosomal (in)stability: a matter of life and death. *Mol Oncol*, 5, 324-35.
- KROKAN, H. E. & BJØRÅS, M. 2013. Base excision repair. *Cold Spring Harb Perspect Biol*, 5, a012583.
- LANDRIEU, P., BAETS, J. & DE JONGHE, P. 2013. Chapter 146 - Hereditary motor-sensory, motor, and sensory neuropathies in childhood. In: DULAC, O., LASSONDE, M. & SARNAT, H. B. (eds.) *Handbook of Clinical Neurology*. Elsevier.
- LAWRENCE, K. S. & ENGBRECHT, J. 2015. The spindle assembly checkpoint: More than just keeping track of the spindle. *Trends in cell & molecular biology*, 10, 141-150.
- LEAW, C. L., REN, E. C. & CHOONG, M. L. 2004. Hcc-1 is a novel component of the nuclear matrix with growth inhibitory function. *Cell Mol Life Sci*, 61, 2264-73.
- LEE, H. O., DAVIDSON, J. M. & DURONIO, R. J. 2009. Endoreplication: polyploidy with purpose. *Genes Dev*, 23, 2461-77.
- LI, Z., SUN, B., CLEWELL, R. A., ADELEYE, Y., ANDERSEN, M. E. & ZHANG, Q. 2013. Dose-Response Modeling of Etoposide-Induced DNA Damage Response. *Toxicological Sciences*, 137, 371-384.

- LINGLE, W. L. & SALISBURY, J. L. 1999. Altered centrosome structure is associated with abnormal mitoses in human breast tumors. *Am J Pathol*, 155, 1941-51.
- LIU, S. L., MAY, J. R., HELGESON, L. A. & NOLEN, B. J. 2013. Insertions within the actin core of actin-related protein 3 (Arp3) modulate branching nucleation by Arp2/3 complex. *J Biol Chem*, 288, 487-97.
- LOK, T. M., WANG, Y., XU, W. K., XIE, S., MA, H. T. & POON, R. Y. C. 2020. Mitotic slippage is determined by p31comet and the weakening of the spindle-assembly checkpoint. *Oncogene*, 39, 2819-2834.
- LUO, Y., AHMAD, E. & LIU, S.-T. 2018. MAD1: Kinetochore Receptors and Catalytic Mechanisms. *Frontiers in Cell and Developmental Biology*, 6.
- MA, M., RODRIGUEZ, A. & SUGIMOTO, K. 2020. Activation of ATR-related protein kinase upon DNA damage recognition. *Current Genetics*, 66, 327-333.
- MAIATO, H., GOMES, A. M., SOUSA, F. & BARISIC, M. 2017. Mechanisms of Chromosome Congression during Mitosis. *Biology*, 6, 13.
- MALUMBRES, M. & BARBACID, M. 2009. Cell cycle, CDKs and cancer: a changing paradigm. *Nature Reviews Cancer*, 9, 153-166.
- MAO, Z., BOZZELLA, M., SELUANOV, A. & GORBUNOVA, V. 2008. DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells. *Cell Cycle*, 7, 2902-6.
- MARTINEZ-CONTRERAS, R., CLOUTIER, P., SHKRETA, L., FISETTE, J. F., REVIL, T. & CHABOT, B. 2007. hnRNP proteins and splicing control. *Adv Exp Med Biol*, 623, 123-47.
- MATSON, J. P. & COOK, J. G. 2017. Cell cycle proliferation decisions: the impact of single cell analyses. *Febs j*, 284, 362-375.
- MAYR, C. & BARTEL, D. P. 2009. Widespread Shortening of 3'UTRs by Alternative Cleavage and Polyadenylation Activates Oncogenes in Cancer Cells. *Cell*, 138, 673-684.
- MCINTOSH, J. R., MOLODTSOV, M. I. & ATAULLAKHANOV, F. I. 2012. Biophysics of mitosis. *Quarterly reviews of biophysics*, 45, 147-207.
- MENNELLA, V., KESZTHELYI, B., MCDONALD, K. L., CHHUN, B., KAN, F., ROGERS, G. C., HUANG, B. & AGARD, D. A. 2012. Subdiffraction-resolution fluorescence microscopy reveals a domain of the centrosome critical for pericentriolar material organization. *Nat Cell Biol*, 14, 1159-68.
- MERINO, D., KELLY, G. L., LESSENE, G., WEI, A. H., ROBERTS, A. W. & STRASSER, A. 2018. BH3-Mimetic Drugs: Blazing the Trail for New Cancer Medicines. *Cancer Cell*, 34, 879-891.
- MIERZWA, B. & GERLICH, DANIEL W. 2014. Cytokinetic Abscission: Molecular Mechanisms and Temporal Control. *Developmental Cell*, 31, 525-538.
- MIETTINEN, T. P. & BJÖRKLUND, M. 2017. Mitochondrial Function and Cell Size: An Allometric Relationship. *Trends in Cell Biology*, 27, 393-402.
- MIKULE, K., DELAVAL, B., KALDIS, P., JURCYZK, A., HERGERT, P. & DOXSEY, S. 2007. Loss of centrosome integrity induces p38-p53-p21-dependent G1-S arrest. *Nat Cell Biol*, 9, 160-70.
- MILLER, A. L. 2011. The contractile ring. *Current biology : CB*, 21, R976-R978.
- NEGRINI, S., GORGOLIS, V. G. & HALAZONETIS, T. D. 2010. Genomic instability--an evolving hallmark of cancer. *Nat Rev Mol Cell Biol*, 11, 220-8.
- NIGG, E. A. & RAFF, J. W. 2009. Centrioles, Centrosomes, and Cilia in Health and Disease. *Cell*, 139, 663-678.
- NORMAND, G. & KING, R. W. 2010. Understanding cytokinesis failure. *Adv Exp Med Biol*, 676, 27-55.
- OBENG, E. A., STEWART, C. & ABDEL-WAHAB, O. 2019. Altered RNA Processing in Cancer Pathogenesis and Therapy. *Cancer Discov*, 9, 1493-1510.
- PAULOVICH, A. G., TOCZYSKI, D. P. & HARTWELL, L. H. 1997. When checkpoints fail. *Cell*, 88, 315-21.
- PERRIN-VIDOZ, L., SINILNIKOVA, O. M., STOPPA-LYONNET, D., LENOIR, G. M. & MAZOYER, S. 2002. The nonsense-mediated mRNA decay pathway triggers degradation of most BRCA1 mRNAs bearing premature termination codons. *Hum Mol Genet*, 11, 2805-14.
- PETSALAKI, E. & ZACHOS, G. 2021. The Abscission Checkpoint: A Guardian of Chromosomal Stability. *Cells*, 10.

- PIHAN, G. A., PUROHIT, A., WALLACE, J., KNECHT, H., WODA, B., QUESENBERRY, P. & DOXSEY, S. J. 1998. Centrosome defects and genetic instability in malignant tumors. *Cancer Res*, 58, 3974-85.
- PODHORECKA, M., SKLADANOWSKI, A. & BOZKO, P. 2010. H2AX Phosphorylation: Its Role in DNA Damage Response and Cancer Therapy. *J Nucleic Acids*, 2010.
- PONOMARENKO, E. A., POVERENNAYA, E. V., ILGISONIS, E. V., PYATNITSKIY, M. A., KOPYLOV, A. T., ZGODA, V. G., LISITSA, A. V. & ARCHAKOV, A. I. 2016. The Size of the Human Proteome: The Width and Depth. *Int J Anal Chem*, 2016, 7436849.
- PRYOR, A., TUNG, L., YANG, Z., KAPADIA, F., CHANG, T. H. & JOHNSON, L. F. 2004. Growth-regulated expression and G0-specific turnover of the mRNA that encodes URH49, a mammalian DEXH/D box protein that is highly related to the mRNA export protein UAP56. *Nucleic Acids Res*, 32, 1857-65.
- PÜHRINGER, T., HOHMANN, U., FIN, L., PACHECO-FIALLOS, B., SCHELLHAAS, U., BRENNECKE, J. & PLASCHKA, C. 2020. Structure of the human core transcription-export complex reveals a hub for multivalent interactions. *Elife*, 9.
- RAMANATHAN, A., ROBB, G. B. & CHAN, S.-H. 2016. mRNA capping: biological functions and applications. *Nucleic acids research*, 44, 7511-7526.
- READ, A. & SCHRÖDER, M. 2021. The Unfolded Protein Response: An Overview. *Biology (Basel)*, 10.
- RIEDER, C. L. & PALAZZO, R. E. 1992. Colcemid and the mitotic cycle. *J Cell Sci*, 102 (Pt 3), 387-92.
- RODRÍGUEZ-REAL, G., DOMÍNGUEZ-CALVO, A., PRADOS-CARVAJAL, R., BAYONA-FELIÚ, A., GOMES-PEREIRA, S., BALESTRA, F. R. & HUERTAS, P. 2023. Centriolar subdistal appendages promote double-strand break repair through homologous recombination. *EMBO Rep*, 24, e56724.
- SATO, N., MIZUMOTO, K., NAKAMURA, M., NAKAMURA, K., KUSUMOTO, M., NIIYAMA, H., OGAWA, T. & TANAKA, M. 1999. Centrosome abnormalities in pancreatic ductal carcinoma. *Clin Cancer Res*, 5, 963-70.
- SATO, Y., TSUCHIYA, H., YAMAGATA, A., OKATSU, K., TANAKA, K., SAEKI, Y. & FUKAI, S. 2019. Structural insights into ubiquitin recognition and Ufd1 interaction of Npl4. *Nature Communications*, 10, 5708.
- SCHÄRER, O. D. 2013. Nucleotide excision repair in eukaryotes. *Cold Spring Harb Perspect Biol*, 5, a012609.
- SCHEER, U. 2014. Historical roots of centrosome research: discovery of Boveri's microscope slides in Würzburg. *Philos Trans R Soc Lond B Biol Sci*, 369.
- SCHOBERER, J., SHIN, Y. J., VAVRA, U., VEIT, C. & STRASSER, R. 2018. Analysis of Protein Glycosylation in the ER. *Methods Mol Biol*, 1691, 205-222.
- SCHOLEY, J. M., BRUST-MASCHER, I. & MOGILNER, A. 2003. Cell division. *Nature*, 422, 746-752.
- SCHULZE, W. M., STEIN, F., RETTEL, M., NANAQ, M. & CUSACK, S. 2018. Structural analysis of human ARS2 as a platform for co-transcriptional RNA sorting. *Nature Communications*, 9, 1701.
- SHEN, M., ZHANG, R., JIA, W., ZHU, Z., ZHAO, X., ZHAO, L., HUANG, G. & LIU, J. 2021. Nuclear scaffold protein p54nrb/NONO facilitates the hypoxia-enhanced progression of hepatocellular carcinoma. *Oncogene*, 40, 4167-4183.
- SHERR, C. J. & ROBERTS, J. M. 1999. CDK inhibitors: positive and negative regulators of G1-phase progression. *Genes Dev*, 13, 1501-12.
- SHI, R., LIAO, C. & ZHANG, Q. 2021. Hypoxia-Driven Effects in Cancer: Characterization, Mechanisms, and Therapeutic Implications. *Cells*, 10.
- SHI, Y. 2017. Mechanistic insights into precursor messenger RNA splicing by the spliceosome. *Nature Reviews Molecular Cell Biology*, 18, 655-670.
- SINHA, D., DUIJF, P. H. G. & KHANNA, K. K. 2019. Mitotic slippage: an old tale with a new twist. *Cell Cycle*, 18, 7-15.
- SMITTENAAR, C. R., PETERSEN, K. A., STEWART, K. & MOITT, N. 2016. Cancer incidence and mortality projections in the UK until 2035. *British journal of cancer*, 115, 1147-1155.

- SNYDER, A. K., WILLIAMS, C. R., JOHNSON, A., O'DONNELL, M. & BLOOM, L. B. 2004. Mechanism of loading the Escherichia coli DNA polymerase III sliding clamp: II. Uncoupling the beta and DNA binding activities of the gamma complex. *J Biol Chem*, 279, 4386-93.
- SOLIMAN, T. N., KEIFENHEIM, D., PARKER, P. J. & CLARKE, D. J. 2023. Cell cycle responses to Topoisomerase II inhibition: Molecular mechanisms and clinical implications. *J Cell Biol*, 222.
- SPIRA, F., CUYLEN-HAERING, S., MEHTA, S., SAMWER, M., REVERSAT, A., VERMA, A., OLDENBOURG, R., SIXT, M. & GERLICH, D. W. 2017. Cytokinesis in vertebrate cells initiates by contraction of an equatorial actomyosin network composed of randomly oriented filaments. *eLife*, 6, e30867.
- STADING, R., GASTELUM, G., CHU, C., JIANG, W. & MOORTHY, B. 2021. Molecular mechanisms of pulmonary carcinogenesis by polycyclic aromatic hydrocarbons (PAHs): Implications for human lung cancer. *Semin Cancer Biol*, 76, 3-16.
- STEIGEMANN, P., WURZENBERGER, C., SCHMITZ, M. H. A., HELD, M., GUIZETTI, J., MAAR, S. & GERLICH, D. W. 2009. Aurora B-Mediated Abscission Checkpoint Protects against Tetraploidization. *Cell*, 136, 473-484.
- STEINLE, H., BEHRING, A., SCHLENSAK, C., WENDEL, H. P. & AVCI-ADALI, M. 2017. Concise Review: Application of In Vitro Transcribed Messenger RNA for Cellular Engineering and Reprogramming: Progress and Challenges. *STEM CELLS*, 35, 68-79.
- STOTHARD, P. 2000. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *BioTechniques*, 28, 1102-1104.
- STRÄSSER, K., MASUDA, S., MASON, P., PFANNSTIEL, J., OPPIZZI, M., RODRIGUEZ-NAVARRO, S., RONDÓN, A. G., AGUILERA, A., STRUHL, K., REED, R. & HURT, E. 2002. TREX is a conserved complex coupling transcription with messenger RNA export. *Nature*, 417, 304-8.
- SUGIURA, T., SAKURAI, K. & NAGANO, Y. 2007. Intracellular characterization of DDX39, a novel growth-associated RNA helicase. *Exp Cell Res*, 313, 782-90.
- SUNDARAM, G. M., QUAH, S. & SAMPATH, P. 2018. Cancer: the dark side of wound healing. *The FEBS Journal*, 285, 4516-4534.
- TANIGUCHI, I. & OHNO, M. 2008. ATP-dependent recruitment of export factor Aly/REF onto intronless mRNAs by RNA helicase UAP56. *Mol Cell Biol*, 28, 601-8.
- THOMPSON, S. L., BAKHOUM, S. F. & COMPTON, D. A. 2010. Mechanisms of chromosomal instability. *Curr Biol*, 20, R285-95.
- TRUONG, L. N. & WU, X. 2011. Prevention of DNA re-replication in eukaryotic cells. *Journal of molecular cell biology*, 3, 13-22.
- VALIEVA, Y., IVANOVA, E., FAYZULLIN, A., KURKOV, A. & IGRUNKOVA, A. 2022. Senescence-Associated β -Galactosidase Detection in Pathology. *Diagnostics (Basel)*, 12.
- VAN VUGT, M. A. T. M., GARDINO, A. K., LINDING, R., OSTHEIMER, G. J., REINHARDT, H. C., ONG, S.-E., TAN, C. S., MIAO, H., KEEZER, S. M., LI, J., PAWSON, T., LEWIS, T. A., CARR, S. A., SMERDON, S. J., BRUMMELKAMP, T. R. & YAFFE, M. B. 2010. A mitotic phosphorylation feedback network connects Cdk1, Plk1, 53BP1, and Chk2 to inactivate the G(2)/M DNA damage checkpoint. *PLoS biology*, 8, e1000287-e1000287.
- VARADI, M., ANYANGO, S., DESHPANDE, M., NAIR, S., NATASSIA, C., YORDANOVA, G., YUAN, D., STROE, O., WOOD, G., LAYDON, A., ŽÍDEK, A., GREEN, T., TUNYASUVUNAKOOL, K., PETERSEN, S., JUMPER, J., CLANCY, E., GREEN, R., VORA, A., LUTFI, M., FIGURNOV, M., COWIE, A., HOBBS, N., KOHLI, P., KLEYWEGT, G., BIRNEY, E., HASSABIS, D. & VELANKAR, S. 2022. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research*, 50, D439-D444.
- VARGAS-RONDÓN, N., VILLEGAS, V. E. & RONDÓN-LAGOS, M. 2017. The Role of Chromosomal Instability in Cancer and Therapeutic Responses. *Cancers*, 10, 4.
- VERMEULEN, K., VAN BOCKSTAELE, D. R. & BERNEMAN, Z. N. 2003. The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer. *Cell Prolif*, 36, 131-49.

- VILAR, E. & GRUBER, S. B. 2010. Microsatellite instability in colorectal cancer—the stable evidence. *Nature Reviews Clinical Oncology*, 7, 153-162.
- VIPHAKONE, N., CUMBERBATCH, M. G., LIVINGSTONE, M. J., HEATH, P. R., DICKMAN, M. J., CATTO, J. W. & WILSON, S. A. 2015. Luszp4 defines a new mRNA export pathway in cancer cells. *Nucleic acids research*, 43, 2353-2366.
- VITRE, B. D. & CLEVELAND, D. W. 2012. Centrosomes, chromosome instability (CIN) and aneuploidy. *Curr Opin Cell Biol*, 24, 809-15.
- WARE, M. D., DESILVA, D., SINILNIKOVA, O. M., STOPPA-LYONNET, D., TAVTIGIAN, S. V. & MAZOYER, S. 2006. Does nonsense-mediated mRNA decay explain the ovarian cancer cluster region of the BRCA2 gene? *Oncogene*, 25, 323-8.
- WEINBERG, R. A. 2014. Coming Full Circle—From Endless Complexity to Simplicity and Back Again. *Cell*, 157, 267-271.
- WEIR, H. K., ANDERSON, R. N., COLEMAN KING, S. M., SOMAN, A., THOMPSON, T. D., HONG, Y., MOLLER, B. & LEADBETTER, S. 2016. Heart Disease and Cancer Deaths - Trends and Projections in the United States, 1969-2020. *Prev Chronic Dis*, 13, E157.
- WICKRAMASINGHE, V. O. & LASKEY, R. A. 2015. Control of mammalian gene expression by selective mRNA export. *Nature Reviews Molecular Cell Biology*, 16, 431-442.
- WICKRAMASINGHE, V. O. & VENKITARAMAN, A. R. 2016. RNA Processing and Genome Stability: Cause and Consequence. *Mol Cell*, 61, 496-505.
- WU, P. 2020. Inhibition of RNA-binding proteins with small molecules. *Nature Reviews Chemistry*, 4, 441-458.
- YADAV, R. K., CHAE, S. W., KIM, H. R. & CHAE, H. J. 2014. Endoplasmic reticulum stress and cancer. *J Cancer Prev*, 19, 75-88.
- YAMAZAKI, T., FUJIWARA, N., YUKINAGA, H., EBISUYA, M., SHIKI, T., KURIHARA, T., KIOKA, N., KAMBE, T., NAGAO, M., NISHIDA, E. & MASUDA, S. 2010. The closely related RNA helicases, UAP56 and URH49, preferentially form distinct mRNA export machineries and coordinately regulate mitotic progression. *Mol Biol Cell*, 21, 2953-65.
- YEE, A. S., PAULSON, E. K., MCDEVITT, M. A., RIEGER-CHRIST, K., SUMMERHAYES, I., BERASI, S. P., KIM, J., HUANG, C. Y. & ZHANG, X. 2004. The HBP1 transcriptional repressor and the p38 MAP kinase: unlikely partners in G1 regulation and tumor suppression. *Gene*, 336, 1-13.
- ZHANG, N., KAUR, R., AKHTER, S. & LEGERSKI, R. J. 2009. Cdc5L interacts with ATR and is required for the S-phase cell-cycle checkpoint. *EMBO reports*, 10, 1029-1035.