UNIVERSITY OF BIRMINGHAM

Pathogenic mutations in components of the SMC5/6 complex cause segmented chromosomes and mosaic variegated hyperploidy

by

LAURA JANE GRANGE

A thesis submitted to the University of Birmingham for the degree of DOCTOR OF PHILOSOPHY

Institute of Cancer and Genomic Sciences

College of Medical and Dental Sciences

University of Birmingham

September 2022

Abstract

The structural maintenance of chromosomes (SMC) complexes are a highly conserved family of ring shaped ATPases which have many important functions in maintaining genome stability. Whilst cohesin (SMC1/3) and condensin (SMC2/4) have well characterised roles in cellular division, the precise functions of SMC5/6 are less well-defined, although the complex is believed to function primarily in DNA repair and replication. SMC5/6 comprises of several subunits, including SMC6, SMC5 and a number of non-SMC elements (NSEs), six of which have been identified in yeast (Nse1-6). Whilst only four NSE subunits have been identified in human cells (NSE1-4), SLF1 and SLF2, functional paralogues of Nse5 and Nse6, have been shown to promote the recruitment of SMC5/6 to damaged chromatin. The importance of SMC5/6 for human development is highlighted by that fact that mutations in SMC5/6 complex components are associated with two separate childhood syndromes. Individuals with NSE2 variants present with primordial dwarfism, extreme insulin resistance, and primary gonadal failure, whilst patients with mutations in NSE3 exhibit structural chromosome abnormalities and pulmonary disease. In this thesis I will explore the function of SMC5/6, and its impact on human disease, further, by characterising patients with mutations in two SMC5/6 associated genes: SLF2 and SMC5. Using patient derived lymphoblastoid and fibroblast cell lines, alongside U-2-OS SLF2 CRISPR mutants, I will show that SLF2 and SMC5 promote proper replication and repair and suggest novel functions for these factors in mitosis and promoting efficient replication through G4-quadruplex DNA secondary structures.

Acknowledgements

Firstly, huge thanks go to my supervisor, Dr John Reynolds, who guided me through this PhD with a seemingly unending reserve of patience and understanding. Thank you for your all your insight and hard work, I could never have achieved this without you.

Thanks also go to all my wonderful lab mates, Rob, Louise, Sophie, Anil, and Gavin, as well as my friends and colleagues in the IBR west, especially Lizzie, Karan, Amalia, Liza, and Clare. I'm so grateful for all your support over the last four years. In return for all the laughs, I wish I had something wittier to write. Instead, I just want to say how much I genuinely adore you all. I couldn't have asked for a kinder, funnier, or more helpful bunch of people to work alongside.

Massive thanks also go out to my long-suffering family, my mum, my dad, and my sister, who encouraged and supported me when I was at my lowest.

And lastly, but perhaps mostly, my eternal gratitude goes to Will. From beginning to end, you have been my rock. It's meant the world to me.

Contents

Abbreviations

53BP1	=	P53-binding protein
9-1-1	=	RAD9-RAD1-HUS1
ABC	=	ATP-binding-cassette
ALT	=	Alternative lengthening of telomeres
AP	=	Apurinic/apyrimidinic
APB	=	ALT-associated PML bodies
APC/C	=	Anaphase-promoting complex/cyclosome
APH	=	Aphidicolin
APIM	=	AlkB homologue 2 PCNA-interacting motif
ATM	=	ATM serine/threonine kinase
ATP	=	Adenosine 5′-triphosphate
ATR	=	Ataxia Telangiectasia and Rad3-related protein
ATRIP	=	ATR-interacting protein
ATRX	=	Alpha-thalassemia/mental retardation, X-linked
BIR	=	Break induced replication
BLM	=	BLM RecQ like helicase
BOD1L	=	Biorientation of chromosomes in cell division 1 like 1
bp	=	Base pair
BRCA1	=	Breast cancer gene 1
BRCA2	=	Breast cancer gene 2
BrdU	=	Bromodeoxyuridine
BTR	=	BLM-Top3α-RMI1-RMI2
BUBR1	=	Bub1-related kinase
BUB1	=	Benzimidazole protein 1
BUB1B	=	BUB1 mitotic checkpoint serine/threonine kinase B
BUB3	=	BUB3 mitotic checkpoint protein
CAP	=	Chromosome-associated polypeptide
Cas9	=	CRISPR-associated protein 9
CC	=	Coiled-coil
CDC6	=	Cell division cycle 6
CDC20	=	Cell division cycle 20
CDC25a	=	Cell division cycle 25a
CDC45	=	Cell division cycle 45
CDK	=	Cyclin-dependent kinase
cDNA	=	Complementary DNA
CDT1	=	Chromatin licensing and DNA replication factor 1
CENATAC	=	Centrosomal AT-AC splicing factor
CENPA	=	Centromere protein A
CENPF	=	Centromere protein F
CEP57	=	Centrosomal protein 57
CFS	=	Common fragile site
cGAS-STING	=	Cyclic GMP-AMP synthase-stimulator of interferon genes

CHK1	=	Checkpoint kinase 1
CIN	=	Chromosome instability
CldU	=	5-chloro-2′-deoxyuridine
CMG	=	CDC45-MCM-GINS
CPD	=	Cyclobutane pyrimidine dimer
CRISPR	=	Clustered Regularly Interspaced Short Palindromic Repeats
CtIP	=	CtBP-interacting protein
DAPI	=	4′,6-diamidino-2-phenylindole
DDK	=	Dbf4/Drf1-dependent kinase
DDX11	=	DEAD/H-Box helicase 11
dHJ	=	Double Holliday junction
D-loop	=	Displacement-loop
DNA	=	Deoxyribonucleic acid
DNA2	=	DNA replication helicase/nuclease 2
DNA-PKcs	=	DNA-dependent protein kinase catalytic subunit
dNTP	=	Deoxynucleoside triphosphate
DSB	=	Double strand break
DSBR	=	Double strand break repair
dsDNA	=	Double-stranded DNA
EDTA	=	Ethylenediaminetetraacetic acid
EdU	=	5-ethynyl-2'-deoxyuridine
EME1	=	Essential meiotic structure-specific endonuclease 1
ERCC1	=	Excision repair cross-complementing group 1
ESCO2	=	Establishment of cohesion 1 homolog 2
EU	=	5-ethynyluridine
EXO1	=	Exonuclease 1
FA	=	Fanconi anaemia
FACs	=	Fluorescence-activated cell sorting
FAN1	=	FANCD2 and FANCI associated nuclease 1
FANCD2	=	FA complementation group D2
FANCI	=	FA Complementation Group I
FANCL	=	FA Complementation Group L
FANCM	=	FA complementation group M
FBS	=	Fetal bovine serum
FEN1	=	Flap structure-specific endonuclease 1
FISH	=	Fluorescence in situ hybridization
Fob1	=	Fork blocking less
GAPDH	=	Glyceraldehyde-3-phosphate dehydrogenase
GEN1	=	GEN1 Holliday junction 5' flap endonuclease
GFP	=	Green fluorescent protein
GINS1	=	GINS complex subunit 1
GTP	=	Guanosine-5'-triphosphate
HEK	=	Human embryonic kidney

Hgt	=	Height
HJ	=	Holliday junction
HR	=	Homologous recombination
hTERT	=	Human telomerase reverse transcriptase
HTLF	=	Helicase-like transcription factor
ICL	=	Interstrand cross-link
ID2	=	FANCI-FANCD2
IdU	=	5-iodo-2′-deoxyuridine
IPTG	=	Isopropyl β- d-1-thiogalactopyranoside
LB	=	Lysogeny broth
LCL	=	Lymphoblastoid cell lines
LOH	=	Loss of heterozygosity
MAD2	=	Mitotic arrest deficient 2
MCC	=	Mitotic checkpoint complex
MCM	=	Minichromosome maintenance
MCM10	=	Minichromosome maintenance 10 replication initiation factor
MCPH1	=	Microcephalin
MDC1	=	Mediator of DNA damage checkpoint 1
Mec1	=	Mitosis entry checkpoint
MEF	=	Mouse embryonic fibroblast
mESCs	=	Mouse embryonic stem cells
MiDAS	=	Mitotic DNA synthesis
MIN	=	Microsatellite instability
MMC	=	Mitomycin-C
MMR	=	Mismatch repair
Mph1	=	Mps1p-like pombe homolog
MRE11	=	Meiotic recombination 11
MRN	=	MRE11-RAD50-NBS1
mRNA	=	Messenger RNA
MUS81	=	Methyl methane sulfonate ultraviolet sensitive gene clone 81
MVA	=	Mosaic Variegated Aneuploidy
MVH	=	Mosaic Variegated Hyperploidy
Myc	=	Myc proto-oncogene
NA	=	Not available
NBS1	=	Nijmegen breakage syndrome 1 mutated gene
NHEJ	=	Nonhomologous end joining
NIN	=	Nucleotide instability
NS		Not significant
NSE	=	Non-structural maintenance of chromosomes element
OFC	=	Orbitofrontal cortex
ORC	=	Origin recognition complex
PAM	=	Protospacer adjacent motif

PBS	=	Phosphate-buffered saline
PCNA	=	Proliferating Cell Nuclear Antigen
PCNT	=	Pericentrin
PCR	=	Polymerase chain reaction
PCS	=	Premature Chromatid Seperation
PDS5	=	Precocious dissociation of sisters 5
PIC	=	Preinitiation complex
PICH	=	PLK1-interacting checkpoint helicase
PLK1	=	Polo like kinase 1
PML	=	Promyelocytic leukaemia
PNA	=	Peptide nucleic acid
Pol	=	Polymerase
POT1	=	Protection Of Telomeres 1
pre-RC	=	Pre-replication complex
PrimPol	=	Primase And DNA Directed Polymerase
RAD21	=	Radiation sensitive 21
RAD51	=	Radiation sensitive 54
Rad53	=	Radiation sensitive 53
RAD54	=	Radiation sensitive 54
rDNA	=	Ribosomal DNA
RECQL4	=	RecQ Like Helicase 4
RFB	=	Replication fork barrier
RFC	=	Replication factor C
RMI1	=	RecQ mediated genome instability 1
RMI2	=	RecQ mediated genome instability 2
RNA	=	Ribonucleic acid
RNF168	=	Ring finger protein 168
RNF8	=	Ring finger protein 8
RPA	=	Replication protein A
RPE1	=	Human retinal pigment epithelial
rpm	=	Rotations per minute
RTEL1	=	Regulator of telomere elongation helicase 1
RT-PCR	=	Reverse transcription polymerase chain reaction
SAC	=	Spindle assembly checkpoint
SCE	=	Sister chromatid exchange
SD	=	Standard deviation
SDSA	=	Synthesis dependent strand annealing
SDS-PAGE	=	Sodium dodecyl-sulfate polyacrylamide gel electrophoresis
sgRNA	=	Single guide RNA
Sgs1	=	Slow growth suppressor 1
SLF1	=	SMC5-SMC6 complex localization factor 1
SLF2	=	SMC5-SMC6 complex localization factor 2
SLX1	=	SLX1 structure-specific endonuclease subunit
SLX4	=	SLX4 structure-specific endonuclease subunit

SMARCAL1	=	SWI/SNF related, matrix associated, actin dependent regulator Of chromatin, Subfamily A Like 1
SMC	=	Structural Maintenance of Chromosomes
SOC	=	Super Optimal broth with Catabolite repression
S-phase	=	Synthesis phase
SSA	=	Single strand annealing
SSB	=	Single strand break
SSC	=	saline-sodium citrate
ssDNA	=	Single-stranded DNA
STR	=	Sgs1-Top3-Rmi1
SUMO	=	Small ubiquitin-like modifier
TBST	=	Tris-buffered saline, 0.1% Tween 20
TDP1	=	Tyrosyl-DNA phosphodiesterase 1
TDP2	=	Tyrosyl-DNA phosphodiesterase 2
TERRA	=	Telomeric repeat-containing RNA
TIPIN	=	TIMELESS interacting protein
TLS	=	Translesion synthesis
Top1	=	Topoisomerase type 1
Top2	=	Topoisomerase type 2
Top2α	=	Topoisomerase II α
Top3α	=	Topoisomerase III α
TOPBP1	=	DNA topoisomerase II binding protein 1
TOPO	=	Topoisomerase
TRIP13	=	Thyroid hormone receptor interactor 13
t-SCE	=	Telomere sister chromatid exchange
U	=	Unit
U-2-OS	=	Human bone osteosarcoma epithelial cell
UBE2T	=	Ubiquitin Conjugating Enzyme E2 T
UBZ	=	ubiquitin-binding zinc finger
UFB	=	Ultrafine anaphase bridge
UV	=	Ultraviolet
v/v	=	Volume per volume
WABS	=	Warsaw breakage syndrome
WAPL	=	Wings apart-like protein homolog
Wgt	=	Weight
WRN	=	Werner syndrome ATP-dependent helicase
WT	=	Wildtype
w/v	=	Weight per volume
XPF	=	Xeroderma pigmentosum group F-complementing protein
ZRANB3	=	Zinc finger RANBP2-type containing 3
γH2AX	=	Histone H2A variant X phosphorylated on S139

List of figures

List of tables

Chapter 1:

Introduction

1.1. Genomic instability

In order to maintain genomic integrity, a cell must coordinate the faithful replication and transmission of two meters of DNA (Grosberg, 2012), whilst also protecting it from the constant attack of mutagenic genotoxins. The accuracy of these processes is essential, as errors can lead to genome alterations which, if left uncorrected, could result in cell dysfunction or death (Lindahl, 1993, Aguilera and Gomez-Gonzalez, 2008, Lopez-Bergami and Ronai, 2011). Genomic instability refers to an increased frequency of genome alterations and can be divided into three main subtypes which describe their DNA consequences. The first is nucleotide instability (NIN), resulting in the increased deletion, insertion or substitution of single bases- otherwise known as point mutations. These are typically caused by errors in replication or DNA repair (Pikor et al., 2013). Whilst these mutations generate only minor sequence changes which are often neutral, nonsynonymous (and even synonymous) mutations can have hugely detrimental impacts on gene expression and protein function. Consequently, point mutations are associated with >60% of cases of inherited, genetic human disease (Stenson et al., 2009). Another form of genomic instability is microsatellite instability (MIN), which occurs when repetitive elements of DNA are lengthened or shortened due to the formation of spontaneous replicative lesions which are then misrepaired. MIN is elevated in 15% of colorectal cancers and has been causally associated with tumorigenesis in many other cancer subtypes (Hampel et al., 2005, Vilar and Gruber, 2010, Li et al., 2020b). The third form of genomic instability is chromosome instability (CIN). This represents changes in both chromosome structure and number and includes the insertion, deletion, translocation and inversion of DNA sequences, as well as losses and gains of whole chromosomes. The causes of this

type of instability are incredibly diverse but principally include errors in replication and repair, as well as mitotic dysfunction (Taylor et al., 2019). Altered chromosome architecture is a feature observed in >75% of solid tumours and is frequently implicated in pathogenesis (Knouse et al., 2017, Taylor et al., 2018). Ultimately, genomic instability can cause deleterious consequences on a cellular and organism level and is directly associated, often causally, with human disease. As such, the cell must coordinate numerous pathways which monitor and manage this instability to ensure proper cellular function.

1.2. Replication

1.2.1. An overview of Eukaryotic Replication

DNA replication must occur accurately, efficiently and exactly once per cell cycle in order to ensure the faithful transmission of our genome to daughter cells. The initiation of this tightly controlled process begins at regions called replication origins, which become primed for replication, or 'licensed', during late mitosis and early G1. This occurs when the origin of recognition complex (ORC1-6) binds to the DNA replication origin and recruits CDC6 and CDT1 (Siddiqui et al., 2013, Parker et al., 2017). These factors cooperate to load two inactive MCM2-7 helicases, which form double hexamers, and complete the pre-replication complex (pre-RC) (Remus et al., 2009, Fernandez-Cid et al., 2013, Limas and Cook, 2019). During late G1 through to S-phase, after the formation of the pre-RC, licensing factors become downregulated, thereby preventing the re-licensing of origins later in the cell cycle. For example, CDT1 can be inactivated either by proteolysis or upon interaction with Geminin, a small protein which inhibits the interaction between CDT1 and MCM2-7 (Hodgson et

al., 2002, Arias and Walter, 2007, Klotz-Noack et al., 2012). Furthermore, whilst low cyclin-dependent kinase (CDK) activity in late M phase and early G1 stabilises licensing factors, high CDK activity promotes progression into S-phase, replication initiation and the degradation of ORC1-6 and CDC6 (Vas et al., 2001, Mimura et al., 2004, Arias and Walter, 2007, Siddiqui et al., 2013). At the onset of S-phase, CDK, alongside Dbf4/Drf1-dependent kinase (DDK) and the recruitment several other factors, including CDC45, GINS, TOPBP1, Treslin and RECQL4, facilitates the conversion of the pre-RC to the preinitiation complex (PIC) (Im et al., 2009, Aparicio et al., 2009, Kumagai et al., 2010, Ilves et al., 2010, Kumagai and Dunphy, 2017), with CDC45, GINS and MCM2-7 forming the active CMG helicase (CDC45-MCM-GINS) (Aparicio et al., 2009, Ilves et al., 2010, Douglas et al., 2018). Finally, the recruitment of numerous other replication factors, including RFC, PCNA, RPA, TIPIN-TIMELESS, alongside the replicative DNA polymerases, allow for the assembly of two functional replisomes that unwind the DNA duplex bi-directionally and initiate replication, a process termed origin firing (O'Donnell and Li, 2016, Yeeles et al., 2017, Douglas et al., 2018, Jones et al., 2021, Tanaka and Ogawa, 2022, Baris et al., 2022)..

To ensure that replication is completed rapidly, multiple origins are licensed and initiate replication during each cell cycle. Super-resolution fluorescent microscopy has revealed that most DNA replication occurs in ‘factories’- discrete nuclear sites where replication likely originates from a cluster of origins (Hozak and Cook, 1994, Cseresnyes et al., 2009, Gillespie and Blow, 2010). Without multiple origins being fired simultaneously from different replication factories, it is estimated that it would take at least 20 days for a human cell to complete S-phase (Mechali, 2010). However, of the 50,000 replication origins across the human genome that assemble the pre-

RC, only a small minority initiate replication (Woodward et al., 2006, Ibarra et al., 2008, Mechali, 2010). In fact, very few replication origins are thought of as constitutive, in that they fire every cell cycle, with most origins considered flexible, firing stochastically in different cells, or dormant, firing rarely. In addition to this, origins can also be categorized in terms of when they initiate replication, either defined as early, mid or late firing (Mechali, 2010, Kaykov and Nurse, 2015, Valenzuela, 2012, Boos and Ferreira, 2019). The temporal and spatial separation of origin firing allows the cell to initiate replication flexibly. In response to different physiological conditions, such as replication stress, this permits additional, dormant origins across the genome to be fired later into S-phase, even when licensing factors have been downregulated (Woodward et al., 2006, Ibarra et al., 2008, Valenzuela, 2012). As will be discussed, this helps replication to be completed properly.

1.2.2. Replication Initiation and Elongation

The replisome comprises the CMG helicase and polymerases, alongside numerous accessory proteins (O'Donnell and Li, 2016, Jones et al., 2021). Once the helicase unwinds the DNA double helix at replication origins, RPA coats the ssDNA, preventing it from reannealing or signaling DNA repair pathways (Burgers and Kunkel, 2017). In eukaryotes, at least three DNA polymerases (α, δ and ε) are then required for the synthesis of new DNA in unperturbed conditions (Figure 1.1) These enzymes move 3’ to 5’ across the DNA, replicating the parental DNA by catalyzing the addition of nucleotides at the 3’ end of polynucleotide daughter chains. However, polymerases are unable to initiate polynucleotide synthesis and so require a primer from which to begin replication (Lujan et al., 2016, Guilliam and Yeeles, 2020). Primer synthesis is accomplished by the polymerase α-primase complex, comprised of a polymerase

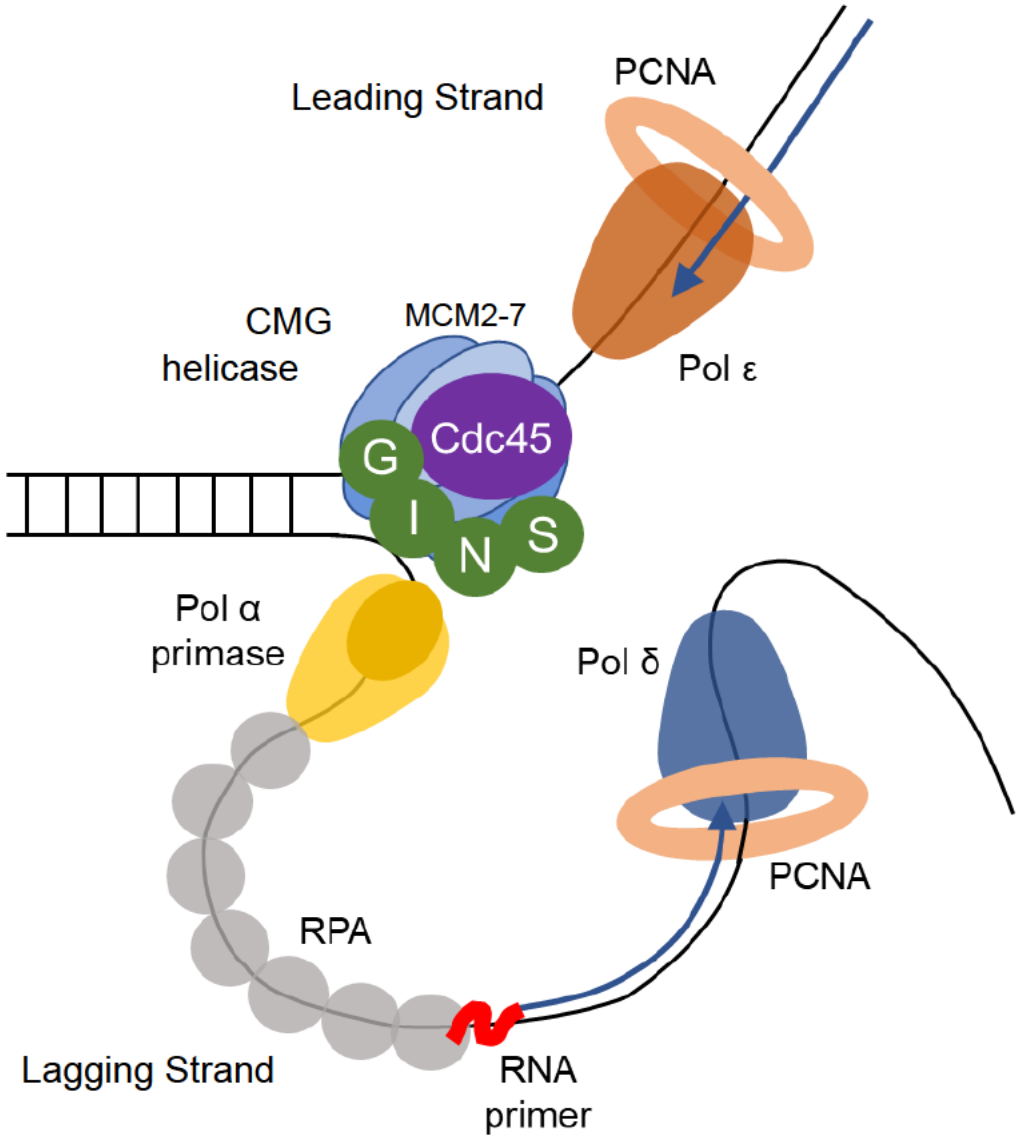


Figure 1.1. Schematic of the DNA replication fork.

Active CMG helicases (CDC45-MCM-GINS) promote the unwinding of duplex DNA and translocate across the DNA biodirectionally, with replication initiating via DNA polymerase α-primase priming. The CMG helicase associates with the leading strand, which is elongated by polymerase ε. In contrast, on the lagging strand the polymerase α-primase generates short RNA primers which are synthesised discontinuously by polymerase δ. The PCNA sliding clamp increases the processivity of the polymerases.

subunit (p180) and a regulatory factor (B-subunit), alongside two primase subunits (p49 and p58) which synthesize short 10-12 nucleotide RNA primers on both strands of the denatured duplex (Copeland and Wang, 1993, Muzi-Falconi et al., 2003, Nunez-Ramirez et al., 2011). Due to the limited processivity of polymerase α, the primers are extended by only 20-40 nucleotides by the polymerase subunits of the enzyme (Johansson and Macneill, 2010, Perera et al., 2013, Henninger and Pursell, 2014, Aria and Yeeles, 2018). The RFC-PCNA complex then loads PCNA onto the parental and daughter DNA strands before RFC is ejected, allowing for the recruitment of DNA polymerases which elongate the primers (Yao and O'Donnell, 2012, Schrecker et al., 2022). PCNA acts as a sliding clamp, increasing polymerase processivity (Prelich et al., 1987, Mondol et al., 2019).

Once DNA replication has been initiated, the replisomes continue across the genome bidirectionally, driven by the unwinding of the DNA double helix. This movement forms structures called replication forks which represent the active region of DNA synthesis, comprised of the replication machinery and the intersection between dsDNA and denatured DNA (Figure 1.1) (Leman and Noguchi, 2013, Burgers and Kunkel, 2017). Due to the antiparallel nature of the DNA duplex, the synthesis of daughter strands at the replication fork is asymmetric (Lujan et al., 2016). The leading strand is replicated parallel to the movement of the replication fork by polymerase ε, which associates directly with the CMG helicase (Georgescu et al., 2014, Daigaku et al., 2015). However, the lagging strand is synthesized in the opposite direction, and as such is replicated discontinuously. To achieve this, multiple DNA sequences of around 100-200 nucleotides, called Okazaki fragments, are primed by polymerase α-primase, elongated by polymerase δ and then ligated together (Ogawa and Okazaki, 1980,

Zheng and Shen, 2011, Georgescu et al., 2014, Daigaku et al., 2015). However, before a continuous DNA strand can be generated, the RNA-DNA primers must be removed and replaced with deoxyribonucleotides in a process which is dependent on the interaction of polymerase δ with several endonucleases. Once each Okazaki fragment has been sufficiently extended, polymerase δ encounters the 5' terminus of the previously synthesized fragment and renders it single-stranded via displacement synthesis. The displaced RNA-DNA flap structure is then removed by either DNA2 or FEN1, or additional exonucleolytic digestion pathways, leaving DNA nicks which are filled by the polymerase and sealed by DNA ligase (Ayyagari et al., 2003, Garg et al., 2004, Liu et al., 2017, Sun et al., 2022).

1.2.3. Topoisomerases and the Resolution of Topological Stress

As DNA is unwound, the parental duplex ahead of the replisome can overwind, resulting in positive supercoiling. If left to persist, these topological structures can impede the movement of the replisome, prevent the denaturing of the DNA helix and result in under-replicated regions of the genome (Deweese and Osheroff, 2009, Keszthelyi et al., 2016). To counteract supercoiling, the fork can rotate clockwise relative to the progression of replication. However, this generates pre-catenanes, in which the newly replicated DNA helices behind the replisome become intertwined (Postow et al., 2001, Cebrian et al., 2015). Once replication has completed, pre-catenanes are converted into catenanes as the chromatids interlink together (Dewar and Walter, 2017). If left unmanaged, these links can prevent the proper segregation of sister chromatids and lead to DNA damage (DiNardo et al., 1984, Bermejo et al., 2008).

DNA topoisomerases are essential enzymes which resolve torsional stress by generating transient DNA breaks. In eukaryotes, topoisomerases are divided into two major classes: type 1 and type 2 (Top1 and Top2, respectively) (Bush et al., 2015, Keszthelyi et al., 2016, Pommier et al., 2022). Both types clamp around DNA and catalyze the transient breakage of phosphodiester bonds via nucleophilic attack (Redinbo et al., 2000, Champoux, 2001, McKie et al., 2020). However, they each use this mechanism to alter DNA topology in distinct ways. In humans, Top1 generates breaks on a single strand of the DNA duplex and becomes covalently bonded to the 3' end of the cleaved site. This single stranded nick can relieve supercoiling by allowing the 5' end of the break to freely rotate (Leppard and Champoux, 2005). In contrast, Top2 can resolve catenation as well as supercoiling by catalyzing transient double strand breaks (DSBs) through which an intact duplex can pass, altering the writhe of the DNA (Lucas et al., 2001, McClendon et al., 2005, Bush et al., 2015). Whilst Top1 doesn't require an energy source, Top2 must hydrolyse ATP and bind metal ion cofactors in order to catalyze the cleavage reaction before covalently bonding to the 5' end of the cleaved site (Baird et al., 1999, Bates and Maxwell, 2007, Sissi and Palumbo, 2009). As replication terminates, the action of Top2 becomes increasingly important (Fachinetti et al., 2010). Topoisomerase enzymes become sterically excluded from converging forks, meaning that obstructive positive supercoiling must be managed by compensatory rotation, generating cantenanes dependent on Top2 activity (Baxter and Diffley, 2008, Keszthelyi et al., 2016, Dewar and Walter, 2017, Heintzman et al., 2019).

After breaks are generated, the covalent bond between the topoisomerase and the DNA protects the DNA from degradation and stores energy to later facilitate repair of

the cleaved site. After the alleviation of torsional stress, Top1 and Top2 catalyse the religation of the DNA nick and DNA DSB, respectively, and the enzymes are released from the DNA (Deweese and Osheroff, 2009, Nitiss, 2009, Bush et al., 2015). However, under certain circumstances, the topoisomerase reaction mechanism can become ‘abortive’- a state in which the topoisomerase becomes covalently linked to the DSB, meaning religation of the break is not possible. In this case, the restoration of the DNA backbone is initiated by the proteasomal degradation of the topoisomerase (Mao et al., 2001, Sun et al., 2020), which leaves behind short peptide fragments. These are then directly removed by phosphodiesterases which hydrolyse the bond between the degraded topoisomerase and the DNA (Pommier et al., 2014). Tyrosyl-DNA phosphodiesterase 1 (TDP1) acts on Top1 (Yang et al., 1996), whilst Tyrosyl-DNA phosphodiesterase 2 (TDP2) acts on Top2 (Cortes Ledesma et al., 2009), generating breaks free for repair via DNA single strand break repair or NHEJ, respectively (Strande et al., 2012, Gomez-Herreros et al., 2013).

1.2.4. Replication Termination

Replication termination is an incredibly complex yet relatively poorly understood process which occurs throughout S-phase when neighboring replication forks meet (Dewar and Walter, 2017). The forks approach one another from opposite directions, encountering high levels of torsional stress. Supercoiled structures that build up between the forks may impede the progress of replication, so must first be resolved by compensatory rotation or the activity of topoisomerases (Baxter and Diffley, 2008, Keszthelyi et al., 2016, Dewar and Walter, 2017, Heintzman et al., 2019). Once the DNA has been relaxed, the approaching forks converge and the replisomes bypass each other, encountering downstream Okazaki fragments and processing them.

Studies on *Xenopus* egg extracts have demonstrated that this process is highly efficient, with the rate of synthesis not significantly dropping throughout termination (Dewar et al., 2015). This is facilitated by the replication fork architecture, in which the CMG helicase is tightly associated with polymerase ε on the leading strand (Langston et al., 2014, Gambus, 2017). Whilst studies in yeast have indicated that polymerase α-primase is weakly associated with the helicase (Bai et al., 2017), there is little evidence that any of the other components of the lagging strand replication machinery interact. During convergence, this means that the bulky CMG-polymerase ε-associated replisomes progress on opposing leading strands and can bypass each other without fork collision or stalling (Dewar and Walter, 2017, Gambus, 2017).

Once the replication machinery encounters the 5' terminus of the downstream lagging strand, the final Okazaki primer must be displaced, filled and ligated. However, research has not yet determined how the primer is processed, and which polymerase is responsible (Gambus, 2017). On its own, polymerase ε does not interact with either of the Okazaki fragment maturation endonucleases, DNA2 and FEN1 (Bae and Seo, 2000, Liu et al., 2017, Devbhandari et al., 2017). This would suggest that either polymerase δ or the endonucleases themselves are recruited by leading strand replisome components in the context of replication termination, or that polymerase ε displaces the RNA-DNA primer via an additional, unknown mechanism (Gambus, 2017).

Once replication has been completed, the replisome must dissociate from the DNA. To prevent premature dissociation and replication stalling, this process is tightly regulated and is believed to only occur when the CMG helicase loads onto dsDNA (Dewar et al., 2015, Dewar and Walter, 2017). In mammalian cells, the initiation of

replisome disassembly has been shown to be the polyubiquitylation of the MCM7 subunit of the CMG helicase by a cullin 2 E3-ligase called $CUL2^{LRR1}$ (Moreno et al., 2014, Sonneville et al., 2017, Dewar and Walter, 2017, Villa et al., 2021). These modifications are then recognised by the ubiquitylation dependent segregase p97, which uses ATPase activity to remove the CMG helicase and its associated components from the chromatin (Sonneville et al., 2017, Villa et al., 2021).

1.2.5. Re-replication

Cells that cannot downregulate licensing factors during late G1 risk inducing re-replication. For instance, in mammalian cells the deregulation of CDT1 or the loss of the CDT1 inhibitory factor Geminin enables the pre-RC to assemble again at origins during S-phase and re-initiate replication (Kerns et al., 2007, Truong and Wu, 2011, Klotz-Noack et al., 2012, Zhou et al., 2020). As an unscheduled event, replication forks that result from re-replication are slow moving and cannot replicate entire chromosomes (Nguyen et al., 2001). However, even limited re-replication can produce gene amplifications, whilst the re-replication of centromeres risks chromosome breakage and missegregation, which could result in aneuploid daughter cells (Arias and Walter, 2007, Green et al., 2010, Neelsen et al., 2013, Hanlon and Li, 2015). Additionally, extensive re-replication generates increased levels of DSBs via ‘head-to-tail’ fork collisions between original and re-replicated forks, as well as persistently stalled forks (Davidson et al., 2006, Truong and Wu, 2011, Alexander et al., 2015).

Studies have suggested that ATR (a checkpoint kinase that will be discussed in more detail in the next section) signalling acts to sense re-replication and protect the cell

from further dysfunction (Vaziri et al., 2003, Davidson et al., 2006, Liu et al., 2007). Accordingly, re-replication is readily induced by licensing defects in tumour cells lines, where replication checkpoints are defective (Truong and Wu, 2011). Re-licensed and fired origins produce aberrant and unscheduled forks which readily accumulate ssDNA through replication uncoupling, when the progression of the polymerase is blocked but the helicase continues to unwind the DNA duplex (Liu et al., 2007). This, alongside signals from stalled forks, activates ATR and triggers checkpoint responses (Davidson et al., 2006, Liu et al., 2007, Truong and Wu, 2011). However, ATR also phosphorylates downstream replication factors, including RPA and MCM2, to supress further re-replication (Vaziri et al., 2003, Liu et al., 2007, Truong and Wu, 2011).

1.2.6. Replication Stress

Obstructions to, or errors in, the replication process can lead to replication stress; a principal source of genomic instability. Replication stress is loosely defined as the slowing or stalling of the replication machinery and has numerous causes (Zeman and Cimprich, 2014, Gaillard et al., 2015, Techer et al., 2017), as will be discussed later. If replication stress persists, DNA damage can accumulate as stalled replication forks collapse and form DSBs. Improper replication may also result in under-replicated regions or intertwined DNA. As the cell cycle progresses from S-phase into G2 and mitosis, these replicative errors can lead to nuclease processing of aberrant forks or incomplete chromatid segregation. With additional rounds of divisions these errors may amplify and further increase mitotic errors and chromosome breakage (Zeman and Cimprich, 2014, Gelot et al., 2015). To protect against this, stalled replication forks trigger the replication stress response and S-phase checkpoints, which centre around the activation of the ATR kinase (Flynn and Zou, 2011).

The generation of ssDNA is critical to the initiation of the replication stress response and can occur after replication uncoupling (Byun et al., 2005). When the entire replisome pauses, enzymatic remodelling by helicases and nucleases also generates ssDNA which could potentially activate ATR, although many fork remodelling factors, including HTLF and SMARCAL1, have yet to be associated with ATR activation (Blastyak et al., 2010, Nam and Cortez, 2011, Betous et al., 2012). ssDNA is rapidly bound by RPA, which prevents the degradation of the DNA and formation of secondary structures, whilst also acting as a scaffold for the assembly of additional factors (Bhat and Cortez, 2018). This includes ATRIP, which is recruited to the fork via a direct interaction with RPA and brings with it its binding partner, ATR (Zou and Elledge, 2003, Namiki and Zou, 2006). However, ssDNA is not sufficient for ATR activation (Byun et al., 2005), with studies suggesting that the loading of polymerase α-primase and the synthesis of an RNA primer on the lagging strand is also necessary (Michael et al., 2000, MacDougall et al., 2007). The RAD17-RFC complex recognises the primer-template junction and is loaded onto the chromatin, promoting the recruitment of the RAD9-HUS1-RAD1 (9-1-1) complex (Zou and Elledge, 2003, Bermudez et al., 2003). TOPBP1 interacts with Rad9 and is localised to the fork by 9-1-1. This is required for TOPBP1 to interact with ATRIP-ATR and stimulate the full activity of the kinase (Kumagai et al., 2006, Delacroix et al., 2007, Liu et al., 2011b). The indirect interaction of ATR and RPA, alongside the basal kinase activity of the enzyme, allows ATR to autophosphorylate on T1989 (Liu et al., 2011b, Nam et al., 2011). The importance of this event to the full activation of ATR is still debated (Blackford and Jackson, 2017), with studies suggesting that it is critical to enable TOPBP1 to stably engage with ATR (Liu et al., 2011b). ATR stimulation is achieved

via TOPBP1's Activation Domain (AAD), although the mechanism behind this is not fully understood (Zhou et al., 2013). It is hypothesised that TOPBP1 may alter the conformation of the ATR kinase to allow it easier access to substrates (Mordes et al., 2008).

In the case of acute replication stress, 5 to 10-fold decreased replication fork progression leads to the activation of ATR-induced checkpoint responses (Koundrioukoff et al., 2013). However, in certain cases of mild to moderate replication stress, despite the loading of ATR pathway sensors and mediators onto DNA, the phosphorylation of important effectors can remain below a certain signal threshold, essentially being undetectable. This leads to a failure to stimulate ATR-induced checkpoint activities (Wang et al., 2011b, Koundrioukoff et al., 2013, Gelot et al., 2015).

1.2.7. ATR Activation and Checkpoint Activity

After activation by stalled replication forks, ATR phosphorylates downstream targets to initiate mechanistically distinct checkpoints. The S-phase replication checkpoint blocks origin firing and prevents S-phase progression (Chastain et al., 2006, Seiler et al., 2007, Karnani and Dutta, 2011, Saldivar et al., 2017), whilst the G2/M checkpoint prevents premature mitotic onset (Saldivar et al., 2017). Although numerous ATR substrates are involved in these regulatory processes, the mediator kinase CHK1 is known to be critical (Zhang and Hunter, 2014, Blackford and Jackson, 2017). ATR induced phosphorylation activates CHK1 and allows it to dissociate from the chromatin where it can phosphorylate the CDC25a phosphatase, targeting it for proteasomal degradation. As a result, CDC25a is prevented from dephosphorylating

CDK/Cyclin complexes, thereby locking them in an inactive form (Xiao et al., 2003, Patil et al., 2013). Although other mechanisms may be facilitatory, inhibition of the CDK1/Cyclin B mitotic initiation complex via CDC25a degradation is implicated in the activation of G2/M checkpoint during replication stress. This ensures that replication is faithfully completed before mitosis can be initiated (Zeng et al., 1998, Mailand et al., 2002, Shen and Huang, 2012).

During origin firing, the association of CDC45 with the pre-RC requires the phosphorylation of the replisome component Treslin by the CDK2/Cyclin E complex. As such, the inhibition of CDK2/Cyclin E via CDC25a degradation, as well as the inhibitory phosphorylation of Treslin by CHK1, prevents CDC45 from loading onto chromatin (Guo et al., 2015, Kelly et al., 2022). This blocks the initiation of replication in new replication factories and prevents additional replication stress from being generated by late-firing regions of the genome which haven't undergone replication yet. However, depending on the severity of the replication stress, local origins near the stalled fork can continue to be active, thus allowing dormant 'back-up' origins to initiate and replicate nearby regions where synthesis has been impeded (Ge and Blow, 2010, Yekezare et al., 2013). The mechanism behind this spatial organisation of origin firing is not currently understood.

ATR signalling and checkpoint activity has also been suggested to function in the stabilisation of stalled replication forks. The inhibition of origin firing through ATR-CHK1 signalling contributes to this by limiting the generation of further ssDNA, thus preventing the depletion of nuclear pools of RPA which would render forks unprotected and cause them to collapse (Toledo et al., 2013). In addition, ATR signalling has historically been considered to promote replisome stability. Where

dormant origins are unable to initiate and compensate for stalled forks, maintaining the association of replisome components becomes crucial for the stalled fork to restart. However, evidence for ATR's role in this regulation remains indirect and contradictory. For instance, immunoprecipitation studies using *Saccharomyces cerevisiae* have shown that $Mec1^{ATR}$ and $Rad53^{CHK1}$ deficient cells treated with replication stress inducing drugs undergo dissociation of the helicase and polymerases (Lucca et al., 2004, Cobb et al., 2005). However, more recent investigations have observed that these components actually remain stably associated at the majority of replication forks and have reasoned that previous contradictions to these observations may be explained by checkpoint activities at the very earliest firing origins (De Piccoli et al., 2012, Liu et al., 2021). These more recent findings have proposed that, for most origins, ATR induced phosphorylation of replisome components does not affect fork stability (Dungrawala et al., 2015), but more likely influences replisome function (De Piccoli et al., 2012), possibly to prevent unwanted replication fork resection or unwinding (Saldivar et al., 2017, Liu et al., 2021).

1.2.8. Mechanisms of Replication Restart

As previously mentioned, a small minority of licenced origins fire during S-phase, with most serving as 'back-up' origins which can initiate at later timepoints, during replication stress, to replicate through stalled regions (Woodward et al., 2006, Ibarra et al., 2008, Mechali, 2010). The probability of dormant origin firing increases with time as fork slowing or stalling reduces the rate of passive replication through dormant origins, allowing them time to initiate (Courtot et al., 2018). In agreement with this, computer modelling has shown that passive dormant firing produces similar levels of

origin activation as seen *in vivo* (Blow and Ge, 2009). However, active pathways to dormant origin activation also exist. In response to mild replication stress, FANCI, a component of the Fanconi anaemia (FA) pathway, activates adjacent, dormant origins, probably as a regulator of DDK activity. Conversely, during acute replication stress, ATR becomes activated and phosphorylates FANCI. This inhibits FANCI, blocking nearby origin firing and allowing collapsed forks time to repair or activate DDR pathways (Chen et al., 2015, Chen et al., 2019).

To restart replication, lesions responsible for replication fork stalling can be bypassed by specialised polymerases in a process called translesion synthesis (TLS). After the replisome pauses, the replicative polymerase dissociates and is substituted by one of the Y-family TLS polymerases which incorporates a nucleotide opposite the lesion (Waters et al., 2009, Guilliam and Doherty, 2017). These polymerases have larger active sites capable of accommodating bulky adducts, but usually specialise in bypassing a specific type of obstruction. For instance, polymerase η is particularly efficient at inserting nucleotides opposite cyclobutane pyrimidine dimers (Yoon et al., 2009, Inomata et al., 2021). A further polymerase switching event causes the Y-family polymerase to be displaced by polymerase ζ, which efficiently extends mismatched primers *in vitro* (Johnson et al., 2001, Haracska et al., 2003). In yeast, polymerase ζ can extend 200bp beyond the lesion until the replicative polymerase reengages, although the same activity has yet to be shown in human cells (Waters et al., 2009, Kochenova et al., 2015). Whilst studies on TLS have historically focused on Y-family polymerases, more recent discoveries have identified a further eukaryotic primase, PrimPol. Like Y-family polymerases, the polymerase activity of PrimPol means it may be able to enact TLS to traverse lesions. However, its more significant function

involves its additional primase activities, which allows it to skip past bulky obstructions, re-anneal dNTP primers to downstream regions and continue with replication, leaving ssDNA gaps which can be repaired post-replicatively (Guilliam et al., 2015, Kobayashi et al., 2016, Guilliam and Doherty, 2017, Piberger et al., 2020). Whilst the functions of TLS polymerases and PrimPol prevent fork collapse, they are considered as 'error prone'. Their lack of proof-reading domains mean they also lack fidelity and, as such, incorporate an incorrect nucleotide on average every 10 to 10,000 replicated bases, as opposed to every 10^6 to 10^8 for replicative polymerases (Waters et al., 2009, Guilliam et al., 2015, Guilliam and Doherty, 2017).

To facilitate replication restart, stalled forks may also be remodelled. This is achieved by fork reversal, in which the fork moves backwards, opposite to the direction of DNA synthesis. The nascent strands anneal together whilst the parental strands reanneal, generating a four-way junction known a 'chickenfoot' structure (Figure 1.2). This reduces the generation of ssDNA, thus stabilising the fork, and places the impeding lesion back into dsDNA where DDR pathways can access it. As such, fork reversal is believed to be important to promote genome instability and prevent chromosome breakage (Petermann and Helleday, 2010, Ray Chaudhuri et al., 2012, Neelsen and Lopes, 2015, Meng and Zhao, 2017, Mutreja et al., 2018). Conversely, however, in certain contexts this reversal can act as an inducer of genomic instability, particularly in highly repetitive regions where the structure can promote misalignment (McMurray, 2010). Additionally, when stabilising factors such as BRCA2 and BOD1L become depleted reversed forks become degraded, leading to chromosomal breakage (Higgs et al., 2015, Higgs and Stewart, 2016, Mijic et al., 2017, Kolinjivadi et al., 2017, Lemacon et al., 2017).

There are thought to be a number of different molecular outcomes to fork reversal which mediate replication restart, including lesion removal, template switching and the formation of Holliday junctions (HJs) (Figure 1.2) (Petermann and Helleday, 2010, Neelsen and Lopes, 2015, Meng and Zhao, 2017). However, it is only relatively recently that fork reversal has been shown to be a frequent physiological response to numerous types of replication stress (Zellweger et al., 2015). As such, mechanisms for replication restart via fork regression and the factors necessary for these processes are poorly understood, although many proteins have been implicated. Whilst yet to be confirmed *in vivo, in vitro* fork reversal is promoted by RecQ helicases, including WRN and BLM, which unwind the stalled fork (Machwe et al., 2006, Ralf et al., 2006), as well as recombinases RAD51 and RAD54, which catalyze regression through branch migration activities (Bugreev et al., 2011, Zellweger et al., 2015). Studies have also suggested that various translocases, including SMARCAL1, ZRANB3 and HLTF (Poole and Cortez, 2017), recognise different fork structures and promote fork regression and restart in different contexts, whilst a number of factors, like BRCA1 and BRCA2, prevent inappropriate resection of the 'chicken foot' structure (Schlacher et al., 2011, Quinet et al., 2017, Mijic et al., 2017). Whilst electron microscopy analysis has demonstrated no link between the phosphorylation of CHK1 and fork reversal (Zellweger et al., 2015), ATR has been shown to phosphorylate the SMARCAL1 translocase, regulating levels of the protein, limiting its fork regression activities and preventing the aberrant degradation of the stalled fork (Couch et al., 2013). Consistent with this, ATR inhibition is associated with reduced levels of DNA damage-induced fork reversal (Mutreja et al., 2018). Thus, the role the replication

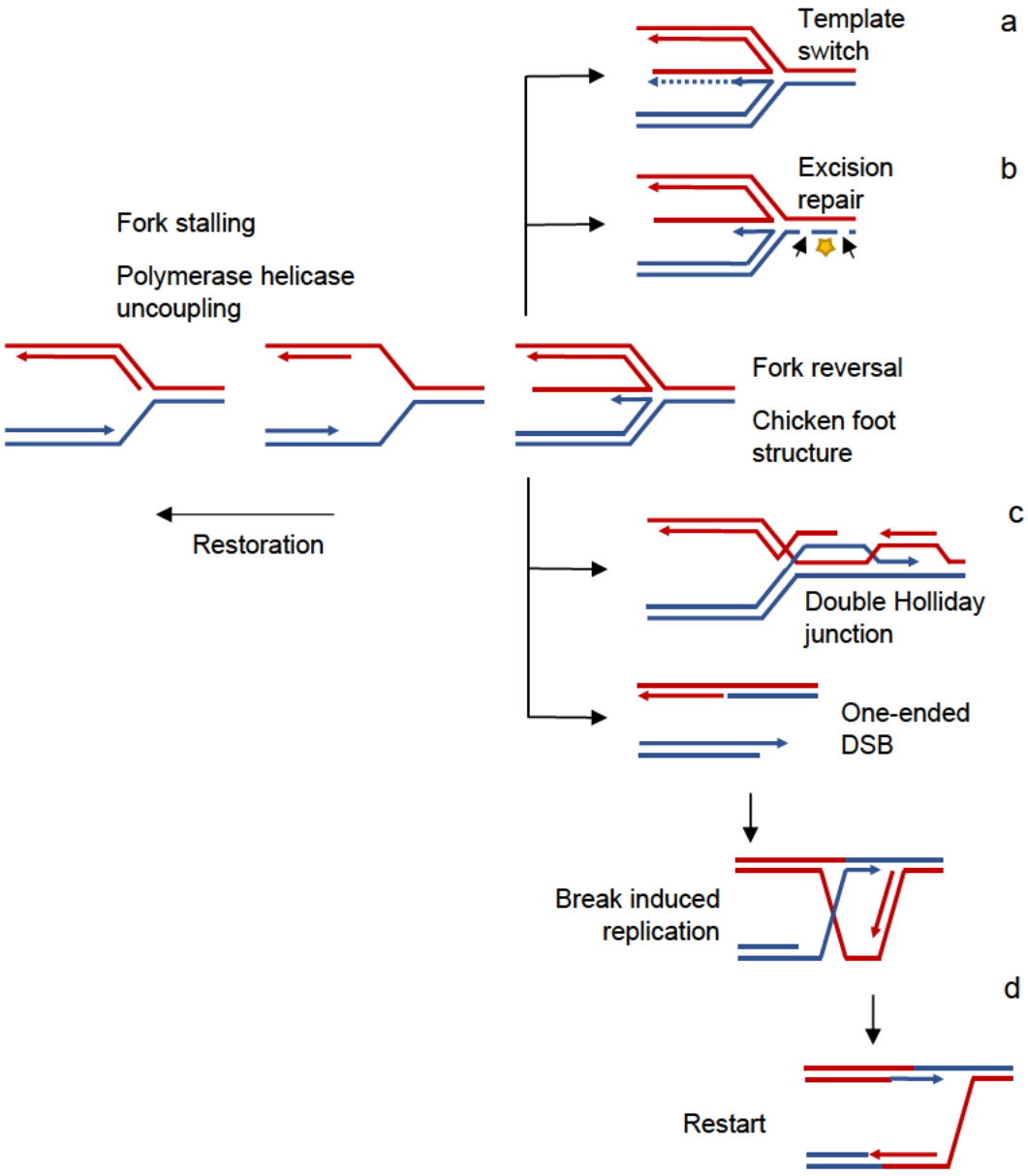


Figure 1.2. Mechanisms of replication fork restart following fork reversal.

Persistent stalled replication forks risk fork collapse and the generation of DNA damage. To avoid this, forks can be restarted via fork reversal, generating a 'chicken foot' structure. (a) If the nascent 3' end of the break is recessed comparatively to the 5' end, then synthesis can reprime, using this as a template, causing a template-switch and replication restart following branch migration. (b) Alternatively, fork reversal may protect the stalled fork from degradation whilst DNA repair pathways, such as base excision repair, remove the impeding lesion. (c) Reversed forks may also form Holliday junctions, enabling restart via HR, (d) whilst single ended-DSBs induced by nucleolytic attack can allow for break induced replication (BIR). Figure adapted from (Petermann and Helleday, 2010).

checkpoint and ATR activation on the regulation of fork reversal is, as of yet, poorly understood (Berti et al., 2020).

1.2.9. Collapsed Replication forks

Replication fork collapse can describe several different processes in which the replication fork loses the ability to perform DNA synthesis (Zeman and Cimprich, 2014). This can occur if replication fork stalling is prolonged, with some studies suggesting that this can lead to the dissociation or displacement of the replicative machinery (Lucca et al., 2003, Cobb et al., 2005, Cortez, 2015). Alternatively, persistently stalled or reversed forks may generate DSBs through the action of endonucleases, including the MUS81-EME1 complex (Regairaz et al., 2011, Fugger et al., 2013, Amangyeld et al., 2014). An unrepaired single stranded break may also be directly converted into a single-ended DSB when the nick is traversed by the replication fork (Cannan and Pederson, 2016). During S-phase, free ends of this type trigger repair by homologous recombination (HR), as will be discussed.

Collapsed forks are very deleterious. Not only do they risk the incomplete synthesis of the genome, but DSB formation can lead to chromosomal aberrations and, if left unrepaired, can result in cellular senescence or cell death (Jeggo and Lobrich, 2007). Accordingly, cell lines with replication checkpoint deficiencies exhibit increased levels of DSB formation and chromosome breakage (Cha and Kleckner, 2002, Trenz et al., 2006). However, controlled and timely DSB formation and subsequent HR has also been shown to be important for replication restart in certain contexts (Petermann and Helleday, 2010, Rickman and Smogorzewska, 2019). Evidence for this comes from observations that mammalian cells display a fork restart defect after treatment with

replication stress inducing drugs (Hanada et al., 2007, Regairaz et al., 2011, Pepe and West, 2014). However, MUS81-mediated cleavage has also been associated with an amplification of the replication checkpoint signal and the elimination of cells experiencing significantly elevated levels of replication stress (Fugger et al., 2013). Therefore, it seems likely that fork reversal and DSBs arising from fork collapse represent a trade-off between the promotion of replication restart and the increased risk of genomic instability.

1.3. Causes of Replication Stress

Replication stress can arise as a result of several pathological and physiological processes. This includes the misincorporation of ribonucleotides, oncogene activation and the depletion of dNTPs or replication factors (Zeman and Cimprich, 2014). Here I will discuss in detail a few further, common causes.

1.3.1. DNA lesions

Unrepaired DNA lesions are a well-documented cause of replication stress, with thousands of endogenously and exogenously generated lesions produced every day in human cells (the formation and repair of different types of DNA damage is explored in more detail in a later section) (Gaillard et al., 2015). Unrepaired DNA lesions, or single stand breaks (SSBs) generated as repair intermediates, can present barriers to replication fork progression. For instance, base lesions, bulky DNA adducts, interstrand crosslinks or AP sites can directly block the replisome, leading to fork stalling and, where leading strand synthesis is blocked, risking helicase and polymerase uncoupling (Byun et al., 2005, Minca and Kowalski, 2011, Fu et al., 2011, Hung et al., 2020). In the case of single strand lesions, it has been estimated that at

least 1% of SSBs escape the activity of DNA repair pathways, even in healthy, unperturbed human cells (Vilenchik and Knudson, 2003). During S-phase, these impediments can cause replication fork stalling or collapse, generating approximately 50 single ended DSBs per cell cycle (Vilenchik and Knudson, 2003, Cannan and Pederson, 2016).

1.3.2. Conflicts between transcription and replication

Both replicative and transcriptional machineries use the DNA duplex as a template for the synthesis of nucleotides. During S-phase, these processes occur concurrently, with R-loops, structures formed when newly transcribed RNA hybridises with the denatured template strand, and transcription complexes acting as potential barriers to replication (Azvolinsky et al., 2009, Gan et al., 2011, Kotsantis et al., 2016, Hamperl et al., 2017, Gomez-Gonzalez and Aguilera, 2019, Lalonde et al., 2021). Like DNA polymerases, RNA polymerases synthesise nucleotides in a 3' to 5' direction, thereby encountering the replisome codirectionally on the leading strand and head-on on the lagging strand. Whilst studies have suggested that both types of collisions may act as obstacles for replication, it is widely established that head-on collisions have the most deleterious impact, leading to replication fork stalling, DSB generation and hyper-recombination (Prado and Aguilera, 2005, Gan et al., 2011, Hamperl and Cimprich, 2016, Hamperl et al., 2017). The reason behind this difference is yet to be fully investigated in eukaryotes, but may be explained by positive supercoiling generated ahead of the converging complexes (Hamperl and Cimprich, 2016), or topological stress arising from tethering RNA transcripts to the nuclear pore impeding the progression of the replication fork (Bermejo et al., 2011, Zeman and Cimprich, 2014).

1.3.3. ‘Difficult-to-replicate’ regions

Certain regions of the genome are inherently difficult to replicate. Numerous factors contribute to these problems, including repetitive elements and regions of highly active transcription, both of which promote the formation of complex secondary structures. If the frequency of these structures is increased or they are not properly resolved during S-phase, they can impede the movement of the replisome. As such, these regions exhibit increased levels of fork stalling or fork collapse (Bhowmick and Hickson, 2017, Ozer and Hickson, 2018, Kaushal and Freudenreich, 2019). They also frequently give rise to replication stress-associated ultra-fine anaphase bridges (UFBs), which form when sister chromatids remain physically associated due to unresolved replication intermediates or under-replicated DNA (Chan et al., 2009, Naim and Rosselli, 2009, Sarbajna et al., 2014, Oestergaard and Lisby, 2017). The genome is comprised of numerous ‘difficult-to-replicate’ loci, of which ribosomal DNA (rDNA), telomeres and common fragile sites (CFSs) are the most well defined (Oestergaard and Lisby, 2017, Ozer and Hickson, 2018).

Ribosome biogenesis is an incredibly demanding process. As such, rDNA represents the most heavily transcribed region of the human genome, encompassing several genomic loci consisting of hundreds of guanine-rich, repetitive genes (Warmerdam and Wolthuis, 2019). To minimise head-on collisions between the replication fork and the transcriptional machinery, eukaryotic rDNA repeats are separated by replication fork barriers (RFBs) which inhibit the movement of replication forks progressing in the opposite direction to transcription (Kobayashi et al., 1998, Kobayashi, 2003, Castan et al., 2017, Gomez-Gonzalez and Aguilera, 2019). In *Saccharomyces cerevisiae*, whilst most forks at RFBs remain stabilised (Brewer et al., 1992), presumably awaiting

rescue by a converging fork, a minority of forks collapse (Fritsch et al., 2010), with fork collapse increasing significantly when the barrier becomes dysfunctional (Kobayashi, 2003). The active transcription of ribosomal genes means these regions are also prone to forming obstructive DNA:RNA R-loop hybrids (Hamperl et al., 2017, Warmerdam and Wolthuis, 2019), whilst their high GC content promotes the formation of G4-quadruplexes: a complex architecture in which repetitive guanine motifs stack together and become stabilised (described in more detail below) (Datta et al., 2021). As such, elevated levels of replication instability at these regions contributes to aberrant recombination and the hypervariability of ribosomal gene copy numbers (Kobayashi et al., 1998, Salim et al., 2017).

Like rDNA, telomeres are very large and repetitive. These specialised nucleoprotein structures occur at each chromosome end and consist of a tandem array of hexanucleotide repeats, bound by a six-protein complex called shelterin. This assembly protects the ends of the chromosomes from degradation and aberrant repair (Denchi and de Lange, 2007, Pitt and Cooper, 2010, Sfeir and de Lange, 2012, Lim and Cech, 2021). Whilst heterochromatic, telomeres can be transcribed by RNA polymerase II to produce long non-coding RNA, TERRA, which forms DNA:RNA R-loop hybrids at chromosome ends and contributes to telomere protection (Balk et al., 2013, Lopez de Silanes et al., 2014, Montero et al., 2016). Telomeres also have a high GC skew, which facilitates G4-quadruplex formation (Paeschke et al., 2005), and, together with TERRA-associated R-loops and the tightly bound shelterin complex, lead to obstruction of the replisome and fork stalling (Ohki and Ishikawa, 2004, Pan et al., 2019, Matmati et al., 2020, Feretzaki et al., 2020). In addition to this, telomeres frequently form T-loop structures in which the terminal ssDNA overhang

invades the upstream, double stranded telomeric DNA (Griffith et al., 1999, Doksani et al., 2013). It is hypothesised that these structures protect the chromosome end from aberrant repair (Griffith et al., 1999, Van Ly et al., 2018), with super-resolution imaging revealing an association between linearized telomeres and the activation of the DNA damage response (Van Ly et al., 2018). Components of the shelterin complex recruit the helicase RTEL1 to telomeres during S-phase to promote T-loop unwinding and facilitate unimpeded replication, but dysfunctions in this pathway can cause the T-loop structure to persist and become obstructive during replication (Vannier et al., 2012, Sarek et al., 2016).

Fragile sites represent large, repetitive regions of the genome, spanning hundreds to thousands of kilobases, which are characterised by the presence of replication stress-induced chromosomal aberrations. These regions are characterised as ‘common’ or ‘rare’ depending on their prevalence in the general population (Zlotorynski et al., 2003, Oestergaard and Lisby, 2017). Whilst CFSs exist in all individuals, rare fragile sites are observed in only 5% of the population (Zlotorynski et al., 2003). Numerous intrinsic factors are believed to contribute to the fragility of these sites. Notably, these repetitive regions are often very AT-rich, which makes them highly flexible, with reduced helix stability, and are thereby prone to form stable, complex secondary structures (Mishmar et al., 1998, Irony-Tur Sinai et al., 2019, Kaushal et al., 2019, Li and Wu, 2020). In addition, CFSs are frequently located in large, actively transcribed genes which take a full cell-cycle to transcribe (Helmrich et al., 2011, Pentzold et al., 2018). This makes collisions between replication and transcription machineries inevitable and can increase the frequency of obstructive R-loops (Helmrich et al., 2011). These factors, alongside a characteristic lack of replication origins which could

otherwise rescue stalled forks, increase fork collapse at these regions (Helmrich et al., 2011, Miotto et al., 2016, Neil et al., 2018, Sugimoto et al., 2018, Gomez-Gonzalez and Aguilera, 2019). Accordingly, CFSs are associated with breakpoints and genome deletions in cancers (Hellman et al., 2002, Finnis et al., 2005, Bignell et al., 2010, Glover et al., 2017).

1.3.4. R-loops

R-loops form when newly transcribed RNA hybridises with the denatured template strand, causing the non-template strand to be displaced. Regions with a GC-skew and strand asymmetry in base pair distribution increase the likelihood for R-loop formation, particularly when the cytosine-rich strand serves as a transcription template (Allison and Wang, 2019). These structures have been shown to have important cellular functions, particularly at gene promoters and terminators where they can influence epigenetic mechanisms of gene regulation (Ginno et al., 2013, Niehrs and Luke, 2020). However, when they form aberrantly, they can, as discussed, impede replication, and are associated with genome instability (Gan et al., 2011, Hamperl and Cimprich, 2014, Hamperl et al., 2017, Parajuli et al., 2017, Allison and Wang, 2019). Increased supercoiling behind the transcription fork also promotes the formation of R-loops by increasing the opportunity for interactions between the RNA and the template strand, with head-on-collisions between transcription and replication machineries likely exacerbating R-loop formation by increasing this torsional stress (El Hage et al., 2010, Kuzminov, 2018, Allison and Wang, 2019). In addition to this, the displaced non-template is prone to forming additional secondary structures which may stabilise the DNA:RNA hybrid and act as further barriers to replication (Hamperl and Cimprich, 2014).

The cell protects against R-loops in numerous ways. For instance, topoisomerases have been observed to alleviate torsional stress and prevent the formation of R-loops at rDNA (El Hage et al., 2010). RNA-binding protein complexes which associate with the nascent RNA and couple transcription with swift nuclear export are also believed to be important for regulating R-loops, preventing the transcript from interacting with the template DNA. Consistent with this, the depletion of the THO/TREX transcription export complex leads to increased levels of hyperrecombination which can be reduced upon R-loop disruption (Huertas and Aguilera, 2003, Gomez-Gonzalez and Aguilera, 2007). When they do form, R-loops are more stable than dsDNA so must be enzymatically dissembled. This can be achieved by RNase H enzymes, which eliminate ribonucleotides within DNA:RNA hybrids through 5'-3' exonuclease activities (Cerritelli and Crouch, 2009, Parajuli et al., 2017), although R-loops have also been suggested to be removed via helicases, which unwind the structure (Mischo et al., 2011, Chang et al., 2017). The FA complex, which enables the replication fork to bypass ICLs, has also been associated with R-loop resolution and the proper progression of the replication fork through transcription-replication conflicts (Schwab et al., 2015, Liang et al., 2019).

1.3.5. G4-quadruplexes

G-quartets assemble when four guanines interact in a cyclical arrangement, stabilised by Hoogsteen bonds. These quartets can then stack together in three or more layers, called G4-quadruplexes (Figure 1.3). These structures can be extremely polymorphic, with their precise architecture depending on strand length and direction (Rhodes and Lipps, 2015, Lerner and Sale, 2019). The opportunity for G4-quadruplex formation is increased when ssDNA becomes available, with the bonds between duplex strands

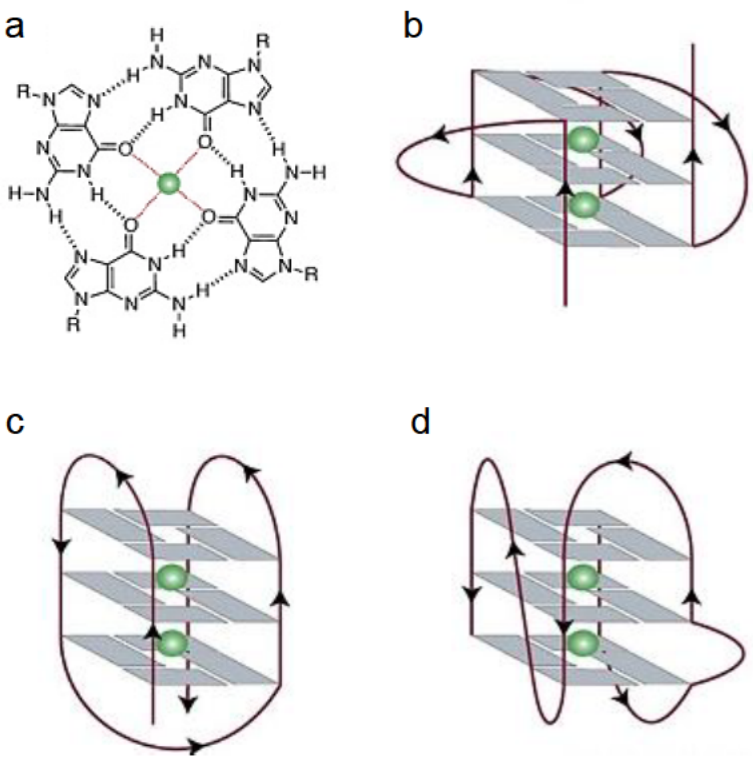


Figure 1.3. Schematic of G-quartets and G4-quadruplexes

The structure of a G-quartet (a), the basic structural motif of G4-quadruplexes, formed by four Hoogsteen bonded guanines and stabilised by a central cation (shown in green). G-quartets can stack in various arrangements to form G4-quadruplexes. Depicted are (b) parallel, (c) anti-parallel, and (d) hybrid intramolecular G4-quadruplexes, with arrows indicating strand direction. Figure adapted from (Spiegel et al., 2020).

no longer impeding the interactions between repetitive guanines (Lerner and Sale, 2019). Accordingly, G4-quadruplexes are thought to preferentially form at the 3' overhang of telomeres (Tang et al., 2007), as well as ssDNA generated from replication and transcription bubbles (Zhang et al., 2014).

However, these structures can also form on dsDNA, particularly at gene regulatory elements (Huppert and Balasubramanian, 2007, Hänsel-Hertsch et al., 2016). In fact, 42% of human genes contain at least one G4 motif at their promoters (Huppert and Balasubramanian, 2007). The stability of G4-quadruplexes depends on a number of factors, including the number of G-quartets and their topology (Pandey et al., 2013). However, these structures have also been shown to be spatially and temporally regulated by specific helicases, which unwind them *in vivo*. In human cells, a number of helicases have been associated with the resolution of G4-quadruplexes, including WRN, BLM, FANCJ and ATRX (Sauer and Paeschke, 2017), with WRN and BLM-mediated unwinding of these structures implicated in gene expression changes (Johnson et al., 2010, Nguyen et al., 2014, Tang et al., 2016). This is consistent with genome wide gene expression studies which show significant changes in numerous G4-promoter enriched genes after the addition of G4-quadruplex stabilising ligands (Verma et al., 2009). Together, this suggests that G4-quadruplexes have distinct functional roles in controlling gene transcription at regulatory gene elements where they may act to recruit transcription associated factors, exclude nucleosomes or block the RNA polymerase (Kim, 2019). Alongside the 3' overhang, the guanine-rich nature of telomeres is also believed to contribute to an increased formation of G4-quadruplexes at chromosome ends (Granotier et al., 2005, Müller et al., 2010, Bryan, 2020). Like T-loops, it may be that the architecture of G4-quadruplexes prevents the

telomeric ends associating with DNA repair proteins or nucleases which could degrade the terminal sequences or cause aberrant fusions. However, experimental evidence for this function is very limited (Bryan, 2020). Instead, it has been proposed that G4-quadruplex formation has dynamic roles in the regulation of telomere extension. On one hand, it has been suggested that these structures may act antagonistically to POT1, a ssDNA binding protein which promotes the recruitment of telomerase, the enzyme responsible for telomeric extension (Zaug et al., 2005, Wang et al., 2011a). However, more recent research suggests that G4-quadruplex formation in telomerase product DNA may positively regulate telomere extension by increasing telomerase processivity (Jansson et al., 2019, Patrick et al., 2020).

Although G4-quadruplexes appear to have numerous important cellular functions, they have also been shown to contribute to genome instability. Extensive *in vitro* experimentation has demonstrated that persistent G4-quadruplexes can stall replication forks (Castillo Bosch et al., 2014, Edwards et al., 2014, Lee et al., 2021), interrupting replication on both the leading and lagging strands (Lerner and Sale, 2019). In support of a similar function *in vivo*, genomic fragility and replication stalling is exacerbated upon treatment with G4-quadruplex stabilising drugs and the depletion of G4-quadruplex unwinding helicases in many cellular models (Kruisselbrink et al., 2008, London et al., 2008, Dahan et al., 2018, Obi et al., 2020). Whilst replication forks can encounter and become stalled by pre-formed structures, G4-quadruplexes can also form within the replisome as ssDNA is generated (Lerner and Sale, 2019). Recent work has demonstrated that these persistent structures can prevent RPA from binding to the replisome's ssDNA, leading to the silencing of replication stress signalling and replication fork collapse (Lee et al., 2021). Consistent with G4-

quadruplexes being more frequent at telomeric sequences, telomere replication appears to be particularly impacted by the presence of G4-structures (Crabbe et al., 2004, Rizzo et al., 2009, Zhang et al., 2019), with breakage at these loci leading to dramatic telomere loss and deleterious fusions upon G4-quadruplex stabilisation (Tahara et al., 2006, Zhang et al., 2019).

1.4. Mitosis

Mitosis describes a dynamic phase of the cell cycle in which newly duplicated genetic material is divided equally into daughter cells. As will be discussed in more detail in later sections of the introduction, this requires the actions of condensin, which condenses the DNA into chromosomes (Paul et al., 2019), and cohesin, which topologically entraps sister duplex DNA together after replication, dissociating from the chromatin in a timely manner to enable sister chromatids to segregate successfully (Peters et al., 2008).

In higher eukaryotes, mitosis begins with prophase, when the CDK1 kinase phosphorylates condensin II, causing the condensin complex to load onto the chromatin and trigger chromosome compaction (Abe et al., 2011). At the same time, cohesion starts to dissociate from the sister duplex DNA, an action initiated via Polo-like kinase 1 (PLK1)-mediated phosphorylation (Sumara et al., 2002, Hauf et al., 2005). Centrosomes, which are replicated during S-phase, travel to opposite ends of the cell where they form microtubule-based bipolar spindles (Conduit et al., 2015), whilst late prophase sees kinases such as CDK1 and PLK1 trigger the breakdown of the nuclear envelope (Güttinger et al., 2009, Linder et al., 2017, Martino et al., 2017). In the second phase of mitosis, prometaphase, the microtubules extend from the

centrosomes, undergoing successive rounds of GTP-dependent polymerisation and depolymerisation, alternating between growth and shrinkage (Gudimchuk and McIntosh, 2021). Extension of a subset of microtubules into the nuclear space facilitates their attachment to the chromosomes' kinetochores- protein complexes assembled at the centromere (Rusan et al., 2002, Gudimchuk and McIntosh, 2021). Proper microtubule-kinetochore attachment is vital to make sure that sister chromatids can become properly bioriented, so iterative rounds of error correction, mediated in part by Aurora B, are employed to destabilise incorrect attachments (Cimini et al., 2006).

During this period, inappropriate or early chromosome segregation is blocked by the spindle assembly checkpoint (SAC), which detects these improper microtubule-kinetochore attachments. Several proteins have been identified as components of the mitotic checkpoint complex (MCC) (BUBR1, BUB3, CDC20 and MAD2), whose concerted actions, alongside the regulatory proteins PLK1, Aurora B, TRIP13 and CDK1, impose the SAC by inhibiting the anaphase-promoting complex/cyclosome (APC/C) (Liu and Zhang, 2016, Ma and Poon, 2016). As chromosomes become amphitelically attached to the microtubules, they line up along the spindle equator-forming the metaphase plate. Once the microtubules are correctly attached, the metaphase to anaphase transition is initiated (Etemad et al., 2015). SAC-mediated inhibition is alleviated, allowing the APC/C to activate and polyubiquitylate its mitotic substrates, including CDK1/Cyclin B and securin, targeting them for degradation (Thornton and Toczyski, 2003). The loss of securin and disruption of CDK1/Cyclin B relieves the inhibition of seperase, which can then bind and cleave centromeric

cohesin, causing it too to dissociate from the sister chromosomes (Gorr et al., 2005, Rosen et al., 2019).

The complete removal of cohesin triggers anaphase (Wirth et al., 2006), in which the microtubules depolymerise and shorten, generating pulling forces which separate the sister chromatids to opposite sides of the cell (Gudimchuk and McIntosh, 2021). During telophase, the chromatids reach the spindle poles and begin to decondense (Antonin and Neumann, 2016), whilst the nuclear envelope also starts to reassemble, forming two daughter nuclei (Güttinger et al., 2009). This is followed by cytokinesis, in which the cytoplasm around the two nuclei furrows, eventually dividing the cell in two (Bringmann and Hyman, 2005, Barr and Gruneberg, 2007). Whilst mitosis is highly regulated, numerous errors, including improper chromosome condensation, SAC signalling or microtubule-kinetochore attachments, can prevent the proper biorientation of chromosomes, which can cause them to missegregate into daughter cells and give rise to aneuploidies (Matsuura et al., 2006, Samoshkin et al., 2009, Bakhoum et al., 2009, Potapova and Gorbsky, 2017, Carvalhal et al., 2022). Equally, if chromosome disjunction fails to occur successfully then cytokinesis can abort, producing binucleated, polyploid cells (Potapova and Gorbsky, 2017).

To ensure proper chromosome segregation, the cell employs checks to make sure that replication has successfully completed prior to mitosis. As previously discussed, ATR induced CHK1 signalling disrupts the CDK1/Cyclin B complex and triggers the G2/M checkpoint, preventing cells with damaged or under-replicated DNA from entering mitosis and allowing time for replication and repair (Zhang and Hunter, 2014). However, mild or moderate replication stress can escape checkpoint activation, leading to under-replicated DNA persisting through to mitosis (Wang et al., 2011b,

Koundrioukoff et al., 2013, Gelot et al., 2015). Whilst the precise links between replicative dysfunction and mitosis are undefined, mild replication stress has been shown to induce segregation defects by promoting the premature disengagement of centrioles, the core centrosome components, leading to transient multipolar spindles (Wilhelm et al., 2019a). In addition to this, centrosome duplications have been observed upon replication stress in HR-deficient cells (Daboussi et al., 2005, Wilhelm et al., 2014).

Persistent under-replicated DNA, recombination intermediates and catenated chromosomes that are not resolved before the metaphase to anaphase transition can also generate chromatin bridges or UFBs, the latter of which form at regions of genomic fragility (centromeres, telomeres and chromosome fragile sites) (Chan et al., 2009, Fernández-Casañas and Chan, 2018, Chan et al., 2018). It is not yet known how chromatin bridges may resolve, although immunofluorescent staining of UFBs reveals that these structures become decorated with BLM and PICH helicases, which, alongside Top2, are implicated in their resolution. However, unresolved bridging can lead to subsequent missegregation events when the interlinked sister chromatids attempt to separate (Chan et al., 2007, Chan and West, 2018). Furthermore, under the stress of the mitotic apparatus and cell cleavage, these structures can break, giving rise to DNA damage and chromosome fusions which can then enter into breakage-fusion-bridge cycles (Umbreit et al., 2020).

1.5. DNA damage

It is estimated that human cells are subject to approximately 70,000 DNA damaging events per day (Tubbs and Nussenzweig, 2017). These events induce lesions which

change the basic structure of DNA and, if left unrepaired, can impede important cellular processes, including replication, and introduce deleterious mutations that threaten cell viability. Common types of lesions include the chemical alteration or misincorporation of DNA bases, as well as breaks in the DNA phosphate backbone (Chatterjee and Walker, 2017).

DNA damaging events can be triggered by either endogenous or exogenous sources. Those classed as endogenous primarily occur due to normal metabolic processes in which DNA engages in hydrolytic or oxidative reactions (Tubbs and Nussenzweig, 2017, Chatterjee and Walker, 2017). For instance, reactive oxygen species, by-products of cellular respiration, can attack DNA bases and the DNA backbone, leading to several types of lesions including oxidized bases, as well as single- and double-strand breaks (Lindahl, 1993, Shokolenko et al., 2009, Poetsch, 2020). Random errors in DNA replication and repair, or the abortive activity of topoisomerases, can also introduce DNA damage via normal cellular mechanisms (Kunkel, 2009, Gomez-Herreros et al., 2014, Gomez-Herreros, 2019).

In contrast, DNA damage can also occur exogenously following exposure to environmental, physical, and chemical agents. This includes ionizing radiation (IR), which occurs naturally from rocks and soils, as well as medical devices. Comprised of X-rays and gamma rays, IR can impact DNA directly by inducing base lesions as well as single- and double-strand breaks. Further, indirect damage is also caused by IR-induced free radicals which can attack DNA to produce base damage (Gulston et al., 2002, Spitz et al., 2004, Reisz et al., 2014, Chatterjee and Walker, 2017). Ultraviolet (UV) radiation, a component of sunlight, generates covalent linkages between adjacent pyrimidines, forming various types of bulky lesions, such as

cyclobutane pyrimidine dimers and pyrimidine (6 – 4) photoproducts (Rastogi et al., 2010, Hung et al., 2020). Finally, chemical agents, including aromatic amines and polycyclic aromatic hydrocarbons, both found in cigarette smoke, form persistent bulky DNA adducts (Barnes et al., 2018).

At significant levels, the accumulation of DNA damage can induce cellular senescence or death (Surova and Zhivotovsky, 2013). To avoid this, the cell employs numerous, and often overlapping, DNA damage detection, signalling and repair pathways which aim to respond to the many types of DNA lesions that can arise. These responses typically induce cell-cycle arrest, preventing mitosis until the damaged DNA has been repaired, thus ensuring the faithful inheritance of genomic information (Sancar et al., 2004, Chatterjee and Walker, 2017). In this section I will discuss in more detail some of the major DNA repair pathways.

1.5.1. Double strand break repair

1.5.1.1. Double strand breaks

DSBs represent breakages in the sugar-phosphate backbone of both DNA strands. These lesions are considered to be particularly genotoxic if unrepaired, risking mutagenesis, chromosome rearrangements, and aberrant recombination (Jackson, 2002). Alongside exogenous damages, they can form as by-products of cellular processes, such as abortive topoisomerase action or the collision of the replication machinery with SSBs, subsequently generating single ended DSBs (Cannan and Pederson, 2016, Mehta and Haber, 2014). Whilst rare, a single DSB can be sufficient to trigger cell death (Jackson and Bartek, 2009). Therefore, it’s extremely important for these lesions to be repaired. To do so, the cell employs a number of double strand

break repair (DSBR) mechanisms, of which nonhomologous end joining (NHEJ) and HR represent two major repair pathways.

The choice between NHEJ and HR is determined primarily by cell-cycle stage (Saleh-Gohari and Helleday, 2004, Brandsma and Gent, 2012). NHEJ is a relatively simple mode of DSBR which occurs throughout the cell-cycle, being particularly important during G1, in which blunt DNA ends are ligated together (Lieber, 2010). During this process, the DSB is first sensed by the Ku70/Ku80 heterodimer, which binds to either free end of the break (Walker et al., 2001). Ku70/80 recruits the NHEJ complex, including the Artemis:DNA-PKcs nuclease and the ligation complex (DNA ligase IV-XRCC4-XLF) (Drouet et al., 2005, Yano et al., 2009). Together, Ku70/80 and DNA-PKcs form a synaptic complex which bridges the DNA DSB ends together and activates DNA-PKcs's kinase activities (DeFazio et al., 2002, Chen et al., 2021, Yue et al., 2020), allowing the kinase to autophosphorylate and activate Artemis (Ma et al., 2002, Gu et al., 2010, Yue et al., 2020). To produce blunt DNA ends, DNA overhangs can then be processed, either via Artemis mediated nucleolytic cleavage or DNA polymerase gap filling (Ma et al., 2005, Lieber, 2010, Ramsden, 2011). Finally, these blunt ends are joined by the ligation complex and the NHEJ machinery dissociates from the DNA (Lieber, 2010).

The other major DSBR pathway, HR, is restricted during G1, but is active during S/G2 phase, when DNA has been replicated, as it uses homologous sister chromatids as templates for repair (Krejci et al., 2012). Whilst Ku70/Ku80 still senses and binds to the break, the complex is replaced by MRN (MRE11-RAD50-NBS1), which removes the complex via endonucleolytic cuts (Myler et al., 2017). MRN then promotes the resection of the DSB ends, thus preventing the simple re-ligation of the lesion and

committing pathway choice to HR (Brandsma and Gent, 2012, Shibata et al., 2014). This action is regulated by a number of mechanisms, including the activities of CDKs which target components of the HR resection machinery (Ferretti et al., 2013).

1.5.1.2. Initiation of homologous recombination

Once recruited to the lesion, MRN tethers the broken DSB ends together (Williams et al., 2008b). Following the detection of the break, CtIP, together with RAD50, promotes the endonuclease activity of MRE11, causing cleavage of the 5'-terminated DNA strand (Sartori et al., 2007, Anand et al., 2016). The cleaved site is then bidirectionally resected, 3'-5' by MRE11's exonuclease activity and 5'-3' by the nuclease activities of EXO1 or DNA2, the latter in complex with BLM or WRN (Nimonkar et al., 2011, Sturzenegger et al., 2014). This generates long 3'-ssDNA overhangs, often kilobases in length (Daley et al., 2015). MRN also recruits ATM via an interaction with NBS1, thus initiating a signalling cascade, resulting in cell-cycle arrest via intermediates like p53 (Uziel et al., 2003).

The single stranded DNA overhangs are rapidly coated by RPA, which prevents further nuclease activity and inhibits the formation of secondary structures (Chen et al., 2013, Dueva and Iliakis, 2020). RPA is subsequently displaced by the RAD51 recombinase, an action promoted by the pro-recombinogenic factor BRCA2 (Davies et al., 2001, Liu et al., 2010, Dueva and Iliakis, 2020). The recombinase bound ssDNA forms the presynaptic filament, the 3' end of which then probes for homologous sequences on the intact, complementary sister chromatid (Li and Heyer, 2008, Liu et al., 2011a). Once located, the 3' ended tail invades the DNA duplex, generating a displacement-loop (D-loop). Using the sister chromatid as a template, polymerase δ

mediated DNA synthesis is primed from the 3' end of the invading strand and restores the DNA sequence lost by resection (Li and Heyer, 2008, Liu et al., 2011a).

The D-loop intermediate structure generated by strand invasion can be resolved by various pathways to reinstate the DNA duplexes, as will be discussed. Pathway choice determines if genetic information is exchanged between the interacting chromatids, generating either a noncrossover or crossover product. Whilst crossovers are important during meiosis to promote proper chromosome segregation and genetic diversity (Cole et al., 2010), elevated levels of these events during mitosis risks a loss of heterozygosity (LOH) (LaRocque et al., 2011). As such, HR is regulated to favour noncrossover events (Sanchez et al., 2021).

1.5.1.3. Double Holliday junction resolution

Following D-loop formation, the capture of the second end of the DSB can generate a double Holliday junction (dHJ) if DNA ligation occurs, meaning that sister chromatids become linked (Nimonkar and Kowalczykowski, 2009). dHJs are primarily resolved by 'dissolution', in which only noncrossover products are produced (Figure 1.4) (Matos and West, 2014). This process is mediated by the 'dissolvasome', or BTR complex, which is comprised of the BLM helicase (Sgs1 in yeast) and topoisomerase 3α, alongside the 'RecQ-mediated instability factors', RMI1 and RMI2 (Wu and Hickson, 2003, Manthei and Keck, 2013). To initiate the process, BLM catalyses the migration of the HJs towards each other where they collapse to form a hemicatenated structure (Karow et al., 2000, Manthei and Keck, 2013). This is then decatenated by topoisomerase 3α in a BLM-dependent manner. Whilst the functional role of the helicase at this point is still poorly understood, it could be that BLM is important for

mediating strand passage, or for unwinding the hemicatenated structure into a topoisomerase-accessible conformation (Wu et al., 2005, Plank et al., 2006, Bizard and Hickson, 2014). dHJs which escape the BTR complex can be processed later in the cell cycle by ‘resolution’, a pathway which uses structure specific endonucleases, termed resolvases, to cleave the HJs in an unbiased manner (Matos and West, 2014, Wyatt and West, 2014). This means that, depending on the orientation of the cleavage, either noncrossover or crossover events can be produced (Figure 1.4) (Wyatt and West, 2014). In mammalian cells, several resolvases have been identified. SLX1-SLX4 has been shown to cooperate with MUS81-EME1 to initiate resolution, with SLX1 cleaving the DNA first, followed by the endonuclease activity of MUS81 on the opposite strand (Wyatt et al., 2013). In a distinct pathway, the GEN1 resolvase also cleaves HJs (Rass et al., 2010, Wyatt et al., 2013), although epistasis experiments have suggested that the enzyme acts as a backup pathway for the activities of SLX1-SLX4/MUS81-EME1, which are more efficient (Wechsler et al., 2011, Garner et al., 2013).

1.5.1.4. Synthesis dependent strand annealing

Synthesis dependent strand annealing (SDSA) is thought to represent a major mechanism of DSBR in mammalian cells, which may be because the pathway limits potentially deleterious crossover events. (Zapotoczny and Sekelsky, 2017). During SDSA, after strand invasion and extension, the migrating D-loop structure becomes unstable, unwinds, and the invading ssDNA disassociates before second end capture can occur. Thus, dHJs are not generated and SDSA avoids the action of crossover-promoting resolvases, biasing repair towards noncrossover events and avoiding LOH

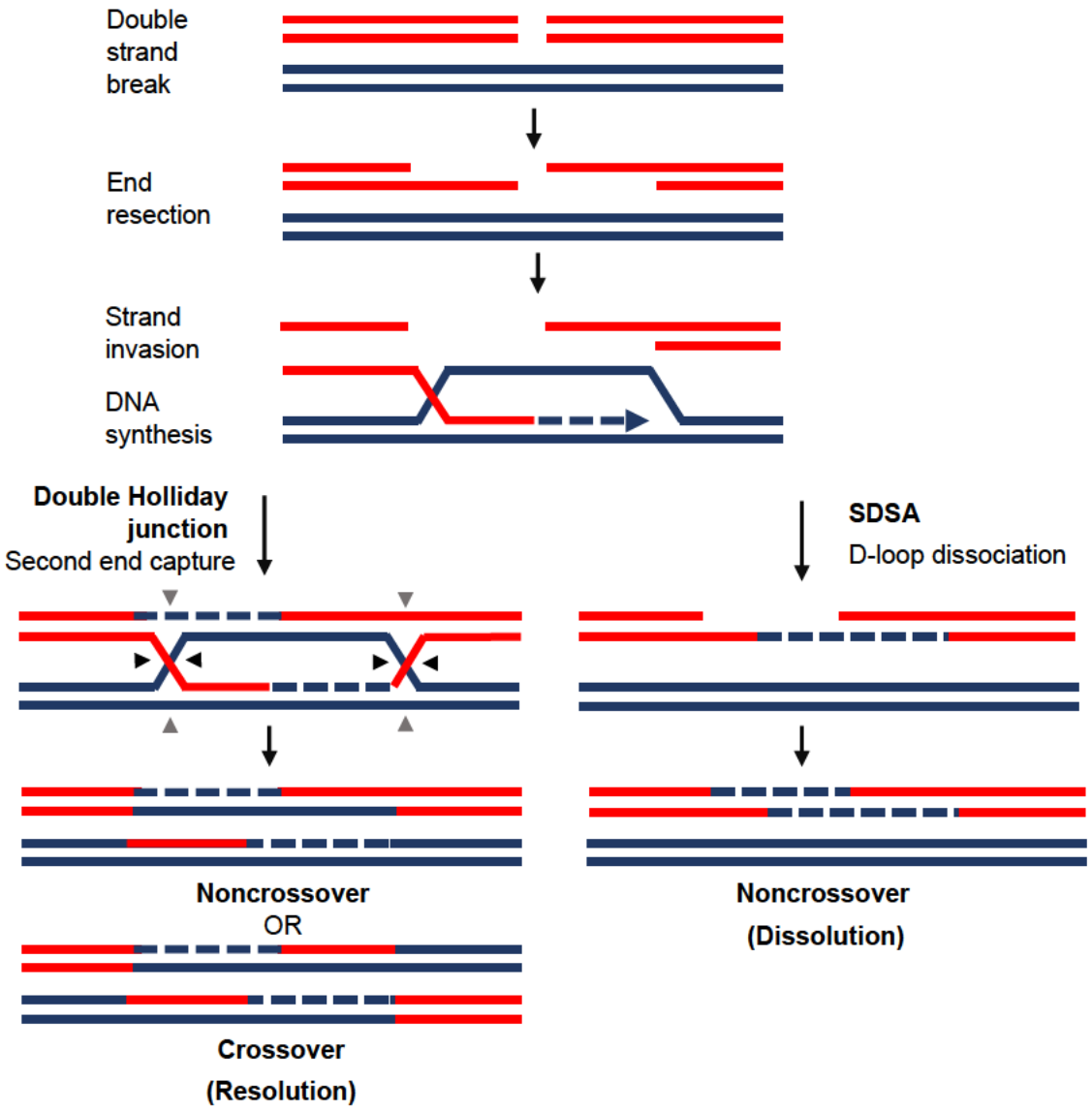


Figure 1.4. Molecular model of crossover and noncrossover pathways.

During HR, DSBs are detected and resected to produce 3' overhangs. This is followed by strand invasion, in which RAD51-ssDNA 3' presynaptic filaments invade homologous DNA on the opposing DNA duplex, generating a D-loop from which DNA synthesis can prime. During synthesis dependent strand annealing (SDSA), the elongated 3′ end of the invading DNA dissociates from the D-loop and reanneals with the second free 3′ end of the break site. The resulting ssDNA gaps are filled, generating a non-crossover product. In contrast, D-loops can also be captured by the second DSB end and ligated to produce a double Holliday junction structure, which can then be resolved by a dissolution or resolution sub-pathway. During dissolution, the Holliday junctions migrate towards each other and the hemicatenated sister chromatids are then separated in a BLM-TOP3a-RMI1 dependent process which produces non-crossover products. However, during resolution, the nonbiased cleavage of the Holliday junctions by structure-specific endonucleases, termed resolvases, generate either crossover or non-crossover events.

(Morrical, 2015). Instead, the newly synthesised and dissociated DNA anneals to thecomplementary end of the break where the resulting SSBs can be repaired with gap-filling DNA synthesis and ligation (Verma and Greenberg, 2016, Wright et al., 2018). Whilst the precise dynamics of downstream SDSA are poorly understood, studies in fruit flies have suggested that FANCM and BLM helicases may be integral for this mode of repair, with the loss of either factor inducing SDSA defects, accompanied by deletions on either side of the DSB in BLM mutants (McVey et al., 2004, Romero et al., 2016). These observations are consistent with a role for both proteins in promoting SDSA via D-loop disassembly.

1.5.1.5. Break induced replication

Most models of HR involve two ended-DSBs, which enable second end capture and reannealing of the invading ssDNA. Contrastingly, break induced replication (BIR) acts as a repair pathway when only one DSB end is present (Malkova and Ira, 2013, Kramara et al., 2018), such as single-ended DSBs which are generated when a replication fork collapses. During this process, the intermediate D-loop structure migrates across the DNA, allowing repair synthesis to continue until it meets with the next converging replication fork (Saini et al., 2013). Whilst BIR is important for replication restart, particularly in regions that lack replication origins, it is also highly mutagenic, with BIR-associated mutations 1000-fold higher compared to normal S-phase replication (Deem et al., 2011, Saini et al., 2013). This is likely caused, at least in part, by replication within the translocating D-loop, which leads to the increased dissociation of polymerase δ from the template DNA, subsequent template switching, strand slippage and frameshift mutations (Smith et al., 2007, Deem et al., 2011, Malkova and Ira, 2013, Jalan et al., 2019). The rapid dissociation of newly synthesised

DNA from the migrating structure also interferes with MMR, which is no longer able to sense and correct nucleotide misincorporations (Deem et al., 2011, Sakofsky et al., 2012).

1.5.1.6. Single strand annealing

Single strand annealing (SSA) describes the repair of DSBs flanked by long, homologous sequences (Bhargava et al., 2016). These sequences are resected and annealed together, thus avoiding the need for a homologous sister chromatid template. As described for previous modes of HR, resection of the DNA ends is mediated by the nuclease activities of MRE11, EXO1 and DNA2. During SSA, RAD52 then promotes the annealing of the resulting homologous 3' ssDNA overhangs (Van Dyck et al., 2001), leaving the non-homologous ends as free tails that need to be removed. Their cleavage is believed to be catalysed by the ERCC1/XPF complex, a structure-specific endonuclease, before the gaps are filled by DNA polymerase (Sargent et al., 2000, Al-Minawi et al., 2008).

It's likely that SSA is important for promoting DSBR in situations where resection has been initiated, but no homologous template is available. However, the pathway is error-prone, generating a deletion rearrangement between flanking homologous regions (Bhargava et al., 2016). Therefore, to prevent excessive SSA-mediated mutagenesis, RAD51 acts as a negative regulator of the pathway. By binding to the 3' ssDNA overhangs, the recombinase limits the DNA annealing activity of RAD52 and promotes other mechanisms of repair (Stark et al., 2004, So et al., 2022).

1.5.2 The Fanconi anaemia pathway

DNA interstrand crosslinks (ICLs) covalently link opposing DNA strands. During S-phase, these lesions act as potent blocks to the replication machinery and transcription machinery. As such, their removal is important for maintaining genome stability, and is coordinated by the Fanconi anaemia (FA) pathway, comprising of 22 identified FA or FA-like proteins (FANCA-W) that work together alongside several HR factors (Deans and West, 2011, Ceccaldi et al., 2016, Michl et al., 2016). The FA pathway is activated when converging replication forks stall due to an impeding ICL (Zhang et al., 2015a), with the lesion then recognised by the FANCM–FAAP24–MHF1–MHF2 anchor complex (Ceccaldi et al., 2016). This anchor complex then recruits the 9-protein FA core complex, including FANCL, an E3 ubiquitin ligase, to the chromatin (Kim et al., 2008). Together with the E2 enzyme UBE2T, FANCL mediates the mono-ubiquitination of FANCI and FANCD2, a heterodimer otherwise known as the ID2 complex (Smogorzewska et al., 2007, Longerich et al., 2009, Li et al., 2020c). Whilst little is conclusively understood about the structure and function of the other FA core complex proteins, some studies have suggested that these subunits are also required for the successful mono-ubiquitination of the ID2 complex (Garcia-Higuera et al., 2001, Kottemann and Smogorzewska, 2013, Rajendra et al., 2014).

Mono-ubiquitinated FANCD2 acts as a platform for DNA repair factors, including nucleases which initiate the removal of the ICL. Nucleases implicated in this process include XPF, MUS81, SLX4, and FAN1 (Yamamoto et al., 2011, Kim et al., 2013, Klein Douwel et al., 2014, Chaudhury et al., 2014). Together, these factors are thought to cleave DNA surrounding the ICL on one of the two parental strands, ‘unhooking’ the lesion on one end so it remains attached to the other. This then

enables the replication machinery to bypass the damage using TLS, with ligation restoring the DNA duplex (Kim and D'Andrea, 2012). The DSB generated on the other DNA duplex can then be processed as it would during HR, via resection, strand invasion and homology searching (Michl et al., 2016). Having said this, research has also suggested that up to 60% of replication events occurring around ICLs are actually mediated via replication traverse, a pathway promoted by the DNA translocase, FANCM. In these situations, replication encounters the lesion and restarts on the distal side of the ICL, with repair pathways presumably removing the ICL from the DNA duplex post replication (Huang et al., 2013).

1.6. The SMC5/6 complex

1.6.1 SMC complexes

Highly conserved across all kingdoms of life, the structural maintenance of chromosomes (SMC) complexes are a family of ATPases which coordinate chromosome organisation. In eukaryotes, three distinct SMC complexes have been identified: condensin (SMC2/4), cohesin (SMC1/3) and SMC5/6 (Hirano, 2002, Uhlmann, 2016). All three SMC complexes share a ring-shaped structure, comprised of a core heterodimer of SMC proteins alongside several additional subunits (Figure 1.5). Both condensin and cohesin have well defined roles, establishing DNA condensation and chromosome cohesion, respectively. However, the precise functions of SMC5/6 are far less well understood (Hirano, 2002, Uhlmann, 2016, Aragon, 2018). SMC monomers are long molecules with nucleotide binding motifs at the termini, known as Walker A and B box-sequence motifs. These are separated by two coil-coiled regions and a central globular hinge domain. Each individual SMC

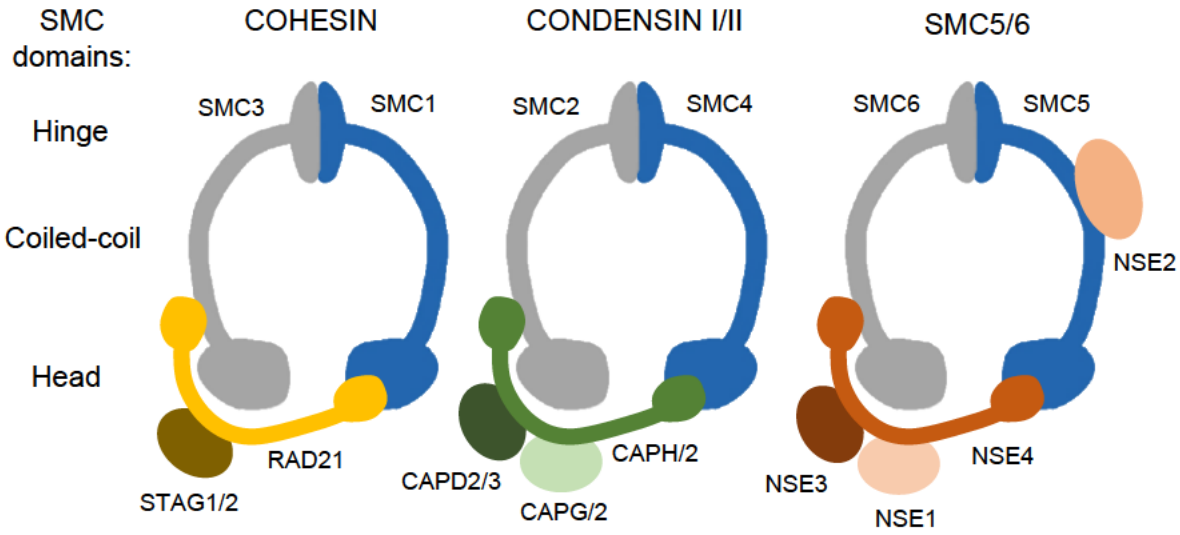


Figure 1.5. The basic structures of the human SMC complexes.

Cohesin (SMC1/3), condensin (SMC2/4) and SMC5/6 belong to a family of ATPases, all of which comprise of a heterodimer of two SMC proteins. Each SMC protein consists an ATPase head domain, and a globular hinge domain, where the SMC proteins dimerise, separated by an intramolecular coiled-coil. The head domains are bridged by additional accessory subunits which are thought to aid the complex's interaction with DNA. SMC5/6 has four non-SMC elements (NSE1-4), including NSE2, which docks onto the coiled-coil of SMC5 and has distinct SUMOylation activities. The architecture of the SMC complexes enables them to form a characteristic open, ring-like structure upon ATP binding, facilitating their interaction with DNA.

protein self-folds at the hinge, forming an ATP-binding cassette (ABC)-type ATPase head domain where the Walker A and Walker B motifs interact (Lowe et al., 2001). This is separated from the distant hinge domain by the ~ 50 nm flexible, anti-parallel arrangement of coiled-coils. Dimerization between SMC proteins occurs via the interaction of hydrophobic residues between hinge domains (Hirano and Hirano, 2002, Haering et al., 2002, Yatskevich et al., 2019), whilst the ATPase heads are asymmetrically connected by an additional 'kleisin'-type subunit (Schleiffer et al., 2003, Nasmyth and Haering, 2005, Yatskevich et al., 2019). The binding of ATP to the Walker B box- sequence motifs also initiates dimerization by stimulating the transient interaction of both head domains (Hirano, 2002, Lammens et al., 2004).

The ring-shaped structure of SMC complexes enables them to interact with DNA. Prior to ATP binding, the coiled-coil regions of cohesin, condensin and SMC5/6 have been shown to interact from hinge to head, forming a closed rod-like structure (Lee et al., 2020, Yu et al., 2021). ATP binding engages the head domains and is thought to disrupt the interaction of the coiled coils, promoting an open formation (Nasmyth and Haering, 2005, Lee et al., 2020, Adamus et al., 2020). Recent cryo-EM studies of condensin and cohesin have shown that these protein rings clamp DNA, in an interaction that requires the binding of ATP, as well as the interaction of chromatin

with DNA-sensing residues on the HEAT domains of SMC accessory subunits (Collier et al., 2020, Lee et al., 2022). SMC5/6 is believed to employ a similar mechanism of DNA interaction, with the interactions of NSE3 and NSE4 with DNA promoting the association of the complex with chromatin (Zabrady et al., 2016, Yu et al., 2022). For cohesin, ATP binding and subsequent hydrolysis promotes a conformational change in the complex, allowing it to transiently open at either the SMC1/3 hinge region or

the SMC3/RAD21 interface. These changes generate DNA transport gates through which DNA can enter and become topologically entrapped, and then later exit (Gruber et al., 2006, Murayama and Uhlmann, 2015).

1.6.2. Condensin (SMC2/4)

In preparation for cell division, DNA is compacted into chromosomes. Alongside the essential activity of Top2, which decatenates sister chromatids before mitosis (Uemura et al., 1987, Nielsen et al., 2020), this is mediated by the action of condensins. Most higher eukaryotes, including humans, have two isoforms: condensin I and II. Both complexes consist of five subunits- a core heterodimer of SMC2 and SMC4, alongside three accessory proteins (Condensin I: CAP-D2, CAP-G, and CAP-H, condensin II: CAP-D3, CAP-G2, and CAP-H2) (Figure 1.5) (Hirano, 2012). These accessory subunits dictate the localisation and function of the complexes. Whilst condensin I is restricted to the cytoplasm during interphase, loading to chromatin only after the nuclear envelope breaks down during prometaphase, condensin II is nuclear and initiates chromosome condensation during prophase (Ono et al., 2004, Thadani et al., 2012). The association of condensin with chromatin is regulated by a number of mitotic factors. For instance, the interaction of condensin I with chromatin is mediated by Aurora B (Kagami et al., 2017), whilst the phosphorylation of CAP-H2 by PLK1 regulates CAP-H2 expression and condensin II function (Kagami et al., 2017).

In vitro experiments in yeast have suggested that condensin drives changes in chromosome architecture through loop extrusion, in which the condensin complex encircles a single DNA duplex, reeling the strand through its ring to create loops and

simultaneously linking distal elements (Goloborodko et al., 2016, Ganji et al., 2018). Loop extrusion has been observed to be mediated differently by each condensin, with condensin II promoting the formation of far larger loops of around 400kbp in length which condensin I then divides into further smaller, 80kbp loops (Gibcus et al., 2018). Combined, a nested-loop structure is formed, promoting extensive chromosome compaction. Whilst it is not yet understood how condensin drives loop extrusion, *in vitro* single-molecule studies have demonstrated that the complex translocates along DNA, with ATP hydrolysis acting as a mechanochemical motor (Terakawa et al., 2017).

1.6.3. Cohesin (SMC1/3)

The cohesin complex uses its ring-shaped structure to encircle sister chromatids together (Haering et al., 2008), mediating their cohesion from DNA replication until their eventual separation during anaphase. In doing so, the complex prevents premature chromatid separation and ensures proper biorientation during segregation, thereby avoiding subsequent mis-segregation events and aneuplodies (Brooker and Berkowitz, 2014). However, the complex has also been implicated in the regulation of several other cellular processes, including condensation, transcription and DNA repair (Mehta et al., 2013). In human cells, cohesin is enriched proximally to DSBs, where the co-entrapment of sister duplexes promotes efficient HR whilst also preventing the joining of distal breaks (Potts et al., 2006). However, as with condesin, *in vitro* experiments have demonstrated that cohesin behaves as an ATPase-dependent loop extruding factor, during interphase (Davidson et al., 2019, Golfier et al., 2020). Whilst little is understood of cohesin's role in transcription, it may be that it

relates to the complex's ability to mediate physical interactions between distal intrachromosomal elements (Kagey et al., 2010, Mehta et al., 2013).

Cohesin is comprised of core SMC1 and SMC3 proteins, alongside the 'kleisin'-like protein RAD21 which links the ATPase heads (Brooker and Berkowitz, 2014). RAD21 acts as a docking protein for several other subunits, including WAPL, STAG1/2 and PDS5 (Figure 1.5) (Cheng et al., 2020). The concerted action of the NIPBL/MAU2 loading complex, alongside ATP hydrolysis, targets cohesin to chromatin and mediates the entrapment of DNA within the complex (Arumugam et al., 2003, Hinshaw et al., 2015, Collier et al., 2020).

Cohesin interacts with DNA throughout the cell cycle. During G1, cohesin loading is dynamic, with a large fraction of complexes remaining only transiently associated with DNA, presumably a single duplex. Cohesin is unloaded and reloaded repeatedly, associating with chromatin for less than 25 minutes at a time (Gerlich et al., 2006). This rapid dissociation is coordinated by the WAPL-PDS5 complex (Gandhi et al., 2006), possibly by promoting cohesin's ATP hydrolysis activity

and the opening of cohesin's DNA transport gate, facilitating DNA exit (Elbatsh et al., 2016). In order to establish sister chromatid cohesion during S-phase, however, the association of cohesin with DNA changes. Firstly, cohesin topologically entraps sister chromatids together. To achieve this, it's thought that some complexes are accommodated by the migration of the replication fork through its ring, whilst others are loaded *de novo*, unloading and reloading behind the replication machinery (Srinivasan et al., 2020). In addition to this, the cohesin complex become protected from WAPL's activity. The precise mechanism by which this happens is not yet known,

although experimental evidence suggests that this requires the establishment of cohesion 1 homologue 1 and 2 (ESCO1 and ESCO2) acetyltransferases, which acetylate SMC3 (Nishiyama et al., 2010, Kawasumi et al., 2017), as well as the S-phase expression of soronin, which prevents WAPL and PDS5 binding in vertebrates (Nishiyama et al., 2010, Ladurner et al., 2016). Together, this stabilises cohesin, allowing the complex to reside on chromatin for longer (Nishiyama et al., 2010, Ladurner et al., 2016, Kawasumi et al., 2017).

During mitosis, cohesin must be removed in a timely manner to enable proper chromosome segregation. In vertebrates, the majority of cohesin complexes are permanently removed during prophase via PLK1 and Aurora B mediated phosphorylation (Sumara et al., 2002, Gimenez-Abian et al., 2004, Nishiyama et al., 2013, Haarhuis et al., 2014). However, cohesin bound around heterochromatin, including centromeric regions, remains tightly associated, holding the sister chromatids together as they biorientate themselves (Gimenez-Abian et al., 2004, Haarhuis et al., 2014). Finally, during the metaphase-to-anaphase transition, the remaining cohesin is removed by the separase caspase (Hauf et al., 2001). This is triggered by the activity of the APC/C which signals for separase's inhibitory partner, secruin, to be degraded. Once active, separase cleaves RAD21, causing the cohesin ring to open and dissociate (Hauf et al., 2001, Luo and Tong, 2021). In contrast to this two-step removal of cohesin in vertebrates, in yeast most, if not all, cohesin is retained on chromosome arms until the onset of anaphase, when it is removed via separase (Haarhuis et al., 2014).

1.6.4. SMC5/6

Of the SMC complexes, SMC5/6 is the least well understood. The complex was first identified in fission yeast as an important factor in the repair of IR-induced DNA damage (Lehmann et al., 1995). However, since its discovery, research has implicated its involvement in numerous cellular processes. Its major function is thought to be in HR and replication, although it has emerging roles in the maintenance of rDNA and telomere elongation (Aragon, 2018).

The complex is comprised of a heterodimer of SMC5 and SMC6, alongside 4 non-SMC elements (NSE1-4). NSE4, the kleisin subunit which connects the complex's head domains, forms a subcomplex with NSE1 and NSE3. Whilst biochemical activities for NSE3 and NSE4 have not yet been identified (Aragon, 2018), NSE1 contains a RING-like domain characteristic of ubiquitin E3 ligases. However, evidence regarding the activity of the subunit is controversial, with no *in vivo* targets yet identified (Pebernard et al., 2008a, Doyle et al., 2010, Kolesar et al., 2022). In contrast, the SUMO E3 ligase activity of NSE2, which docks onto the middle of SMC5's coiled-coil arm, is well established, through which the complex can SUMOylate SMC6 and SMC5, whilst also targeting downstream substrates (Zhao and Blobel, 2005, Potts and Yu, 2005, Andrews et al., 2005, Kliszczak et al., 2012, Bonner et al., 2016). In yeast, additional subunits Nse5 and Nse6 form a heterodimeric subcomplex that interacts with the arms of Smc5 and Smc6 (Adamus et al., 2020), and is believed to promote the loading of Smc5/6 to DNA (Oravcova et al., 2019, Hallett et al., 2021b). Interestingly, yeast 3D modelling of Smc5/6 has shown that Nse2, Nse5 and Nse6 all bind in close proximity to Smc5 and Smc6's SUMOylated regions (Yu et al., 2021). A collaborative role for Nse5/6 and Nse2 is suggested by

the observation that mutations in Nse5's SUMO-interacting motifs are associated with decreased levels of Smc5 and Smc6 Nse2-dependent SUMOylation (Yu et al., 2021).

Whilst Nse1-4 are highly conserved across all kingdoms of life, Nse5 and Nse6 have only weak sequence similarity to their mammalian functional paralogs, SMC5/6 localisation factor 1 and 2 (SLF1 and SLF2, respectively), which were identified in a proteomic screen (Raschle et al., 2015). Using knockdown experiments, both SLF1 and SLF2 were shown to coordinate a linear pathway responsible for recruiting SMC5/6 to damaged chromatin. Upon DNA damage, it is known that ubiquitin ligases such as RNF8, MDC1 and RNF168, signal DNA lesions by catalysing the poly-ubiquitylation of histones (Stewart et al., 2009, Coster and Goldberg, 2010, Bohgaki et al., 2013, Uckelmann and Sixma, 2017). The UBZ domain of RAD18 recognises these modifications and is recruited to the damage. RAD18 then, via an interaction with SLF1, acts to recruit SLF2 and the rest of the SMC5/6 complex, enabling its subsequent interactions with DNA (Raschle et al., 2015).

As with condensin and cohesin, SMC5/6 interacts with DNA by encircling chromatin within its proteinaceous ring (Yu et al., 2022). Work in *Saccharomyces cerevisiae* has proposed that SMC5/6 works similarly to cohesin, by linking sister chromatids together in an intermolecular fashion to promote proper replication and segregation (Kanno et al., 2015). Although the precise mechanism by which this may occur isn't well studied, the topological interaction of Smc5/6 with DNA has been shown to be promoted by ATP binding and hydrolysis, with yeast Smc5/6 mutants lacking ATPase activity showing significantly reduced chromatin binding (Kanno et al., 2015, Etheridge et al., 2021). Nse5/6 has been shown to modulate this activity by either preventing the interaction between ATP and the SMC5/6 head domains or the

engagement of ATP bound heads. This inhibition is released when the complex interacts with DNA (Taschner et al., 2021). Interestingly, the SUMOylation activity of Nse2 also relies on ATP hydrolysis, which promotes the activity of the subunit via a conformational change in the Smc5 coiled-coil domains where the protein is docked (Bermudez-Lopez et al., 2015). Thus, the Nse5/6 heterodimer may regulate the interaction of Smc5/6 with DNA, as well as the complexes' SUMO ligase activities, by mediating timely ATP hydrolysis. However, analogous functions for SLF1/2 in ATP hydrolysis or NSE2-mediated SUMOylation have yet to be identified.

1.6.4.1. Functions of SMC5/6 in recombination and replication fork stability

Studies in yeast and human cell lines have demonstrated that defects in SMC5/6 are associated with elevated levels of unresolved DNA damage and chromosome breaks, suggesting that SMC5/6 is important for efficient DNA repair (Torres-Rosell et al., 2005). Early experiments in *Schizosaccharomyces pombe* suggested that the complex is epistatic with factors involved in HR, including rhp51^{RAD51}, which promotes homology searching. This is indicative of a role for the complex in HR, although the precise mechanism behind this function remains elusive (McDonald et al., 2003, Murray and Carr, 2008).

Epistasis experiments in both yeast and human cells have indicated that SMC5/6 and cohesin work in the same pathway (Potts et al., 2006, Ström et al., 2007). Although the loss of SMC5/6 does little to impact cohesin's genome wide localisation (Potts et al., 2006, Ström et al., 2007), the complex has been shown to promote the recruitment of cohesin to DSBs (Potts et al., 2006), where SMC5/6 and cohesin both localise during G2/M and are required for repair (Betts Lindroos et al., 2006). Experimental

evidence suggests that the interplay between these two complexes is mediated by SMC5/6's NSE2 SUMO-ligase, which SUMOylates cohesin's RAD21^{Scc1} subunit (Potts et al., 2006, McAleenan et al., 2012). Consistent with this, Nse2 knockout yeast exhibit decreased levels of Scc1^{RAD21} SUMOylation, as well a reduction in Scc1^{RAD21} binding to DNA damage (McAleenan et al., 2012). Thus, early models exploring SMC5/6's role in HR focus on the complex's function in promoting sister chromatid cohesion during HR strand invasion (Potts et al., 2006, Almedawar et al., 2012). In agreement with this, the loss of SMC5/6 complex components has been associated with hypo-recombination and enhanced gene targeting in several cellular models (Potts et al., 2006, De Piccoli et al., 2006). Furthermore, whilst Nse2's SUMO-ligase activity is not essential in yeast, SUMO mutants exhibit hypersensitivity to DNA damaging agents, slow growth, and elevated levels of mis-segregation, consistent with the SUMOylation activity of the complex being important for proper DNA repair (Stephan et al., 2011b). In fact, the loss of NSE2's SUMO-ligase functions is associated with the development of human disease, giving rise to a genome instability syndrome characterised by primordial dwarfism and insulin resistance (Payne et al., 2014).

These results, however, appear to conflict with additional studies which have observed elevated levels of SCEs in cell lines depleted of SMC5/6 components (Stephan et al., 2011a, Kliszczak et al., 2012). Whilst these findings may be explained by variations between model organisms and experimental set ups, research has suggested that they reflect an additional layer of NSE2-mediated regulation on HR. Alongside cohesin, Smc5/6's SUMO-ligase subunit regulates the activity of Sgs1–Top3–Rmi1 (STR) (BLM–Top3α–RMI1–RMI2 (BTR) in mammalian cells), a complex

that promotes dissolution during HR. In yeast, in response to the accumulation of recombination intermediates, Smc5 interacts with Sgs1, localising it to DNA, and SUMOylates all three STR subunits, stimulating the complex's enzymatic activities (Bonner et al., 2016, Bermudez-Lopez et al., 2016). Disruption of the Smc5-Sgs1 interaction, as well as the mutation of Sgs1's SUMOylation sites, results in the accumulation of X-shaped DNA structures, suggesting that the Nse2 SUMO ligase promotes the resolution of recombination via $Sgs1^{BLM}$ (Bermudez-Lopez et al., 2016, Bonner et al., 2016).

Smc5/6 has also been shown to be recruited to collapsed replication forks, where the complex is believed to promote fork rescue (Betts Lindroos et al., 2006). Consistent with this, the loss of S-phase Smc5/6 has been shown to be synthetic lethal/sick with factors that promote replication restart, including Mus81 and the STR complex, implicating the Smc5/6 in HR-mediated replication fork recovery (Menolfi et al., 2015). Furthermore, two-dimensional gel electrophoresis has demonstrated that, even under unperturbed conditions, yeast Smc5/6 mutants are slow growing and accumulate X-shaped DNA structures at collapsed replication forks, likely representing unresolved recombination intermediates, including reversed forks, dHJs and hemicatenanes (Branzei et al., 2006, Sollier et al., 2009, Agashe et al., 2021). During normal replication, recombinant structures promote the restart of the replication fork. However, if allowed to accumulate, these covalent linkages can escape the G2/M checkpoint and give rise to under-replicated DNA. Consistent with a collaborative or facilitatory role for Nse2 in Sgs1-mediated replication restart and repair at damaged replication forks, the association between Sgs1 and Smc5 increases upon S-phase

stalling, with Sgs1 SUMO mutants also accumulating X-shaped structures at damaged forks (Bermudez-Lopez et al., 2016).

As previously discussed, the genome contains numerous ‘difficult-to-replicate’ regions which the replication fork struggles to traverse. Studies have shown that Smc5/6 localises to many repetitive loci, including telomeres, centromeres, transfer DNA, and rDNA arrays (Torres-Rosell et al., 2005, Betts Lindroos et al., 2006, Menolfi et al., 2015), with the loss of Smc5/6 complex components associated with elevated levels of X-shaped structures and chromosome missegregation at these loci (Torres-Rosell et al., 2005, Betts Lindroos et al., 2006). Thus, it has been proposed that Smc5/6 may be important for the negative regulation of HR in certain contexts. At these ‘difficult-to-replicate’, repetitive regions, for instance, it appears that Smc5/6 not only promotes proper replication-linked recombination, but also ensures faithful mitosis by mediating sister chromatid interactions.

At rDNA specifically, fork stalling can be promoted by obstructive R-loops, replication-transcription collisions and RFBs (Kobayashi, 2003, García-Muse and Aguilera, 2016), which are established in yeast by the replication fork block protein Fob1 (Kobayashi, 2003). Fork stalling in these tandem repetitive regions can be particularly deleterious as copy number repeats can be lost during HR-mediated DSB repair (Warmerdam et al., 2016). Studies in *Saccharomyces cerevisiae* have demonstrated that Smc5/6 plays a critical role in the regulation of recombination at rDNA arrays, located on chromosome 12. Cells depleted of Smc5 and Smc6 via auxin inducible degron tag, which causes rapid protein degradation, exhibited a replication defect specifically in chromosome 12, in the first cell cycle (Peng et al., 2018). Loss of Smc5 and Smc6 is also associated with an increase in persistent recombination

intermediates at these sites, which can be rescued by the individual or combined deletions of Fob1 and Mph1, a prorecombinogenic factor (Peng et al., 2018). Together, this suggests that Smc5/6 is important for preventing potentially deleterious recombinogeneic events via Mph1 at chromosome 12's RFBs in yeast. Interestingly, the SUMO-independent interaction of Smc5 and Mph1 is also implicated in the suppression of recombination intermediates outside of rDNA arrays, as well as the promotion of proper centromeric segregation (Chen et al., 2009) Thus, it may be that Smc5/6 promotes proper replication via Mph1 in an Nse2 independent manner, with this function being particularly important at fragile sites. One hypothesis is that Smc5/6 regulates Mph1 upstream of Sgs1's functions in HR, by preventing the replication fork regression activities of the Mph1 helicase (Palecek, 2018).

Having said this, SMC5/6 has also been linked to other specific HR regulatory functions at rDNA. In yeast, for instance, the proper repair of rDNA DSBs has been shown to be promoted by the Smc5/6 dependent relocalisation of these breaks outside of the nucleolus, where rDNA is usually compartmentalised. Accordingly, mutations in Smc6 are associated with aberrant HR foci inside the nucleolus, as well as rDNA hyperrecombination (Torres-Rosell et al., 2007). Similarly, work in *Drosophila* has suggested that SMC5/6 promotes the relocalisation of heterochromatic DSBs to the nuclear periphery and away from repetitive sequences which may generate chromosome rearrangements upon aberrant HR (Chiolo et al., 2011, Ryu et al., 2015, Caridi et al., 2018). Thus, it may be that SMC5/6 has important roles in mediating repair by spatially regulating HR.

1.6.4.2. Functions of SMC5/6 at telomeres

SMC5/6 is enriched at some telomeres; repetitive nucleoprotein structures, 5-15kb long, which cap chromosome ends and protect them from aberrant degradation or repair (Torres-Rosell et al., 2005, Betts Lindroos et al., 2006, Moradi-Fard et al., 2016). In most germline cells, telomerase catalyses the extension of these regions to prevent their loss over successive rounds of division (Tomita, 2018). However, in certain tumours, including glioblastomas and osteosarcomas, telomeric repeats are maintained by an alternate pathway, known as the alternative lengthening of telomeres (ALT) (Dilley and Greenberg, 2015). When ALT is activated, telomeres become localised to ALT-associated promyelocytic leukaemia (PML) bodies (APBs), where they are thought to be extended via HR between telomeric repeats. This generates considerable heterogeneity in chromosome lengths (Cesare and Reddel, 2010).

SMC5/6 is implicated heavily in the regulation of ALT, where the complex promotes the association between telomeres and PML bodies (Chung et al., 2011). The complex has been shown to localise to these bodies, where NSE2 promotes the SUMOylation of several telomeric proteins, including TRF1 and TRF2. Mutation of either protein's SUMOylation sites prevents NSE2-mediated SUMOylation and inhibits the recruitment of TRF1 and TRF2 to PML bodies, where they are required for APB formation (Potts and Yu, 2007). Accordingly, the loss of SMC5, SMC6 or NSE2 leads to reduced levels of sister chromatid exchange events in ALT-positive MEF and SUSM1 cell lines. This is associated with increased levels of end-to-end fusions and telomere shortening (Potts and Yu, 2007).

Whilst the function of SMC5/6 at telomeres in ALT-negative cells is less understood, Nse2 mutants in *Saccharomyces cerevisiae* demonstrate abnormal telomere clustering and nucleolar structures, suggesting that SMC5/6-mediated SUMOylation may be important for telomere organisation in noncancerous cells, also (Zhao and Blobel, 2005, Moradi-Fard et al., 2016). In addition to this, a recent BioID study using ALT-negative HEK293 cells has demonstrated that SLF2 interacts with ATRX, a chromatin remodeller required for the deposition of histones at telomeres, through which it mediates telomere replication and HR events (Scott et al., 2021). Considering that SLF2 and SMC5's subcellular localisation overlaps strongly at telomeres (Scott et al., 2021), this may suggest that SMC5/6 has emerging functions in regulating recombination in these regions in telomerase positive cells. However, whether this relates to SMC5/6's putative functions in the nuclear organisation of telomeres is unknown.

1.6.5. Pathogenic diseases associated with the SMC complexes

1.6.5.1 Diseases associated with condensin

Autosomal recessive primary microcephaly describes a rare genetic condition characterised by a significant reduction in head circumference (less than 3 SDs below that of the age and sex matched mean), which can present alongside a broad spectrum of additional, typically neurodevelopmental, phenotypes, including seizures and mental retardation. Microcephaly is associated with defects in a number of genes required for faithful replication and chromosome segregation (Jayaraman et al., 2018, Tingler et al., 2022), although pathogenic mutations that cause microcephaly were first identified in Microcephalin (MCPH1) (Jackson et al., 1998). Since this time,

MCPH1 has been shown to be a negative regulator of condensin II (Trimborn et al., 2006), binding to the complex's CAPG2 subunit and regulating the SMC2-CAPH2 interface to prevent aberrant association of condensin II with DNA (Houlard et al., 2021). Through this, MCPH1 can limit chromosome condensation until prophase and mediate timely decondensation post mitosis. Accordingly, metaphase spreads generated from patients with MCPH1-type microcephaly exhibit premature chromosome condensation (Trimborn et al., 2004).

Further cases of microcephaly and neurodevelopmental disorder have also been reported in patients with *de novo* or autosomal recessive deletions across condensin I or II component genes, including CAPH, CAPD2, CAPH2, CAPD3 and CAPG2 (Chen et al., 2010, Martin et al., 2016, Khan et al., 2019). Similar to dysfunctions in MCPH1, defects in these genes impair chromosome compaction, leading to decatenation failure and errors during chromosome segregation. Together, these dysfunctions have been shown to impact cell division, reducing cell proliferation and increasing cell death in the cerebral cortex, resulting in reduced brain size (Martin et al., 2016). Interestingly, premature chromosome condensation associated with defects in MCPH1 may also be associated with the disruption of gene expression during G2, also impeding proper brain development (Pang et al., 2022).

1.6.5.2 Diseases associated with cohesin

Mutations in cohesin and cohesin-related genes lead to pathologies collectively referred to as cohesinopathies (Piché et al., 2019). Of these, the most well characterised is Cornelia de Lange syndrome, caused primarily by mutations in NIPBL, a component of the cohesin loading complex (Hinshaw et al., 2015), although

defects in RAD21, SMC3 and SMC1 have also been associated with Cornelia de Lange syndrome-like probands (Ansari et al., 2014). Patients present with microcephaly, short stature and mental retardation, alongside other varying phenotypes including frequent limb, craniofacial, heart and gastrointestinal defects (Kline et al., 2018, Piché et al., 2019). Similarly, Roberts syndrome is caused by mutations in ESCO2, an acetyltransferase which promotes the retention of cohesin on chromatin. Like Cornelia de Lange syndrome, patients present with severe growth deficiency, microcephaly, mental retardation and limb malformations, although heart and gastrointestinal defects are not usually associated with the condition (Vega et al., 2005b, Piché et al., 2019).

Less well characterized cohesinopathies include Warsaw Breakage Syndrome (WABS), caused by mutations in DDX11, a helicase which promotes the association of cohesin with chromatin during replication (van Schie et al., 2020, Faramarz et al., 2020), and facilitates DSB repair via HR (Abe et al., 2018). Individuals with WABS present similarly to those with Roberts and Cornelia de Lange syndrome, although cell lines derived from these patients also exhibit extreme chromosome breakage in response to treatment with mitomycin C (MMC), a DNA cross-linking agent (van der Lelij et al., 2010a, Piché et al., 2019). This cellular phenotype is unique among cohesinopathies, instead comparable to patients with mutations in FA pathway proteins, and likely reflects DDX11's additional functions in DNA repair (van der Lelij et al., 2010b, Piché et al., 2019).

Consistent with cohesion dysfunction, cell lines derived from patients with Roberts syndrome exhibit railroad chromosomes (Schulz et al., 2008, da Costa Almeida et al., 2020), in which the centromere prematurely separates, alongside mitotic failure and

decreased proliferation. Similarly, Cornelia de Lange syndrome and WABS are associated with premature chromatid separation, in which cohesion is prematurely lost from both the centromeres and arms (Kaur et al., 2005, van Schie et al., 2020). However, as of yet there is no direct evidence that cohesion defects lead to the clinical phenotypes observed in these syndromes. Considering that cohesion has important functions in both DSBR and transcription, it may be that these clinical phenotypes arise due to defects in other cellular pathways (Piché et al., 2019). For instance, in the case of Cornelia de Lange syndrome, genome-wide transcriptional microarrays of patient cell lines have uncovered changes in the expression of specific genes (Liu et al., 2009, Garcia et al., 2021), with animal modelling implicating transcriptional changes in disease progression (Kawauchi et al., 2016). These observations have led to claims that Cornelia de Lange syndrome should be redefined as a transcriptomopathy, and question whether similar cellular defects drive the disease phenotype of other conditions currently classed as cohesinopathies.

1.6.5.3 Diseases associated with SMC5/6

Currently, there are two known syndromes associated with the SMC5/6 complex. The first has been observed in two patients with compound heterozygosity for frameshift mutations in NSE2, giving rise to a disease characterised by primordial dwarfism, insulin resistance and primary gonadal failure. Both patients, who exhibited a dramatic decrease in NSE2 protein expression, showed a mild increase in SCEs following UV irradiation of patient LCLs, and displayed HU induced nuclear abnormalities which could be rescued upon re-expression of the wild-type protein (Payne et al., 2014). As is well documented for cases such as ATR-seckel syndrome, diseases associated with increases in replication stress are frequently accompanied by global growth

defects, with slower replication and increased cell death. Thus, these patient phenotypes are consistent with the proposed role for SMC5/6-mediated SUMOylation activity in replication and fork restart. Increases in SCEs in patient cell lines has also mimicked previously explained research implicating SMC5/6 SUMOylation in limiting hyper-recombination (Payne et al., 2014).

The second of the known SMC5/6-linked syndromes arises from mutations in NSE3, a member of the melanoma-associated antigen (MAGE) protein family with unknown functions. Nine patients with homozygous and compound heterozygous NSE3 mutations have been identified, all of whom suffered from immune deficiencies and paediatric pulmonary disease (van der Crabben et al., 2016, Willemse et al., 2021). Cell lines derived from these patients exhibited significantly reduced expression of NSE3, as well as SMC5 and SMC6, thus destabilising the complex. In the patient fibroblasts, this was associated with elevated markers of DNA damage, alongside a delay in recovery from S-phase damage (van der Crabben et al., 2016).

General aim of thesis

The major aim of this thesis was to question how the SMC5/6 complex promotes genome stability and prevents human disease. Prior to this study, SMC5/6 had only been linked with the development of two genetic syndromes with very different clinical presentations, caused by mutations in NSE2 and NSE3. Firstly, by identifying and characterising mutations in SLF2 and SMC5 in patients presenting with microcephaly, short stature, cardiac abnormalities and anaemia, I hoped to further understand the clinical spectrum of syndromes associated with SMC5/6, as well as the significance of these specific factors to normal human development. The SMC5/6 complex is thought to have important cellular functions in replication fork stability and HR-mediated repair, as well as emerging roles in telomere maintenance. Therefore, my second aim was to investigate how the loss of SLF2 and SMC5 impacts these putative functions. This was carried out by studying the cellular phenotypes of fibroblast and lymphoblastoid cell lines derived from SLF2 and SMC5 patients, as well as SLF2 U-2-OS CRISPR mutant cells. Finally, I aimed to investigate a possible function for SMC5/6 in promoting replication through certain 'difficult-to-replicate' regions. Although RNF168, a ubiquitin ligase involved in the recruitment of SMC5/6 to damaged chromatin, has been implicated in signalling the presence of stabilised G4-quadruplexes, a role for SMC5/6 in relation to these structures has yet to be shown. To test for this, the G4-stabilising drug CX-5461 was used to investigate the impact of persistent G4-quadruplexes on replication and chromosome stability.

Chapter Two:

Materials and methods

2.1 Mammalian cell culture

2.1.1 Maintenance of cell lines

All cell lines were maintained at 37°C and 5% CO_2 in a humidified tissue culture incubator. Lymphoblastoid cell lines (LCLs) were cultured in suspension in RPMI-1640 with L-glutamine (Gibco, 21875034), supplemented with 10% Fetal Bovine Serum (FBS) (Gibco, 10270106) and 5% Penicillin/Streptomycin (PS) (Sigma-Aldrich, P4333). Fibroblasts immortalised with human telomerase reverse transcriptase (hTERT) were grown and maintained in DMEM with 4.5g/L D-glucose, L-glutamine and pyruvate (Gibco, 41966029) supplemented with 20% FBS and 5% PS. Both RPE1 and U-2-OS cell lines were obtained from ATCC. RPE1s were cultured in DMEM with 4.5g/L D-glucose, L-glutamine and pyruvate supplemented with 10% FBS and 5% PS, whilst U-2-OS cells were cultured in McCoys 5A (modified) media with L-glutamine supplemented with 20% FBS and 5% Gibco PS.

Complementation of fibroblast and U-2-OS cell-lines was carried by re-expressing Myc-tagged WT-SLF2 or untagged WT-SMC5 using the pLVX-IRES-Neo lentiviral vector (Takara Bio).

For routine cell culture maintenance and harvesting cells for experiments, adherent cells were disassociated from tissue culture plasticware by trypsination using TrypLE™ Express trypsin (Thermofisher Scientific, 12605010). Cell concentrations of LCLs and U-2-OS cells were determined using the Countess™ 3 Automated Cell Counter (Invitrogen). For human fibroblasts, cell concentrations were calculated using a hemocytometer. For experiments with exogenous DNA damage treatments, patient-derived cells were incubated with Mitomycin-C (MMC) (Sigma-Aldrich, M4287), CX-

5461 (Selleck Chemicals, S2684), Wiskostatin (Abcam, 141085), aphidicolin (APH) (Sigma-Aldrich, A4487), or BMH21 (Bio-Techne, 5417), at the concentration stated in the figure legends.

For long term storage of cell lines, cells were pelleted by centrifugation at 1,200 rpm and resuspended at around 5.0- 7.5x10^6 cells/ml in FBS containing 10% DMSO (Sigma-Aldrich, D650). 1ml aliquots of the cellular suspension were transferred to cyrovials and stored short term at -80°C using a Mr. Frosty™ freezing container (Thermofisher Scientific, 5100-0001) for at least 24 hours before permanent storage in liquid nitrogen. To thaw cells after storage in liquid nitrogen, cryovials were warmed in a 37°C water bath. The cell suspension mix was then added to 10ml of appropriate media before being pelleted by centrifugation at 1,200 rpm and resuspended in fresh media.

2.1.2 Plasmid transfection

For transfection of plasmid DNA into cells, 2.5x10^5 cells were first seeded into 6cm dishes. After 24 hours, cells were transfected in 1ml of antibiotic-free media with 9µg of DNA in a 3:1 ratio of FuGENE® HD Transfection Reagent (Promega, E2311) to DNA, prepared in 1ml of Opti-MEM reduced serum medium (Gibco, 31985062). After 6 hours, the Opti-MEM was replaced with fresh media. Cells were sorted around 48 hours later.

2.1.3 Fluorescence-activated cell sorting

Cells to be sorted based on the presence of GFP expression were first resuspended in PBS with 1% FBS at a concentration of approximately 1x10^6 /ml. The cells were then filtered through a 50µM mesh (CellTrics, 04-0042-2317) and sorted into 96-well

dishes based on GFP expression using the BD FACS Aria Cell-Sorting System. U-2-OS cells were recovered in McCoys 5A (modified), and RPE1s with DMEM with 4.5g/L D-glucose, both supplemented with 20% FBS and 5% Gibco PS.

2.2 Chromosome analysis

2.2.1 Metaphase Spreads

LCLs were seeded at a density of $5x10^5$ cells/ml in a volume of 10ml of media. For the analysis of human fibroblasts or U-2-OS cell lines, $1x10^6$ cells were seeded in a 10cm tissue culture dish. 48 hours after seeding, karyoMAX colcemid (Gibco, 15212-012) was added at a final concentration of 0.2 µg/ml for 4 hours. Following this incubation, LCLs were collected and adherent cells were harvested via trypsinisation. All types of cells were pelleted by centrifugation at 1,200 rpm before being resuspended with PBS and pelleted again. The cells were then subjected to hypotonic shock by gently resuspending with hypotonic buffer (10 mM KCl, 15% FCS) and incubating at 37˚C for 30 minutes. Cells were fixed using methanol/acetic acid (ratio 3:1) fixative solution. First, 1ml of methanol/acetic acid fixative solution was added to the hypotonic cell suspension, which was then centrifuged at 1,200 rpm. Most of the supernatant was removed, leaving the pellet in approximately 1ml of solution, which was resuspended gently with approximately 10ml of methanol/acetic acid fixative solution. This cell suspension was again centrifuged at 1,200 rpm and resuspended with another approximately 10ml of methanol/acetic acid fixative solution. Finally, the cell suspension was pelleted and stored in 1ml of methanol/acetic acid fixative solution at -20˚C for at least 24 hours.

To prepare metaphase spreads, fixed cell suspensions were dropped onto glass microscope slides and allowed to air dry overnight, before being stained for 15 minutes in Giemsa-modified solution (5% v/v in ddH_2O; Scientific Laboratory Supplies GS500) and destained in water for 5 minutes. After allowing to dry overnight, coverslips were mounted onto the microscope slides using Entellan™ mounting medium (Sigma-Aldrich, 107961). The slides were analysed via brightfield microscopy using a Nikon Eclipse Ni microscope and a 100x oil immersion objective.

For the analysis of cohesion fatigue in LCLs, the metaphase spread protocol was followed as above. However, instead of using karyoMAX colcemid, 25µM MG132 (Sigma-Aldrich, M7449) was added 4 hours before harvesting.

To prepare metaphase spreads from peripheral blood samples, whole blood was diluted in RPMI 1640 and 180µg/ml phytohaemagglutinin (Thermo Fisher Scientific, R30852701) was added for 48-72 hours at 37°C. The metaphase spread protocol was followed as above, although hypotonic shock was reduced to a 10-minute 37°C incubation using 75mM KCl as hypotonic buffer.

2.2.2 Sister Chromatid Exchange Analysis

To study sister chromatid exchange events, LCLs were seeded at $5x10^5$ cells/ml and incubated with 10 µM BrdU (Thermo Fisher Scientific, B23151) for 48 hours (two cell population doubling times). 3 hours prior to harvesting, karyoMAX colcemid (Gibco, 15212-012) was added to LCLs at a final concentration of 0.2µg/ml. Following this, cells were pelleted by centrifugation at 1,200 rpm and washed with PBS. After being spun at 1,200 rpm, cells were gently resuspended in 75mM KCl hypotonic buffer, to a total volume of approximately 10ml and incubated at 37˚C for 20 minutes. After

incubation in hypotonic buffer, cells were spun at 1,200 rpm and were fixed by resuspending in ethanol/acetic acid (ratio 3:1) fixative solution. Fixed cell suspensions were pelleted by centrifugation at 1,200 rpm, resuspended in approximately 10ml of ethanol/acetic acid fixative solution, before being pelleted again. Fixed cell suspensions were stored at -20˚C in 1ml of in ethanol/acetic acid fixative solution.

After being at -20˚C for at least 24 hours, metaphase spreads were prepared by dropping the cell suspension onto glass microscope slides and allowed to air dry overnight. To differentially stain the two sister chromatids, slides were first incubated in Hoescht 33258 (10µg/ml in ddH_2O; Sigma-Aldrich-Aldrich, 861405) for 20 minutes followed by a 5 min incubation in 20x SSC buffer (3M NaCl; 300mM trisodium citrate, pH7) and then exposure to UVA light for an hour. Slides were then incubated in 20x SSC buffer for 1 hour at 60˚C, washed in water, and stained for 15 minutes in Giemsa-modified solution (5% v/v in ddH_2O; Scientific Laboratory Supplies, GS500) and destained in ddH_2O for 3 minutes.

2.2.3 Fluorescence in situ hybridization

Centromeric fluorescence in situ hybridization (FISH) was carried out on metaphases prepared from peripheral blood lymphocytes harvested as described above. However, after being dropped onto glass microscope slides and air dried for 24 hours, the microscope slides were rehydrated with PBS, dehydrated in a series of ethanol dilutions (70%, 90%, 100%) and air dried again. The slides were prewarmed to 37˚C and then incubated with CENPB-Alexa488-conjugated PNA probe (PNA Bio, F3001; 1:1000) made up according to the manufacturer's instruction in hybridization solution (20 mM Tris, pH7.4, 60% formamide, 0.5% blocking reagent [Roche, 11096176001]). Incubation with the PNA probe was conducted first at 85°C for 10 minutes before the

slides were transferred to a dark, humidified chamber for a further 2 hours. The slides were then washed twice in wash buffer (70% formamide, 10mM Tris), each time for 15 minutes with gentle agitation at room temperature. This was followed by three PBS washes, each for 5 minutes with gentle agitation at room temperature. The slides were then mounted using ProLong™ Gold Antifade Mountant (ThermoFisher Scientific, P36930) and sealed with nail varnish. Images were visualised using a Nikon Eclipse Ni microscope with NIS-Elements software (Nikon Instruments) and captured using a 100x oil immersion objective.

2.3 Analysis of cellular extracts by immunoblotting

2.3.1 Preparation of total cellular protein extracts

Whole cell extracts were obtained by lysis cell pellets in either RIPA buffer (150mM NaCl, 5mM EDTA, 50mM TRIS, 1% NP-40, 0.5% Sodium deoxycholate, 0.1% SDS) or UTB (8M Urea, 50mm TRIS, 150mM B Mercaptoethanol). For lysis in RIPA buffer, cell pellets were resuspended and then rotated for 30 minutes at 4˚C. For lysis in UTB buffer, cell pellets were resuspended and sonicated. In both cases, samples were then centrifuged at 16,000 rpm for 20 minutes. The supernatant was then retained, and its protein concentration determined by Bradford Assay (Bio-Rad, #5000006).

2.3.2 SDS- Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Protein samples were denatured at 95 ˚C in Laemmli sample buffer (2% SDS, 10% glycerol, 5% 2-mercaptoethanol, 0.002% bromphenol blue and 0.0625 M Tris HCl, pH approx. 6.8) for 5 minutes. Denatured samples were loaded onto an 8% SDS polyacrylamide gel (44ml final volume: 11ml 30% acrylamide, 25ml ddH2O, 4ml 1M Tris/Bicine, 1ml 10% SDS, 0.1ml TEMED, 0.2ml 10% APS) and subjected to

electrophoresis. A full range Amersham™ ECL™ Rainbow™ Marker (Sigma-Aldrich-Aldrich, GERPN800E) was used as a protein ladder.

2.3.3 Immunoblot analysis

Following SDS-PAGE, proteins were transferred onto a nitrocellulose membrane via wet transfer using a TE42 Standard Transfer Tank (Hoefer) in transfer buffer (20mM Tris, 150mM glycine, 10% v/v methanol) for 18 hours. Non-permanent staining with 0.1 % (w/v) Ponceau S in 5% acetic acid (Biotium, 22001) was used to verify protein transfer before destaining with a 5 minute wash in 1x TBST (10mM Tris, 140mM sodium chloride, 0.1% v/v Tween 20, pH 7.9). The nitrocellulose membrane was then subjected to blocking by incubation with 5% (w/v) milk in 1x TBST for an hour. Membranes were then incubated with the appropriate primary antibody overnight with gentle agitation (Table 2.1). All antibodies were diluted in 5% (w/v) milk in 1x TBST. Following incubation in primary antibody, the membrane was washed three times in 1x TBST for 5 min each, and then the membrane was incubated with the relevant HRP-linked secondary antibody for 1hr at room temperature (Table 2.1). Following three further 5-minute washes in 1x TBST, the membrane was incubated in Immobilon Western HRP Substrate (Millipore, WBKLS0500) for 1 minute and

Antibody (And host animal)	Purpose	Concentration	Manufacturer
Primary antibodies			
anti-SLF2 (Rabbit)	Immunoblotting	1:1,000	Gifted by Niels Mailand
anti-NBS1 (Mouse)	Immunoblotting	1:10,000	Genetex: GTX70224
anti-GAPDH (Rabbit)	Immunoblotting	1:1,000	Genetex: GTX100118
anti-Myc (Mouse)	Immunoblotting	1:1,000	BD Biosciences: 551101
anti-Mitosin (Mouse)	Immunofluorescence	1:500	BD Biosciences: 610768
anti-53BP1 (Rabbit)	Immunofluorescence	1:1,000	Novus Biologicals: NB100-304
anti-CENPA (Mouse)	Immunofluorescence on interphase cells and metaphase spreads	1:750	Abcam: Ab13939
anti-α-Tubulin (Mouse)	Immunofluorescence	1:4,000	Sigma: B-5-1-2
anti-PLK1 (Mouse)	Immunofluorescence	1:200	Abcam: Ab17056
anti-RAD51 (Rabbit)	Immunofluorescence	1:500	Merck: PC130
anti-Sodium/potassium-transporting ATPase (Rabbit)	Immunofluorescence	1:500	Abcam: Ab76020
anti-SMC2 (Rabbit)	Immunofluorescence on metaphase spreads	1:200	Novus Biologicals: NB100-373
anti-Aurora B (Rabbit)	Immunofluorescence on interphase cells	1:500	Bethyl Laboratories: A300-431A

	and metaphase spreads		
anti-PCNT1 (Rabbit)	Immunofluorescence	1:100	Abcam: Ab4448
anti-TopIIα (Rabbit)	Immunofluorescence on metaphase spreads	1:200	Merck: SAB4502998
anti-BrdU (recognises CldU) (Rat) Clone BU1/75, ICR1	DNA fibres	1:500	Abcam: Ab6326
anti-BrdU (recognises IdU) (Mouse) Clone B44, BD	DNA fibres	1:500	Biosciences: 347583
Secondary antibodies			
anti-Rabbit-HRP	Immunoblotting	1:3000	Agilent Dako: P0260
anti-Mouse-HRP	Immunoblotting	1:1000	Agilent Dako: P0447
anti-Rabbit- Alexa Fluor 488	Immunofluorescence on interphase cells and metaphase spreads	1:1000	Invitrogen: 11070
anti-Mouse- Alexa Fluor 594	Immunofluorescence on interphase cells and metaphase spreads	1:1000	Invitrogen: 11032
anti-Rat- Alexa Fluor 555	DNA fibres	1:500	Invitrogen: 21434
anti-Mouse- Alexa Fluor 488	DNA fibres	1:500	Invitrogen: 11029

Table 2.1

The source, type and concentration of antibodies used for immunoblotting and immunofluorescence.

signal was visualised by exposure to Amersham Hyperfilm MP (GE Healthcare, 28906845).

2.4 Immunofluorescence analysis

2.4.1 Immunofluorescence analysis of interphase and mitotic cells

For the analysis of adherent cells, such as human fibroblasts or U-2-OS CRISPR cell lines, $1x10^6$ cells were seeded in a 10cm dish onto glass coverslips and harvested 48 hours later. 22x22mm square coverslips were used for the analysis of interphase cells and Poly-L-Lysine coated 12mm round coverslips (Scientific Laboratory Supplies Ltd, 354085) were used for the analysis of mitotic cells. Interphase cells were pre-extracted for 5 minutes with ice cold extraction buffer (25 mM HEPES, pH 7.4, 50 mM NaCl, 1 mM EDTA, 3 mM MgCl2, 300 mM sucrose and 0.5% Triton X-100) and fixed with 4% paraformaldehyde in PBS for 10 minutes. For analysis of mitotic cells in adherent cell lines, cells were fixed with ice cold 100% methanol for 15 minutes.

For the analysis of LCLs, approximately 200µl of a $1x10^6$ cells/ml suspension was pipetted onto Poly-L-Lysine coated coverslips (Scientific Laboratory Supplies Ltd, 354085) where cells were allowed to adhere to the coverslips for 20 minutes before being fixed with methanol, as above.

In all cases, after fixation the coverslips were washed for with PBS three times, and then blocked for 1 hour with 10% FBS (Gibco, 10270106) in PBS at room temperature. The coverslips were then incubated in the appropriate primary antibodies, diluted in 2% FBS in PBS (as detailed in Table 2.1), for 1 hour being washed three times with PBS. Coverslips were then incubated with the relevant

secondary antibodies conjugated to either Alexa Fluor-488 or -594, diluted in 2% FBS in PBS (Table 2.1). The coverslips were washed a further three times in PBS before being mounted onto microscope slides using Vectashield Antifade Mounting Medium with DAPI (Vectorlabs, H-1200-10) and sealed using nail varnish. Microscope slides were viewed using fluorescent microscopy using a Nikon Eclipse Ni microscope with NIS-Elements software (Nikon Instruments), and images were captured using either a 60x or a 100x oil immersion objective and analysed using FIJI version 2.1.0/153.c (Schindelin et al., 2012).

2.4.2 Immunofluorescence microscopy of metaphase spreads

To perform immunofluorescent imaging of chromosomes within metaphase spreads, LCLs were first incubated with 0.2 µg/ml KaryoMAX colcemid (Gibco, 15212-012) for 3 hours. Cells were then harvested by centrifugation at 1,200 rpm and subjected to hypotonic shock with 75mM KCl for 20 minutes at 37 ˚C. The cells were resuspended at approximately $5x10^5$ cells/ml in PBS and 125µl of the cell suspension was spun onto glass coverslips at 1150 rpm for 10 minutes using the Thermo Scientific™ Cytospin™ 4 Cytocentrifuge. The cells were immediately incubated in 1x KCM buffer (100mM KCl, 30mM $CaCl_2$, 50mM $MgCl_2$) for 10 minutes before being fixed with 100% methanol for 20 minutes, permeabilised with 0.5% triton in PBS for 20 minutes and finally blocked in 3% BSA in PBS, all on ice. Antibody staining and PBS washes were carried out as for interphase and mitotic cells as detailed in the previous section, although incubation with primary antibodies was carried out overnight (Table 2.1). Coverslips were mounted onto microscope slides using ProLong™ Gold Antifade Mountant (ThermoFisher Scientific, P36930) and sealed with nail varnish. Images were visualised in the same manner as above.

2.5 Molecular cloning

2.5.1 sgRNA design

In order to design guide RNAs for use in the CRISPR-Cas-9 system, the CHOPCHOP v2 web tool was first used to locate pairs of CRISPR target sites within SLF1 and SLF2 genes (Labun et al., 2019). Using these sequences, CLC Main Workbench 7 (QIAGEN) was then used to design guide oligonucleotides which would target these sites (Table 2.2). The complementary forward and reverse oligonucleotides were purchased from Sigma-Aldrich-Aldrich and annealed at 16˚C. The oligonucleotides designed for SLF1 and SLF2 targeting are detailed in Table 2.

2.5.2 Ligation

Pairs of CRISPR Cas-9 sgRNA oligos were ligated into the GFP All-in-One Cas9^{D10A} nickase vector (Gifted from Prof. Stephen Jackson) separately, at the plasmid's unique BbsI and BsaI restriction sites. 1µg of plasmid DNA was digested using either BbsI (NEB, R0539S) or BsaI-HF®v2 (NEB, R3733S) restriction enzymes at 37°C for 1 hour, and then incubated with 1U of calf-intestinal alkaline phosphatase (CIP) (ThermoFisher Scientific, 18009019) for 1 hour to prevent the cut vector from self-ligating. The CIP was heat inactivated at 80°C for 15 minutes and the vector was purified using the QIAquick PCR Purification Kit (QIAGEN, 28104). sgRNA oligonucleotides were hybridised together by heating the forward and reverse oligos at 95°C for 5 minutes, followed by slow cooling to room temperature. The sgRNAs were then ligated into the GFP All-in-One Cas9D10A nickase vector using a 1:3 molar ratio of vector to insert. 50ng of plasmid was incubated with the appropriate amount of sgRNA oligo and 1U of T4 DNA Ligase (NEB, M0202S) in a total volume of 20µl overnight at room temperature.

Target and Construct	Oligonucleotide	Sequence (5'-3')
SLF1-CRISPR plasmid 1 (1.1)	sgRNA 1 Forward	ACCGGACAACTTTCCATCTGTGGA
	sgRNA 1 Reverse	AAACTCCACAGATGGAAAGTTGTC
	sgRNA 2 Forward	ACCGAGCGAAGTGATTCTCTTATA
	sgRNA 2 Reverse	AAACTATAAGAGAATCACTTCGCT
SLF1-CRISPR plasmid 2 (1.2)	sgRNA 1 Forward	ACCGTGTCAGTCACAAAGGAATTT
	sgRNA 1 Reverse	AAACAAATTCCTTTGTGACTGACA
	sgRNA 2 Forward	ACCGAGGAAAGTGGATACTAACCA
	sgRNA 2 Reverse	AAACTGGTTAGTATCCACTTTCCT
SLF2-CRISPR plasmid 1 (2.1)	sgRNA 1 Forward	ACCGAGTTTCATCACTCGGTTCCT
	sgRNA 1 Reverse	AAACAGGAACCGAGTGATGAAACT
	sgRNA 2 Forward	ACCGGGCTTGGCACCTTCAAATTC
	sgRNA 2 Reverse	AAACGAATTTGAAGGTGCCAAGCC
SLF2-CRISPR plasmid 2 (2.2)	sgRNA 1 Forward	ACCGCCGATTCTAATGTATCTTCA
	sgRNA 1 Reverse	AAACTGAAGATACATTAGAATCGG
	sgRNA 2 Forward	ACCGGTGCCATATTCTGAGCGCAA
	sgRNA 2 Reverse	AAACTTGCGCTCAGAATATGGCAC

Table 2.2

The DNA sequence of complimentary forward and reverse oligonucleotides used to generate CRISPR nickase vectors targeting either SLF1 or SLF2.

2.5.3 Transformation into chemically competent bacteria

1-10ng of plasmid DNA was added to 50µl of chemically competent 5-alpha *E. coli* (NEB, C2987H) cells which were placed on ice for 30 minutes. The cells were then heat shocked at 42°C for 45 seconds before being placed back onto ice for a further 5 minutes. The cells were then incubated with 950µl SOC in a 37 °C shaking incubator for 1 hour. 20-200µl of the transformed mixture was spread onto warmed LB agar plates containing the relevant antibiotic and the plates were incubated overnight 30°C.

After transformation, single bacterial colonies were picked from LB agar plates. To generate a master stock of plasmid, each colony was grown in 200ml LB overnight in a 37°C shaking incubator and plasmid DNA was purified using the PureLink™ HiPure Plasmid Maxiprep Kit (Invitrogen, K210006). To prepare plasmids for sequencing, individual colonies were grown in 5ml LB overnight in a 37°C shaking incubator before purification with the QIAprep Spin Miniprep Kit (QIAGEN, 19064). DNA concentration was determined using the UV5Nano Spectrophotometer (Metter Toledo).

2.5.4 Genomic DNA extraction

Cells were lysed in lysis butter (100 mM Tris/HCl pH 8.5, 5 mM EDTA, 0.2% SDS, 200 mM NaCl, 100 µg Proteinase K/ml) and the DNA was precipitated with isopropanol before being resuspended in 10 mM Tris/HCl, 0.1 mM EDTA, pH 7.5.

2.5.5 PCR analysis

SLF2 fragments were amplified from genomic DNA using Q5® Hot Start High-Fidelity Master Mix (NEB, M0494S). 20ng of template DNA was amplified by PCR using 0.5µM of relevant primers (Table 2.3) with an extension time of 1 min for 45 cycles.

2.5.6 RT-PCR analysis

Total RNA was extracted from cell lines using RNeasy Mini kit (Qiagen, 74004) according to the manufacturer's instructions. DNA was removed by treatment with DNase I (Qiagen, 74004), and cDNA was generated using Superscript II (Invitrogen, 18064071), primed with oligo-dT (Invitrogen, 18418012). RT-PCR was carried out with relevant primers using Phusion Hot Start II (Thermo Life Science, F549L) (Table 2.3).

2.5.7 DNA agarose gel electrophoresis

DNA agarose gels were prepared using 1% w/v agarose (Sigma-Aldrich-Aldrich, A9539) in 1x TAE buffer (Invitrogen, 24710030) with 0.5µg/ml SYBR Safe stain (Invitrogen, S33102). DNA samples were made up with DNA gel loading dye (Thermo Scientific, R0611) and ddH_2O, loaded onto the gel alongside a 1Kb Plus DNA Ladder (Invitrogen, 10787018), and subjected to electrophoresis. DNA was then visualised using UV transillumination.

2.5.8 DNA sequencing

Genomic, cDNA and plasmid DNA was sequenced by Source BioScience, using standard or specific primers, as detailed in Table 2.3. Specific primers were designed using NCBI PrimerBlast alongside CLC Main Workbench 7 (QIAGEN) and purchased from Sigma-Aldrich.

Sequencing primers	
SLF2 cloning primers	
Forward	5’-AACCGCUAGCUCCACGACAAA-3’
Reverse	5'-CAAGACGTGATTCATGAAGTA-3'
SMC5 cloning primers	
Forward	5'-GTTGTTGCGGCCGCGCCACCATG GCGACTCCGAGCAAGAAGACGTC-3’
Reverse	5’-GTTGTTGCGGCCGCAAGAAGCTT GAGTGAATGTAATACGGCGGC-3’
CRISPR mutagenesis genomic sequencing primers	
SLF2 Forward	5’-TTTCTGCAACCAGGTAGTCCT-3’
SLF2 Reverse	5’-AGTTCCGATAATCCACCCCTT-3’
SLF1 Forward	5’-ACAAGGAAAACACTTGCATTCG-3’
SLF1 Reverse	5’-ACCCTTAATTTGTGTGCATGTT-3’
SLF2 cDNA sequencing primers	
760 Forward	5’-AAGGAGCAAATGGAGCAGAGAA-3’
1624 Forward	5’-TGCGCTCAGAATATGGCACT-3’
2556 Forward	5’-GTCTGATGTAGCAGCTGTGTT-3’
2961 Forward	5’-TGAACTCTCCAGTCATCCCCA-3’
1768 Reverse	5’-GGCTTTATCTGAAGGTGCTGC-3’
2575 Reverse	5’-ACACAGCTGCTACATCAGACA-3’
3437 Reverse	5’-CTGGCGACCAAGTCTTTCAC-3’
TOPO vector sequencing primers	
M13 Forward	5´-GTAAAACGACGGCCAG-3´
M13 Reverse	5´-CAGGAAACAGCTATGAC-3´

Table 2.3

The DNA sequences of primers used for sequencing or cloning.

2.5.9 TOPO™ Cloning

PCR products were A-tailed using Taq DNA Polymerase (NEB, M0267) according to NEB's standard protocol, and then TOPO™ cloned into the TOPO® vector using the TOPO™ TA Cloning™ Kit for Sequencing (Invitrogen, 450030). The TOPO™ plasmids were transformed into chemically competent 5-alpha *E. coli* (NEB, C2987H) as described above. To identify bacterial colonies possessing plasmids with the desired PCR products, bacteria were spread onto X-gal/IPTG LB agar plates (supplemented with 20 mg/ml X-gal, 1mM IPTG and 50 µg/ml ampicillin) and screened by blue-white screening. After incubation overnight at 30°C, blue colonies (indicating that the TOPO® vector had re-ligated without a PCR product) were avoided, whilst white colonies (indicating that the TOPO® vector contained a PCR product) were picked and, miniprepped and sequenced across the insert site using the primers detailed in Table 2.3.

2.6 DNA fibre spreading assay

For the analysis of DNA fibres derived from patient-derived fibroblasts or U-2-OS cells, $1x10^6$ cells were seeded in 10cm dishes 48 hours prior to harvesting so that they reached around 70% confluency. Cells were incubated with 25mM CldU for 30 minutes, washed with media containing 250mM IdU (with or without drug treatments) and incubated with 250mM IdU (with or without drug treatments) for a further 30 minutes. Cells were then harvested by trypsinization, pelleted by centrifugation at 1,200rpm and washed with PBS. For patient-derived LCLs, cells were seeded at $5x10^5$ cells/ml and harvested 24-48 hours later. Cells were incubated with 25mM CldU (with or without drug treatments) for 20 min, washed with media containing 250mM

IdU (with or without drug treatments), and incubated with 250mM IdU for 20 minutes. As before, cells were then pelleted via centrifugation and washed with PBS.

Cells were resuspended at a concentration of $5x10^5$ cells/ml in PBS. 2µl of cellular suspension was pipetted directly onto a glass microscope slide and lysed by the addition of 7µl lysis buffer (200 mM Tris-HCl, pH 7.5, 50 mM EDTA, 0.5% SDS). The mixture was allowed to air dry for 4 minutes before the DNA fibres were spread down the slide by gravity. The slides were incubated in methanol/acetic acid (ratio 3:1) fixative solution, before being air dried for 15 minutes and stored at 4 °C until staining.

To immunostain, slides were rinsed with H_2O and the DNA denatured with 2.5M HCl for 75 minutes. Slides were then rinsed with PBS and incubated in blocking solution (1% BSA, 0.1% Tween20, in PBS) for 30 minutes. For primary antibody staining, CldU was detected with rat anti-BrdU antibody, and IdU with mouse anti-BrdU antibody, diluted in blocking solution, as detailed in Table 2.1. After incubation for 1 hour with primary antibodies, the slides were rinsed with PBS and fixed with 4% paraformaldehyde for 10 minutes. Following a further wash in PBS, and then blocking solution, the slides were incubated for 1.5 hours in secondary antibodies conjugated to Alexa-Fluor-594 and Alexa-Fluor-488, as detailed in Table 2.1, diluted in blocking solution. Finally, the slides were rinsed with PBS and blocking solution before being mounted by Anti-Fade Fluorescence Mounting Medium (Abcam; ab104135). DNA fibres were visualised by fluorescent microscopy using a Nikon Eclipse Ni microscope with either a 100x oil immersion objective, images were captured with NIS-Elements software (Nikon Instruments), and analysed using FIJI version 2.1.0/153.c (Schindelin et al., 2012).

2.7 LCL proliferation assay

LCLs were seeded at a density of $0.25x10^6$ cells/ml and incubated with increasing concentrations of CX-5461 (Selleck Chemicals, S2684). When the untreated cells reached a concentration of $2.0x10^6$ cells/ml (after three population doublings), the cell concentration of the treated cells was counted. The viability of the treated cells was expressed as a percentage of the untreated cell count

2.8 Statistical Analyses

Statistical analyses were carried out as detailed in the figure legends. N refers to number of independent experiments. Any error bars represent the standard error of the mean (SEM). A p-value of less than 0.05 represents a statistically significant result. * = $P<0.05$; ** = $P<0.01$, *** = $P<0.001$.

Chapter Three:

SLF2 and SMC5 mutations are associated with human disease and cellular dysfunction

3.1 Introduction

The SMC5/6 complex comprises of a heterodimer of SMC5 and SMC6, alongside additional subunits NSE1, NSE2 (also called MMS21), NSE3 and NSE4, all of which are highly conserved from yeast to higher eukaryotes (Fousteri and Lehmann, 2000, Taylor et al., 2001, Taylor et al., 2008, Aragon, 2018, Hallett et al., 2021a). The complex has numerous roles in genome stability and chromosome biology, including the stabilisation and recovery of stalled replication forks and the regulation of HR (Aragon, 2018, Palecek, 2018). The importance of SMC5/6 is highlighted by the fact that inherited mutations in two components of complex, namely NSE2 and NSE3, are associated with two autosomal recessive diseases (Payne et al., 2014, van der Crabben et al., 2016, Willemse et al., 2021).

Two NSE2 variants, p.Ser116Leufs*18 and p.Ala234Glufs*4, have been causally linked to two cases of primordial dwarfism in which patients presented with extreme insulin resistant diabetes, including severe fatty liver and dyslipidemia, alongside primary gonadal failure (Payne et al., 2014). NSE2 possesses E3 SUMO-ligase activity via the presence of a C-terminal Siz/PIAS -RING (SP-RING) domain, enabling the SMC5/6 complex to promote auto-SUMOylation, as well as target several downstream substrates (Potts and Yu, 2005, Zhao and Blobel, 2005, Andrews et al., 2005). The identified mutations led to a drastic decrease in total protein levels of NSE2 in patient-derived lymphoblastoid cell lines (LCLs), although protein expression levels of SMC5 and SMC6 were only mildly decreased (Payne et al., 2014). Characterisation of the individual mutations revealed that the p.Ser116Leufs*18 was completely defective in auto-SUMOylation activity, comparable to a ligase dead mutant, but the p.Ala234Glufs*4 possessed wildtype (WT) levels of auto-

SUMOylation activity. Consistent with the known role for SMC5/6 in replication and recombination, patient-derived LCLs exhibited delayed S-phase progression and increased levels of UV-induced sister chromatid exchanges. Furthermore, increased levels of HU-induced nuclear abnormalities in patient-derived fibroblasts, such as micronuclei and nucleoplasmic bridges, could be rescued by re-expression of WT NSE2 but not a ligase dead mutant, demonstrating that functional E3 SUMO-ligase activity was required to protect genome stability (Payne et al., 2014).

More recently, mutations in NSE3 have been identified in two separate studies, together describing nine patients from four families, all of whom who died during childhood with chronic lung disease (van der Crabben et al., 2016, Willemse et al., 2021). The immunological phenotypes of the five patients identified by Willemse *et al* were variable. However, all four individuals identified by van der Crabben *et al* had immune deficiencies, presenting with B and T cell abnormalities. Karyotypic analysis of three of these patients also revealed that a small proportion of cells had chromosome abnormalities, including structural changes and *de novo* supernumerary marker chromosomes; small, structurally abnormal chromosomes which originate from copied regions from other chromosomes. Of the four patients reported by van der Crabben *et al*, further analysis revealed that two individuals from one family were homozygous for p.Leu264Phe, whilst the other two patients, identified from an additional family, were compound heterozygous for p.Pro209Leu and p.Leu264Phe. Both p.Leu264Phe and p.Pro209Leu mutants were found to significantly destabilise NSE3, as well as SMC5 and SMC6. Cells from affected individuals exhibited hypersensitivity to a range of DNA damaging agents, as well as delayed clearance of yH2AX foci in G2, but not G1, cells after IR, suggesting the presence of a HR defect.

Furthermore, NSE3 patient-derived fibroblasts showed a delayed recovery of replication following a HU-induced S-phase arrest, similar to the phenotype seen in NSE2 mutated patient cells.

It is striking that mutations in two subunits of the same complex can result in such differing clinical presentations. One notable difference between the two diseases is that mutations in NSE3, which together with NSE1 and NSE4 forms a bridge between the SMC5 and SMC6 subunits (Palecek et al., 2006, Hudson et al., 2011, Guerineau et al., 2012), causes a much greater destabilisation of the SMC5/6 complex than loss/mutation of NSE2 (Payne et al., 2014, van der Crabben et al., 2016). As the SMC5/6 complex has functions that do not require E3 SUMO-ligase activity, this may partly explain the differences in clinical features (Andrews et al., 2005, Rai et al., 2011).

In yeast, additional Smc5/6 subunits Nse5 and Nse6 form an obligate heterodimer which facilitates the loading of the complex onto DNA and promotes its functions in DNA repair (Pebernard et al., 2006, Duan et al., 2009b). Whilst these subunits show little sequence similarity across other species, they are regarded as functionally equivalent to mammalian SLF1 and SLF2 (Raschle et al., 2015). Work in human cell lines has shown that the SMC5/6 complex is localised to sites of DNA damage via a protein-protein interaction cascade pathway involving RAD18, SLF1 and SLF2 (Raschle et al., 2015). The recruitment of this pathway requires RNF8, MDC1, and RNF168, which promote the ubiquitination of histones around DNA lesions. This ubiquitination then acts as a docking site for DNA repair factors (Schwertman et al., 2016), including RAD18 which binds ubiquitin via its UBZ domain (Huang et al., 2009). Via an interaction between SLF1 and RAD18, SMC5/6-SLF2-SLF1 is then recruited

to the site of DNA damage (Raschle et al., 2015). At present, very little is known about the function of these proteins, other than in the recruitment of SMC5/6 within the SMC5/6-SLF2-SLF1-RAD18 protein pathway.

Interestingly, mutations in RNF168 have also been linked with human disease. RNF168 variants were first identified in a single individual with RIDDLE syndrome; a genetic condition characterised by radiosensitivity, immunodeficiency, dysmorphic features, and learning difficulties (Stewart et al., 2007, Stewart et al., 2009), a markedly different clinical presentation to that seen in patients with mutations in NSE2 and NSE3. Cells derived from patients with RIDDLE syndrome show defective recruitment of 53BP1 and BRCA1 to DNA DSBs, a hypersensitivity to IR and cell cycle checkpoint defects (Stewart et al., 2009). Since this discovery, three further patients with defects in RNF168 have been identified, from two unrelated families, presenting with radiosensitivity and immunodeficiency, alongside telangiectasia and ataxia (Devgan et al., 2011, Pietrucha et al., 2017). Interestingly, one of these patients also exhibited microcephaly and pulmonary failure (Devgan et al., 2011), phenotypes common to patients with mutations in NSE3 (van der Crabben et al., 2016, Willemse et al., 2021).

3.1.2 Aims

The identification of distinct clinical diseases associated with mutations in NSE2 and NSE3 raises the question of whether mutations in the other components of the SMC5/6 complex are associated with human disease, and, if so, what the clinical presentations of such diseases would be. Therefore, the purpose of this chapter was to determine if mutations in different subunits of the SMC5/6 complex contribute to a

common clinical and cellular phenotype, and to better understand the spectrum of clinical phenotypes associated with defects in the SMC5/6 complex.

Following the identification of SLF2 variants in a previously undiagnosed patient with microcephaly and short stature via whole exome sequencing, the first aim of this chapter was to identify other patients with mutations in SLF2 and/or other components of the SMC5/6 complex. The second aim was to analyse the impact of any identified mutations on both a protein and cellular level using patient-derived fibroblasts and LCLs, with a particular focus on the replication stress response due to the critical role of the SMC5/6 complex in maintaining genome stability during S-phase.

3.2 Results

3.2.1 Clinical and genetic features of SLF2 and SMC5 patients

Whole exome sequencing was performed on an initial undiagnosed patient who presented with microcephaly and short stature, alongside other clinical phenotypes, such as recurrent bacterial infections, immunodeficiency, learning difficulties and speech delays. The patient also exhibited café-au-lait spots, brown oval birthmarks that are observed in several genetic conditions associated with defects in DNA damage repair, including Ataxia–Telangiectasia and Fanconi anemia (FA) (Lalor et al., 2020). Whilst no *de novo* mutations were observed, analysis under a recessive model of inheritance identified heterozygous or homozygous mutations in several genes. A detailed consideration of these genes within previous studies and genetic databases suggested that the homozygous frameshift SLF2 variant (c.1006dup; p.Arg336LysfsTer27) was likely to be responsible for the clinical phenotype that was observed. Having concluded this, a search was extended to try to find other patients

with mutations in the same gene. Using the GeneMatcher web tool (Sobreira et al., 2015), six additional patients from five different families were identified, all of whom possessed biallelic mutations in SLF2 and also presented with microcephaly and short stature (Figure 3.1, Table 3.1, Figure 3.2a). As SLF2's currently only known function is within the SMC5/6 complex, GeneMatcher was used to try and identify further patients with mutations within this complex. Four patients from three unrelated families were found, all of whom were confirmed to harbour biallelic mutations in the SMC5 gene (Table 3.1, Figure 3.2b). As with the SLF2 patients, these individuals presented with microcephaly and short stature (Figure 3.1, Table 3.2).

Of particular interest, all the SMC5 patients, alongside two of the SLF2 patients (SLF2-P4-2, SLF2-P5), had cardiac abnormalities, including atrial and ventricular septal defects, heart murmurs and supravalvular pulmonic stenosis- a narrowing of the pulmonary artery (Table 3.3). Whilst cardiac problems are often observed in human diseases associated with the defective repair of DNA interstrand crosslinks, principally FA complementation groups, this clinical presentation is atypical when considering the spectrum of other DNA repair or replication stress linked conditions. Similarly, patients with FA frequently exhibit gross anaemia and bone marrow failure, a clinical finding that is far less common amongst other replication stress associated diseases (Oostra et al., 2012, Reynolds and Stewart, 2018). Strikingly, five out of the eleven individuals presented with anaemia (SLF2-P3, SLF2-P4-1, SLF2-P4-2, SLF2-P5, SMC5-P9-2), with SMC5-P9-2 eventually developing myelodysplastic syndrome, a type of blood cancer which affects the bone marrow (Hofmann and Koeffler, 2005). As the surviving SLF2 and SMC5 patients are all very young, it remains to be seen

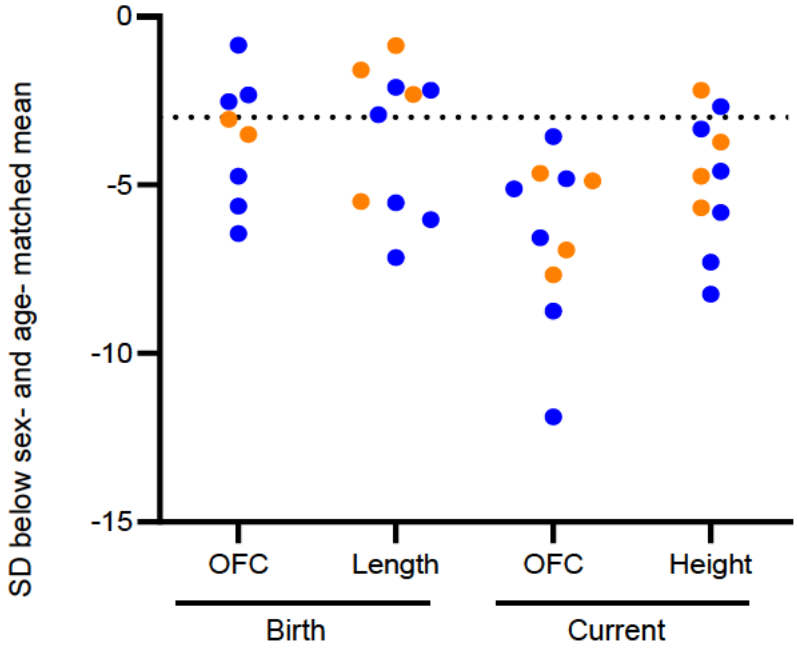


Figure 3.1. Microcephalic dwarfism in patients with mutations in SLF2 and SMC5.

The head circumference (occipital frontal circumference; OFC) and height of SLF2 and SMC5 patients at birth and at age of last examination, presented as z-scores (standard deviations (SD) relative to age and sex matched mean). SMC5 patients are in orange and SLF2 patients are in blue. The dashed line (-3 SD) indicates the cut off for normal distribution.

Individual	Ancestry	Gene	Mutation 1	Mutation 2
P1	UK	SLF2	c.1006dup; p.Arg336LysX27	c.1006dup; p.Arg336LysX27
P2	France	SLF2	c.2444C>G; p.Ser815X	c.3486G>C (ss); p.Gln1162His
P3	Netherlands	SLF2	c.3330G>A (ss); p.Arg1110Arg	c.3330G>A (ss); p.Arg1110Arg
P4-1	Japan	SLF2	c.2582A>T; p.Asn861Ile	c.2719dup; p.Ser907Phefs5X
P4-2	Japan	SLF2	c.2582A>T; p.Asn861Ile	c.2719dup; p.Ser907Phefs5X
P5	Saudi Arabia	SLF2	c.2347_2348del; p.Asp783Serfs53X	c.2347_2348del; p.Asp783Serfs53X
P6	German	SLF2	c.568C>T; p.Arg190X	c.568C>T; p.Arg190X
P7	Spain	SMC5	c.1110_1112del; p.Arg372del	c.1273C>T; p.Arg425X
P8	US	SMC5	c.2970C>G, p.His990Asp	c.2970C>G, p.His990Asp
P9-1	US	SMC5	c.2970C>G, p.His990Asp	c.2970C>G, p.His990Asp
P9-2	US	SMC5	c.2970C>G, p.His990Asp	c.2970C>G, p.His990Asp

Table 3.1: Biallelic SLF2 and SMC5 mutations in 11 individuals.

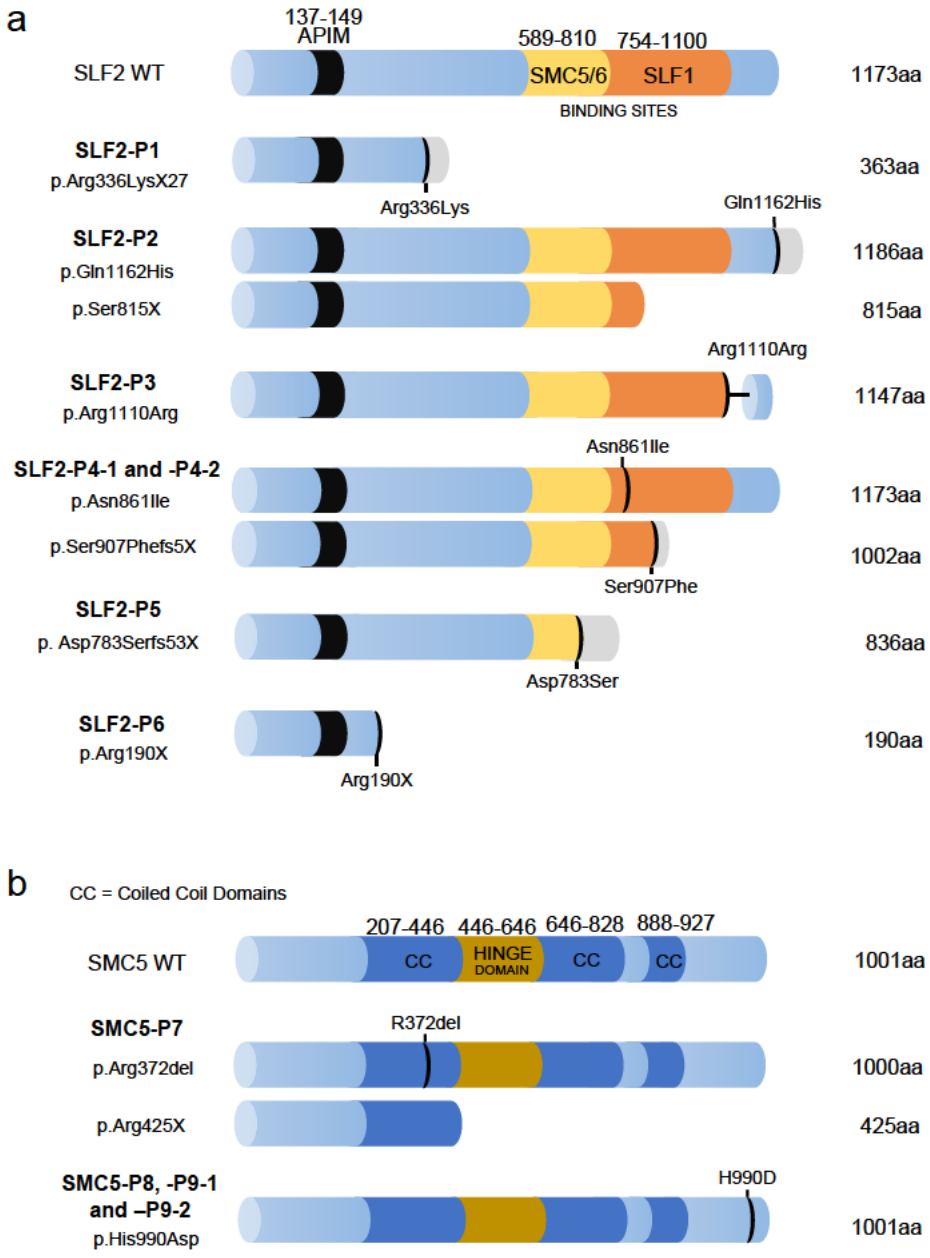


Figure 3.2. SLF2 and SMC5 mutations found in individuals with microcephalic dwarfism

(a) Cartoon schematics of full-length wildtype SLF2 and SLF2 patient variants, including known/predicted domains and interacting sites. APIM, atypical PCNA binding motif. Nonsense sequences are depicted in grey. A single allele is shown for homozygous mutations, and two for compound heterozygous mutations.

(b) Cartoon schematics of full-length wildtype SMC5 and SMC5 patient variants. CC, coiled-coil region. A single allele is shown for homozygous mutations, and two for compound heterozygous mutations.

		Birth				Current Exam			
Patient	Gene	Gestn (weeks + days)	Wgt/g (SD)	OFC/cm (SD)	Lgth/cm (SD)	Age at exam	Wgt/kg (SD)	OFC/cm (SD)	Hgt/cm (SD)
P1	*SLF2*	40+3	2280g (-2.78)	NA	NA	8 y 10 m	22kg (-1.70)	47cm (-5.12)	116.5cm (-2.67)
P2	*SLF2*	38	2230g (-2.09)	31cm (-2.33)	45cm (-2.2)	19 y 4 m	45.2kg (-3.00)	49cm (-4.82)	154cm (-3.34)
P3	*SLF2*	40	1380g (-4.91)	27cm (-6.44)	40cm (-5.53)	10 y 11 m	13.2kg (-11.23)	35cm (-11.88) at 6.5 y	104.5cm (-5.82)
P4-1	*SLF2*	33	1270g (-1.86)	26.5cm (-2.53)	36cm (-2.92)	7 y 0 m	9.0kg (-3.40)	43.3cm (-6.59)	83.8cm (-7.39)
P4-2	*SLF2*	37+2	1412g (-3.85)	27cm (-4.75)	37cm (-6.03)	1 y 3 m	4.475kg (-5.00)	36.9cm (-8.85)	55.5cm (-8.39)
P5	*SLF2*	35+6	845g (-5.07)	25cm (-5.63)	34cm (-7.16)	0 y 2 m	1.88kg (-3.00) at 217 days	NA	NA
P6	*SLF2*	39	2060g (-2.89)	33.5cm (-0.86)	46cm (-2.11)	0 y 4 m	3580g	38cm (-3.81)	53cm (-4.82)
P7	*SMC5*	38+5	1480g (-4.25)	30cm (-3.06)	39cm (-5.5)	8 y 4 m	13.3kg (-2.70)	43.5cm (-5.66) at 7y 6m	97.4cm (-5.7)
P8	*SMC5*	38	2409.7g (-1.45)	29cm (-3.51)	47cm (-0.87)	4 y 9 m	12.5kg (-2.92)	46.5cm (-4.23)	97.2cm (-2.25)
P9-1	*SMC5*	42	2551g (-2.87)	NA	49.53 cm (-1.59)	5 y 4 m	10.7kg (-5.33)	45.5cm (-4.88)	90cm (-4.7)
P9-2	*SMC5*	34	1644g (-1.56)	NA.	41.275 cm (-2.32)	10 y	16.6kg (-4.33)	45cm (-6.94)	114.3cm (-3.8)

Table 3.2: Measurements detailing the severity of microcephaly and short stature in patients with mutations in SLF2 and SMC5.

Length (Lgth), height (Hgt) weight (Wgt) and occipital frontal circumference (OFC) at birth and last exam are shown. The number of standard deviations (SD) from age and sex matched means are shown in brackets.

Patient	Gene	Central nervous system	Haematology and Immunology	Abnormalities of organs	Facial abnormalities	Deceased
P1	*SLF2*	Developmental delay. Learning difficulties. Mild neonatal hypotonia.	Recurrent infections. Combined immunodeficiency with impaired T cell proliferative response to CD3 and immune dysregulated T cell spectratype.	None	Microcephaly. Irregular teeth. Feeding difficulties. Narrow, long face.	-
P2	*SLF2*	Mild developmental delay. Learning difficulties.	NA	None	Microcephaly. Small teeth. Feeding difficulties. Large nose, small ears.	-
P3	*SLF2*	Moderate to severe developmental delay. Learning difficulties.	Recurrent infections. Thrombopenia, leukopenia and anemia	Ocular abnormalities. Liver problems.	Microcephaly. Feeding difficulties in early life.	Deceased aged 12y 2m
P4-1	*SLF2*	Mild developmental delay. Learning difficulties.	Anaemia during neonata period. No anaemia currently.	None	Microcephaly Calcification abnormality of teeth. Narrow, long face.	-
P4-2	*SLF2*	Mild developmental delay.	Sideroblastic anaemia. Died with interstitial pneumonia triggered by infection at the age of 1y 3m.	Cardiac abnormalities; atrial septal defect.	Microcephaly	Deceased aged 1y 3m
P5	*SLF2*	Moderate developmental delay. Moderate hypotonia/hypertonia.	Anaemia, required 3 weekly transfusions. Neutropenia, resolved later.	Cardiac abnormalities; ventricular septal defect (closed spontaneously). Ocular abnormalities.	Microcephaly Toothless. Swallowing abnormalities. Long philtrum and small chin.	Deceased
P6	*SLF2*	Severe developmental delay. Hypotonia.	NA.	None	Microcephaly	Deceased
P7	*SMC5*	Learning difficulties due to attention impairment.	NA.	Cardiac abnormalities; neonatal persistent ductus arteriosus and pulmonary stenosis (self-corrected).	Microcephaly. Upper incisors diastema. Delayed adult tooth eruption. Narrow, long face, slight epicanthal folds, beaked prominent nose.	-
P8	*SMC5*	Mild motor delays related to blindness, no intellectual disability. History of seizures.	Low relative lymphocytes, elevated relative eosinophils and low absolute lymphocytes at last exam.	Cardiac abnormalities; V bratory systolic ejection murmur Ocular abnormalities. Legally blind. Kidney abnormalities	Microcephaly. Bilateral microphthalmia.	-
P9-1	*SMC 5*	Behavioural anomalies: anxiety, startled easily as an infant.	Low platelets. low relative lymphocytes, low absolute lymphocytes and elevated relative eosinophils at last exam.	Cardiac abnormalities; Innocent heart murmur, normal ECG and echo.	Microcephaly.	-
P9-2	*SMC5*	Mild motor delays resolved, no intellectual disability.	Anaemia. Neutropenia. Myelodysplastic syndrome,	Cardiac abnormalities; Mild supravalvar pulmonary stenosis. All resolved. Pulmonary flow murmur, likely innocent. Ocular abnormalities.	Microcephaly. Micrognathia.	Deceased

Table 3.3: Details of the clinical symptoms present in patients with mutations in SLF2 and SMC5.

whether the anaemia present in the other affected individuals will also progress and have similar, serious consequences. In contrast to cardiac abnormalities and anaemia, immunodeficiency is seen more frequently in patients with mutations in DNA replication genes, such as NBS1, BLM, GINS1, polymerase δ and MCM10 (Gennery et al., 2000, Wolska-Kusnierz et al., 2015, Cottineau et al., 2017, Schoenaker et al., 2018, Conde et al., 2019, Mace et al., 2020, Redmond et al., 2022). At least five of the eleven SLF2 (SLF2-P1, SLF2-P3, SLF2-P4-2) and SMC5 (SMC5-P9-1, SMC5-P9-2) patients exhibited evidence of immunodeficiency, such as recurrent bacterial infections and neutropenia, and SLF2-P4-2 died due to interstitial pneumonia triggered by infection (Table 3.3).

3.2.2 SLF2 and SMC5 mutations impact protein stability

To begin to understand how the SLF2 variants may impact the function of the protein and the integrity of the SMC5/6-SLF2-SLF1-RAD18 pathway, I first considered the patient mutations in the context of SLF2's interacting domains; the minimal regions required for binding to SMC5/6 (residues 589 - 810) and SLF1 (residues 754-1100), as mapped by our collaborator, Prof. Neils Mailand (Grange et al., 2022) (Figure 3.2a). Five SLF2 variants, p.Arg336Lysfs25X (present as a homozygous variant in SLF2-P1), p.Ser815X (present in trans with the p.Gln1162His mutation in SLF2-P2), p.Ser907Phefs5X (present in trans with pAsn861Ile in SLF2-P4-1 and SLF2-P4-2), p.Asp783Serfs53X and p.Arg190X (present as homozygous variants in SLF2-P5 and SLF2-P6, respectively) all led to premature stop codons which truncated key SLF1 and/or SMC5 binding sites. This suggested that even if these proteins were expressed and stable, the loss of these regions would likely abrogate the interaction between SLF1 and the SMC5/6 complex. pAsn861Ile, a missense mutation present in SLF2-

P4-1 and SLF2-P4-2, also occurred within SLF2's putative SLF1 binding region. Whilst it's difficult to predict the impact of this change, multiple sequence alignments showed that the residue was highly conserved in higher eukaryotes, suggesting that the alteration may likely be deleterious (Figure 3.3).

I next sought to investigate the SLF2 variants, p.Gln1162His (present in SLF2-P2) and p.Arg1110Arg (present as a homozygous variant in SLF2-P3). Work done by our collaborator Prof. Cedric Le Caignec, using splice-site analysis software, predicted that the c.3486G>C (p.Gln1162His) mutation, present in patient SLF2-P2, would impact splicing at the exon 19 donor site in the most abundant SLF2 mRNA transcript (NM_018121). RT-PCR and western blotting indicated that the isoform-specific splice defect led to decreased mRNA and protein stability (Prof. Cedric Le Caignec; data not shown). Similarly, the c.3330G>A (p.Arg1110Arg) variant in SLF2-P3 presented with a nonsense mutation adjacent to exon 17's donor site. cDNA sequencing confirmed that this change also impacted splicing, leading to an in-frame deletion of exon 17 and the loss of 23 residues within a highly conserved region that lies proximal to the predicted SLF1 binding region (Figure 3.3, Figure 3.4).

Having considered SLF2 patient mutations in the context of the minimal binding regions of SLF2-interacting partners, western blots were carried out using cell extracts derived from available patient LCLs to determine if these variants impacted protein stability. In all of the SLF2 patients that were analysed, there was either a significant decrease or a complete loss of full-length SLF2 protein expression (Figure 3.5a). This suggested that all the tested mutations destabilised the SLF2 protein and were therefore also likely to lead to disruption of the SMC5/6-SLF2-SLF1-RAD18 protein

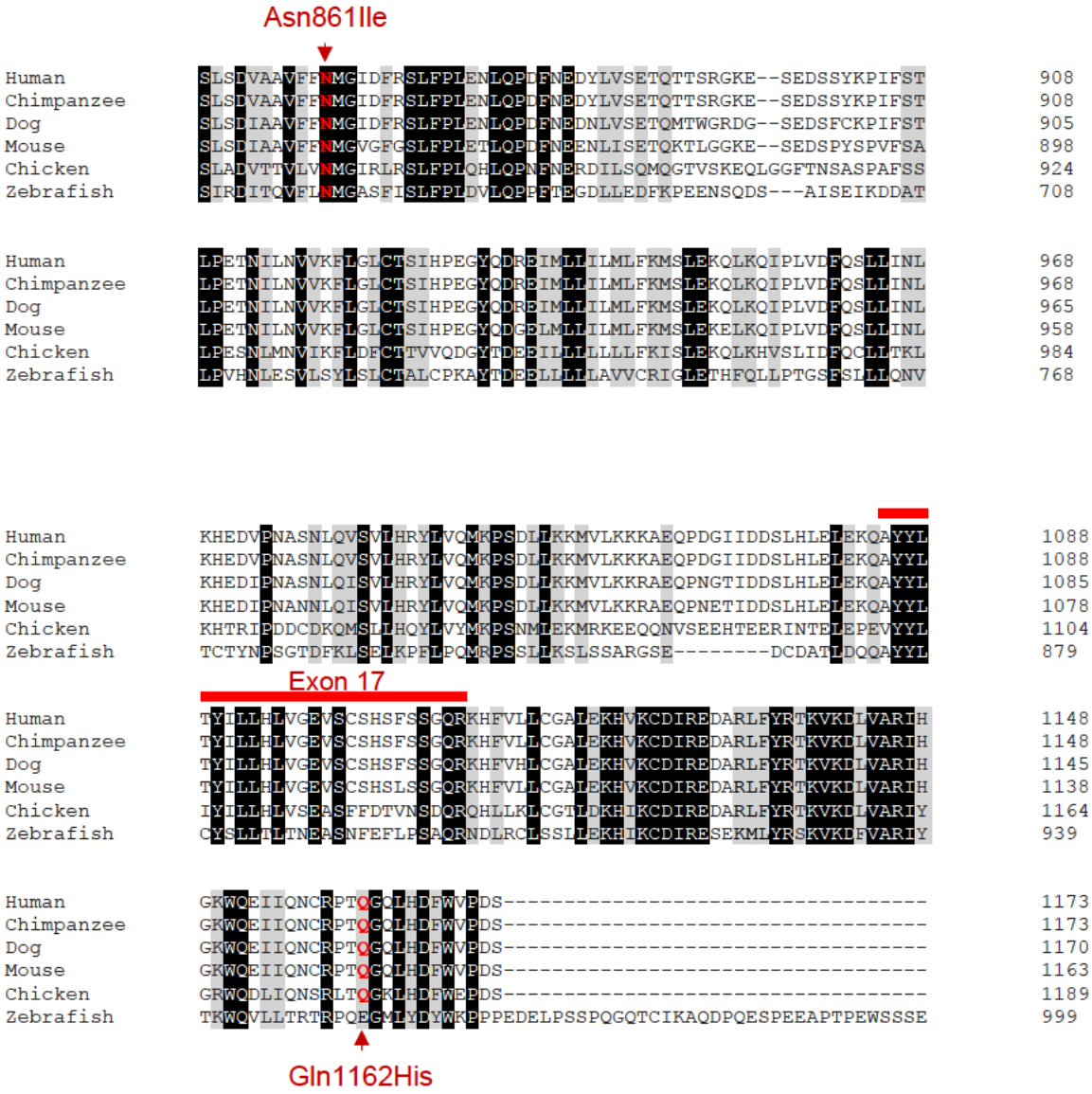


Figure 3.3. Multiple sequence alignments of SLF2 showing the conservation of patient point mutations

Amino acid alignment of the SLF2 protein from different species showing the degree of evolutionary conservation of disease causing SLF2 point mutations, generated using Clustal Omega. The red highlight and arrows indicate the location of missense variants present in SLF2-P2 (p.Gln1162His) and SLF2-P4-1 and SLF2-P4-2 (p.Asn861Ile). The red line indicates the region that exon17 encompasses (which is lost due to splicing defects in the p.Arg1110Arg variant in SLF2-P3). Black shading indicates positions which have a single, fully conserved residue, grey shading indicate conservation between groups of strongly similar properties.

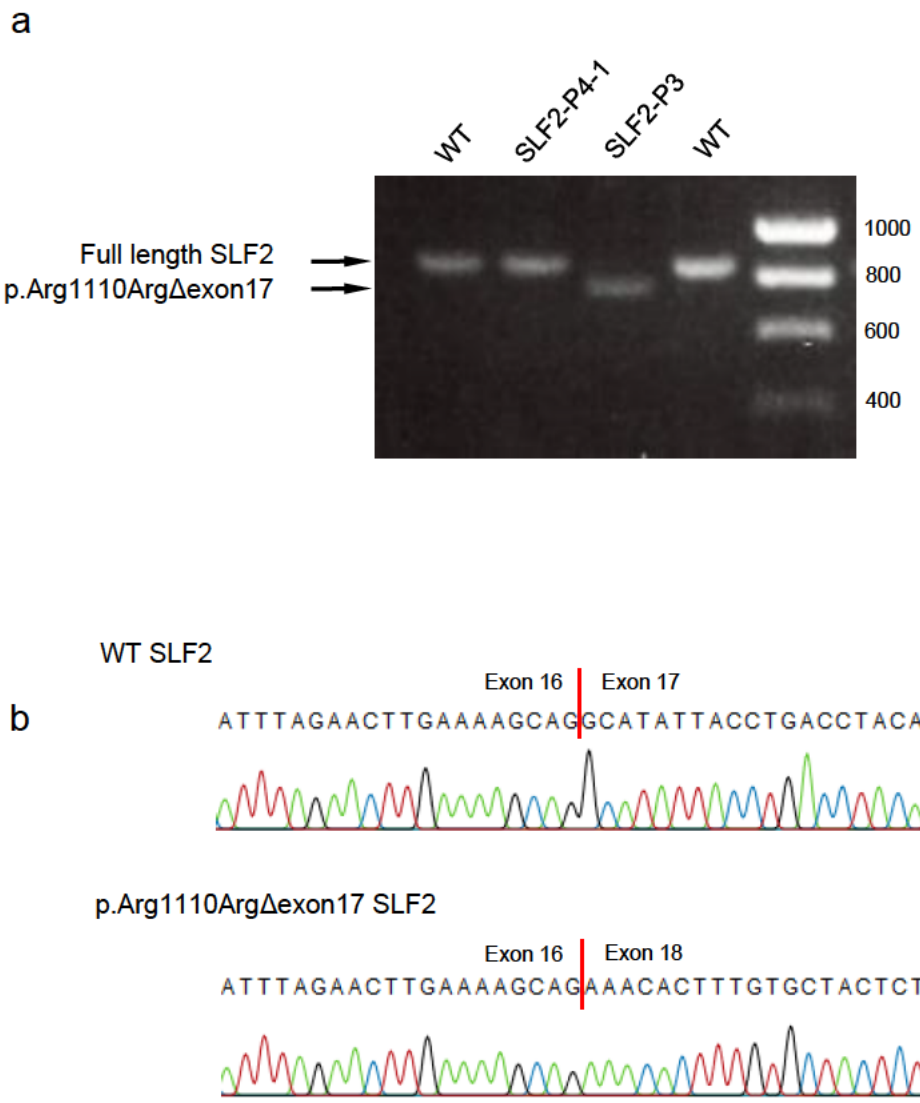


Figure 3.4. Mutations in SLF2-P3 disrupt mRNA splicing

(a) Agarose gel electrophoresis showing PCR amplification products of cDNA extracted from SLF2-P4-1, SLF2-P3 and unrelated, healthy individuals (WT).
(b) cDNA sequencing indicating the loss of exon 17 in SLF2-P3. Chromatograms generated using QIAGEN CLC Genomics Workbench.

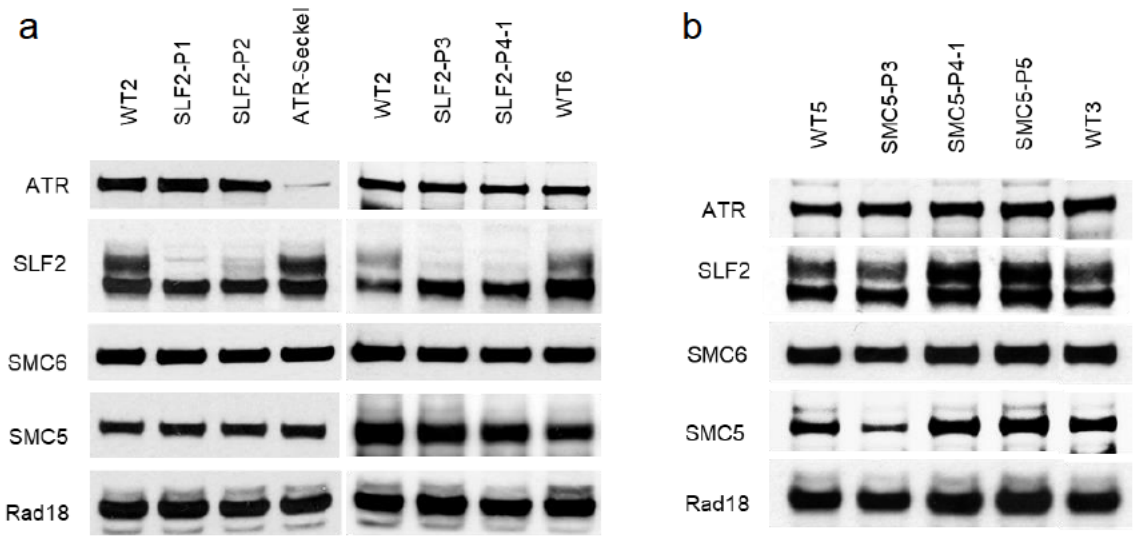


Figure 3.5. Protein expression of SLF2 and SMC5 in patient-derived lymphoblastoid cell lines

(a) Representative immunoblots of whole cell extracts from SLF2 patient lymphoblastoid cell lines (LCLs), using the indicated antibodies, depicting the stability of SLF2 and its interacting proteins. Whole cell extracts from unrelated, healthy individuals (WT2 and WT6) and an ATR Seckel syndrome patient are also included as controls. ATR is used as a loading control.

(b) Representative immunoblots of whole cell extracts from SMC5 patient LCLs, using the indicated antibodies, depicting the stability of SMC5 and its interacting proteins. Whole cell extracts from unrelated, healthy individuals (WT3 and WT5) are also included as controls. ATR is used as a loading control.

Experiments in (a) and (b) were carried out by Prof. Grant S Stewart and are representative of two independent biological replicates in which similar results were achieved.

pathway. This includes the p.Gln1162His and p.Arg1110Arg variants, which despite residing outside of the SLF1 and SMC5/6 binding regions, are still likely to impact on SLF2 function due to their impact on SLF2 protein stability.

In contrast, western blotting of SMC5 patient LCLs revealed that the p.His990Asp SMC5 homozygous variant present in SMC5-P8 , -P9-1 and -P9-2 had no observable impact on full length protein expression (Figure 3.5b). Only cell extracts generated from SMC5-P7 patient LCLs exhibited a mild decrease in the levels of SMC5 protein. In all cases, western blotting also showed that, despite any changes in SLF2 or SMC5 expression, levels of SMC6 and RAD18 were unaffected in patient LCLs, indicating that mutations in neither protein resulted in loss of stability of other components of the pathway.

Patient SMC5-P7 presented with two variants: p.Arg425X and p.Arg372del. p.Arg425X possesses a premature stop codon, causing a severe protein truncation that is likely to significantly disrupt function as it deletes the protein's hinge and secondary coiled coil regions. This is most likely the predominant cause of the decrease in levels of full length SMC5 protein seen in SMC5-P7 patient-derived LCLs. Conversely, p.Arg372del represents only a single amino acid deletion. Whilst the residue is conserved in higher eukaryotes (such as human, chimpanzee, dog, mouse and zebrafish), it is amongst a region of weak overall conservation (Figure 3.6). To better understand the impact of this deletion, I turned to existing structural models of the SMC5/6 complex. Lys368, which is the *Saccharomyces cerevisiae* functional equivalent of human SMC5 Arg372, is proximal to Smc5's Nse2's binding site in yeast (Figure 3.6). This presents the possibility that the loss of Arg372 in humans disrupts

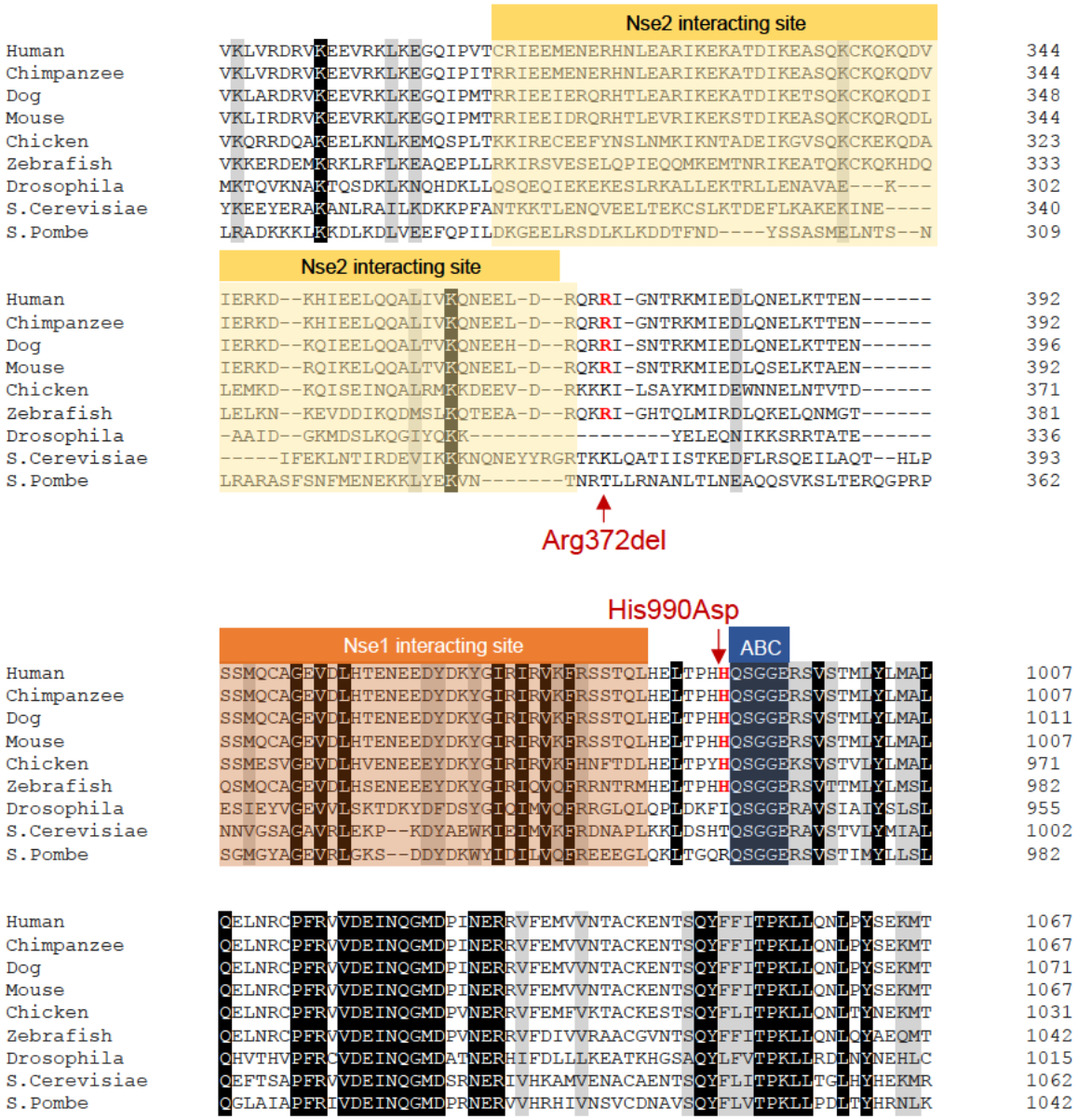


Figure 3.6. Multiple sequence alignments of SMC5 showing the conservation of patient point mutations

Amino acid alignment of the SMC5 protein from different species showing the degree of evolutionary conservation of disease causing SLF2 point mutations, generated using Clustal Omega. The red highlight and arrows indicate the location of missense variants present in SMC5-P7 (p.Arg372del) and SMC5-P8, SMC5-P9-1 and SMC5-P9-2 (p.His990Asp). SMC5's *Saccharomyces cerevisiae* NSE2 and NSE1 interacting sites and ABC-signature motif are highlighted. Black shading indicates positions which have a single, fully conserved residue, grey shading indicates conservation between groups of strongly similar properties.

NSE2's interaction with SMC5, and/or the SUMO-ligase activities of SMC5/6 (Duan et al., 2009a, Hallett et al., 2021a).

In addition to this, SMC5's structure showed that His990, a residue with high conservation in higher eukaryotes, is positioned directly adjacent to the ATP-binding cassette (ABC) transporter signature motif (QSGGE), a region important for promoting ATP binding (Figure 3.6). The alteration of the overall charge of the region, by replacing an aromatic amino acid with a negatively charged one in the His990Asp point mutation (SMC5-P8, SMC5-P9-1 and SMC5-P9-2), could potentially disrupt the ATPase activity of SMC5/6. In turn, this could affect the complex's ability to both load and disassociate chromatin (Duan et al., 2009a, Hallett et al., 2021a, Etheridge et al., 2021). Therefore, it is likely that all three SMC5 patient mutations will impact the function of the SMC5/6 complex to some extent.

3.2.3 SLF2 and SMC5 patient cells exhibit spontaneous S-phase associated DNA damage and genomic instability

Considering that the primary function of SMC5/6 is believed to occur during DNA replication (Venegas et al., 2020), it was pertinent to understand how the SLF2 and SMC5 variants would impact replication dynamics. To question this, the DNA fibre assay was utilised. Using this method, asynchronous cells are sequentially pulsed with thymidine nucleotide analogues, CldU and IdU, which become incorporated into nascent DNA (Nieminuszczy et al., 2016) (Figure 3.7a). The cells are then lysed, and the DNA is spread onto microscope slides before two different antibodies raised against thymidine that differentially bind to CldU and IdU are used to immunostain the regions of the DNA fibres that contained newly replicated DNA (CldU in red and IdU

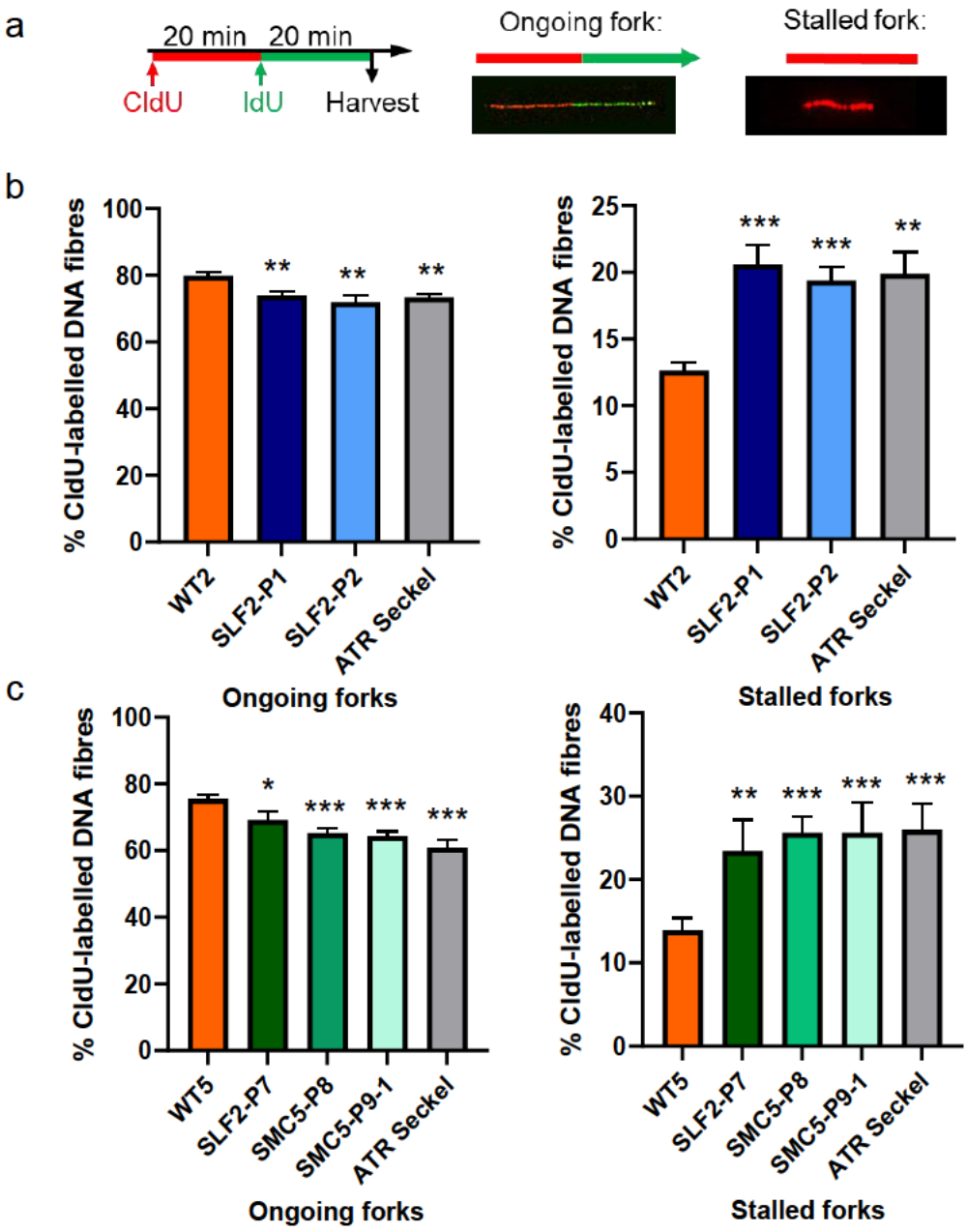


Figure 3.7. SLF2 and SMC5 patient lymphoblastoid cell lines exhibit replicative dysfunction

(a) Experimental outline of the DNA fibre assay. Lymphoblastoid cell lines (LCLs) were sequentially labelled with two different thymidine analogs, CldU (red) and IdU (green) for 20 mins each, as shown.

DNA fibre analysis of (b) SLF2 and (c) SMC5 patient LCLs, in comparison to a LCLs derived from an unrelated, healthy individual (WT2 and WT5) and an ATR seckel syndrome patient. The percentage of ongoing forks (red and green tracks) and stalled forks (red-only tracks) were quantified. n=4 independent experiments. A minimum of 750 fork structures were counted. Statistical differences were determined with an unpaired, 2-tailed, student's t-test. Experiments carried out by Dr John Reynolds and Dr Robert Hollingworth.

in green). This staining technique can indicate the direction of replication and can also reveal additional DNA structures, such as stalled forks (a red only track) and new origins (a green only track).

The DNA fibre assay, which was performed on patient LCLs, demonstrated spontaneous increases in the levels of stalled replication forks, as well as decreases in ongoing forks, across all SLF2 and SMC5 patients in comparison to WT controls (Figures 3.7b and 3.7c). Not only did this indicate that SLF2 and the integrity of the SMC5/6-SLF2-SLF1-RAD18 complex is important for efficient replication, but it also demonstrated that the mutations in SMC5 were sufficient to adversely affect these functions, too. In particular, the data supported the prediction that the homozygous His990Asp mutation would negatively impact the function of SMC5/6 despite not affecting protein expression.

Importantly, these findings were verified in complemented cell lines by generating patient-derived fibroblasts infected with a lentivirus encoding either Myc-tagged WT SLF2/SMC5 or an empty vector (Figure 3.8). However, SMC5 patient fibroblasts became inviable upon the re-expression of the Myc-tagged SMC5 protein and were instead complemented with untagged SMC5 (Figure 3.8b). DNA fibre analysis using these cells showed that levels of stalled forks significantly decreased upon complementation with either WT SLF2 or SMC5, demonstrating that mutations in these genes are responsible for the replicative dysfunction observed in these patient cell lines (Figure 3.9). Having established that SLF2 and SMC5 patient derived cell lines exhibit problems with replication, I sought to understand the cellular impact of this dysfunction by analysing additional markers of replication stress. Since prolonged

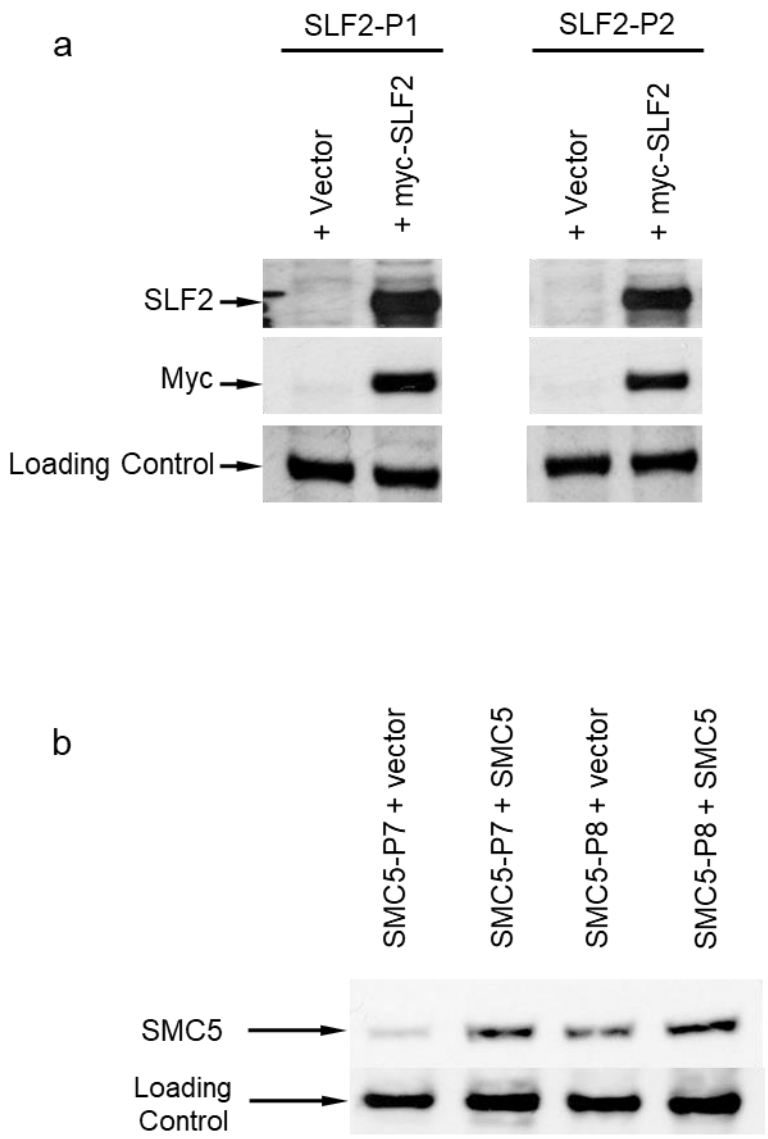


Figure 3.8. Complementation of SLF2 and SMC5 patient derived fibroblasts

Representative immunoblots of whole cell extracts from (a) SLF2 and (b) SMC5 patient fibroblasts complemented with either an empty lentiviral expression vector or a lentiviral expression vector encoding WT Myc-tagged SLF2 or WT untagged SMC5. The loading control was a non-specific band recognised by either the SLF2 or SMC5 antibody.

These cell lines were created by, and immunoblotting was carried out by, Prof. Grant S Stewart.

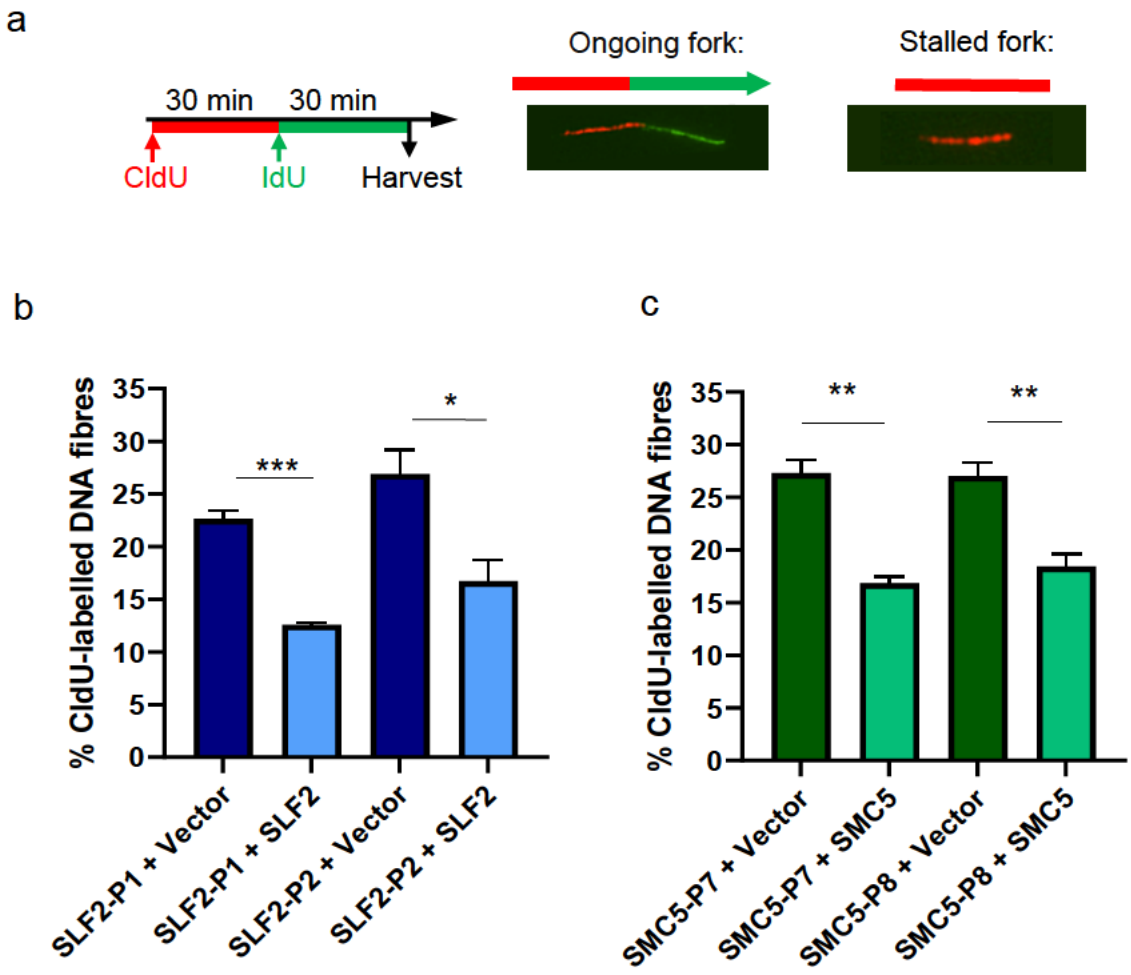


Figure 3.9. SLF2 and SMC5 patient derived fibroblasts exhibit elevated levels of stalled forks which reduce upon complementation

(a) Experimental outline of the DNA fibre assay. Fibroblasts were sequentially labelled with two different thymidine analogs, CldU (red) and IdU (green) for 30 mins each, as shown.

DNA fibre analysis of (b) SLF2 and (c) SMC5 patient fibroblasts complemented with a lentivirus encoding wildtype (WT) SLF2, WT SMC5, or an empty vector. The percentage of stalled forks (red-only tracks) were quantified. n=3 independent experiments. A minimum of 350 fork structures were counted. Statistical differences were determined with an unpaired, 2-tailed, student's t-test. Experiments carried out by Dr John Reynolds.

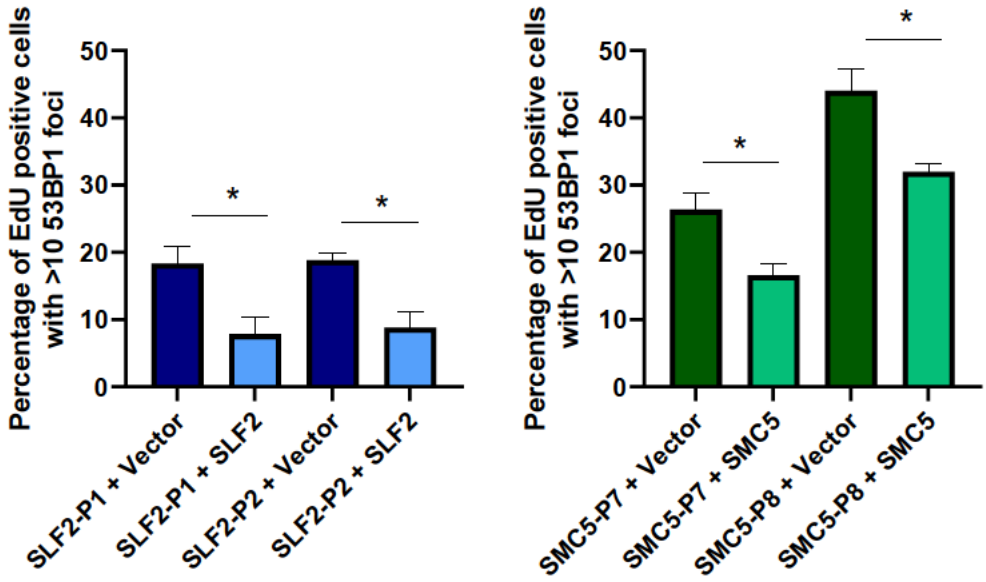
replication fork stalling can lead to fork collapse, I first questioned if there were elevated levels of DSBs present during replication in the patient cell lines. Using 53BP1 foci as a marker of DSB formation and EdU DNA incorporation as a positive indicator of S-phase, I observed a significant decrease in the proportion of highly damaged cells exhibiting >10 foci in both WT SLF2 and WT SMC5 complemented fibroblasts (Figure 3.10). Consistent with decreased levels of DNA damage and genome instability, the patient fibroblasts also exhibited a significant reduction in the average number of micronuclei per cell upon the re-expression of WT protein (Figure 3.11).

Elevated levels of replication stress can lead to under-replicated regions of DNA persisting throughout the cell cycle. This is primarily caused by the stalling of converging forks lacking licensed origins in between them; a phenomena termed double fork stalling (Bertolin et al., 2020). Under-replicated DNA is not detected by G2/M checkpoint machinery (Bertolin et al., 2020), meaning it has the potential to persist through mitosis, leading to further defects such as chromatin bridges and aberrant chromosome segregation. To prevent this, mitotic DNA synthesis (MiDAS) completes synthesis at under-replicated regions after the cells have initiated prophase (Minocherhomji et al., 2015, Bertolin et al., 2020). To understand if the increased levels of replication stress caused by the SLF2 and SMC5 variants also led to elevated levels of under-replicated DNA, I analysed the levels of MiDAS in the patient LCLs (Figure 3.12). Using EdU positive foci as markers of replication, I saw significantly elevated levels of MiDAS in all patient SLF2 and SMC5 patient LCLs in the absence of exogenous DNA damage. However, after 24 hours treatment with a

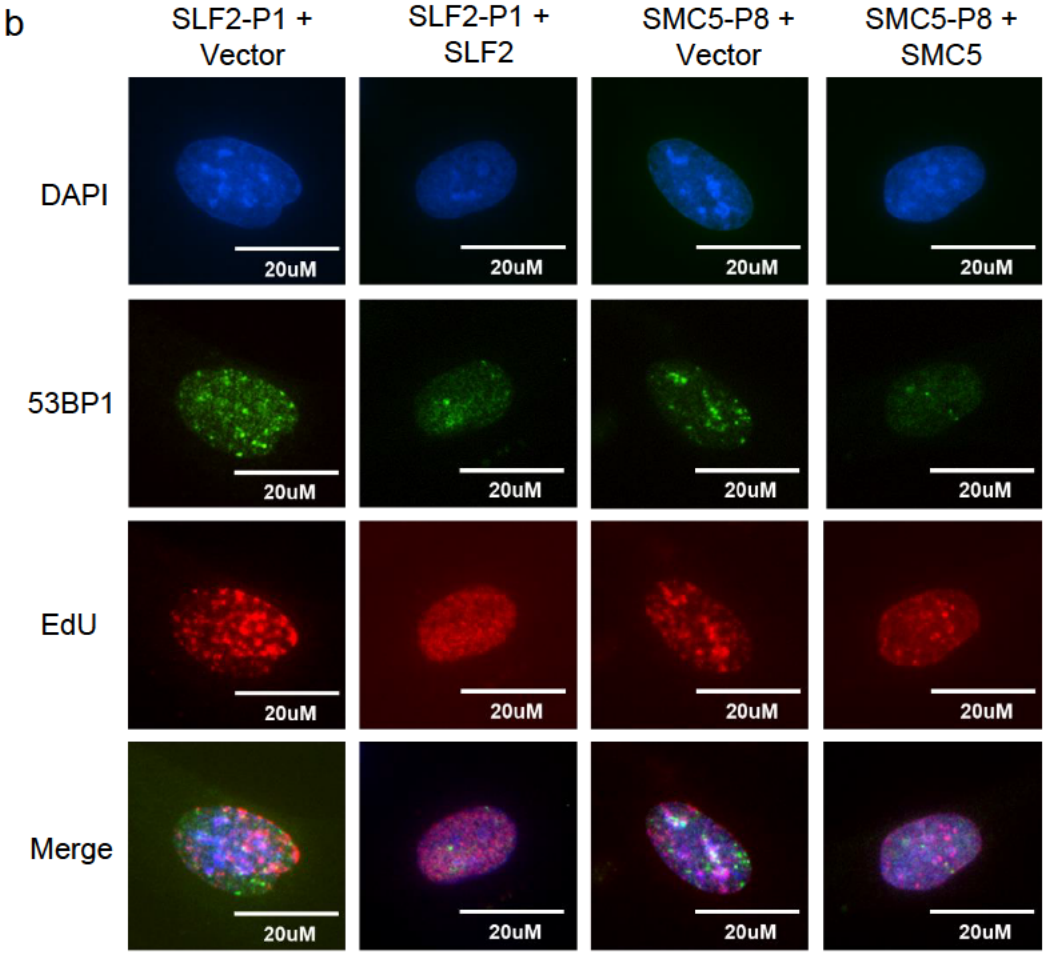
Figure 3.10. SLF2 and SMC5 patient derived fibroblasts exhibit elevated levels of 53BP1 foci which reduce upon complementation (data shown overleaf)

(a) Quantification of the percentage of S-phase cells with >10 53BP1 foci in SLF2 (left) and SMC5 (right) patient fibroblasts complemented with a lentivirus encoding WT SLF2, WT SMC5, or an empty vector. Cells were treated with 10 μM EdU (a thymidine analogue) for 45 min before fixation and immunostaining with an antibody specific to 53BP1. DNA containing EdU was detected with click chemistry using Alexa Fluor-488 picolyl azide. S-phase cells were identified as those that were positive for EdU incorporation. n=3 independent experiments. A minimum of 900 cells were counted. Statistical differences were determined with an unpaired, 2-tailed, student's t-test.
(b) Representative immunofluorescent images of patient derived fibroblasts, complemented with a lentivirus encoding WT SLF2, WT SMC5, or an empty vector from (a). DAPI shown in blue, 53BP1 in green and EdU in red.

a


Percentage of EdU positive cells with >10 53BP1 foci
50
40
30
20
10
0
*
*
SLF2-P1 + Vector
SLF2-P1 + SLF2
SLF2-P2 + Vector
SLF2-P2 + SLF2
Percentage of EdU positive cells with >10 53BP1 foci
50
40
30
20
10
0
*
*
SMC5-P7 + Vector
SMC5-P7 + SMC5
SMC5-P8 + Vector
SMC5-P8 + SMC5

b


SLF2-P1 + Vector
SLF2-P1 + SLF2
SMC5-P8 + Vector
SMC5-P8 + SMC5
DAPI
53BP1
EdU
Merge
20uM

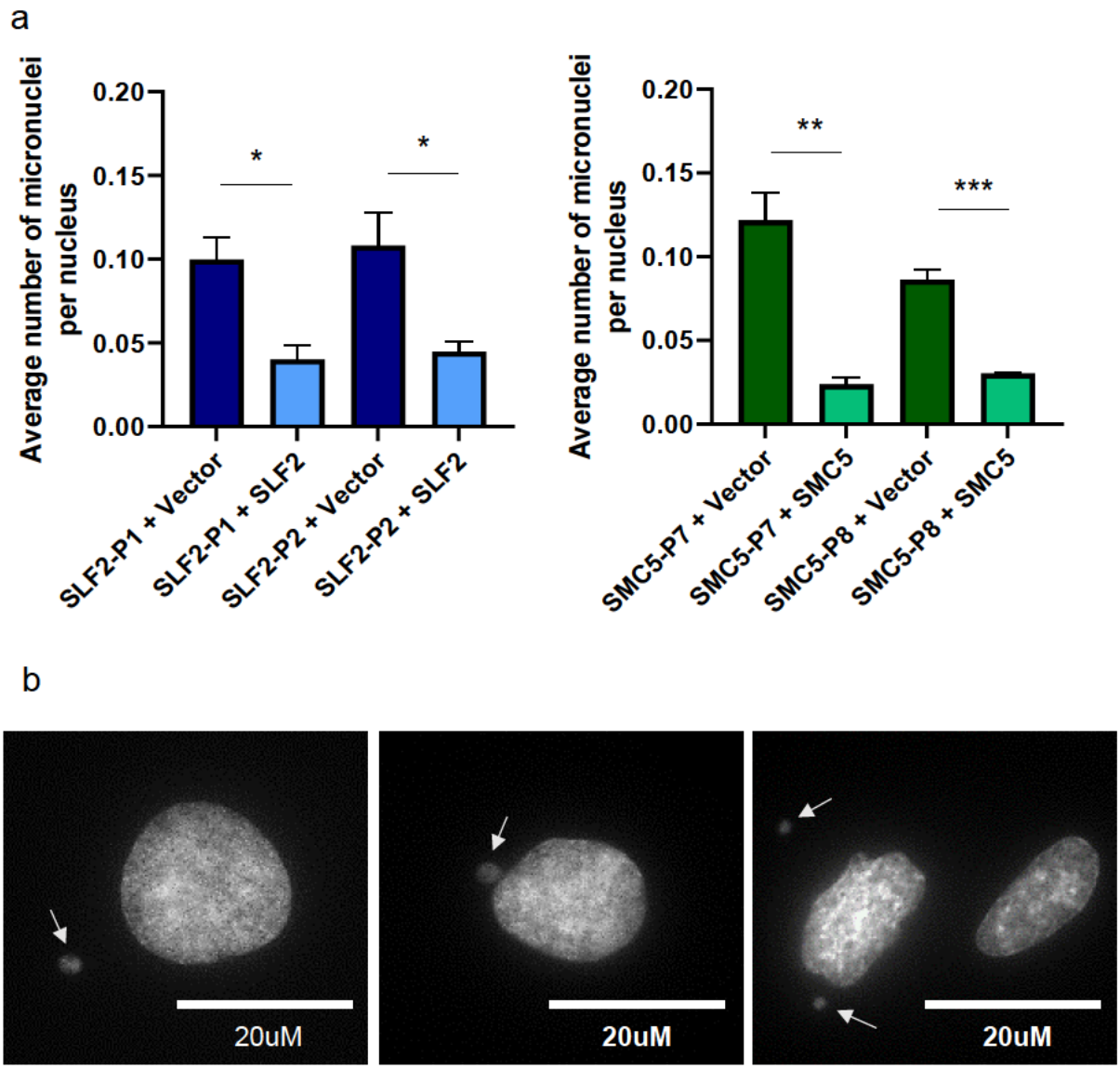


Figure 3.11. SLF2 and SMC5 patient derived fibroblasts exhibit elevated numbers of micronuclei which reduce upon complementation with WT protein

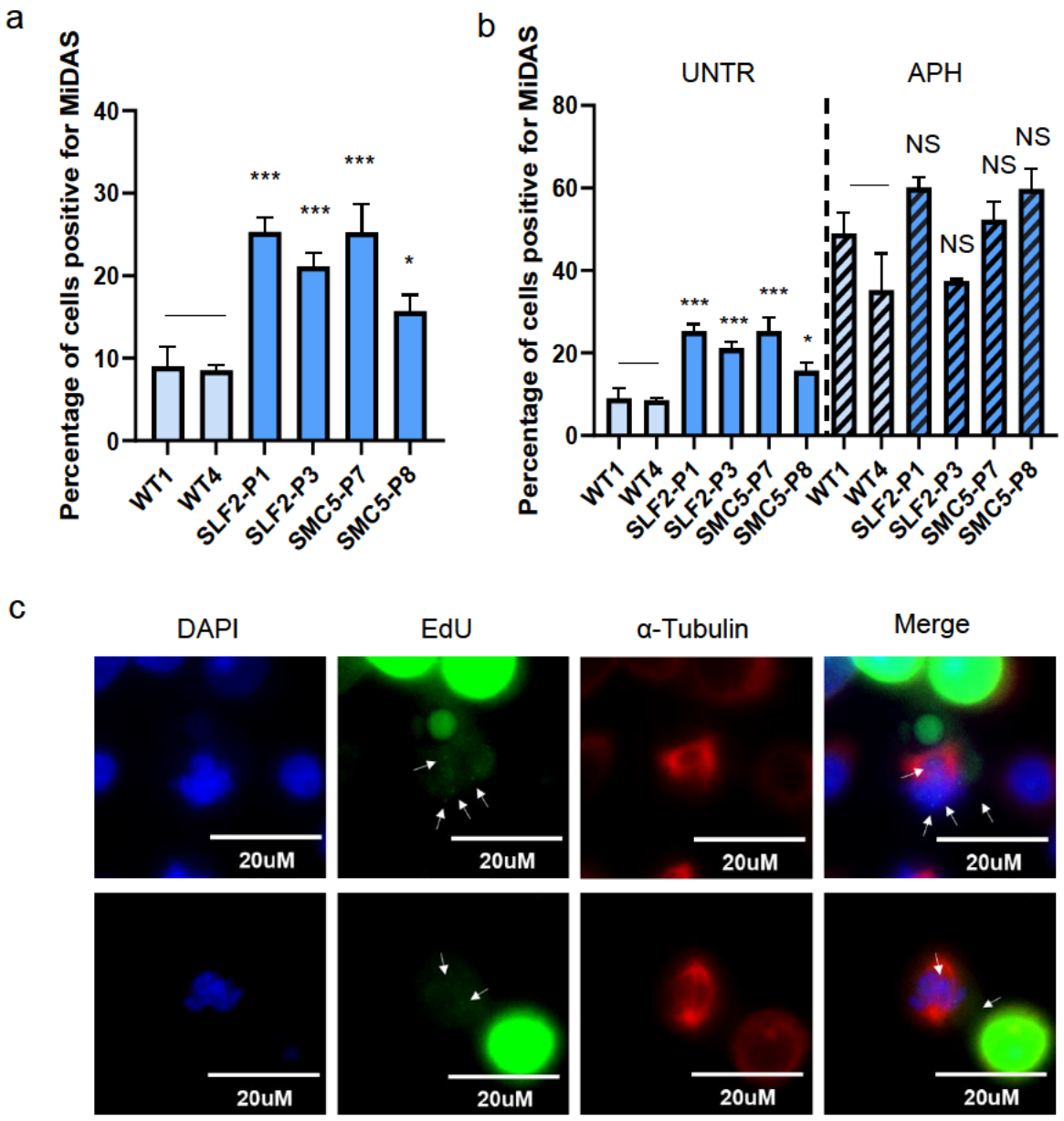
(a) Quantification of the average number of micronuclei per cell in SLF2 (left) and SMC5 (right) patient fibroblasts complemented with a lentivirus encoding WT SLF2, WT SMC5, or an empty vector. n=3 independent experiments. A minimum of 2,500 cells were counted. Statistical differences were determined with an unpaired, 2-tailed, student's t-test.
(b) Representative fluorescent microscopy images of patient fibroblast nuclei exhibiting micronuclei (indicated by white arrows), visualised using DAPI.

Figure 3.12. SLF2 and SMC5 patient derived lymphoblastoid cell lines exhibit elevated levels of spontaneous MiDAS (data shown overleaf)

(a) Quantification of the percentage of mitotic cells exhibiting spontaneous mitotic DNA synthesis (MiDAS) in lymphoblastoid cell lines (LCLs) derived from patients with mutations in SLF2 and SMC5, as well as healthy, unrelated individuals (WT1 and WT4). Cells were treated with 10 μM EdU for 45 min before fixation. MiDAS was visualised by mitotic EdU incorporation after labelling with Click chemistry. n=3 independent experiments. A minimum of 300 mitotic cells were counted. Statistical differences were determined with an unpaired, 2-tailed, student's t-test.

(b) The untreated (UNTR) experiment as shown in (a), alongside quantification of the percentage of MiDAS positive patient LCLs after 24 hours treatment with 250 nM aphidicolin (APH). Cells were treated and stained as explained in (a). n=3 independent experiments. A minimum of 300 mitotic cells were counted. Statistical differences were determined with an unpaired, 2-tailed, student's t-test.

(c) Representative immunofluorescent images of SLF2-P8 LCLs indicating the presence of EdU positive MiDAS foci with white arrows. DAPI shown in blue, EdU in green and α-Tubulin in red.


a
Percentage of cells positive for MiDAS
40
30
20
10
0
***
***
***
*
WT1
WT4
SLF2-P1
SLF2-P3
SMC5-P7
SMC5-P8
b
Percentage of cells positive for MiDAS
UNTR
APH
80
60
40
20
0
***
***
***
*
NS
NS
NS
WT1
WT4
SLF2-P1
SLF2-P3
SMC5-P7
SMC5-P8
WT1
WT4
SLF2-P1
SLF2-P3
SMC5-P7
SMC5-P8
c
DAPI
EdU
α-Tubulin
Merge
20uM
20uM
20uM
20uM
20uM
20uM
20uM
20uM

low dose of aphidicolin (APH), whilst levels of MiDAS were elevated, the patient lines no longer exhibited increases relative to WT controls, indicating that this cellular phenotype was predominantly seen spontaneously in untreated cells. I then went on to question if the elevated levels of replication stress led to DNA damage or under replicated DNA being inherited by daughter cells. Using the patient fibroblasts, I analysed levels of 53BP1 bodies in G1 and observed, as with the micronuclei and 53BP1 foci, that the re-expression of wild-type SLF2 or SMC5 significantly reduced these markers of replication stress (Figure 3.13). Together, this suggested that replicative dysfunction induced by mutations in these proteins contributes to elevated levels of genome instability and under replicated regions of DNA that persist after cell division.

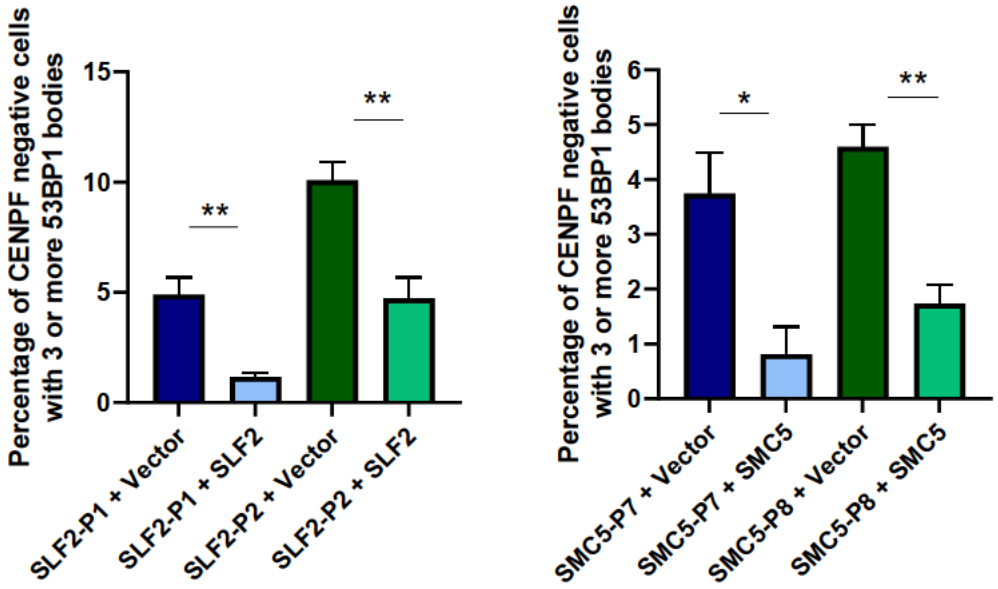
3.3 Discussion

The data presented in this chapter demonstrate that mutations in SLF2 and SMC5, components of the SMC5/6-SLF2-SLF1-RAD18 complex, are associated with human disease, with all patients exhibiting differing severities of microcephaly (SD: -3.57 to -11.88) and short stature (SD: -2.19 to -8.24). Several patients also presented with cardiac abnormalities, anaemia and/or immunodeficiency. Notably, SLF2 and SMC5 patients have very similar clinical presentations, suggesting that common molecular defects underly both diseases. Individuals with mutations in SLF2 and SMC5 also appear to share clinical features with patients with FA and Warsaw Breakage Syndrome, both driven by mutations in genes which promote proper replication. This may reflect overlapping functions for these factors in DNA replication.

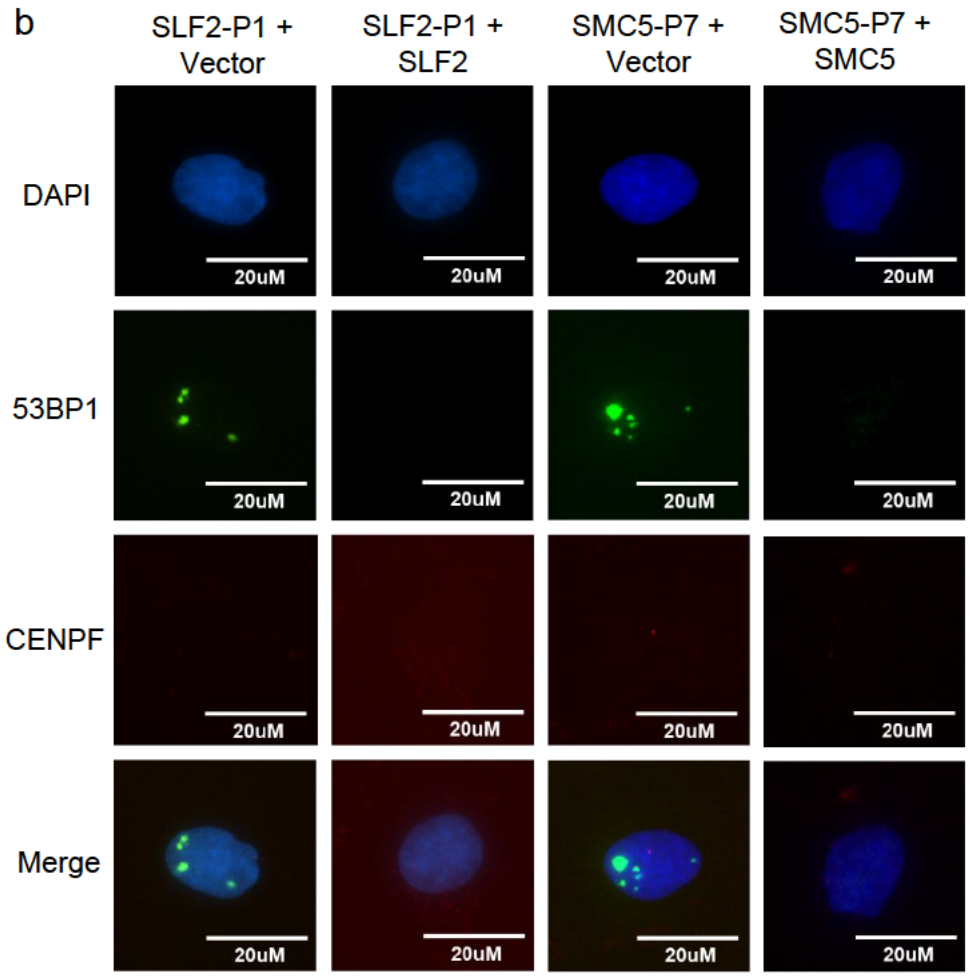
Figure 3.13. SLF2 and SMC5 patient derived fibroblasts exhibit elevated levels of 53BP1 bodies which reduce upon complementation with WT protein (data shown overleaf)

(a) Quantification of the percentage of G1-phase cells with >3 53BP1 bodies in SLF2 (left) and SMC5 (right) patient fibroblasts complemented with a lentivirus encoding WT SLF2, WT SMC5, or an empty vector. G1-phase cells were identified as those negative for CENPF staining. n=3 independent experiments. A minimum of 750 cells were counted. Statistical differences were determined with an unpaired, 2-tailed, student's t-test.
(b) Representative immunofluorescent images of patient derived fibroblasts, complemented with a lentivirus encoding WT SLF2, WT SMC5, or an empty vector. DAPI shown in blue, 53BP1 in green and CENPF in red.

a


Percentage of CENPF negative cells with 3 or more 53BP1 bodies
15
10
5
0
**
**
SLF2-P1 + Vector
SLF2-P1 + SLF2
SLF2-P2 + Vector
SLF2-P2 + SLF2
Percentage of CENPF negative cells with 3 or more 53BP1 bodies
6
5
4
3
2
1
0
*
**
SMC5-P7 + Vector
SMC5-P7 + SMC5
SMC5-P8 + Vector
SMC5-P8 + SMC5

b


SLF2-P1 + Vector
SLF2-P1 + SLF2
SMC5-P7 + Vector
SMC5-P7 + SMC5
DAPI
53BP1
CENPF
Merge
20uM

Conversely, however, the diseases caused by mutations in SLF2 and SMC5 had strikingly different clinical presentations compared to patients with mutations in NSE3, who suffered from lung damage (van der Crabben et al., 2016, Willemse et al., 2021), and patients with mutations in NSE2, who presented with primordial dwarfism and extreme insulin diabetes (Payne et al., 2014). It is not yet clear why mutations in four subunits of the same complex can present with such differing diseases. It is possible that this represents non-redundant roles for the different subunits in genome stability, or the effects of specific mutations. Despite the lack of impact on the stability of the SMC5/6 complex as a whole, cell extracts prepared from patient-derived cell lines show that the mutations in SLF2 completely destabilise the protein. In line with previous studies which describe the linear nature of the SMC5/6-SLF2-SLF1-RAD18 protein pathway (Raschle et al., 2015), this loss is expected to disrupt the recruitment of SMC5/6 to sites of DNA damage. In contrast, mutations in SMC5 still allow for protein expression. Consistent with this, components of the SMC5/6 holocomplex are required for the viability of human cells (Venegas et al., 2020), suggesting that a complete loss of the protein may be incompatible with human life. Instead, the mutations in SMC5 appear likely to impact the interactions of components within the complex, such as the recruitment of NSE2, or the function of the complex's ATPase activity.

All patient cell lines derived from individuals with mutations in SLF2 and SMC5 exhibited elevated levels of spontaneous replication stress which contributes to further increases in genomic instability, suggesting that the SMC5/6-SLF2-SLF1-RAD18 protein pathway is important for maintaining faithful replication. This is consistent with previous studies which demonstrate a role for SMC5/6 in repairing

collapsed replication forks and promoting efficient replication (Ampatzidou et al., 2006, Bermudez-Lopez et al., 2010, Menolfi et al., 2015, Aragon, 2018, Palecek, 2018, Agashe et al., 2021). Importantly, this replication stress phenotype was corrected by the expression of WT SLF2 or WT SMC5, demonstrating that the identified mutations were the cause of the replication abnormalities and genome instability.

Chapter 4:

Generating SLF2 loss of function model cell lines

4.1 Introduction

The data presented in the previous chapter demonstrates that cell lines derived from patients with mutations in SLF2 or SMC5 exhibit increases in replication stress and genomic instability. Whilst the identified SMC5 variants retain significant levels of protein expression, mutations in SLF2 lead to either a significant decrease or a complete loss of full length SLF2 protein. Although the SMC5/6 complex is required for viability in human cells and *Saccharomyces cerevisiae* (Zhao and Blobel, 2005, Venegas et al., 2020), this data suggests that a complete knockout of SLF2, or its binding partner SLF1, may be achievable.

The *Streptococcus pyogenes* type II CRISPR-Cas9 system was first identified as an antiviral immune response in bacteria and archaea and has since been co-opted for the purposes of genome engineering (Wiedenheft et al., 2012, Li et al., 2020a). By inducing targeted DSBs, which are then repaired by NHEJ or HR, the Cas9 endonuclease is able to introduce mutations at specific loci. Error-prone NHEJ can lead to insertions/deletions (indels) and subsequent nonsense mutations, whilst homology cassettes can be incorporated into the genome during HR (Li et al., 2020a).

Several studies, however, have demonstrated that Cas9 is able to bind and cleave non-specific genomic sites, thereby risking experimental off-target effects where mutations are introduced in unintended genes (Zhang et al., 2015b, Alkan et al., 2018). To avoid this a CRISPR-Cas9 nickase system can be utilised, in which pairs of enzymes generate a DSB by nicking single-stranded DNA at neighbouring sites. Not only do the long overhangs generated by this approach provide a greater control over precision gene integration, but off-target effects are minimised as two adjacent

Cas9 cleavage events are required for a DSB to occur (Mali et al., 2013). In a study using HEK 293FT cells, a Cas9 nickase using two guide RNAs achieved 100-fold greater targeting specificity when compared to wild type Cas9 with a single guide (Ran et al., 2013). To streamline this technique further, all-in-one Cas9 nickase plasmids have been designed, including a single Cas9 gene and pairs of guide RNAs (Chiang et al., 2016). This avoids the need to co-transfect multiple plasmids, which may otherwise reduce transfection efficiency and the efficacy of the intended mutagenesis.

4.1.2 Chapter Aims

The aim of this chapter was to use CRISPR technology to generate a complete loss of SLF1 or SLF2 function in two commonly used immortalised cell lines. This would enable the analysis of gene knockouts in an easily manipulatable system, whilst eliminating the phenotypic variation that inevitably occurs between non-isogenic, patient derived cell lines. Whilst patients with mutations in SLF1 have yet to be identified, loss of function SLF1 cell lines would allow me to further understand the role of the SMC5/6-SLF2-SLF1-RAD18 pathway in maintaining genome stability. The impact of the loss of SLF2 by CRISPR editing was confirmed via the complementation of the cell lines and the subsequent analysis of their cellular phenotypes in the context of replication and genomic instability.

4.2 Results

4.2.1 Generating SLF1 and SLF2 CRISPR nickase plasmids

Pairs of guide RNAs targeting SLF1 and SLF2 were designed using CHOP CHOP, a webtool for selecting target sites for CRISPR/Cas9 directed mutagenesis (Labun et al., 2019). To maximise the chances of generating a successful CRISPR system, two pairs of guide RNAs were designed for each gene, with each pair targeting neighbouring loci 15-19bp away from each other (Figure 4.1). With the exception of SLF1 guide RNA pair 1, which targets sequences spanning the boundary of intron 3 and exon 4 (Figure 4.1a), these sites were located in coding regions in the centre of exons (SLF1 exon 4 or SLF2 exon 5). Nonsense mutations at these loci would severely truncate either protein and completely disrupt SLF1 and SMC5/6 binding sites within SLF2, whilst hopefully being distanced enough from the 3' end of the gene to avoid gene expression via downstream, secondary start codons.

The chosen guide RNA pairs were ligated into the purpose built GFP All-in-One Cas9^{D10A} nickase vector at unique BbsI and BsaI restriction sites (Figure 4.2), producing two SLF1 and two SLF2 targeting CRISPR plasmids. This vector possesses an enhanced green fluorescent protein (GFP)-open reading frame to allow GFP expression to be used as a marker of successful transfection into cells. These vectors are referred to henceforth as SLF1 plasmid 1 and plasmid 2 or SLF2 plasmid 1 and plasmid 2. In each case, guide incorporation was verified via diagnostic restriction site digestions and confirmed by sanger sequencing.

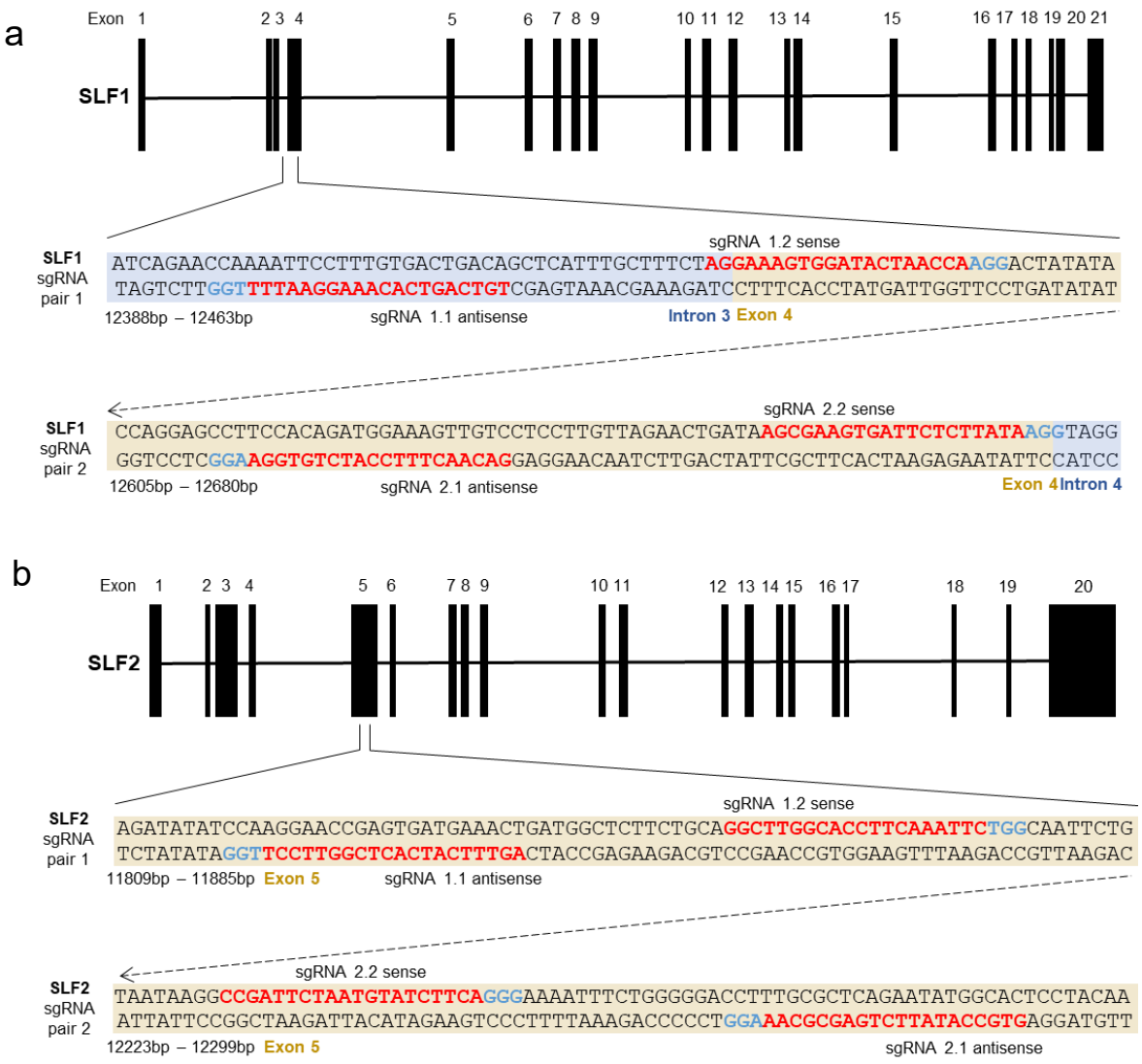


Figure 4.1 SLF1 and SLF2 CRISPR target locations and sequences

Schematic of the human SLF1 (a) and SLF2 (b) genomic locus. Filled rectangles indicate coding exons and black horizontal lines denote introns. CRISPR target sites are enlarged. Pairs of sense and antisense guide RNA (sgRNA) sequences are highlighted by red text and the location of protospacer adjacent motifs (PAM) are indicated by blue text.

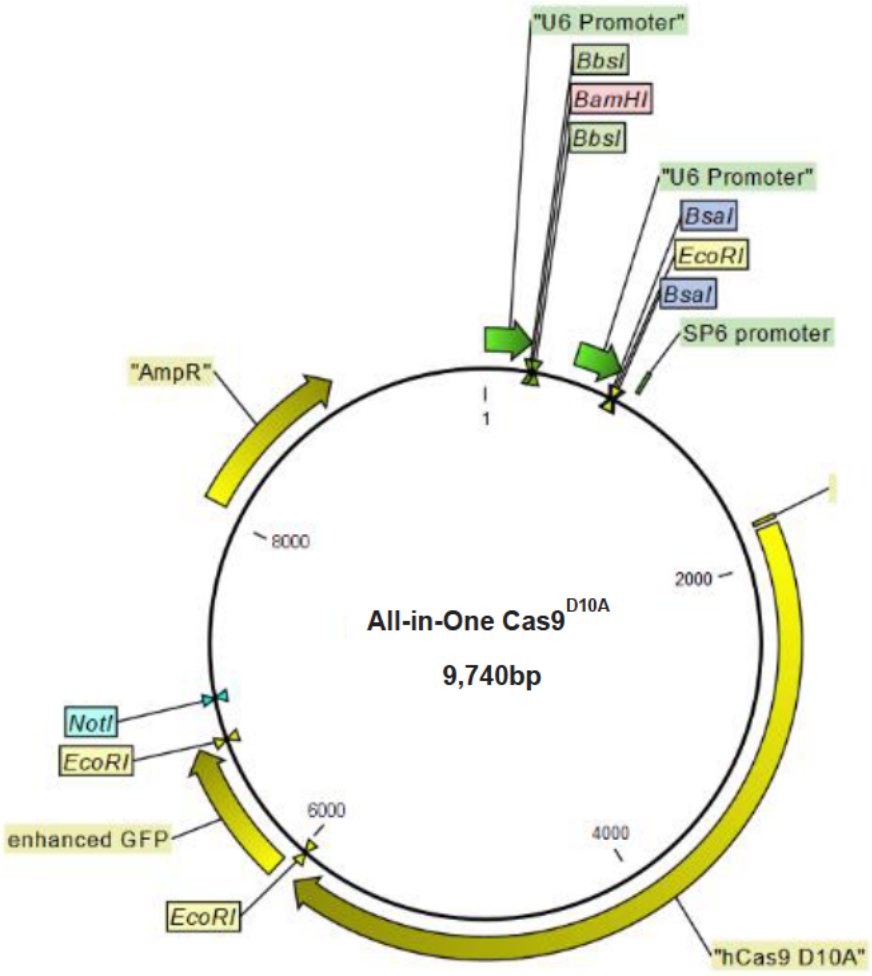


Figure 4.2 The All-in-One Cas9^{D10A} nickase vector, illustrated using CLC workbench.

5′ overhangs for sgRNA cloning were generated by digestion with BbsI and BsaI restriction enzymes. Guide incorporation was verified using EcoRI and BamHI and NotI diagnostic digests before sanger sequencing. Green-fluorescent protein (GFP) was used as a marker of successful vector transformation.

4.2.2 U-2-OS cell lines transfected with SLF2 CRISPR plasmids exhibit mutations at target sites which impact protein expression

The SLF1 and SLF2 CRISPR plasmids were transfected into U-2-OS and RPE1 cell lines using a 3:1 ratio of FuGene to DNA. U-2-OS, a human bone osteosarcoma epithelial cell line, was chosen due to their ease of culturing and manipulation. In addition, hTERT RPE1, a retinal pigment epithelial cell line, was also selected due to its normal karyotype, which I hypothesised would be the best genetic background for eventually investigating any possible mitotic dysfunctions.

48 hours after transfection, GFP positive single cells were sorted into 96 well dishes using fluorescence activated cell sorting. Culturing media contained 20% FBS to aid with recovery. After two weeks of growth, around 20% of the sorted U-2-OS cells had formed colonies. However, this was far more infrequent in the RPE1 cells, which often appeared large and with prominent stress granules. 25 U-2-OS colonies and 15 RPE1 colonies transfected with each plasmid were propagated (100 U-2-OS colonies and 60 RPE1 colonies in total), their genomic DNA was extracted and an approximately 1.4 kb region spanning the SLF1 or SLF2 target sites was sequenced.

Genomic sequencing analysis of the SLF2 gene in U-2-OS clones revealed that those transfected with SLF2 plasmid 2 exhibited only unedited, WT sequence. Of those transfected with SLF2 plasmid 1, five clones (clones 2, 8, 9, 14 and 22) presented with multiple traces that began around the guide RNA targeting sites, suggesting that at least one SLF2 allele had been mutated (Figure 4.3a). Clones 2, 9 and 14 retained a WT sequence that could be read amongst the multiple traces. This indicated that CRISPR mutagenesis had not been successful at every target site and that SLF2

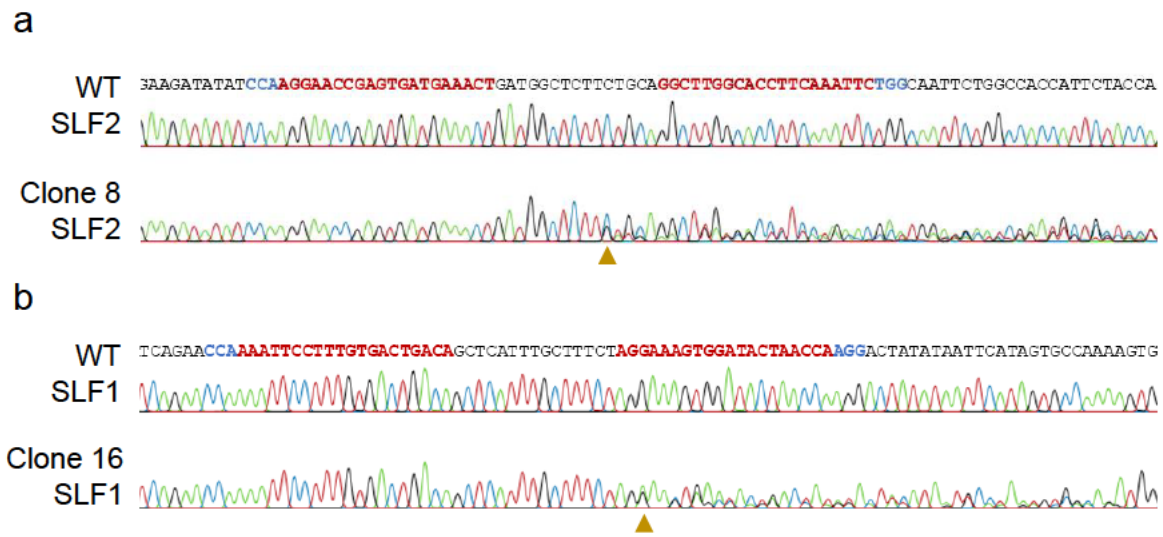


Figure 4.3 Screening using genomic sequencing identifies U-2-OS CRISPR clones with mutations at SLF1 and SLF2 target sites

Representative genomic DNA sequence chromatograms showing multiple traces downstream of SLF1 (a) and SLF2 (b) nickase-targeting sites in CRISPR-edited U-2-OS clones. CRISPR clone sequences are aligned against the relevant wildtype (WT) gene. For each WT sequence, positions of the single guide RNAs (sgRNA) are highlighted by red text and the location of the protospacer adjacent motif (PAM) is indicated by blue text. For each CRISPR clone, the start of multiple sequencing traces is indicated by the yellow arrow.

protein function may be retained. Conversely, no WT sequence could be read for clones 8 or 22, suggesting that each SLF2 allele had been successfully targeted and edited.

Similarly, genomic sequencing analysis of the SLF1 gene in U-2-OS cell lines demonstrated that only select clones transfected with SLF1 plasmid 1 exhibited DNA alterations. 7 clones showed multiple sequencing traces starting around the guide RNA target regions (Figure 4.3b). However, only three presented with traces lacking WT sequence: clones 6, 12 and 16. Genomic analysis also revealed that none of the RPE1 clones exhibited sequence changes indicative of CRISPR-mediated mutations in either SLF1 or SLF2. Therefore, I focused on validating the U-2-OS CRISPR clones further.

To understand how the SLF2 CRISPR mutations impact SLF2 expression, I used immunoblotting to investigate whether the SLF2 CRISPR cell lines exhibited decreased protein stability. The equivalent experiment could not be attempted for the SLF1 CRISPR cell lines due to the lack of a suitable SLF1 antibody. Using cell extracts prepared from all five SLF2 clones, I observed reductions in SLF2 protein expression when compared to controls (Figure 4.4). Although a low level of SLF2 expression was evident for all the U-2-OS SLF2 CRISPR cell lines, protein abundance was lowest in clones 8 and 22. Interestingly, SLF2 protein size was observed to have changed in both, appearing fractionally larger in clone 8 and smaller in clone 22. This suggested that the expressed allele(s) in these two clones contained in-frame indels which changed the protein's size and might also affect its function.

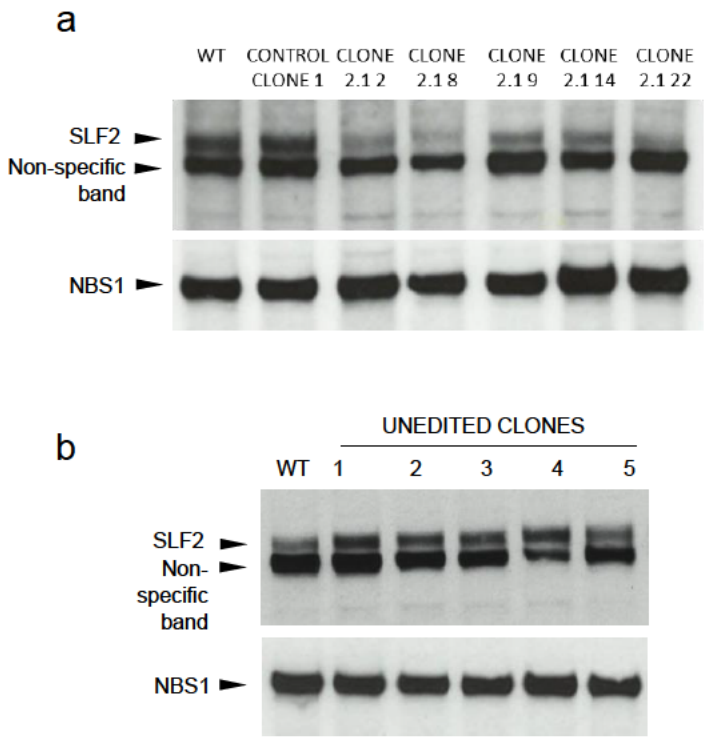


Figure 4.4 U-2-OS SLF2 CRISPR clones exhibit reduced SLF2 protein expression

(a) Representative immunoblot analysis of SLF2 expression in U-2-OS SLF2 CRISPR clones. Wildtype parental U-2-OS (WT) and an unedited U-2-OS cell line previously transfected with SLF2 plasmid 2 (control clone 1) are presented as controls. NBS1 was used as a loading control. The depicted experiment is representative of two independent repeats in which similar results were achieved.

(b) Immunoblot analysis of SLF2 expression in wildtype parental U-2-OS (WT) alongside several unedited U-2-OS cell lines previously transfected with SLF2 plasmid 2 (control clones 1 to 5) but confirmed to have WT genomic SLF2 sequence. NBS1 was used as a loading control. The depicted experiment is representative of two independent repeats in which similar results were achieved.

4.2.3 Identifying the SLF1 and SLF2 variants present in the CRISPR clones

To investigate the possible in-frame mutations in SLF2 clone 8 and 22, I used TOPO cloning to sequence each of the clone's SLF2 alleles individually. SLF2 is located on chromosome 10, of which U-2-OS cell lines have an estimated 3 copies according to the COSMIC database (Tate et al., 2019). Consistent with this, TOPO cloning indicated that clone 8 had three unique SLF2 alleles, all with duplications. g.11831_11849dupGATGAAACTGATGGCTCTT (p.Ser40fs3X) and g.11823_1232dupAACCGAGTGATGAAACTGATGGCTCTTCTGCAGGCTTGGCA CCTTCAAA (p.Asn411Lysfs3X) were predicted to lead to premature terminations, whilst g.11824_11871dupACCGAGTGATGAAACTGATGGCTCTTCTGCAGGCTTG GCACCTTCAAA (p.Ser410_Asn411insKPSDETDGSSAGLAPS) represented an in-frame duplication of 48bp (Figure 4.5a, Table 4.1).

Contrastingly, only two unique SLF2 sequences were observed for clone 22. Both had deletions, one which led to a premature downstream termination (g.11846_11862delTCTTCTGCAGGCTTGGC; p.Ser403Thrfs14X) and the other which was in-frame deletion (g.11827_11847delGAGTGATGAAACTGATGGCTC; p.Asp398_Ser404del) (Figure 4.5b, Table 4.1). A frequency ratio of 4:1 in-frame to out-of-frame deletions amongst the TOPO cloned sequences suggested that there may have been multiple alleles with the same in-frame mutation (p.Asp398_Ser404del). This could arise if a cleaved allele used another mutated sequence as a template during HR. Therefore, whilst all copies of SLF2 were successfully edited in both clones, in each case at least one allele contained an in-frame indel which may still be expressed.

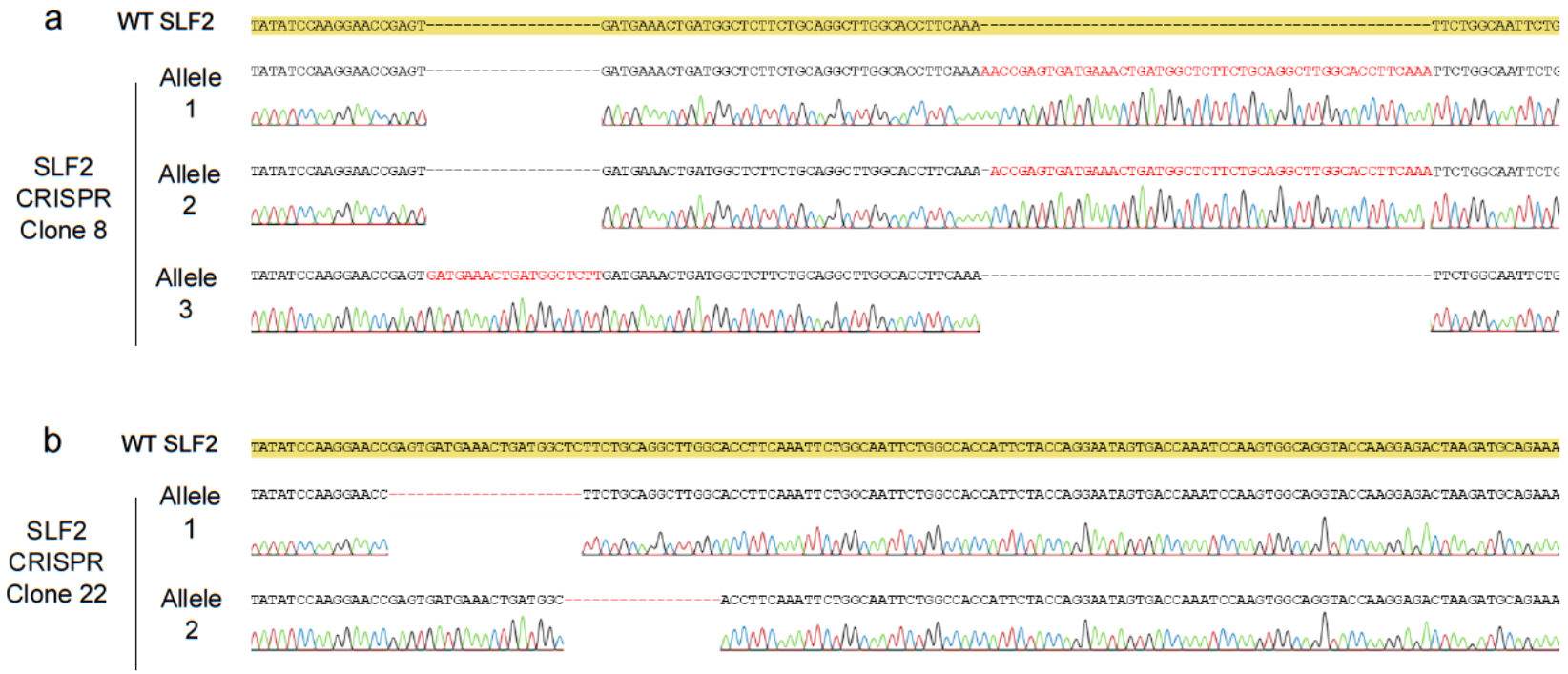


Figure 4.5 Genomic sequencing of U-2-OS SLF2 CRISPR cell lines identifies clones in which each SLF2 allele has been mutated

Genomic sequencing of U-2-OS SLF2 CRISPR clones 8 **(a)** and 22 **(b)**, assembled using CLC workbench. TOPO cloning was used to separate each allele and sequence them individually. Dashed red lines indicate deletions and red sequences indicate insertions. WT SLF2 is aligned in yellow.

Clone #	Allele #	Mutation
SLF2 Cl. 8	1	g.11823_1232dupAACCGAGTGATGAAACTGATGGCTCTTCTGCAGGCTTGGCACCTTCAAA, p.Asn411Lysfs3X
	2	g.11824_11871dupACCGAGTGATGAAACTGATGGCTCTTCTGCAGGCTTGGCACCTTCAAA, p.Ser410_Asn411insKPSDETDGSSAGLAPS
	3	g. 11831 11849dupGATGAAACTGATGGCTCTT, p.Ser403X
SLF2 Cl. 22	1	g.11827_11847delGAGTGATGAAACTGATGGCTC, p.Asp398_Ser404del
	2	g.11846 11862delTCTTCTGCAGGCTTGGC, p.Ser403Thrfs14X

Table 4.1 Genomic and predicted proteomic descriptions of each unique SLF2 variant identified in U-2-OS SLF2 mutant CRISPR clones 8 and 22.

These mutations occurred upstream of the putative SLF1 and SMC5/6 binding regions, so it was difficult to predict whether they would affect the protein's function. TOPO cloning also confirmed the presence of WT SLF2 alleles in clones 2, 9 and 14. The lack of WT SLF2 in clones 8 and 22, alongside their reduced levels of protein expression, meant that these cell lines were chosen for further experimentation.

TOPO cloning was also used to analyse the individual SLF1 alleles present in the three SLF1 CRISPR clones predicted to lack WT sequences. COSMIC predicted 3 copies of the gene on chromosome 5 (Tate et al., 2019). In line with this, three unique SLF1 alleles were identified in clone 12, two of which were duplications (g.12259_12278dupCATTTGCTTTCTAGGAAAGT) and (g.12240_12279dupTCCTTTGTGACTGACAGCTCATTTGCTTTCTAGGAAAGTG) and another which represented a deletion (g.12237_12278delAATTCCTTTGTGACTGACAGCTCATTTGCTTTCTAGGAAAGTGGATACTAACCAA) (Figure 4.6b, Table 4.2). However, because these mutations disrupted SLF1's 5' intron 3 and exon 4 boundary, involving the duplication or deletion of the acceptor splice site, further investigation was required to understand their impact upon RNA splicing.

Further TOPO cloning revealed that clone 6 had two unique SLF1 alleles, both with deletions. Again, the first occurred across the boundary of intron 3 and exon 4, with currently unknown impacts on splicing and protein sequence (g.12265_12273delCTTTCTAGG) (Figure 4.6a, Table 4.2). A secondary single base pair deletion, 15bp downstream of the start of exon 4, was predicted to lead to a frameshift and premature termination (g.12287del, p.Thr69Profs4X). A frequency

Figure 4.6 Genomic sequencing of U-2-OS SLF1 CRISPR cell lines identifies clones in which each SLF1 allele has been mutated

Genomic sequencing of U-2-OS SLF1 CRISPR clones 6, 12 and 16, assembled using CLC workbench. TOPO cloning was used to separate each allele and sequence them individually. Dashed red lines indicate deletions and red sequences indicate insertions. WT SLF1 exon 4 is highlighted in yellow, with SLF1 intron 3 highlighted blue.

Clone #	Allele #	Mutation
SLF1 Cl. 6	1	g.12265_12273delCTTTCTAGG
	2	g.12287del p.Thr69Profs4X
SLF1 Cl. 12	1	g.12259_12278dupCATTTGCTTTCTAGGAAAGT
	2	g.12240_12279dupTCCTTTGTGACTGACAGCTCATTTGCTTTCTAGGAAAGTG
	3	g.12237_12278delAATTCCTTTGTGACTGACAGCTCATTTGCTTTCTAGGAAAGT GGATACTAACCAA
SLF1 Cl. 16	1	g.12273del
	2	g.12276_12283delAGTGGATA p.Lys65Asnfs8X

Table 4.2 Genomic and, where appropriate, proteomic descriptions of each unique SLF1 variant identified in U-2-OS SLF1 CRISPR clones 6, 12 and 16.

ratio of 4:1 amongst the TOPO cloned sequences suggested that the cell line was likely to have at least two alleles with the g.12265_12273delCTTTCTAGG intron 3/exon 4 boundary deletion compared to the g.12287del mutation.

Finally, two unique SLF1 variants were characterised in clone 16. A single base pair deletion at the very start of exon 4 appeared likely to impact RNA splicing (g.12273del), whilst a downstream deletion of 8 base pairs introduced a putative premature stop codon (g.12276_12283delAGTGGATA, p.Lys65Asnfs8X) (Figure 4.6c, Table 4.2). Having been observed in an almost 1:1 ratio, it was difficult to establish if either was present in more than one allele.

To understand how these mutations might impact RNA splicing, I attempted to sequence cDNA around the CRISPR target sites. However, this was largely unsuccessful, with PCR attempts either failing to amplify the template cDNA or returning nonspecific sequencing results. Considering the inherent difficulties in validating the impacts of SLF1 CRISPR mutations on either an RNA or protein level, due to lack of available SLF1 antibody, I chose to focus my ongoing efforts on characterising the SLF2 clones.

4.2.4 SLF2 CRISPR cell lines phenocopy the replicative issues and genomic instability associated with SLF2 and SMC5 patient cell lines

To facilitate further investigation, SLF2 CRISPR mutant clones 8 and 22, referred to henceforth as SLF2 CRISPR clones A and B, were complemented with Myc-SLF2 cloned into pLVX-IRES-neo. Western blots confirmed the presence of Myc-SLF2 in the complemented cell lines, although in excess comparatively to endogenous SLF2 (Figure 4.7). To determine if the SLF2 CRISPR mutagenesis led to replicative dysfunction I used the DNA fibre assay to monitor replication dynamics (Figure 4.8).

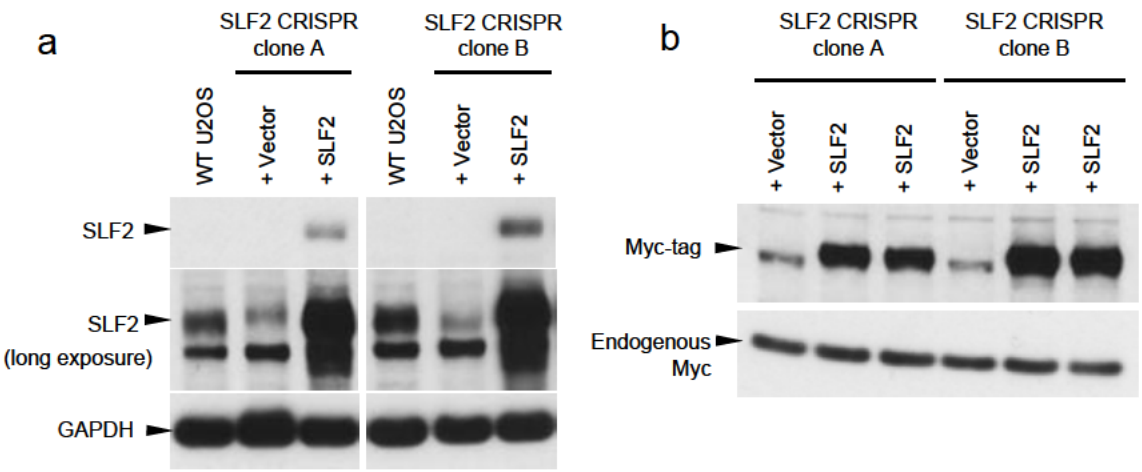


Figure 4.7 Complementation of U-2-OS SLF2 mutant CRISPR clones

(a) Representative immunoblot analysis of SLF2 expression in U-2-OS SLF2 CRISPR clones A and B infected with lentiviruses encoding Myc-tagged wildtype (WT) SLF2 or an empty vector. GAPDH was used as a loading control.

(b) Representative immunoblot analysis of Myc-SLF2 expression in U-2-OS SLF2 CRISPR cell lines infected with lentiviruses encoding Myc-tagged WT SLF2 (two tagged cell lines shown for each clone) or an empty vector. Endogenous c-Myc was used as a loading control.

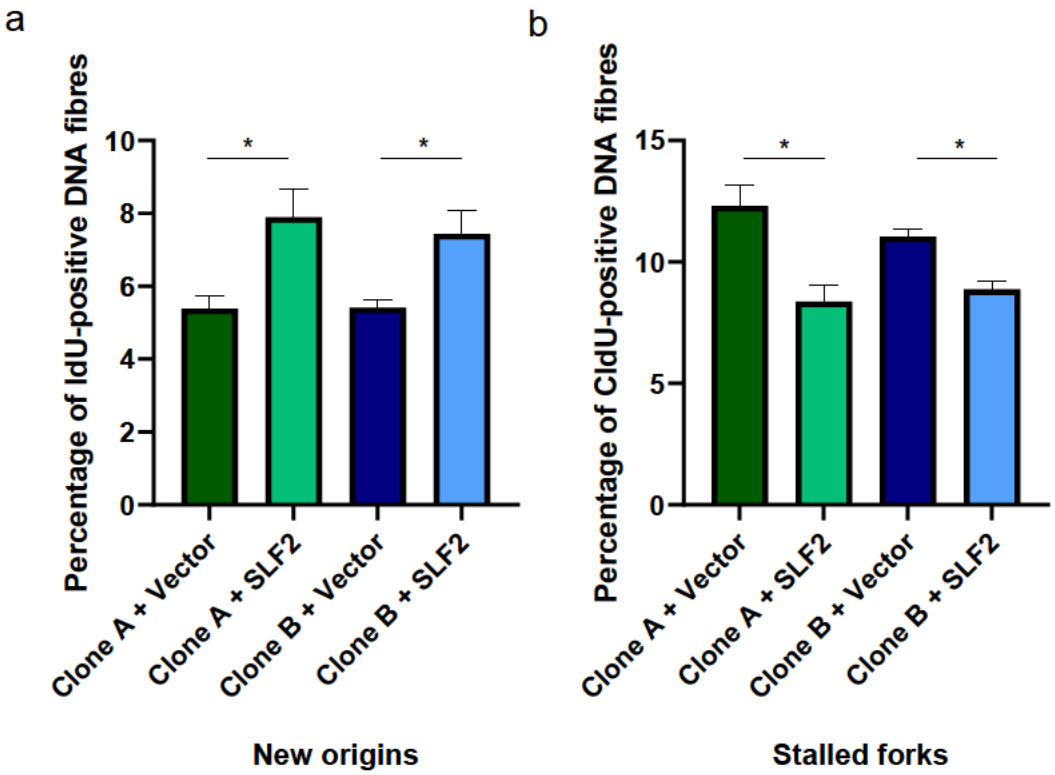


Figure 4.8 U-2-OS SLF2 CRISPR clones exhibit replicative dysfunction

DNA fibre analysis of U-2-OS SLF2 CRISPR clones A and B infected with lentiviruses encoding WT SLF2 or an empty vector. The percentage of new origins (a) and stalled forks (b) were quantified. n=4 independent experiments. A minimum of 1,600 fork structures were counted. Statistical differences were determined with an unpaired, 2-tailed, student's t-test. Error bars denote SEM.

For both clones, consistent with the results obtained from SLF2 patient derived LCLs and fibroblasts, I observed significant increases in the levels of new replication origins (Figure 4.8a), alongside significantly decreased levels of stalled forks (Figure 4.8b), when complemented with WT SLF2. Similarly, the WT SLF2 complemented CRISPR mutant clones exhibited significantly reduced levels of micronuclei (Figure 4.9), suggesting that CRISPR-mediated decrease in SLF2 expression contributed to increased levels of DNA damage and genome instability. This demonstrated that the targeted knockdown of SLF2 in U-2-OS cells phenocopies the increase in replication stress observed in the patient cell lines.

4.3 Discussion

The data in this chapter demonstrates that U-2-OS cell lines with targeted CRISPR mutations in SLF2 phenocopy the increases in replication stress and genomic instability observed in SLF2 and SMC5 patient derived cell lines. Considering that SLF2's predominant known function is in the recruitment of the SMC5/6 complex (Raschle et al., 2015), this provides further evidence that SLF2, and hence the integrity of the SMC5/6-SLF2-SLF1-RAD18 pathway, is important for efficient replication.

To further question this hypothesis, U-2-OS cell lines with CRISPR-mediated SLF1 mutations were generated. However, because many of these mutations spanned SLF1's 5' intron 3-exon 4 boundary, more work would be needed to fully understand their impact on splicing and protein expression. Multiple PCR primers, spanning the length of SLF1, would be needed to successfully amplify and sequence the clones' cDNA to determine if differential splicing has occurred. Considering that a lack of

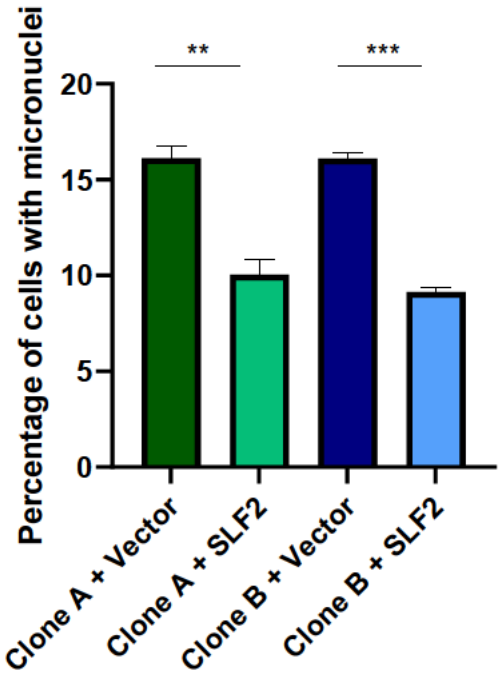


Figure 4.9 U-2-OS SLF2 CRISPR clones exhibit elevated numbers of micronuclei which reduce upon complementation

The percentage of cells with micronuclei in U-2-OS SLF2 CRISPR clones infected with lentiviruses encoding WT SLF2 or an empty vector. n=3 independent experiments. Error bars denote SEM. A minimum of 1,700 cells were counted. Statistical differences were determined with an unpaired, 2-tailed, student's t-test.

SLF1 antibodies prevented analysis of protein expression via immunoblotting, and given the difficulties encountering in amplifying the SLF1 mutant cDNA, it was decided not to continue to pursue the generation of the SLF1 CRISPR knockout clones.

Whilst the SLF2 CRISPR mutant clones were easier to validate, a complete knockout of the gene seemed to be difficult to achieve, as at least one expressed SLF2 variant was retained in each identified CRISPR cell line. Although it can't be ruled out that a complete knockout can be found by screening more potential clones, or by using different sets of CRISPR guides, one possibility is that a complete loss of SLF2 expression may be incompatible with cellular viability, a hypothesis that would be consistent with previous studies which have shown that null mutations in SMC5/6 subunits are embryonically lethal in mice and lead to death and senescence in cellular models (Ju et al., 2013, Jacome et al., 2015, Pryzhkova and Jordan, 2016, Venegas et al., 2020). Considering that SLF2 and SLF1 have been shown to function together as a heterodimer (Raschle et al., 2015), this may also suggest that complete loss of function SLF1 cell lines are also inviable and that the CRISPR clones generated here retain SLF1 protein expression.

Intriguingly, CRISPR mutagenesis was only achieved in U-2-OS cell lines. RPE1 cells generated very few colonies after single cell sorting, of which none were observed to have been successfully mutated by the CRISPR nickase system. It is conceivable that the reduced numbers of colonies reflect a loss of cellular viability due to the sorting process itself. Numerous factors contribute to cell recovery post sorting, including jet pressure, nozzle size and rescuing medium, and it is possible that these parameters required further optimisation. Considering that only 12% of analysed U-2-OS clones exhibited SLF2 mutations, and far fewer RPE1 CRISPR clones were

propagated and sequenced, the failure to generate an RPE1 CRISPR mutants may simply be because not enough cell lines could be recovered and analysed.

On the other hand, the inability of the RPE1 cell line to propagate efficiently after sorting may be caused by the successful mutagenesis of SLF2 and could indicate that the cells aren't able to tolerate the disruption to SLF2 expression in the same way as U-2-OS cells. Whilst RPE1 cells are karyotypically normal, the tumoral U-2-OS cell line is chromosomally highly altered with an increased number of SLF2 alleles (Tate et al., 2019). It is possible that U-2-OS, unlike RPE1, can lose expression of multiple copies of the gene before SLF2 expression reduces to levels which are cellularly lethal. SLF2 mutagenesis may also be tolerated in U-2-OS due to an attenuated p53 response and unstable G1/S checkpoint. This contrasts with RPE1, which retains WT p53 status and possesses fully functioning intact checkpoints (Chao et al., 2017). Previous studies have indicated that replicative dysfunction induced by a depletion of SMC5/6 components induces chromosome missegregation alongside other mitotic defects, including chromatin bridges (Gallego-Paez et al., 2014, Venegas et al., 2020). If the G1/S checkpoint is typically activated following this mitotic dysfunction and contributes to the loss of cellular viability observed when components of the complex are depleted, U-2-OS's genetic background could mean that the cells are able to avoid cell-cycle arrest and senescence. Consistent with this, the only previous study to generate an SLF2 knockout did so in Jurkat T-lymphocytes, a cell line which lacks p53 (Dupont et al., 2021).

A recent paper by Ian Hickson, however, shows that an inactivation of *TP53* failed to rescue the lethality of auxin-degron induced NSE4 or SMC6 degradation. Instead, the study hypothesised that alternatively pathways, such as cGAS-STING, trigger cell-

cycle arrest. (Venegas et al., 2020). Whilst further investigation is required to ask if the cGAS-STING pathway is involved in the cellular inviability induced by SMC5/6 depletion, previous work has shown that U-2-OS are STING-deficient and thereby unable to activate cGAS-mediated type I interferon expression (Chen et al., 2020). Hence, it may be that SLF2 depletion is better tolerated in U-2-OS cells due to deregulated cell signalling pathways.

Chapter 5:

Exploring the impact of disease associated mutations in the SMC5/6-SLF1/2-RAD18 pathway on mitosis

5.1 Introduction

In the previous two chapters I have demonstrated, using both patient derived cell lines and U-2-OS CRISPR mutants, that mutations in SLF2 and SMC5 are associated with increased levels of endogenous replication stress and genomic instability, consistent with the known roles of the SMC5/6 complex in S-phase (Ampatzidou et al., 2006, Bermudez-Lopez et al., 2010, Menolfi et al., 2015, Aragon, 2018, Palecek, 2018, Venegas et al., 2020, Agashe et al., 2021). As I will be discussing in this chapter, an unexpected phenotype was discovered upon the initial analysis of metaphase spreads prepared from peripheral blood lymphocytes from patient SLF2-P1. A substantial proportion of these peripheral blood lymphocyte metaphases exhibited large gains in chromosome number. These large increases were variable, with each abnormal metaphase having a different total number of chromosomes. This was particularly interesting, not only as significant changes in karyotype are not commonly observed in DNA replication disorders, but because this phenotype is reminiscent of a rare syndrome called mosaic variegated aneuploidy (MVA) which has been linked to mitotic dysfunction (Hanks et al., 2004, Matsuura et al., 2006, Snape et al., 2011, Yost et al., 2017, de Wolf et al., 2021, Santos-Simarro et al., 2021, Carvalhal et al., 2022).

Patients with MVA have diverse clinical phenotypes, but commonly present with growth deficiency, microcephaly and a predisposition to certain cancers. Studies have shown that affected individuals exhibit aneuploidies in over 5% of cells originating from different tissues, although these usually represent gains or losses of a single chromosome. However, similar to the phenotype seen in metaphases from patient SLF2-P1, the identity of the lost/gained chromosome/s is not uniform between the

aneuploid cells in MVA patients. Currently, MVA has only been observed in patients with mutations in five genes: CENATAC, BUB1B, CEP57, TRIP13, and BUB1 (Hanks et al., 2004, Matsuura et al., 2006, Snape et al., 2011, Yost et al., 2017, Santos-Simarro et al., 2021, de Wolf et al., 2021, Carvalhal et al., 2022). With the exception of CENATAC, a very recently identified minor splicesome component with unknown targets (de Wolf et al., 2021), these genes encode kinetochore proteins which localise to centromeres and have functions in the spindle assembly checkpoint, microtubular stabilisation and the recruitment of additional mitotic factors. Mutations in these genes are implicated in mitotic deregulation, including the perturbation of centrosome maturation, aberrant spindles, and cohesion defects, giving rise to the observed changes in chromosome numbers (Hanks et al., 2004, Matsuura et al., 2006, Snape et al., 2011, Yost et al., 2017, de Wolf et al., 2021, Santos-Simarro et al., 2021, Carvalhal et al., 2022).

The depletion of SMC5/6 complex components has been associated with issues in both replication and mitosis, including increases in the presence of lagging chromosomes and anaphase bridges in mitosis (Gallego-Paez et al., 2014, Venegas et al., 2020). Although one study has demonstrated that SMC5 depletion in mESC's is associated with an increase in polyploid cells (Pryzhkova and Jordan, 2016), few studies have demonstrated a conclusive link between SMC5/6 dysfunction and the development of aneuploidy. Consistent with this, chromosome gains or losses were not reported in cell lines obtained from patients with mutations in either NSE2 or NSE3 (Payne et al., 2014, van der Crabben et al., 2016). However, karyotyping of peripheral blood lymphocytes derived from individuals with mutations in NSE3 did reveal variable numbers of *de novo* supernumerary marker chromosomes; small, structurally

abnormal chromosomes which originate from copied regions from other chromosomes (van der Crabben et al., 2016).

Recent experiments in which SMC6 has been rapidly degraded in different phases of the cell cycle using an auxin-inducible degron tag have suggested that mitotic dysfunction associated with the loss of SMC5/6 is due to its functions in interphase, rather than mitosis, and mitotic abnormalities only arise when SMC5/6 subunits are depleted prior to S-phase (Venegas et al., 2020). This may indicate that mitotic abnormalities are linked solely to the functions of SMC5/6 in replication. However, this remains a point of contention as SMC5/6 has functional links to numerous mitotic factors. For instance, studies have suggested that the complex is important for regulating Top2α localisation at centromeric chromatin (Gomez et al., 2013, Gallego-Paez et al., 2014). In RPE1 cells, SMC5 depletion has been observed to induce aberrant Top2α distribution, alongside increases in centromeric PICH UFBs, suggesting that SMC5/6 functions alongside the topoisomerase to limit topological DNA entanglements (Gallego-Paez et al., 2014). The depletion of the SMC5/6 complex is also associated with a decrease in condensin at centromeric regions in mitosis, as well as the aberrant distribution of kinetochore factors PLK1 and Aurora B (Gallego-Paez et al., 2014, Pryzhkova and Jordan, 2016). Whilst these latter two proteins have functions in the SAC, proper microtubule orientation and the timely dissociation of cohesin from chromatid arms during prophase, they are also known to coordinate the binding and supercoiling activity of condensin (Maresca, 2011, Thadani et al., 2012, Combes et al., 2017). This may implicate improper chromosome condensation, as well as checkpoint regulation, as drivers of missegregation in SMC5/6 depleted cell lines. However, discrepancies between cell-types and systems,

as well as a fundamental lack of mechanist insight, means that the link between SMC5/6 and faithful mitosis is poorly understood.

5.1.2 Chapter Aims

In this chapter I aim to further explore the karyotypic abnormalities observed in the peripheral blood lymphocytes derived from SLF2-P1 and determine if these phenotypes are common to other SLF2 and SMC5 patients. Furthermore, as SMC5/6 has been linked to numerous mitotic processes that can give rise to abnormalities during cell division if impaired, including chromosome cohesion and condensation, I will investigate whether any of these putative functions of SMC5/6 are defective in the patient LCLs and fibroblasts, and whether these problems could give rise to changes in chromosome number.

5.2 Results

5.2.1 SLF2 and SMC5 patient cells exhibit 'mosaic variegated hyperploidy' and mitotic abnormalities

To understand the extent of the aneuploidy observed in SLF2-P1, and its frequency amongst the other SLF2 and SMC5 patients, I quantified the number of chromosomes observed in metaphase spreads prepared from patient derived peripheral blood lymphocytes. In the blood samples that were obtained from four patients (SLF2-P1, SLF2-P2, SLF2-P3 and SMC5-P7), a significant subset of metaphases from all patients exhibited increases in total chromosome number (Figure 5.1a). The numbers of chromosomes gained varied dramatically and were often very large, with up to 180 chromosomes observed per metaphase. Analysis of peripheral blood lymphocyte

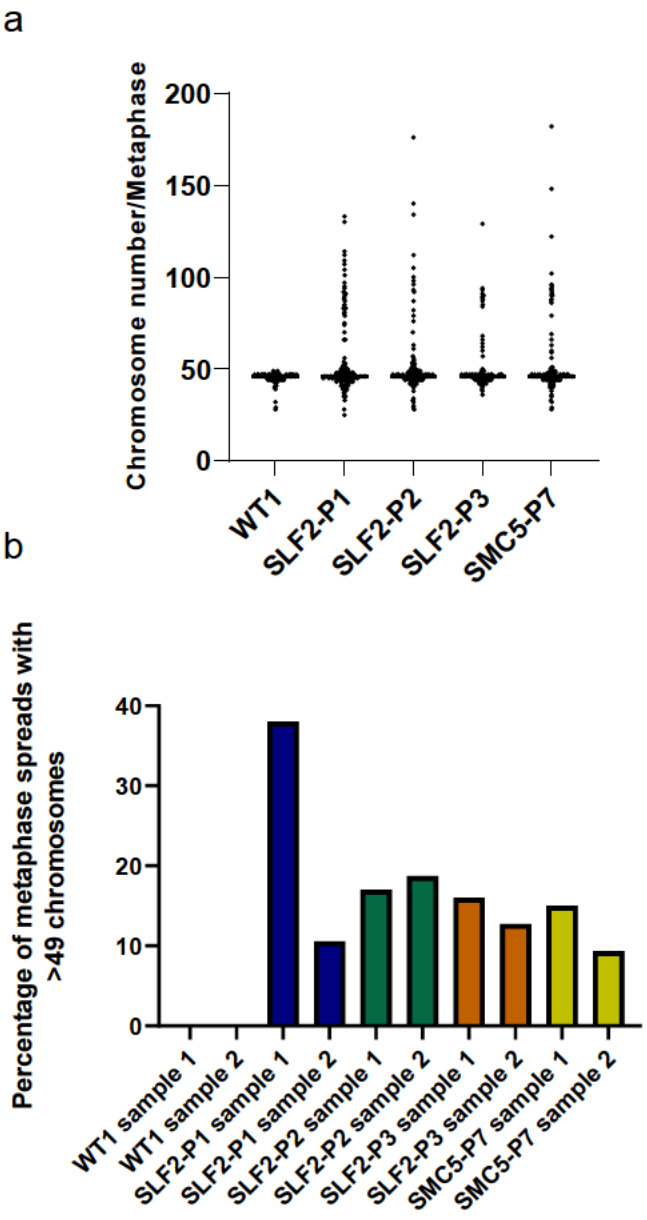


Figure 5.1. A proportion of metaphases from SLF2 and SMC5 patient peripheral blood samples exhibit increased numbers of chromosomes

(a) Quantification of the numbers of chromosomes per metaphase in peripheral blood lymphocytes from patients with mutations in SLF2 or SMC5, as well as an unrelated, healthy individual (WT1). In total >185 metaphases were counted from 2 independent blood samples.
(b) The percentage of metaphases analysed in (a) with >49 chromosomes for each separate blood sample.

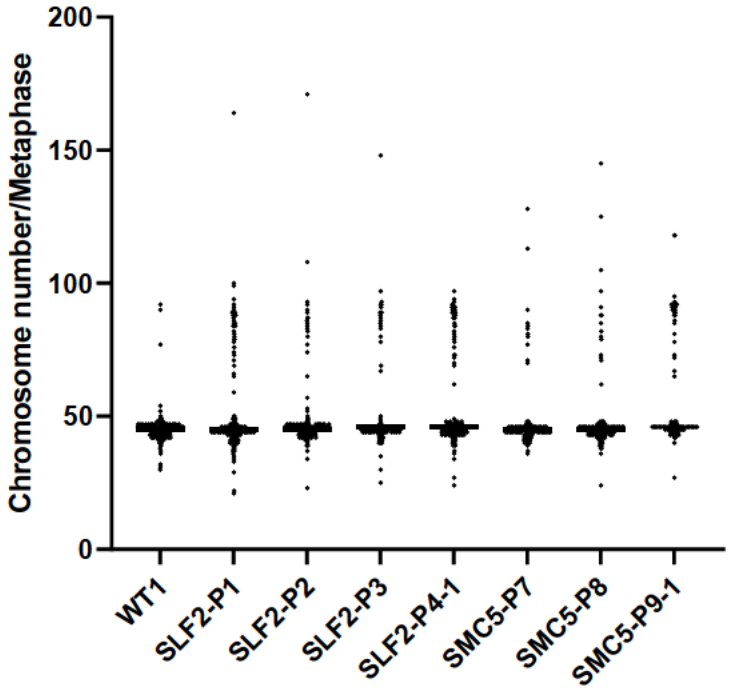
metaphases from a normal individual suggested that experimental variance accounted for a range of 43-49 chromosomes in healthy cells. Whilst I failed to observe any metaphases prepared from WT control cell lines with chromosome numbers exceeding 49, >14% of metaphases from each patient sample possessed more than 49 chromosomes (Figure 5.1b), suggesting that they were likely to be biologically significant. Interestingly, a subset of metaphases from SLF2-P3 and SMC5-P7 patients had total numbers of chromosomes that clustered around 92, indicating a whole genome duplication event may have occurred for these cells.

As it was not possible to obtain a blood sample for all SLF2 and SMC5 patients, but we were instead fortunate to have immortalised LCLs from these patients, I also quantified the number of chromosomes observed in metaphase spreads prepared from all patient LCLs. As with the peripheral blood lymphocytes, a distinct proportion of patient metaphases exhibited variable gains in chromosome number (Figure 5.2). Again, these increases were frequently large, with a total chromosome number of up to 171. For all patient LCLs, the proportion of metaphase spreads with a total chromosome number exceeding 49 was significantly higher than in WT lines (Figure 5.2b). Considering the presence of variable gains in chromosome numbers in a subset of metaphases from SLF2 and SMC5 patient cells, and the fact that the large numbers of chromosomes gained made this phenotype distinct from the previously reported MVA, we named this karyotypic abnormality ‘mosaic variegated hyperploidy’ (MVH).

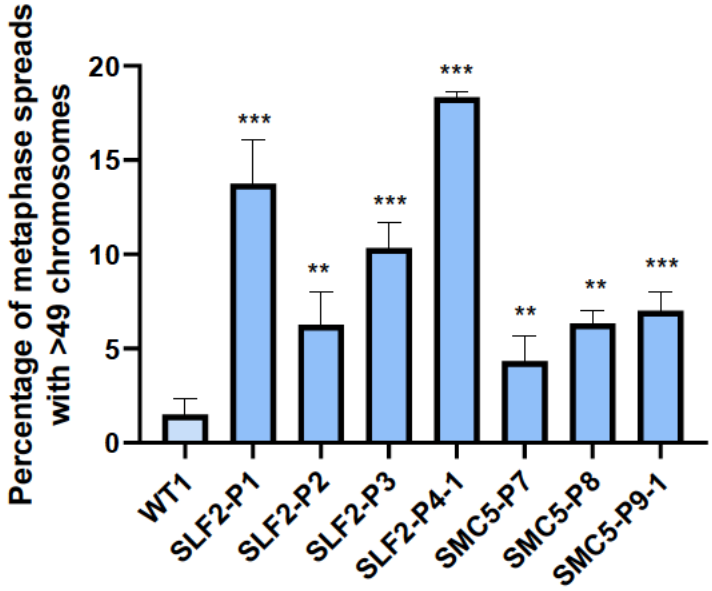
Figure 5.2. A proportion of metaphases from SLF2 and SMC5 lymphoblastoid cell lines exhibit increased numbers of chromosomes (data shown overleaf)

(a) Quantification of the numbers of chromosomes per metaphase in lymphoblastoid cell lines derived from patients with mutations in SLF2 or SMC5, as well as an unrelated, healthy individual (WT). n=3 independent experiments. 300 metaphases were counted.
(b) The average percentage of metaphases analysed in (a) with >49 chromosomes. Statistical analysis was performed using an unpaired, 2-tailed, student's t-test.
(c) Representative images of metaphase spreads from SLF2-P1 and control LCLs. The number of chromosomes in shown in red.

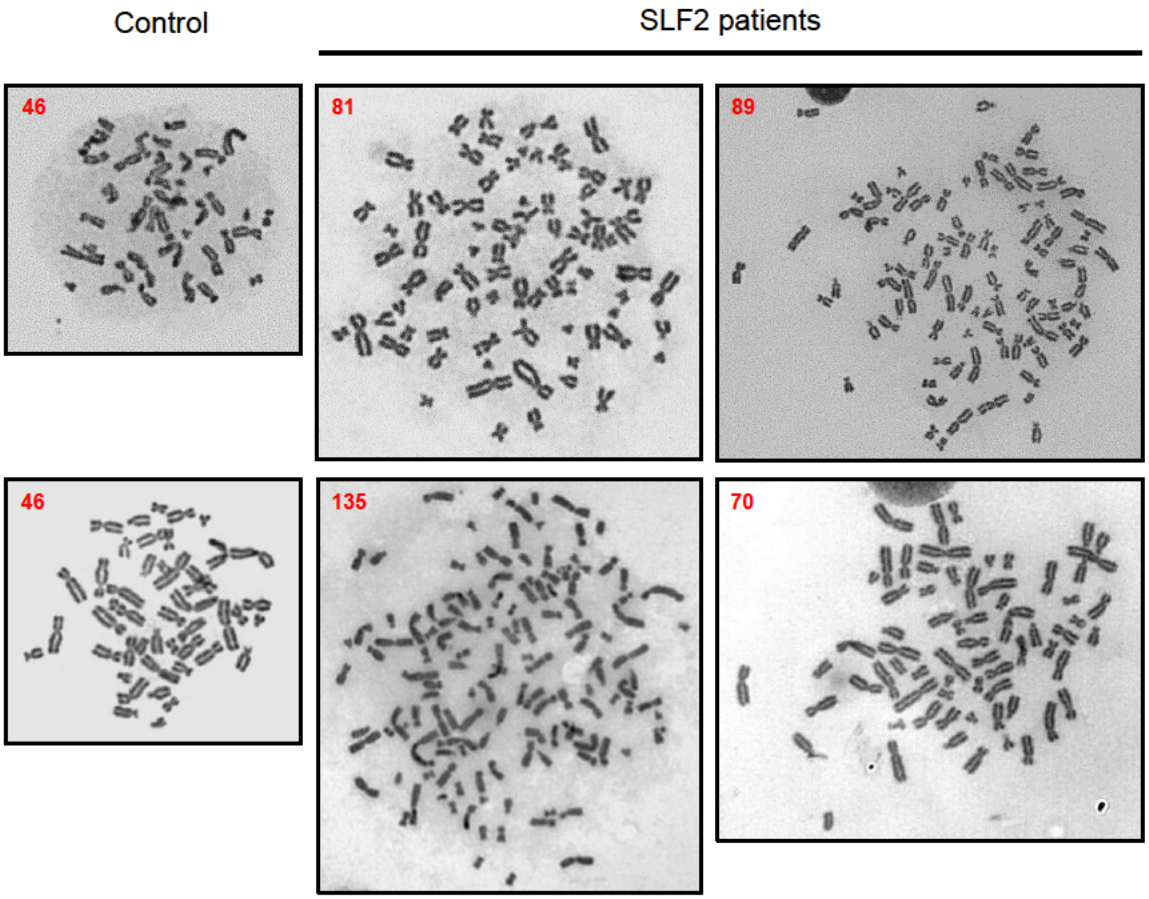
a


Chromosome number/Metaphase
200
150
100
50
0
WT1
SLF2-P1
SLF2-P2
SLF2-P3
SLF2-P4-1
SMC5-P7
SMC5-P8
SMC5-P9-1

b


Percentage of metaphase spreads with >49 chromosomes
20
15
10
5
0
***
**
***
***
**
**
***
WT1
SLF2-P1
SLF2-P2
SLF2-P3
SLF2-P4-1
SMC5-P7
SMC5-P8
SMC5-P9-1

C



As seen with metaphases prepared from SLF2-P3 and SMC5-P7 blood samples, there seemed to be a substantial proportion of metaphases prepared from patient LCLs that had a total chromosome number clustering around 92 (Figure 5.3a). The presence of an elevated number of tetraploid cells may suggest that polyploidy, which maintains genomic balance, is more beneficial for survival in comparison to other aneuploidies (Birchler and Veitia, 2012). Alternatively, it may be indicative of a cytokinetic defect, in which mitosis is completed but the cell does not divide. This would double the chromosomal content of the cell but also lead to centrosome amplifications and further missegregation (Normand and King, 2010). To investigate if cytokinetic defects were present in SLF2-P1 and SMC5-P8 LCLs, I used immunofluorescence to stain for the sodium potassium ATPase, which was used as a marker of the plasma membrane, and quantified the proportion of interphase cells with multiple nuclei as visualised by DAPI (Figure 5.3b, Figure 5.3c). Cells treated with Wiskostatin, an inhibitor of cytokinesis, were used as a positive control to induce binucleated cells. Quantifications confirmed that there were very low levels of spontaneous binucleated cells in both WT and patient LCLs, which didn't increase when exogenous replication stress was added via 24 hours treatment with APH. Therefore, I concluded that there was no obvious indication of a cytokinetic defect in SLF2 and SMC5 patient cell lines, and that the increases in chromosome number observed in patient cell lines were unlikely to be caused by a failure of cell division.

Having established that SLF2 and SMC5 patient derived cell lines exhibit MVH, I sought to understand if they presented with additional mitotic abnormalities. To question this, I used immunofluorescence microscopy to quantify the levels of DAPI-positive anaphase bridges and lagging chromosomes in the patient derived

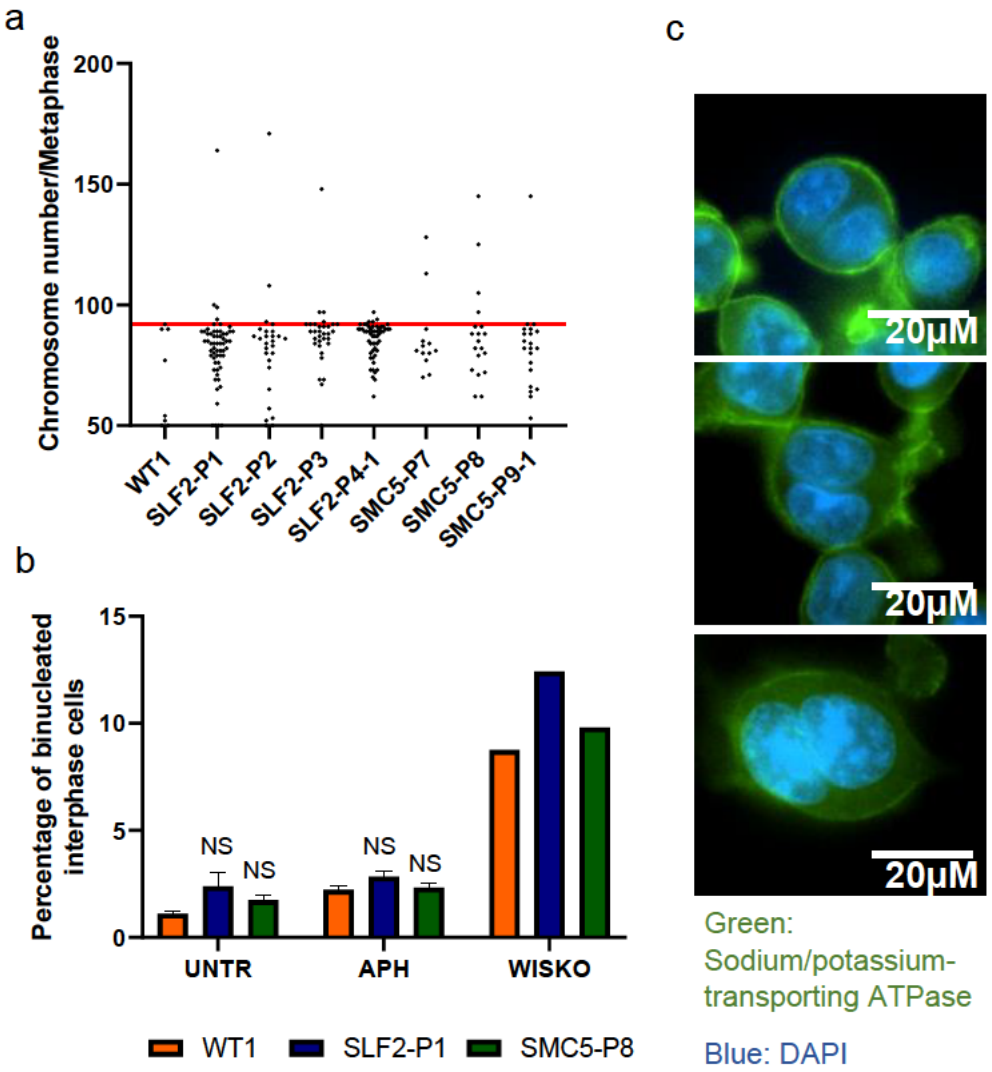


Figure 5.3. SLF2 and SMC5 patient lymphoblastoid cell lines do not exhibit increased levels of binucleated cells

(a) Graph representing the numbers of metaphases in SLF2 and SMC5 patient lymphoblastoid cell lines (LCLs), and an unrelated healthy (WT1) control cell line, that possess >49 chromosomes (data is from the same experiment as Figure 5.2a). 92 chromosomes per metaphase is indicated by the red line. n=3 independent experiments. 300 metaphases were counted.

(b) Quantification of the average percentage of binucleated cells in SLF2-P1, SMC5-P8 and WT1 LCLs, with and without 24 hours treatment with 250 nM aphidicolin (APH). n=3 independent experiments. >900 cells were counted (300 cells per repeat). Statistical analysis was performed using an unpaired, 2-tailed, student's t-test. Cells that were treated with 5μM wiskostatin for 24 hours treatment were used as a positive control for binucleation.

(c) Representative immunofluorescent images of the SLF2-P1 LCL showing binucleated cells. DAPI is shown in blue and Sodium/potassium-transporting ATPase in green.

fibroblasts. Whilst anaphase bridging was a rare event, meaning I could not robustly determine whether levels decreased upon the re-expression of WT protein, lagging chromosomes were observed to significantly decrease in all patient lines upon complementation (Figure 5.4). In support of this, SLF2 CRISPR mutant cell lines exhibited levels of lagging chromosomes which significantly reduced with WT SLF2 re-expression (Figure 5.5). Together, this shows that the disruption of SLF2 or SMC5 increases levels of spontaneous chromosome missegregation.

To understand the extent to which mitotic issues contribute to the genome instability observed in the patient cell lines, I considered the causes of the elevated levels of micronuclei seen in SLF2 and SMC5 patient cells, as described in chapter three. Using immunofluorescence microscopy, I analysed the micronuclear content of the complemented fibroblasts. These extranuclear bodies primarily form from chromosome fragments or lagging chromosomes that are not properly segregated into daughter nuclei (Fenech et al., 2011). To attempt to differentiate between the two scenarios, I immunostained for 53BP1 and CENPA. Micronuclei positive for 53BP1 foci would mark the presence of unrepaired DNA DSBs and suggest that DNA damage contributed to their formation. Conversely, micronuclei positive for CENPA, a nucleosome protein which directs the formation of kinetochores, would indicate the presence of either partial or whole chromosomes containing centromeres that hadn't segregated with the rest of the genetic material.

Whilst the majority of micronuclei in all cell lines were CENPA negative, indicating the presence of chromosome fragments, a substantial proportion of micronuclei were CENPA positive, which decreased significantly upon the re-expression of WT SLF2 and SMC5 protein (Figure 5.6a). This suggests that despite a majority of micronuclei

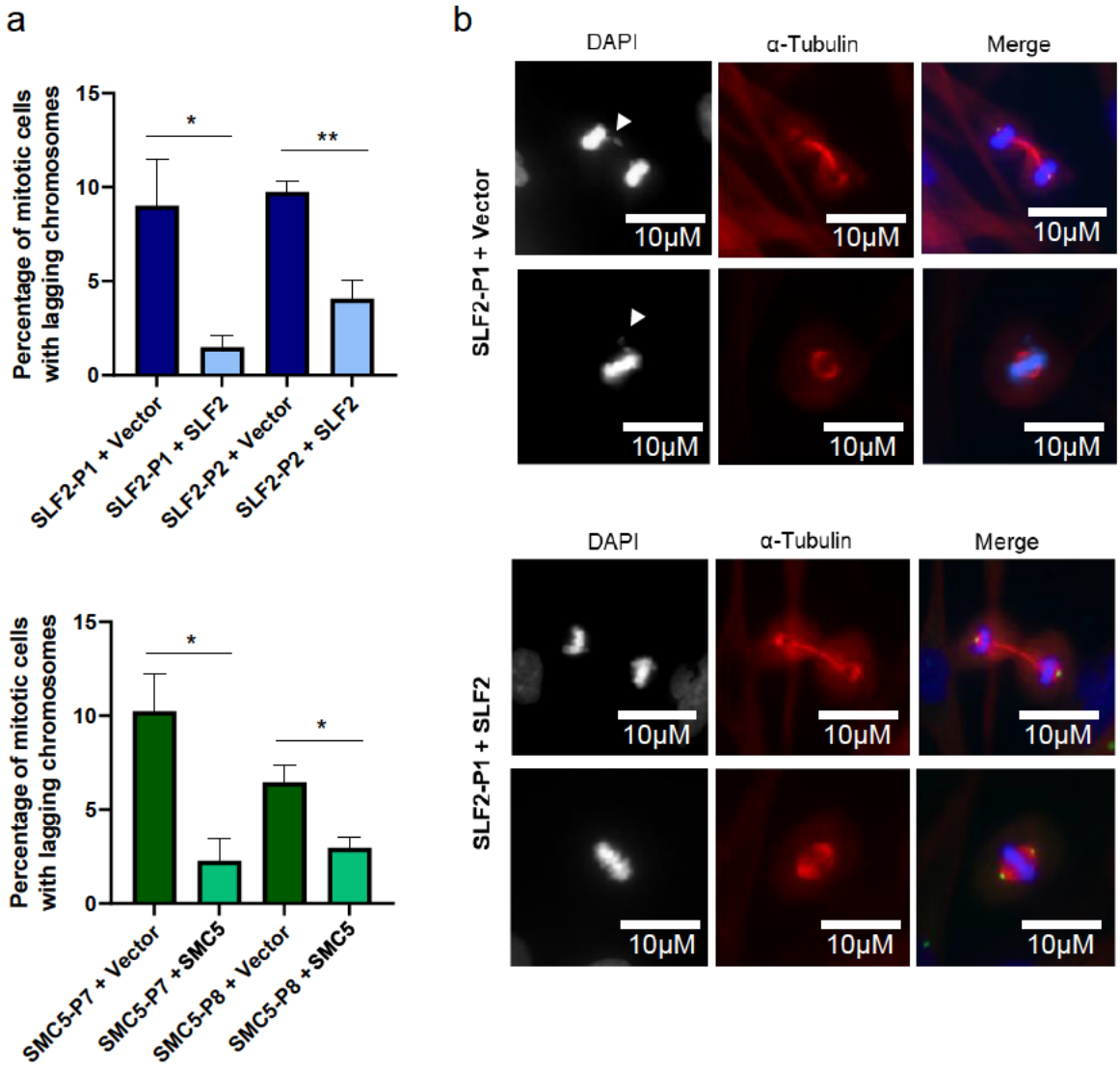


Figure 5.4. SLF2 and SMC5 patient derived fibroblasts exhibit elevated levels of lagging chromosomes which reduce upon complementation

(a) Quantification of the average number of mitotic cells with lagging chromosomes in SLF2 (above) and SMC5 (below) patient fibroblasts complemented with a lentivirus encoding WT SLF2, WT SMC5, or an empty vector. n=3 independent experiments. >250 mitotic cells were counted. Statistical differences were determined with an unpaired, 2-tailed, student's t-test.

(b) Representative immunofluorescent images of SLF2-P1 derived fibroblasts, complemented with a lentivirus encoding WT SLF2, or an empty vector. DAPI shown in blue and α-tubulin in red. Lagging chromosomes are indicated by white arrows.

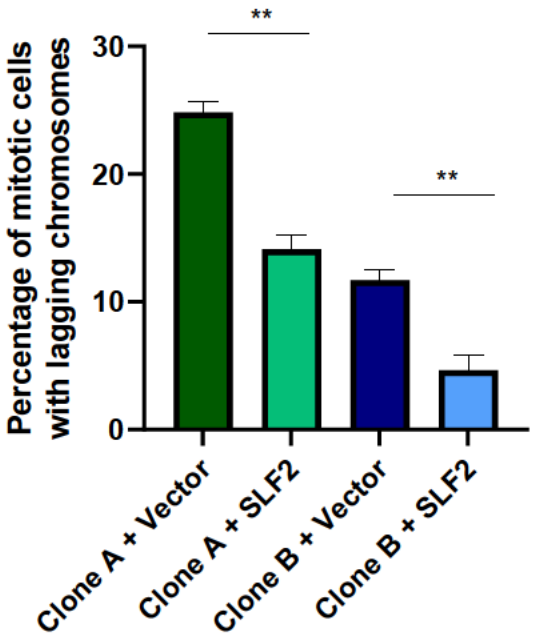


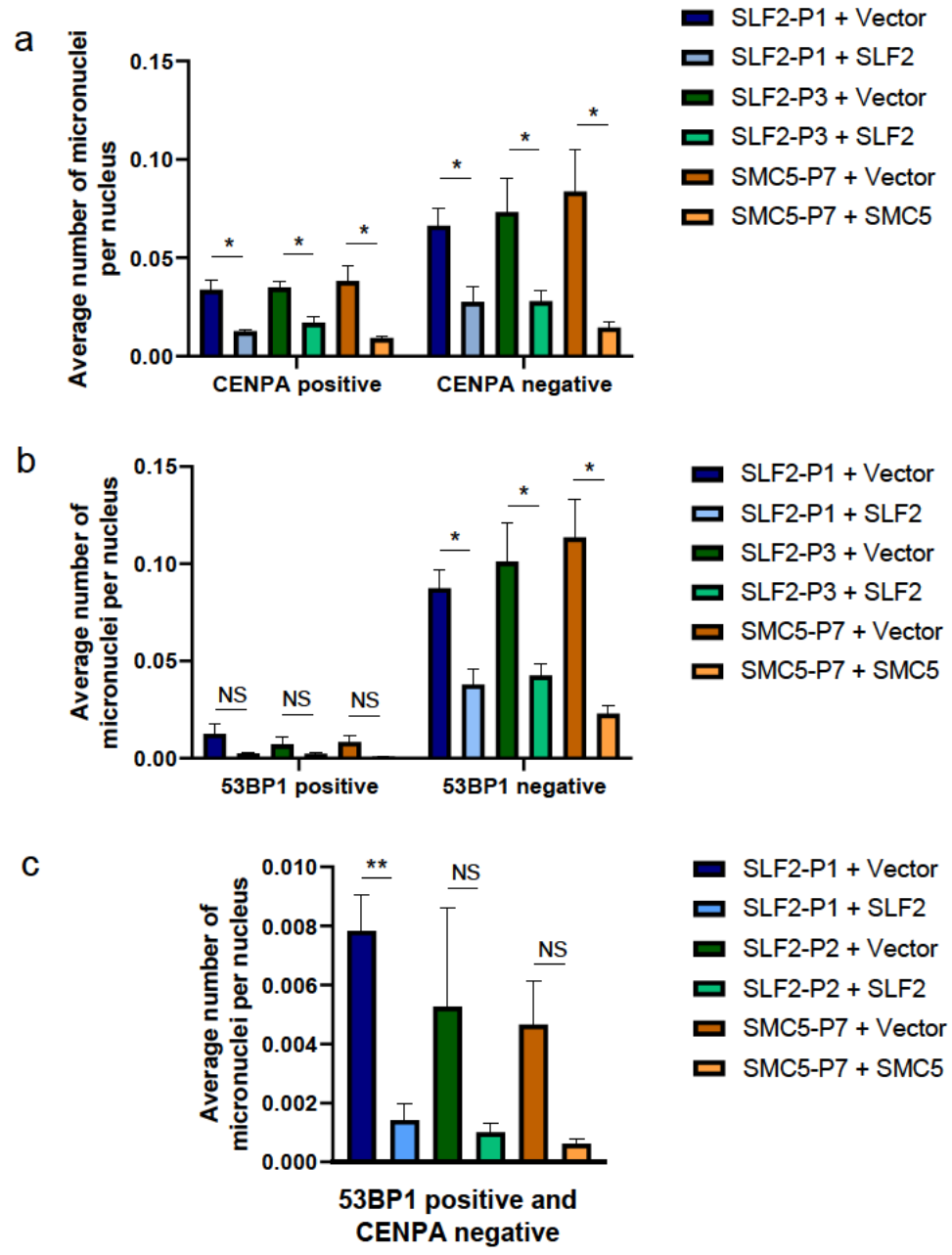
Figure 5.5. SLF2 CRISPR U-2-OS cell lines exhibit elevated levels of lagging chromosomes which reduce upon complementation

Quantification of the average number of mitotic cells with lagging chromosomes in U-2-OS SLF2 CRISPR cell lines complemented with a lentivirus encoding WT SLF2 or an empty vector. n=3 independent experiments. >250 mitotic cells were counted. Statistical differences were determined with an unpaired, 2-tailed, student's t-test.

Figure 5.6. Investigating the source of micronuclei in SLF2 and SMC5 patient fibroblasts (shown overleaf)

(a) Quantification of the average number of CENPA positive and negative micronuclei in SLF2 and SMC5 patient fibroblasts complemented with a lentivirus encoding WT SLF2, WT SMC5, or an empty vector. n=3 independent experiments. >185 micronuclei were counted.
(b) From the data presented in (a), quantification of the average number of 53BP1 positive and negative micronuclei in SLF2 and SMC5 patient fibroblasts complemented with a lentivirus encoding WT SLF2, WT SMC5, or an empty vector.
(c) From the data presented in (a), quantification of the average number of 53BP1 positive CENPA negative micronuclei in SLF2 and SMC5 patient fibroblasts complemented with a lentivirus encoding WT SLF2, WT SMC5, or an empty vector.
(d) Representative immunofluorescent images indicating 53BP1 positive (left, 53BP1 in green) and CENPA positive (right, CENPA in red) micronuclei. DAPI shown in blue.

Statistical differences were determined with an unpaired, 2-tailed, student's t-test.


a
Average number of micronuclei per nucleus
0.15
0.10
0.05
0.00
CENPA positive
CENPA negative
SLF2-P1 + Vector
SLF2-P1 + SLF2
SLF2-P3 + Vector
SLF2-P3 + SLF2
SMC5-P7 + Vector
SMC5-P7 + SMC5
b
Average number of micronuclei per nucleus
NS
53BP1 positive
53BP1 negative
c
Average number of micronuclei per nucleus
0.010
0.008
0.006
0.004
0.002
0.000
**
NS
53BP1 positive and CENPA negative
SLF2-P2 + Vector
SLF2-P2 + SLF2
d
20μM
Red: CENPA
Green: 53BP1
Blue: DAPI

likely forming due to acentric fragments, missegregation of chromosomes with a centromere is a significant driver of micronuclei formation. Further analysis revealed that most of the micronuclei were also negative for 53BP1 staining, with very few observed to be 53BP1 positive (Figure 5.6b). This indicated that there were few unrepaired DNA DSBs present in the micronuclei at the time of the experiment. However, the small level of CENPA negative, 53BP1 positive micronuclei, which likely represent acentric chromosome fragments with unrepaired DSBs, decreased in fibroblasts complemented with WT SLF2 or SMC5 (Figure 5.6c). Considering the levels of micronuclei that lack centromeres, it is probable that unrepaired DNA damage also contributes to the formation of micronuclei in SLF2 and SMC5 patient cell lines. The lack of 53BP1 staining in the majority of the acentric micronuclei may be due to the defective signalling of DSBs in micronuclei, or it may be that the micronuclei formed from an unrepaired DSB which has since been repaired. Thus, it is likely that the micronuclei in SLF2 and SMC5 patient cells arise from a combination of unrepaired DNA damage and chromosome missegregation.

5.2.2 SLF2 and SMC5 patient cells exhibit premature chromatid separation

Alongside the increased levels of MVH, analysis of metaphases prepared from patient peripheral blood samples also showed elevated numbers of cells with “railroad” chromosomes, in which sister chromatids align in parallel and fail to constrict at the centromere (Figure 5.7). A phenotype often observed in cohesinopathies, this chromosome abnormality reflects premature sister chromatid separation (PCS)

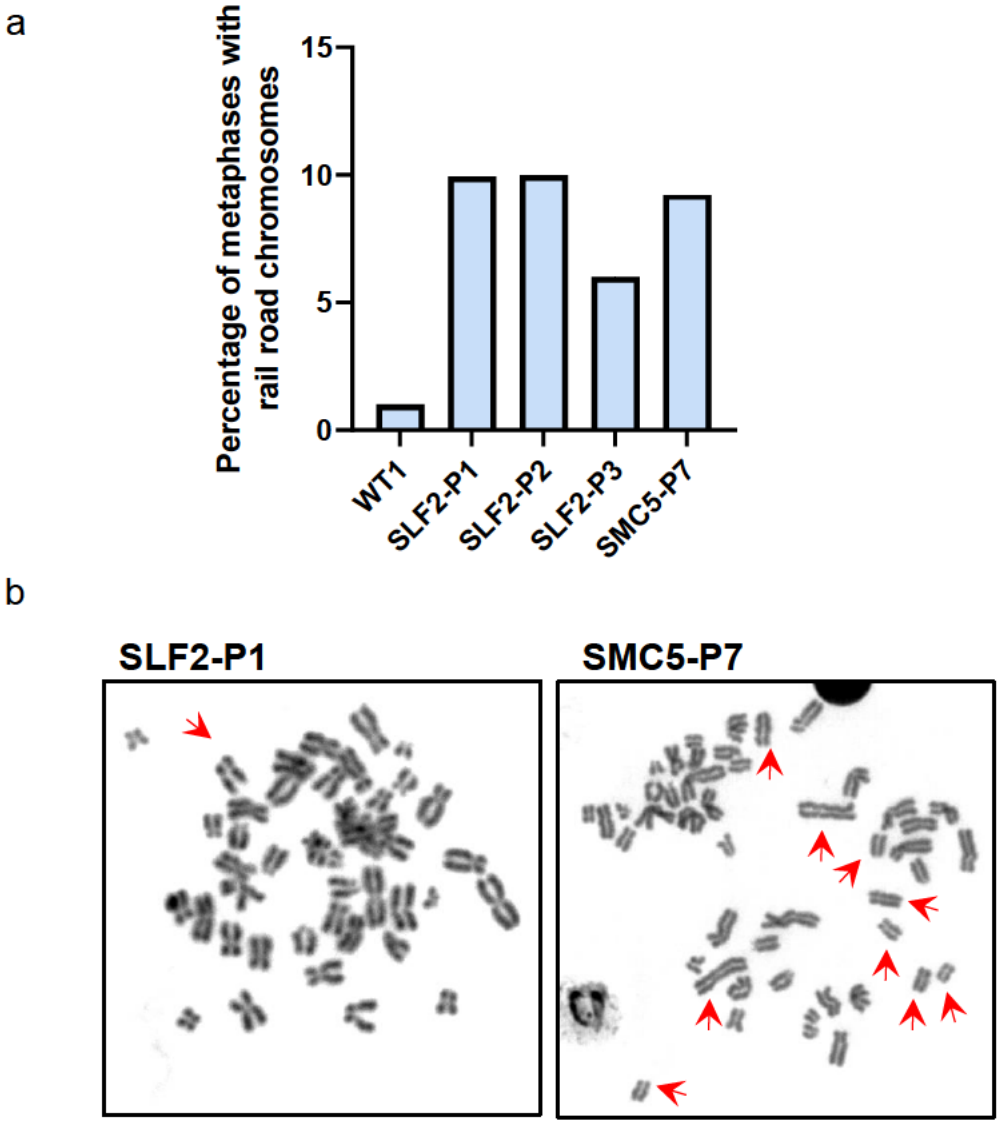


Figure 5.7. SLF2 and SMC5 patient peripheral blood lymphocytes exhibit elevated levels of railroad chromosomes

(a) Quantification of the average percentage of metaphases with railroad chromosomes in peripheral blood lymphocytes derived from patients with mutations in SLF2 or SMC5, as well as an unrelated, healthy individual (WT). >380 metaphases were counted in total from 2 independent blood samples.
(b) Representative images of metaphase spreads from SLF2-P1 and SMC5-P7. Railroad chromosomes are indicated by the red arrows.

usually related to cohesion defects (Kaur et al., 2005, Schulz et al., 2008, de Lange et al., 2015, van Schie et al., 2020, da Costa Almeida et al., 2020). Whilst aneuploidies aren't consistently observed in patients with cohesinopathies, cohesion defects can induce chromosome missegregation (Barbero, 2013). As such, mild gains and losses of chromosomes have been reported in studies concerning individuals with Roberts Syndrome (Jabs et al., 1991, Gerkes et al., 2010); a widely studied cohesinopathy caused by a defect in sister chromatid cohesion due to mutations in the establishment of the cohesion 1 homologue 2 (ESCO2) gene (Jabs et al., 1991, Vega et al., 2005a, Gerkes et al., 2010), which is critical for the establishment of sister chromatid cohesion during S-phase (Alomer et al., 2017). Interestingly, investigations in *Saccharomyces cerevisiae* have demonstrated an association between cohesin and SMC5/6, with SMC5/6 observed to be enriched at centromeric heterochromatin cohesin binding sites (Pebernard et al., 2008b, Jeppsson et al., 2014). Additionally, experiments in human cells have demonstrated that SMC5/6 recruits cohesion to DSBs and promotes HR (Potts et al., 2006), with SMC5/6 acting to stimulate the SUMOylation of cohesin's RAD21^{Scc1} subunit (Potts et al., 2006, McAleenan et al., 2012). Taken together, this may suggest that SMC5/6 is also important for maintaining proper sister chromatid cohesion during the cell cycle.

Having said this, railroad chromosomes were not evident in the patient LCL metaphase spreads. Considering that immortalised cell lines may present cohesion defects differently to primary cells, I sought to explore if the patient LCLs instead exhibited elevated levels of cohesion fatigue: asynchronous chromatid separation that occurs when the length of metaphase is extended during mitosis. During this mitotic delay, pulling forces originating from microtubule-kinetochore interactions eventually

disrupt the cohesin complex, although the precise mechanism through which this occurs is unknown (Daum et al., 2011, Gorbsky, 2013). Studies have demonstrated that compromised cohesin accelerates this process (Sapkota et al., 2018), indicating that cellular susceptibility to cohesion fatigue can be used as a readout of cohesin function. To test this, I analysed metaphase spreads to compare the extent of cohesion fatigue in the patient LCLs to WT controls at specific time points. Using MG132 to induce mitotic arrest during metaphase, I observed a significant increase in the proportion of SLF2 and SMC5 patient metaphases which exhibited cohesion fatigue 4 hours post drug treatment (Figure 5.8). This suggested that the patient LCLs have cohesion defects which can be observed if mitosis is delayed.

Next, I questioned if the disruption of SLF2 or SMC5 impacted the localisation of mitotic factors associated with SMC5/6. I hypothesised that the improper recruitment of these factors might contribute to cohesion defects. I first investigated the localisation of kinases Aurora B and PLK1, both of which are implicated in the proper regulation of cohesin and condensin and are required for faithful chromosome segregation (Gimenez-Abian et al., 2004, Dai et al., 2006, Lipp et al., 2007, Abe et al., 2011). Previous investigations have shown that a conditional loss of SMC5 in mESCs leads to the loss of PLK1 at pericentromeric regions. Similarly, Aurora B is redistributed from the centromere to the distal ends of chromosome arms (Pryzhkova and Jordan, 2016). However, using immunofluorescence to compare SLF2 and SMC5 patient and WT LCLs, I failed to observe a gross difference in PLK1 localisation in mitotic cells (Figure 5.9). This localisation was also unaffected by increased levels of replication stress, as induced by 24 hours chronic treatment with APH.

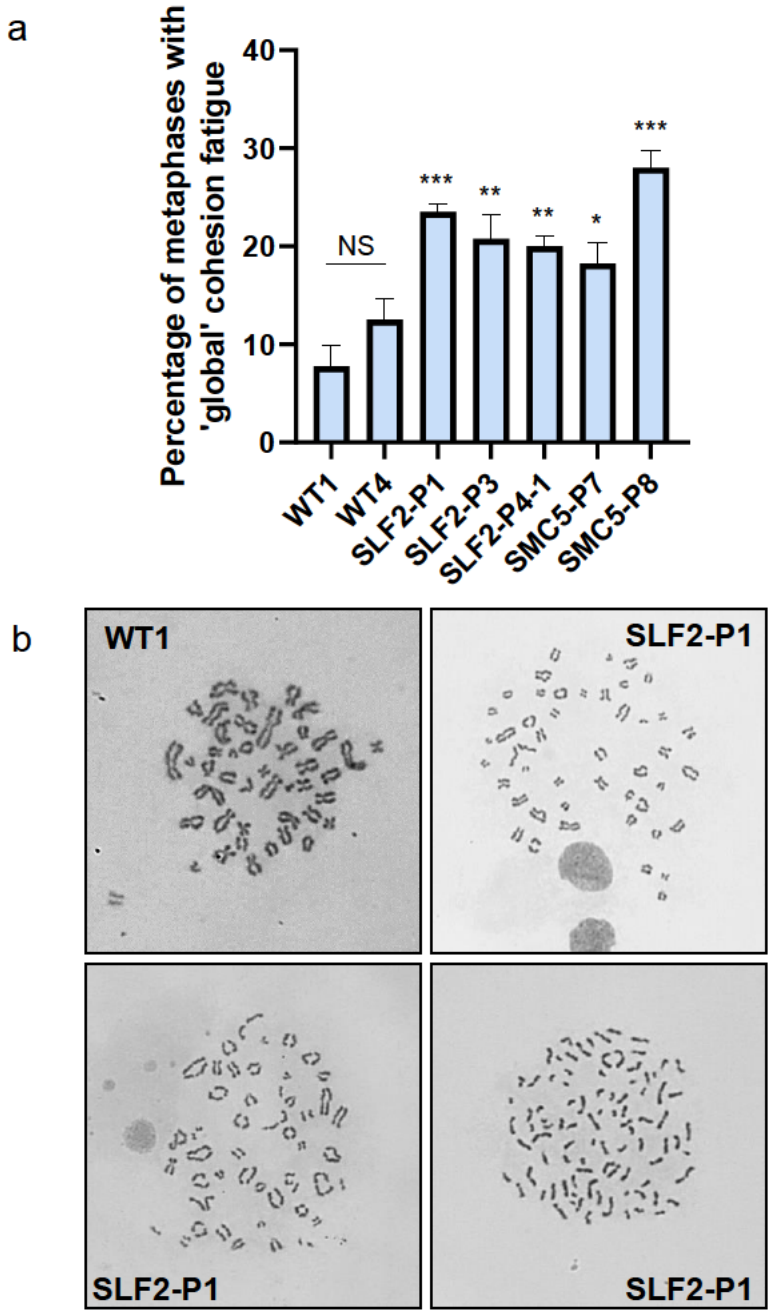


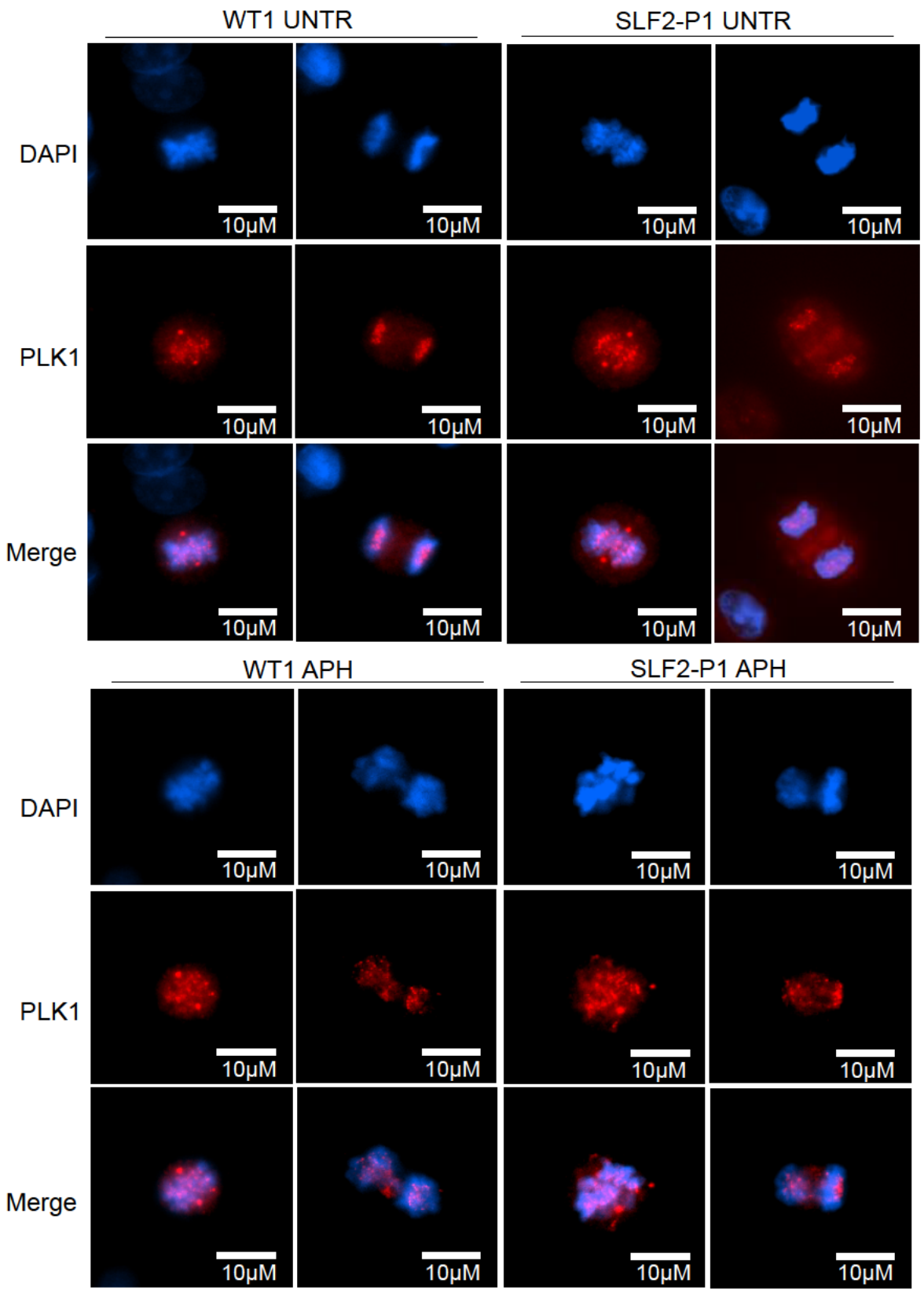
Figure 5.8. SLF2 and SMC5 patient lymphoblastoid cell lines exhibit elevated levels of cohesion fatigue

(a) Quantification of the average percentage of metaphases exhibiting cohesion fatigue across all chromosomes ('global') in lymphoblastoid cell lines (LCLs) derived from patients with mutations in SLF2 or SMC5, as well as from unrelated, healthy individuals (WT1 and WT4). n=4 independent experiments. 400 metaphases were counted. Statistical differences were determined with an unpaired student's, 2-tailed, t-test.

(b) Representative images of metaphase spreads demonstrating various levels of cohesion fatigue in the SLF2-P1 LCL compared to a wildtype control (WT1).

Figure 5.9. SLF2 and SMC5 patient lymphoblastoid cell lines normal distribution of PLK1 (data shown overleaf)

Immunofluorescent images showing the distribution of PLK1 in patient lymphoblastoid cell lines (LCLs). Analysis was performed on LCLs derived from SLF2-P1 and SMC5-P8, alongside an unrelated, healthy individual (WT1), with representative images from SLF2-P1 and WT1 shown. Cells were treated with and without 250nM aphidicolin (APH) for 24 hours. DAPI shown in blue and PLK1 in red. Individual cells are shown in both metaphase and early/late anaphase. These images are representative of two independent biological replicates that showed similar results.


WT1 UNTR
SLF2-P1 UNTR
DAPI
PLK1
Merge
WT1 APH
SLF2-P1 APH
10μM

To study the distribution of Aurora B, I first considered its distribution in mitotic cells. Again, the kinase appeared to be similarly recruited to mitotic chromosomes in WT and SLF2 and SMC5 patient LCLs, in both untreated and APH treated experiments (Figure 5.10a). To examine Aurora B's localisation across chromosome arms in more detail, I cytospan LCLs treated with the mitotic inhibitor colcemid onto slides and immunostained the resulting metaphase spreads (Figure 5.10b). The kinase's co-localisation with CENPA confirmed that Aurora B is localised to centromeric regions in WT LCLs. Again, distribution of the kinase appeared unchanged between SLF2 and SMC5 patient and WT cell lines, with very little signal on the chromosome arms. To question if there was a broader issue with the distribution of condensin, rather than its regulatory factors, I also immunostained metaphase spreads for SMC2 (Figure 5.11). However, as before, immunostaining did not reveal any gross mislocalisation in the patient LCLs, with SMC2 recruited consistently across chromosome arms.

A previous study has shown that the depletion of SMC5/6 is associated with the improper redistribution of Top2α from centromeric regions to chromosome arms. This was causatively linked to mitotic abnormalities including lagging chromosomes and chromatin bridges (Gallego-Paez et al., 2014). The functions of Top2α are known to be important for faithful cohesion, with mutations that affect the SUMOylation of Top2α leading to defects in centromeric cohesion and the spindle checkpoint in yeast (Bachant et al., 2002). Therefore, to question if increases in cohesion fatigue and mitotic issues correlate with Top2α dysfunction in patient LCLs, I immunostained metaphase spreads to observe the enzyme's mitotic distribution (Figure 5.12). Again, the Top2α signal appeared to be unchanged in SLF2 and SMC5 patient cell lines

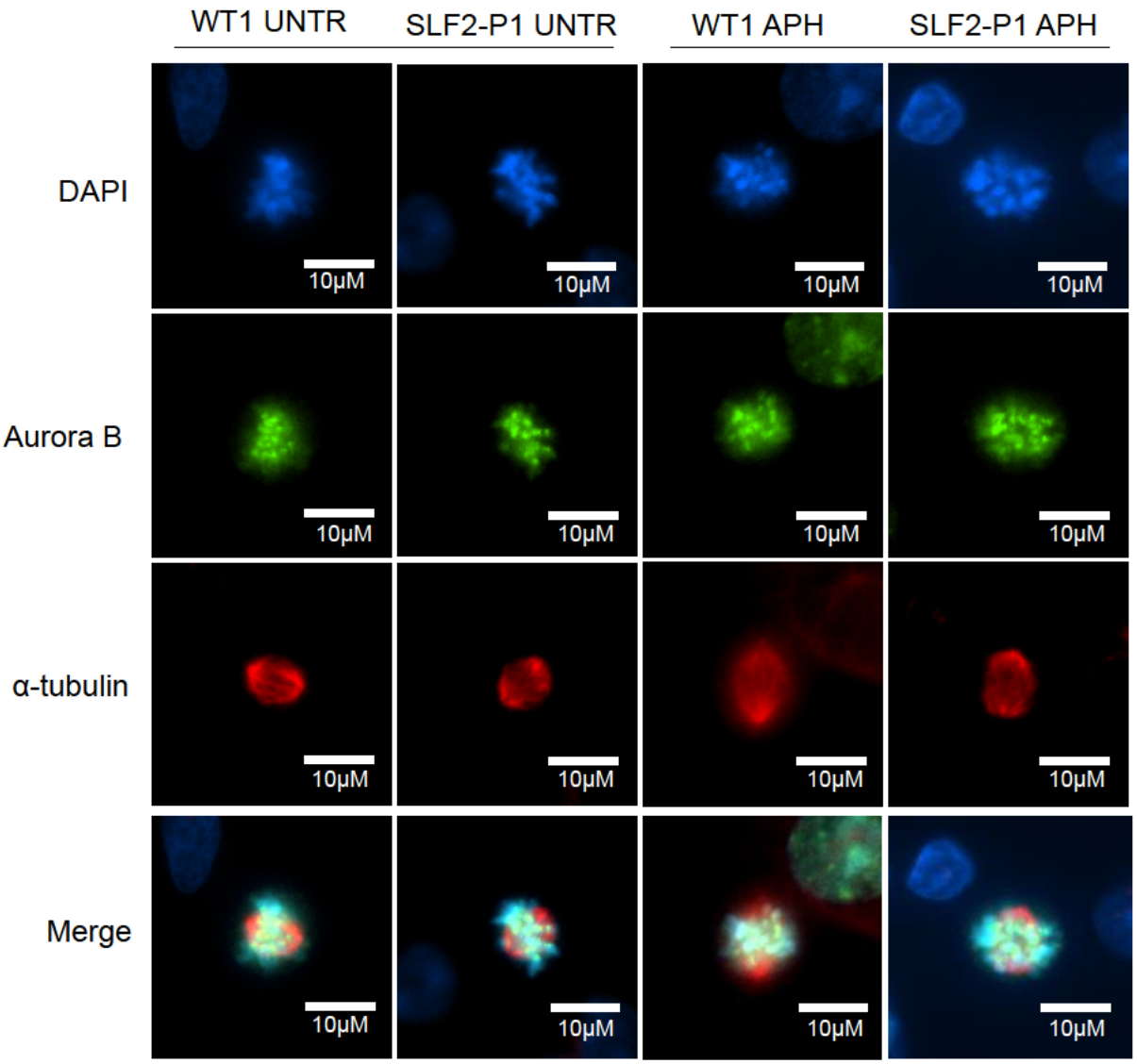
Figure 5.10. SLF2 and SMC5 patient lymphoblastoid cell lines exhibit unperturbed distribution of Aurora B (data shown overleaf)

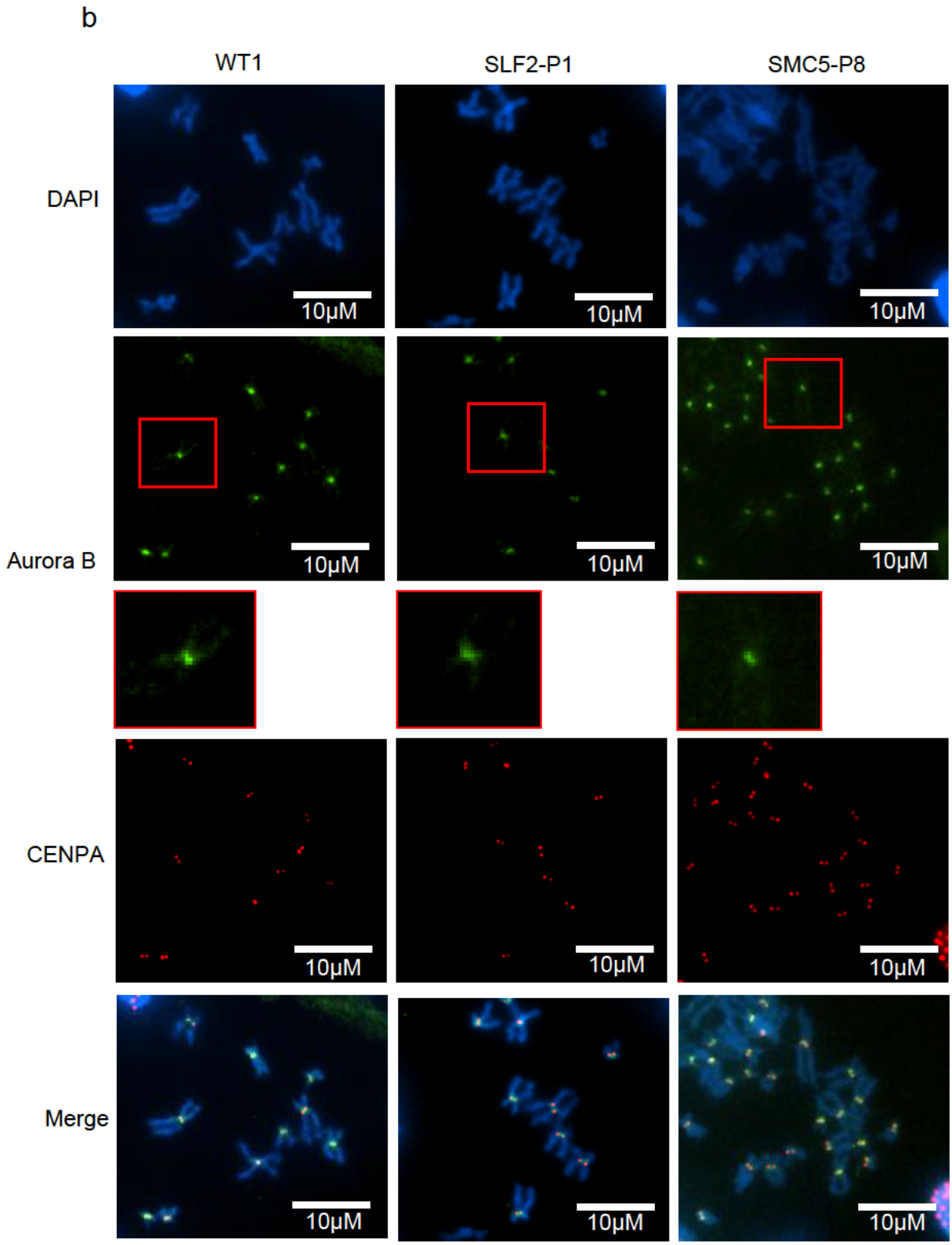
(a) Immunofluorescent images showing the distribution of Aurora B in patient lymphoblastoid cell lines (LCLs). Analysis was performed on SLF2-P1 and SMC5-P8, alongside a wildtype (WT1) LCL, with representative images from SLF2-P1 and WT1 shown. Cells were treated with and without 250nM aphidicolin (APH) for 24 hours. DAPI shown in blue and PLK1 is in red. Cells shown in anaphase.

(b) Immunofluorescent images showing the distribution of Aurora B in metaphase spreads generated from patient LCLs. Analysis was performed on LCLs derived from SLF2-P1, SLF2-P3, SLF2-P4, SMC5-P7 and SMC5-P8, alongside an unrelated, healthy individual (WT1), with representative images from SLF2-P1 and WT1 shown. DAPI shown in blue, Aurora B in green and α-tubulin in red.

In both (a) and (b) images are representative of two independent biological replicates that showed similar results.

a


WT1 UNTR
SLF2-P1 UNTR
WT1 APH
SLF2-P1 APH
DAPI
Aurora B
α-tubulin
Merge
10μM


b
WT1
SLF2-P1
SMC5-P8
DAPI
Aurora B
CENPA
Merge
10μM

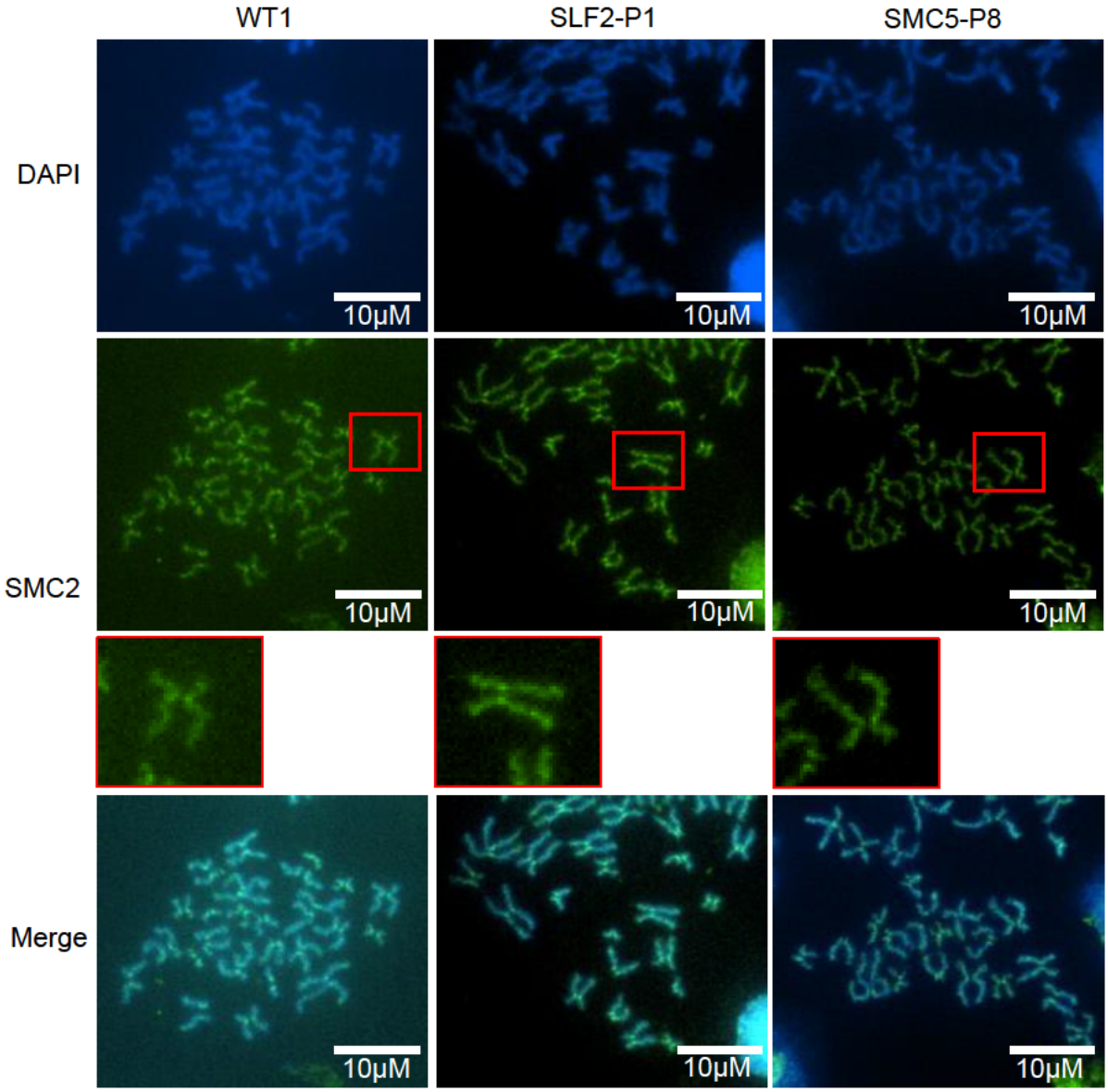


Figure 5.11. SLF2 and SMC5 patient lymphoblastoid cell lines exhibit wildtype distribution of SMC2.

Immunofluorescent images showing the distribution of SMC2 in metaphase spreads generated from patient lymphoblastoid cell lines (LCLs). Analysis was performed on LCLs derived from SLF2-P1 and SMC5-P8, alongside an unrelated, healthy individual (WT1), with representative images from SLF2-P1 and WT1 shown. DAPI shown in blue and SMC2 in green. Images are representative of two independent biological replicates that showed similar results.

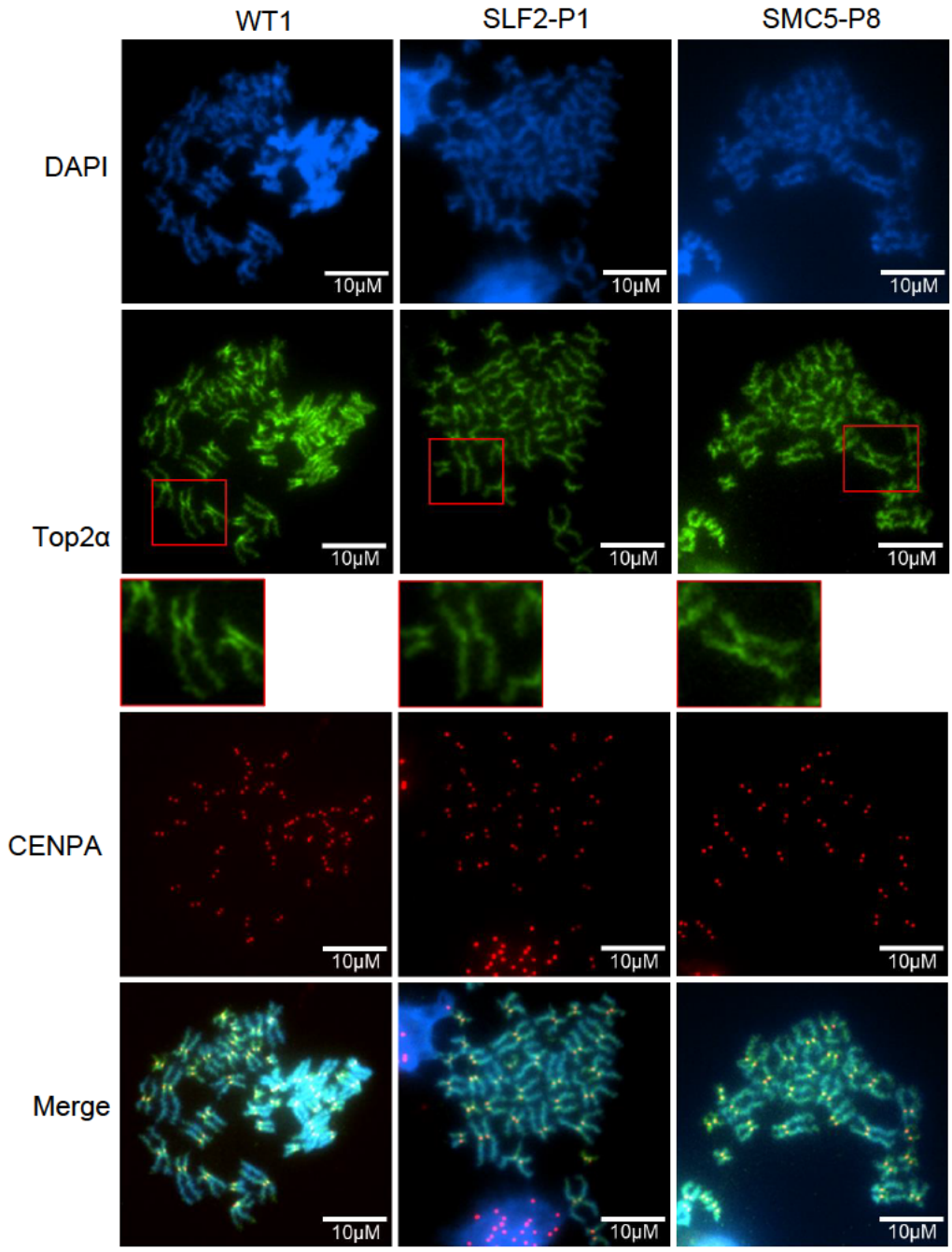


Figure 5.12. SLF2 and SMC5 patient lymphoblastoid cell lines exhibit wildtype distribution of Top2α.

Immunofluorescent images showing the distribution of Top2α in metaphase spreads generated from patient lymphoblastoid cell lines (LCLs). Analysis was performed on LCLs derived from SLF2-P1, SLF2-P3, SLF2-P4, SMC5-P7 and SMC5-P8, alongside an unrelated, healthy individual (WT), with representative images from SLF2-P1 and WT1 shown. DAPI shown in blue, Top2α in green and CENPA in red. Images are representative of two independent biological replicates that showed similar results.

when compared with WT controls, with increased staining at the centromere and lower levels along the chromosome arms. In summary, a brief exploration of the localisation of some of the key mitotic factors previously linked to SMC5/6 failed to implicate a specific protein or pathway in the increased levels of cohesion fatigue observed in the patient LCLs.

5.2.3 Mitotic abnormalities associated with SLF2 and SMC5 mutations are linked to exogenous replicative stress

Whilst performing previous analyses, I observed an induction of multipolar spindles in SLF2 and SMC5 patient LCLs treated with APH. Quantifications confirmed that whilst there was no observable spontaneous defect in spindle formation in mitotic patient LCLs, patient cell lines treated with low levels of APH for 24 hours exhibited significantly elevated levels of multipolar spindles in mitotic cells in comparison to WT controls (Figure 5.13). Indeed, exposure to APH induced multipolar spindles in more than 50% of all mitotic cells in the SLF2-P1 LCL and more than 20% of mitotic cells in SLF2-P3 and SMC5-P8 LCLs.

Multipolar spindles can form due to several defects, including centrosome duplication or fragmentation. Using immunofluorescence microscopy for the centrosome factor PCNT1, I aimed to investigate the presence of either in S/G2 cells. PCNT1 immunostaining revealed a significant increase in centrosome numbers in SLF2 and SMC5 patient LCLs treated with APH for 24 hours (Figure 5.14). In contrast, there was no substantial differences in the levels of spontaneous or APH-induced centrosome fragmentation across all the SLF2 and SMC5 patient LCLs when compared to WT controls (Figure 5.15). Together, this indicated that the APH

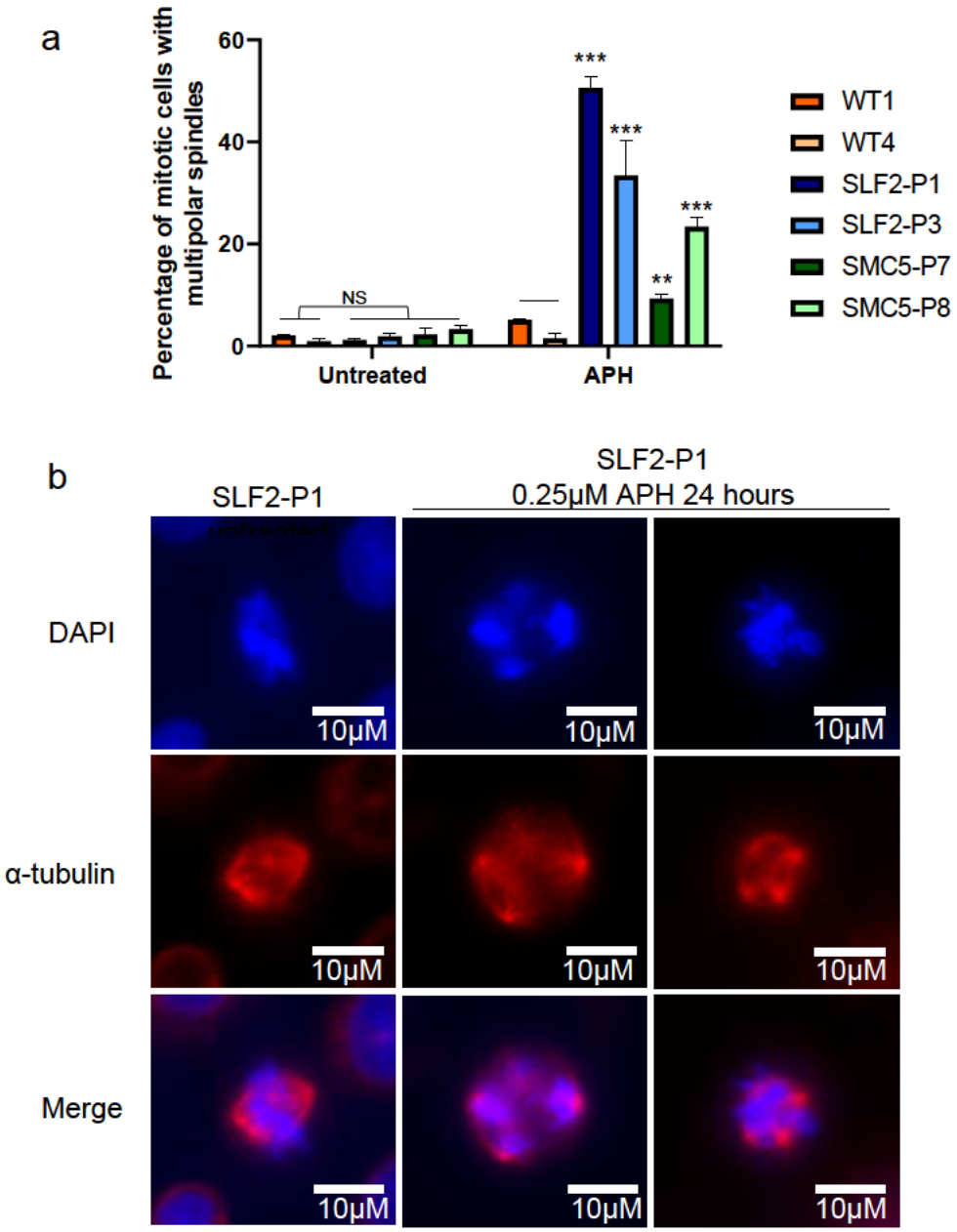


Figure 5.13. SLF2 and SMC5 patient lymphoblastoid cell lines exhibit elevated levels of aphidicolin induced multipolar spindles in mitotic cells

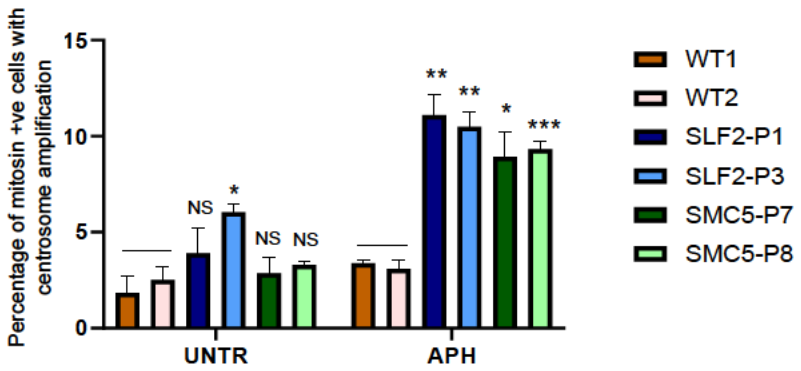
(a) Quantification of the average percentage of mitotic cells with multipolar spindles in lymphoblastoid cell lines (LCLs) derived from patients with mutations in SLF2 and SMC5, as well as unrelated, healthy individuals (WT1 and WT4). Cells were treated with and without 250nM aphidicolin (APH) for 24 hours. n=3 independent experiments. >300 metaphases were counted. Statistical differences were determined with an unpaired, 2-tailed, student's t-test.

(b) Immunofluorescent images of mitotic LCLs from SLF2-P1 demonstrating the induction of multipolar spindles after APH treatment. DAPI (blue) and α-tubulin (red).

Figure 5.14. SLF2 and SMC5 patient lymphoblastoid cell lines exhibit elevated levels of aphidicolin induced centrosome amplifications in S/G2 cells (data shown overleaf)

(a) Quantification of the average percentage of S/G2 (mitosin positive) cells with centrosome amplifications in lymphoblastoid cell lines (LCLs) derived from patients with mutations in SLF2 and SMC5, as well as unrelated, healthy individuals (WT1 and WT2). Cells were treated with and without 250nM aphidicolin (APH) for 24 hours. n=3 independent experiments. >300 metaphases were counted. Statistical differences were determined with an unpaired, 2-tailed, student's t-test.
(b) Representative immunofluorescent images of mitotic lymphoblastoid cells from SLF2-P1 demonstrating the induction of centrosome duplications after APH treatment. DAPI shown in blue, PCNT1 in green and mitosin in red.

a


Percentage of mitosin +ve cells with centrosome amplification
15
10
5
0
NS
*
NS
NS
**
**
*
***
UNTR
APH
WT1
WT2
SLF2-P1
SLF2-P3
SMC5-P7
SMC5-P8

b

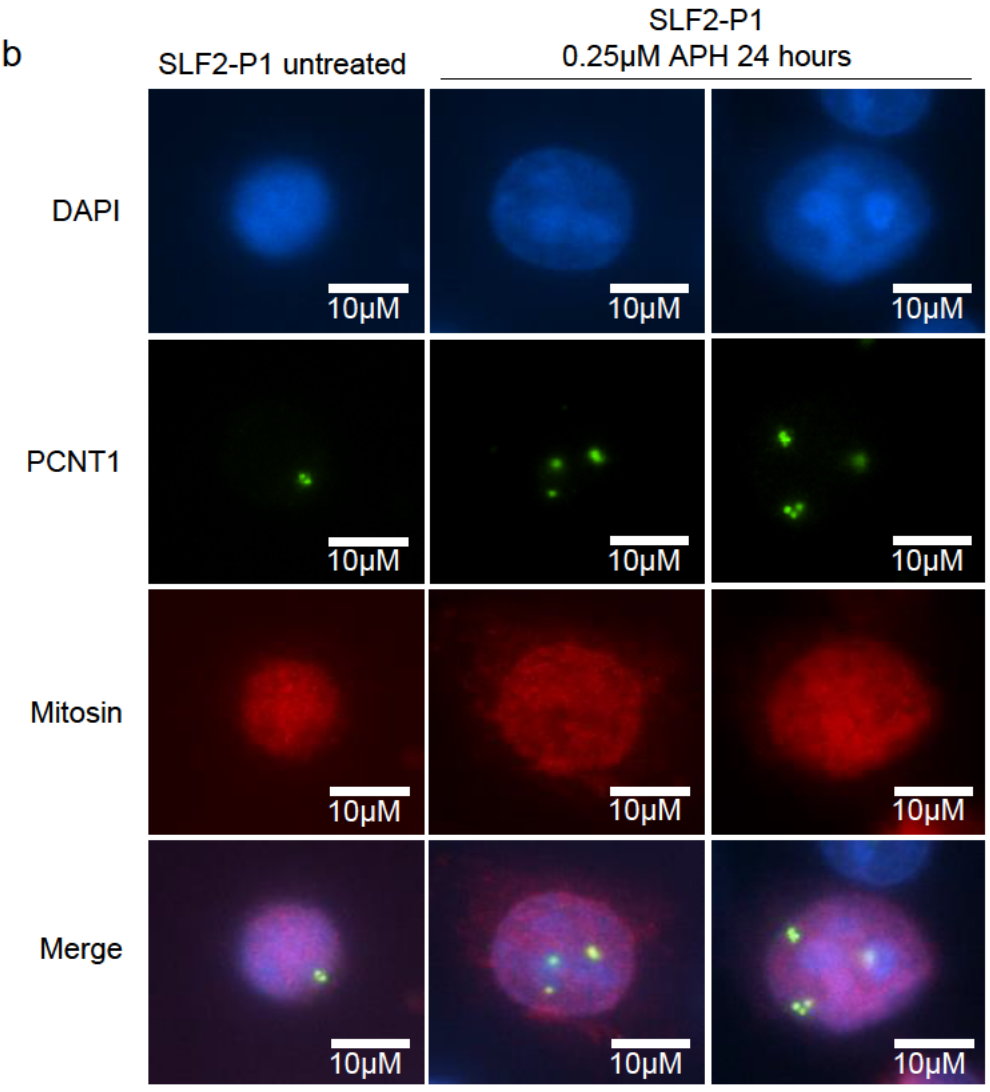
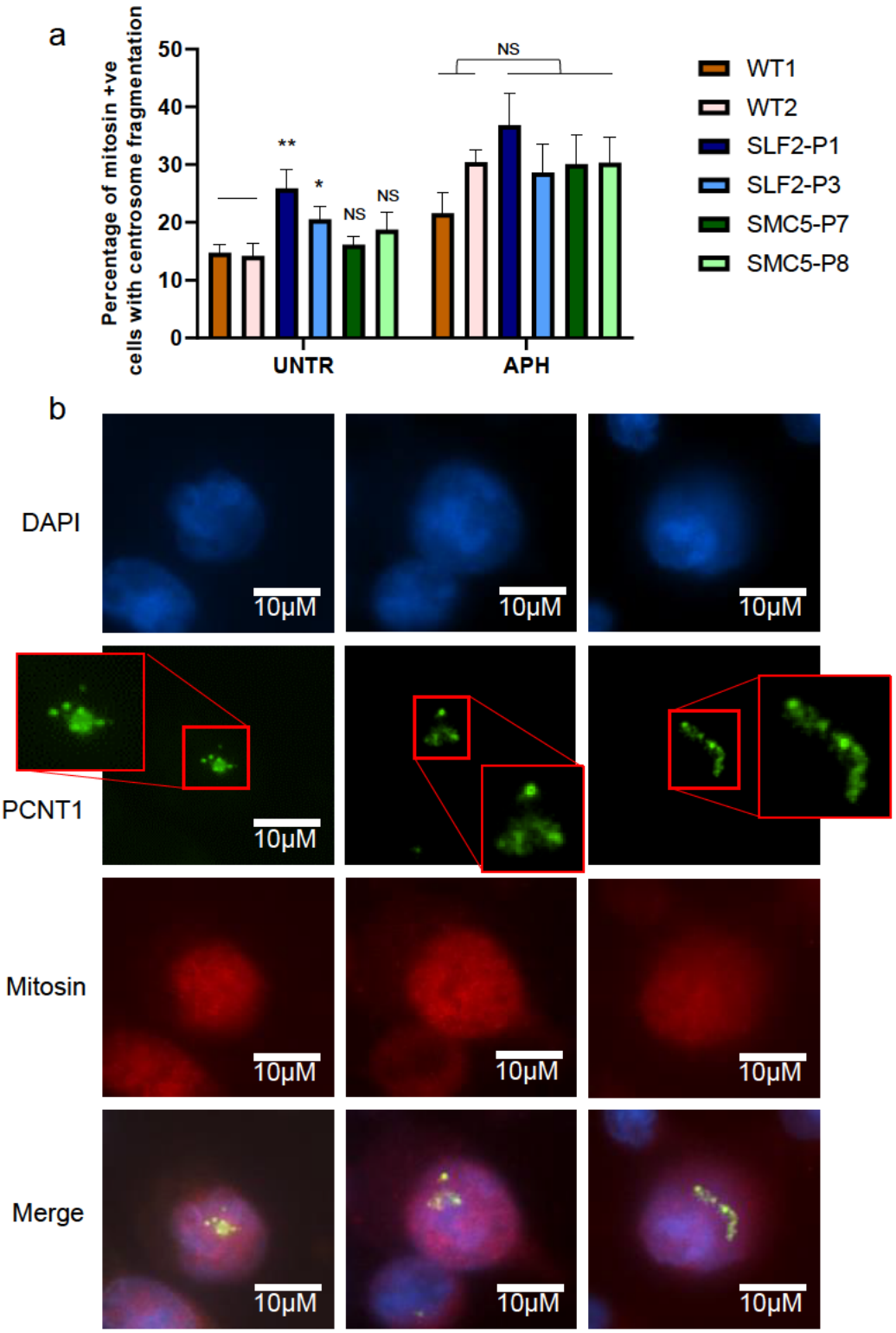

SLF2-P1 untreated
SLF2-P1
0.25μM APH 24 hours
DAPI
PCNT1
Mitosin
Merge
10μM

Figure 5.15. SLF2 and SMC5 patient lymphoblastoid cell lines exhibit wildtype levels of centrosome fragmentation (data shown overleaf)

(a) Quantification of the average percentage of S/G2 (mitosin positive) cells with fragmented centrosomes in lymphoblastoid cell lines (LCLs) derived from patients with mutations in SLF2 and SMC5. Cells were treated with and without 250nM aphidicolin (APH) for 24 hours. n=3 independent experiments. >300 metaphases were counted. Statistical differences were determined with an unpaired, 2-tailed, student's t-test.

(b) Representative immunofluorescent images of the SLF2-P1 LCL exhibiting centrosome fragmentation. DAPI is shown in blue, PCNT1 in green and mitosin in red,


a
Percentage of mitosin +ve cells with centrosome fragmentation
50
40
30
20
10
0
**
*
NS
NS
NS
UNTR
APH
WT1
WT2
SLF2-P1
SLF2-P3
SMC5-P7
SMC5-P8
b
DAPI
PCNT1
Mitosin
Merge
10μM

induced multipolar spindles observed in the SMC5 and SLF2 patient cell lines may arise due to centrosome duplication in S-phase.

Mutations in RNF168 are also associated with mitotic problems

The elevated levels of MVH observed in the SLF2 and SMC5 patient LCLs and peripheral blood lymphocytes is atypical for a replication stress disorder. To understand if the hyperploidy is common to patients with mutations in other genes involved in the SMC5/6-SLF2-SLF1-RAD18 protein pathway, I quantified the number of chromosomes in metaphase spreads derived from a RIDDLE patient LCL (Figure 5.16a). This revealed that the RIDDLE cell line also had an elevated proportion of cells with variable gains in chromosome number. 21% of RIDDLE patient metaphase spreads exhibited >49 chromosomes, in comparison to less than 3% in an unrelated WT control and ATR-seckel syndrome patient (Figure 5.16b).

Having established that the RIDDLE cell line exhibits MVH, I used immunofluorescence to explore if mutations in RNF168 were also associated with increases in multipolar spindles which may drive missegregation. Analysis revealed similarly low levels of spontaneous multipolar mitotic spindles in both untreated WT and RIDDLE patient cell lines (Figure 5.17). However, after 24 hours treatment with APH, levels significantly increased in the RIDDLE LCL whilst WT controls appeared unaffected. Interestingly, the ATR-seckel syndrome patient LCL exhibited low levels of multipolar spindles which were unaffected by APH treatment, remaining comparable to WT levels.

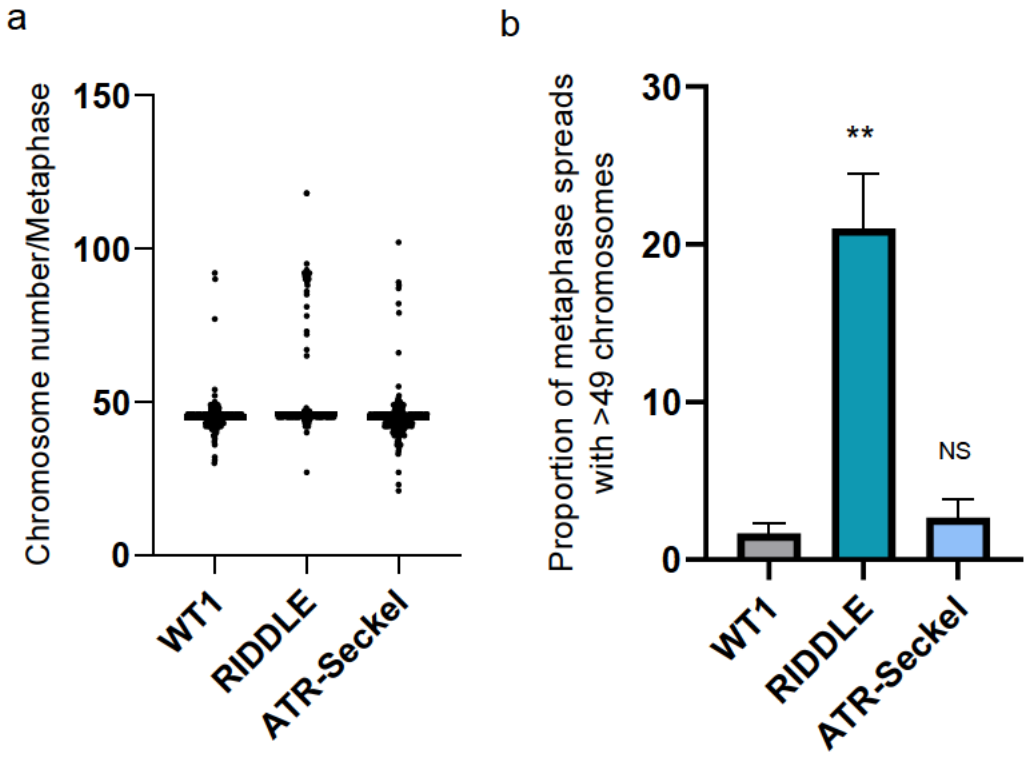


Figure 5.16. RIDDLE patient lymphoblastoid cell lines exhibit elevated levels of cells with mosaic variegated hyperploidy

(a) Quantification of the numbers of chromosomes per metaphase in lymphoblastoid cell lines (LCLs) derived from RIDDLE and ATR-seckel patients, as well as an unrelated, healthy individual (WT1). n=3 independent experiments. 300 metaphases were counted.
(b) The average proportion of metaphases analysed in (a) with >49 chromosomes. Statistical differences were determined with an unpaired, 2-tailed, student's t-test.

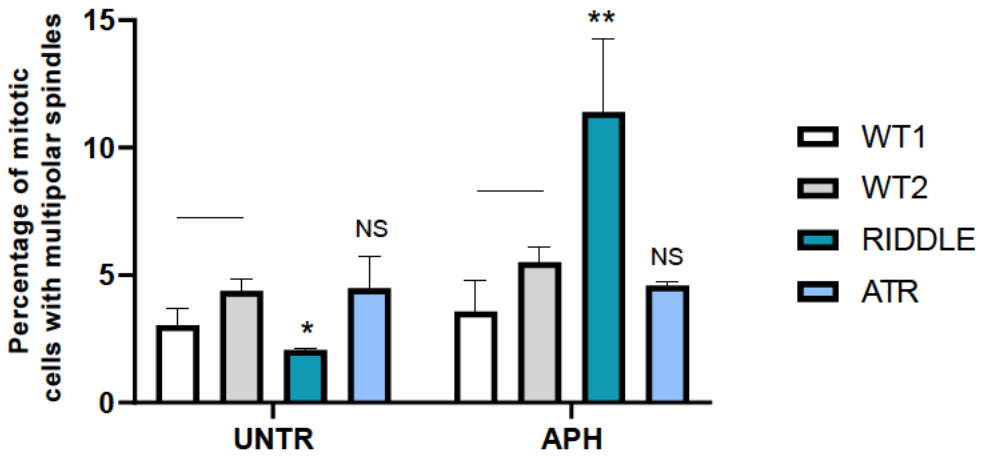


Figure 5.17. RIDDLE patient lymphoblastoid cell lines exhibit elevated levels of aphidicolin induced multipolar spindles

Quantification of the average percentage of mitotic cells with multipolar spindles in lymphoblastoid cell lines (LCLs) derived from RIDDLE and ATR-seckel patients, as well as an unrelated, healthy individual (WT1), treated with and without 250nM aphidicolin (APH) for 24 hours. n=3 independent experiments. >300 metaphases were counted. Statistical differences were determined with an unpaired, 2-tailed, student's t-test.

5.3 Discussion

The data presented in this chapter demonstrates that patient cell lines with mutations in SLF2 and SMC5 exhibit elevated levels of mitotic issues including lagging chromosomes, APH-induced multipolar spindles and variable increases in chromosome numbers. This ‘mosaic variegated hyperploidy’ was also observed in the RIDDLE patient LCL, in which mutations in RNF168 prevent the ubiquitylation of damaged chromatin (Stewart et al., 2007, Stewart et al., 2009), and are, therefore, expected to preclude the subsequent recruitment of RAD18. Taken together, this indicates that the SMC5/6-SLF2-SLF1-RAD18 pathway is important for promoting proper cellular division and maintaining diploidy.

Whilst chromosome gains were the focus of my studies, it seems plausible that the cellular defects that contribute to these increases also lead to chromosome losses. However, during the process of preparing metaphase spreads it is not possible to know whether individual chromosomes have drifted or become inadvertently washed off the microscope slide. Since it is difficult to be confident of whether any observed chromosomes losses are real, I only quantified chromosome gains. It is also probable that there is a limit to the number of chromosomes that can be lost before cell survival is impacted, thereby limiting the prevalence of cells with drastically smaller numbers of chromosomes.

A key question presented by this data was whether the mitotic abnormalities present in the SMC5 and SLF2 patient cell lines are dependent on SMC5/6’s functions in replication or mitosis. Increases in replication stress are certainly associated with chromosome missegregation, although the mechanisms by which it causes these

issues are unclear, with different types of replication stress leading to varying mitotic consequences (Wilhelm et al., 2019b, Wilhelm et al., 2020). However, the hyperploidy observed here is particularly striking and typically not associated with dysfunctional replication, as evidenced by its absence in the LCL derived from an ATR-seckel syndrome patient with mutations in ATR. This may suggest that the mitotic defects seen in SLF2 and SMC5 patient cells have not solely been caused by a defect in replication or an increase in replication stress.

The high frequency of hyperploid SLF2 and SMC5 patient cells exhibiting around 92 chromosomes suggests that a cytokinetic defect may contribute to chromosome gains, although preliminary investigations failed to observe an increase in the number of binucleated cells. As previously mentioned, this may suggest that tetraploid cells arise independently of a cytokinetic defect, but survive better than cells exhibiting other chromosome numbers due to the maintenance of genomic balance (Birchler and Veitia, 2012). However, a lot is unknown regarding the dynamics of the hyperploidy observed in the SLF2, SMC5 and RIDDLE patient cell lines. From chromosome quantifications alone it's very difficult to understand how frequently chromosome gains occur, for how long hyperploid cells survive and if their population fluctuates with time. As such, it's possible that any putative cytokinetic defect occurs infrequently and is difficult to observe in an experiment. In addition, my experiment failed to consider how long binuclear cells resulting from cytokinetic defects may exist as two separate nuclei within a single cell. If the nuclei merge following subsequent rounds of mitosis, then tetraploid cells could be overlooked by my analysis. Thus, even very rare and difficult to observe cytokinetic defects could be responsible for the significant levels of hyperploidy and tetraploid cells.

Further putative mitotic defects are suggested by the numerous superficial similarities between individuals with MVA and those with mutations in SLF2 and SMC5, with cell lines derived from both groups of patients exhibiting variable changes in chromosome numbers in just a proportion of cells. Since MVA is associated with mitotic dysfunction, this suggested that problems in mitosis might contribute to MVH, also. Consistent with this, recent research has indicated that SMC5/6 is associated with several mitotic factors. This includes Aurora B, which contributes to centromeric cohesion protection by promoting the localisation of the kinetochore kinase Bub1, a protein recently linked with MVA (Carvalhal et al., 2022). However, whilst the depletion of SMC5/6 has been observed to lead to the mislocalisation of Aurora B, alongside other mitotic factors (Pryzhkova and Jordan, 2016), I was unable to observe a change/defect in their distribution on chromosomes in the SLF2 and SMC5 patient cell lines. The severity of the gains in chromosome number observed in the patient cell lines meant that I was anticipating gross changes in the localisation of these factors, although the hypomorphic nature of the patient mutations and/or cell line differences may mean that any impact may be more subtle. As such, a microscope that allows for higher resolution imaging might reveal defects in the distribution of mitotic factors that I have been unable to identify. It's also possible that protein function rather than localisation has been affected and that the phosphorylation or recruitment of downstream factors should be investigated.

Whilst the distribution of specific mitotic factors associated with SMC5/6 appeared unchanged, patient cell lines presented with sister chromatid cohesion abnormalities, including rail-road chromosomes in the peripheral blood lymphocytes and cohesion fatigue in the LCLs. This indicated that defects in SLF2 and SMC5 contribute to

premature sister chromatid separation. In support of this, SLF2 and SMC5 patients present with similar clinical phenotypes to the cohesinopathy Warsaw Breakage Syndrome (WABS), including microcephaly, growth restriction and heart defects. as well as elevated levels of railroad chromosomes in patient cell lines (van Schie et al., 2020). WABS arises due to mutations in DDX11, a gene encoding a DNA helicase (also called ChlR1) (van der Lelij et al., 2010a, Wu et al., 2012). Interestingly, DDX11 has been observed to be involved in the rescue of stalled replication forks, as well as the association of cohesin to ongoing replisomes (Cali et al., 2016, Abe et al., 2018, Cortone et al., 2018). However, its DNA unwinding enzymatic functions are not essential for cohesin recruitment, indicating that these roles are distinct from one another (Samora et al., 2016). Recently, SMC5 and DDX11 were shown to act jointly to promote cell growth in the presence of cisplatin, suggesting that these factors function together within a distinct DNA repair pathway (Rossi et al., 2020). This may suggest that the clinical similarities between WABS and patients with SMC5 and SLF2 mutations arise due to common DNA repair defects. However, considering the cohesion defects observed in the SMC5 and SLF2 patient derived cell lines, it's possible that SMC5/6, like DDX11, is a multifunctional protein with roles in both repair and cohesin recruitment that contribute to this disease phenotype when disrupted. Whether these factors work together to promote sister chromatid cohesion remains to be seen.

Whilst mutations in DDX11 have not yet been associated with aneuploidies in human cell lines, other cohesinopathies like Roberts syndrome, caused by mutations in ESCO2 (Vega et al., 2005a, Gerkes et al., 2010), report varying karyotypic abnormalities, with some patient derived cells presenting with aneuploidies and some

without. However, these changes rarely reflect large chromosome gains (Jabs et al., 1991, Van Den Berg and Francke, 1993). Further to this, whilst cohesion defects are common to all SLF2 and SMC5 patients presented in this study, only MVA patients with mutations in BUB1 and BUB1B have been associated with spontaneous premature sister chromatid separation (Matsuura et al., 2006, Carvalhal et al., 2022). Though this work directly associates cohesion defects in SMC5/6-associated factors with mitotic abnormalities, it is currently not possible to infer the extent to which chromatid cohesion defects contribute to the hyperploidy observed in the SLF2 and SMC5 patient cell lines. However, as the cohesion defects observed here are relatively mild, with spontaneous railroad chromosomes only occurring in around 10% of peripheral blood lymphocytes metaphases and not in the LCLs, it seems unlikely that it is responsible for the majority of the karyotypic abnormalities, although it may still contribute to a lesser extent. Instead, it would be interesting to ask if, as with the multipolar spindles, additional replication stress exacerbates this phenotype.

After 24 hours of APH treatment, elevated levels of multipolar spindles were observed in the SLF2 and SMC5 patient LCLs. My data suggested that these were likely caused by the over-duplication of centrosomes. For each SLF2 and SMC5 patient LCL, the proportion of mitotic cells with replication stress induced multipolar spindles correlated well with the observed levels of MVH; being highest in the SLF2 patients (in SLF2-P1 in particular) and lower in the SMC5 patients. Taken together, this may suggest that replication stress induced multipolar spindles contribute significantly to MVH and that the SMC5/6-SLF2-SLF1-RAD18 pathway is important for promoting proper centrosome duplication in certain conditions. If this is a primary driver of chromosome gains, then this would suggest that missegregation may not occur spontaneously but

is instead induced in conditions of elevated replication stress. To test this hypothesis, it would be interesting to ask if other sources of replication stress induce multipolar spindles or if prolonged APH treatment led to elevated levels of hyperploid cells.

In normal conditions, the progression of centrosome duplication is tightly coupled with that of DNA replication, thereby maintaining consistent centrosome numbers. However, strong DNA replication stress can lead to an extended S-phase delay that allows for centrosome overduplication and could therefore lead to multi-polar spindles, lagging chromosomes and missegregation (Meraldi et al., 1999). This could indicate that elevated levels of replication stress, induced by defects in SLF2 or SMC5 and exacerbated by mild APH treatment, leads to an extended S-phase and subsequent centrosome over-duplication. However, whilst the RIDDLE patient LCL exhibited multipolar spindles after 24 hours incubation with the same dose of APH, the ATR-seckel syndrome LCL did not, indicating that this is not a phenotype that is common to all cell lines with replicative dysfunction. This could simply suggest that mutations in the SMC5/6-SLF2-SLF1-RAD18 pathway, accompanied by APH treatment, induce higher levels of replication stress which exceeds the threshold for S-phase delay. However, it could also demonstrate a specific centrosome function for SMC5/6 that it is yet to be discovered. Interestingly, HR proteins are associated with centrosomes, with experimental evidence showing that supernumerary centrosomes are frequently induced in HR deficient cell lines upon treatment with replication stress inducing drugs (Wilhelm et al., 2014). Therefore, it may be that SMC5/6's functions in HR specifically are also important for promoting proper centrosome duplication, whether that be directly or indirectly. To investigate this further, FACs analysis could

be undertaken first to determine if APH treated SLF2 and SMC5 patient LCLs exhibit a significantly extended S-phase.

The aim of this chapter was to determine if SLF2 and SMC5 patient cell lines exhibited mitotic abnormalities that could potentially explain the large alterations in chromosome numbers. Whilst several mitotic defects were identified present it is difficult to understand to what extent they individually contribute to MVH. It will now be important to establish if the mitotic abnormalities arise due to replicative dysfunction or because of specific mitotic roles for SMC5/6. Once this has been determined, it will be possible to specifically target these defects and monitor how these impacts upon the observed MVH phenotype.

Chapter 6:

Investigating the role of SMC5 and SLF2 in promoting replication through regions prone to the formation of DNA secondary structure

6.1 Introduction

The data presented in the previous chapters have demonstrated that patient-derived cell lines with SLF2 and SMC5 mutations present with elevated levels of DNA damage, genome instability and increased levels of spontaneous replication stress. This raises the question of what the source of endogenous replication stress in the SLF2 and SMC5 patient cells is, and whether a defect in the known role of SMC5/6 in regulating HR during replication fork progression contributes to the replication abnormalities.

During S-phase, replication forks that stall and collapse can be rescued by converging forks from nearby dormant origins. However, replication restart can also be initiated by HR, which promotes the invasion of the homologous sister chromatid (Petermann and Helleday, 2010). As previously discussed, SMC5/6 is implicated in HR and promoting faithful replication by preventing fork regression and resolving replication-linked joint molecules (Chen et al., 2009, Bermudez-Lopez et al., 2010, Aragon, 2018, Palecek, 2018, Agashe et al., 2021). Smc5/6 has also been reported to be enriched at repetitive regions, including ribosomal DNA (rDNA) and telomeres (Torres-Rosell et al., 2005, Menolfi et al., 2015), and depletion of the complex has been shown to result in the accumulation of X-shaped DNA structures at specific repetitive regions, as well as missegregation errors during mitotic cell division (Torres-Rosell et al., 2005, Chavez et al., 2010, Chavez et al., 2011, Peng et al., 2018). This suggests that SMC5/6 negatively regulates HR at repetitive regions, where otherwise aberrant recombination could lead to deleterious rearrangements and gene slippage. However, it is difficult to determine whether this is indicative of a specific function for

SMC5/6 at these sites of repetitive DNA or whether this reflects a broader role for SMC5/6 in HR.

There is evidence that SMC5/6 has an active role in the nuclear organisation of recombination at repetitive elements. For instance, experimental evidence from yeast has shown that the SMC5/6 complex facilitates the movement of rDNA HR foci away from other rDNA repeats via exclusion from the nucleolus (Torres-Rosell et al., 2007). Similarly, in human cells which utilise the alternative lengthening of telomeres (ALT) pathway, SMC5/6 is believed to organise the nuclear movement of telomeres to ALT-associated promyelocytic leukemia bodies (APBs), where recombination-mediated telomere elongation is promoted (Potts et al., 2006). These functions appear to indicate that SMC5/6 can promote specific mechanisms which limit deleterious crossover events at specific repetitive regions. Alternatively, however, an increase in recombination intermediates at repetitive elements in the absence of a functional SMC5/6 complex may reflect a genome-wide role for SMC5/6 in regulating HR, exacerbated by the inherent difficulty in replicating through repetitive regions, some of which can readily form complex DNA secondary structures (Kaushal and Freudenreich, 2019). Notably, rDNA and telomeres represent regions of unidirectional replication, in which converging forks are unable to rescue stalled forks. Thus, faithful recombination-mediated replication restart is important to avoid under-replicated DNA at these loci and may explain SMC5/6's association with these sites (Murray and Carr, 2008).

Mutations in the DDX11 helicase cause Warsaw breakage syndrome (WABS), a developmental disorder in which cells present with genome instability phenotypes comparable to those seen in Fanconi Anemia (FA) (van der Lelij et al., 2010a, Wu et

al., 2012). DDX11 has been implicated in replication fork restart and HR-mediated repair, with separable functions in promoting cohesin recruitment (Cali et al., 2016, Abe et al., 2018, Cortone et al., 2018). Studies investigating the helicase's molecular mechanisms have found that, in conjunction with the 9-9-1 clamp loader, the protein mediates the removal of interstrand cross-links in a backup FA-like pathway (Abe et al., 2018). Interestingly, cells lacking DDX11 have shown a sensitivity to CX-5461, a G4-quadruplex stabilising drug (van Schie et al., 2020). This may indicate that the helicase is also required to remove these structures, which can act as impediments to the replication fork (Edwards et al., 2014, Lemmens et al., 2015, Kolesnikova and Curtis, 2019, Lerner and Sale, 2019).. Recently, SMC5 has been shown to function with the DDX11 helicase to promote cellular proliferation and survival in response to the DNA damaging agent, cisplatin (Rossi et al., 2020). This, in combination with the similar defects in the establishment of sister chromatid cohesion in patients with mutations in DDX11, SLF2 and SMC5 (discussed in chapter 5), may implicate the SMC5/6 complex in a number of DDX11-mediated cellular pathways. This may also explain the numerous clinical similarities between WABS patients and those with mutations in SLF2 and SMC5, including microcephaly, growth restriction and heart defects. Whilst it is not yet known whether a loss of SMC5/6 complex components render cells sensitive to G4-quadruplex stabilising agents, it has been shown that RNF168, which is required for the recruitment of the SMC5/6-SLF2-SLF1-RAD18 complex to damaged chromatin, signals the presence of G4-quadruplex structures stabilised by CX-5461 (Masud et al., 2021).

6.1.2 Chapter Aims

The aim of this chapter is to further characterise the defects present in SLF2 and SMC5 patient derived cell lines, with a particular focus on investigating whether the functions of SMC5/6 in HR are affected by the pathogenic mutations. Furthermore, I also aim to understand what the source/s of the spontaneous replication stress in SLF2 and SMC5 patient cells are.

6.2 Results

6.2.1 Mutations in SLF2 or SMC5 are associated with increased levels of chromosomal aberrations, rare chromosomal morphologies and telomeric dysfunction

To further understand the extent of the genome instability caused by mutations in SLF2 and SMC5, levels of chromosome instability were measured by quantifying the number of chromosomal aberrations (chromosome gaps and breaks, chromatid gaps and breaks, and chromosome radials) present in the SLF2 and SMC5 patient LCLs (Figure 6.1). In comparison to WT controls, SLF2 and SMC5 patient cell lines exhibited significantly increased levels of spontaneously induced chromosomal aberrations, comparable to levels observed in LCLs derived from an ATR-seckel syndrome patient. Interestingly, in contrast to the ATR-seckel syndrome LCL, levels of chromosomal aberrations did not appear to be significantly exacerbated in the patient LCLs in response to APH and MMC treatment. Increased levels of chromosomal aberrations were also seen in metaphase spreads prepared from patient peripheral blood lymphocytes (Figure 6.2). Importantly, levels of chromosome breakage were significantly reduced in patient fibroblasts and U-2-OS SLF2 CRISPR mutant cell lines complemented with WT SLF2 and SMC5 protein compared to empty.

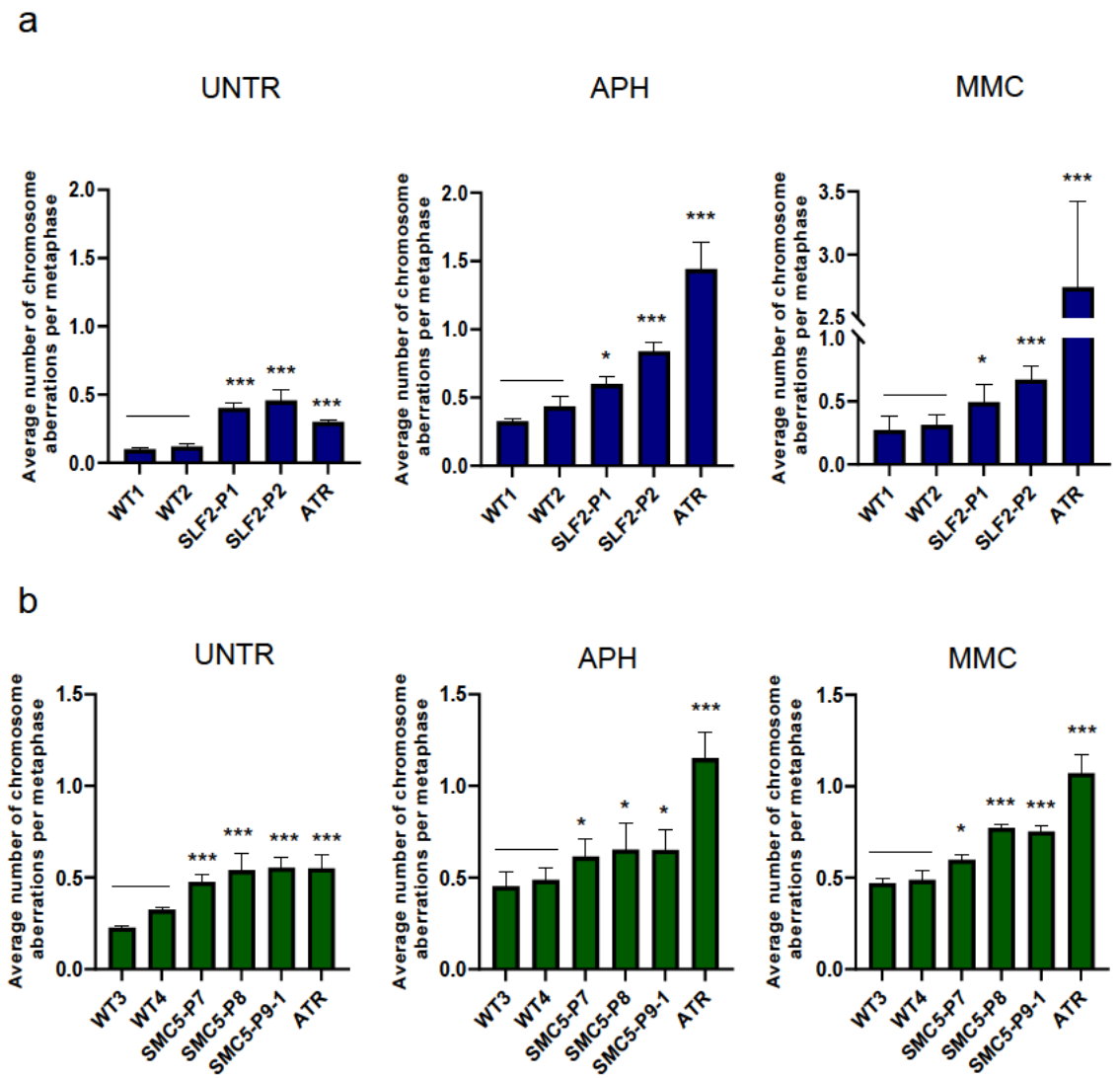


Figure 6.1. SLF2 and SMC5 patient lymphoblastoid cell lines exhibit elevated levels of chromosome aberrations

Quantification of the average number of chromosome aberrations (chromosome gaps and breaks, chromatid gaps and breaks, and chromosome radials) in metaphase spreads from SLF2 (a) and SMC5 (b) patient lymphoblastoids cell lines, compared to those derived from unrelated, healthy individuals (WT1, WT2, WT3 and WT4). Cells were either untreated or exposed to 500 nM aphidicolin (APH) or 50ng/ml mitomycin-C (MMC) for 24 h. n=3 independent experiments. In total >140 metaphases were counted. Statistical differences were determined with an unpaired, 2-tailed, student's t-test. Experiments carried out by Dr John Reynolds.

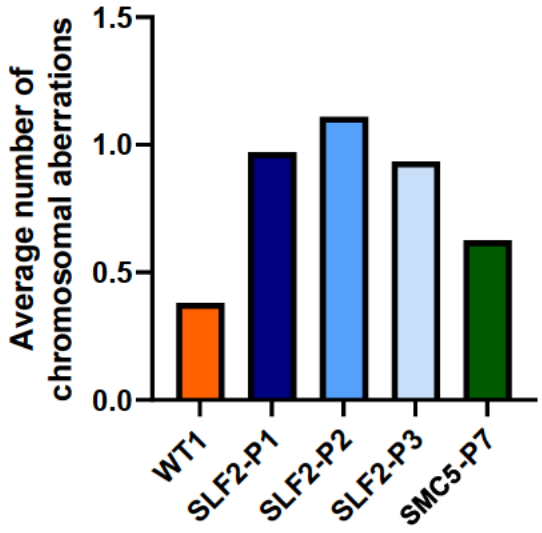


Figure 6.2. SLF2 and SMC5 patient peripheral blood lymphocytes exhibit elevated levels of chromosome aberrations

Quantification of the average number of chromosome aberrations (chromosome gaps and breaks, chromatid gaps and breaks, and chromosome radials) in metaphase spreads from SLF2 and SMC5 patient peripheral blood metaphases, as well as those produced from a healthy, wildtype individual (WT1). 200 metaphases were counted in total from 2 independent blood samples.

vector controls (Figure 6.3, Figure 6.4). Taken together, this indicates that mutations in SLF2 and SMC5 cause increased levels of spontaneous chromosomal aberrations

Whilst investigating the levels of chromosomal aberrations present in blood peripheral lymphocytes metaphases, it became evident that that a subset of chromosomes in the patient samples exhibited unusual chromosome morphologies. I categorised these into two classes; ‘Type 1’ chromosomes which presented with one or more distinct gaps and/or breaks across both chromatids, and ‘Type 2’ chromosomes, which seem to have two constricted centromeric regions, thereby resembling dicentric chromosomes (Figure 6.5a). Quantifications revealed that there was an increased proportion of both ‘Type 1’ and ‘Type 2’ chromosomes in the SLF2 and SMC5 patient blood peripheral lymphocytes metaphases (Figure 6.5b). The majority of these consisted of ‘Type 1’ chromosomes with a single break.

To confirm if the ‘Type 2’ chromosomes present in the patient peripheral blood lymphocytes were dicentric, I used fluorescence in situ hybridization (FISH) to visualise centromeres (Figure 6.6). In this approach, fluorophore-coupled nucleotide probes are hybridized to specific, complementary DNA sequences to aid the detection of certain loci. SLF2 and SMC5 metaphase preparations were dropped onto slides, denatured with formamide, and hybridized with a CENPB-Alexa-488 peptide nucleic acid (PNA) probe. Immunofluorescence imaging of SLF2-P2 and SLF2-P3 metaphase spreads revealed numerous ‘Type 2’ chromosomes with dual centromeric staining.

Dicentric chromosomes can arise when two chromosomes fuse, each with their own centromere. This can occur when dysfunctional telomeres enable aberrant repair

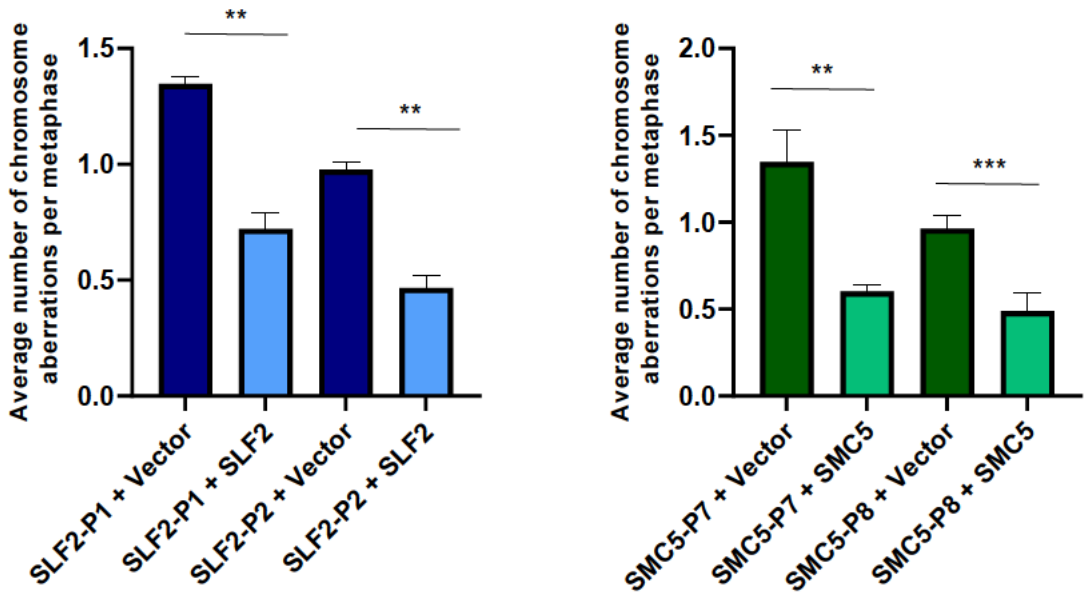


Figure 6.3. SLF2 and SMC5 patient derived fibroblasts exhibit elevated levels of chromosome aberrations

Quantification of the average number of chromosome aberrations (chromosome gaps and breaks, chromatid gaps and breaks, and chromosome radials) in SLF2 (left) and SMC5 (right) patient fibroblasts complemented with a lentivirus encoding WT SLF2, WT SMC5, or an empty vector. n=3 independent experiments. >90 metaphases were counted. Statistical differences were determined with an unpaired, 2-tailed, student's t-test.

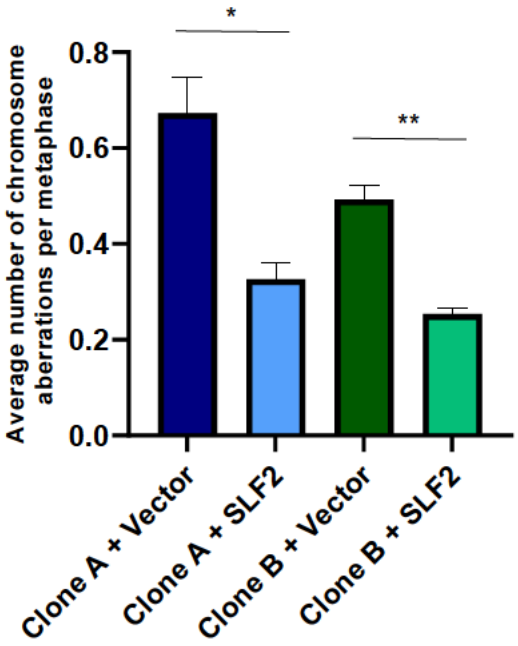


Figure 6.4. SLF2 CRISPR U-2-OS cell lines exhibit elevated levels of chromosome aberrations

Quantification of the average number of chromosome aberrations in U-2-OS SLF2 CRISPR clone A and clone B complemented with a lentivirus encoding WT SLF2, WT SMC5, or an empty vector. n=3 independent experiments. 150 metaphases were counted. Statistical differences were determined with an unpaired, 2-tailed, student's t-test.

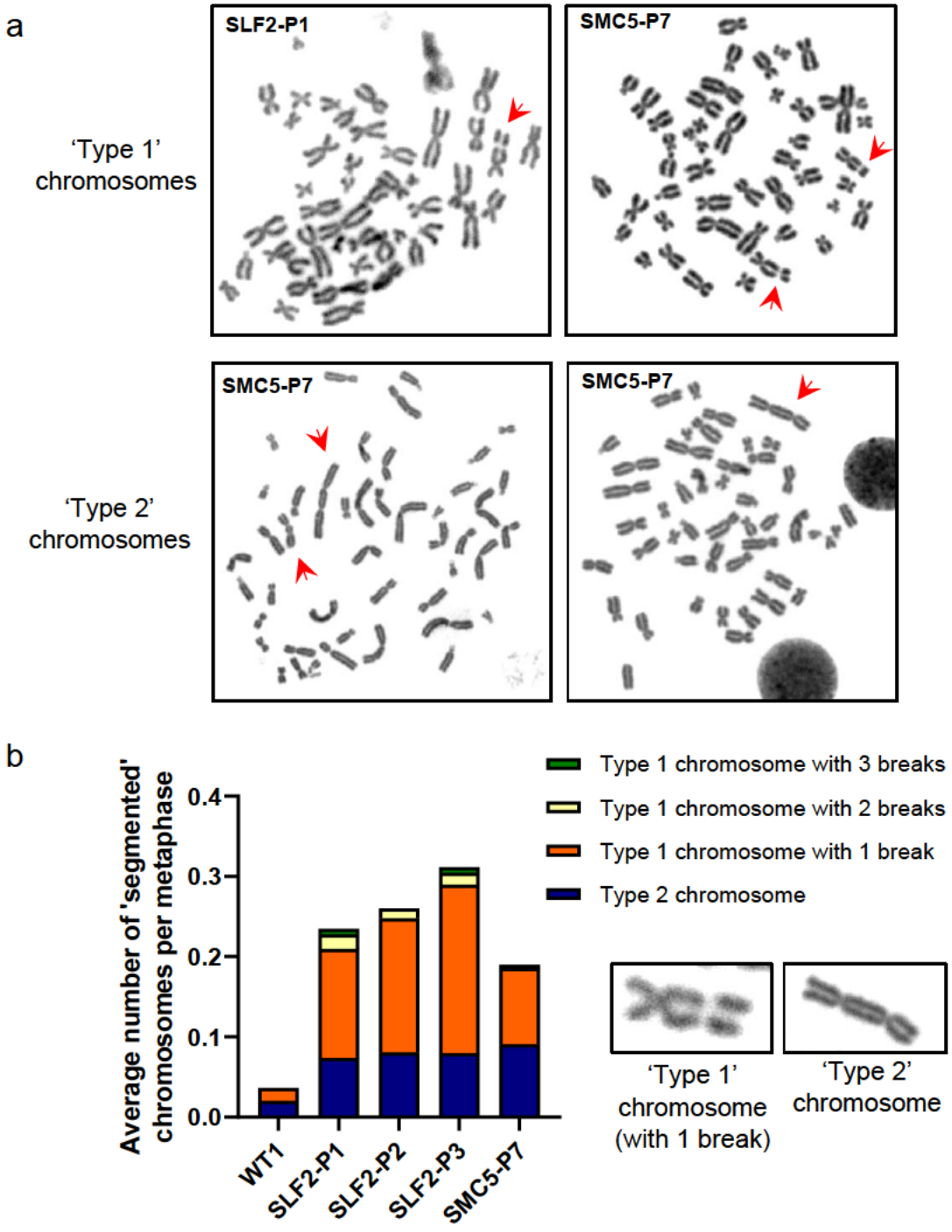


Figure 6.5. SLF2 and SMC5 patient peripheral blood samples exhibit unusual chromosome abnormalities

(a) Representative images of 'Type 1' and 'Type 2' chromosomes, indicated by red arrows.

(b) Quantification of the average number of 'Type 1' and 'Type 2' chromosomes in patient peripheral blood lymphocytes generated from samples attained from SLF2 and SMC5 patients, as well as an unrelated, healthy individual (WT1). 250 metaphases were counted in total from 2 independent blood samples.

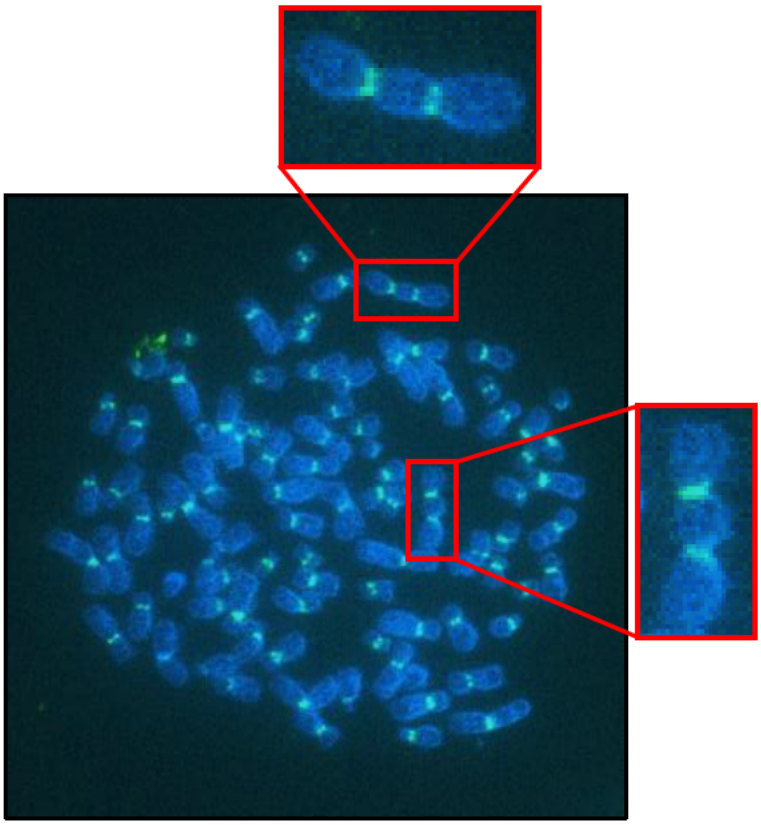


Figure 6.6. SLF2 and SMC5 patient peripheral blood samples exhibit dicentric chromosomes

Representative image of fluorescence in situ hybridization (FISH) performed with a centromere specific probe on SLF2-P3 peripheral blood lymphocytes. Dicentric chromosomes are highlighted in red.

between deprotected chromosome ends (van Steensel et al., 1998, Pardo and Marcand, 2005, Muraki et al., 2012). Considering that Smc5/6 is enriched at telomeres and is associated with an accumulation of unresolved DNA structures in these regions (Torres-Rosell et al., 2005, Chavez et al., 2010), I hypothesised that the 'Type 2' chromosomes may occur due to a role for the complex in telomeric maintenance. To explore this further in the patient LCLs, I performed telomeric FISH using a TelC-Cy5 PNA probe to visualise and quantify telomere abnormalities, such as signal-free ends and/or fragile telomeres, which typically present with elongated signals or multi-telomeric signals (Figure 6.7a). Previous work has suggested that this fragility may arise due to improper chromatin condensation around under-replicated DNA or stalled replication forks at telomeres (Sfeir et al., 2009). Initial quantifications revealed that the proportion of fragile and signal-free telomeres ranged considerably between LCLs, particularly between WT controls (Figure 6.7b). Thus, there appeared to be a large degree of variation which was inherent to non-isogenic cell lines. To avoid this, I conducted the same telomere FISH experiments using one of the complemented U-2-OS SLF2 CRISPR cell lines (Figure 6.8). Immunofluorescence analysis using SLF2 CRISPR mutant clone A did not reveal a difference in the levels of signal-free chromosome ends between cells complemented with empty vector and those complemented with WT SLF2. In contrast, cells complemented with WT protein exhibited a mild, but significant, reduction in the levels of fragile, dysfunctional telomeres in comparison to the vector only line, indicating that SLF2 may be important for maintaining telomeric integrity in U-2-OS cells.

To question if the putative telomeric problems in SLF2 deficient cell lines were linked

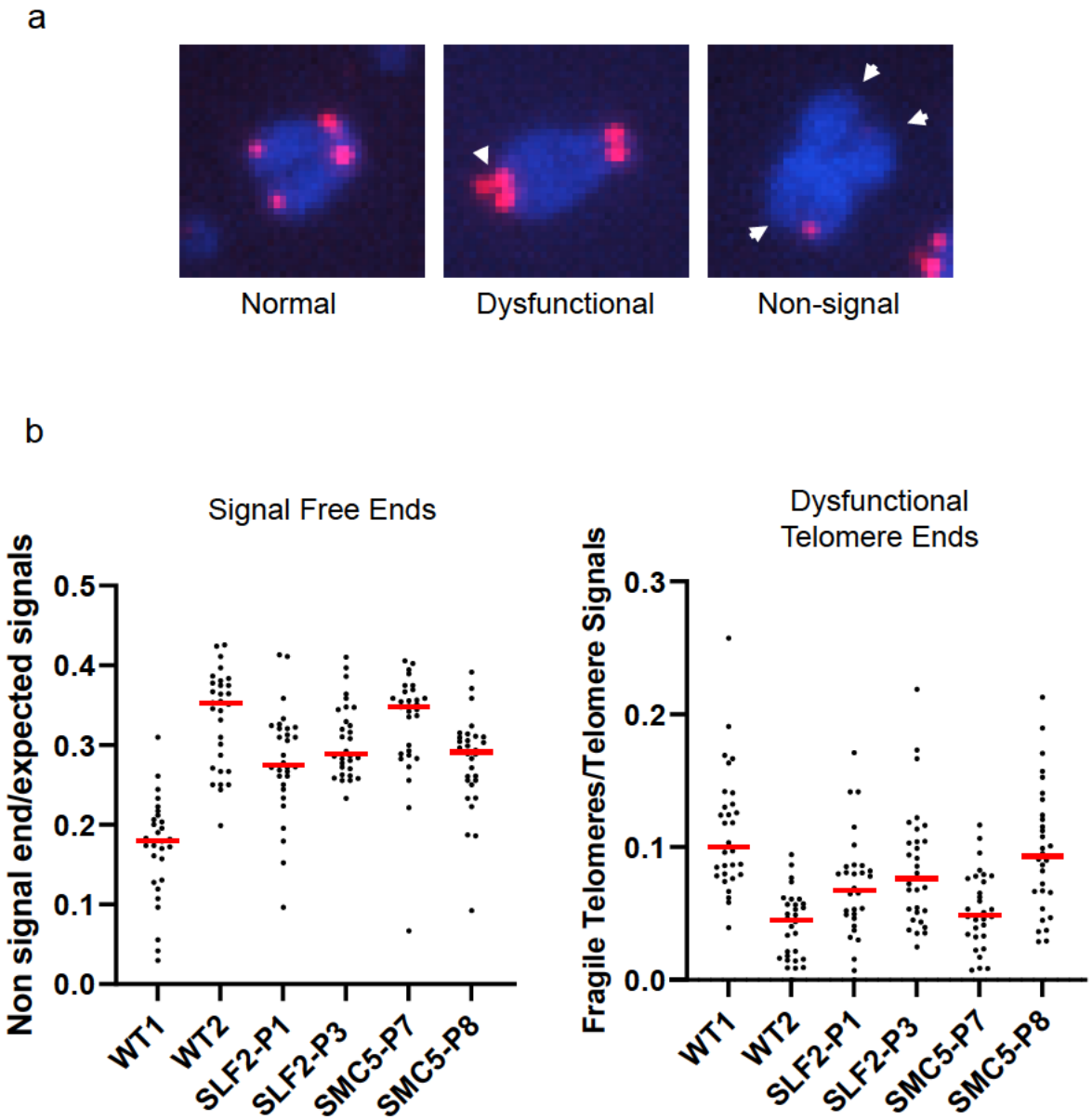


Figure 6.7. Levels of telomeric dysfunction vary considerably between lymphoblastoid cell lines

(a) Representative images of telomere defects, indicated by white arrows.
(b) The number of signal free (left) and dysfunctional (right) telomere ends in lymphoblastoid cell lines derived from SLF2 and SMC5 patients, as well as an unrelated, healthy individual (WT1 and WT2), visualised using fluorescence in situ hybridization (FISH) with a telomere specific probe. The red line denotes the median. 30 metaphases were counted from one experiment.

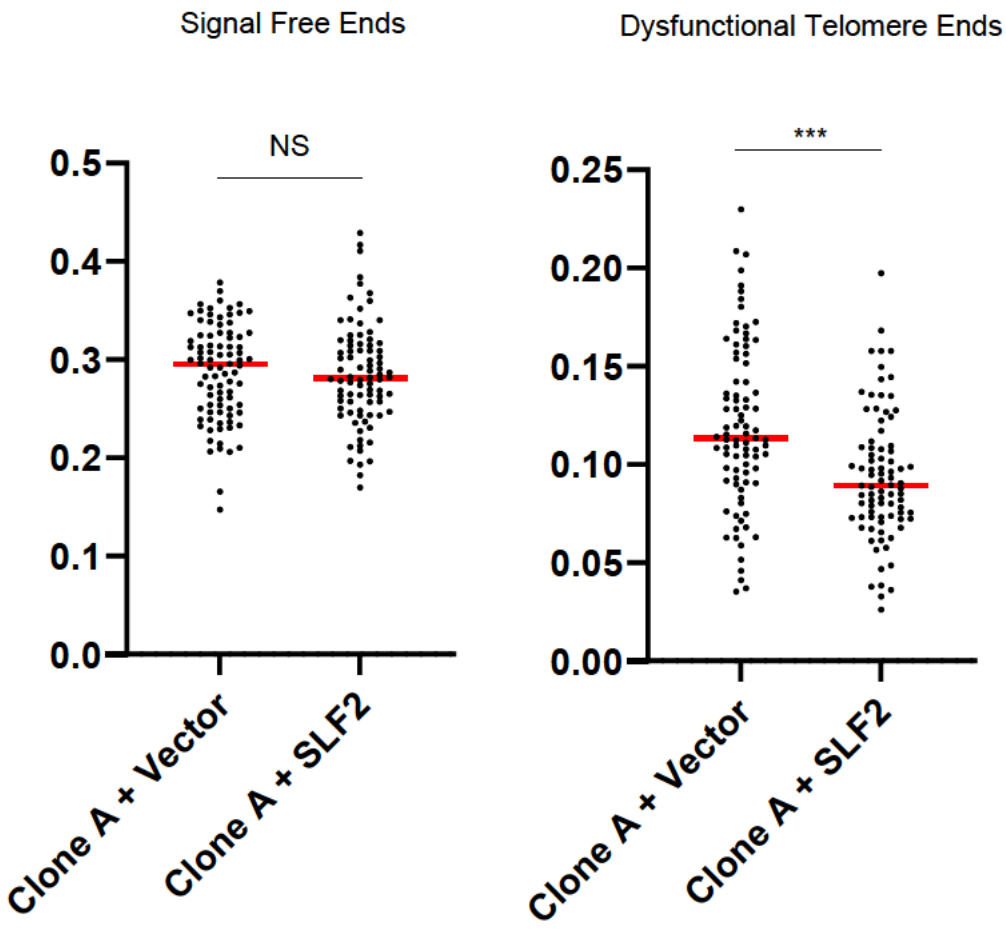


Figure 6.8. SLF2 CRISPR U-2-OS clone A exhibits elevated levels of dysfunctional telomeres

The number of signal free (left) and dysfunctional (right) telomere ends in SLF2 CRISPR U-2-OS clone A complemented with a lentivirus encoding wildtype SLF2 or an empty vector, visualised using fluorescence in situ hybridization (FISH) with a telomere specific probe. The red line denotes the median. n=3 independent experiments. >80 metaphases were counted. Statistical differences were determined with a Mann Whitney Rank Sum test.

to dysfunctional HR, Dr. Ryan Baxley, from the University of Minnesota, used chromosome orientation FISH (CO-FISH) to measure the frequency of telomere sister chromatid exchanges (t-SCEs) in the SLF2 patient LCLs (Baxley et al., 2021). By using strand specific telomeric PNA probes, this technique enables the user to visualise exchanges in genetic information. LCLs derived from SLF2-P1, SLF2-P3, SLF2-P4-1, and SMC5-P8 all showed significantly elevated levels of telomeric sister chromatid exchanges (t-SCEs) (Grange et al., 2022) compared to three WT control LCLs. Together this suggests a role for SLF2 in maintaining genome stability at telomeres by suppressing HR. This would be consistent with the known role of the SMC5/6 complex in regulating recombination at repetitive genomic regions.

6.2.2 SLF2 and SMC5 patient cells exhibit increased levels of recombination intermediates

Whilst a telomere maintenance defect could potentially explain the presence of the 'type 2' dicentric chromosomes, it is not as obvious as to why SLF2 and SMC5 patients exhibit increased levels of 'Type 1' segmented chromosomes. These 'Type 1' chromosomes were reminiscent of the segmented chromosomes seen in cells which are defective in the resolution of recombination intermediates, via the co-depletion of SLX4 and BLM, SLX4 and GEN1, or MUS81 and GEN1 (Garner et al., 2013, Sarbajna et al., 2014, Chan et al., 2018). Accordingly, I next sought to question if the elevated levels of t-SCEs seen in SLF2 and SMC5 patients reflected a genome-wide increase in crossover events. To do so, I first quantified the levels of global sister chromatid exchanges (SCEs) in metaphase spreads prepared from patient LCLs (Figure 6.9). Analysis showed that both SLF2 and SMC5 patient cell lines exhibited

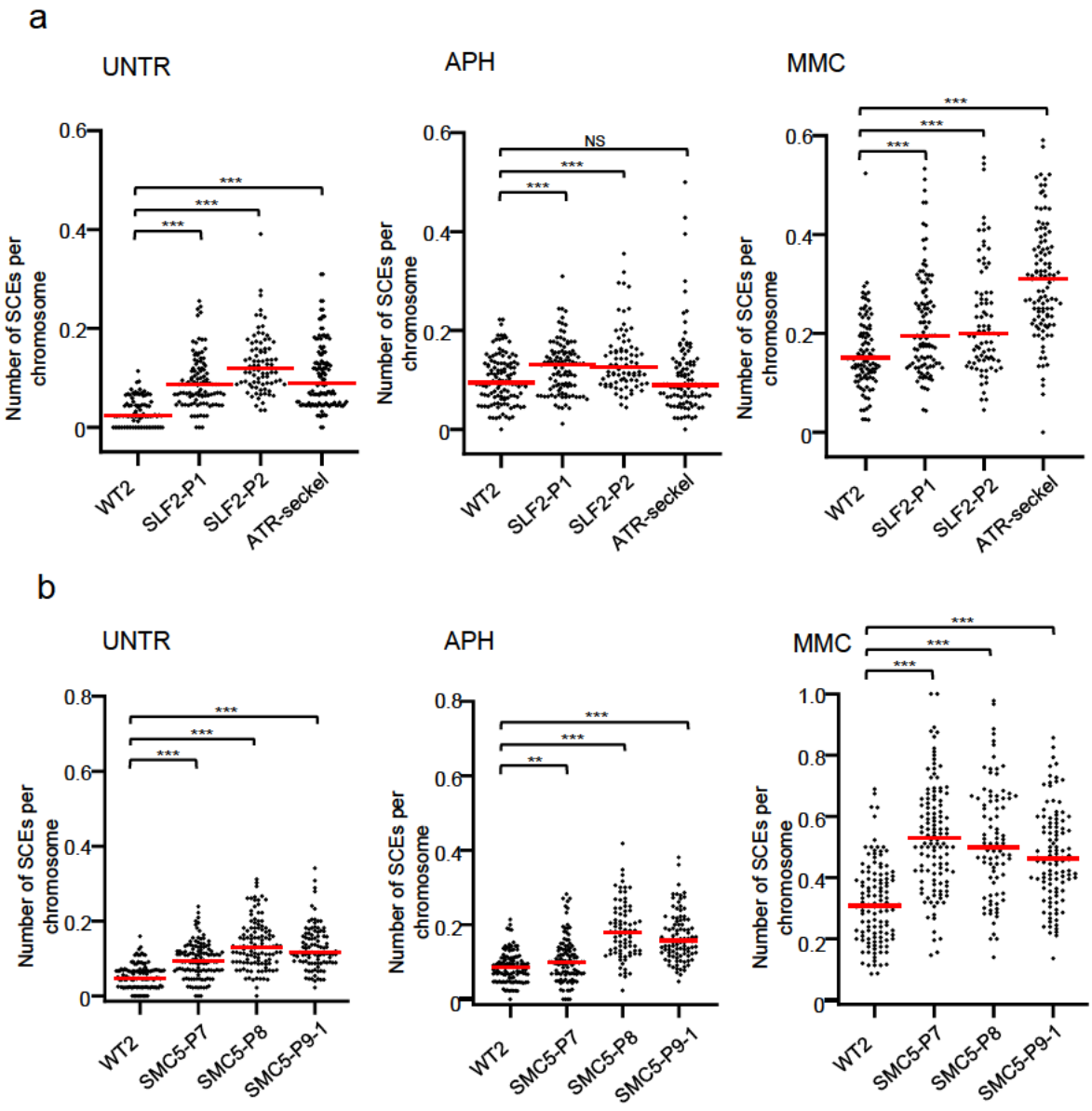


Figure 6.9. SLF2 and SMC5 patient lymphoblastoid cell lines exhibit elevated levels of sister chromatid exchanges

The number of sister chromatid exchanges in metaphase spreads produced from SLF2 (a) and SMC5 (b) patient lymphoblastoid cell lines. Cells were either untreated or exposed to 500 nM aphidicolin (APH) or 50ng/ml mitomycin-C (MMC) for 24 h. n=3 independent experiments. In total >100 metaphases were counted. Statistical differences were determined with a Mann Whitney Rank Sum test.

spontaneous increases in SCEs in the absence of exogenous replication stress, comparable to the levels observed for the ATR-seckel syndrome LCL. After treatment with MMC and APH, levels of SCEs were not notably exacerbated in the either SLF2 or SMC5 patient cell lines when compared to WT controls. Contrastingly, the ATR-seckel syndrome LCL exhibited a significantly elevated number of crossover events in response to MMC treatment.

To further confirm the increase in HR events in the patient cell lines, I performed immunofluorescence analysis on SLF2 and SMC5 patient fibroblasts to quantify the levels of spontaneous RAD51 foci in S/G2 cells (Figure 6.10). This analysis showed that fibroblasts complemented with an empty vector exhibited significantly increased levels RAD51 foci compared to cell lines complemented with WT SLF2 or SMC5. Taken together, these observations demonstrate that both SLF2 and SMC5 patient-derived cell lines exhibit increased levels of recombination, both globally and also at telomeres, consistent with the known roles of the SMC5/6 complex in regulating recombination at stalled replication forks and repetitive genomic regions. Furthermore, this also suggests that, similar to cells lacking recombination resolution and dissolution pathways, the 'Type 1' segmented chromosomes present in SLF2 and SMC5 patient cells may result from increased levels of unresolved recombination intermediates (Garner et al., 2013, Sarbajna et al., 2014, Chan et al., 2018).

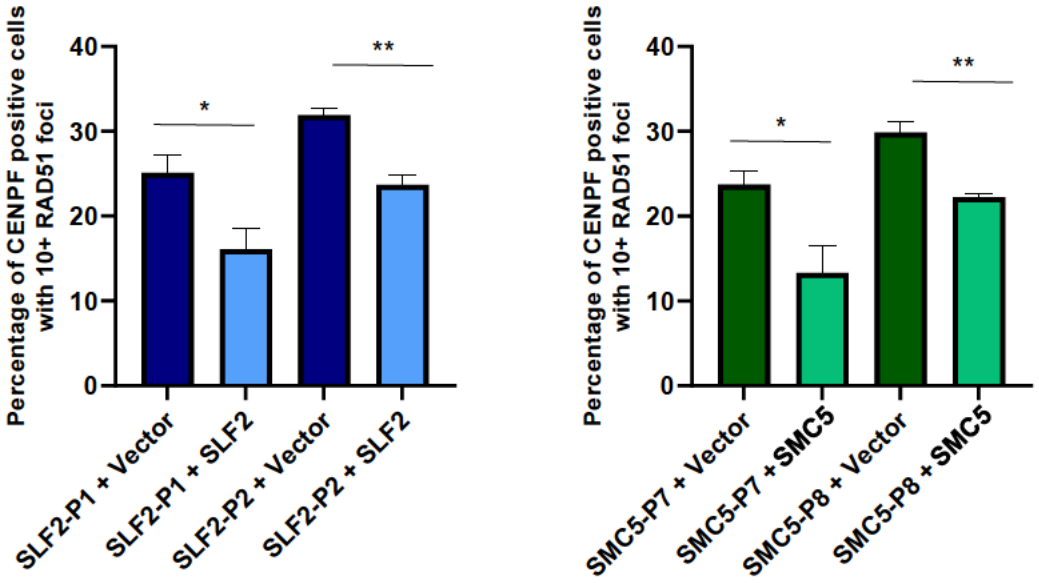
6.2.3 SLF2 and SMC5 patient cells exhibit increased sensitivity to G4 quadruplex stabilising agents

As discussed, SMC5/6 is implicated in the proper resolution of recombination intermediates at difficult to replicate regions of the genome, including repetitive regions such as rDNA and telomeres (Torres-Rosell et al., 2005, Menolfi et al., 2015).

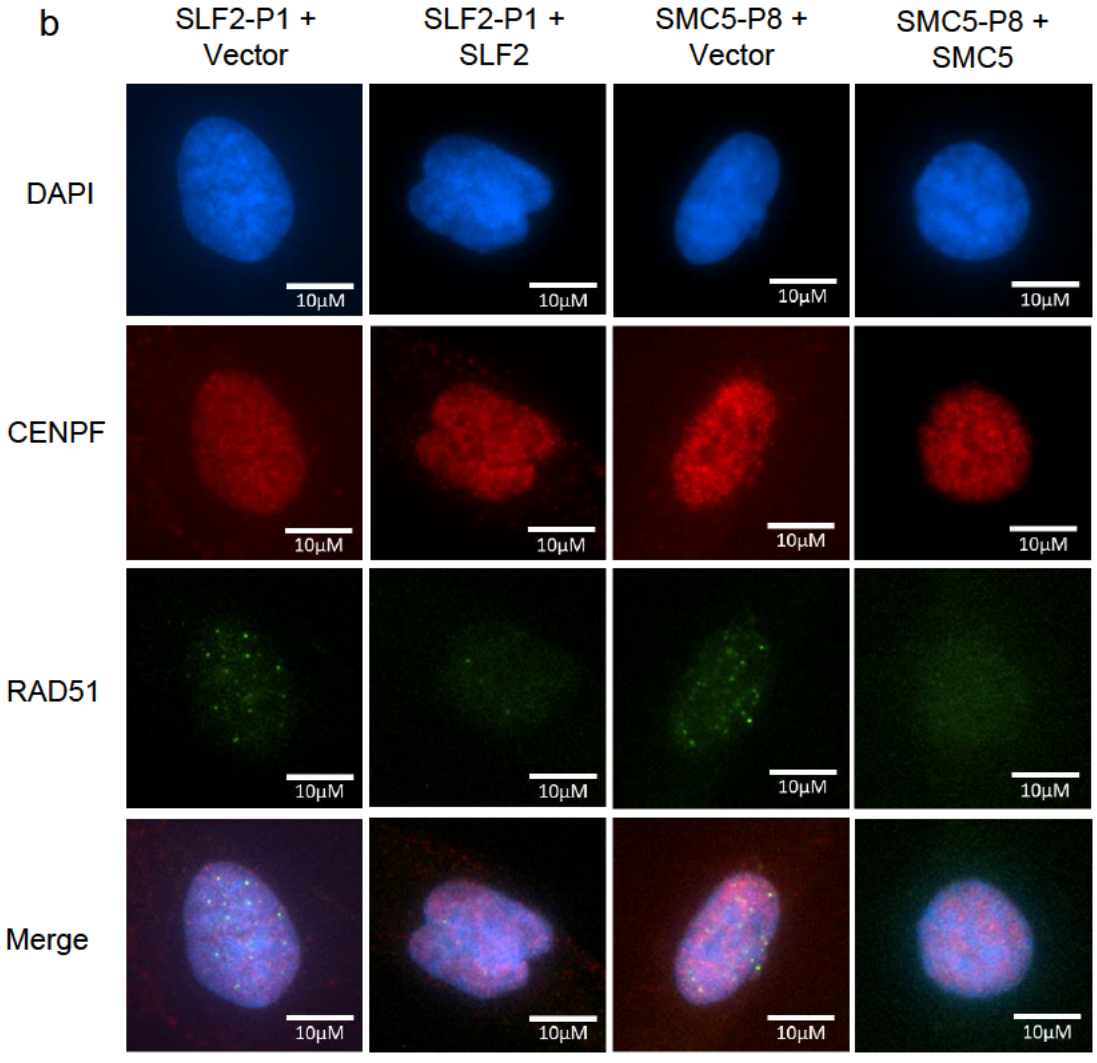
Figure 6.10. SLF2 and SMC5 patient fibroblasts exhibit elevated levels of RAD51 foci (data shown overleaf)

(a) The average percentage of S/G2-phase cells (CENPF positive) with >10 RAD51 foci in SLF2 (left) and SMC5 (right) patient fibroblasts complemented with a lentivirus encoding WT SLF2, WT SMC5, or an empty vector. n=3 independent experiments. In total >140 metaphases were counted. Statistical differences were determined with an unpaired, 2-tailed, student's t-test.
(b) Representative immunofluorescent images of patient derived fibroblasts, complemented with a lentivirus encoding WT SLF2, WT SMC5, or an empty vector. DAPI shown in blue, CENPF in red and RAD51 in green.

a


Percentage of CENPF positive cells with 10+ RAD51 foci
40
30
20
10
0
*
**
SLF2-P1 + Vector
SLF2-P1 + SLF2
SLF2-P2 + Vector
SLF2-P2 + SLF2
Percentage of CENPF positive cells with 10+ RAD51 foci
40
30
20
10
0
*
**
SMC5-P7 + Vector
SMC5-P7 + SMC5
SMC5-P8 + Vector
SMC5-P8 + SMC5

b


SLF2-P1 + Vector
SLF2-P1 + SLF2
SMC5-P8 + Vector
SMC5-P8 + SMC5
DAPI
CENPF
RAD51
Merge
10μM

Consistent with this role for the SMC5/6-SLF2-SLF1-RAD18 pathway, knockdown of RNF168 has recently been shown to sensitize HCT116 cells to the G4-quadruplex stabilising agent CX-5461 (Masud et al., 2021). G4-quadruplexes form at guanine-rich regions, including telomeres where there is emerging evidence that the structures contribute to end protection (Smith et al., 2011, Jurikova et al., 2020). However, defects in the resolution of G4-quadruplexes could lead to difficulties in replicating through regions that are prone to forming these secondary structures and give rise to elevated levels of DNA damage, under-replicated regions of DNA and recombination events, including problems at telomeric regions (Edwards et al., 2014, Lemmens et al., 2015, Kolesnikova and Curtis, 2019, Lerner and Sale, 2019, Zhang et al., 2019).

Considering this, I sought to investigate whether difficulties in replicating through G4-quadruplexes were a source of endogenous replication stress in SLF2 and SMC5 patient cells. To answer this, I questioned if treatment with CX-5461, a G4-quadruplex stabilising agent, induced elevated levels of DNA damage. I first prepared metaphase spreads from patient LCLs and quantified the levels of chromosome aberrations (chromosome gaps and breaks, chromatid gaps and breaks, and chromosome radials) (Figure 6.11). Analysis revealed that CX-5461 treatment induced elevated levels of aberrations in SLF2-P1 and SMC5-P8 LCLs when compared to WT cells. Increased levels of CX-5461 induced chromosome aberrations were also observed in U-2-OS SLF2 CRISPR mutant clone A containing empty vector in comparison to WT SLF2 expressing cells (Figure 6.12).

Interestingly, despite ‘Type 1’ segmented chromosomes being far less frequent in the LCLs in comparison to the peripheral blood lymphocytes, exposure to CX-5461 induced increases in ‘Type 1’ segmented chromosomes in SLF2 and SMC5 patient

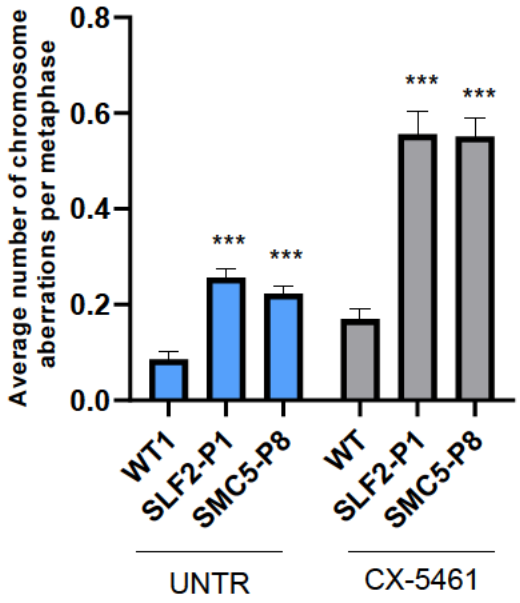


Figure 6.11. SLF2 and SMC5 patient lymphoblastoid cell lines exhibit elevated levels of chromosome aberrations in response to CX-5461 treatment

Quantification of the average number of chromosome aberrations (chromosome gaps and breaks, chromatid gaps and breaks, and chromosome radials) in metaphase spreads prepared from lymphoblastoid cell lines derived from patients with mutations in SLF2, SMC5, or an unrelated, healthy individual (WT1). Cells were either untreated or exposed to 250 nM CX-5461 for 24 hours. In total >150 metaphases were counted from at least 3 independent experiments. Statistical differences were determined with an unpaired, 2-tailed, student's t-test.

Figure 6.12. SLF2 CRISPR U-2-OS clone 1 exhibits elevated levels chromosome aberrations in response to CX-5461 treatment

Quantification of the average number of chromosome aberrations (chromosome gaps and breaks, chromatid gaps and breaks, and chromosome radials) in U-2-OS SLF2 CRIPSR clone A, complemented with a lentivirus encoding wildtype SLF2 or an empty vector. n=3 independent experiments. 150 metaphases were counted. Statistical differences were determined with an unpaired, 2-tailed, student's t-test.

LCLs, whilst levels were unchanged in the WT control (Figure 6.13). This may indicate that difficulties in replicating through G4-quadruplex regions not only contribute to elevated levels of DNA damage, but may also give rise to the unusual ‘Type 1’ chromosome gaps in SLF2 and SMC5 patient cell lines.

We then questioned if the increase in chromosome instability in the patient cell lines following CX-5461 treatment was associated with elevated replication stress. To do so, the DNA fibre assay was utilised. In untreated cells (Figure 6.14a, Figure 6.14b), SLF2-P1 and SMC5-P8 fibroblasts showed no difference in the IdU/CldU ratio between cell lines complemented with and without WT SLF2/SMC5 protein. A ratio close to one indicated that replication had progressed at a similar rate when incubated with the first label (CldU) and second label (IdU). However, when CX-5461 was added to the second IdU label, patient fibroblasts complemented with empty vector exhibited a decrease in the ratio of IdU/CldU tract length, showing that replication had slowed or stalled (Figure 6.14b). In cell lines complemented with either WT SLF2 or SMC5, the IdU/CldU ratio was restored to approximately one, suggesting that SLF2 and SMC5 are required for proper replication through G4-quadruplex structures.

To question if the CX-5461-induced replicative defects led to increased replication stress persisting throughout the cell cycle, I also used patient fibroblasts to quantify levels of CX-5461 induced 53BP1 bodies in G1 (Figure 6.14c). The proportion of fibroblasts with 10+ bodies was consistently low in untreated experiments. However, treatment with CX-5461 for 24 hours increased these levels, with patient fibroblasts complemented with empty vector exhibiting a significant increase in the proportion of cells with greater than 10 53BP1 foci when compared to cells complemented with WT SLF2 or SMC5. Next, I sought to determine if the increased levels of CX-5461-

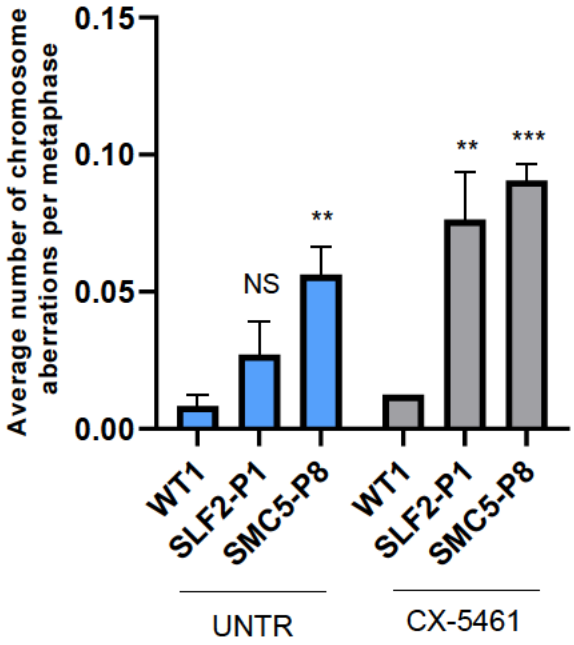


Figure 6.13. SLF2 and SMC5 patient lymphoblastoid cell lines exhibit elevated levels of 'Type I' chromosomes in response to CX-5461 treatment

Quantification of the average number of 'Type 1' segmented chromosomes in metaphase spreads prepared from lymphoblastoid cell lines derived from patients with mutations in SLF2 and SMC5, as well as an unrelated, healthy individual (WT1). Cells were either untreated or exposed to 250 nM CX-5461 for 24 hours. In total >320 metaphases were counted from at least 4 independent experiments. Statistical differences were determined with an unpaired, 2-tailed, student's t-test.

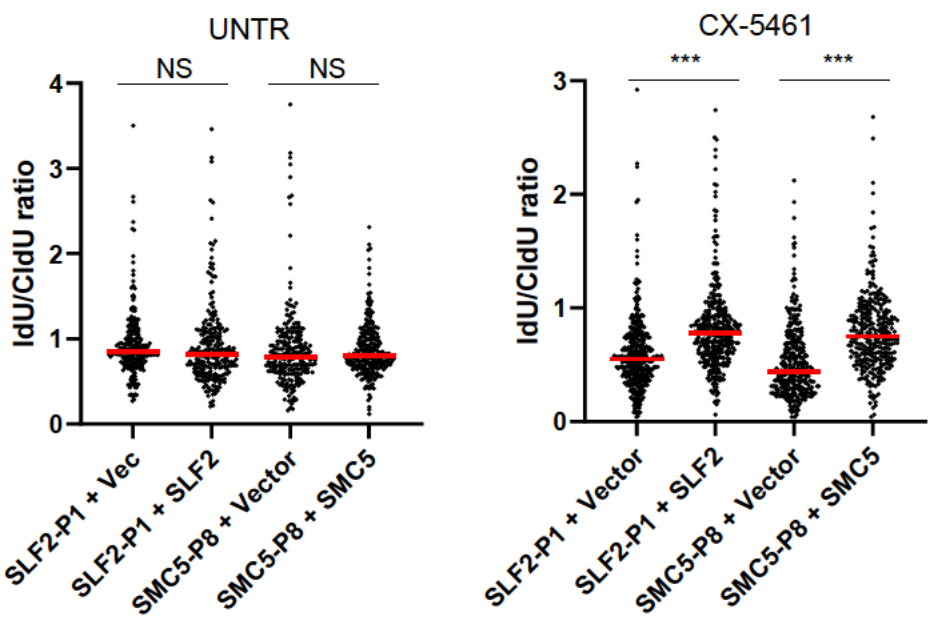
Figure 6.14. SLF2 and SMC5 patient derived fibroblasts exhibit fork slowing and elevated levels of 53BP1 bodies in response to CX-5461 treatment (data shown overleaf)

(a) Experimental outline of the DNA fibre assay. Fibroblasts were sequentially labelled with two different thymidine analogs, CldU (red) and IdU (green) for 30 mins each, as shown, with or without 250nM CX-5461 in the IdU label.
(b) Quantification of the ratio of IdU tract length to CldU tract length in untreated and CX-5461 treated SLF2 and SMC5 patient fibroblasts complemented with a lentivirus encoding WT SLF2, WT SMC5, or an empty vector. Cells were pulse labelled with CldU for 30 minutes and then pulse-labelled with IdU, with or without 250nm CX-5461. n=3 independent experiments. >250 ongoing fork structures were counted. Statistical differences were determined with a Mann Whitney Rank Sum test. Experiments carried out by Dr John Reynolds.
(c) Quantification of the percentage of cells with >10 53BP1 bodies in SLF2 and SMC5 patient fibroblasts complemented with a lentivirus encoding WT SLF2, WT SMC5, or an empty vector. Cells were either untreated or exposed to 250nm CX-5461 for 24 hours. n=3 independent experiments. 300 cells were counted. Statistical differences were determined with an unpaired, 2-tailed, student's t-test.

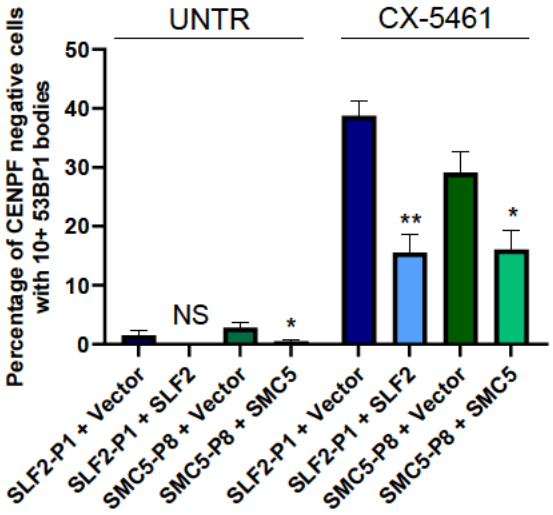
a


30 min
30 min
CldU
IdU
Harvest
+/- CX-5461
Ongoing fork:

b


UNTR
CX-5461
NS
NS
***
***
IdU/CldU ratio
0
1
2
3
4
SLF2-P1 + Vec
SLF2-P1 + SLF2
SMC5-P8 + Vector
SMC5-P8 + SMC5
SLF2-P1 + Vector
SLF2-P1 + SLF2
SMC5-P8 + Vector
SMC5-P8 + SMC5

c


UNTR
CX-5461
Percentage of CENPF negative cells with 10+ 53BP1 bodies
0
10
20
30
40
50
NS
*
**
*
SLF2-P1 + Vector
SLF2-P1 + SLF2
SMC5-P8 + Vector
SMC5-P8 + SMC5
SLF2-P1 + Vector
SLF2-P1 + SLF2
SMC5-P8 + Vector
SMC5-P8 + SMC5

induced replication stress had an impact on the viability of the patient cells (Figure 6.15). WT, SLF2-P1 and SLF2-P8 lymphoblastoid cell lines were continuously exposed to increasing concentrations of CX-5461. After three population doublings in the untreated cells, the cell numbers of both untreated and treated cells were counted, and cell viability of the treated cells was expressed as a percentage of untreated cells. This analysis showed that, in comparison to the WT control, the SLF2 and SMC5 patient cell lines exhibit decreased cellular viability with increasing concentrations of CX-5461.

Taken together, these observations confirm that, in the absence of functional SLF2 and SMC5, exposure to CX-5461 causes increased levels of chromosome instability, replication stress and a loss of cellular viability due to replication abnormalities. This is consistent with a novel role for the SMC5/6 complex in the resolution of G4-quadruplex lesions. However, as CX-5461 was first identified as an inhibitor of rDNA transcription (Drygin et al., 2011), it is possible that the phenotypes that I have described are due to the inhibition of RNA polymerase I and not the stabilisation of G4-quadruplexes. To control for this, I also investigated the impact of the small molecule DNA intercalator, BMH21, on the levels of chromosomal aberrations in the patient LCLs (Figure 6.16a). BMH21 binds to rDNA and inhibits RNA polymerase I transcription (Peltonen et al., 2014). 24 hours incubation with the drug induced elevated numbers of aberrations in all LCLs, with no differences between SLF2 and SMC5 patient cell lines compared to WT controls. Furthermore, whilst exposure to a titration of BMH21 concentrations led to elevated levels of G1 53BP1 bodies in SLF2-P1 and SMC5-P8 fibroblasts, there were also no differences between cell lines

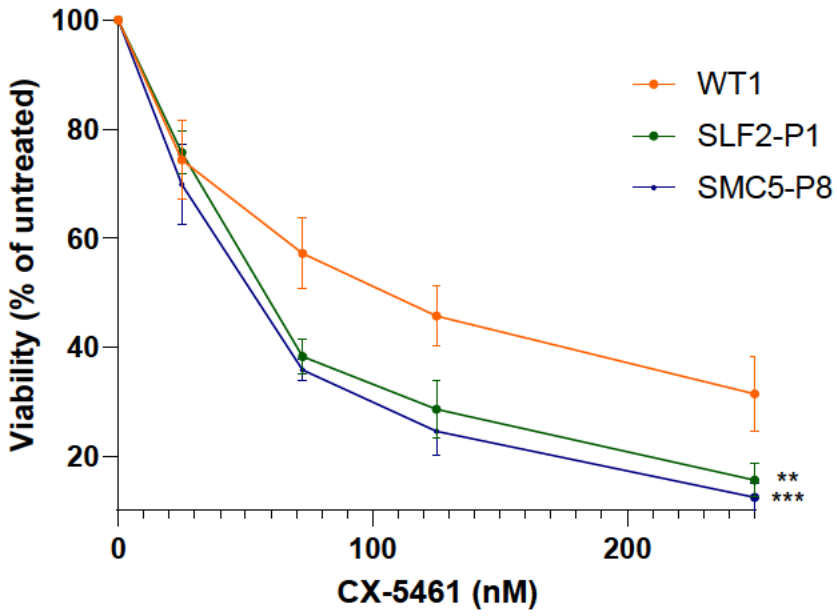


Figure 6.15. SLF2 and SMC5 patient LCLs exhibit decreased cellular viability upon exposure to CX-5461 treatment

The cellular viability of lymphoblastoid cell lines derived from SLF2 and SMC5 patients, as well as an unrelated, healthy individual (WT1), in response to increasing doses of CX-5461 treatment. Cells were continuously exposed to increasing concentrations of CX-5461. After the untreated cells have undergone three population doubling times, cells were counted and plotted as percentage of untreated cells. n=3 independent experiments. Statistical differences were determined using a 2-way ANOVA. Error bars denote SEM.

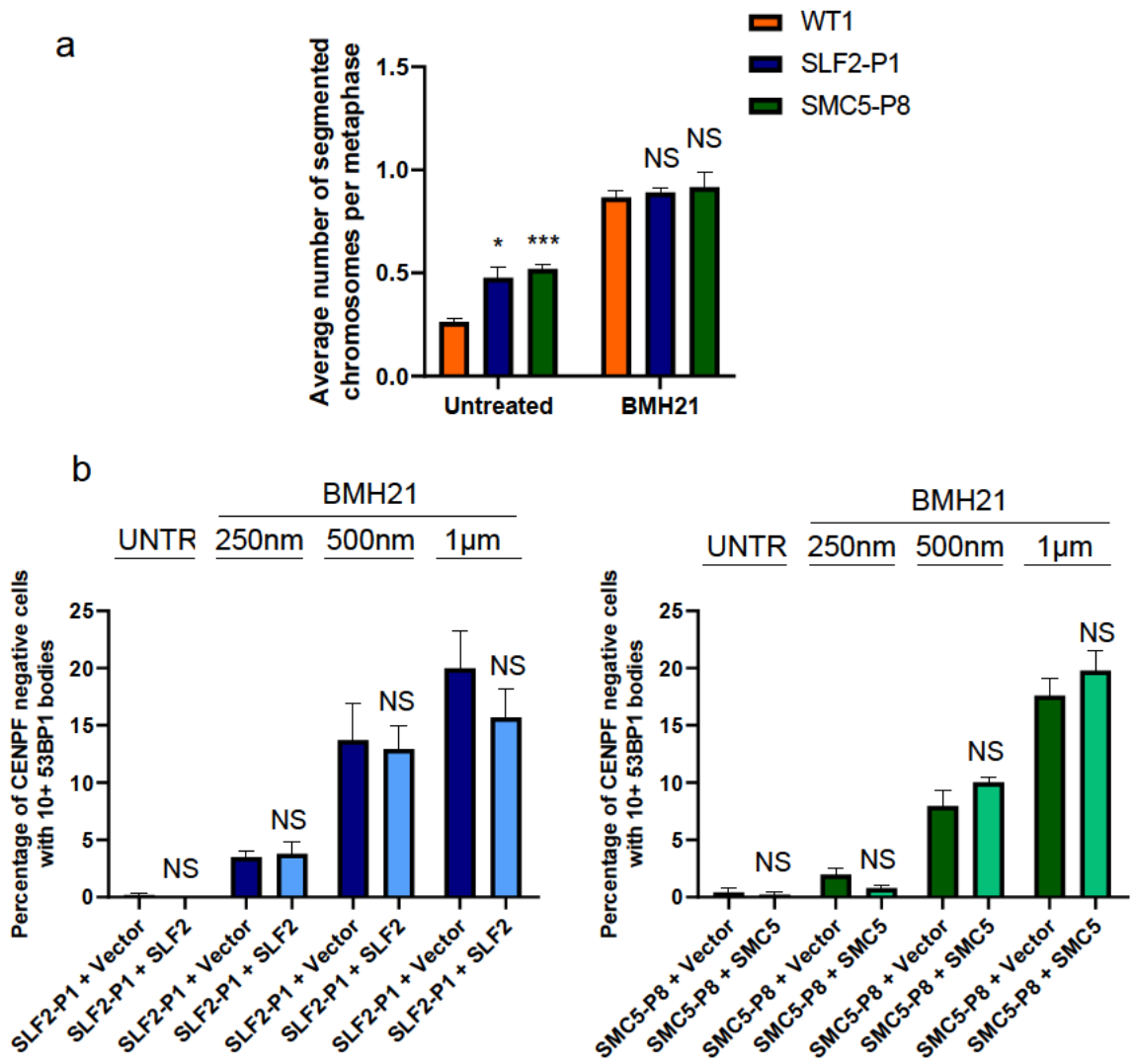


Figure 6.16. SLF2 and SMC5 patient cells do not exhibit elevated levels of chromosome aberrations or 53BP1 bodies upon exposure to BMH21

(a) Quantification of the average number of chromosome aberrations (chromosome gaps and breaks, chromatid gaps and breaks, and chromosome radials) in metaphase spreads prepared lymphoblastoid cell lines derived from SLF2 and SMC5 patients, as well as an unrelated, healthy individual (WT1). Cells were either untreated or exposed to 1µm BMH21 for 24 hours. In total >150 metaphases were counted from 3 independent experiments. Statistical differences were determined with an unpaired, 2-tailed, student's t-test.

(b) Quantification of the average percentage of cells with 10+ 53BP1 bodies in SLF2 (left) and SMC5 (right) patient fibroblasts complemented with a lentivirus encoding WT SLF2, WT SMC5, or an empty vector. Cells were exposed to titration of BMH21 (untreated, 250nm, 500nm, 1µm) for 24 hours. n=3 independent experiments. 300 cells were counted. Statistical differences were determined with an unpaired, 2-tailed, student's t-test.

complemented with WT protein or empty vector at any of the tested drug dosages (Figure 6.16b). EU staining was used to confirm that the doses of BMH21 treatment used on the fibroblast patient cell lines were sufficient to inhibit rDNA transcription, as previously reported (Figure 6.17) (van Schie et al., 2020). Finally, I confirmed that when incubated with 1μM BMH21, a dose that inhibits rDNA transcription, SLF2 and SMC5 patient fibroblasts possessing an empty vector could still replicate as efficiently as cells expressing WT SLF2 and SMC5 protein (Figure 6.18). This confirms that the replication stress phenotypes in cells treated with CX-5461 are more likely to be due to the drug's G4-quadruplex stabilising ability, rather than an indirect effect of inhibiting RNA polymerase I.

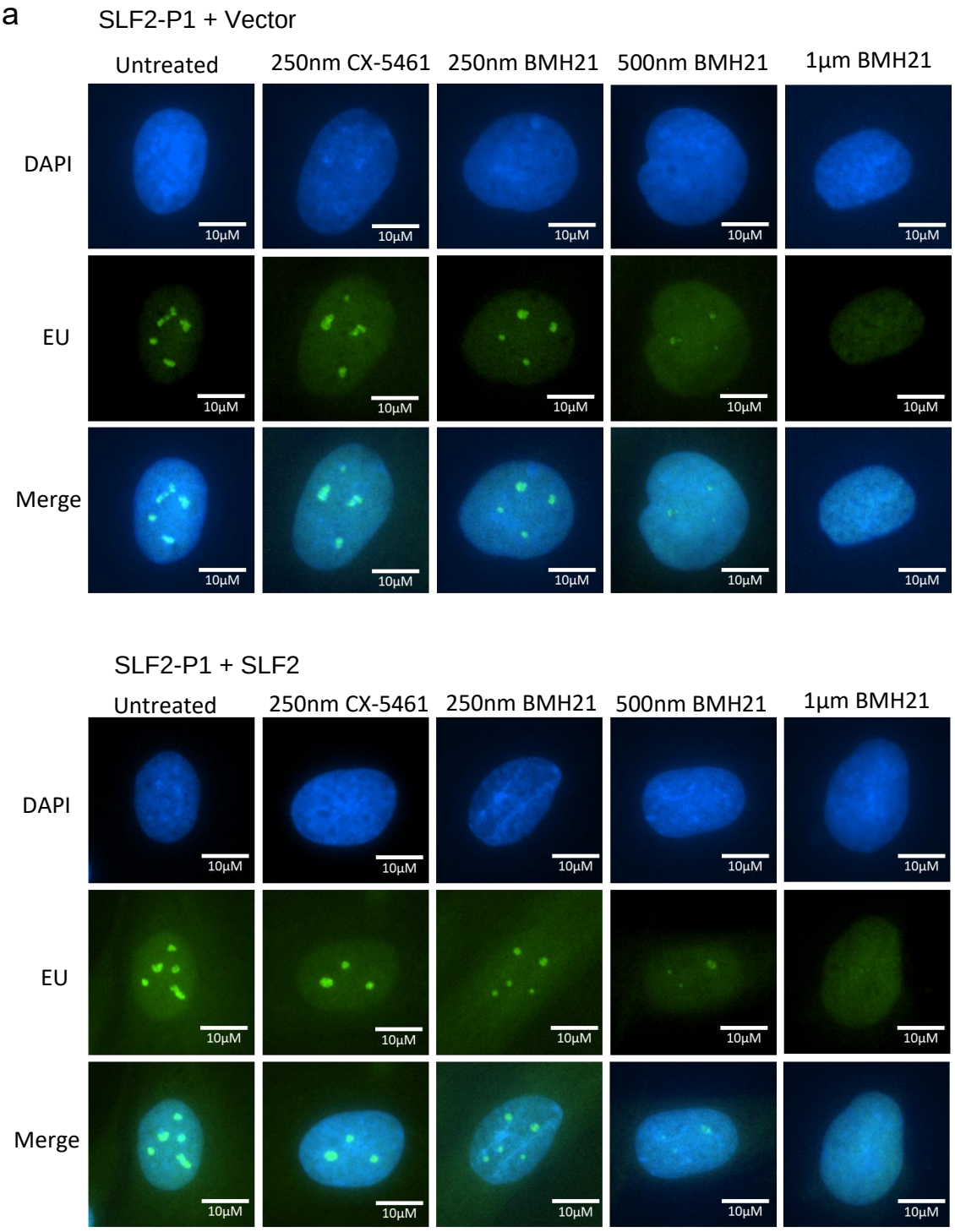
6.3 Discussion

The data presented in this chapter demonstrates that cell lines derived from patients with mutations in SLF2 or SMC5 exhibit elevated levels of chromosome instability, sister chromatid exchanges and RAD51 foci, which supports previous findings that SMC5/6 is important for promoting proper HR, resolving recombination intermediates, and limiting aberrant crossover events (Potts et al., 2006, Ampatzidou et al., 2006, Chen et al., 2009, Bermudez-Lopez et al., 2010, Agashe et al., 2021). My observations suggest that the SMC5/6-SLF2-SLF1-RAD18 pathway is important for these functions, and that HR defects contribute to the chromosomal damage observed in the patient cell lines. Several publications using yeast have proposed that Smc5/6's role in role HR may centre on its regulation of the STR 'dissolvasome', a complex involved in processing recombination intermediates, via an interaction with Sgs1 (BLM's homolog) and the SUMOylation of STR components (Bonner et al., 2016, Bermudez-Lopez et al., 2016). Considering this, it would be interesting to know

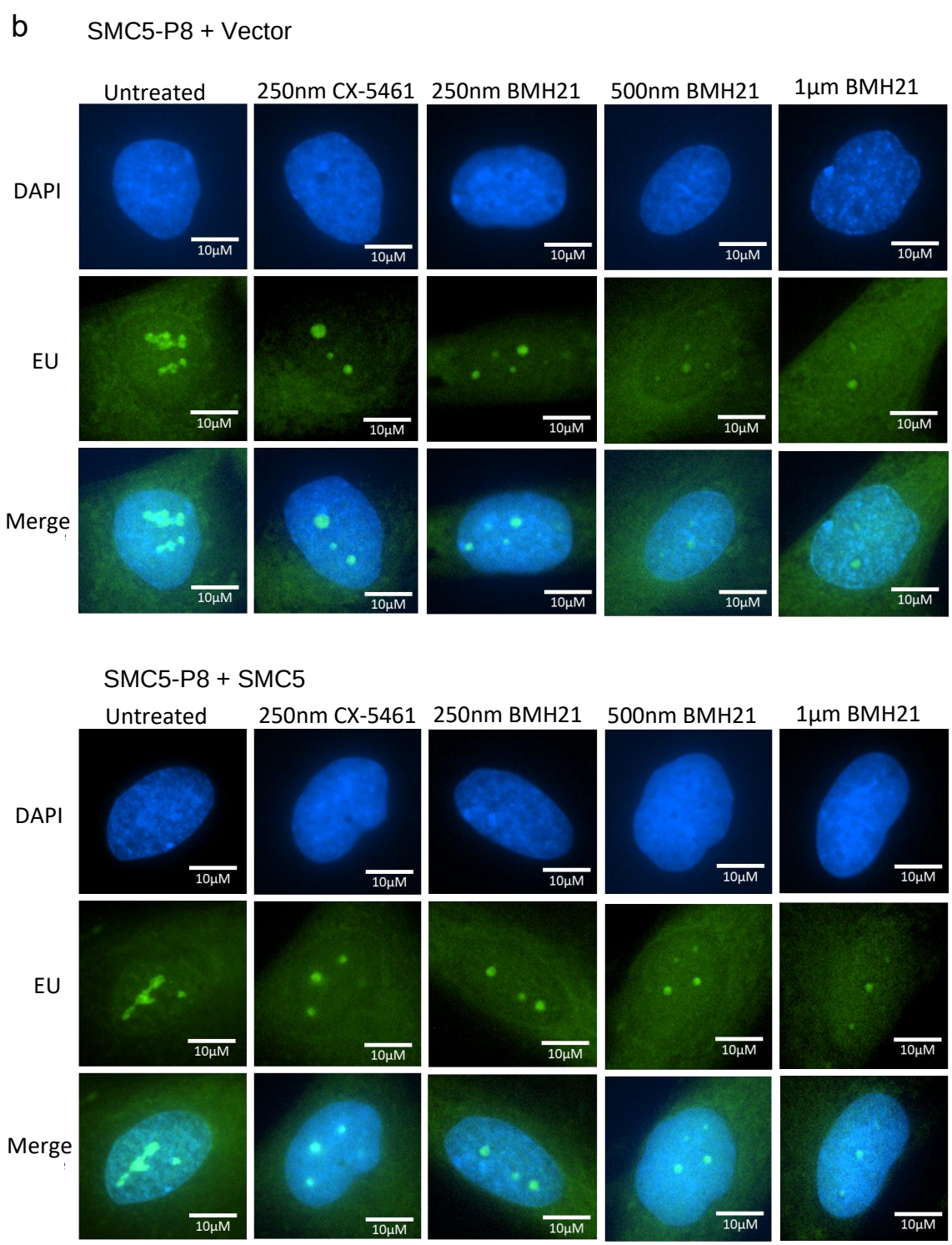
Figure 4.17. The impact of CX-5461 and BMH21 treatment on RNA transcription (data shown overleaf)

Immunofluorescent images of SLF2 (a) and SMC5 (b) patient fibroblasts, complemented with a lentivirus encoding wildtype (WT) SLF2, WT SMC5, or an empty vector, indicating the impact CX-5461 and BMH21 treatments on RNA transcription. Cells were either untreated or exposed to 250nm of CX-5461 or a titration of BMH21 (250nm, 500nm, 1μm) for 24 hours. RNA transcription was monitored using EU labelling. 1 mM of EU was added to the cells 30 minutes before fixation and antibody staining. Images are representative of two independent experiments which produced similar results.

a


SLF2-P1 + Vector
Untreated
250nm CX-5461
250nm BMH21
500nm BMH21
1µm BMH21
DAPI
EU
Merge
10µM
SLF2-P1 + SLF2
Untreated
250nm CX-5461
250nm BMH21
500nm BMH21
1µm BMH21
DAPI
EU
Merge
10µM

b


SMC5-P8 + Vector
Untreated
250nm CX-5461
250nm BMH21
500nm BMH21
1µm BMH21
DAPI
EU
Merge
10µM
SMC5-P8 + SMC5
Untreated
250nm CX-5461
250nm BMH21
500nm BMH21
1µm BMH21
DAPI
EU
Merge
10µM

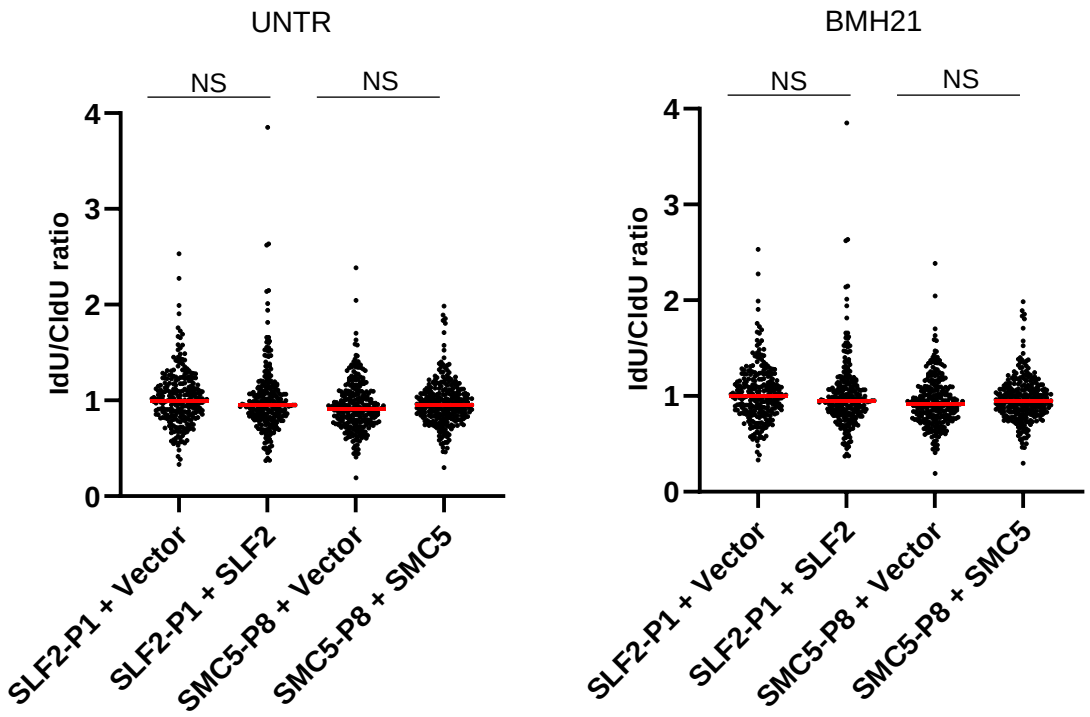


Figure 6.18. SLF2 and SMC5 patient derived fibroblasts do not exhibit fork slowing in response to BMH21 treatment

Quantification of the ratio of IdU tract length to CldU tract length in untreated and BMH21 treated SLF2 and SMC5 patient fibroblasts complemented with a lentivirus encoding WT SLF2, WT SMC5, or an empty vector. Cells were pulse-labelled with CldU for 30 minutes and then pulse-labelled with IdU, with or without 1μM BMH21. n=2 independent experiments. >250 ongoing fork structures were counted. Statistical differences were determined with a Mann Whitney Rank Sum test.

whether SMC5/6 also functions to SUMOylate and/or regulate BLM in humans in the same way as in yeast, and whether defects in this regulation contributes to the increased levels of recombination events that I observed in SLF2 and SMC5 patient cell lines. To begin to question this, immunoprecipitation experiments could be used to first investigate whether SMC5/6 and BLM interact in human cell lines.

Alongside increased levels of chromosomal gaps and breaks, patient peripheral blood lymphocytes exhibited elevated levels of unusual aberration types, categorised here as ‘Type 1’ and ‘Type 2’ segmented chromosomes. As previously discussed, the more common ‘Type 1’ chromosome gaps/breaks were reminiscent of chromosomal aberrations observed in previous publications characterising resolvase deficient cells (Garner et al., 2013, Sarbajna et al., 2014, Chan et al., 2018). Studies investigating these structures in GEN1 knockout cell lines, as well as SLX4-null cells depleted of BLM, were able to rescue elevated levels of chromosome segmentation via the exogenous expression of the bacterial resolvase RusA. On the basis of this, the authors proposed that this segmentation was caused by defects in chromatid condensation around sites of unresolved chromosome intermediates (Garner et al., 2013, Chan et al., 2018). The phenotypic similarities between these observations of chromosome segmentation are a further indication that mutations in SLF2 and SMC5 contribute to problems in HR, with ‘Type 1’ chromosome gaps/breaks possibly arising due to the accumulation of unresolved recombination intermediates. However, my work fails to establish if this is due to an elevated number of HR events being initiated, or whether the increased observation of recombination events are due to delayed/defective processing during HR.

Due to the aberrations' scarcity in immortalised patient-derived cell lines, further study of the 'Type 1' segmented chromosomes was difficult. However, treatment with CX-5461 induced elevated levels of segmentation in patient LCLs. As a stabiliser of G4-quadruplexes, this observation may suggest that SMC5/6 is particularly important for ensuring proper HR at certain regions with complex secondary structures. Whilst G4-quadruplexes have regulatory functions in several biological processes, these structures are also believed to contribute to genomic instability by impeding replication forks (Edwards et al., 2014, Lemmens et al., 2015, Kolesnikova and Curtis, 2019, Lerner and Sale, 2019). Consistent with this, CX-5461 treatment led to elevated levels of replication stress. In *Saccharomyces cerevisiae*, SMC5/6 mutants show an accumulation of X-shaped structures at collapsed replication forks, suggesting that the complex is required to restart stalled forks and prevent the accumulation of recombination intermediates (Ampatzidou et al., 2006). Thus, it may be that G4-quadruplexes represent difficult-to-replicate regions that specifically require SMC5/6 to either resolve these impediments or promote HR-mediated restart following fork collapse. To question this further, it would be interesting to ask if CX-5461 treatment induces elevated levels of SCEs and RAD51 foci, or if the CX-5461 induced recombination or replication defects are rescued by RusA resolvase expression.

Metaphase spread analysis also revealed elevated numbers of putative telomere fusions, referred to as 'Type 2' dicentric chromosomes, in the patient peripheral blood lymphocytes. This suggested that the SMC5/6-SLF2-SLF1-RAD18 pathway is important for maintaining telomeric integrity. In addition to this, U-2-OS SLF2 CRISPR clone A exhibited increased levels of telomere dysfunction. Whilst this may be consistent with observations of telomere fusions in patient cells, EBV-transformed cell

lines maintain telomere length via the stimulation of telomerase, whereas U-2-OS utilise the alternative lengthening of telomeres (ALT) pathway, which is unique to cancer cell lines (Dilley and Greenberg, 2015). Although ALT is poorly understood, it is thought that the process elongates telomeres via HR in ALT-associated PML bodies (APBs) (Yeager et al., 1999, Cesare and Reddel, 2008). Previous studies have suggested that SMC5/6 has specific functions in promoting ALT, with the complex localising to APBs in ALT-positive cell lines where it is required for telomere recombination. Consistent with this, loss of SMC5/6 components in SUSM1 ALT cells are associated with progressive telomere shortening (Potts and Yu, 2007). Therefore, it may be that the telomere dysfunction in the U-2-OS SLF2 CRISPR mutant cell line reflects a role for the SMC5/6-SLF2-SLF1-RAD18 pathway in the HR-mediated regulation of telomeres via ALT.

The increased levels of t-SCEs in SLF2 and SMC5 patient-derived LCLs is consistent with studies that suggest that SMC5/6 has functions at telomeres in telomerase-positive cells, being enriched at chromosome ends and involved in telomere organisation (Zhao and Blobel, 2005, Torres-Rosell et al., 2005, Moradi-Fard et al., 2016). Considering that SLF2 is associated with ATRX, a chromatin remodeller which mediates HR at telomeres (Scott et al., 2021), it may be that the SMC5/6-SLF2-SLF1-RAD18 pathway has specific functions in repair or replication at these regions. Therefore, a loss of SMC5/6 pathway components could lead to elevated levels of telomeric DNA breakage and aberrant recombination between deprotected chromosome ends, giving rise to putative ‘Type 2’ dicentric chromosomes. Considering that G4-quadruplexes are believed to accumulate at telomeres (Bryan, 2020), a role for SMC5/6 at these regions is further implied by the sensitivity of SLF2

and SMC5 patient cell lines to CX-5461. However, it's difficult to understand if this reflects a specific role for SMC5/6 at chromosome ends, or instead a genome wide function for the complex at G4-quadruplexes or other complex secondary structures.

There are approximately 375,000 sites in the human genome, including promoter regions, which are predicted to be capable of forming G4-quadruplexes (Todd et al., 2005), and may also require the function of SMC5/6-mediated HR for repair or replication restart. DNA fibre experiments revealed a significant decrease in the speed of replication forks progressing in the presence of CX-5461 in patient fibroblasts. Whilst even mild genotoxic stress has been shown to lead to fork slowing (Zellweger et al., 2015), these findings imply that a large proportion of forks are impacted by CX-5461 treatment. This could suggest that SMC5/6 is important for proper replication, and possibly HR-mediated restart, through G4-quadruplexes at numerous loci, not just telomeres. To explore this further, it would be interesting to compare levels of whole chromosome SCEs to levels of t-SCEs in response to CX-5461 treatment in the patient cell lines. The telomere specific G4-stabilising agent, telomestatin (Kim et al., 2002), could also be used to investigate if HR or replication defects are similarly induced when only telomeric secondary structures are promoted. Telomeric DNA fibre replication assays could then be utilised to question if G4-quadruplex stabilisation is associated with elevated levels of telomeric replication fork stalling specifically (Baxley et al., 2021).

Whilst the work discussed in this chapter suggests a functional role for the SMC5/6-SLF2-SLF1-RAD18 pathway in the repair of or replication through regions with G4-quadruplex structures, there are caveats to this conclusion. I was not able to successfully optimise the use of anti-DNA/RNA G4-quadruplex antibodies for

immunofluorescence and was therefore unable to confirm if there are increased levels of endogenous G4 structures in SLF2 and SMC5 patient cell lines or even that CX-5461 treatment led to an increase in G4-quadruplex structures. This is important if I am to establish that the stabilisation of these structures causes the cellular sensitivity to the drug, rather than alternative functions of CX-5461. For these reasons, it would also be pertinent to verify my findings using additional G4-stabilising agents, such as pyridostatin. CX-5461 also has been shown to cause the formation of abortive Top2 DSBs (Olivieri et al., 2020, Bruno et al., 2020, Bossaert et al., 2021). Thus, it may be that SLF2 and SMC5 deficient cell lines are particularly sensitive to persistent DNA catenation rather than, or as well as, G4-quadruplexes. This may be explained by the possible association of SMC5/6 with condensin, as discussed in the previous chapter, a factor involved in resolving these entanglements. It may be that CX-5461-mediated inhibition of Top2 alongside a possible deregulation of condensin in SLF2 or SMC5 deficient cells leads to significant increases in DNA damage. Whilst a recent study in yeast suggested that SMC5/6 is not required to minimise Top2-mediated DNA entanglements (Dyson et al., 2021), it would be interesting to explore this in the patient cell lines. Detailed experimentation using additional topoisomerase inhibitors would determine if the defects presented in this chapter are related to DNA catenation.

Nevertheless, in this chapter I present several novel exciting findings, including the discovery of the unusual, segmented chromosome aberration phenotype present in metaphases prepared from SLF2 and SMC5 patient blood samples. The hypothesised role of SMC5/6 in the promotion of replication through G4-quadruplex structures would be newly discovered function for the SMC5/6 complex, and future

research on this would likely shed more light on the role of this complex in maintaining genome stability during DNA replication.

Chapter 7:

Discussion

The SMC complexes represent a highly conserved family of ring shaped ATPases, including cohesin (SMC2/4) and condensin (SMC1/3), both of which are known to play well characterised roles in mitosis to ensure faithful cellular division (Nasmyth and Haering, 2009, Aragon, 2018, Makrantoni and Marston, 2018). The human SMC5/6 complex comprises of a heterodimer of SMC5 and SMC6, alongside a number of accessory subunits (NSE1-4) (Taylor et al., 2001, Taylor et al., 2008, Hallett et al., 2021a). Although less well delineated, SMC5/6 is believed to function principally in DNA replication and repair (Ampatzidou et al., 2006, Bermudez-Lopez et al., 2010, Menolfi et al., 2015, Aragon, 2018, Palecek, 2018, Agashe et al., 2021). Whilst it is not yet understood how the individual subunits of SMC5/6 mechanistically contribute to these roles, NSE2 exhibits SUMOylation activities through which it can autoSUMOylate SMC5/6 complex components (Zhao and Blobel, 2005, Potts and Yu, 2005, Andrews et al., 2005, Kliszczak et al., 2012, Sole-Soler and Torres-Rosell, 2020). In humans, SLF2 has been shown to interact with the SMC5/6 complex and facilitate its recruitment to damaged chromatin as part of the linear SMC5/6-SLF2-SLF1-RAD18 protein pathway (Raschle et al., 2015).

The research described in this thesis identifies 11 patients with mutations in two SMC5/6 related factors, SLF2 and SMC5, and characterises the cellular defects present in cell lines and peripheral blood samples derived from these patients. These individuals exhibit microcephaly and short stature, alongside frequent cardiac defects, and anaemia. Consistent with SMC5/6's known functions in genome stability, cell lines derived from these patients showed elevated levels of DNA damage and replication stress. However, the analysis of chromosome spreads also revealed an

interesting chromosomal instability phenotype, characterised by segmented and dicentric chromosomes, alongside ‘mosaic variegated hyperploidy’ (MVH).

The presence of such unusual chromosome abnormalities in the SLF2 and SMC5 patient cells is unique and provides a potential distinctive diagnostic criterion for patients with this syndrome. This is particularly pertinent because the clinical phenotypes caused by genome instability and replication stress-linked conditions are often broad and overlapping (Zeman and Cimprich, 2014, Terabayashi and Hanada, 2018). Alongside whole exome sequencing, metaphase spread analysis could therefore provide a very specific test that could narrow down the candidate genes causally associated with this condition at early stages of diagnosis.

7.2.2 Clinical differences in patients associated with mutations in the SMC5/6-SLF2-SLF1-RAD18 pathway

The link between elevated levels of replication stress, microcephaly and reduced body size has a potentially simple explanation. The rapid expansion of pluripotent stem cells during embryogenesis, as well as neural progenitor cells during neurogenesis, induces elevated levels of replication stress which then require faithful replication-associated DNA repair pathways to maintain genome stability. Accordingly, patients with mutations in genes associated with replication or the replication stress response exhibit significant increases in genome instability. This can lead to elevated levels of cell death and reduced cellular proliferation *in utero* which can hinder growth and brain development at critical developmental timepoints (Murga et al., 2009, Alcantara and O'Driscoll, 2014, Reynolds et al., 2017). However, replication stress linked human diseases have also been associated with severe defects in other tissues. For instance, mutations in FA family proteins, which are required for the repair of DNA

inter-strand crosslinks, give rise to bone marrow failure, which can, in severe cases, develop into acute myeloid leukaemia (Zeman and Cimprich, 2014, Peake and Noguchi, 2022). Experimentation using mouse models has suggested that a loss of FA pathway components causes bone marrow failure by inducing stem cell defects in the early stages early of embryogenesis, when hematopoietic stem cells and progenitor cells amplify (Ceccaldi et al., 2012, Domenech et al., 2018). Together, this suggests that the replication stress response is vital *in utero* for the expansion of different rapidly dividing cell types. Considering that both embryonic and foetal cardiomyocytes, cardiac muscle cells, undergo rapid cell division (Zhao et al., 2020), it may be that the SMC5/6-SLF2-SLF1-RAD18 pathway is required to manage replication stress in developing heart tissues, leading to cardiac defects when either SLF2 or SMC5 is mutated. However, it's still not well understood why different genes involved in the replication stress response may be specifically required in certain tissues at specific embryological timepoints.

The characterisation of SLF2 and SMC5 patients contributes to a spectrum of three distinct diseases now associated with the dysfunction of the SMC5/6 complex. As previously discussed, mutations in the E3 SUMO-ligase NSE2 subunit are associated with 2 unrelated patients with severe primordial dwarfism, extreme insulin resistance and primary gonadal failure (Payne et al., 2014). Contrastingly, 9 individuals with NSE3 variants, from 4 unrelated families, have been reported to present with fatal pediatric pulmonary disease, often accompanied by immune deficiency (van der Crabben et al., 2016, Willemse et al., 2021). Thus, patient mutations within different subunits of SMC5/6 give rise to distinct clinical presentations. This is despite their similar cellular defects, including elevated levels of DNA damage markers and S-

phase defects, all of which align with SMC5/6's proposed roles in promoting proper DNA repair and replication. However, considering that microcephaly is a hallmark of replication stress diseases, it is curious that a reduction in head circumference was not reported in NSE2 and NSE3 patients.

The differences between these clinical presentations may be explained by the varying impacts of the patient mutations on the biochemical activities of SMC5/6, specifically the E3 SUMO-ligase activity of NSE2, which has been shown to be important for the DNA repair functions of the complex (Zhao and Blobel, 2005, Potts and Yu, 2005, Andrews et al., 2005, Kliszczak et al., 2012). Whilst the impact of the NSE3 and SLF2 patient mutations on this activity have yet to be explored, the two NSE2 variants studied by Payne *et al* were predicted to result in a loss-of-function, leading to a significant decrease in protein stability and, in the case of one variant, a complete loss of SMC5/6 auto-SUMOylation (Payne et al., 2014). Considering that the SUMO-ligase activity of Nse2 is believed to be dependent on its interaction with Smc5/6 (Bermudez-Lopez et al., 2015), where it docks onto the coiled-coil arm of Smc5 (Zhao and Blobel, 2005, Duan et al., 2009a), mutations in SMC5 could hypothetically disrupt the complex's SUMO functions, also. For instance, SMC5-P7's p.Arg372del variant presented with an amino acid deletion proximal to NSE2's binding site in yeast, which may prevent its interaction with SMC5. Contrastingly, the p.His990Asp mutation (SMC5-P8 and -P9) lies just upstream of the ATP-binding cassette (ABC) in SMC5, and could, therefore, be predicated to interfere with the binding of ATP. Whilst this may not impede NSE2's interaction with the SMC5/6 complex, research in yeast has suggested that ATPase-defective Smc5 is unable to auto-SUMOylate (Bermudez-Lopez et al., 2015). This suggests that at least some of Smc5/6's SUMO-ligase

functions are dependent on its ATPase activity and could be disrupted by the p.His990Asp variant in humans. Certainly, a mild disruption of NSE2 function in the hypomorphic SLF2 and SMC5 patient derived cell lines, compared to the putative loss-of-function mutations in the NSE2 patients, may provide a rationale for the contrasting disease presentations between the two syndromes.

Whilst SMC5/6's SUMO-ligase function is certainly important for faithful repair (Zhao and Blobel, 2005, Potts and Yu, 2005, Andrews et al., 2005, Kliszczak et al., 2012), NSE2's C-terminal RING domain, which encodes this activity, is dispensable for cellular survival (Bermudez-Lopez et al., 2015). Experiments in *Saccharomyces cerevisiae*, for instance, have shown that whilst the deletion of Nse2's SP-RING motif is not lethal, the loss of other Smc5/6 subunits are, suggesting that the complex has other essential functions that are independent of Nse2's SUMOylation activities (Zhao and Blobel, 2005). Thus, it is conceivable that the instability of the complex as a whole is associated with a different disease phenotype when compared with the loss of the SUMO-ligase. Considering that the identified p.Pro209Leu and p.Leu264Phe NSE3 variants were shown to disrupt the interactions between SMC5/6 subunits, this may explain the extreme fatal pulmonary disease observed in NSE3 patients. Furthermore, cell lines derived from patients with these variants showed that the protein expression of NSE3, SMC5 and SMC6 was significantly reduced, indicating that the missense mutations in NSE3 significantly destabilise the SMC5/6 complex (van der Crabben et al., 2016). Together with NSE1 and NSE4, NSE3 forms a bridge between the SMC5 and SMC6 subunits (Palecek et al., 2006, Hudson et al., 2011, Guerineau et al., 2012), a function that may explain why mutations in NSE3 impact the stability of multiple SMC5/6 complex components. In contrast, the patient mutations in NSE2

were shown to be associated with only a mild decrease of SMC5 and SMC6 expression (Payne et al., 2014), whilst the SLF2 and SMC5 variants failed to affect the protein levels of other complex components. Thus, it appears that the identified SMC5/6 patient mutations have varying impacts on the stability and overall function of the complex.

When considering the impact of the SMC5/6 associated patient mutations, it is also important to note that the precise biochemical activities and cellular functions of SMC5/6's individual subunits have not yet been fully elucidated, many of which may be interrelated. For instance, whilst SLF2 has been shown to promote the recruitment of SMC5/6 to chromatin at sites of DNA damage (Raschle et al., 2015), it is yet unknown whether this role is dependent on other SMC5/6 complex components which could also potentially disrupt this function if mutated. Certainly, multiple SMC5/6 subunits have been shown to bind DNA via several domains, likely conferring multivalent DNA binding properties (Roy et al., 2015, Zabrady et al., 2016, Alt et al., 2017, Yu et al., 2021). Furthermore, the loading of the complex onto damaged chromatin is intricately linked with SMC5/6's ATPase activity which is believed to drive a conformational change in the complex, allowing it to interact with DNA (Lammens et al., 2004, Kanno et al., 2015, Adamus et al., 2020). Interestingly, recent research in yeast has implicated Nse5/6, the functional paralogues of SLF1/2, in the regulation of the complex's ATPase activity (Taschner et al., 2021), suggesting a possible dynamic role for these factors in the interaction of the SMC5/6 complex and chromatin. Finally, the NSE1 subunit contains a RING domain which is suggestive of ubiquitin ligase activity, although evidence for this activity is mixed. Whilst NSE1 on its own supports only very weak ubiquitin ligase activity (Pebernard et al., 2008a,

Doyle et al., 2010, Kolesar et al., 2022), some *in vitro* experiments have suggested that this can be greatly stimulated by the interaction of the subunit with NSE3 and NSE4 (Doyle et al., 2010, Kolesar et al., 2022). Considering ubiquitin modifications play a significant role in the cellular movement of proteins (Haglund and Dikic, 2005), it's possible that this activity is important for the repair and replication functions of SMC5/6. In human cells, however, no *in vivo* ubiquitin targets for NSE1 have yet been identified, meaning the subunit's functions are largely unknown. Taken together, this demonstrates that SMC5/6 complex components have diverse but often overlapping functions, and suggests that it may, therefore, be difficult to make firm conclusions about the impact of the loss of specific SMC5/6 complex activities on disease presentations.

Loss of function mutations in RNF168, an E3 ubiquitin ligase required for the recruitment of the SMC5/6-SLF2-SLF1-RAD18 protein pathway to DNA damage (Raschle et al., 2015), have also been associated with human disease. RNF168 variants were first identified in a single patient exhibiting radiosensitivity, immune deficiency, dysmorphic features and learning difficulties, a clinical presentation defined as RIDDLE syndrome (Stewart et al., 2007, Stewart et al., 2009). 3 further patients with RNF168 mutations, from two unrelated families, have since been identified. Consistent with the first RIDDLE syndrome patient, these patients presented with radiosensitivity and immunodeficiency. However, all 3 individuals also exhibited telangiectasia and ataxia (Devgan et al., 2011, Pietrucha et al., 2017), with one patient presenting with microcephaly and pulmonary failure (Devgan et al., 2011). Therefore, whilst cellular radiosensitivity and immunodeficiency appear to be hallmarks of syndromes associated with RNF168 mutations, RNF168 patients also

present with numerous other varying clinical phenotypes. Some of these phenotypes are shared by patients with mutations in either SLF2, SMC5, NSE2 or NSE3, indicating that they may develop due to similar defects in the stability or recruitment of the SMC5/6-SLF2-SLF1-RAD18 pathway. It is yet unknown why RNF168 patients are so phenotypically variable, although it's possible that this could also be caused by the varying impacts of RNF168 mutations on SMC5/6 functions. The clinical discrepancies between patients with mutations in RNF168 and those with mutations in SMC5/6 complex components could be explained, at least in part, by RNF168's additional roles outside of SMC5/6 localisation. For instance, RNF168 also promotes the recruitment of 53BP1, a key DSBR factor, to sites of DNA damage (Stewart et al., 2009, Bohgaki et al., 2013). Interestingly, the fact that a cell line from a RIDDLE patient with mutations in RNF168 also exhibits MVH, to a very similar extent to SMC5 and SLF2 patients, despite the clear differences in clinical symptoms, may also suggest that MVH doesn't significantly contribute to the development of either syndrome.

However, it's also possible that the clinical phenotypes as yet reported in patients with mutations in SMC5/6 and RNF168 do not represent a consistent or complete clinical spectrum due to the small number of patients identified so far. Instead, it's likely that a more complete understanding of the clinical presentations of these syndromes, and any differences between them, will only be reached once many more additional patients are identified. Equally, it's important to note that many of the SLF2 and SMC5 patients investigated in this research are also young and may go on to develop symptoms common to other SMC5/6 associated diseases, such as lung disease or cardiac defects, later in life.

7.2.3 Identifying the endogenous lesion in SMC5 and SLF2 derived cell lines

SLF2 and SMC5 patient derived cell lines exhibited elevated levels of replication stress, DNA damage markers and SCEs, consistent with a role for the SMC5/6-SLF2-SLF1-RAD18 pathway in replication fork stability pathways and HR mediated repair. It's possible that these functions are linked, with SMC5/6 promoting HR-mediated replication fork restart at stalled or damaged forks. However, considering that levels of chromosome aberrations and SCEs didn't significantly increase with the addition of DNA damaging agents, it seems likely that the replicative dysfunction present in SLF2 and SMC5 patient cell lines affects just a subset of replication forks. Consistent with this, exposure to CX-5461, a drug which stabilises G4-quadruplexes, led to elevated levels of DNA damage and replicative dysfunction, alongside decreased levels of cell viability. This suggests that SMC5/6 promotes proper replication by unwinding these DNA secondary structures or promoting the restart of replication forks stalled at G4-quadruplexes, either of which would be a novel role for the complex. However, further work would be required to clarify whether these putative functions are specific to G4-quadruplexes, or whether SMC5/6 promotes proper replication at other loci associated with different types of DNA secondary structures, such as hairpins and i-Motifs.

Investigations in yeast have suggested that SMC5/6 interacts with and promotes the SUMOylation of $Sgs1^{BLM}$, as well as the other members of the Sgs1-Top3-Rmi1 'dissolvasome'- a complex which resolves recombination intermediates via dissolution to produce non-cross over products of HR (Bonner et al., 2016, Bermudez-Lopez et al., 2016). This could suggest that the SMC5/6-SLF2-SLF1-RAD18 pathway

promotes faithful HR in human cells by regulating BLM-mediated dissolution. Consistent with this, yeast Sgs1^{BLM} mutants, including SUMO mutants, and Nse2 SUMO-ligase dead mutants exhibit elevated levels of joint molecules following methyl methanesulfonate (MMS) treatment (Branzei et al., 2006, Bonner et al., 2016, Bermudez-Lopez et al., 2016), which induces elevated levels of replication stress (Tercero and Diffley, 2001). However, the precise mechanism by which SMC5/6 may regulate BLM is contested, with experiments in yeast and human cell lines, as well as mouse models, demonstrating that a depletion of both Sgs1/BLM and SMC5/6 complex components leads to either synthetic lethality or sickness, suggesting that their roles are not entirely epistatic (Raschle et al., 2015, Menolfi et al., 2015, Jacome et al., 2015). This may explain why mutations in either BLM or SMC5/6 give rise to overlapping yet distinct disease phenotypes, despite their similar roles in promoting HR. For instance, similarly to patients with mutations in SLF2 and SMC5, Bloom syndrome, caused by BLM loss-of-function mutations, is characterised by short stature and microcephaly. In contrast, however, Bloom patients also typically exhibit photosensitive facial skin lesions and an increased risk of cancer. (Arora et al., 2014, Cunniff et al., 2017). Furthermore, whilst mutations in BLM, SMC5 and SLF2 are all associated with increased SCEs, the hyperrecombination observed in BLM defective cell lines far exceeds the levels associated with SLF2 and SMC5 defects (Chaganti et al., 1974).

One possible explanation for the clinical and cellular discrepancies between patients with Bloom syndrome and individuals with mutations in SLF2 or SMC5 could be that SMC5/6 has an active role in regulating certain loci-specific BLM functions. If human SMC5/6 does indeed regulate BLM-mediated HR at damaged replication forks at

specific loci, this would be consistent with the hypothesis that SMC5/6 is required at just a subset of replication forks. Interestingly, research has implicated BLM in the unwinding G4-quadruplexes throughout the genome during replication. Current models predict that, without this function, the replication machinery can stall, leading to fork collapse and aberrant recombination (Mohaghegh et al., 2001, Bachrati et al., 2006, Drosopoulos et al., 2015, van Wietmarschen et al., 2018). Consistent with this, in both human and murine BLM knockout models, SCEs are enriched at G4 motifs in actively transcribed genes, where G4-quadruplex structures are believed to act as targets of transcriptional regulation (van Wietmarschen et al., 2018). It's not yet understood how the persistence of these replication impeding structures may contribute to the general replication stress observed in BLM knockout lines, but it's possible that they play a significant role, alongside the disruption of BLM-mediated dissolution at DSBs.

Whilst the Smc5/6-mediated SUMOylation of $Sgs1^{BLM}$ has yet only been implicated in the proper resolution of HJs and D-loops at damaged replication forks (Bonner et al., 2016, Bermudez-Lopez et al., 2016, Agashe et al., 2021), this research has been primarily undertaken in yeast, where G4-quadruplex motifs are predicted to be significantly less abundant than in humans (Wu et al., 2021). Therefore, it may be that the unwinding of G4-quadruplexes is far more important for faithful replication in mammalian cells, and represents a far more prominent function for BLM in humans. If SMC5/6 does have a role in promoting BLM-dependent resolution of G4-quadruplexes and HR intermediates, it may explain why SMC5 and SLF2 patient derived cell lines exhibit a sensitivity to CX-5461, a G4-stabilising drug. It would also implicate these structures as an endogenous lesion that the patient cell lines

specifically struggle to traverse during replication, rather than a general ‘difficult-to-replicate’ structure inducing SMC5/6-mediated replication restart (Figure 7.1).

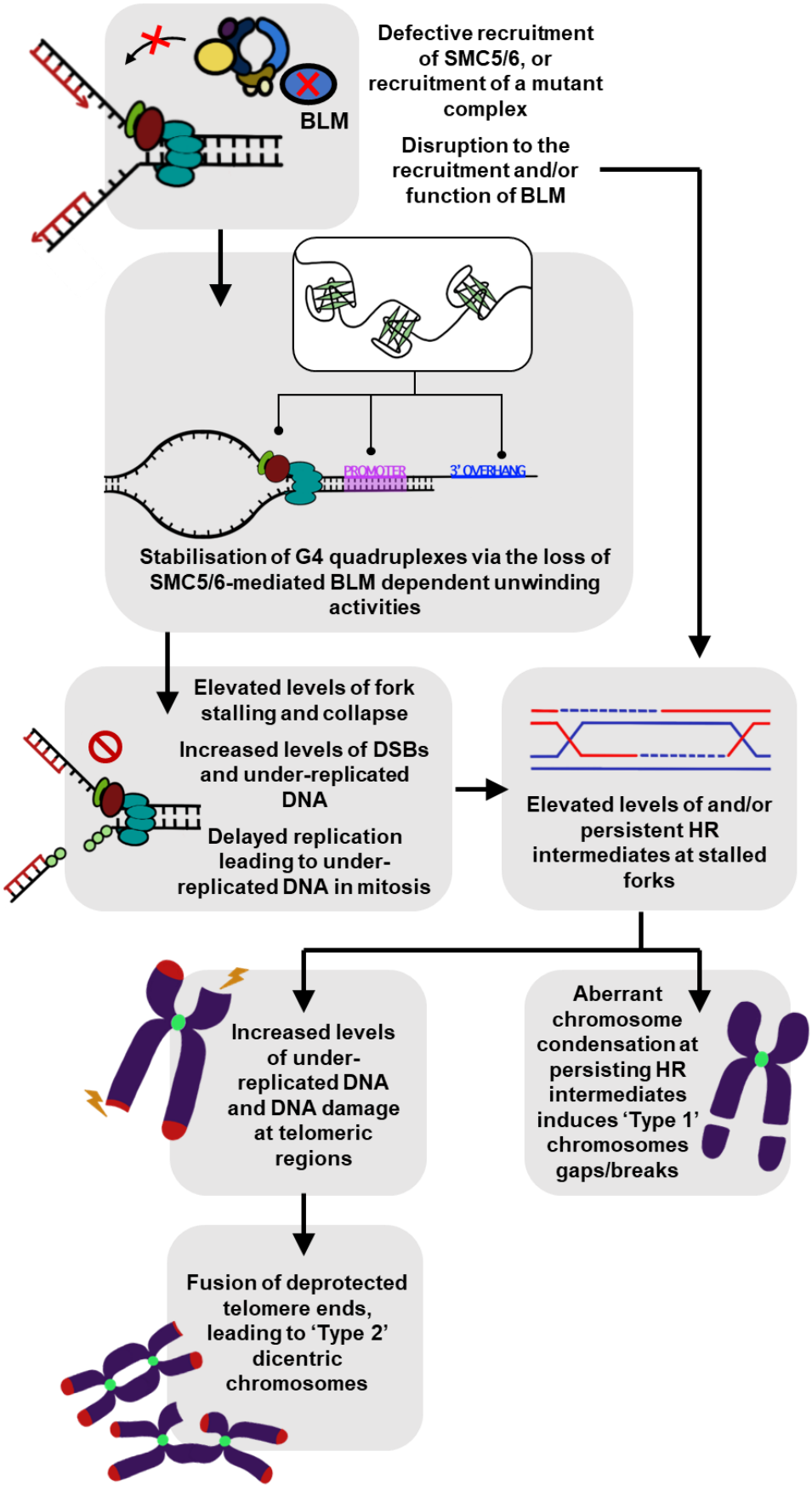
Previous studies have demonstrated a causative link between defective HR and distinct chromosome abnormalities, reminiscent of the segmented ‘Type 1’ chromosome gaps/breaks present in cells with mutations in SLF2 or SMC5 (Garner et al., 2013, Sarbajna et al., 2014, Chan et al., 2018). Again, this implicates SMC5/6 in the proper resolution of HR intermediates. Considering this, it is interesting that these chromosome abnormalities have not yet been associated with other human diseases, specifically those related to defective HR. However, it is important to note that, whilst this phenotype appeared in experimental cellular models depleted of HR resolvases and dissolution factors (Garner et al., 2013, Sarbajna et al., 2014, Chan et al., 2018), and in the context of hypomorphic mutations in human disease these aberrations may be far rarer. It may also be that the chromosome abnormalities are more easily observable in specific cell lines, suggesting that these structures can be overlooked in certain models due to cell-type differences. Interestingly, levels of segmented ‘Type 1’ chromosomes gaps/breaks were elevated in SLF2 and SMC5 patient LCLs after CX-5461 treatment. Whilst it’s possible that this is an indirect effect, occurring simply because CX-5461 induces elevated levels of fork stalling, this supports the hypothesis that SMC5/6 is specifically required to either avoid or resolve HR-intermediates at G4-quadruplexes (Figure 7.1). As previously discussed, G4-quadruplexes also preferentially form at telomeric regions, which are G-rich and present with a single-stranded 3’ overhang (Tang et al., 2007). Thus, the putative stabilisation of telomeric G4-quadruplexes in patient cell lines may also explain the

Figure 7.1 A model for the induction of 'Type 1' and 'Type 2' chromosomes abnormalities in SLF2 and SMC5 patient cell lines

In normal, healthy cells, the SMC5/6 complex is recruited to damaged replication forks in an SLF2 dependent manner, where it promotes replication fork stability and fork restart. However, the SMC5 and SLF2 patient mutations described in this study are predicted to impede the proper recruitment of the SMC5/6 complex and/or disrupt the complex's biochemical activities in patient cells. As studies in yeast report that Sgs1 (a homologue of mammalian BLM) interacts with SMC5/6 to promote Sgs1's activities and localisation to sites of damage, I hypothesise that SLF2 and SMC5 patient mutations also disrupt the proper function of BLM in patient cell lines.

Due to the role of BLM in resolving G4-quadruplexes, the loss of SMC5/6-mediated BLM activity in SLF2 and SMC5 patient cells could lead to the stabilisation and persistence of G4-quadruplexes within the replisome, at gene regulatory elements and at telomeres. This increase in G4-quadruplexes would obstruct DNA replication, leading to elevated levels of stalled forks, DNA damage, and regions of under-replicated DNA persisting through to mitosis. The increased levels of persisting G4-quadruplexes, replication fork stalling, and the disruption of SMC5/6-mediated BLM regulation could all contribute to the elevated levels of HR intermediates observed in SMC5 and SLF2 patient cells.

I propose that the 'Type 1' chromosome gaps and breaks result from defective condensation of chromatin surrounding these G4-quadruplex associated HR intermediates. Further to this, I hypothesis that the stabilisation/persistence of G4-quadruplex structures at telomeres in patient cell lines leads to replication stress, DNA damage and HR defects at chromosome ends. This would ultimately lead to telomeric degradation and the subsequent deprotection of telomeres, resulting in chromosome ends fusing to form 'Type 2' dicentric chromosomes.


BLM
Defective recruitment of SMC5/6, or recruitment of a mutant complex
Disruption to the recruitment and/or function of BLM
PROMOTER
3' OVERHANG
Stabilisation of G4 quadruplexes via the loss of SMC5/6-mediated BLM dependent unwinding activities
Elevated levels of fork stalling and collapse
Increased levels of DSBs and under-replicated DNA
Delayed replication leading to under-replicated DNA in mitosis
Elevated levels of and/or persistent HR intermediates at stalled forks
Increased levels of under-replicated DNA and DNA damage at telomeric regions
Aberrant chromosome condensation at persisting HR intermediates induces 'Type 1' chromosomes gaps/breaks
Fusion of deprotected telomere ends, leading to 'Type 2' dicentric chromosomes

elevated levels of telomeric dysfunction in these regions, subsequent fusions, and the generation of ‘Type 2’ dicentric chromosomes (Figure 7.1).

Whether SMC5/6’s putative functions in HR represent a primary function for the complex is yet unknown. However, since the loss of SMC5/6 subunits are lethal in *Saccharomyces cerevisiae*, whilst the depletion of HR factors are not, it may be that the complex has additional functions outside of HR which are important for cellular survival (Zhao and Blobel, 2005).

7.2.4 Does SMC5/6 have a specific role in mitosis?

As discussed in chapter 5, metaphase spread analysis revealed that SLF2 and SMC5 patient peripheral blood lymphocytes and LCLs exhibit chromosome gains in a subset of cells. Immunofluorescence analysis of these patient cell lines revealed a number of mitotic defects that could likely contribute to chromosome gains, including lagging chromosomes and multipolar spindles, linked to centrosome over-duplication in S/G2. Considering that the centrosome over-duplication and multipolar spindles were induced by exogenous replication stress, it may be that these mitotic issues arise due to SMC5/6-associated replication defects, such as S-phase arrest. Although further work would be needed to question if the lagging chromosomes observed in SLF2 and SMC5 patient lines are also exacerbated by replicative dysfunction, experiments in yeast have shown that treatment with MMS induces chromosome segregation defects in Smc5/6 mutants cells, including SUMO-ligase dead Nse2 mutants. (Bermudez-Lopez et al., 2010). This is consistent with previous work which has demonstrated that chromosome missegregation and UFBs are conferred by the loss of SMC5/6 prior to S-phase specifically (Venegas et al., 2020).

In addition to this, increased levels of railroad chromosomes and cohesion fatigue in both SLF2 and SMC5 patient peripheral blood and LCLs suggests that mutations in either factor induces cohesion defects which may also contribute to changes in chromosome numbers. These phenotypes are frequently observed in some cohesinopathies, including WABS and Roberts syndrome, multisystem disorders driven by mutations in the cohesin network (van der Lelij et al., 2009, van der Lelij et al., 2010a, van der Lelij et al., 2010b, van Schie et al., 2020). Whilst this could imply a regulatory role for SMC5/6 in cohesin recruitment or function, a loss of sister chromatid cohesion is also commonly observed in cancer cell lines lacking cohesin-related mutations, where the phenomenon is causally associated with increases in replication stress. Accordingly, a number of replication-stress inducing drugs, including hydroxyurea, cisplatin, and etoposide, have been shown to induce elevated levels of chromatid separation during mitosis (Kukolj et al., 2017, Benedict et al., 2020). Some studies have attributed this to replication-stress induced metaphase arrest in which chromosomes are exposed to the pulling forces of the mitotic spindle for a prolonged period, risking cohesion fatigue (Kukolj et al., 2017, Masamsetti et al., 2019). However, one recent study using RPE1 cell lines with oncogenic mutations argued against this. The authors used a combination of DNA fibre and comet assays to show that cohesin removal was required for successful replication coupled DSBR during periods of elevated replication stress, suggesting that, under these conditions, the increased loss of cohesin at stalled or damaged replication forks may lead to a premature loss of sister chromatid cohesion (Benedict et al., 2020). Thus, there may be a distinction between cohesion defects induced by mutations in or related to cohesin, and those caused by increases in replication stress, with further work

required to distinguish between either driver in the context of the mutations in SLF2 and SMC5.

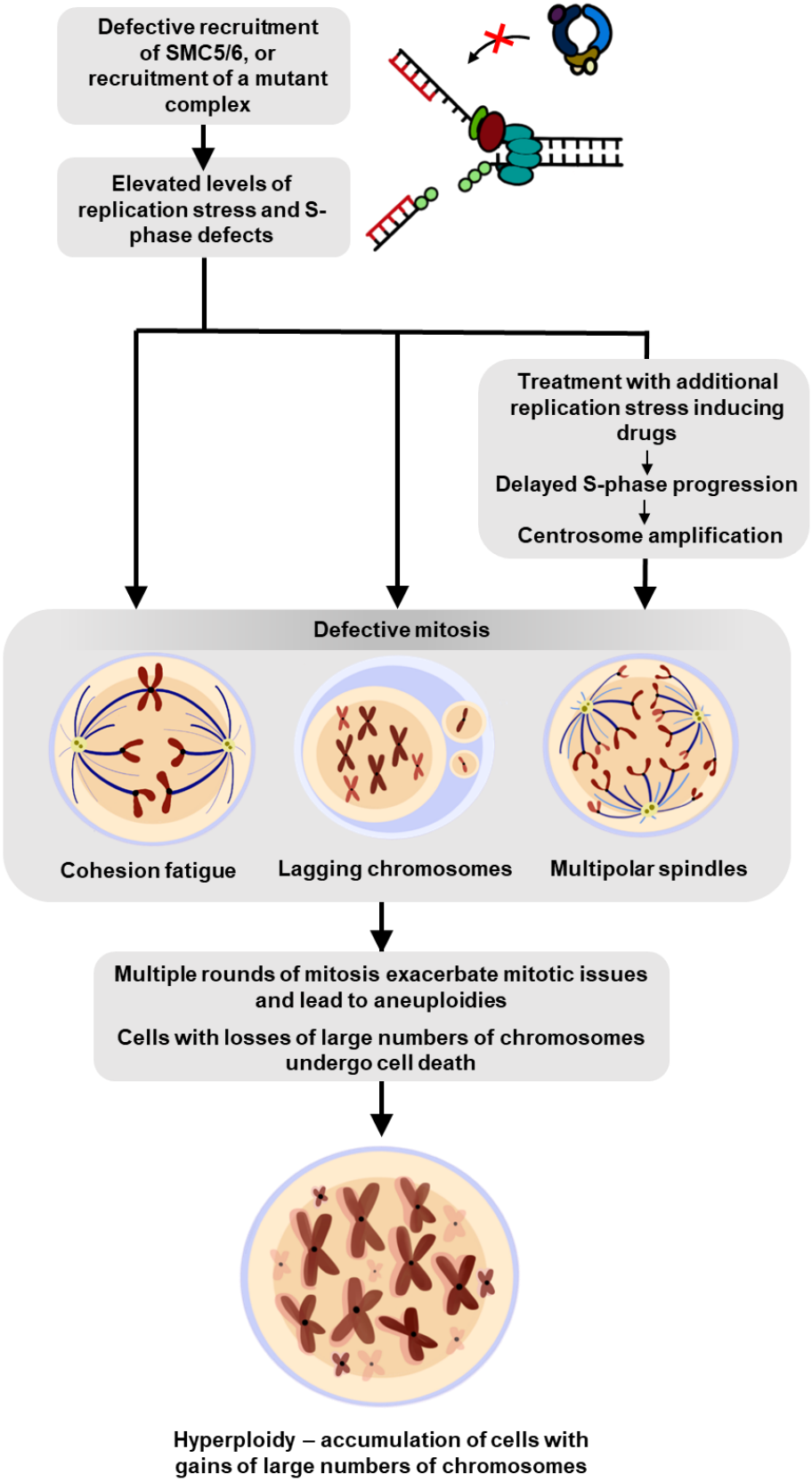
Acknowledging this, it is possible that the MVH observed in the SLF2 and SMC5 patient cell lines is the result of several sources of mitotic dysfunction driven by increases in replication stress (Figure 7.2). However, this raises questions as to why other genome instability and replication-stress linked disorders do not present with significant levels hyperploidy, despite similar defects in repair and replication. Interestingly, MVH has yet only been reported in patients with mutations in SLF2, SMC5 and RNF168, rather than other SMC5/6 complex components. This could suggest that these factors are required for a novel function, whether that be replication-linked or mitotic, which is important to maintain chromosome stability. Considering that replication-stress induced multipolar spindles are common to both syndromes, it could be that spindle abnormalities play a critical role in driving these chromosome gains (Figure 7.2). However, further analysis will be required to determine whether any other SMC5 and SLF2 associated mitotic problems, such as cohesion defects, are also associated with the RIDDLE syndrome LCL, or other cell lines exhibiting defects in RNF168. Furthermore, analysis of the SLF2 and SMC5 patient derived blood samples and cell lines has indicated that the proportion of hyperploid cells can vary significantly between individual patients and samples. Therefore, it may be that MVH is associated with mutations in other SMC5/6 complex components but that this has yet to be reported due to the small patient cohorts and the scarcity of hyperploid cells.

Figure 7.2. A model for the generation of hyperploid cells in SLF2 and SMC5 patient cell lines

Patient mutations in SLF2 and SMC5 lead to the defective recruitment or activity of the SMC5/6 complex at damaged replication forks, resulting in elevated levels of replication stress and other S-phase associated defects. This results in several defects during mitosis, such as lagging chromosomes, which mis-segregate and form micronuclei, and a premature loss of cohesin, leading to cohesion fatigue. A mechanism for SMC5/6-associated cohesion fatigue has yet to be elucidated but may be caused by prolonged metaphase arrest due to elevated levels of replication-stress, or the removal of cohesin at replication forks during HR-mediated restart.

I propose that delayed S-phase progression in SLF2 and SMC5 patient cell lines treated with replication stress inducing drugs causes centrosome amplification in S-phase, ultimately leading to mitotic cells with multipolar spindles. This, together with lagging chromosomes and cohesin fatigue, induces considerable chromosome mis-segregation and ultimately large alterations in chromosome number.

I hypothesise that over successive rounds of mitosis, cells with gains in chromosome numbers may accumulate, whilst cells that experience significant chromosome losses undergo cell death, generating a significant proportion of cells with hyperploidy. Whilst this study did not associate the disruption of SLF2 and SMC5 with the gross mis-localisation of mitotic factors, it is possible be that SMC5/6 has undiscovered specific miotic functions which also contribute to such significant chromosome gains.


Defective recruitment of SMC5/6, or recruitment of a mutant complex
Elevated levels of replication stress and S-phase defects
Treatment with additional replication stress inducing drugs
Delayed S-phase progression
Centrosome amplification
Defective mitosis
Cohesion fatigue
Lagging chromosomes
Multipolar spindles
Multiple rounds of mitosis exacerbate mitotic issues and lead to aneuploidies
Cells with losses of large numbers of chromosomes undergo cell death
Hyperploidy – accumulation of cells with gains of large numbers of chromosomes

7.2.5 Considering a potential link between SMC5/6 and cancer

Replication stress and genome stability are both considered major drivers of cancer (Negrini et al., 2010, Macheret and Halazonetis, 2015). However, the relationship between replication stress and tumorigenesis is not straightforward. Unlike other genome instability disorders, replication stress-linked conditions are not typically characterised by the development of cancers (Zeman and Cimprich, 2014, Terabayashi and Hanada, 2018). Instead, as previously discussed, these disorders frequently present with neurodevelopmental issues and growth defects, suggesting that elevated levels of replication stress drive disease progression via increased levels of cell death during development (Murga et al., 2009, Alcantara and O'Driscoll, 2014, Reynolds et al., 2017). However, seeing that replication stress is a key driver of DNA damage, mitotic issues, and general genome instability, it is curious that these conditions do not present with cancers more frequently (Zeman and Cimprich, 2014, Gelot et al., 2015).

Considering that aneuploidy is a hallmark of most tumours (Taylor et al., 2018), it is also interesting that cancer is not yet associated with mutations in SLF2 or SMC5, particularly as the extreme chromosome gains associated with SLF2 and SMC5 dysfunction extend far beyond that seen in typical replication stress disorders. However, as with replication stress, our understanding of the role of aneuploidy in cancer development is complicated. Whilst aneuploidy is believed to be generally deleterious, with studies demonstrating that gains of single chromosomes impair cell proliferation (Torres et al., 2007, Williams et al., 2008a, Stingele et al., 2012), some studies have noted that, in certain conditions, copy number alterations of specific genes can drive increases in cell survival and tumorigenesis (Ben-David et al., 2014).

This may explain why MVA syndrome, which confers random changes in chromosome numbers, is associated with an increased risk of some cancers (Kajii et al., 2001, Yost et al., 2017). Whilst much remains unknown regarding the production of hyperploid cells in the SLF2 and SMC5 patients, such as the frequency with which they are generated, how long they persist, and whether the affected cells accumulate chromosomes over several rounds of cellular division, interesting questions are raised regarding how MVH may contribute to the disease phenotype observed in the SLF2 and SMC5 patients and whether these patients, like those with MVA, are likely to develop cancer in the future.

7.2.6 Future work

The phenotypic analysis of the SLF2 and SMC5 patient cell lines presented in this thesis links the SMC5/6-SLF2-SLF1-RAD18 pathway with several known cellular functions, including replication and HR. Interestingly, however, my research also suggests a number of possible novel functions for SMC5/6 in mitosis and the regulation of G4-quadruplexes. Further work will now be required to mechanistically explore these putative roles and confirm their dysfunction in the patient cell lines.

As discussed, the segmented ‘Type 1’ chromosome gap/break abnormalities observed in SLF2 and SMC5 patient cell lines are reminiscent of aberrations associated with persistent HR intermediates (Garner et al., 2013, Sarbajna et al., 2014), suggesting that the SMC5/6-SLF2-SLF1-RAD18 pathway has important roles in resolving these structures. However, my work has yet to conclusively verify that these aberrations are linked to defective HR. In previous studies, the exogenous expression of RusA, a bacterial HJ resolvase, and the subsequent rescue of

segmented chromosomes, has been used to confirm the presence of persisting HJ intermediates (Garner et al., 2013, Chan et al., 2018). Thus, future work could explore whether the SLF2 and SMC5 patient mutations induce HR defects, and whether these defects drive the formation of the segmented ‘Type 1’ chromosome abnormalities, by questioning if the expression of RusA is sufficient to reduce elevated levels of these structures. Similarly, experiments could then question if RusA expression rescues other cellular defects associated with mutations in SLF2 and SMC5, such as the elevated levels of fork stalling. These observations would causatively link these phenotypes with HR dysfunction. It may also be interesting to ask if the resolution of persistent HR intermediates, which may otherwise contribute to chromosome missegregation, leads to a reduction in the number of cells exhibiting chromosome gains. To question this, the chromosome number of patient derived cells transfected with RusA could be analysed after extended periods of cell culture. As has been previously discussed, it's possible that SMC5/6’s putative HR function is conferred via the NSE2-mediated SUMOylation of BLM and the BTR complex. To explore this further in the context of this disease model, it would be interesting to determine if the SUMOylation activity of SMC5/6 is compromised in patient cells, and if BLM’s SUMO modifications are disrupted. Considering SMC5/6 may regulate BLM during HR-mediated fork restart at damaged replication forks, proximity ligation assay (PLA) could also be used to question if BLM’s localisation to replication forks is disrupted in the SLF2 and SMC5 patient cell lines.

The presence of MVH in the patient cell lines represents a unique cellular phenotype associated with mutations in SLF2 and SMC5. However, it also raises critical questions regarding whether these chromosome gains are the result of elevated

replication stress, or instead a novel function for the SMC5/6 complex in mitosis. Whilst my research failed to observe any gross mislocalisation of mitotic factors previously associated with SMC5/6, future research should work to establish if the dysfunction in SLF2 or SMC5 instead disrupts the enzymatic functions of these factors or deregulates any of their downstream substrates. As discussed in this thesis, the generation of hyperploid cells could theoretically arise due to numerous cellular defects, including a loss of sister chromatid cohesion, multipolar spindles and lagging chromosomes. To understand more regarding the role of these putative dysfunctions, it would be interesting to use live cell imaging to monitor the progression of patient derived cells through mitosis. This would give a detailed insight into the aberrant mitotic phenotypes associated with defects in SLF2 and SMC5 and an understanding of how these abnormalities may lead to chromosome gains. Alternatively, future experiments could focus on exploring the possible relationship between SMC5/6's replicative functions and hyperploidy. By treating SLF2 and SMC5 patient cells with replication-stress inducing drugs for elongated periods of time, experiments could question if elevated levels of replication-stress are associated with an increased number of hyperploid cells, suggesting a causative link between the two. Considering that the levels of SCEs and chromosome aberrations did not significantly increase in response to replication stress inducing drugs it may be that the replicative functions of SMC5/6 are specific to a subset of replication forks. Thus, exacerbating the possible link between replication stress and chromosome gains may be easier if endogenous lesions causing the replicative dysfunction can be directly targeted.

My research proposes that SMC5/6 may have specific functions related to G4-quadruplexes, with treatment with CX-5461, a G4-stabilising agent, inducing

significantly elevated levels of chromosomal damage and fork stalling, alongside decreased cellular viability. Future work should use G4-quadruplex antibodies to confirm both the endogenous persistence of these structures, as well as their CX-5461-mediated stabilisation, via immunofluorescence. As discussed in chapter 6, it would be important to verify that the CX-5461-induced cellular defects in the patient cells are caused by the G4-stabilising activities of the drug, rather than its functions as a Top2 poison or an RNA polymerase I inhibitor, by testing the sensitivity of the cells to other G4-quadruplex stabilising agents (e.g. pyridostatin), Top2 poisons (e.g. etoposide and doxorubicin) and RNA polymerase I inhibitors (e.g. BMH-21). The link between G4-quadruplexes and HR defects in the patient cell lines could then be explored further by determining if treatment with G4-quadruplex stabilising drugs induces significantly elevated levels of SCEs. Since G4-quadruplexes are believed to form frequently at telomeres (Bryan, 2020), it would also be pertinent to ask if G4 structures represent the endogenous lesions responsible for the 'Type 2' dicentric chromosomes observed in the SMC5 and SLF2 patient cell lines. This could hypothetically occur if these 'difficult-to-replicate' lesions induce elevated levels of DNA damage at chromosome ends, leading to their deprotection. An interesting way to query the relationship between SMC5/6 and these loci would be to use BrdU-seq to investigate sites of G2/M synthesis and question if these regions of incomplete S-phase replication are enriched at telomeres or G4-structures (Wang and Saponaro, 2021).

My research contributes to a broad clinical spectrum of SMC5/6 associated syndromes, linked to defects in different subunits. In order to mechanistically understand the cellular phenotypes observed in SLF2 and SMC5 patient cell lines, as

well as their distinction from other SMC5/6 related conditions, it will be important to investigate how the patient variants impact the structure and biochemical activities of the complex. A first step to understanding this could be to use western blotting to determine the protein expression levels of NSE1-4 in the patient cell lines. This will reveal if mutations in SLF2 or SMC5 disrupt the stability, and therefore likely impacting their activity, of the other complex subunits. By pulling down SMC5 or SMC6 and probing for NSE subunits, immunoprecipitation studies could also be used to assess precisely if/how the patient mutations impact the interaction of certain complex components and the assembly of SMC5/6. Another way of determining this could be to use BioID or mass spectrometry experiments to compare SMC5 or SLF2 interactors in WT and patient cell lines. Any differences could also help to identify potential downstream pathways responsible for the SMC5/6-associated cellular dysfunction.

In order to specifically assess the impact of the patient variants on the SUMOylation activity of NSE2, which docks onto the arm of SMC5 and could hypothetically be disrupted by SMC5 mutations (Duan et al., 2009a), the auto-SUMOylation activity of NSE2 should be investigated. Tagged SUMO could be expressed in the patient fibroblasts, with immunoprecipitation experiments then used to pull down SUMO and probe for NSE2 (and vice versa). This would determine whether the interaction between SUMO and NSE2 is decreased in patient cell lines in comparison to WT controls. The SUMOylation of various other factors, including SMC5, SMC6 and BLM, or other SMC5/6 interacting proteins identified by the previously discussed BioID and mass spectrometry studies, could also be assessed to question if this is impacted by

the patient variants. These interactions could then be confirmed *in vitro* using purified constructs.

As discussed above, it will be particularly important to investigate the possible SMC5/6-mediated SUMOylation of BLM, which I hypothesise could be disrupted in the patient cells lines and thereby responsible for the HR-associated defects. If BLM was shown to be SUMOylated by SMC5/6, the site(s) of BLM's SUMOylation could be determined, either by mass spectrometry or the mutation of predicted SUMOylation sites. If identified, mutant BLM which cannot be SUMOylated by SMC5/6 could be exogenously expressed in WT cell lines to question if this disruption is sufficient to induce similar phenotypes to the ones seen in the patient cell lines. It may then also be interesting to ascertain when in the cell cycle BLM, or any other identified factor, may be SUMOylated by SMC5/6, and confirm whether this is linked specifically to S-phase functions. By arresting patient cell lines at different points in the cell cycle, pulling down BLM and probing for exogenously expressed and tagged SUMO, in untreated and SMC5/6 siRNA treated experiments, it would be possible to identify when SMC5/6-mediated BLM SUMOylation occurs. The same experiments could also be conducted in conjunction with different drug treatments to question if this SUMOylation occurs in response to replication stress generally, or if it is exacerbated by specific drugs, such as G4-stabilising agents.

Since SMC5 and SMC6 have important ATPase activities which modulate the interaction of the complex with DNA, the ATP binding and hydrolysis activity of purified preparations of WT and patient variant SMC5/6 complexes could be compared. Previous studies have assayed SMC5/6's ATPase activities by measuring the proportional oxidation of NADH (Taschner et al., 2021). By utilising these assays to

monitor levels of ATP turnover at saturating ATP concentrations, these investigations could also determine whether the putative decreases in ATPase activity associated with patient mutations are likely to be impacted by defective ATP hydrolysis, or instead by decreased ATP binding by SMC5/6. These purified complexes could also be incubated with DNA constructs and analysed using an electrophoretic mobility shift assay to understand the extent to which defective SMC5/6 ATPase activity may disrupt the complex binding to DNA. Using this method, it would be possible to question how the complex interacts with DNA, and whether this interaction is enhanced at, or specific to, certain structures, for instance double strand breaks, replication forks, regressed forks or G4-quadruplexes, and whether these interactions were impacted differently by the patient variants and differences in ATPase activity.

The impact of these mutations on the development of the patients' clinical and cellular phenotypes could then be confirmed using animal models. A conditional SMC5 knockout mouse model has been generated already, exhibiting reduced brain size and sensorimotor issues (Atkins et al., 2020). In addition to this, collaborators of my work (Dr Erica Davis) have recently generated SLF2 and SMC5 knockout zebrafish which phenocopy microcephaly and reduced body size. SMC5 knockout zebrafish could be rescued with WT SMC5, but not the patient variants, thus confirming the pathogenicity of these mutations (Grange et al., 2022). It may now be interesting to question if these phenotypes can be rescued by the concomitant deletion of p53, or other factors which are associated with SMC5/6. Since zebrafish are considered to be excellent models of heart development (Francoeur and Sen, 2021), it's possible that future work could use these models to explore the association of cardiac defects with SMC5/6 dysfunction, as well as replication stress and MVH. Neuronal zebrafish

cells could be isolated to monitor the presentation of these abnormalities in this specific cell type, also. To question this even further, it may be interesting to develop brain organoid models. Induced pluripotent stem cells could be generated from patient and WT fibroblasts, then differentiated into neurons and grown into organoids, with the size of these organoids and the cellular defects of the neurons compared between the different genetic backgrounds.

7.2.7 Summary

The research presented in this thesis characterises a new human syndrome related to mutations in SLF2 and SMC5. The cellular phenotypes associated with these individuals provide further evidence for the role of SMC5/6 in DNA replication and repair and suggests functions for the complex in telomere maintenance and replication through G4-quadruplex structures. However, major questions remain regarding the precise role of the SMC5/6 complex in replication and mitosis and the cause of the MVH in the patient cell lines. By expanding the range of clinical phenotypes associated with SMC5/6 dysfunction, this work also questions why mutations in different SMC5/6 complex components give rise to such contrasting clinical presentations. These questions will hopefully be answered with more research, the discovery of more patients, the identification of new syndromes associated with other SMC5/6 subunits and the further use of animal models.

References

ABE, S., NAGASAKA, K., HIRAYAMA, Y., KOZUKA-HATA, H., OYAMA, M., AOYAGI, Y., OBUSE, C. & HIROTA, T. 2011. The initial phase of chromosome condensation requires Cdk1-mediated phosphorylation of the CAP-D3 subunit of condensin II. *Genes Dev,* 25**,** 863-74.

ABE, T., OOKA, M., KAWASUMI, R., MIYATA, K., TAKATA, M., HIROTA, K. & BRANZEI, D. 2018. Warsaw breakage syndrome DDX11 helicase acts jointly with RAD17 in the repair of bulky lesions and replication through abasic sites. *Proc Natl Acad Sci U S A,* 115**,** 8412-8417.

ADAMUS, M., LELKES, E., POTESIL, D., GANJI, S. R., KOLESAR, P., ZABRADY, K., ZDRAHAL, Z. & PALECEK, J. J. 2020. Molecular Insights into the Architecture of the Human SMC5/6 Complex. *J Mol Biol,* 432**,** 3820-3837.

AGASHE, S., JOSEPH, C. R., REYES, T. A. C., MENOLFI, D., GIANNATTASIO, M., WAIZENEGGER, A., SZAKAL, B. & BRANZEI, D. 2021. Smc5/6 functions with Sgs1-Top3-Rmi1 to complete chromosome replication at natural pause sites. *Nat Commun,* 12**,** 2111.

AGUILERA, A. & GOMEZ-GONZALEZ, B. 2008. Genome instability: a mechanistic view of its causes and consequences. *Nat Rev Genet,* 9**,** 204-17.

AL-MINAWI, A. Z., SALEH-GOHARI, N. & HELLEDAY, T. 2008. The ERCC1/XPF endonuclease is required for efficient single-strand annealing and gene conversion in mammalian cells. *Nucleic Acids Res,* 36**,** 1-9.

ALCANTARA, D. & O'DRISCOLL, M. 2014. Congenital microcephaly. *Am J Med Genet C Semin Med Genet,* 166C**,** 124-39.

ALEXANDER, J. L., BARRASA, M. I. & ORR-WEAVER, T. L. 2015. Replication fork progression during re-replication requires the DNA damage checkpoint and double-strand break repair. *Curr Biol,* 25**,** 1654-60.

ALKAN, F., WENZEL, A., ANTHON, C., HAVGAARD, J. H. & GORODKIN, J. 2018. CRISPR-Cas9 off-targeting assessment with nucleic acid duplex energy parameters. *Genome Biol,* 19**,** 177.

ALLISON, D. F. & WANG, G. G. 2019. R-loops: formation, function, and relevance to cell stress. *Cell Stress,* 3**,** 38-46.

ALMEDAWAR, S., COLOMINA, N., BERMUDEZ-LOPEZ, M., POCINO-MERINO, I. & TORRES-ROSELL, J. 2012. A SUMO-dependent step during establishment of sister chromatid cohesion. *Curr Biol,* 22**,** 1576-81.

ALOMER, R. M., DA SILVA, E. M. L., CHEN, J., PIEKARZ, K. M., MCDONALD, K., SANSAM, C. G., SANSAM, C. L. & RANKIN, S. 2017. Esco1 and Esco2 regulate distinct cohesin functions during cell cycle progression. *Proc Natl Acad Sci U S A,* 114**,** 9906-9911.

ALT, A., DANG, H. Q., WELLS, O. S., POLO, L. M., SMITH, M. A., MCGREGOR, G. A., WELTE, T., LEHMANN, A. R., PEARL, L. H., MURRAY, J. M. & OLIVER, A. W. 2017. Specialized interfaces of Smc5/6 control hinge stability and DNA association. *Nat Commun,* 8**,** 14011.

AMANGYELD, T., SHIN, Y. K., LEE, M., KWON, B. & SEO, Y. S. 2014. Human MUS81-EME2 can cleave a variety of DNA structures including intact Holliday junction and nicked duplex. *Nucleic Acids Res,* 42**,** 5846-62.

AMPATZIDOU, E., IRMISCH, A., O'CONNELL, M. J. & MURRAY, J. M. 2006. Smc5/6 is required for repair at collapsed replication forks. *Mol Cell Biol,* 26**,** 9387-401.

ANAND, R., RANJHA, L., CANNAVO, E. & CEJKA, P. 2016. Phosphorylated CtIP Functions as a Co-factor of the MRE11-RAD50-NBS1 Endonuclease in DNA End Resection. *Mol Cell,* 64**,** 940-950.

ANDREWS, E. A., PALECEK, J., SERGEANT, J., TAYLOR, E., LEHMANN, A. R. & WATTS, F. Z. 2005. Nse2, a component of the Smc5-6 complex, is a SUMO ligase required for the response to DNA damage. *Mol Cell Biol,* 25**,** 185-96.

ANSARI, M., POKE, G., FERRY, Q., WILLIAMSON, K., ALDRIDGE, R., MEYNERT, A. M., BENGANI, H., CHAN, C. Y., KAYSERILI, H., AVCI, Ş., HENNEKAM, R. C. M., LAMPE, A. K., REDEKER, E., HOMFRAY, T., ROSS, A., FALKENBERG SMELAND, M., MANSOUR, S., PARKER, M. J., COOK, J. A., SPLITT, M., FISHER, R. B., FRYER, A., MAGEE, A. C., WILKIE, A., BARNICOAT, A., BRADY, A. F., COOPER, N. S., MERCER, C., DESHPANDE, C., BENNETT, C. P., PILZ, D. T., RUDDY, D., CILLIERS, D., JOHNSON, D. S., JOSIFOVA, D., ROSSER, E., THOMPSON, E. M., WAKELING, E., KINNING, E., STEWART, F., FLINTER, F., GIRISHA, K. M., COX, H., FIRTH, H. V., KINGSTON, H., WEE, J. S., HURST, J. A., CLAYTON-SMITH, J., TOLMIE, J., VOGT, J., TATTON–BROWN, K., CHANDLER, K., PRESCOTT, K., WILSON, L., BEHNAM, M., MCENTAGART, M., DAVIDSON, R., LYNCH, S.-A., SISODIYA, S., MEHTA, S. G., MCKEE, S. A., MOHAMMED, S., HOLDEN, S., PARK, S.-M., HOLDER, S. E., HARRISON, V., MCCONNELL, V., LAM, W. K., GREEN, A. J., DONNAI, D., BITNER-GLINDZICZ, M., DONNELLY, D. E., NELLÅKER, C., TAYLOR, M. S. & FITZPATRICK, D. R. 2014. Genetic heterogeneity in Cornelia de Lange syndrome (CdLS) and CdLS-like phenotypes with observed and predicted levels of mosaicism. *Journal of Medical Genetics,* 51**,** 659.

ANTONIN, W. & NEUMANN, H. 2016. Chromosome condensation and decondensation during mitosis. *Current Opinion in Cell Biology,* 40**,** 15-22.

APARICIO, T., GUILLOU, E., COLOMA, J., MONTOYA, G. & MENDEZ, J. 2009. The human GINS complex associates with Cdc45 and MCM and is essential for DNA replication. *Nucleic Acids Res,* 37**,** 2087-95.

ARAGON, L. 2018. The Smc5/6 Complex: New and Old Functions of the Enigmatic Long-Distance Relative. *Annu Rev Genet,* 52**,** 89-107.

ARIA, V. & YEELES, J. T. P. 2018. Mechanism of Bidirectional Leading-Strand Synthesis Establishment at Eukaryotic DNA Replication Origins. *Mol Cell*.

ARIAS, E. E. & WALTER, J. C. 2007. Strength in numbers: preventing rereplication via multiple mechanisms in eukaryotic cells. *Genes Dev,* 21**,** 497-518.

ARORA, H., CHACON, A. H., CHOUDHARY, S., MCLEOD, M. P., MESHKOV, L., NOURI, K. & IZAKOVIC, J. 2014. Bloom syndrome. *Int J Dermatol,* 53**,** 798-802.

ARUMUGAM, P., GRUBER, S., TANAKA, K., HAERING, C. H., MECHTLER, K. & NASMYTH, K. 2003. ATP Hydrolysis Is Required for Cohesin's Association with Chromosomes. *Current Biology,* 13**,** 1941-1953.

ATKINS, A., XU, M. J., LI, M., ROGERS, N. P., PRYZHKOVA, M. V. & JORDAN, P. W. 2020. SMC5/6 is required for replication fork stability and faithful chromosome segregation during neurogenesis. *Elife,* 9.

AYYAGARI, R., GOMES, X. V., GORDENIN, D. A. & BURGERS, P. M. 2003. Okazaki fragment maturation in yeast. I. Distribution of functions between FEN1 AND DNA2. *J Biol Chem,* 278**,** 1618-25.

AZVOLINSKY, A., GIRESI, P. G., LIEB, J. D. & ZAKIAN, V. A. 2009. Highly transcribed RNA polymerase II genes are impediments to replication fork progression in Saccharomyces cerevisiae. *Mol Cell,* 34**,** 722-34.

BACHANT, J., ALCASABAS, A., BLAT, Y., KLECKNER, N. & ELLEDGE, S. J. 2002. The SUMO-1 isopeptidase Smt4 is linked to centromeric cohesion through SUMO-1 modification of DNA topoisomerase II. *Mol Cell,* 9**,** 1169-82.

BACHRATI, C. Z., BORTS, R. H. & HICKSON, I. D. 2006. Mobile D-loops are a preferred substrate for the Bloom's syndrome helicase. *Nucleic Acids Res,* 34**,** 2269-79.

BAE, S. H. & SEO, Y. S. 2000. Characterization of the enzymatic properties of the yeast dna2 Helicase/endonuclease suggests a new model for Okazaki fragment processing. *J Biol Chem,* 275**,** 38022-31.

BAI, L., YUAN, Z., SUN, J., GEORGESCU, R., O'DONNELL, M. E. & LI, H. 2017. Architecture of the Saccharomyces cerevisiae Replisome. *Adv Exp Med Biol,* 1042**,** 207-228.

BAIRD, C. L., HARKINS, T. T., MORRIS, S. K. & LINDSLEY, J. E. 1999. Topoisomerase II drives DNA transport by hydrolyzing one ATP. *Proc Natl Acad Sci U S A,* 96**,** 13685-90.
BAKHOUM, S. F., GENOVESE, G. & COMPTON, D. A. 2009. Deviant Kinetochore Microtubule Dynamics Underlie Chromosomal Instability. *Current Biology,* 19**,** 1937-1942.
BALK, B., MAICHER, A., DEES, M., KLERMUND, J., LUKE-GLASER, S., BENDER, K. & LUKE, B. 2013. Telomeric RNA-DNA hybrids affect telomere-length dynamics and senescence. *Nat Struct Mol Biol,* 20**,** 1199-205.
BARBERO, J. L. 2013. Genetic basis of cohesinopathies. *Appl Clin Genet,* 6**,** 15-23.
BARIS, Y., TAYLOR, M. R. G., ARIA, V. & YEELES, J. T. P. 2022. Fast and efficient DNA replication with purified human proteins. *Nature,* 606**,** 204-210.
BARNES, J. L., ZUBAIR, M., JOHN, K., POIRIER, M. C. & MARTIN, F. L. 2018. Carcinogens and DNA damage. *Biochem Soc Trans,* 46**,** 1213-1224.
BARR, F. A. & GRUNEBERG, U. 2007. Cytokinesis: Placing and Making the Final Cut. *Cell,* 131**,** 847-860.
BATES, A. D. & MAXWELL, A. 2007. Energy coupling in type II topoisomerases: why do they hydrolyze ATP? *Biochemistry,* 46**,** 7929-41.
BAXLEY, R. M., LEUNG, W., SCHMIT, M. M., MATSON, J. P., YIN, L., ORAM, M. K., WANG, L., TAYLOR, J., HEDBERG, J., ROGERS, C. B., HARVEY, A. J., BASU, D., TAYLOR, J. C., PAGNAMENTA, A. T., DREAU, H., CRAFT, J., ORMONDROYD, E., WATKINS, H., HENDRICKSON, E. A., MACE, E. M., ORANGE, J. S., AIHARA, H., STEWART, G. S., BLAIR, E., COOK, J. G. & BIELINSKY, A. K. 2021. Bi-allelic MCM10 variants associated with immune dysfunction and cardiomyopathy cause telomere shortening. *Nat Commun,* 12**,** 1626.
BAXTER, J. & DIFFLEY, J. F. X. 2008. Topoisomerase II Inactivation Prevents the Completion of DNA Replication in Budding Yeast. *Molecular Cell,* 30**,** 790-802.
BEN-DAVID, U., ARAD, G., WEISSBEIN, U., MANDEFRO, B., MAIMON, A., GOLAN-LEV, T., NARWANI, K., CLARK, A. T., ANDREWS, P. W., BENVENISTY, N. & CARLOS BIANCOTTI, J. 2014. Aneuploidy induces profound changes in gene expression, proliferation and tumorigenicity of human pluripotent stem cells. *Nat Commun,* 5**,** 4825.
BENEDICT, B., VAN SCHIE, J. J. M., OOSTRA, A. B., BALK, J. A., WOLTHUIS, R. M. F., RIELE, H. T. & DE LANGE, J. 2020. WAPL-Dependent Repair of Damaged DNA Replication Forks Underlies Oncogene-Induced Loss of Sister Chromatid Cohesion. *Dev Cell,* 52**,** 683-698 e7.
BERMEJO, R., BRANZEI, D. & FOIANI, M. 2008. Cohesion by topology: sister chromatids interlocked by DNA. *Genes Dev,* 22**,** 2297-301.
BERMEJO, R., CAPRA, T., JOSSEN, R., COLOSIO, A., FRATTINI, C., CAROTENUTO, W., COCITO, A., DOKSANI, Y., KLEIN, H., GOMEZ-GONZALEZ, B., AGUILERA, A., KATOU, Y., SHIRAHIGE, K. & FOIANI, M. 2011. The replication checkpoint protects fork stability by releasing transcribed genes from nuclear pores. *Cell,* 146**,** 233-46.
BERMUDEZ-LOPEZ, M., CESCHIA, A., DE PICCOLI, G., COLOMINA, N., PASERO, P., ARAGON, L. & TORRES-ROSELL, J. 2010. The Smc5/6 complex is required for dissolution of DNA-mediated sister chromatid linkages. *Nucleic Acids Res,* 38**,** 6502-12.
BERMUDEZ-LOPEZ, M., POCINO-MERINO, I., SANCHEZ, H., BUENO, A., GUASCH, C., ALMEDAWAR, S., BRU-VIRGILI, S., GARI, E., WYMAN, C., REVERTER, D., COLOMINA, N. & TORRES-ROSELL, J. 2015. ATPase-dependent control of the Mms21 SUMO ligase during DNA repair. *PLoS Biol,* 13**,** e1002089.
BERMUDEZ-LOPEZ, M., VILLORIA, M. T., ESTERAS, M., JARMUZ, A., TORRES-ROSELL, J., CLEMENTE-BLANCO, A. & ARAGON, L. 2016. Sgs1's roles in DNA end resection, HJ dissolution, and crossover suppression require a two-step SUMO regulation dependent on Smc5/6. *Genes Dev,* 30**,** 1339-56.
BERMUDEZ, V. P., LINDSEY-BOLTZ, L. A., CESARE, A. J., MANIWA, Y., GRIFFITH, J. D., HURWITZ, J. & SANCAR, A. 2003. Loading of the human 9-1-1 checkpoint complex onto DNA by the

checkpoint clamp loader hRad17-replication factor C complex in vitro. *Proc Natl Acad Sci U S A,* 100**,** 1633-8.
BERTI, M., CORTEZ, D. & LOPES, M. 2020. The plasticity of DNA replication forks in response to clinically relevant genotoxic stress. *Nat Rev Mol Cell Biol,* 21**,** 633-651.
BERTOLIN, A. P., HOFFMANN, J. S. & GOTTIFREDI, V. 2020. Under-Replicated DNA: The Byproduct of Large Genomes? *Cancers (Basel),* 12.
BETOUS, R., MASON, A. C., RAMBO, R. P., BANSBACH, C. E., BADU-NKANSAH, A., SIRBU, B. M., EICHMAN, B. F. & CORTEZ, D. 2012. SMARCAL1 catalyzes fork regression and Holliday junction migration to maintain genome stability during DNA replication. *Genes Dev,* 26**,** 151-62.
BETTS LINDROOS, H., STRÖM, L., ITOH, T., KATOU, Y., SHIRAHIGE, K. & SJÖGREN, C. 2006. Chromosomal Association of the Smc5/6 Complex Reveals that It Functions in Differently Regulated Pathways. *Molecular Cell,* 22**,** 755-767.
BHARGAVA, R., ONYANGO, D. O. & STARK, J. M. 2016. Regulation of Single-Strand Annealing and its Role in Genome Maintenance. *Trends Genet,* 32**,** 566-575.
BHAT, K. P. & CORTEZ, D. 2018. RPA and RAD51: fork reversal, fork protection, and genome stability. *Nat Struct Mol Biol,* 25**,** 446-453.
BHOWMICK, R. & HICKSON, I. D. 2017. The "enemies within": regions of the genome that are inherently difficult to replicate. *F1000Res,* 6**,** 666.
BIGNELL, G. R., GREENMAN, C. D., DAVIES, H., BUTLER, A. P., EDKINS, S., ANDREWS, J. M., BUCK, G., CHEN, L., BEARE, D., LATIMER, C., WIDAA, S., HINTON, J., FAHEY, C., FU, B., SWAMY, S., DALGLIESH, G. L., TEH, B. T., DELOUKAS, P., YANG, F., CAMPBELL, P. J., FUTREAL, P. A. & STRATTON, M. R. 2010. Signatures of mutation and selection in the cancer genome. *Nature,* 463**,** 893-8.
BIRCHLER, J. A. & VEITIA, R. A. 2012. Gene balance hypothesis: connecting issues of dosage sensitivity across biological disciplines. *Proc Natl Acad Sci U S A,* 109**,** 14746-53.
BIZARD, A. H. & HICKSON, I. D. 2014. The dissolution of double Holliday junctions. *Cold Spring Harb Perspect Biol,* 6**,** a016477.
BLACKFORD, A. N. & JACKSON, S. P. 2017. ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. *Mol Cell,* 66**,** 801-817.
BLASTYAK, A., HAJDU, I., UNK, I. & HARACSKA, L. 2010. Role of double-stranded DNA translocase activity of human HLTF in replication of damaged DNA. *Mol Cell Biol,* 30**,** 684-93.
BLOW, J. J. & GE, X. Q. 2009. A model for DNA replication showing how dormant origins safeguard against replication fork failure. *EMBO Rep,* 10**,** 406-12.
BOHGAKI, M., BOHGAKI, T., EL GHAMRASNI, S., SRIKUMAR, T., MAIRE, G., PANIER, S., FRADET-TURCOTTE, A., STEWART, G. S., RAUGHT, B., HAKEM, A. & HAKEM, R. 2013. RNF168 ubiquitylates 53BP1 and controls its response to DNA double-strand breaks. *Proc Natl Acad Sci U S A,* 110**,** 20982-7.
BONNER, J. N., CHOI, K., XUE, X., TORRES, N. P., SZAKAL, B., WEI, L., WAN, B., ARTER, M., MATOS, J., SUNG, P., BROWN, G. W., BRANZEI, D. & ZHAO, X. 2016. Smc5/6 Mediated Sumoylation of the Sgs1-Top3-Rmi1 Complex Promotes Removal of Recombination Intermediates. *Cell Rep,* 16**,** 368-378.
BOOS, D. & FERREIRA, P. 2019. Origin Firing Regulations to Control Genome Replication Timing. *Genes,* 10.
BOSSAERT, M., PIPIER, A., RIOU, J. F., NOIROT, C., NGUYEN, L. T., SERRE, R. F., BOUCHEZ, O., DEFRANCQ, E., CALSOU, P., BRITTON, S. & GOMEZ, D. 2021. Transcription-associated topoisomerase 2alpha (TOP2A) activity is a major effector of cytotoxicity induced by G-quadruplex ligands. *Elife,* 10.
BRANDSMA, I. & GENT, D. C. 2012. Pathway choice in DNA double strand break repair: observations of a balancing act. *Genome Integr,* 3**,** 9.

BRANZEI, D., SOLLIER, J., LIBERI, G., ZHAO, X., MAEDA, D., SEKI, M., ENOMOTO, T., OHTA, K. & FOIANI, M. 2006. Ubc9- and mms21-mediated sumoylation counteracts recombinogenic events at damaged replication forks. *Cell,* 127**,** 509-22.

BREWER, B. J., LOCKSHON, D. & FANGMAN, W. L. 1992. The arrest of replication forks in the rDNA of yeast occurs independently of transcription. *Cell,* 71**,** 267-76.

BRINGMANN, H. & HYMAN, A. A. 2005. A cytokinesis furrow is positioned by two consecutive signals. *Nature,* 436**,** 731-734.

BROOKER, A. S. & BERKOWITZ, K. M. 2014. The roles of cohesins in mitosis, meiosis, and human health and disease. *Methods Mol Biol,* 1170**,** 229-66.

BRUNO, P. M., LU, M., DENNIS, K. A., INAM, H., MOORE, C. J., SHEEHE, J., ELLEDGE, S. J., HEMANN, M. T. & PRITCHARD, J. R. 2020. The primary mechanism of cytotoxicity of the chemotherapeutic agent CX-5461 is topoisomerase II poisoning. *Proc Natl Acad Sci U S A,* 117**,** 4053-4060.

BRYAN, T. M. 2020. G-Quadruplexes at Telomeres: Friend or Foe? *Molecules,* 25.

BUGREEV, D. V., ROSSI, M. J. & MAZIN, A. V. 2011. Cooperation of RAD51 and RAD54 in regression of a model replication fork. *Nucleic Acids Res,* 39**,** 2153-64.

BURGERS, P. M. J. & KUNKEL, T. A. 2017. Eukaryotic DNA Replication Fork. *Annu Rev Biochem,* 86**,** 417-438.

BUSH, N. G., EVANS-ROBERTS, K. & MAXWELL, A. 2015. DNA Topoisomerases. *EcoSal Plus,* 6.

BYUN, T. S., PACEK, M., YEE, M. C., WALTER, J. C. & CIMPRICH, K. A. 2005. Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint. *Genes Dev,* 19**,** 1040-52.

CALI, F., BHARTI, S. K., DI PERNA, R., BROSH, R. M., JR. & PISANI, F. M. 2016. Tim/Timeless, a member of the replication fork protection complex, operates with the Warsaw breakage syndrome DNA helicase DDX11 in the same fork recovery pathway. *Nucleic Acids Res,* 44**,** 705-17.

CANNAN, W. J. & PEDERSON, D. S. 2016. Mechanisms and Consequences of Double-Strand DNA Break Formation in Chromatin. *J Cell Physiol,* 231**,** 3-14.

CARIDI, C. P., D'AGOSTINO, C., RYU, T., ZAPOTOCZNY, G., DELABAERE, L., LI, X., KHODAVERDIAN, V. Y., AMARAL, N., LIN, E., RAU, A. R. & CHIOLO, I. 2018. Nuclear F-actin and myosins drive relocalization of heterochromatic breaks. *Nature,* 559**,** 54-60.

CARVALHAL, S., BADER, I., ROOIMANS, M. A., OOSTRA, A. B., BALK, J. A., FEICHTINGER, R. G., BEICHLER, C., SPEICHER, M. R., VAN HAGEN, J. M., WAISFISZ, Q., VAN HAELST, M., BRUIJN, M., TAVARES, A., MAYR, J. A., WOLTHUIS, R. M. F., OLIVEIRA, R. A. & DE LANGE, J. 2022. Biallelic BUB1 mutations cause microcephaly, developmental delay, and variable effects on cohesion and chromosome segregation. *Sci Adv,* 8**,** eabk0114.

CASTAN, A., HERNANDEZ, P., KRIMER, D. B. & SCHVARTZMAN, J. B. 2017. The abundance of Fob1 modulates the efficiency of rRFBs to stall replication forks. *Nucleic Acids Res,* 45**,** 10089-10102.

CASTILLO BOSCH, P., SEGURA-BAYONA, S., KOOLE, W., VAN HETEREN, J. T., DEWAR, J. M., TIJSTERMAN, M. & KNIPSCHEER, P. 2014. FANCJ promotes DNA synthesis through G-quadruplex structures. *The EMBO Journal,* 33**,** 2521-2533.

CEBRIAN, J., CASTAN, A., MARTINEZ, V., KADOMATSU-HERMOSA, M. J., PARRA, C., FERNANDEZ-NESTOSA, M. J., SCHAERER, C., HERNANDEZ, P., KRIMER, D. B. & SCHVARTZMAN, J. B. 2015. Direct Evidence for the Formation of Precatenanes during DNA Replication. *J Biol Chem,* 290**,** 13725-35.

CECCALDI, R., PARMAR, K., MOULY, E., DELORD, M., KIM, J. M., REGAIRAZ, M., PLA, M., VASQUEZ, N., ZHANG, Q. S., PONDARRE, C., PEFFAULT DE LATOUR, R., GLUCKMAN, E., CAVAZZANA-CALVO, M., LEBLANC, T., LARGHERO, J., GROMPE, M., SOCIE, G., D'ANDREA, A. D. & SOULIER, J. 2012. Bone marrow failure in Fanconi anemia is triggered by an exacerbated

p53/p21 DNA damage response that impairs hematopoietic stem and progenitor cells. *Cell Stem Cell,* 11**,** 36-49.
CECCALDI, R., SARANGI, P. & D'ANDREA, A. D. 2016. The Fanconi anaemia pathway: new players and new functions. *Nat Rev Mol Cell Biol,* 17**,** 337-49.
CERRITELLI, S. M. & CROUCH, R. J. 2009. Ribonuclease H: the enzymes in eukaryotes. *FEBS J,* 276**,** 1494-505.
CESARE, A. J. & REDDEL, R. R. 2008. Telomere uncapping and alternative lengthening of telomeres. *Mech Ageing Dev,* 129**,** 99-108.
CESARE, A. J. & REDDEL, R. R. 2010. Alternative lengthening of telomeres: models, mechanisms and implications. *Nat Rev Genet,* 11**,** 319-30.
CHA, R. S. & KLECKNER, N. 2002. ATR homolog Mec1 promotes fork progression, thus averting breaks in replication slow zones. *Science,* 297**,** 602-6.
CHAGANTI, R. S., SCHONBERG, S. & GERMAN, J. 1974. A manyfold increase in sister chromatid exchanges in Bloom's syndrome lymphocytes. *Proc Natl Acad Sci U S A,* 71**,** 4508-12.
CHAMPOUX, J. J. 2001. DNA topoisomerases: structure, function, and mechanism. *Annu Rev Biochem,* 70**,** 369-413.
CHAN, K.-L., NORTH, P. S. & HICKSON, I. D. 2007. BLM is required for faithful chromosome segregation and its localization defines a class of ultrafine anaphase bridges. *The EMBO Journal,* 26**,** 3397-3409.
CHAN, K. L., PALMAI-PALLAG, T., YING, S. & HICKSON, I. D. 2009. Replication stress induces sister-chromatid bridging at fragile site loci in mitosis. *Nat Cell Biol,* 11**,** 753-60.
CHAN, Y. W., FUGGER, K. & WEST, S. C. 2018. Unresolved recombination intermediates lead to ultra-fine anaphase bridges, chromosome breaks and aberrations. *Nat Cell Biol,* 20**,** 92-103.
CHAN, Y. W. & WEST, S. C. 2018. A new class of ultrafine anaphase bridges generated by homologous recombination. *Cell Cycle,* 17**,** 2101-2109.
CHANG, E. Y., NOVOA, C. A., ARISTIZABAL, M. J., COULOMBE, Y., SEGOVIA, R., CHATURVEDI, R., SHEN, Y., KEONG, C., TAM, A. S., JONES, S. J. M., MASSON, J. Y., KOBOR, M. S. & STIRLING, P. C. 2017. RECQ-like helicases Sgs1 and BLM regulate R-loop-associated genome instability. *J Cell Biol,* 216**,** 3991-4005.
CHAO, H. X., POOVEY, C. E., PRIVETTE, A. A., GRANT, G. D., CHAO, H. Y., COOK, J. G. & PURVIS, J. E. 2017. Orchestration of DNA Damage Checkpoint Dynamics across the Human Cell Cycle. *Cell Syst,* 5**,** 445-459 e5.
CHASTAIN, P. D., 2ND, HEFFERNAN, T. P., NEVIS, K. R., LIN, L., KAUFMANN, W. K., KAUFMAN, D. G. & CORDEIRO-STONE, M. 2006. Checkpoint regulation of replication dynamics in UV-irradiated human cells. *Cell Cycle,* 5**,** 2160-7.
CHATTERJEE, N. & WALKER, G. C. 2017. Mechanisms of DNA damage, repair, and mutagenesis. *Environ Mol Mutagen,* 58**,** 235-263.
CHAUDHURY, I., STROIK, D. R. & SOBECK, A. 2014. FANCD2-controlled chromatin access of the Fanconi-associated nuclease FAN1 is crucial for the recovery of stalled replication forks. *Mol Cell Biol,* 34**,** 3939-54.
CHAVEZ, A., AGRAWAL, V. & JOHNSON, F. B. 2011. Homologous recombination-dependent rescue of deficiency in the structural maintenance of chromosomes (Smc) 5/6 complex. *J Biol Chem,* 286**,** 5119-25.
CHAVEZ, A., GEORGE, V., AGRAWAL, V. & JOHNSON, F. B. 2010. Sumoylation and the structural maintenance of chromosomes (Smc) 5/6 complex slow senescence through recombination intermediate resolution. *J Biol Chem,* 285**,** 11922-30.
CHEN, C.-P., LIN, S.-P., CHERN, S.-R., TSAI, F.-J., WU, P.-C., LEE, C.-C., CHEN, Y.-T., CHEN, W.-L. & WANG, W. 2010. A de novo 7.9 Mb deletion in 22q13.2→qter in a boy with autistic features, epilepsy, developmental delay, atopic dermatitis and abnormal immunological findings. *European Journal of Medical Genetics,* 53**,** 329-332.

CHEN, H., CHEN, H., ZHANG, J., WANG, Y., SIMONEAU, A., YANG, H., LEVINE, A. S., ZOU, L., CHEN, Z. & LAN, L. 2020. cGAS suppresses genomic instability as a decelerator of replication forks. *Sci Adv,* 6.

CHEN, H., LISBY, M. & SYMINGTON, L. S. 2013. RPA coordinates DNA end resection and prevents formation of DNA hairpins. *Mol Cell,* 50**,** 589-600.

CHEN, X., XU, X., CHEN, Y., CHEUNG, J. C., WANG, H., JIANG, J., DE VAL, N., FOX, T., GELLERT, M. & YANG, W. 2021. Structure of an activated DNA-PK and its implications for NHEJ. *Mol Cell,* 81**,** 801-810 e3.

CHEN, Y. H., CHOI, K., SZAKAL, B., ARENZ, J., DUAN, X., YE, H., BRANZEI, D. & ZHAO, X. 2009. Interplay between the Smc5/6 complex and the Mph1 helicase in recombinational repair. *Proc Natl Acad Sci U S A,* 106**,** 21252-7.

CHEN, Y. H., JONES, M. J., YIN, Y., CRIST, S. B., COLNAGHI, L., SIMS, R. J., 3RD, ROTHENBERG, E., JALLEPALLI, P. V. & HUANG, T. T. 2015. ATR-mediated phosphorylation of FANCI regulates dormant origin firing in response to replication stress. *Mol Cell,* 58**,** 323-38.

CHEN, Y. H., KEEGAN, S., KAHLI, M., TONZI, P., FENYO, D., HUANG, T. T. & SMITH, D. J. 2019. Transcription shapes DNA replication initiation and termination in human cells. *Nat Struct Mol Biol,* 26**,** 67-77.

CHENG, H., ZHANG, N. & PATI, D. 2020. Cohesin subunit RAD21: From biology to disease. *Gene,* 758**,** 144966.

CHIANG, T. W., LE SAGE, C., LARRIEU, D., DEMIR, M. & JACKSON, S. P. 2016. CRISPR-Cas9(D10A) nickase-based genotypic and phenotypic screening to enhance genome editing. *Sci Rep,* 6**,** 24356.

CHIOLO, I., MINODA, A., COLMENARES, S. U., POLYZOS, A., COSTES, S. V. & KARPEN, G. H. 2011. Double-strand breaks in heterochromatin move outside of a dynamic HP1a domain to complete recombinational repair. *Cell,* 144**,** 732-44.

CHUNG, I., LEONHARDT, H. & RIPPE, K. 2011. De novo assembly of a PML nuclear subcompartment occurs through multiple pathways and induces telomere elongation. *Journal of Cell Science,* 124**,** 3603-3618.

CIMINI, D., WAN, X., HIREL, C. B. & SALMON, E. D. 2006. Aurora Kinase Promotes Turnover of Kinetochore Microtubules to Reduce Chromosome Segregation Errors. *Current Biology,* 16**,** 1711-1718.

COBB, J. A., SCHLEKER, T., ROJAS, V., BJERGBAEK, L., TERCERO, J. A. & GASSER, S. M. 2005. Replisome instability, fork collapse, and gross chromosomal rearrangements arise synergistically from Mec1 kinase and RecQ helicase mutations. *Genes Dev,* 19**,** 3055-69.

COLE, F., KEENEY, S. & JASIN, M. 2010. Evolutionary conservation of meiotic DSB proteins: more than just Spo11. *Genes Dev,* 24**,** 1201-7.

COLLIER, J. E., LEE, B.-G., ROIG, M. B., YATSKEVICH, S., PETELA, N. J., METSON, J., VOULGARIS, M., GONZALEZ LLAMAZARES, A., LÖWE, J. & NASMYTH, K. A. 2020. Transport of DNA within cohesin involves clamping on top of engaged heads by Scc2 and entrapment within the ring by Scc3. *eLife,* 9**,** e59560.

COMBES, G., ALHARBI, I., BRAGA, L. G. & ELOWE, S. 2017. Playing polo during mitosis: PLK1 takes the lead. *Oncogene,* 36**,** 4819-4827.

CONDE, C. D., PETRONCZKI, O. Y., BARIS, S., WILLMANN, K. L., GIRARDI, E., SALZER, E., WEITZER, S., ARDY, R. C., KROLO, A., IJSPEERT, H., KIYKIM, A., KARAKOC-AYDINER, E., FORSTER-WALDL, E., KAGER, L., PICKL, W. F., SUPERTI-FURGA, G., MARTINEZ, J., LOIZOU, J. I., OZEN, A., VAN DER BURG, M. & BOZTUG, K. 2019. Polymerase delta deficiency causes syndromic immunodeficiency with replicative stress. *J Clin Invest,* 129**,** 4194-4206.

CONDUIT, P. T., WAINMAN, A. & RAFF, J. W. 2015. Centrosome function and assembly in animal cells. *Nature Reviews Molecular Cell Biology,* 16**,** 611-624.

COPELAND, W. C. & WANG, T. S. 1993. Enzymatic characterization of the individual mammalian primase subunits reveals a biphasic mechanism for initiation of DNA replication. *J Biol Chem,* 268**,** 26179-89.

CORTES LEDESMA, F., EL KHAMISY, S. F., ZUMA, M. C., OSBORN, K. & CALDECOTT, K. W. 2009. A human 5'-tyrosyl DNA phosphodiesterase that repairs topoisomerase-mediated DNA damage. *Nature,* 461**,** 674-8.

CORTEZ, D. 2015. Preventing replication fork collapse to maintain genome integrity. *DNA Repair (Amst),* 32**,** 149-157.

CORTONE, G., ZHENG, G., PENSIERI, P., CHIAPPETTA, V., TATE, R., MALACARIA, E., PICHIERRI, P., YU, H. & PISANI, F. M. 2018. Interaction of the Warsaw breakage syndrome DNA helicase DDX11 with the replication fork-protection factor Timeless promotes sister chromatid cohesion. *PLoS Genet,* 14**,** e1007622.

COSTER, G. & GOLDBERG, M. 2010. The cellular response to DNA damage: a focus on MDC1 and its interacting proteins. *Nucleus,* 1**,** 166-78.

COTTINEAU, J., KOTTEMANN, M. C., LACH, F. P., KANG, Y. H., VELY, F., DEENICK, E. K., LAZAROV, T., GINEAU, L., WANG, Y., FARINA, A., CHANSEL, M., LORENZO, L., PIPEROGLOU, C., MA, C. S., NITSCHKE, P., BELKADI, A., ITAN, Y., BOISSON, B., JABOT-HANIN, F., PICARD, C., BUSTAMANTE, J., EIDENSCHENK, C., BOUCHERIT, S., ALADJIDI, N., LACOMBE, D., BARAT, P., QASIM, W., HURST, J. A., POLLARD, A. J., UHLIG, H. H., FIESCHI, C., MICHON, J., BERMUDEZ, V. P., ABEL, L., DE VILLARTAY, J. P., GEISSMANN, F., TANGYE, S. G., HURWITZ, J., VIVIER, E., CASANOVA, J. L., SMOGORZEWSKA, A. & JOUANGUY, E. 2017. Inherited GINS1 deficiency underlies growth retardation along with neutropenia and NK cell deficiency. *J Clin Invest,* 127**,** 1991-2006.

COUCH, F. B., BANSBACH, C. E., DRISCOLL, R., LUZWICK, J. W., GLICK, G. G., BETOUS, R., CARROLL, C. M., JUNG, S. Y., QIN, J., CIMPRICH, K. A. & CORTEZ, D. 2013. ATR phosphorylates SMARCAL1 to prevent replication fork collapse. *Genes Dev,* 27**,** 1610-23.

COURTOT, L., HOFFMANN, J. S. & BERGOGLIO, V. 2018. The Protective Role of Dormant Origins in Response to Replicative Stress. *Int J Mol Sci,* 19.

CRABBE, L., VERDUN, R. E., HAGGBLOM, C. I. & KARLSEDER, J. 2004. Defective telomere lagging strand synthesis in cells lacking WRN helicase activity. *Science,* 306**,** 1951-3.

CSERESNYES, Z., SCHWARZ, U. & GREEN, C. M. 2009. Analysis of replication factories in human cells by super-resolution light microscopy. *BMC Cell Biol,* 10**,** 88.

CUNNIFF, C., BASSETTI, J. A. & ELLIS, N. A. 2017. Bloom's Syndrome: Clinical Spectrum, Molecular Pathogenesis, and Cancer Predisposition. *Mol Syndromol,* 8**,** 4-23.

DA COSTA ALMEIDA, C. B., WELTER, A. T., ABECH, G. D., BRANDÃO, G. R., FLORES, J. A. M., SCHÜLE, B., FRANCKE, U., FIEGENBAUM, M., ZEN, P. R. G. & ROSA, R. F. M. 2020. Report of the Phenotype of a Patient with Roberts Syndrome and a Rare ESCO2 Variant. *J Pediatr Genet,* 9**,** 58-62.

DABOUSSI, F., THACKER, J. & LOPEZ, B. S. 2005. Genetic interactions between RAD51 and its paralogues for centrosome fragmentation and ploidy control, independently of the sensitivity to genotoxic stresses. *Oncogene,* 24**,** 3691-3696.

DAHAN, D., TSIRKAS, I., DOVRAT, D., SPARKS, M. A., SINGH, S. P., GALLETTO, R. & AHARONI, A. 2018. Pif1 is essential for efficient replisome progression through lagging strand G-quadruplex DNA secondary structures. *Nucleic Acids Res,* 46**,** 11847-11857.

DAI, J., SULLIVAN, B. A. & HIGGINS, J. M. 2006. Regulation of mitotic chromosome cohesion by Haspin and Aurora B. *Dev Cell,* 11**,** 741-50.

DAIGAKU, Y., KESZTHELYI, A., MULLER, C. A., MIYABE, I., BROOKS, T., RETKUTE, R., HUBANK, M., NIEDUSZYNSKI, C. A. & CARR, A. M. 2015. A global profile of replicative polymerase usage. *Nat Struct Mol Biol,* 22**,** 192-198.

DALEY, J. M., NIU, H., MILLER, A. S. & SUNG, P. 2015. Biochemical mechanism of DSB end resection and its regulation. *DNA Repair (Amst),* 32**,** 66-74.
DATTA, A., POLLOCK, K. J., KORMUTH, K. A. & BROSH, R. M., JR. 2021. G-Quadruplex Assembly by Ribosomal DNA: Emerging Roles in Disease Pathogenesis and Cancer Biology. *Cytogenet Genome Res,* 161**,** 285-296.
DAUM, J. R., POTAPOVA, T. A., SIVAKUMAR, S., DANIEL, J. J., FLYNN, J. N., RANKIN, S. & GORBSKY, G. J. 2011. Cohesion fatigue induces chromatid separation in cells delayed at metaphase. *Curr Biol,* 21**,** 1018-24.
DAVIDSON, I. F., BAUER, B., GOETZ, D., TANG, W., WUTZ, G. & PETERS, J.-M. 2019. DNA loop extrusion by human cohesin. *Science,* 366**,** 1338-1345.
DAVIDSON, I. F., LI, A. & BLOW, J. J. 2006. Deregulated replication licensing causes DNA fragmentation consistent with head-to-tail fork collision. *Mol Cell,* 24**,** 433-43.
DAVIES, A. A., MASSON, J. Y., MCILWRAITH, M. J., STASIAK, A. Z., STASIAK, A., VENKITARAMAN, A. R. & WEST, S. C. 2001. Role of BRCA2 in control of the RAD51 recombination and DNA repair protein. *Mol Cell,* 7**,** 273-82.
DE LANGE, J., FARAMARZ, A., OOSTRA, A. B., DE MENEZES, R. X., VAN DER MEULEN, I. H., ROOIMANS, M. A., ROCKX, D. A., BRAKENHOFF, R. H., VAN BEUSECHEM, V. W., KING, R. W., DE WINTER, J. P. & WOLTHUIS, R. M. F. 2015. Defective sister chromatid cohesion is synthetically lethal with impaired APC/C function. *Nat Commun,* 6**,** 8399.
DE PICCOLI, G., CORTES-LEDESMA, F., IRA, G., TORRES-ROSELL, J., UHLE, S., FARMER, S., HWANG, J. Y., MACHIN, F., CESCHIA, A., MCALEENAN, A., CORDON-PRECIADO, V., CLEMENTE-BLANCO, A., VILELLA-MITJANA, F., ULLAL, P., JARMUZ, A., LEITAO, B., BRESSAN, D., DOTIWALA, F., PAPUSHA, A., ZHAO, X., MYUNG, K., HABER, J. E., AGUILERA, A. & ARAGON, L. 2006. Smc5-Smc6 mediate DNA double-strand-break repair by promoting sister-chromatid recombination. *Nat Cell Biol,* 8**,** 1032-4.
DE PICCOLI, G., KATOU, Y., ITOH, T., NAKATO, R., SHIRAHIGE, K. & LABIB, K. 2012. Replisome stability at defective DNA replication forks is independent of S phase checkpoint kinases. *Mol Cell,* 45**,** 696-704.
DE WOLF, B., OGHABIAN, A., AKINYI, M. V., HANKS, S., TROMER, E. C., VAN HOOFF, J. J. E., VAN VOORTHUIJSEN, L., VAN ROOIJEN, L. E., VERBEEREN, J., UIJTTEWAAL, E. C. H., BALTISSEN, M. P. A., YOST, S., PILOQUET, P., VERMEULEN, M., SNEL, B., ISIDOR, B., RAHMAN, N., FRILANDER, M. J. & KOPS, G. 2021. Chromosomal instability by mutations in the novel minor spliceosome component CENATAC. *EMBO J,* 40**,** e106536.
DEANS, A. J. & WEST, S. C. 2011. DNA interstrand crosslink repair and cancer. *Nat Rev Cancer,* 11**,** 467-80.
DEEM, A., KESZTHELYI, A., BLACKGROVE, T., VAYL, A., COFFEY, B., MATHUR, R., CHABES, A. & MALKOVA, A. 2011. Break-induced replication is highly inaccurate. *PLoS Biol,* 9**,** e1000594.
DEFAZIO, L. G., STANSEL, R. M., GRIFFITH, J. D. & CHU, G. 2002. Synapsis of DNA ends by DNA-dependent protein kinase. *EMBO J,* 21**,** 3192-200.
DELACROIX, S., WAGNER, J. M., KOBAYASHI, M., YAMAMOTO, K. & KARNITZ, L. M. 2007. The Rad9-Hus1-Rad1 (9-1-1) clamp activates checkpoint signaling via TopBP1. *Genes Dev,* 21**,** 1472-7.
DENCHI, E. L. & DE LANGE, T. 2007. Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1. *Nature,* 448**,** 1068-71.
DEVBHANDARI, S., JIANG, J., KUMAR, C., WHITEHOUSE, I. & REMUS, D. 2017. Chromatin Constrains the Initiation and Elongation of DNA Replication. *Mol Cell,* 65**,** 131-141.
DEVGAN, S. S., SANAL, O., DOIL, C., NAKAMURA, K., NAHAS, S. A., PETTIJOHN, K., BARTEK, J., LUKAS, C., LUKAS, J. & GATTI, R. A. 2011. Homozygous deficiency of ubiquitin-ligase ring-finger protein RNF168 mimics the radiosensitivity syndrome of ataxia-telangiectasia. *Cell Death Differ,* 18**,** 1500-6.

DEWAR, J. M., BUDZOWSKA, M. & WALTER, J. C. 2015. The mechanism of DNA replication termination in vertebrates. *Nature,* 525**,** 345-50.
DEWAR, J. M. & WALTER, J. C. 2017. Mechanisms of DNA replication termination. *Nat Rev Mol Cell Biol,* 18**,** 507-516.
DEWEESE, J. E. & OSHEROFF, N. 2009. The DNA cleavage reaction of topoisomerase II: wolf in sheep's clothing. *Nucleic Acids Res,* 37**,** 738-48.
DILLEY, R. L. & GREENBERG, R. A. 2015. ALTernative Telomere Maintenance and Cancer. *Trends Cancer,* 1**,** 145-156.
DINARDO, S., VOELKEL, K. & STERNGLANZ, R. 1984. DNA topoisomerase II mutant of Saccharomyces cerevisiae: topoisomerase II is required for segregation of daughter molecules at the termination of DNA replication. *Proc Natl Acad Sci U S A,* 81**,** 2616-20.
DOKSANI, Y., WU, J. Y., DE LANGE, T. & ZHUANG, X. 2013. Super-resolution fluorescence imaging of telomeres reveals TRF2-dependent T-loop formation. *Cell,* 155**,** 345-356.
DOMENECH, C., MAILLARD, L., ROUSSEAU, A., GUIDEZ, F., PETIT, L., PLA, M., CLAY, D., GUIMIOT, F., SANFILIPPO, S., JACQUES, S., DE LA GRANGE, P., ROBIL, N., SOULIER, J. & SOUYRI, M. 2018. Studies in an Early Development Window Unveils a Severe HSC Defect in both Murine and Human Fanconi Anemia. *Stem Cell Reports,* 11**,** 1075-1091.
DOUGLAS, M. E., ALI, F. A., COSTA, A. & DIFFLEY, J. F. X. 2018. The mechanism of eukaryotic CMG helicase activation. *Nature,* 555**,** 265-268.
DOYLE, J. M., GAO, J., WANG, J., YANG, M. & POTTS, P. R. 2010. MAGE-RING protein complexes comprise a family of E3 ubiquitin ligases. *Mol Cell,* 39**,** 963-74.
DROSOPOULOS, W. C., KOSIYATRAKUL, S. T. & SCHILDKRAUT, C. L. 2015. BLM helicase facilitates telomere replication during leading strand synthesis of telomeres. *J Cell Biol,* 210**,** 191-208.
DROUET, J., DELTEIL, C., LEFRANCOIS, J., CONCANNON, P., SALLES, B. & CALSOU, P. 2005. DNA-dependent protein kinase and XRCC4-DNA ligase IV mobilization in the cell in response to DNA double strand breaks. *J Biol Chem,* 280**,** 7060-9.
DRYGIN, D., LIN, A., BLIESATH, J., HO, C. B., O'BRIEN, S. E., PROFFITT, C., OMORI, M., HADDACH, M., SCHWAEBE, M. K., SIDDIQUI-JAIN, A., STREINER, N., QUIN, J. E., SANIJ, E., BYWATER, M. J., HANNAN, R. D., RYCKMAN, D., ANDERES, K. & RICE, W. G. 2011. Targeting RNA polymerase I with an oral small molecule CX-5461 inhibits ribosomal RNA synthesis and solid tumor growth. *Cancer Res,* 71**,** 1418-30.
DUAN, X., SARANGI, P., LIU, X., RANGI, G. K., ZHAO, X. & YE, H. 2009a. Structural and functional insights into the roles of the Mms21 subunit of the Smc5/6 complex. *Mol Cell,* 35**,** 657-68.
DUAN, X., YANG, Y., CHEN, Y. H., ARENZ, J., RANGI, G. K., ZHAO, X. & YE, H. 2009b. Architecture of the Smc5/6 Complex of Saccharomyces cerevisiae Reveals a Unique Interaction between the Nse5-6 Subcomplex and the Hinge Regions of Smc5 and Smc6. *J Biol Chem,* 284**,** 8507-15.
DUEVA, R. & ILIAKIS, G. 2020. Replication protein A: a multifunctional protein with roles in DNA replication, repair and beyond. *NAR Cancer,* 2**,** zcaa022.
DUNGRAWALA, H., ROSE, K. L., BHAT, K. P., MOHNI, K. N., GLICK, G. G., COUCH, F. B. & CORTEZ, D. 2015. The Replication Checkpoint Prevents Two Types of Fork Collapse without Regulating Replisome Stability. *Mol Cell,* 59**,** 998-1010.
DUPONT, L., BLOOR, S., WILLIAMSON, J. C., CUESTA, S. M., SHAH, R., TEIXEIRA-SILVA, A., NAAMATI, A., GREENWOOD, E. J. D., SARAFIANOS, S. G., MATHESON, N. J. & LEHNER, P. J. 2021. The SMC5/6 complex compacts and silences unintegrated HIV-1 DNA and is antagonized by Vpr. *Cell Host Microbe,* 29**,** 792-805 e6.
DYSON, S., SEGURA, J., MARTINEZ-GARCIA, B., VALDES, A. & ROCA, J. 2021. Condensin minimizes topoisomerase II-mediated entanglements of DNA in vivo. *EMBO J,* 40**,** e105393.

EDWARDS, D. N., MACHWE, A., WANG, Z. & ORREN, D. K. 2014. Intramolecular telomeric G-quadruplexes dramatically inhibit DNA synthesis by replicative and translesion polymerases, revealing their potential to lead to genetic change. *PLoS One,* 9**,** e80664.

EL HAGE, A., FRENCH, S. L., BEYER, A. L. & TOLLERVEY, D. 2010. Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. *Genes Dev,* 24**,** 1546-58.

ELBATSH, AHMED M. O., HAARHUIS, JUDITH H. I., PETELA, N., CHAPARD, C., FISH, A., CELIE, PATRICK H., STADNIK, M., RISTIC, D., WYMAN, C., MEDEMA, RENÉ H., NASMYTH, K. & ROWLAND, BENJAMIN D. 2016. Cohesin Releases DNA through Asymmetric ATPase-Driven Ring Opening. *Molecular Cell,* 61**,** 575-588.

ETEMAD, B., KUIJT, T. E. F. & KOPS, G. J. P. L. 2015. Kinetochore–microtubule attachment is sufficient to satisfy the human spindle assembly checkpoint. *Nature Communications,* 6**,** 8987.

ETHERIDGE, T. J., VILLAHERMOSA, D., CAMPILLO-FUNOLLET, E., HERBERT, A. D., IRMISCH, A., WATSON, A. T., DANG, H. Q., OSBORNE, M. A., OLIVER, A. W., CARR, A. M. & MURRAY, J. M. 2021. Live-cell single-molecule tracking highlights requirements for stable Smc5/6 chromatin association in vivo. *Elife,* 10.

FACHINETTI, D., BERMEJO, R., COCITO, A., MINARDI, S., KATOU, Y., KANOH, Y., SHIRAHIGE, K., AZVOLINSKY, A., ZAKIAN, V. A. & FOIANI, M. 2010. Replication termination at eukaryotic chromosomes is mediated by Top2 and occurs at genomic loci containing pausing elements. *Mol Cell,* 39**,** 595-605.

FARAMARZ, A., BALK, J. A., VAN SCHIE, J. J. M., OOSTRA, A. B., GHANDOUR, C. A., ROOIMANS, M. A., WOLTHUIS, R. M. F. & DE LANGE, J. 2020. Non-redundant roles in sister chromatid cohesion of the DNA helicase DDX11 and the SMC3 acetyl transferases ESCO1 and ESCO2. *PLoS One,* 15**,** e0220348.

FENECH, M., KIRSCH-VOLDERS, M., NATARAJAN, A. T., SURRALLES, J., CROTT, J. W., PARRY, J., NORPPA, H., EASTMOND, D. A., TUCKER, J. D. & THOMAS, P. 2011. Molecular mechanisms of micronucleus, nucleoplasmic bridge and nuclear bud formation in mammalian and human cells. *Mutagenesis,* 26**,** 125-32.

FERETZAKI, M., POSPISILOVA, M., VALADOR FERNANDES, R., LUNARDI, T., KREJCI, L. & LINGNER, J. 2020. RAD51-dependent recruitment of TERRA lncRNA to telomeres through R-loops. *Nature,* 587**,** 303-308.

FERNÁNDEZ-CASAÑAS, M. & CHAN, K. L. 2018. The Unresolved Problem of DNA Bridging. *Genes (Basel),* 9.

FERNANDEZ-CID, A., RIERA, A., TOGNETTI, S., HERRERA, M. C., SAMEL, S., EVRIN, C., WINKLER, C., GARDENAL, E., UHLE, S. & SPECK, C. 2013. An ORC/Cdc6/MCM2-7 complex is formed in a multistep reaction to serve as a platform for MCM double-hexamer assembly. *Mol Cell,* 50**,** 577-88.

FERRETTI, L. P., LAFRANCHI, L. & SARTORI, A. A. 2013. Controlling DNA-end resection: a new task for CDKs. *Front Genet,* 4**,** 99.

FINNIS, M., DAYAN, S., HOBSON, L., CHENEVIX-TRENCH, G., FRIEND, K., RIED, K., VENTER, D., WOOLLATT, E., BAKER, E. & RICHARDS, R. I. 2005. Common chromosomal fragile site FRA16D mutation in cancer cells. *Hum Mol Genet,* 14**,** 1341-9.

FLYNN, R. L. & ZOU, L. 2011. ATR: a master conductor of cellular responses to DNA replication stress. *Trends Biochem Sci,* 36**,** 133-40.

FOUSTERI, M. I. & LEHMANN, A. R. 2000. A novel SMC protein complex in Schizosaccharomyces pombe contains the Rad18 DNA repair protein. *EMBO J,* 19**,** 1691-702.

FRANCOEUR, N. & SEN, R. 2021. Advances in Cardiac Development and Regeneration Using Zebrafish as a Model System for High-Throughput Research. *J Dev Biol,* 9.

FRITSCH, O., BURKHALTER, M. D., KAIS, S., SOGO, J. M. & SCHAR, P. 2010. DNA ligase 4 stabilizes the ribosomal DNA array upon fork collapse at the replication fork barrier. *DNA Repair (Amst),* 9**,** 879-88.

FU, Y. V., YARDIMCI, H., LONG, D. T., HO, T. V., GUAINAZZI, A., BERMUDEZ, V. P., HURWITZ, J., VAN OIJEN, A., SCHARER, O. D. & WALTER, J. C. 2011. Selective bypass of a lagging strand roadblock by the eukaryotic replicative DNA helicase. *Cell,* 146**,** 931-41.

FUGGER, K., CHU, W. K., HAAHR, P., KOUSHOLT, A. N., BECK, H., PAYNE, M. J., HANADA, K., HICKSON, I. D. & SORENSEN, C. S. 2013. FBH1 co-operates with MUS81 in inducing DNA double-strand breaks and cell death following replication stress. *Nat Commun,* 4**,** 1423.

GAILLARD, H., GARCIA-MUSE, T. & AGUILERA, A. 2015. Replication stress and cancer. *Nat Rev Cancer,* 15**,** 276-89.

GALLEGO-PAEZ, L. M., TANAKA, H., BANDO, M., TAKAHASHI, M., NOZAKI, N., NAKATO, R., SHIRAHIGE, K. & HIROTA, T. 2014. Smc5/6-mediated regulation of replication progression contributes to chromosome assembly during mitosis in human cells. *Mol Biol Cell,* 25**,** 302-17.

GAMBUS, A. 2017. Termination of Eukaryotic Replication Forks. DNA Replication. *Advances in Experimental Medicine and Biology***,** 163-187.

GAN, W., GUAN, Z., LIU, J., GUI, T., SHEN, K., MANLEY, J. L. & LI, X. 2011. R-loop-mediated genomic instability is caused by impairment of replication fork progression. *Genes Dev,* 25**,** 2041-56.

GANDHI, R., GILLESPIE, P. J. & HIRANO, T. 2006. Human Wapl Is a Cohesin-Binding Protein that Promotes Sister-Chromatid Resolution in Mitotic Prophase. *Current Biology,* 16**,** 2406-2417.

GANJI, M., SHALTIEL, I. A., BISHT, S., KIM, E., KALICHAVA, A., HAERING, C. H. & DEKKER, C. 2018. Real-time imaging of DNA loop extrusion by condensin. *Science,* 360**,** 102-105.

GARCIA-HIGUERA, I., TANIGUCHI, T., GANESAN, S., MEYN, M. S., TIMMERS, C., HEJNA, J., GROMPE, M. & D'ANDREA, A. D. 2001. Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway. *Mol Cell,* 7**,** 249-62.

GARCÍA-MUSE, T. & AGUILERA, A. 2016. Transcription-replication conflicts: how they occur and how they are resolved. *Nat Rev Mol Cell Biol,* 17**,** 553-63.

GARCIA, P., FERNANDEZ-HERNANDEZ, R., CUADRADO, A., COCA, I., GOMEZ, A., MAQUEDA, M., LATORRE-PELLICER, A., PUISAC, B., RAMOS, F. J., SANDOVAL, J., ESTELLER, M., MOSQUERA, J. L., RODRIGUEZ, J., PIE, J., LOSADA, A. & QUERALT, E. 2021. Disruption of NIPBL/Scc2 in Cornelia de Lange Syndrome provokes cohesin genome-wide redistribution with an impact in the transcriptome. *Nat Commun,* 12**,** 4551.

GARG, P., STITH, C. M., SABOURI, N., JOHANSSON, E. & BURGERS, P. M. 2004. Idling by DNA polymerase delta maintains a ligatable nick during lagging-strand DNA replication. *Genes Dev,* 18**,** 2764-73.

GARNER, E., KIM, Y., LACH, F. P., KOTTEMANN, M. C. & SMOGORZEWSKA, A. 2013. Human GEN1 and the SLX4-associated nucleases MUS81 and SLX1 are essential for the resolution of replication-induced Holliday junctions. *Cell Rep,* 5**,** 207-15.

GE, X. Q. & BLOW, J. J. 2010. Chk1 inhibits replication factory activation but allows dormant origin firing in existing factories. *J Cell Biol,* 191**,** 1285-97.

GELOT, C., MAGDALOU, I. & LOPEZ, B. S. 2015. Replication stress in Mammalian cells and its consequences for mitosis. *Genes (Basel),* 6**,** 267-98.

GENNERY, A. R., CANT, A. J. & JEGGO, P. A. 2000. Immunodeficiency associated with DNA repair defects. *Clin Exp Immunol,* 121**,** 1-7.

GEORGESCU, R. E., LANGSTON, L., YAO, N. Y., YURIEVA, O., ZHANG, D., FINKELSTEIN, J., AGARWAL, T. & O'DONNELL, M. E. 2014. Mechanism of asymmetric polymerase assembly at the eukaryotic replication fork. *Nat Struct Mol Biol,* 21**,** 664-70.

GERKES, E. H., VAN DER KEVIE-KERSEMAEKERS, A. M., YAKIN, M., SMEETS, D. F. & VAN RAVENSWAAIJ-ARTS, C. M. 2010. The importance of chromosome studies in Roberts syndrome/SC phocomelia and other cohesinopathies. *Eur J Med Genet,* 53**,** 40-4.

GERLICH, D., KOCH, B., DUPEUX, F., PETERS, J.-M. & ELLENBERG, J. 2006. Live-Cell Imaging Reveals a Stable Cohesin-Chromatin Interaction after but Not before DNA Replication. *Current Biology,* 16**,** 1571-1578.

GIBCUS, J. H., SAMEJIMA, K., GOLOBORODKO, A., SAMEJIMA, I., NAUMOVA, N., NUEBLER, J., KANEMAKI, M. T., XIE, L., PAULSON, J. R., EARNSHAW, W. C., MIRNY, L. A. & DEKKER, J. 2018. A pathway for mitotic chromosome formation. *Science,* 359.

GILLESPIE, P. J. & BLOW, J. J. 2010. Clusters, factories and domains: The complex structure of S-phase comes into focus. *Cell Cycle,* 9**,** 3218-26.

GIMENEZ-ABIAN, J. F., SUMARA, I., HIROTA, T., HAUF, S., GERLICH, D., DE LA TORRE, C., ELLENBERG, J. & PETERS, J. M. 2004. Regulation of sister chromatid cohesion between chromosome arms. *Curr Biol,* 14**,** 1187-93.

GINNO, P. A., LIM, Y. W., LOTT, P. L., KORF, I. & CHEDIN, F. 2013. GC skew at the 5' and 3' ends of human genes links R-loop formation to epigenetic regulation and transcription termination. *Genome Res,* 23**,** 1590-600.

GLOVER, T. W., WILSON, T. E. & ARLT, M. F. 2017. Fragile sites in cancer: more than meets the eye. *Nat Rev Cancer,* 17**,** 489-501.

GOLFIER, S., QUAIL, T., KIMURA, H. & BRUGUÉS, J. 2020. Cohesin and condensin extrude DNA loops in a cell cycle-dependent manner. *eLife,* 9**,** e53885.

GOLOBORODKO, A., IMAKAEV, M. V., MARKO, J. F. & MIRNY, L. 2016. Compaction and segregation of sister chromatids via active loop extrusion. *Elife,* 5.

GOMEZ-GONZALEZ, B. & AGUILERA, A. 2007. Activation-induced cytidine deaminase action is strongly stimulated by mutations of the THO complex. *Proc Natl Acad Sci U S A,* 104**,** 8409-14.

GOMEZ-GONZALEZ, B. & AGUILERA, A. 2019. Transcription-mediated replication hindrance: a major driver of genome instability. *Genes Dev,* 33**,** 1008-1026.

GOMEZ-HERREROS, F. 2019. DNA Double Strand Breaks and Chromosomal Translocations Induced by DNA Topoisomerase II. *Front Mol Biosci,* 6**,** 141.

GOMEZ-HERREROS, F., ROMERO-GRANADOS, R., ZENG, Z., ALVAREZ-QUILON, A., QUINTERO, C., JU, L., UMANS, L., VERMEIRE, L., HUYLEBROECK, D., CALDECOTT, K. W. & CORTES-LEDESMA, F. 2013. TDP2-dependent non-homologous end-joining protects against topoisomerase II-induced DNA breaks and genome instability in cells and in vivo. *PLoS Genet,* 9**,** e1003226.

GOMEZ-HERREROS, F., SCHUURS-HOEIJMAKERS, J. H., MCCORMACK, M., GREALLY, M. T., RULTEN, S., ROMERO-GRANADOS, R., COUNIHAN, T. J., CHAILA, E., CONROY, J., ENNIS, S., DELANTY, N., CORTES-LEDESMA, F., DE BROUWER, A. P., CAVALLERI, G. L., EL-KHAMISY, S. F., DE VRIES, B. B. & CALDECOTT, K. W. 2014. TDP2 protects transcription from abortive topoisomerase activity and is required for normal neural function. *Nat Genet,* 46**,** 516-21.

GOMEZ, R., JORDAN, P. W., VIERA, A., ALSHEIMER, M., FUKUDA, T., JESSBERGER, R., LLANO, E., PENDAS, A. M., HANDEL, M. A. & SUJA, J. A. 2013. Dynamic localization of SMC5/6 complex proteins during mammalian meiosis and mitosis suggests functions in distinct chromosome processes. *J Cell Sci,* 126**,** 4239-52.

GORBSKY, G. J. 2013. Cohesion fatigue. *Curr Biol,* 23**,** R986-R988.

GORR, I. H., BOOS, D. & STEMMANN, O. 2005. Mutual Inhibition of Separase and Cdk1 by Two-Step Complex Formation. *Molecular Cell,* 19**,** 135-141.

GRANGE, L. J., REYNOLDS, J. J., ULLAH, F., ISIDOR, B., SHEARER, R. F., LATYPOVA, X., BAXLEY, R. M., OLIVER, A. W., GANESH, A., COOKE, S. L., JHUJH, S. S., MCNEE, G. S., HOLLINGWORTH, R., HIGGS, M. R., NATSUME, T., KHAN, T., MARTOS-MORENO, G. A., CHUPP, S., MATHEW, C. G., PARRY, D., SIMPSON, M. A., NAHAVANDI, N., YUKSEL, Z., DRASDO, M., KRON, A., VOGT, P.,

JONASSON, A., SETH, S. A., GONZAGA-JAUREGUI, C., BRIGATTI, K. W., STEGMANN, A. P. A., KANEMAKI, M., JOSIFOVA, D., UCHIYAMA, Y., OH, Y., MORIMOTO, A., OSAKA, H., AMMOUS, Z., ARGENTE, J., MATSUMOTO, N., STUMPEL, C., TAYLOR, A. M. R., JACKSON, A. P., BIELINSKY, A. K., MAILAND, N., LE CAIGNEC, C., DAVIS, E. E. & STEWART, G. S. 2022. Pathogenic variants in SLF2 and SMC5 cause segmented chromosomes and mosaic variegated hyperploidy. *Nat Commun,* 13**,** 6664.

GRANOTIER, C., PENNARUN, G., RIOU, L., HOFFSCHIR, F., GAUTHIER, L. R., DE CIAN, A., GOMEZ, D., MANDINE, E., RIOU, J. F., MERGNY, J. L., MAILLIET, P., DUTRILLAUX, B. & BOUSSIN, F. D. 2005. Preferential binding of a G-quadruplex ligand to human chromosome ends. *Nucleic Acids Res,* 33**,** 4182-90.

GREEN, B. M., FINN, K. J. & LI, J. J. 2010. Loss of DNA replication control is a potent inducer of gene amplification. *Science,* 329**,** 943-6.

GRIFFITH, J. D., COMEAU, L., ROSENFIELD, S., STANSEL, R. M., BIANCHI, A., MOSS, H. & DE LANGE, T. 1999. Mammalian telomeres end in a large duplex loop. *Cell,* 97**,** 503-14.

GROSBERG, A. Y. 2012. How two meters of DNA fit into a cell nucleus: Polymer models with topological constraints and experimental data. *Polymer Science Series C,* 54**,** 1-10.

GRUBER, S., ARUMUGAM, P., KATOU, Y., KUGLITSCH, D., HELMHART, W., SHIRAHIGE, K. & NASMYTH, K. 2006. Evidence that Loading of Cohesin Onto Chromosomes Involves Opening of Its SMC Hinge. *Cell,* 127**,** 523-537.

GU, J., LI, S., ZHANG, X., WANG, L. C., NIEWOLIK, D., SCHWARZ, K., LEGERSKI, R. J., ZANDI, E. & LIEBER, M. R. 2010. DNA-PKcs regulates a single-stranded DNA endonuclease activity of Artemis. *DNA Repair (Amst),* 9**,** 429-37.

GUDIMCHUK, N. B. & MCINTOSH, J. R. 2021. Regulation of microtubule dynamics, mechanics and function through the growing tip. *Nature Reviews Molecular Cell Biology,* 22**,** 777-795.

GUERINEAU, M., KRIZ, Z., KOZAKOVA, L., BEDNAROVA, K., JANOS, P. & PALECEK, J. 2012. Analysis of the Nse3/MAGE-binding domain of the Nse4/EID family proteins. *PLoS One,* 7**,** e35813.

GUILLIAM, T. A. & DOHERTY, A. J. 2017. PrimPol-Prime Time to Reprime. *Genes (Basel),* 8.

GUILLIAM, T. A., JOZWIAKOWSKI, S. K., EHLINGER, A., BARNES, R. P., RUDD, S. G., BAILEY, L. J., SKEHEL, J. M., ECKERT, K. A., CHAZIN, W. J. & DOHERTY, A. J. 2015. Human PrimPol is a highly error-prone polymerase regulated by single-stranded DNA binding proteins. *Nucleic Acids Res,* 43**,** 1056-68.

GUILLIAM, T. A. & YEELES, J. T. P. 2020. An updated perspective on the polymerase division of labor during eukaryotic DNA replication. *Critical Reviews in Biochemistry and Molecular Biology,* 55**,** 469-481.

GULSTON, M., FULFORD, J., JENNER, T., DE LARA, C. & O'NEILL, P. 2002. Clustered DNA damage induced by gamma radiation in human fibroblasts (HF19), hamster (V79-4) cells and plasmid DNA is revealed as Fpg and Nth sensitive sites. *Nucleic Acids Res,* 30**,** 3464-72.

GUO, C., KUMAGAI, A., SCHLACHER, K., SHEVCHENKO, A., SHEVCHENKO, A. & DUNPHY, W. G. 2015. Interaction of Chk1 with Treslin negatively regulates the initiation of chromosomal DNA replication. *Mol Cell,* 57**,** 492-505.

GÜTTINGER, S., LAURELL, E. & KUTAY, U. 2009. Orchestrating nuclear envelope disassembly and reassembly during mitosis. *Nature Reviews Molecular Cell Biology,* 10**,** 178-191.

HAARHUIS, J. H., ELBATSH, A. M. & ROWLAND, B. D. 2014. Cohesin and its regulation: on the logic of X-shaped chromosomes. *Dev Cell,* 31**,** 7-18.

HAERING, C. H., FARCAS, A.-M., ARUMUGAM, P., METSON, J. & NASMYTH, K. 2008. The cohesin ring concatenates sister DNA molecules. *Nature,* 454**,** 297-301.

HAERING, C. H., LOWE, J., HOCHWAGEN, A. & NASMYTH, K. 2002. Molecular architecture of SMC proteins and the yeast cohesin complex. *Mol Cell,* 9**,** 773-88.

HAGLUND, K. & DIKIC, I. 2005. Ubiquitylation and cell signaling. *EMBO J,* 24**,** 3353-9.

HALLETT, S. T., HARRY, I. C., SCHELLENBERGER, P., ZHOU, L., CRONIN, N. B., BAXTER, J., ETHERIDGE, T. J., MURRAY, J. M. & OLIVER, A. W. 2021a. Cryo-EM structure of the Smc5/6 holo-complex. *bioRxiv*, 2021.11.25.470006.

HALLETT, S. T., SCHELLENBERGER, P., ZHOU, L., BEURON, F., MORRIS, E., MURRAY, J. M. & OLIVER, A. W. 2021b. Nse5/6 is a negative regulator of the ATPase activity of the Smc5/6 complex. *Nucleic Acids Res,* 49**,** 4534-4549.

HAMPEL, H., FRANKEL, W. L., MARTIN, E., ARNOLD, M., KHANDUJA, K., KUEBLER, P., NAKAGAWA, H., SOTAMAA, K., PRIOR, T. W., WESTMAN, J., PANESCU, J., FIX, D., LOCKMAN, J., COMERAS, I. & DE LA CHAPELLE, A. 2005. Screening for the Lynch syndrome (hereditary nonpolyposis colorectal cancer). *N Engl J Med,* 352**,** 1851-60.

HAMPERL, S., BOCEK, M. J., SALDIVAR, J. C., SWIGUT, T. & CIMPRICH, K. A. 2017. Transcription-Replication Conflict Orientation Modulates R-Loop Levels and Activates Distinct DNA Damage Responses. *Cell,* 170**,** 774-786 e19.

HAMPERL, S. & CIMPRICH, K. A. 2014. The contribution of co-transcriptional RNA:DNA hybrid structures to DNA damage and genome instability. *DNA Repair (Amst),* 19**,** 84-94.

HAMPERL, S. & CIMPRICH, K. A. 2016. Conflict Resolution in the Genome: How Transcription and Replication Make It Work. *Cell,* 167**,** 1455-1467.

HANADA, K., BUDZOWSKA, M., DAVIES, S. L., VAN DRUNEN, E., ONIZAWA, H., BEVERLOO, H. B., MAAS, A., ESSERS, J., HICKSON, I. D. & KANAAR, R. 2007. The structure-specific endonuclease Mus81 contributes to replication restart by generating double-strand DNA breaks. *Nat Struct Mol Biol,* 14**,** 1096-104.

HANKS, S., COLEMAN, K., REID, S., PLAJA, A., FIRTH, H., FITZPATRICK, D., KIDD, A., MEHES, K., NASH, R., ROBIN, N., SHANNON, N., TOLMIE, J., SWANSBURY, J., IRRTHUM, A., DOUGLAS, J. & RAHMAN, N. 2004. Constitutional aneuploidy and cancer predisposition caused by biallelic mutations in BUB1B. *Nat Genet,* 36**,** 1159-61.

HANLON, S. L. & LI, J. J. 2015. Re-replication of a Centromere Induces Chromosomal Instability and Aneuploidy. *PLOS Genetics,* 11**,** e1005039.

HÄNSEL-HERTSCH, R., BERALDI, D., LENSING, S. V., MARSICO, G., ZYNER, K., PARRY, A., DI ANTONIO, M., PIKE, J., KIMURA, H., NARITA, M., TANNAHILL, D. & BALASUBRAMANIAN, S. 2016. G-quadruplex structures mark human regulatory chromatin. *Nature Genetics,* 48**,** 1267-1272.

HARACSKA, L., PRAKASH, S. & PRAKASH, L. 2003. Yeast DNA polymerase zeta is an efficient extender of primer ends opposite from 7,8-dihydro-8-Oxoguanine and O6-methylguanine. *Mol Cell Biol,* 23**,** 1453-9.

HAUF, S., ROITINGER, E., KOCH, B., DITTRICH, C. M., MECHTLER, K. & PETERS, J.-M. 2005. Dissociation of Cohesin from Chromosome Arms and Loss of Arm Cohesion during Early Mitosis Depends on Phosphorylation of SA2. *PLOS Biology,* 3**,** e69.

HAUF, S., WAIZENEGGER, I. C. & PETERS, J.-M. 2001. Cohesin Cleavage by Separase Required for Anaphase and Cytokinesis in Human Cells. *Science,* 293**,** 1320-1323.

HEINTZMAN, D. R., CAMPOS, L. V., BYL, J. A. W., OSHEROFF, N. & DEWAR, J. M. 2019. Topoisomerase II Is Crucial for Fork Convergence during Vertebrate Replication Termination. *Cell Rep,* 29**,** 422-436 e5.

HELLMAN, A., ZLOTORYNSKI, E., SCHERER, S. W., CHEUNG, J., VINCENT, J. B., SMITH, D. I., TRAKHTENBROT, L. & KEREM, B. 2002. A role for common fragile site induction in amplification of human oncogenes. *Cancer Cell,* 1**,** 89-97.

HELMRICH, A., BALLARINO, M. & TORA, L. 2011. Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes. *Mol Cell,* 44**,** 966-77.

HENNINGER, E. E. & PURSELL, Z. F. 2014. DNA polymerase epsilon and its roles in genome stability. *IUBMB Life,* 66**,** 339-51.

HIGGS, M. R., REYNOLDS, J. J., WINCZURA, A., BLACKFORD, A. N., BOREL, V., MILLER, E. S., ZLATANOU, A., NIEMINUSZCZY, J., RYAN, E. L., DAVIES, N. J., STANKOVIC, T., BOULTON, S. J., NIEDZWIEDZ, W. & STEWART, G. S. 2015. BOD1L Is Required to Suppress Deleterious Resection of Stressed Replication Forks. *Mol Cell,* 59**,** 462-77.

HIGGS, M. R. & STEWART, G. S. 2016. Protection or resection: BOD1L as a novel replication fork protection factor. *Nucleus,* 7**,** 34-40.

HINSHAW, S. M., MAKRANTONI, V., KERR, A., MARSTON, A. L. & HARRISON, S. C. 2015. Structural evidence for Scc4-dependent localization of cohesin loading. *eLife,* 4**,** e06057.

HIRANO, M. & HIRANO, T. 2002. Hinge-mediated dimerization of SMC protein is essential for its dynamic interaction with DNA. *EMBO J,* 21**,** 5733-44.

HIRANO, T. 2002. The ABCs of SMC proteins: two-armed ATPases for chromosome condensation, cohesion, and repair. *Genes Dev,* 16**,** 399-414.

HIRANO, T. 2012. Condensins: universal organizers of chromosomes with diverse functions. *Genes Dev,* 26**,** 1659-78.

HODGSON, B., LI, A., TADA, S. & BLOW, J. J. 2002. Geminin becomes activated as an inhibitor of Cdt1/RLF-B following nuclear import. *Curr Biol,* 12**,** 678-83.

HOFMANN, W. K. & KOEFFLER, H. P. 2005. Myelodysplastic syndrome. *Annu Rev Med,* 56**,** 1-16.

HOULARD, M., CUTTS, E. E., SHAMIM, M. S., GODWIN, J., WEISZ, D., PRESSER AIDEN, A., LIEBERMAN AIDEN, E., SCHERMELLEH, L., VANNINI, A. & NASMYTH, K. 2021. MCPH1 inhibits Condensin II during interphase by regulating its SMC2-Kleisin interface. *Elife,* 10.

HOZAK, P. & COOK, P. R. 1994. Replication factories. *Trends Cell Biol,* 4**,** 48-52.

HUANG, J., HUEN, M. S., KIM, H., LEUNG, C. C., GLOVER, J. N., YU, X. & CHEN, J. 2009. RAD18 transmits DNA damage signalling to elicit homologous recombination repair. *Nat Cell Biol,* 11**,** 592-603.

HUANG, J., LIU, S., BELLANI, M. A., THAZHATHVEETIL, A. K., LING, C., DE WINTER, J. P., WANG, Y., WANG, W. & SEIDMAN, M. M. 2013. The DNA translocase FANCM/MHF promotes replication traverse of DNA interstrand crosslinks. *Mol Cell,* 52**,** 434-46.

HUDSON, J. J., BEDNAROVA, K., KOZAKOVA, L., LIAO, C., GUERINEAU, M., COLNAGHI, R., VIDOT, S., MAREK, J., BATHULA, S. R., LEHMANN, A. R. & PALECEK, J. 2011. Interactions between the Nse3 and Nse4 components of the SMC5-6 complex identify evolutionarily conserved interactions between MAGE and EID Families. *PLoS One,* 6**,** e17270.

HUERTAS, P. & AGUILERA, A. 2003. Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Mol Cell,* 12**,** 711-21.

HUNG, K. F., SIDOROVA, J. M., NGHIEM, P. & KAWASUMI, M. 2020. The 6-4 photoproduct is the trigger of UV-induced replication blockage and ATR activation. *Proc Natl Acad Sci U S A,* 117**,** 12806-12816.

HUPPERT, J. L. & BALASUBRAMANIAN, S. 2007. G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res,* 35**,** 406-13.

IBARRA, A., SCHWOB, E. & MENDEZ, J. 2008. Excess MCM proteins protect human cells from replicative stress by licensing backup origins of replication. *Proc Natl Acad Sci U S A,* 105**,** 8956-61.

ILVES, I., PETOJEVIC, T., PESAVENTO, J. J. & BOTCHAN, M. R. 2010. Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins. *Mol Cell,* 37**,** 247-58.

IM, J. S., KI, S. H., FARINA, A., JUNG, D. S., HURWITZ, J. & LEE, J. K. 2009. Assembly of the Cdc45-Mcm2-7-GINS complex in human cells requires the Ctf4/And-1, RecQL4, and Mcm10 proteins. *Proc Natl Acad Sci U S A,* 106**,** 15628-32.

INOMATA, Y., ABE, T., TSUDA, M., TAKEDA, S. & HIROTA, K. 2021. Division of labor of Y-family polymerases in translesion-DNA synthesis for distinct types of DNA damage. *PLoS One,* 16**,** e0252587.

IRONY-TUR SINAI, M., SALAMON, A., STANLEIGH, N., GOLDBERG, T., WEISS, A., WANG, Y. H. & KEREM, B. 2019. AT-dinucleotide rich sequences drive fragile site formation. *Nucleic Acids Res,* 47**,** 9685-9695.

JABS, E. W., TUCK-MULLER, C. M., CUSANO, R. & RATTNER, J. B. 1991. Studies of mitotic and centromeric abnormalities in Roberts syndrome: implications for a defect in the mitotic mechanism. *Chromosoma,* 100**,** 251-61.

JACKSON, A. P., MCHALE, D. P., CAMPBELL, D. A., JAFRI, H., RASHID, Y., MANNAN, J., KARBANI, G., CORRY, P., LEVENE, M. I., MUELLER, R. F., MARKHAM, A. F., LENCH, N. J. & WOODS, C. G. 1998. Primary Autosomal Recessive Microcephaly (MCPH1) Maps to Chromosome 8p22-pter. *The American Journal of Human Genetics,* 63**,** 541-546.

JACKSON, S. P. 2002. Sensing and repairing DNA double-strand breaks. *Carcinogenesis,* 23**,** 687-96.

JACKSON, S. P. & BARTEK, J. 2009. The DNA-damage response in human biology and disease. *Nature,* 461**,** 1071-8.

JACOME, A., GUTIERREZ-MARTINEZ, P., SCHIAVONI, F., TENAGLIA, E., MARTINEZ, P., RODRIGUEZ-ACEBES, S., LECONA, E., MURGA, M., MENDEZ, J., BLASCO, M. A. & FERNANDEZ-CAPETILLO, O. 2015. NSMCE2 suppresses cancer and aging in mice independently of its SUMO ligase activity. *EMBO J,* 34**,** 2604-19.

JALAN, M., OEHLER, J., MORROW, C. A., OSMAN, F. & WHITBY, M. C. 2019. Factors affecting template switch recombination associated with restarted DNA replication. *Elife,* 8.

JANSSON, L. I., HENTSCHEL, J., PARKS, J. W., CHANG, T. R., LU, C., BARAL, R., BAGSHAW, C. R. & STONE, M. D. 2019. Telomere DNA G-quadruplex folding within actively extending human telomerase. *Proceedings of the National Academy of Sciences,* 116**,** 9350-9359.

JAYARAMAN, D., BAE, B.-I. & WALSH, C. A. 2018. The Genetics of Primary Microcephaly. *Annual Review of Genomics and Human Genetics,* 19**,** 177-200.

JEGGO, P. A. & LOBRICH, M. 2007. DNA double-strand breaks: their cellular and clinical impact? *Oncogene,* 26**,** 7717-9.

JEPPSSON, K., CARLBORG, K. K., NAKATO, R., BERTA, D. G., LILIENTHAL, I., KANNO, T., LINDQVIST, A., BRINK, M. C., DANTUMA, N. P., KATOU, Y., SHIRAHIGE, K. & SJOGREN, C. 2014. The chromosomal association of the Smc5/6 complex depends on cohesion and predicts the level of sister chromatid entanglement. *PLoS Genet,* 10**,** e1004680.

JOHANSSON, E. & MACNEILL, S. A. 2010. The eukaryotic replicative DNA polymerases take shape. *Trends Biochem Sci,* 35**,** 339-47.

JOHNSON, J. E., CAO, K., RYVKIN, P., WANG, L. S. & JOHNSON, F. B. 2010. Altered gene expression in the Werner and Bloom syndromes is associated with sequences having G-quadruplex forming potential. *Nucleic Acids Res,* 38**,** 1114-22.

JOHNSON, R. E., HARACSKA, L., PRAKASH, S. & PRAKASH, L. 2001. Role of DNA polymerase eta in the bypass of a (6-4) TT photoproduct. *Mol Cell Biol,* 21**,** 3558-63.

JONES, M. L., BARIS, Y., TAYLOR, M. R. G. & YEELES, J. T. P. 2021. Structure of a human replisome shows the organisation and interactions of a DNA replication machine. *EMBO J,* 40**,** e108819.

JU, L., WING, J., TAYLOR, E., BRANDT, R., SLIJEPCEVIC, P., HORSCH, M., RATHKOLB, B., RACZ, I., BECKER, L., HANS, W., ADLER, T., BECKERS, J., ROZMAN, J., KLINGENSPOR, M., WOLF, E., ZIMMER, A., KLOPSTOCK, T., BUSCH, D. H., GAILUS-DURNER, V., FUCHS, H., DE ANGELIS, M. H., VAN DER HORST, G. & LEHMANN, A. R. 2013. SMC6 is an essential gene in mice, but a hypomorphic mutant in the ATPase domain has a mild phenotype with a range of subtle abnormalities. *DNA Repair (Amst),* 12**,** 356-66.

JURIKOVA, K., GAJARSKY, M., HAJIKAZEMI, M., NOSEK, J., PROCHAZKOVA, K., PAESCHKE, K., TRANTIREK, L. & TOMASKA, L. 2020. Role of folding kinetics of secondary structures in telomeric G-overhangs in the regulation of telomere maintenance in Saccharomyces cerevisiae. *J Biol Chem,* 295**,** 8958-8971.

KAGAMI, Y., ONO, M. & YOSHIDA, K. 2017. Plk1 phosphorylation of CAP-H2 triggers chromosome condensation by condensin II at the early phase of mitosis. *Sci Rep,* 7**,** 5583.

KAGEY, M. H., NEWMAN, J. J., BILODEAU, S., ZHAN, Y., ORLANDO, D. A., VAN BERKUM, N. L., EBMEIER, C. C., GOOSSENS, J., RAHL, P. B., LEVINE, S. S., TAATJES, D. J., DEKKER, J. & YOUNG, R. A. 2010. Mediator and cohesin connect gene expression and chromatin architecture. *Nature,* 467**,** 430-5.

KAJII, T., IKEUCHI, T., YANG, Z. Q., NAKAMURA, Y., TSUJI, Y., YOKOMORI, K., KAWAMURA, M., FUKUDA, S., HORITA, S. & ASAMOTO, A. 2001. Cancer-prone syndrome of mosaic variegated aneuploidy and total premature chromatid separation: report of five infants. *Am J Med Genet,* 104**,** 57-64.

KANNO, T., BERTA, D. G. & SJOGREN, C. 2015. The Smc5/6 Complex Is an ATP-Dependent Intermolecular DNA Linker. *Cell Rep,* 12**,** 1471-82.

KARNANI, N. & DUTTA, A. 2011. The effect of the intra-S-phase checkpoint on origins of replication in human cells. *Genes Dev,* 25**,** 621-33.

KAROW, J. K., CONSTANTINOU, A., LI, J. L., WEST, S. C. & HICKSON, I. D. 2000. The Bloom's syndrome gene product promotes branch migration of holliday junctions. *Proc Natl Acad Sci U S A,* 97**,** 6504-8.

KAUR, M., DESCIPIO, C., MCCALLUM, J., YAEGER, D., DEVOTO, M., JACKSON, L. G., SPINNER, N. B. & KRANTZ, I. D. 2005. Precocious sister chromatid separation (PSCS) in Cornelia de Lange syndrome. *Am J Med Genet A,* 138**,** 27-31.

KAUSHAL, S. & FREUDENREICH, C. H. 2019. The role of fork stalling and DNA structures in causing chromosome fragility. *Genes Chromosomes Cancer,* 58**,** 270-283.

KAUSHAL, S., WOLLMUTH, C. E., DAS, K., HILE, S. E., REGAN, S. B., BARNES, R. P., HAOUZI, A., LEE, S. M., HOUSE, N. C. M., GUYUMDZHYAN, M., ECKERT, K. A. & FREUDENREICH, C. H. 2019. Sequence and Nuclease Requirements for Breakage and Healing of a Structure-Forming (AT)n Sequence within Fragile Site FRA16D. *Cell Rep,* 27**,** 1151-1164 e5.

KAWASUMI, R., ABE, T., ARAKAWA, H., GARRE, M., HIROTA, K. & BRANZEI, D. 2017. ESCO1/2's roles in chromosome structure and interphase chromatin organization. *Genes Dev,* 31**,** 2136-2150.

KAWAUCHI, S., SANTOS, R., MUTO, A., LOPEZ-BURKS, M. E., SCHILLING, T. F., LANDER, A. D. & CALOF, A. L. 2016. Using mouse and zebrafish models to understand the etiology of developmental defects in Cornelia de Lange Syndrome. *Am J Med Genet C Semin Med Genet,* 172**,** 138-45.

KAYKOV, A. & NURSE, P. 2015. The spatial and temporal organization of origin firing during the S-phase of fission yeast. *Genome Research,* 25**,** 391-401.

KELLY, R. L., HUEHLS, A. M., VENKATACHALAM, A., HUNTOON, C. J., MACHIDA, Y. J. & KARNITZ, L. M. 2022. Intra-S phase checkpoint kinase Chk1 dissociates replication proteins Treslin and TopBP1 through multiple mechanisms during replication stress. *J Biol Chem,* 298**,** 101777.

KERNS, S. L., TORKE, S. J., BENJAMIN, J. M. & MCGARRY, T. J. 2007. Geminin prevents rereplication during xenopus development. *J Biol Chem,* 282**,** 5514-21.

KESZTHELYI, A., MINCHELL, N. E. & BAXTER, J. 2016. The Causes and Consequences of Topological Stress during DNA Replication. *Genes (Basel),* 7.

KHAN, T. N., KHAN, K., SADEGHPOUR, A., REYNOLDS, H., PERILLA, Y., MCDONALD, M. T., GALLENTINE, W. B., BAIG, S. M., ALLORI, A., ANGRIST, M., ASHLEY, P., BIDEGAIN, M., BOYD, B., CHAMBERS, E., COPE, H., COTTEN, C. M., CURINGTON, T., DAVIS, E. E., ELLESTAD, S., FISHER, K., FRENCH, A., GALLENTINE, W., GOLDBERG, R., HILL, K., KANSAGRA, S., KATSANIS, N., KATSANIS, S., KURTZBERG, J., MARCUS, J., MCDONALD, M., MIKATI, M., MILLER, S., MURTHA, A., PERILLA, Y., PIZOLI, C., PURVES, T., ROSS, S., SADEGHPOUR, A., SMITH, E., WIENER, J., DAVIS, E. E. & KATSANIS, N. 2019. Mutations in NCAPG2 Cause a Severe

Neurodevelopmental Syndrome that Expands the Phenotypic Spectrum of Condensinopathies. *The American Journal of Human Genetics,* 104**,** 94-111.
KIM, H. & D'ANDREA, A. D. 2012. Regulation of DNA cross-link repair by the Fanconi anemia/BRCA pathway. *Genes Dev,* 26**,** 1393-408.
KIM, J. M., KEE, Y., GURTAN, A. & D'ANDREA, A. D. 2008. Cell cycle-dependent chromatin loading of the Fanconi anemia core complex by FANCM/FAAP24. *Blood,* 111**,** 5215-22.
KIM, M. Y., VANKAYALAPATI, H., SHIN-YA, K., WIERZBA, K. & HURLEY, L. H. 2002. Telomestatin, a potent telomerase inhibitor that interacts quite specifically with the human telomeric intramolecular g-quadruplex. *J Am Chem Soc,* 124**,** 2098-9.
KIM, N. 2019. The Interplay between G-quadruplex and Transcription. *Curr Med Chem,* 26**,** 2898-2917.
KIM, Y., SPITZ, G. S., VETURI, U., LACH, F. P., AUERBACH, A. D. & SMOGORZEWSKA, A. 2013. Regulation of multiple DNA repair pathways by the Fanconi anemia protein SLX4. *Blood,* 121**,** 54-63.
KLEIN DOUWEL, D., BOONEN, R. A., LONG, D. T., SZYPOWSKA, A. A., RASCHLE, M., WALTER, J. C. & KNIPSCHEER, P. 2014. XPF-ERCC1 acts in Unhooking DNA interstrand crosslinks in cooperation with FANCD2 and FANCP/SLX4. *Mol Cell,* 54**,** 460-71.
KLINE, A. D., MOSS, J. F., SELICORNI, A., BISGAARD, A.-M., DEARDORFF, M. A., GILLETT, P. M., ISHMAN, S. L., KERR, L. M., LEVIN, A. V., MULDER, P. A., RAMOS, F. J., WIERZBA, J., AJMONE, P. F., AXTELL, D., BLAGOWIDOW, N., CEREDA, A., COSTANTINO, A., CORMIER-DAIRE, V., FITZPATRICK, D., GRADOS, M., GROVES, L., GUTHRIE, W., HUISMAN, S., KAISER, F. J., KOEKKOEK, G., LEVIS, M., MARIANI, M., MCCLEERY, J. P., MENKE, L. A., METRENA, A., O'CONNOR, J., OLIVER, C., PIE, J., PIENING, S., POTTER, C. J., QUAGLIO, A. L., REDEKER, E., RICHMAN, D., RIGAMONTI, C., SHI, A., TÜMER, Z., VAN BALKOM, I. D. C. & HENNEKAM, R. C. 2018. Diagnosis and management of Cornelia de Lange syndrome: first international consensus statement. *Nature Reviews Genetics,* 19**,** 649-666.
KLISZCZAK, M., STEPHAN, A. K., FLANAGAN, A. M. & MORRISON, C. G. 2012. SUMO ligase activity of vertebrate Mms21/Nse2 is required for efficient DNA repair but not for Smc5/6 complex stability. *DNA Repair (Amst),* 11**,** 799-810.
KLOTZ-NOACK, K., MCINTOSH, D., SCHURCH, N., PRATT, N. & BLOW, J. J. 2012. Re-replication induced by geminin depletion occurs from G2 and is enhanced by checkpoint activation. *J Cell Sci,* 125**,** 2436-45.
KNOUSE, K. A., DAVOLI, T., ELLEDGE, S. J. & AMON, A. 2017. Aneuploidy in Cancer: Seq-ing Answers to Old Questions. *Annual Review of Cancer Biology,* 1**,** 335-354.
KOBAYASHI, K., GUILLIAM, T. A., TSUDA, M., YAMAMOTO, J., BAILEY, L. J., IWAI, S., TAKEDA, S., DOHERTY, A. J. & HIROTA, K. 2016. Repriming by PrimPol is critical for DNA replication restart downstream of lesions and chain-terminating nucleosides. *Cell Cycle,* 15**,** 1997-2008.
KOBAYASHI, T. 2003. The replication fork barrier site forms a unique structure with Fob1p and inhibits the replication fork. *Mol Cell Biol,* 23**,** 9178-88.
KOBAYASHI, T., HECK, D. J., NOMURA, M. & HORIUCHI, T. 1998. Expansion and contraction of ribosomal DNA repeats in Saccharomyces cerevisiae: requirement of replication fork blocking (Fob1) protein and the role of RNA polymerase I. *Genes Dev,* 12**,** 3821-30.
KOCHENOVA, O. V., DAEE, D. L., MERTZ, T. M. & SHCHERBAKOVA, P. V. 2015. DNA polymerase ζ-dependent lesion bypass in Saccharomyces cerevisiae is accompanied by error-prone copying of long stretches of adjacent DNA. *PLoS Genet,* 11**,** e1005110.
KOLESAR, P., STEJSKAL, K., POTESIL, D., MURRAY, J. M. & PALECEK, J. J. 2022. Role of Nse1 Subunit of SMC5/6 Complex as a Ubiquitin Ligase. *Cells,* 11.
KOLESNIKOVA, S. & CURTIS, E. A. 2019. Structure and Function of Multimeric G-Quadruplexes. *Molecules,* 24.

KOLINJIVADI, A. M., SANNINO, V., DE ANTONI, A., ZADOROZHNY, K., KILKENNY, M., TECHER, H., BALDI, G., SHEN, R., CICCIA, A., PELLEGRINI, L., KREJCI, L. & COSTANZO, V. 2017. Smarcal1-Mediated Fork Reversal Triggers Mre11-Dependent Degradation of Nascent DNA in the Absence of Brca2 and Stable Rad51 Nucleofilaments. *Mol Cell,* 67**,** 867-881 e7.

KOTSANTIS, P., SILVA, L. M., IRMSCHER, S., JONES, R. M., FOLKES, L., GROMAK, N. & PETERMANN, E. 2016. Increased global transcription activity as a mechanism of replication stress in cancer. *Nat Commun,* 7**,** 13087.

KOTTEMANN, M. C. & SMOGORZEWSKA, A. 2013. Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. *Nature,* 493**,** 356-63.

KOUNDRIOUKOFF, S., CARIGNON, S., TECHER, H., LETESSIER, A., BRISON, O. & DEBATISSE, M. 2013. Stepwise activation of the ATR signaling pathway upon increasing replication stress impacts fragile site integrity. *PLoS Genet,* 9**,** e1003643.

KRAMARA, J., OSIA, B. & MALKOVA, A. 2018. Break-Induced Replication: The Where, The Why, and The How. *Trends Genet,* 34**,** 518-531.

KREJCI, L., ALTMANNOVA, V., SPIREK, M. & ZHAO, X. 2012. Homologous recombination and its regulation. *Nucleic Acids Res,* 40**,** 5795-818.

KRUISSELBRINK, E., GURYEV, V., BROUWER, K., PONTIER, D. B., CUPPEN, E. & TIJSTERMAN, M. 2008. Mutagenic Capacity of Endogenous G4 DNA Underlies Genome Instability in FANCJ-Defective C. elegans. *Current Biology,* 18**,** 900-905.

KUKOLJ, E., KAUFMANN, T., DICK, A. E., ZEILLINGER, R., GERLICH, D. W. & SLADE, D. 2017. PARP inhibition causes premature loss of cohesion in cancer cells. *Oncotarget,* 8**,** 103931-103951.

KUMAGAI, A. & DUNPHY, W. G. 2017. MTBP, the partner of Treslin, contains a novel DNA-binding domain that is essential for proper initiation of DNA replication. *Mol Biol Cell,* 28**,** 2998-3012.

KUMAGAI, A., LEE, J., YOO, H. Y. & DUNPHY, W. G. 2006. TopBP1 activates the ATR-ATRIP complex. *Cell,* 124**,** 943-55.

KUMAGAI, A., SHEVCHENKO, A., SHEVCHENKO, A. & DUNPHY, W. G. 2010. Treslin collaborates with TopBP1 in triggering the initiation of DNA replication. *Cell,* 140**,** 349-59.

KUNKEL, T. A. 2009. Evolving views of DNA replication (in)fidelity. *Cold Spring Harb Symp Quant Biol,* 74**,** 91-101.

KUZMINOV, A. 2018. When DNA Topology Turns Deadly - RNA Polymerases Dig in Their R-Loops to Stand Their Ground: New Positive and Negative (Super)Twists in the Replication-Transcription Conflict. *Trends Genet,* 34**,** 111-120.

LABUN, K., MONTAGUE, T. G., KRAUSE, M., TORRES CLEUREN, Y. N., TJELDNES, H. & VALEN, E. 2019. CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res,* 47**,** W171-W174.

LADURNER, R., KREIDL, E., IVANOV, M. P., EKKER, H., IDARRAGA-AMADO, M. H., BUSSLINGER, G. A., WUTZ, G., CISNEROS, D. A. & PETERS, J.-M. 2016. Sororin actively maintains sister chromatid cohesion. *The EMBO Journal,* 35**,** 635-653.

LALONDE, M., TRAUNER, M., WERNER, M. & HAMPERL, S. 2021. Consequences and Resolution of Transcription-Replication Conflicts. *Life (Basel),* 11.

LALOR, L., DAVIES, O. M. T., BASEL, D. & SIEGEL, D. H. 2020. Cafe au lait spots: When and how to pursue their genetic origins. *Clin Dermatol,* 38**,** 421-431.

LAMMENS, A., SCHELE, A. & HOPFNER, K. P. 2004. Structural biochemistry of ATP-driven dimerization and DNA-stimulated activation of SMC ATPases. *Curr Biol,* 14**,** 1778-82.

LANGSTON, L. D., ZHANG, D., YURIEVA, O., GEORGESCU, R. E., FINKELSTEIN, J., YAO, N. Y., INDIANI, C. & O'DONNELL, M. E. 2014. CMG helicase and DNA polymerase epsilon form a functional 15-subunit holoenzyme for eukaryotic leading-strand DNA replication. *Proc Natl Acad Sci U S A,* 111**,** 15390-5.

LAROCQUE, J. R., STARK, J. M., OH, J., BOJILOVA, E., YUSA, K., HORIE, K., TAKEDA, J. & JASIN, M. 2011. Interhomolog recombination and loss of heterozygosity in wild-type and Bloom syndrome helicase (BLM)-deficient mammalian cells. *Proc Natl Acad Sci U S A,* 108**,** 11971-6.

LEE, B.-G., RHODES, J. & LÖWE, J. 2022. Clamping of DNA shuts the condensin neck gate. *Proceedings of the National Academy of Sciences,* 119**,** e2120006119.

LEE, B. G., MERKEL, F., ALLEGRETTI, M., HASSLER, M., CAWOOD, C., LECOMTE, L., O'REILLY, F. J., SINN, L. R., GUTIERREZ-ESCRIBANO, P., KSCHONSAK, M., BRAVO, S., NAKANE, T., RAPPSILBER, J., ARAGON, L., BECK, M., LOWE, J. & HAERING, C. H. 2020. Cryo-EM structures of holo condensin reveal a subunit flip-flop mechanism. *Nat Struct Mol Biol,* 27**,** 743-751.

LEE, W. T. C., YIN, Y., MORTEN, M. J., TONZI, P., GWO, P. P., ODERMATT, D. C., MODESTI, M., CANTOR, S. B., GARI, K., HUANG, T. T. & ROTHENBERG, E. 2021. Single-molecule imaging reveals replication fork coupled formation of G-quadruplex structures hinders local replication stress signaling. *Nature Communications,* 12**,** 2525.

LEHMANN, A. R., WALICKA, M., GRIFFITHS, D. J., MURRAY, J. M., WATTS, F. Z., MCCREADY, S. & CARR, A. M. 1995. The rad18 gene of Schizosaccharomyces pombe defines a new subgroup of the SMC superfamily involved in DNA repair. *Mol Cell Biol,* 15**,** 7067-80.

LEMACON, D., JACKSON, J., QUINET, A., BRICKNER, J. R., LI, S., YAZINSKI, S., YOU, Z., IRA, G., ZOU, L., MOSAMMAPARAST, N. & VINDIGNI, A. 2017. MRE11 and EXO1 nucleases degrade reversed forks and elicit MUS81-dependent fork rescue in BRCA2-deficient cells. *Nat Commun,* 8**,** 860.

LEMAN, A. R. & NOGUCHI, E. 2013. The replication fork: understanding the eukaryotic replication machinery and the challenges to genome duplication. *Genes (Basel),* 4**,** 1-32.

LEMMENS, B., VAN SCHENDEL, R. & TIJSTERMAN, M. 2015. Mutagenic consequences of a single G-quadruplex demonstrate mitotic inheritance of DNA replication fork barriers. *Nat Commun,* 6**,** 8909.

LEPPARD, J. B. & CHAMPOUX, J. J. 2005. Human DNA topoisomerase I: relaxation, roles, and damage control. *Chromosoma,* 114**,** 75-85.

LERNER, L. K. & SALE, J. E. 2019. Replication of G Quadruplex DNA. *Genes (Basel),* 10.

LI, H., YANG, Y., HONG, W., HUANG, M., WU, M. & ZHAO, X. 2020a. Applications of genome editing technology in the targeted therapy of human diseases: mechanisms, advances and prospects. *Signal Transduct Target Ther,* 5**,** 1.

LI, K., LUO, H., HUANG, L., LUO, H. & ZHU, X. 2020b. Microsatellite instability: a review of what the oncologist should know. *Cancer Cell Int,* 20**,** 16.

LI, L., TAN, W. & DEANS, A. J. 2020c. Structural insight into FANCI-FANCD2 monoubiquitination. *Essays Biochem,* 64**,** 807-817.

LI, S. & WU, X. 2020. Common fragile sites: protection and repair. *Cell Biosci,* 10**,** 29.

LI, X. & HEYER, W. D. 2008. Homologous recombination in DNA repair and DNA damage tolerance. *Cell Res,* 18**,** 99-113.

LIANG, Z., LIANG, F., TENG, Y., CHEN, X., LIU, J., LONGERICH, S., RAO, T., GREEN, A. M., COLLINS, N. B., XIONG, Y., LAN, L., SUNG, P. & KUPFER, G. M. 2019. Binding of FANCI-FANCD2 Complex to RNA and R-Loops Stimulates Robust FANCD2 Monoubiquitination. *Cell Rep,* 26**,** 564-572 e5.

LIEBER, M. R. 2010. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev Biochem,* 79**,** 181-211.

LIM, C. J. & CECH, T. R. 2021. Publisher Correction: Shaping human telomeres: from shelterin and CST complexes to telomeric chromatin organization. *Nat Rev Mol Cell Biol,* 22**,** 299.

LIMAS, J. C. & COOK, J. G. 2019. Preparation for DNA replication: the key to a successful S phase. *FEBS Lett,* 593**,** 2853-2867.

LINDAHL, T. 1993. Instability and decay of the primary structure of DNA. *Nature,* 362**,** 709-15.

LINDER, M. I., KÖHLER, M., BOERSEMA, P., WEBERRUSS, M., WANDKE, C., MARINO, J., ASHIONO, C., PICOTTI, P., ANTONIN, W. & KUTAY, U. 2017. Mitotic Disassembly of Nuclear Pore Complexes Involves CDK1- and PLK1-Mediated Phosphorylation of Key Interconnecting Nucleoporins. *Developmental Cell,* 43**,** 141-156.e7.

LIPP, J. J., HIROTA, T., POSER, I. & PETERS, J. M. 2007. Aurora B controls the association of condensin I but not condensin II with mitotic chromosomes. *J Cell Sci,* 120**,** 1245-55.

LIU, B., HU, J., WANG, J. & KONG, D. 2017. Direct Visualization of RNA-DNA Primer Removal from Okazaki Fragments Provides Support for Flap Cleavage and Exonucleolytic Pathways in Eukaryotic Cells. *J Biol Chem,* 292**,** 4777-4788.

LIU, E., LEE, A. Y., CHIBA, T., OLSON, E., SUN, P. & WU, X. 2007. The ATR-mediated S phase checkpoint prevents rereplication in mammalian cells when licensing control is disrupted. *J Cell Biol,* 179**,** 643-57.

LIU, J., DOTY, T., GIBSON, B. & HEYER, W. D. 2010. Human BRCA2 protein promotes RAD51 filament formation on RPA-covered single-stranded DNA. *Nat Struct Mol Biol,* 17**,** 1260-2.

LIU, J., EHMSEN, K. T., HEYER, W. D. & MORRICAL, S. W. 2011a. Presynaptic filament dynamics in homologous recombination and DNA repair. *Crit Rev Biochem Mol Biol,* 46**,** 240-70.

LIU, J., ZHANG, Z., BANDO, M., ITOH, T., DEARDORFF, M. A., CLARK, D., KAUR, M., TANDY, S., KONDOH, T., RAPPAPORT, E., SPINNER, N. B., VEGA, H., JACKSON, L. G., SHIRAHIGE, K. & KRANTZ, I. D. 2009. Transcriptional dysregulation in NIPBL and cohesin mutant human cells. *PLoS Biol,* 7**,** e1000119.

LIU, S., SHIOTANI, B., LAHIRI, M., MARECHAL, A., TSE, A., LEUNG, C. C., GLOVER, J. N., YANG, X. H. & ZOU, L. 2011b. ATR autophosphorylation as a molecular switch for checkpoint activation. *Mol Cell,* 43**,** 192-202.

LIU, S. T. & ZHANG, H. 2016. The mitotic checkpoint complex (MCC): looking back and forth after 15 years. *AIMS Mol Sci,* 3**,** 597-634.

LIU, Y., WANG, L., XU, X., YUAN, Y., ZHANG, B., LI, Z., XIE, Y., YAN, R., ZHENG, Z., JI, J., MURRAY, J. M., CARR, A. M. & KONG, D. 2021. The intra-S phase checkpoint directly regulates replication elongation to preserve the integrity of stalled replisomes. *Proc Natl Acad Sci U S A,* 118.

LONDON, T. B., BARBER, L. J., MOSEDALE, G., KELLY, G. P., BALASUBRAMANIAN, S., HICKSON, I. D., BOULTON, S. J. & HIOM, K. 2008. FANCJ is a structure-specific DNA helicase associated with the maintenance of genomic G/C tracts. *J Biol Chem,* 283**,** 36132-9.

LONGERICH, S., SAN FILIPPO, J., LIU, D. & SUNG, P. 2009. FANCI binds branched DNA and is monoubiquitinated by UBE2T-FANCL. *J Biol Chem,* 284**,** 23182-6.

LOPEZ-BERGAMI, P. & RONAI, Z. 2011. Cell Death in Response to Genotoxic Stress and DNA Damage. *In:* REED, J. C. (ed.) *Apoptosis Physiology and Pathology.* Cambridge University Press.

LOPEZ DE SILANES, I., GRANA, O., DE BONIS, M. L., DOMINGUEZ, O., PISANO, D. G. & BLASCO, M. A. 2014. Identification of TERRA locus unveils a telomere protection role through association to nearly all chromosomes. *Nat Commun,* 5**,** 4723.

LOWE, J., CORDELL, S. C. & VAN DEN ENT, F. 2001. Crystal structure of the SMC head domain: an ABC ATPase with 900 residues antiparallel coiled-coil inserted. *J Mol Biol,* 306**,** 25-35.

LUCAS, I., GERME, T., CHEVRIER-MILLER, M. & HYRIEN, O. 2001. Topoisomerase II can unlink replicating DNA by precatenane removal. *EMBO J,* 20**,** 6509-19.

LUCCA, C., VANOLI, F., COTTA-RAMUSINO, C., ARAKI, H. & FOIANI, M. 2003. The Rad53-mediated checkpoint response is required to stabilize replisome-fork association when DNA replication is pausing. *Yeast,* 20**,** S104-S104.

LUCCA, C., VANOLI, F., COTTA-RAMUSINO, C., PELLICIOLI, A., LIBERI, G., HABER, J. & FOIANI, M. 2004. Checkpoint-mediated control of replisome-fork association and signalling in response to replication pausing. *Oncogene,* 23**,** 1206-13.

LUJAN, S. A., WILLIAMS, J. S. & KUNKEL, T. A. 2016. DNA Polymerases Divide the Labor of Genome Replication. *Trends Cell Biol,* 26**,** 640-654.
LUO, S. & TONG, L. 2021. Structure and Function of the Separase-Securin Complex. *Subcell Biochem,* 96**,** 217-232.
MA, HOI T. & POON, RANDY YAT C. 2016. TRIP13 Regulates Both the Activation and Inactivation of the Spindle-Assembly Checkpoint. *Cell Reports,* 14**,** 1086-1099.
MA, Y., PANNICKE, U., SCHWARZ, K. & LIEBER, M. R. 2002. Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. *Cell,* 108**,** 781-94.
MA, Y., SCHWARZ, K. & LIEBER, M. R. 2005. The Artemis:DNA-PKcs endonuclease cleaves DNA loops, flaps, and gaps. *DNA Repair (Amst),* 4**,** 845-51.
MACDOUGALL, C. A., BYUN, T. S., VAN, C., YEE, M. C. & CIMPRICH, K. A. 2007. The structural determinants of checkpoint activation. *Genes Dev,* 21**,** 898-903.
MACE, E. M., PAUST, S., CONTE, M. I., BAXLEY, R. M., SCHMIT, M. M., PATIL, S. L., GUILZ, N. C., MUKHERJEE, M., PEZZI, A. E., CHMIELOWIEC, J., TATINENI, S., CHINN, I. K., AKDEMIR, Z. C., JHANGIANI, S. N., MUZNY, D. M., STRAY-PEDERSEN, A., BRADLEY, R. E., MOODY, M., CONNOR, P. P., HEAPS, A. G., STEWARD, C., BANERJEE, P. P., GIBBS, R. A., BOROWIAK, M., LUPSKI, J. R., JOLLES, S., BIELINSKY, A. K. & ORANGE, J. S. 2020. Human NK cell deficiency as a result of biallelic mutations in MCM10. *J Clin Invest,* 130**,** 5272-5286.
MACHERET, M. & HALAZONETIS, T. D. 2015. DNA replication stress as a hallmark of cancer. *Annu Rev Pathol,* 10**,** 425-48.
MACHWE, A., XIAO, L., GRODEN, J. & ORREN, D. K. 2006. The Werner and Bloom syndrome proteins catalyze regression of a model replication fork. *Biochemistry,* 45**,** 13939-46.
MAILAND, N., PODTELEJNIKOV, A. V., GROTH, A., MANN, M., BARTEK, J. & LUKAS, J. 2002. Regulation of G(2)/M events by Cdc25A through phosphorylation-dependent modulation of its stability. *Embo j,* 21**,** 5911-20.
MAKRANTONI, V. & MARSTON, A. L. 2018. Cohesin and chromosome segregation. *Curr Biol,* 28**,** R688-R693.
MALI, P., AACH, J., STRANGES, P. B., ESVELT, K. M., MOOSBURNER, M., KOSURI, S., YANG, L. & CHURCH, G. M. 2013. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol,* 31**,** 833-8.
MALKOVA, A. & IRA, G. 2013. Break-induced replication: functions and molecular mechanism. *Curr Opin Genet Dev,* 23**,** 271-9.
MANTHEI, K. A. & KECK, J. L. 2013. The BLM dissolvasome in DNA replication and repair. *Cell Mol Life Sci,* 70**,** 4067-84.
MAO, Y., DESAI, S. D., TING, C. Y., HWANG, J. & LIU, L. F. 2001. 26 S proteasome-mediated degradation of topoisomerase II cleavable complexes. *J Biol Chem,* 276**,** 40652-8.
MARESCA, T. J. 2011. Cell division: aurora B illuminates a checkpoint pathway. *Curr Biol,* 21**,** R557-9.
MARTIN, C. A., MURRAY, J. E., CARROLL, P., LEITCH, A., MACKENZIE, K. J., HALACHEV, M., FETIT, A. E., KEITH, C., BICKNELL, L. S., FLUTEAU, A., GAUTIER, P., HALL, E. A., JOSS, S., SOARES, G., SILVA, J., BOBER, M. B., DUKER, A., WISE, C. A., QUIGLEY, A. J., PHADKE, S. R., WOOD, A. J., VAGNARELLI, P. & JACKSON, A. P. 2016. Mutations in genes encoding condensin complex proteins cause microcephaly through decatenation failure at mitosis. *Genes Dev,* 30**,** 2158-2172.
MARTINO, L., MORCHOISNE-BOLHY, S., CHEERAMBATHUR, D. K., VAN HOVE, L., DUMONT, J., JOLY, N., DESAI, A., DOYE, V. & PINTARD, L. 2017. Channel Nucleoporins Recruit PLK-1 to Nuclear Pore Complexes to Direct Nuclear Envelope Breakdown in C. elegans. *Developmental Cell,* 43**,** 157-171.e7.
MASAMSETTI, V. P., LOW, R. R. J., MAK, K. S., O'CONNOR, A., RIFFKIN, C. D., LAMM, N., CRABBE, L., KARLSEDER, J., HUANG, D. C. S., HAYASHI, M. T. & CESARE, A. J. 2019. Replication stress

induces mitotic death through parallel pathways regulated by WAPL and telomere deprotection. *Nat Commun,* 10**,** 4224.
MASUD, T., SOONG, C., XU, H., BIELE, J., BJORNSON, S., MCKINNEY, S. & APARICIO, S. 2021. Ubiquitin-mediated DNA damage response is synthetic lethal with G-quadruplex stabilizer CX-5461. *Sci Rep,* 11**,** 9812.
MATMATI, S., LAMBERT, S., GELI, V. & COULON, S. 2020. Telomerase Repairs Collapsed Replication Forks at Telomeres. *Cell Rep,* 30**,** 3312-3322 e3.
MATOS, J. & WEST, S. C. 2014. Holliday junction resolution: regulation in space and time. *DNA Repair (Amst),* 19**,** 176-81.
MATSUURA, S., MATSUMOTO, Y., MORISHIMA, K., IZUMI, H., MATSUMOTO, H., ITO, E., TSUTSUI, K., KOBAYASHI, J., TAUCHI, H., KAJIWARA, Y., HAMA, S., KURISU, K., TAHARA, H., OSHIMURA, M., KOMATSU, K., IKEUCHI, T. & KAJII, T. 2006. Monoallelic BUB1B mutations and defective mitotic-spindle checkpoint in seven families with premature chromatid separation (PCS) syndrome. *Am J Med Genet A,* 140**,** 358-67.
MCALEENAN, A., CORDON-PRECIADO, V., CLEMENTE-BLANCO, A., LIU, I. C., SEN, N., LEONARD, J., JARMUZ, A. & ARAGON, L. 2012. SUMOylation of the alpha-kleisin subunit of cohesin is required for DNA damage-induced cohesion. *Curr Biol,* 22**,** 1564-75.
MCCLENDON, A. K., RODRIGUEZ, A. C. & OSHEROFF, N. 2005. Human topoisomerase IIalpha rapidly relaxes positively supercoiled DNA: implications for enzyme action ahead of replication forks. *J Biol Chem,* 280**,** 39337-45.
MCDONALD, W. H., PAVLOVA, Y., YATES, J. R. & BODDY, M. N. 2003. Novel Essential DNA Repair Proteins Nse1 and Nse2 Are Subunits of the Fission Yeast Smc5-Smc6 Complex*. *Journal of Biological Chemistry,* 278**,** 45460-45467.
MCKIE, S. J., MAXWELL, A. & NEUMAN, K. C. 2020. Mapping DNA Topoisomerase Binding and Cleavage Genome Wide Using Next-Generation Sequencing Techniques. *Genes (Basel),* 11.
MCMURRAY, C. T. 2010. Mechanisms of trinucleotide repeat instability during human development. *Nat Rev Genet,* 11**,** 786-99.
MCVEY, M., LAROCQUE, J. R., ADAMS, M. D. & SEKELSKY, J. J. 2004. Formation of deletions during double-strand break repair in Drosophila DmBlm mutants occurs after strand invasion. *Proc Natl Acad Sci U S A,* 101**,** 15694-9.
MECHALI, M. 2010. Eukaryotic DNA replication origins: many choices for appropriate answers. *Nat Rev Mol Cell Biol,* 11**,** 728-38.
MEHTA, A. & HABER, J. E. 2014. Sources of DNA double-strand breaks and models of recombinational DNA repair. *Cold Spring Harb Perspect Biol,* 6**,** a016428.
MEHTA, G. D., KUMAR, R., SRIVASTAVA, S. & GHOSH, S. K. 2013. Cohesin: Functions beyond sister chromatid cohesion. *FEBS Letters,* 587**,** 2299-2312.
MENG, X. & ZHAO, X. 2017. Replication fork regression and its regulation. *FEMS Yeast Res,* 17.
MENOLFI, D., DELAMARRE, A., LENGRONNE, A., PASERO, P. & BRANZEI, D. 2015. Essential Roles of the Smc5/6 Complex in Replication through Natural Pausing Sites and Endogenous DNA Damage Tolerance. *Mol Cell,* 60**,** 835-46.
MERALDI, P., LUKAS, J., FRY, A. M., BARTEK, J. & NIGG, E. A. 1999. Centrosome duplication in mammalian somatic cells requires E2F and Cdk2-cyclin A. *Nat Cell Biol,* 1**,** 88-93.
MICHAEL, W. M., OTT, R., FANNING, E. & NEWPORT, J. 2000. Activation of the DNA replication checkpoint through RNA synthesis by primase. *Science,* 289**,** 2133-7.
MICHL, J., ZIMMER, J. & TARSOUNAS, M. 2016. Interplay between Fanconi anemia and homologous recombination pathways in genome integrity. *EMBO J,* 35**,** 909-23.
MIJIC, S., ZELLWEGER, R., CHAPPIDI, N., BERTI, M., JACOBS, K., MUTREJA, K., URSICH, S., RAY CHAUDHURI, A., NUSSENZWEIG, A., JANSCAK, P. & LOPES, M. 2017. Replication fork reversal triggers fork degradation in BRCA2-defective cells. *Nat Commun,* 8**,** 859.

MIMURA, S., SEKI, T., TANAKA, S. & DIFFLEY, J. F. 2004. Phosphorylation-dependent binding of mitotic cyclins to Cdc6 contributes to DNA replication control. *Nature,* 431**,** 1118-23.
MINCA, E. C. & KOWALSKI, D. 2011. Replication fork stalling by bulky DNA damage: localization at active origins and checkpoint modulation. *Nucleic Acids Res,* 39**,** 2610-23.
MINOCHERHOMJI, S., YING, S., BJERREGAARD, V. A., BURSOMANNO, S., ALELIUNAITE, A., WU, W., MANKOURI, H. W., SHEN, H., LIU, Y. & HICKSON, I. D. 2015. Replication stress activates DNA repair synthesis in mitosis. *Nature,* 528**,** 286-90.
MIOTTO, B., JI, Z. & STRUHL, K. 2016. Selectivity of ORC binding sites and the relation to replication timing, fragile sites, and deletions in cancers. *Proc Natl Acad Sci U S A,* 113**,** E4810-9.
MISCHO, H. E., GOMEZ-GONZALEZ, B., GRZECHNIK, P., RONDON, A. G., WEI, W., STEINMETZ, L., AGUILERA, A. & PROUDFOOT, N. J. 2011. Yeast Sen1 helicase protects the genome from transcription-associated instability. *Mol Cell,* 41**,** 21-32.
MISHMAR, D., RAHAT, A., SCHERER, S. W., NYAKATURA, G., HINZMANN, B., KOHWI, Y., MANDEL-GUTFROIND, Y., LEE, J. R., DRESCHER, B., SAS, D. E., MARGALIT, H., PLATZER, M., WEISS, A., TSUI, L. C., ROSENTHAL, A. & KEREM, B. 1998. Molecular characterization of a common fragile site (FRA7H) on human chromosome 7 by the cloning of a simian virus 40 integration site. *Proc Natl Acad Sci U S A,* 95**,** 8141-6.
MOHAGHEGH, P., KAROW, J. K., BROSH, R. M., JR., BOHR, V. A. & HICKSON, I. D. 2001. The Bloom's and Werner's syndrome proteins are DNA structure-specific helicases. *Nucleic Acids Res,* 29**,** 2843-9.
MONDOL, T., STODOLA, J. L., GALLETTO, R. & BURGERS, P. M. 2019. PCNA accelerates the nucleotide incorporation rate by DNA polymerase delta. *Nucleic Acids Res,* 47**,** 1977-1986.
MONTERO, J. J., LOPEZ DE SILANES, I., GRANA, O. & BLASCO, M. A. 2016. Telomeric RNAs are essential to maintain telomeres. *Nat Commun,* 7**,** 12534.
MORADI-FARD, S., SARTHI, J., TITTEL-ELMER, M., LALONDE, M., CUSANELLI, E., CHARTRAND, P. & COBB, J. A. 2016. Smc5/6 Is a Telomere-Associated Complex that Regulates Sir4 Binding and TPE. *PLoS Genet,* 12**,** e1006268.
MORDES, D. A., GLICK, G. G., ZHAO, R. & CORTEZ, D. 2008. TopBP1 activates ATR through ATRIP and a PIKK regulatory domain. *Genes Dev,* 22**,** 1478-89.
MORENO, S. P., BAILEY, R., CAMPION, N., HERRON, S. & GAMBUS, A. 2014. Polyubiquitylation drives replisome disassembly at the termination of DNA replication. *Science,* 346**,** 477-81.
MORRICAL, S. W. 2015. DNA-pairing and annealing processes in homologous recombination and homology-directed repair. *Cold Spring Harb Perspect Biol,* 7**,** a016444.
MÜLLER, S., KUMARI, S., RODRIGUEZ, R. & BALASUBRAMANIAN, S. 2010. Small-molecule-mediated G-quadruplex isolation from human cells. *Nat Chem,* 2**,** 1095-8.
MURAKI, K., NYHAN, K., HAN, L. & MURNANE, J. P. 2012. Mechanisms of telomere loss and their consequences for chromosome instability. *Front Oncol,* 2**,** 135.
MURAYAMA, Y. & UHLMANN, F. 2015. DNA Entry into and Exit out of the Cohesin Ring by an Interlocking Gate Mechanism. *Cell,* 163**,** 1628-40.
MURGA, M., BUNTING, S., MONTANA, M. F., SORIA, R., MULERO, F., CANAMERO, M., LEE, Y., MCKINNON, P. J., NUSSENZWEIG, A. & FERNANDEZ-CAPETILLO, O. 2009. A mouse model of ATR-Seckel shows embryonic replicative stress and accelerated aging. *Nat Genet,* 41**,** 891-8.
MURRAY, J. M. & CARR, A. M. 2008. Smc5/6: a link between DNA repair and unidirectional replication? *Nat Rev Mol Cell Biol,* 9**,** 177-82.
MUTREJA, K., KRIETSCH, J., HESS, J., URSICH, S., BERTI, M., ROESSLER, F. K., ZELLWEGER, R., PATRA, M., GASSER, G. & LOPES, M. 2018. ATR-Mediated Global Fork Slowing and Reversal Assist Fork Traverse and Prevent Chromosomal Breakage at DNA Interstrand Cross-Links. *Cell Rep,* 24**,** 2629-2642 e5.

MUZI-FALCONI, M., GIANNATTASIO, M., FOIANI, M. & PLEVANI, P. 2003. The DNA polymerase alpha-primase complex: multiple functions and interactions. *ScientificWorldJournal,* 3**,** 21-33.

MYLER, L. R., GALLARDO, I. F., SONIAT, M. M., DESHPANDE, R. A., GONZALEZ, X. B., KIM, Y., PAULL, T. T. & FINKELSTEIN, I. J. 2017. Single-Molecule Imaging Reveals How Mre11-Rad50-Nbs1 Initiates DNA Break Repair. *Mol Cell,* 67**,** 891-898 e4.

NAIM, V. & ROSSELLI, F. 2009. The FANC pathway and BLM collaborate during mitosis to prevent micro-nucleation and chromosome abnormalities. *Nat Cell Biol,* 11**,** 761-8.

NAM, E. A. & CORTEZ, D. 2011. ATR signalling: more than meeting at the fork. *Biochem J,* 436**,** 527-36.

NAM, E. A., ZHAO, R., GLICK, G. G., BANSBACH, C. E., FRIEDMAN, D. B. & CORTEZ, D. 2011. Thr-1989 phosphorylation is a marker of active ataxia telangiectasia-mutated and Rad3-related (ATR) kinase. *J Biol Chem,* 286**,** 28707-28714.

NAMIKI, Y. & ZOU, L. 2006. ATRIP associates with replication protein A-coated ssDNA through multiple interactions. *Proceedings of the National Academy of Sciences,* 103**,** 580-585.

NASMYTH, K. & HAERING, C. H. 2005. The structure and function of SMC and kleisin complexes. *Annu Rev Biochem,* 74**,** 595-648.

NASMYTH, K. & HAERING, C. H. 2009. Cohesin: its roles and mechanisms. *Annu Rev Genet,* 43**,** 525-58.

NEELSEN, K. J. & LOPES, M. 2015. Replication fork reversal in eukaryotes: from dead end to dynamic response. *Nat Rev Mol Cell Biol,* 16**,** 207-20.

NEELSEN, K. J., ZANINI, I. M., MIJIC, S., HERRADOR, R., ZELLWEGER, R., RAY CHAUDHURI, A., CREAVIN, K. D., BLOW, J. J. & LOPES, M. 2013. Deregulated origin licensing leads to chromosomal breaks by rereplication of a gapped DNA template. *Genes Dev,* 27**,** 2537-42.

NEGRINI, S., GORGOULIS, V. G. & HALAZONETIS, T. D. 2010. Genomic instability--an evolving hallmark of cancer. *Nat Rev Mol Cell Biol,* 11**,** 220-8.

NEIL, A. J., LIANG, M. U., KHRISTICH, A. N., SHAH, K. A. & MIRKIN, S. M. 2018. RNA-DNA hybrids promote the expansion of Friedreich's ataxia (GAA)n repeats via break-induced replication. *Nucleic Acids Res,* 46**,** 3487-3497.

NGUYEN, G. H., TANG, W., ROBLES, A. I., BEYER, R. P., GRAY, L. T., WELSH, J. A., SCHETTER, A. J., KUMAMOTO, K., WANG, X. W., HICKSON, I. D., MAIZELS, N., MONNAT, R. J., JR. & HARRIS, C. C. 2014. Regulation of gene expression by the BLM helicase correlates with the presence of G-quadruplex DNA motifs. *Proc Natl Acad Sci U S A,* 111**,** 9905-10.

NGUYEN, V. Q., CO, C. & LI, J. J. 2001. Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms. *Nature,* 411**,** 1068-73.

NIEHRS, C. & LUKE, B. 2020. Regulatory R-loops as facilitators of gene expression and genome stability. *Nat Rev Mol Cell Biol,* 21**,** 167-178.

NIELSEN, C. F., ZHANG, T., BARISIC, M., KALITSIS, P. & HUDSON, D. F. 2020. Topoisomerase IIalpha is essential for maintenance of mitotic chromosome structure. *Proc Natl Acad Sci U S A,* 117**,** 12131-12142.

NIEMINUSZCZY, J., SCHWAB, R. A. & NIEDZWIEDZ, W. 2016. The DNA fibre technique - tracking helicases at work. *Methods,* 108**,** 92-8.

NIMONKAR, A. V., GENSCHEL, J., KINOSHITA, E., POLACZEK, P., CAMPBELL, J. L., WYMAN, C., MODRICH, P. & KOWALCZYKOWSKI, S. C. 2011. BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. *Genes Dev,* 25**,** 350-62.

NIMONKAR, A. V. & KOWALCZYKOWSKI, S. C. 2009. Second-end DNA capture in double-strand break repair: how to catch a DNA by its tail. *Cell Cycle,* 8**,** 1816-7.

NISHIYAMA, T., LADURNER, R., SCHMITZ, J., KREIDL, E., SCHLEIFFER, A., BHASKARA, V., BANDO, M., SHIRAHIGE, K., HYMAN, A. A., MECHTLER, K. & PETERS, J.-M. 2010. Sororin Mediates Sister Chromatid Cohesion by Antagonizing Wapl. *Cell,* 143**,** 737-749.

NISHIYAMA, T., SYKORA, M. M., HUIS IN 'T VELD, P. J., MECHTLER, K. & PETERS, J. M. 2013. Aurora B and Cdk1 mediate Wapl activation and release of acetylated cohesin from chromosomes by phosphorylating Sororin. *Proc Natl Acad Sci U S A,* 110**,** 13404-9.

NITISS, J. L. 2009. DNA topoisomerase II and its growing repertoire of biological functions. *Nat Rev Cancer,* 9**,** 327-37.

NORMAND, G. & KING, R. W. 2010. Understanding cytokinesis failure. *Adv Exp Med Biol,* 676**,** 27-55.

NUNEZ-RAMIREZ, R., KLINGE, S., SAUGUET, L., MELERO, R., RECUERO-CHECA, M. A., KILKENNY, M., PERERA, R. L., GARCIA-ALVAREZ, B., HALL, R. J., NOGALES, E., PELLEGRINI, L. & LLORCA, O. 2011. Flexible tethering of primase and DNA Pol alpha in the eukaryotic primosome. *Nucleic Acids Res,* 39**,** 8187-99.

O'DONNELL, M. & LI, H. 2016. The Eukaryotic Replisome Goes Under the Microscope. *Curr Biol,* 26**,** R247-56.

OBI, I., RENTOFT, M., SINGH, V., JAMROSKOVIC, J., CHAND, K., CHORELL, E., WESTERLUND, F. & SABOURI, N. 2020. Stabilization of G-quadruplex DNA structures in Schizosaccharomyces pombe causes single-strand DNA lesions and impedes DNA replication. *Nucleic Acids Res,* 48**,** 10998-11015.

OESTERGAARD, V. H. & LISBY, M. 2017. Transcription-replication conflicts at chromosomal fragile sites-consequences in M phase and beyond. *Chromosoma,* 126**,** 213-222.

OGAWA, T. & OKAZAKI, T. 1980. Discontinuous DNA replication. *Annu Rev Biochem,* 49**,** 421-57.

OHKI, R. & ISHIKAWA, F. 2004. Telomere-bound TRF1 and TRF2 stall the replication fork at telomeric repeats. *Nucleic Acids Res,* 32**,** 1627-37.

OLIVIERI, M., CHO, T., ALVAREZ-QUILON, A., LI, K., SCHELLENBERG, M. J., ZIMMERMANN, M., HUSTEDT, N., ROSSI, S. E., ADAM, S., MELO, H., HEIJINK, A. M., SASTRE-MORENO, G., MOATTI, N., SZILARD, R. K., MCEWAN, A., LING, A. K., SERRANO-BENITEZ, A., UBHI, T., FENG, S., PAWLING, J., DELGADO-SAINZ, I., FERGUSON, M. W., DENNIS, J. W., BROWN, G. W., CORTES-LEDESMA, F., WILLIAMS, R. S., MARTIN, A., XU, D. & DUROCHER, D. 2020. A Genetic Map of the Response to DNA Damage in Human Cells. *Cell,* 182**,** 481-496 e21.

ONO, T., FANG, Y., SPECTOR, D. L. & HIRANO, T. 2004. Spatial and temporal regulation of Condensins I and II in mitotic chromosome assembly in human cells. *Mol Biol Cell,* 15**,** 3296-308.

OOSTRA, A. B., NIEUWINT, A. W., JOENJE, H. & DE WINTER, J. P. 2012. Diagnosis of fanconi anemia: chromosomal breakage analysis. *Anemia,* 2012**,** 238731.

ORAVCOVA, M., GADALETA, M. C., NIE, M., REUBENS, M. C., LIMBO, O., RUSSELL, P. & BODDY, M. N. 2019. Brc1 Promotes the Focal Accumulation and SUMO Ligase Activity of Smc5-Smc6 during Replication Stress. *Mol Cell Biol,* 39.

OZER, O. & HICKSON, I. D. 2018. Pathways for maintenance of telomeres and common fragile sites during DNA replication stress. *Open Biol,* 8.

PAESCHKE, K., SIMONSSON, T., POSTBERG, J., RHODES, D. & LIPPS, H. J. 2005. Telomere end-binding proteins control the formation of G-quadruplex DNA structures in vivo. *Nat Struct Mol Biol,* 12**,** 847-54.

PALECEK, J., VIDOT, S., FENG, M., DOHERTY, A. J. & LEHMANN, A. R. 2006. The Smc5-Smc6 DNA repair complex. bridging of the Smc5-Smc6 heads by the KLEISIN, Nse4, and non-Kleisin subunits. *J Biol Chem,* 281**,** 36952-9.

PALECEK, J. J. 2018. SMC5/6: Multifunctional Player in Replication. *Genes (Basel),* 10.

PAN, X., CHEN, Y., BIJU, B., AHMED, N., KONG, J., GOLDENBERG, M., HUANG, J., MOHAN, N., KLOSEK, S., PARSA, K., GUH, C. Y., LU, R., PICKETT, H. A., CHU, H. P. & ZHANG, D. 2019.

FANCM suppresses DNA replication stress at ALT telomeres by disrupting TERRA R-loops. *Sci Rep,* 9**,** 19110.
PANDEY, S., AGARWALA, P. & MAITI, S. 2013. Effect of Loops and G-Quartets on the Stability of RNA G-Quadruplexes. *The Journal of Physical Chemistry B,* 117**,** 6896-6905.
PANG, D., YU, S. & YANG, X. 2022. A mini-review of the role of condensin in human nervous system diseases. *Front Mol Neurosci,* 15**,** 889796.
PARAJULI, S., TEASLEY, D. C., MURALI, B., JACKSON, J., VINDIGNI, A. & STEWART, S. A. 2017. Human ribonuclease H1 resolves R-loops and thereby enables progression of the DNA replication fork. *J Biol Chem,* 292**,** 15216-15224.
PARDO, B. & MARCAND, S. 2005. Rap1 prevents telomere fusions by nonhomologous end joining. *EMBO J,* 24**,** 3117-27.
PARKER, M. W., BOTCHAN, M. R. & BERGER, J. M. 2017. Mechanisms and regulation of DNA replication initiation in eukaryotes. *Crit Rev Biochem Mol Biol,* 52**,** 107-144.
PATIL, M., PABLA, N. & DONG, Z. 2013. Checkpoint kinase 1 in DNA damage response and cell cycle regulation. *Cell Mol Life Sci,* 70**,** 4009-21.
PATRICK, E. M., SLIVKA, J. D., PAYNE, B., COMSTOCK, M. J. & SCHMIDT, J. C. 2020. Observation of processive telomerase catalysis using high-resolution optical tweezers. *Nat Chem Biol,* 16**,** 801-809.
PAUL, M. R., HOCHWAGEN, A. & ERCAN, S. 2019. Condensin action and compaction. *Curr Genet,* 65**,** 407-415.
PAYNE, F., COLNAGHI, R., ROCHA, N., SETH, A., HARRIS, J., CARPENTER, G., BOTTOMLEY, W. E., WHEELER, E., WONG, S., SAUDEK, V., SAVAGE, D., O'RAHILLY, S., CAREL, J. C., BARROSO, I., O'DRISCOLL, M. & SEMPLE, R. 2014. Hypomorphism in human NSMCE2 linked to primordial dwarfism and insulin resistance. *J Clin Invest,* 124**,** 4028-38.
PEAKE, J. D. & NOGUCHI, E. 2022. Fanconi anemia: current insights regarding epidemiology, cancer, and DNA repair. *Hum Genet*.
PEBERNARD, S., PERRY, J. J., TAINER, J. A. & BODDY, M. N. 2008a. Nse1 RING-like domain supports functions of the Smc5-Smc6 holocomplex in genome stability. *Mol Biol Cell,* 19**,** 4099-109.
PEBERNARD, S., SCHAFFER, L., CAMPBELL, D., HEAD, S. R. & BODDY, M. N. 2008b. Localization of Smc5/6 to centromeres and telomeres requires heterochromatin and SUMO, respectively. *EMBO J,* 27**,** 3011-23.
PEBERNARD, S., WOHLSCHLEGEL, J., MCDONALD, W. H., YATES, J. R., 3RD & BODDY, M. N. 2006. The Nse5-Nse6 dimer mediates DNA repair roles of the Smc5-Smc6 complex. *Mol Cell Biol,* 26**,** 1617-30.
PELTONEN, K., COLIS, L., LIU, H., TRIVEDI, R., MOUBAREK, M. S., MOORE, H. M., BAI, B., RUDEK, M. A., BIEBERICH, C. J. & LAIHO, M. 2014. A targeting modality for destruction of RNA polymerase I that possesses anticancer activity. *Cancer Cell,* 25**,** 77-90.
PENG, X. P., LIM, S., LI, S., MARJAVAARA, L., CHABES, A. & ZHAO, X. 2018. Acute Smc5/6 depletion reveals its primary role in rDNA replication by restraining recombination at fork pausing sites. *PLoS Genet,* 14**,** e1007129.
PENTZOLD, C., SHAH, S. A., HANSEN, N. R., LE TALLEC, B., SEGUIN-ORLANDO, A., DEBATISSE, M., LISBY, M. & OESTERGAARD, V. H. 2018. FANCD2 binding identifies conserved fragile sites at large transcribed genes in avian cells. *Nucleic Acids Res,* 46**,** 1280-1294.
PEPE, A. & WEST, S. C. 2014. MUS81-EME2 promotes replication fork restart. *Cell Rep,* 7**,** 1048-55.
PERERA, R. L., TORELLA, R., KLINGE, S., KILKENNY, M. L., MAMAN, J. D. & PELLEGRINI, L. 2013. Mechanism for priming DNA synthesis by yeast DNA polymerase alpha. *Elife,* 2**,** e00482.
PETERMANN, E. & HELLEDAY, T. 2010. Pathways of mammalian replication fork restart. *Nat Rev Mol Cell Biol,* 11**,** 683-7.
PETERS, J. M., TEDESCHI, A. & SCHMITZ, J. 2008. The cohesin complex and its roles in chromosome biology. *Genes Dev,* 22**,** 3089-114.

PIBERGER, A. L., BOWRY, A., KELLY, R. D. W., WALKER, A. K., GONZÁLEZ-ACOSTA, D., BAILEY, L. J., DOHERTY, A. J., MÉNDEZ, J., MORRIS, J. R., BRYANT, H. E. & PETERMANN, E. 2020. PrimPol-dependent single-stranded gap formation mediates homologous recombination at bulky DNA adducts. *Nature Communications,* 11**,** 5863.

PICHÉ, J., VAN VLIET, P. P., PUCÉAT, M. & ANDELFINGER, G. 2019. The expanding phenotypes of cohesinopathies: one ring to rule them all! *Cell Cycle,* 18**,** 2828-2848.

PIETRUCHA, B., HEROPOLITANSKA-PLISZKA, E., GEFFERS, R., ENSSEN, J., WIELAND, B., BOGDANOVA, N. V. & DORK, T. 2017. Clinical and Biological Manifestation of RNF168 Deficiency in Two Polish Siblings. *Front Immunol,* 8**,** 1683.

PIKOR, L., THU, K., VUCIC, E. & LAM, W. 2013. The detection and implication of genome instability in cancer. *Cancer Metastasis Rev,* 32**,** 341-52.

PITT, C. W. & COOPER, J. P. 2010. Pot1 inactivation leads to rampant telomere resection and loss in one cell cycle. *Nucleic Acids Res,* 38**,** 6968-75.

PLANK, J. L., WU, J. & HSIEH, T. S. 2006. Topoisomerase IIIalpha and Bloom's helicase can resolve a mobile double Holliday junction substrate through convergent branch migration. *Proc Natl Acad Sci U S A,* 103**,** 11118-23.

POETSCH, A. R. 2020. The genomics of oxidative DNA damage, repair, and resulting mutagenesis. *Comput Struct Biotechnol J,* 18**,** 207-219.

POMMIER, Y., HUANG, S. Y., GAO, R., DAS, B. B., MURAI, J. & MARCHAND, C. 2014. Tyrosyl-DNA-phosphodiesterases (TDP1 and TDP2). *DNA Repair (Amst),* 19**,** 114-29.

POMMIER, Y., NUSSENZWEIG, A., TAKEDA, S. & AUSTIN, C. 2022. Human topoisomerases and their roles in genome stability and organization. *Nat Rev Mol Cell Biol,* 23**,** 407-427.

POOLE, L. A. & CORTEZ, D. 2017. Functions of SMARCAL1, ZRANB3, and HLTF in maintaining genome stability. *Crit Rev Biochem Mol Biol,* 52**,** 696-714.

POSTOW, L., CRISONA, N. J., PETER, B. J., HARDY, C. D. & COZZARELLI, N. R. 2001. Topological challenges to DNA replication: conformations at the fork. *Proc Natl Acad Sci U S A,* 98**,** 8219-26.

POTAPOVA, T. & GORBSKY, G. J. 2017. The Consequences of Chromosome Segregation Errors in Mitosis and Meiosis. *Biology* [Online], 6.

POTTS, P. R., PORTEUS, M. H. & YU, H. 2006. Human SMC5/6 complex promotes sister chromatid homologous recombination by recruiting the SMC1/3 cohesin complex to double-strand breaks. *EMBO J,* 25**,** 3377-88.

POTTS, P. R. & YU, H. 2005. Human MMS21/NSE2 is a SUMO ligase required for DNA repair. *Mol Cell Biol,* 25**,** 7021-32.

POTTS, P. R. & YU, H. 2007. The SMC5/6 complex maintains telomere length in ALT cancer cells through SUMOylation of telomere-binding proteins. *Nat Struct Mol Biol,* 14**,** 581-90.

PRADO, F. & AGUILERA, A. 2005. Impairment of replication fork progression mediates RNA polII transcription-associated recombination. *EMBO J,* 24**,** 1267-76.

PRELICH, G., TAN, C. K., KOSTURA, M., MATHEWS, M. B., SO, A. G., DOWNEY, K. M. & STILLMAN, B. 1987. Functional identity of proliferating cell nuclear antigen and a DNA polymerase-delta auxiliary protein. *Nature,* 326**,** 517-20.

PRYZHKOVA, M. V. & JORDAN, P. W. 2016. Conditional mutation of Smc5 in mouse embryonic stem cells perturbs condensin localization and mitotic progression. *J Cell Sci,* 129**,** 1619-34.

QUINET, A., LEMACON, D. & VINDIGNI, A. 2017. Replication Fork Reversal: Players and Guardians. *Mol Cell,* 68**,** 830-833.

RAI, R., VARMA, S. P., SHINDE, N., GHOSH, S., KUMARAN, S. P., SKARIAH, G. & LALORAYA, S. 2011. Small ubiquitin-related modifier ligase activity of Mms21 is required for maintenance of chromosome integrity during the unperturbed mitotic cell division cycle in Saccharomyces cerevisiae. *J Biol Chem,* 286**,** 14516-30.

RAJENDRA, E., OESTERGAARD, V. H., LANGEVIN, F., WANG, M., DORNAN, G. L., PATEL, K. J. & PASSMORE, L. A. 2014. The genetic and biochemical basis of FANCD2 monoubiquitination. *Mol Cell,* 54**,** 858-69.

RALF, C., HICKSON, I. D. & WU, L. 2006. The Bloom's syndrome helicase can promote the regression of a model replication fork. *J Biol Chem,* 281**,** 22839-46.

RAMSDEN, D. A. 2011. Polymerases in nonhomologous end joining: building a bridge over broken chromosomes. *Antioxid Redox Signal,* 14**,** 2509-19.

RAN, F. A., HSU, P. D., LIN, C. Y., GOOTENBERG, J. S., KONERMANN, S., TREVINO, A. E., SCOTT, D. A., INOUE, A., MATOBA, S., ZHANG, Y. & ZHANG, F. 2013. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell,* 154**,** 1380-9.

RASCHLE, M., SMEENK, G., HANSEN, R. K., TEMU, T., OKA, Y., HEIN, M. Y., NAGARAJ, N., LONG, D. T., WALTER, J. C., HOFMANN, K., STORCHOVA, Z., COX, J., BEKKER-JENSEN, S., MAILAND, N. & MANN, M. 2015. DNA repair. Proteomics reveals dynamic assembly of repair complexes during bypass of DNA cross-links. *Science,* 348**,** 1253671.

RASS, U., COMPTON, S. A., MATOS, J., SINGLETON, M. R., IP, S. C., BLANCO, M. G., GRIFFITH, J. D. & WEST, S. C. 2010. Mechanism of Holliday junction resolution by the human GEN1 protein. *Genes Dev,* 24**,** 1559-69.

RASTOGI, R. P., RICHA, KUMAR, A., TYAGI, M. B. & SINHA, R. P. 2010. Molecular mechanisms of ultraviolet radiation-induced DNA damage and repair. *J Nucleic Acids,* 2010**,** 592980.

RAY CHAUDHURI, A., HASHIMOTO, Y., HERRADOR, R., NEELSEN, K. J., FACHINETTI, D., BERMEJO, R., COCITO, A., COSTANZO, V. & LOPES, M. 2012. Topoisomerase I poisoning results in PARP-mediated replication fork reversal. *Nat Struct Mol Biol,* 19**,** 417-23.

REDINBO, M. R., CHAMPOUX, J. J. & HOL, W. G. 2000. Novel insights into catalytic mechanism from a crystal structure of human topoisomerase I in complex with DNA. *Biochemistry,* 39**,** 6832-40.

REDMOND, M. T., SCHERZER, R. & PRINCE, B. T. 2022. Novel Genetic Discoveries in Primary Immunodeficiency Disorders. *Clin Rev Allergy Immunol,* 63**,** 55-74.

REGAIRAZ, M., ZHANG, Y. W., FU, H., AGAMA, K. K., TATA, N., AGRAWAL, S., ALADJEM, M. I. & POMMIER, Y. 2011. Mus81-mediated DNA cleavage resolves replication forks stalled by topoisomerase I-DNA complexes. *J Cell Biol,* 195**,** 739-49.

REISZ, J. A., BANSAL, N., QIAN, J., ZHAO, W. & FURDUI, C. M. 2014. Effects of ionizing radiation on biological molecules--mechanisms of damage and emerging methods of detection. *Antioxid Redox Signal,* 21**,** 260-92.

REMUS, D., BEURON, F., TOLUN, G., GRIFFITH, J. D., MORRIS, E. P. & DIFFLEY, J. F. 2009. Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. *Cell,* 139**,** 719-30.

REYNOLDS, J. J., BICKNELL, L. S., CARROLL, P., HIGGS, M. R., SHAHEEN, R., MURRAY, J. E., PAPADOPOULOS, D. K., LEITCH, A., MURINA, O., TARNAUSKAITE, Z., WESSEL, S. R., ZLATANOU, A., VERNET, A., VON KRIEGSHEIM, A., MOTTRAM, R. M., LOGAN, C. V., BYE, H., LI, Y., BREAN, A., MADDIREVULA, S., CHALLIS, R. C., SKOULOUDAKI, K., ALMOISHEER, A., ALSAIF, H. S., AMAR, A., PRESCOTT, N. J., BOBER, M. B., DUKER, A., FAQEIH, E., SEIDAHMED, M. Z., AL TALA, S., ALSWAID, A., AHMED, S., AL-AAMA, J. Y., ALTMULLER, J., AL BALWI, M., BRADY, A. F., CHESSA, L., COX, H., FISCHETTO, R., HELLER, R., HENDERSON, B. D., HOBSON, E., NURNBERG, P., PERCIN, E. F., PERON, A., SPACCINI, L., QUIGLEY, A. J., THAKUR, S., WISE, C. A., YOON, G., ALNEMER, M., TOMANCAK, P., YIGIT, G., TAYLOR, A. M., REIJNS, M. A., SIMPSON, M. A., CORTEZ, D., ALKURAYA, F. S., MATHEW, C. G., JACKSON, A. P. & STEWART, G. S. 2017. Mutations in DONSON disrupt replication fork stability and cause microcephalic dwarfism. *Nat Genet,* 49**,** 537-549.

REYNOLDS, J. J. & STEWART, G. S. 2018. DNA Replication and Inherited Human Disease. *In:* GRAND, R. J. A. & REYNOLDS, J. J. (eds.) *DNA Repair and Replication: Mechanisms and Clinical Significance.* CRC Press.

RHODES, D. & LIPPS, H. J. 2015. G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Res,* 43**,** 8627-37.

RICKMAN, K. & SMOGORZEWSKA, A. 2019. Advances in understanding DNA processing and protection at stalled replication forks. *J Cell Biol,* 218**,** 1096-1107.

RIZZO, A., SALVATI, E., PORRU, M., D'ANGELO, C., STEVENS, M. F., D'INCALCI, M., LEONETTI, C., GILSON, E., ZUPI, G. & BIROCCIO, A. 2009. Stabilization of quadruplex DNA perturbs telomere replication leading to the activation of an ATR-dependent ATM signaling pathway. *Nucleic Acids Res,* 37**,** 5353-64.

ROMERO, N. E., MATSON, S. W. & SEKELSKY, J. 2016. Biochemical Activities and Genetic Functions of the Drosophila melanogaster Fancm Helicase in DNA Repair. *Genetics,* 204**,** 531-541.

ROSEN, L. E., KLEBBA, J. E., ASFAHA, J. B., GHENT, C. M., CAMPBELL, M. G., CHENG, Y. & MORGAN, D. O. 2019. Cohesin cleavage by separase is enhanced by a substrate motif distinct from the cleavage site. *Nature Communications,* 10**,** 5189.

ROSSI, F., HELBLING-LECLERC, A., KAWASUMI, R., JEGADESAN, N. K., XU, X., DEVULDER, P., ABE, T., TAKATA, M., XU, D., ROSSELLI, F. & BRANZEI, D. 2020. SMC5/6 acts jointly with Fanconi anemia factors to support DNA repair and genome stability. *EMBO Rep,* 21**,** e48222.

ROY, M. A., DHANARAMAN, T. & D'AMOURS, D. 2015. The Smc5-Smc6 heterodimer associates with DNA through several independent binding domains. *Sci Rep,* 5**,** 9797.

RUSAN, N. M., TULU, U. S., FAGERSTROM, C. & WADSWORTH, P. 2002. Reorganization of the microtubule array in prophase/prometaphase requires cytoplasmic dynein-dependent microtubule transport. *J Cell Biol,* 158**,** 997-1003.

RYU, T., SPATOLA, B., DELABAERE, L., BOWLIN, K., HOPP, H., KUNITAKE, R., KARPEN, G. H. & CHIOLO, I. 2015. Heterochromatic breaks move to the nuclear periphery to continue recombinational repair. *Nat Cell Biol,* 17**,** 1401-11.

SAINI, N., RAMAKRISHNAN, S., ELANGO, R., AYYAR, S., ZHANG, Y., DEEM, A., IRA, G., HABER, J. E., LOBACHEV, K. S. & MALKOVA, A. 2013. Migrating bubble during break-induced replication drives conservative DNA synthesis. *Nature,* 502**,** 389-92.

SAKOFSKY, C. J., AYYAR, S. & MALKOVA, A. 2012. Break-induced replication and genome stability. *Biomolecules,* 2**,** 483-504.

SALDIVAR, J. C., CORTEZ, D. & CIMPRICH, K. A. 2017. The essential kinase ATR: ensuring faithful duplication of a challenging genome. *Nat Rev Mol Cell Biol,* 18**,** 622-636.

SALEH-GOHARI, N. & HELLEDAY, T. 2004. Conservative homologous recombination preferentially repairs DNA double-strand breaks in the S phase of the cell cycle in human cells. *Nucleic Acids Res,* 32**,** 3683-8.

SALIM, D., BRADFORD, W. D., FREELAND, A., CADY, G., WANG, J., PRUITT, S. C. & GERTON, J. L. 2017. DNA replication stress restricts ribosomal DNA copy number. *PLoS Genet,* 13**,** e1007006.

SAMORA, C. P., SAKSOUK, J., GOSWAMI, P., WADE, B. O., SINGLETON, M. R., BATES, P. A., LENGRONNE, A., COSTA, A. & UHLMANN, F. 2016. Ctf4 Links DNA Replication with Sister Chromatid Cohesion Establishment by Recruiting the Chl1 Helicase to the Replisome. *Mol Cell,* 63**,** 371-84.

SAMOSHKIN, A., ARNAOUTOV, A., JANSEN, L. E., OUSPENSKI, I., DYE, L., KARPOVA, T., MCNALLY, J., DASSO, M., CLEVELAND, D. W. & STRUNNIKOV, A. 2009. Human condensin function is essential for centromeric chromatin assembly and proper sister kinetochore orientation. *PLoS One,* 4**,** e6831.

SANCAR, A., LINDSEY-BOLTZ, L. A., UNSAL-KACMAZ, K. & LINN, S. 2004. Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annu Rev Biochem,* 73**,** 39-85.

SANCHEZ, A., REGINATO, G. & CEJKA, P. 2021. Crossover or non-crossover outcomes: tailored processing of homologous recombination intermediates. *Curr Opin Genet Dev,* 71**,** 39-47.
SANTOS-SIMARRO, F., PACIO, M., CUETO-GONZALEZ, A. M., MANSILLA, E., VALENZUELA-PALAFOLL, M. I., LOPEZ-GRONDONA, F., LLEDIN, M. D., SCHUFFELMANN, C., DEL POZO, A., SOLIS, M., VALLCORBA, P., LAPUNZINA, P., MENENDEZ SUSO, J. J., SICCHA, S. M., MONTEJO, J. M., MENA, R., JIMENEZ-RODRIGUEZ, C., GARCIA-MINAUR, S. & PALOMARES-BRALO, M. 2021. Mosaic Variegated Aneuploidy syndrome 2 caused by biallelic variants in CEP57, two new cases and review of the phenotype. *Eur J Med Genet,* 64**,** 104338.
SAPKOTA, H., WASIAK, E., DAUM, J. R. & GORBSKY, G. J. 2018. Multiple determinants and consequences of cohesion fatigue in mammalian cells. *Mol Biol Cell,* 29**,** 1811-1824.
SARBAJNA, S., DAVIES, D. & WEST, S. C. 2014. Roles of SLX1-SLX4, MUS81-EME1, and GEN1 in avoiding genome instability and mitotic catastrophe. *Genes Dev,* 28**,** 1124-36.
SAREK, G., VANNIER, J. B., PANIER, S., PETRINI, J. H. J. & BOULTON, S. J. 2016. TRF2 Recruits RTEL1 to Telomeres in S Phase to Promote T-Loop Unwinding. *Mol Cell,* 61**,** 788-789.
SARGENT, R. G., MESERVY, J. L., PERKINS, B. D., KILBURN, A. E., INTODY, Z., ADAIR, G. M., NAIRN, R. S. & WILSON, J. H. 2000. Role of the nucleotide excision repair gene ERCC1 in formation of recombination-dependent rearrangements in mammalian cells. *Nucleic Acids Res,* 28**,** 3771-8.
SARTORI, A. A., LUKAS, C., COATES, J., MISTRIK, M., FU, S., BARTEK, J., BAER, R., LUKAS, J. & JACKSON, S. P. 2007. Human CtIP promotes DNA end resection. *Nature,* 450**,** 509-14.
SAUER, M. & PAESCHKE, K. 2017. G-quadruplex unwinding helicases and their function in vivo. *Biochem Soc Trans,* 45**,** 1173-1182.
SCHINDELIN, J., ARGANDA-CARRERAS, I., FRISE, E., KAYNIG, V., LONGAIR, M., PIETZSCH, T., PREIBISCH, S., RUEDEN, C., SAALFELD, S., SCHMID, B., TINEVEZ, J. Y., WHITE, D. J., HARTENSTEIN, V., ELICEIRI, K., TOMANCAK, P. & CARDONA, A. 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods,* 9**,** 676-82.
SCHLACHER, K., CHRIST, N., SIAUD, N., EGASHIRA, A., WU, H. & JASIN, M. 2011. Double-strand break repair-independent role for BRCA2 in blocking stalled replication fork degradation by MRE11. *Cell,* 145**,** 529-42.
SCHLEIFFER, A., KAITNA, S., MAURER-STROH, S., GLOTZER, M., NASMYTH, K. & EISENHABER, F. 2003. Kleisins: a superfamily of bacterial and eukaryotic SMC protein partners. *Mol Cell,* 11**,** 571-5.
SCHOENAKER, M. H. D., HENRIET, S. S., ZONDERLAND, J., VAN DEUREN, M., PAN-HAMMARSTROM, Q., POSTHUMUS-VAN SLUIJS, S. J., PICO-KNIJNENBURG, I., WEEMAES, C. M. R. & H, I. J. 2018. Immunodeficiency in Bloom's Syndrome. *J Clin Immunol,* 38**,** 35-44.
SCHRECKER, M., CASTANEDA, J. C., DEVBHANDARI, S., KUMAR, C., REMUS, D. & HITE, R. K. 2022. Multistep loading of a DNA sliding clamp onto DNA by replication factor C. *Elife,* 11.
SCHULZ, S., GERLOFF, C., LEDIG, S., LANGER, D., VOLLETH, M., SHIRNESHAN, K. & WIEACKER, P. 2008. Prenatal diagnosis of Roberts syndrome and detection of an ESCO2 frameshift mutation in a Pakistani family. *Prenat Diagn,* 28**,** 42-5.
SCHWAB, R. A., NIEMINUSZCZY, J., SHAH, F., LANGTON, J., LOPEZ MARTINEZ, D., LIANG, C. C., COHN, M. A., GIBBONS, R. J., DEANS, A. J. & NIEDZWIEDZ, W. 2015. The Fanconi Anemia Pathway Maintains Genome Stability by Coordinating Replication and Transcription. *Mol Cell,* 60**,** 351-61.
SCHWERTMAN, P., BEKKER-JENSEN, S. & MAILAND, N. 2016. Regulation of DNA double-strand break repair by ubiquitin and ubiquitin-like modifiers. *Nat Rev Mol Cell Biol,* 17**,** 379-94.
SCOTT, W. A., DHANJI, E. Z., DYAKOV, B. J. A., DRESERIS, E. S., ASA, J. S., GRANGE, L. J., MIRCETA, M., PEARSON, C. E., STEWART, G. S., GINGRAS, A. C. & CAMPOS, E. I. 2021. ATRX proximal protein associations boast roles beyond histone deposition. *PLoS Genet,* 17**,** e1009909.

SEILER, J. A., CONTI, C., SYED, A., ALADJEM, M. I. & POMMIER, Y. 2007. The intra-S-phase checkpoint affects both DNA replication initiation and elongation: single-cell and -DNA fiber analyses. *Mol Cell Biol,* 27**,** 5806-18.

SFEIR, A. & DE LANGE, T. 2012. Removal of shelterin reveals the telomere end-protection problem. *Science,* 336**,** 593-7.

SFEIR, A., KOSIYATRAKUL, S. T., HOCKEMEYER, D., MACRAE, S. L., KARLSEDER, J., SCHILDKRAUT, C. L. & DE LANGE, T. 2009. Mammalian telomeres resemble fragile sites and require TRF1 for efficient replication. *Cell,* 138**,** 90-103.

SHEN, T. & HUANG, S. 2012. The role of Cdc25A in the regulation of cell proliferation and apoptosis. *Anticancer Agents Med Chem,* 12**,** 631-9.

SHIBATA, A., MOIANI, D., ARVAI, A. S., PERRY, J., HARDING, S. M., GENOIS, M. M., MAITY, R., VAN ROSSUM-FIKKERT, S., KERTOKALIO, A., ROMOLI, F., ISMAIL, A., ISMALAJ, E., PETRICCI, E., NEALE, M. J., BRISTOW, R. G., MASSON, J. Y., WYMAN, C., JEGGO, P. A. & TAINER, J. A. 2014. DNA double-strand break repair pathway choice is directed by distinct MRE11 nuclease activities. *Mol Cell,* 53**,** 7-18.

SHOKOLENKO, I., VENEDIKTOVA, N., BOCHKAREVA, A., WILSON, G. L. & ALEXEYEV, M. F. 2009. Oxidative stress induces degradation of mitochondrial DNA. *Nucleic Acids Res,* 37**,** 2539-48.

SIDDIQUI, K., ON, K. F. & DIFFLEY, J. F. 2013. Regulating DNA replication in eukarya. *Cold Spring Harb Perspect Biol,* 5.

SISSI, C. & PALUMBO, M. 2009. Effects of magnesium and related divalent metal ions in topoisomerase structure and function. *Nucleic Acids Res,* 37**,** 702-11.

SMITH, C. E., LLORENTE, B. & SYMINGTON, L. S. 2007. Template switching during break-induced replication. *Nature,* 447**,** 102-5.

SMITH, J. S., CHEN, Q., YATSUNYK, L. A., NICOLUDIS, J. M., GARCIA, M. S., KRANASTER, R., BALASUBRAMANIAN, S., MONCHAUD, D., TEULADE-FICHOU, M. P., ABRAMOWITZ, L., SCHULTZ, D. C. & JOHNSON, F. B. 2011. Rudimentary G-quadruplex-based telomere capping in Saccharomyces cerevisiae. *Nat Struct Mol Biol,* 18**,** 478-85.

SMOGORZEWSKA, A., MATSUOKA, S., VINCIGUERRA, P., MCDONALD, E. R., 3RD, HUROV, K. E., LUO, J., BALLIF, B. A., GYGI, S. P., HOFMANN, K., D'ANDREA, A. D. & ELLEDGE, S. J. 2007. Identification of the FANCI protein, a monoubiquitinated FANCD2 paralog required for DNA repair. *Cell,* 129**,** 289-301.

SNAPE, K., HANKS, S., RUARK, E., BARROS-NUNEZ, P., ELLIOTT, A., MURRAY, A., LANE, A. H., SHANNON, N., CALLIER, P., CHITAYAT, D., CLAYTON-SMITH, J., FITZPATRICK, D. R., GISSELSSON, D., JACQUEMONT, S., ASAKURA-HAY, K., MICALE, M. A., TOLMIE, J., TURNPENNY, P. D., WRIGHT, M., DOUGLAS, J. & RAHMAN, N. 2011. Mutations in CEP57 cause mosaic variegated aneuploidy syndrome. *Nat Genet,* 43**,** 527-9.

SO, A., DARDILLAC, E., MUHAMMAD, A., CHAILLEUX, C., SESMA-SANZ, L., RAGU, S., LE CAM, E., CANITROT, Y., MASSON, J. Y., DUPAIGNE, P., LOPEZ, B. S. & GUIROUILH-BARBAT, J. 2022. RAD51 protects against nonconservative DNA double-strand break repair through a nonenzymatic function. *Nucleic Acids Res,* 50**,** 2651-2666.

SOBREIRA, N., SCHIETTECATTE, F., VALLE, D. & HAMOSH, A. 2015. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. *Hum Mutat,* 36**,** 928-30.

SOLE-SOLER, R. & TORRES-ROSELL, J. 2020. Smc5/6, an atypical SMC complex with two RING-type subunits. *Biochem Soc Trans,* 48**,** 2159-2171.

SOLLIER, J., DRISCOLL, R., CASTELLUCCI, F., FOIANI, M., JACKSON, S. P. & BRANZEI, D. 2009. The Saccharomyces cerevisiae Esc2 and Smc5-6 proteins promote sister chromatid junction-mediated intra-S repair. *Mol Biol Cell,* 20**,** 1671-82.

SONNEVILLE, R., MORENO, S. P., KNEBEL, A., JOHNSON, C., HASTIE, C. J., GARTNER, A., GAMBUS, A. & LABIB, K. 2017. CUL-2(LRR-1) and UBXN-3 drive replisome disassembly during DNA replication termination and mitosis. *Nat Cell Biol,* 19**,** 468-479.

SPIEGEL, J., ADHIKARI, S. & BALASUBRAMANIAN, S. 2020. The Structure and Function of DNA G-Quadruplexes. *Trends Chem,* 2**,** 123-136.
SPITZ, D. R., AZZAM, E. I., LI, J. J. & GIUS, D. 2004. Metabolic oxidation/reduction reactions and cellular responses to ionizing radiation: a unifying concept in stress response biology. *Cancer Metastasis Rev,* 23**,** 311-22.
SRINIVASAN, M., FUMASONI, M., PETELA, N. J., MURRAY, A. & NASMYTH, K. A. 2020. Cohesion is established during DNA replication utilising chromosome associated cohesin rings as well as those loaded de novo onto nascent DNAs. *eLife,* 9**,** e56611.
STARK, J. M., PIERCE, A. J., OH, J., PASTINK, A. & JASIN, M. 2004. Genetic steps of mammalian homologous repair with distinct mutagenic consequences. *Mol Cell Biol,* 24**,** 9305-16.
STENSON, P. D., MORT, M., BALL, E. V., HOWELLS, K., PHILLIPS, A. D., THOMAS, N. S. & COOPER, D. N. 2009. The Human Gene Mutation Database: 2008 update. *Genome Med,* 1**,** 13.
STEPHAN, A. K., KLISZCZAK, M., DODSON, H., COOLEY, C. & MORRISON, C. G. 2011a. Roles of vertebrate Smc5 in sister chromatid cohesion and homologous recombinational repair. *Mol Cell Biol,* 31**,** 1369-81.
STEPHAN, A. K., KLISZCZAK, M. & MORRISON, C. G. 2011b. The Nse2/Mms21 SUMO ligase of the Smc5/6 complex in the maintenance of genome stability. *FEBS Lett,* 585**,** 2907-13.
STEWART, G. S., PANIER, S., TOWNSEND, K., AL-HAKIM, A. K., KOLAS, N. K., MILLER, E. S., NAKADA, S., YLANKO, J., OLIVARIUS, S., MENDEZ, M., OLDREIVE, C., WILDENHAIN, J., TAGLIAFERRO, A., PELLETIER, L., TAUBENHEIM, N., DURANDY, A., BYRD, P. J., STANKOVIC, T., TAYLOR, A. M. & DUROCHER, D. 2009. The RIDDLE syndrome protein mediates a ubiquitin-dependent signaling cascade at sites of DNA damage. *Cell,* 136**,** 420-34.
STEWART, G. S., STANKOVIC, T., BYRD, P. J., WECHSLER, T., MILLER, E. S., HUISSOON, A., DRAYSON, M. T., WEST, S. C., ELLEDGE, S. J. & TAYLOR, A. M. 2007. RIDDLE immunodeficiency syndrome is linked to defects in 53BP1-mediated DNA damage signaling. *Proc Natl Acad Sci U S A,* 104**,** 16910-5.
STINGELE, S., STOEHR, G., PEPLOWSKA, K., COX, J., MANN, M. & STORCHOVA, Z. 2012. Global analysis of genome, transcriptome and proteome reveals the response to aneuploidy in human cells. *Mol Syst Biol,* 8**,** 608.
STRANDE, N. T., WATERS, C. A. & RAMSDEN, D. A. 2012. Resolution of complex ends by Nonhomologous end joining - better to be lucky than good? *Genome Integr,* 3**,** 10.
STRÖM, L., KARLSSON, C., LINDROOS, H. B., WEDAHL, S., KATOU, Y., SHIRAHIGE, K. & SJÖGREN, C. 2007. Postreplicative Formation of Cohesion Is Required for Repair and Induced by a Single DNA Break. *Science,* 317**,** 242-245.
STURZENEGGER, A., BURDOVA, K., KANAGARAJ, R., LEVIKOVA, M., PINTO, C., CEJKA, P. & JANSCAK, P. 2014. DNA2 cooperates with the WRN and BLM RecQ helicases to mediate long-range DNA end resection in human cells. *J Biol Chem,* 289**,** 27314-27326.
SUGIMOTO, N., MAEHARA, K., YOSHIDA, K., OHKAWA, Y. & FUJITA, M. 2018. Genome-wide analysis of the spatiotemporal regulation of firing and dormant replication origins in human cells. *Nucleic Acids Res,* 46**,** 6683-6696.
SUMARA, I., VORLAUFER, E., STUKENBERG, P. T., KELM, O., REDEMANN, N., NIGG, E. A. & PETERS, J.-M. 2002. The Dissociation of Cohesin from Chromosomes in Prophase Is Regulated by Polo-like Kinase. *Molecular Cell,* 9**,** 515-525.
SUN, H., MA, L., TSAI, Y. F., ABEYWARDANA, T., SHEN, B. & ZHENG, L. 2022. Okazaki fragment maturation: DNA flap dynamics for cell proliferation and survival. *Trends Cell Biol*.
SUN, Y., MILLER JENKINS, L. M., SU, Y. P., NITISS, K. C., NITISS, J. L. & POMMIER, Y. 2020. A conserved SUMO pathway repairs topoisomerase DNA-protein cross-links by engaging ubiquitin-mediated proteasomal degradation. *Sci Adv,* 6.
SUROVA, O. & ZHIVOTOVSKY, B. 2013. Various modes of cell death induced by DNA damage. *Oncogene,* 32**,** 3789-97.

TAHARA, H., SHIN-YA, K., SEIMIYA, H., YAMADA, H., TSURUO, T. & IDE, T. 2006. G-Quadruplex stabilization by telomestatin induces TRF2 protein dissociation from telomeres and anaphase bridge formation accompanied by loss of the 3' telomeric overhang in cancer cells. *Oncogene,* 25**,** 1955-66.

TANAKA, S. & OGAWA, S. 2022. Dimerization of Firing Factors for Replication Origin Activation in Eukaryotes: A Crucial Process for Simultaneous Assembly of Bidirectional Replication Forks? *Biology (Basel),* 11.

TANG, J., KAN, Z.-Y., YAO, Y., WANG, Q., HAO, Y.-H. & TAN, Z. 2007. G-quadruplex preferentially forms at the very 3' end of vertebrate telomeric DNA. *Nucleic Acids Research,* 36**,** 1200-1208.

TANG, W., ROBLES, A. I., BEYER, R. P., GRAY, L. T., NGUYEN, G. H., OSHIMA, J., MAIZELS, N., HARRIS, C. C. & MONNAT, R. J., JR. 2016. The Werner syndrome RECQ helicase targets G4 DNA in human cells to modulate transcription. *Hum Mol Genet,* 25**,** 2060-2069.

TASCHNER, M., BASQUIN, J., STEIGENBERGER, B., SCHAFER, I. B., SOH, Y. M., BASQUIN, C., LORENTZEN, E., RASCHLE, M., SCHELTEMA, R. A. & GRUBER, S. 2021. Nse5/6 inhibits the Smc5/6 ATPase and modulates DNA substrate binding. *EMBO J,* 40**,** e107807.

TATE, J. G., BAMFORD, S., JUBB, H. C., SONDKA, Z., BEARE, D. M., BINDAL, N., BOUTSELAKIS, H., COLE, C. G., CREATORE, C., DAWSON, E., FISH, P., HARSHA, B., HATHAWAY, C., JUPE, S. C., KOK, C. Y., NOBLE, K., PONTING, L., RAMSHAW, C. C., RYE, C. E., SPEEDY, H. E., STEFANCSIK, R., THOMPSON, S. L., WANG, S., WARD, S., CAMPBELL, P. J. & FORBES, S. A. 2019. COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res,* 47**,** D941-D947.

TAYLOR, A. M., SHIH, J., HA, G., GAO, G. F., ZHANG, X., BERGER, A. C., SCHUMACHER, S. E., WANG, C., HU, H., LIU, J., LAZAR, A. J., CANCER GENOME ATLAS RESEARCH, N., CHERNIACK, A. D., BEROUKHIM, R. & MEYERSON, M. 2018. Genomic and Functional Approaches to Understanding Cancer Aneuploidy. *Cancer Cell,* 33**,** 676-689 e3.

TAYLOR, A. M. R., ROTHBLUM-OVIATT, C., ELLIS, N. A., HICKSON, I. D., MEYER, S., CRAWFORD, T. O., SMOGORZEWSKA, A., PIETRUCHA, B., WEEMAES, C. & STEWART, G. S. 2019. Chromosome instability syndromes. *Nat Rev Dis Primers,* 5**,** 64.

TAYLOR, E. M., COPSEY, A. C., HUDSON, J. J., VIDOT, S. & LEHMANN, A. R. 2008. Identification of the proteins, including MAGEG1, that make up the human SMC5-6 protein complex. *Mol Cell Biol,* 28**,** 1197-206.

TAYLOR, E. M., MOGHRABY, J. S., LEES, J. H., SMIT, B., MOENS, P. B. & LEHMANN, A. R. 2001. Characterization of a novel human SMC heterodimer homologous to the Schizosaccharomyces pombe Rad18/Spr18 complex. *Mol Biol Cell,* 12**,** 1583-94.

TECHER, H., KOUNDRIOUKOFF, S., NICOLAS, A. & DEBATISSE, M. 2017. The impact of replication stress on replication dynamics and DNA damage in vertebrate cells. *Nat Rev Genet,* 18**,** 535-550.

TERABAYASHI, T. & HANADA, K. 2018. Genome instability syndromes caused by impaired DNA repair and aberrant DNA damage responses. *Cell Biol Toxicol,* 34**,** 337-350.

TERAKAWA, T., BISHT, S., EEFTENS, J. M., DEKKER, C., HAERING, C. H. & GREENE, E. C. 2017. The condensin complex is a mechanochemical motor that translocates along DNA. *Science,* 358**,** 672-676.

TERCERO, J. A. & DIFFLEY, J. F. 2001. Regulation of DNA replication fork progression through damaged DNA by the Mec1/Rad53 checkpoint. *Nature,* 412**,** 553-7.

THADANI, R., UHLMANN, F. & HEEGER, S. 2012. Condensin, chromatin crossbarring and chromosome condensation. *Curr Biol,* 22**,** R1012-21.

THORNTON, B. R. & TOCZYSKI, D. P. 2003. Securin and B-cyclin/CDK are the only essential targets of the APC. *Nature Cell Biology,* 5**,** 1090-1094.

TINGLER, M., PHILIPP, M. & BURKHALTER, M. D. 2022. DNA Replication proteins in primary microcephaly syndromes. *Biol Cell,* 114**,** 143-159.

TODD, A. K., JOHNSTON, M. & NEIDLE, S. 2005. Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Res,* 33**,** 2901-7.
TOLEDO, L. I., ALTMEYER, M., RASK, M. B., LUKAS, C., LARSEN, D. H., POVLSEN, L. K., BEKKER-JENSEN, S., MAILAND, N., BARTEK, J. & LUKAS, J. 2013. ATR prohibits replication catastrophe by preventing global exhaustion of RPA. *Cell,* 155**,** 1088-103.
TOMITA, K. 2018. How long does telomerase extend telomeres? Regulation of telomerase release and telomere length homeostasis. *Curr Genet,* 64**,** 1177-1181.
TORRES-ROSELL, J., MACHIN, F., FARMER, S., JARMUZ, A., EYDMANN, T., DALGAARD, J. Z. & ARAGON, L. 2005. SMC5 and SMC6 genes are required for the segregation of repetitive chromosome regions. *Nat Cell Biol,* 7**,** 412-9.
TORRES-ROSELL, J., SUNJEVARIC, I., DE PICCOLI, G., SACHER, M., ECKERT-BOULET, N., REID, R., JENTSCH, S., ROTHSTEIN, R., ARAGON, L. & LISBY, M. 2007. The Smc5-Smc6 complex and SUMO modification of Rad52 regulates recombinational repair at the ribosomal gene locus. *Nat Cell Biol,* 9**,** 923-31.
TORRES, E. M., SOKOLSKY, T., TUCKER, C. M., CHAN, L. Y., BOSELLI, M., DUNHAM, M. J. & AMON, A. 2007. Effects of aneuploidy on cellular physiology and cell division in haploid yeast. *Science,* 317**,** 916-24.
TRENZ, K., SMITH, E., SMITH, S. & COSTANZO, V. 2006. ATM and ATR promote Mre11 dependent restart of collapsed replication forks and prevent accumulation of DNA breaks. *EMBO J,* 25**,** 1764-74.
TRIMBORN, M., BELL, S. M., FELIX, C., RASHID, Y., JAFRI, H., GRIFFITHS, P. D., NEUMANN, L. M., KREBS, A., REIS, A., SPERLING, K., NEITZEL, H. & JACKSON, A. P. 2004. Mutations in microcephalin cause aberrant regulation of chromosome condensation. *Am J Hum Genet,* 75**,** 261-6.
TRIMBORN, M., SCHINDLER, D., NEITZEL, H. & HIRANO, T. 2006. Misregulated chromosome condensation in MCPH1 primary microcephaly is mediated by condensin II. *Cell Cycle,* 5**,** 322-6.
TRUONG, L. N. & WU, X. 2011. Prevention of DNA re-replication in eukaryotic cells. *J Mol Cell Biol,* 3**,** 13-22.
TUBBS, A. & NUSSENZWEIG, A. 2017. Endogenous DNA Damage as a Source of Genomic Instability in Cancer. *Cell,* 168**,** 644-656.
UCKELMANN, M. & SIXMA, T. K. 2017. Histone ubiquitination in the DNA damage response. *DNA Repair,* 56**,** 92-101.
UEMURA, T., OHKURA, H., ADACHI, Y., MORINO, K., SHIOZAKI, K. & YANAGIDA, M. 1987. DNA topoisomerase II is required for condensation and separation of mitotic chromosomes in S. pombe. *Cell,* 50**,** 917-25.
UHLMANN, F. 2016. SMC complexes: from DNA to chromosomes. *Nat Rev Mol Cell Biol,* 17**,** 399-412.
UMBREIT, N. T., ZHANG, C.-Z., LYNCH, L. D., BLAINE, L. J., CHENG, A. M., TOURDOT, R., SUN, L., ALMUBARAK, H. F., JUDGE, K., MITCHELL, T. J., SPEKTOR, A. & PELLMAN, D. 2020. Mechanisms generating cancer genome complexity from a single cell division error. *Science,* 368**,** eaba0712.
UZIEL, T., LERENTHAL, Y., MOYAL, L., ANDEGEKO, Y., MITTELMAN, L. & SHILOH, Y. 2003. Requirement of the MRN complex for ATM activation by DNA damage. *EMBO J,* 22**,** 5612-21.
VALENZUELA, M. S. 2012. Initiation of DNA Replication in the Human Genome. *Hereditary Genet,* Suppl 1.
VAN DEN BERG, D. J. & FRANCKE, U. 1993. Roberts syndrome: a review of 100 cases and a new rating system for severity. *Am J Med Genet,* 47**,** 1104-23.

VAN DER CRABBEN, S. N., HENNUS, M. P., MCGREGOR, G. A., RITTER, D. I., NAGAMANI, S. C., WELLS, O. S., HARAKALOVA, M., CHINN, I. K., ALT, A., VONDROVA, L., HOCHSTENBACH, R., VAN MONTFRANS, J. M., TERHEGGEN-LAGRO, S. W., VAN LIESHOUT, S., VAN ROOSMALEN, M. J., RENKENS, I., DURAN, K., NIJMAN, I. J., KLOOSTERMAN, W. P., HENNEKAM, E., ORANGE, J. S., VAN HASSELT, P. M., WHEELER, D. A., PALECEK, J. J., LEHMANN, A. R., OLIVER, A. W., PEARL, L. H., PLON, S. E., MURRAY, J. M. & VAN HAAFTEN, G. 2016. Destabilized SMC5/6 complex leads to chromosome breakage syndrome with severe lung disease. *J Clin Invest,* 126**,** 2881-92.

VAN DER LELIJ, P., CHRZANOWSKA, K. H., GODTHELP, B. C., ROOIMANS, M. A., OOSTRA, A. B., STUMM, M., ZDZIENICKA, M. Z., JOENJE, H. & DE WINTER, J. P. 2010a. Warsaw breakage syndrome, a cohesinopathy associated with mutations in the XPD helicase family member DDX11/ChlR1. *Am J Hum Genet,* 86**,** 262-6.

VAN DER LELIJ, P., GODTHELP, B. C., VAN ZON, W., VAN GOSLIGA, D., OOSTRA, A. B., STELTENPOOL, J., DE GROOT, J., SCHEPER, R. J., WOLTHUIS, R. M., WAISFISZ, Q., DARROUDI, F., JOENJE, H. & DE WINTER, J. P. 2009. The cellular phenotype of Roberts syndrome fibroblasts as revealed by ectopic expression of ESCO2. *PLoS One,* 4**,** e6936.

VAN DER LELIJ, P., OOSTRA, A. B., ROOIMANS, M. A., JOENJE, H. & DE WINTER, J. P. 2010b. Diagnostic Overlap between Fanconi Anemia and the Cohesinopathies: Roberts Syndrome and Warsaw Breakage Syndrome. *Anemia,* 2010**,** 565268.

VAN DYCK, E., STASIAK, A. Z., STASIAK, A. & WEST, S. C. 2001. Visualization of recombination intermediates produced by RAD52-mediated single-strand annealing. *EMBO Rep,* 2**,** 905-9.

VAN LY, D., LOW, R. R. J., FROLICH, S., BARTOLEC, T. K., KAFER, G. R., PICKETT, H. A., GAUS, K. & CESARE, A. J. 2018. Telomere Loop Dynamics in Chromosome End Protection. *Mol Cell,* 71**,** 510-525 e6.

VAN SCHIE, J. J. M., FARAMARZ, A., BALK, J. A., STEWART, G. S., CANTELLI, E., OOSTRA, A. B., ROOIMANS, M. A., PARISH, J. L., DE ALMEIDA ESTEVES, C., DUMIC, K., BARISIC, I., DIDERICH, K. E. M., VAN SLEGTENHORST, M. A., MAHTAB, M., PISANI, F. M., TE RIELE, H., AMEZIANE, N., WOLTHUIS, R. M. F. & DE LANGE, J. 2020. Warsaw Breakage Syndrome associated DDX11 helicase resolves G-quadruplex structures to support sister chromatid cohesion. *Nat Commun,* 11**,** 4287.

VAN STEENSEL, B., SMOGORZEWSKA, A. & DE LANGE, T. 1998. TRF2 protects human telomeres from end-to-end fusions. *Cell,* 92**,** 401-13.

VAN WIETMARSCHEN, N., MERZOUK, S., HALSEMA, N., SPIERINGS, D. C. J., GURYEV, V. & LANSDORP, P. M. 2018. BLM helicase suppresses recombination at G-quadruplex motifs in transcribed genes. *Nat Commun,* 9**,** 271.

VANNIER, J. B., PAVICIC-KALTENBRUNNER, V., PETALCORIN, M. I., DING, H. & BOULTON, S. J. 2012. RTEL1 dismantles T loops and counteracts telomeric G4-DNA to maintain telomere integrity. *Cell,* 149**,** 795-806.

VAS, A., MOK, W. & LEATHERWOOD, J. 2001. Control of DNA rereplication via Cdc2 phosphorylation sites in the origin recognition complex. *Mol Cell Biol,* 21**,** 5767-77.

VAZIRI, C., SAXENA, S., JEON, Y., LEE, C., MURATA, K., MACHIDA, Y., WAGLE, N., HWANG, D. S. & DUTTA, A. 2003. A p53-dependent checkpoint pathway prevents rereplication. *Mol Cell,* 11**,** 997-1008.

VEGA, H., WAISFISZ, Q., GORDILLO, M., SAKAI, N., YANAGIHARA, I., YAMADA, M., VAN GOSLIGA, D., KAYSERILI, H., XU, C., OZONO, K., JABS, E. W., INUI, K. & JOENJE, H. 2005a. Roberts syndrome is caused by mutations in ESCO2, a human homolog of yeast ECO1 that is essential for the establishment of sister chromatid cohesion. *Nat Genet,* 37**,** 468-70.

VEGA, H., WAISFISZ, Q., GORDILLO, M., SAKAI, N., YANAGIHARA, I., YAMADA, M., VAN GOSLIGA, D., KAYSERILI, H., XU, C., OZONO, K., WANG JABS, E., INUI, K. & JOENJE, H. 2005b. Roberts

syndrome is caused by mutations in ESCO2, a human homolog of yeast ECO1 that is essential for the establishment of sister chromatid cohesion. *Nature Genetics,* 37**,** 468-470.
VENEGAS, A. B., NATSUME, T., KANEMAKI, M. & HICKSON, I. D. 2020. Inducible Degradation of the Human SMC5/6 Complex Reveals an Essential Role Only during Interphase. *Cell Rep,* 31**,** 107533.
VERMA, A., YADAV, V. K., BASUNDRA, R., KUMAR, A. & CHOWDHURY, S. 2009. Evidence of genome-wide G4 DNA-mediated gene expression in human cancer cells. *Nucleic Acids Res,* 37**,** 4194-204.
VERMA, P. & GREENBERG, R. A. 2016. Noncanonical views of homology-directed DNA repair. *Genes Dev,* 30**,** 1138-54.
VILAR, E. & GRUBER, S. B. 2010. Microsatellite instability in colorectal cancer-the stable evidence. *Nat Rev Clin Oncol,* 7**,** 153-62.
VILENCHIK, M. M. & KNUDSON, A. G. 2003. Endogenous DNA double-strand breaks: production, fidelity of repair, and induction of cancer. *Proc Natl Acad Sci U S A,* 100**,** 12871-6.
VILLA, F., FUJISAWA, R., AINSWORTH, J., NISHIMURA, K., LIE, A. L. M., LACAUD, G. & LABIB, K. P. 2021. CUL2(LRR1) , TRAIP and p97 control CMG helicase disassembly in the mammalian cell cycle. *EMBO Rep,* 22**,** e52164.
WALKER, J. R., CORPINA, R. A. & GOLDBERG, J. 2001. Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature,* 412**,** 607-14.
WANG, H., NORA, G. J., GHODKE, H. & OPRESKO, P. L. 2011a. Single molecule studies of physiologically relevant telomeric tails reveal POT1 mechanism for promoting G-quadruplex unfolding. *J Biol Chem,* 286**,** 7479-89.
WANG, H., SHEN, P. & ZHU, W. 2011b. Damage and Replication Stress Responses. *In:* SELIGMANN, H. (ed.) *DNA Replication - Current Advances.*
WANG, J. & SAPONARO, M. 2021. Protocol for analysis of G2/M DNA synthesis in human cells. *STAR Protoc,* 2**,** 100570.
WARMERDAM, D. O., VAN DEN BERG, J. & MEDEMA, R. H. 2016. Breaks in the 45S rDNA Lead to Recombination-Mediated Loss of Repeats. *Cell Rep,* 14**,** 2519-27.
WARMERDAM, D. O. & WOLTHUIS, R. M. F. 2019. Keeping ribosomal DNA intact: a repeating challenge. *Chromosome Res,* 27**,** 57-72.
WATERS, L. S., MINESINGER, B. K., WILTROUT, M. E., D'SOUZA, S., WOODRUFF, R. V. & WALKER, G. C. 2009. Eukaryotic translesion polymerases and their roles and regulation in DNA damage tolerance. *Microbiol Mol Biol Rev,* 73**,** 134-54.
WECHSLER, T., NEWMAN, S. & WEST, S. C. 2011. Aberrant chromosome morphology in human cells defective for Holliday junction resolution. *Nature,* 471**,** 642-6.
WIEDENHEFT, B., STERNBERG, S. H. & DOUDNA, J. A. 2012. RNA-guided genetic silencing systems in bacteria and archaea. *Nature,* 482**,** 331-8.
WILHELM, T., MAGDALOU, I., BARASCU, A., TÉCHER, H., DEBATISSE, M. & LOPEZ, B. S. 2014. Spontaneous slow replication fork progression elicits mitosis alterations in homologous recombination-deficient mammalian cells. *Proc Natl Acad Sci U S A,* 111**,** 763-8.
WILHELM, T., OLZIERSKY, A.-M., HARRY, D., DE SOUSA, F., VASSAL, H., ESKAT, A. & MERALDI, P. 2019a. Mild replication stress causes chromosome mis-segregation via premature centriole disengagement. *Nature Communications,* 10**,** 3585.
WILHELM, T., OLZIERSKY, A. M., HARRY, D., DE SOUSA, F., VASSAL, H., ESKAT, A. & MERALDI, P. 2019b. Mild replication stress causes chromosome mis-segregation via premature centriole disengagement. *Nat Commun,* 10**,** 3585.
WILHELM, T., SAID, M. & NAIM, V. 2020. DNA Replication Stress and Chromosomal Instability: Dangerous Liaisons. *Genes (Basel),* 11.
WILLEMSE, B. W. M., VAN DER CRABBEN, S. N., KERSTJENS-FREDERIKSE, W. S., TIMENS, W., VAN MONTFRANS, J. M., LINDEMANS, C. A., BOELENS, J. J., HENNUS, M. P. & VAN HAAFTEN, G.

2021. New insights in phenotype and treatment of lung disease immuno-deficiency and chromosome breakage syndrome (LICS). *Orphanet J Rare Dis,* 16**,** 137.

WILLIAMS, B. R., PRABHU, V. R., HUNTER, K. E., GLAZIER, C. M., WHITTAKER, C. A., HOUSMAN, D. E. & AMON, A. 2008a. Aneuploidy affects proliferation and spontaneous immortalization in mammalian cells. *Science,* 322**,** 703-9.

WILLIAMS, R. S., MONCALIAN, G., WILLIAMS, J. S., YAMADA, Y., LIMBO, O., SHIN, D. S., GROOCOCK, L. M., CAHILL, D., HITOMI, C., GUENTHER, G., MOIANI, D., CARNEY, J. P., RUSSELL, P. & TAINER, J. A. 2008b. Mre11 dimers coordinate DNA end bridging and nuclease processing in double-strand-break repair. *Cell,* 135**,** 97-109.

WIRTH, K. G., WUTZ, G., KUDO, N. R., DESDOUETS, C., ZETTERBERG, A., TAGHYBEEGLU, S., SEZNEC, J., DUCOS, G. M., RICCI, R., FIRNBERG, N., PETERS, J. M. & NASMYTH, K. 2006. Separase: a universal trigger for sister chromatid disjunction but not chromosome cycle progression. *J Cell Biol,* 172**,** 847-60.

WOLSKA-KUSNIERZ, B., GREGOREK, H., CHRZANOWSKA, K., PIATOSA, B., PIETRUCHA, B., HEROPOLITANSKA-PLISZKA, E., PAC, M., KLAUDEL-DRESZLER, M., KOSTYUCHENKO, L., PASIC, S., MARODI, L., BELOHRADSKY, B. H., CIZNAR, P., SHCHERBINA, A., KILIC, S. S., BAUMANN, U., SEIDEL, M. G., GENNERY, A. R., SYCZEWSKA, M., MIKOLUC, B., KALWAK, K., STYCZYNSKI, J., PIECZONKA, A., DRABKO, K., WAKULINSKA, A., GATHMANN, B., ALBERT, M. H., SKARZYNSKA, U., BERNATOWSKA, E., INBORN ERRORS WORKING PARTY OF THE SOCIETY FOR EUROPEAN, B., MARROW, T. & THE EUROPEAN SOCIETY FOR IMMUNE, D. 2015. Nijmegen Breakage Syndrome: Clinical and Immunological Features, Long-Term Outcome and Treatment Options - a Retrospective Analysis. *J Clin Immunol,* 35**,** 538-49.

WOODWARD, A. M., GOHLER, T., LUCIANI, M. G., OEHLMANN, M., GE, X., GARTNER, A., JACKSON, D. A. & BLOW, J. J. 2006. Excess Mcm2-7 license dormant origins of replication that can be used under conditions of replicative stress. *J Cell Biol,* 173**,** 673-83.

WRIGHT, W. D., SHAH, S. S. & HEYER, W. D. 2018. Homologous recombination and the repair of DNA double-strand breaks. *J Biol Chem,* 293**,** 10524-10535.

WU, F., NIU, K., CUI, Y., LI, C., LYU, M., REN, Y., CHEN, Y., DENG, H., HUANG, L., ZHENG, S., LIU, L., WANG, J., SONG, Q., XIANG, H. & FENG, Q. 2021. Genome-wide analysis of DNA G-quadruplex motifs across 37 species provides insights into G4 evolution. *Commun Biol,* 4**,** 98.

WU, L., CHAN, K. L., RALF, C., BERNSTEIN, D. A., GARCIA, P. L., BOHR, V. A., VINDIGNI, A., JANSCAK, P., KECK, J. L. & HICKSON, I. D. 2005. The HRDC domain of BLM is required for the dissolution of double Holliday junctions. *EMBO J,* 24**,** 2679-87.

WU, L. & HICKSON, I. D. 2003. The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature,* 426**,** 870-4.

WU, Y., SOMMERS, J. A., KHAN, I., DE WINTER, J. P. & BROSH, R. M., JR. 2012. Biochemical characterization of Warsaw breakage syndrome helicase. *J Biol Chem,* 287**,** 1007-21.

WYATT, H. D., SARBAJNA, S., MATOS, J. & WEST, S. C. 2013. Coordinated actions of SLX1-SLX4 and MUS81-EME1 for Holliday junction resolution in human cells. *Mol Cell,* 52**,** 234-47.

WYATT, H. D. & WEST, S. C. 2014. Holliday junction resolvases. *Cold Spring Harb Perspect Biol,* 6**,** a023192.

XIAO, Z., CHEN, Z., GUNASEKERA, A. H., SOWIN, T. J., ROSENBERG, S. H., FESIK, S. & ZHANG, H. 2003. Chk1 mediates S and G2 arrests through Cdc25A degradation in response to DNA-damaging agents. *J Biol Chem,* 278**,** 21767-73.

YAMAMOTO, K. N., KOBAYASHI, S., TSUDA, M., KURUMIZAKA, H., TAKATA, M., KONO, K., JIRICNY, J., TAKEDA, S. & HIROTA, K. 2011. Involvement of SLX4 in interstrand cross-link repair is regulated by the Fanconi anemia pathway. *Proc Natl Acad Sci U S A,* 108**,** 6492-6.

YANG, S. W., BURGIN, A. B., JR., HUIZENGA, B. N., ROBERTSON, C. A., YAO, K. C. & NASH, H. A. 1996. A eukaryotic enzyme that can disjoin dead-end covalent complexes between DNA and type I topoisomerases. *Proc Natl Acad Sci U S A,* 93**,** 11534-9.

YANO, K., MOROTOMI-YANO, K., ADACHI, N. & AKIYAMA, H. 2009. Molecular mechanism of protein assembly on DNA double-strand breaks in the non-homologous end-joining pathway. *J Radiat Res,* 50**,** 97-108.

YAO, N. Y. & O'DONNELL, M. 2012. The RFC clamp loader: structure and function. *Subcell Biochem,* 62**,** 259-79.

YATSKEVICH, S., RHODES, J. & NASMYTH, K. 2019. Organization of Chromosomal DNA by SMC Complexes. *Annual Review of Genetics,* 53**,** 445-482.

YEAGER, T. R., NEUMANN, A. A., ENGLEZOU, A., HUSCHTSCHA, L. I., NOBLE, J. R. & REDDEL, R. R. 1999. Telomerase-negative immortalized human cells contain a novel type of promyelocytic leukemia (PML) body. *Cancer Res,* 59**,** 4175-9.

YEELES, J. T. P., JANSKA, A., EARLY, A. & DIFFLEY, J. F. X. 2017. How the Eukaryotic Replisome Achieves Rapid and Efficient DNA Replication. *Mol Cell,* 65**,** 105-116.

YEKEZARE, M., GOMEZ-GONZALEZ, B. & DIFFLEY, J. F. 2013. Controlling DNA replication origins in response to DNA damage - inhibit globally, activate locally. *J Cell Sci,* 126**,** 1297-306.

YOON, J. H., PRAKASH, L. & PRAKASH, S. 2009. Highly error-free role of DNA polymerase eta in the replicative bypass of UV-induced pyrimidine dimers in mouse and human cells. *Proc Natl Acad Sci U S A,* 106**,** 18219-24.

YOST, S., DE WOLF, B., HANKS, S., ZACHARIOU, A., MARCOZZI, C., CLARKE, M., DE VOER, R., ETEMAD, B., UIJTTEWAAL, E., RAMSAY, E., WYLIE, H., ELLIOTT, A., PICTON, S., SMITH, A., SMITHSON, S., SEAL, S., RUARK, E., HOUGE, G., PINES, J., KOPS, G. & RAHMAN, N. 2017. Biallelic TRIP13 mutations predispose to Wilms tumor and chromosome missegregation. *Nat Genet,* 49**,** 1148-1151.

YU, Y., LI, S., SER, Z., KUANG, H., THAN, T., GUAN, D., ZHAO, X. & PATEL, D. J. 2022. Cryo-EM structure of DNA-bound Smc5/6 reveals DNA clamping enabled by multi-subunit conformational changes. *Proc Natl Acad Sci U S A,* 119**,** e2202799119.

YU, Y., LI, S., SER, Z., SANYAL, T., CHOI, K., WAN, B., KUANG, H., SALI, A., KENTSIS, A., PATEL, D. J. & ZHAO, X. 2021. Integrative analysis reveals unique structural and functional features of the Smc5/6 complex. *Proc Natl Acad Sci U S A,* 118.

YUE, X., BAI, C., XIE, D., MA, T. & ZHOU, P. K. 2020. DNA-PKcs: A Multi-Faceted Player in DNA Damage Response. *Front Genet,* 11**,** 607428.

ZABRADY, K., ADAMUS, M., VONDROVA, L., LIAO, C., SKOUPILOVA, H., NOVAKOVA, M., JURCISINOVA, L., ALT, A., OLIVER, A. W., LEHMANN, A. R. & PALECEK, J. J. 2016. Chromatin association of the SMC5/6 complex is dependent on binding of its NSE3 subunit to DNA. *Nucleic Acids Res,* 44**,** 1064-79.

ZAPOTOCZNY, G. & SEKELSKY, J. 2017. Human Cell Assays for Synthesis-Dependent Strand Annealing and Crossing over During Double-Strand Break Repair. *G3 (Bethesda),* 7**,** 1191-1199.

ZAUG, A. J., PODELL, E. R. & CECH, T. R. 2005. Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension in vitro. *Proc Natl Acad Sci U S A,* 102**,** 10864-9.

ZELLWEGER, R., DALCHER, D., MUTREJA, K., BERTI, M., SCHMID, J. A., HERRADOR, R., VINDIGNI, A. & LOPES, M. 2015. Rad51-mediated replication fork reversal is a global response to genotoxic treatments in human cells. *J Cell Biol,* 208**,** 563-79.

ZEMAN, M. K. & CIMPRICH, K. A. 2014. Causes and consequences of replication stress. *Nat Cell Biol,* 16**,** 2-9.

ZENG, Y., FORBES, K. C., WU, Z., MORENO, S., PIWNICA-WORMS, H. & ENOCH, T. 1998. Replication checkpoint requires phosphorylation of the phosphatase Cdc25 by Cds1 or Chk1. *Nature,* 395**,** 507-10.

ZHANG, J.-Y., ZHENG, K.-W., XIAO, S., HAO, Y.-H. & TAN, Z. 2014. Mechanism and Manipulation of DNA:RNA Hybrid G-Quadruplex Formation in Transcription of G-Rich DNA. *Journal of the American Chemical Society,* 136**,** 1381-1390.

ZHANG, J., DEWAR, J. M., BUDZOWSKA, M., MOTNENKO, A., COHN, M. A. & WALTER, J. C. 2015a. DNA interstrand cross-link repair requires replication-fork convergence. *Nat Struct Mol Biol,* 22**,** 242-7.

ZHANG, M., WANG, B., LI, T., LIU, R., XIAO, Y., GENG, X., LI, G., LIU, Q., PRICE, C. M., LIU, Y. & WANG, F. 2019. Mammalian CST averts replication failure by preventing G-quadruplex accumulation. *Nucleic Acids Res,* 47**,** 5243-5259.

ZHANG, X. H., TEE, L. Y., WANG, X. G., HUANG, Q. S. & YANG, S. H. 2015b. Off-target Effects in CRISPR/Cas9-mediated Genome Engineering. *Mol Ther Nucleic Acids,* 4**,** e264.

ZHANG, Y. & HUNTER, T. 2014. Roles of Chk1 in cell biology and cancer therapy. *Int J Cancer,* 134**,** 1013-23.

ZHAO, M. T., YE, S., SU, J. & GARG, V. 2020. Cardiomyocyte Proliferation and Maturation: Two Sides of the Same Coin for Heart Regeneration. *Front Cell Dev Biol,* 8**,** 594226.

ZHAO, X. & BLOBEL, G. 2005. A SUMO ligase is part of a nuclear multiprotein complex that affects DNA repair and chromosomal organization. *Proc Natl Acad Sci U S A,* 102**,** 4777-82.

ZHENG, L. & SHEN, B. 2011. Okazaki fragment maturation: nucleases take centre stage. *J Mol Cell Biol,* 3**,** 23-30.

ZHOU, Y., POZO, P. N., OH, S., STONE, H. M. & COOK, J. G. 2020. Distinct and sequential re-replication barriers ensure precise genome duplication. *PLoS Genet,* 16**,** e1008988.

ZHOU, Z. W., LIU, C., LI, T. L., BRUHN, C., KRUEGER, A., MIN, W., WANG, Z. Q. & CARR, A. M. 2013. An essential function for the ATR-activation-domain (AAD) of TopBP1 in mouse development and cellular senescence. *PLoS Genet,* 9**,** e1003702.

ZLOTORYNSKI, E., RAHAT, A., SKAUG, J., BEN-PORAT, N., OZERI, E., HERSHBERG, R., LEVI, A., SCHERER, S. W., MARGALIT, H. & KEREM, B. 2003. Molecular basis for expression of common and rare fragile sites. *Mol Cell Biol,* 23**,** 7143-51.

ZOU, L. & ELLEDGE, S. J. 2003. Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science,* 300**,** 1542-8.

Published Work

nature communications

Article

https://doi.org/10.1038/s41467 022 34349 8

Pathogenic variants in *SLF2* and *SMC5* cause segmented chromosomes and mosaic variegated hyperploidy

Received: 11 November 2021

Accepted: 21 October 2022

Published online: 04 November 2022

Check for updates

A list of authors and their affiliations appears at the end of the paper

Embryonic development is dictated by tight regulation of DNA replication, cell division and differentiation. Mutations in DNA repair and replication genes disrupt this equilibrium, giving rise to neurodevelopmental disease characterized by microcephaly, short stature and chromosomal breakage. Here, we identify biallelic variants in two components of the RAD18-SLF1/2-SMC5/6 genome stability pathway, *SLF2* and *SMC5*, in 11 patients with microcephaly, short stature, cardiac abnormalities and anemia. Patient-derived cells exhibit a unique chromosomal instability phenotype consisting of segmented and dicentric chromosomes with mosaic variegated hyperploidy. To signify the importance of these segmented chromosomes, we have named this disorder Atelís (meaning - incomplete) Syndrome. Analysis of Atelís Syndrome cells reveals elevated levels of replication stress, partly due to a reduced ability to replicate through G-quadruplex DNA structures, and also loss of sister chromatid cohesion. Together, these data strengthen the functional link between SLF2 and the SMC5/6 complex, highlighting a distinct role for this pathway in maintaining genome stability.

Despite the fundamental nature of DNA replication and cell division, inherited variants in genes involved in these processes are an underlying cause of human disease. Whilst these syndromes usually display unique clinical features that define them diagnostically, they typically exhibit common neurodevelopmental deficits, such as severe microcephaly and pre- and post-natal growth retardation[1–3]. As such, many of these syndromes can be collectively referred to as microcephalic dwarfism (MD) disorders. This constellation of conditions includes Meier-Gorlin Syndrome, Seckel Syndrome Spectrum Disorders, Bloom Syndrome and Microcephalic Osteodysplastic Primordial Dwarfism type II and can be broadly classified as having deficiencies in one of three cellular processes: DNA replication, DNA repair, and mitotic cell division[1–4]. Although mechanistically distinct, the common clinical phenotypes exhibited by these diseases are thought to result from a reduction in cellular proliferation and/or excessive cell death in the developing embryo, which reduces the number of cells available to maintain normal foetal growth[5]. Cells from these patients often exhibit signs of increased genome instability, such as micronuclei and/or elevated chromosome breakage. A distinct subgroup of these syndromes exhibit rare cytogenetic anomalies, for example, mosaic variegated aneuploidy syndrome (MVA)[6–8] caused by variants in the spindle assembly checkpoint genes *BUB1B*, *CEP57* and *TRIP13*, or railroad chromosomes and premature chromatid separation (PCS) associated with Warsaw Breakage Syndrome (WABS) and Cornelia de Lange syndrome, caused by variants in the helicase DDX11 and components of SMC1/3 cohesin complex respectively[9,10]. Whilst, the presence of these chromosomal abnormalities is a useful diagnostic tool they can also help dissect the cellular mechanisms underlying the disease pathology.

Here, we report 11 patients with a neurodevelopmental disorder overlapping clinically with MVA and Fanconi Anemia (FA) with pathogenic variants in SLF2 and SMC5, two components of the recently discovered RAD18-SLF1/2-SMC5/6 genome stability pathway[11]. The precise function of the SMC5/6 complex remains enigmatic, however, it has been linked to a number of fundamental processes, including DNA transcription, DNA replication, DNA repair and

e-mail:

chromosome segregation[12,13]. Evidence suggests that the primary function of this complex occurs during DNA replication to stabilize stalled forks, suppress the activity of pro-recombination factors and promote efficient replication through difficult-to-replicate and/or repetitive regions of the genome, such as rDNA and telomeres[14]. In contrast, the function of SLF1 and SLF2 remain unclear, other than a reported role in recruiting the SMC5/6 complex to sites of DNA damage[11].

Analysis of SLF2 and SMC5 patient-derived cell lines revealed spontaneous replication stress and multiple mitotic abnormalities that give rise to a unique, diagnostically relevant, genome instability phenotype consisting of segmented, dicentric and rail-road chromosomes, and mosaic variegated hyperploidy (MVH). The underlying basis for this chromosomal instability is not fully understood, but our data suggest that it may arise, in part, from the failed resolution of aberrant DNA structures during S-phase, such as G-quadruplexes (G4), potentially leading to a combination of under-replicated DNA and unresolved recombination intermediates persisting through to mitosis. Together, these data demonstrate that despite a hitherto unknown role as a core component of the SMC5/6 complex, SLF2 is essential for the SMC5/6 cohesin-like complex to maintain genome stability by regulating both DNA replication and cell division.

Results

Patients with microcephaly and short stature have biallelic *SLF2* (*FAM178A*) and *SMC5* variants

Whole exome sequencing (WES) was carried out on seven patients (P1, P2, P3, P4-1, P4-2, P5 and P6) from five families, presenting with microcephaly, short stature, mild to severe developmental delay and spontaneous chromosome breakage. After aligning WES reads to the reference genome, variant calling, and filtering for rare variants (MAF <0.005), analysis under a recessive model of inheritance identified biallelic variants in *SLF2* (*FAM178A*) in all seven patients. All identified *SLF2* variants segregated amongst family members (with the exception of patients P1 and P5 where parental material was unavailable) and were present at a frequency of <0.5% in the gnomAD database (Fig. 1a, c; Supplementary Data 1–7; Supplementary Fig. 1a). Comparative genomic hybridization (CGH) array analysis carried out on gDNA from patient P5 confirmed the homozygosity of the identified *SLF2* variant.

Given that SLF2 had been identified previously as part of the RAD18-SLF1/2-SMC5/6 genome stability pathway[11], we hypothesized that variants in other components of this pathway may also give rise to a similar neurodevelopmental disorder. By querying gene matching platforms, four patients exhibiting microcephaly and growth retardation that had undergone WES were identified to carry biallelic variants in *SMC5*: patient P7 (c.1110 1112del; p.Arg372del, c.1273C>T; p.Arg425Ter) and patients P8, P9-1 and P9-2 (c.2970C>G; p.His990Asp) (Fig. 1a, c; Supplementary Data 1; Supplementary Data 8–10; Supplementary Fig. 1b). All variants were verified by Sanger sequencing, segregated amongst family members in an autosomal recessive paradigm and were present at a frequency of <0.5% in gnomAD.

SLF2 and *SMC5* variants give rise to neurodevelopmental abnormalities, cardiac defects and anemia

All individuals with *SLF2* and *SMC5* variants presented with a similar clinical phenotype, including marked microcephaly (−3.57 to −11.88 SD) and a reduction in height (-2.19 to -8.24 SD) (Fig. 1b; Supplementary Data 1). Moreover, the majority of patients also exhibited a developmental delay along with learning difficulties. Mild skeletal defects (i.e., clinodactyly), skin hyperpigmentation and ocular abnormalities were present in several patients (Supplementary Data 1). Notably, two of seven SLF2 patients (P4-1, P5) and all four SMC5 patients (P7, P8, P9-1 and P9-2) displayed cardiac defects (Supplementary Data 1), such as atrial or ventricular defects, a phenotype commonly observed in patients with cohesinopathies[15,16] but not DNA replication disorders. Furthermore, five of eleven patients (P3, P4-1, P4-2, P5, P9-2) also developed anemia, with one of these patients (P9-2) subsequently developing myelodysplastic syndrome (Supplementary Data 1). This, coupled with other clinical features, could potentially result in future cases being mistakenly diagnosed with an atypical form of FA in the absence of a clear genetic diagnosis using WES. This is particularly relevant since components of the SMC5/6 complex have been previously shown to functionally interact with the FA pathway to repair DNA damage[17]. Only one patient (P3) developed severe pulmonary disease similar to patients with variants in the SMC5/6 complex subunit NSMCE3[18,19], whereas insulin-resistant diabetes and metabolic dysfunction, which are characteristic to patients with *NSMCE2* variants were absent among this cohort[20]. Collectively, these clinical and genetic observations support the premise that variants in *SLF2* and *SMC5* cause microcephaly and short stature associated with cardiac defects and the development of anemia.

SLF2 and *SMC5* variants compromise protein stability, interactions with other components of the RAD18-SLF1/2-SMC5/6 pathway and recruitment to sites of DNA damage

To determine the pathogenicity of the identified patient variants, we carried out western blotting on extracts from SLF2 patient-derived cell lines (SLF2-P1, SLF2-P2, SLF2-P3 and SLF2-P4-1) to ascertain if SLF2 protein abundance or stability was compromised. Notably, all four of the SLF2-mutant patient cell lines examined exhibited a reduction or absence of detectable full length SLF2 protein whilst maintaining wild type (WT) levels of RAD18, SMC5, and SMC6 protein (Fig. 2a). SLF1 protein level was not tested due to the absence of an available antibody.

We next investigated the *SLF2* variants in patients P2 and P3 in more detail. Analysis of cDNA from the SLF2-P3 cell line demonstrated that the synonymous homozygous variant c.3330G>A (p.Arg1110Arg), disrupted splicing leading to an in-frame deletion of exon 17 (Supplementary Fig. 2a, b). We then analysed the impact of the c.3486G>C (p.Gln1162His) variant, present in patient P2, on splicing. Multiple *SLF2* transcripts are annotated in the human genome and although c.3486G>C (p.Gln1162His) introduces a nonsynonymous change in the two longest transcripts (NM 018121 and NM 001136123), it only affects mRNA splicing of the most abundant *SLF2* transcript (NM 018121) by impairing the exon 19 splice donor splice site (Supplementary Figs. 2c, 3a–e). The p.(Gln1162His) variant also displayed compromised protein stability when expressed transiently indicating that this variant disrupts both mRNA and protein stability (Supplementary Fig. 3f). Together, these data suggest that most of the identified *SLF2* variants have an adverse effect on protein stability.

In contrast, analysis of SMC5 patient cell lines revealed that the homozygous p.(His990Asp) variant present in patients P8, P9-1 and P9-2 had little detectable impact on the protein stability of SMC5, or RAD18, SLF2, and SMC6 (Fig. 2b). Only a cell line derived from patient P7 exhibited a reduced abundance of SMC5 protein, presumably due to the presence of a nonsense variant (p.Arg425Ter) on one of the *SMC5* alleles. As loss of Smc5 is embryonically lethal[21], it is possible that the *SMC5* variants are hypomorphic and that significant disruption of SMC5 protein stability to the extent observed with the SLF2 variants is incompatible with life.

SLF1 and SLF2 have been identified as bridging factors between RAD18 and the SMC5/6 complex at sites of stalled replication[11]. To address whether the SLF2 and SMC5 variants compromised their ability to bind components of the RAD18-SLF1/2-SMC5/6 pathway, we initially mapped the binding sites of RAD18, SLF1 and SMC6 on SLF2. Using co-immunoprecipitation analysis with tagged proteins, we determined that the binding of RAD18 and SLF1 to SLF2 requires the C-terminal 471 amino acids (aa702-1173), which also overlapped with the SMC6 binding site located at amino acids 589–810 (Supplementary

a

Individual	Ancestry	Gene	Mutation 1	Allele Frequency	Polyphen-2 score	Mutation 2	Allele Frequency	Polyphen-2 score
P1	UK	*SLF2*	c.1006dup; p.(Arg336LysfsTer27)	-	NA	c.1006dup; p.(Arg336LysfsTer27)	-	NA
P2	France	*SLF2*	c.2444C>G; p.(Ser815Ter)	-	NA	c.3486G>C (ss); p.(Gln1162His)	-	0.999
P3	Netherlands	*SLF2*	c.3330G>A (ss) p.(Arg1110ArgΔexon17)	-	NA	c.3330G>A (ss); p.(Arg1110ArgΔexon17)	-	NA
P4-1	Japan	*SLF2*	c.2582A>T; p.(Asn861Ile)	3.98e-6	0.996	c.2719dup; p.(Ser907PhefsTer5)	-	NA
P4-2	Japan	*SLF2*	c.2582A>T; p.(Asn861Ile)	3.98e-6	0.966	c.2719dup; p.(Ser907PhefsTer5)	-	NA
P5	Saudi Arabia	*SLF2*	c.2347_2348del; p.(Asp783SerfsTer53)	-	NA	c.2347_2348del; p.(Asp783SerfsTer53)	-	NA
P6	German	*SLF2*	c.568C>T; p.(Arg190Ter)	3.19e-5	NA	c.568C>T; p.(Arg190Ter)	3.19e-5	NA
P7	Spain	*SMC5*	c.1110_1112del; p.(Arg372del)	-	NA	c.1273C>T; p.(Arg425Ter)	3.99e-6	NA
P8	US	*SMC5*	c.2970C>G; p.(His990Asp)	-	1.000	c.2970C>G; p.(His990Asp)	-	1.000
P9-1	US	*SMC5*	c.2970C>G; p.(His990Asp)	-	1.000	c.2970C>G; p.(His990Asp)	-	1.000
P9-2	US	*SMC5*	c.2970C>G; p.(His990Asp)	-	1.000	c.2970C>G; p.(His990Asp)	-	1.000

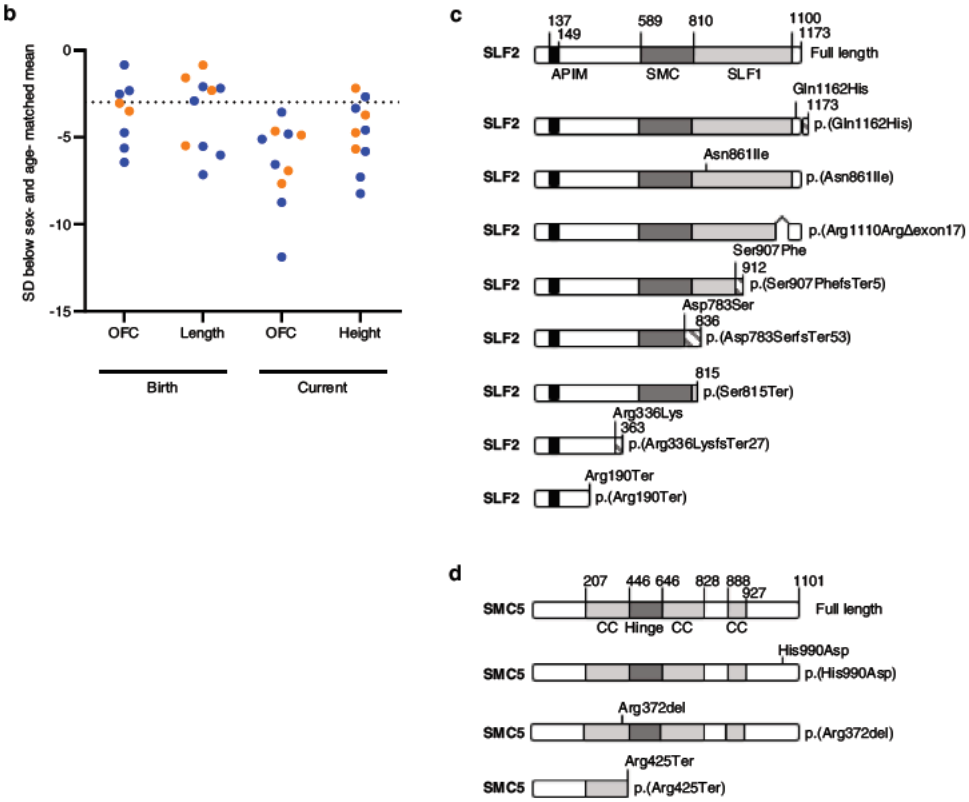


Fig. 4a–d). All patient-associated variants in SLF2, with the exception of p.(Gln1162His), are located within or truncate the SLF1/RAD18 binding domain of SLF2 (Fig. 1c). Consistent with SLF1 binding being essential for SLF2 to mediate bridging between RAD18 and the SMC5/6 complex, co-immunoprecipitation studies using extracts from hydroxyurea (HU) treated SLF2 patient-derived LCLs revealed a failure of all cell lines tested to co-purify SMC6 with RAD18 (Fig. 2c). Furthermore, all SLF2-mutant proteins, with the exception of p.(Gln1162His), failed to or exhibited a reduced ability to, be recruited to sites of DNA damage induced by laser micro-irradiation (Supplementary Fig. 4e).

We next extended the co-immunoprecipitation analysis to include SMC5 patient LCLs (Fig. 2d). The interaction between RAD18 and SMC6

Fig. 1 | *SLF2* and *SMC5* variants cause severe microcephaly and short stature. **a** Table listing biallelic *SLF2* and *SMC5* variants in 11 individuals. ss, splice site cre ated or destroyed by variant. ‘ ’ denotes that the allele variant was not present in the gnomAD database. Scores predicting the pathogenicity of the identified mis sense variants in *SLF2* and *SMC5* were generated using Polyphen 2 (http://genetics. bwh.harvard.edu/pph2/). NA Not applicable. **b** Length and head circumference (occipital frontal circumference; OFC) at birth and at the age of last exam as z scores (s.d. from population mean for age and sex; SD). Dashed line at 3 SD indicates cut off for normal population distribution. Orange values indicate SMC5 patients and blue values indicate SLF2 patients. **c** Schematic of full length WT SLF2 protein and SLF2 patient variants. APIM, atypical PCNA binding motif. SMC, SMC5/6 binding region. SLF1, SLF1 binding region. **d** Schematic of full length WT SMC5 protein and SMC5 patient variants. CC coiled coil region.

in SMC5-P8 and SMC5-P9-1 cells was observed to be at WT levels, suggesting that p.(His990Asp) had no discernible impact on the integrity of the RAD18-SLF1/2-SMC5/6 complex, whereas the association of RAD18 with SMC6 was partially affected in SMC5-P7 cells. However, both the p.(Arg372del) and p.(His990Asp) SMC5 mutants failed to re-localize efficiently to sites of laser micro-irradiation induced damage, with the latter being more severely affected (Supplementary Fig. 4f). These observations indicate that whilst these variants largely do not appear to compromise their binding to components of the RAD18-SLF1-SLF2-SMC5/6 pathway, they do affect their re-localization to and/or retention at sites of damage.

To gain insight into why the SMC5 mutants affected stability of the SMC5/6 complex at sites of damage, we carried out co-immunoprecipitation analysis to assess if these mutations affected binding to other components of the complex. Interestingly, whilst the p.(His990Asp) mutation did not significantly affect binding to other components of the SMC5/6 complex, the p.(Arg372del) significantly compromised binding to SLF2, SMC6 and NSMCE2 (Fig. 2e). Moreover, endogenous NSMCE2 exhibited reduced binding to SMC5 in cells from patient SMC5-P7 (Fig. 2f). Consistent with these observations, the Nse2 binding site on yeast Smc5 lies in close proximity to Lys368, which is the yeast functional equivalent of human SMC5 Arg372 (Supplementary Fig. 5). This suggests that the failure of the p.(Arg372del) mutant SMC5 to be recruited to sites of laser damage may be due to this mutation compromising the binding of other key components of the SMC5/6 complex.

To explore the possibility that the p.(His990Asp) may have a deleterious impact on the structure of the SMC5/6 complex, we compared the AlphaFold model for human SMC5 to the X-ray crystal structures for the head domain of *Pyrococcus furiosus* Rad50 (*Pf*.Rad50) in both the unliganded and ATP-bound forms[22]. Notably, His990 lies just upstream of the ATP-binding cassette (ABC) signature motif of Smc5 (Supplementary Fig. 6a), a region of the protein implicit in both binding ATP and mediating the complex set of conformational changes that occur when SMC proteins bind nucleotide[23]. Interestingly, His990 sits in a position functionally equivalent to Phe791 of *Pf*.Rad50 - a residue known to interact directly with the adenine moiety of bound ATP[22]. Whilst mutation of His990 to aspartic acid would appear to be tolerated and unlikely to cause any gross-misfolding of the protein, as judged by the lack of steric clashes produced by the mutation (Supplementary Fig. 6b), it removes an aromatic amino acid and replaces it with one carrying a negative charge. As such, this would alter the overall charge of a region that normally functions to accept the adenine moiety. Therefore, it is likely that the p.(His990Asp) mutation perturbs the ability of the complex to either bind or turnover ATP, in turn affecting its association with, or retention on chromatin[24].

Cell cycle arrest and increased apoptosis in the developing brain underlies the development of microcephaly in zebrafish lacking *slf2* and *smc5*

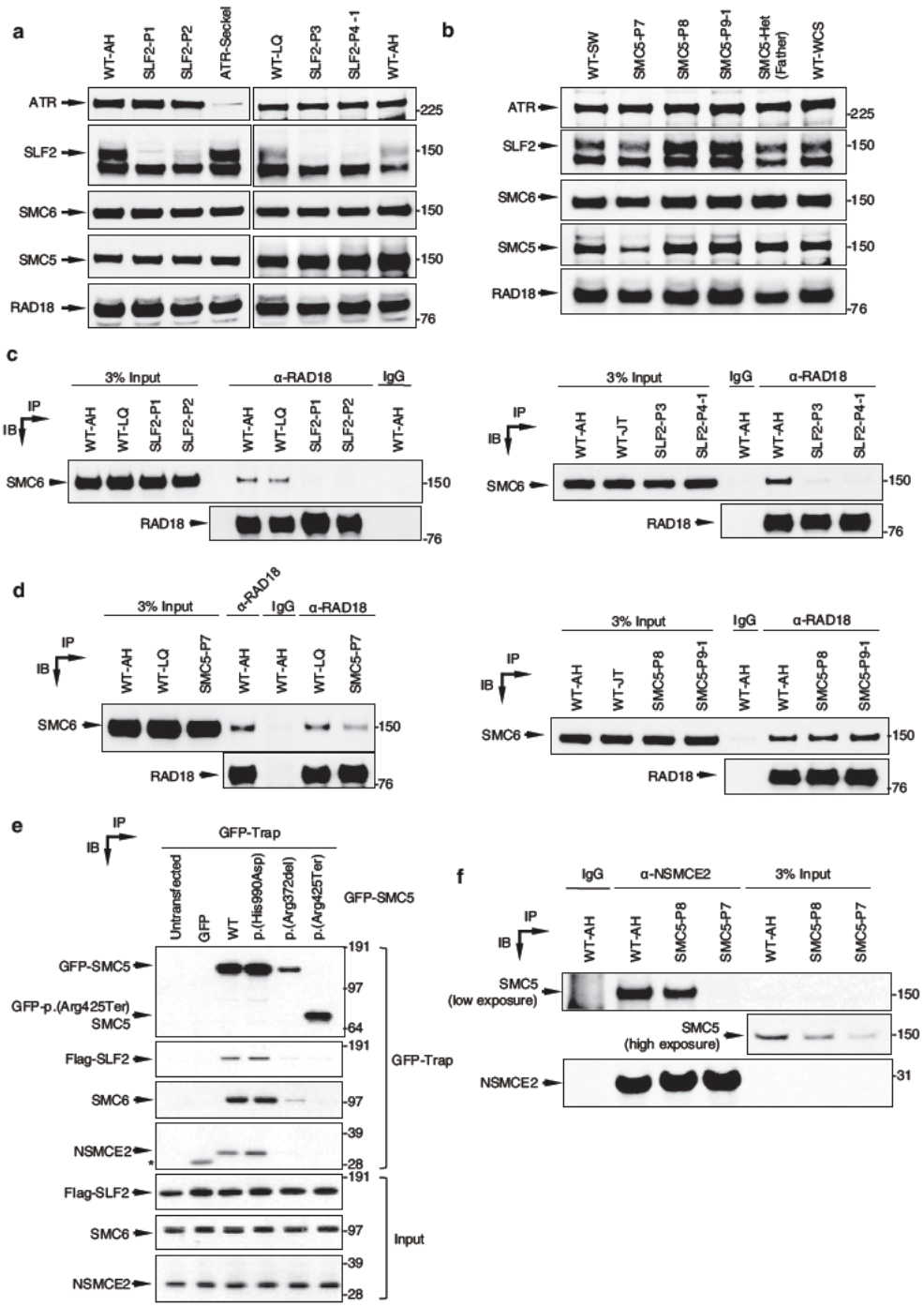
To gain insight into how SLF2 and SMC5 patient-associated variants affect neurodevelopment, we utilized CRISPR-Cas9 genome editing to ablate the single zebrafish orthologs of each of *slf2* and *smc5* in zebrafish embryos. Single guide (sg) RNAs targeting the primary isoforms of *slf2* and *smc5* (Supplementary Fig. 7a, f) were injected, with or without recombinant Cas9 protein, into *-1.4col1a1:egfp* reporter embryos at the single-cell stage, which were allowed to develop until 3 days post-fertilization (dpf) (Supplementary Fig. 7b, c, g, h). This reporter allows visualization of craniofacial patterning during embryonal development[25]. Bright field lateral images were acquired to measure head size and ventral fluorescent images of GFP-positive cells allowed visualization of the pharyngeal skeleton. Similar to the clinical phenotype exhibited by SLF2 and SMC5 patients, zebrafish embryos lacking *slf2* and *smc5* displayed a significant reduction in head size and aberrant craniofacial patterning, as indicated by a broadening of the angle of the ceratohyal cartilage; a major mandibular structure (Fig. 3a–f). Furthermore, unlike *smc5*, which is an essential gene[21], we were able to generate stable F2 *slf2* mutants possessing a frameshifting 8 bp deletion allele in *slf2* exon 7 (c.515 522del; p.Ser172 Ser174fsTer191; Supplementary Fig. 7d, e). Consistent with our observations from F0 embryos injected with sgRNA and Cas9, stable F2 *slf2* null mutants also exhibited microcephaly and aberrant craniofacial patterning (Fig. 3g).

To validate these findings, we used morpholinos (MO) to suppress the expression of *slf2* and *smc5* in zebrafish embryos. Splice blocking MO targeting the single zebrafish ortholog of each gene, *slf2* (exon 11) and *smc5* (exon 3), were designed and depletion of *slf2* and *smc5* mRNA was confirmed by RT-PCR after injection into zebrafish larvae (Supplementary Fig. 8a–d). MO were injected into *-1.4col1a1:egfp* reporter embryos at the single-cell stage. Injected embryos were reared to 3 dpf and then bright field images were acquired to measure head size and ventral fluorescent images of GFP-positive cells to visualize the pharyngeal skeleton. Comparable to our observations from the zebrafish embryos lacking *slf2* and *smc5*, zebrafish embryos depleted of *slf2* and *smc5* using MO also displayed a significant reduction in head size and aberrant craniofacial patterning in the pharyngeal skeleton (Supplementary Figs. 8e–h, 9a–f), which could both be rescued by re-expression of WT human *SLF2* or *SMC5* mRNA.

To confirm the pathogenicity of the *SMC5* disease associated variants we utilized our *smc5* morphant zebrafish model to ascertain whether the three patient-associated *SMC5* variants could rescue the developmental abnormalities caused by loss of *smc5* expression. Neither the p.(Arg425Ter), p.(Arg372del) nor p.(His990Asp) variants could complement the reduced head size and increased ceratohyal angle resulting from *smc5* depletion (Supplementary Fig. 9g–i), reinforcing that they confer a loss of function effect. In contrast, both the head size and ceratohyal angle could be restored to normal following expression of WT human *SMC5* or a polymorphic *SMC5* variant, p.(Arg733Gln), identified from gnomAD.

To investigate the two principal underlying causes of microcephaly, slowed cell cycle progression and/or increased apoptosis in the developing brain[2,26–28], fixed whole-mount *slf2* and *smc5* depleted zebrafish embryos were stained with markers of cell cycle stage (G2/M: phospho-histone H3 serine-10) and apoptosis (TUNEL). F0 CRISPR embryos injected with either *slf2* or *smc5* sgRNA with recombinant Cas9 (Fig. 4) exhibited a pronounced increase in both phospho-histone H3 and TUNEL staining in the developing brain when compared to control zebrafish. Importantly, this phenotype was recapitulated in zebrafish embryos transfected with *slf2* or *smc5* MO, which could be complemented by re-expression of the orthologous WT human mRNA (Supplementary Fig. 10). Together, these in vivo data confirm that a functional RAD18-SLF1/2-SMC5/6 pathway is required for normal development of the brain and cartilaginous structures, and compromising this pathway triggers a G2/M cell cycle arrest and the onset of apoptosis leading to microcephaly.

https://doi.org/10.1038/s41467 022 34349 8



SLF2/SMC5 mutant patient-derived cell lines exhibit increased spontaneous replication stress

Although the SMC5/6 complex has been implicated in regulating numerous DNA repair and replication pathways, it is thought that its primary function is to promote efficient replication[14,29]. Therefore, we used DNA fiber analysis to study the impact of SLF2 and SMC5 variants on replication dynamics. All *SLF2* and *SMC5* mutant LCLs examined exhibited a significant increase in spontaneous replication fork stalling and fork asymmetry comparable to that observed in an LCL derived from an ATR-Seckel Syndrome patient (Fig. 5a–d). Importantly, this increased spontaneous replication fork stalling was also observed in patient-derived fibroblasts and could be suppressed by re-expressing

Fig. 2 | Impact of patient associated variants on the stability of SLF2 and SMC5 protein and the integrity of the SMC5/6 complex. a Representative immunoblot analysis of cell extracts from lymphoblastoid (LCL) cell lines derived from patients with variants in *SLF2*. WT AH and WT LQ (WT wild type) indicate unrelated heathy individuals. **b** Representative immunoblot analysis of cell extracts from LCLs derived from patients with variants in *SMC5*. WT SW and WT WCS indicate unrelated heathy individuals. **c**, **d** Whole cell extracts prepared from WT cell lines, SLF2 patient LCLs (**c**) or SMC5 patient LCLs (**d**) were subjected to immunopreci pitation with the indicated antibodies, and inputs and immunoprecipitates (IP) were analysed by immunoblotting (IB). **e** U 2 OS cells expressing Flag SLF2 were transfected with WT or mutant GFP SMC5. GFP SMC5 was precipitated from cell extracts using GFP Trap beads and co precipitated proteins were detected using immunoblotting with the indicated antibodies. *represents a cross reaction of the NSMCE2 antibody to GFP. **f** Whole cell extracts prepared from WT cell lines or SMC5 patient LCLs were subjected to immunoprecipitation with the indicated antibody, and inputs and immunoprecipitates were analysed by immunoblotting. Immunoblotting and immunoprecipitation experiments in (**a**, **b**, **c**, **d**, **f**) are repre sentative of two independent experiments with similar results. Panel e is repre sentative of three independent experiments with similar results.

WT *SLF2* or *SMC5* (Fig. 5e, f; Supplementary Fig. 11a, b). Unlike the ATR-Seckel cell line, all the SLF2-mutant LCLs and one of the SMC5 mutant LCLs exhibited WT levels of replication fork speed (Supplementary Fig. 11c, d). In contrast, LCLs carrying the homozygous p.(His990Asp) exhibited a moderate reduction in replication fork speed.

To confirm these observations, we used CRISPR-Cas9 gene editing in U-2 OS cells to generate SLF2 knockout clones. Despite several attempts we were unable to generate complete SLF2 knockout clones. Rather, we generated two hypomorphic (HM) clones, each with one expressed mutant allele of *SLF2* in conjunction with one or more truncating mutant alleles: SLF2 HM cl.1 (p.Asn411Lysins16, p.Ser403-Ter, p.Asn411LysfsTer3) and SLF2 HM cl.2 (p.Asp398 Ser404del, p.Ser403ThrfsTer14). These clones were subsequently complemented by re-expressing WT *SLF2* (Supplementary Fig. 12). Importantly, DNA fiber analysis of these SLF2 HM clones demonstrated that the vector complemented SLF2 HM cell lines exhibited significantly elevated levels of spontaneous fork stalling compared to the WT SLF2 complemented clones (Fig. 5g).

Since spontaneous replication stress exhibited by cells can be attributed to defective ATR-dependent DNA damage signaling, we used DNA fiber analysis and western blotting to monitor activation of the ATR-dependent stress response[30,31]. In contrast to the ATR-Seckel syndrome cell line, all the SLF2 or SMC5 patient cell lines were capable of activating ATR or the intra-S phase checkpoint in response to HU and MMC (Supplementary Figs. 11e, f, 13) indicating that dysregulation of the ATR stress response pathway does not account for the observed DNA replication defects. This is consistent with previous work demonstrating that loss of the SMC5/6 pathway does not affect activation of the ATR-dependent DDR[17].

We next investigated the cellular impact of the increased spontaneous replication fork instability observed in the patient cell lines using different markers of replication stress. Significantly, both SLF2 and SMC5 patient cell lines exhibited elevated signs of spontaneous replication stress including the presence of DNA double strand breaks (DSBs) in S-phase cells (53BP1 foci in EdU positive cells), an increased frequency of mitotic cells undergoing mitotic DNA synthesis (MiDAS), elevated levels of 53BP1 G1 bodies and the formation of micronuclei (Fig. 6a–d, Supplementary Fig. 14a–d)[17,29]. Crucially, all these phenotypes could be complemented by re-expressing either WT *SLF2* or *SMC5* (Fig. 6). Moreover, the U-2 OS SLF2 HM cell lines also exhibited elevated levels of micronuclei compared to the corrected WT SLF2 expressing clones (Fig. 6e).

Hypomorphic variants in *SLF2* and *SMC5* are associated with mitotic abnormalities, segmented chromosomes, cohesion defects and mosaic variegated hyperploidy

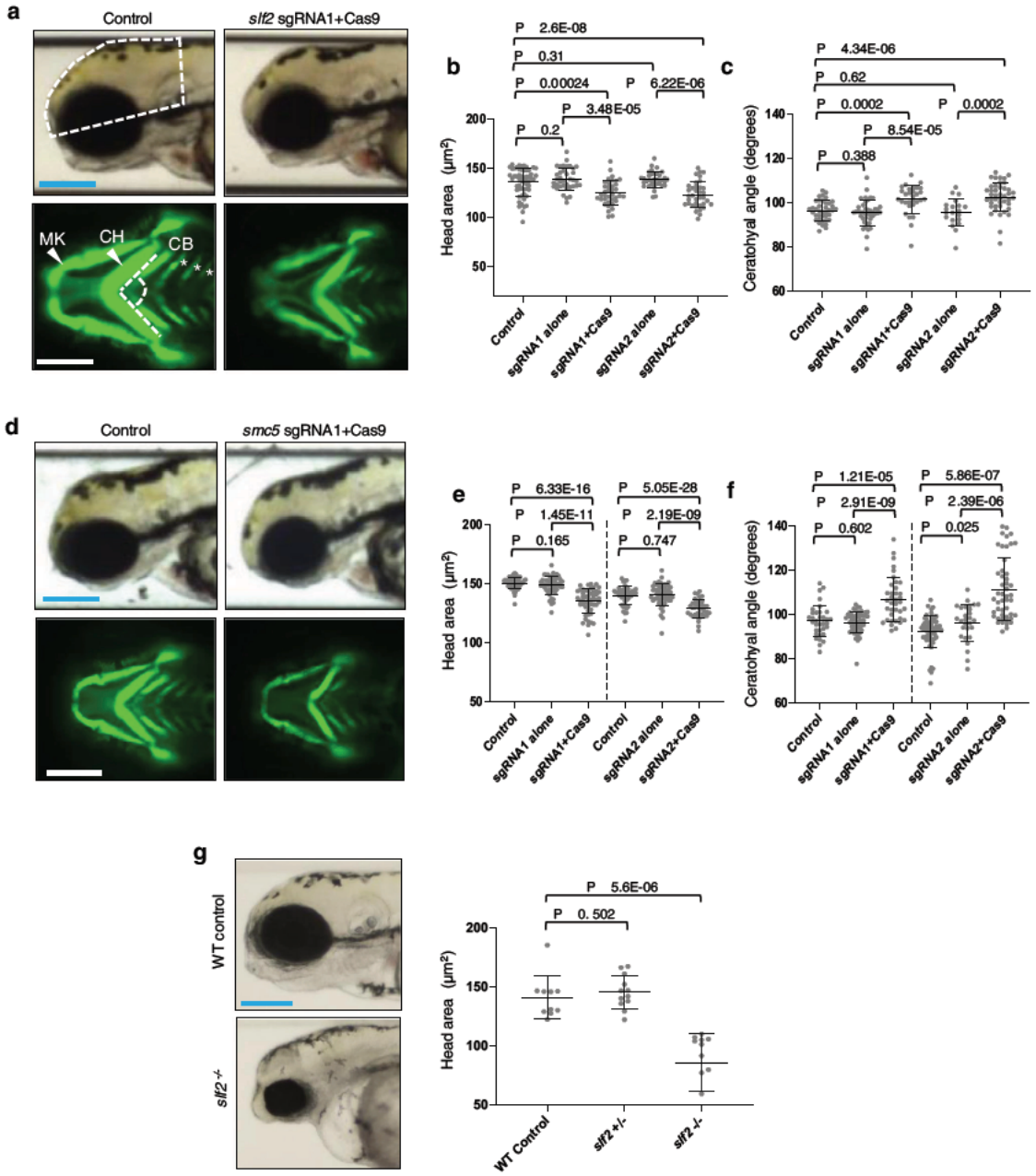
Consistent with the elevated levels of spontaneous replication stress, LCLs derived from SLF2 and SMC5 mutant patients all exhibited increased levels of chromosomal aberrations (such as chromosome and chromatid gaps/breaks and chromosome radials) comparable to that observed in an ATR-Seckel Syndrome LCL (Fig. 6f, g). Notably, this phenotype was not significantly exacerbated by exposure to either APH or MMC, unlike LCLs from an ATR-Seckel Syndrome patient (Supplementary Fig. 15a, b). Importantly, the elevated spontaneous levels of chromosomal aberrations in the SLF2/SMC5 patient fibroblasts and the U-2 OS SLF2 HM cells, were rescued by re-expression of either WT *SLF2* or *SMC5* (Fig. 6h, i).

In addition to the spontaneous chromosomal aberrations, metaphase spread analysis of both the peripheral blood and patient-derived LCLs of SLF2 and SMC5 patients revealed that a significant subset of cells exhibited large increases in chromosome numbers, with some metaphases having >100 chromosomes (Fig. 7a; Supplementary Figs. 16a, b, 17a). Unlike MVA, which typically involves the loss/gain of small numbers of chromosomes, the cytogenetic abnormality observed in SLF2 and SMC5 patient cells predominantly involved huge chromosomal gains. Therefore, we have termed this cytogenetic abnormality mosaic variegated hyperploidy (MVH), i.e., chromosome number >46.

To investigate the cause of the MVH, we explored whether SLF2 or SMC5 patient-derived cell lines exhibited spontaneous mitotic abnormalities. Both SLF2 and SMC5 patient fibroblast cell lines, and U-2 OS SLF2 HM cells, displayed a significant increase in mitotic cells with lagging chromosomes in empty vector complemented cells compared to cells re-expressing WT protein (Fig. 7b–d), consistent with previous reports[17,29,32]. Additionally, when we examined the origins of these lagging chromosomes/micronuclei using CENPA as a marker of centromeres, it was evident that a significant proportion of the micronuclei were positive for CENPA, suggesting that they could have resulted from failed mitotic segregation (Supplementary Fig. 16c, d). This is supportive of the RAD18-SLF1/2-SMC5/6 pathway playing an important role in promoting proper chromosomal segregation.

Since SMC5/6 forms a cohesin-like complex and has been implicated in facilitating centromeric and sister chromatid cohesion[21,32–35], we analysed metaphase spreads from SLF2 and SMC5 patient-derived cells for the presence of cohesion defects. SLF2 and SMC5 peripheral blood lymphocytes showed loss of sister chromatid cohesion as evidenced by the presence of rail-road chromosomes (Fig. 7e; Supplementary Fig. 17b). Moreover, SLF2 and SMC5 patient-derived LCLs exhibited PCS after treatment with the proteasome inhibitor MG132, which is known to induce cohesion fatigue by preventing the metaphase-to-anaphase transition[36] (Fig. 7f). Together, these observations suggest that the MVH characteristic to SLF2 and SMC5 patient cells may also be caused by PCS resulting from cohesion fatigue.

However, given the extent of the karyotypic abnormalities it seemed plausible that other cellular defects may contribute to the large increases in chromosome number seen in SLF2 and SMC5 mutant cell lines in addition to PCS. Replication stress can trigger centrosome amplification via fragmentation of the pericentriolar material (PCM)[37] or premature centriole disengagement, which can lead to mitotic arrest and aneuploidy-induced cell death and microcephaly[38]. To investigate whether centrosome abnormalities could contribute to the cellular pathology associated with SLF2 and SMC5 dysfunction, patient-derived cell lines were subjected to immunofluorescence with antibodies to PCNT1 (a component of the PCM) and mitosin/CENPF (marker of S/G2 cells) before and after incubation with aphidicolin



(APH). Notably, following APH exposure a significant proportion of S/G2 cells possessed more than two centrosomes (Fig. 7g). We also observed that APH treatment had a profound effect on mitosis with >10–50% of SLF2 and SMC5 patient-derived LCLs exhibiting multi-polar spindles during mitosis (Fig. 7h, Supplementary Fig. 16e). This increase in centrosome number and multi-polar spindles is not due to higher levels of replication stress in the APH treated patient cells as quantification of APH-induced G1 53BP1 bodies revealed no difference between empty vector and WT SLF2/SMC5 complemented cells (Fig. 7i). Therefore, it is likely that the MVH observed in SLF2 and SMC5 patient cells arises as a consequence of multiple defects including unresolved replication stress, PCS, chromosome mis-segregation and centrosome amplification.

SLF2/SMC5 mutant cells are unable to replicate efficiently in the presence of stabilized G-quadruplex structures

During our analysis of metaphase spreads of peripheral blood lymphocytes from SLF2 and SMC5 patients, we noted that among the increased levels of spontaneous chromosomal damage, two distinct types of chromosome abnormality were evident (Fig. 8a;

Fig. 3 | Loss of *slf2* and *smc5* in zebrafish give rise to microcephaly and aberrant craniofacial patterning. a Top: Representative lateral bright field images acquired at 3 days post fertilization (dpf); white dashed shape depicts head size measured. Bottom: Representative ventral images of GFP signal from the anterior region of *1.4col1a1:egfp* transgenic reporter larvae at 3 dpf. The white dashed lines show the ceratohyal angle. **b** Quantification of lateral head size measurements. Larvae were injected with two independent sgRNAs targeting *slf2* with or without Cas9; $n = 3$ independent experiments (left to right; 56, 37, 37, 36, 36 larvae/batch). **c** Quantification of the ceratohyal angle. Larvae were injected with two independent *slf2* sgRNAs: $n = 3$ independent experiments (left to right; 39, 42, 30, 20, 44 larvae/batch). **d** Top: Representative lateral bright field images at 3 dpf. Bottom: Repre sentative ventral images of GFP signal in the anterior region of *1.4col1a1:egfp smc5* sgRNA1 transgenic larvae at 3 dpf. **e** Quantification of lateral head size measure ments in 3 dpf larvae (as shown in panel **a**); $n = 3$ independent experiments (left to right; 50, 50, 52, 46, 53, 38 larvae/batch). The chart shows two independent experiments for sgRNA1 and sgRNA2 with a vertical line grouping independent controls with test conditions. **f** Quantification of the ceratohyal angle. Larvae were injected with two independent *smc5* sgRNAs: $n = 3$ independent experiments (left to right; 34, 53, 37, 62, 28, 48 larvae/batch). The chart shows two independent experiments for sgRNA1 and sgRNA2 with a vertical line grouping independent controls with test conditions. **g** Left: Representative lateral bright field images of WT control and *slf2*$^{-/-}$ mutants at 3 dpf. Right: Quantification of lateral head size measurements in 3 dpf WT control and *slf2*$^{-/-}$ mutant larvae (as shown in **a**); $n = 3$ independent experiments (left to right; 10, 12, 12 larvae/batch). In (**a**, **b**): (top left) white dashed shape depicts head size measured; (bottom left) white dashed lines show the ceratohyal angle measured. MK Meckel's cartilage, CH ceratohyal carti lage (indicated with arrowheads, respectively), and CB ceratobranchial arches (asterisks). Scale bars represent 300 µm, with equivalent sizing across panels. Error bars represent standard deviation of the mean. Statistical differences were deter mined with an unpaired Student's *t* test (two sided).

Supplementary Fig. 18). The first type of abnormal chromosome, which we termed segmented chromosomes, contained one or more chromosome gaps/breaks along the body of the chromosome (type 1). Type 1 segmented chromosomes with two or more gaps/breaks were particularly evident in SLF2-P1 and SLF2-P3, whilst most of the segmented chromosomes in SLF2-P2 and SMC5-P7 possessed one gap/break. The second type of abnormal chromosomal structure resembled a dicentric chromosome, which was confirmed by the presence of two centromeres using centromere-specific FISH probes (type 2) (Fig. 8b).

The type 1 segmented chromosomes were reminiscent of the chromosomal abnormalities resulting from combined inactivation of GEN1 and either MUS81 or SLX4, suggesting that they may be caused by an inability to resolve recombination intermediates[39,40]. Accordingly, both SLF2 and SMC5 patient-derived cell lines exhibited elevated levels of recombination as indicated by increased levels of spontaneous RAD51 foci and sister chromatid exchanges (SCEs) in the patient-derived fibroblasts and LCLs respectively (Fig. 8c, Supplementary Figs. 19a, b, 15c, d). This is in line with previous work demonstrating a role for the SMC5/6 complex in resolving recombination intermediates[41–44]. We also observed an increased frequency of telomeric SCEs in SLF2-mutant LCLs (Supplementary Fig. 19c), which could, in part, contribute to the generation of the observed dicentric chromosomes. To investigate whether the spontaneous chromosomal aberrations observed in SLF2/SMC5 mutant cells could arise as a consequence of the presence of unresolved HR intermediates, we examined the effect of stably expressing the bacterial Holliday junction resolvase, RusA, in patient-derived cell lines on genome stability[40]. In line with SLF2 and SMC5 dysfunction causing unresolved HR intermediates to accumulate and this leading to increased genome instability, expression of WT RusA increased the level of spontaneous chromosome aberrations in SLF2/SMC5 mutant cells lines complemented with an empty vector but not with WT SLF2 or SMC5 (Supplementary Fig. 19d, e).

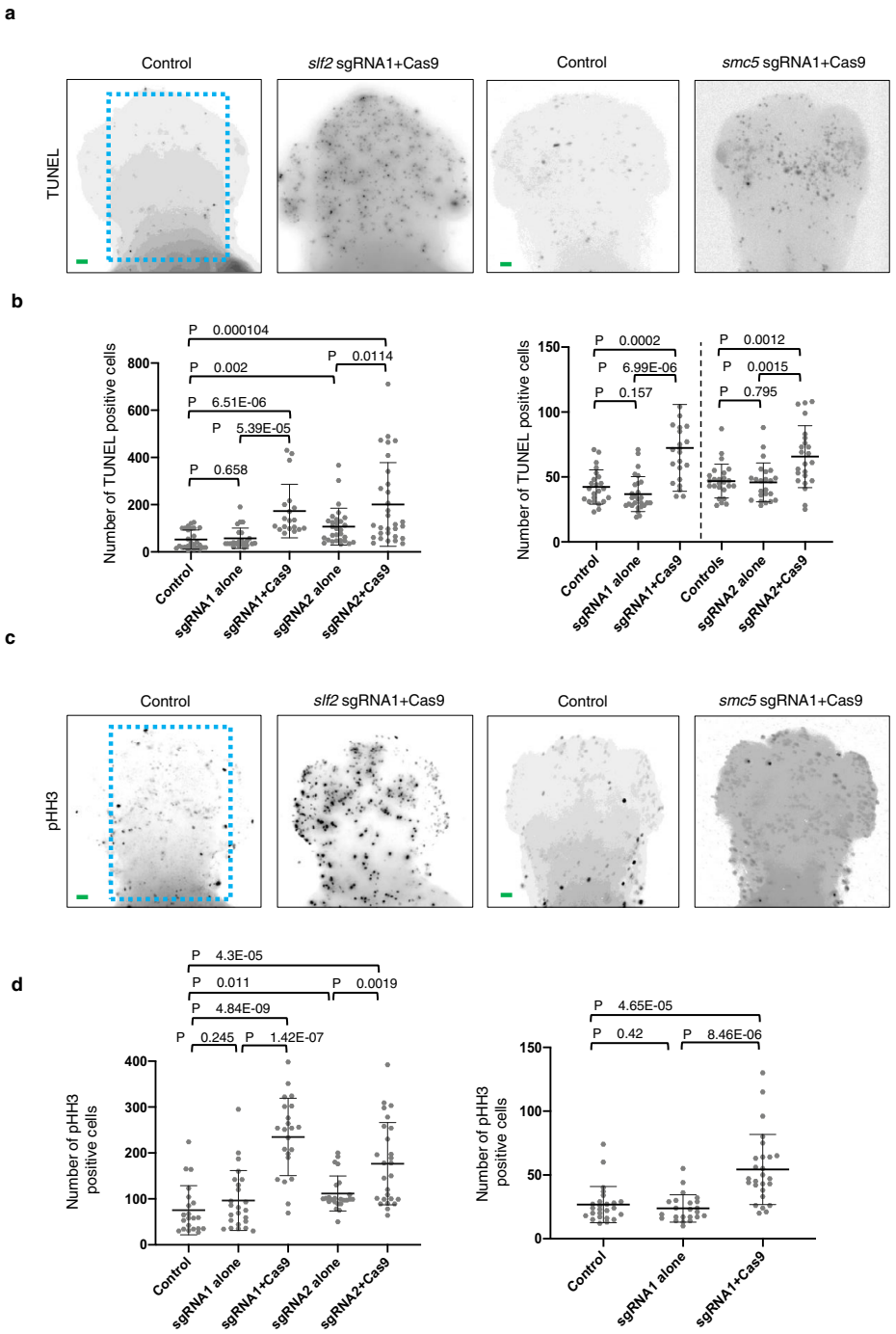
It is known that the SMC5/6 complex is important for the dissolution of replication stress-induced recombination, especially at repetitive regions prone to forming secondary structures and natural replication pause site intermediates[41,43–46]. This is consistent with our observations that the replication stress phenotype observed in SLF2/SMC5 mutant cells was not markedly exacerbated by exposure to MMC, APH and HU (Fig. 5; Supplementary Figs. 11, 13). Recently, it has been shown that RNF168, which promotes the recruitment of the RAD18-SLF1/2-SMC5/6 pathway to damaged replication forks, is important for signaling the presence of G-quadruplex (G4) DNA structures stabilized by the RNA polymerase I inhibitor, CX5461[47]. Since cells deficient in BRCA1/2 and the cohesin-associated helicase DDX11 are also hypersensitive to this agent[48,49] and DDX11 was shown to function with SMC5/6 to repair DNA damage[17,50,51], we hypothesized that the RAD18-SLF1/2-SMC5/6 pathway might play a role in suppressing replication stress at sites of stabilized G4 structures. To test this possibility, we first investigated the effects of CX5461 on DNA replication using DNA fiber analysis. This revealed that whilst WT SLF2 and SMC5 expressing patient fibroblasts could replicate normally in the presence of CX5461, SLF2 and SMC5 patient fibroblasts complemented with an empty vector exhibited a significant reduction in replication fork speed when incubated with this G4-stabilizing compound (Fig. 8d). Additionally, SLF2 and SMC5 patient-derived fibroblasts, LCLs and U-2 OS SLF2 HM cells treated with CX5461 exhibited increased levels of G1-phase 53BP1 bodies and chromosome aberrations (Fig. 8e, Supplementary Fig. 20a, c). In keeping with this, LCLs from SLF2-P1 and SMC5-P8 displayed an increased sensitivity to CX5461 (Fig. 8f). Strikingly, we also observed that CX5461 treatment induced a significant increase in the levels of type 1 segmented chromosomes in the SLF2 and SMC5 patient LCLs, but not in the WT LCLs (Supplementary Fig. 20b). These data suggest a role for SLF2 and the SMC5/6 complex in resolving replication stress at sites of stabilized G4 structures.

Whilst CX5461 is known to inhibit RNA polymerase I and stabilize G-quadruplexes, more recently it has also been identified as a TOP2 poison[52,53]. Given the pleiotropic nature of CX5461, we sought to identify which genotoxic lesion induced by CX5461 was causing the increased replication stress in cells deficient in components of the SMC5/6 complex. In this respect, we carried out DNA fiber and chromosomal aberration analysis on patient-derived cell lines following exposure to pyridostatin (a G-quadruplex stabilizer), etoposide (a TOP2 poison) and BMH21 (an RNA polymerase I inhibitor). Interestingly, only exposure to pyridostatin caused a significant reduction in replication progression and an increase in the levels of chromosome aberrations in SLF2 and SMC5 mutant cell lines (Fig. 8g, Supplementary Fig. 20d).

Taken together, these observations support the notion that the spontaneous replication stress and chromosomal instability displayed by cells from patients with SLF2/SMC5 mutations is caused, in part, by an inability to resolve a specific subset of replication-associated recombination intermediates arising at sites of G4 structures.

Discussion

Disrupting the delicate balance between stem cell proliferation and differentiation profoundly affects embryonic development, particularly body growth and brain development. Rapidly proliferating pluripotent stem cells exhibit constitutively high levels of replication stress and as such are heavily reliant on replication-associated DNA damage response pathways to maintain genome stability. Unsurprisingly, patients with pathogenic variants in genes encoding components of the replisome, the DNA damage response (DDR) and factors that maintain sister chromatid cohesion exhibit developmental abnormalities including severe microcephaly and dwarfism. Furthermore, variants in centrosome components and regulators of the microtubule-spindle network can also result in these developmental abnormalities



by affecting the orientation of the spindle pole and/or triggering excessive cell death through the generation of aneuploid cells[1]. However, it is often difficult to determine whether the cellular pathology underlying the development of these neurodevelopmental disorders results primarily from the presence of aberrant replication or defective mitosis[38,54,55].

Here we report the clinical and genetic characterization of 11 patients with biallelic variants in two components of the newly described RAD18-SLF1/2-SMC5/6 DDR pathway, *SLF2* and *SMC5*, exhibiting microcephaly, short stature, cardiac defects and anemia. However, in contrast to FA and other known disorders, cells from these patients exhibit a unique chromosomal instability phenotype,

Fig. 4 | Loss of *slf2* and *smc5* induces apoptosis and altered cell cycle progres sion in zebrafish larvae. a Representative dorsal inverted fluorescent images showing TUNEL positive cells in control and *slf2* F0 mutants at 2 dpf (left two panels), and control and *smc5* F0 mutants at 3 dpf (right two panels). The blue dashed line indicates the region of interest (ROI) quantified. Embryos of the same developmental stage and similar magnification were evaluated for all *slf2* and *smc5* conditions. **b** Left: Quantification of TUNEL positive cells in the ROI of control and *slf2* F0 mutants at 2 dpf shown in panel **a** (left to right; 27, 23, 19, 29, 30 embryos/condition were analysed from 3 independent experiments). Right: Quantification of TUNEL positive cells in control and *smc5* F0 mutants at 3 dpf in the ROI as shown in panel **a** (left to right; 37, 27, 22, 25, 23, 23 embryos/condition were analysed from 3 independent experiments). The chart shows two independent experiments for sgRNA1 and sgRNA2 with a vertical line grouping independent controls with test conditions. **c** Representative dorsal inverted fluorescent images showing phospho histone H3 (pHH3) positive cells in control and *slf2* F0 mutants at 2 dpf (left two panels), and control and *slf2* F0 mutants at 3 dpf (right two panels). Embryos of the same developmental stage and similar magnification were evaluated for all *slf2* and *smc5* conditions. **d** Left: Quantification of pHH3 positive cells of control and *slf2* F0 mutants at 2 dpf in the ROI as shown in panel **a** (left to right; 21, 24, 22, 24, 26 embryos/condition were analysed from 3 independent experiments). Right: Quantification of pHH3 positive cells in the ROI in control and *smc5* F0 mutants at 3 dpf as shown in panel **a** (left to right; 25, 23, 26 embryos/condition were analysed from 3 independent experiments). For all panels: Statistical differences were determined with an unpaired Student's *t* test (two sided). Error bars represent standard deviation of the mean. Scale bars, 30 μm with equivalent sizing across panels.

hallmarked by segmented and dicentric chromosomes and mosaic variegated hyperploidy, arising from a combination of replication stress- and mitosis-associated cellular pathologies. Given that the segmented chromosomes seen in SLF2 and SMC5 patient cells represent a chromosome instability phenotype not previously associated with any known DNA repair or replication deficiency disorder, we have named this syndrome, Atelís Syndrome (ATS), after the Greek word for incomplete to signify the importance of these atelic or segmented chromosomes as a diagnostic marker of the disease.

The SMC5/6 complex has been shown to have many functions in the cell, including regulating homologous recombination (HR)-dependent DNA repair, stabilizing and restarting stalled replication forks, maintaining replication through highly repetitive regions of the genome, maintaining rDNA stability, elongating telomeres by ALT and controlling the topology of unusual DNA structures[12,14,56,57]. In contrast, little is known about the functions of SLF1 and SLF2, which were identified during a large proteomic screen of proteins associated with damaged replication forks[11]. However, it has been suggested that SLF1 and SLF2 are functional orthologs of the yeast Nse5 and Nse6 proteins, respectively, which are important for localizing the SMC5/6 complex to DNA damage and regulating its ATPase activity[11,58–60].

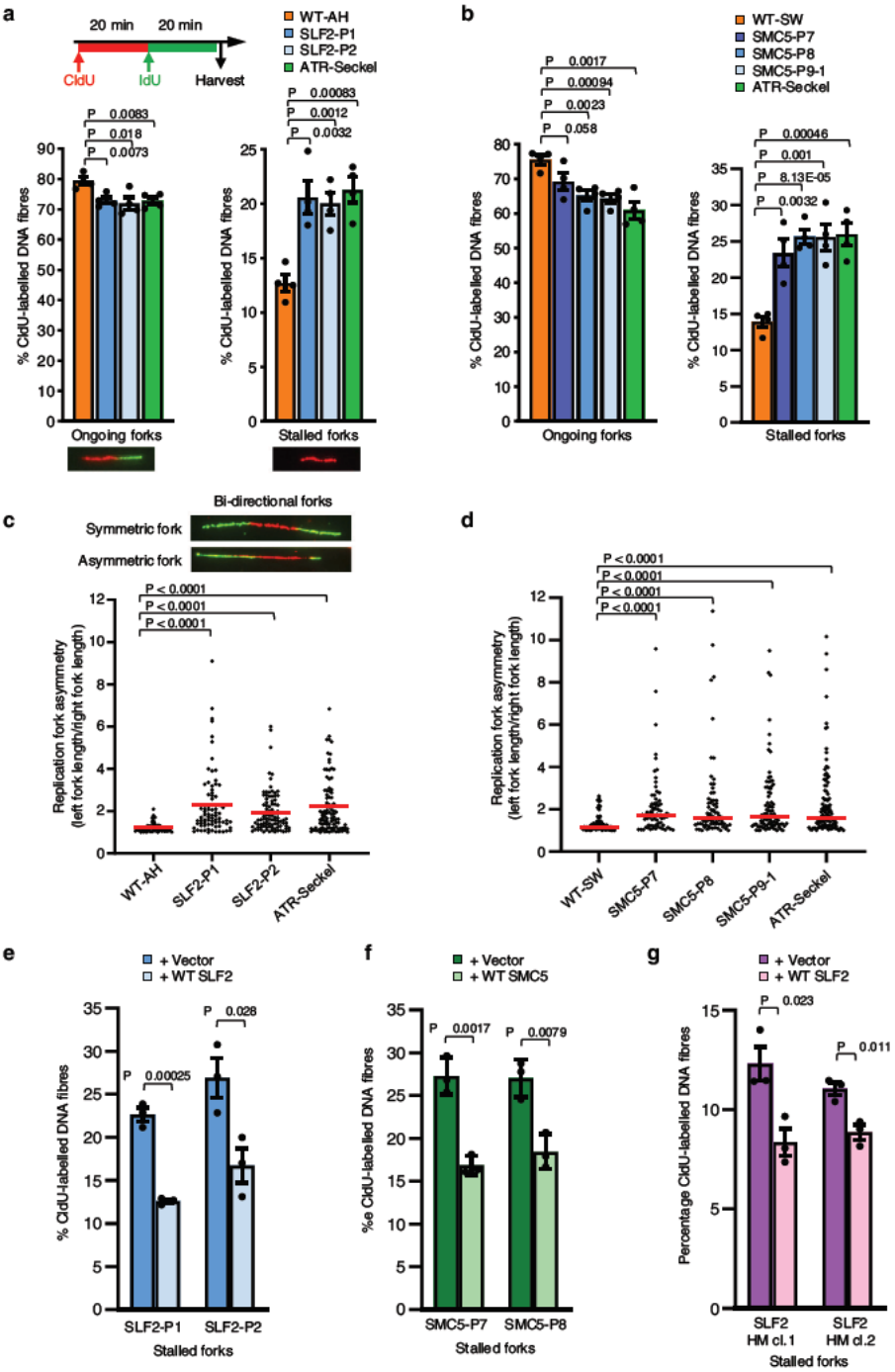
Pursuant to the role of the SLF1/2-SMC5/6 complex in maintaining replication fork stability, we demonstrate that cells from ATS patients exhibit elevated levels of spontaneous replication stress, although this was not exacerbated significantly following exposure to replication stress-inducing agents (HU, MMC or APH). This suggests that the clinical phenotype resulting from variants in *SLF2* and *SMC5* may not simply arise from elevated levels of replication stress, but rather from deficits with a subset of replication forks, such as those replicating through difficult-to-replicate regions of the genome or encountering specific types of endogenous DNA lesions. Consistent with this hypothesis, ATS cells fail to replicate efficiently in the presence of stabilized G4 structures and accumulate chromosomal damage, suggesting that the RAD18-SLF1/2-SMC5/6 pathway functions to resolve replication intermediates occurring at these lesions. Since G4 structures have been shown to be enriched at telomere repeat sequences[61], a defect in the ability to replicate through these lesions could result in genome instability at telomeres, potentially explaining the presence of dicentric chromosomes in ATS patient cells.

ATS patients exhibit overlapping clinical and cellular features with WABS patients, including microcephaly, growth restriction, skin hyperpigmentation, ocular abnormalities and heart defects. Moreover, cell lines derived from both ATS and WABS patients exhibit loss of sister chromatid cohesion and premature chromatid separation[49]. Interestingly, the loss of sister chromatid cohesion in WABS cell lines is exacerbated upon exposure to replication stress-inducing genotoxins, including G4 stabilizing agents[49]. Notably, cells from *Ddx11* null mice display loss of sister chromatid cohesion, chromosome segregation errors and aneuploidy, which has been shown to induce a G2/M cell cycle delay and apoptosis[62]. This suggests that a failure to resolve specific endogenous DNA lesions, such as G4 structures, in ATS cells may directly compromise cohesion, or exacerbate a pre-existing cohesion defect, thus giving rise to chromosome segregation defects and aneuploidy that triggers cell death in highly proliferative tissues, such as the developing brain.

It is clear that the RAD18-SLF1/2-SMC5/6 pathway plays additional cellular roles beyond promoting replication through G4 lesions. In yeast, the smc5/6 complex restrains recombination at programmed fork pause sites, for example, in the rDNA locus[43,44,63] and, in mammalian cells, SMC5/6 is involved in suppressing HR at highly repetitive sequences, e.g., rDNA, centromeres and telomeres[14,63]. Consistent with this, ATS cells exhibit elevated levels of RAD51 foci in S-phase cells and spontaneous SCEs and tSCEs. Interestingly, segmented chromosomes have been observed in cells that have a combined defect in both the Holliday junction dissolution and resolution pathways[64], indicating that the gaps in the type 1 segmented chromosomes may result from a failure to dissolve/resolve recombination intermediates[41].

Cells from *NSMCE2* and *NSMCE3* mutant patients are not known to display segmented or dicentric chromosomes, and whilst *NSMCE3* patient-derived cells exhibit aneuploidy and structural chromosome abnormalities, hyperploidy to the extent seen in ATS cells was not reported[18,20]. This indicates that neither NSMCE2 nor NSMCE3 subunits are essential for this SMC5/6 function, or that the hypomorphic variants in these genes retain sufficient function to suppress these chromosomal phenotypes. Consistent with the latter scenario, *Nsmce2* transgenic mice lacking SUMO E3 ligase activity developed normally, whereas a complete loss of *Nsmce2* resulted in early embryonic lethality associated with chromosome segregation defects[65]. Notably MEFs derived from the *Nsmce2* knockout mice exhibited increased spontaneous replication stress and genome instability due to a failure to detangle recombination intermediates similar to ATS patient cell lines (e.g., elevated levels of BRCA1 foci, increased sister chromatid and telomeric SCEs and chromosomal segregation errors)[65] indicating that ATS represents a more severe form of SMC5/6 dysfunction.

Interestingly, the clinical phenotype exhibited by patients with variants in the SMC5/6 complex components *NSMCE2* and *NSMCE3* are different from each other, with the former being associated with microcephalic primordial dwarfism and insulin resistance[20] and the latter being associated with severe pulmonary disease and immunodeficiency[18,19]. It is unclear why these clinical presentations are different, especially as the cellular phenotype resulting from *NSMCE2* and *NSMCE3* variants are similar[18,20]. One possible important cellular difference between the two disorders is that the patient-associated missense variants in *NSMCE3* result in the destabilization of the SMC5/6 complex to a much greater extent than the nonsense variants present in *NSMCE2* patients[18,20]. It is notable that the clinical phenotype of ATS patients more closely resembles that of *NSMCE2* patients than *NSMCE3* patients, and like *NSMCE2* patient variants, *SLF2* and *SMC5* patient variants do not destabilize the SMC5/6 complex to any significant degree.



Taken together, we have demonstrated that variants in two components of the RAD18-SLF1/2-SMC5/6 pathway give rise to a FA/MVA-like disorder, termed Atelis Syndrome, with clinical and cellular features overlapping with WABS, MVA, NSMCE2 variants and FA. In vivo ablation of *slf2* and *smc5* in zebrafish recapitulate patient phenotypes including microcephaly and craniofacial patterning defects, likely due to concomitant cell cycle defects and apoptosis. We show that cells from ATS patients display a unique and complex chromosomal instability phenotype consisting of atelic (segmented) and dicentric chromosomes coupled with MVH, which should allow for cytogenetic diagnosis of patients with this disorder.

Fig. 5 | Patient derived cell lines from individuals with biallelic *SLF2* or *SMC5* variants exhibit increased levels of spontaneous replication fork instability. **a** Top: Schematic representation for DNA fiber analysis in untreated cells. The indicated cell lines were pulse labeled with CldU for 20 min, then pulse labeled with IdU for 20 min. Bottom: DNA fiber analysis of SLF2 patient derived LCLs or LCLs from a WT individual. The percentage of ongoing forks (left) or stalled forks (right) was quantified. $n = 4$ independent experiments. A minimum of 1500 fork structures were counted. **b** DNA fiber analysis of SMC5 patient derived LCLs or WT LCLs. Quantification of the levels of ongoing forks (left) or stalled forks (right). $n = 4$ independent experiments. A minimum of 750 fork structures were counted. **c**, **d** Quantification of replication fork asymmetry of WT, SLF2 patient (**c**) or SMC5 patient LCLs (**d**). $n = 4$ independent experiments. A minimum of 75 fork structures were counted. Red lines denote median values. A Mann Whitney rank sum test was performed for statistical analysis. Replication fork asymmetry represents the ratio of the left to right fork track lengths of bidirectional replication forks. **e**, **f** DNA fiber analysis of SLF2 (**e**) and SMC5 (**f**) mutant fibroblast cell lines infected with lenti viruses encoding WT SLF2, WT SMC5, or an empty vector. The percentage of ongoing forks (left) or stalled forks (right) in untreated cells was quantified. A minimum of 350 fork structures in total were counted over 3 independent experiments. **g** DNA fiber analysis of U 2 OS SLF2 CRISPR hypomorphic (HM) cells infected with lentiviruses encoding WT SLF2 or an empty vector. The percentage of stalled forks in untreated cells was quantified. A minimum of 1000 fork structures in total were counted over 3 independent experiments. For (**a**, **b**, **e**, **f**, **g**); a Student's *t* test (two sided, equal variance) was performed for statistical analysis and error bars denote SEM.

Methods

Research subjects

Informed consent was obtained from all participating families to take clinical samples and to publish clinical information in accordance with local approval regulations and in compliance with the Declaration of Helsinki principles. This study was approved by the West Midlands, Coventry and Warwickshire Research Ethics Committee (REC: 20/WM/0098), the Scottish Multicentre Research Ethics Committee (REC: 05/MRE00/74), the Lancaster General Hospital Institutional Review Board and the Institutional Review Boards of Yokohama City University Graduate School of Medicine (ID: A190800001) and Jichi Medical University (ID: G21-V06). A collaboration to study the pathological significance of the identified *SLF2* and *SMC5* variants was established via GeneMatcher[66].

Exome sequencing

Genomic DNA from affected children and family members was extracted from peripheral blood using standard methods. Whole exome capture and sequencing was performed as described to a minimum of *30x* coverage[67]. Exome sequencing for families 8 and 9 was conducted in collaboration with the Regeneron Genetics Center as previously described[68]. Briefly, DNA was sheared (Covaris S2), exome capture performed using the Agilent SureSelect v5 enrichment kit according to manufacturer's instructions, and libraries were sequenced with 125 bp read-pairs using the Illumina HiSeq 2500 V4 platform. All analyses were performed as described[69]. Variants were confirmed by bidirectional capillary dye-terminator sequencing and annotated using the reference sequences, GenBank: NM 018121.4, NM 001136123.2 and NM 015110.4. Capillary sequencing was performed in the MRC Human Genetics Unit, Edinburgh, UK, the University of Birmingham, UK, the Bioscientia Institute for Medical Diagnostics, Germany, the Rare Disease Genomics Department, Yokohama City University Hospital, Japan and the Regeneron Genetics Center, Regeneron Pharmaceuticals Inc., USA.

Cell lines

Patient-derived lymphoblastoid cell lines (LCLs) were generated from peripheral blood samples with Epstein Barr virus (EBV) transformation using standard methods and were maintained in RPMI-1640 medium (Life Technologies) supplemented with 10% FBS, L-glutamine and penicillin-streptomycin. The ATR-Seckel LCL used in this study was reported previously[31]. Dermal primary fibroblasts were grown from skin-punch biopsies and maintained in Dulbecco's modified Eagle's medium (DMEM; Thermo Fisher Scientific) supplemented with 20% FCS, 5% L-glutamine and 5% penicillin-streptomycin. Primary fibroblasts were immortalized with a lentivirus expressing human telomerase reverse transcriptase (hTERT) that was generated by transfecting 293FT cells (Thermo Fisher Scientific) with the plasmids: pLV-hTERT-IRES-hygro (Addgene #85140), psPax2 (Addgene #12260) and pMD2.G (Addgene #12259). Selection was performed using Hygromycin (Thermo Fisher Scientific) at 70 µg/ml. All LCLs were routinely grown in RPMI-1640 (Thermo Fisher Scientific) supplemented with 10% FCS, 5% L-glutamine and 5% penicillin-streptomycin. Patient cell lines were validated using Sanger sequencing and immunoblotting. Fibroblast and U-2 OS cell complementation was carried out using the pLVX-IRES-Neo lentiviral vector (Takara Bio) encoding *2x*Myc-tagged *SLF2* or untagged *SMC5*.

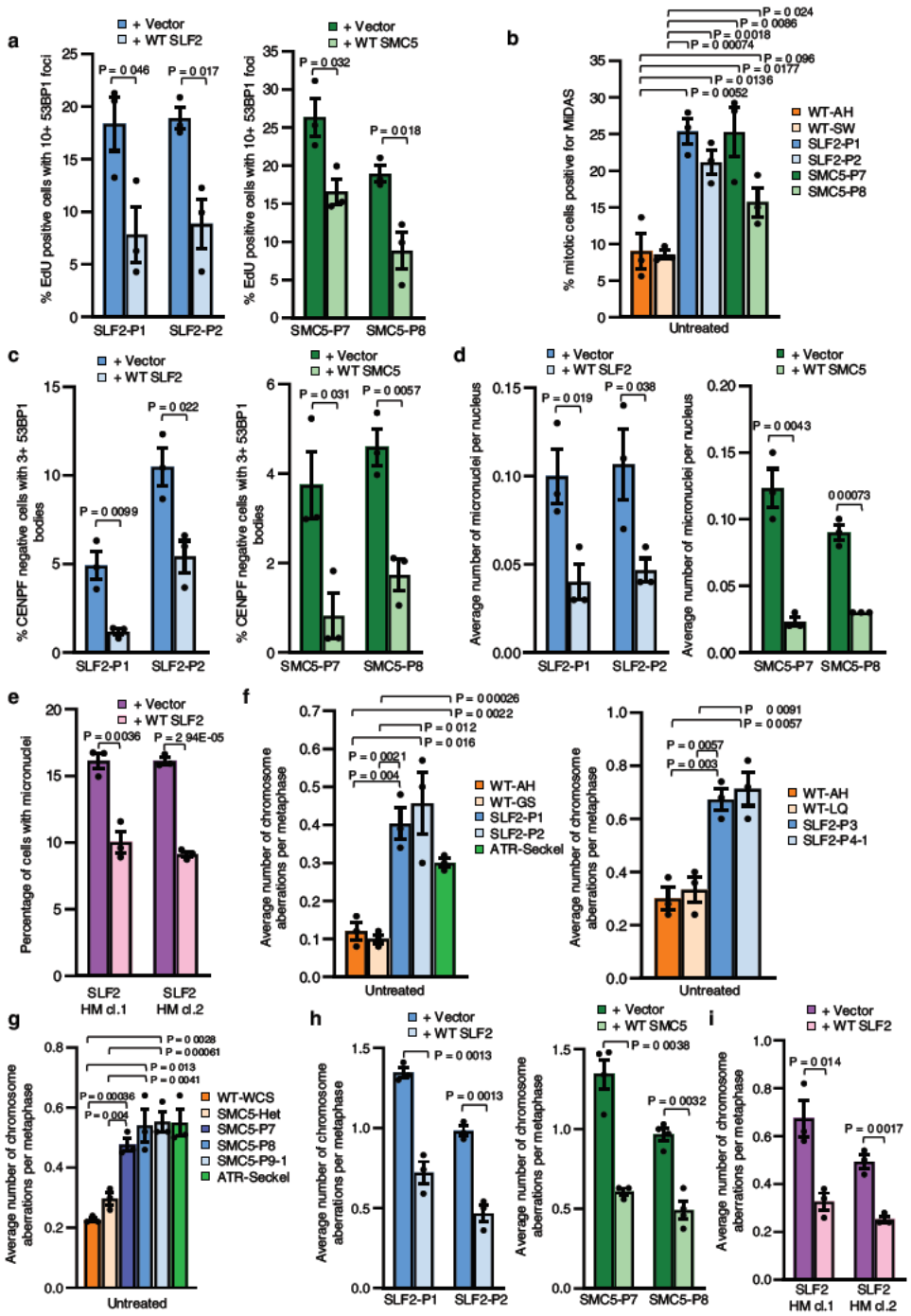
293FT cells (Thermo Fisher Scientific) were maintained in DMEM supplemented with 10% FBS, 5% l-glutamine and 5% penicillin-streptomycin and U-2-OS cells were cultured in McCoy's 5A medium, supplemented with 10% FBS, and 5% penicillin/streptomycin. 293FT cells were transiently transfected with GFP-BLM or GFP expression vectors using Lipofectamine 2000 (Thermo Fisher Scientific). U-2 OS cells were transiently transfected with SLF2/SMC5 expression vectors using FuGENE 6 Transfection Reagent (E2692, Promega) or Lipofectamine 3000 Reagent (L3000015, Thermo Fisher Scientific) where indicated. Stable GFP-SMC5 cell lines were generated by G418 selection and low expressing clones were selected based on GFP expression. All cell lines were routinely tested for mycoplasma.

Western blotting

Whole-cell extracts were obtained by sonication in UTB buffer (8 M urea, 50 mM Tris, 150 mM β-mercaptoethanol) and analyzed by SDS–PAGE following standard procedures. Protein samples were run on 6–12% acrylamide gels with SDS–PAGE and transferred onto a nitrocellulose membrane. Immunoblotting was performed using antibodies to: RAD18 (Fortis Life Sciences, A301-340A; 1:1000), SMC5 (Fortis Life Sciences, A300-236A; 1:500), SMC6 (Fortis Life Sciences, A300-237A; 1:2000), SLF2 (generated in house; 1:1000)[11], GAPDH (Genetex, GTX100118; 1:1000), Myc (Abcam, ab32; 1:1000), GFP (SCBT, sc-9996; 1:1000), HA (SCBT, sc-7392; 1:1000), α-Tubulin (Sigma-Aldrich, T9026; 1:20,000), ATR (Fortis Life Sciences, A300-137A; 1:1,000), phospho-ATR (Thr1989) (GeneTex, GTX128145; 1:500), FANCD2 (SCBT, sc-20022; 1:1,000), CHK1 (SCBT, sc-8408; 1:1,000), phospho-CHK1 (Ser345) (Cell Signaling Technology, 2341; 1:100), NBS1 (Genetex, GTX70224; 1:10,000); phospho-NBS1 (Ser343) (Abcam, 47272; 1:500); SMC1 (Fortis Life Sciences, A300-055A; 1:1,000); phospho-SMC1 (Ser966) (Fortis Life Sciences, A300-050A; 1:1,000); HA (Abcam, Ab9110; 1:1000). Loading controls for all blots were derived from re-probing the same membrane, except for the phospho-antibody immunoblots, for which paired gels were run simultaneously and blotted in parallel for phosphorylated and total proteins.

Co-immunoprecipitation and GFP-Trap pull-downs

For GFP-Trap pulldown experiments with 293FT cells, cells transfected with plasmids using Lipofectamine 2000, were treated with 2 mM HU for 16 h and harvested. Cells were incubated in lysis buffer (150 mM NaCl, 50 mM Tris-HCl pH 7.5, 2 mM $MgCl_2$, 1% NP40, 90 U/ml Benzonase (Novagen) and EDTA-free protease inhibitor cocktail [Roche]) for 30 min with rotation at 4 °C. Cell lysates were then pre-cleared at 65,000 × *g* at 4 °C for 30 min. For GFP-Trap, 3–5 mg of lysate was



incubated with GFP-Trap agarose beads (ChromoTek) for 5 h at 4 °C. The resulting GFP-Trap complexes were washed with wash buffer (150 mM NaCl, 50 mM Tris-HCl pH 7.5, 0.5% NP40, and complete protease inhibitor cocktail [Roche]) and analysed by SDS–PAGE.

For immunoprecipitations from patient-derived LCLs, 3 mg of lysate (prepared with the same lysis buffer as above) was immunoprecipitated with 5 µg of antibody (RAD18; Fortis Life Sciences, A301-340A or NSMCE2; Fortis Life Sciences, A304-129A) and protein A-sepharose beads (GE Healthcare). Complexes were washed with wash buffer (as described above) and analysed by SDS–PAGE. Experiments were carried out in the presence of Benzonase nuclease to exclude the possibility of interactions being mediated by DNA.

Fig. 6 | SLF2 and SMC5 patient cells exhibit S phase associated DNA damage. **a** Percentage of cells positive for EdU staining with >10 53BP1 foci in SLF2 and SMC5 mutant fibroblast cell lines infected with lentiviruses encoding WT SLF2, WT SMC5, or an empty vector. A minimum of 900 EdU positive cells across 3 independent experiments were counted. **b** SLF2 and SMC5 patient fibroblast cell lines were pulsed with 10 μM EdU for 45 min, fixed, and mitotic DNA synthesis was visualized by mitotic EdU incorporation following labeling with click chemistry. The percen tage of mitotic cells with EdU foci was quantified. A minimum of 300 mitotic cells were counted. $n = 3$ independent experiments. **c** Immunofluorescent microscopy analysis to quantify the percentage of G1 phase cells (CENPF negative cells) with >3 53BP1 bodies in WT SLF2, WT SMC5, or an empty vector expressing SLF2 and SMC5 patient fibroblasts. $n = 3$ independent experiments. A minimum of 750 G1 phase cells were counted. **d** Levels of micronuclei in cells from (**c**). $n = 3$ independent experiments. A minimum of 2500 cells were counted. **e** Levels of micronuclei in U 2 OS SLF2 CRISPR HM cells infected with lentiviruses encoding WT SLF2 or an empty vector. $n = 3$ independent experiments. A minimum of 1700 cells were counted. **f**, **g** Quantification of the average number of chromosomal aberrations per meta phase (which includes chromatid/chromosome gaps, breaks, fragments and chromosomes radials) in WT, SLF2 patient (**f**), or SMC5 patient LCLs (**g**). $n = 3$ independent experiments. A minimum of 140 metaphases were counted. **h** Average number of chromosomal aberrations per metaphase (chromatid/chromosome gaps, breaks, fragments and chromosome radials) in SLF2 and SMC5 mutant fibroblast cell lines infected with lentiviruses encoding WT SLF2, WT SMC5, or an empty vector was quantified. $n = 3$ independent experiments. A minimum of 90 metaphases were counted. **i** Average number of chromosomal aberrations (chro matid/chromosome gaps, breaks, fragments and chromosome radials) per meta phase in U 2 OS SLF2 CRISPR HM cell lines expressing either WT SLF2 or an empty vector. $n = 3$ independent experiments. A minimum of 100 metaphases were counted. In all cases, a Student's t test (two sided, equal variance) was performed for statistical analysis and error bars denote SEM.

For immunoprecipitations from U-2 OS cells, cell lysates were generated using EBC buffer (150-mM NaCl; 50-mM Tris, pH 7.5; 1-mM EDTA; 0.5% IGEPAL CA-630). Lysates were subject to Co-IP using Strep-Tactin Sepharose (IBA GmbH) prior to immunoblot using the following antibodies: GFP (sc-9996, SCBT; 1:1000), HA (sc-7392, SCBT; 1:1000), RAD18 (A301-340A, Fortis Life Sciences; 1:1000), SMC6 (A300-237A, Fortis Life Sciences; 1:2000), SMC5 (Fortis Life Sciences, A300-236A; 1:500), NSMCE2 (Fortis Life Sciences, A304-129A; 1:500), α-Tubulin (T9026, Sigma-Aldrich; 1:20000).

Laser micro-irradiation

U-2 OS cells were grown on coverslips and sensitized to laser induced DSB formation using 5-Bromo-2-deoxyuridine (B9285-50MG, Sigma-Aldrich) for 24 h. GFP-SLF2 expression vectors were transiently transfected 24 h prior and GFP-SMC5 stable expressing cells were used for micro-irradiation. Laser micro-irradiation induced DSB formation was performed as previously described[70] with 1 h allowed for recovery. Cells were pre-extracted using CSK buffer (100 mM NaCl, 10 mM HEPES, 3 mM $MgCl_2$, 300 mM Sucrose, 0.25% Triton-X-100, 1 mM PMSF) prior to fixation in formalin buffer (AMPQ43182, VWR) for 15 mins at room temperature (RT).

Fixed coverslips were blocked with 5% Bovine Serum Albumin (A7906, Sigma-Aldrich) for 1 h prior to staining with anti-γ-H2AX (Ser139) (1:1000, 05-636, Merck) and anti-GFP (1:500, PABG1, Chromotek) overnight at 4 °C. After PBS washes cells were stained with Alexa Fluor secondary antibodies and 4',6-Diamidino-2-Phenylindole (DAPI, D1306, Molecular Probes) for 30 min at RT. After further washing, coverslips were dried completely and mounted for imaging using Mowiol (81381, Sigma-Aldrich).

Zebrafish husbandry and embryo maintenance

All zebrafish experiments were performed according to protocols approved by the Duke University and Northwestern University institutional animal care and use committees (IACUC). Wild type (WT: ZDR or NIH) adults or transgenic *−1.4col1a1:egfp*[25] adults were maintained on an AB background and subjected to natural matings to generate embryos for microinjection and/or phenotyping. Embryos were grown in egg water (0.3 g/L NaCl, 75 mg/L $CaSO_4$, 37.5 mg/L $NaHCO_3$, 0.003% methylene blue) at 28 °C until assessment. Zebrafish sex is unknown until animals are ~3 months old. Therefore, in the larvae at <5 days post fertilization, it is not possible to know how many males and females are present, and there should be no sex-dependent effects at this stage. However, adults that were used to generate embryos were crossed in a 1 male to 1 female ratio.

CRISPR-Cas9 genome editing of zebrafish embryos

Reciprocal translated BLAST of human *SLF2* (NP 060591.3) and *SMC5* (NP 055925.2) was performed against the zebrafish genome and found a single ortholog corresponding to either protein (transcripts targeted: *slf2*: ENSDART00000136689.3, *smc5*: ENSDART00000122170.4). To identify CRISPR/Cas9 single guide RNA (sgRNA) targets in both genes, CHOPCHOPv2[71] (and http://chopchop.cbu.uib.no) was used. sgRNAs were generated using the GeneArt precision gRNA synthesis kit (Thermo Fisher Scientific) according to the manufacturer's instructions (Supplementary Table 1). 1 nl of cocktail containing 100 pg sgRNA with or without 200 pg of Cas9 protein (PNA Bio) was injected into the cell of single-cell staged zebrafish embryos. To estimate the percentage mosaicism of genome-edited cells, genomic DNA from individual embryos was extracted at 2 days post fertilization (dpf; two controls and ten founder [F0] embryos per sgRNA). PCR was used to amplify the sgRNA targeted region using flanking primers and heteroduplex analysis was performed using polyacrylamide gel electrophoresis (PAGE). PCR products were denatured, reannealed slowly, and migrated on a 20% polyacrylamide gel (Thermo Fisher Scientific). PCR products from five embryos per sgRNA were randomly selected from the heteroduplex analysis, cloned into a TOPO-TA vector (Thermo Fisher Scientific) and sequenced using BigDye terminator 3.1 chemistry (Applied Biosystems). To isolate stable *slf2* mutants, F0 animals were crossed to WT ZDR adults and heterozygous F1 mutants bearing the c.515 522del (p.Ser172 Ser174fs191Ter) variant were identified. Mutant F1 adult siblings were inter-crossed to generate homozygous F2 animals for phenotyping. *slf2* mRNA expression level was monitored by qRT-PCR (QuantStudio, Thermo Fisher Scientific) using SYBR Green detection kit (Thermo Fisher Scientific) with normalization to *β-actin*.

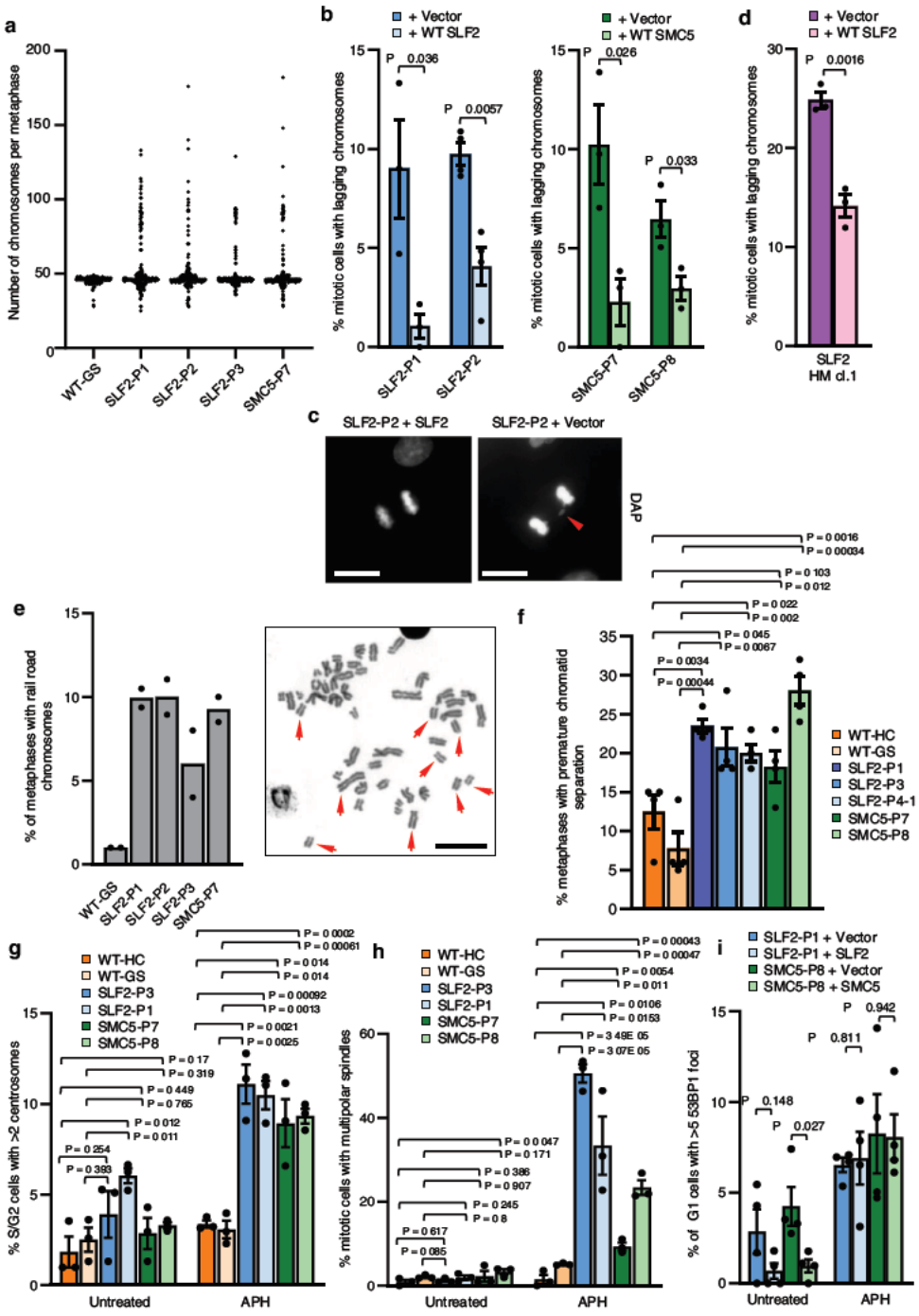
Transient suppression of *slf2* and *smc5* in zebrafish embryos

Splice blocking morpholinos (MOs) were designed to target the *slf2* exon 11 (e11i11) and *smc5* exon 3 (e3i3) splice donor sites (Gene Tools; Supplementary Table 1)). Each gene was transiently suppressed independently by injecting 1 nl at different doses (3, 6, and 9 ng) into one to four cell staged zebrafish embryos. To validate MO efficiency, total RNA was extracted from pools of 2 dpf embryos (25 animals/condition; controls and MO-injected) using Trizol (Thermo Fisher Scientific) according to manufacturer's instructions. cDNA was synthesized with the QuantiTect Reverse Transcription kit (Qiagen), RT-PCR of the MO target locus was performed, and PCR products were separated on a 1% agarose gel. Resulting PCR bands were gel purified with the QIAquick gel extraction kit (Qiagen) and cloned into the TOPO-TA cloning vector (Thermo Fisher Scientific). Purified plasmids from resulting colonies ($n = 4$/PCR product) were sequenced using BigDye 3.1 terminator chemistry according to standard protocols.

Molecular cloning and site-directed mutagenesis of human *SLF2* and *SMC5* constructs for expression of human proteins in zebrafish

Full length Gateway-compatible *SLF2* (NM 018121.4) and *SMC5* (NM 015110.4) open reading frame (ORF) entry vectors were obtained (Horizon). WT ORFs of both genes were inserted into a pCS2+ Gateway

https://doi.org/10.1038/s41467 022 34349 8



destination vector using LR clonase II (Thermo Fisher Scientific). *SMC5* variants identified in either affected individuals (p.His990Asp, p.Arg372del, p.Arg425Ter) or in gnomAD (dbSNP ID: rs59648118, p.(Arg733Gln); 16 homozygotes of 140,814 individuals, negative control) were inserted using site-directed mutagenesis as described (Supplementary Table 1)[72]. After full ORF sequence confirmation of all WT and mutant plasmids, each construct was linearized with NotI and in vitro transcription was performed with the mMessage mMachine SP6 Transcription kit (Thermo Fisher Scientific) according to manufacturer's instructions. 150 pg *SLF2* mRNA with 6 ng *slf2* MO and 150 pg *SMC5* mRNA with 9 ng *smc5* MO was used for in vivo complementation assays.

Fig. 7 | SLF2 and SMC5 patient cells exhibit mosaic variegated hyperploidy, mitotic abnormalities and sister chromatid cohesion defects. a Quantification of the numbers of chromosomes per metaphase in peripheral blood lymphocytes from SLF2 or SMC5 patients, or an unrelated WT individual. 200 metaphases were counted in total from 2 independent blood samples. **b** Average number of mitotic cells with mis segregated lagging chromosomes in SLF2 and SMC5 mutant fibro blast cell lines infected with lentiviruses encoding WT SLF2, WT SMC5, or an empty vector. $n = 3$ independent experiments for SLF2 P1, SMC5 P7 and SMC5 P8, and $n = 4$ independent experiments for SLF2 P2. A minimum of 250 mitotic cells were counted. **c** Representative images of mitotic cells from (**b**) with lagging chromo somes (scale bar: 10 μM). **d** Average number of mitotic cells with mis segregated lagging chromosomes in U 2 OS SLF2 CRISPR HM cells infected with lentiviruses encoding WT SLF2 or an empty vector. $n = 3$ independent experiments. A minimum of 190 mitotic cells were counted. **e** Left: percentage of metaphases with rail road chromosomes in peripheral blood lymphocytes from SLF2 or SMC5 patients, or an unrelated WT individual. A minimum of 380 metaphases were counted in total from 2 independent blood samples. Right: Representative images of metaphases (scale bar: 10 μM). **f** Percentage of metaphases with premature chromatid separation following 4 h treatment with 25 μM MG132 in SLF2 and SMC5 patient LCLs. $n = 4$ independent experiments. 200 total metaphases were counted. **g** Percentage of S/G2 cells (CENPF positive cells) with >2 centrosomes with or without 24 h exposure to 250 nM APH. $n = 3$ independent experiments. A minimum of 900 CENPF positive cells were counted. **h** Percentage of mitotic cells in SLF2 and SMC5 mutant LCLs with multi polar spindles in untreated cells and cells exposed to 250 nM APH for 24 h. A minimum of 300 mitotic cells were counted over 3 independent experi ments. **i** The percentage of G1 phase cells (CENPF negative cells) with >5 53BP1 bodies in SLF2 and SMC5 mutant fibroblast cell lines, with or without 24 h exposure to 500 nM APH. $n = 4$ independent experiments. A minimum of 390 G1 phase cells were counted. In all cases, a Student's t test (two sided, equal variance) statistical test was performed and error bars denote SEM.

Live imaging of zebrafish larvae

Images of tricaine-anesthetized larvae at 3 dpf were captured using the Vertebrate Automated Screening Technology (VAST) Bioimager (Union Biometrica) mounted to an AXIO Imager.M2m microscope (Zeiss) with a *10x* objective lens. Larvae were passed sequentially through a 600 μm capillary on the detection platform. Each larva was detected by software on the computer screen and oriented automatically for lateral and ventral side images with a pre-provided template setting in the software. VAST software (version 1.2.6.7) operated in automatic imaging mode with a 70% minimum similarity threshold, as described[73]. Bright field lateral images were captured with the VAST onboard camera and a fluorescent signal from ventrally positioned larvae with an Axiocam 503 monochrome camera (Zeiss) and ZenPro software (Zeiss).

TUNEL assay and phospho-histone H3 (pHH3) immunostaining in zebrafish larvae

Terminal deoxynucleotidyl transferase biotin-dUTP nick end labeling (TUNEL) assays or pHH3 immunostaining on whole-mount embryos were performed as described[27,74,75]. Embryos were dechorionated at 2 dpf (*slf2* and *smc5*) or 3 dpf (*smc5*) and fixed overnight in 4% paraformaldehyde (PFA) at 4 °C. Embryos were then dehydrated in methanol at −20 °C for 2 h and gradually rehydrated in methanol in PBS and 0.1% Tween (PBST) in the following percent volume/volume ratios: 75/25; 50/50; 25/75 for 10 min each at RT. Embryos were bleached for 12 min in a solution of 9 ml PBST + 1 ml H_2O_2 + 0.05 g KOH before proteinase K treatment and fixation in 4% PFA for 20 min at RT. For TUNEL, embryos were then incubated in equilibration buffer for 1 h and treated overnight with TdT enzyme at 37 °C in a humidified incubator. Following treatment with digoxigenin (ApopTag red in situ apoptosis detection kit, Sigma-Aldrich) for 2 h, embryos were washed *3x* with PBST (10 min each) and processed for imaging. For pHH3 staining, embryos were washed *3x* (10 min each) with PBST and incubated in blocking solution (IF buffer [1% BSA in PBST] + 10% FBS]) for 1 h. Embryos were then treated with primary antibody diluted in 1% BSA overnight: anti-pHH3 (SCBT, sc-374669: 1:500) at 4 °C. Following staining with a secondary antibody: Alexa Fluor 488 goat anti-rabbit IgG (Thermo Fisher Scientific, A11008: 1:500) diluted in 1% BSA for 2 h at RT, embryos were washed *2x* (10 min each) with IF buffer and processed for imaging. For both TUNEL and pHH3 stained embryos, a z-stacked fluorescent signal of the dorsal aspect was captured with a Nikon AZ100 microscope facilitated by a Nikon camera controlled by Nikon NIS Elements Software.

Zebrafish image analysis

ImageJ (NIH) was used to measure lateral head size, ceratohyal angle and count cells (TUNEL or pHH3) in the specified head region. Raw images were exported as TIF files and contrast and brightness were adjusted using identical settings for all images across the experiments. To measure head size, a straight line was drawn from the posterior otolith to the tip of the mouth (line a), the dorsal head area outlined (line b), and the arbitrary shape closed with a line perpendicular to line a (line c). Ceratohyal angle was measured with the angle tool. To count TUNEL or pHH3 positive cells, the image-based tool for counting nuclei (ICTN) plugin for ImageJ was used. A consistent region between the two eyes was selected that spanned the most anterior region of the head to the most anterior region of the yolk.

Immunofluorescence in human cells

Patient-derived fibroblasts or U-2 OS CRISPR HM cells were seeded onto coverslips at least 48 h before extraction and fixation. Cells were pre-extracted for 5 min on ice with ice-cold extraction buffer (25 mM HEPES [pH 7.4], 50 mM NaCl, 1 mM EDTA, 3 mM $MgCl_2$, 300 mM sucrose, and 0.5% Triton X-100) and then fixed with 4% paraformaldehyde (PFA) for 10 min. For immunofluorescence involving patient-derived LCLs, cells were seeded onto Poly-L-Lysine coated coverslips 20 min before fixation with ice-cold methanol for 20 min. For immunofluorescence using cells treated with exogenous DNA damage, patient-derived fibroblasts or LCLs cells were incubated with 500 nM APH, 50 ng/ml MMC or 250 μM CX5461 (Selleck Chemicals, S2684), as indicated in the figure legends, 24 h before fixation.

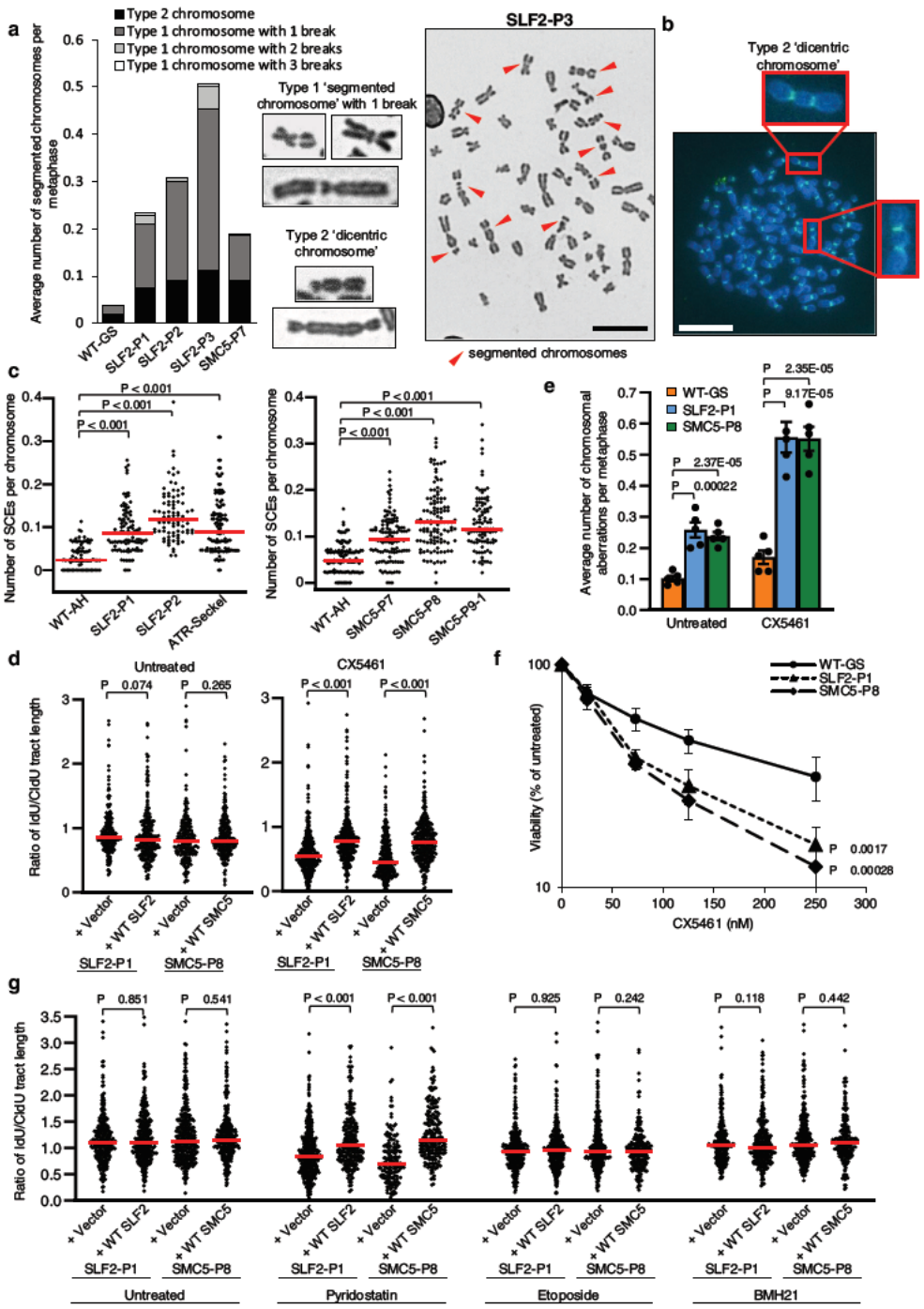
Fixed cells were then stained with primary antibodies specific to γH2AX (Sigma-Aldrich, 05-636; 1:1,000), CENPA (Abcam, Ab13939; 1:750), 53BP1 (Novus Biologicals, NB100−304; 1:1,000), CENPF/Mitosin (Abcam, Ab5; 1:500 and BD Transduction Laboratories, 610768; 1:500), α-Tubulin (Sigma-Aldrich, B-5-1−2; 1:4000), PCNT (Abcam, Ab4448; 1:100), and RAD51 (Merck, PC130; 1:500), and with secondary antibodies: anti-rabbit IgG Alexa Fluor 488 (Thermo Fisher Scientific, A11070; 1:1000) and anti-mouse IgG Alexa Fluor 594 (Thermo Fisher Scientific, A11032; 1:1000). Cells were then stained with DAPI and visualized with a *100x* oil-immersion objective lens on a Nikon Eclipse Ni microscope.

To visualize DNA replication, cells were incubated in medium containing 10 μM EdU for 30−45 min before harvesting. EdU immunolabeling was performed using the Click-iT EdU Imaging Kit (Thermo Fisher Scientific, C10337) according to the manufacturer's protocol.

DNA fiber spreading assay

Patient-derived fibroblasts or U-2 OS cells were seeded at least 48 h prior to harvesting. Cells were incubated with 25 mM CldU for 30 min, washed with media containing 250 mM IdU (with or without 250 μM CX5461, 1 μM pyridostatin, 50 nM etoposide or 1 μM BMH21), incubated with 250 mM IdU (with or without 250 μM CX5461, 1 μM pyridostatin, 50 nM etoposide or 1 μM BMH21) for 30 min, and harvested by trypsinization. For patient-derived LCLs, untreated cells were incubated with 25 mM CldU for 20 min, washed with media containing 250 mM IdU, before being incubated with 250 mM IdU for 20 min and harvested. LCLs were incubated with 50 ng/ml MMC for 24 h prior to pulse

https://doi.org/10.1038/s41467 022 34349 8



labeling with 25 mM CldU for 20 min and then 250 mM IdU for 20 min. For all incubation or washing steps, 50 ng/ml MMC was present in the media. For cells treated with HU, after being incubated with 25 mM CldU for 20 min, LCLs were incubated with media containing 2 mM HU for 2 h, before being washed in media containing 250 mM IdU, then incubated with 250 mM IdU for 20 min and harvested.

Following harvesting, cells were washed with PBS and resuspended to a concentration of 500,000 cells/ml in PBS, and then lysed in lysis buffer (200 mM Tris-HCl [pH 7.5], 50 mM EDTA, 0.5% SDS) directly on glass microscope slides. DNA fibers were spread down the slide by gravity, fixed in methanol/acetic acid (3:1) and denatured with 2.5 M HCl. The thymidine analogs, CldU and IdU, were detected via rat

Fig. 8 | Variants in the RAD18 SLF1/2 SMC5/6 complex compromise the ability of cells to replicate in the presence of stabilized G4 quadruplex structures. **a** Left: Average number of segmented chromosomes per metaphase in peripheral blood lymphocytes (PBLs) from SLF2 or SMC5 patients, or an unrelated WT indi vidual. 250 total metaphases were counted from 2 independent blood samples. Middle: Representative images of 'type 1' and 'type 2' segmented chromosomes. Right: Representative image of a metaphase exhibiting segmented chromosomes from SLF2 P3 PBLs (scale bar: 10 μM). **b** Representative image of FISH with a centromere specific probe showing dicentric chromosomes in a metaphase pre pared from SLF2 P3 PBLs (scale bar: 10 μM). **c** Average number of sister chromatid exchanges in metaphase spreads from SLF2 and SMC5 patient derived LCLs. $n = 3$ independent experiments. A minimum of 100 metaphases were counted. **d** Quantification of the IdU:CldU track length ratio in untreated and CX451 treated SLF2 and SMC5 patient fibroblast cells. Cell lines were pulse labeled first with CldU for 30 min, followed by IdU, with or without 250 nM CX5461, for 30 min. $n = 3$ independent experiments. A minimum of 250 ongoing fork structures were counted. **e** Average number of chromosomal aberrations (chromatid/chromosome gaps, breaks, fragments and chromosome radials) per metaphase in SLF2 and SMC5 patient derived LCLs with and without 24 h exposure to 250 nM CX5461. $n = 5$ independent experiments. A minimum of 350 metaphases were counted. Student's t test (two sided, equal variance) was performed. Error bars denote SEM. **f** LCL proliferation assay. WT and SLF2 and SMC5 patient derived LCLs were cultured in increasing concentrations of CX5461 for the time untreated cells took to undergo three population doublings. Cell viability following CX5461 treatment was calcu lated as a percentage of the number of untreated cells. $n = 4$ independent experi ments. Error bars denote SEM. A two way ANOVA statistical test was performed. **g** Quantification of IdU:CldU track length ratio in untreated, pyridostatin , etopo side and BMH21 treated SLF2 and SMC5 mutant fibroblast cells. Cell lines were pulse labeled first with CldU for 30 min, followed by IdU with or without 1μM pyridostatin, 50 nM etoposide or 1 μM BMH21, for 30 min. $n = 3$ independent experiments. A minimum of 150 ongoing forks were counted. For (**c**, **d**, **g**), red lines denote median values, and a Mann Whitney rank sum statistical test was performed.

anti-BrdU antibody (clone BU1/75, ICR1; Abcam, ab6326; 1:500) and mouse anti-BrdU antibody (clone B44; BD Biosciences, 347583; 1:500) respectively, and secondary antibodies conjugated to Alexa Fluor 594 or Alexa Fluor 488 (Thermo Fisher Scientific). Labeled DNA fibers were visualized with a Nikon Eclipse Ni microscope with *100x* oil-immersion objective lenses, and images were acquired with NIS Elements software (Nikon Instruments). Replication fork structures and CldU and IdU track lengths were then quantified with ImageJ software (US NIH).

Metaphase spreads

Giemsa-stained metaphase spreads from patient-derived cell lines or U-2 OS CRISPR SLF2 HM cells were prepared by adding of 0.2 mg/ml colcemid (KaryoMAX, Life Technologies) and incubating for 3 h. The cells were then harvested by trypsinization, subjected to hypotonic shock for 30 min at 37 °C in hypotonic buffer (10 mM KCl, 15% FCS), and fixed in ethanol/acetic-acid solution (3:1). The cells were dropped onto microscope slides, stained for 15 min in Giemsa-modified solution (Sigma-Aldrich; 5% vol/vol in water), and washed in water for 5 min. For analysis of cohesion fatigue in SLF2 patient LCLs, the metaphase spread protocol was followed as above. However, instead of adding colcemid, 25μM MG132 (Sigma-Aldrich, M7449) was added 4 h before harvesting.

To prepare Giemsa-stained metaphase spreads from peripheral blood, whole blood was diluted in RPMI-1640 and 180 μg/ml PHA (Thermo Fisher Scientific) was added for 48–72 h at 37 °C. 4 h prior to harvesting 0.2 mg/ml colcemid was added. The cells were pelleted and then subjected to hypotonic shock for 10 min at 37 °C in hypotonic buffer (0.075M KCl). Finally, the cells were then fixed in methanol/acetic-acid solution (3:1) and processed as described above.

Fluorescence in situ hybridization

For Fluorescence In Situ Hybridization (FISH) was carried out on peripheral blood lymphocytes metaphases using a peptide nucleic acid (PNA) pan-centromere FISH probe conjugated to Alexa Fluor 488 (5'-ATTCGTTGGAAACGGGA-3', PNA Bio, F3004 CENPB-Alexa488). Briefly, the PNA FISH probes was made up as per the manufacturer's instructions. Metaphase spreads were harvested from patient blood samples as above, and metaphases were dropped onto acetic-acid humidified microscope slides. 24 h later, the slides were rehydrated in PBS, dehydrated in an ethanol series (70%, 95%, 100%) and air dried. The slides were pre-warmed to 37 °C and before being incubated with hybridization buffer (20 mM Tris, pH7.4, 60% formamide, 0.5% blocking reagent [Roche Blocking Reagent, 11096176001], 1% v/v PNA probe) for 10 min at 85 °C. The slides were then incubated in a dark, humidified chamber at RT for 2 h, before being washed in wash buffer (70% formamide, 10-mM Tris) and dehydrated in an ethanol series (70%, 95%, 100%). The slides were then air dried and fixed with prolong gold DAPI mounting medium (ProLong Gold Antifade Mountant with DAPI, P36935).

Sister chromatid exchange analysis

For sister chromatid exchange analysis, LCLs were incubated with 10 μM BrdU for 48 h before incubating with 0.2 μg/ml demecolcine for 3 h. Cells were then resuspended in 0.075M KCl and incubated at 37 °C for 1 h, fixed in methanol/acetic acid (3:1) and dropped onto microscope slides. The slides were then incubated in 10 μg/ml Hoescht for 20 min and exposed to UVA light for 1 h in 2× SSC buffer. Slides were incubated in 2× SSC buffer for 1 h at 60 °C and stained with 5% Giemsa. For metaphase spread analysis of cells treated with exogenous DNA damage, patient-derived LCLs cells were incubated with 500 nM APH or 50 ng/ml MMC 24 h before harvesting.

For analyses of telomere sister chromatid exchange, LCLs were cultured in the presence of BrdU:BrdC (final concentration of 7.5 mM BrdU (MP Biomedicals, 100166) and 2.5 mM BrdC (Sigma-Aldrich, B5002)) for 10 h prior to harvesting. KaryoMAX colcemid (Gibco, 15212-012) was added at a concentration of 0.1 μg/mL during the last 2 h. Cells were collected and washed in 75 mM KCl. Cells were then fixed *3x* in methanol:acetic acid (3:1) by adding fixative solution dropwise with constant gentle agitation by vortex. Following fixation, cells were dropped onto microscope slides and metaphase spreads were allowed to dry overnight. Next, slides were rehydrated in *1x* PBS and then treated with 0.5 mg/ml RNase A (Sigma-Aldrich, R5125) for 30 min at 37 °C. Next, slides were treated with 0.5 μg/ml Hoescht 33258 (Sigma-Aldrich, 861405) in *2x* SSC for 15 min at RT, UV-irradiated, and digested with ExoIII (NEB M0206L) for at least 30 min at 37 °C. Slides were then washed once in *1x* PBS and dehydrated in an ethanol series (70%, 90%, 100%) and air dried. FISH was performed using a TelC-Alexa488-conjugated PNA probe (PNA Bio, F1004; 1:1,000) followed by a TelG-Cy3-conjugated PNA probe (PNA Bio, F1006; 1:1,000) diluted in hybridization solution (10 mM Tris-HCl pH 7.2; 70% formamide; 0.5% blocking reagent (Roche, 11096176001)) each for 2 h at RT. Next, slides were washed at RT twice for 30 min in PNA wash A (70% formamide, 0.1% BSA, 10 mM Tris pH 7.2) and *3x* for 5 min in PNA wash B (100 mM Tris pH 7.2, 150 mM NaCl, 0.1% Tween-20). The second PNA wash B contained DAPI (Life Technologies, D1306) at a 1:1000 concentration. Slides were then dehydrated and dried as described above prior to mounting with Vectashield (Vectalabs, H1000). Slides were imaged using a Zeiss Spinning Disk confocal microscope. Image analyses were blinded and used FIJI version 2.1.0/153.c. Statistical analysis was performed using GraphPad Prism version 9.4.1.

LCL proliferation assays

LCL proliferation assays were carried out as previously reported[49]. Briefly, LCLs were seeded at a concentration of 0.25×10^6 cells per ml in

25 cm^2 flasks and incubated with an increasing concentration of CX5461. The treated cells were counted when the untreated cells had reached a concentration of 2.0×10^6 cells per ml (approximately three population doubling times). The viability of the cells was expressed as a percentage of the untreated cell count.

Plasmids, mutagenesis and sequencing primers

Total RNA was extracted from cell lines using RNeasy Mini kit (Qiagen) according to the manufacturer's instructions. DNA was removed by treatment with DNase I (Qiagen), and cDNA was generated using Superscript II and primed with oligo-dT (Thermo Fisher Scientific). PCR was carried out using Phusion Hot Start II (Thermo Fisher Scientific). *2x*Myc-SLF2 or untagged SMC5 lentiviral expression constructs were generated by cloning a PCR-generated cDNA into the NotI site of pLVX-IRES-neo (Takara Bio). The SLF2 and SMC5 ORFs were verified by sequencing using the primers in Supplementary Table 2.

Full length SLF2 cDNA was also cloned into pcDNA4/TO (Thermo Fisher Scientific) and deletion constructs were generated using KOD Hot Start DNA polymerase (Merck) according to manufacturer's instructions. The following primer sets in Supplementary Table 3 were used to generate the SLF2 deletion constructs and SLF2 'minimal binding region' (MBR) constructs. GFP-SLF2 is previously described[11]. Full length SMC5 cDNA was amplified and cloned into pEGFP-C1 (Takara Bio) using KpnI/BamHI. SLF2/SMC5 mutagenesis was achieved using the Q5 Site-Directed Mutagenesis Kit (E0554S, NEB) according to manufacturer's instructions. The following primer sets in Supplementary Table 3 were used to generate mutant expression vectors. SLF2 p.Gln1162His variant was generated using gene synthesis (Thermo Fisher Scientific).

Lentiviral plasmids encoding the bacterial Holliday junction resolvase RusA were a kind gift from Agata Smorgorzewska[40].

RT-PCR analysis of patient cells

RT-PCR of SLF2 was performed using transcript specific primers (Supplementary Table 4) to assess the mRNA levels of the two longest annotated *SLF2* transcripts (NM 018121.4 and NM 001136123.2) in patient whole blood RNA (Paxgene) or commercially-obtained human cDNA panels: Human Universal QUICK-Clone II (Clontech), which is pool of cDNA obtained from 35 different healthy adult or fetal tissues; and Human multiple tissue cDNA (MTC) panel I (Clontech). PCR product was migrated on a 1% agarose gel for 40 min at 100 V.

CRISPR-Cas9 genome editing of U-2 OS cells

Pairs of SLF2 targeting guide RNAs (sgRNA 1, 5′-AGTTTCATCACTCGGTTCCT-3′; sgRNA 2, 5′-GGCTTGGCACCTTCAAATTC-3′) were designed using the CHOPCHOP web tool (version 2)[71,76] and hybridized and ligated into the purpose built AIO-GFP All-in-One Cas9D10A nickase vector at unique BbsI and BsaI sites. These constructs were transfected into U-2 OS cells using FuGENE transfection reagent according to manufacturer's instructions (3:1 ratio of FuGENE to DNA). Cells were sorted for high GFP expression by fluorescence-activated cell sorting (FACS) into 96-well dishes and recovered in McCoys 5A media supplemented with 20% FBS and 5% penicillin-streptomycin. After 3 weeks, 25 colonies were chosen to be propagated and screened for successful gene editing. After propagating, potential clones were lysed in lysis butter (100 mM Tris/HCl pH 8.5, 5 mM EDTA, 0.2% SDS, 200 mM NaCl, 100 μg Proteinase K/ml) and the DNA was precipitated with isopropanol and resuspended in 10 mM Tris/HCl, 0.1 mM EDTA, pH 7.5. Screening of genomic DNA from clones was achieved by sequencing a region of SLF2 surrounding the Cas9 nickase cut sites (Reverse primer, 5′-AGTTCCGATAATCCACCCCTT-3′; Forward primer, 5′-TTTCTGCAACCAGGTAGTCCT-3′). Following secondary screening of five clones by Western blotting, two SLF2 CRISPR HM clones were chosen (renamed as cl.1 and cl.2) and were characterized further by inserting the amplified region of SLF2 described above into TOPO-TA vectors. 20 TOPO-TA vector clones were then sequenced for both cl.1 and cl.2 to identify all SLF2-mutant alleles and ensure no WT allele was present. The HM clones cl.1 and cl.2 were then complemented by 2xMyc-tagged SLF2 cloned into pLVX-IRES-neo (Takara Bio).

Statistical analysis

Statistical analyses were performed as indicated in the figure legends. A p value of <0.05 indicates significance. The number of independent experimental replicates is denoted in the figure legends. In all cases, independent experiments represent distinct samples, and not the same sample measured repeatedly.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The datasets generated during WES that support this study are available from the corresponding authors upon reasonable request. Informed consents from patients do not cover the deposition of sequencing data from the patient samples, but data can be shared for research purposes with permission of the patient or his/her legal guardian. Gene variant frequency was obtained from the gnomAD database (https://gnomad.broadinstitute.org/). Accession codes for genes/proteins analysed within this study are: Human SLF2 (NM 018121.4, NM 001136123.2, NP 060591.3), Human SMC5 (NM 015110.4, NP 055925.2), zebrafish slf2 (XM 002664123.6, XP 002664169.3), zebrafish smc5 (NM 001193541.1, NP 001180470.1). Plasmids obtained from Addgene (https://www.addgene.org/) used in this study: pLV-hTERT-IRES-hygro (Addgene #85140), psPax2 (Addgene #12260) and pMD2.G (Addgene #12259). PDB files used within this study to model the structural impact of SMC5 patient variants: Saccharomyces cerevisiae Smc5 3HTK [https://doi.org/10.2210/pdb3HTK/pdb], Pyrococcus furiosus RAD50 1F2Thttps://doi.org/10.2210/pdb1F2T/pdb] and 1FTU [https://doi.org/10.2210/pdb1FTU/pdb]. AlphaFold models used to facilitate structural predictions: human SMC5 (AF-Q8IY18-F1). Source data are provided with this paper.

References

1. Alcantara, D. & O'Driscoll, M. Congenital microcephaly. *Am. J. Med. Genet C. Semin. Med. Genet.* **166C**, 124 139 (2014).
2. O'Driscoll, M. The pathological consequences of impaired genome integrity in humans; disorders of the DNA replication machinery. *J. Pathol.* **241**, 192 207 (2017).
3. Grand, R. J. A. & Reynolds, J.J. (eds.). *DNA Repair and Replication: Mechanisms and Clinical Significance*, 348 (Garland Science, 2018).
4. Schmit, M. & Bielinsky, A.K. Congenital Diseases of DNA Replication: Clinical Phenotypes and Molecular Mechanisms. *Int. J. Mol. Sci.* **22** 911 (2021).
5. Klingseisen, A. & Jackson, A. P. Mechanisms and pathways of growth failure in primordial dwarfism. *Genes Dev.* **25**, 2011 2024 (2011).
6. Hanks, S. et al. Constitutional aneuploidy and cancer predisposition caused by biallelic mutations in BUB1B. *Nat. Genet.* **36**, 1159 1161 (2004).
7. Snape, K. et al. Mutations in CEP57 cause mosaic variegated aneuploidy syndrome. *Nat. Genet.* **43**, 527 529 (2011).
8. Yost, S. et al. Biallelic TRIP13 mutations predispose to Wilms tumor and chromosome missegregation. *Nat. Genet.* **49**, 1148 1151 (2017).
9. Pisani, F. M. Spotlight on Warsaw Breakage Syndrome. *Appl Clin. Genet.* **12**, 239 248 (2019).
10. Avagliano, L. et al. Chromatinopathies: A focus on Cornelia de Lange syndrome. *Clin. Genet.* **97**, 3 11 (2020).

11. Raschle, M. et al. DNA repair. Proteomics reveals dynamic assembly of repair complexes during bypass of DNA cross links. *Science* **348**, 1253671 (2015).
12. Aragon, L. The Smc5/6 Complex: New and Old Functions of the Enigmatic Long Distance Relative. *Annu Rev. Genet.* **52**, 89 107 (2018).
13. Sole Soler, R. & Torres Rosell, J. Smc5/6, an atypical SMC complex with two RING type subunits. *Biochem Soc. Trans.* **48**, 2159 2171 (2020).
14. Palecek, J.J. SMC5/6: Multifunctional Player in Replication. *Genes* **10**, 7 (2018).
15. Schuster, K. et al. A neural crest origin for cohesinopathy heart defects. *Hum. Mol. Genet.* **24**, 7005 7016 (2015).
16. Piché, J., Piet Van Vliet, P., Pucéat, M. & Andelfinger, G. The expanding phenotypes of cohesinopathies: one ring to rule them all! *Cell Cycle* **18**, 2828 2848 (2019).
17. Rossi, F. et al. SMC5/6 acts jointly with Fanconi anemia factors to support DNA repair and genome instability. *EMBO Rep.* **21**, e48222 (2020).
18. van der Crabben, S. N. et al. Destabilized SMC5/6 complex leads to chromosome breakage syndrome with severe lung disease. *J. Clin. Investig.* **126**, 2881 2892 (2016).
19. Willemse, B. W. M. et al. New insights in phenotype and treatment of lung disease immuno deficiency and chromosome breakage syn drome (LICS). *Orphanet J. Rare Dis.* **16**, 137 (2021).
20. Payne, F. et al. Hypomorphism in human NSMCE2 linked to pri mordial dwarfism and insulin resistance. *J. Clin. Invest* **124**, 4028 4038 (2014).
21. Pryzhkova, M. V. & Jordan, P. W. Conditional mutation of Smc5 in mouse embryonic stem cells perturbs condensin localization and mitotic progression. *J. Cell Sci.* **129**, 1619 1634 (2016).
22. Hopfner, K. P. et al. Structural biology of Rad50 ATPase: ATPdriven conformational control in DNA double strand break repair and the ABC ATPase superfamily. *Cell* **101**, 789 800 (2000).
23. Hassler, M. et al. Towards a unified model of SMC complex func tion. *Curr. Biol.* **28**, R1266 R1281 (2018).
24. Etheridge, T. J. et al. Live cell single molecule tracking highlights requirements for stable Smc5/6 chromatin association in vivo. *Elife* **10**, e68579 (2021).
25. Kague, E. et al. Skeletogenic fate of zebrafish cranial and trunk neural crest. *PLoS ONE* **7**, e47394 (2012).
26. Atkins, A. et al. SMC5/6 is required for replication fork stability and faithful chromosome segregation during neurogenesis. *Elife* **9**, e61171 (2020).
27. Ansar, M. et al. Bi allelic Variants in DYNC1I2 Cause Syndromic Microcephaly with Intellectual Disability, Cerebral Malformations, and Dysmorphic Facial Features. *Am. J. Hum. Genet.* **104**, 1073 1087 (2019).
28. Khan, T. N. et al. Mutations in NCAPG2 Cause a Severe Neurodevelopmental Syndrome that Expands the Phenotypic Spectrum of Condensinopathies. *Am. J. Hum. Genet.* **104**, 94 111 (2019).
29. Venegas, A. B., Natsume, T., Kanemaki, M. & Hickson, I. D. Inducible degradation of the human SMC5/6 complex reveals an essential role only during interphase. *Cell Rep.* **31**, 107533 (2020).
30. Ogi, T. et al. Identification of the first ATRIP deficient patient and novel mutations in ATR define a clinical spectrum for ATR ATRIP Seckel Syndrome. *PLoS Genet.* **8**, e1002945 (2012).
31. Reynolds, J. J. et al. Mutations in DONSON disrupt replication fork stability and cause microcephalic dwarfism. *Nat. Genet.* **49**, 537 549 (2017).
32. Copsey, A. et al. Smc5/6 coordinates formation and resolution of joint molecules with chromosome morphology to ensure meiotic divisions. *PLoS Genet.* **9**, e1004071 (2013).
33. Behlke Steinert, S., Touat Todeschini, L., Skoufias, D. A. & Margolis, R. L. SMC5 and MMS21 are required for chromosome cohesion and mitotic progression. *Cell Cycle* **8**, 2211 2218 (2009).
34. Stephan, A. K. et al. Roles of vertebrate Smc5 in sister chromatid cohesion and homologous recombination. *Mol. Cell Biol.* **31**, 1369 1381 (2011).
35. Lin, S. J. et al. An acetyltransferase independent function of Eso1 regulates centromere cohesion. *Mol. Cell Biol.* **27**, 4002 4010 (2016).
36. Daum, J. R. et al. Cohesion fatigue induces chromatid separation in cells delayed at metaphase. *Curr. Biol.* **21**, 1018 1024 (2011).
37. Loffler, H., Fechter, A., Liu, F. Y., Poppelreuther, S. & Kramer, A. DNA damage induced centrosome amplification occurs via excessive formation of centriolar satellites. *Oncogene* **32**, 2963 2972 (2013).
38. Marthiens, V. et al. Centrosome amplification causes microcephaly. *Nat. Cell Biol.* **15**, 731 740 (2013).
39. Sarbajna, S., Davies, D. & West, S. C. Roles of SLX1 SLX4, MUS81 EME1, and GEN1 in avoiding genome instability and mitotic cata strophe. *Genes Dev.* **28**, 1124 1136 (2014).
40. Garner, E., Kim, Y., Lach, F. P., Kottemann, M. C. & Smogorzewska, A. Human GEN1 and the SLX4 associated nucleases MUS81 and SLX1 are essential for the resolution of replication induced Holliday junctions. *Cell Rep.* **5**, 207 215 (2013).
41. Agashe, S. et al. Smc5/6 functions with Sgs1 Top3 Rmi1 to com plete chromosome replication at natural pause sites. *Nat. Commun.* **12**, 2111 (2021).
42. Branzei, D. et al. Ubc9 and mms21 mediated sumoylation coun teracts recombinogenic events at damaged replication forks. *Cell* **127**, 509 522 (2006).
43. Menolfi, D. et al. Essential roles of the Smc5/6 complex in replica tion through natural pausing sites and endogenous DNA damage tolerance. *Mol. Cell* **60**, 835 846 (2015).
44. Torres Rosell, J. et al. SMC5 and SMC6 genes are required for the segregation of repetitive chromosome regions. *Nat. Cell Biol.* **7**, 412 419 (2005).
45. Bermudez Lopez, M. et al. Sgs1's roles in DNA end resection, HJ dissolution, and crossover suppression require a two step SUMO regulation dependent on Smc5/6. *Genes Dev.* **30**, 1339 1356 (2016).
46. Bonner, J. N. et al. Smc5/6 Mediated Sumoylation of the Sgs1 Top3 Rmi1 Complex Promotes Removal of Recombination Intermediates. *Cell Rep.* **16**, 368 378 (2016).
47. Masud, T. et al. Ubiquitin mediated DNA damage response is syn thetic lethal with G quadruplex stabilizer CX 5461. *Sci. Rep.* **11**, 9812 (2021).
48. Xu, H. et al. CX 5461 is a DNA G quadruplex stabilizer with selective lethality in BRCA1/2 deficient tumours. *Nat. Commun.* **8**, 14432 (2017).
49. van Schie, J. J. M. et al. Warsaw Breakage Syndrome associated DDX11 helicase resolves G quadruplex structures to support sister chromatid cohesion. *Nat. Commun.* **11**, 4287 (2020).
50. Jegadesan, N. K. & Branzei, D. DDX11 loss causes replication stress and pharmacologically exloitable DNA repair defects. *Proc. Natl Acad. Sci. USA* **118**, e2024258118 (2021).
51. Lerner, L. K. et al. Timeless couples G quadruplex detection with processing by DDX11 helicase during DNA replication. *EMBO J.* **39**, e104185 (2020).
52. Bruno, P. M. et al. The primary mechanism of cytotoxicity of the chemotherapeutic agent CX 5461 is topoisomerase II poisoning. *Proc. Natl Acad. Sci. USA* **117**, 4053 4060 (2020).
53. Pan, M. et a. The chemotherapeutic CX 5461 primarily targets TOP2B and exhibits selective activity in high risk neuroblastoma. *Nat. Commun.* **12**, 6468 (2021).
54. Fernandez Casanas, M. & Chan, K. L. The Unresolved Problem of DNA Bridging. *Genes* **9**, 623 (2018).

55. Wilhelm, T. et al. Mild replication stress causes chromosome mis segregation via premature centriole disengagement. *Nat. Com mun.* **10**, 3585 (2019).
56. Gutierrez Escribano, P. et al. Purified Smc5/6 Complex Exhibits DNA Substrate Recognition and Compaction. *Mol. Cell* **80**, 1039 1054 e6 (2020).
57. Serrano, D. et al. The Smc5/6 Core Complex Is a Structure Specific DNA Binding and Compacting Machine. *Mol. Cell* **80**, 1025 1038.e5 (2020).
58. Adamus, M. et al. Molecular Insights into the Architecture of the Human SMC5/6 Complex. *J. Mol. Biol.* **432**, 3820 3837 (2020).
59. Hallett, S. T. et al. Nse5/6 is a negative regulator of the ATPase activity of the Smc5/6 complex. *Nucleic Acids Res.* **49**, 4534 4549 (2021).
60. Taschner, M. et al. Nse5/6 inhibits the Smc5/6 ATPase and mod ulates DNA substrate binding. *EMBO J.* **40**, e107807 (2021).
61. Spiegel, J., Adhikari, S. & Balasubramanian, S. The Structure and Function of DNA G Quadruplexes. *Trends Chem.* **2**, 123 136 (2020).
62. Inoue, A. et al. Loss of ChlR1 helicase in mouse causes lethality due to the accumulation of aneuploid cells generated by cohesion defects and placental malformation. *Cell Cycle* **6**, 1646 1654 (2007).
63. Peng, X. P. et al. Acute Smc5/6 depletion reveals its primary role in rDNA replication by restraining recombination at fork pausing sites. *PLoS Genet.* **14**, e1007129 (2018).
64. Wechsler, T., Newman, S. & West, S. C. Aberrant chromosome morphology in human cells defective for Holliday junction resolu tion. *Nature* **471**, 642 646 (2011).
65. Jacome, A. et al. NSMCE2 suppresses cancer and aging in mice independently of its SUMO ligase activity. *EMBO J.* **34**, 2604 2619 (2015).
66. Sobreira, N., Schiettecatte, F., Valle, D. & Hamosh, A. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. *Hum. Mutat.* **36**, 928 930 (2015).
67. Logan, C. V. et al. DNA Polymerase Epsilon Deficiency Causes IMAGe Syndrome with Variable Immunodeficiency. *Am. J. Hum. Genet.* **103**, 1038 1044 (2018).
68. Strauss, K. A. et al. Genomic diagnostics within a medically underserved population: efficacy and implications. *Genet. Med.* **20**, 31 41 (2018).
69. Murray, J. E. et al. Mutations in the NHEJ component XRCC4 cause primordial dwarfism. *Am. J. Hum. Genet.* **96**, 412 424 (2015).
70. Poulsen, M., Lukas, C., Lukas, J., Bekker Jensen, S. & Mailand, N. Human RNF169 is a negative regulator of the ubiquitin dependent response to DNA double strand breaks. *J. Cell Biol.* **197**, 189 199 (2012).
71. Labun, K., Montague, T. G., Gagnon, J. A., Thyme, S. B. & Valen, E. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Res.* **44**, W272 W276 (2016).
72. Niederriter, A. R. et al. In vivo modeling of the morbid human genome using Danio rerio. *J. Vis. Exp.* **78**, e50338 (2013).
73. Isrie, M. et al. Mutations in Either TUBB or MAPRE2 Cause Cir cumferential Skin Creases Kunze Type. *Am. J. Hum. Genet.* **97**, 790 800 (2015).
74. Frosk, P. et al. A truncating mutation in CEP55 is the likely cause of MARCH, a novel syndrome affecting neuronal mitosis. *J. Med Genet.* **54**, 490 501 (2017).
75. Stankiewicz, P. et al. Haploinsufficiency of the Chromatin Remo deler BPTF Causes Syndromic Developmental and Speech Delay, Postnatal Microcephaly, and Dysmorphic Features. *Am. J. Hum. Genet.* **101**, 503 515 (2017).
76. Labun, K. et al. CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res.* **47**, W171 W174 (2019).

Acknowledgements

We would like to thank the parents and affected individuals from the Atelís Syndrome families for taking part in this study and generously donating tissue samples. We would also like to thank Agata Smorgor zewska, Kasper Fugger and Stephen West for providing plasmids expressing RusA, Gen1, Mus81 and Eme1. G.S.S., R.H., G.S.M., S.L.C., A.G. are funded by a CR UK Program Grant (C17183/A23303). L.J.G. is supported by a joint funded University of Birmingham and CR UK Ph.D studentship (C17422/A25154). J.J.R. and A.M.R.T. are supported by the University of Birmingham. S.S.J. is supported by a project grant funded by the Great Ormond Street Hospital Charity and Sparks (V5019). R.F.S. and N.Mai are supported by Novo Nordisk Foundation (NNF14CC0001) and Independent Research Fund Denmark (9040 00038B). D.P. and A.P.J. are supported by a European Union Horizon 2020 research and innovation program European Research Council (ERC) Advanced Grant (grant agreement 788093) and by a Medical Research Council (MRC) Unit core grant (U127580972). N.Mat is supported by AMED (JP21ek0109486, JP21ek0109549, and JP21ek0109493). Y.U. is sup ported by JSPS KAKENHI (JP21K15907). A.K.B. and R.M.B. are supported by grants from the National Institutes of Health (R01 GM134681 and R35 GM141805). F.U. is funded by the Higher Education Commission of Pakistan under the International research support initiative program (IRSIP). E.E.D. is supported by US National Institutes of Health grant R01 MH106826 and is the Ann Marie and Francis Klocke MD Research Scholar.

Author contributions

L.J.G., J.J.R., R.F.S., R.M.B., R.H., G.S.M., S.L.C., M.R.H., N.Mai, A K.B., and G.S.S. designed and performed the cell biology experiments. F.U., X.L., T.K., and E.E.D. generated, performed, and supervised analysis on the *slf2* and *smc5* knockdown and knockout zebrafish. L.J.G., A.G., and S.S.J. designed and generated the *SLF2* CRISPR knockout constructs and U 2 OS cell lines. A.W.O carried out structural modeling of patient associated mutations in SMC5. T.N. and N.Mat, designed and generated aid tagged SLF2 degron cell lines. B.I., G.A.M M., S.Ch, C.G.M., D.P., M.A.S., N.N., Z.Y., M.D., A.K., P.V., A.J., S.A.S., C.G J., K.W.B., A.P.A.S., M.K., D.J., Y.U., Y.O., A.M., H.O., Z.A., J.A., C.T.R.M.S., A.M.R.T., A.P.J., and C.LeC. provided patient samples, performed next generation sequen cing, carried out bioinformatic analysis of next generation sequencing data and performed other molecular genetic experiments. J.J.R., A.M.R.T., N.Mai, A K.B., A.P.J., E.E.D. and G.S.S. wrote the paper. G.S.S. planned and supervised the study.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467 022 34349 8.

Correspondence and requests for materials should be addressed to CedricLe Caignec, Erica E. Davis or Grant S. Stewart.

Peer review information *Nature Communications* thanks Dana Branzei and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Laura J. Grange[1,30], John J. Reynolds[1,30], Farid Ullah[2,3,30], Bertrand Isidor[4,30], Robert F. Shearer[5], Xenia Latypova[4], Ryan M. Baxley[6], Antony W. Oliver[7], Anil Ganesh[1], Sophie L. Cooke[1], Satpal S. Jhujh[1], Gavin S. McNee[1], Robert Hollingworth[1], Martin R. Higgs[1], Toyoaki Natsume[8], Tahir Khan[9], Gabriel Á. Martos-Moreno[10], Sharon Chupp[11], Christopher G. Mathew[12], David Parry[13], Michael A. Simpson[14], Nahid Nahavandi[15], Zafer Yüksel[15], Mojgan Drasdo[15], Anja Kron[15], Petra Vogt[15], Annemarie Jonasson[15], Saad Ahmed Seth[16], Claudia Gonzaga-Jauregui[17,18], Karlla W. Brigatti[19], Alexander P. A. Stegmann[20], Masato Kanemaki[21], Dragana Josifova[22], Yuri Uchiyama[23,24], Yukiko Oh[25], Akira Morimoto[25], Hitoshi Osaka[25], Zineb Ammous[11], Jesús Argente[10,26], Naomichi Matsumoto[23], Constance T.R.M. Stumpel[27], Alexander M. R. Taylor[1], Andrew P. Jackson[13], Anja-Katrin Bielinsky[6], Niels Mailand[5], Cedric Le Caignec[28] ✉, Erica E. Davis[2,29] ✉ & Grant S. Stewart[1] ✉

[1]Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham, UK. [2]Advanced Center for Genetic and Translational Medicine (ACT-GeM), Stanley Manne Children's Research Institute, Ann & Robert H Lurie Children's Hospital of Chicago, Chicago, IL, USA. [3]National Institute for Biotechnology and Genetic Engineering (NIBGE-C), Faisalabad, Pakistan Institute of Engineering and Applied Sciences (PIEAS), Islamabad, Pakistan. [4]Service de Génétique Médicale, CHU Nantes, Nantes Cedex 1, France. [5]Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. [6]Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, Minneapolis, MN, USA. [7]Genome Damage and Stability Centre, Science Park Road, University of Sussex, Falmer, Brighton, UK. [8]Department of Chromosome Science, National Institute of Genetics, Research Organization of Information and Systems (ROIS), Mishima, Shizuoka, Japan. [9]Center for Human Disease Modeling, Duke University Medical Center, Durham, NC, USA. [10]Hospital Infantil Universitario Niño Jesús, CIBER de fisiopatología de la obesidad y nutrición (CIBEROBN), Instituto de Salud Carlos III, Universidad Autónoma de Madrid, Madrid, Spain. [11]The Community Health Clinic, Topeka, IN, USA. [12]Sydney Brenner Institute for Molecular Bioscience, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa. [13]MRC Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine, Western General Hospital, The University of Edinburgh, Edinburgh, Scotland. [14]Department of Medical and Molecular Genetics, Faculty of Life Science and Medicine, Guy's Hospital, King's College London, London, UK. [15]Bioscientia Institute for Medical Diagnostics, Human Genetics, Ingelheim, Germany. [16]King Fahad Military Medical Complex, Dhahran, Saudi Arabia. [17]Regeneron Genetics Center, Regeneron Pharmaceuticals Inc., Tarrytown, NY, USA. [18]International Laboratory for Human Genome Research, Universidad Nacional Autónoma de México, Querétaro, México. [19]Clinic for Special Children, Strasburg, PA, USA. [20]Department of Clinical Genetics, Maastricht University Medical Center, Maastricht, The Netherlands. [21]Department of Genetics, The Graduate University for Advanced Studies (SOKENDAI), Mishima, Shizuoka, Japan. [22]Clinical Genetics Department, Guy's Hospital, London, UK. [23]Department of Rare Disease Genomics, Yokohama City University Hospital, Yokohama, Japan. [24]Department of Human Genetics, Yokohama City University Graduate School of Medicine, Yokohama, Japan. [25]Department of Paediatrics, Jichi Medical University School of Medicine, Tochigi, Japan. [26]IMDEA Alimentación/IMDEA Food, Madrid, Spain. [27]Department of Clinical Genetics and GROW-School for Oncology and Developmental Biology, Maastricht University Medical Center, Maastricht, The Netherlands. [28]Centre Hospitalier Universitaire Toulouse, Service de Génétique Médicale and ToNIC, Toulouse NeuroImaging Center, Inserm, UPS, Université de Toulouse, Toulouse, France. [29]Department of Pediatrics; Department of Cell and Developmental Biology, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA. [30]These authors contributed equally: Laura J. Grange, John J. Reynolds, Farid Ullah, Bertrand Isidor. ✉e-mail: [redacted]

PLOS GENETICS

RESEARCH ARTICLE

ATRX proximal protein associations boast roles beyond histone deposition

William A. Scott[1,2], Erum Z. Dhanji[1,2], Boris J. A. Dyakov[2,3], Ema S. Dreseris[1], Jonathon S. Asa[1,2], Laura J. Grange[4], Mila Mirceta[1,2], Christopher E. Pearson[1,2], Grant S. Stewart[4], Anne-Claude Gingras[2,3], Eric I. Campos[1,2]*

1 Genetics & Genome Biology program, The Hospital for Sick Children, Toronto, ON, Canada, **2** Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada, **3** Lunenfeld Tanenbaum Research Institute, Mount Sinai Hospital, Sinai Health System, Toronto, ON, Canada, **4** Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham, United Kingdom

*

Citation: Scott WA, Dhanji EZ, Dyakov BJA, Dreseris ES, Asa JS, Grange LJ, et al. (2021) ATRX proximal protein associations boast roles beyond histone deposition. PLoS Genet 17(11): e1009909. https://doi.org/10.1371/journal.pgen.1009909

Editor: John M. Greally, Albert Einstein College of Medicine, UNITED STATES

Received: October 7, 2021

Accepted: October 23, 2021

Published: November 15, 2021

Data Availability Statement: Data have been deposited as a complete submission to the MassIVE repository and assigned the accession number MSV000088090. The ProteomeXchange accession is PXD028451. Raw datasets used to generate data plots are included in the Supplemental Files.

Funding: EIC is supported by the Canadian Institutes of Health Research (CIHR, PJT 159683), the Natural Sciences and Engineering Research Council of Canada (NSERC, RGPIN 2016 05559), the Cancer Research Society (CRS), and the

Abstract

The ATRX ATP-dependent chromatin remodelling/helicase protein associates with the DAXX histone chaperone to deposit histone H3.3 over repetitive DNA regions. Because ATRX-protein interactions impart functions, such as histone deposition, we used proximity-dependent biotinylation (BioID) to identify proximal associations for ATRX. The proteomic screen captured known interactors, such as DAXX, NBS1, and PML, but also identified a range of new associating proteins. To gauge the scope of their roles, we examined three novel ATRX-associating proteins that likely differed in function, and for which little data were available. We found CCDC71 to associate with ATRX, but also HP1 and NAP1, suggesting a role in chromatin maintenance. Contrastingly, FAM207A associated with proteins involved in ribosome biosynthesis and localized to the nucleolus. ATRX proximal associations with the SLF2 DNA damage response factor help inhibit telomere exchanges. We further screened for the proteomic changes at telomeres when ATRX, SLF2, or both proteins were deleted. The loss caused important changes in the abundance of chromatin remodelling, DNA replication, and DNA repair factors at telomeres. Interestingly, several of these have previously been implicated in alternative lengthening of telomeres. Altogether, this study expands the repertoire of ATRX-associating proteins and functions.

Author summary

ATRX is a protein that is needed to keep repetitive DNA regions organized. It does so in part by binding the DAXX histone chaperone to deposit histone proteins on DNA and assemble structures known as nucleosomes. While important, ATRX has additional functions that remain understudied. To better understand its various biological roles, we first identified the other proteins that are found in its proximity. ATRX associating proteins were implicated in a range of functions, in addition to histone deposition. Our results suggest that ATRX associating proteins likely help compact DNA after it is assembled into nucleosomes, and also promote its stability.

Garron Family Cancer Centre (GFCC). GSS is funded by CR UK (C17183/A23303). ACG is supported by Canada Foundation for Innovation (CFI), Genome Canada, Ontario Genomics OGI 139. WAS received a scholarship from CIHR, and WAS and BJAD received fellowships from NSERC. LJG is supported by a joint funded University of Birmingham and CR UK Ph.D. studentship (C17422/A25154). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

We then examined the effect of ATRX on telomeres (repetitive DNA regions at the end of chromosomes). ATRX and at least one of its associating proteins suppressed spurious DNA exchanges at telomeres. To understand why, we then identified proteomic changes that occur at telomeres when ATRX was deleted. Loss of ATRX altered the enrichment of a surprising number of proteins at telomeres, including several DNA damage response and chromatin remodelling proteins.

Introduction

The alpha thalassemia/intellectual disability, X linked (ATRX) protein is an SNF2 type ATP dependent chromatin remodeller/helicase that maintains chromatin over repetitive DNA regions, such as pericentric heterochromatin and telomeres. The protein promotes chromatin compaction [1 3] and prompt DNA damage repair [4, 5]. ATRX deregulation is intimately linked to disease. Germline *ATRX* mutations cause ATR X Syndrome, a rare X linked disorder characterized by alpha thalassemia and intellectual disability [6]. Loss of functional ATRX is also frequently observed in cancers that utilize alternative lengthening of telomeres (ALT) to extend telomeres and evade replicative senescence [7, 8]. Moreover, ectopic ATRX expression in ALT+, ATRX null cells sharply suppresses ALT activity [9].

ALT occurs in 4 15% of cancers, though much higher rates are observed in malignancies of mesenchymal and neuroepithelial origin [10]. ALT is characterized by long and heterogeneous telomeres [11] that are typically associated with a loss of chromatin compaction [12]; presence of telomeric DNA in PML bodies (known as ALT associated PML bodies, or APBs) [13]; high levels of telomere exchange and synthesis [14 16]; and an accumulation of extrachromosomal, circular telomeric DNA [17]. Recent evidence suggests that ALT is an adaptive response caused by a failure to reactivate telomerase, changes in chromatin compaction, and high levels of DNA damage at telomeres [18 20]. Loss of ATRX is likely an early event in ALT [21], and progressive downstream changes on heterochromatin that enable replication stress at telo meres likely select cells that adopt the ALT phenotype [20]. Cells that use this telomere mainte nance mechanism then experience 'self perpetuating' replication stress and double stranded DNA breaks at telomeres, further sustaining the phenotype [22, 23].

ALT can take different routes (i.e., RAD51 dependent or independent) but is akin to break induced replication and leads to telomere synthesis [18, 19, 24]. This is thought to occur within APBs, where telomeres and DNA repair factors amass [13]. The formation of APBs, and ALT, are influenced by some of these very same DNA repair proteins, including the MRE11 RAD50 NBS1 (MRN) complex [25 27]. The MRN complex initiates DNA end resec tion for homology directed repair [28] and its deregulation promotes ALT activity in some can cer cells [9, 29]. Affinity purification mass spectrometry (AP MS) experiments found ATRX to co purify with the MRN repair complex [30, 31]. ATRX has been proposed to sequester the com plex away from telomeres, thereby limiting homology directed repair associated with ALT [9].

ATRX binds several other proteins. It associates with heterochromatin through its ATRX DNMT3 DNMT3L (ADD) domain that binds trimethylated lysine 9 on histone H3 (H3K9me3) in the absence of H3K4 modifications [32 34]. A PxVxL like motif also binds het erochromatin protein 1 (HP1) and influences the localization of the ATRX protein [34, 35]. ATRX forms a complex with the histone chaperone Death Domain associated Protein 6 (DAXX) to deposit the replication independent histone variant H3.3 [36, 37], thereby main taining nucleosome density and genomic integrity over repetitive DNA elements [1 3, 38, 39]. ATRX also influences sister chromatid cohesion at telomeres. It antagonizes macroH2A1.1

deposition, whose accumulation on chromatin impedes tankyrase 1 dependent resolution of sister telomere cohesion when entering mitosis [40]. As a result, cells that lack ATRX have high levels of intra chromosomal telomeric sister chromatid exchanges (t SCEs). In inter phase, ATRX promotes sister chromatid cohesion and cells lacking ATRX have high levels of inter chromosomal telomere exchanges [15].

ATRX clearly plays multiple roles and is influenced by its protein protein interactions. Comprehensive biochemical characterizations of ATRX protein protein interactions have been invaluable, but are limited to a few stable interactions (i.e., DAXX, the MRN complex) that remain associated after harsh extractions [30, 31, 41]. We therefore screened ATRX pro tein protein associations using proximity dependent biotinylation (BioID) [42, 43], which tags proteins (prey) that are proximal to a protein of interest (bait) in live cells. Our experiments confirmed robust associations between ATRX and DAXX, as well as with the MRN complex. In addition to established interactions, we also identified uncharacterized proximal associa tions with proteins involved in rRNA processing, chromatin remodelling, and the DNA dam age response. We then examined the role of three poorly characterized proteins that likely differed in function. Our proteomic analyses ascribed roles in ribosome biogenesis, chromatin compaction, and telomere stability for FAM207A, CCDC71, and SLF2, respectively, under scoring the diversity of the cellular roles played by ATRX.

SLF2 (SMC5/6 localization factor 2) was previously identified in a proteomic screen for fac tors that accumulate at DNA crosslinks, along with SLF1. Together these proteins link the RAD18 DNA damage response protein and the cohesin like structural maintenance of chro mosome 5 and 6 (SMC5/6) complex [44]. SMC5/6 are conserved proteins that facilitate DNA repair, replication, and mitotic segregation, including over repetitive DNA regions [45, 46]. Recent work using magnetic tweezers showed that SMC5/6 captures and compacts DNA ter tiary structures in an ATP dependent manner [47, 48]. In the context of ALT, deregulation of SMC5/6, and E3 SUMO ligase activity conferred by an accessory protein, facilitates APB for mation, homology directed repair, and ALT [49, 50]. We found ATRX to slightly increase the amount of SLF2 on telomeres. While a loss of either ATRX or SLF2 had a limited effect on telo mere exchanges, the loss of both proteins showed aberrant telomere exchanges. This indicates a role for these proteins in suppressing telomere recombination. To better understand how ATRX and SLF2 impact telomeres, we performed another BioID screen on the RAP1 telomeric (shelterin complex) protein. A loss of ATRX, SLF2, or both proteins revealed important changes in the binding of chromatin remodelling, DNA replication, and repair proteins at telomeres. Taken together, our data highlight a diverse array of functions by which ATRX influences chromatin and telomeres.

Results

BioID reveals novel and diverse associations with ATRX

We theorized that a proximity based protein labeling methodology in live cells would capture ATRX protein protein associations that are not represented when using biochemical means. We therefore made use of proximity dependent biotinylation (BioID) [42, 43] to identify pro teins that directly or indirectly associate with the ATRX protein (Fig 1A). ATRX was fused to a FLAG tagged BirA* biotin ligase at either the N or C terminus, and expressed from isogenic, tetracycline inducible HEK293 Flp In T REx cells (herein denoted as HEK293 Flp In). This produced nuclear bait proteins capable of biotinylating associating proteins in living cells, regardless of whether they are direct or indirect, or static or dynamic. Cells were lysed under stringent conditions, and the proteins purified using streptavidin beads (S1(A) S1(C) Fig). To identify ATRX associating proteins, we rigorously filtered the ATRX BioID against a total of

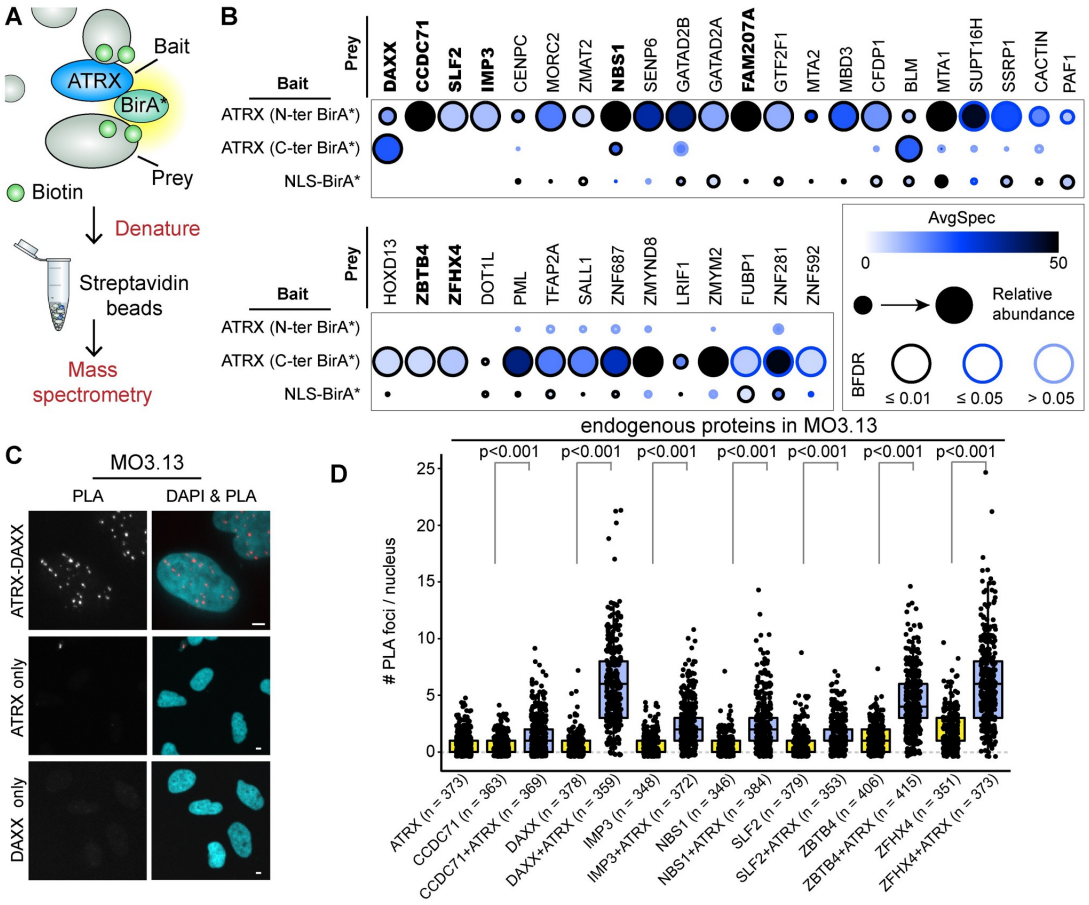


Fig 1. Proximity-dependent biotin identification (BioID) of ATRX-associating proteins. A: Experimental pipeline. The BirA* biotin ligase was fused to ATRX to biotinylate proximal proteins. Labeled proteins were captured on streptavidin beads and analyzed by mass spectrometry. B: Dot plot showing prey proteins identified with ATRX-BirA* that were enriched over endogenous biotinylation (untransfected) and unspecific pan-cellular biotinylation (BirA* alone—BFDR ≤ 5%, SAINT [53]). Data represent two biological replicates. Proteins in boldface remained statistically enriched when the nuclear localization signal (NLS)-BirA* control was used to further filter the ATRX BioID data. C-D: Proximity-ligation assay (PLA) showing proximal associations between endogenous proteins in the MO3.13 human glial cell model. PLA data plots account for ~100 nuclei in three independent experiments, with the exact number of nuclei assessed indicated in brackets. The p-values were obtained using a 2-sided Student's t-test with unequal variance. Scale bars = 4 μm.

https://doi.org/10.1371/journal.pgen.1009909.g001

20 negative controls consisting of untransfected cells and isogenic HEK293 Flp In cells expressing BirA* alone. This identified 36 ATRX proximal associations with a Bayesian false discovery rate (BFDR) ≤ 5% and at least 5 peptides detected across biological duplicates (Fig 1B). BirA* fused to a nuclear localization sequence (NLS) was included for comparison. The DAXX histone chaperone, NBS1 component of the MRN complex, PML protein, BLM heli case, IMP3 U3 snoRNP protein, and the SUPT16H and SSRP1 subunits of the FACT histone chaperone were common to our BioID and published AP MS experiments in HeLa and HEK293T cells [30, 31, 41, 51], or the curated BioGRID database [52]. Affinity purified ATRX predominantly captured DAXX, the MRN complex, and the FACT histone chaperone by mass

spectrometry [30, 31, 41]. BioID expands this repertoire by considering proximal associations and biochemically labile interactions. Gene ontology analysis of the ATRX BioID showed an enrichment of proteins relating to chromosome organization, chromatin remodelling, and transcriptional regulation (S1(D) Fig).

To identify the most robust associations, we further filtered our ATRX associating proteins against proteins identified with the NLS BirA* dataset (Fig 1B and S1 Data). CCDC71, DAXX, FAM207A, IMP3, NBS1, SLF2, ZBTB4, and ZFHX4 remained enriched over all three negative controls (untransfected, BirA* alone, NLS BirA*). Subcellular fractionation experiments were performed, and while some tubulin was detected in the nuclear fraction, the ATRX associating proteins that we probed for were all present in the nucleus (S1(E) Fig). To visualize the endog enous protein associations, a proximity ligation assay (PLA [53]) was performed. PLA gener ates a fluorescent signal when antibodies that recognize associating proteins are in close proximity to one another (Fig 1C and 1D and S1(F)–S1(I) Fig). Endogenous proteins were fol lowed in the human oligodendrocytic MO3.13 cell model [54] because ATRX is highly expressed in brain tissue and is often deregulated in brain tumours. PLA experiments probing for any one protein alone served as a control. PLA signals were obtained for all but one of the eight proteins listed above. Endogenous FAM207A was excluded from our endogenous PLA analyses because commercial antibodies did not recognize the protein by western blotting.

The endogenous PLA signals for CCDC71 and SLF2 in MO3.13 cells were done using the same antibodies previously used by the Human Protein Atlas [55], but we found these two antibodies also recognized additional proteins by western blotting. The endogenous PLA sig nals likely reflect true associations, because proteins that cross-react with the antibodies would also need to be in close proximity to generate a signal. Nevertheless, to further ascertain these associations, PLA signals were further obtained between endogenous ATRX and tagged, exog enous CCDC71, FAM207A, and SLF2 (S1(G)–S1(I) Fig). The FLAG tag increased the abun dance of PLA background signals, but stark differences in signal intensity were observed when protein associations occurred. Because we lacked PLA data for the endogenous FAM207A pro tein, we further validated this interaction *in vitro*. Recombinant FAM207A was expressed using a coupled *in vitro* transcription/translation reaction in the presence of biotinylated lysine. FAM207A was immobilized on streptavidin beads, washed, and incubated with recom binant ATRX expressed in SF9 insect cells. Further washing of the beads showed that the two proteins bind one another (S1(J) Fig).

Altogether, the data show that ATRX forms several uncharacterized associations with a diverse group of proteins. To gauge the breadth of functions imparted by ATRX proximal associations, we overviewed three novel ATRX-associating partners for whom there is limited information namely, FAM207A, CCDC71, and SLF2.

FAM207A is associated with ribosome biogenesis

An estimated 273 genes orthologous to *FAM207A* (*SLX9*) exist across 256 metazoan species [56]. Subcellular fractionation of MO3.13 and HEK293 Flp-In cells expressing the exogenous protein shows varying levels of FAM207A across the fractions of the two lineages (Fig 2A and S2(A) Fig). This may be due to the exogenous protein expression, but possibly also reflects tis sue-specific differences. Immunolabeling of detergent-extracted cells showed that FAM207A strongly localized to the cell nucleoli (Fig 2B). To further ascertain this, cells were then co-labeled for FLAG-FAM207A and IMP3, which is not only an ATRX-associating protein (Fig 1B), but also an established nucleolar protein [57]. Both proteins perfectly co-localized with one another (S2(C) Fig). Curiously, the ATRX-FAM207A PLA signals were not confined to the nucleolus (S1(H) Fig). Because PLA signals form when there is a protein–protein

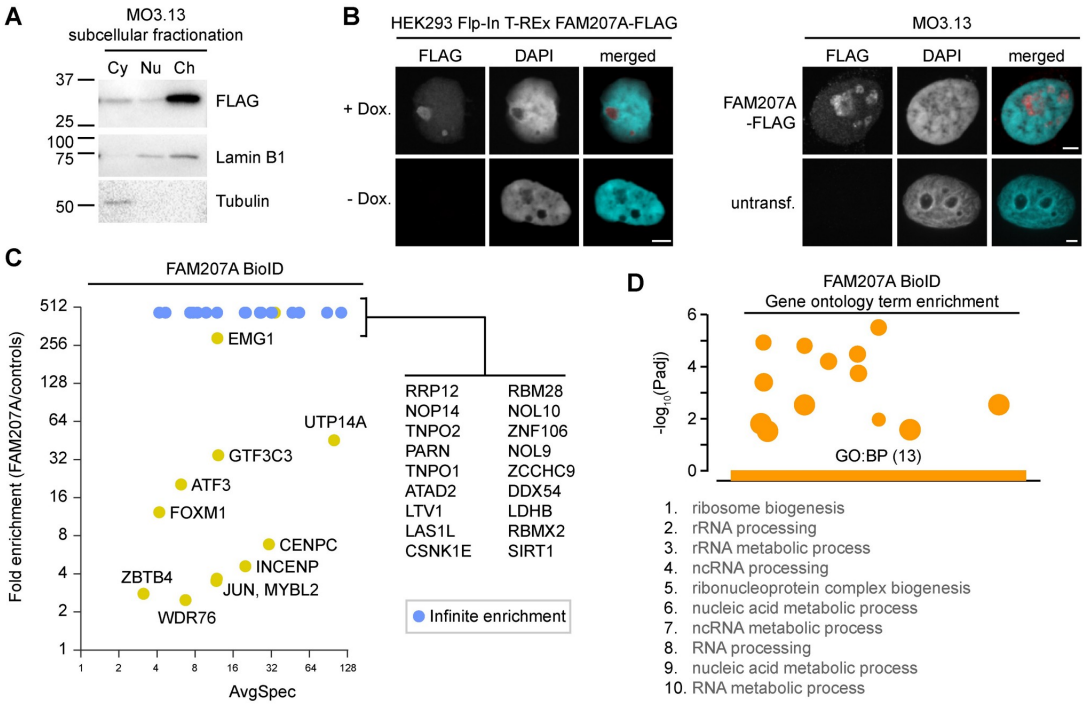


Fig 2. FAM207A is a nucleolar protein involved in ribosome biogenesis. A: Subcellular fractionation of MO3.13 cells expressing FLAG-FAM207A. B: Immunolabeling of exogenous FAM207A in HEK293 Flp-In and MO3.13 cells. C: Fold enrichment of FAM207A BioID prey (BFDR $\leq$ 5%, SAINT [103]) over negative controls. Data were obtained from biological duplicates. D: Gene ontology terms represented in the FAM207A BioID. Cy—cytoplasm; Nu—nucleus; Ch—chromatin. Scale bars = 4 μm.

https://doi.org/10.1371/journal.pgen.1009909.g002

interaction, the signals outside the nucleolus could suggest role(s) for ATRX FAM207A distinct from those in ribosome biogenesis. It is also worth noting that the Human Protein Atlas database suggests that the protein resides in the Golgi apparatus [58]. However, despite high spectral counts for the protein in our BioID, we were not able to detect the endogenous protein by western blotting when we used the same antibody, even under low stringency conditions (e.g., no tween, high antibody concentration).

A prior affinity purification of the nucleolar NOC4 protein identified FAM207A as a likely component of pre ribosomal particles [59] and further work suggests a role in the export of 40S pre ribosomes [60]. Our FAM207A BioID results confirm an association with proteins involved in ribosome maturation (Fig 2C and 2D and S2(D) Fig). Twenty nine proteins were enriched in the FAM207A BioID (BFDR $\leq$ 1%) in comparison to the BirA* and NLS BirA* controls (Fig 2C). A much larger number of proteins were retained when NLS BirA* was used as a comparative bait instead of a control (S2(D) Fig and S1 Data). We believe that the large number of associated proteins reflects the protein's involvement in macromolecular assemblies. While ATRX was detected in the FAM207A BioID with a relatively high average spectral count of 27.5, it was filtered out by the negative controls in the SAINT analysis. This perhaps indicates that ATRX only constitutes a minor proportion of the protein associations made by FAM207A, since we otherwise detected the ATRX FAM207A association through the ATRX BioID and PLAs. Gene ontology groups relating to ribosome biogenesis and rRNA processing

predominated (Fig 2D), as previously suggested [59]. FAM207A, however, also associated with proteins involved in chromosome and transcriptional regulation, again suggesting more than one biological function.

CCDC71 is a chromatin-bound protein that associates with HP1

Current databases and our own experimental data implicate CCDC71 in chromatin organiza tion. CCDC71 belongs to the coiled coil domain 71/71L family found within vertebrates (Fig 3A). In addition to a coiled coil region, the protein harbours an NLS and several regions that are predicted to be disordered (Fig 3B). The mouse protein contains two canonical PxVxL sequences, a motif found in proteins that bind HP1 [35]. Variation within the sequence is tol erated to various degrees [35, 61], and both the human and the consensus sequence from 254 putative orthologs contain PxVxL like sequences at these two locations (Fig 3C). A canonical PxVxL sequence is also found in the human CCDC71L protein, a shorter family member that shares a high degree of identity within the N terminal half of CCDC71 and its C terminal extremity (S3(A) Fig). It was, therefore, not surprising to see that the exogenous protein was almost exclusively found in the chromatin fraction of MO3.13 and HEK293 Flp In cells (Fig 3D and S3(B) Fig).

To our knowledge, there are no reports on CCDC71, but data gathered by the Human Pro tein Atlas [55] suggest that the protein locates to the nuclear periphery in HeLa cells, with some cell types showing a more nuclear signal. Immunolabeling in HEK293 Flp In cells showed pan nuclear granular signals (Fig 3E), a pattern also seen in the ATRX CCDC71 PLA experiments in HEK293 Flp In cells (S1(G) Fig). Putative CCDC71 protein protein interac tions reported in BioGrid [52] suggest a role in chromatin compaction. They include histone H3.3, the EED subunit of the polycomb repressive complex 2 (PRC2), the RbAp46 histone chaperone, and the HDAC1 histone deacetylase. We could indeed co immunoprecipitate H3.3, EZH2 (the catalytic subunit of PRC2), and HDAC1 with CCDC71 (S3(C) Fig). To obtain a clearer portrait of CCDC71 functions and its proximal protein associations, a BioID analysis was also performed for this protein. While HDAC1 was represented, the most preva lent associations included ATRX, all three HP1 isoforms, the NAP1 histone chaperone, and components of the NuRD chromatin remodelling protein complex (also represented in the ATRX BioID; Fig 3F, S3(E) Fig and S1 Data). As expected, a gene ontology analysis of the CCDC71 BioID results included terms related to chromosome and chromatin organization (Fig 3G). This shows that ATRX and CCDC71 share important functions in chromatin regulation.

SLF2 prevents spurious telomere recombination

The SLF2 proteome was previously described [44], but the association with ATRX is novel. It was found that SLF1/2 loads the cohesin like SMC5/6 proteins at sites of DNA damage [44]. A genetic interaction between *ATRX* and *SMC5* and *SMC6* was recently reported [62] and dereg ulation of SMC5/6 in ALT positive cells promotes APB formation and telomere exchanges [49]. To determine if ATRX influences SLF1/2 and SMC5/6 recruitment to telomeres, we dis rupted ATRX expression in HEK293 Flp In cells using CRISPR Cas9 (Fig 4A), and quantified SLF2 occupancy at telomeres. Exogenous SLF2 and endogenous SMC5 were imaged by immu nofluorescence and telomeric regions identified using fluorescence *in situ* hybridization (IF FISH). SLF2 and SMC5 formed foci whose signals strongly overlapped with one another (S4(A) and S4(B) Fig). Most SLF2 and SMC5 foci formed outside telomeres, but approximately 20% of the signals coincided with those from the telomeric FISH probe (Fig 4B). The propor tion of SLF2/SMC5 overlap with telomeres decreased in the ATRX null cells. G quadruplex

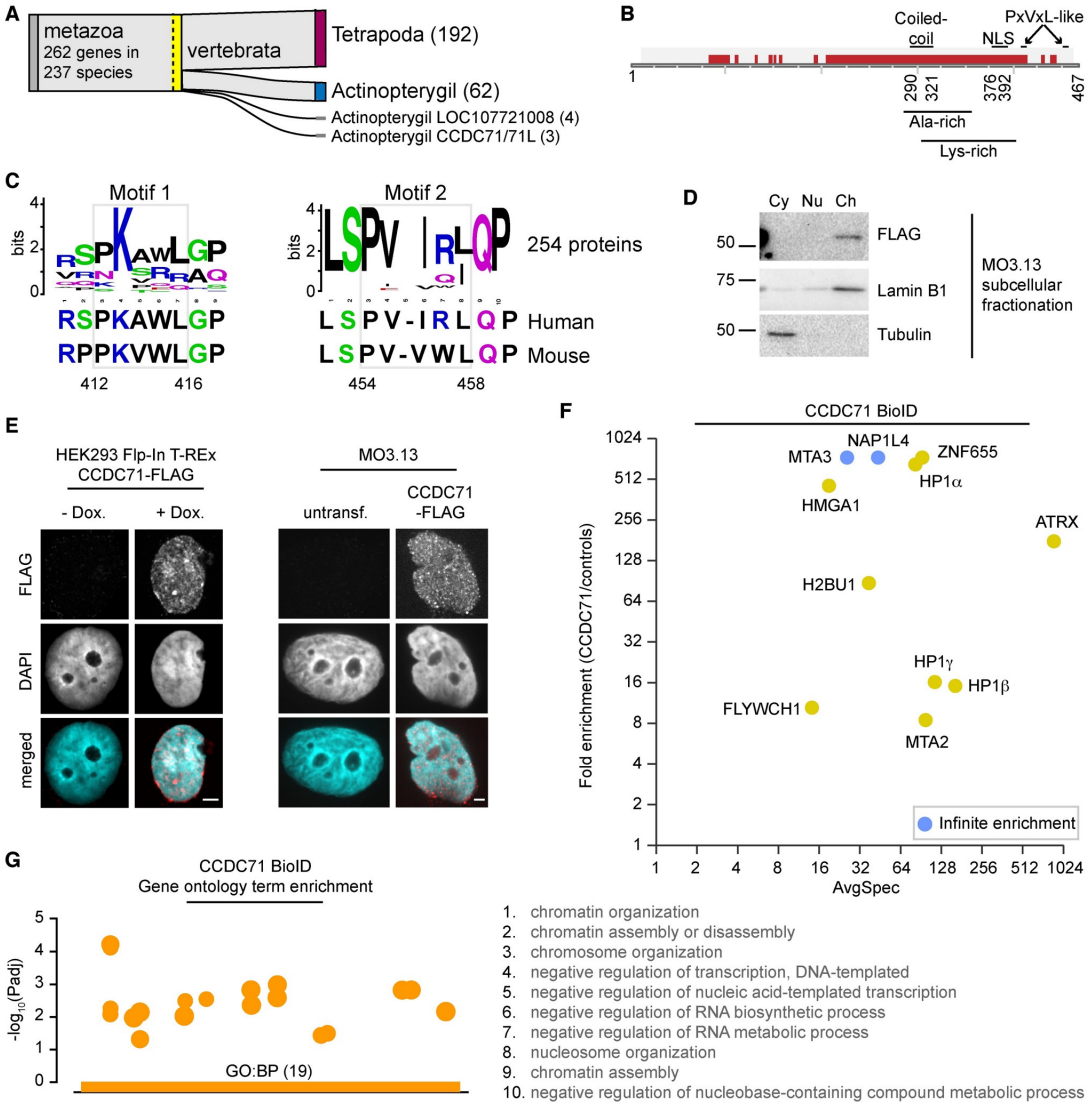


Fig 3. CCDC71 is a chromatin-bound protein that associates with HP1. A: Hierarchical catalogue of orthologs (OrthoDB v10.1) [56] diagram showing 262 putative genes within the *CCDC71/71L* family. B: Graphical depiction of the CCDC71 primary structure and features identified using Pfam [110] and Prosite [111]. Putative PxVxL-like motifs are also shown. The CCDC71/71L domain is shown in light grey and putative disordered regions in dark red. C: Sequence analysis of 254 CCDC71 orthologs aligned around the two PxVxL-like motifs. The sequence alignment was generated using COBALT [112] and visualized using WebLogo [113]. D: Subcellular fractionation of MO3.13 cells expressing FLAG-CCDC71. E: Immunolabeling of exogenous CCDC71 in HEK293 Flp-In and MO3.13 cells. F: Fold enrichment of CCDC71 BioID prey (BFDR $\leq$ 5%, SAINT [103]) over negative controls. Data were obtained from biological duplicates. G: Gene ontology terms represented in the CCDC71 BioID. Scale bars = 4 μm.

https://doi.org/10.1371/journal.pgen.1009909.g003

(G4) structures can form at telomeres, where they cause DNA damage if not properly resolved. The analysis was therefore repeated in cells exposed to the G4 stabilizing small molecule, pyridostatin [63], but the treatment had no impact on the SLF2/SMC5 telomere signal overlap (Fig 4B).

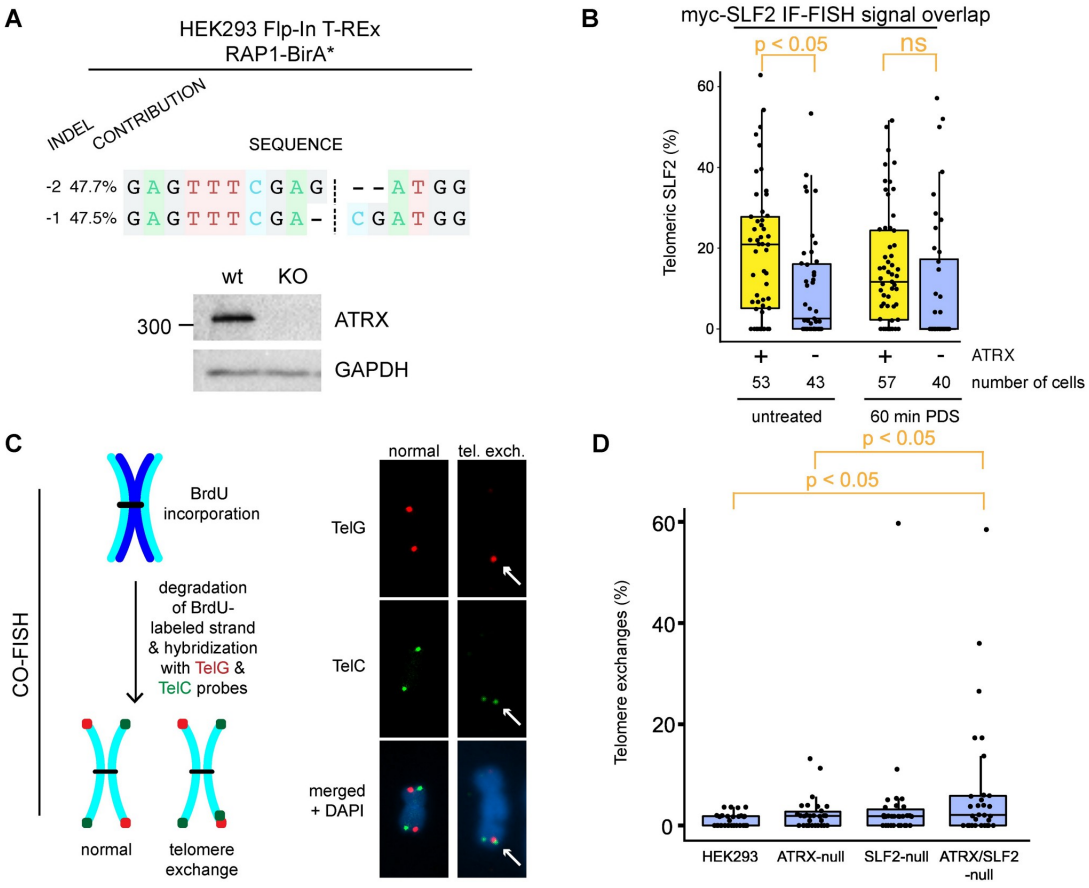


Fig 4. Loss of ATRX and SLF2 enables telomere exchanges. A: DNA sequencing and western blotting confirming the disruption of ATRX expression in CRISPR-Cas9-edited HEK293 Flp-In T-REx cells inducibly expressing RAP1-BirA*. B: Relative signal overlap between immunolabeled SLF2 and telomere fluorescence *in situ* hybridization (IF-FISH), in ATRX-expressing or -null HEK293 Flp-In cells. Cells were exposed to 10 μM pyridostatin (PDS, G4 stabilizer) for 60 min. At least 40 cells per condition were analyzed using a CellProfiler colocalization pipeline [106]. ATRX loss decreased SLF2 recruitment to telomeres in the absence of pyridostatin. C: Schematic of dual-colour chromosome-orientation fluorescence *in situ* hybridization (CO-FISH) [64] using PNA probes to label C- and G-rich telomere strands (left). Examples of normal CO-FISH signals and of a chromosome end with a telomere exchange are shown (right). D: Relative amount of telomere exchanges in ATRX, SLF2, and ATRX/SLF2 KO HEK293 Flp-In cells. At least 30 mitotic spreads per condition were counted and the percent of telomeric exchanges plotted. p-values were obtained using a 2-sided Student's t-test with unequal variance.

https://doi.org/10.1371/journal.pgen.1009909.g004

ATRX expression restrains telomere exchanges [9, 15, 40]. We therefore wanted to deter mine if ATRX and SLF2 together repress t SCEs. To investigate this, we disrupted SLF2 in ATRX expressing and null HEK293 Flp In cells (S4(C) and S4(D) Fig), and performed chro mosome orientation fluorescent *in situ* hybridization (CO FISH) [64] to visualize and quan tify telomere exchanges (Fig 4C and 4D). As recently reported [15], we observed increased levels of t SCEs (dual signal on matching sister chromatid ends), but also a high proportion of exchanges on telomere ends of single chromatids in our ATRX null cells (Fig 4D). This can indicate non allelic exchanges and the events were therefore grouped as 'telomere exchanges,' as previously done [15]. Low p values were obtained when comparing telomere exchange rates

between the single null cells and the parental HEK293 Flp In (e.g., 0.059 and 0.15 for ATRX and SLF2 null cells, respectively), but significance was reached in double ATRX/SLF2 null cells (Fig 4D). This suggests that both proteins inhibit spurious exchanges. However, the lack of a clear synergistic effect may signify that the proteins do so independently of one another. In the ALT positive (ATRX null) U2OS cells, an *SLF2* gene disruption again caused a signifi cant increase in telomere exchanges (S4(E) Fig). Altogether, this shows that ATRX has some influence on SLF2 recruitment, and a loss of both proteins permits high levels of telomere exchanges.

ATRX and SLF2 depletion alters protein enrichment at telomeres

A loss of ATRX causes gradual changes on telomeres [20], so we wanted to determine if its loss, or that of SLF2, caused immediate observable changes in protein enrichment at telomeres. To examine this, we performed BioID on the RAP1 (TERF2IP) subunit of the shelterin com plex, in parental HEK293 Flp In cells and compared the results to that of the same BioID done in isogenic SLF2 null, ATRX null, or cells stably expressing ATRX shRNA (Fig 5 and S5 Fig). The same BioID strategy, but using TERF1 as a bait, previously identified changes in the telo meric proteome of ALT negative and ALT positive cells [65]. While we did not enrich for all the shelterin components, TERF2 was one of the most abundant prey proteins identified in the RAP1 BioID (S1 Data). Importantly, 38% of the 343 proteins identified by our RAP1 BioID in HEK293 Flp In cells were also identified by the previously reported TERF1 BioID in HeLa cells [65], and 55% our of preys were also identified by other proteomic techniques used to identify telomeric proteins (namely, PICh [66], QTIP [67], C BERST [68], and TERF1 BioID [65]; S5(C) Fig). A similar overlap was seen when using NLS BirA* as an additional control in our BioID, but we first omitted the control from this analysis to identify as many changes as possible (data were still filtered against untransfected cells and BirA*). As expected, RAP1 BioID experiments performed in parental and ATRX or SLF2 depleted HEK293 Flp In cells were far more similar to one another than our other BioID experiments with other baits (Fig 5A).

The ATRX knockout (Fig 5C) and knockdown (S5(D) Fig) caused surprisingly vast changes at telomeres, with proteins involved in chromatin maintenance, DNA replication, and repair showing both gain and loss of abundance. The relative abundance of some the proteins that associated with ATRX (such as NBS1 and PML) also changed in relative abundance at telo meres, but the most prominent changes involved other non associating proteins. Loss of ATRX therefore carries indirect downstream changes that alter telomeres. We then examined the changes that occur in SLF2 null and isogenic SLF2/ATRX double null cells (Fig 6). In com parison to ATRX, SLF2 depletion had a far more limited effect on telomeric proteins, although chromatin remodelling and DNA repair proteins again changed in abundance at telomeres (Fig 6A). The double null cells largely recapitulated the effects seen in the individual gene knockouts, suggesting that most of the effects are additive.

The increase in PIN2/TERF1 interacting telomerase inhibitor 1 (PINX1) at telomeres upon SLF2 loss is interesting because the protein inhibits telomerase activity [69] and promotes TERF1 binding to telomeres [70]. Proteomic changes caused by a loss of ATRX and/or SLF2 therefore likely destabilize telomeres.

Discussion

Isolation of ATRX binding proteins yields biochemically stable interactions with DAXX and components of the MRN protein complex [30, 31, 41], yet ATRX binds additional proteins. The discrepancy is likely due to the need for harsh extraction methods to analyze its

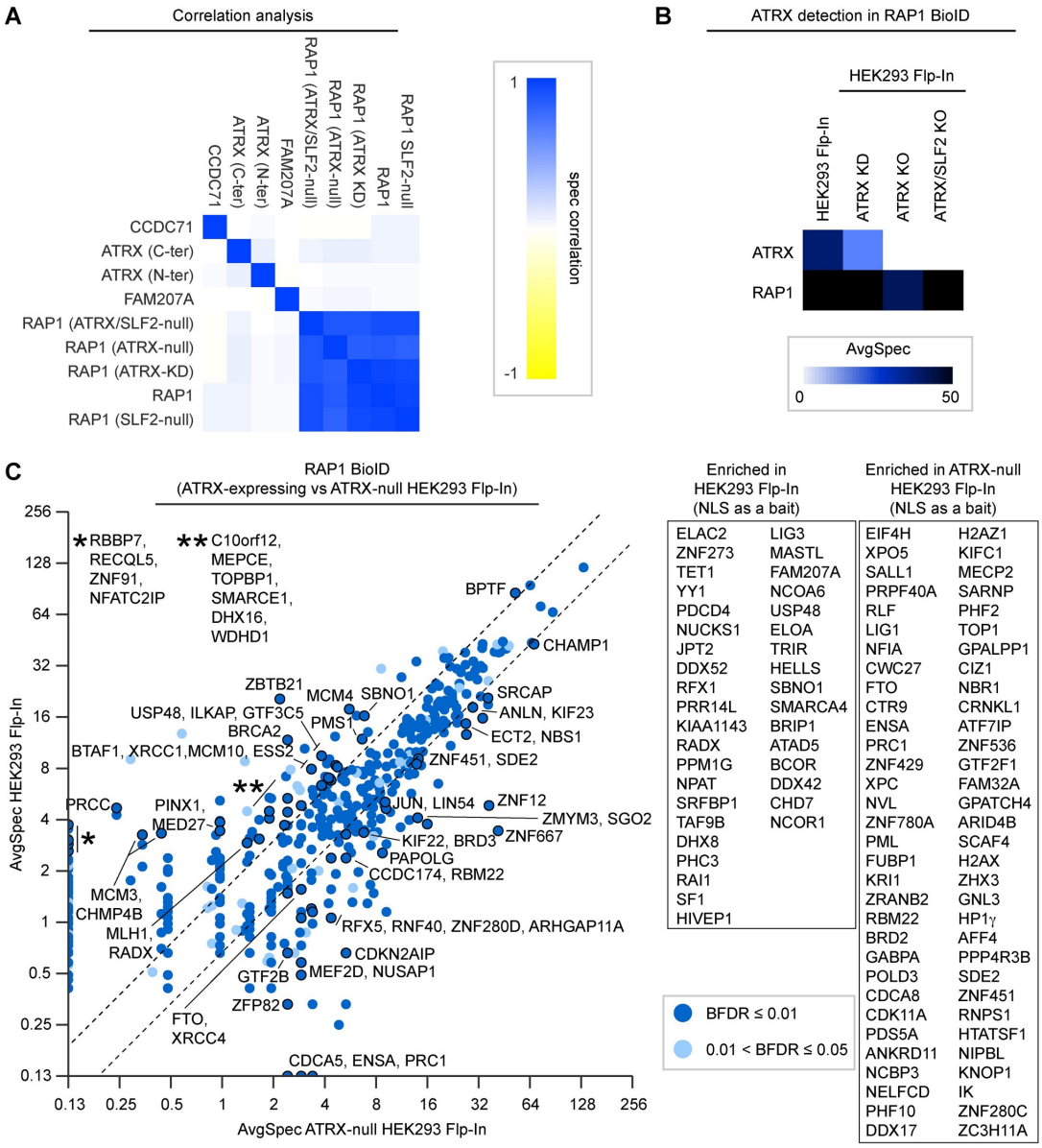


Fig 5. Loss of ATRX changes the enrichment of a subset of proteins at telomeres. A: Correlation between BioID bait proteins. RAP1 BioID experiments strongly correlated with one another independently of ATRX expression. B: Relative ATRX detection by RAP1 BioID in ATRX-expressing and isogenic cells where ATRX expression was knocked-down (KD) using shRNA or knocked-out (KO) using CRISPR-Cas9. C: Fold change of prey proteins identified by RAP1 BioID (BFDR ≤ 5%, SAINT [103]) in ATRX-expressing vs. ATRX-null HEK293 Flp-In cells. Labeled proteins had ≥2.5 average spectra and ≥1.5-fold change between conditions. Proteins labeled in the scatter plot remained significant when NLS-BirA* was used as a control. Inset box represents additional proteins identified with NLS-BirA* as a bait (less stringent).

https://doi.org/10.1371/journal.pgen.1009909.g005

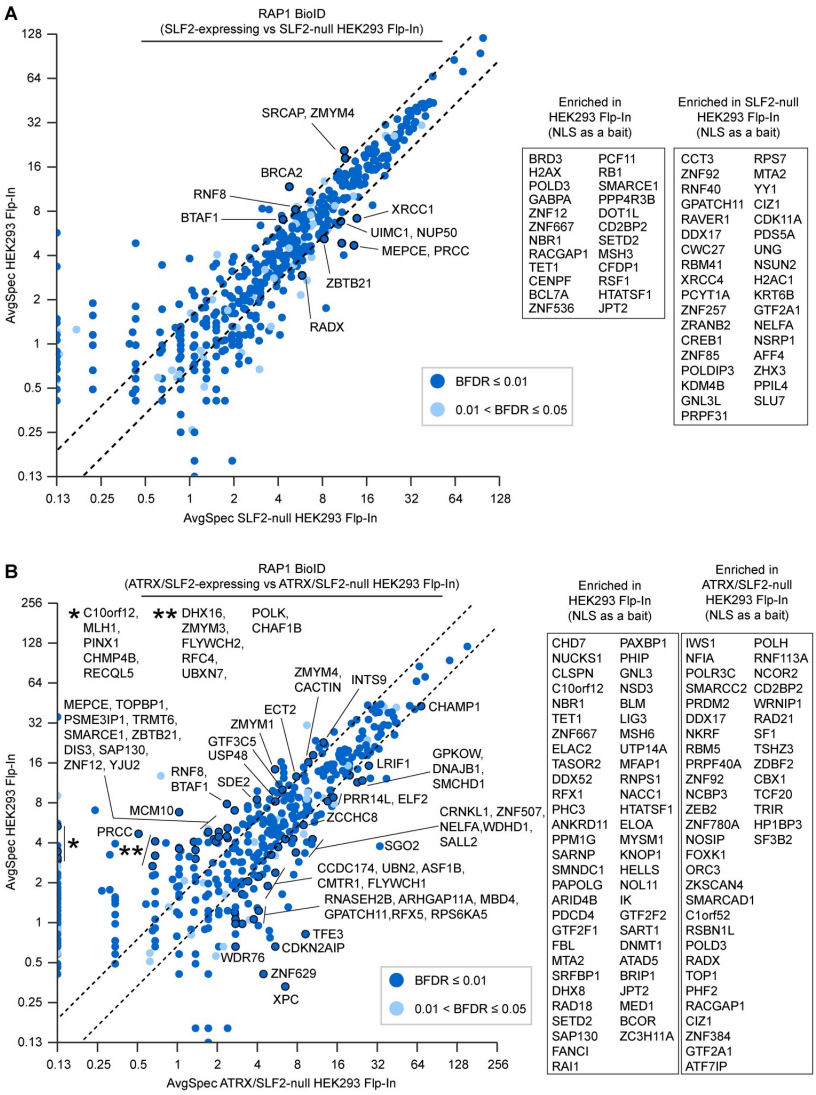


Fig 6. Changes at telomeres caused by a loss of SLF2 and ATRX. Fold change of prey proteins identified by RAP1 BioID (BFDR ≤ 5%, SAINT [103]) in A: SLF2-null, and B: ATRX/SLF2-null vs. ATRX- and SLF2-expressing HEK293 Flp-In cells. Labeled proteins had ≥2.5 average spectra and ≥1.5-fold change between conditions. Proteins labeled in the scatter plot remained significant when NLS-BirA* was used as a control. Inset box represents additional proteins identified with NLS-BirA* as a bait (less stringent).

https://doi.org/10.1371/journal.pgen.1009909.g006

interactions on chromatin. With the advent of proximity based labeling techniques [43], we were able to surmount this technical limit and identify 36 ATRX associating proteins (Fig 1B). Our BioID results are high confidence since the data were stringently filtered against 20 BioID control replicates. To identify the strongest associations (whether direct or indirect), we fur ther filtered the data against an additional 14 biological replicates of NLS BirA*, where

indicated. While DAXX and NBS1 were enriched in the BioID experiments performed with BirA* at either the N or C terminus of the protein, there was a surprising degree of variation between the two baits. The differences could be due to spatial constraints within these associations, but also hindrance or misfolding caused by the bulky biotin ligase. However, both protein termini are predicted to be of very low complexity and unstructured [71], making misfolding of BirA* tagged ATRX less likely. Spatial specificity via BioID has also been reported for other proteins, where data obtained by biotin ligases at the N and C terminus of the bait protein were seen as complementary [72, 73]. We were indeed able to visualize the high confidence associations by PLAs regardless of the BirA* position (Fig 1C and 1D), suggesting that the differences are most likely due to spatial constraints on the ~300 kDa protein. It is interesting to note that ATRX has two DAXX interacting regions: a weak N terminal one (spanning a.a. 321–865) and a stronger C terminal one via ATRX's DAXX Binding Motif (a.a. 1189–1326) [74] and that our BioID results reflect that difference. The interaction with PML is also mediated by the C terminal extremity of ATRX [75], again recapitulated in the BioID.

Interestingly, the previous affinity purified ATRX mass spectrometry experiments, and our ATRX BioID dataset, lacked reported interactions with HP1 [35, 76, 77], macroH2A.1 [78, 79], and MeCP2 [80, 81]. These interactions had been established by immunoprecipitation/western, co localization, or *in vitro* experiments. While HP1β (CBX1), HP1γ (CBX3), and macroH2A.1 were detected in the ATRX BioID, they were filtered out because they were also labeled by our negative controls (S1 Data). MeCP2 was, however, only detected by the negative controls. Some of the results could reflect cell type specific differences and/or spatiotemporal limits of the biotin ligase. Interestingly, HP1 was later captured not by ATRX, but rather a novel ATRX associating protein (see below). The specific identification of the NBS1 component of the MRN complex (and lack of MRE11 and RAD50 subunits) does not necessarily mean that ATRX only associates with NBS1, but could reflect the relative position of the biotin ligase over the protein complex. Indeed, others showed that ATRX also co purifies with the other MRN complex subunits [30, 31].

Because the ATRX associations enriched for ribosomal, chromatin, and DNA repair proteins, we then examined three associating proteins for which there is limited information, but that likely reflects these functions. FAM207A, CCDC71, and SLF2 were chosen because there is limited information on the proteins, and associations reported in BioGrid [52] suggested that they had very different biological functions. Our data confirm that FAM207A is predominantly nucleolar (Fig 2B). This is interesting because ATRX does enrich over rDNA repeats across different life kingdoms [76, 82–84]. ATRX depletion in mESCs destabilizes rDNA repeats through a loss of histone deposition and repressive histone marks, and a decrease in rDNA copy numbers and rDNA transcription that renders cells sensitive to RNA polymerase I inhibitors [85]. The FAM207A protein previously co purified with isolated pre ribosomal particles and RNAi mediated depletion of the protein implicated it in 40S precursor maturation [59]. Our data concur and demonstrate that FAM207A associates with proteins involved in ribosome biogenesis (Fig 2C and 2D). However, FAM207A BioID also identified proteins with functions in gene transcription and chromosome organization (S2(D) Fig), and our PLA analysis showed that the ATRX FAM207A association is not exclusive to the nucleolus (S1(H) Fig). In fact, like ATRX, yeast FAM207A (Slx9) is proposed to bind G quadruplex DNA structures that form over repetitive G rich DNA [86] and it will be very interesting to find the exact role of the ATRX FAM207A association.

A second ribosomal protein, IMP3, was also identified as an ATRX interacting protein in our BioID analysis (Fig 1B). IMP3 is a conserved nucleolar rRNA processing protein also implicated in ribosome biogenesis [57]. The ATRX IMP3 interaction was also reported in the BioPlex interactome database [51]. While FAM207A BioID did not capture IMP3, it is possible

that their functions converge. It is also worth noting that a hypomorphic *IMP3* mutant allele in budding yeast sensitised cells to various DNA damaging agents and produced slightly longer telomeres [87]. Additional experiments will be needed to determine if this holds true in other organisms.

CCDC71 is an uncharacterized coiled coil domain containing protein for which there are no associated functions. Work done by the Human Protein Atlas program [55] suggests that the protein enriches at the nuclear periphery in HeLa and MCF7 cells and forms nuclear foci in U2OS cells. While it may be tempting to attribute the difference to the utilization of the ALT pathway, we found the antibody used in those experiments to also recognize several other proteins by western blotting. Our labeling of an epitope tagged exogenous protein showed a pan nuclear granular nuclear signal that was not exclusive of DAPI dense regions (Fig 3E). Curated data suggest that CCDC71 binds histone H3.3 [52], components of the PRC2 complex [88], HDAC1 [89], and RBBP7 [89], linking the protein to chromatin compaction and tran scriptional repression. We confirmed these interactions by immunoprecipitating the exoge nous CCDC71 protein (S3(C) Fig). However, BioID provided important information on the most prominent associations (Fig 3F and S3(E) Fig). Our analysis clearly demonstrates a strong association between CCDC71 and HP1 isoforms, likely mediated by non canonical PxVxL motifs in human cells (and canonical ones in mice) and/or through further associations with the related CCDC71L protein (Fig 3C and 3F, S3(A) and S3(E) Fig). In addition to con firming ATRX as a strong association, the CCDC71 BioID showed clear associations with the NAP1 histone chaperone and components of the NuRD chromatin remodelling complex (which were also seen in the ATRX BioID, Fig 1B).

It is interesting to note that the ZFHX4 transcription factor (identified in the ATRX BioID) also interacts with NuRD to promote stem cell like states [90]. Genetic deletions over the *ZFHX4* locus have been associated with syndromic Peters anomaly [91], ocular abnormalities [92], and intellectual disability [93], and the protein plays a role in maintaining the undifferen tiated, self renewing state of glioblastoma tumour initiating cells [90]. Further studies will be needed to examine the effect of the ATRX CCDC71 HP1, NAP1, and NuRD associations on chromatin. It will also be interesting to determine how these proteins (including those, such as the PRC2 complex, confirmed by immunoprecipitation in S3(C) Fig) functionally work together.

Chromatin maintenance is intimately tied to genome stability; hence, our investigations into the third chosen ATRX proximal association, SLF2. SLF2 associates with SLF1 and RAD18 and loads the cohesin like SMC5/6 complex to promote genomic stability [44]. SLF1/2 knockdown impairs the recruitment of SMC5 to sites of DNA damage and increased the rate of global sister chromatid exchanges [44]. In contrast, SMC5/6 was found to promote the for mation of APBs in ALT cells where it facilitates recombination between telomeres [49]. A tightly regulated process is therefore needed to prevent spurious recombination at telomeres. We saw evidence of ATRX mediated recruitment of SLF2 to telomeres (Fig 4B), and found that the loss of the proteins causes telomere exchanges in HEK293 Flp In cells (Fig 4D). SLF2 depletion was sufficient to induce high levels of telomere exchanges in U2OS cells (S4(E) Fig), showing that SLF1/2 is an important suppressor of recombination at telomeres.

Chromatin organization and maintenance by ATRX is needed to prevent gradual changes that enable the ALT phenotype [20]. To better understand how ATRX and SLF2 influence telo meres, we used RAP1 BioID to obtain information on changes caused by their loss. While few changes were observed upon loss of SLF2, intriguing changes were seen at telomeres approxi mately 15 passages after *ATRX* gene disruption (Figs 5C and 6). The changes seen in the single gene deletions were largely recapitulated in the double null cells. Changes in chromatin remodelling, DNA replication, and repair were particularly notable. This opens the door to

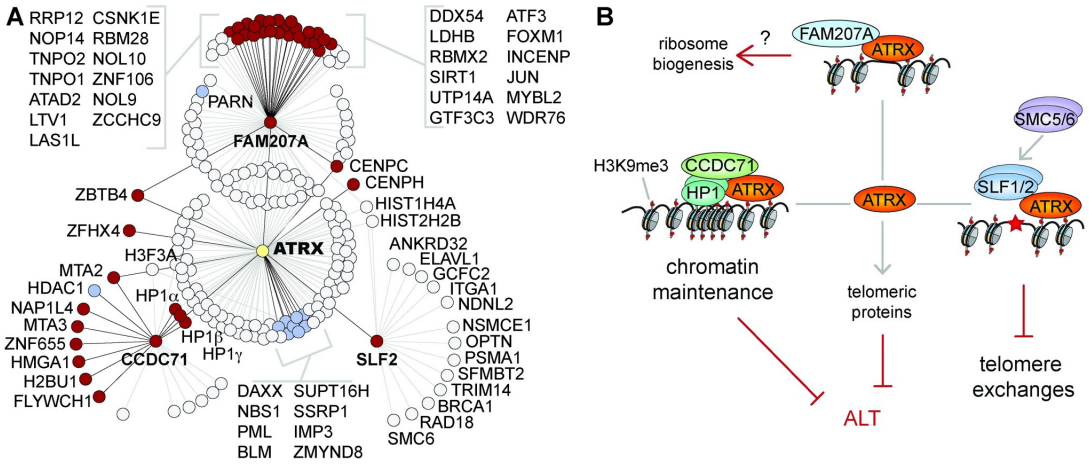


Fig 7. ATRX proteome overview and proposed model. A: Protein associations with ATRX, FAM207A, and CCDC71. Pale grey circles denote previously reported associations, while dark red circles represent newly identified proximal associations. Blue circles represent previously reported associations also identified by our BioID. Image was built using ProHits Viz [104] and the BioGrid database [52], and manually organized. B: Model proposing new biological roles for ATRX (other than histone deposition), mediated through proximal associations with FAM207A, CCDC71, and SLF2.

https://doi.org/10.1371/journal.pgen.1009909.g007

further explorations, especially those pertaining to ALT. For example, while a loss of ATRX does not in itself cause ALT, we saw increased levels of H2AX, PML, and the POLD3 subunit of DNA Polδ, as well as lower levels of BRCA2 in ATRX null cells (Fig 5C); trends seen in ALT cell models [13, 94, 95]. However, some of the strongest proteomic changes on telomeres caused by the loss of ATRX included a gain of a few zinc finger proteins and in the shugoshin 2 (SGO2) protein. While shugoshin proteins have centromeric functions, yeast Sgo2 is also needed for subtelomeric stability [96]. Contrastingly, the loss of ATRX also caused a notable depletion of core replisome proteins, such as MCM2 7 proteins and MCM10, from telomeres. Further deletion of SLF2 caused a number of changes that seemed additive (Fig 6), and this was sufficient to increase the number of telomere exchanges (Fig 4D). It is therefore likely that the changes on telomeric chromatin caused by the loss of ATRX potentiate the ALT phenotype through direct (e.g., loss of protein interactions) and indirect (e.g., downstream) effects.

Clearly ATRX mediates multiple functions well beyond histone deposition with DAXX. Our data show that ATRX associates with a wide range of chromatin regulators with very dif ferent roles, as reflected by new and uncharacterized ATRX associating proteins (Fig 7A). We propose that ATRX associating proteins confer additional biological roles to protect repetitive DNA regions, and further impact chromatin in a manner that transcends ATRX's immediate protein interactions (Fig 7B).

Materials and methods

Antibodies

The following primary antibodies were used: ATRX (Santa Cruz, sc 55584, PLA: 3 μg/mL, WB: 1:1000), ATRX (Santa Cruz, sc 15408, WB: 1:1000), ATRX (Thermo Fisher Scientific, PA5 21348, PLA: 1:1000 1:5000, WB: 1:1000), biotin (Cell Signaling Technology, 5597S, WB: 1:500), CCDC71 (Thermo Fisher Scientific, PA5 61543, PLA: 1:250), DAXX (Santa Cruz, sc 7152, PLA: 1:250, WB: 1:500), EZH2 (Sigma Aldrich, MABE362, WB: 1:1000), FLAG

(Sigma Aldrich, F1804, IP: 2 μg/mg lysate, PLA: 1:1000, WB: 1:1000), GAPDH (Santa Cruz, sc 25778, WB: 1:1000), H3.3 (Abcam, ab176840, WB: 1:5000), HDAC1 (Thermo Fisher Scientific, PA1 860, WB: 1:1000), HP1 (Cell Signaling Technology, 2616S, WB: 1:1000), IMP3 (Thermo Fisher Scientific, PA5 26897, IF: 1:500, PLA: 1:500, WB: 1:1000), LaminB1 (Proteintech, 66095 1 Ig, WB: 1:50000), c myc (Invitrogen, MA1 980, PLA: 1:2000), NBS1 (Novus Biologicals, NB100 143, PLA: 1:200, WB: 1:1000), SLF1 (Sigma Aldrich, SAB2701555, WB: 1:1000), SLF2 (Thermo Fisher Scientific, PA5 66091, PLA: 2 μg/mL), SMC5 (Bethyl Laboratories, A300 236A, IF: 1:1000, WB: 1:1000), beta tubulin (Developmental Studies Hybridoma Bank, E7, WB: 1:10000), ZBTB4 (Novus Biologicals, NBP1 76517, PLA: 1:10000), and ZFHX4 (Sigma Aldrich, HPA023837, PLA: 1:500, WB: 1:500). The following secondary antibodies were used: anti rabbit IgG HRP (Invitrogen, 0031458, WB: 1:10,000), anti mouse kappa light chain HRP (Abcam, AB99617, WB: 1:10,000), anti Rabbit IgG Alexa Fluor 594 (Jackson ImmunoResearch, 111 585 008, IF: 1:250), anti rabbit IgG ATTO 647N (Rockland, 611 156 122S, IF: 1:250), anti mouse IgG ATTO 488 (Rockland, 610 152 121S, IF: 1:250), and anti mouse IgG AlexaFluor 647 (BioLegend, 405322, IF: 1:500).

cDNA constructs

ATRX cDNA was amplified by PCR and the DNA band was excised and purified using the PureLink Gel Extraction Kit (Thermo Fisher Scientific, K210012) and introduced into pDONR223 using Gateway BP Clonase II Enzyme mix (Thermo Fisher Scientific, 8602289719), as well as a pFastBac vector (ThermoFisher Scientific) for expression in insect cells. pDONR223 plasmids containing CCDC71 and FAM207A were generated in the laboratory of Dr. Anne Claude Gingras. Plasmids were prepared using Stable Competent E. coli (NEB, C3040I). Clones were selected from LB agar plates (1% tryptone, 0.5% yeast extract, 170 mM NaCl, pH 7.5) containing 50 μg/mL spectinomycin. *ATRX* was subcloned into a pDEST vector containing an N or C terminal FLAG BirA* using Gateway LR Clonase II Enzyme mix (Thermo Fisher Scientific, 11791020). Transformed cells were selected on LB agar plates with 50 μg/mL ampicillin. The pDEST pcDNA5 RAP1 BirA* FLAG C term plasmid was prepared in the laboratory of Dr. Anne Claude Gingras. Lentiviral particles encoding myc tagged SLF2 were prepared in the laboratory of Dr. Grant Stewart.

Cell lines and cell culture

HEK293 Flp In T REx cells (Thermo Fisher Scientific) were cultured in DMEM (Corning, 10017 CV) supplemented with 10% FBS (Wisent, 080 150), 1X penicillin/streptomycin, 5 μg/mL blasticidin, and 75 μg/mL zeocin. HeLa S3 cells were grown in Joklik MEM (Caisson Labs), and MO3.13 (Cellutions Biosystems) in DMEM. Both were supplemented with bovine serum and penicillin/streptomycin. All cell lines were regularly tested for mycoplasma. Transfections were performed with lipofectamine 3000 (Thermo Fisher Scientific, L3000008) according to manufacturer's instructions. For HEK293 Flp In T REx cells, transfections were performed at a 9:1 ratio of pOG44 to pDEST plasmid. Transfected HEK293 Flp In cells (BirA*, NLS BirA*, ATRX BirA*, and RAP1 BirA*) were cultured in DMEM supplemented with 10% FBS, 1X penicillin/streptomycin, 5 μg/mL blasticidin, and 200 μg/mL hygromycin (for selection, or 100 μg/mL for maintenance). To generate ATRX KO lines, an sgRNA targeting exon 9 (5' AAATTCCGAGTTTCGAGCGA 3') was cloned into pSpCas9(BB) 2A Puro (pX459) [97], a gift from Dr. Feng Zhang, Addgene plasmid 48139. For the SLF2 KO lines, sgRNA guides targeting exon 5 (5' AGTTTCATCACTCGGTTCCT, GGCTTGGCACCTTCAAATTC 3') were cloned into pSpCas9n(BB) 2A GFP (pX461). Double ATRX/SLF2 KO cells were obtained by disrupting the *SLF2* gene in the *ATRX* null cells. Constructs were transfected

into HEK293 Flp In RAP1 BirA* cells as described above. To validate the gene disruptions, genomic DNA was isolated by suspending cell pellets in genomic DNA extraction buffer (20 mM Tris pH 7.6, 150 mM KCl, 1 mM EDTA, 20% glycerol, 0.5% NP 40) supplemented with 1 mM DTT, 25 μg/mL RNase A, and 20 U/mL proteinase K and incubated at 60˚C for 2 hr. DNA was extracted with phenol:chloroform and the disrupted genes were PCR amplified. PCR products were purified using PureLink PCR Purification Kit (Invitrogen, K310002) fol lowing the manufacturer's protocol and prepared for Sanger sequencing using the BigDye Ter minator v3.1 Cycle Sequencing Kit (Thermo Fisher, 4337457). Samples were sequenced on the Applied Biosystems SeqStudio Genetic Analyzer (Thermo Fisher). ATRX null cells were fur ther verified by western blotting. For SLF2, validation at the RNA level was done using Pow erUp SYBR Green Master Mix (Applied Biosystems, A25742) according to manufacturer's instructions, with the following primers: SLF2 RT F 5' AAACACTTTGTGCTACTCTGTGG 3'; SLF2 RT R 5' GTATCCTGGCGACCAAGTCTTTCA 3'; GAPDH F 5' CAATGACCCC TTCATTGACCTC 3'; and GAPDH R 5' GATCTCGCTCCTGGAAGATG 3'. Samples were analyzed with the QuantStudio 3 Real Time PCR System (Applied Biosystems, A28567). Pyri dostatin was purchased from Cayman Chemical (18013).

BioID

Protocol was adapted from previously published work [98]. All BioID runs were performed in biological duplicates as previously described [98], but with results compared to 6 replicates of untransfected HEK293 Flp In T REx, 14 replicates of cells expressing BirA* and, where indicated, also 14 replicates of cells expressing BirA* NLS. Less than 20 cell passages elapsed between the CRISPR Cas9 gene disruptions and the RAP1 BioID experiments. Cells were seeded at 70% confluency and induced with 1 μg/mL tetracycline (or doxycycline concentra tions yielding near equal protein expression across samples) for 25 hr and 50 μM biotin added during the last 8 hr. Cells were pelleted with at least 0.1 g per sample, and snap frozen. Upon processing, cells were lysed in modified RIPA buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1.5 mM $MgCl_2$, 1 mM EGTA, 0.1% SDS, 1% IGEPAL CA 630) with freshly added sodium deoxycholate (0.4%) and protease inhibitors (Sigma Aldrich P8340) at 400 μL/0.1 g cells and solubilized for 20 min, gently rotating at 4˚C (all the following 4˚C incubations were also done with gentle rotation). Samples were sonicated at 25% amplitude for 5 sec on, 3 sec off cycles, for 3 cycles, using a Qsonica sonicator with CL 18 probe. Benzonase (Millipore 70746) was added and incubated for 15 min, 4˚C (1.5 μL or 375 units per sample). Samples were spiked with additional SDS to a final concentration of 0.4% and incubated 15 min, 4˚C. Samples were spun 16,000 x *g* for 20 min and the cleared lysates (supernatant) was transferred to a new tube. Streptavidin sepharose beads (GE 17 5113 01) were washed 3X with modified RIPA buffer with 0.4% SDS, and 30 μl (bed volume) was added to each sample and incubated 3 hr, 4˚C. Samples were washed once with wash buffer (2% SDS, 50 mM Tris pH 7.5), 2X with modified RIPA buffer with 0.4% SDS, and 3X with ABC buffer (50 mM ammonium bicarbonate pH 8.5). Samples were spun, supernatant removed, and on bead trypsin digest of peptides was performed by incubating with 1 μg trypsin dissolved in ABC buffer, rotating overnight at 37˚C. An additional 0.5 μg trypsin per sample was added the next day and sam ples incubated for 2 hr at 37˚C. Samples were gently vortexed, spun down, and supernatant transferred to new tube, with additional washing of beads and collection of supernatant per formed twice (30 μL of HPLC grade water). Fresh 50% formic acid was added to samples to a final concentration of 2% prior to drying by vacuum centrifugation and subsequent storage at 80˚C.

Mass spectrometry acquisition

Each sample (6 μL in 2% formic acid; corresponding to 1/6th of a 15 cm tissue culture dish) was directly loaded at 800 nL/min onto an equilibrated HPLC column (pulled and packed in house). The peptides were eluted from the column over a 90 min gradient generated by a Eksi gent ekspert nanoLC 425 (Eksigent, Dublin CA) nano pump and analysed on a TripleTOF 6600 instrument (AB SCIEX, Concord, Ontario, Canada). The gradient was delivered at 400 nL/min starting from 2% acetonitrile with 0.1% formic acid to 35% acetonitrile with 0.1% for mic acid over 90 min followed by a 15 min clean up at 80% acetonitrile with 0.1% formic acid, and a 15 min equilibration period back to 2% acetonitrile with 0.1% formic acid, for a total of 120 min. To minimize carryover between each sample, the analytical column was washed for 2 hr by running an alternating sawtooth gradient from 35% acetonitrile with 0.1% formic acid to 80% acetonitrile with 0.1% formic acid at a flow rate of 1500 nL/min, holding each gradient concentration for 5 min. Analytical column and instrument performance was verified after each sample by loading 30 fmol bovine serum albumin (BSA) tryptic peptide standard with 60 fmol alpha casein tryptic digest and running a short 30 min gradient. TOF MS mass calibra tion was performed on BSA reference ions before running the next sample to adjust for mass drift and verify peak intensity. Samples were analyzed with two separate injections with instru ment method set to data dependent acquisition (DDA) mode. The DDA method consisted of one 250 milliseconds (ms) MS1 TOF survey scan from 400 1800 Da followed by ten 100 ms MS2 candidate ion scans from 100 1800 Da in high sensitivity mode. Only ions with a charge of 2+ to 5+ that exceeded a threshold of 300 cps were selected for MS2, and former precursors were excluded for 7 seconds after one occurrence.

Data-dependent acquisition data search

Mass spectrometry data generated were stored, searched, and analyzed using ProHits labora tory information management system (LIMS) platform [99]. Within ProHits, WIFF files were converted to an MGF format using the WIFF2MGF converter and to an mzML format using ProteoWizard (V3.0.10702) and the AB SCIEX MS Data Converter (V1.3 beta). The data were then searched using Mascot (V2.3.02) [100] and Comet (V2016.01 rev.2) [101]. The spectra were searched with the human and adenovirus sequences in the RefSeq database (version 57, January 30th, 2013) acquired from NCBI, supplemented with "common contaminants" from the Max Planck Institute (http://maxquant.org/contaminants.zip) and the Global Proteome Machine (GPM; ftp://ftp.thegpm.org/fasta/cRAP/crap.fasta), forward and reverse sequences (labeled "gi|9999" or "DECOY"), sequence tags (BirA, GST26, mCherry and GFP) and strepta vidin, for a total of 72,481 entries. Database parameters were set to search for tryptic cleavages, allowing up to 2 missed cleavages sites per peptide with a mass tolerance of 35 ppm for precur sors with charges of 2+ to 4+ and a tolerance of 0.15 amu for fragment ions. Variable modifica tions were selected for deamidated asparagine and glutamine and oxidized methionine. Results from each search engine were analyzed through TPP (the Trans Proteomic Pipeline, v.4.7 POLAR VORTEX rev 1) via the iProphet pipeline [102].

SAINT analysis

SAINTexpress version 3.6.1 [103] was used as a statistical tool to calculate the probability of potential protein protein associations compared to background contaminants using default parameters, with bait compression set to 2 and control compression set to 4. A 95% FDR iPro phet filter was used. SAINT scores with a Bayesian false discovery rate (BFDR) $\leq$ 1% were con sidered high confidence protein interactions. All non human protein interactors (did not start with "NP" in Prey column) were removed from the SAINT analysis, except for BirA R118G

H0QFJ5. Dot plots were generated using the "Dot plot generator" tool in ProHits viz [104] using SAINTexpress file generated from ProHits. Data were normalized by total abundance. Gene ontology term analysis for biological process (GO:BP) was done using g:Profiler [105].

Proximity ligation assay

Cells were seeded to a density of 50 70% on 12 mm glass coverslips coated in poly L lysine (Sigma Aldrich, P8920). Cells were fixed with a 2% PFA solution (2% PFA, 0.2% Triton X 100, pH 8.2), washed with 1X PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4), permeabilized with 0.5% tergitol (Sigma Aldrich, NP40S), and washed with 1X PBS. PLAs were performed according to manufacturer's instructions using the Duolink *In Situ* Detection Kit (Sigma Aldrich, DUO92013). Confocal microscopy was done at the SickKids Imaging Facility and foci counted using CellProfiler [106]. Signal intensity was obtained using ImageJ, with background signal subtracted for each figure [107].

Cell lysis & subcellular fractionation

All steps below were performed at 4˚C. Subcellular fractionation was done by gently dounce homogenizing PBS washed cell pellets in 2 pellet volumes of cytosolic lysis buffer (20 mM Tris pH 7.6, 10 mM NaCl, 1 mM EDTA, 0.5% IGEPAL, 0.5 mM DTT, protease inhibitors) using a loose pestle. Nuclei were recovered by centrifugation at 500 x *g* for 5 min, and the supernatant containing the cytoplasmic fraction was transferred to a new tube. Nuclei were then resus pended in 3 pellet volumes of nuclear extraction buffer (20 mM Tris pH 7.6, 300 mM NaCl, 1 mM EDTA, 20% glycerol, 0.5 mM DTT) with protease inhibitors and dounce homogenized with a tight pestle. The mixture was spun down at 30,000 x *g* for 15 min, and the supernatant containing nuclear extracts was transferred to a fresh tube. The remaining insoluble pellet was resuspended in 3 volumes of BC100 (20 mM Tris pH 7.6, 100 mM NaCl, 0.2 mM EDTA, 20% glycerol, 0.5 mM DTT) with protease inhibitors. The sample was then sonicated and digested with 75 U/μL MNase in the presence of 5 mM $CaCl_2$ for 2 hr, with gentle rotation. The subcel lular fractions were dialyzed against BC100 and insoluble material removed by centrifugation as per conditions above. Whole cell extracts were generated by incubating PBS washed cell pel lets with 3 pellet volumes of lysis buffer (50 mM Tris pH 7.6, 100 mM KCl, 2 mM EDTA, 0.1% NP 40, 10% glycerol, 0.5 mM DTT). Samples were then sonicated, digested and processed as per the solubilized nuclear fraction above.

Immunoprecipitation

Cells were seeded at 4 x 10^6 cells per plate and induced with 1 μg/mL doxycycline for 24 hr. The cells were harvested by first washing with cold 1X PBS, scraping, and collecting the cells in 1 mL 1X PBS. The cells were spun down at 500 x *g* for 5 min at 4˚C, and lysed in 1X IP lysis buffer (50 mM Tris pH 7.6, 100 mM KCl, 2 mM EDTA, 0.1% IGEPAL, 10% glycerol). The lysate was flash frozen on dry ice for 10 min and then allowed to thaw at room temperature. The lysate was then sonicated 3X for 10 sec at 1.5 amp. After sonication, the lysate was digested with 1 U/μL MNase and the buffer supplemented with 5 mM $CalCl_2$. This was left to incubate overnight with rocking at 4˚C. Four milligrams of protein was incubated with anti FLAG (1 μg/mg lysate) overnight at 4˚C, rocking. The following day, magnetic Protein G beads (Cytiva, 28967070) were washed three times in 1X IP lysis buffer and the lysate and antibody mix was added onto the beads and incubated at 4˚C for 4 hr with rocking. The supernatant was removed, and the beads were washed three times in 1X IP lysis buffer. Proteins were eluted with 40 μL 1X Laemmli (2% SDS, 0.1% bromophenol blue, 10% glycerol, 62.5 mM Tris pH 6.8, 100 mM DTT) and boiled at 95˚C for 15 min.

In vitro transcription/translation & pulldown assays

FLAG ATRX was expressed from baculovirus infected SF9 insect cells as per previous descrip tions [108]. The protein was immobilized on an M2 resin (MilliporeSigma), washed, and eluted with 25 μg/ml 3X FLAG peptide (MilliporeSigma). The TnT Quick Coupled Transcrip tion/Translation System (Promega, L1170) was used to generate FAM207A in accordance with manufacturer's instructions, with the addition of Transcend tRNA (Promega L506A) and 1 μg/mL leupeptin. Recombinant FAM207A was incubated with Strep Tactin MacroPrep resin (IBA LifeSciences 2 1505 010) in BC150 (20 mM Tris pH 7.6, 150 mM NaCl, 0.2 mM EDTA, 20% glycerol) for 30 min at 4˚C prior to the addition of 20 pmol of recombinant ATRX. Proteins were incubated together at room temperature for 1 hr before washing with BC150 + 0.1% NP40, BC300 (20 mM Tris pH 7.6, 300 mM NaCl, 0.2 mM EDTA, 20% glycerol) + 0.1% NP40, and BC150. Samples were spiked with 1X Laemmli buffer and boiled at 95˚C for 10 min prior to Western blotting.

Western blotting

Protein lysates were diluted in 1X Laemmli buffer and boiled at 95˚C for 5 min. For streptavi din pull downs, 2 mM biotin was added to 1X Laemmli prior to boiling. Protein lysates were loaded onto 8 12% Tris Glycine gels and run at 100 V for 90 min in 1X running buffer (25 mM Tris, 192 mM glycine, 0.1% SDS). Proteins were then transferred onto activated PVDF membranes for 2 hr at 270 mA in 1X transfer buffer (200 mM glycine, 25 mM Tris) using a Mini protean wet transfer system (BioRad). All samples were run this way, with the exception of ZFHX4 fractionations, which were loaded onto a 3 8% Tris Acetate gel (Thermo Fisher, EA0375PK2), run at 125V for 1.5 hr in 1X running buffer (Invitrogen, LA0041) supplemented with 500 μL antioxidant (Thermo Fisher, NP0005). Gels were washed in 20% ethanol for 10 min prior to transfer to PVDF with the iBlot (Thermo, IB21001) at 20V for 2 min, 23V for 6 min, and 25V for 4 min. All membranes were blocked in 5% milk in 1X TBST (20 mM Tris, 150 mM NaCl, 0.1% Tween) and incubated with primary antibody diluted in 1X TBST and 0.04% NaN_3 for 1 hr at room temperature, or overnight at 4˚C with gentle rotation. Mem branes were then washed 3X for 5 min with 1X TBST and incubated with HRP conjugated sec ondary antibodies diluted in 1X TBST with 5% milk at room temperature, rocking for 1 hr. Membranes were then washed 3X 5 min with 1X TBST. To image, a 5X luminol solution (100 mM Tris pH 8.8, 1.25 mM luminol, 0.2 mM coumaric acid) was diluted as needed and supple mented with 5 μL 10% H_2O_2 per mL, and poured over the blot. For westerns of the TnT reac tions, samples were treated as described above, with the exception of the luminol solution. Instead, blots were washed treated with streptavidin AP (Promega V5591) according to manu facturer's instructions, rocking at room temperature for 1 hr. Blots were washed twice with 1X TBST and then twice with water. Blots were then incubated in Western Blue Stabilized Sub strate (Promega S384C) according to manufacturer's instructions. Images were captured using a BioRad ChemiDoc XRS+ system.

CO-FISH

CO FISH was performed as previously described [109] with the following modifications. Cells were treated with 10 μM BrdU for 18 hr and 0.2 μg/mL KaryoMAX Colcemid (Thermo Fisher, 15212012) for the final 4 hr. Metaphase spreads were prepared using a CDS 5 Cytogenetic Dry ing chamber (Thermotron) according to standard methods, rehydrated with 1X PBS, and slides were treated with 0.5 mg/mL RNase A (Life Technologies, 12091021) at 37˚C for 10 min. Slides were stained with 0.5 μg/mL bisbenzimide Hoechst 33258 (Sigma Aldrich, 14530) in 2X SSC (300 mM NaCl, 30 mM sodium citrate) for 15 min at room temperature. Slides

were exposed to 365 nm UV, 5.4 x 10^3 J/m2 twice. Next, slides were treated with 10 U/μL exo nuclease III (NEB, M0206L) for 40 min at 37˚C. Cells were washed in 1X PBS and then dehy drated in a series of 70, 95, and 100% ethanol. TelG Cy3 (PNA Bio, F1006) and TelC ALEXA 488 (PNA Bio, F1004) probes were heated at 70˚C for 30 min, diluted in hybridization buffer [20 mM Tris pH 7.2, 60% formamide (Thermo Fisher, 15515026), 0.5% block (Roche, 11096176001)], and applied to slides for 2 hr each at room temperature, sequentially. Slides were washed with Wash Buffer 1 (10 mM Tris pH 7.2, 70% formamide, 0,1% BSA) and Wash Buffer 2 (100 mM Tris pH 7.2, 150 mM NaCl, 0.08% Tween 20) three times each. DAPI (Sigma Aldrich, D9542) was added at final concentration of 1 μg/mL to the second wash with Wash Buffer 2. Then, slides were dehydrated in an ethanol series, as before, and mounted with ProLong Gold Antifade Mountant (Invitrogen, P36930). Slides were imaged with an Olympus BX61 microscope with the CytoPower automated imaging system (Applied Spectral Imaging). Telomere exchanges were manually scored in a blinded fashion, with sample identity only revealed after all samples were scored.

Immunofluorescence

Cells were seeded on coverslips and fixed in a 2% PFA solution for 20 min at room tempera ture after the desired treatment. Cells were washed 3X in 1X PBS and permeabilised with 0.5% tergitol in 1X PBS for 10 min at room temperature. Coverslips were washed 3X in 1X PBS and incubated in blocking buffer (3% BSA, 1% NGS, 1x PBS) for 1 hr at room temperature. Cells were then incubated in primary antibody diluted in blocking buffer for 1.5 hr at room temper ature in a humidity chamber, washed 3X in 1X PBS, and incubated in secondary antibody diluted in blocking buffer for 30 min at room temperature. Coverslips were washed 3X with 1X PBS and treated with 1 μg/mL DAPI diluted in 1X PBS for 10 min at room temperature. Coverslips were washed once more in 1X PBS, mounted onto slides and imaged as described for the PLAs.

IF-FISH

Cells were seeded on slides, fixed in a 2% PFA solution for 10 min at room temperature, washed 3X in 1X PBS, and treated with 0.5% tergitol in 1X PBS for 10 min at room tempera ture. Cells were washed 3X in 1X PBS and incubated in blocking buffer for 1 hr at room tem perature. Cells were then incubated in primary antibody diluted in blocking buffer for 1.5 hr at room temperature in a humidity chamber, washed 3X in 1X PBS, and incubated in secondary antibody diluted in blocking buffer for 30 min at room temperature. Cells were washed 2X with 1X PBS and fixed in 2% PFA in 1X PBS for 10 min at room temperature. Slides were dehydrated in an ethanol series of 70, 95, and 100% for 5 min each. Slides were air dried and TelC 488 probe (PNA Bio, F1004) was heated at 72˚C for 30 min before being diluted 1:500 in hybridization buffer [70% formamide, 0.5% blocking reagent (Roche, 11096176001), 10 mM Tris pH 7.2] and being applied to slides. Slides were heated at 72˚C for 10 min, sealed with rub ber cement, and placed in a humidity chamber to allow probe binding overnight. The next day, slides were washed 2X, 15 min each at room temperature in wash buffer 1 (70% formam ide, 10 mM Tris pH 7.2). Slides were then washed 3X, 5 min each at room temperature in wash buffer 2 (100 mM Tris pH 7.2, 150 mM NaCl, 0.08% Tween 20), with 1 μg/mL of DAPI being added to the second wash. Slides were dehydrated in an ethanol series as before, allowed to air dry, and mounted before being imaged. Colocalization was scored using the Cellprofiler Colocalization pipeline [106] scoring foci that were 2 15 pixels in diameter and greater than 0.15 units of intensity.

Supporting information

S1 Fig. ATRX BioID system and proximal associations. A: Western blot exemplifying stable integration and expression of FLAG tagged BirA* protein fusions in inducible HEK293 Flp In T REx cells (top), and similar expression levels of N and C terminally tagged ATRX (bottom). B: Western blot (top) and Coomassie staining (bottom) of biotinylated proteins captured on streptavidin beads. C: Example of immunofluorescent labeling demonstrating nuclear targeting of the ATRX BirA* fusion constructs. D: Gene ontology analysis of ATRX associating proteins. E: Western blot showing the subcellular distribution of ATRX associating proteins in HeLa S3 cells. F: Examples of the proximity ligation assay (PLA) showing associations between endogenous proteins in MO3.13 cells. G H: PLAs showing associations between endogenous ATRX and exogenous FLAG tagged CCDC71 (G) or FAM207A (H) in HEK293 Flp In T REx cells. I: PLAs showing an association between endogenous ATRX and myc tagged SLF2. Data plots account for ∼100 nuclei in three independent experiments, with the total number of nuclei assessed in brackets. The p values were obtained using a 2 sided Student's t test with unequal variance. J: *In vitro* interaction between recombinant FAM207A and ATRX.
(PDF)

S2 Fig. FAM207A subcellular expression and BioID system. A: Subcellular fractionation of HEK293 Flp In T REx cells expressing FLAG FAM207A. Cy cytoplasm; Nu nucleus; Ch chromatin. B: Western analysis of BirA* constructs used for the FAM207A BioID experiment. Duplicate (induced) lanes are shown. C: Immunolabeling experiment showing co localization of FAM207A and the IMP3 nucleolar protein. Scale bar = 4μm. D: Dot plot showing prey proteins identified with FAM207A BirA* that were enriched over endogenous biotinylation (untransfected), unspecific pan cellular biotinylation (BirA*), and unspecific nuclear biotinylation (NLS BirA*; BFDR ≤ 5%, SAINT [103]). Data represent two biological replicates. Proteins in boldface remained statistically enriched when the nuclear localization signal (NLS) BirA* control was used to further filter the FAM207A BioID data. E: Full western blots for Fig 2A. The red boxes indicate the areas shown in the main figure.
(PDF)

S3 Fig. PxVxL motif in CCDC71L, CCDC71 subcellular fractionation and immunoprecipitation, and BioID system. A: PxVxL motif in human CCDC71L. B: Subcellular fractionation of HEK293 Flp In T REx cells expressing FLAG CCDC71. Cy cytoplasm; Nu nucleus; Ch chromatin. C: CCDC71 immunoprecipitation and verification of associations reported in BioGrid [52]. D: Western analysis of CCDC71 and control BirA* constructs. E: Dot plot showing prey proteins identified with CCDC71 BirA* that were enriched over endogenous biotinylation (untransfected) and unspecific pan cellular biotinylation (BirA*; BFDR ≤ 5%, SAINT [103]). Data represent two biological replicates. Proteins in boldface remained statistically enriched when the nuclear localization signal (NLS) BirA* control was used to further filter the CCDC71 BioID data. F: Full western blots for Fig 3D. The red boxes indicate the areas shown in the main figure.
(PDF)

S4 Fig. Co localization of SLF2 and telomeres, SLF2 null HEK293 Flp In T REx cells, and telomere exchanges in U2OS cells. A: Western blot showing exogenous myc SLF2 expression in lentivirus infected HEK293 Flp In T REx (with inducible RAP1 BirA*, but the latter is not induced). B: Immunolabeling of myc SLF2 (yellow) and SMC5 (red), and fluorescence *in situ* hybridization (IF FISH) using a telomeric (green) probe in ATRX expressing and null HEK293 Flp In cells. C: CRISPR Cas9 mediated *SLF2* gene disruptions in HEK293 Flp In

T REx (RAP1 BirA*), detected by DNA sequencing, and matching qRT PCR (D). E: Telomere exchange rates observed in U2OS cells. At least 30 mitotic spreads per condition were counted and the percent of telomeric exchanges plotted. F: Full western blots for Fig 4A. The red boxes indicate the areas shown in the main figure.
(PDF)

S5 Fig. RAP1 BioID and SLF2 null cells. A: Western blot showing the induction of FLAG and BirA* tagged RAP1 in ATRX expressing and null HEK293 Flp In T REx cells. B: West ern analysis of shRNA mediated ATRX knockdown in HEK293 Flp In T REx cells expressing RAP1 BirA*. Doubly transfected cells (shRNA constructs 1 and 2) were used for the RAP1 BioID. C: Overlap between RAP1 BioID and TERF1 BioID [65], dCas9 APEX2 biotinylation at genomic elements by restricted spatial tagging (C BERST) [68], quantitative telomeric chro matin isolation protocol (QTIP) [67], and proteomics of isolated chromatin segments (PICh) [66]. PICh and QTIP data were grouped to simplify the Venn diagrams. Data consider HEK293 Flp In cells expressing ATRX (unperturbed). A list of proteins commonly identified in at least three different telomere proteomic screens is shown on the right. D: Fold change of prey proteins identified by RAP1 BioID (BFDR $\leq$ 5%, SAINT [103]) in ATRX KD vs. ATRX expressing HEK293 Flp In cells. Labeled proteins had $\geq$2.5 average spectra and $\geq$1.5 fold change between conditions. Proteins labeled in the scatter plot remained significant when NLS BirA* was used as a control. Inset box represents proteins also identified with NLS BirA* as a bait (less stringent).
(PDF)

S1 Data. BioID analysis.
(XLSX)

S2 Data. Data for plots other than BioID.
(XLSX)

Acknowledgments

Proteomics was performed at the Network Biology Collaborative Centre at the Lunenfeld Tanenbaum Research Institute. We would like to thank Steven Erwood for help with the gene editing, as well as Alejandro Saettone and Aparna Srinivasan for their help with LaTeX.

Author Contributions

Conceptualization: William A. Scott, Christopher E. Pearson, Grant S. Stewart, Anne Claude Gingras, Eric I. Campos.

Formal analysis: William A. Scott, Erum Z. Dhanji, Boris J. A. Dyakov, Anne Claude Gingras, Eric I. Campos.

Investigation: William A. Scott, Erum Z. Dhanji, Boris J. A. Dyakov, Ema S. Dreseris, Jona thon S. Asa, Laura J. Grange, Mila Mirceta, Grant S. Stewart, Eric I. Campos.

Supervision: Anne Claude Gingras, Eric I. Campos.

Writing original draft: William A. Scott, Eric I. Campos.

References

1. Wong LH, McGhie JD, Sim M, Anderson MA, Ahn S, Hannan RD, George AJ, Morgan KA, Mann JR, Choo KH. ATRX interacts with H3.3 in maintaining telomere structural integrity in pluripotent

embryonic stem cells. Genome Res. 2010 Mar; 20(3):351 60. https://doi.org/10.1101/gr.101477.109 PMID: 20110566

2. Emelyanov AV, Konev AY, Vershilova E, Fyodorov DV. Protein complex of Drosophila ATRX/XNP and HP1a is required for the formation of pericentric beta heterochromatin in vivo. J Biol Chem. 2010 May 14; 285(20):15027 15037. https://doi.org/10.1074/jbc.M109.064790 PMID: 20154359
3. Voon HP, Hughes JR, Rode C, De La Rosa Velázquez IA, Jenuwein T, Feil R, Higgs DR, Gibbons RJ. ATRX Plays a Key Role in Maintaining Silencing at Interstitial Heterochromatic Loci and Imprinted Genes. Cell Rep. 2015 Apr 21; 11(3):405 18. https://doi.org/10.1016/j.celrep.2015.03.036 PMID: 25865896
4. Juhász S, Elbakry A, Mathes A, Löbrich M. ATRX Promotes DNA Repair Synthesis and Sister Chro matid Exchange during Homologous Recombination. Mol Cell. 2018 Jul 5; 71(1):11 24.e7. https://doi.org/10.1016/j.molcel.2018.05.014 PMID: 29937341
5. Watson LA, Solomon LA, Li JR, Jiang Y, Edwards M, Shin ya K, Beier F, Bérubé NG. Atrx deficiency induces telomere dysfunction, endocrine defects, and reduced life span. J Clin Invest. 2013 May; 123(5):2049 63. https://doi.org/10.1172/JCI65634 PMID: 23563309
6. Gibbons RJ, Picketts DJ, Villard L, Higgs DR. Mutations in a putative global transcriptional regulator cause X linked mental retardation with alpha thalassemia (ATR X syndrome). Cell. 1995 Mar 24; 80(6):837 45. https://doi.org/10.1016/0092 8674(95)90287 2 PMID: 7697714
7. Heaphy CM, de Wilde RF, Jiao Y, Klein AP, Edil BH, Shi C, Bettegowda C, Rodriguez FJ, Eberhart CG, Hebbar S, Offerhaus GJ, McLendon R, Rasheed BA, He Y, Yan H, Bigner DD, Oba Shinjo SM, Marie SK, Riggins GJ, Kinzler KW, Vogelstein B, Hruban RH, Maitra A, Papadopoulos N, Meeker AK. Altered telomeres in tumors with ATRX and DAXX mutations. Science. 2011 Jul 22; 333(6041):425. https://doi.org/10.1126/science.1207313 PMID: 21719641
8. Lovejoy CA, Li W, Reisenweber S, Thongthip S, Bruno J, de Lange T, De S, Petrini JH, Sung PA, Jasin M, Rosenbluh J, Zwang Y, Weir BA, Hatton C, Ivanova E, Macconaill L, Hanna M, Hahn WC, Lue NF, Reddel RR, Jiao Y, Kinzler K, Vogelstein B, Papadopoulos N, Meeker AK; ALT Starr Cancer Consortium. Loss of ATRX, genome instability, and an altered DNA damage response are hallmarks of the alternative lengthening of telomeres pathway. PLoS Genet. 2012; 8(7):e1002772. https://doi.org/10.1371/journal.pgen.1002772 PMID: 22829774
9. Clynes D, Jelinska C, Xella B, Ayyub H, Scott C, Mitson M, Taylor S, Higgs DR, Gibbons RJ. Suppres sion of the alternative lengthening of telomere pathway by the chromatin remodelling factor ATRX. Nat Commun. 2015 Jul 6; 6:7538. https://doi.org/10.1038/ncomms8538 PMID: 26143912
10. Heaphy CM, Subhawong AP, Hong SM, Goggins MG, Montgomery EA, Gabrielson E, Netto GJ, Epstein JI, Lotan TL, Westra WH, Shih IeM, Iacobuzio Donahue CA, Maitra A, Li QK, Eberhart CG, Taube JM, Rakheja D, Kurman RJ, Wu TC, Roden RB, Argani P, De Marzo AM, Terracciano L, Tor benson M, Meeker AK. Prevalence of the alternative lengthening of telomeres telomere maintenance mechanism in human cancer subtypes. Am J Pathol. 2011 Oct; 179(4):1608 15. https://doi.org/10.1016/j.ajpath.2011.06.018 PMID: 21888887
11. Bryan TM, Englezou A, Gupta J, Bacchetti S, Reddel RR. Telomere elongation in immortal human cells without detectable telomerase activity. EMBO J. 1995 Sep 1; 14(17):4240 8. https://doi.org/10.1002/j.1460 2075.1995.tb00098.x PMID: 7556065
12. Episkopou H, Draskovic I, Van Beneden A, Tilman G, Mattiussi M, Gobin M, Arnoult N, Londoño Val lejo A, Decottignies A. Alternative Lengthening of Telomeres is characterized by reduced compaction of telomeric chromatin. Nucleic Acids Res. 2014 Apr; 42(7):4391 405. https://doi.org/10.1093/nar/gku114 PMID: 24500201
13. Yeager TR, Neumann AA, Englezou A, Huschtscha LI, Noble JR, Reddel RR. Telomerase negative immortalized human cells contain a novel type of promyelocytic leukemia (PML) body. Cancer Res. 1999 Sep 1; 59(17):4175 9. PMID: 10485449
14. Londoño Vallejo JA, Der Sarkissian H, Cazes L, Bacchetti S, Reddel RR. Alternative lengthening of telomeres is characterized by high rates of telomeric exchange. Cancer Res. 2004 Apr 1; 64(7):2324 7. https://doi.org/10.1158/0008 5472.CAN 03 4035 PMID: 15059879
15. Lovejoy CA, Takai K, Huh MS, Picketts DJ, de Lange T. ATRX affects the repair of telomeric DSBs by promoting cohesion and a DAXX dependent activity. PLoS Biol. 2020 Jan 2; 18(1):e3000594. https://doi.org/10.1371/journal.pbio.3000594 PMID: 31895940
16. Sobinoff AP, Allen JA, Neumann AA, Yang SF, Walsh ME, Henson JD, Reddel RR, Pickett HA. BLM and SLX4 play opposing roles in recombination dependent replication at human telomeres. EMBO J. 2017 Oct 2; 36(19):2907 2919. https://doi.org/10.15252/embj.201796889 PMID: 28877996
17. Cesare AJ, Griffith JD. Telomeric DNA in ALT cells is characterized by free telomeric circles and het erogeneous t loops. Mol Cell Biol. 2004 Nov; 24(22):9948 57. https://doi.org/10.1128/MCB.24.22.9948 9957.2004 PMID: 15509797

18. Sobinoff AP, Pickett HA. Mechanisms that drive telomere maintenance and recombination in human cancers. Curr Opin Genet Dev. 2020 Feb; 60:25 30. https://doi.org/10.1016/j.gde.2020.02.006 PMID: 32119936

19. Kent T, Gracias D, Shepherd S, Clynes D. Alternative Lengthening of Telomeres in Pediatric Cancer: Mechanisms to Therapies. Front Oncol. 2020 Jan 21; 9:1518. https://doi.org/10.3389/fonc.2019.01518 PMID: 32039009

20. Li F, Deng Z, Zhang L, Wu C, Jin Y, Hwang I, Vladimirova O, Xu L, Yang L, Lu B, Dheekollu J, Li JY, Feng H, Hu J, Vakoc CR, Ying H, Paik J, Lieberman PM, Zheng H. ATRX loss induces telomere dys function and necessitates induction of alternative lengthening of telomeres during human cell immor talization. EMBO J. 2019 Oct 1; 38(19):e96659. https://doi.org/10.15252/embj.201796659 PMID: 31454099

21. Napier CE, Huschtscha LI, Harvey A, Bower K, Noble JR, Hendrickson EA, Reddel RR. ATRX represses alternative lengthening of telomeres. Oncotarget. 2015 Jun 30; 6(18):16543 58. https://doi.org/10.18632/oncotarget.3846 PMID: 26001292

22. Cesare AJ, Kaul Z, Cohen SB, Napier CE, Pickett HA, Neumann AA, Reddel RR. Spontaneous occur rence of telomeric DNA damage response in the absence of chromosome fusions. Nat Struct Mol Biol. 2009 Dec; 16(12):1244 51. https://doi.org/10.1038/nsmb.1725 PMID: 19935685

23. Zhang JM, Genois MM, Ouyang J, Lan L, Zou L. Alternative lengthening of telomeres is a self perpetu ating process in ALT associated PML bodies. Mol Cell. 2021 Mar 4; 81(5):1027 1042.e4. https://doi.org/10.1016/j.molcel.2020.12.030 PMID: 33453166

24. Hoang SM, O'Sullivan RJ. Alternative Lengthening of Telomeres: Building Bridges To Connect Chro mosome Ends. Trends Cancer. 2020 Mar; 6(3):247 260. https://doi.org/10.1016/j.trecan.2019.12.009 PMID: 32101727

25. Wu G, Jiang X, Lee WH, Chen PL. Assembly of functional ALT associated promyelocytic leukemia bodies requires Nijmegen Breakage Syndrome 1. Cancer Res. 2003 May 15; 63(10):2589 95. PMID: 12750284

26. Jiang WQ, Zhong ZH, Henson JD, Neumann AA, Chang AC, Reddel RR. Suppression of alternative lengthening of telomeres by Sp100 mediated sequestration of the MRE11/RAD50/NBS1 complex. Mol Cell Biol. 2005 Apr; 25(7):2708 21. https://doi.org/10.1128/MCB.25.7.2708 2721.2005 PMID: 15767676

27. Wu G, Lee WH, Chen PL. NBS1 and TRF1 colocalize at promyelocytic leukemia bodies during late S/G2 phases in immortalized telomerase negative cells. Implication of NBS1 in alternative lengthening of telomeres. J Biol Chem. 2000 Sep 29; 275(39):30618 22. https://doi.org/10.1074/jbc.C000390200 PMID: 10913111

28. Syed A, Tainer JA. The MRE11 RAD50 NBS1 Complex Conducts the Orchestration of Damage Sig naling and Outcomes to Stress in DNA Replication and Repair. Annu Rev Biochem. 2018 Jun 20; 87:263 294. https://doi.org/10.1146/annurev biochem 062917 012415 PMID: 29709199

29. Zhong ZH, Jiang WQ, Cesare AJ, Neumann AA, Wadhwa R, Reddel RR. Disruption of telomere main tenance by depletion of the MRE11/RAD50/NBS1 complex in cells that use alternative lengthening of telomeres. J Biol Chem. 2007 Oct 5; 282(40):29314 22. https://doi.org/10.1074/jbc.M701413200 PMID: 17693401

30. Leung JW, Ghosal G, Wang W, Shen X, Wang J, Li L, Chen J. Alpha thalassemia/mental retardation syndrome X linked gene product ATRX is required for proper replication restart and cellular resistance to replication stress. J Biol Chem. 2013 Mar 1; 288(9):6342 50. https://doi.org/10.1074/jbc.M112.411603 PMID: 23329831

31. Clynes D, Jelinska C, Xella B, Ayyub H, Taylor S, Mitson M, Bachrati CZ, Higgs DR, Gibbons RJ. ATRX dysfunction induces replication defects in primary mouse cells. PLoS One. 2014 Mar 20; 9(3):e92915. https://doi.org/10.1371/journal.pone.0092915 PMID: 24651726

32. Iwase S, Xiang B, Ghosh S, Ren T, Lewis PW, Cochrane JC, Allis CD, Picketts DJ, Patel DJ, Li H, Shi Y. ATRX ADD domain links an atypical histone methylation recognition mechanism to human mental retardation syndrome. Nat Struct Mol Biol. 2011 Jun 12; 18(7):769 76. https://doi.org/10.1038/nsmb.2062 PMID: 21666679

33. Dhayalan A, Tamas R, Bock I, Tattermusch A, Dimitrova E, Kudithipudi S, Ragozin S, Jeltsch A. The ATRX ADD domain binds to H3 tail peptides and reads the combined methylation state of K4 and K9. Hum Mol Genet. 2011 Jun 1; 20(11):2195 203. https://doi.org/10.1093/hmg/ddr107 PMID: 21421568

34. Eustermann S, Yang JC, Law MJ, Amos R, Chapman LM, Jelinska C, Garrick D, Clynes D, Gibbons RJ, Rhodes D, Higgs DR, Neuhaus D. Combinatorial readout of histone H3 modifications specifies localization of ATRX to heterochromatin. Nat Struct Mol Biol. 2011 Jun 12; 18(7):777 82. https://doi.org/10.1038/nsmb.2070 PMID: 21666677

35. Lechner MS, Schultz DC, Negorev D, Maul GG, Rauscher FJ 3rd. The mammalian heterochromatin protein 1 binds diverse nuclear proteins through a common motif that targets the chromoshadow domain. Biochem Biophys Res Commun. 2005 Jun 17; 331(4):929 37. https://doi.org/10.1016/j.bbrc.2005.04.016 PMID: 15882967

36. Drané P, Ouararhni K, Depaux A, Shuaib M, Hamiche A. The death associated protein DAXX is a novel histone chaperone involved in the replication independent deposition of H3.3. Genes Dev. 2010 Jun 15; 24(12):1253 65. https://doi.org/10.1101/gad.566910 PMID: 20504901

37. Lewis PW, Elsaesser SJ, Noh KM, Stadler SC, Allis CD. Daxx is an H3.3 specific histone chaperone and cooperates with ATRX in replication independent chromatin assembly at telomeres. Proc Natl Acad Sci U S A. 2010 Aug 10; 107(32):14075 80. https://doi.org/10.1073/pnas.1008850107 PMID: 20651253

38. Elsässer SJ, Noh KM, Diaz N, Allis CD, Banaszynski LA. Histone H3.3 is required for endogenous ret roviral element silencing in embryonic stem cells. Nature. 2015 Jun 11; 522(7555):240 244. https://doi.org/10.1038/nature14345 PMID: 25938714

39. Udugama M, Chang FTM, Chan FL, Tang MC, Pickett HA, McGhie JDR, Mayne L, Collas P, Mann JR, Wong LH. Histone variant H3.3 provides the heterochromatic H3 lysine 9 tri methylation mark at telo meres. Nucleic Acids Res. 2015 Dec 2; 43(21):10227 37. https://doi.org/10.1093/nar/gkv847 PMID: 26304540

40. Ramamoorthy M, Smith S. Loss of ATRX Suppresses Resolution of Telomere Cohesion to Control Recombination in ALT Cancer Cells. Cancer Cell. 2015 Sep 14; 28(3):357 69. https://doi.org/10.1016/j.ccell.2015.08.003 PMID: 26373281

41. Hein MY, Hubner NC, Poser I, Cox J, Nagaraj N, Toyoda Y, Gak IA, Weisswange I, Mansfeld J, Buch holz F, Hyman AA, Mann M. A human interactome in three quantitative dimensions organized by stoi chiometries and abundances. Cell. 2015 Oct 22; 163(3):712 23. https://doi.org/10.1016/j.cell.2015.09.053 PMID: 26496610

42. Roux KJ, Kim DI, Raida M, Burke B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. J Cell Biol. 2012 Mar 19; 196(6):801 10. https://doi.org/10.1083/jcb.201112098 PMID: 22412018

43. Scott WA, Campos EI. Interactions With Histone H3 and Tools to Study Them. Front Cell Dev Biol. 2020 Jul 31; 8:701. https://doi.org/10.3389/fcell.2020.00701 PMID: 32850821

44. Räschle M, Smeenk G, Hansen RK, Temu T, Oka Y, Hein MY, Nagaraj N, Long DT, Walter JC, Hof mann K, Storchova Z, Cox J, Bekker Jensen S, Mailand N, Mann M. DNA repair. Proteomics reveals dynamic assembly of repair complexes during bypass of DNA cross links. Science. 2015 May 1; 348(6234):1253671. https://doi.org/10.1126/science.1253671 PMID: 25931565

45. Torres Rosell J, Machín F, Farmer S, Jarmuz A, Eydmann T, Dalgaard JZ, Aragón L. SMC5 and SMC6 genes are required for the segregation of repetitive chromosome regions. Nat Cell Biol. 2005 Apr; 7(4):412 9. https://doi.org/10.1038/ncb1239 PMID: 15793567

46. Aragón L. The Smc5/6 Complex: New and Old Functions of the Enigmatic Long Distance Relative. Annu Rev Genet. 2018 Nov 23; 52:89 107. https://doi.org/10.1146/annurev genet 120417 031353 PMID: 30476445

47. Gutierrez Escribano P, Hormeño S, Madariaga Marcos J, Solé Soler R, O'Reilly FJ, Morris K, Aicart Ramos C, Aramayo R, Montoya A, Kramer H, Rappsilber J, Torres Rosell J, Moreno Herrero F, Ara gon L. Purified Smc5/6 Complex Exhibits DNA Substrate Recognition and Compaction. Mol Cell. 2020 Dec 17; 80(6):1039 1054.e6. https://doi.org/10.1016/j.molcel.2020.11.012 PMID: 33301732

48. Serrano D, Cordero G, Kawamura R, Sverzhinsky A, Sarker M, Roy S, Malo C, Pascal JM, Marko JF, D'Amours D. The Smc5/6 Core Complex Is a Structure Specific DNA Binding and Compacting Machine. Mol Cell. 2020 Dec 17; 80(6):1025 1038.e5. https://doi.org/10.1016/j.molcel.2020.11.011 PMID: 33301731

49. Potts PR, Yu H. The SMC5/6 complex maintains telomere length in ALT cancer cells through SUMOy lation of telomere binding proteins. Nat Struct Mol Biol. 2007 Jul; 14(7):581 90. https://doi.org/10.1038/nsmb1259 PMID: 17589526

50. Min J, Wright WE, Shay JW. Alternative Lengthening of Telomeres Mediated by Mitotic DNA Synthesis Engages Break Induced Replication Processes. Mol Cell Biol. 2017 Sep 26; 37(20):e00226 17. https://doi.org/10.1128/MCB.00226 17 PMID: 28760773

51. Huttlin EL, Bruckner RJ, Navarrete Perea J, Cannon JR, Baltier K, Gebreab F, Gygi MP, Thornock A, Zarraga G, Tam S, Szpyt J, Gassaway BM, Panov A, Parzen H, Fu S, Golbazi A, Maenpaa E, Stricker K, Guha Thakurta S, Zhang T, Rad R, Pan J, Nusinow DP, Paulo JA, Schweppe DK, Vaites LP, Harper JW, Gygi SP. Dual proteome scale networks reveal cell specific remodeling of the human interac tome. Cell. 2021 May 27; 184(11):3022 3040.e28. https://doi.org/10.1016/j.cell.2021.04.011 PMID: 33961781

52. Oughtred R, Rust J, Chang C, Breitkreutz BJ, Stark C, Willems A, Boucher L, Leung G, Kolas N, Zhang F, Dolma S, Coulombe Huntington J, Chatr Aryamontri A, Dolinski K, Tyers M. The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interac tions. Protein Sci. 2021 Jan; 30(1):187 200. https://doi.org/10.1002/pro.3978 PMID: 33070389

53. Gullberg M, Fredriksson S, Taussig M, Jarvius J, Gustafsdottir S, Landegren U. A sense of closeness: protein detection by proximity ligation. Curr Opin Biotechnol. 2003 Feb; 14(1):82 6. https://doi.org/10.1016/S0958 1669(02)00011 3 PMID: 12566006

54. McLaurin J, Trudel GC, Shaw IT, Antel JP, Cashman NR. A human glial hybrid cell line differentially expressing genes subserving oligodendrocyte and astrocyte phenotype. J Neurobiol. 1995 Feb; 26 (2):283 93. https://doi.org/10.1002/neu.480260212 PMID: 7707048

55. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A, Olsson I, Edlund K, Lundberg E, Navani S, Szigyarto CA, Odeberg J, Djureino vic D, Takanen JO, Hober S, Alm T, Edqvist PH, Berling H, Tegel H, Mulder J, Rockberg J, Nilsson P, Schwenk JM, Hamsten M, von Feilitzen K, Forsberg M, Persson L, Johansson F, Zwahlen M, von Heijne G, Nielsen J, Pontén F. Proteomics. Tissue based map of the human proteome. Science. 2015 Jan 23; 347(6220):1260419. https://doi.org/10.1126/science.1260419 PMID: 25613900

56. Kriventseva EV, Kuznetsov D, Tegenfeldt F, Manni M, Dias R, Simão FA, Zdobnov EM. OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. Nucleic Acids Res. 2019 Jan 8; 47(D1):D807 D811. https://doi.org/10.1093/nar/gky1053 PMID: 30395283

57. Granneman S, Gallagher JE, Vogelzangs J, Horstman W, van Venrooij WJ, Baserga SJ, Pruijn GJ. The human Imp3 and Imp4 proteins form a ternary complex with hMpp10, which only interacts with the U3 snoRNA in 60 80S ribonucleoprotein complexes. Nucleic Acids Res. 2003 Apr 1; 31(7):1877 87. https://doi.org/10.1093/nar/gkg300 PMID: 12655004

58. Thul PJ, Åkesson L, Wiking M, Mahdessian D, Geladaki A, Ait Blal H, Alm T, Asplund A, Björk L, Breckels LM, Bäckström A, Danielsson F, Fagerberg L, Fall J, Gatto L, Gnann C, Hober S, Hjelmare M, Johansson F, Lee S, Lindskog C, Mulder J, Mulvey CM, Nilsson P, Oksvold P, Rockberg J, Schut ten R, Schwenk JM, Sivertsson Å, Sjöstedt E, Skogs M, Stadler C, Sullivan DP, Tegel H, Winsnes C, Zhang C, Zwahlen M, Mardinoglu A, Pontén F, von Feilitzen K, Lilley KS, Uhlén M, Lundberg E. A sub cellular map of the human proteome. Science. 2017 May 26; 356(6340):eaal3321. https://doi.org/10.1126/science.aal3321 PMID: 28495876

59. Wyler E, Zimmermann M, Widmann B, Gstaiger M, Pfannstiel J, Kutay U, Zemp I. Tandem affinity purification combined with inducible shRNA expression as a tool to study the maturation of macromolecular assemblies. RNA. 2011 Jan; 17(1):189 200. https://doi.org/10.1261/rna.2325911 PMID: 21097556

60. Faza MB, Chang Y, Occhipinti L, Kemmler S, Panse VG. Role of Mex67 Mtr2 in the nuclear export of 40S pre ribosomes. PLoS Genet. 2012; 8(8):e1002915. https://doi.org/10.1371/journal.pgen.1002915 PMID: 22956913

61. Smothers JF, Henikoff S. The HP1 chromo shadow domain binds a consensus peptide pentamer. Curr Biol. 2000 Jan 13; 10(1):27 30. https://doi.org/10.1016/S0960 9822(99)00260 2 PMID: 10660299

62. Liang J, Zhao H, Diplas BH, Liu S, Liu J, Wang D, Lu Y, Zhu Q, Wu J, Wang W, Yan H, Zeng YX, Wang X, Jiao Y. Genome Wide CRISPR Cas9 Screen Reveals Selective Vulnerability of ATRX Mutant Cancers to WEE1 Inhibition. Cancer Res. 2020 Feb 1; 80(3):510 523. https://doi.org/10.1158/0008 5472.CAN 18 3374 PMID: 31551363

63. Rodriguez R, Müller S, Yeoman JA, Trentesaux C, Riou JF, Balasubramanian S. A novel small molecule that alters shelterin integrity and triggers a DNA damage response at telomeres. J Am Chem Soc. 2008 Nov 26; 130(47):15758 9. https://doi.org/10.1021/ja805615w PMID: 18975896

64. Bechter OE, Zou Y, Walker W, Wright WE, Shay JW. Telomeric recombination in mismatch repair deficient human colon cancer cells after telomerase inhibition. Cancer Res. 2004 May 15; 64(10):3444 51. https://doi.org/10.1158/0008 5472.CAN 04 0323 PMID: 15150096

65. Garcia Exposito L, Bournique E, Bergoglio V, Bose A, Barroso Gonzalez J, Zhang S, Roncaioli JL, Lee M, Wallace CT, Watkins SC, Opresko PL, Hoffmann JS, O'Sullivan RJ. Proteomic Profiling Reveals a Specific Role for Translesion DNA Polymerase n in the Alternative Lengthening of Telomeres. Cell Rep. 2016 Nov 8; 17(7):1858 1871. https://doi.org/10.1016/j.celrep.2016.10.048 PMID: 27829156

66. Déjardin J, Kingston RE. Purification of proteins associated with specific genomic Loci. Cell. 2009 Jan 9; 136(1):175 86. https://doi.org/10.1016/j.cell.2008.11.045 PMID: 19135898

67. Grolimund L, Aeby E, Hamelin R, Armand F, Chiappe D, Moniatte M, Lingner J. A quantitative telomeric chromatin isolation protocol identifies different telomeric states. Nat Commun. 2013; 4:2848. https://doi.org/10.1038/ncomms3848 PMID: 24270157

68. Gao XD, Tu LC, Mir A, Rodriguez T, Ding Y, Leszyk J, Dekker J, Shaffer SA, Zhu LJ, Wolfe SA, Sontheimer EJ. C BERST: defining subnuclear proteomic landscapes at genomic elements with dCas9 APEX2. Nat Methods. 2018 Jun; 15(6):433 436. https://doi.org/10.1038/s41592 018 0006 2 PMID: 29735996

69. Zhou XZ, Lu KP. The Pin2/TRF1 interacting protein PinX1 is a potent telomerase inhibitor. Cell. 2001 Nov 2; 107(3):347 59. https://doi.org/10.1016/S0092 8674(01)00538 4 PMID: 11701125

70. Yoo JE, Oh BK, Park YN. Human PinX1 mediates TRF1 accumulation in nucleolus and enhances TRF1 binding to telomeres. J Mol Biol. 2009 May 22; 388(5):928 40. https://doi.org/10.1016/j.jmb. 2009.02.051 PMID: 19265708

71. 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 SAA, Ballard AJ, 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 AW, Kavukcuoglu K, Kohli P, Hassabis D. Highly accurate protein structure prediction with AlphaFold. Nature. 2021 Aug; 596 (7873):583 589. https://doi.org/10.1038/s41586 021 03819 2 PMID: 34265844

72. Firat Karalar EN, Rauniyar N, Yates JR 3rd, Stearns T. Proximity interactions among centrosome components identify regulators of centriole duplication. Curr Biol. 2014 Mar 17; 24(6):664 70. https:// doi.org/10.1016/j.cub.2014.01.067 PMID: 24613305

73. Van Itallie CM, Aponte A, Tietgens AJ, Gucek M, Fredriksson K, Anderson JM. The N and C termini of ZO 1 are surrounded by distinct proteins and functional protein networks. J Biol Chem. 2013 May 10; 288(19):13775 88. https://doi.org/10.1074/jbc.M113.466193 PMID: 23553632

74. Tang J, Wu S, Liu H, Stratt R, Barak OG, Shiekhattar R, Picketts DJ, Yang X. A novel transcription reg ulatory complex containing death domain associated protein and the ATR X syndrome protein. J Biol Chem. 2004 May 7; 279(19):20369 77. https://doi.org/10.1074/jbc.M401321200 PMID: 14990586

75. Garrick D, Samara V, McDowell TL, Smith AJ, Dobbie L, Higgs DR, Gibbons RJ. A conserved trun cated isoform of the ATR X syndrome protein lacking the SWI/SNF homology domain. Gene. 2004 Feb 4; 326:23 34. https://doi.org/10.1016/j.gene.2003.10.026 PMID: 14729260

76. McDowell TL, Gibbons RJ, Sutherland H, O'Rourke DM, Bickmore WA, Pombo A, Turley H, Gatter K, Picketts DJ, Buckle VJ, Chapman L, Rhodes D, Higgs DR. Localization of a putative transcriptional regulator (ATRX) at pericentromeric heterochromatin and the short arms of acrocentric chromosomes. Proc Natl Acad Sci U S A. 1999 Nov 23; 96(24):13983 8. https://doi.org/10.1073/pnas.96.24.13983 PMID: 10570185

77. Bérubé NG, Smeenk CA, Picketts DJ. Cell cycle dependent phosphorylation of the ATRX protein cor relates with changes in nuclear matrix and chromatin association. Hum Mol Genet. 2000 Mar 1; 9 (4):539 47. https://doi.org/10.1093/hmg/9.4.539 PMID: 10699177

78. Ratnakumar K, Duarte LF, LeRoy G, Hasson D, Smeets D, Vardabasso C, Bönisch C, Zeng T, Xiang B, Zhang DY, Li H, Wang X, Hake SB, Schermelleh L, Garcia BA, Bernstein E. ATRX mediated chro matin association of histone variant macroH2A1 regulates α globin expression. Genes Dev. 2012 Mar 1; 26(5):433 8. https://doi.org/10.1101/gad.179416.111 PMID: 22391447

79. Sarma K, Cifuentes Rojas C, Ergun A, Del Rosario A, Jeon Y, White F, Sadreyev R, Lee JT. ATRX Directs Binding of PRC2 to Xist RNA and Polycomb Targets. Cell. 2014 Nov 20; 159(5):1228. https:// doi.org/10.1016/j.cell.2014.11.010 PMID: 28898627

80. Nan X, Hou J, Maclean A, Nasir J, Lafuente MJ, Shu X, Kriaucionis S, Bird A. Interaction between chromatin proteins MECP2 and ATRX is disrupted by mutations that cause inherited mental retarda tion. Proc Natl Acad Sci U S A. 2007 Feb 20; 104(8):2709 14. https://doi.org/10.1073/pnas. 0608056104 PMID: 17296936

81. Kernohan KD, Jiang Y, Tremblay DC, Bonvissuto AC, Eubanks JH, Mann MR, Bérubé NG. ATRX partners with cohesin and MeCP2 and contributes to developmental silencing of imprinted genes in the brain. Dev Cell. 2010 Feb 16; 18(2):191 202. https://doi.org/10.1016/j.devcel.2009.12.017 PMID: 20159591

82. Gibbons RJ, McDowell TL, Raman S, O'Rourke DM, Garrick D, Ayyub H, Higgs DR. Mutations in ATRX, encoding a SWI/SNF like protein, cause diverse changes in the pattern of DNA methylation. Nat Genet. 2000 Apr; 24(4):368 71. https://doi.org/10.1038/74191 PMID: 10742099

83. Law MJ, Lower KM, Voon HP, Hughes JR, Garrick D, Viprakasit V, Mitson M, De Gobbi M, Marra M, Morris A, Abbott A, Wilder SP, Taylor S, Santos GM, Cross J, Ayyub H, Jones S, Ragoussis J, Rhodes D, Dunham I, Higgs DR, Gibbons RJ. ATR X syndrome protein targets tandem repeats and influences allele specific expression in a size dependent manner. Cell. 2010 Oct 29; 143(3):367 78. https://doi. org/10.1016/j.cell.2010.09.023 PMID: 21029860

84. Duc C, Benoit M, Détourné G, Simon L, Poulet A, Jung M, Veluchamy A, Latrasse D, Le Goff S, Cotter ell S, Tatout C, Benhamed M, Probst AV. Arabidopsis ATRX Modulates H3.3 Occupancy and Fine

Tunes Gene Expression. Plant Cell. 2017 Jul; 29(7):1773 1793. https://doi.org/10.1105/tpc.16.00877 PMID: 28684426

85. Udugama M, Sanij E, Voon HPJ, Son J, Hii L, Henson JD, Chan FL, Chang FTM, Liu Y, Pearson RB, Kalitsis P, Mann JR, Collas P, Hannan RD, Wong LH. Ribosomal DNA copy loss and repeat instability in ATRX mutated cancers. Proc Natl Acad Sci U S A. 2018 May 1; 115(18):4737 4742. https://doi.org/10.1073/pnas.1720391115 PMID: 29669917

86. Götz S, Pandey S, Bartsch S, Juranek S, Paeschke K. A Novel G Quadruplex Binding Protein in Yeast Slx9. Molecules. 2019 May 7; 24(9):1774. https://doi.org/10.3390/molecules24091774 PMID: 31067825

87. Cosnier B, Kwapisz M, Hatin I, Namy O, Hermann Le Denmat S, Morillon A, Rousset JP, Fabret C. A viable hypomorphic allele of the essential IMP3 gene reveals novel protein functions in Saccharomy ces cerevisiae. PLoS One. 2011 Apr 29; 6(4):e19500. https://doi.org/10.1371/journal.pone.0019500 PMID: 21559332

88. Cao Q, Wang X, Zhao M, Yang R, Malik R, Qiao Y, Poliakov A, Yocum AK, Li Y, Chen W, Cao X, Jiang X, Dahiya A, Harris C, Feng FY, Kalantry S, Qin ZS, Dhanasekaran SM, Chinnaiyan AM. The central role of EED in the orchestration of polycomb group complexes. Nat Commun. 2014; 5:3127. https://doi.org/10.1038/ncomms4127 PMID: 24457600

89. Huttlin EL, Bruckner RJ, Paulo JA, Cannon JR, Ting L, Baltier K, Colby G, Gebreab F, Gygi MP, Par zen H, Szpyt J, Tam S, Zarraga G, Pontano Vaites L, Swarup S, White AE, Schweppe DK, Rad R, Erickson BK, Obar RA, Guruharsha KG, Li K, Artavanis Tsakonas S, Gygi SP, Harper JW. Architec ture of the human interactome defines protein communities and disease networks. Nature. 2017 May 25; 545(7655):505 509. https://doi.org/10.1038/nature22366 PMID: 28514442

90. Chudnovsky Y, Kim D, Zheng S, Whyte WA, Bansal M, Bray MA, Gopal S, Theisen MA, Bilodeau S, Thiru P, Muffat J, Yilmaz OH, Mitalipova M, Woolard K, Lee J, Nishimura R, Sakata N, Fine HA, Car penter AE, Silver SJ, Verhaak RG, Califano A, Young RA, Ligon KL, Mellinghoff IK, Root DE, Sabatini DM, Hahn WC, Chheda MG. ZFHX4 interacts with the NuRD core member CHD4 and regulates the glioblastoma tumor initiating cell state. Cell Rep. 2014 Jan 30; 6(2):313 24. https://doi.org/10.1016/j.celrep.2013.12.032 PMID: 24440720

91. Happ H, Schilter KF, Weh E, Reis LM, Semina EV. 8q21.11 microdeletion in two patients with syndro mic peters anomaly. Am J Med Genet A. 2016 Sep; 170(9):2471 5. https://doi.org/10.1002/ajmg.a.37840 PMID: 27378168

92. Adhikari S, Thakur N, Shrestha U, Shrestha MK, Manshrestha M, Thapa B, Poudel M, Kunwar A. Genetic analysis of children with congenital ocular anomalies in three ecological regions of Nepal: a phase II of Nepal pediatric ocular diseases study. BMC Med Genet. 2020 Sep 22; 21(1):185. https://doi.org/10.1186/s12881 020 01116 9 PMID: 32962661

93. Palomares M, Delicado A, Mansilla E, de Torres ML, Vallespín E, Fernandez L, Martinez Glez V, Gar cía Miñaur S, Nevado J, Simarro FS, Ruiz Perez VL, Lynch SA, Sharkey FH, Thuresson AC, Annerén G, Belligni EF, Martínez Fernández ML, Bermejo E, Nowakowska B, Kutkowska Kazmierczak A, Bocian E, Obersztyn E, Martínez Frías ML, Hennekam RC, Lapunzina P. Characterization of a 8q21.11 microdeletion syndrome associated with intellectual disability and a recognizable phenotype. Am J Hum Genet. 2011 Aug 12; 89(2):295 301. https://doi.org/10.1016/j.ajhg.2011.06.012 PMID: 21802062

94. Kwon MS, Lee JJ, Min J, Hwang K, Park SG, Kim EH, Kim BC, Bhak J, Lee H. Brca2 abrogation engages with the alternative lengthening of telomeres via break induced replication. FEBS J. 2019 May; 286(10):1841 1858. https://doi.org/10.1111/febs.14796 PMID: 30811824

95. Dilley RL, Verma P, Cho NW, Winters HD, Wondisford AR, Greenberg RA. Break induced telomere synthesis underlies alternative telomere maintenance. Nature. 2016 Nov 3; 539(7627):54 58. https://doi.org/10.1038/nature20099 PMID: 27760120

96. Tashiro S, Handa T, Matsuda A, Ban T, Takigawa T, Miyasato K, Ishii K, Kugou K, Ohta K, Hiraoka Y, Masukata H, Kanoh J. Shugoshin forms a specialized chromatin domain at subtelomeres that regu lates transcription and replication timing. Nat Commun. 2016 Jan 25; 7:10393. https://doi.org/10.1038/ncomms10393 PMID: 26804021

97. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR Cas9 system. Nat Protoc. 2013 Nov; 8(11):2281 2308. https://doi.org/10.1038/nprot.2013.143 PMID: 24157548

98. Lambert JP, Tucholska M, Go C, Knight JD, Gingras AC. Proximity biotinylation and affinity purification are complementary approaches for the interactome mapping of chromatin associated protein complexes. J Proteomics. 2015 Apr 6; 118:81 94. https://doi.org/10.1016/j.jprot.2014.09.011 PMID: 25281560

99. Liu G, Knight JD, Zhang JP, Tsou CC, Wang J, Lambert JP, Larsen B, Tyers M, Raught B, Bandeira N, Nesvizhskii AI, Choi H, Gingras AC. Data Independent Acquisition analysis in ProHits 4.0. J Proteo mics. 2016 Oct 21; 149:64 68. https://doi.org/10.1016/j.jprot.2016.04.042 PMID: 27132685

100. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Probability based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis. 1999 Dec; 20(18):3551 67. https://doi.org/10.1002/(SICI)1522 2683(19991201)20:18%3C3551::AID ELPS3551%3E3.0.CO;2 2 PMID: 10612281

101. Eng JK, Jahan TA, Hoopmann MR. Comet: an open source MS/MS sequence database search tool. Proteomics. 2013 Jan; 13(1):22 4. https://doi.org/10.1002/pmic.201200439 PMID: 23148064

102. Shteynberg D, Deutsch EW, Lam H, Eng JK, Sun Z, Tasman N, Mendoza L, Moritz RL, Aebersold R, Nesvizhskii AI. iProphet: multi level integrative analysis of shotgun proteomic data improves peptide and protein identification rates and error estimates. Mol Cell Proteomics. 2011 Dec; 10(12): M111.007690. https://doi.org/10.1074/mcp.M111.007690 PMID: 21876204

103. Teo G, Liu G, Zhang J, Nesvizhskii AI, Gingras AC, Choi H. SAINTexpress: improvements and addi tional features in Significance Analysis of INTeractome software. J Proteomics. 2014 Apr 4; 100:37 43. https://doi.org/10.1016/j.jprot.2013.10.023 PMID: 24513533

104. Knight JDR, Choi H, Gupta GD, Pelletier L, Raught B, Nesvizhskii AI, Gingras AC. ProHits viz: a suite of web tools for visualizing interaction proteomics data. Nat Methods. 2017 Jun 29; 14(7):645 646. https://doi.org/10.1038/nmeth.4330 PMID: 28661499

105. Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, Vilo J. g:Profiler: a web server for func tional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res. 2019 Jul 2; 47(W1):W191 W198. https://doi.org/10.1093/nar/gkz369 PMID: 31066453

106. McQuin C, Goodman A, Chernyshev V, Kamentsky L, Cimini BA, Karhohs KW, Doan M, Ding L, Rafelski SM, Thirstrup D, Wiegraebe W, Singh S, Becker T, Caicedo JC, Carpenter AE. CellProfiler 3.0: Next generation image processing for biology. PLoS Biol. 2018 Jul 3; 16(7):e2005970. https://doi.org/10.1371/journal.pbio.2005970 PMID: 29969450

107. Abramoff M.D., Magalhaes P.J., Ram S.J. Image Processing with ImageJ Biophotonics International 2004; 7(11):36 42. PMID: 15084075

108. Campos EI, Smits AH, Kang YH, Landry S, Escobar TM, Nayak S, Ueberheide BM, Durocher D, Ver meulen M, Hurwitz J, Reinberg D. Analysis of the Histone H3.1 Interactome: A Suitable Chaperone for the Right Event. Mol Cell. 2015 Nov 19; 60(4):697 709. https://doi.org/10.1016/j.molcel.2015.08.005 PMID: 26527279

109. Williams ES, Bailey SM. Chromosome orientation fluorescence in situ hybridization (CO FISH). Cold Spring Harb Protoc. 2009 Aug; 2009(8):pdb.prot5269. https://doi.org/10.1101/pdb.prot5269 PMID: 20147245

110. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, Tosatto SCE, Pala din L, Raj S, Richardson LJ, Finn RD, Bateman A. Pfam: The protein families database in 2021. Nucleic Acids Res. 2021 Jan 8; 49(D1):D412 D419. https://doi.org/10.1093/nar/gkaa913 PMID: 33125078

111. Sigrist CJ, de Castro E, Cerutti L, Cuche BA, Hulo N, Bridge A, Bougueleret L, Xenarios I. New and continuing developments at PROSITE. Nucleic Acids Res. 2013 Jan; 41(Database issue):D344 7. https://doi.org/10.1093/nar/gks1067 PMID: 23161676

112. Papadopoulos JS, Agarwala R. COBALT: constraint based alignment tool for multiple protein sequences. Bioinformatics. 2007 May 1; 23(9):1073 9. https://doi.org/10.1093/bioinformatics/btm076 PMID: 17332019

113. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. Genome Res. 2004 Jun; 14(6):1188 90. https://doi.org/10.1101/gr.849004 PMID: 15173120