The Role of Fanconi Anaemia Factors in the Response to Transcription-Replication Conflict

Thesis submitted for the degree of Doctor of Philosophy

Jamie Langton

University College

Department of Oncology

University of Oxford

Trinity Term 2018

The Role of Fanconi Anaemia Factors in the Response to Transcription-Replication Conflict

Jamie Langton, University College

Thesis submitted for the degree of Doctor of Philosophy, Trinity Term 2018

Replication and transcription both require large, multi-protein complexes that bind and translocate DNA to facilitate genome duplication and gene expression. This shared requirement inevitably leads to conflict because, despite numerous cellular adaptations, these molecular machines will utilize the same areas of the DNA template during S-phase. This conflict results in transcription being a source of replication stress and ultimately genomic instability.

How these conflicts are managed and resolved is poorly understood - the current study attempts to investigate the role of the Fanconi anaemia (FA) pathway in resolving replication stress caused by conflict with transcription. The FA pathway is associated with the repair of interstrand cross-links (ICLs) produced by exogenous agents, but it is not clear if these types of lesions are produced endogenously or represent the only substrate of the FA pathway.

Loss of the FA pathway leads to an increased level of transcription associated DNA:RNA hybrids (R-loops) and associated replication defects and damage. The current study demonstrates that these structures are relevant to the pathology of FA, as they

accumulate in cells derived from FA patients and in murine haematopoietic stem cells (HSCs).

This strongly suggests the FA pathway has a specific role in dealing with cell stress arising from transcription and R-loops. In line with this notion, this study demonstrates that depletion of the transcription-coupled repair factor CSB leads to increased transcription-replication conflicts that requires the action of FA factors. In the absence of FA proteins, the loss of CSB leads to replication fork stalling, mitotic catastrophe and reduced proliferation.

Overall, this data supports the growing evidence that the FA pathway is required to respond to conflicts between replication and transcription, as well as transcription-associated R-loops. My data suggests that replication fork collisions with transcription complexes and/or R-loops represents a major substrate of the FA pathway and suggests transcription-replication conflict could contribute to the attrition of HSCs in FA patients.

Acknowledgements

I would first like to thank my supervisors, Prof Wojciech Niedzwiedz and Prof Peter McHugh. My thanks go to Wojciech for his willingness to take on a student with a nonlinear path to the DPhil and for the inception of and guidance in a project that led to the completion of this thesis. I am also very grateful to Peter for adopting me toward the end of my studies and guiding me toward submission and the world of academia beyond.

I owe a great deal to the people within the Niedzwiedz and McHugh labs. Rebekka, Jadwiga and Ronan were excellent tutors in my first fumbling attempts to undertake cellular and molecular biology, thanks to their training I was well placed to complete experiments independently toward the end of my DPhil.

To the members of the McHugh lab: Lonnie for his expertise and sardonicism in the lab and for providing me with great opportunities. I have a deep gratitude for Sanja, HWebb, Ruth and Will for making one of the toughest years of my life also one of the most fun, memorable and exciting. Thank you to HBadd for inspiring me to be a better human every day, and for the ferocious hugs. I also want to thank the members of the Blackford lab: Sam, Carl and Ann-Marie for being great coffee and dinner pals but also being amazing at pub quizzes. I cannot imagine I would have made it through my final year without swapping labs and becoming close to these people.

Thank you also to the Axis of Argyle: Todd, PChest, Lizzie, Ben and Franco. Thank you for being amazing housemates and letting me win so many board games. Thank you to Todd in particular – firstly for being Todd Carter, an endless source of wonder and amusement, but also being a true friend who was always there when needed.

I would also like to thank the many friends I made at UNIV, of which there are too many to name, who broadened my horizons and provided ample friendship and distraction from work.

Finally, none of what I have achieved in life would be remotely possible without the love, stability and the physical, emotional and financial support provided by my parents. At every stage of my life they have provided me with opportunities that they did not have - I can only hope I have made the most of them.

List of Abbreviations

53BP1	p53 binding protein 1
5-EU	5-ethynyl uridine
9-1-1 complex	RAD9-RAD1-HUS1 complex
AID	Activation induced cytosine deaminase
ALDH2/5	aldehyde dehydrogenase 2/5
ALT	Alternative lengthening of telomeres
AP site	Apurinic / apyrimidinic site
AT	Ataxia telangiectasia
ATM	Ataxia telangiectasia mutated
ATR	ATM- and Rad3 related
ATRIP	ATR-interacting protein
BER	Base excision repair
BLM	Bloom syndrome protein
BMF	Bone marrow failure
BRCA1/2	breast cancer type 1/2 susceptibility protein
BrdU	5-bromo-2'-deoxyuridine
Cas9	CRISPR-associated protein 9
CDK	Cyclin dependent kinase
CFS	Common fragile site
CldU	5-chloro-2’-deoxyuridine
cNHEJ	Classical (or canonical) NHEJ
CPD	Cyclobutane pyrimidine dimers
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CS	Cockayne syndrome
CSA	Cockayne syndrome group B protein
CSB	Cockayne syndrome group A protein
CSR	Class switch recombination
DDR	DNA damage response
D-loop	Displacement loop
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic Acid
DNA-PK	DNA-dependent protein kinase
DNA-PKcs	DNA-PK catalytic subunit
DPC	DNA-protein cross-link
DSB	Double strand break
dsDNA	Double strand DNA
ERCC6	Excision Repair Cross-Complementation Group 6

FA	Fanconi anaemia
FAAP	Fanconi anaemia associated protein
FCS	Foetal calf serum
G4	Quadruplex
GG-NER	Global genome NER
gRNA	Guide RNA
HR	Homologous recombination
HRP	Horseradish peroxidase
HSC	haematopoietic stem cell
HU	Hydroxyurea
ICL	Interstrand cross-link
ID2	FANCD2-FANCI heterodimer complex
IdU	5-iodo-2'-deoxyuridine
IF	Immunofluorescent
iPOND	Isolation of protein on nascent DNA
IR	Ionizing Radiation
kb	kilobase
kDa	kilo Dalton
LT-HSC	Long-term HSC
miRNA	Micro RNA
MMC	Mitomycin C
MMR	Mismatch repair
MRN	Mre11-Rad50-Nbs1 complex
mRNA	Messenger RNA
ncRNA	Non-coding RNA
NER	Nucleotide excision repair
NHEJ	Non-homologous end joining
PARP	Poly(ADP)-ribose polymerase
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
PIKK	Phosphatidylionsitol-3-OH-kinase
Pol	Polymerase
rDNA	Ribosomal DNA
RNA	Ribonucleic acid
RNAi	RNA interference
RNAP	RNA polymerase
RNase	Ribonuclease
ROS	Reactive oxygen species
RPA	Replication protein A
rRNA	Ribosomal RNA
SD	Standard deviation

SEM	Standard error of the mean
SETX	Senataxin
siRNA	Small interfering RNA
snRNA	Small nuclear RNA
SSB	Single strand break
ssDNA	Single strand DNA
TC-NER	Transcription-coupled NER
TCR	Transcription-coupled repair
TFIIH	Transcription factor II H
TLS	Translesion synthesis
TOPBP1	Topoisomerase-binding protein 1
tRNA	Transfer RNA
TTD	Trichothiodystrophy
UV	Ultraviolet
XP	Xeroderma pigmentosum

Experimental Contributions

The CRISPR gRNA design, cloning, transfection and cell sorting of the FANCD2 U2OS knockout clones used throughout was performed by Jadwiga Nieminuszczy.

Table of Contents

1. Introduction 1

1.1 Genome stability 1

1.2 Sources of genome instability 2

1.2.1 DNA processes errors 2

1.2.2 DNA damage 3

1.3 The DNA damage response 5

1.3.1 ATM 7

1.3.2 ATR 9

1.3.3 DNA-PK 11

1.4 DNA repair pathways 14

1.4.1 Mismatch repair and base excision repair 14

1.4.2 Nucleotide excision repair 15

1.4.3 Double strand break repair 21

1.4.4 Cross-link repair 27

1.5 The Fanconi anaemia pathway 29

1.5.1 FA pathway activation in response to ICLs 34

1.5.2 FA mediated removal of ICLs and repair of DNA 37

1.5.3 Replication fork protection by the FA pathway 38

1.5.4 The association of the FA pathway and reactive aldehydes 40

1.6 Transcription-replication conflict 41

1.7 R-loops....44

1.7.1 Physiological function of R-loops....49

1.7.2 R-loop driven instability....50

1.8 FA pathway and FA factors implicated in transcription-replication conflict ...52

1.9 Aims of the current study 54

2. Materials and Methods....55

2.1 Cell culture 55

2.2 siRNA knockdown....55

2.3 Antibodies 57

2.4 Western Blotting 57

2.5 Agarose gel electrophoresis....58

2.6 DNA fiber technique....58

2.7 Immunofluorescence staining of DDR markers 59

2.8 R-loop staining by IF and the S9.6 antibody 60

2.9 Anaphase bridge analysis....61

2.10 Cell survival assay....61

2.11 Cell growth assay....62

2.12 Generation of FANCD2 knockout line by CRISPR/Cas9....62

2.13 R-loop detection by S9.6 Slot-blot 63

2.14 Transcription assay....64

2.15 Analysis....64

3. The relevance of R-loops to the pathology of Fanconi anaemia....66

3.1 Introduction 66

3.2 Results 69

3.2.1 Knockout of FANCD2 via CRISPR-Cas9 69

3.2.2 FANCD2$^{-/-}$ knockout clones are sensitive to cisplatin 72

3.2.3 Increased R-loop formation is a general phenomenon of FA deficient cells 73

3.2.4 Increased R-loop formation is present in cells derived from patients and FA haematopoietic stem cells. 78

3.3 Conclusion 84

4. The FA pathway protects against transcription-based cellular stress 86

4.1 Introduction 86

4.2 Results 88

4.2.1 Targeting the TCR pathway by RNA interference of CSB 88

4.2.2 CSB depletion increases R-loop formation 91

4.2.3 Depletion of CSB causes replication-transcription conflict 94

4.2.4 Inducing transcriptional stress in FA deficient cells by knockdown of CSB increases R-loops and DNA damage 96

4.2.5 Loss of CSB in FA deficient cells causes mitotic catastrophe and severely diminished rate of proliferation 101

4.2.6 Depletion of FANCD2 and CSB causes similar growth attrition in HeLa cells 105

4.2.7 The requirement of CSB for viability in FA deficient cells does not extend to other members of the TC-NER pathway 107

4.3 Conclusion 109

5. Discussion 110

5.1 The FA pathway response to transcription-associated replication stress 110

5.2 Targeting transcription-replication conflict in cancer therapy 114

5.3 The contribution of transcription-replication conflict to the pathology of FA 118

5.4 Conclusion 122

6. References 124

1. Introduction

1.1 Genome stability

DNA is a biological information carrier that serves as the hereditary molecule. In isolation, the sequence of nucleic acids contained in DNA remain relatively stable. However, in an active cell DNA is regularly exposed to chemical attack whilst also passively involved in transactions with numerous sets of molecular actors to achieve basic cellular processes (Lindahl, 1993).

The sheer magnitude of exposure and the number of transactions on DNA can result in changes in the underlying sequence. This plasticity is at the heart of the process of evolution and speciation, where random mutations conferring selective advantages has driven the diversification of organisms – allowing terrestrial life to fill ecological niches and withstand environmental change.

However, to maintain the overall advantage of this plasticity the rate of change must be minimal, and the vulnerability of DNA and volatility of transactions performed upon it comes at a price of substantial cellular investment in the active management and protection of the underlying genomic sequence. DNA is constantly under threat of change through spontaneous modifications, processing errors and reactions with genotoxic agents. There are many compounds which can cause damage or changes to DNA, some of which are produced endogenously as products of normal cellular metabolism.

Accordingly, cells have evolved an abundance of proteins, processes and pathways dedicated to maintaining the integrity and fidelity of the genomic sequence. The importance of protecting DNA is evident as instability of the genome is the main etiological factor in a range of diseases, particularly in the development of cancers and the accumulation of DNA mutations is thought to be central to aging (Bohr and Michael Anson, 1995; Hoeijmakers, 2009). Understanding the molecular mechanisms by which cells protect their DNA and maintain the integrity of the genomic sequence is of vital importance in understanding human disease.

1.2 Sources of genome instability

1.2.1 DNA processes errors

Sequence alterations can occur due to errors during DNA replication and it is estimated incorrect nucleotides are incorporated by DNA polymerases every 10^8 bases (Kunkel, 2004). However, most of these mistakes are fixed immediately during replication due to DNA polymerase proofreading, possessed in eukaryotes by Pol δ and Pol ε, which perform the majority of DNA replication. These polymerases recognize the incorrect 3'OH group position of the erroneously incorporated nucleotide and possess a 3'-5' exonuclease activity to excise the incorrect nucleotide (Bernad et al., 1989). This reduces the error rate of replication ~100 fold and substantially reducing the risk of mutation (Loeb, 1991). DNA bases can also spontaneously change through deamination, depurination or modification by alkylation (Ciccia and Elledge, 2010).

1.2.2 DNA damage

DNA damage represents the greatest threat to genome stability and can originate from both endogenous and exogenous sources. External DNA damage can originate from physical or chemical sources. Exposure to ultraviolet (UV, specifically UVA at 320-400 nanometres and UVB at 290-320 nanometres) light from sunlight is a major contributor to the formation of squamous and basal cell carcinomas (Brash et al., 1996). UV radiation can induce the formation covalent linkages between adjacent pyrimidines, forming pyrimidine dimers and 6-4 photoproducts, which block polymerase progression and arrest replication forks (Whitmore et al., 2001). Ionizing radiation (IR) is commonly employed in some medical procedures, such as X-rays or radiotherapy, and can induce oxidation of DNA bases and DNA breaks (Vignard et al., 2013).

A wide range of chemical agents in cancer therapy operate by inducing covalent cross-links between bases (Cheung-Ong et al., 2013). These can be between bases in the same strand (intrastrand) or between bases that lie on the opposite strands (interstrand). Interstrand cross-links (ICLs) are particularly toxic to cells as they prevent the separation of the two strands of DNA thereby blocking replication and transcription (Noll et al., 2006).

DNA damage can arise endogenously due to exposure to the intermediates and end-products of normal cellular metabolism, including aldehydes or reactive oxygen species (ROS). These compounds can react with DNA and introduce a range of lesions, including

base modification, single- and double-strand breaks (SSB and DSBs, respectively) and DNA cross-links (Klungland et al., 1999; O'Brien et al., 2005).

Whilst DNA repair acts to protect genome stability, in some contexts the survival of a cell comes at the price of undergoing DNA repair that results in mutations or deletions of sequence. For example, when homologous sequences aren't available to provide a template for repair, erroneous end-joining mechanisms can be employed which have no means for restoring sequence around a DNA break. In addition, the probability of chromosomal rearrangements is much higher, as end joining mechanisms of repair can catalyse the joining of unrelated DNA molecules.

DNA repair processes require the action of proteins that have enzymatic activities that chemically manipulate and modify DNA. These include nucleases, helicases, polymerases, topoisomerases, recombinases, ligases, glycosylases, demethylases, kinases, and phosphatases. These enzymes also have the potential to significantly alter the DNA sequence if they are able to access and modify DNA without restraint.

Damage and instability can also arise endogenously due to the formation of non-B form DNA structures. Eukaryotic genomes are highly ordered and organised, but the double helical structure of B-form DNA is not always the most thermodynamically stable configuration, particularly when the double helix is unwound during replication or transcription (Zhao et al., 2010). There are many alternative DNA structures that can form, some of which possess biological function, but their aberrant formation or

persistence can cause genome instability by blocking the progression of replication and transcription machineries (See section 1.6 and 1.7) (Bacolla and Wells, 2004; Mirkin, 2008).

1.3 The DNA damage response

DNA damage represents the main threat to genomic stability and occurs frequently in all cells of the human body, estimated as up to 10^5 DNA lesions per day per cell, originating from endogenous and exogenous sources (Lindahl and Barnes, 2000).

The threat to DNA from a broad spectrum of sources has prompted the evolution of a complex and intricate network of pathways dedicated to sensing and repairing DNA lesions. This DNA damage response (DDR) network is vital in preventing the accumulation of deleterious DNA damage that can lead to genome instability and facilitate the progression of cells to a diseased state. The DDR thereby acts to guard the genome and prevent disease which is illustrated by the severe consequences of inherited diseases caused by the loss or defects in DDR factors, which includes neurological degeneration, immune deficiency, premature aging and a significantly increased risk of cancer (Hoeijmakers, 2009; Jackson and Bartek, 2009).

In general, the cellular response to DNA damage is enacted through a signal transduction pathway composed of sensors, transducers and effectors which comprise DDR pathways (Figure 1.1). It is activated by aberrant DNA structures, such as bulky DNA adducts, DNA breaks, or by stress sensed through the replication and transcription

apparatus. When problems in DNA are detected, sensor proteins activate a signalling cascade, predominantly through phosphorylation, that control a wide range of cellular processes that are vital for genome stability that include DNA repair, DNA replication and cell-cycle control. In the event of irreparable damage or the accumulation of a substantial amount of damage, this signalling cascade can steer the cell towards senescence or apoptosis, to avoid the propagation of DNA sequence errors (Roos and Kaina, 2006; te Poele et al., 2002). Cell-cycle control is achieved through three checkpoints at transition stages of the cell cycle to allow for a pause in cell cycle progression to ensure there is sufficient time for repair to occur for accurate genome duplication (Zhou and Elledge, 2000). Overall, the DDR response is integral for an organism to sense and signal problems with its DNA and ensure cells do not progress to a diseased or malignant state.

Fundamental to the DDR in mammalian cells are the master DDR regulators ataxia-telangiectasia mutated (ATM), ATM- and Rad3-related (ATR) and DNA-dependent protein kinase (DNA-PK) (Blackford and Jackson, 2017). These three proteins are members of the phosphatidylinositol-3-OH-kinase(PI3K)-related kinase (PIKK) family of proteins that act as transducers in DDR and are conserved from yeast to mammals. In addition to these PIKK family proteins, poly(ADP-ribose) polymerases (PARPs) signal DNA repair through the addition of multiple ADP-ribose molecules on target substrates (Ray Chaudhuri and Nussenzweig, 2017).

Signalling through these transducers requires sensor proteins to detect damage to localize them to DNA, but the presence of damage is not required for transducers to propagate the DDR, as their signalling cascades can be initiated in the absence of DNA damage by their prolonged binding to chromatin (Soutoglou and Misteli, 2008).

Post-translational modifications are central to cellular signalling and modulates substrate stability, localization, activity and interactions with other proteins. These modifications include phosphorylation, ubiquitination, SUMOylation and poly(ADP)-ribosylation. The PIKKs mentioned above are apical protein kinases that can act alone or in concert to respond to DNA damage by stimulating an extensive signalling network.

1.3.1 ATM

ATM is a large kinase at the apex of the cellular response to DNA double-strand breaks (DSBs) and phosphorylates proteins on serine or threonine residues that are followed by glutamine, named 'SQ/TQ' motifs (Bakkenist and Kastan, 2004). ATM is named after the autosomal recessive disorder ataxia telangiectasia (AT), which is caused by loss or inactivation of ATM (Savitsky et al., 1995). AT is a neurodegenerative disease that occurs in childhood and results in uncoordinated (or ataxic) movement and is characterized by cerebellar degeneration (McKinnon, 2004). In undamaged cells ATM lies dormant in a homodimer configuration that reportedly blocks both kinase domains. Direction interaction of ATM with NBS1 of the MRE11/RAD50/NBS1 (MRN) complex bound to damaged DNA leads to its conversion from dormant dimer to two active monomeric

kinases (Falck et al., 2005). Activated ATM undergoes autophosphorylation and proceeds to rapidly phosphorylate hundreds of other substrates to finely tune the DDR at all levels. The phosphorylation of substrates enhances or represses their activity thereby mediating a broad range of DDR processes including checkpoint activation, apoptosis, senescence and DNA repair (Shiloh and Ziv, 2013).

Although ATM phosphorylates numerous substrates, there are several key proteins that are focal points for further signalling. Two such proteins are p53 and CHK2, which are themselves vital in the DDR and mutations in which are heavily associated with cancer (Apostolou and Papasotiriou, 2017; Joerger and Fersht, 2016). CHK2 is another protein kinase, that has many of its own phosphorylation targets, demonstrating that ATM signalling is not limited to directly phosphorylated targets (Bartek and Lukas, 2003).

p53, arguably the most important tumour suppressor, rapidly accumulates in response to DNA damage and activates a transcription program that results in expression of target genes involved in halting cell cycle progression in G1, as well as apoptosis and senescence (Joerger and Fersht, 2016). ATM stabilizes p53 through limiting its interaction with the ubiquitin ligase MDM2, which promotes its degradation. This is achieved through direct phosphorylation of p53 on multiple sites or by phosphorylation of MDM2 (Cheng and Chen, 2010).

Another target of ATM's kinase activity is serine 139 of histone H2A.X (referred to as γH2AX after phosphorylation), that is phosphorylated heavily in chromatin flanking the

lesion after the formation of DSBs, serving as an indicator of the nearby break and subsequently is interacted with by a large array of repair factors (Rogakou et al., 1999; Rogakou et al., 1998).

1.3.2 ATR

ATR sits at the top of the hierarchy of the cellular response to DNA replication stress and, unlike ATM or DNA-PKcs, is essential in proliferating cells (de Klein et al., 2000). Hypomorphic mutations in ATR cause Seckel syndrome, a rare disorder that shares some characteristics of other DNA repair deficiency syndromes, such as slowed growth, microcephaly and a distinct “bird-like” facial appearance (O'Driscoll et al., 2003).

The requirement for ATR in viability is attributed to the ubiquitous nature of replication stress and the requirement of ATR in regulating replication origin firing (Chen et al., 2015; Marheineke and Hyrien, 2004). Despite this difference, ATM and ATR are biochemically and functionally similar - ATR also phosphorylates numerous substrates and many them are likely shared by ATM as they play general DDR roles, such as histone H2AX (Chanoux et al., 2009).

ATR is activated in a range of DNA damage contexts, but it seems to be particularly strongly activated in response to the formation of single-stranded DNA (ssDNA) that is subsequently coated with replication protein A (RPA) (Zou and Elledge, 2003). ATR forms a heterodimer with its binding partner, ATR-interacting protein (ATRIP), which appears

to be responsible for recognition of RPA coated ssDNA and promoting ATR localization to sites of replication stress or DNA damage (Cimprich and Cortez, 2008). This alone is not enough for ATR activation, as the RAD9-RAD1-HUS1 (9-1-1 complex) is required to recognize and bind the DNA end that is adjacent to the stretch of RPA-ssDNA and recruit topoisomerase-binding protein 1 (TOPBP1), which then activates ATR (Kumagai et al., 2006; Lee et al., 2007).

Similar to ATM, a key substrate of ATR is another protein kinase that extends the signalling network. Phosphorylation of CHK1 is important for degradation of CDC25A, which stabilizes cyclin-dependent kinase 1 (CDK1) causing a cell cycle arrest in G2/M (Zhao et al., 2002). CHK1 also affects replication progression at the level of origin firing and acts to protect fork integrity, providing sufficient time for replication fork rescue and prevents early entry into mitosis (Petermann et al., 2006; Petermann et al., 2010b).

Unique to ATR, in comparison to the other PIKKs, is the role it plays in the repair of DNA cross-links where it appears to respond to replication forks stalled at interstrand cross-links by promoting FANCD2 mono-ubiquitylation. This is the key step in activating the Fanconi anaemia (FA) pathway, responsible for ICL repair (Andreassen et al., 2004). This activation is likely achieved through the ATR dependent phosphorylation of FANCI, the binding partner of FANCD2 (Chen et al., 2015; Ishiai et al., 2008).

ATR is implicated in the in S/G2 transition checkpoint, as treatment of cells with an ATR inhibitor (ATRi) in S-phase resulted in early entry into mitosis and the accumulation of

cyclin B and other pro-mitotic factors (Saldivar et al., 2018). These differentially expressed genes are bound at their promoters by the forkhead box protein M1 (FOXM1) and depletion of FOXM1 ablated cyclin B accumulation in ATRi-treated cells. Saldivar et al. (2018) propose a model wherein ATR is kept in an active state during S-phase through RPA recruited ETAA1, a known ATR activator. Once replication is completed ETAA1 is no longer recruited and ATR becomes inactive, allowing the expression of pro-mitotic factors and allows the S/G2 transition.

1.3.3 DNA-PK

The DNA-dependent protein kinase (DNA-PK) is comprised of DNA-PKcs and the Ku hetero-dimer, which consists of the Ku80 and Ku70 subunits. The discovery of DNA-PK was precipitated by the finding that kinase activity could be stimulated by double strand DNA (dsDNA) by the observation that the addition of dsDNA to extracts from a range of eukaryotic cells induced the phosphorylation of several proteins (Walker et al., 1985). This led to the partial purification of a DNA-dependent protein kinase, of which the phosphorylation was dependent upon a >300 kDa polypeptide that was named the DNA-PK catalytic subunit (DNA-PKcs) (Carter et al., 1990). This subunit activity was found to be dependent upon Ku, which had been shown to preferentially bind to dsDNA ends and led to the notion that DNA-PK may be involved in DNA repair (Gottlieb and Jackson, 1993; Mimori et al., 1981; Mimori and Hardin, 1986).

The subsequent identification of Ku and DNA-PK as the causative factors in radiosensitive, DSB-repair-deficient rodent cell lines (Getts and Stamato, 1994; Peterson et al., 1995) established it as an essential factor for DNA repair. It is now established that DNA-PKcs and the Ku heterodimer are essential for the most common pathway for the repair of double strand breaks by direct joining of DNA ends in the process called non-homologous end joining (NHEJ).

The formation of the DNA-PK complex is dependent upon the initial binding of Ku to DSBs, where it recruits DNA-PKcs which promotes DNA end tethering (Graham et al., 2016) and the recruitment of additional NHEJ factors for repair (detailed in Chapter 1.3.4.1).

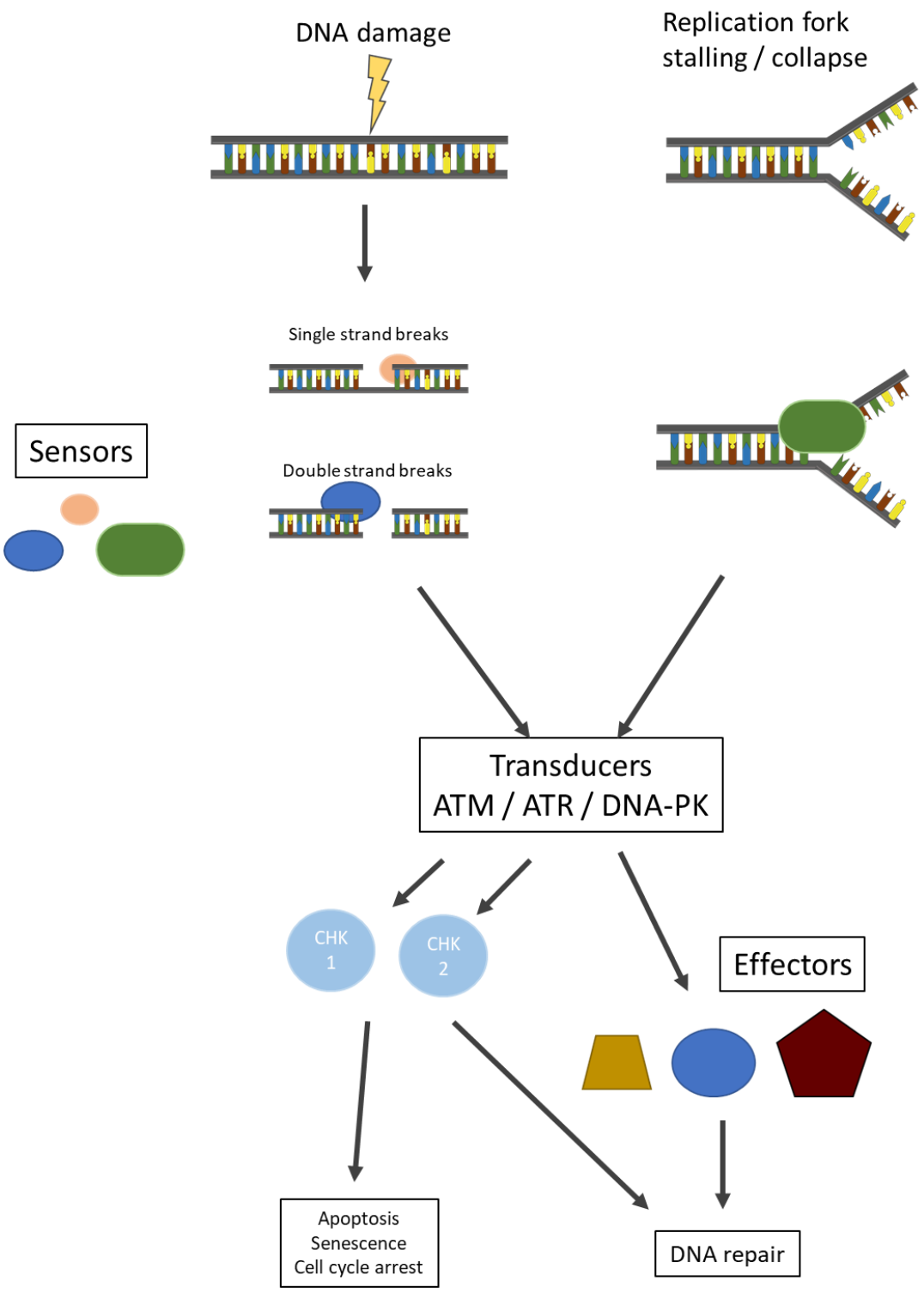


Figure 1.1: An overview of the DNA damage response

The cellular response to DNA damage forms a signalling cascade starting with sensors that activate transducers which propagate the signal to effectors.

1.4 DNA repair pathways

When damage to DNA accumulates in a cell of a multi cellular organism, the DDR can signal to permanently end differentiation and replication by moving the cell towards senescence or by entering a program of cell death. Both routes prevent potentially deleterious and mutated DNA sequences being propagated into future lineages or generations.

However, this is generally the last resort and cells will normally attempt to repair DNA lesions to enable the continuation of normal cellular function and replication. The exact repair response that is enacted is dependent upon the type of DNA damage.

1.4.1 Mismatch repair and base excision repair

Simple mispairing of bases can occur through insertion, deletion or mis-incorporation during replication or recombination processes. These types of lesions occur commonly in all organisms and are repaired through DNA mismatch repair (MMR), which is highly conserved from prokaryotes to eukaryotes. MMR is activated through the binding of MutSα and MutSβ to substrates containing mismatches to recruit MutL which then activates downstream repair (Jiricny, 2006).

Small chemical alterations of DNA bases that cause mild disturbances to the helical DNA structure are repaired through the highly conserved base excision repair (BER) pathway. BER responds to absent bases, or bases that have been chemically modified in several ways, including deamination, oxidization, alkylation and even absent bases. DNA

glycosylases recognize these abnormal bases and hydrolyse the bond linking the base to the sugar-phosphate backbone. This generates an abasic, or apurinic/apyrimidinic site (AP-site) in DNA, which can then be further processed by an AP-endonuclease, an exonuclease, a DNA polymerase and a ligase to restore the original DNA sequence (Jacobs and Schär, 2012).

1.4.2 Nucleotide excision repair

More complex lesions that disturb the normal structure of the DNA double helix are normally repaired by the nucleotide excision repair (NER) pathway. One of the main characteristics of the NER pathway is that because it detects distortions in the DNA structure, rather than relying on a more specific enzymatic recognition via a lesion-binding pocket, it can recognize a very wide range of substrates (Schärer, 2013).

In general, NER initiates repair by detecting bulky lesions followed by organizing the excision of a short single-strand of DNA that contains the lesion. The undamaged strand thereafter serves as a template for DNA polymerase to synthesize a complementary strand that is finally ligated to complete repair (Figure 1.2).

NER is divided into two sub-pathways that are initiated depending upon where in the genome the DNA damage occurs. Global-genome NER (GG-NER) can repair DNA anywhere in the genome (Petruseva et al., 2014), whereas transcription-coupled NER (TC-NER or TCR) responds rapidly to lesions that occur in the transcribed strand of active

genes that block the progression of RNA polymerase (RNAP) (Hanawalt and Spivak, 2008). Whilst both pathways utilize the core NER factors to perform the excision process, they differ in how they detect damage to initiate repair.

1.4.2.1 Transcriptional-coupled nucleotide excision repair

TC-NER is initiated by the obstruction of RNAP progression, and the DDR signal originates from RNAP itself, which is a highly sensitive and specific detector of damage (Lindsey-Boltz and Sancar, 2007). Whilst the exact molecular mechanisms of TC-NER are not well defined, the rescue of RNAP and subsequent repair of DNA requires proteins implicated in the human disease, Cockayne Syndrome (CS).

The current model of TC-NER proposes that the Cockayne syndrome B (CSB also known as ERCC6) protein travels with the elongation complex and transiently interacts with RNAP to ensure RNA synthesis is completed properly (Tantin et al., 1997). When transcriptional arrest occurs due to DNA damage, RNAP signalling causes this interaction to be stabilised and prolonged during which downstream repair is orchestrated (van den Boom et al., 2004).

The involvement of TC-NER is well established for RNA polymerase II, which generates all the different protein-coding messenger RNA (mRNA) from thousands of different genes as well as many non-coding RNAs (ncRNAs) such as small nuclear- (snRNA) and micro RNAs (miRNA).

Mammalian cells also encode two other RNAPs, RNAP I and RNAP III, that have more specialized functions. RNAP III is responsible for the transcription of some ncRNAs but the most abundant products are transfer RNAs (tRNAs) that, like RNAP I products, are arranged into tandem repeats and fundamental components of the cellular protein synthesis machinery (Dieci et al., 2007). These RNAP III transcribed genes are small in comparison to RNAP II and there is a lack of evidence for a similar TC-NER response in these transcribed areas of the genome (Dammann and Pfeifer, 1997).

RNAP I is responsible for the transcription of ribosomal DNA (rDNA), which encodes ribosomal RNA (rRNA) that comprises the main structural and catalytic component of the ribosome. Synthesis of rRNA occurs in the nucleolus, where rDNA is organised as large tandem repeats that are transcribed throughout the cell cycle (Richard et al., 2008). In contrast to RNAP III, there are several studies that support a transcription-coupled DDR at active rDNA genes, although its mechanism remains unclear.

In yeast, TC-NER has been implicated in the repair of UV-induced cyclobutane-pyrimidine dimers (CPD) in transcribed strands of active rDNA genes (Conconi et al., 2002; Meier et al., 2002), suggesting a similar signalling response from RNAP I upon stalling at DNA damage. In mammalian cells, CSB is reported to be present in the nucleolus in complex with RNAP I and can rescue rDNA synthesis when transfected into CSB-deficient cells (Bradsher et al., 2002).

1.4.2.2 Global-genome nucleotide excision repair

GG-NER is initiated by the heterotrimeric XPC complex comprised of XPC, RAD23B and Centrin-2, which binds DNA that has lesions which compromise the proper base pairing of opposite or adjacent bases (Araki et al., 2001). XPC appears to bind the strand that lies opposite to the lesion, indicating it detects the helix distortion rather than the lesion itself and explains its capacity to address a wide range of lesions (Lee et al., 2014). Certain UV-induced lesions such as CPD that do not dramatically distort the duplex are first recognized by the UV-DDB complex, but this ultimately facilitates XPC loading onto DNA (Matsumoto et al., 2015). Whilst XPC can detect these lesions, it does so with substantially lower affinity and UV-DDB is the first of the two to accumulate at these sites of damage (Wakasugi et al., 2002).

At this point, the GG- and TC-NER pathways converge on the recruitment of protein complexes that include helicases and nucleases to excise the lesion, followed by repair and ligation. This first requires the recruitment of transcription factor II (TFIIH), a ten-subunit complex that includes the 5'-3' helicase, XPD, which opens a nucleotide bubble 20-30 base pairs in length that encompasses the lesion (Coin et al., 2007). Following unwinding, RPA, XPA and XPG are recruited.

RPA is a major ssDNA binding protein that is required for the protection of ssDNA intermediate structures found in processes during repair, replication and recombination. In NER, RPA binds the ssDNA opposite the lesion and protects it from

degradation whilst also serving as a marker that helps co-ordinate the excision and repair events (Schärer, 2013; Spivak, 2015).

XPA binds at the 5' side of the bubble and interacts with a number of other factors including the endonuclease ERCC1-XPF that, along with the other endonuclease XPG, perform the incision step of NER. XPG interaction with TFIIH provides some structural support, but does not act as an endonuclease until the first incision is made by ERCC1-XPF, which acts prior to XPG and provides it with a viable substrate (Fagbemi et al., 2011; Hohl et al., 2003; Schärer, 2013).

These dual incisions occur either side of the lesion and allow the release of the damaged strand, allowing access for DNA replication machinery to synthesize a new strand by using the complementary, undamaged strand as a template. Repair is completed when a final ligation step is performed by either ligase I in replicating cells, or ligase IIIα (with its cofactor XRCC1) in non-replication cells (Schärer, 2013; Spivak, 2015).

Deficiencies in proteins involved in NER result in several inherited human diseases, including Xeroderma pigmentosa (XP), Cockayne syndrome (CS), trichothiodystrophy (TTD) and Fanconi anaemia (FA). XP genes encode proteins that are widely involved in NER and comprise seven complementation groups from A through to G (XPA to XPG). XP patients have heterogenous clinical features that include severe sensitivity to UV light, skin abnormalities and a significant increase in susceptibility to cancers, particularly skin cancers.

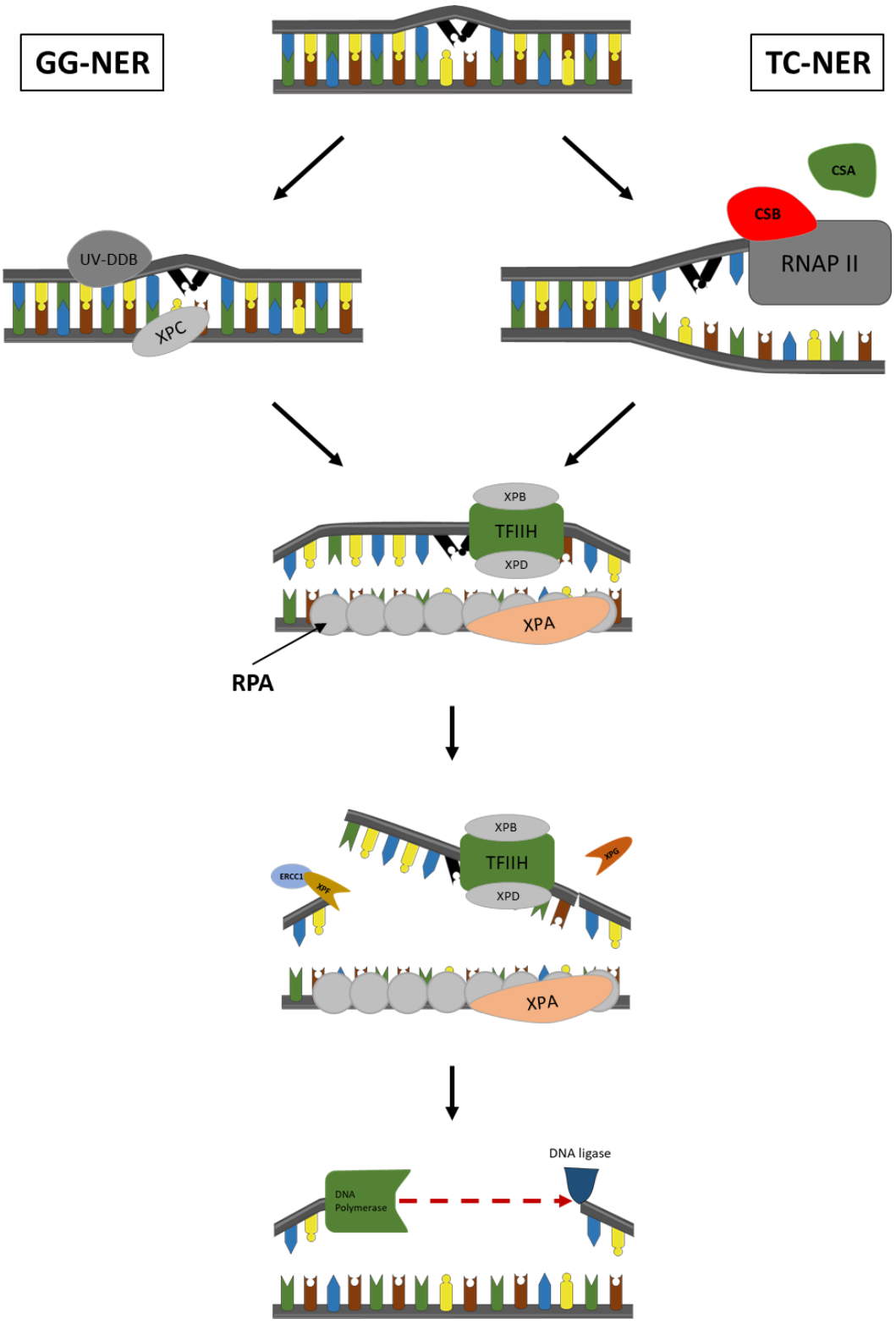


Figure 1.2: An overview of GG-NER and TC-NER

Helix distorting lesions are recognized by XPC in non-transcribed regions resulting in GG-NER, or by the stalling of RNAP in TC-NER.

The XP complementation groups occur at different frequencies and severities, with XPA and XPC being the most frequent (Hengge and Emmert, 2008). XPF exemplifies the wide range of contexts that NER machinery is involved in, possessing a role in multiple DNA repair processes. Mutated XPF can cause several NER-related diseases and one recorded individual with an XPF mutation exhibits clinical features of three different disorders; CS, XP and FA (Kashiyama et al., 2013).

1.4.3 Double strand break repair

DNA double strand breaks (DSBs) are considered to be one of the most harmful DNA lesions and are a significant threat to the integrity and stability of the genome. The potent toxicity of these lesions has been shown in yeast, where a single unrepaired DSB in a dispensable plasmid can be sufficient to arrest growth and trigger cell death (Bennett et al., 1993). The formation of DSBs can lead to severe chromosomal rearrangements such as translocations, deletions and amplifications. These gross chromosomal abnormalities can drive cellular progression to malignancy.

Although repair of DSBs can occur through several mechanisms, cells typically employ one of two mechanisms, which differ in their ability to retain the fidelity of the underlying DNA sequence. The most common pathway that repairs the majority of DSBs is classical non-homologous end joining (cNHEJ) which, as mentioned previously, is dependent upon the binding of DSB ends by Ku70/80 and the recruitment of DNA-PKcs.

1.4.3.1 Classical non-homologous end joining

cNHEJ occurs throughout the cell cycle and is the dominant DSB repair pathway in G0 and G1 phases and acts to quickly ligate DSB blunt ends together, regardless of sequence homology (Figure 1.3). This has the benefit of quickly securing DNA ends in their approximate genomic loci and preventing severe chromosomal translocation and other rearrangements (Difilippantonio et al., 2000).

cNHEJ can proceed with no modification of DNA ends, but a small amount of end processing may be required for cNHEJ to perform the ligation, including removal of damaged nucleotides. This end processing can result in additions or deletions of nucleotides and coupled with the non-specific nature of end joining it was long assumed that repair through cNHEJ resulted in a substantial number of sequence alterations (Bahmed et al., 2011). However, with the emergence of alternate end-joining as a distinct process from cNHEJ, a more recent study suggests that error rate is not inherent to the cNHEJ machinery but dictated by structure of DNA ends (Betermier et al., 2014).

After DSB binding by Ku70/80 and DNA-PKcs, additional cNHEJ factors, such as PAXX, XRCC4, XLF and DNA ligase IV are recruited in order to closely align DNA ends to facilitate their ligation (Ahnesorg et al., 2006; Grawunder et al., 1997; Li et al., 1995; Ochi et al., 2015).

1.4.3.2 Homologous recombination

In contrast to cNHEJ, the homologous recombination (HR) repair pathway requires the presence of homologous duplex DNA to serve as a template for DNA synthesis (Figure 1.3). As such it is restricted to mid to late S and G2 phases of replicating cells, where sister chromatids are available to serve as a template (Heyer et al., 2010). HR involves significant processing of DSB ends, requiring the resection of the strands by exonucleases to generate long 3' ssDNA overhangs. Resection is a complex and highly regulated process that involves several nucleases including the MRN complex, BLM, CtIP, EXO1 and DNA2 (Symington and Gautier, 2011).

After the generation of the 3' ssDNA by exonucleases the overhangs become coated in RPA to protect against further degradation, prevent the formation of DNA secondary structures and antagonizes alternative, resection-dependent and error prone repair (Chen et al., 2013; Deng et al., 2014). RPA is subsequently displaced by the recombinase protein RAD51 to form a RAD51-ssDNA nucleoprotein complex that is referred to as the presynaptic filament (Zhao et al., 2017). The loading of RAD51 onto DNA and displacement of RPA involves the breast cancer susceptibility proteins BRCA1 and BRCA2, PALB2 as well as several RAD51 paralogues (Daley and Sung, 2014; Liu et al., 2010).

Once RAD51-ssDNA nucleoprotein filaments are formed it engages in a search for homologous duplex DNA to carry out strand exchange. How these recombinases from the RAD51 family search through the vast quantities of DNA contained within the cell

and find appropriate stretches of DNA to act as a template has remained enigmatic since the inception of HR. However, one strand of the donor duplex DNA is destabilized by the presynaptic complex, opening the duplex and allowing bases to be sampled for homology. Yeast (*S. cerevisiae*) Rad51 homology search requires a minimum length of 8 nucleotides of microhomology for stable dsDNA binding (Qi et al., 2015).

Once appropriate homology is found, a displacement loop (D-loop) is formed, and DNA synthesis can proceed to replace the DNA that was lost around the former break site. The displaced strand of the D-loop can provide a template for further DNA synthesis in the second 3' resected end. This process results in an intermediate structure known as Holliday junction, the dissolution of which is depend upon topoisomerase IIIα (TOPIIIα) in complex with BLM and hRMI1 (Bizard and Hickson, 2014; Wu and Hickson, 2003).

Ultimately, the completed process of HR mediated repair should leave two completed strands that have retained the correct sequence that existed before the DSB occurred.

1.4.3.3 DSB repair pathway choice

Since both HR and cNHEJ are both able to repair DSBs, these processes are often vying for the repair of the same lesion. This competition between pathways for the same lesions is evident from measuring HR rates in Ku70-, XRCC4- and DNA-PKcs-deficient cell lines where repair via HR is elevated in comparison to wildtype cells (Pierce et al., 2001).

The resection of DSB ends appears to be the critical determinant for repair pathway choice, as resection results in greatly reduced efficiency of cNHEJ and commits the break to repair by other means (Chapman et al., 2012). Accordingly, the resection process is heavily regulated and upstream DNA repair proteins act to secure DSB ends in order to facilitate or prevent resection.

One such protein is 53BP1, named originally as a binding partner of the tumour suppressor p53 (Iwabuchi et al., 1994). Despite being uncovered over two decades ago the significance of its association with p53 is not well defined, however 53BP1 is understood to be central to cellular response to DSBs and repair pathway choice. Upon detection of DSBs, 53BP1 is phosphorylated by ATM and rapidly accumulates on chromatin that surrounds the break site forming microscopically visible nuclear foci (Anderson et al., 2001; Schultz et al., 2000).

During G1, the 53BP1 protein along with its effectors, RIF1 and PTIP, are required for cNHEJ and act to prevent the resection of DSB ends thereby promoting NHEJ (Chapman et al., 2013; Zimmermann et al., 2013). DSB ends are also protected by a multi-protein complex called Shieldin (composed of RINN1/2/3 and REV7), members of which have important interactions with 53BP1 and are required for NHEJ (Gupta et al., 2018).

In S/G2 phases of the cell cycle RINN1, 53BP1 and RIF1 are antagonized by BRCA1 and CtIP. This interference with DSB end protection allows access to exonucleases and for repair to completed by resection dependent processes (Daley and Sung, 2014).

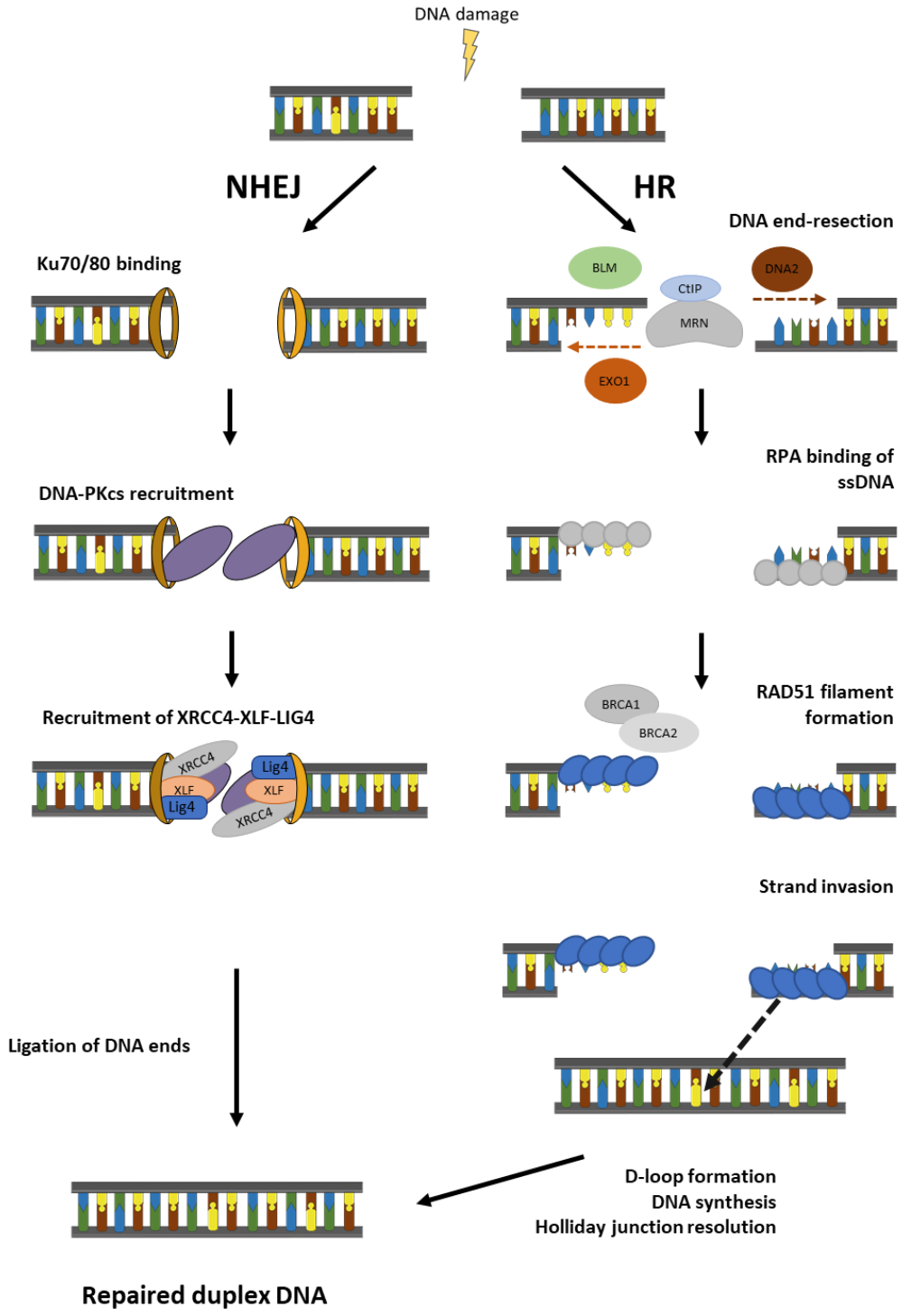


Figure 1.3 DSB repair by cNHEJ or HR

The two main modes of DSB repair differ in their processing of DNA ends. cNHEJ undertakes little end processing to ligate ends together. HR repair involves substantial resection of ends followed by homology search for a template for DNA synthesis.

1.4.4 Cross-link repair

Cross-linking of DNA occurs when covalent bonds form between nucleotides. Some types of DNA cross-links have been mentioned previously – such as thymine dimers that can be caused by UV irradiation. These are intrastrand cross-links, that form between nucleotides on the same strand and are primarily repaired via NER in mammals. Covalent bonds that form links between opposite strands of duplex DNA are referred to as interstrand cross-links (ICLs). These are complex lesions that affect both strands and prevent the unwinding of DNA, causing complications for essential aspects of cellular metabolism that require the separation of strands such as transcription and replication.

1.4.4.1 Sources of ICLs

Many chemotherapeutic drugs used in the treatment of cancer target the fast proliferation of malignant cells by reacting with DNA to form adducts and invoke the DDR (Lord and Ashworth, 2012). Several of the commonly used agents are ICL-inducing compounds, that introduce DNA cross-links to hinder replication sufficiently to trigger apoptosis. The use of cross-linking agents in the treatment of cancer dates back to world war II, where exposure of civilians and soldiers to mustard gas indicated the lymphotoxic effect of these agents, leading to clinical trials in the treatment of lymphoma patients with nitrogen mustards (Kohn, 1996). Nitrogen mustard derivatives remain a mainstay of antitumour therapy and have been shown to form ICLs by reacting with the N^7 position of two guanines residues on opposite DNA strands (Guainazzi and Schärer, 2010).

However, ICL forming agents are not limited to exogenous compounds used in cancer chemotherapy as some by-products of cellular metabolism can form ICLs. Aldehydes are pervasive in the environment and a major component in dietary sources as well as common by- and end-products of various cellular processes (O'Brien et al., 2005). Members of this class of compounds, such as those produced as products of lipid peroxidation like acrolein, crotonaldehyde and malondialdehyde have been shown to be able to induce ICLs (Niedernhofer et al., 2003; Stone et al., 2008), but the extent to which they occur in cellular DNA is unclear.

1.4.4.2 Repair of ICLs

Bacteria and eukaryotes have developed mechanisms of ICL repair that generally requires the concerted effort of factors involved in several repair pathways including NER, translesion synthesis (TLS) and HR. This repair has been described through two distinct mechanisms, which are intrinsically tied to the stage of the cell cycle in which the lesion is detected.

The best characterised response occurs during replication, but there is evidence of a replication-independent response to ICLs that likely occurs when ICLs block transcription and requires the action of TC-NER and the CSB protein (Enoiu et al., 2012). Replication-dependent ICL repair has emerged as an intricate network of proteins that have mostly been characterised by the study of a rare inherited disease called Fanconi anaemia (FA; Figure 1.4). Whilst there are putative homologs of the major members of this pathway in bacteria and yeast, a larger and more complex pathway is present in

higher eukaryotes (McHugh et al., 2012). FA factors are now categorized into a complex pathway that detects and repairs ICLs when replication forks collide with them during S-phase.

1.5 The Fanconi anaemia pathway

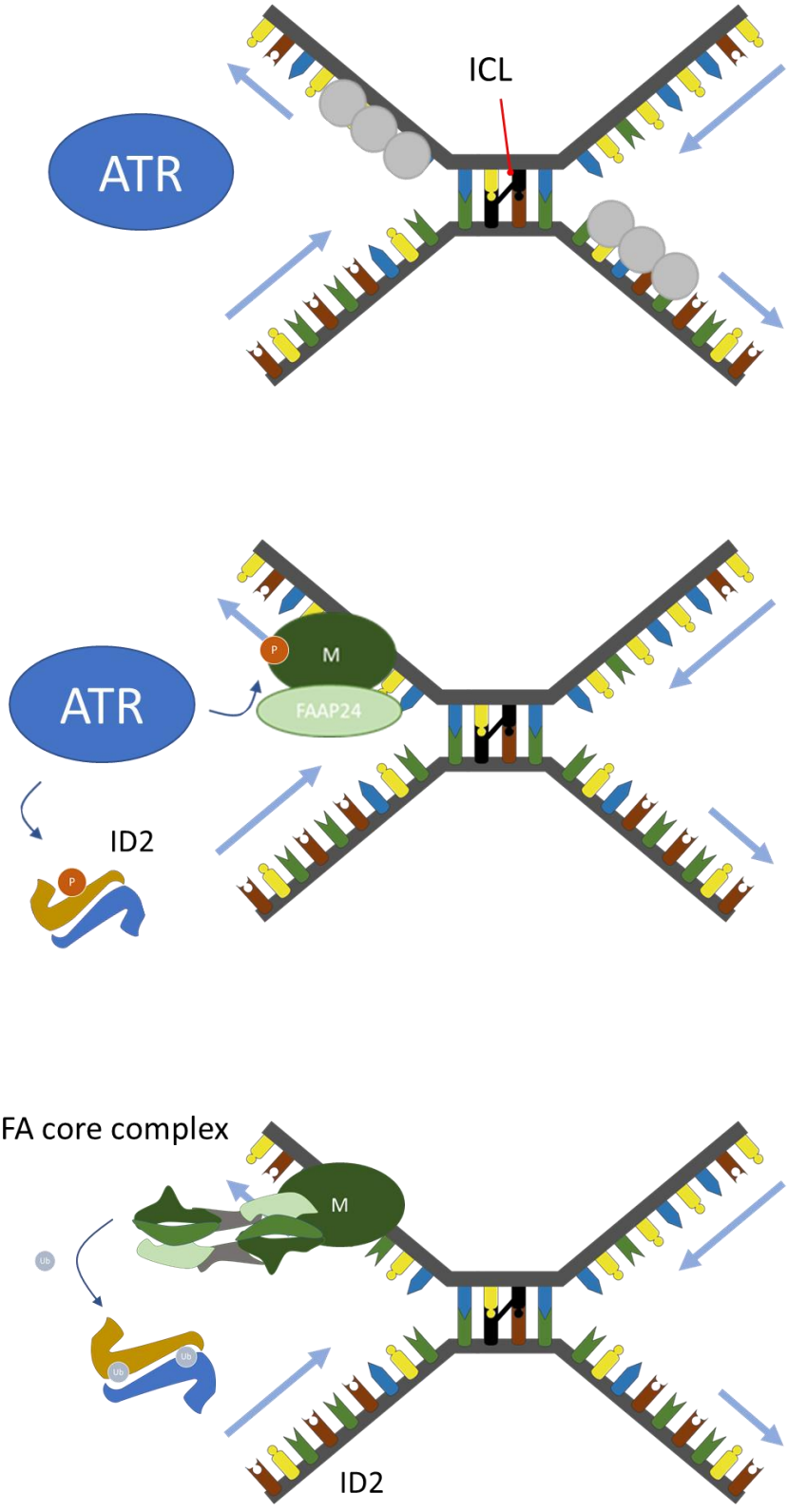
Fanconi anaemia (FA) is a rare genetic disorder characterised by increased genome instability, early onset haematological complications, developmental abnormalities and a predisposition to cancer. It is caused by the biallelic loss of one of 22 proteins (Organised in complementation groups FANCA-FANCW; See Table 1.1) which constitute a pathway that is commonly referred to as the FA-BRCA pathway, owing to the direct involvement of BRCA1 and BRCA2, the breast cancer susceptibility genes (Ceccaldi et al., 2016). The mode of inheritance for all subtypes of FA is autosomal recessive, apart from FA-B, which is X-linked (Meetei et al., 2004).

FA occurs at around 1/100,000 births, but significantly higher in certain populations such as Ashkenazi Jews (1/30,000) and Afrikaners (1/22,000), making it one of the most common inherited bone marrow failure syndromes (Tipping et al., 2001; Whitney et al., 1993; Yagasaki et al., 2003). Diagnosis is not always straightforward, as the clinical symptoms of FA are diverse, and some patients present with few or mild abnormalities. Many FA patients remain undiagnosed until the development of haematological irregularities, usually in the latter half of the first decade of life, where all mature haematopoietic lineages can eventually be comprised (pancytopenia) owing to the progressive depletion of progenitor cells (Auerbach, 2009).

Cells derived from patients with FA are sensitive to DNA damaging agents, particularly to those that introduce ICLs. Exposure of these cells to compounds such as mitomycin C (MMC) results in observable chromosome breakage (Sasaki and Tonomura, 1973), that is substantial enough to serve as a diagnostic marker (Castella et al., 2011).

It is increasingly understood that the FA pathway and its individual members have multifaceted roles in a wide array of genome stability and DNA repair processes. However, much of the focus of research concerning the FA pathway has been on its role in the cellular response to ICLs which is fundamental to better understanding the treatment and outcomes of cancer with chemotherapeutic agents.

The FA pathway is divided into three groups; the core complex, the heterodimer FANCI-FANCD2 (ID2) complex and downstream repair factors. In general, the FA pathway responds to ICLs by the formation of a multi-protein complex, referred to as the core complex, which acts as an E3 ligase to monoubiquitinate two other members of the pathway, FANCD2 and FANCI. These form the ID2 complex, which subsequently orchestrates repair via DNA incision and lesion unhooking followed by lesions bypass and finally repair.



(Continued overleaf)

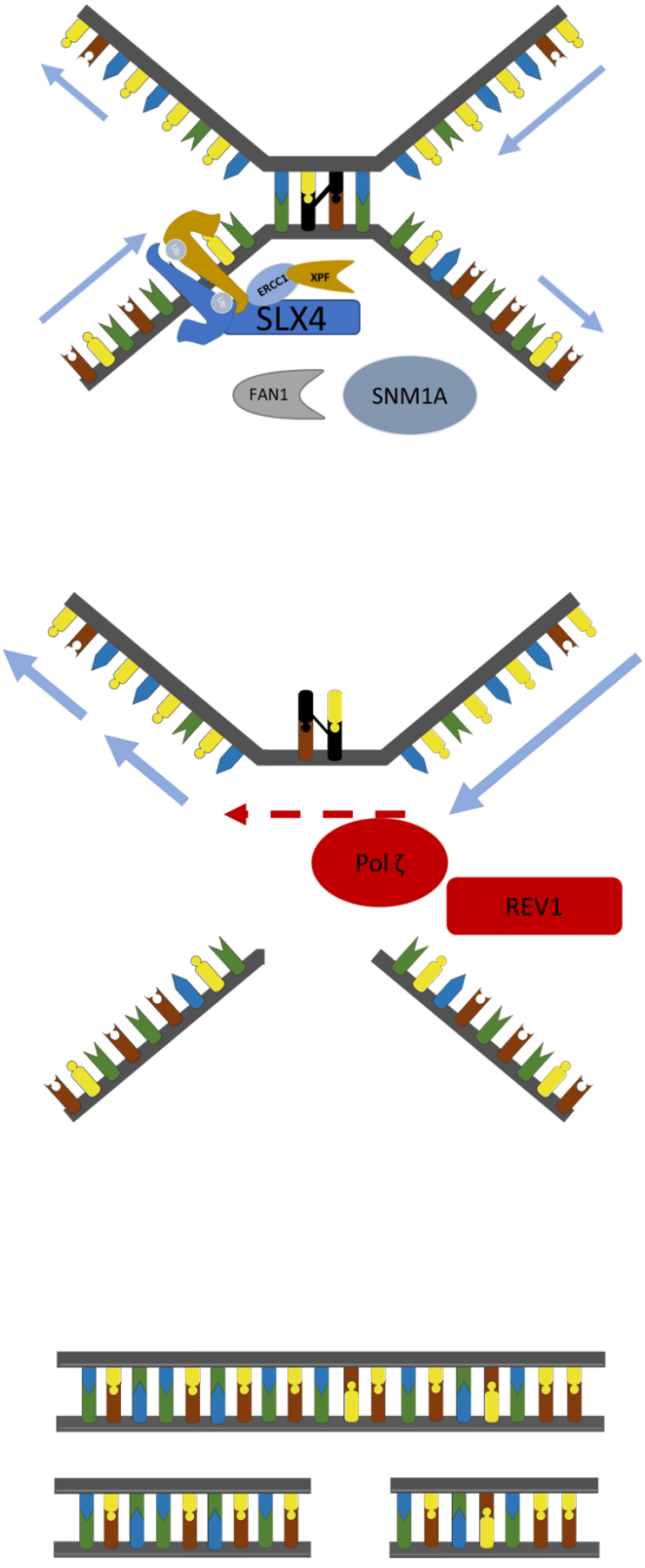


Fig 1.4: An overview of ICL repair by the FA pathway

Complementation Group	Aliases	FA Pathway Function
FANCA		Core complex
FANCB		Core complex
FANCC		Core complex
FANCD1	BRCA2	Downstream; HR and fork stabilization
FANCD2		Binds FANCI; monoubiquitinated
FANCE		Core complex
FANCF		Core complex
FANCG	XRCC9	Core complex
FANCI		Binds FANCD2; monoubiquitinated
FANCJ	BRIP1	Downstream; Helicase, HR and TLS
FANCL		Core complex; E3 Ligase
FANCM		Core complex; Damage recognition, translocase
FANCN	PALB2	Downstream; HR
FANCO	RAD51C	Downstream; HR
FANCP	SLX4	Downstream; DNA repair scaffold
FANCQ	XPF, ERCC4	Downstream; DNA repair, endonuclease in NER
FANCR	RAD51	Downstream; HR and fork stabilization
FANCS	BRCA1	Downstream; HR and fork stabilization
FANCT	UBE2T	Core complex; E2 Conjugating
FANCU	XRCC2	Downstream; HR
FANCV	REV7	Downstream; TLS
FANCW	RFWD3	HR; ubiquitin ligase

Table 1.1 Members of the FA pathway

1.5.1 FA pathway activation in response to ICLs

The activation of the FA pathway response is similar to other DDR pathways, in that it requires recognition of a lesion followed by multi-step activation via post-translational modification.

Detection of ICLs during S-phase is dependent upon the convergence and subsequent stalling of two replication forks at the ICL, resulting in replication stress and the generation of local ssDNA (Zhang et al., 2015). These stretches of ssDNA are bound by RPA and ATRIP, resulting in ATR recruitment and checkpoint activation (Zou and Elledge, 2003). ATR signalling results in the phosphorylation of several FA proteins, including FANCM (Andreassen et al., 2004; Pichierri and Rosselli, 2004; Singh et al., 2013; Wang et al., 2007).

FANCM is a DNA-binding protein with translocase motifs that has FA-independent functions in replication fork protection and ATR signalling (Collis et al., 2008; Schwab et al., 2010). In response to ICL induced replication stress, FANCM is phosphorylated by ATR and in conjunction with its binding partner FAAP24, it anchors the FA core complex to DNA (Kim et al., 2008). FANCM's translocase function is vital for the remodelling of replication fork structures to promote fork regression (Luke-Glaser et al., 2010; Schwab et al., 2010)

The central step in the activation of the FA pathway is the monoubiquitination of the FANCD2-FANCI heterodimer by the core complex. Through genetic approaches and

protein purification the current understanding of the core complex is that it is composed of eight FA pathway members (FANCA, B, C, E, F, G, L and M) and five associated proteins (FAAP; Fanconi anaemia-associated protein) of which no patient mutations are yet described (FAAP20, FAAP100, FAAP24 and FAAP10) (Ceccaldi et al., 2016). This core complex, in conjunction with the E2 ubiquitin conjugating enzyme UBE2T, assigned FANCT (Hira et al., 2015), acts as a large ubiquitin ligase to monoubiquitinate both FANCI and FANCD2 interdependently.

The monoubiquitination of FANCD2 was identified as the activation step of pathway before the E3 ligase was identified. BRCA1 was first suggested to be the enzyme responsible for this due to its association with FANCD2 in foci and the possession of a ring domain with known ubiquitin ligase activity (Folias et al., 2002; Garcia-Higuera et al., 2001). However, BRCA-1 mutant cells displayed low levels of FANCD2 ubiquitination and further studies definitively showed BRCA1 was not required for monoubiquitination of FANCD2, but did play a role in the localization of FANCD2 to sites of damage (Vandenberg et al., 2003). Eventually, discovery of PHF9 as the gene responsible for Fanconi anaemia group L, which has E3-ubiquitin ligase activity, identified FANCL as the E3 ligase (Meetei et al., 2003).

Recent biochemical purification and *in vitro* analysis of some of these core complex proteins suggests that interactions within the core complex allow it to be further divided into three subcomplexes (van Twest et al., 2017). A subcomplex comprised of FANCB:FAAP100:FANCL forms a dimer in which FANCB assists with the spatial

arrangement of two FANCL molecules. A second subcomplex of FANCC:FANCE:FANCF acts as a bridge between the first complex and both FANCI and FANCD2, coordinating the ubiquitination of both proteins in the ID2 heterodimer. A third sub-complex that appears to be dispensable for the ubiquitination process is composed of FANCA:FANCG:FAAP20 but does show an ability to bind the FANB:FAAP100:FANCL complex.

The formation of the core complex appears to function solely to activate the ID complex, as FANCD2 and FANCI were the only targets whilst no other potential substrates have (so far) been identified, including in a mass spectrometry-based screen (Renaudin et al., 2014). The individual members of the core complex may have other independent functions, but few biochemical activities have been ascribed due to the lack of evolutionarily conserved domains. The deubiquitination of the ID complex is also integral for a proper ICL repair response and is performed by the ubiquitin-specific protease USP:UAF1 (Nijman et al., 2005). The importance of this process is evident as knockout USP1 or UAF1 in mice results in an FA-like phenotype (Kim et al., 2009).

The eviction of the ID2 complex from DNA is also an important regulatory step and is under the control of post-translation modification by SUMO E3 ligases and depends upon ATR. SUMOylated ID2 is then polyubiquitinated by a SUMO-targeted ubiquitin ligase, promoting the removal of ID2 from the site of damage (Gibbs-Seymour et al., 2015).

Ultimately, the formation and localization of the core complex to DNA results in the phosphorylation of the ID2 complex followed by ubiquitination, which allows FANCD2 to bind DNA at sites of DNA damage (Andreassen et al., 2004; Ishiai et al., 2008).

1.5.2 FA mediated removal of ICLs and repair of DNA

Once the ubiquitinated FANCD2 is loaded onto chromatin it recruits SLX4 (assigned FANCP), which associates with activated FANCD2 through its ubiquitin-binding zinc-finger domain (Yamamoto et al., 2011). SLX4 acts as a nuclease scaffold, recruiting and activating several structure specific endonucleases, including the ERCC1-XPF (XPF is assigned FANCQ) heterodimer which promotes the unhooking incision, allowing access to the SNM1A exonuclease to digest past the crosslink (Abdullah et al., 2017; Klein Douwel et al., 2014). This process completes the unhooking and introduces a DSB in one of the two duplexes.

Unhooking leaves the crosslinked nucleotide attached to the other strand, which is then bypassed by TLS polymerases which possess a larger binding pocket than replicative helicases that facilitates bulky or modified bases such as ICL lesions. After the ICL is bypassed, extension of opposite strand occurs via REV1 and DNA polymerase ζ and results in the restoration of one DNA duplex (Budzowska et al., 2015). The TLS process is vital to the overall ICL repair (including replication-independent) as mutations in DNA polymerase ζ subunits or REV1 results in hypersensitivity to ICL inducing agents (Sarkar et al., 2006; Suzuki et al., 2016).

With one restored DNA duplex, the DSB break caused by the nucleolytic incision step can be repaired through homologous recombination repair, of which many proteins involved are also implicated in FA (FANCJ, BRCA1/2, RAD51; See table 1.1).

Repair of the double strand break by HR is promoted by FANCD2 through direct interaction with CtIP, tethering it to damaged DNA, promoting its ability to stimulate resection and thereby committing DSB repair to homologous recombination (Murina et al., 2014; Unno et al., 2014). In addition, FANCD2 suppresses NHEJ by preventing 53BP1 association through its interaction with the acetyltransferase TIP60 (Hejna et al., 2008). Localization of TIP60 to chromatin allows it to acetylate histone 4 (H4K16), blocking 53BP1 access to its docking site, H4K20me2 (Renaud et al., 2016).

1.5.3 Replication fork protection by the FA pathway

As well as orchestrating the repair of ICLs, the FA pathway has a role in suppressing genomic instability upon fork stalling that is not dependent upon ICL formation. Hydroxyurea (HU) is a ribonucleotide reductase that inhibits replication by the depletion of deoxyribonucleotide pools (Snyder, 1984) and treatment of cells with HU appears to activate FA pathway and factors without introducing ICLs (Chen et al., 2016; Howlett et al., 2005). Sensitivity to HU is not due to a defect in general HR repair, as FA patients are not severely defective in the repair of DSBs (Nakanishi et al., 2011).

The protection of stalled forks is an important factor in maintaining genome stability as unscheduled nucleolytic degradation can result in loss of sequence, whilst prolonged

stalling leads to collapse of forks and the formation of DSBs (Petermann et al., 2010a). Several studies have implicated FA factors at sites of damage and stalled replication forks with proposed roles in preventing degradation and promoting fork restart.

FANCD2 and FANCI are enriched at HU stalled forks, as ascertained by iPOND (isolation of protein on nascent DNA) screening, which detects proteins bound to nascent DNA in the vicinity of the replisome (Lossaint et al., 2013). Stalled forks contain intermediate replication structures that are liable to be degraded by the action of nucleases. Instability of nascent DNA in cells deficient in FA factors RAD51, BRCA1/2, or FANCD2 can be rescued by the depletion or inhibition of MRE11, indicating they act to prevent unscheduled degradation (Schlacher et al., 2011; Schlacher et al., 2012).

The restart of stalled forks after blockage of DNA synthesis is promoted by FANCD2 through its recruitment or interaction with BLM and FAN1. FAN1 is a nuclease that is recruited to ICLs in S-phase by ubiquitinated FANCD2 and is also recruited to non-ICL stalled forks (Shereda et al., 2010). FANCD2, independent of its binding partner FANCI, interacts with FAN1 and BLM to both promote fork restart and suppress origin firing (Chaudhury et al., 2014; Lachaud et al., 2016). This function of fork restart and suppression of origin firing is also shared by BRCA2 and FANCJ (Chaudhury et al., 2013).

1.5.4 The association of the FA pathway and reactive aldehydes

The extreme sensitivity of FA deficient cells to exogenous ICL inducing agents such as those used in chemotherapy fuelled the notion that the FA pathway had evolved to respond to, and FA is caused by, endogenously produced ICLs.

As mentioned in 1.4.5.1, reactive aldehydes are common end- and by-products of metabolism that can introduce a range of DNA lesions, including ICLs, but until recently no link was made between aldehydes and FA. Substantial evidence has now accumulated that in the absence of formaldehyde or acetaldehyde detoxifying enzymes (or when these detoxification pathways are overwhelmed), DNA damage accumulates that requires the action of the FA pathway.

One of the early indications of this association was that avian cells that are deficient in FA proteins were highly sensitive to low doses of formaldehyde (Ridpath et al., 2007). This was followed by a substantial amount of research into the genetic relationship between aldehyde detoxifying pathways and FA in mice. Deletion of acetaldehyde or formaldehyde detoxifying enzymes (*Aldh2* and *Adh5*, respectively) in conjunction with *Fancd2* deletion results in a substantial reduction in viability (Langevin et al., 2011; Rosado et al., 2011). Any double *Fancd2* and *Aldh2/5* deficient mice that were viable mostly succumbed to leukaemia and those that did not developed spontaneous bone marrow failure (Garaycoechea et al., 2012). Until this point, mouse models of FA had not been able to recapitulate the human phenotype, and this was the first instance of

FA deficient mice displaying a profound reduction in haematopoietic stem cells, as observed in FA patients (Auerbach, 2009).

Although aldehydes can induce the formation of ICLs, whether they interact with DNA *in vivo* in this way remains unclear, and so the exact lesion that requires the action of FA that is introduced by aldehydes remains to be determined. Aldehydes also introduced DNA-protein cross-links (DPCs), which also require the action of some FA factors for repair and block the same basic cellular processes as ICLs, replication and transcription.

One of the characteristics of aldehydes that is shared by many of the exogenous compounds that FA deficient cells display a striking sensitivity to, is the wide range of lesions they induce. Cisplatin, for example, introduces a small proportion of ICLs in comparison to other lesions, most of which are bulky intrastrand cross-links that are more commonly associated with blocking the progression of transcription complexes and resolution by NER (Schärer, 2005). Interestingly, one of the emerging functions of the FA pathway is its response to transcription-based replication stress.

1.6 Transcription-replication conflict

Transcription has long been understood to be a source of replication stress, as these processes both require the binding and translocation of large protein complexes to DNA to facilitate genome duplication and gene expression. The transcription machinery

therefore represents an endogenous impediment to replication fork progression that must be dislodged for replication to continue unhindered (French, 1992).

In bacteria, transcription and replication occur simultaneously on the circular chromosome, without spatial or temporal separation of the processes. The bacterial replisome operates more than 10 times faster than RNA polymerase (Helmrich et al., 2013) making encounters between the two machineries inevitable. These conflicts can be resolved by pathways that rescue stalled forks (Trautinger et al., 2005) and failure to properly orchestrate this conflict can result in genome instability (Vilette et al., 1995).

Collisions between replication and transcription can occur in two orientations; head-on or co-directional. Much evidence has been produced that head-on collisions have a larger negative impact on DNA replication forks than co-directional ones (Liu and Alberts, 1995; Mirkin and Mirkin, 2005; Olavarrieta et al., 2002). Bacterial genomes appear to have adapted to this by orientating most genes co-directionally with replication origins (Brewer, 1988). However, co-directional conflicts at highly transcribed regions such as rRNA genes can also stall replication *in vivo* (Merrikh et al., 2011). These regions, along with tRNA transcription units, are physically blocked by barriers that prevent the entry of replication forks that would result in head-on collisions and allows completion of transcription before replication can occur (Brewer, 1988; Labib and Hodgson, 2007; Lopez-estrano et al., 1998).

Despite the threat of transcription-replication conflict, bacteria manage to grow rapidly under ideal growth conditions where there is a high demand for both efficient genome duplication and RNA synthesis, suggesting robust processes are in place to deal with collisions (Rudolph et al., 2007).

Eukaryotic cells have several adaptations that temporally and spatially separate replication and transcription. Firstly, replication and transcription are temporally separated by the cell cycle and active transcription predominantly occurs during the G1 phase (Meryet-Figuiere et al., 2014). Genes that are transcribed during S-phase appear to coordinate their replication and transcription to avoid overlap (Meryet-Figuiere et al., 2014; Vieira et al., 2004). For example, the single-copy gene that encodes cyclin B1 is replicated early during S-phase but is not transcriptionally active until late S-phase (Helmrich et al., 2011).

Despite these adaptations, eukaryotic replication and transcription machineries still meet under certain conditions, particularly at genes that are heavily transcribed and require the binding of multiple RNAP complexes per gene (Azvolinsky et al., 2009; Tuduri et al., 2009). Mammalian cells encode dozens of genes that are very long in length (>800 kb) and these require extensive transcription that occurs over a period of more than one cell cycle and will inevitably encounter the replisome during S-Phase (Helmrich et al., 2011).

This conflict results in long genes containing common fragile sites (CFSs), loci that are recognized as preferentially exhibiting breaks and chromosomal rearrangements (Durkin and Glover, 2007) and instability at CFSs in long genes is dependent upon the active expression of the underlying gene in replicating cells (Helmrich et al., 2011). Whether or not replication and transcription machineries physically collide is unclear, as the DNA supercoiling ahead of the complexes may preclude this, however this topological constraint also represent a significant barrier to replication fork progression (Kuzminov, 2018). The DNA-bound elongation complex is not the only structure produced during RNA synthesis that acts as a block to replication fork progression. Transcription can result in the production of long stretches of DNA:RNA hybrids, called R-loops, that are now significantly associated with replication blockage, DNA damage and genomic instability.

1.7 R-loops

R-loops are DNA secondary structures are composed of a nascent RNA strand that has annealed to the complementary template, forming a DNA:RNA hybrid that displaces that other strand of the duplex (Figure 1.5). Short DNA:RNA tracts form commonly during cellular processes where DNA is unwound, such as during replication where they prime DNA synthesis or during transcription where a ~8 base pair hybrid forms inside the active site of RNA polymerase during RNA synthesis.

However, these short DNA:RNA hybrids form within the confines of specific catalytic processes of chain elongation and are usually transient. Longer formations of DNA:RNA

hybrids are generated when nascent RNA transcripts leave elongating RNA polymerases and thread back to re-anneal to their complementary template DNA strand. These extended tracts of DNA:RNA hybrids along with the displaced duplex strand are referred to as R-loops and these structures are associated with DNA damage and genome instability (Gan et al., 2011), although R-loop formation alone does not appear to be sufficient to cause genome instability and an additional modification to chromatin is required (Garcia-Pichardo et al., 2017).

R-loop formation is facilitated by the topological features present during transcription, particularly the subtractive twisting in the wake of the elongation complex. This negative supercoiling can facilitate the transient unwinding of the DNA duplex, releasing the template DNA strand to be annealed by the nascent RNA transcript (Drolet et al., 2003). Preferential R-loop formation in this region has been highlighted in several studies that modify or deplete topoisomerase enzymes, which regulate supercoiling (Drolet et al., 1995; El Hage et al., 2010).

Where the displaced strand of DNA in the formation of an R-loop is G-rich in sequence, another secondary structure can form, termed G-quadruplex (G4) DNA (Duquette et al., 2004). G4 DNA can arise due to the ability of guanine to hydrogen bond with itself, forming planar rings of four guanines that are stabilized through Hoogsteen bonds. In the presence of a suitable monovalent cation these G-tetrads can stack with each other, forming the structures termed G-quadruplexes (Burge et al., 2006).

The minimum nucleic acid requirement for the formation of a G-quadruplex is four sets of 3-5 guanines separated by 2-4 other bases. Formation of G4 DNA structures can threaten genome stability as they represent a block to DNA and RNA polymerases and may contribute to R-loop associated damage and instability when their formation is facilitated by the formation of a hybrid (Sen and Gilbert, 1988; Tornaletti et al., 2008; Woodford et al., 1994).

Once formed, the thermodynamic stability of DNA:RNA hybrids is greater than that of DNA:DNA (Roberts and Crothers, 1992). This is thought to be due to the conformation of the hybrid being an intermediate between the A-form of dsRNA and the B-form of duplex DNA (Shaw and Arya, 2008). This overall stability means disassociating R-loops to restore duplex DNA does not occur naturally and requires the interference of catalytic enzymes, the absence of which causes R-loops to persist and accumulate (Sollier and Cimprich, 2015).

Accordingly, there are a range of ribonucleases and helicases that function to remove R-loops from DNA. Degradation of DNA:RNA hybrids can be achieved through specific cleavage of the RNA moiety, as performed by the ribonuclease H (RNase H) group of enzymes (Stein and Hausen, 1969), at least one of which are present in all organisms. Eukaryotes express two types of RNase H enzymes, RNase H1 and RNase H2, which are larger and more complex than their prokaryotic counterparts. Both of these types have the ability to degrade DNA:RNA hybrids, but have different substrate preferences with RNase H1 enzymes requiring at least four ribonucleotide-containing base pairs to be

present in a substrate, whilst RNase H2 can cleave single ribonucleotides (Cerritelli and Crouch, 2009).

R-loop formation is also prevented by the engagement of the nascent RNA by processing factors after it exits the RNA polymerase. Early work by the Aguilera group demonstrated that loss of factors of the THO/TREX protein complex that are involved at the interface of mRNA processing and export (Sträßer et al., 2002) leads a R-loop formation dependent transcription-associated hyperrecombination phenotype and impairment of replication fork progression (Huertas and Aguilera, 2003; Wellinger et al., 2006).

The formation of R-loops at certain sites of the genome can play a protective role. More than half of human promoters contain runs of a cytosine followed by a guanine (Saxonov et al., 2006). These CpG islands are important regulatory regions as methylation of the cytosine results in silencing of the gene (Illingworth and Bird, 2009). This abundance of cytosines and guanines can result in a GC skew that confers the ability to preferentially form R-loops at these sites and by doing so protect CpG islands from methylation and therefore prevent transcriptional silencing (Ginno et al., 2012).

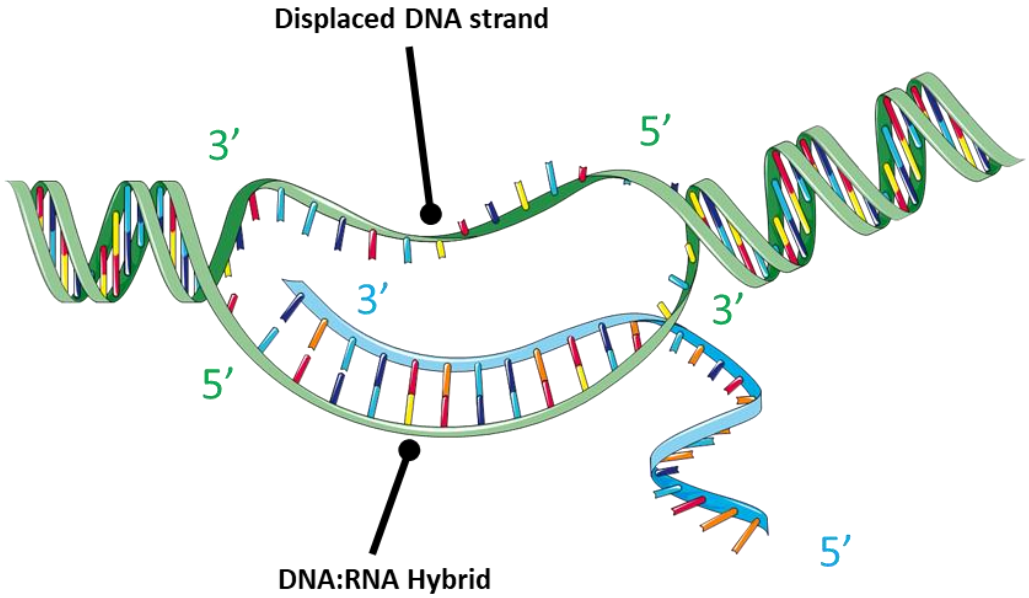


Figure 1.5: The structure of an R-loop

The formation of a DNA:RNA hybrid displaces one of the strands of duplex DNA. The hybrid pictured above is short, as commonly formed inside the active site of RNA polymerases. However, the nascent RNA can thread back after exiting the polymerase to form longer DNA:RNA hybrids that potentially span many hundreds of base pairs. It is these longer structures that are commonly referred to as R-loops that are associated DNA damage and genome instability

1.7.1 Physiological function of R-loops

There are several physiological processes in cells in which the formation of R-loops as natural intermediate structures is required. Class-switch recombination (CSR) is a process that allows switching to different Ig isotypes without changing the antigen specificity. R-loops are key intermediates in CSR that are formed upon transcription through switch regions (Reaban et al., 1994), stabilizing ssDNA regions that can be altered by activation-induced cytidine deaminase (AID) (Chaudhuri and Alt, 2004; Reaban and Griffin, 1990). How dC to dU transition in switch regions lead CSR to is not well understood, but DSB formation is required and it is possible that the repair of deaminated cytosine bases to uracil is processed by BER, leading to DNA nicks that are ultimately processed to DSBs during replication (Catalan et al., 2003).

The formation of R-loops appears to be required for efficient transcription termination, as their formation is especially prevalent at pause sites downstream of gene poly-A signals. This R-loop formation potentiates pausing at gene termination, to allow exonuclease degradation of the RNA attached to the polymerase. This R-loop appears to be resolved by the purported helicase Senataxin (SETX), as the depletion of SETX leads to improper transcription termination and RNA:DNA hybrid accumulation, however no biochemical data as of yet has shown hybrid unwinding activity of SETX (Skourti-Stathaki et al., 2011).

R-loops are formed on mitochondrial DNA (mtDNA) where they appear to play an important role in mtDNA replication, in a similar manner to ColE1 plasmid replication

(Itoh and Tomizawa, 1980). An RNA molecule transcribed from a promoter upstream of the mitochondrial heavy chain origin forms a stable and persistent R-loop that terminates at the origin and serves as a primer for initiation of mtDNA replication (Falkenberg, 2018; Xu and Clayton, 1996).

Despite R-loop formation being a requirement for some cellular functions, persistence or accumulation of these structures is heavily associated with DNA damage and genome instability. Much of the evidence comes from studies which deplete factors involved in the processing, resolution or degradation of R-loops.

1.7.2 R-loop driven instability

Several studies have produced evidence that indicates R-loops may naturally form at sites of DBSs. Laser micro-irradiation can be targeted at specific areas of a cell, resulting in DNA breaks (Holton et al., 2017). By employing a catalytically inactive form of RNaseHI fused to a fluorescent tag, Britton et al. (2014) were able to track RNaseHI localization to the site of irradiation, suggesting RNA:DNA hybrid formation at sites of DSBs.

SETX is reported to be recruited to DSB ends in transcriptionally active chromatin, presumably to resolve R-loops and perhaps to ensure that premature termination occurs in broken genes to allow access for repair machineries (Aymard et al., 2014; Cohen et al., 2018). Depletion of SETX results in increased R-loops, impaired RAD51 foci formation and increased accrual of 53bp1 at DSBs using the DSB inducible via AsiSI

(DIvA) cell line (Cohen et al., 2018). Loss of SETX can also result in large deletions around DSBs, that are attributed to R-loop processing by NER factors (Brustel et al., 2018; Sollier et al., 2014).

The genomic instability found at CFSs of long genes (described in Section 1.6) is also dependent upon the formation of R-loops, which causes the retention of FANCD2 at CFSs, both of which could be ablated by RNaseHI overexpression (Helmrich et al., 2011; Okamoto et al., 2018).

R-loop formation is a factor that contributes to the pathology of several human diseases. In Aicardi-Goutières syndrome mutations in enzymes that degrade nucleic acids, including RNase H2, resulting in an accumulation of DNA:RNA hybrids that appear to result in lower levels of DNA methylation potentially leading to alterations in gene activity and genome instability (Lim et al., 2015).

Aberrant expansion of repeated nucleotide sequences that results in the gene silencing can result in disease, as observed in Friedreich ataxia and Fragile X syndrome. In both these cases, the formation of R-loops is observed over the extended repeat sequences and contributes to transcriptional silencing (Groh et al., 2014).

The association of R-loops with DNA damage and genome instability suggests that aberrant R-loop formation could be a factor in the development of cancer. Increased R-loop formation is associated with Ewing sarcoma and hybrid accumulation is due to the

loss of EWSR1 function (Gorthi et al., 2018). R-loops have been implicated as the main source of DNA damage in cells that accumulate estrogen, higher levels of which are correlated with an increased risk of breast cancer (Stork et al., 2016).

1.8 FA pathway and FA factors implicated in transcription-replication conflict

Numerous FA factors are associated with the prevention of transcription and R-loop mediated genomic instability (Chang and Stirling, 2017). BRCA1 (FANCS) has long been understood to regulate transcription (Mullan et al., 2006) which prompted investigation into a possible role in preventing R-loop mediated instability. BRCA1 binds DNA at R-loop-associated transcription termination regions, where it recruits SETX to prevent R-loop formation and subsequent DNA damage (Hatchi et al., 2015). BRCA2 (FANCD1) is also associated with control of transcription and has been ascribed a role in preventing R-loop formation, as its depletion results in the accumulation of these structures and associated damage (Bhatia et al., 2014; Shivji et al., 2018).

The studies described above link the roles of BRCA1 and BRCA2 in preventing R-loop associated instability to cancer. Breast cancers lacking BRCA1 had an enrichment of insertions and deletions near to the mapped sites of BRCA1 bound termination regions (Hatchi et al., 2015). In BRCA2-defective pancreatic adenocarcinoma cells, the overexpression of RNaseHI could reduce γH2AX signals, indicating removal of R-loops results in a reduction of cancer cell DNA damage and instability (Bhatia et al., 2014). The association of BRCA1 and BRCA2 with R-loops prompted several studies in the possible role of the FA pathway in mediating transcription associated replication stress.

In one such study, the association of FA with transcription was discovered by the colocalization of FANCD2 with elongating RNAP II by the proximity ligation assay (Schwab et al., 2015). This association and the activation of the FA pathway could be reduced by general inhibition of transcription with cordycepin. This study also demonstrated that depletion of FANCD2 and other FA core-complex members results in the accumulation of R-loops and associated DNA damage, which could be rescued by transcription inhibition or overexpression of RNaseHI. In addition, *FANCD2*$^{-/-}$ avian DT40 cells that are re-complemented with the FANCD2^{K563R} variant, which are monoubiquitination-defective, also accumulate DNA:RNA hybrids further suggesting FA pathway activation is required to limit R-loop formation. Biochemical studies also showed that purified FANCM, but not a translocase-dead mutant, could disrupt DNA:RNA hybrids *in vitro.* This function of FANCM has been supported by further investigation (Hodson et al., 2018).

Strikingly, formaldehyde treatment was shown to further increase R-loop formation in FANCD2 depleted cells, suggesting the lesions that endogenous aldehydes introduce may mediate their instability via transcription and R-loop formation.

These findings were further confirmed in a parallel study that employed DNA immunoprecipitation and quantitative PCR (DIP-qPCR) to demonstrate the accumulation of R-loop associated damage in a variety of cells lines deficient in FANCD2 and FANCA, that could also be reduced upon treatment with RNaseHI (García-Rubio et

al., 2015). Taken together, these studies suggest a role for the FA pathway in preventing the accumulation of R-loops by mediating transcription-replication conflict.

1.9 Aims of the current study

Given the role of FA in mediating transcription-replication conflict, this study aimed to employ CRISPR-cas9 mediated gene editing to assess the impact of R-loop formation and associated defects in FA deficient cells.

Whilst there is support for R-loop accumulation in FA deficient model cell lines, it is unclear if is relevant to the phenotype in patients and how this relates to the pathology of FA. To address this, the DNA:RNA hybrids load was assessed in cells derived from patients with FA as well as FA deficient murine haematopoietic stem cells.

In addition, the sensitivity of FA deficient cells to transcription-based replication stress was probed by depleting cells of factors involved in transcription coupled repair.

2. Materials and Methods

2.1 Cell culture

Patient derived lymphoblastoid cell lines were maintained in 1640RPMI (Thermo Fisher) with 10-20% FCS (foetal calf serum) at 37°C under 5% carbon dioxide in an upright T25 tissue culture vented flask with 10-20 ml medium. Cells were split at 200,000 – 500,000 viable cells per mL.

U2OS and HeLa cells were maintained in DMEM (Thermo Fisher) with 10% FCS at 37°C under 5% carbon dioxide in a T75 tissue culture vented flask with the cell sheet down to allow attachment. If required, cells were cultured with 1% penicillin-streptomycin (Thermo Fisher) to avoid contamination.

In order to passage cells, media was first aspirated followed by a wash with PBS. Cells were incubated for 3-5 minutes with trypsin (0.25%, no phenol red) until detachment. Inactivation of trypsin was achieved by washing cells with 5 mL of media. Cells were then spun down at 750 x *g* and resuspended at an appropriate volume for re-seeding or experimental procedures. In general, cells were passaged until 80% confluency and split 1:10 to 1:15

2.2 siRNA knockdown

Small interfering RNA (siRNA) knockdown was achieved with a double pulse of 20 nM siRNA targeting the desired protein (see table 2.1). An siRNA sequence targeting Luciferase was used as a control in all experiments (Referred to as siCTRL in text and on

graphs). Optimal knockdown was found to be 48 hours after second pulse for majority of knockdown targets in U2OS and HeLa cells. A mixture of 90 µL Opti-MEM was mixed with 3 µL of 20 µM siRNA and 12 µL HiPerfect (Qiagen) transfection reagent and incubated for 10-15 minutes at room temperature.

For siRNA knockdown experiments targeting two different proteins at the same time, 20 nM of each target siRNA was added to the Hi-Perfect / Opti-MEM mixture. Single knockdown transfection mixes in these experiments contained 20 nM of target siRNA along with 20 nM control siRNA. Wildtype controls were transfected 40 nM of siCTRL. In this arrangement all replicates received 40 nM total siRNA.

Transfection was performed in a six-centimetre cell culture dish was seeded with 0.5 – 0.7 x 10^6 cells in 2.4 mL growth media before adding the incubated siRNA/transfection reagent mix dropwise. This first siRNA treatment was incubated for 24 hours. The following day the media was aspirated, and fresh media added before a second treatment of siRNA was performed as on day 1. Following the second treatment, transfected cells were incubated for 4-6 hours before splitting as adequate for the required experiment.

siRNA Target	**Sequence / Cat. Number & Company**
CSB	UGUAUCUCGUAAUCAGUGG
CTRL	CGUACGCGGAAUACUUCGA
FANCD2	GGUCAGAGCUGUAUUAUUC
Luciferase	12935-146 (Invitrogen)
XPC	AM16708 – 146633 (Ambion)
XPF	AM16708 – s4800 (Ambion)

Table 2.1: siRNA sequences or commercial products used for RNAi

2.3 Antibodies

Anti-FANCD2 (sc-20022, Santa Cruz), anti-CSB (ab66598, Abcam), anti-α-Tubulin (T5168, Sigma), anti-MCM2 (ab4461, Abcam), anti-H3 (ab1791, Abcam), anti-H2AX (05-636, Millipore), anti-CSA (GTX100145, Insight Biotechnology), anti-cyclin A (sc-751, Santa Cruz), anti-53BP1 (MAB3802, Millipore), anti-KU70 (611892, BD Biosciences), anti-XPC (D-10, Santa Cruz), anti-BrdU (mouse) (347580, BD Biosciences), anti-BrdU (rat) (ab6326, Abcam), anti-BLM (A300-110A, Bethyl), anti-XPF (ab76948, Abcam).

2.4 Western Blotting

Cultured cells were washed with PBS and harvested by either scraping or treatment with trypsin. Cells were pelleted at 750 x *g* and transferred to eppendorfs to be pelleted washed a further time by resuspension in PBS. Cells were lysed in sample loading buffer (9M Urea, 150mM 2-mercaptoethanol, 50mM Tris-HCl pH 7.3) and sonicated to shear genomic DNA.

Proteins were separated in an appropriate running buffer on 6-15% acrylamide (UN3426, AppliChem) gradient gel or on a range of commercial pre-cast alternatives as appropriate for the target size; 3-8% Tris-Acetate (EA0375BOX , Invitrogen), 7% Tris-Acetate (EA0358BOX, Invitrogen), 4-12 % Bis-Tris (NP0322BOX, Invitrogen). Gels were run at 150-180 V for 1-1.5 hours depending on desired separation, as indicated by Spectra Multicolour High Range protein ladder (26625, Thermo Scientific).

After separation, proteins were transferred to a nitrocellulose membrane (10600016, GE Healthcare) in a tris-glycine running buffer containing 20% methanol. Transfer was run at either 90 V for 90 minutes at room temperature, or overnight at 30 V at 4°C.

Antigens were visualized with HRP conjugated secondary antibodies (Dako; 1:2000 anti-mouse, 1:4000 anti-rabbit) and Immobilon enhanced chemiluminescence reagents (Millipore).

2.5 Agarose gel electrophoresis

Agarose gel electrophoresis was a method used to validate various steps of cloning processes. A 1% gel (w/v) was prepared in TAE (0.03 M Tris-acetate, 0.001 M EDTA), with the addition of ethidium bromide (1 µg/ml) for visualization.

2.6 DNA fiber technique

Cells were labelled with IdU (Sigma) at 25 µM for 20 minutes followed by labelling with CldU (Sigma) at 250 µM for 20 minutes. Cells were washed with ice cold PBS and resuspended at a concentration of 7.5 x 10^5 cells.

A glass slide was spotted with 2 µL of cell suspension and left to dry for 5 minutes followed by the addition of 7 µL of lysis solution (50 mM EDTA and 0.5% SDS in 200 mM Tris-HCl, pH 7.5). Slides were tilted to 15° to allow the fibers to spread along the slide.

Once dried, slides were immersed in methanol/acetic acid (3:1) in a staining jar and incubated for 10 minutes.

Slides were washed in ddH_2O and then immersed in 2.5 M HCl for 80 minutes followed by 3 washes in PBS. Slides were blocked with 5% BSA for 20 minutes and incubated with primary antibodies: 1:25 anti-BrdU (mouse) and 1:400 anti-BrdU (rat). After three washes in PBS slides were incubated with secondary antibodies: 1:500 sheep anti-mouse Cy3 (Merck) and 1:400 goat anti-rat Alexa Fluor 488 (Invitrogen). Slides were washed with PBS and coverslip was mounted with Vectashield. Images were acquired with a Zeiss 510 Meta laser-scanning confocal microscope.

2.7 Immunofluorescence staining of DDR markers

Cells were grown on cover slips until appropriate confluency for analysis and harvested by washing three times in PBS for 5 minutes. Fixation was performed with 4% PFA-PBS for 10 minutes. Slides were washed and incubated for one hour in blocking buffer (5% BSA) and overnight at 4°C in primary antibody.

The following day slides were washed in 0.1% Tween-PBS and incubated with fluorochrome conjugated secondary antibody for two hours at room temperature. Slides were washed and dried before adding Vectashield (with DAPI staining) and coverslip. Images were acquired with a Zeiss 510 Meta laser-scanning confocal microscope. ImageJ was used for picture processing and quantification of foci.

2.8 R-loop staining by IF and the S9.6 antibody

Cells were grown on cover slips in 6 cm dishes until they were 70% confluent and harvested by washing three times in PBS for 5 minutes. For RNaseH (M0297, NEB) treatment, cells were pre-permeabilized by incubation with 0.1% Triton-X in PBS for 1 minute and washed in PBS for 5 minutes. RNaseH was mixed at indicated concentrations with the reaction buffer provided by the manufacturer and incubated on cells at 37°C for 10 minutes before being washed with PBS as above.

A stock of 16% methanol-free formaldehyde (Thermo Fisher) was diluted with PBS for a final concentration of 4% and added to cells for 10 minutes at room temperature. Coverslips were washed with PBS three times and incubated for one hour in blocking buffer (5% BSA, 0.2% Milk, 0.005% NaAz in PBS) followed by an overnight incubation at 4°C with S9.6 in blocking buffer at a concentration of 1:60.

After overnight incubation cells were washed 3 times in PBS followed by incubation with Alexa-fluor 488 (Invitrogen; 1:500) conjugated secondary antibody in blocking buffer for 1 hour at room temperature in blocking buffer followed by 3 washes of PBS. Once dried, Vectashield (with DAPI staining) and coverslip were applied. Images were acquired with a Zeiss 510 Meta laser-scanning confocal microscope. ImageJ was used for picture processing and quantification of pan nuclear staining intensity. Nucleolar signal was not excluded in analysis.

2.9 Anaphase bridge analysis

Cells were cultured in a T175 flask until 80% confluence. Enrichment of ana/telophase cells was achieved by mitotic shake-off (5 hard strikes of the T175), to detach mitotic cells which have rounded up, but not cells in other stages of the cell cycle. Detached cells were harvested into a 50 mL Falcon tube and spun down at 700 x *g* for 3 minutes at 4°C. Media was removed and cells were resuspended in a small amount (~100 µL) of remaining media. Cells were pipetted onto a Polysine adhesion slide (J2800AMNZ, Thermo Scientific) inside a circle marked with a hydrophobic pen. Cells were spun onto slide at 700 x *g* for 3 minutes at 4°C.

Slides were subsequently incubated in ice cold methanol at -20°C for 20 minutes before washing three times with PBS. Cells were then blocked with 10% FBS-PBS for 1 hour at room temperature. For DAPI bridges, cells were stained with α-Tubulin. For Ultra-Fine bridges (UFBs) cells were stained with BLM. Primary antibodies were incubated in 0.1% FBS-PBS for 1 hour at room temperature followed by 3 washes in PBS. Secondary antibodies were incubated for 45 minutes at room temperature in 0.1% FBS-PBS followed by 5 washes of PBS. Once dried Vectashield (with DAPI staining) and coverslip were applied. Bridge structures were counted, and images were acquired on a DeltaVision microscope.

2.10 Cell survival assay

Cell survival assay were carried out in opaque 96-well microplates (Corning). 2500 cells per well were seeded into 96 well plates and allowed to adhere overnight. Cisplatin was

added at indicated concentrations and incubated for 48 hrs at 37°C. After incubation the existing media was removed and 100 µL phenol red free media was added with 20 µL of 0.15 mg/mL Resazurin to each well and incubated for 2 hours.

2.11 Cell growth assay

Cell growth assay were performed as above, but without any drug treatment. Cells were treated every 12 or 24 hours with 0.15 mg/mL Resazurin and incubated for two hours. The amount of resorufin was subsequently quantified by measuring fluorescence spectra at 560 nm excitation / 590 nm emission. Cell survival (as determined by the absorbance of resorufin) was calculated as a percentage of untreated cells. Cell growth was calculated as fold increase in fluorescent signal.

2.12 Generation of FANCD2 knockout line by CRISPR/Cas9

The following guide RNA (gRNA) sequences targeting the fourth exon of FANCD2 were selected using the optimized CRISPR Design tool (http://crispr.mit.edu; (Hsu et al., 2013); gRNA1, TTTGTCTTGTGAGCGTCTGC; gRNA2, GGAGTCTTACATTGAGGATG).

DNA oligonucleotides were purchased from Integrated DNA Technologies and cloned into the pX335-GFP vector (Cong et al., 2013) to generate targeting constructs that were subsequently co-transfected in an equimolar ratio into U2OS cells using Lipofectamine 2000. 24 hours after transfection, cells were sorted using a MoFlo cell sorter (Beckman Coulter) for cells expressing Cas9 nickase (GFP-positive cells) and left to recover for 6

days before sorting for single cells and allowing colonies to form. FANCD2 expression was analysed by western blotting.

Disruption of the targeted exon was confirmed by isolation of genomic DNA from wildtype or *FANCD2*$^{-/-}$ U2OS cells using DNeasy Blood and Tissue Kit (Qiagen). The genomic region of interest was amplified with primers designed to flank the proposed cut sites (Table 2.2). Amplified DNA was confirmed to be the expected size by agarose gel electrophoresis and prepared for sequencing with MinElute PCR Purification Kit (Qiagen). Sequencing was performed with a primer designed inside the amplified region (Table 2.2).

Primer	Sequence
FANCD2_fwd	TCATCAGGCAAGAAACTTGG
FANCD2_rev	CCACAAACCTCTTCCATCAA
FANCD2_seq	CTATGGTAGGAAACTGGTGACC

Table 2.2 PCR and sequencing primers for FANCD2 analysis

2.13 R-loop detection by S9.6 Slot-blot

Cultured cells were lysed with DNA lysis buffer (100 mM Tris-HCl pH 8.5, 5mM EDTA, 0.2% SDS, 100 mM NaCl) and treated overnight with Proteinase-K (03115879001, Roche) at 55°C. Genomic DNA was extracted using isopropanol / ethanol precipitation and re-suspend overnight at room temperature in TE.

A total of 500ng was transferred into a Hybond N+ membrane (RPN303B, Amersham Biosciences) using a pump and slot blot equipment (BioRad). DNA was fixed to the membrane by baking at 80°C for 2 hours, blocked for 1 hour in 5% milk TBST and

incubated with the S9.6 antibody overnight at 4°C or 1 hour at room temperature followed by incubation with infrared dye secondary antibody (926-32210, Li-Cor). Membranes were scanned using a quantitative fluorescent imaging system (Odyssey, Li-Cor Biosciences).

S9.6 antibody was removed using 0.5 mg/ml Proteinase-K overnight at 55°C and then DNA was denatured by washing in 0.4M NaOH, 0.6M NaCl and neutralized by washing in 1.5M NaCl, 0.5M Tris pH 7.4. The membrane was then blotted for single strand DNA (MAB3034, Millipore) for 1 hour at room temperature. Band intensity was quantified using ImageJ analysis.

2.14 Transcription assay.

Transcription levels after siRNA treatment was assessed with the Click-it RNA Imaging Kit (Invitrogen) and Alexa Fluor 488 (Invitrogen). Cells were grown on cover slips and following treatment with siRNA. 48 hours after the second siRNA pulse (for peak knockdown), cells were incubated with 1 mM EU (Invitrogen) for 1 h and then fixed with 4% paraformaldehyde (Thermo Fisher) for 10 min. Click reactions were performed in accordance with the manufacturer's instructions.

2.15 Analysis

Statistical analysis was performed in GraphPad Prism software. Tests are described in figure legends. Where direct comparison of only two groups was required, a Two tailed

Mann-Whitney test was employed. Where comparison of three or more groups was required, a one-way ANOVA test was used to compare all means.

*p ≤ 0.05; **p ≤ 0.01; ***p ≤ 0.001, ****p ≤ 0.0001.

3. The relevance of R-loops to the pathology of Fanconi anaemia

3.1 Introduction

In the 90 years since Guido Fanconi first described a “familial infantile pernicious-like anaemia” the understanding of the disease which bears his name has grown dramatically. It is now known that Fanconi anaemia (FA) is caused by the loss of one of at least 22 genes that comprise a pathway, named the FA pathway, that is widely associated with maintaining genome stability by repairing damaged DNA.

One of the best understood functions of the FA pathway is the response to chemotherapy compounds that induce, among other lesions, DNA interstrand crosslinks. This class of cytotoxic drugs, such as cisplatin and mitomycin C (MMC), are not produced endogenously and are not commonly encountered in the environment, so therefore are not likely to be the driving force of the recent evolution of this large and multi-faceted pathway.

There is however, a growing body of evidence that aldehydes are one of the endogenously produced agents that cause lesions that require FA pathway. These are reactive compounds which are natural by-products of cellular metabolism that can cause a range of lesions, including DNA and DNA-protein crosslinks. When aldehydes accumulate they can overwhelm the various dedicated detoxification pathways, causing DNA damage that requires the action of the DNA repair pathways (Wang et al., 2000).

In mouse models, the absence of a key enzyme in detoxification of acetaldehyde, called Aldh2, causes a significant elevation in DNA repair by homologous recombination and this can be exacerbated by the addition of ethanol, which is converted into acetaldehyde during its metabolism (Garaycoechea et al., 2018). FA deficient cells are hypersensitive to formaldehyde, a simple aldehyde that is a common intermediate or product of cellular metabolism (Ridpath et al., 2007). When aldehyde detoxifying enzymes are lost in conjunction with an FA protein, such as FANCD2, there is a substantial loss of viability and mice that do survive eventually develop bone marrow failure (BMF) and other haematological abnormalities, phenocopying the symptoms found in patients (Garaycoechea et al., 2018; Langevin et al., 2011).

It is reasonable, then, to assume that when aldehydes escape processing by detoxification pathways, they can react with DNA and cause lesions that subsequently require the action of the FA pathway for repair. However the loss of Rev1, which is an important factor in translesion synthesis (TLS), does not produce the same lack of viability (Oberbeck et al., 2014). Considering that TLS is an important step in the repair of ICLs, this data suggests that the requirement to repair the lesion(s) introduced by the substrate of Aldh2 may not be the same as the repair of canonical chemotherapy-induced crosslinks.

The specific nature of these aldehyde-induced lesions remains unknown but many of the proposed lesions that aldehydes can generate will block the progression of transcription complexes if they occur in transcribed regions. Recently, several FA

proteins, including BRCA1/2, have also been associated with preventing DNA damage and genomic instability arising from transcription associated DNA:RNA hybrid structures called R-loops (Bhatia et al., 2014; Hatchi et al., 2015). This has led to studies demonstrating that the FA pathway appears to suppress genome instability arising from R-loops (García-Rubio et al., 2015; Hodson et al., 2018; Schwab et al., 2015).

In cancer cell lines, knockdown of the FA pathway leads to an increased level of DNA:RNA hybrids and associated replication defects and damage, which can be rescued by inhibition of transcription or enzymatic removal of R-loops by overexpression of RNase H (García-Rubio et al., 2015). In addition, one member of the FA pathway, FANCM, can provide the enzymatic mechanism by which R-loops are resolved *in vitro* (Hodson et al., 2018).

The aim of the current study was to employ CRISPR-Cas9 gene editing to knockout FANCD2 and assess R-loop load and associated defects in FA deficient cells. In addition, this study aimed to assess the relevance of DNA:RNA hybrids to the pathology of FA by probing R-loop formation in cells derived from patients and FA-deficient mouse haematopoietic stem cells.

3.2 Results

3.2.1 Knockout of FANCD2 via CRISPR-Cas9

The expression and activation of FANCD2 is pivotal to the overall function of the FA pathway, and loss of FANCD2 activity is a major cause of FA as 80% of associated biallelic germline mutations are attributed to compromised FANCD2 monoubiquitination (Nepal et al., 2017).

FANCD2 has several protein coding isoforms (Zerbino et al., 2018) and in order to generate a FANCD2 knockout across all functional protein encoding transcripts the CRISPR-Cas9 genome editing technique was employed with guide RNAs (gRNAs) designed to target the fourth exon, which is shared across most isoforms, (Figure 3.1A) using the optimized CRISPR design tool (http://crispr.mit.edu (Hsu et al., 2013)). DNA oligonucleotides were cloned into the pX335-GFP vector (Cong et al., 2013) to generate targeting constructs that were co-transfected in an equimolar ratio into U2OS and HeLa cells using Lipofectamine 2000 (The CRISPR gRNA design, cloning, transfection and cell sorting performed by Jadwiga Nieminuszczy).

GFP positive colonies were identified during FACS and given six days to recover before being sorted into single wells in multiple 96 well plates. Individual cells that formed colonies were expanded and screened by western blot for the expression of FANCD2 protein. Two U2OS and one HeLa derived single cell population showed complete loss of FANCD2 protein expression (Figure 3.1B).

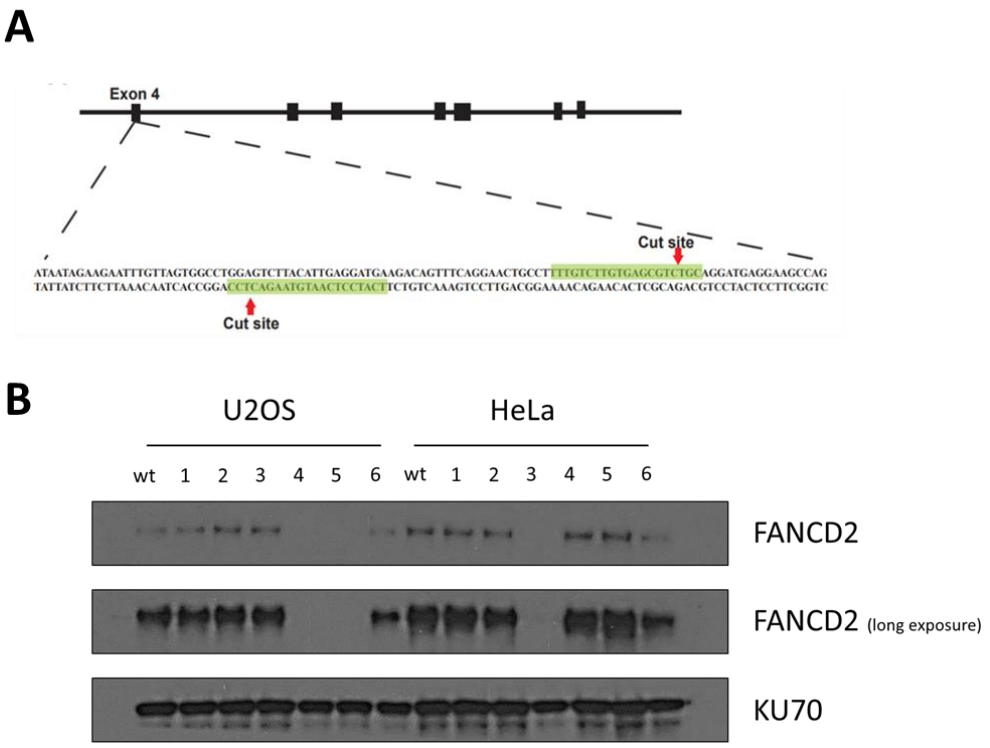


Figure 3.1: Strategy and western blot analysis of FANCD2 knockout by CRIPSR-Cas9 gene editing.

(A) Overview of FANCD2 knockout strategy using the CRISPR-Cas9 nickase (Schwab et al., 2015). Schematic representation of the human FANCD2 genomic locus with guide RNA sequences highlighted in green and predicted cut sites marked by red arrows.

(B) Western blot analysis of FANCD2 expression whole cell lysates of GFP positive single cell populations from U2OS and HeLa cells targetted with FANCD2 gRNAs. KU70 serves as loading control.

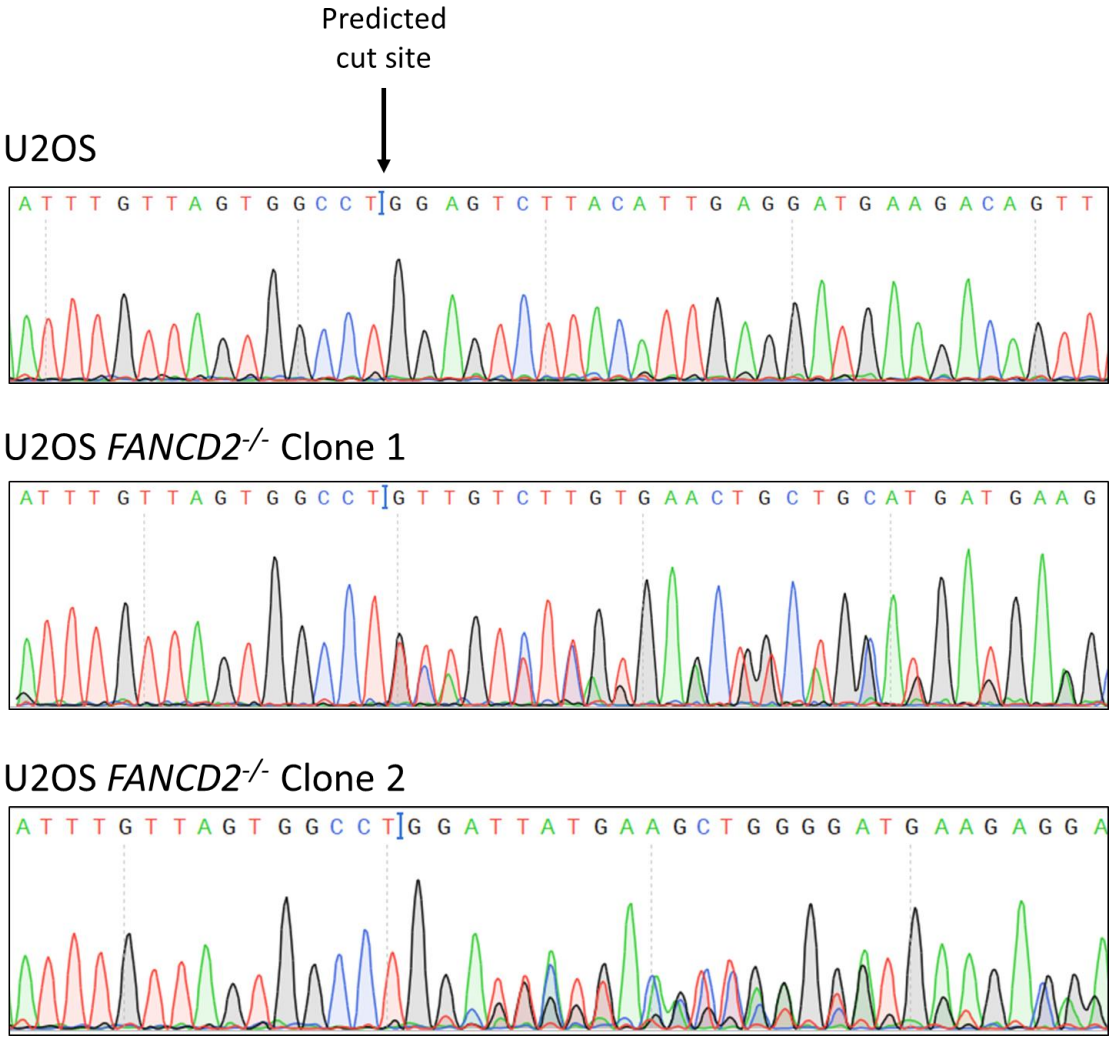


Figure 3.2: Genomic DNA sequencing of the FANCD2 locus.

Chromatogram of sequencing data of U2OS wildtype and two purported clones that lack FANCD2 protein expression. The predicted cut site is shown, and chromatogram data thereafter indicates alterations in multiple alleles.

The disruption of wildtype FANCD2 gene sequence was determined by sequencing with a primer designed 300 base pairs downstream of exon 4. The sequencing chromatogram from both FANCD2 knockout clones shows a loss of normal genomic sequence directly after the propose cut site (Figure 3.2). This suggests there is loss of nucleotides at the cut site resulting in a frameshift. The multiple nucleotide calls after this point suggests there are different alterations to the sequence in each allele.

3.2.2 FANCD2$^{-/-}$ knockout clones are sensitive to cisplatin

Cells that are deficient for FA pathway function are sensitive to a range of DNA damaging compounds but are particularly sensitive to agents that induce DNA cross-links (Sasaki and Tonomura, 1973). There are several agents that are capable of inducing interstrand cross-links (ICL) and these are commonly used as anti-tumour drugs. Cisplatin forms several adducts on DNA, of which ICLs represents a low proportion but are thought to be predominantly responsible for its toxicity. Cisplatin forms ICLs between guanines in the sequence d(CpG) and the incorporation of this causes a major distortion in DNA that blocks the progression of polymerases (Dronkert and Kanaar, 2001).

To determine if these FANCD2 U2OS knockout cells lacked FA pathway function, their response to a range of doses of cisplatin was examined by the resazurin viability assay. Two U2OS clones that lacked FANCD2 expression were plated at an optimal seeding density (2500 cells/well) in a 96 well plate and exposed to an increasing dose of cisplatin. The two U2OS derived FANCD2$^{-/-}$ clones displayed sensitivity to cisplatin at similar doses

(IC_{50} of 713 nM for wildtype U2OS, whilst the knockout clones are 326 nM and 283 nM, respectively) in comparison to their parental cell line (Figure 3.3).

3.2.3 Increased R-loop formation is a general phenomenon of FA deficient cells

When replication occurs at actively transcribing regions there is the potential for localised replication stress, compromising genome stability (Helmrich et al., 2011). Both the transcribing RNA polymerase (RNAP) and transcription-based DNA:RNA hybrids (R-loops) can block replication fork progression (García-Muse and Aguilera, 2016). Whilst stalled elongation complexes can be rescued or degraded through transcription-coupled repair (Vermeulen and Fousteri, 2013), aberrant R-loops may be retained and remain undetected until encountered by the replication fork. Whilst the FA pathway is typically associated with repair of ICLs, it functions to protect replication forks from degradation that have stalled in a variety of contexts (Schlacher et al., 2012) and is activated in response to transcription conflicts (Schwab et al., 2015).

To determine if the FA pathway has a role in preventing or resolving aberrant R-loop accumulation, wildtype and *FANCD2*$^{-/-}$ U2OS cells were assayed for DNA:RNA hybrid load with the S9.6 antibody, which specifically recognizes DNA:RNA hybrids (Boguslawski et al., 1986). Genomic DNA was isolated from cells and blotted onto a nitrocellulose membrane before being probed with the S9.6 antibody. In comparison to wildtype, both FANCD2 knockout clones exhibit increased levels of DNA:RNA hybrid formation (Figure 3.4).

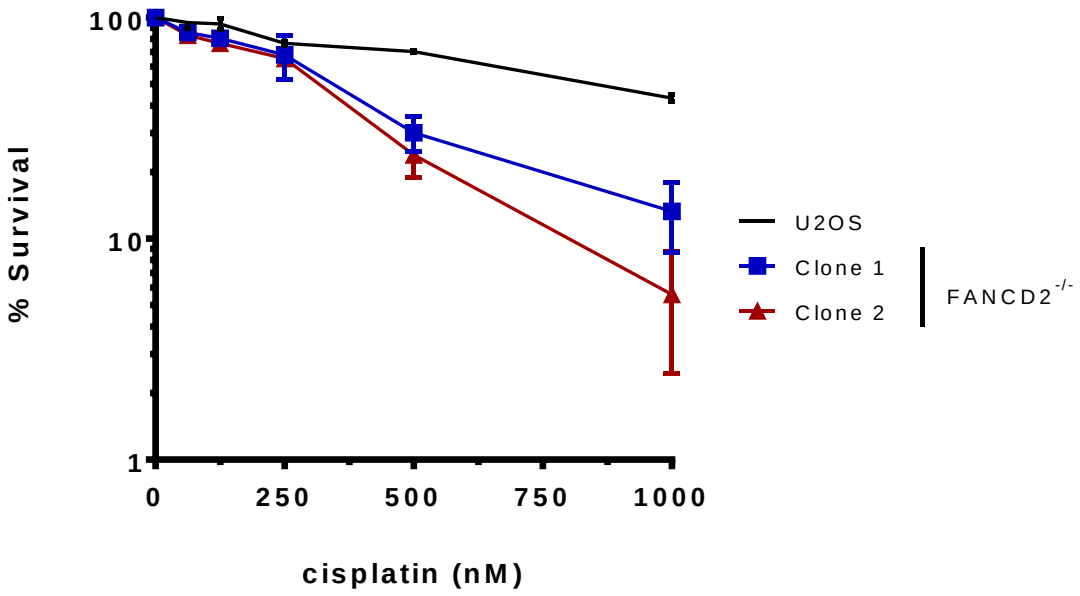


Figure 3.3: Sensitivity of FANCD2 knockout clones to cisplatin

Survival assay of control U2OS and FANCD2$^{-/-}$ *clones 1 and 2 treated with the indicated doses of cisplatin. IC_{50} of 713 nM for wildtype U2OS, whilst the knockout clones are 326 nM and 283 nM, respectively. Error bars show SD from triplicates.*

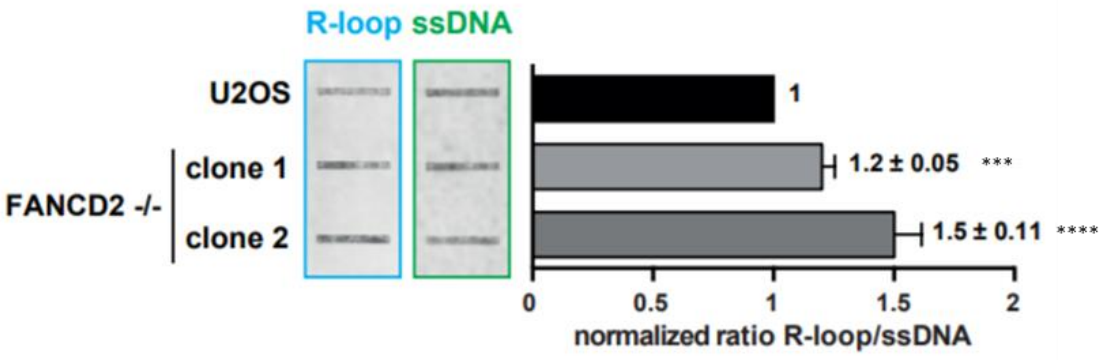


Figure 3.4: R-loop formation in FANCD2 knockout U2OS clones

*Slot-blot of genomic DNA from control and FANCD2-depleted cells probed with the S9.6 antibody. Denatured DNA was probed with an antibody against single-stranded DNA (ssDNA) to determine DNA loading. The graph shows the ratio of DNA:RNA hybrid IR fluorescence intensity divided by single stranded DNA IR fluorescence intensity from three independent experiments. (n=3. Two tailed Mann-Whitney test ****p < 0.0001)*

Whilst the slot-blot method of probing DNA:RNA hybrids is robust, it lacks the sensitivity of other assays such as immunofluorescent (IF) staining followed by imaging by confocal microscopy. In an attempt to achieve better sensitivity in detecting R-loops, a number of optimization experiments were conducted (data not shown) to determine immunofluorescent staining conditions and methods such as cell density, antibody concentration, fixation and permeabilization.

In cells, DNA:RNA hybrids can be removed by specific enzymatic degradation of the RNA strand by RNase H. A genuine DNA:RNA hybrid signal should be reduced by treatment with RNase H. Accordingly, U2OS cells were cultured on glass slides, pre-permeabilized and treated with RNase H at the indicated doses before being immunostaining overnight with the S9.6 antibody. The mean IF signal of the S9.6 antibody could be significantly reduced upon low concentration RNase H treatment (5 U/mL) and further reduced by increasing concentration (Figure 3.5). These data demonstrate that IF staining of U2OS cells with the S9.6 produces a quantifiable signal of nuclear DNA:RNA hybrid formation. However, this RNase H treatment demonstrating reduction of signal should have been included in every experiment throughout this thesis (in this and subsequent chapters) wherever the S9.6 signal is quantified to be sure the signal is genuinely from DNA:RNA hybrids. In addition, a treatment with RNase A could have been included to quantify the amount of signal that may be produced by antibody binding to ss/dsRNA, for which the S9.6 antibody has been demonstrated to have an affinity for, albeit lower than DNA:RNA hybrids (König et al., 2017).

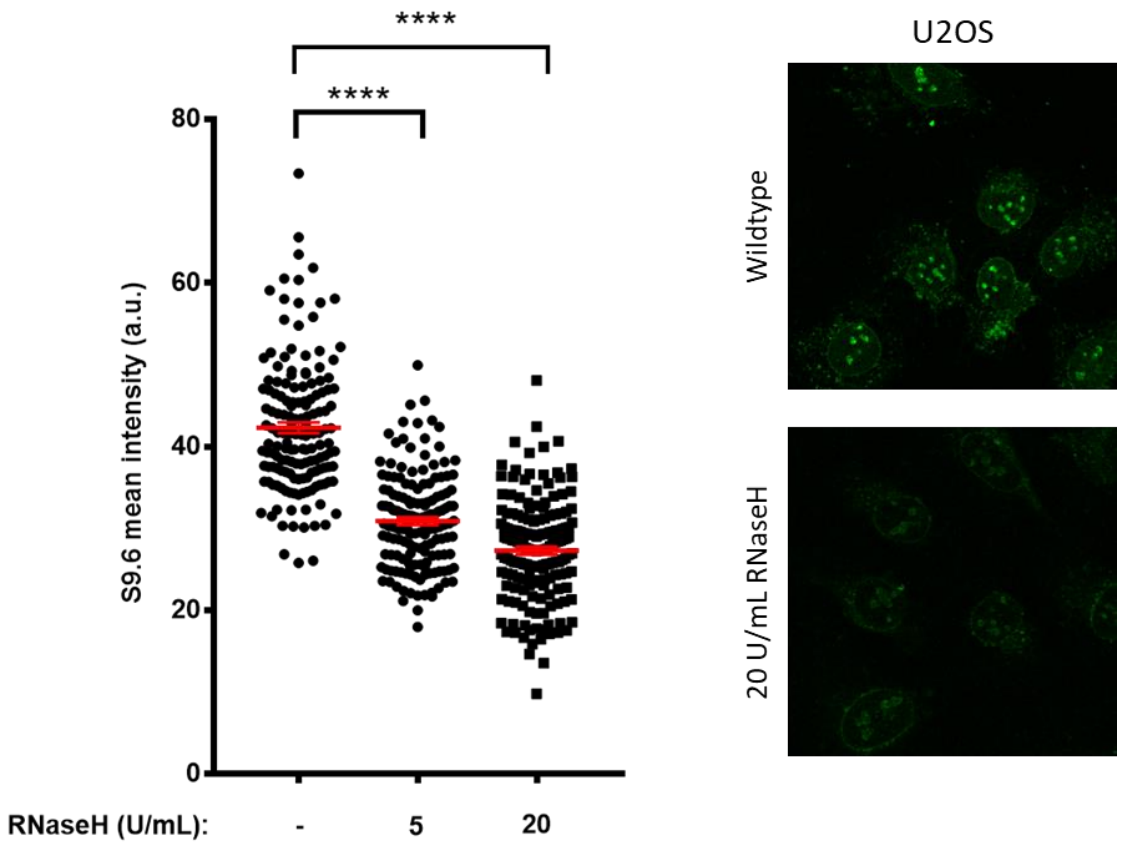


Figure 3.5: DNA:RNA hybrid signal reduction by treatment with RNase H

*Mean immunofluorescence intensity of S9.6 signal in U2OS cells treated with RNase H. RNase H signal is ablated upon treatment with RNase H prior to fixation. Red line and arrow bar represent mean ±SEM. (n=3. Two tailed Mann-Whitney test ****p < 0.0001).*

3.2.4 Increased R-loop formation is present in cells derived from patients and FA haematopoietic stem cells.

The quantification of the *FANCD2*$^{-/-}$ DNA:RNA hybrid load in Figure 3.4 supports other R-loop quantifications performed in U2OS and HeLa cells where FA factor deficiency was achieved through RNAi targeting and depletion of FANCD2 as well as FANCA, FANCL and FANCM (García-Rubio et al., 2015; Schwab et al., 2015). Whilst these cell lines are effective models for a wide range of genetic manipulations and experimental assays, they are derived from cancer cell lines and the phenotype they exhibit may not be reflective of the situation in patients.

To assess whether the function of the FA pathway in the prevention or resolution of R-loops could be relevant to its pathology, FA patient-derived lymphoblastoid cells deficient in core complex proteins FANCA and FANCF (CV1725 and CV1785) along with FANCD2 (CV1665) were obtained, along with a wildtype control (Smetsers et al., 2012).

The DNA:RNA hybrid load of these cells was ascertained by performing immunostaining experiments with the S9.6 antibody and confocal microscopy. In line with the immunoblotting and immunostaining experiments in model cells lines, FA deficient cells from these differing FA complementation groups all exhibit a marked increase in the nuclear fluorescence intensity (Figure 3.6), indicating an increased DNA:RNA hybrid load. The hybrid load in FANCA samples, while increased, appears to be slightly lower than of FANCD2 or FANCF. This is perhaps surprising as FANCA and FANCF both work in the FA core complex, but may represent their differing function or absolute

requirements for monoubiquitination of FANCD2/I as biochemical studies of the function of the core complex observed that the sub-complex of FANCA:FANCG:FAAP20 was dispensable for the ubiquitination process (van Twest et al., 2017).

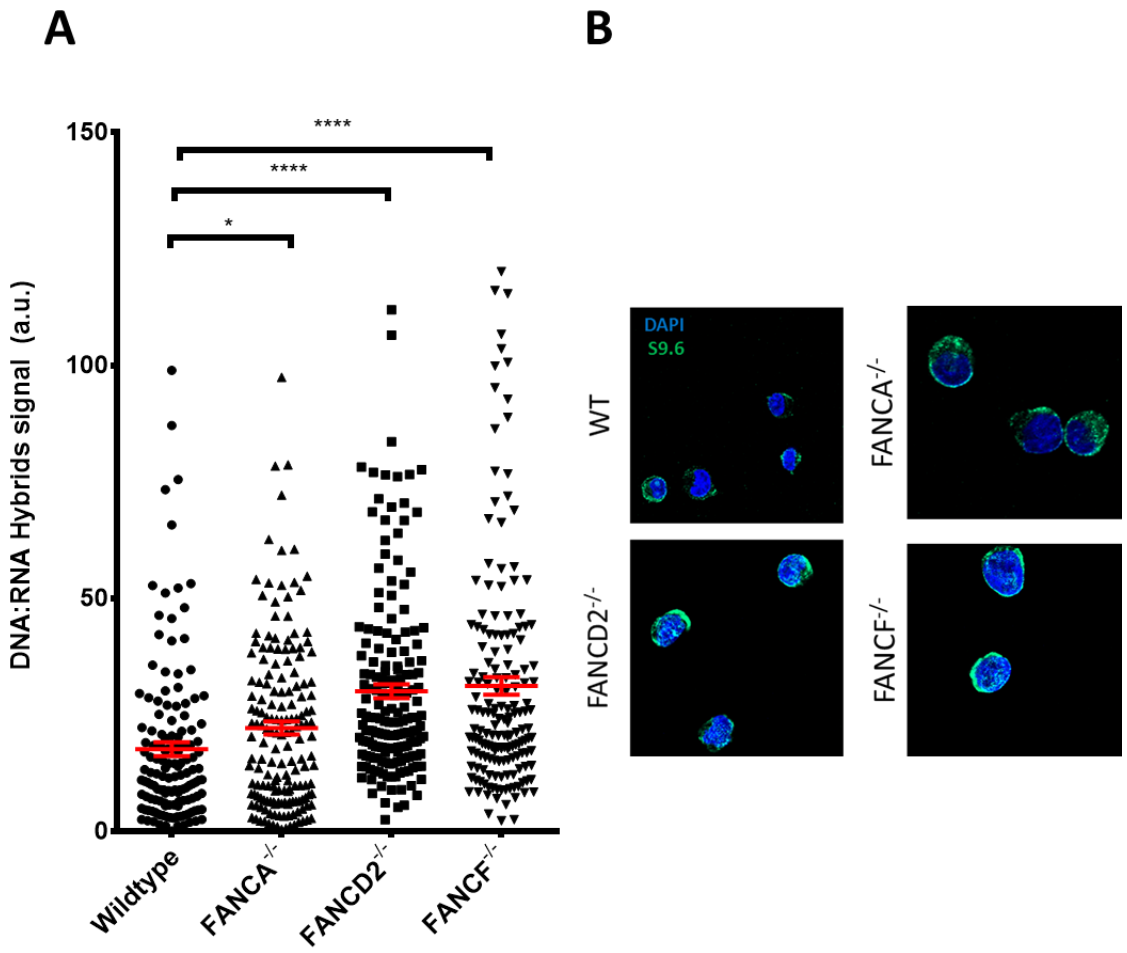


Figure 3.6: DNA:RNA hybrid load in lymphoblastoid cells derived from patients with FA

*(A) Mean immunofluorescence intensity of S9.6 signal of one wildtype and three patient-derived lymphoblastoid cell lines deficient in FANCD2, FANCA and FANCF. Red line and arrow bar represent mean ±SEM. n=3 (Two tailed Mann-Whitney test ****p < 0.0001). (B) Representative staining images from wildtype and FA deficient cells. DAPI is used to mark nucleus.*

The main clinical features of FA are haematological abnormalities and are found in nearly all patients, including thrombocytopenia (abnormally low platelet count), neutropenia (low neutrophil count) and progressive pancytopenia (defects in multiple lineages) (Auerbach, 2009).

Whilst all mature haematopoietic lineages are eventually compromised, FA patients do not present with BMF until late in the first decade of life suggesting a progressive depletion of progenitor cells. Indeed, the ability of FA patient derived haematopoietic progenitors to grow is severely compromised *in vitro* (Bagnara et al., 1992; Daneshbod-Skibba et al., 1980), suggesting that the loss of FA pathway function affects blood production at the top of the hierarchy of haematopoiesis. This is supported by FA patient data demonstrating decreased numbers of haematopoietic stem cells (HSCs) and progenitors, even before clinical presentation of bone marrow failure (Kelly et al., 2007). Similarly, there are reduced numbers of these cells in the umbilical cord blood taken from new-born FA patients (Auerbach et al., 1990). If R-loop prevention or resolution is a key function of the FA pathway that contributes to the pathology of FA, then the increased DNA:RNA hybrid load observed in model systems and FA patient derived lymphoblastoid cells should also be present in HSCs.

In FA deficient mouse models, accumulation of DNA damage in HSCs appears to be a direct consequence of the exit from a quiescent state in response to physiological stress which eventually leads to collapse of the haematopoietic system (Walter et al., 2015). In order to determine if this exit from dormancy produces an abnormal R-loops levels similar to DNA damage, long-term HSCs (LT-HSC) were isolated (lineage negative: c-Kit^{+},

Sca-1^{+}, CD48^{-}, CD150^{-}, CD34^{-}) by the Milsom group from wildtype or *Fanca*$^{-/-}$ mice that were untreated or had been treated with polyinosinic:polycytidylic acid (pI:pC). The administration of pI:pC leads to a rapid transition of LT-HSCs from quiescence into an active cell cycle *in vivo* (See Extended Figure 1F from Walter et al. (2015) for cell cycle data).

pI:pC is a double-stranded RNA that is composed a polymer of inosinic acid paired with a polymer of cytidylic acid and is structurally similar to double stranded RNA present in some viruses. This acts as an immunostimulant that mimics viral infection and induces a type I interferon response, producing physiological stress and the upregulation of blood cell production that is achieved through the cycling of HSCs to produce multi-potent progenitors. LT-HSCs were fixed and assayed by immunofluorescence with the S9.6 antibody to quantify their DNA:RNA hybrid load.

Interestingly, LT-HSCs from untreated *Fanca*$^{-/-}$ mice display a similar DNA:RNA hybrid load to those isolated from wildtype mice. This is perhaps unsurprising as the majority of these LT-HSCs from both mice are likely to be dormant and non-cycling. However, upon treatment with pI:pC this population of cells exit senescence and begin to cycle. Under these conditions, both wildtype and *Fanca*$^{-/-}$ mice exhibit increased DNA:RNA hybrids in their LT-HSCs, but notably the FA deficient mice have significantly more (Figure 3.7).

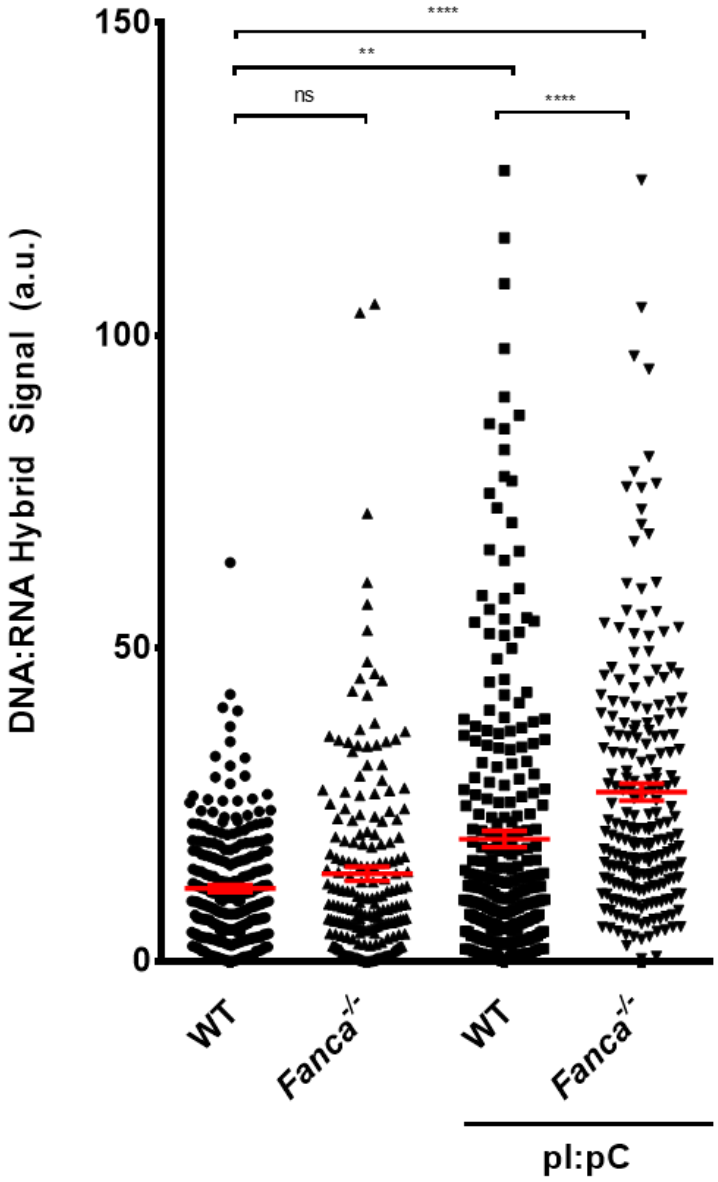


Figure 3.7: DNA:RNA hybrid load in wildtype and *Fanca*$^{-/-}$ mouse LT-HSCs

Mean immunofluorescence intensity of S9.6 signal of wildtype or Fanca$^{-/-}$ *from control or pI:pC treated long-term HSCs. Red line and arrow bar represent mean ±SEM. (n=3. Two tailed Mann-Whitney test ****p < 0.0001)*

3.3 Conclusion

There are now numerous studies showing that common cancer cell lines that are made deficient for FA pathway members through RNAi depletion display increased levels of DNA:RNA hybrids. This phenotype is also present in our *FANCD2*$^{-/-}$ CRISPR knockout, indicating this phenomenon is not due to a non-specific effect of siRNA transfection and knockdown in transformed cell lines.

My study has also established that increased R-loop formation appears to be a general phenomenon in FA patient derived cell lines. Increased hybrid formation was observed across three different complementation groups that are compromised in expression of FANCD2 or their ability to form a core complex and therefore lacking FANCD2 monoubiquitination and pathway activation. This immunofluorescence data is now supported by another study assaying different patient derived, transformed *FANCD2*$^{-/-}$ (PD20) and FANCA$^{-/-}$ (HSC72) cell lines by employing DIP-qPCR to measure hybrid formation at specific genomic loci (García-Rubio et al., 2015).

Lastly, when LT-HSCs of FA deficient mice are required to exit dormancy and begin cycling, they form higher levels of R-loops than their wildtype counterparts. Damage and instability occurring due to these structures may contribute to the attrition of these cells and therefore be a major contributor to the pathology of FA.

Taken together, these data support a role for the FA pathway in preventing excessive R-loop formation. When the FA pathway is lost R-loops accumulate and persist

abnormally, which appears to be a major contributor to the observed increase in DNA damage and genomic instability of FA deficient cells.

Cells that accrue too much DNA damage fail to properly duplicate their genomes, preventing replication and often activating apoptosis pathways. Due to this, R-loop associated damage could be a major contributor to the loss of LT-HSC and the development of bone marrow failure in patients with FA.

4. The FA pathway protects against transcription-based cellular stress

4.1 Introduction

Considering that the FA pathway is activated by transcription, it is reasonable to postulate that interference with efficient transcription in cells lacking a functional FA pathway may lead to replication fork stalling and collapse at sites of dysfunctional transcription. The collapse of replication forks leads to the generation of highly toxic single-ended double-strand breaks and subsequently, under-replicated or damaged DNA entering mitosis. At certain thresholds this DNA damage becomes intolerable and cells enter apoptosis (Roos and Kaina, 2006).

Whilst biallelic deactivation of FA proteins causes the eponymous disease, monoallelic loss of some of these proteins leads to an increased risk of cancer, and somatic mutations of FA proteins are frequently found in a range of tumours (Nalepa and Clapp, 2018; Shen et al., 2015).

This type of genetic alteration within a cancer cell differentiates them from healthy cells and potentially increases their relative requirement for other pathways or molecular actors to compensate for this deficiency. Two genes are said to be “synthetic lethal” if mutation of either alone is compatible with viability, but loss of both leads to cell death. Identifying such interactions has led to the development of successful novel therapies, particularly by targeting deficiencies in DNA damage repair (Lord and Ashworth, 2017).

Transcription-coupled nucleotide excision repair (TC-NER or TCR) is a pathway that specifically deals with repair of lesions in actively transcribing genes and is vital for maintaining efficient transcription. CSB is the main orchestrator of TCR and is responsible for recognizing stalled RNA polymerase (RNAP), initiating repair and recruiting the nucleotide excision repair (NER) machinery to remove DNA lesions that block transcription progression.

Interestingly, in the absence of DNA:RNA helicases such as Senataxin (SETX) or Aquarius (AQR) that process and resolve R-loops, the TCR pathway appears to be able to process R-loops into DSBs (Sollier et al., 2014). Loss of CSB and the TCR pathway could induce transcription-based replication stress by the persistence of two blocks to replication forks; stalled RNA polymerases and aberrant R-loops.

This suggests that the loss of CSB will increase the number of transcription-associated structures that subsequently require the action of the FA pathway. The following study aimed to explore the consequences of disrupting transcription-coupled repair in cells that lack a functional FA pathway by targeting the CSB protein.

4.2 Results

4.2.1 Targeting the TCR pathway by RNA interference of CSB

CSB deficient cells show a severely deficient recovery of RNA synthesis after DNA damage (Andrade-Lima et al., 2015), suggesting that the loss of TCR results in prolonged stalling of RNAPs when they encounter obstacles on DNA. The extended stalling of RNAP increases the potential for these DNA bound complexes to become a block to replication fork progression. In addition, the stalling of RNAP promotes R-loop formation which can also block replication. In this manner, a compromised DNA repair response to stalled transcription complexes should result in increased replication-transcription conflict.

In order to reduce cellular CSB protein levels, an siRNA previously established to target CSB (Sakai et al., 2012) was transfected into U2OS cells producing a substantial reduction in CSB protein levels, as measured by western blot (Figure 4.1).

Frequent stalling of RNA polymerase causes an overall decline in transcription, and CSB expression is required to maintain normal levels of RNA synthesis (Epanchintsev et al., 2017; Vermeulen and Fousteri, 2013). Consistent with its role in ensuring efficient transcription, overall levels of RNA synthesis are decreased upon CSB knockdown (Figure 4.2).

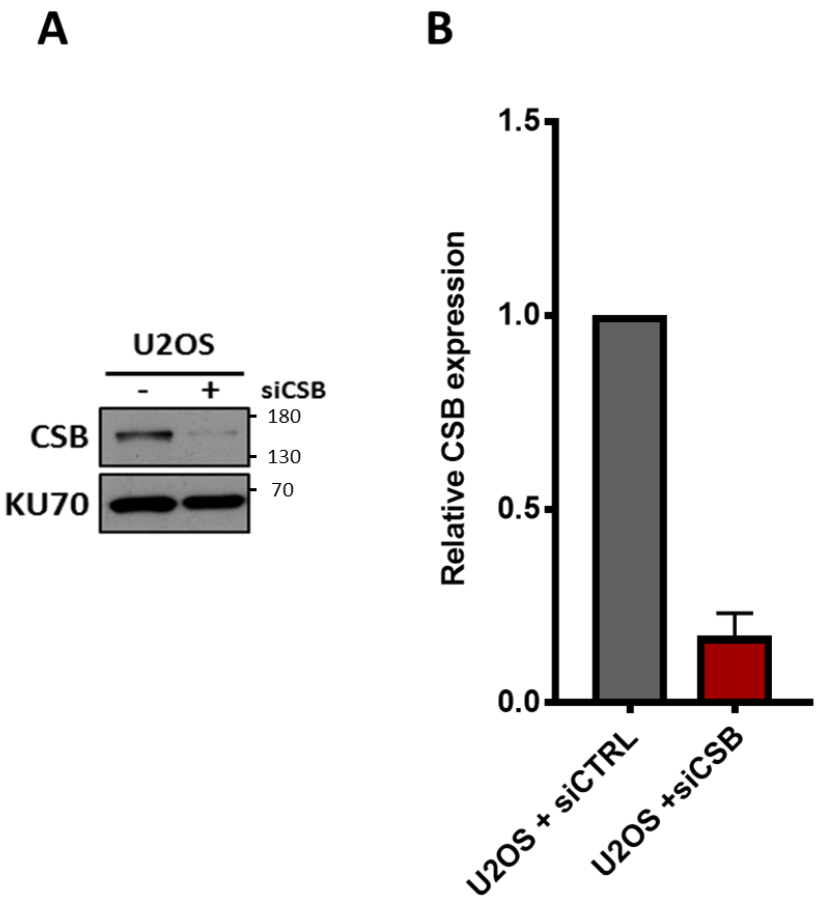


Figure 4.1: siRNA knockdown of CSB in U2OS cells.

(A) Western blot showing reduction in CSB expression in wildtype U2OS cells upon treatment with 20 nM siRNA targeting either luciferase (control) or CSB. Cells were pulsed twice with siRNA and harvested 24 hours after second pulse. Ku70 serves as loading control. (B) Quantification of (A) by densitometry.

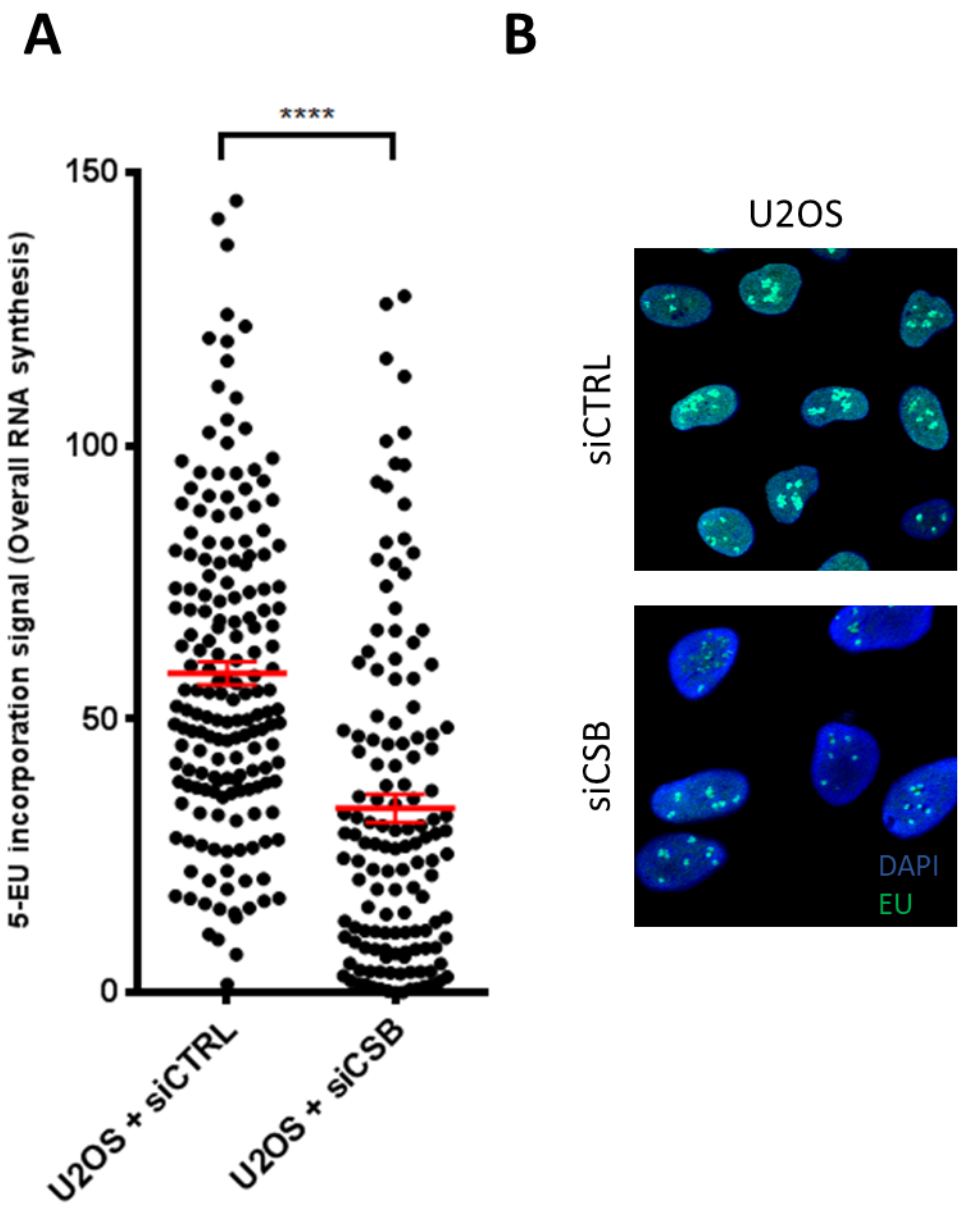


Fig 4.2: Reduction in RNA synthesis upon depletion of CSB.

*Mean (±SEM) global RNA synthesis levels in wildtype U2OS cells as determined by incorporation of 5-ethynyl uridine. (n=3, at least 50 cells per replicate. Two tailed Mann-Whitney test ****p < 0.0001)*

4.2.2 CSB depletion increases R-loop formation

Despite this decrease in overall transcription, the loss of CSB induces an increase in DNA:RNA hybrid signal (Fig 4.3A). This suggests that this transcription loss is at least partly associated with RNA polymerase stalling, leading to the stabilisation of R-loops within the negatively supercoiled genomic regions directly behind the elongation complex (Drolet et al., 2003). The signal from the S9.6 antibody is not limited to the nucleus however and is also present in the cytoplasm, which may represent some form of hybrid in the cytoplasm, mitochondrial R-loops or perhaps other RNA species (such as dsRNA) that the antibody has a lower affinity for.

This increase in R-loop levels is not limited to G1 as the DNA:RNA hybrid signal is also increased in cells during S/G2 (Figure 4.3B), further indicating the potential for transcriptional stress occurring during replication.

Consistent with previous studies, the loss of CSB caused a reduction in γH2AX levels (Figure 4.4A) that is associated with the reduction in the processing of aberrant R-loops into double-strand breaks (Sollier et al., 2014). Despite this reduction in γH2AX formation, overall levels of genomic instability appear to increase as these CSB deficient cells form more ultra-fine bridges (Figure 4.4B), which indicates cellular difficulty in separating chromosomes due to inefficient replication or damaged DNA (Chan et al., 2009; Liu et al., 2014).

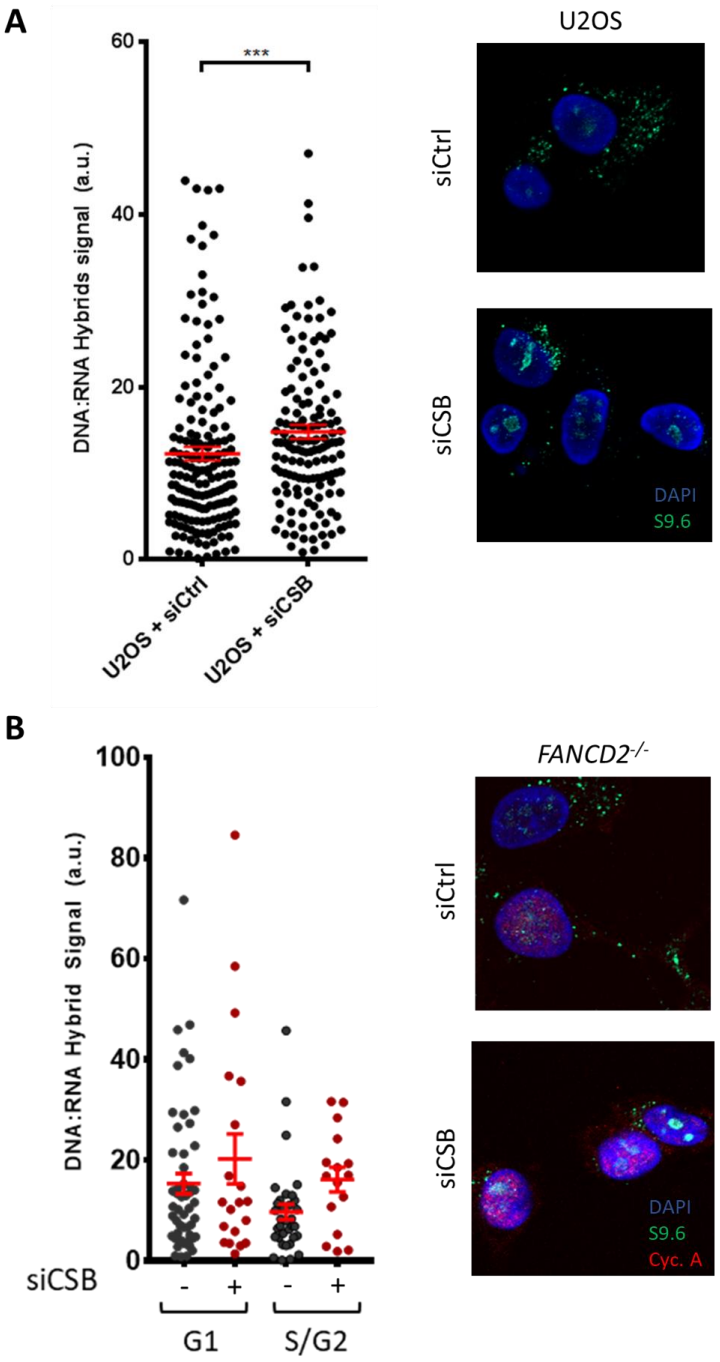


Figure 4.3: DNA:RNA hybrids are increased upon depletion of CSB.

*(A) Quantification of mean (±SEM) fluorescent intensity of individual nuclei from control and CSB depleted U2OS cells. (n=3, at least 50 cells per replicate. Two tailed Mann-Whitney test ***p < 0.001). (B) As in (A) but in conjunction with positive (S/G2) or negative (G1) cyclin A staining. (n=1, at least 20 cells per replicate).*

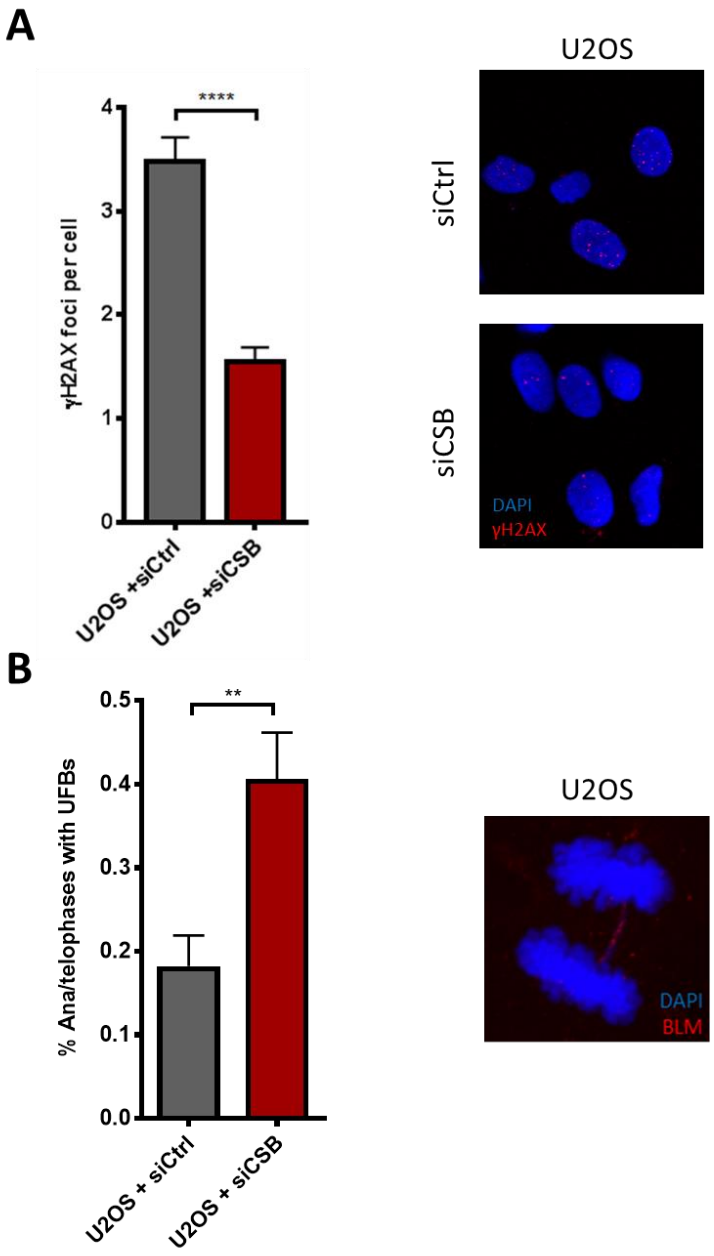


Fig 4.4: Depletion of CSB causes DNA damage and genome instability.

*(A) Quantified immunofluorescent analysis of mean (±SEM) γH2AX foci per cell from control and CSB depleted cells (n=3, at least 50 cells per replicate, Two tailed Mann-Whitney test ****p < 0.0001). (B) Quantification of the % of ana/telophase cells displaying BLM coated ultra-fine bridges (n=3, at least 50 cells per replicate, Two tailed Mann-Whitney test. **p < 0.01).*

4.2.3 Depletion of CSB causes replication-transcription conflict

In order to assess whether loss of TCR had an impact on replication fork progression, the DNA fibre technique was employed, which allows *in vivo* analysis of the DNA replication programme (Nieminuszczy et al., 2016). Visualization of actively progressing replication forks is achieved by consecutive labelling with two thymidine analogues, 5-chloro-2'-deoxyuridine (CldU) and 5-iodo-2'-deoxyuridine (IdU). Each label is added for 20 minutes to allow their incorporation into nascent DNA strands. Once labelled, cells are lysed, and DNA is stretched on a microscope slide before labelling with antibodies directed against halogenated nucleotides. Individual fibres can then be observed with fluorescence microscopy, allowing visualization of the progression of replication forks and quantification of replication fork stall/collapse (Figure 4.5).

Normally progressing replication forks can be observed as consecutive red and green tracts that have a length ratio close to 1 – indicating that the replisome proceeded unhindered through both labelling periods. Fibres that contain only red tracts represent termination events and stalled or collapsed forks.

Cells depleted of CSB had greater numbers of stalled/collapsed fork structures, suggesting that the transcription stress induced by loss of CSB causes replication forks to come into conflict with stalled RNA polymerase and/or their associated R -loops (Figure 4.5C).

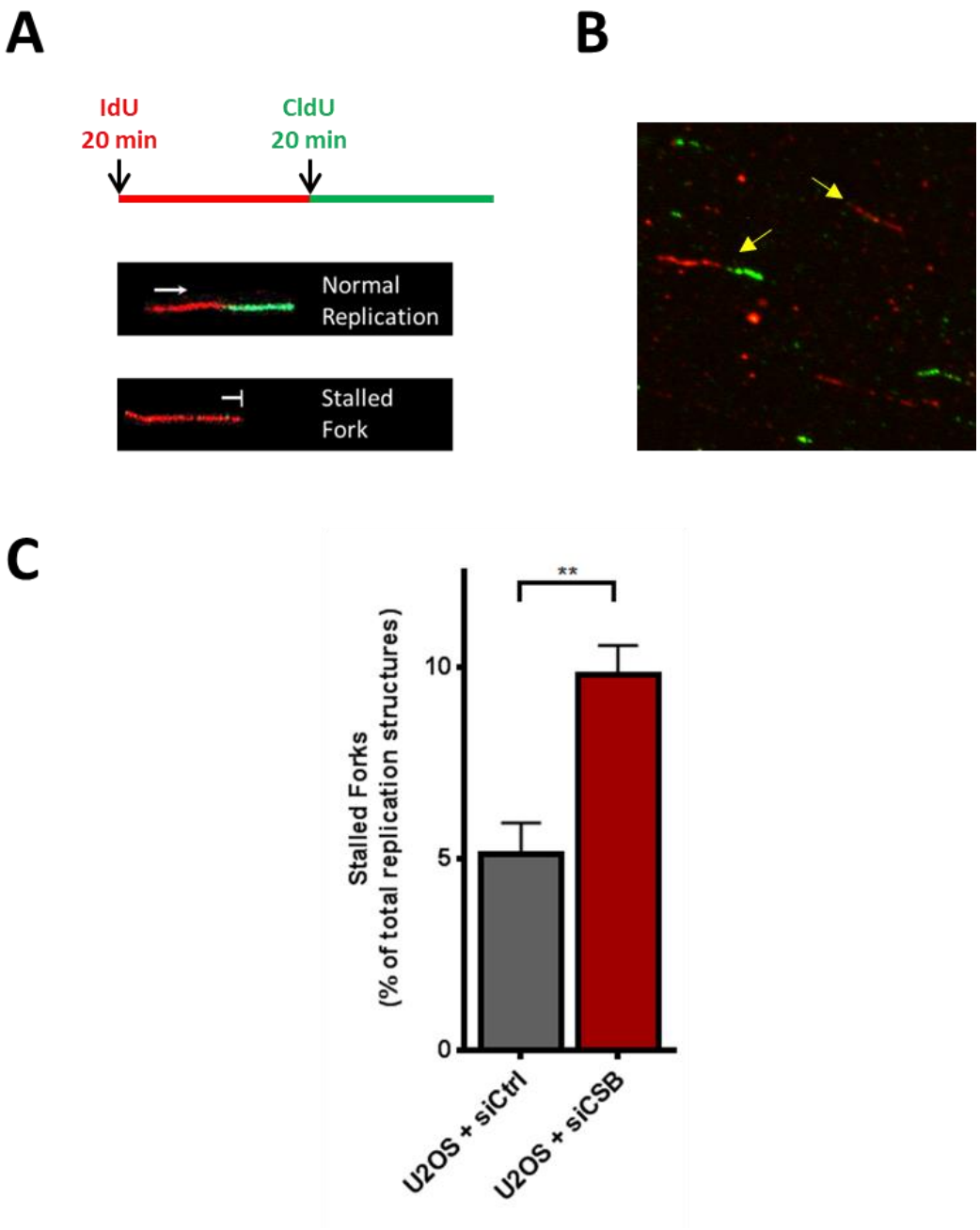


Figure 4.5: Loss of CSB causes replication fork stalling

*(A) Overview of DNA fibre fluorography technique for detection of normal and stalled replication structures. (B) Example of labelled DNA structures. (C) Mean (±SEM) stalled replication forks as a percentage of all replication structures (n=3, at least 80 structures per biological repeat, Two tailed Mann-Whitney test **p < 0.01).*

4.2.4 Inducing transcriptional stress in FA deficient cells by knockdown of CSB increases R-loops and DNA damage

The expression of CSB is required for the rescue of stalled RNAP II and efficient transcription (Figure 4.2) and its depletion appears to cause increased fork stalling (Figure 4.5) in U2OS cells, indicating increased transcription-replication conflict.

Since the FA pathway is activated in response to transcription, cells that are defective in their FA pathway response may be sensitive to the loss of proteins that prevent transcription associated cell stress. To investigate this, U2OS cells were made doubly deficient for both the FA and TCR pathways by siRNA treatment of CSB in the *FANCD2*$^{-/-}$ CRISPR knockout described in Chapter 3.

Interestingly, loss of CSB in (otherwise undamaged) FANCD2 proficient cells seems to cause an increase in FANCD2 ubiquitination and therefore activation of the FA pathway, further supporting transcriptional stress as an activator of FA. Additionally, CSB expression in FANCD2$^{-/-}$ U2OS cells is greater than in parental U2OS cells (Figure 4.6). It is tempting to speculate that in the absence of one of these proteins, the other is required to deal with the additional stress caused by replication-transcription machinery conflicts.

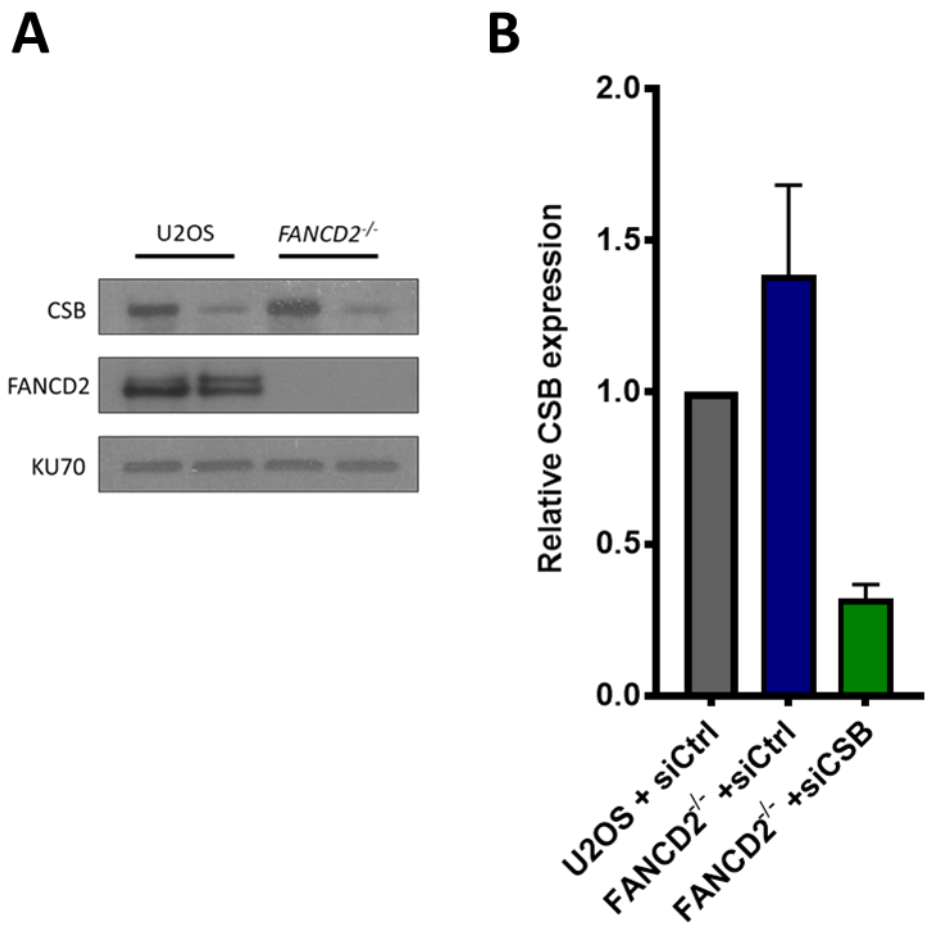


Figure 4.6: siRNA depletion of CSB in *FANCD2*$^{-/-}$ cells

(A) Western blot analysis of protein expression following CRISPR knockout of FANCD2 and treatment with 20 nM siRNA targeting either luciferase (control) or CSB. Cells were pulsed twice with siRNA and harvested 24 hours after second pulse. Ku70 serves as loading control.

(B) Quantification of (A) by densitometry.

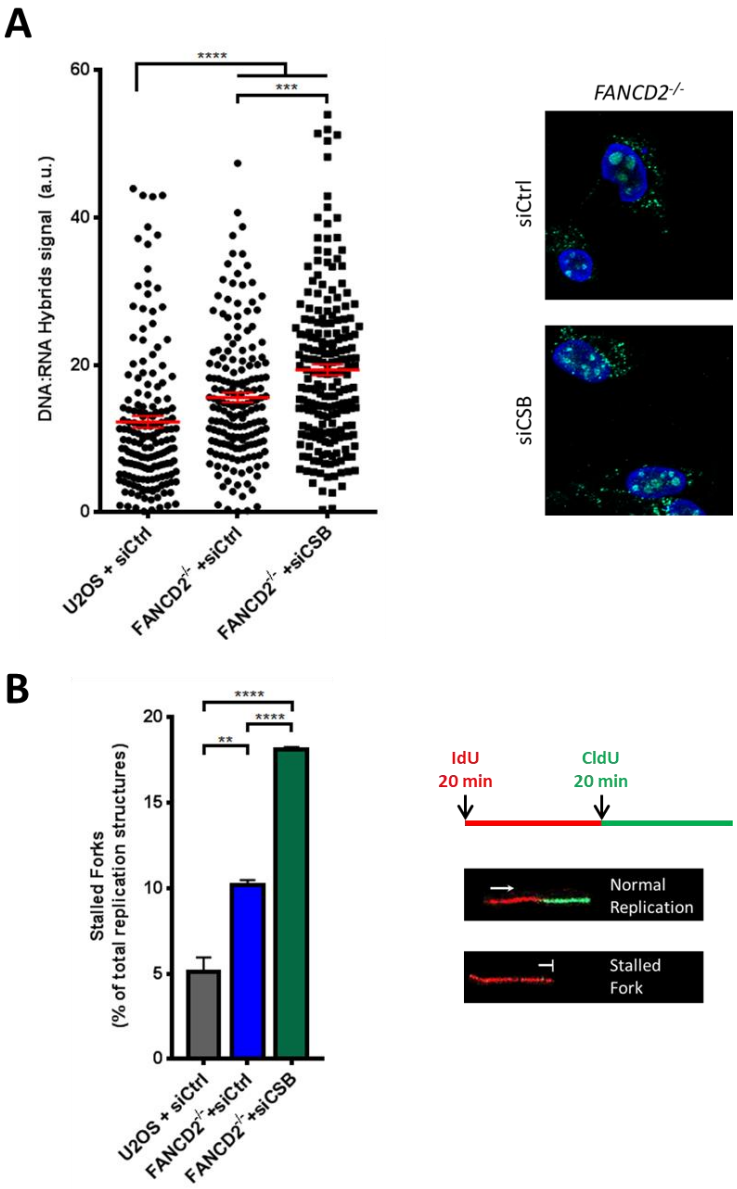


Figure 4.7: R-loops and DNA fibre analysis

*(A) Quantified mean (±SEM) immunofluorescent intensity with DNA:RNA hybrid specific antibody S9.6 (n=3, at least 50 cells per biological repeat, One-way ANOVA ****p < 0.0001).*

*(B) DNA fibre analysis of mean (±SEM) percentage of stalled fork structures in control, FANCD2 knockout or doubly deficient U2OS cells (n=3, at least 80 structures per biological repeat, One-way ANOVA ****p < 0.0001).*

The loss of both CSB and FANCD2 increases both DNA:RNA hybrid accumulation and the level of replication fork stalling, as measured by DNA fibre analysis (Figure 4.7).

The failure to stabilize or restart replication forks often leads to the induction of DNA damage (Alexander and Orr-Weaver, 2016; Cortez, 2015). One of the earliest detectable responses to replication stress and DNA damage is the phosphorylation of H2AX at serine 139 (γH2AX) by ATM or ATR, master regulators of the DNA damage response (Chanoux et al., 2009; Shiloh and Ziv, 2013).

Despite the reduction in γH2AX in the single knockdown of CSB attributed to loss of DNA breakage by NER machinery, immunofluorescent analysis of γH2AX foci formation in double deficient siCSB and *FANCD2*$^{-/-}$ cells indicates that overall DNA damage is increased (Figure 4.7A). This suggests that despite the reduction in γH2AX due to the lack of DNA incisions by NER nucleases, the instability caused by transcription-replication conflict due to the loss of both FANCD2 and CSB results in the accumulation of more DNA damage.

The formation of 53bp1 foci is a more specific indicator of the presence of DNA double-strand breaks than γH2AX (Schultz et al., 2000). Immunofluorescent analysis of 53bp1 foci in cells that express cyclin-A (indicative of cells in S/G2), indicates that there are increased double strand break events during S/G2 in the cells lacking both FANCD2 and CSB (Figure 4.7B). The same increase is not observed in cyclin-A negative cells (G1), indicating this damage is dependent upon replication.

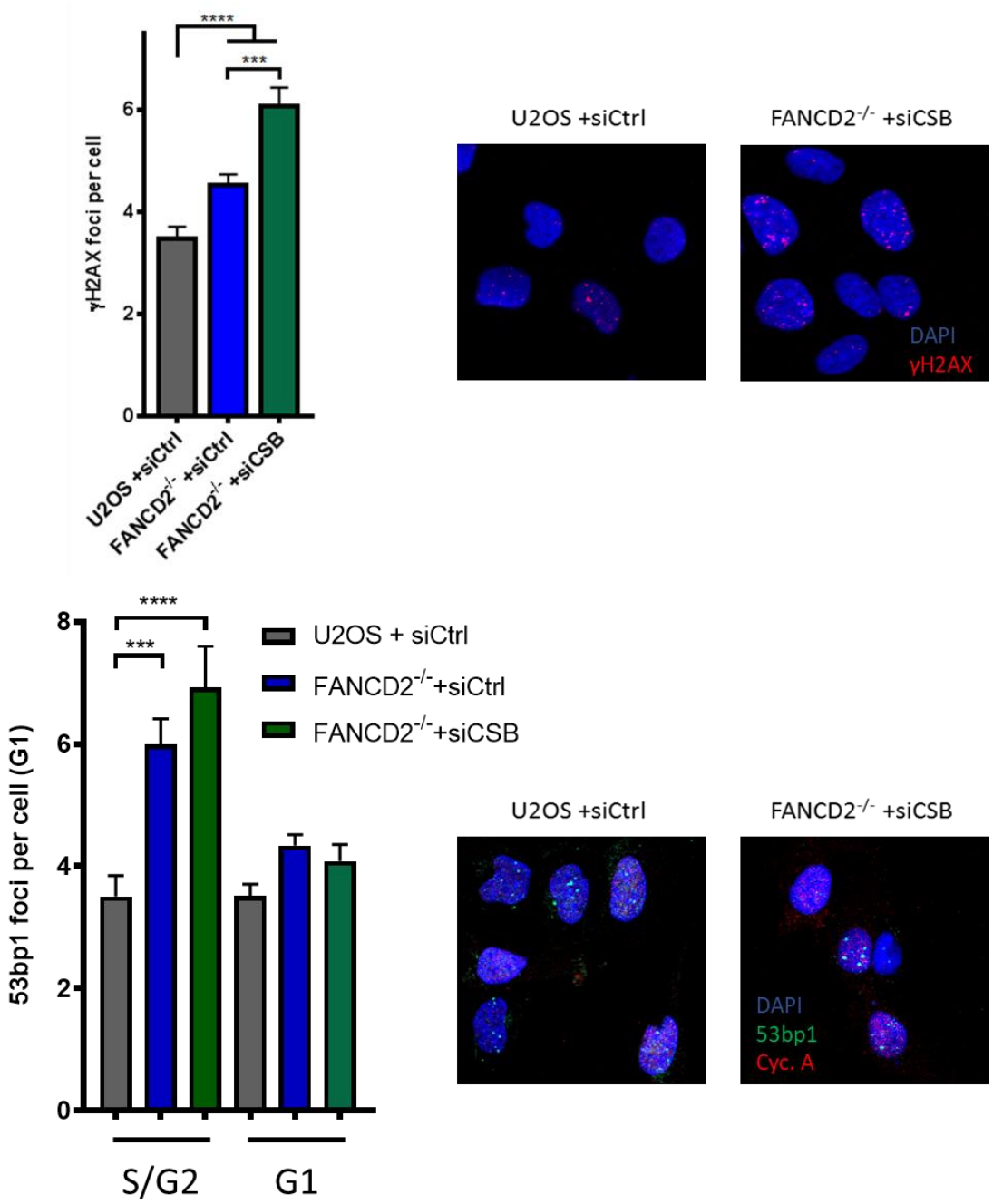


Figure 4.8: DNA damage in double deficient cells

*(A) Immunofluorescent analysis of mean (±SEM) γH2AX foci per cell (n=3, at least 100 cells per biological repeat, One-way ANOVA ****p < 0.0001).*

*(B) Mean (±SEM) 53bp1 foci formation in cyclin-A positive (S/G2) or negative (G1) cells. (n=3, at least 30 cells per biological repeat, One-way ANOVA ****p < 0.0001).*

4.2.5 Loss of CSB in FA deficient cells causes mitotic catastrophe and severely diminished rate of proliferation

When cells struggle to complete genome duplication efficiently, they can enter mitosis with damaged or under-replicated DNA. This has a severe impact on the ability of cells to properly segregate their chromosomes, leading to the formation of so-called chromatin, or anaphase, bridges (Gisselsson et al., 2000).

The inability to fully repair DNA damage and properly segregate chromosomes can lead to the formation of micronuclei. These are fragments of improperly segregated chromosomes that have broken off and exist outside the nucleus that are extensively used as biomarkers of chromosomal damage and genomic instability (Iarmarcovai et al., 2008).

Analysis of micronuclei was achieved by immunofluorescent staining of DAPI and α-tubulin. Whilst DAPI stains the DNA, tubulin serves as a marker of the cytoplasm, allowing the identification of small, DAPI stained micronuclei that are discrete from the nucleus which can then be counted. Anaphase bridge structures were counted after enrichment of ana/telophase cells by performing mitotic shake off (Bizard et al., 2018).

Accordingly, analysis of both anaphase bridges and micronuclei reveals U2OS cells deficient for both FANCD2 and CSB have a significantly higher proportion of these structures, indicating a high level of genomic instability (Figure 4.9).

Improper segregation of chromosomes prevents faithful duplication of the entire genome and leaves daughter cells without a full complement of genetic information, which can severely restrict future proliferative capability . Accordingly, double deficient cells have a severely diminished rate of proliferation (Figure 4.10) as measured by a resazurin viability assay (Riss et al., 2004).

The investigation of R-loop levels, DNA damage and genome instability in CSB depleted cells were performed separately at the start of this chapter but should have been included in the subsequent. Without their presence in Figure 4.6 through to 4.9 the comparison between these depletions is made more difficult.

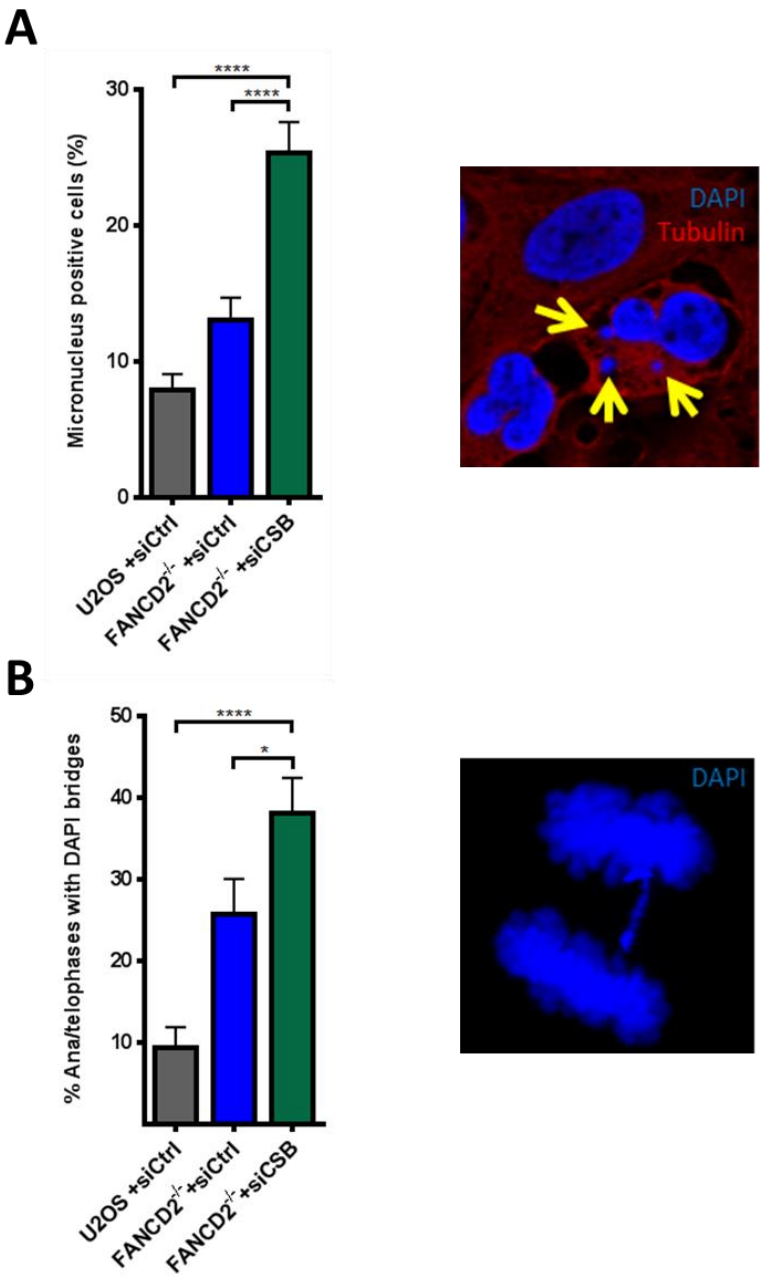


Figure 4.9: Micronuclei and anaphase bridge are indicators of genomic instability

*(A) Mean (±SEM) percentage of micronuclei positive cells. (n=3, at least 100 cells per biological repeat, One-way ANOVA ****p < 0.0001).*

*(B) Mean (±SEM) percentage of ana/telophase cells with chromatin bridges visible by DAPI staining of DNA (n=3, at least 50 cells per biological repeat, Two tailed Mann-Whitney test ****p < 0.0001).*

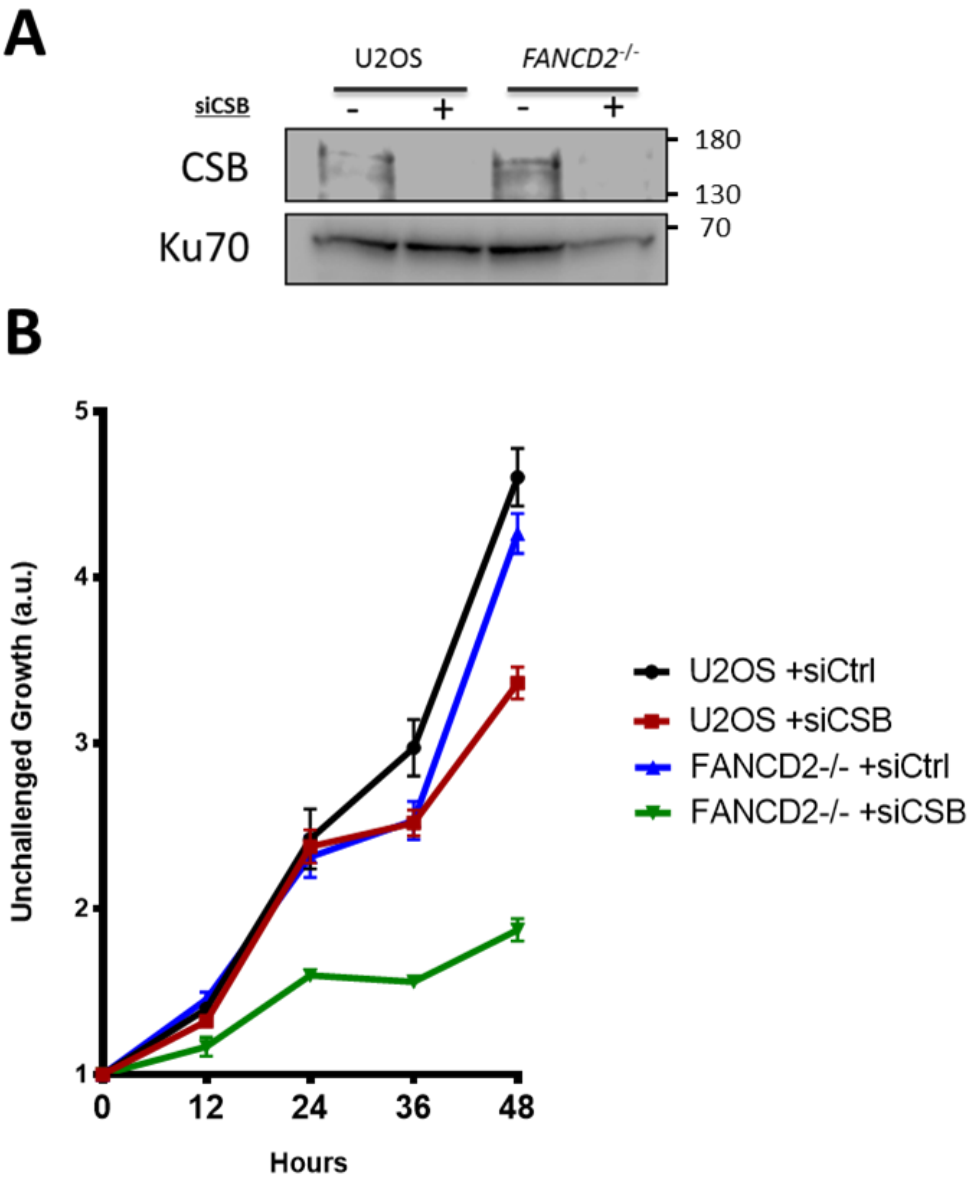


Figure 4.10: Depletion of CSB in FANCD2$^{-/-}$ cells causes a growth defect

(A) Western blot of CSB knockdown in wildtype and FANCD2$^{-/-}$ *U2OS cells*

(B) Cell growth of wildtype and FANCD2$^{-/-}$ U2OS cells transfected with 20 nM siRNA targeting CSB. Error bars show SD from triplicates.

4.2.6 Depletion of FANCD2 and CSB causes similar growth attrition in HeLa cells

Whilst U2OS cells are an effective model for genetic manipulation, their overall genomic stability is compromised and are one of many cancer cell lines that elongate telomeres by the alternative lengthening of telomeres (ALT) pathway (Deeg et al., 2016).

U2OS cells lack ATRX, which functions to suppress R-loop formation in transcribed telomeric repeats (Nguyen et al., 2017) and the growth defect upon CSB depletion in ALT cancer cell lines could be attributed to R-loop disruption at telomeres. In contrast, HeLa cells do not employ ALT and express functional ATRX and when depleted of FANCD2 and CSB by the transfection of two siRNA's, these cells also grow much slower in comparison to siRNA control and when FANCD2 and CSB are targeted individually (Figure 4.11).

Overall, this data demonstrates a reproducible growth defect in two different cancer cell lines using two combinations of gene depletion strategies.

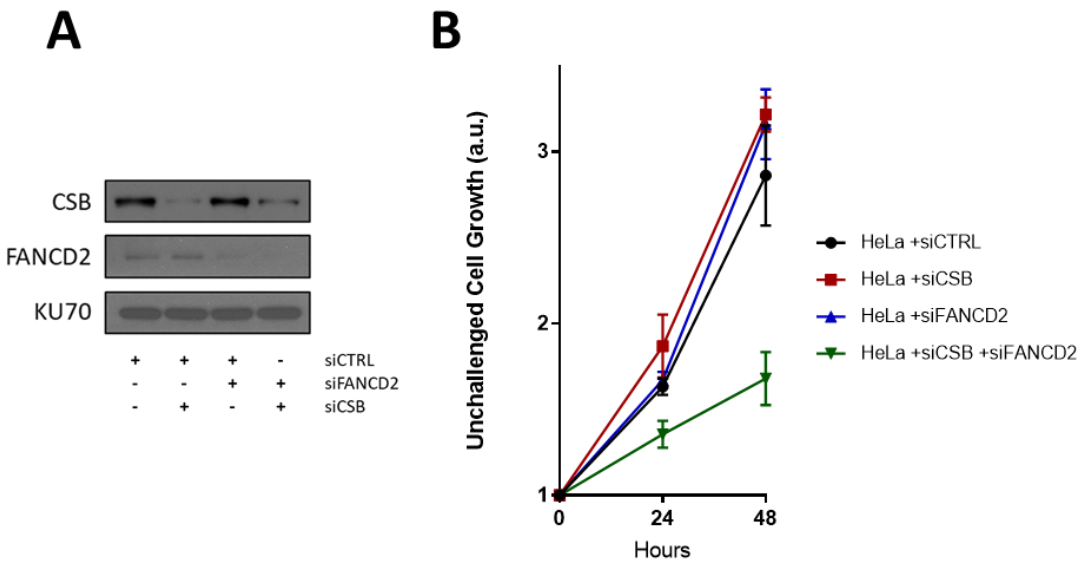


Figure 4.11: Double siRNA targeting of FANCD2 and CSB in HeLa cells

(A) Western blot demonstrating siRNA knockdown of FANCD2 and CSB. HeLa cells treated with siRNA targeting FANCD2 and CSB. Single knockdown transfections contained 20 nM of target siRNA plus 20nM control siRNA targeting luciferase. The double knockdown contained 20 nM of each siRNA.

(B) Quantification of cell growth after siRNA treatments. Error bars show SD from triplicates.

4.2.7 The requirement of CSB for viability in FA deficient cells does not extend to other members of the TC-NER pathway

Recently, the biochemical studies investigating the yeast CSB homolog, Rad26 suggests that the translocase function of this protein may promote the bypass of RNAP through certain lesions (Xu et al., 2017). However, larger obstacles require the recruitment of numerous other factors to overcome RNAP stalling. This includes NER factors to excise lesions and repair DNA to allow transcription to continue. R-loops have previously been reported to be excised by the NER machinery when they persist due to loss of DNA:RNA helicases (Sollier et al., 2014).

Although bulky, helix-distorting lesions that block transcription are predominantly repaired by TCR upon transcription stalling, global-genome NER (GG-NER) can also repair such lesions. GG-NER operates on all regions of DNA and senses damage through constant scanning of the genome by XPC, resulting in the recruitment of NER factors when damage is detected (Petruseva et al., 2014). To assess the requirement for NER factors in FA deficient cells, siRNA for XPC (involved specifically in GG-NER) and XPF (required in all NER processes and replication-coupled ICL repair) were used for depletion in *FANCD2*$^{-/-}$ cells in a similar manner to CSB. Neither XPC or XPF depletion resulted in a similar growth defect and proliferated comparably to wildtype U2OS and *FANCD2*$^{-/-}$ cells treated with a non-specific control siRNA (Figure 4.12).

This data suggests that it is a specific function of CSB, that does not require downstream NER actions, that results in the prevention of transcription becoming a source of replication stress.

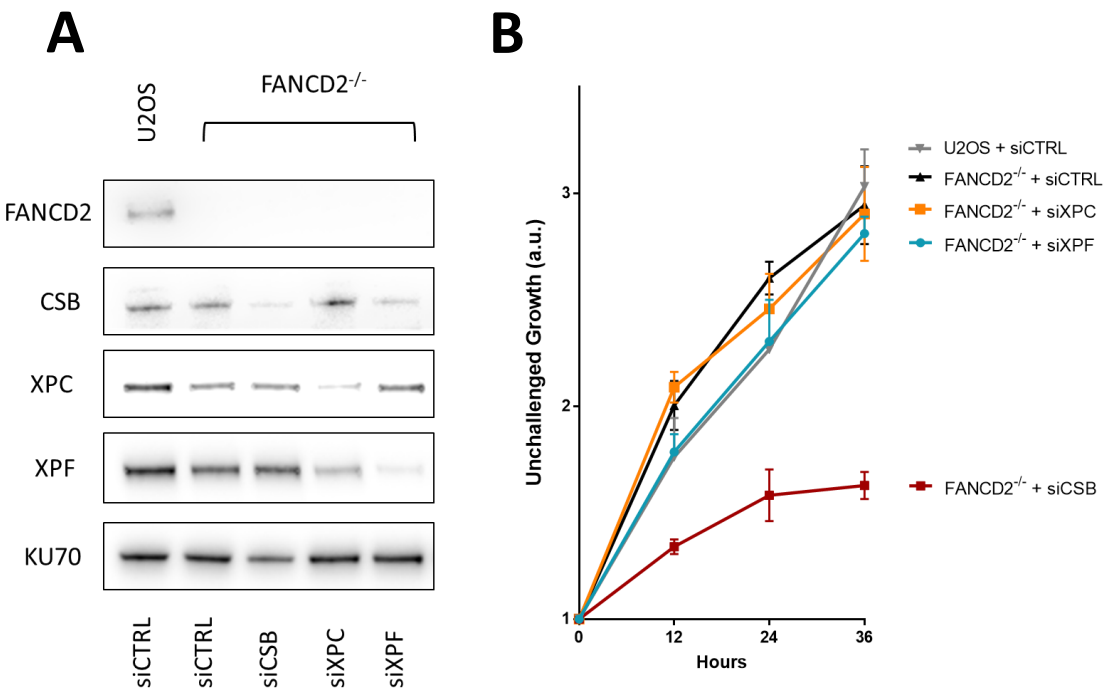


Fig 4.12: Loss of NER factors does not phenocopy the growth defect of CSB knockdown

(A) Western blot of CSB, XPC and XPF knockdown with double pulse of 20 nM siRNA. Cells harvested 24 hours after second pulse.

(B) Cell growth in FANCD2$^{-/-}$ U2OS cells after knockdown of NER factors XPC (specific to GG-NER) and XPF (required for all NER processes) in comparison to CSB. Error bars represent SD from triplicates.

4.3 Conclusion

Transcription is an endogenous source of replication stress as elongating RNA polymerases inevitably come into conflict with replication forks during S-phase (García-Muse and Aguilera, 2016; Helmrich et al., 2013). The activation of the FA pathway by FANCD2 monoubiquitination can be ablated by inhibition of transcription or enzymatic removal of R-loops, suggesting the FA pathway is normally activated when replication forks stall at sites of stalled transcription or R-loops (García-Rubio et al., 2015; Madireddy et al., 2016; Schwab et al., 2015).

This study describes the consequences of the loss of the CSB protein, the main orchestrator of TCR, which include replication fork stalling, R-loop formation and activation of the FA pathway. In the absence of the FA pathway, loss of CSB leads to significant replication-transcription conflict that causes persistent R-loops and stalled replication, leading to fork collapse and double-strand breaks. Under-replicated and damaged DNA persists to mitosis, where it presents a profound problem for the successful separation of chromosomes. This effect appears to be specific to CSB as depletion of downstream NER factors that usually work to excise lesions in TCR do not produce the same effect.

Taken together, this data supports recent findings that transcription and R-loops are important endogenous sources of replication stress and DNA damage that requires the concerted action of the FA pathway.

5. Discussion

Cellular viability is compromised in the absence of efficient DNA replication and transcription, as both processes require access to the DNA sequence to synthesize polynucleotides and ribonucleotides. When these complexes operate at the same loci, the transcription complex can block the progression of the replication fork and it is likely that RNA polymerase (RNAP) must be released from chromatin to allow the passage of the replisome, as is the case in bacteria (Pomerantz and O'Donnell, 2010).

How this process is managed remains unclear, but failure to properly mediate these conflicts leads to DNA damage, genomic instability and replication defects (Helmrich et al., 2013). *S. cerevisiae* cells that express mutant RNAPs that are not easily dissociated from chromatin accumulate the hallmarks of DNA damage, genome instability and replication defects (Felipe-Abrio et al., 2015). This suggests that replication forks that encounter and stall at elongation complexes must be stabilised and protected whilst the remodelling, bypass or disassembly of the transcription machinery is orchestrated. Loss of factors involved in replication fork stability or involved in removal and repair of stalled transcription complexes would likely aggravate the detrimental outcomes of transcription-replication conflicts.

5.1 The FA pathway response to transcription-associated replication stress

The FA pathway has several roles in responding to cellular stress that is caused by the stalling of replication forks. FA factors are required to prevent the unscheduled degradation of intermediate structures after replication stalling (Schlacher et al., 2012)

and also ensure that stalled forks are restarted after obstacles are removed (Chaudhury et al., 2013; Lachaud et al., 2016; Raghunandan et al., 2015).

General inhibition of transcription by several compounds (including cordycepin, flavopiridol and DRB) with varying mechanisms of action, results in a decrease in observable FANCD2 foci and a reduction in the mono-ubiquitinated form of FANCD2, demonstrating that the FA pathway is activated in response to active transcription (Schwab et al., 2015). The loss of the FA pathway results in the accumulation of transcription-associated R-loops, which can block replication fork progression and the removal of these structures can be facilitated by the translocase activity of FANCM (García-Rubio et al., 2015; Hodson et al., 2018; Schwab et al., 2015).

The current study has employed CRISPR-mediated gene editing to knockout FANCD2 and supports previous findings that R-loops accumulate upon the depletion of FA factors by RNAi. In addition, this study demonstrates this R-loop accumulation is not limited to cancer cell lines, as lymphoblastoid cells derived from several patients from multiple FA complementation groups also display a higher R-loop load. This increased R-loop load is also relevant to the main cellular compartment affected in individuals with FA, the haematopoietic stem cell (HSC) compartment.

One of the key questions for the FA field is why loss of the FA pathway results in the specific depletion of the HSC compartment. The requirement for LT-HSCs to repopulate blood cell lineages after physiological stress (for example, infections and blood loss)

results in rapid transition from quiescence into an active cell cycle (Walter et al., 2015) and would involve a burst in transcriptional activity that could cause substantial transcription-replication conflict when cells enter S-phase.

When murine quiescent LT-HSCs are damaged with ionizing radiation (IR), they display a resistance to DNA damage in comparison with cycling cells (Mohrin et al., 2010). In contrast to this, human LT-HSCs appear to be sensitive to IR and have enhanced p53 mediated apoptosis (Milyavsky et al., 2010). This inherent resistance of murine HSCs to damage and the difference in their human counterparts may partially explain why FA deficient mice do not recapitulate the human phenotype of bone marrow failure.

Given the FA pathway role in responding to transcription stress and preventing R-loop accumulation, it was postulated that FA deficient cells may be sensitive to the loss of processes that promote efficient transcription and prevent RNAP stalling. Accordingly, disrupting the cellular ability to rescue stalled RNAP through the depletion of CSB results in replication fork stalling and genomic instability - likely from increasing the occurrence of transcription-replication conflict. The accurate resolution of these conflicts requires the action of the FA pathway to stabilise and protect stalled replication forks.

In the absence of the FA pathway, loss of CSB severely diminishes the proliferative capacity of cells. The absence of fork protection and restart mechanisms results in the degradation and/or collapse of forks and the formation of DSBs. The increase in RNAP obstacles appears to delay replication fork progression, causing under-replicated DNA

to enter mitosis and the subsequent improper segregation of chromosomes due to the formation of chromatin bridges. Interestingly, chromosome fragile sites (CFSs) are frequently interlinked by BLM-associated ultra-fine DNA bridges (Chan et al., 2007). Several CFSs occur in long genes where transcription-replication conflict is inevitable as they take more than one cell cycle to transcribe, and FA factors and the formation of R-loops are both associated with CFSs and the formation of UFBs (Zhang et al., 2018). This association may, in part, be due to the requirement for FA factors at sites of transcription-replication conflict at CFSs.

A major aim of the current study that did not reach completion was to investigate whether loss of other members of FA pathway besides FANCD2 shared a similar growth defect when lost in conjunction with CSB. A considerable effort was made to produce a CRISPR knockout of CSB to undertake experiments with reverse the genetic manipulation; that is, knockout of CSB in conjunction with RNAi of FA factors. This would have yielded a model system in which to probe the requirement of the different subsets of the FA pathway in mediating transcription-replication conflict. However, attempts to genetically manipulate U2OS, HeLa and RPE-1 cells failed to produce a convincing CSB knockout from over 500 GFP-positive single clones, of which ~70 had viable growth and were screened by western blot. To the author's knowledge, there exists no study which includes a diploid cell knockout of CSB via CRISPR-cas9 gene editing. This may indicate that at least some CSB expression is essential for cancer or transformed cell line viability.

5.2 Targeting transcription-replication conflict in cancer therapy

The accumulation of DNA damage and the reduction of growth upon depletion of CSB in FANCD2$^{-/-}$ cells suggests that the perturbation of processes associated with rescue of RNAP or efficient transcription could hint at novel strategies to prevent the proliferation of cancer cells, as FA factors are often mutated in wide range of tumours (Nalepa and Clapp, 2018; Shen et al., 2015). However, these experiments have not been performed in untransformed cells and as such the selectivity for cancer cells is not yet clear.

Trabectedin (Ecteinascidin 743 or ET-743, marketed as Yondelis) is a tetrahydroisoquinoline alkaloid that shows a broad activity spectrum toward tumour cell lines (Poindessous et al., 2003) and has been employed clinically in the treatment of soft tissue and bone sarcomas as well as ovarian cancer (Le Cesne et al., 2005; Sessa et al., 2005).

Cells deficient in HR factors are sensitive to ET-743, whilst Ku80 or DNA-PK deficient cells are not (Soares et al., 2007). Human cells with deficiencies in FA factors, including FANCA, FANCC, FANCF, FANCG and FANCD1 are also sensitive to ET-743 treatment (Casado et al., 2008). ET-743 binds DNA in the minor groove, distorting the helix by bending it toward the major groove whilst also appearing to mimic an ICL by spanning both strands (Hurley and Zewail-Foote, 2001). The sensitivity of FA deficient cell lines to ET-743 was attributed to the ICL-like characteristic of the adduct.

However, cells that are deficient in a range of TC-NER factors, including CSB, are 2-10 times less sensitive to ET-743 than wildtype (Takebayashi et al., 2001), in direct contrast to other ICL-inducing agents such as cisplatin (Furuta et al., 2002).

This resistance is not observed in XPC deficient cells indicating that TC-NER, but not GG-NER, is required for the cytotoxic action of this compound. The resistance of TC-NER deficient cells can be reverted by the complementation of wildtype factors (Damia et al., 2001).

This unique property of ET-743 suggests that cytotoxicity of the drug is dependent upon the response and failed action of the TC-NER machinery. Treatment with trabectedin results in proteasome-dependent degradation of RNAP, an aspect of TC-NER that occurs as a last resort when it fails to properly repair DNA lesions (Wilson et al., 2013). Given the role of FA deficient cells in responding to transcription-associated stress, the cytotoxicity of ET-743 in FA deficient cells could be the result of transcription-replication conflict at sites where this compound has bound DNA and trapped TC-NER machinery (Aune et al., 2008). The lack of TC-NER repair may cause prolonged RNAP stalling and R-loop formation and could be assayed by the quantification of RNA:DNA hybrids in treated cells. Additionally, if R-loop formation is increased upon treatment, whether sensitivity of FA (and other HR) deficient cells could be rescued by RNase H overexpression could indicate if R-loops are relevant to the toxicity of this drug.

Whilst the response of the FA pathway to exogenous compounds that introduce ICLs is well characterised, there is little direct indication that a substantial number of these lesions form *in vivo* from endogenous sources. In addition, ICLs usually represent a small proportion of the lesions that many of these compounds introduce but are usually attributed with the main cytotoxic mechanism of these drugs (Schärer, 2005). For example, whilst platinum-based compounds do introduce ICLs, these structures represent less than 10% of total lesions, with some estimates as low as 1-5% (Blommaert et al., 1995; Fichtinger-Schepman et al., 1985; Schärer, 2005).

Lesions that link bases on the same DNA strand are called intrastrand cross-links, which represent the majority of lesions introduce by drugs such as cisplatin (Damsma et al., 2007). Intrastrand cross-links induced by cisplatin (and other sources such as UV) block RNAP II progression and lead to stable polymerase stalling (Jung and Lippard, 2006), and are mainly repaired by NER (Schärer, 2013).

Given the above and considering the role of the FA pathway in responding to transcription associated stress, many of the compounds that the FA pathway is sensitive to may, at least in part, exert their toxicity by stalling RNAPs and causing substantial replication-transcription conflict that, in the absence of the FA pathway, leads to replication fork collapse and/or degradation, resulting in DNA damage and genomic instability.

Compounds like ET-743 that prevent TCR and/or block the progression of RNAP and increase the rate of transcription-replication conflicts could be a novel strategy to target tumours that lack functional FA pathways.

5.3 The contribution of transcription-replication conflict to the pathology of FA

Reactive aldehydes are endogenously produced compounds that cause DNA damage that requires the action of the FA pathway (Garaycoechea et al., 2018). Physiologically relevant aldehydes can react with DNA to produce a variety of adducts including ICLs, intrastrand cross-links (Matsuda et al., 1998) and DNA-protein cross-links (Shaham et al., 1996), however the rates of occurrence are not clear.

Since the FA pathway was revealed to function in DNA repair, the question of its molecular pathogenesis has lingered – how does deficiency in a DNA repair pathway result in the specific and progressive loss of the HSC compartment? Substantial evidence has accumulated that reactive aldehydes are drivers of bone marrow failure in FA-deficient mice (Langevin et al., 2011), but the precise lesion(s) that the FA pathway acts upon is not yet defined.

However, treatment of FA deficient cells with formaldehyde significantly increases the formation of R-loops (Schwab et al., 2015), indicating that aldehyde treatment increases transcription stress in these cells and could do so by creating intrastrand DNA cross-links. Even so, why this results in the specific loss of HSCs is not clear, but there is some indication that haematological issues can arise through the disturbance of normal ribosome biogenesis.

Ribosomes are critical for living organisms as they facilitate the translation of mRNA to protein in gene expression. Ribosomes are composed from ribosomal RNA (rRNA) and

ribosomal proteins and their organization is conserved among most eukaryotic organisms. In *S. cerevisiae*, the cellular demand for ribosomes is enormous, occupying 50% of the total protein mass in a cell and 60% of total transcription – an estimated ~2000 ribosomes are produced per minute per cell (Warner, 1999).

The RNA components of ribosomes are critical for its function. The rDNA is encoded by ribosomal DNA (rDNA) which, in humans, is found on five rDNA clusters on the short arms of acrocentric chromosomes 13, 14, 15, 21 and 22 and organized into tandem repeats of ~70 units (Sakai et al., 1995). Due to the high demand for ribosomes rDNA is constantly transcribed and is therefore vulnerable to transcription-replication conflict. rDNA is one of the most unstable regions in the genome due to deleterious recombination among repeats (Kobayashi, 2014). The repetitive rDNA arrays of bacteria, yeast and mammals have all been shown to be hotspots for R-loop formation and are highlighted in immunofluorescent S9.6 signal (El Hage et al., 2010; Massé et al., 1997; Nadel et al., 2015).

Mutations in genes involved in ribosomal biogenesis are the genetic cause of several inherited anaemias that share clinical symptoms with FA. These include dyskeratosis congenita (DC), cartilage-hair hypoplasia (CHH), Diamond-Blackfan anaemia (DBA), and Schwachman-Diamond syndrome (SDS). The main clinical symptom of these diseases is bone marrow failure but, similar to FA, display diverse characteristics such as abnormal skin pigmentation, short stature, developmental defects and a predisposition to haematological malignancies (Liu and Ellis, 2006). The attrition of haematopoietic

progenitor cells in these diseases is widely associated with accelerated apoptotic death (Dror and Freedman, 2001; Knudson et al., 2005; Perdahl et al., 1994; Yel et al., 1999).

There is some association of FA factors with ribosome biogenesis. FANCI has been found to be enriched in the nucleolus and its depletion results in reduced transcription by RNAP I (Sondalle et al., 2016). Whilst no bone marrow failure is reported in Cockayne syndrome patients, mutations in a gene with significant homology called ERCC6-like-2 (ERCC6L2), result in similar clinical features to CS, but includes bone marrow failure (Tummala et al., 2014; Zhang et al., 2016).

The depletion of factors involved in rRNA transcription and/or processing triggers acetylation of p53 and the induction of apoptosis (Kumazawa et al., 2015). Interestingly, p53 mediated apoptosis is implicated in driving the loss in number and function of HSCs in *Aldh2*$^{-/-}$ */ Fancd2*$^{-/-}$ mice, and HSC attrition can be rescued in these mice by the further deletion of *Trp53*, at the expense of genome stability (Garaycoechea et al., 2018). This finding might suggest that aldehyde-induced lesions might result in rDNA instability and activation of p53.

To this author's knowledge, there is no detailed investigation into the stability of rDNA in FA deficient HSCs. Given the above, it is tempting to speculate that the loss of FA pathway activity could result in instability at rDNA due to aldehyde-lesion induced transcription-replication conflict. This instability could result in the failure of cycling HSCs to meet the huge cellular demand for ribosomal biogenesis and the induction of

p53-mediated apoptosis, leading to the eventual depletion of HSCs and the loss of mature lineages in FA.

Evidence for rDNA instability in the attrition of stem cells could initially be assessed in in yeast, where rDNA stability of FA deficient strains could be monitored over successive generations. This could be performed unchallenged, but also in the presence of compounds of interest such as aldehydes or agents that induce transcription blocking lesions. However, this type of instability at rDNA may only be observable in FA deficient HSCs – it would be interesting to look at the rDNA stability of FA deficient mouse LT-HSCs treated similarly to the experiments in Chapter 3.2.4. In the current study, the quantification of the S9.6 signal that excludes or is limited to the nucleolar signal may have provided some indication as to whether this compartment is specifically affected in FA deficient cell lines.

5.4 Conclusion

The data described within appears to support the notion that replication fork collisions with transcription complexes or R-loops represents a major substrate of the FA pathway. This may encourage the speculation that transcription-replication conflict can contribute not only to the DNA damage and genomic instability that results in progression to malignancy but could also result in the depletion of HSCs through perturbed ribosome biogenesis, making a substantial contribution to the molecular pathogenesis of FA. Figure 5.1 describes that the loss of FA pathway may result in the loss of protection of replication forks that have encountered RNA polymerase or R-loops. Without RNAP removal or R-loop resolution, replication forks stall and are vulnerable to degradation and collapse, resulting in DNA damage and genome instability.

However, many of the conclusions drawn from the data are made without some important controls, including the proper assessment of the extent to which the S9.6 signal is due to non-specific binding of other RNA structures. In addition, although the phenotypes observed through the knockdown of CSB may, in part, indicate that FA is involved in resolving genomic instability and damage arising from transcription, this limited by the fact that CSB is also a DNA repair factor. Similar experiments involving transcription associated factors, already known to be involved in preventing R-loop formation and associated damage and instability, that aren't involved in DNA repair would be prudent to support the current study.

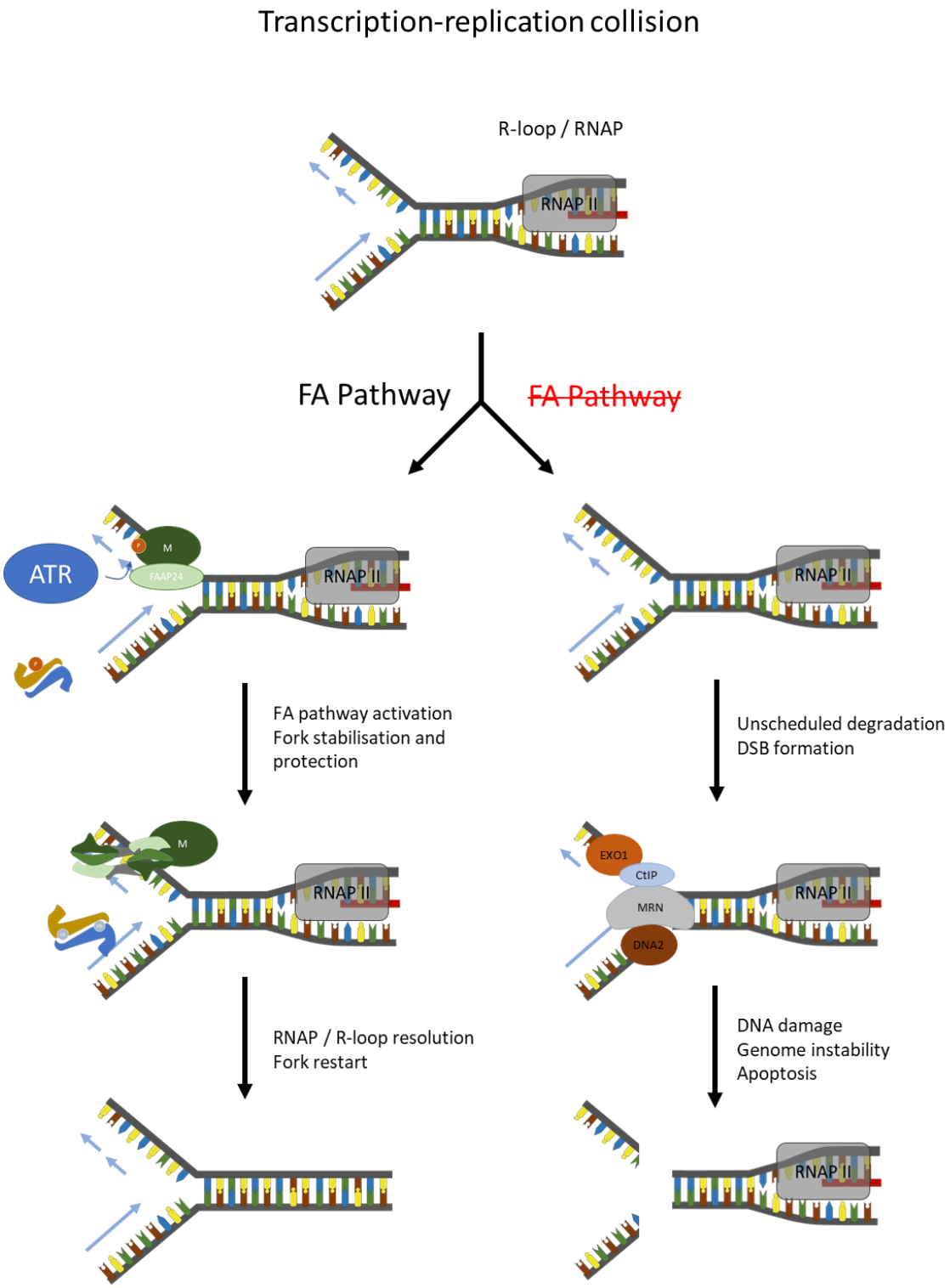


Figure 5.1: The FA pathway protects replication forks stalled at transcription

RNAP or R-loop blockage of the replisome activates the FA pathway to stabilise and protect replication forks from degradation and to restart forks after R-loop resolution and/or RNA polymerase removal

6. References

Abdullah, U.B., J.F. McGouran, S. Brolih, D. Ptchelkine, A.H. El-Sagheer, T. Brown, and P.J. McHugh. 2017. RPA activates the XPF-ERCC1 endonuclease to initiate processing of DNA interstrand crosslinks. *Embo J*. 36:2047-2060.

Ahnesorg, P., P. Smith, and S.P. Jackson. 2006. XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining. *Cell*. 124:301-313.

Alexander, J.L., and T.L. Orr-Weaver. 2016. Replication fork instability and the consequences of fork collisions from rereplication. *Gene Dev*. 30:2241-2252.

Anderson, L., C. Henderson, and Y. Adachi. 2001. Phosphorylation and rapid relocalization of 53BP1 to nuclear foci upon DNA damage. *Mol Cell Biol*. 21:1719-1729.

Andrade-Lima, L.C., A. Veloso, M.T. Paulsen, C.F.M. Menck, and M. Ljungman. 2015. DNA repair and recovery of RNA synthesis following exposure to ultraviolet light are delayed in long genes. *Nucleic Acids Res*. 43:2744-2756.

Andreassen, P.R., A.D. D'Andrea, and T. Taniguchi. 2004. ATR couples FANCD2 monoubiquitination to the DNA-damage response. *Genes Dev*. 18:1958-1963.

Apostolou, P., and I. Papasotiriou. 2017. Current perspectives on CHEK2 mutations in breast cancer. *Breast Cancer : Targets and Therapy*. 9:331-335.

Araki, M., C. Masutani, M. Takemura, A. Uchida, K. Sugasawa, J. Kondoh, Y. Ohkuma, and F. Hanaoka. 2001. Centrosome Protein Centrin 2/Caltractin 1 Is Part of the Xeroderma Pigmentosum Group C Complex That Initiates Global Genome Nucleotide Excision Repair. *Journal of Biological Chemistry*. 276:18665-18672.

Auerbach, A.D. 2009. Fanconi Anemia and its Diagnosis. *Mutation research*. 668:4-10.

Auerbach, A.D., Q. Liu, R. Ghosh, M.S. Pollack, G.W. Douglas, and H.E. Broxmeyer. 1990. Prenatal identification of potential donors for umbilical cord blood transplantation for Fanconi anemia. *Transfusion*. 30:682-687.

Aune, G.J., K. Takagi, O. Sordet, J. Guirouilh-Barbat, S. Antony, V.A. Bohr, and Y. Pommier. 2008. Von Hippel-Lindau-coupled and transcription-coupled nucleotide excision repair-dependent degradation of RNA polymerase II in response to trabectedin. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 14:6449-6455.

Aymard, F., B. Bugler, C.K. Schmidt, E. Guillou, P. Caron, S. Briois, J.S. Iacovoni, V. Daburon, K.M. Miller, S.P. Jackson, and G. Legube. 2014. Transcriptionally active chromatin recruits homologous recombination at DNA double strand breaks. *Nat Struct Mol Biol*. 21:366-374.

Azvolinsky, A., P.G. Giresi, J.D. Lieb, and V.A. Zakian. 2009. Highly Transcribed RNA Polymerase II Genes Are Impediments to Replication Fork Progression in Saccharomyces cerevisiae. *Mol Cell*. 34:722-734.

Bacolla, A., and R.D. Wells. 2004. Non-B DNA Conformations, Genomic Rearrangements, and Human Disease. *Journal of Biological Chemistry*. 279:47411-47414.

Bagnara, G., P. Strippoli, L. Bonsi, M. Brizzi, G. Avanzi, F. Timeus, U. Ramenghi, G. Piaggio, J. Tong, and M. Podesta. 1992. Effect of stem cell factor on colony growth from acquired and constitutional (Fanconi) aplastic anemia. *Blood*. 80:382-387.

Bahmed, K., A. Seth, K.C. Nitiss, and J.L. Nitiss. 2011. End-processing during non-homologous end-joining: a role for exonuclease 1. *Nucleic Acids Res*. 39:970-978.

Bakkenist, C.J., and M.B. Kastan. 2004. Initiating cellular stress responses. *Cell*. 118:9-17.

Bartek, J., and J. Lukas. 2003. Chk1 and Chk2 kinases in checkpoint control and cancer. *Cancer cell*. 3:421-429.

Bennett, C.B., A.L. Lewis, K.K. Baldwin, and M.A. Resnick. 1993. Lethality induced by a single site-specific double-strand break in a dispensable yeast plasmid. *Proc Natl Acad Sci U S A*. 90:5613-5617.

Bernad, A., L. Blanco, J.M. Lazaro, G. Martin, and M. Salas. 1989. A conserved 3'----5' exonuclease active site in prokaryotic and eukaryotic DNA polymerases. *Cell*. 59:219-228.

Betermier, M., P. Bertrand, and B.S. Lopez. 2014. Is non-homologous end-joining really an inherently error-prone process? *PLoS Genet*. 10:e1004086.

Bhatia, V., S.I. Barroso, M.L. García-Rubio, E. Tumini, E. Herrera-Moyano, and A. Aguilera. 2014. BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. *Nature*. 511:362.

Bizard, A.H., and I.D. Hickson. 2014. The Dissolution of Double Holliday Junctions. *Cold Spring Harbor Perspectives in Biology*. 6.

Bizard, A.H., C.F. Nielsen, and I.D. Hickson. 2018. Detection of Ultrafine Anaphase Bridges. *Methods in molecular biology (Clifton, N.J.)*. 1672:495-508.

Blackford, A.N., and S.P. Jackson. 2017. ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. *Mol Cell*. 66:801-817.

Blommaert, F.A., H.C. van Dijk-Knijnenburg, F.J. Dijt, L. den Engelse, R.A. Baan, F. Berends, and A.M. Fichtinger-Schepman. 1995. Formation of DNA adducts by the anticancer drug carboplatin: different nucleotide sequence preferences in vitro and in cells. *Biochemistry*. 34:8474-8480.

Boguslawski, S.J., D.E. Smith, M.A. Michalak, K.E. Mickelson, C.O. Yehle, W.L. Patterson, and R.J. Carrico. 1986. Characterization of monoclonal antibody to DNA · RNA and its application to immunodetection of hybrids. *Journal of Immunological Methods*. 89:123-130.

Bohr, V.A., and R. Michael Anson. 1995. DNA damage, mutation and fine structure DNA repair in aging. *Mutation Research/DNAging*. 338:25-34.

Bradsher, J., J. Auriol, L. Proietti de Santis, S. Iben, J.L. Vonesch, I. Grummt, and J.M. Egly. 2002. CSB is a component of RNA pol I transcription. *Mol Cell*. 10:819-829.

Brash, D.E., A. Ziegler, A.S. Jonason, J.A. Simon, S. Kunala, and D.J. Leffell. 1996. Sunlight and sunburn in human skin cancer: p53, apoptosis, and tumor promotion. *The journal of investigative dermatology. Symposium proceedings*. 1:136-142.

Brewer, B.J. 1988. When polymerases collide: Replication and the transcriptional organization of the E. coli chromosome. *Cell*. 53:679-686.

Britton, S., E. Dernoncourt, C. Delteil, C. Froment, O. Schiltz, B. Salles, P. Frit, and P. Calsou. 2014. DNA damage triggers SAF-A and RNA biogenesis factors exclusion from chromatin coupled to R-loops removal. *Nucleic Acids Res*. 42:9047-9062.

Brustel, J., Z. Kozik, N. Gromak, V. Savic, and S.M.M. Sweet. 2018. Large XPF-dependent deletions following misrepair of a DNA double strand break are prevented by the RNA:DNA helicase Senataxin. *Scientific Reports*. 8:3850.

Budzowska, M., T.G. Graham, A. Sobeck, S. Waga, and J.C. Walter. 2015. Regulation of the Rev1-pol zeta complex during bypass of a DNA interstrand cross-link. *Embo J*. 34:1971-1985.

Burge, S., G.N. Parkinson, P. Hazel, A.K. Todd, and S. Neidle. 2006. Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res*. 34:5402-5415.

Carter, T., I. Vancurova, I. Sun, W. Lou, and S. DeLeon. 1990. A DNA-activated protein kinase from HeLa cell nuclei. *Mol Cell Biol*. 10:6460-6471.

Casado, J.A., P. Rio, E. Marco, V. Garcia-Hernandez, A. Domingo, L. Perez, J.C. Tercero, J.J. Vaquero, B. Albella, F. Gago, and J.A. Bueren. 2008. Relevance of the Fanconi anemia pathway in the response of human cells to trabectedin. *Molecular cancer therapeutics*. 7:1309-1318.

Castella, M., R. Pujol, E. Callen, M.J. Ramirez, J.A. Casado, M. Talavera, T. Ferro, A. Munoz, J. Sevilla, L. Madero, E. Cela, C. Belendez, C.D. de Heredia, T. Olive, J.S. de Toledo, I. Badell, J. Estella, A. Dasi, A. Rodriguez-Villa, P. Gomez, M. Tapia, A. Molines, A. Figuera, J.A. Bueren, and J. Surralles. 2011. Chromosome fragility in patients with Fanconi anaemia: diagnostic implications and clinical impact. *Journal of medical genetics*. 48:242-250.

Catalan, N., F. Selz, K. Imai, P. Revy, A. Fischer, and A. Durandy. 2003. The block in immunoglobulin class switch recombination caused by activation-induced cytidine deaminase deficiency occurs prior to the generation of DNA double strand breaks in switch mu region. *Journal of immunology (Baltimore, Md. : 1950)*. 171:2504-2509.

Ceccaldi, R., P. Sarangi, and A.D. D'Andrea. 2016. The Fanconi anaemia pathway: new players and new functions. *Nature Reviews Molecular Cell Biology*. 17:337.

Cerritelli, S.M., and R.J. Crouch. 2009. Ribonuclease H: the enzymes in eukaryotes. *The FEBS journal*. 276:1494-1505.

Chan, K.L., P.S. North, and I.D. Hickson. 2007. BLM is required for faithful chromosome segregation and its localization defines a class of ultrafine anaphase bridges. *Embo J*. 26:3397-3409.

Chan, K.L., T. Palmai-Pallag, S. Ying, and I.D. Hickson. 2009. Replication stress induces sister-chromatid bridging at fragile site loci in mitosis. *Nat Cell Biol*. 11:753-760.

Chang, E.Y.-C., and P.C. Stirling. 2017. Replication Fork Protection Factors Controlling R-Loop Bypass and Suppression. *Genes*. 8:33.

Chanoux, R.A., B. Yin, K.A. Urtishak, A. Asare, C.H. Bassing, and E.J. Brown. 2009. ATR and H2AX Cooperate in Maintaining Genome Stability under Replication Stress. *The Journal of Biological Chemistry*. 284:5994-6003.

Chapman, J.R., P. Barral, J.B. Vannier, V. Borel, M. Steger, A. Tomas-Loba, A.A. Sartori, I.R. Adams, F.D. Batista, and S.J. Boulton. 2013. RIF1 is essential for 53BP1-dependent nonhomologous end joining and suppression of DNA double-strand break resection. *Mol Cell*. 49:858-871.

Chapman, J.R., M.R. Taylor, and S.J. Boulton. 2012. Playing the end game: DNA double-strand break repair pathway choice. *Mol Cell*. 47:497-510.

Chaudhuri, J., and F.W. Alt. 2004. Class-switch recombination: interplay of transcription, DNA deamination and DNA repair. *Nature Reviews Immunology*. 4:541.

Chaudhury, I., A. Sareen, M. Raghunandan, and A. Sobeck. 2013. FANCD2 regulates BLM complex functions independently of FANCI to promote replication fork recovery. *Nucleic Acids Res*. 41:6444-6459.

Chaudhury, I., D.R. Stroik, and A. Sobeck. 2014. FANCD2-Controlled Chromatin Access of the Fanconi-Associated Nuclease FAN1 Is Crucial for the Recovery of Stalled Replication Forks. *Molecular and Cellular Biology*. 34:3939-3954.

Chen, H., M. Lisby, and L.S. Symington. 2013. RPA coordinates DNA end resection and prevents formation of DNA hairpins. *Mol Cell*. 50:589-600.

Chen, X., L. Bosques, P. Sung, and G.M. Kupfer. 2016. A novel role for non-ubiquitinated FANCD2 in response to hydroxyurea-induced DNA damage. *Oncogene*. 35:22-34.

Chen, Y.H., M.J. Jones, Y. Yin, S.B. Crist, L. Colnaghi, R.J. Sims, 3rd, E. Rothenberg, P.V. Jallepalli, and T.T. Huang. 2015. ATR-mediated phosphorylation of FANCI regulates dormant origin firing in response to replication stress. *Mol Cell*. 58:323-338.

Cheng, Q., and J. Chen. 2010. Mechanism of p53 stabilization by ATM after DNA damage. *Cell Cycle*. 9:472-478.

Cheung-Ong, K., G. Giaever, and C. Nislow. 2013. DNA-Damaging Agents in Cancer Chemotherapy: Serendipity and Chemical Biology. *Chemistry & Biology*. 20:648-659.

Ciccia, A., and S.J. Elledge. 2010. The DNA Damage Response: Making It Safe to Play with Knives. *Mol Cell*. 40:179-204.

Cimprich, K.A., and D. Cortez. 2008. ATR: an essential regulator of genome integrity. *Nature Reviews Molecular Cell Biology*. 9:616.

Cohen, S., N. Puget, Y.-L. Lin, T. Clouaire, M. Aguirrebengoa, V. Rocher, P. Pasero, Y. Canitrot, and G. Legube. 2018. Senataxin resolves RNA:DNA hybrids forming at DNA double-strand breaks to prevent translocations. *Nat Commun*. 9:533.

Coin, F., V. Oksenych, and J.M. Egly. 2007. Distinct roles for the XPB/p52 and XPD/p44 subcomplexes of TFIIH in damaged DNA opening during nucleotide excision repair. *Mol Cell*. 26:245-256.

Collis, S.J., A. Ciccia, A.J. Deans, Z. Hořejší, J.S. Martin, S.L. Maslen, J.M. Skehel, S.J. Elledge, S.C. West, and S.J. Boulton. 2008. FANCM and FAAP24 Function in ATR-Mediated Checkpoint Signaling Independently of the Fanconi Anemia Core Complex. *Mol Cell*. 32:313-324.

Conconi, A., V.A. Bespalov, and M.J. Smerdon. 2002. Transcription-coupled repair in RNA polymerase I-transcribed genes of yeast. *Proc Natl Acad Sci U S A*. 99:649-654.

Cong, L., F.A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P.D. Hsu, X. Wu, W. Jiang, L.A. Marraffini, and F. Zhang. 2013. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*. 339:819-823.

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

Daley, J.M., and P. Sung. 2014. 53BP1, BRCA1, and the choice between recombination and end joining at DNA double-strand breaks. *Mol Cell Biol*. 34:1380-1388.

Damia, G., S. Silvestri, L. Carrassa, L. Filiberti, G.T. Faircloth, G. Liberi, M. Foiani, and M. D'Incalci. 2001. Unique pattern of ET-743 activity in different cellular systems with defined deficiencies in DNA-repair pathways. *International journal of cancer*. 92:583-588.

Dammann, R., and G.P. Pfeifer. 1997. Lack of gene- and strand-specific DNA repair in RNA polymerase III-transcribed human tRNA genes. *Mol Cell Biol*. 17:219-229.

Damsma, G.E., A. Alt, F. Brueckner, T. Carell, and P. Cramer. 2007. Mechanism of transcriptional stalling at cisplatin-damaged DNA. *Nature Structural &Amp; Molecular Biology*. 14:1127.

Daneshbod-Skibba, G., J. Martin, and N.T. Shahidi. 1980. Myeloid and erythroid colony growth in non-anaemic patients with Fanconi's anaemia. *British journal of haematology*. 44:33-38.

de Klein, A., M. Muijtjens, R. van Os, Y. Verhoeven, B. Smit, A.M. Carr, A.R. Lehmann, and J.H. Hoeijmakers. 2000. Targeted disruption of the cell-cycle checkpoint gene ATR leads to early embryonic lethality in mice. *Current biology : CB*. 10:479-482.

Deeg, K.I., I. Chung, C. Bauer, and K. Rippe. 2016. Cancer Cells with Alternative Lengthening of Telomeres Do Not Display a General Hypersensitivity to ATR Inhibition. *Frontiers in Oncology*. 6:186.

Deng, S.K., B. Gibb, M.J. de Almeida, E.C. Greene, and L.S. Symington. 2014. RPA antagonizes microhomology-mediated repair of DNA double-strand breaks. *Nat Struct Mol Biol*. 21:405-412.

Dieci, G., G. Fiorino, M. Castelnuovo, M. Teichmann, and A. Pagano. 2007. The expanding RNA polymerase III transcriptome. *Trends in Genetics*. 23:614-622.

Difilippantonio, M.J., J. Zhu, H.T. Chen, E. Meffre, M.C. Nussenzweig, E.E. Max, T. Ried, and A. Nussenzweig. 2000. DNA repair protein Ku80 suppresses chromosomal aberrations and malignant transformation. *Nature*. 404:510-514.

Drolet, M., S. Broccoli, F. Rallu, C. Hraiky, C. Fortin, E. Masse, and I. Baaklini. 2003. The problem of hypernegative supercoiling and R-loop formation in transcription. *Frontiers in bioscience : a journal and virtual library*. 8:d210-221.

Drolet, M., P. Phoenix, R. Menzel, E. Massé, L.F. Liu, and R.J. Crouch. 1995. Overexpression of RNase H partially complements the growth defect of an Escherichia coli delta topA mutant: R-loop formation is a major problem in the absence of DNA topoisomerase I. *Proceedings of the National Academy of Sciences*. 92:3526-3530.

Dronkert, M.L.G., and R. Kanaar. 2001. Repair of DNA interstrand cross-links. *Mutation Research/DNA Repair*. 486:217-247.

Dror, Y., and M.H. Freedman. 2001. Shwachman-Diamond syndrome marrow cells show abnormally increased apoptosis mediated through the Fas pathway. *Blood*. 97:3011-3016.

Duquette, M.L., P. Handa, J.A. Vincent, A.F. Taylor, and N. Maizels. 2004. Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Gene Dev*. 18:1618-1629.

Durkin, S.G., and T.W. Glover. 2007. Chromosome Fragile Sites. *Annual review of genetics*. 41:169-192.

El Hage, A., S.L. French, A.L. Beyer, and D. Tollervey. 2010. Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. *Genes Dev*. 24:1546-1558.

Enoiu, M., J. Jiricny, and O.D. Scharer. 2012. Repair of cisplatin-induced DNA interstrand crosslinks by a replication-independent pathway involving transcription-coupled repair and translesion synthesis. *Nucleic Acids Res*. 40:8953-8964.

Epanchintsev, A., F. Costanzo, M.A. Rauschendorf, M. Caputo, T. Ye, L.M. Donnio, L. Proietti-de-Santis, F. Coin, V. Laugel, and J.M. Egly. 2017. Cockayne's Syndrome A and B Proteins Regulate Transcription Arrest after Genotoxic Stress by Promoting ATF3 Degradation. *Mol Cell*. 68:1054-1066.e1056.

Fagbemi, A.F., B. Orelli, and O.D. Schärer. 2011. Regulation of endonuclease activity in human nucleotide excision repair. *DNA Repair*. 10:722-729.

Falck, J., J. Coates, and S.P. Jackson. 2005. Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage. *Nature*. 434:605-611.

Falkenberg, M. 2018. Mitochondrial DNA replication in mammalian cells: overview of the pathway. *Essays in biochemistry*. 62:287-296.

Felipe-Abrio, I., J. Lafuente-Barquero, M.L. Garcia-Rubio, and A. Aguilera. 2015. RNA polymerase II contributes to preventing transcription-mediated replication fork stalls. *Embo J*. 34:236-250.

Fichtinger-Schepman, A.M., J.L. van der Veer, J.H. den Hartog, P.H. Lohman, and J. Reedijk. 1985. Adducts of the antitumor drug cis-diamminedichloroplatinum(II) with DNA: formation, identification, and quantitation. *Biochemistry*. 24:707-713.

Folias, A., M. Matkovic, D. Bruun, S. Reid, J. Hejna, M. Grompe, A. D'Andrea, and R. Moses. 2002. BRCA1 interacts directly with the Fanconi anemia protein FANCA. *Human Molecular Genetics*. 11:2591-2597.

French, S. 1992. Consequences of replication fork movement through transcription units in vivo. *Science*. 258:1362-1365.

Furuta, T., T. Ueda, G. Aune, A. Sarasin, K.H. Kraemer, and Y. Pommier. 2002. Transcription-coupled Nucleotide Excision Repair as a Determinant of Cisplatin Sensitivity of Human Cells. *Cancer research*. 62:4899-4902.

Gan, W., Z. Guan, J. Liu, T. Gui, K. Shen, J.L. Manley, and X. Li. 2011. R-loop-mediated genomic instability is caused by impairment of replication fork progression. *Gene Dev*. 25:2041-2056.

Garaycoechea, J.I., G.P. Crossan, F. Langevin, M. Daly, M.J. Arends, and K.J. Patel. 2012. Genotoxic consequences of endogenous aldehydes on mouse haematopoietic stem cell function. *Nature*. 489:571-575.

Garaycoechea, J.I., G.P. Crossan, F. Langevin, L. Mulderrig, S. Louzada, F. Yang, G. Guilbaud, N. Park, S. Roerink, S. Nik-Zainal, M.R. Stratton, and K.J. Patel. 2018. Alcohol and endogenous aldehydes damage chromosomes and mutate stem cells. *Nature*. 553:171-177.

Garcia-Higuera, I., T. Taniguchi, S. Ganesan, M.S. Meyn, C. Timmers, J. Hejna, M. Grompe, and A.D. D'Andrea. 2001. Interaction of the Fanconi Anemia Proteins and BRCA1 in a Common Pathway. *Mol Cell*. 7:249-262.

García-Muse, T., and A. Aguilera. 2016. Transcription–replication conflicts: how they occur and how they are resolved. *Nature Reviews Molecular Cell Biology*. 17:553.

Garcia-Pichardo, D., J.C. Canas, M.L. Garcia-Rubio, B. Gomez-Gonzalez, A.G. Rondon, and A. Aguilera. 2017. Histone Mutants Separate R Loop Formation from Genome Instability Induction. *Mol Cell*. 66:597-609.e595.

García-Rubio, M.L., C. Pérez-Calero, S.I. Barroso, E. Tumini, E. Herrera-Moyano, I.V. Rosado, and A. Aguilera. 2015. The Fanconi Anemia Pathway Protects Genome Integrity from R-loops. *PLOS Genetics*. 11:e1005674.

Getts, R.C., and T.D. Stamato. 1994. Absence of a Ku-like DNA end binding activity in the xrs double-strand DNA repair-deficient mutant. *J Biol Chem*. 269:15981-15984.

Gibbs-Seymour, I., Y. Oka, E. Rajendra, B.T. Weinert, L.A. Passmore, K.J. Patel, J.V. Olsen, C. Choudhary, S. Bekker-Jensen, and N. Mailand. 2015. Ubiquitin-SUMO circuitry controls activated fanconi anemia ID complex dosage in response to DNA damage. *Mol Cell*. 57:150-164.

Ginno, Paul A., Paul L. Lott, Holly C. Christensen, I. Korf, and F. Chédin. 2012. R-Loop Formation Is a Distinctive Characteristic of Unmethylated Human CpG Island Promoters. *Mol Cell*. 45:814-825.

Gisselsson, D., L. Pettersson, M. Höglund, M. Heidenblad, L. Gorunova, J. Wiegant, F. Mertens, P. Dal Cin, F. Mitelman, and N. Mandahl. 2000. Chromosomal breakage-fusion-bridge events cause genetic intratumor heterogeneity. *Proceedings of the National Academy of Sciences*. 97:5357-5362.

Gorthi, A., J.C. Romero, E. Loranc, L. Cao, L.A. Lawrence, E. Goodale, A.B. Iniguez, X. Bernard, V.P. Masamsetti, S. Roston, E.R. Lawlor, J.A. Toretsky, K. Stegmaier, S.L. Lessnick, Y. Chen, and A.J.R. Bishop. 2018. EWS–FLI1 increases transcription to cause R-loops and block BRCA1 repair in Ewing sarcoma. *Nature*. 555:387.

Gottlieb, T.M., and S.P. Jackson. 1993. The DNA-dependent protein kinase: requirement for DNA ends and association with Ku antigen. *Cell*. 72:131-142.

Graham, T.G., J.C. Walter, and J.J. Loparo. 2016. Two-Stage Synapsis of DNA Ends during Non-homologous End Joining. *Mol Cell*. 61:850-858.

Grawunder, U., M. Wilm, X. Wu, P. Kulesza, T.E. Wilson, M. Mann, and M.R. Lieber. 1997. Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. *Nature*. 388:492-495.

Groh, M., M.M.P. Lufino, R. Wade-Martins, and N. Gromak. 2014. R-loops Associated with Triplet Repeat Expansions Promote Gene Silencing in Friedreich Ataxia and Fragile X Syndrome. *PLOS Genetics*. 10:e1004318.

Guainazzi, A., and O.D. Schärer. 2010. Using synthetic DNA interstrand crosslinks to elucidate repair pathways and identify new therapeutic targets for cancer chemotherapy. *Cellular and molecular life sciences : CMLS*. 67:3683-3697.

Gupta, R., K. Somyajit, T. Narita, E. Maskey, A. Stanlie, M. Kremer, D. Typas, M. Lammers, N. Mailand, A. Nussenzweig, J. Lukas, and C. Choudhary. 2018. DNA Repair Network Analysis Reveals Shieldin as a Key Regulator of NHEJ and PARP Inhibitor Sensitivity. *Cell*. 173:972-988.e923.

Hanawalt, P.C., and G. Spivak. 2008. Transcription-coupled DNA repair: two decades of progress and surprises. *Nat Rev Mol Cell Biol*. 9:958-970.

Hatchi, E., K. Skourti-Stathaki, S. Ventz, L. Pinello, A. Yen, K. Kamieniarz-Gdula, S. Dimitrov, S. Pathania, Kristine M. McKinney, Matthew L. Eaton, M. Kellis, Sarah J. Hill, G. Parmigiani, Nicholas J. Proudfoot, and David M. Livingston. 2015. BRCA1 Recruitment to Transcriptional Pause Sites Is Required for R-Loop-Driven DNA Damage Repair. *Mol Cell*. 57:636-647.

Hejna, J., M. Holtorf, J. Hines, L. Mathewson, A. Hemphill, M. Al-Dhalimy, S.B. Olson, and R.E. Moses. 2008. Tip60 Is Required for DNA Interstrand Cross-link Repair in the Fanconi Anemia Pathway. *The Journal of Biological Chemistry*. 283:9844-9851.

Helmrich, A., M. Ballarino, E. Nudler, and L. Tora. 2013. Transcription-replication encounters, consequences and genomic instability. *Nat Struct Mol Biol*. 20:412-418.

Helmrich, A., M. Ballarino, and L. Tora. 2011. Collisions between Replication and Transcription Complexes Cause Common Fragile Site Instability at the Longest Human Genes. *Mol Cell*. 44:966-977.

Hengge, U.R., and S. Emmert. 2008. Clinical features of xeroderma pigmentosum. *Advances in experimental medicine and biology*. 637:10-18.

Heyer, W.D., K.T. Ehmsen, and J. Liu. 2010. Regulation of homologous recombination in eukaryotes. *Annual review of genetics*. 44:113-139.

Hira, A., K. Yoshida, K. Sato, Y. Okuno, Y. Shiraishi, K. Chiba, H. Tanaka, S. Miyano, A. Shimamoto, H. Tahara, E. Ito, S. Kojima, H. Kurumizaka, S. Ogawa, M. Takata, H. Yabe, and M. Yabe. 2015. Mutations in the Gene Encoding the E2 Conjugating Enzyme UBE2T Cause Fanconi Anemia. *The American Journal of Human Genetics*. 96:1001-1007.

Hodson, C., J.J. Rourke, S. van Twest, V.J. Murphy, E. Dunn, and A.J. Deans. 2018. FANCM-family branchpoint translocases remove co-transcriptional R-loops. *bioRxiv*.

Hoeijmakers, J.H.J. 2009. DNA Damage, Aging, and Cancer. *New England Journal of Medicine*. 361:1475-1485.

Hohl, M., F. Thorel, S.G. Clarkson, and O.D. Scharer. 2003. Structural determinants for substrate binding and catalysis by the structure-specific endonuclease XPG. *J Biol Chem*. 278:19500-19508.

Holton, N.W., J.F. Andrews, and N.R. Gassman. 2017. Application of Laser Micro-irradiation for Examination of Single and Double Strand Break Repair in Mammalian Cells. *Journal of Visualized Experiments : JoVE*:56265.

Howlett, N.G., T. Taniguchi, S.G. Durkin, A.D. D'Andrea, and T.W. Glover. 2005. The Fanconi anemia pathway is required for the DNA replication stress response

and for the regulation of common fragile site stability. *Hum Mol Genet*. 14:693-701.

Hsu, P.D., D.A. Scott, J.A. Weinstein, F.A. Ran, S. Konermann, V. Agarwala, Y. Li, E.J. Fine, X. Wu, O. Shalem, T.J. Cradick, L.A. Marraffini, G. Bao, and F. Zhang. 2013. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotech*. 31:827-832.

Huertas, P., and A. Aguilera. 2003. Cotranscriptionally Formed DNA:RNA Hybrids Mediate Transcription Elongation Impairment and Transcription-Associated Recombination. *Mol Cell*. 12:711-721.

Hurley, L.H., and M. Zewail-Foote. 2001. The antitumor agent ecteinascidin 743: characterization of its covalent DNA adducts and chemical stability. *Advances in experimental medicine and biology*. 500:289-299.

Iarmarcovai, G., S. Bonassi, A. Botta, R.A. Baan, and T. Orsière. 2008. Genetic polymorphisms and micronucleus formation: A review of the literature. *Mutation Research/Reviews in Mutation Research*. 658:215-233.

Illingworth, R.S., and A.P. Bird. 2009. CpG islands – 'A rough guide'. *FEBS Letters*. 583:1713-1720.

Ishiai, M., H. Kitao, A. Smogorzewska, J. Tomida, A. Kinomura, E. Uchida, A. Saberi, E. Kinoshita, E. Kinoshita-Kikuta, T. Koike, S. Tashiro, S.J. Elledge, and M. Takata. 2008. FANCI phosphorylation functions as a molecular switch to turn on the Fanconi anemia pathway. *Nat Struct Mol Biol*. 15:1138-1146.

Itoh, T., and J. Tomizawa. 1980. Formation of an RNA primer for initiation of replication of ColE1 DNA by ribonuclease H. *Proc Natl Acad Sci U S A*. 77:2450-2454.

Iwabuchi, K., P.L. Bartel, B. Li, R. Marraccino, and S. Fields. 1994. Two cellular proteins that bind to wild-type but not mutant p53. *P Natl Acad Sci USA*. 91:6098-6102.

Jackson, S.P., and J. Bartek. 2009. The DNA-damage response in human biology and disease. *Nature*. 461:1071-1078.

Jacobs, A.L., and P. Schär. 2012. DNA glycosylases: in DNA repair and beyond. *Chromosoma*. 121:1-20.

Jiricny, J. 2006. The multifaceted mismatch-repair system. *Nat Rev Mol Cell Biol*. 7:335-346.

Joerger, A.C., and A.R. Fersht. 2016. The p53 Pathway: Origins, Inactivation in Cancer, and Emerging Therapeutic Approaches. *Annual Review of Biochemistry*. 85:375-404.

Jung, Y., and S.J. Lippard. 2006. RNA Polymerase II Blockage by Cisplatin-damaged DNA: STABILITY AND POLYUBIQUITYLATION OF STALLED POLYMERASE. *Journal of Biological Chemistry*. 281:1361-1370.

Kashiyama, K., Y. Nakazawa, D.T. Pilz, C. Guo, M. Shimada, K. Sasaki, H. Fawcett, J.F. Wing, S.O. Lewin, L. Carr, T.S. Li, K. Yoshiura, A. Utani, A. Hirano, S. Yamashita, D. Greenblatt, T. Nardo, M. Stefanini, D. McGibbon, R. Sarkany, H. Fassihi, Y. Takahashi, Y. Nagayama, N. Mitsutake, A.R. Lehmann, and T. Ogi. 2013. Malfunction of nuclease ERCC1-XPF results in diverse clinical manifestations and causes Cockayne syndrome, xeroderma pigmentosum, and Fanconi anemia. *American journal of human genetics*. 92:807-819.

Kelly, P.F., S. Radtke, C. von Kalle, B. Balcik, K. Bohn, R. Mueller, T. Schuesler, M. Haren, L. Reeves, J.A. Cancelas, T. Leemhuis, R. Harris, A.D. Auerbach, F.O. Smith, S.M. Davies, and D.A. Williams. 2007. Stem cell collection and gene transfer in Fanconi anemia. *Molecular therapy : the journal of the American Society of Gene Therapy*. 15:211-219.

Kim, J.M., Y. Kee, A. Gurtan, and A.D. D'Andrea. 2008. Cell cycle–dependent chromatin loading of the Fanconi anemia core complex by FANCM/FAAP24. *Blood*. 111:5215-5222.

Kim, J.M., K. Parmar, M. Huang, D.M. Weinstock, C.A. Ruit, J.L. Kutok, and A.D. D'Andrea. 2009. Inactivation of murine Usp1 results in genomic instability and a Fanconi anemia phenotype. *Developmental cell*. 16:314-320.

Klein Douwel, D., Rick A.C.M. Boonen, David T. Long, Anna A. Szypowska, M. Räschle, Johannes C. Walter, and P. Knipscheer. 2014. XPF-ERCC1 Acts in Unhooking DNA Interstrand Crosslinks in Cooperation with FANCD2 and FANCP/SLX4. *Mol Cell*. 54:460-471.

Klungland, A., I. Rosewell, S. Hollenbach, E. Larsen, G. Daly, B. Epe, E. Seeberg, T. Lindahl, and D.E. Barnes. 1999. Accumulation of premutagenic DNA lesions in mice defective in removal of oxidative base damage. *Proc Natl Acad Sci U S A*. 96:13300-13305.

Knudson, M., S. Kulkarni, Z.K. Ballas, M. Bessler, and F. Goldman. 2005. Association of immune abnormalities with telomere shortening in autosomal-dominant dyskeratosis congenita. *Blood*. 105:682-688.

Kobayashi, T. 2014. Ribosomal RNA gene repeats, their stability and cellular senescence. *Proceedings of the Japan Academy. Series B, Physical and biological sciences*. 90:119-129.

Kohn, K.W. 1996. Beyond DNA cross-linking: history and prospects of DNA-targeted cancer treatment--fifteenth Bruce F. Cain Memorial Award Lecture. *Cancer research*. 56:5533-5546.

König, F., T. Schubert, and G. Längst. 2017. The monoclonal S9.6 antibody exhibits highly variable binding affinities towards different R-loop sequences. *PLOS ONE*. 12:e0178875.

Kumagai, A., J. Lee, H.Y. Yoo, and W.G. Dunphy. 2006. TopBP1 Activates the ATR-ATRIP Complex. *Cell*. 124:943-955.

Kumazawa, T., K. Nishimura, N. Katagiri, S. Hashimoto, Y. Hayashi, and K. Kimura. 2015. Gradual reduction in rRNA transcription triggers p53 acetylation and apoptosis via MYBBP1A. *Scientific Reports*. 5:10854.

Kunkel, T.A. 2004. DNA Replication Fidelity. *Journal of Biological Chemistry*. 279:16895-16898.

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 in genetics : TIG*. 34:111-120.

Labib, K., and B. Hodgson. 2007. Replication fork barriers: pausing for a break or stalling for time? *EMBO Reports*. 8:346-353.

Lachaud, C., A. Moreno, F. Marchesi, R. Toth, J.J. Blow, and J. Rouse. 2016. Ubiquitinated Fancd2 recruits Fan1 to stalled replication forks to prevent genome instability. *Science*. 351:846-849.

Langevin, F., G.P. Crossan, I.V. Rosado, M.J. Arends, and K.J. Patel. 2011. Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice. *Nature*. 475:53-58.

Le Cesne, A., J.Y. Blay, I. Judson, A. Van Oosterom, J. Verweij, J. Radford, P. Lorigan, S. Rodenhuis, I. Ray-Coquard, S. Bonvalot, F. Collin, J. Jimeno, E. Di Paola, M. Van Glabbeke, and O.S. Nielsen. 2005. Phase II study of ET-743 in advanced soft tissue sarcomas: a European Organisation for the Research and Treatment of Cancer (EORTC) soft tissue and bone sarcoma group trial. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 23:576-584.

Lee, J., A. Kumagai, and W.G. Dunphy. 2007. The Rad9-Hus1-Rad1 Checkpoint Clamp Regulates Interaction of TopBP1 with ATR. *Journal of Biological Chemistry*. 282:28036-28044.

Lee, Y.-C., Y. Cai, H. Mu, S. Broyde, S. Amin, X. Chen, J.-H. Min, and N.E. Geacintov. 2014. The relationships between XPC binding to conformationally diverse DNA adducts and their excision by the human NER system: Is there a correlation? *DNA Repair*. 19:55-63.

Li, Z., T. Otevrel, Y. Gao, H.L. Cheng, B. Seed, T.D. Stamato, G.E. Taccioli, and F.W. Alt. 1995. The XRCC4 gene encodes a novel protein involved in DNA double-strand break repair and V(D)J recombination. *Cell*. 83:1079-1089.

Lim, Y.W., L.A. Sanz, X. Xu, S.R. Hartono, and F. Chédin. 2015. Genome-wide DNA hypomethylation and RNA:DNA hybrid accumulation in Aicardi–Goutières syndrome. *eLife*. 4:e08007.

Lindahl, T. 1993. Instability and decay of the primary structure of DNA. *Nature*. 362:709.

Lindahl, T., and D.E. Barnes. 2000. Repair of endogenous DNA damage. *Cold Spring Harbor symposia on quantitative biology*. 65:127-133.

Lindsey-Boltz, L.A., and A. Sancar. 2007. RNA polymerase: The most specific damage recognition protein in cellular responses to DNA damage? *P Natl Acad Sci USA*. 104:13213-13214.

Liu, B., and B.M. Alberts. 1995. Head-on collision between a DNA replication apparatus and RNA polymerase transcription complex. *Science*. 267:1131-1137.

Liu, J., T. Doty, B. Gibson, and W.-D. Heyer. 2010. Human BRCA2 protein promotes RAD51 filament formation on RPA-covered single-stranded DNA. *Nature Structural &Amp; Molecular Biology*. 17:1260.

Liu, J.M., and S.R. Ellis. 2006. Ribosomes and marrow failure: coincidental association or molecular paradigm? *Blood*. 107:4583-4588.

Liu, Y., C.F. Nielsen, Q. Yao, and I.D. Hickson. 2014. The origins and processing of ultra fine anaphase DNA bridges. *Current Opinion in Genetics & Development*. 26:1-5.

Loeb, L.A. 1991. Mutator phenotype may be required for multistage carcinogenesis. *Cancer research*. 51:3075-3079.

Lopez-estrano, C., J.B. Schvartzman, D.B. Krimer, and P. Hernandez. 1998. Co-localization of polar replication fork barriers and rRNA transcription terminators in mouse rDNA. *Journal of molecular biology*. 277:249-256.

Lord, C.J., and A. Ashworth. 2012. The DNA damage response and cancer therapy. *Nature*. 481:287.

Lord, C.J., and A. Ashworth. 2017. PARP inhibitors: Synthetic lethality in the clinic. *Science*. 355:1152-1158.

Lossaint, G., M. Larroque, C. Ribeyre, N. Bec, C. Larroque, C. Décaillet, K. Gari, and A. Constantinou. 2013. FANCD2 Binds MCM Proteins and Controls Replisome

Function upon Activation of S Phase Checkpoint Signaling. *Mol Cell*. 51:678-690.

Luke-Glaser, S., B. Luke, S. Grossi, and A. Constantinou. 2010. FANCM regulates DNA chain elongation and is stabilized by S-phase checkpoint signalling. *The EMBO Journal*. 29:795-805.

Madireddy, A., Settapong T. Kosiyatrakul, Rebecca A. Boisvert, E. Herrera-Moyano, María L. García-Rubio, J. Gerhardt, Elizabeth A. Vuono, N. Owen, Z. Yan, S. Olson, A. Aguilera, Niall G. Howlett, and Carl L. Schildkraut. 2016. FANCD2 Facilitates Replication through Common Fragile Sites. *Mol Cell*. 64:388-404.

Marheineke, K., and O. Hyrien. 2004. Control of Replication Origin Density and Firing Time in Xenopus Egg Extracts: ROLE OF A CAFFEINE-SENSITIVE, ATR-DEPENDENT CHECKPOINT. *Journal of Biological Chemistry*. 279:28071-28081.

Massé, E., P. Phoenix, and M. Drolet. 1997. DNA Topoisomerases Regulate R-loop Formation during Transcription of the rrnB Operon in Escherichia coli. *Journal of Biological Chemistry*. 272:12816-12823.

Matsuda, T., M. Kawanishi, T. Yagi, S. Matsui, and H. Takebe. 1998. Specific tandem GG to TT base substitutions induced by acetaldehyde are due to intra-strand crosslinks between adjacent guanine bases. *Nucleic Acids Res*. 26:1769-1774.

Matsumoto, S., E.S. Fischer, T. Yasuda, N. Dohmae, S. Iwai, T. Mori, R. Nishi, K. Yoshino, W. Sakai, F. Hanaoka, N.H. Thoma, and K. Sugasawa. 2015. Functional regulation of the DNA damage-recognition factor DDB2 by ubiquitination and interaction with xeroderma pigmentosum group C protein. *Nucleic Acids Res*. 43:1700-1713.

McHugh, P.J., T.A. Ward, and M. Chovanec. 2012. A prototypical Fanconi anemia pathway in lower eukaryotes? *Cell Cycle*. 11:3739-3744.

McKinnon, P.J. 2004. ATM and ataxia telangiectasia. *EMBO Rep*. 5:772-776.

Meetei, A.R., J.P. de Winter, A.L. Medhurst, M. Wallisch, Q. Waisfisz, H.J. van de Vrugt, A.B. Oostra, Z. Yan, C. Ling, C.E. Bishop, M.E. Hoatlin, H. Joenje, and W. Wang. 2003. A novel ubiquitin ligase is deficient in Fanconi anemia. *Nature genetics*. 35:165-170.

Meetei, A.R., M. Levitus, Y. Xue, A.L. Medhurst, M. Zwaan, C. Ling, M.A. Rooimans, P. Bier, M. Hoatlin, G. Pals, J.P. de Winter, W. Wang, and H. Joenje. 2004. X-linked inheritance of Fanconi anemia complementation group B. *Nature genetics*. 36:1219-1224.

Meier, A., M. Livingstone-Zatchej, and F. Thoma. 2002. Repair of active and silenced rDNA in yeast: the contributions of photolyase and transcription-couples nucleotide excision repair. *J Biol Chem*. 277:11845-11852.

Merrikh, H., C. Machon, W.H. Grainger, A.D. Grossman, and P. Soultanas. 2011. Co-directional replication-transcription conflicts lead to replication restart. *Nature*. 470:554-557.

Meryet-Figuiere, M., B. Alaei-Mahabadi, M.M. Ali, S. Mitra, S. Subhash, G.K. Pandey, E. Larsson, and C. Kanduri. 2014. Temporal separation of replication and transcription during S-phase progression. *Cell Cycle*. 13:3241-3248.

Milyavsky, M., O.I. Gan, M. Trottier, M. Komosa, O. Tabach, F. Notta, E. Lechman, K.G. Hermans, K. Eppert, Z. Konovalova, O. Ornatsky, E. Domany, M.S. Meyn, and J.E. Dick. 2010. A Distinctive DNA Damage Response in Human Hematopoietic Stem Cells Reveals an Apoptosis-Independent Role for p53 in Self-Renewal. *Cell Stem Cell*. 7:186-197.

Mimori, T., M. Akizuki, H. Yamagata, S. Inada, S. Yoshida, and M. Homma. 1981. Characterization of a high molecular weight acidic nuclear protein recognized by autoantibodies in sera from patients with polymyositis-scleroderma overlap. *The Journal of clinical investigation*. 68:611-620.

Mimori, T., and J.A. Hardin. 1986. Mechanism of interaction between Ku protein and DNA. *J Biol Chem*. 261:10375-10379.

Mirkin, E.V., and S.M. Mirkin. 2005. Mechanisms of Transcription-Replication Collisions in Bacteria. *Molecular and Cellular Biology*. 25:888-895.

Mirkin, S.M. 2008. Discovery of alternative DNA structures: a heroic decade (1979-1989). *Frontiers in bioscience : a journal and virtual library*. 13:1064-1071.

Mohrin, M., E. Bourke, D. Alexander, M.R. Warr, K. Barry-Holson, M.M. Le Beau, C.G. Morrison, and E. Passegué. 2010. Hematopoietic Stem Cell Quiescence Promotes Error-Prone DNA Repair and Mutagenesis. *Cell Stem Cell*. 7:174-185.

Mullan, P.B., J.E. Quinn, and D.P. Harkin. 2006. The role of BRCA1 in transcriptional regulation and cell cycle control. *Oncogene*. 25:5854-5863.

Murina, O., C. von Aesch, U. Karakus, Lorenza P. Ferretti, Hella A. Bolck, K. Hänggi, and Alessandro A. Sartori. 2014. FANCD2 and CtIP Cooperate to Repair DNA Interstrand Crosslinks. *Cell Reports*. 7:1030-1038.

Nadel, J., R. Athanasiadou, C. Lemetre, N.A. Wijetunga, P. Ó Broin, H. Sato, Z. Zhang, J. Jeddeloh, C. Montagna, A. Golden, C. Seoighe, and J.M. Greally. 2015. RNA:DNA hybrids in the human genome have distinctive nucleotide characteristics, chromatin composition, and transcriptional relationships. *Epigenetics & Chromatin*. 8:46.

Nakanishi, K., F. Cavallo, L. Perrouault, C. Giovannangeli, M.E. Moynahan, M. Barchi, E. Brunet, and M. Jasin. 2011. Homology-directed Fanconi anemia pathway cross-link repair is dependent on DNA replication. *Nat Struct Mol Biol*. 18:500-503.

Nalepa, G., and D.W. Clapp. 2018. Fanconi anaemia and cancer: an intricate relationship. *Nat Rev Cancer*.

Nepal, M., R. Che, C. Ma, J. Zhang, and P. Fei. 2017. FANCD2 and DNA Damage. *International Journal of Molecular Sciences*. 18:1804.

Nguyen, D.T., H.P.J. Voon, B. Xella, C. Scott, D. Clynes, C. Babbs, H. Ayyub, J. Kerry, J.A. Sharpe, J.A. Sloane-Stanley, S. Butler, C.A. Fisher, N.E. Gray, T. Jenuwein, D.R. Higgs, and R.J. Gibbons. 2017. The chromatin remodelling factor ATRX suppresses R-loops in transcribed telomeric repeats. *EMBO reports*. 18:914-928.

Niedernhofer, L.J., J.S. Daniels, C.A. Rouzer, R.E. Greene, and L.J. Marnett. 2003. Malondialdehyde, a product of lipid peroxidation, is mutagenic in human cells. *J Biol Chem*. 278:31426-31433.

Nieminuszczy, J., R.A. Schwab, and W. Niedzwiedz. 2016. The DNA fibre technique – tracking helicases at work. *Methods*. 108:92-98.

Nijman, S.M.B., T.T. Huang, A.M.G. Dirac, T.R. Brummelkamp, R.M. Kerkhoven, A.D. D'Andrea, and R. Bernards. 2005. The Deubiquitinating Enzyme USP1 Regulates the Fanconi Anemia Pathway. *Mol Cell*. 17:331-339.

Noll, D.M., T.M. Mason, and P.S. Miller. 2006. Formation and Repair of Interstrand Cross-Links in DNA. *Chemical reviews*. 106:277-301.

O'Brien, P.J., A.G. Siraki, and N. Shangari. 2005. Aldehyde sources, metabolism, molecular toxicity mechanisms, and possible effects on human health. *Critical reviews in toxicology*. 35:609-662.

O'Driscoll, M., V.L. Ruiz-Perez, C.G. Woods, P.A. Jeggo, and J.A. Goodship. 2003. A splicing mutation affecting expression of ataxia–telangiectasia and Rad3–related protein (ATR) results in Seckel syndrome. *Nature genetics*. 33:497.

Oberbeck, N., F. Langevin, G. King, N. de Wind, Gerry P. Crossan, and Ketan J. Patel. 2014. Maternal Aldehyde Elimination during Pregnancy Preserves the Fetal Genome. *Mol Cell*. 55:807-817.

Ochi, T., A.N. Blackford, J. Coates, S. Jhujh, S. Mehmood, N. Tamura, J. Travers, Q. Wu, V.M. Draviam, C.V. Robinson, T.L. Blundell, and S.P. Jackson. 2015. DNA repair. PAXX, a paralog of XRCC4 and XLF, interacts with Ku to promote DNA double-strand break repair. *Science*. 347:185-188.

Okamoto, Y., W.M. Iwasaki, K. Kugou, K.K. Takahashi, A. Oda, K. Sato, W. Kobayashi, H. Kawai, R. Sakasai, A. Takaori-Kondo, T. Yamamoto, M.T. Kanemaki, M. Taoka, T. Isobe, H. Kurumizaka, H. Innan, K. Ohta, M. Ishiai, and M. Takata. 2018. Replication stress induces accumulation of FANCD2 at central region of large fragile genes. *Nucleic Acids Res*. 46:2932-2944.

Olavarrieta, L., P. Hernandez, D.B. Krimer, and J.B. Schvartzman. 2002. DNA knotting caused by head-on collision of transcription and replication. *Journal of molecular biology*. 322:1-6.

Perdahl, E., B. Naprstek, W. Wallace, and J. Lipton. 1994. Erythroid failure in Diamond-Blackfan anemia is characterized by apoptosis. *Blood*. 83:645-650.

Petermann, E., A. Maya-Mendoza, G. Zachos, D.A. Gillespie, D.A. Jackson, and K.W. Caldecott. 2006. Chk1 requirement for high global rates of replication fork progression during normal vertebrate S phase. *Mol Cell Biol*. 26:3319-3326.

Petermann, E., M.L. Orta, N. Issaeva, N. Schultz, and T. Helleday. 2010a. Hydroxyurea-Stalled Replication Forks Become Progressively Inactivated and Require Two Different RAD51-Mediated Pathways for Restart and Repair. *Mol Cell*. 37:492-502.

Petermann, E., M. Woodcock, and T. Helleday. 2010b. Chk1 promotes replication fork progression by controlling replication initiation. *Proc Natl Acad Sci U S A*. 107:16090-16095.

Peterson, S.R., A. Kurimasa, M. Oshimura, W.S. Dynan, E.M. Bradbury, and D.J. Chen. 1995. Loss of the catalytic subunit of the DNA-dependent protein kinase in DNA double-strand-break-repair mutant mammalian cells. *Proc Natl Acad Sci U S A*. 92:3171-3174.

Petruseva, I.O., A.N. Evdokimov, and O.I. Lavrik. 2014. Molecular Mechanism of Global Genome Nucleotide Excision Repair. *Acta Naturae*. 6:23-34.

Pichierri, P., and F. Rosselli. 2004. The DNA crosslink-induced S-phase checkpoint depends on ATR–CHK1 and ATR–NBS1–FANCD2 pathways. *The EMBO Journal*. 23:1178-1187.

Pierce, A.J., P. Hu, M. Han, N. Ellis, and M. Jasin. 2001. Ku DNA end-binding protein modulates homologous repair of double-strand breaks in mammalian cells. *Gene Dev*. 15:3237-3242.

Poindessous, V., F. Koeppel, E. Raymond, M. Comisso, S.J. Waters, and A.K. Larsen. 2003. Marked activity of irofulven toward human carcinoma cells: comparison with cisplatin and ecteinascidin. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 9:2817-2825.

Pomerantz, R.T., and M. O'Donnell. 2010. Direct restart of a replication fork stalled by a head-on RNA polymerase. *Science*. 327:590-592.

Qi, Z., S. Redding, Ja Y. Lee, B. Gibb, Y. Kwon, H. Niu, William A. Gaines, P. Sung, and Eric C. Greene. 2015. DNA Sequence Alignment by Microhomology Sampling during Homologous Recombination. *Cell*. 160:856-869.

Raghunandan, M., I. Chaudhury, S.L. Kelich, H. Hanenberg, and A. Sobeck. 2015. FANCD2, FANCJ and BRCA2 cooperate to promote replication fork recovery independently of the Fanconi Anemia core complex. *Cell Cycle*. 14:342-353.

Ray Chaudhuri, A., and A. Nussenzweig. 2017. The multifaceted roles of PARP1 in DNA repair and chromatin remodelling. *Nature Reviews Molecular Cell Biology*. 18:610.

Reaban, M.E., and J.A. Griffin. 1990. Induction of RNA-stabilized DNA conformers by transcription of an immunoglobulin switch region. *Nature*. 348:342-344.

Reaban, M.E., J. Lebowitz, and J.A. Griffin. 1994. Transcription induces the formation of a stable RNA.DNA hybrid in the immunoglobulin alpha switch region. *J Biol Chem*. 269:21850-21857.

Renaud, E., A. Barascu, and F. Rosselli. 2016. Impaired TIP60-mediated H4K16 acetylation accounts for the aberrant chromatin accumulation of 53BP1 and RAP80 in Fanconi anemia pathway-deficient cells. *Nucleic Acids Res*. 44:648-656.

Renaudin, X., J.H. Guervilly, S. Aoufouchi, and F. Rosselli. 2014. Proteomic analysis reveals a FANCA-modulated neddylation pathway involved in CXCR5 membrane targeting and cell mobility. *J Cell Sci*. 127:3546-3554.

Richard, G.-F., A. Kerrest, and B. Dujon. 2008. Comparative Genomics and Molecular Dynamics of DNA Repeats in Eukaryotes. *Microbiology and Molecular Biology Reviews : MMBR*. 72:686-727.

Ridpath, J.R., A. Nakamura, K. Tano, A.M. Luke, E. Sonoda, H. Arakawa, J.M. Buerstedde, D.A. Gillespie, J.E. Sale, M. Yamazoe, D.K. Bishop, M. Takata, S. Takeda, M. Watanabe, J.A. Swenberg, and J. Nakamura. 2007. Cells deficient in the FANC/BRCA pathway are hypersensitive to plasma levels of formaldehyde. *Cancer research*. 67:11117-11122.

Riss, T.L., R.A. Moravec, A.L. Niles, S. Duellman, H.A. Benink, T.J. Worzella, and L. Minor. 2004. Cell Viability Assays. *In* Assay Guidance Manual. G.S. Sittampalam, N.P. Coussens, K. Brimacombe, A. Grossman, M. Arkin, D. Auld, C. Austin, J. Baell, B. Bejcek, T.D.Y. Chung, J.L. Dahlin, V. Devanaryan, T.L. Foley, M. Glicksman, M.D. Hall, J.V. Hass, J. Inglese, P.W. Iversen, S.D. Kahl, S.C. Kales, M. Lal-Nag, Z. Li, J. McGee, O. McManus, T. Riss, O.J. Trask, Jr., J.R. Weidner, M. Xia, and X. Xu, editors. Eli Lilly & Company and the National Center for Advancing Translational Sciences, Bethesda (MD).

Roberts, R.W., and D.M. Crothers. 1992. Stability and properties of double and triple helices: dramatic effects of RNA or DNA backbone composition. *Science*. 258:1463-1466.

Rogakou, E.P., C. Boon, C. Redon, and W.M. Bonner. 1999. Megabase Chromatin Domains Involved in DNA Double-Strand Breaks in Vivo. *The Journal of Cell Biology*. 146:905-916.

Rogakou, E.P., D.R. Pilch, A.H. Orr, V.S. Ivanova, and W.M. Bonner. 1998. DNA Double-stranded Breaks Induce Histone H2AX Phosphorylation on Serine 139. *Journal of Biological Chemistry*. 273:5858-5868.

Roos, W.P., and B. Kaina. 2006. DNA damage-induced cell death by apoptosis. *Trends in molecular medicine*. 12:440-450.

Rosado, I.V., F. Langevin, G.P. Crossan, M. Takata, and K.J. Patel. 2011. Formaldehyde catabolism is essential in cells deficient for the Fanconi anemia DNA-repair pathway. *Nat Struct Mol Biol*. 18:1432-1434.

Rudolph, C.J., P. Dhillon, T. Moore, and R.G. Lloyd. 2007. Avoiding and resolving conflicts between DNA replication and transcription. *DNA Repair*. 6:981-993.

Sakai, A., R. Sakasai, Y. Kakeji, H. Kitao, and Y. Maehara. 2012. PARP and CSB modulate the processing of transcription-mediated DNA strand breaks. *Genes & Genetic Systems*. 87:265-272.

Sakai, K., T. Ohta, S. Minoshima, J. Kudoh, Y. Wang, P.J. de Jong, and N. Shimizu. 1995. Human ribosomal RNA gene cluster: identification of the proximal end containing a novel tandem repeat sequence. *Genomics*. 26:521-526.

Saldivar, J.C., S. Hamperl, M.J. Bocek, M. Chung, T.E. Bass, F. Cisneros-Soberanis, K. Samejima, L. Xie, J.R. Paulson, W.C. Earnshaw, D. Cortez, T. Meyer, and K.A. Cimprich. 2018. An intrinsic S/G2 checkpoint enforced by ATR. *Science*. 361:806-810.

Sarkar, S., A.A. Davies, H.D. Ulrich, and P.J. McHugh. 2006. DNA interstrand crosslink repair during G1 involves nucleotide excision repair and DNA polymerase zeta. *Embo J*. 25:1285-1294.

Sasaki, M.S., and A. Tonomura. 1973. A high susceptibility of Fanconi's anemia to chromosome breakage by DNA cross-linking agents. *Cancer research*. 33:1829-1836.

Savitsky, K., A. Bar-Shira, S. Gilad, G. Rotman, Y. Ziv, L. Vanagaite, D.A. Tagle, S. Smith, T. Uziel, S. Sfez, M. Ashkenazi, I. Pecker, M. Frydman, R. Harnik, S.R. Patanjali, A. Simmons, G.A. Clines, A. Sartiel, R.A. Gatti, L. Chessa, O. Sanal, M.F. Lavin, N.G. Jaspers, A.M. Taylor, C.F. Arlett, T. Miki, S.M. Weissman, M. Lovett, F.S. Collins, and Y. Shiloh. 1995. A single ataxia telangiectasia gene with a product similar to PI-3 kinase. *Science*. 268:1749-1753.

Saxonov, S., P. Berg, and D.L. Brutlag. 2006. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc Natl Acad Sci U S A*. 103:1412-1417.

Schärer, O.D. 2005. DNA Interstrand Crosslinks: Natural and Drug-Induced DNA Adducts that Induce Unique Cellular Responses. *ChemBioChem*. 6:27-32.

Schärer, O.D. 2013. Nucleotide Excision Repair in Eukaryotes. *Cold Spring Harbor Perspectives in Biology*. 5:a012609.

Schlacher, K., N. Christ, N. Siaud, A. Egashira, H. Wu, and M. Jasin. 2011. Double-strand break repair-independent role for BRCA2 in blocking stalled replication fork degradation by MRE11. *Cell*. 145:529-542.

Schlacher, K., H. Wu, and M. Jasin. 2012. A Distinct Replication Fork Protection Pathway Connects Fanconi Anemia Tumor Suppressors to RAD51-BRCA1/2. *Cancer cell*. 22:106-116.

Schultz, L.B., N.H. Chehab, A. Malikzay, and T.D. Halazonetis. 2000. P53 Binding Protein 1 (53bp1) Is an Early Participant in the Cellular Response to DNA Double-Strand Breaks. *The Journal of Cell Biology*. 151:1381.

Schwab, R., J. Nieminuszczy, F. Shah, J. Langton, D. Lopez Martinez, C.-C. Liang, Martin A. Cohn, Richard J. Gibbons, Andrew J. Deans, and W. Niedzwiedz. 2015. The Fanconi Anemia Pathway Maintains Genome Stability by Coordinating Replication and Transcription. *Mol Cell*. 60:351-361.

Schwab, R.A., A.N. Blackford, and W. Niedzwiedz. 2010. ATR activation and replication fork restart are defective in FANCM-deficient cells. *The EMBO Journal*. 29:806-818.

Sen, D., and W. Gilbert. 1988. Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature*. 334:364-366.

Sessa, C., F. De Braud, A. Perotti, J. Bauer, G. Curigliano, C. Noberasco, F. Zanaboni, L. Gianni, S. Marsoni, J. Jimeno, M. D'Incalci, E. Dall'o, and N. Colombo. 2005. Trabectedin for women with ovarian carcinoma after treatment with platinum and taxanes fails. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 23:1867-1874.

Shaham, J., Y. Bomstein, A. Meltzer, Z. Kaufman, E. Palma, and J. Ribak. 1996. DNA-protein crosslinks, a biomarker of exposure to formaldehyde—in vitro and in vivo studies. *Carcinogenesis*. 17:121-126.

Shaw, N.N., and D.P. Arya. 2008. Recognition of the unique structure of DNA:RNA hybrids. *Biochimie*. 90:1026-1039.

Shen, Y., Y.-H. Lee, J. Panneerselvam, J. Zhang, L.W.M. Loo, and P. Fei. 2015. Mutated Fanconi anemia pathway in non-Fanconi anemia cancers. *Oncotarget*. 6:20396-20403.

Shereda, R.D., Y. Machida, and Y.J. Machida. 2010. Human KIAA1018/FAN1 localizes to stalled replication forks via its ubiquitin-binding domain. *Cell Cycle*. 9:3977-3983.

Shiloh, Y., and Y. Ziv. 2013. The ATM protein kinase: regulating the cellular response to genotoxic stress, and more. *Nat Rev Mol Cell Biol*. 14:197-210.

Shivji, M.K.K., X. Renaudin, C.H. Williams, and A.R. Venkitaraman. 2018. BRCA2 Regulates Transcription Elongation by RNA Polymerase II to Prevent R-Loop Accumulation. *Cell Rep*. 22:1031-1039.

Singh, T.R., A.M. Ali, M. Paramasivam, A. Pradhan, K. Wahengbam, M.M. Seidman, and A.R. Meetei. 2013. ATR-Dependent Phosphorylation of FANCM at Serine 1045 is Essential for FANCM Functions. *Cancer research*.

Skourti-Stathaki, K., Nicholas J. Proudfoot, and N. Gromak. 2011. Human Senataxin Resolves RNA/DNA Hybrids Formed at Transcriptional Pause Sites to Promote Xrn2-Dependent Termination. *Mol Cell*. 42:794-805.

Smetsers, S., J. Muter, C. Bristow, L. Patel, K. Chandler, D. Bonney, R.F. Wynn, A.D. Whetton, A.M. Will, D. Rockx, H. Joenje, G. Strathdee, J. Shanks, E. Klopocki, J.J. Gille, J. Dorsman, and S. Meyer. 2012. Heterozygote FANCD2 mutations associated with childhood T Cell ALL and testicular seminoma. *Familial cancer*. 11:661-665.

Snyder, R.D. 1984. The role of deoxynucleoside triphosphate pools in the inhibition of DNA-excision repair and replication in human cells by hydroxyurea. *Mutation Research/DNA Repair Reports*. 131:163-172.

Soares, D.G., A.E. Escargueil, V. Poindessous, A. Sarasin, A. de Gramont, D. Bonatto, J.A. Henriques, and A.K. Larsen. 2007. Replication and homologous recombination repair regulate DNA double-strand break formation by the antitumor alkylator ecteinascidin 743. *Proc Natl Acad Sci U S A*. 104:13062-13067.

Sollier, J., and K.A. Cimprich. 2015. Breaking bad: R-loops and genome integrity. *Trends in Cell Biology*. 25:514-522.

Sollier, J., C.T. Stork, M.L. Garcia-Rubio, R.D. Paulsen, A. Aguilera, and K.A. Cimprich. 2014. Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. *Mol Cell*. 56:777-785.

Sondalle, S.B., S. Longerich, P. Sung, and S.J. Baserga. 2016. A nucleolar role for the Fanconi anemia pathway protein, FANCI, in rDNA transcription. *The FASEB Journal*. 30:1048.1041-1048.1041.

Soutoglou, E., and T. Misteli. 2008. Activation of the cellular DNA damage response in the absence of DNA lesions. *Science*. 320:1507-1510.

Spivak, G. 2015. Nucleotide excision repair in humans. *DNA Repair*. 36:13-18.

Stein, H., and P. Hausen. 1969. Enzyme from calf thymus degrading the RNA moiety of DNA-RNA Hybrids: effect on DNA-dependent RNA polymerase. *Science*. 166:393-395.

Stone, M.P., Y.J. Cho, H. Huang, H.Y. Kim, I.D. Kozekov, A. Kozekova, H. Wang, I.G. Minko, R.S. Lloyd, T.M. Harris, and C.J. Rizzo. 2008. Interstrand DNA cross-links induced by alpha,beta-unsaturated aldehydes derived from lipid peroxidation and environmental sources. *Accounts of chemical research*. 41:793-804.

Stork, C.T., M. Bocek, M.P. Crossley, J. Sollier, L.A. Sanz, F. Chédin, T. Swigut, and K.A. Cimprich. 2016. Co-transcriptional R-loops are the main cause of estrogen-induced DNA damage. *eLife*. 5:e17548.

Sträßer, K., S. Masuda, P. Mason, J. Pfannstiel, M. Oppizzi, S. Rodriguez-Navarro, A.G. Rondón, A. Aguilera, K. Struhl, R. Reed, and E. Hurt. 2002. TREX is a conserved complex coupling transcription with messenger RNA export. *Nature*. 417:304.

Suzuki, T., P. Gruz, M. Honma, N. Adachi, and T. Nohmi. 2016. The role of DNA polymerase zeta in translesion synthesis across bulky DNA adducts and cross-links in human cells. *Mutat Res*. 791-792:35-41.

Symington, L.S., and J. Gautier. 2011. Double-Strand Break End Resection and Repair Pathway Choice. *Annual review of genetics*. 45:247-271.

Takebayashi, Y., P. Pourquier, D.B. Zimonjic, K. Nakayama, S. Emmert, T. Ueda, Y. Urasaki, A. Kanzaki, S.I. Akiyama, N. Popescu, K.H. Kraemer, and Y. Pommier. 2001. Antiproliferative activity of ecteinascidin 743 is dependent upon transcription-coupled nucleotide-excision repair. *Nature medicine*. 7:961-966.

Tantin, D., A. Kansal, and M. Carey. 1997. Recruitment of the putative transcription-repair coupling factor CSB/ERCC6 to RNA polymerase II elongation complexes. *Molecular and Cellular Biology*. 17:6803-6814.

te Poele, R.H., A.L. Okorokov, L. Jardine, J. Cummings, and S.P. Joel. 2002. DNA Damage Is Able to Induce Senescence in Tumor Cells in Vitro and in Vivo. *Cancer research*. 62:1876-1883.

Tipping, A.J., T. Pearson, N.V. Morgan, R.A. Gibson, L.P. Kuyt, C. Havenga, E. Gluckman, H. Joenje, T. de Ravel, S. Jansen, and C.G. Mathew. 2001. Molecular and genealogical evidence for a founder effect in Fanconi anemia families of the Afrikaner population of South Africa. *Proceedings of the National Academy of Sciences*. 98:5734-5739.

Tornaletti, S., S. Park-Snyder, and P.C. Hanawalt. 2008. G4-forming Sequences in the Non-transcribed DNA Strand Pose Blocks to T7 RNA Polymerase and Mammalian RNA Polymerase II. *Journal of Biological Chemistry*. 283:12756-12762.

Trautinger, B.W., R.P. Jaktaji, E. Rusakova, and R.G. Lloyd. 2005. RNA polymerase modulators and DNA repair activities resolve conflicts between DNA replication and transcription. *Mol Cell*. 19:247-258.

Tuduri, S., L. Crabbe, C. Conti, H. Tourriere, H. Holtgreve-Grez, A. Jauch, V. Pantesco, J. De Vos, A. Thomas, C. Theillet, Y. Pommier, J. Tazi, A. Coquelle, and P. Pasero. 2009. Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. *Nat Cell Biol*. 11:1315-1324.

Tummala, H., M. Kirwan, A.J. Walne, U. Hossain, N. Jackson, C. Pondarre, V. Plagnol, T. Vulliamy, and I. Dokal. 2014. ERCC6L2 mutations link a distinct bone-marrow-failure syndrome to DNA repair and mitochondrial function. *American journal of human genetics*. 94:246-256.

Unno, J., A. Itaya, M. Taoka, K. Sato, J. Tomida, W. Sakai, K. Sugasawa, M. Ishiai, T. Ikura, T. Isobe, H. Kurumizaka, and M. Takata. 2014. FANCD2 Binds CtIP and Regulates DNA-End Resection during DNA Interstrand Crosslink Repair. *Cell Reports*. 7:1039-1047.

van den Boom, V., E. Citterio, D. Hoogstraten, A. Zotter, J.M. Egly, W.A. van Cappellen, J.H. Hoeijmakers, A.B. Houtsmuller, and W. Vermeulen. 2004. DNA damage stabilizes interaction of CSB with the transcription elongation machinery. *J Cell Biol*. 166:27-36.

van Twest, S., V.J. Murphy, C. Hodson, W. Tan, P. Swuec, J.J. O'Rourke, J. Heierhorst, W. Crismani, and A.J. Deans. 2017. Mechanism of Ubiquitination and Deubiquitination in the Fanconi Anemia Pathway. *Mol Cell*. 65:247-259.

Vandenberg, C.J., F. Gergely, C.Y. Ong, P. Pace, D.L. Mallery, K. Hiom, and K.J. Patel. 2003. BRCA1-Independent Ubiquitination of FANCD2. *Mol Cell*. 12:247-254.

Vermeulen, W., and M. Fousteri. 2013. Mammalian Transcription-Coupled Excision Repair. *Cold Spring Harbor Perspectives in Biology*. 5.

Vieira, K.F., P.P. Levings, M.A. Hill, V.J. Crusselle, S.-H.L. Kang, J.D. Engel, and J. Bungert. 2004. Recruitment of Transcription Complexes to the β-Globin Gene Locus in Vivo and in Vitro. *Journal of Biological Chemistry*. 279:50350-50357.

Vignard, J., G. Mirey, and B. Salles. 2013. Ionizing-radiation induced DNA double-strand breaks: a direct and indirect lighting up. *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology*. 108:362-369.

Vilette, D., S.D. Ehrlich, and B. Michel. 1995. Transcription-induced deletions in Escherichia coli plasmids. *Molecular microbiology*. 17:493-504.

Wakasugi, M., A. Kawashima, H. Morioka, S. Linn, A. Sancar, T. Mori, O. Nikaido, and T. Matsunaga. 2002. DDB Accumulates at DNA Damage Sites Immediately after UV Irradiation and Directly Stimulates Nucleotide Excision Repair. *Journal of Biological Chemistry*. 277:1637-1640.

Walker, A.I., T. Hunt, R.J. Jackson, and C.W. Anderson. 1985. Double-stranded DNA induces the phosphorylation of several proteins including the 90 000 mol. wt. heat-shock protein in animal cell extracts. *Embo J*. 4:139-145.

Walter, D., A. Lier, A. Geiselhart, F.B. Thalheimer, S. Huntscha, M.C. Sobotta, B. Moehrle, D. Brocks, I. Bayindir, P. Kaschutnig, K. Muedder, C. Klein, A. Jauch, T. Schroeder, H. Geiger, T.P. Dick, T. Holland-Letz, P. Schmezer, S.W. Lane, M.A. Rieger, M.A.G. Essers, D.A. Williams, A. Trumpp, and M.D. Milsom. 2015. Exit from dormancy provokes DNA-damage-induced attrition in haematopoietic stem cells. *Nature*. 520:549.

Wang, M., E.J. McIntee, G. Cheng, Y. Shi, P.W. Villalta, and S.S. Hecht. 2000. Identification of DNA Adducts of Acetaldehyde. *Chemical Research in Toxicology*. 13:1149-1157.

Wang, X., R.D. Kennedy, K. Ray, P. Stuckert, T. Ellenberger, and A.D. D'Andrea. 2007. Chk1-Mediated Phosphorylation of FANCE Is Required for the Fanconi Anemia/BRCA Pathway. *Molecular and Cellular Biology*. 27:3098-3108.

Warner, J.R. 1999. The economics of ribosome biosynthesis in yeast. *Trends in biochemical sciences*. 24:437-440.

Wellinger, R.E., F. Prado, and A. Aguilera. 2006. Replication Fork Progression Is Impaired by Transcription in Hyperrecombinant Yeast Cells Lacking a Functional THO Complex. *Molecular and Cellular Biology*. 26:3327-3334.

Whitmore, S.E., C.S. Potten, C.A. Chadwick, P.T. Strickland, and W.L. Morison. 2001. Effect of photoreactivating light on UV radiation-induced alterations in human skin. *Photodermatology, photoimmunology & photomedicine*. 17:213-217.

Whitney, M.A., H. Saito, P.M. Jakobs, R.A. Gibson, R.E. Moses, and M. Grompe. 1993. A common mutation in the FACC gene causes Fanconi anaemia in Ashkenazi Jews. *Nature genetics*. 4:202.

Wilson, M.D., M. Harreman, and J.Q. Svejstrup. 2013. Ubiquitylation and degradation of elongating RNA polymerase II: the last resort. *Biochimica et biophysica acta*. 1829:151-157.

Woodford, K.J., R.M. Howell, and K. Usdin. 1994. A novel K(+)-dependent DNA synthesis arrest site in a commonly occurring sequence motif in eukaryotes. *J Biol Chem*. 269:27029-27035.

Wu, L., and I.D. Hickson. 2003. The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature*. 426:870.

Xu, B., and D.A. Clayton. 1996. RNA-DNA hybrid formation at the human mitochondrial heavy-strand origin ceases at replication start sites: an implication for RNA-DNA hybrids serving as primers. *The EMBO journal*. 15:3135-3143.

Xu, J., I. Lahiri, W. Wang, A. Wier, M.A. Cianfrocco, J. Chong, A.A. Hare, P.B. Dervan, F. DiMaio, A.E. Leschziner, and D. Wang. 2017. Structural basis for the initiation of eukaryotic transcription-coupled DNA repair. *Nature*. 551:653-657.

Yagasaki, H., T. Oda, D. Adachi, T. Nakajima, T. Nakahata, S. Asano, and T. Yamashita. 2003. Two common founder mutations of the fanconi anemia group g gene FANCG/XRCC9 in the Japanese population. *Human Mutation*. 21:555-555.

Yamamoto, K.N., S. Kobayashi, M. Tsuda, H. Kurumizaka, M. Takata, K. Kono, J. Jiricny, S. Takeda, and K. Hirota. 2011. Involvement of SLX4 in interstrand cross-link repair is regulated by the Fanconi anemia pathway. *Proc Natl Acad Sci U S A*. 108:6492-6496.

Yel, L., S. Aggarwal, and S. Gupta. 1999. Cartilage-hair hypoplasia syndrome: increased apoptosis of T lymphocytes is associated with altered expression of Fas (CD95), FasL (CD95L), IAP, Bax, and Bcl2. *Journal of clinical immunology*. 19:428-434.

Zerbino, D.R., P. Achuthan, W. Akanni, M R. Amode, D. Barrell, J. Bhai, K. Billis, C. Cummins, A. Gall, C.G. Girón, L. Gil, L. Gordon, L. Haggerty, E. Haskell, T. Hourlier, O.G. Izuogu, S.H. Janacek, T. Juettemann, J.K. To, M.R. Laird, I. Lavidas, Z. Liu, J.E. Loveland, T. Maurel, W. McLaren, B. Moore, J. Mudge, D.N. Murphy, V. Newman, M. Nuhn, D. Ogeh, C.K. Ong, A. Parker, M. Patricio, H.S. Riat, H. Schuilenburg, D. Sheppard, H. Sparrow, K. Taylor, A. Thormann, A. Vullo, B. Walts, A. Zadissa, A. Frankish, S.E. Hunt, M. Kostadima, N. Langridge, F.J. Martin, M. Muffato, E. Perry, M. Ruffier, D.M. Staines, S.J. Trevanion, B.L. Aken, F. Cunningham, A. Yates, and P. Flicek. 2018. Ensembl 2018. *Nucleic Acids Res*. 46:D754-D761.

Zhang, B.N., A. Bueno Venegas, I.D. Hickson, and W.K. Chu. 2018. DNA replication stress and its impact on chromosome segregation and tumorigenesis. *Seminars in cancer biology*.

Zhang, J., J.M. Dewar, M. Budzowska, A. Motnenko, M.A. Cohn, and J.C. Walter. 2015. DNA interstrand cross-link repair requires replication-fork convergence. *Nat Struct Mol Biol*. 22:242-247.

Zhang, S., C. Pondarre, G. Pennarun, H. Labussiere-Wallet, G. Vera, B. France, M. Chansel, I. Rouvet, P. Revy, B. Lopez, J. Soulier, P. Bertrand, I. Callebaut, and J.-P. de Villartay. 2016. A nonsense mutation in the DNA repair factor Hebo causes mild bone marrow failure and microcephaly. *The Journal of Experimental Medicine*.

Zhao, H., J.L. Watkins, and H. Piwnica-Worms. 2002. Disruption of the checkpoint kinase 1/cell division cycle 25A pathway abrogates ionizing radiation-induced S and G2 checkpoints. *Proc Natl Acad Sci U S A*. 99:14795-14800.

Zhao, J., A. Bacolla, G. Wang, and K.M. Vasquez. 2010. Non-B DNA structure-induced genetic instability and evolution. *Cellular and molecular life sciences : CMLS*. 67:43-62.

Zhao, W., J.B. Steinfeld, F. Liang, X. Chen, D.G. Maranon, C. Jian Ma, Y. Kwon, T. Rao, W. Wang, C. Sheng, X. Song, Y. Deng, J. Jimenez-Sainz, L. Lu, R.B. Jensen, Y. Xiong, G.M. Kupfer, C. Wiese, E.C. Greene, and P. Sung. 2017. BRCA1–BARD1 promotes RAD51-mediated homologous DNA pairing. *Nature*. 550:360.

Zhou, B.-B.S., and S.J. Elledge. 2000. The DNA damage response: putting checkpoints in perspective. *Nature*. 408:433.

Zimmermann, M., F. Lottersberger, S.B. Buonomo, A. Sfeir, and T. de Lange. 2013. 53BP1 Regulates DSB Repair Using Rif1 to Control 5' End Resection. *Science*. 339:700-704.

Zou, L., and S.J. Elledge. 2003. Sensing DNA Damage Through ATRIP Recognition of RPA-ssDNA Complexes. *Science*. 300:1542-1548.