Investigating the Role of ERCC6L2 in Haematopoiesis and Lymphocyte Development



Pia Isabella Reichl

St Edmund Hall

Thesis submitted for the degree of
Doctor of Philosophy in Medical Sciences
at the University of Oxford

Trinity Term 2025

ACKNOWLEDGEMENTS

Ah yes, the part I've been looking forward to writing most! Like all PhDs, mine was a lot of hard work that required blood (*mostly* not mine), sweat, and tears. I was, however, uniquely privileged in the people that helped me along the way:

First of all, a big thank you to Ross! I'm deeply grateful that you took a chance on me - especially when I got rejected in my first PhD application and you took me on as an RA, to support me whilst we tried again. We might have butted heads more than either of us would have liked, but I could not imagine doing my PhD anywhere else! Your knowledge of the field is honestly inspiring and your approach to science something I hope to take with me.

Another big thank you goes to Adam! You've been a phenomenal secondary supervisor and I enjoyed every single conversation where I got to bug you about all things haematopoiesis. My first results chapter simply would not exist without your help. A special thank you for also sitting down with me and talking about career options in my last year!

To Ash, thank you so, so much! Not only are you one of my favourite humans, you also taught me everything I know about mouse work. Your nonchalant mastery of all you do in the lab and the science behind it is something that inspires me every day.

To Jean, thank you for always being ready to lend a helping hand and completing our little mouseketeer trio!

To Mary, my sibling in science, it was a true joy to grow as a scientist together with you!

To the rest of the Chapman lab (My, Asger, Lucia, Li-Yao, Patrik, Monica, Sarah, Alex). Thank you for being such wonderful lab mates! Even on the shitty days, when it felt like nothing about this project was working and science was personally against me, I was still always looking forward to coming in because you guys are just so delightful - smart, helpful, but also just fantastic humans to hang out with. I cherished every pub trip, coffee, laugh, and whinge - all of which there were many.

To my family (Mama, Papa, Pati), the knowledge that I could always come home and you'd support me no matter what, has given me the confidence to try out scary things (like moving to a new country not once but twice). I love coming home and decidedly not talking about science; it reminds me that there's so much more to life than my tiny bubble of academia.

To my Oxford friends: Sanja, Cat, Piyush, Sarah, and JB. Thank you for sharing your love, laughter, and chaos with me. My time in the UK would have been pretty damn miserable without you.

To the "Harte Kern" - Eli, Oli, Jasmine - thank you for being such wonderful friends even when we are scattered across Europe. Spending New Year's together is one of my highlights of every year.

To Manu, you are my home. Thank you for always being there and making my world brighter. Thank you for listening to me whinge constantly and lending me your calmness and strength. My biggest streak of luck has been getting to live life by your side. I quite literally wouldn't even be in Oxford if it weren't for you. I love you.

Lastly, a thank you to the mice that gave their lives for this research. As important as I think this part of science is, none of it was done lightly.

My DPhil studies have been funded by a Radcliffe Department of Medicine studentship and additional funding was provided by a Lister Institute of Preventative Medicine grant.

DECLARATION OF AUTHORSHIP

I hereby declare that this thesis has been solely written by myself and has not been submitted for any degree at the University of Oxford, or elsewhere. The results presented herein are my own work, unless otherwise specified in the text.

ABSTRACT

Investigating the role of ERCC6L2 in haematopoiesis and lymphocyte development

Pia Isabella Reichl

MRC Molecular Haematology Unit, MRC Weatherall Institute of Molecular Medicine, Radcliffe Department of Medicine and St Edmund Hall, University of Oxford.
Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Medical Sciences, Trinity 2025.

Mutations in *ERCC6L2* cause an inherited bone marrow failure syndrome in humans. However, the mechanism of how ERCC6L2 supports haematopoietic stem cell (HSC) function remains poorly understood. ERCC6L2 has been proposed to be involved in DNA double strand break (DSB) repair via non homologous end-joining (NHEJ) as well as DNA interstrand crosslink repair via the Fanconi anaemia pathway. Using genetically modified mouse models, I show that ERCC6L2 is dispensable for murine blood development and that HSC attrition unlikely to be caused by endogenous aldehydes, which reveals a distinct disease aetiology that separates it from Fanconi anaemia. Instead, I show that ERCC6L2 functions as a NHEJ factor. Specifically, while it is dispensable for the repair of DSBs during V(D)J recombination, ERCC6L2 supports NHEJ during class switch recombination (CSR) at staggered DSB ends. This function is dependent on 53BP1, yet independent of the shieldin complex, revealing a parallel pathway to support NHEJ during CSR. Unbiased mass spectrometry and *in-silico* modelling reveal that ERCC6L2 exerts its function in a complex with KU and MRI. Finally, considering the similar structure of staggered DSBs at telomere ends, I postulate that ERCC6L2 may exert its stem cell protective function at telomeres.

CONTENTS

Acknowledgements **v**

Declaration of Authorship **iii**

Abstract **v**

Contents **vii**

List of Figures **xi**

List of Tables **xiii**

1 Introduction **1**

1.1 Double strand break repair 1

1.1.1 DSB recognition, signalling and repair pathway choice 2

1.1.2 Homologous recombination (HR) 3

1.1.3 Non-homologous end joining (NHEJ) 5

1.2 Haematopoiesis and its dependence on DNA repair 6

1.2.1 Haematopoietic differentiation - from stem cell to mature blood 6

1.2.2 Inherited bone marrow failure syndromes 8

1.3 Lymphocyte development and its dependence on NHEJ 12

1.3.1 Lymphocyte development 13

1.3.2 V(D)J recombination . . . 15
1.3.3 Class switch recombination (CSR) . . . 17
1.4 ERCC6L2 - DNA excision repair protein ERCC6-like 2 . . . 20
1.4.1 ERCC6L2 syndrome . . . 20
1.4.2 Molecular function . . . 22
1.5 Aims of this thesis . . . 27

2 Materials and Methods **29**

2.1 Mice and primary mouse tissue isolation methods . . . 29
2.1.1 Mouse strains . . . 29
2.1.2 Mouse strain generation . . . 30
2.1.3 Mouse genotyping . . . 30
2.1.4 Mouse embryonic fibroblast isolation . . . 31
2.1.5 *v-Abl* pre-B cell generation . . . 31
2.1.6 Splenic B-cell isolation and stimulation . . . 32
2.1.7 HSPC/c-Kit$^+$ enrichment . . . 33
2.1.8 Competitive bone marrow transplantation . . . 34
2.2 Cell lines and culture conditions . . . 34
2.3 CRISPR/Cas9 genome editing . . . 35
2.4 Drug sensitivity assay . . . 36
2.5 Flow Cytometry . . . 36
2.5.1 Peripheral blood chimerism . . . 37
2.5.2 Bone marrow engraftment chimerism . . . 37
2.5.3 Haematopoietic stem and progenitor cells . . . 38
2.5.4 HSPC culture . . . 39
2.5.5 B cell development - bone marrow . . . 40
2.5.6 B cell development - spleen . . . 42
2.5.7 T cell development . . . 43
2.5.8 B cell isotype switching . . . 44
2.5.9 V(D)J recombination assay . . . 45
2.5.10 Micronucleus assay . . . 46

2.5.11 CSR assay in CH12-F3 cells . . . 46
2.6 Proteomics and Mass Spectrometry . . . 47
2.6.1 Cell lysate preparation and complex immunoprecipiation . . . 47
2.6.2 sTRAP digestion and proteomic analysis . . . 49
2.7 SDS-PAGE-Western Blot . . . 51
2.8 AlphaFold modelling . . . 52
2.9 Reverse-transcription qPCR . . . 52
2.10 Genome-wide CRISPR/Cas9 knockout screen in primary HSPCs . . . 53
2.10.1 Titration of sgRNA library lentivirus . . . 53
2.10.2 Cell culture . . . 53
2.10.3 NGS sample preparation and sequencing . . . 54
2.10.4 Computational analysis . . . 55

3 Defining the role of ERCC6L2 in Haematopoiesis 57
3.1 Introduction . . . 57
3.2 Results . . . 58
3.2.1 *Ercc6l2* knockout mice have normal numbers of HS(P)Cs . . . 58
3.2.2 ERCC6L2 is dispensable for bone marrow engraftment and blood repopulation 59
3.2.3 ERCC6L2 syndrome is not driven by aldehyde-induced DNA damage . . . 62
3.2.4 A quest for answers - genome-wide CRISPR screen in *Ercc6l2*$^{-/-}$ HSPCs . . . 65
3.3 Concluding remarks . . . 72

4 Defining the role of ERCC6L2 in lymphocyte development 75
4.1 Introduction . . . 75
4.2 Results . . . 76
4.2.1 ERCC6L2 is dispensable for lymphocyte differentiation . . . 76
4.2.2 ERCC6L2 supports CSR independently of the shieldin complex . . . 82
4.3 Concluding remarks . . . 88

5 Investigating the molecular function of ERCC6L2 91
5.1 Introduction . . . 91
5.2 Results . . . 92

5.2.1 ERCC6L2 is a NHEJ factor 92
5.2.2 Proteomics reveal the ERCC6L2 interactome 96
5.2.3 ERCC6L2 interacts with KU via MRI 97
5.3 Concluding remarks 101

6 Discussion **105**
6.1 Why does the bone marrow not fail in *Ercc6l2*$^{-/-}$ mice? 105
6.2 ERCC6L2 as an accessory cNHEJ factor 108
6.3 ERCC6L2 as a telomere maintenance factor? 110
6.4 Conclusions and final model 113

References **115**

LIST OF FIGURES

1.1 Steps of HR. . . . 4
1.2 Steps of NHEJ. . . . 5
1.3 Simplified model of haematopoietic differentiation. . . . 7
1.4 Steps of the FA pathway. . . . 9
1.5 B cell development. . . . 14
1.6 T cell development. . . . 15
1.7 V(D)J recombination. . . . 16
1.8 Class switch recombination. . . . 18
1.9 Schematic representation of the human ERCC6L2 protein. . . . 24

2.1 The genomic structure of the *Ercc6l2* locus. . . . 31
2.2 Representative genotyping gels. . . . 33
2.3 Peripheral blood FACS gating scheme. . . . 38
2.4 Bone marrow engraftment FACS gating scheme . . . 40
2.5 HS(P)C FACS gating scheme. . . . 42
2.6 HSPC culture FACS gating scheme. . . . 43
2.7 Hardy fraction FACS gating scheme . . . 44
2.8 Splenic B cell FACS gating scheme. . . . 45
2.9 T cell development FACS gating scheme . . . 45
2.10 Isotype switching FACS gating scheme. . . . 47

2.11 *v-Abl* V(D)J recombination FACS gating scheme. 48
2.12 Micronucleus assay FACS gating scheme. 48
2.13 CH12-F3 isotype switching FACS gating scheme. 49

3.1 *Ercc6l2*$^{-/-}$ mice have normal numbers of HS(P)Cs. 58
3.2 Primary competitive bone marrow transplant. 61
3.3 Secondary competitive bone marrow transplant. 62
3.4 Combined genetic loss of *Ercc6l2* and *Adh5* does not lead to synergistic loss of HS(P)Cs. 63
3.5 Combined genetic loss of *Ercc6l2* and *Adh5* does not lead to synergistic genomic instability 64
3.6 Summary of set-up and quality control for genome-wide CRISPR screen 66
3.7 CRISPR screen - DDR hits 68
3.8 CRISPR screen results - telomere maintenance hits 70

4.1 ERCC6L2 is dispensable for B cell development - even in absence of 53BP1. . . . 77
4.2 ERCC6L2 is dispensable for mature B cells - even in abscence of 53BP1 78
4.3 ERCC6L2 is dispensable for T cell development - even in absence of 53BP1. . . . 80
4.4 ERCC6L2 participates in extra-chromosomal V(D)J recombination. 81
4.5 ERCC6L2 supports NHEJ during CSR downstream of 53BP1. 83
4.6 Combined loss of *Ercc6l2* and *Shld2* does not affect B cell development. 84
4.7 Combined loss of *Ercc6l2* and *Shld2* does not affect mature splenic B cells. . . . 86
4.8 ERCC6L2 supports NHEJ during CSR independently of the shieldin complex. . . 87

5.1 ERCC6L2 is a NHEJ factor. 93
5.2 Affinity purification-MS reveals the ERCC6L2 interactome. 95
5.3 AlphaFold3 model of the ERCC6L2-MRI-KU-complex. 98
5.4 ERCC6L2 interacts with KU via MRI to support NHEJ during CSR. 99

6.1 Final model of the ERCC6L2's function 114

LIST OF TABLES

2.1 *Ercc6l2* gRNA sequences for mouse strain generation. 30

2.2 Primer sequences for genotyping 32

2.3 gRNA sequences for genome editing. 36

2.4 List of antibodies for peripheral blood analysis. 37

2.5 List of antibodies for the biotin-conjugated lineage antibody stain. 39

2.6 List of antibodies for bone marrow engraftment chimerism analysis. 39

2.7 List of antibodies for bone marrow HSC analysis. 39

2.8 List of antibodies for HSPC culture analysis. 41

2.9 List of antibodies for B cell development analysis in the bone marrow. 41

2.10 List of antibodies for B cell development analysis in the spleen. 42

2.11 List of antibodies for T cell development analysis. 44

2.12 List of biotin-conjugated antibodies for B cell isotype switching analysis. 46

2.13 List of secondary antibodies for B cell isotype switching analysis. 46

2.14 List of primary antibodies used for Western blotting. 51

2.15 List of secondary antibodies used for Western blotting. 51

2.16 Primer sequences for RT-qPCR. 53

2.17 Primer sequences for ddPCR. 54

2.18 Thermocycler conditions for amplification of lentiviral gRNA integrations. 55

2.19 Primer sequences for CRISPR library PCR. 55

3.1 Results from MAGeCK count algorithm. 65

5.1 ERCC6L2 interactome detected via LC-MS^2. 97

Chapter

1 INTRODUCTION

Our genomes are constantly under attack from genotoxins stemming from both endogenous and exogenous sources. Consequently, we have evolved an intricately connected network of different DNA repair pathways, use of which depends on not only the type of damage, but also other cues such as cell cycle stage, severity of damage, and even cell type. Collectively, this is known as the DNA damage response (DDR) [1].

If DNA damage is not - or inaccurately - repaired, this can lead to a variety of cellular outcomes ranging from senescence or apoptosis, to oncogenic transformation. Genome instability is one of the hallmarks of cancer [2] and genetic defects in DNA repair factors lead to a vast array of disease, a small selection of which will be discussed in more detail later in this thesis.

1.1 Double strand break repair

The cleavage of the phosphodiester backbone on both strands of DNA is known as a DNA double strand break (DSB). DSBs are arguably the most toxic DNA lesion, as they can lead to complete loss of genetic information and genome rearrangements. Like all DNA lesions, DSBs can arise

from both exogenous sources (e.g. ionizing radiation (IR), chemotherapeutic (such as cisplatin or etoposide), CRISPR/Cas9-mediated genome engineering) or endogenous sources (e.g. replication fork collapse, replication-transcription conflicts, abortive activity of topoisomerase II (TOP2), or DNA modifying enzymes (e.g. AID, RAG1/2, SPO11)) [3].

In response, cells have evolved two main pathways to repair DSBs: homologous recombination (HR) and non-homologous end-joining (NHEJ) - although several other specialised DSB repair pathways exist [4]. How cells decide which of those DSB repair pathways to use (i.e. DSB repair pathway choice) is a still evolving, highly active field of research. However, whilst the intricacies are still being debated, it is clear that the two main determinants of pathway choice are cell cycle stage and DNA end structure/chemistry [5]. Notably, both HR and NHEJ are actively repressed during mitosis, with the predominant DSB repair pathway during M phase being microhomology-mediated end joining (MMEJ) [6, 7].

1.1.1 DSB recognition, signalling and repair pathway choice

When a DSB occurs, the MRN (MRE11-RAD50-NBS1) complex rapidly recognizes the break, and recruits and activates the kinase ATM [8]. ATM then initiates a complex signalling cascade by phosphorylating a plethora of downstream targets, such as p53, CHK2, and importantly, serine 139 (Ser139) on the histone variant H2AX [9, 10]. Phosphorylated H2AX (termed γH2AX) is recognized and bound by the scaffolding protein MDC1 [11]. MDC1 itself subsequently gets phosphorylated by ATM, allowing for the binding of the E3 ligase RNF8 [12–14]. RNF8 ubiquitinates histones H2A, H2AX, and, importantly, lysine 63 (K63) on histone H1. The E3 ligase RNF168 recognises and binds this mark and subsequently deposits ubiquitin on lysine 13 and 15 (K13/K15) [15, 16]. Together, RNF8 and RNF168 amplify the chromatin ubiquitylation signal to a threshold that allows for the recruitment of 53BP1 and BRCA1 [17, 18].

The tumour suppressor proteins 53BP1 and BRCA1 are at the heart of an intricate process determining the pathway being chosen for DSB repair. It has been proposed that 53BP1 and BRCA1 mainly control this decision by regulating the extent of DNA end resection at DSBs, a crucial early step in homology-directed repair [5].

In particular, 53BP1 recognises two histone marks: dimethylated lysine 20 on histone 4 (H4K20me2) and H2AK15ub, deposited by RNF168 in response to DNA damage [19, 20]. Notably, DNA replication leads to a dilution of H4K20me2 on chromatin, thereby disfavouring 53BP1 recruitment when DSBs can be repaired by HR [21]. 53BP1 accumulation on chromatin inhibits DNA end resection, thus favouring NHEJ [22]. This inhibition critically depends on 53BP1's downstream effector RIF1 [23, 24]. In contrast, the shieldin complex (consisting of REV7, SHLD1, SHLD2, and SHLD3; discussed in detail in Section1.3.3), which is another downstream effector of 53BP1, contribute to HR suppression to a much lesser extent, demonstrating distinct functions of 53BP1 depending on the repair context [25].

During S and G_2 phase, BRCA1 - together with its obligate binding partner BARD1 - antagonizes 53BP1 and its effectors on chromatin, ultimately favouring HR. [5] BRCA1-BARD1 can engage the combination of DNA-damage (H2AK15ub) and post-replication (H4K20me0) histone PTM states, accumulating at damage proximal chromatin and excluding 53BP1 [26, 27]. Additionally, recent work reports that BRCA1-BARD1 has further pro-HR functions independent of its antagonism of 53BP1, particularly through the BRCA1-C complex (BRCA1-BARD1, MRN, and CtIP), which promotes long-range resection [28, 29].

1.1.2 Homologous recombination (HR)

During HR, resection of the 5' end of the DSB is initiated by CtIP and the MRN complex creating a short 3' overhang (short range resection) (Fig. 1.1). This overhang can then be further extended by the exonucleases EXO1, or the endonuclease DNA2 acting on ssDNA unwound by

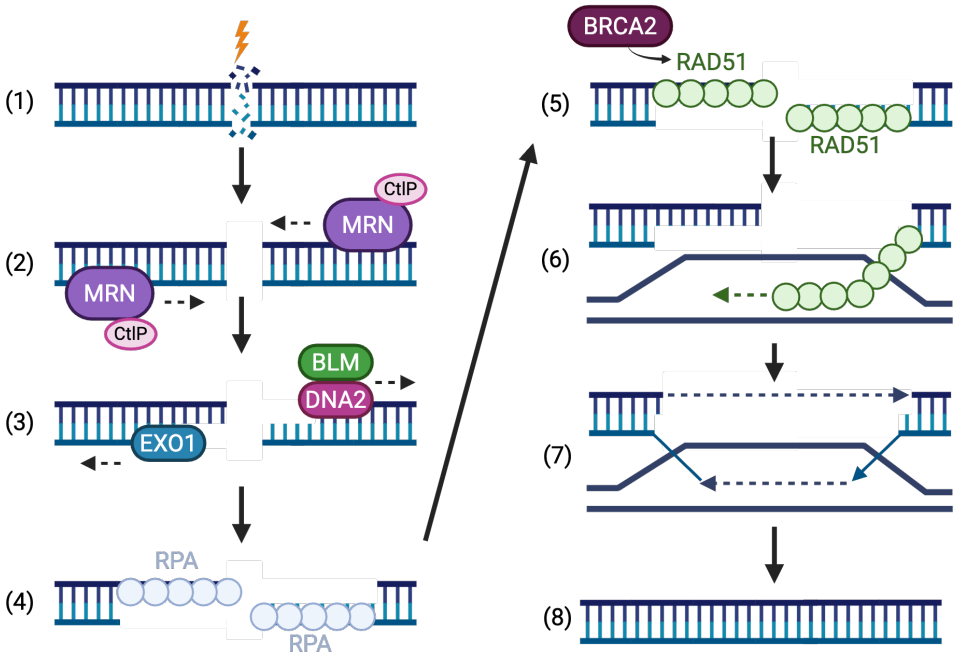


Figure 1.1: **Steps of HR.**(1) DSB generation (2) short-range resection (3) long-range resection (4) RPA coating of ssDNA (5) RAD51 loading (D) D-loop formation and DNA synthesis (7) SDSA or Holliday junction resolution or dissolution (8) repaired DNA

the helicases BLM and WRN (long-range resection). The single-stranded overhang is then coated by RPA with high affinity. BRCA1-BARD1 then mediates the recruitment of PALB2 to BRCA2, which then replaces RPA with RAD51 filaments [30]. Rad51 nucleofilaments (i.e. Rad51-coated DNA) execute the homology search, resulting in strand invasion of the repair template, forming a structure called displacement loop (D-loop). DNA synthesis elongates the resected 3' overhang using the invaded strand as a template. In mitotic cells, synthesis-dependent strand annealing (SDSA) is thought to be the dominant pathway to used to process the D loop structure, as avoiding crossover minimises the risk of loss of heterozygosity. But re-ligation of the DNA ends of the DSB can also lead to the generation of a double Holliday junction that can be dissolved by the BTR (BLM-TOP3A-RMI1/2) complex or resolved by SLX4/SLX1, Mus81/EME1 or GEN1 [31, 32].

HR is considered to be an error-free DSB repair pathway as it uses a repair template to restore the original DNA sequence. However, due to its dependence on a repair template (usually the

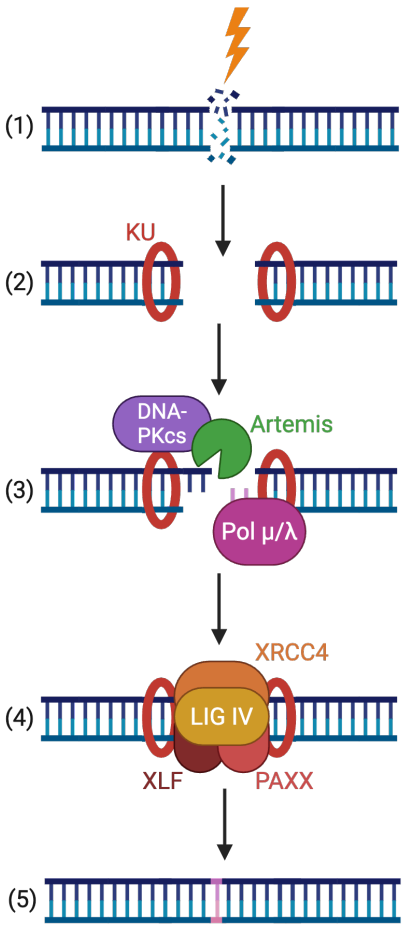


Figure 1.2: **Steps of NHEJ.** (1) DSB generation, (2) detection of DSB by KU, (3) DNA end processing as required (4) ligation of break by the Lig IV/XRCC4 complex supported by XLF and/or PAXX, (5) repaired DNA, insertions or deletions are possible

sister chromatid), HR is restricted to the S and G_2 stages of the cell cycle [27, 33]. This also means it is unavailable for DSB repair in quiescent or post-mitotic cells.

1.1.3 Non-homologous end joining (NHEJ)

During NHEJ, the KU70/KU80 heterodimer (Ku) is the first factor to bind the DSB (Fig. 1.2). DNA-PKcs is then recruited to the site by KU (forming the DNA-PK holoenzyme) and subsequently undergoes autophosphorylation [34]. The DNA-PK holoenzyme is required to tether the broken DNA ends together in the long-range synaptic complex [35]. Then, if the end structures do not allow for direct ligation, ends are processes in the form of excision, modification, or addition of nucleotides: DNA-PKcs recruits and activates Artemis, which endolytically removes 5' and 3' DNA overhangs to create an end structure compatible for ligation by XRCC4/Lig IV. [36] In

contrast to Artemis, DNA polymerases Pol μ and Pol λ can add nucleotides to the DNA ends, helping to reduce the extent of deletions at NHEJ substrates by helping to fill-in gaps in non-complementary ends [37]. Once they ends are compatible, DNA-PK switches from the long-range synaptic complex to the short-range complex, closely aligning the DNA ends. This change in the synaptic complex is aided by the NHEJ accessory factors XLF or PAXX. Once the DNA ends are close proximity, this allows for re-ligation by the XRCC4-Lig IV complex [38].

Due to the requirement for end processing, repair by NHEJ can lead to insertions or deletions at the break site, leading to point mutations. Additionally, NHEJ can lead to genome rearrangements or chromosome fusions as any two break sites can be ligated together. Still, NHEJ is the predominant DSB repair pathway in mammals and accounts for the repair of nearly all DSBs outside of S and G_2 phase [39].

1.2 Haematopoiesis and its dependence on DNA repair

1.2.1 Haematopoietic differentiation - from stem cell to mature blood

Haematopoiesis is defined as the continuous production of various blood lineages to sustain homeostasis of the blood system. The mature blood lineages being produced during this process (i.e. erythrocytes, megakaryocytes/platelets, monocytes, macrophages, neutrophils, and lymphocytes) carry out vastly different functions and have a wide range of half-lives, hence requiring both high levels of plasticity and proliferative potential [40].

At the apex of haematopoiesis reside haematopoietic stem cells (HSCs), a rare, largely quiescent cell population located in the bone marrow of adult mammals. Functionally, HSCs are defined by their twin capabilities of self-renewal and multipotent differentiation, meaning whilst they are capable of producing more HSCs, they are also able to produce all mature blood lineages.

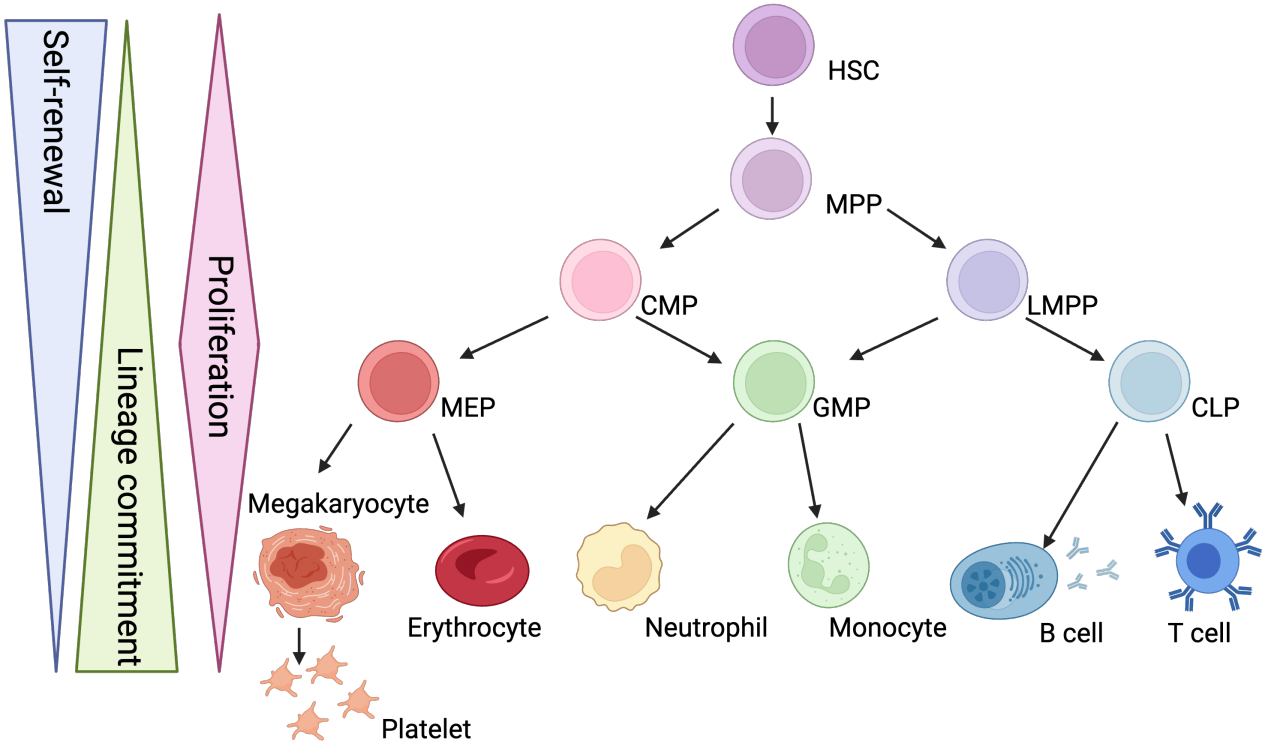


Figure 1.3: **Simplified model of haematopoietic differentiation.** During blood development, cells undergo a cascade of differentiation steps to ultimately produce all lineages of mature blood cells. Schematic was generated using BioRender. HSC, haematopoietic stem cell; MPP, multipotent progenitor; CMP, common myeloid progenitor; LMPP, lymphoid-primed multipotent progenitor; MEP, megakaryocyte-erythroid progenitor; GMP, granulocyte-macrophage progenitor; CLP, common lymphoid progenitor

During haematopoiesis, HSCs give rise to progenitors that progressively become more restricted to one or several lineages, whilst simultaneously losing their self-renewal capacity (Fig. 1.3) [41]. In practical terms, this means that whereas HSCs can be serially transplanted and continue to output all blood lineages, progenitors have restricted lineage differentiation capacity and are usually lost within a few weeks after transplantation [42].

Importantly, whilst simple models of haematopoiesis imply discrete steps leading from HSCs to mature blood cells, blood differentiation is actually a continuous process [43]. The exact relationships between HSCs and various types of progenitors, as well as the functional heterogeneity within HSCs themselves remains a highly active field of research [40].

1.2.2 Inherited bone marrow failure syndromes

Inherited bone marrow failure syndromes (IBMFS) are a heterogeneous group of genetic disorders characterized by the inability of bone marrow to produce an adequate number of mature blood cells (i.e. cytopenia). Cytopenias caused by IBMFS can be both uni- or multi-lineage, meaning they can affect erythrocytes (anaemia), megakaryocytes (thrombocytopenia), and lymphocytes (lymphocytopenia), both individually or in combination. This loss of mature blood cells is ultimately due to a loss of haematopoietic stem and progenitor cells in the bone marrow. IBMFS also predispose patients to cancer and are associated with a significant premature mortality [44].

Mechanistically, IBMFS can be broadly categorized into three groups: (1) DNA repair defects, (2) telomere maintenance defects, and (3) ribosome biogenesis defects. Examples of (1) and (2) are discussed in detail hereafter. Ribosome biogenesis defects (3) are of less relevance for this thesis and won't be discussed further.

1.2.2.1 Fanconi anaemia

Fanconi anaemia (FA) is the most common IBMF. It is clinically heterogeneous, with common symptoms including short stature, skin hypopigmentation, hypogonadism, and developmental delays. However, up to one third of patiens with FA do not have any congenital abnormalities [45]. The two most common features of FA are the progressive development of bone marrow failure - usually during the first two decades of life - and a markedly increased risk of developing cancer. FA patients develop cancer at an earlier age than the general populace and are particularly susceptible to develop haematological malignancies and solid tumours. Genetically, FA is caused by germline mutations in one of at least 19 genes involved in the Fanconi anaemia pathway (*FANCA* to *FANCT*) [46].

The Fanconi anaemia pathway is the major DNA interstrand crosslink repair pathway in eukaryotes. DNA interstrand crosslinks (ICLs) are aberrant covalent bonds between the two strands

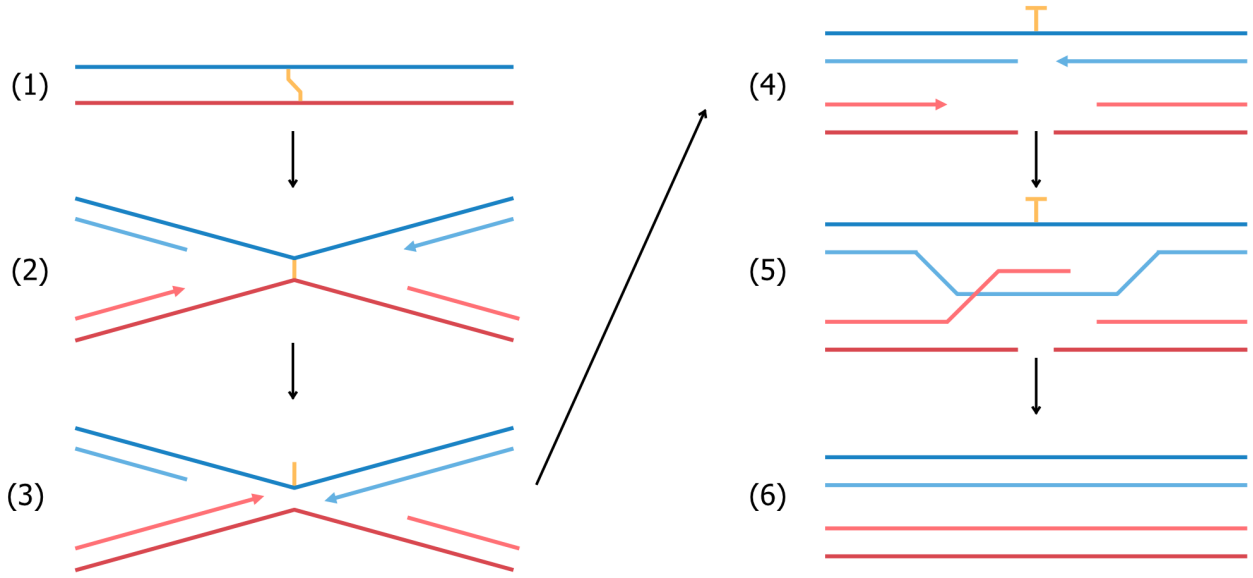


Figure 1.4: **Steps of the Fanconi anaemia pathway.** (1) ICL (2) lesion recognition & fork convergence (3) nucleolytic incision & unhooking (4) lesion bypass and translesion synthesis (5) HR & NER (6) repaired DNA

of DNA, preventing the progression of transcription and replication by blocking the unwinding of the DNA double-strand [47]. The FA pathway removes ICLs encountered during replication through a unique combination of nucleotide excision repair, translesion synthesis, and homologous recombination.

Briefly, FANCM gets activated through phosphorylation by ATR and binds to the lesion (Fig. 1.4). FANCM then functions as a landing hub for the FA core complex (compromising 14 FANC proteins). In turn, the core complex ubiquitinates the FANCD2-FANCI heterodimer, thereby activating it. FAND2-I recruits several structure-specific nucleases (such as XPF-ERCC1, MUS81-EME1, and SLX1) via the nuclease scaffold protein SLX4, to excise the lesion thereby "unhooking" the ICL from one of the DNA strands. This results in the formation of a DSB in one of the daughter strands and a continuous DNA strand still containing the nucleotide adduct in the other. The adduct is subsequently removed, likely via nucleotide excision repair (NER). The DSB subsequently gets repaired via HR (discussed in detail in Section 1.1.2). For the second daughter strand, low-fidelity translesion synthesis polymerases (such as Pol ζ) carry out "lesion bypass" by incorporating nucleotides opposite the adduct at the cost of potentially introducing

mutations. Ultimately, the FA pathway enables the progression of DNA replication with two intact daughter strands [46].

Seminal work over the last decade has shown that endogenous aldehydes - such as formaldehyde - are a potent natural source of ICLs, requiring the FA pathway for repair [48, 49]. Interestingly, mice lacking genes involved in the FA pathway (e.g. *Fancd2*) have superficially normal blood development and do not spontaneously develop bone marrow failure - a stark contrast to the phenotype seen in humans. However, mice lacking both genes involved in the FA pathway and aldehyde detoxification (such as *Adh5* or *Aldh2*), rapidly succumb to aplastic anaemia caused by acute BMF during early adulthood, phenocopying Fanconi anaemia in humans [50, 51]. This suggests that endogenous aldehydes also drive BMF in Fanoni anaemia, a hypothesis bolstered by findings that defective *ALDH2* variants lead to accelerated BMF in FA patients [52]. More recently, Dingler *et al.* demonstrated that a FA-like phenotype can be induced in mice deficient for both *Aldh2* and *Adh5*, two enzymes essential for the detoxification of endogenous formaldehyde [53]. Notably, although these mice retained full proficiency in repairing ICLs through the canonical FA pathway, the complete absence of aldehyde clearance overwhelmed their DNA repair capacity, thereby revealing a strictly metabolic, aldehyde-driven route to Fanconi anaemia. Indeed, they were subsequently able to identify patients suffering from BMF of previously unknown aetiology that harboured mutations in both *ALDH2* and *ADH5*, demonstrating that this mechanism holds true in humans.

1.2.2.2 DNA ligase IV syndrome

DNA ligase IV syndrome - also known as LIG4 or Ligase 4 syndrome - is an extremely rare autosomal recessive disorder caused by mutations in the *LIG4* gene. The disease is characterized by congenital microcephaly, neurodevelopmental delays, abnormal "bird-like" facial features, bone marrow hypoplasia, immunodeficiency, and hypersensitivity to ionizing radiation [54]. Ligase IV is a core component of the NHEJ pathway, required for the re-ligation of the phosphate-backbone

of DNA ends (see Section 1.1.3). In mice, complete loss of LIG4 is embryonically lethal and mutations described in human patients are also hypomorphic, leading to significantly impaired NHEJ, although some activity is retained [55]. The immunodeficiency seen in patients is most likely caused by a combination of the loss of HSPCs due to progressive accumulation of DSBs, and defective V(D)J recombination and in T and B lymphocytes and CSR, since both processes rely on NHEJ. V(D)J recombination and CSR, and their reliance on NHEJ are discussed in Section 1.3.

1.2.2.3 Cerunnos/XLF-deficiency

Similarly to patients with mutations in *LIG4*, patients with mutations in *XLF/Cernunnos* suffer from a wide array of clinical features caused by defective NHEJ. In particular, the main clinical features include growth retardation, microcephaly, immunodeficiency, and hypersensitivity to ionizing radiation [56]. XRCC4-like factor (XLF) is another core component of the NHEJ pathway, supporting ligation by the ligase IV complex (see Section 1.1.3). However, mice lacking *Xlf* only have a modest decrease in the number of mature lymphocytes and can perform nearly normal V(D)J recombination [57]. More recent studies suggest that the lymphcytopenia in human XLF-deficiency syndrome is actually caused by premature aging of HSCs, instead of a V(D)J recombination defect [58].

1.2.2.4 Dyskeratosis congenita

Classical dyskeratosis congenita (DC) is caused by mutations in *DKC1* and characterized by the dermatological triad of abnormal skin pigmentation, oral leukoplakia, and nail dystrophy. Loss of dyskerin, the protein encoded by *DKC1*, destabilizes the telomerase complex, leading to premature shortening of patients' telomeres [44]. Nowadays, the term DC is generally used interchangeably with telomere biology disorders or telomeropathies, and encompasses disease cause by a wide

range of genes required for telomere maintenance (e.g. *CTC1*, *RTEL*, *TERT*, *TERC*, *TINF2*, among others) [59].

Briefly, telomeres are hexanucleotide tandem repeats (TTAGGG) at the end of vertebrate chromosomes solving two problems arising with linear chromosomes: the end-replication and the end-protection problem: Firstly, canonical DNA replication is unable to fully replicate the ends of linear DNA, leading to loss of genetic information - the end-replication problem [60]. In somatic tissues, telomeres shorten progressively at each cell division, protecting non-telomeric DNA from degradation. When telomeres themselves become too short, this leads to cellular senescence, a phenomenon also called the "Hayflick limit" [61, 62]. To counteract telomere attrition, certain cell types (such as germ cells) are able to synthesize new telomeric repeats using the enzyme telomerase, thereby resolving the attrition of the leading strand [63]. Meanwhile, CST-Polα-primase solves a second end-replication problem: During replication, lagging strand DNA synthesis is unable to completely duplicate the C-strand. CST-Polα-primase maintains the C-strand by mediating fill-in synthesis, thereby preventing premature shortening of the telomere [64].

Secondly, telomeres protect chromosome ends from being recognised as DSBs and activating the DDR - the end-protection problem [65]. This is achieved by the shelterin complex (consisting of POT1, RAP1, ,TIN2, TPP1, TRF1, TRF2). Shelterin binds to the end of the telomere and sequesters the DNA end in a lariat-structure called the t-loop, thereby avoiding a DNA-damage response [66].

1.3 Lymphocyte development and its dependence on NHEJ

1.3.1 Lymphocyte development

The (jawed) vertebrate adaptive immune system is based on two major cell lineages: B (for bone marrow-derived) and T (for thymus-derived) lymphocytes. The defining feature of lymphocytes is their ability to produce tailor-made antigen-binding receptors, allowing for highly specific recognition of pathogen-derived antigens, as well as long-term immune memory - all of this whilst tolerating self-antigens [67]. In order to be able to recognize virtually all pathogens, a staggering amount of different antigen receptors have to be produced. To this end, lymphocytes use programmed DSBs to rearrange whole segments of their antigen receptor genes during both V(D)J recombination and class switch recombination (CSR). Additionally, during somatic hypermutation, point mutations are generated through the mutagenic repair of cytosine deamination events in the *IgH* and *IgL* variable regions, generating slightly modified antigen binding receptors. During affinity maturation, B cells expressing receptors with increased antigen affinity are then selected [68].

Successful lymphocyte development is intrinsically linked with the proper assembly of the antigen binding receptor during several rounds of V(D)J recombination. Generally, cells undergo cycles of antigen receptor rearrangement, positive and negative selection based on recognising pathogens, whilst tolerating self antigens, and subsequent bursts of proliferation [67]. Furthermore, somatic hypermutation introduces point mutations within the *IgH* and *IgL* variable regions as a consequence of mutagenic repair following cytosine deamination. This process generates slightly modified antigen-binding receptors, which then undergo selection for enhanced antigen affinity, in a process called an antibody affinity maturation [68].

1.3.1.1 B cells

B cells produce immunoglobulins (Igs) as their antigen receptors. Immunoglobulins can be presented on the cell surface or secreted as antibodies, thereby mediating adaptive humoral immunity.

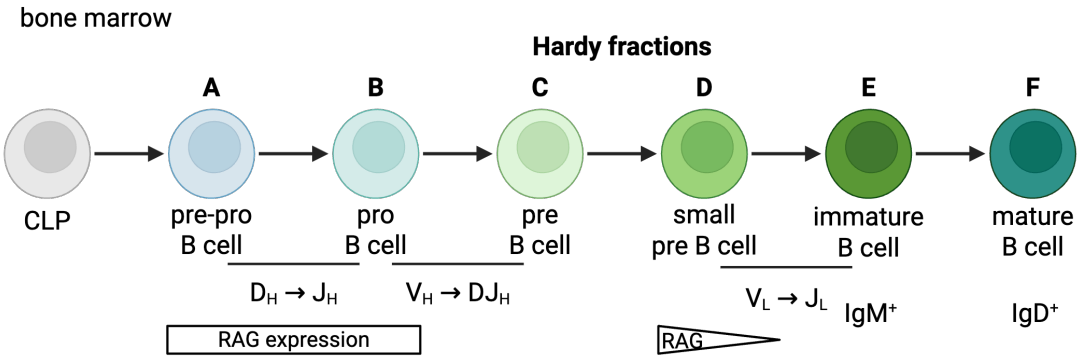


Figure 1.5: **Steps of B cell development.** B cells develop from common lymphoid progenitors into mature B cells undergoing subsequent stages of IgH and IgL rearrangements during differentiation. B cell development stages and their corresponding Hardy fractions (determined by flow cytometry) are listed in bold. Expression pattern of the RAG endonuclease complex during development is shown. Schematic was created using BioRender. CLP, common lymphoid progenitor

Structurally, Igs are compromised of the two copies of the Ig heavy (IgH) chain and two copies of the Ig light chain (IgL, i.e. Igκ or Igλ) chain. Both the heavy and light chain loci can undergo V(D)J recombination, however the light chain does not possess diversity segments [69].

To produce mature B cells, cells undergo several B cell progenitor stages which are closely linked to successful antigen receptor rearrangements (Fig. 1.5). In the bone marrow, common lymphoid progenitors differentiate into mature B cells in several distinct stages consisting of pre-pro B cells, pro B cells, pre B cells, small pre B cells, and immature B cells. Multicolour flow cytometry allows for the flow cytometric differentiation of the progenitor stages - so-called "Hardy fractions" [70, 71] - using panels of antibodies raised against distinct cell surface antigens.

1.3.1.2 T cells

T cells produce T cell receptors (TCRs) that are displayed on the cell surface. They can either kill infected cells directly (cytotoxic T cells) or support B and T cells by further activating them (helper T cells), thereby mediating cell-based adaptive immunity [67]. T cells development starts when common lymphoid progenitors differentiate into double-negative (DN, $CD4^-$ $CD8^-$) T cells in the thymus (Fig. 2.9). During the DN stage, cells must then successfully rearrange their TCRβ

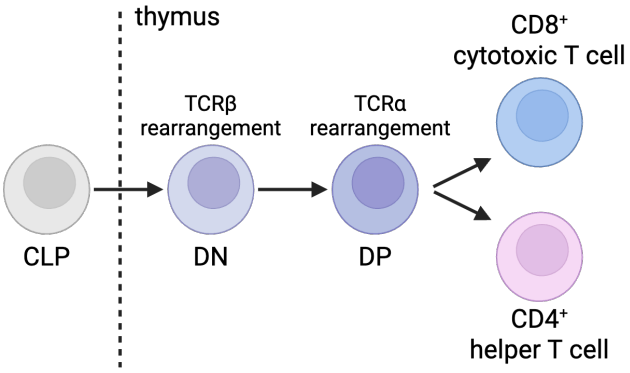


Figure 1.6: **Steps of T cell development.** T cells develop from common lymphoid progenitors into mature T cells undergoing subsequent stages of TCR rearrangements during differentiation. Schematic was created using BioRender. CLP, common lymphoid progenitor; DN, double-negative; DP, double-positive

chain [72]. Once they pass this hurdle, cells expand in a proliferative burst and become double-positive (DP, $CD4^+$ $CD8^+$) T cells. DP T cells then rearrange their TCRα chain, eventually expressing TCR$\alpha\beta$ complexes on their cell surface. Finally, cells undergo positive and negative selection, becoming either $CD8^+$ cytotoxic T cells or $CD4^+$ helper T cells.

1.3.2 V(D)J recombination

The adaptive immune response of vertebrates is critically dependent on the vast repertoire of different antigen receptors. To achieve antigen receptor diversity, lymphocytes undergo a process called V(D)J recombination [69]. During V(D)J recombination, specific sections of the antigen receptor genes, called variable (V), diversity (D), and joining (J) segments, are rearranged (Fig. 1.7a). V, D, and J segments each have multiple copies that get randomly selected for rearrangement, thereby creating a astounding number of potential combinations. Highlighting the importance of V(D)J recombination, mutations in genes encoding for factors centrally involved in this process (such as RAG1/2, or Artemis) cause severe combined immunodeficiency (SCID), due to the absence of mature lymphocytes [73, 74]. V(D)J recombination is highly similar between B and T cells with the major differences being the location of the cells (B cells undergo V(D)J recombination in the bone marrow and T cells in the thymus) and the rearranged loci (*Igh* and *Igl*/*Igk* in B cells and TCRα and TCRβ in T cells [**?**].

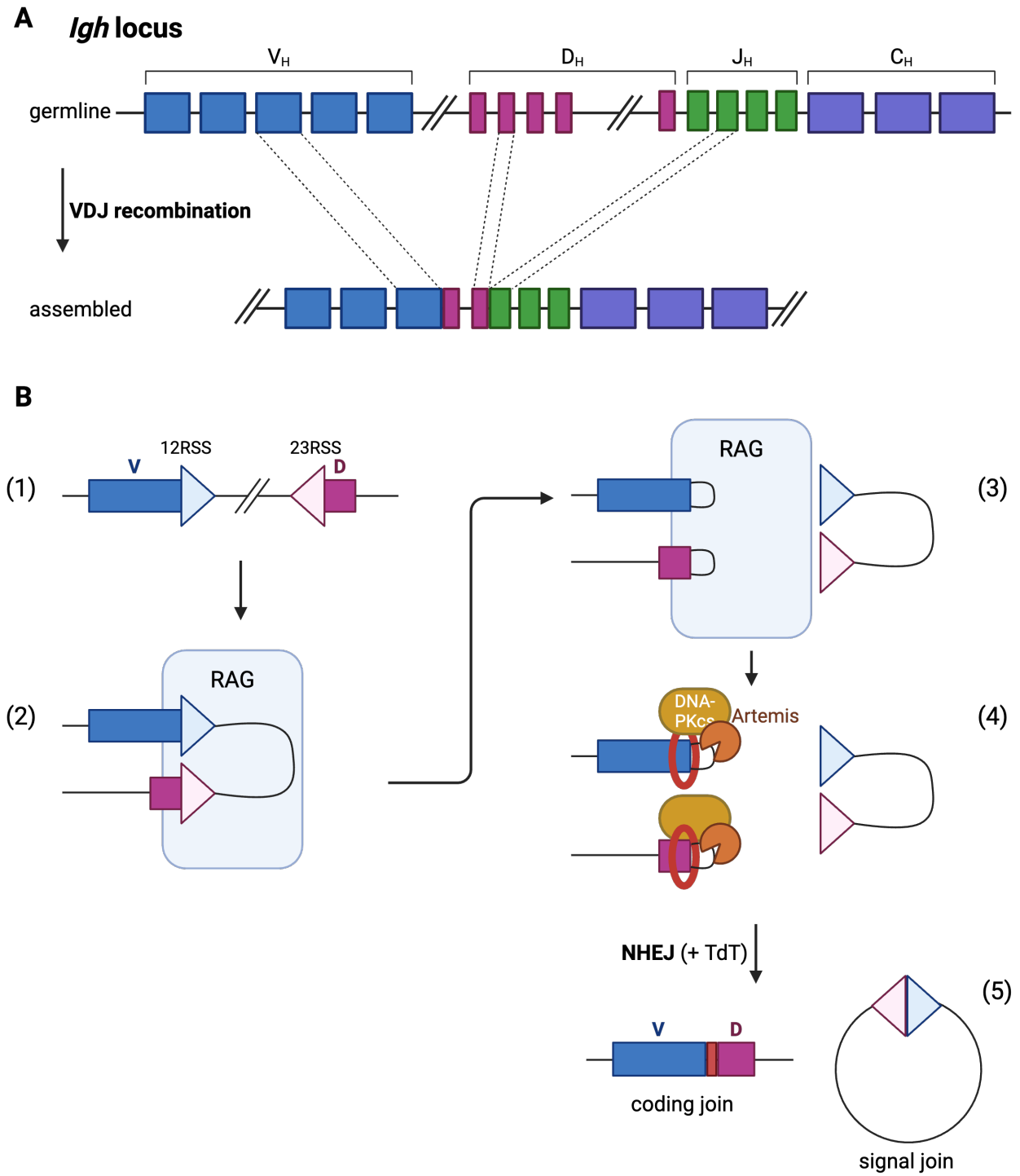


Figure 1.7: **V(D)J recombination.** **A** Structure of the *Igh* variable region. Variable (V), diversity (D), and J (joining) segments are represented as blue, pink, and green rectangles and the constant region as purple rectangles. **B** Steps of V(D)J recombination. (1) germline *Igh* locus (2) synapsis of 12/23 RSSs by RAG recombinase (3) hairpin formation (4) hairpin opening by Artemis-DNA-PKcs (5) assembled VD segment with insertions and signal joint. Schematic was created using BioRender.

Mechanistically, V(D)J recombination is initiated by the RAG recombinase complex formed by RAG1 and RAG2 (Fig. 1.7b) [34, 69]. RAG binds specific DNA sequences flanking each V,D, or J gene segment, called recombination signal sequences (RSSs). RSSs consist of conserved heptamer or nonamer elements separated by a spacer region of either 12 or 23 bases, which are consequently called 12RSS and 23RSS. During V(D)J recombination, there's a strong preference for recombination between one 12RSS and one 23RSS. Once RAG has bound an RSS, it then captures the second RSS, forming the synaptic complex. DNA is then cleaved in a two-step process. First, RAG introduces a nick between the RSS and the gene segment. The resulting 3' hydroxyl group then attacks the other DNA strand, forming a blunt DSB on the signal end and a sealed hairpin on the coding end. The blunt DSB can be directly re-ligated by NHEJ forming the circular signal joint. KU can bind the DNA ends and recruits DNA-PKcs and Artemis. The endonuclease Artemis opens the hairpin and the resulting DSB can subsequently be repaired via NHEJ. End processing during NHEJ might include further end resection by Artemis or nucleotide additions by Polμ, Polλ, or TdT. TdT in particular is only expressed in lymphocytes and adds nucleotides in a template-independent manner, further increasing the junction diversity.

1.3.3 Class switch recombination (CSR)

Immunoglobulins can be divided into different antibody classes/isotypes, namely IgM, IgD, IgG, IgE, and IgA. These differ in the constant region of their (IgH) heavy chain and have distinct tissue distributions and combat different pathogens. To change their isotype, mature B cells can undergo immunoglobulin class switch recombination, where the C_H exon cluster is exchanged from Cμ (for IgM) to Cγ (IgG), Cϵ (IgE), or Cα (IgA) (Fig. 2.10a). Importantly, IgD is not generated through CSR, but instead assembled using alternative splicing of the germline transcript, which encodes IgM [75].

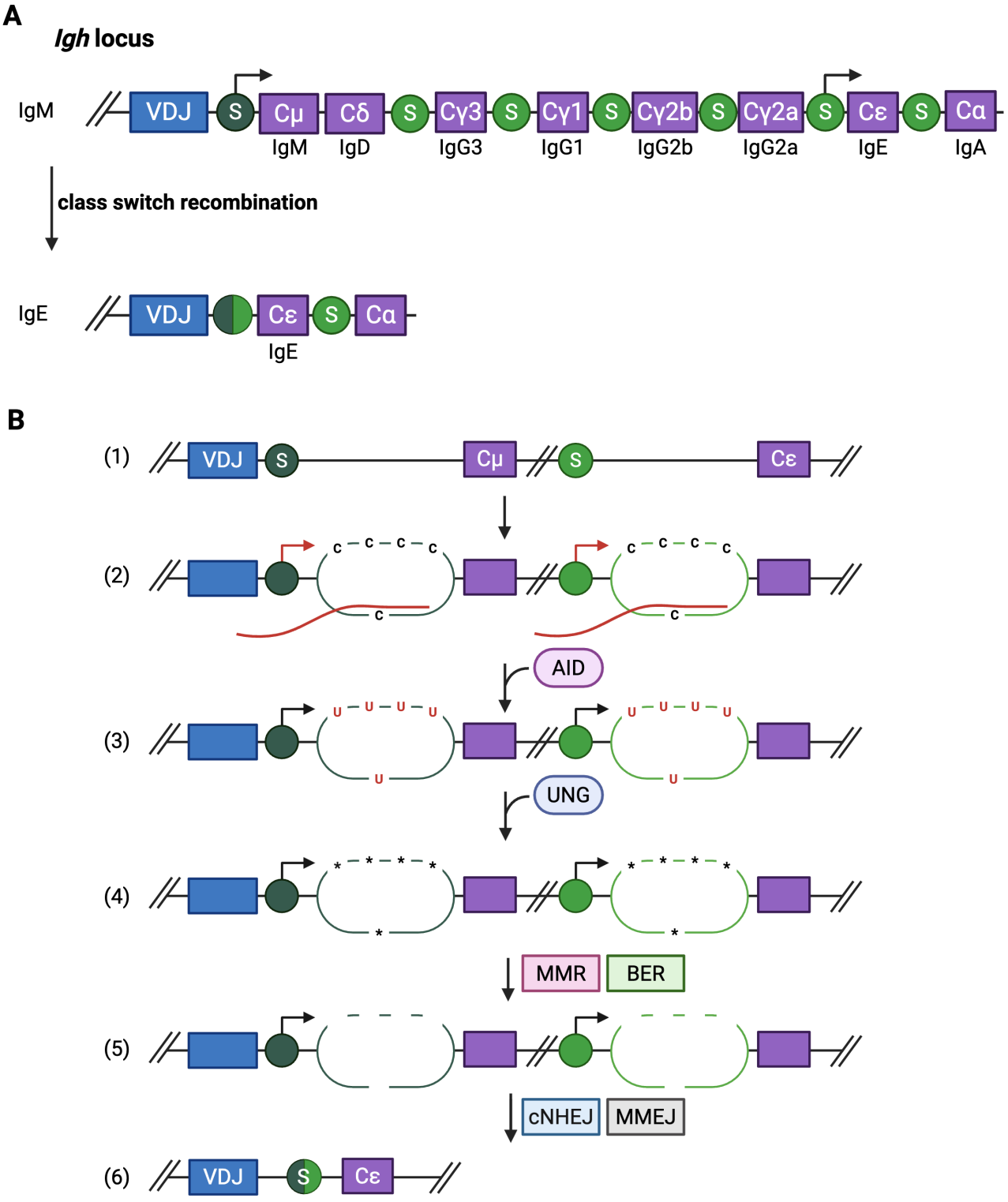


Figure 1.8: **Class switch recombination.** **A** Structure of the constant region of the *Igh* gene. The donor switch region is represented as a dark green rectangle, the acceptor switch regions as a light green rectangle. Each constant exon sets encoding for a different isotype is represented as purple rectangles. **B** Steps of CSR. (1) germline *Igh* constant region (2) germline transcript transcription (GLT) (3) AID recruitment and cytosine deamination (4) recognition and removal of uracil by UNG (5) excision of abasic sites by base excision repair (BER) and/or mismatch repair (MMR) (6) DSB repair by cNHEJ or microhomology-mediated repair (MMEJ). Schematic was created using BioRender.

Mechanistically, upon B cell activation, the *Igh* locus becomes transcribed at switch (S) regions, producing non-coding transcripts in a process called germline transcription (GLT) (Fig. 2.10b). GLT exposes ssDNA on which AID then deaminates cytosines to uracil. U:G mismatches are then recognised by the glycosylase UNG and uracils are removed. The resulting abasic sites are excised by APE endonucleaes and/or MRE11/RAD50, leading to the formation of single strand breaks (SSBs). If the SSBs are in close proximity on opposite strands, they naturally form a staggered DSB. Otherwise, the mismatch repair (MMR) pathway further processes the lesion, subsequently converting it into a DSB. In particular, DSBs are generated in the donor Sμ region and the downstream acceptor S region - which are subsequently joined in a deletional manner [75–77]. These breaks are then largely repaired by NHEJ. However, in absence of NHEJ, microhomology-mediated end-joining (MMEJ) can repair a proportion of the breaks leading to productive CSR [78, 79].

Crucially, repair of DSBs in CSR critically depends on 53BP1, with 53BP1-deficiency leading to even stronger CSR defects than loss of core NHEJ proteins [80, 81]. Seminal work over the last decade has discovered a new sub-pathway of DNA end joining - the 53BP1-RIF1-shieldin-CST axis - that seems to have co-evolved with CSR [82]. In particular, during CSR, 53BP1-RIF1 recruits the shieldin complex, compromising of SHLD1, SHLD2, SHLD3, and REV7 [83–85]. Shieldin can bind ssDNA overhangs (as found in DSBs generated during CSR) and subsequently recruits the CST (CTC1, STN1, TEN1) complex in an ATM-dependent manner [25, 86, 87]. In turn, CST, together with Polα primase facilitates fill-in synthesis of the ssDNA overhang, ultimately producing a DNA substrate conducive to repair by NHEJ.

1.4 ERCC6L2 - DNA excision repair protein ERCC6-like 2

In 2014, exome sequencing of two unrelated patients suffering from acute bone marrow failure revealed bi-allelic truncating mutations in ERCC excision repair 6 like 2 (also known as RAD26L, or HEBO (HElicase mutated in BOne marrow failure) (*ERCC6L2*) [88]. The patients suffered from tri-lineage bone marrow failure caused by hypocellular bone marrow, microcephaly, and developmental delays characterised by learning disabilities. Sanger sequencing of the patient's families revealed that, in both cases, both parents were heterozygous carriers the disease-causing variants. Accordingly, the disease is an autosomal-recessive trait, with the mutations not acquired somatically, but instead transmitted through the germline, thereby revealing *ERCC6L2* mutations as a novel cause for an inherited bone marrow failure syndrome.

Since its discovery, ERCC6L2 has been generously attributed to a plethora of cellular processes. The human phenotype associated with its dysfunction has changed markedly as well. For clarity, the current understanding of the human disease and its proposed molecular function have been split into two sections.

1.4.1 ERCC6L2 syndrome

In a comprehensive study in 2023, 53 subjects with bi-allelic germ line variants of *ERCC6L2* were identified [89]. Of those, 32 were diagnosed with BMF, and 15 with a haematological malignancy (i.e. myelodysplastic syndrome (MDS) or acute myeloid leukaemia (AML)), and 5 were asymptomatic siblings. Generally, changes in blood counts were mild, but biopsies revealed severe bone marrow hypoplasia. The median age of patients at referral to a haematologist was 18 years and almost all patients initially presented with various degrees of thrombocytopenia, but cytopenias of other blood lineages were also common.

Whilst initial reports suggested extra-haematological features including microcephaly and neurological conditions, these symptoms fail to apply to the significantly larger cohort of patients identified to date [88, 90, 91]. Instead, early cases were identified in individuals from consanguineous families (i.e the parents are second cousins or more closely related), raising the possibility that these symptoms are caused by other genetic aberrations [92].

Patients with ERCC6L2-disease have a high risk of progressing from BMF to MDS, and subsequently AML, and progression coincides with clonal expansion of loss-of-function *TP53* mutant cells [89, 93]. *TP53* mutations are likely a form of somatic restoration, "restoring" poor HSPC fitness by sacrificing quality control by p53, whilst also heightening the risk of malignant transformation. Once patients progress to AML, the prognosis is dismal, with all patients reported to date ultimately succumbing to the disease [94]. Of particular note and perhaps in connection with the poor prognosis, ERCC6L2-disease possesses another peculiar feature: patients almost exclusively progress to acute erythroid leukaemia (FAB classification: AML M6) [95].

In this thesis, the terms AML M6 and acute erythroleukaemia are used interchangeably. Whilst there are distinctions between the WHO and FAB classification systems, these are of little relevance to this piece of research. Acute erythroleukaemia is defined by $> 50\%$ predominance of erythroid precursors in the bone marrow. It is an exceedingly rare subtype of AML (1% of overall *de novo* AML) that typically occurs in older patients [96]. The most frequently mutated gene in AML M6 is *TP53*, with other leukaemic driver mutations being notably less enriched compared to other AML subtypes (although, notably, currently used sequencing panels do not contain *ERCC6L2*) [97]. Prognosis for erythroleukaemia patients is exceptionally poor, with median overall survival of patients reported to be 6 months [98].

Douglas *et al.* reported that 4/10 AML M6 cases they tested carried homozygous *ERCC6L2* mutations in comparison with 0/165 cases with AML of all other subtypes [95]. Strikingly, the

median age of diagnosis of AML M6 was much earlier in *ERCC6L2*-mutated patients compared to other AML M6 patients (49 years vs. 67 years).

In summary, germline *ERCC6L2* variants cause an inherited bone marrow failure syndrome associated with progression to a highly specific, high-risk leukaemia subtype of the erythroid lineage, characterised by *TP53* mutations. Notably, whilst the number of patients with ERCC6L2 syndrome is currently small, the disease burden is likely to be underestimated. Previously, genetic testing has been limited to paediatric patients, whereas adolescents (the median age of onset in ERCC6L2 disease) were not tested. The rapid cost reduction in sequencing technologies should democratise diagnosis and might uncover a host of previously unidentified patients with *ERCC6L2* mutations.

1.4.2 Molecular function

DNA repair pathway association

ERCC6L2's molecular function has been under considerable scrutiny since its discovery. Initial experiments quickly suggested that it has a role in DNA repair, but which exact DNA repair pathway remained under discussion [88]. Several groups found that *ERCC6L2*-depleted cells were sensitive to mitomycin C, a potent DNA crosslinker [88, 90]. However, this sensitivity was markedly milder than what could be observed in cells deficient in bona fide FA repair factors, such as FANCD2.

Additionally, it was additionally reported that *ERCC6L2*-depleted cells had an increased sensitivity to irofulven - which causes DNA adducts that depend on transcription coupled-nucleotide excision repair (TC-NER) [88, 99]. Furthermore, some degree of sequence homology between ERCC6L2 and *ERCC6*/CSB was noted. During TC-NER, CSB is recruited to sites of stalled RNA polymerase II and subsequently serves as a platform for the recruitment of further downstream repair factors, ultimately allowing for the removal of the transcription-blocking lesion

[100]. Taken together, the authors concluded that ERCC6L2's function was most likely to be in TC-NER.

At the same time, Zhang *et al.* isolated fibroblasts from a patient suffering from *ERCC6L2* syndrome and evaluated their sensitivity to a number of genotoxic agents [90]. The patient fibroblasts were found to behave identically to controls in their response to: PARP inhibition - excluding a role in HR, methyl methanesulfonate (MMS) - excluding a role in both FA repair and MMR, and, notably, UV exposure - excluding a role in both TC-NER and global genomic (GG)-NER. Instead, patient fibroblasts were found to be sensitive to phleomycin and ionizing radiation - which is indicative of a role in DSB repair. However, fibroblasts of patients with *XLF/Cernunnos* mutations were markedly more sensitive to both genotoxins.

Indeed, in 2020, three separate papers classified ERCC6L2 as a novel factor in NHEJ. Olivieri *et al.* undertook several genome-wide CRISPR screens with 27 genotoxic agents and found that ERCC6L2 clustered with other factors in NHEJ, such as LIG4, XLF, and 53BP1 [101]. In a similar approach, Liu *et al.* also found that ERCC6L2 behaved like other factors in NHEJ in several chemogenic CRISPR screens, as well as a CRISPR screen looking for modulators of CSR [102]. Meanwhile, Francica *et al.* identified ERCC6L2 as a new factor modulating the response to ionizing radiation [103].

In line with a role in NHEJ, ERCC6L2's impact on V(D)J recombination and CSR was investigated. All groups noted that *Ercc6l2*-deficient cells had a marked CSR defect [101–103]. In addition, Liu *et al.* showed that whilst *Ercc6l2* was dispensable for V(D)J recombination by itself, combined loss with *Xlf* lead to synergistic impairment in an *in vitro* V(D)J assay [102]. Intriguingly, they also found that ERCC6L2 was required for directional (i.e. deletion of the segment between breaks) recombination of AID-induced breaks during CSR, with breaks being repaired without a preference for deletion in *Ercc6l2*$^{-/-}$ B cells.

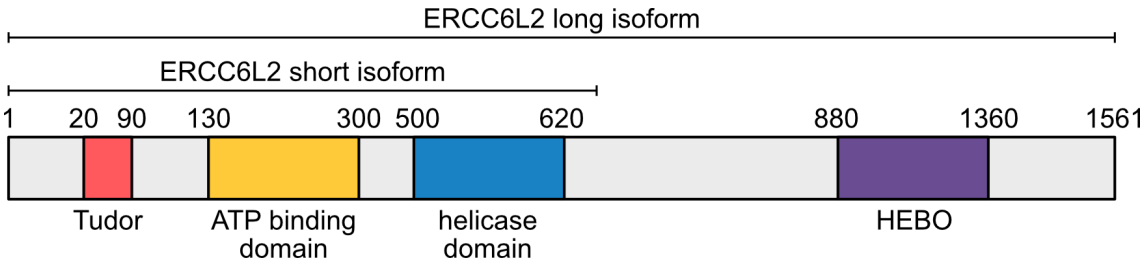


Figure 1.9: **Schematic representation of the human ERCC6L2 protein.** Predicted structured domains are highlighted in colour and their approximate amino acid positions are listed. Tudor domain - putative function: reading methyl marks, helicase domain - function: unwind DNA or translocate along DNA, HEBO domain - putative function: interaction with partner protein.

Expression and isoforms

ERCC6L2 is ubiquitously expressed, showing no tissue specificity [88, 104]. Two different isoforms have been identified - a 80 kDa short isoform, and a 171 kDa long isoform. Whilst both isoforms seem to be expressed, only the long isoform supports nuclear localisation and a function in DNA repair [90]. Indeed, due to an error in the automatic annotation of the human *ERCC6L2* gene on the Ensembl database, only the short transcript was identified at first, meaning initial complementation experiments were only done with the short isoform that does not localise to the nucleus [88].

Structure

Structurally, ERCC6L2 belongs to the Snf2 family of helicase-like proteins [105]. Snf2 family members are often the catalytic subunits of ATP-dependent chromatin remodelling complexes that regulate the structure of chromatin, thereby influencing a broad range of nuclear processes, such as transcription, replication, and DNA repair. They are motor proteins using ATP hydrolysis as an energy source to move along DNA, enabling separation of DNA from nucleosomes or proteins [106].

Close to the N-terminus, ERCC6L2 contains a Tudor domain (Fig. 1.9) [90]. Tudor domains recognize methylated lysine and arginine residues, which are often associated with modified histones marks [107]. For example, 53BP1 contains a tandem Tudor domain that recognizes histone H4 dimethylated at lysine 20 (H4K20me2), allowing 53BP1 to bind to chromatin at DSB sites [19].

However, the Tudor domain of ERCC6L2 lacks some of the aromatic residues required for binding of methyl marks, suggesting it is not used for reading chromatin states [90]. Structure/function studies have also shown that the Tudor domain is dispensable for ERCC6L2's function in CSR [102]. Following after the Tudor domain are the ATP binding and helicase domain required for its function as a DNA translocase. Importantly, experiments using *Ercc6l2*^{D270N} mice - which have a helicase catalytic-dead version of ERCC6L2 - showed that its enzymatic function is crucial for CSR, supporting a molecular mechanism involving DNA translocation [102]. The C-terminal half of the protein is predicted to be intrinsically disordered and less conserved. The only conserved domain in the C-terminus is the HEBO domain unique to ERCC6L2, that is thought to either behave as a globular domain or fold upon interaction with a partner [90].

Localisation

Ectopically overexpressed ERCC6L2 fused to GFP has been used to study its recruitment to sites of DNA damage after microirradiation. ERCC6L2 was found to be recruited with the same kinetics observed in NBS1, suggesting that it is an early responder to DNA damage [90]. In fact, in cells lacking *NBS1*, ERCC6L2 recruitment to DNA damage was abolished. This could suggest direct recruitment by NBS1, but could also be explained by the role of the MRN complex in activating ATM-signalling in the DSB response signalling cascade [8]. In a similar approach, recruitment of ERCC6L2-GFP to sites of micro-irradation was independent of KU70, H2AX, NBS1, XLF, and PARP1/2 - contradicting the previous results [102]. Instead, recruitment of XRCC4 and XLF was slightly decreased in *Ercc6l2*$^{-/-}$ cells, suggesting ERCC6L2 could facilitate ligase complex loading.

In another study, YPF-tagged ERCC6L2 was shown to form discrete nuclear foci that co-localize with CENP-A, the defining constitutent of the centromeric chromatin [108]. Loss of ERCC6L2 caused the loss of centromeric factors and changes in centromeric DNA replication.

Interacting partners

To identify interacting proteins, co-immunoprecipitates of ERCC6L2-GFP have been analysed using affinity-based mass spectrometry [99]. ERCC6L2 was found to interact with a multitude of proteins, involved in DNA repair, RNA binding, and mitochondrial function. The authors found a striking interaction with DNA-PKcs. However, due to the use of (artificially) overexpressed ERCC6L2-GFP, the improbable complexity and abundance of interacting proteins, and the low-stringency approach to the analysis (i.e. total spectral counts, not unique spectral counts were used for analysis, and protein size was not considered when interpreting spectral counts), the results should be interpreted cautiously.

More convincingly, ERCC6L2 has been reported to interact with MRI/CYREN in an IP-MS approach and a genome-wide yeast two-hybrid screen [109, 110]. This interaction - as well as an interaction with KU - was subsequently confirmed by immunoblotting of co-precipitates of endogenous, HA-tagged ERCC6L2 [102]. MRI (also called CYREN) has been reported to interact with KU and stimulate NHEJ [109, 111]. In contrast, others have reported that MRI inhibits NHEJ specifically in the S and G_2 phase of the cell cycle [112].

1.5 Aims of this thesis

This thesis aims to elucidate the molecular function of ERCC6L2 in both haematopoiesis and lymphocyte development. ERCC6L2 was initially discovered as bi-allelic mutations cause a novel inherited bone marrow failure syndrome. Nevertheless, it is currently unclear how ERCC6L2 supports HSC maintenance and function. ERCC6L2 has been implicated in NHEJ, TC-NER, and the FA pathway. Additionally, several groups have demonstrated that ERCC6L2 participates in CSR, and it has also been implicated in V(D)J recombination, but its exact function during these processes remains unknown. Specifically, I endeavoured to answer the following questions:

1. The function of ERCC6L2 in haematopoiesis. To address this, an *Ercc6l2* knockout mouse model was generated and its haematopoietic stem cell compartment was examined both immunophenotypically and functionally.

2. The function of ERCC6L2 in lymphocyte development. Using several genetic mouse models, I investigated ERCC6L2's function and hierarchy in both processes.

3. The DNA repair pathway association of ERCC6L2. With ERCC6L2 implicated in several DNA repair pathways, I re-examined its participation in these pathways using genetic mouse models and cell-based methods.

4. The interactome of ERCC6L2. Using an unbiased approach, I investigated which proteins ERCC6L2 interacts with and subsequently analysed the interactions' functional relevance.

Chapter

MATERIALS AND METHODS 2

2.1 Mice and primary mouse tissue isolation methods

2.1.1 Mouse strains

All mice used for this thesis except for recipient and competitor mice for bone marrow transplantation experiments were generated on a C57BL/6J background. All experiments involved age-matched 8-16-week-old animals on an inbred C57BL/6J background and were approved by the University of Oxford Ethical Review Committee and performed under a UK Home Office Licence in compliance with animal use guidelines and ethical regulations. Animals were house in individually ventilated cages, maintained on a 12h light-dark cycle, and provided with food and water ad libitum.

53bp1$^{-/-}$ mice (MGI: 2654201, C57BL/6J) were generated and described elsewhere [113]. *Adh5*$^{-/-}$ mice (MGI: 3033711, C57BL/6J) were generated and described elsewhere [114]. *Ercc6l2*$^{-/-}$ mice were generated by collaborator Dr. B. Davies using CRISPR/Cas9 technology detailed in Section

Target exon	Protospacer sequence
Ercc6l2 exon 2	GGTACTTGCGAGATTACCAAAGG
Ercc6l2 exon 16	ACAGCAGCTTAAAGCGTATATGG

Table 2.1: gRNA sequences used for CRISPR/Cas9-mediated knockout of *Ercc6l2*.

2.1.2. Rosa26$^{\text{CAG-Cas9-eGFP}}$ mice (MGI: 5583839, C57BL/6J) were generated and described elsewhere [115]. To achieve Cas9 expression in all tissues, mice were crossed with PGK-Cre-positive mice (MGI: 2178050, C57BL/6, JAX stock 020811) and PGK-Cre was subsequently removed during breeding. *Shld2*$^{-/-}$ mice (MGI: 5428631, C57BL/6J) were generated and described elsewhere [25].

2.1.2 Mouse strain generation

Ercc6l2$^{-/-}$ mice were generated in Dr. B. Davies' laboratory using CRISPR/Cas9 technology. gRNAs targeting exons 2 and 16 of the *Ercc6l2* gene were designed using crispor.tefor.net and selected on the basis of genome-wide specificity. gRNA sequences are listed in Table 2.1 and gRNA targeting strategy is shown in Figure 2.1. sgRNA ribonucleoprotein complexes containing both gRNAs were electroporated into C57BL/6J zygotes and a founder with compound deletions spanning exons 2 to 16 was generated (*Ercc6l2*^{em1} 52,455 bp and *Ercc6l2*^{em2} 52,601 bp, respectively.) The founder was bred with C57BL/6J mice and both alleles were transmitted to the F1 generation. Only the *Ercc6l2*^{em1} allele was used for this thesis.

2.1.3 Mouse genotyping

All mice used in this thesis were genotyped via polymerase chain reaction (PCR) amplification of the genomic region of interest. DNA was extracted from ear biopsies by lysis using alkaline lysis buffer (25 mM NaOH in de-ionized water, pH 12) for 15-30 min at 95°C. Samples were then neutralized using neutralizing buffer (40 mM Tris-HCl in de-ionized water, pH 5) and 1μL of the solution was used as the DNA template in a standard 20/muL PCR reaction using OneTaq 2X Master Mix (NEB) as per manufacturer's instruction. Primers are listed in Table 2.2. PCR

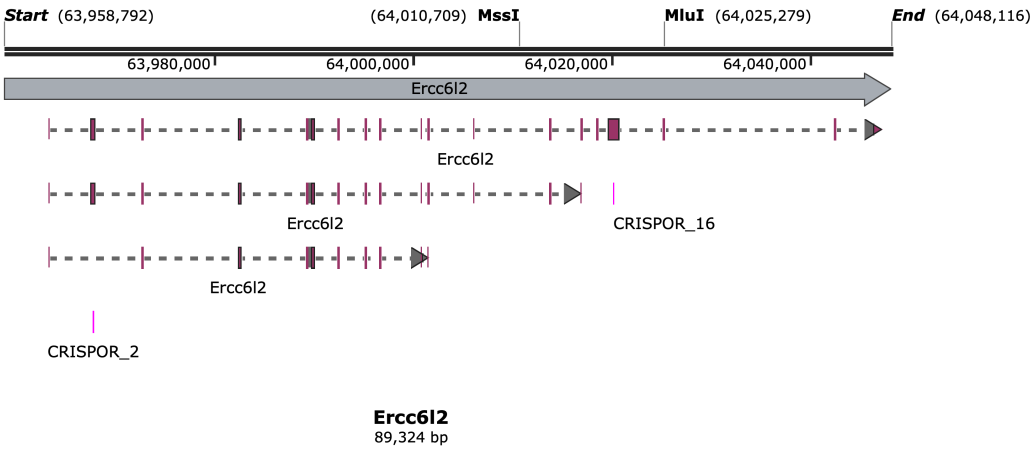


Figure 2.1: The genomic structure of the *Ercc6l2* locus. The gRNA binding sites (CRISPOR_2 and CRISPOR_16) are indicated.

products were then run on a 2% (w/v) agarose gel at 130 V for 30 min. Representative images of genotyping results are shown in Fig. 2.2.

2.1.4 Mouse embryonic fibroblast isolation

Pregnant females were sacrificed 13.5 days post confirmation of vaginal plug. Uteri containing embryos were extracted and immediately transferred to ice-cold PBS. Individual embryos were then extracted from the uterus, embryos were decapitated, heart and liver were removed, and the tail was taken for subsequent genotyping. The remaining embryo was minced and incubated in trypsin (Gibco) for 10 min at 37°C. Trypsin was then quenched by the addition of medium and the cell suspension was transferred to 10 cm dishes and cultured at 37°C, 5% CO_2, and 5% O_2 until immortalization. Cells were immortalized using SV40 large T antigen as per standard protocol.

2.1.5 *v-Abl* pre-B cell generation

v-Abl pre-B cells were generated as described in [116]. Brielfy, mice were sacrificed and one femur and tibia were dissected out. Bone marrow was collected by flushing the bones with PBS using a 23-gauge needle. The cell suspension was then counted using a Pentra ES60 Hematology Analyzer

Genotype	Primer sequence	Ann T (°C)
53bp1	GATTTTGCCTTACCCAGTTCCCGA CTACCCGGTAGAATTGACCTGCA CCTTCTATTCTACTTCAGGCTCTGG	60
Adh5	TCTCTTGATTCCCACTTTGTG AATGGACGAGTGGAGATTTC AGTCTGGAGTCACTCATTTG	60
Cas9	CGACCACAGCGACGTGGACAAGC CACGAGAGCTTTCAGGGTCA	63
Ercc6l2	TGGTGTCCGGGAGAAAGATG GTACTGAGCATAAAACAATGAGACA TGACTTGGGGAACATGTCAA	63
Rosa26	AAGGGAGCTGCAGTGGAGTA CCGAAAATCTGTGGGAAGTC	62
Shld2	GACCACTTTTCAGCCTCCAG AGAATCCTGCCCTTCTCTCC ATGGCGAGCTCAGACCATAA	62

Table 2.2: Primer sequences for locus-specific amplification. Annealing temperatures (Ann T) are listed.

(Horiba) and plated at 2×10^6cells/mL in complete *v-Abl* media with 20% FBS in 6-well plates. Cells were spinoculated with 1 mL pMSCV-v-Abl retroviral supernatant and 5μg/mL polybrene at 2000 rpm for 1 h at 30°C. Once pre-B cells were transformed, cells were further spinoculated with pMX-Bcl2-hCD271 retroviral supernatant as described above. After 2-3 days, cells were positively enriched for hCD271 expression using anti-hCD271 microbeads (130-099-023, Miltenyi Biotec) on MACS columns (Miltenyi Biotec) as per manufacturer's protocol. After several days of recovery, cells were then further spinoculated with pMX-RSS-IRES-INV/GFP-hCD4 retroviral supernatant as described above, and positively enriched with anti-hCD4 microbeads (130-097-048, Miltenyi Biotec) on MACS columns (Miltenyi Biotec) as per manufacturer's protocol.

2.1.6 Splenic B-cell isolation and stimulation

Murine splenic B cells were isolated from erythrocyte-lysed single-cell suspensions by magnetic negative selection using a B cell isolation kit (130-090-862, Miltenyi Biotec). Isolated B cells were then seeded at 3×10^5 cells/mL in RPMI-1640 medium (Sigma) supplemented with 10% FBS (Gibco), 1%Penicillin-Streptomycin (10,000U/mL, Sigma), 2mM L-Glutamine (Gibco), 1X MEM nonessential amino acids (Gibco), 1mM sodium pyruvate (Gibco), and 50μM β-mercaptoethanol

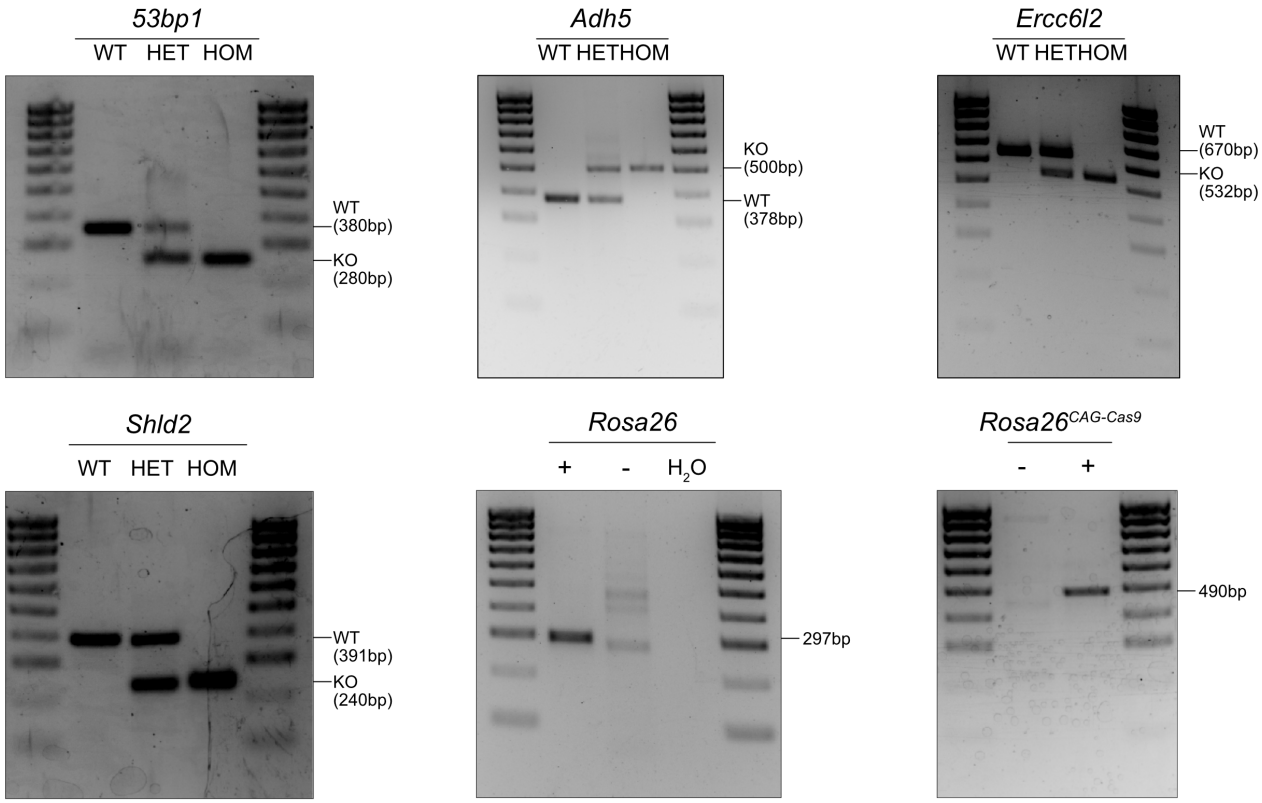


Figure 2.2: Representative image of PCR amplicons from genomic DNA from ear biopsies from mice of the indicated genotype. Identity of amplicons and their size are indicated.

at 37°C, and 5% CO_2 in a humidified atmosphere. B cells were stimulated with LPS (Sigma, 5μg/mL), mouse recombinant IL-4 (10ng/mL, PeproTech), and agonist anti-CD40 antibody (0.5μg/mL, Miltenyi Biotec). Four days after seeding, stimulated B cells were analyzed using flow cytometry.

2.1.7 HSPC/c-Kit$^+$ enrichment

HSPCs/c-Kit$^+$ cells were isolated as described in [117]. Briefly, mice were sacrificed and femurs, tibias, pelvis, and spine were dissected out. Bone marrow was collected by gently crushing the bones with a mortar and pestle and resuspended in PBS. The cell suspension was then counted using a Pentra ES60 Hematology Analyzer (Horiba) and cells were stained with 2μL/10^8 cells c-Kit-APC (17-1171-83, invitrogen) for 30 min at 4°C. Subsequently, cells were washed with ice-cold PBS and incubated with 2μL/10^8 cells anti-APC microbeads (130-090-855, Miltenyi Biotec) for 15 min at 4°C. Cells were then washed with ice-cold PBS and c-Kit$^+$ cells were positively

enriched using MACS columns (Miltenyi Biotec) as per manufacturer's protocol. Cells were then cultured as described in Section 2.2.

2.1.8 Competitive bone marrow transplantation

For mixed bone marrow chimera experiments, wild-type B6.SJL CD45.1 mice (MGI: 4819849, B6.SJL-Ptprc^{a}Pepc^{b}/BoyJ) were irradiated with two doses of 5.5 Gy spaced 4 h apart, and subsequently injected with 1×10^{6} whole bone marrow cells (approximately 1:1 mixture) from wild-type CD45.1/CD45.2 mice (B6.SJL×C57BL/6) and *Ercc6l2*$^{-/-}$ CD45.2 mice (C57BL/6). Baytril 10% antibiotic solution (BMS, University of Oxford) was added to the drinking water (1.2 mL per 250 mL sterile water) for 4 weeks after irradiation. Donor chimerism was tracked by collecting peripheral blood every 4 weeks for 16 weeks and analyzing the ratio of CD45.1/CD45.2 expression via flow cytometry (see 2.5.1 for the detailed protocol). After 16 weeks, mice were sacrificed and bone marrow donor chimerism was determined by analyzing the ratio of CD45.1/CD45.2 expression via flow cytometry (see 2.5.2 for the detailed protocol). Secondary bone marrow chimera experiments were performed by injecting 1×10^{6} pooled whole bone marrow cells from primary recipient into lethally irradiated wild-type B6.SJL C45.1 mice. Mice were treated and donor chimerism was determined as described above.

2.2 Cell lines and culture conditions

Human embryonic kidney 293T cells (HEK 293T) and mouse embryonic fibroblasts (MEF) were cultured in DMEM (Sigma) supplemented with 10% FBS (Gibco), 1% Penicillin-Streptomycin (10,000 U/mL, Sigma), and 2 mM L-Glutamine (Gibco) at 37°C and 5% CO_2 in a humidified atmosphere.

CH12-F3 cells were cultured in RPMI-1640 medium (Sigma) supplemented with 10% FBS (Gibco), 5% NCTC-109 medium (Sigma), 1% Penicillin-Streptomycin (10,000 U/mL, Sigma), 2 mM L-Glutamine (Gibco), and 50 μM β-mercaptoethanol (Gibco) at 37°C and 5% CO_2 in a humidified atmosphere.

Viral Abelson kinase (*v-Abl*) transformed pre-B cell lines were cultured in RPMI-1640 medium (Sigma) supplemented with 10% FBS (Gibco), 1% Penicillin-Streptomycin (10,000 U/mL, Sigma), 2mM L-Glutamine (Gibco), and 100 μM β-mercaptoethanol (Gibco) at 37°C and 5% CO_2 in a humidified atmosphere.

Primary murine haematopoietic stem and progenitor cells (HSPC) were cultured in Ham's F-12 medium (Gibco), 1 mg/mL PVA (MP Biomedicals), 10 mM HEPES (Gibco), 1% Insulin-Transferrin-Selenium-Ethanolamine (Gibco), 1% Penicillin-Streptomycin-Glutamine (Gibco), 100 ng/mL mouse TPO (Miltenyi Biotec), 10 ng/mL mouse SCF (Peprotech) at 37°C, 5% CO_2, and 3% O_2 in a humidified atmosphere. Complete medium changes were performed every 2-3 days.

2.3 CRISPR/Cas9 genome editing

Gene-specific gRNAs were integrated into pSpCas9(BB)-2A-GFP (pX458, Addgene #48138). Protospacer sequences are listed in Table 2.3. 2.5μg of the plasmid was then electroporated into 5×10^6 cells using a Lonza 4D-Nucleofector according to manufacturer's protocol for each cell line. 24h after electroporation, the brightest 10% GFP^+ cells were bulk-sorted using a Sony SH800 cell sorter. Cells were left to recover for 2 days, then the sorted population was seeded into 96-well plates at a density of 0.5 cells/well to isolate isogenic clones. Clones bearing bi-allelic indel mutations were identified using native PAGE resolution of PCR amplicons of the edited loci and gene disruption was subsequently confirmed via Sanger sequencing.

Gene	Protospacer sequence	Targeted exon
Ercc6l2	GGAGTCTGCGCGGGACTGCG	exon 1 (for knock-in)
Mri	AGTACACGGTCTCCGTGGCT	exon 2
Xrcc4	TTTGTTATTACACTTACTGA	exon 1

Table 2.3: Protospacer sequences of CRISPR/Cas9 genome editing. Targeted exons of each gene are listed.

For knock-in of the TwinStrep-HA tag on the N-terminus of *Ercc6l2* in CH12-F3 cells, cells were treated as above with the addition of 2.5μg plasmid containing the HDR template during nucleofection, and cells were recovered in media containing the DNA-PK inhibitor NU7441 (1μM, S2638, Cambridge Bioscience) for 24h before cell sorting to increase HDR rates.

2.4 Drug sensitivity assay

CH12-F3 or MEF cells were seeded in triplicate per drug concentration in 96-well plates at a density of 1250 or 600 cells per well, respectively. Then, etoposide or cisplatin (Merck) was added to the final indicated concentrations. Four (CH12-F3) or seven (MEF) days after cell seeding, growth medium with resazurin was added to a final concentration of 10 μg/mL resazurin. Plates were then returned to the incubator for 2-4 h or until the growth medium in the untreated control wells developed a pink colour. Relative fluorescence was measured with a BMG LABTECH CLARIOStar plate reader. After subtraction of the background fluorescence, relative survival compared to the untreated control was calculated.

2.5 Flow Cytometry

All flow cytometry experiment were performed using a LSR Fortessa (BD Biosciences), a LSR Fortessa-X20 (BD Biosciences), or an Attune NxT flow cytometer (Thermo Fisher). Analysis was

Antibody	Manufacturer	Cat #	Dilution
eFluor450-CD45.2	invitrogen	48-0454-82	1:333
PE-CD11b	invitrogen	12-0112-82	1:2000
PE-Ly6G	invitrogen	12-5931-82	1:2000
PE/Cy7-CD45.1	BioLegend	110730	1:400
APC-CD4	invitrogen	17-0042-83	1:2000
APC-CD8	invitrogen	17-0081-83	1:2000
APC-eFluor780-B220	invitrogen	47-0452-82	1:1000

Table 2.4: List of antibodies for peripheral blood analysis. Working dilutions are listed.

performed using FlowJo (version 10.10.0, BD Bioscience). FACS buffer used in all experiments was 2% bovine serum albumin (Sigma) in PBS with 0.025% NaN_3.

2.5.1 Peripheral blood chimerism

Flow cytometry was performed as described in [118]. Briefly, 50μL of peripheral blood was collected from the tail vein of recipient mice into lithium heparin-coated tubes (Sarstedt). Erythrocytes were subsequently lysed for 15 min at room temperature using ACK lysis buffer (Chem-Cruz). Lysis was repeated once, then the remaining mononuclear cells were incubated in 100μL of antibody stain described in Table 2.4 for 30 min at 4°C. Cells were twice washed in FACS buffer, resuspended in FACS buffer with propidium iodide (0.5μg/mL, Sigma) and analyzed on the flow cytometer. The gating strategy is shown in Fig. 2.3.

2.5.2 Bone marrow engraftment chimerism

Flow cytometry was performed as described in [118]. Briefly, at the end point, recipient mice were sacrificed and one femur and tibia were dissected out. Bone marrow was collected by flushing the bones with PBS using a 23-gauge needle. The cell suspension was then counted using a Pentra ES60 Hematology Analyzer (Horiba) and $5-10\times10^6$ cells were stained with 300μL of the biotin-conjugated lineage antibody stain (Table 2.5) for 30 min at 4°C, washed once with PBS, and then stained with 300μL of the bone marrow engraftment antibody stain (Table 2.6) for 90 min at 4°C, and washed once in PBS. Subsequently, cells were resuspended in FACS buffer with

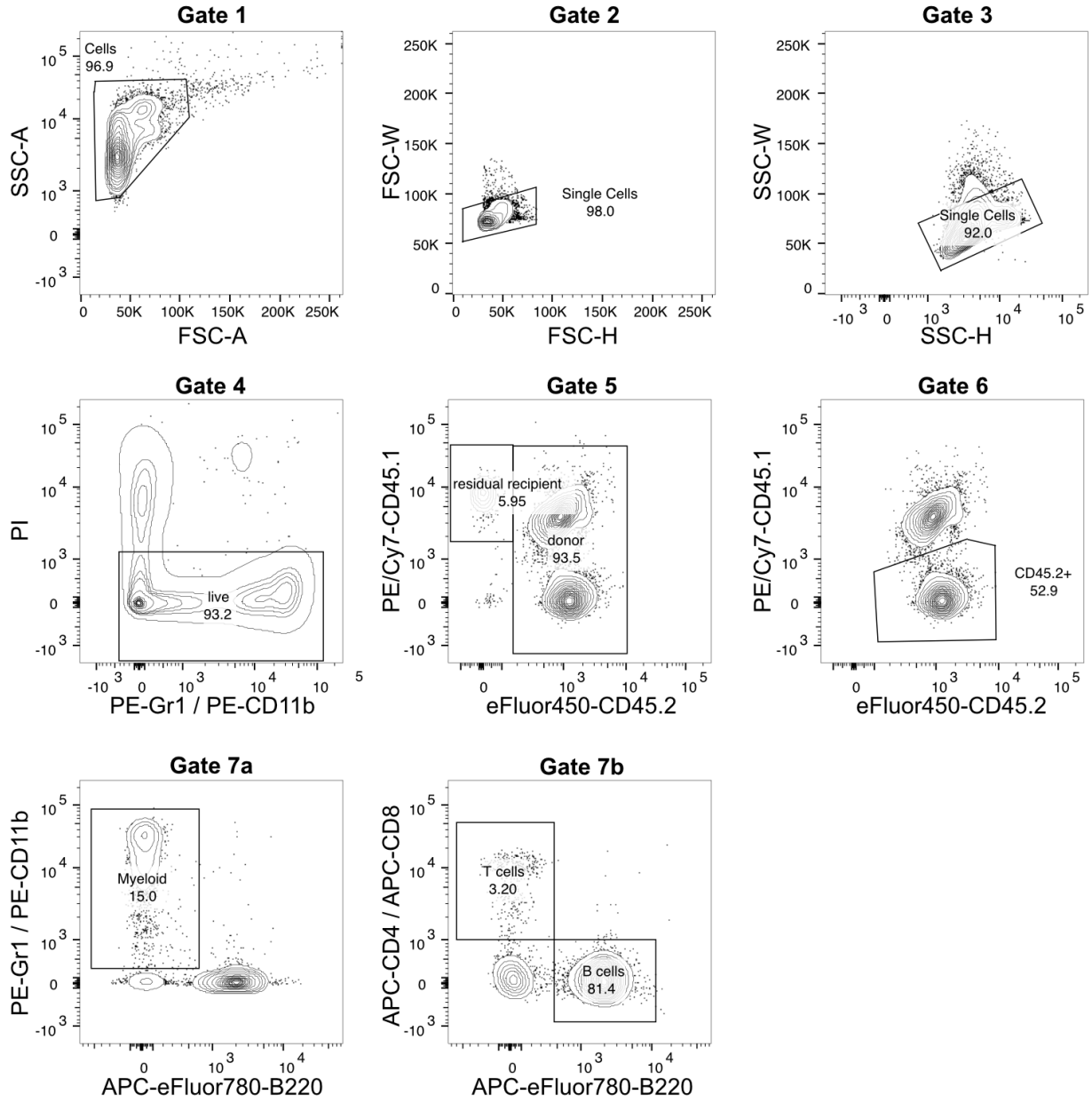


Figure 2.3: Representative gating scheme for flow cytometric analysis of peripheral blood chimerism.

propidium iodide (0.5μg/mL, Sigma) and analyzed on the flow cytometer. The gating strategy is shown in Fig. 2.4.

2.5.3 Haematopoietic stem and progenitor cells

Flow cytometry was performed as described in [118]. Briefly, mice were sacrificed and one femur and tibia were dissected out. Bone marrow was collected by flushing the bones with PBS using a 23-gauge needle. The cell suspension was then counted using a Pentra ES60 Hematology Analyzer (Horiba) and $5-10\times10^6$ cells were stained with 300 μL of the biotin-conjugated lineage antibody

Antibody	Manufacturer	Cat #	Dilution
Biotin-Ly6G	invitrogen	13-5931-85	1:700
Biotin-Ter119	invitrogen	13-5921-85	1:700
Biotin-CD4	invitrogen	13-0042-85	1:2800
Biotin-CD8	invitrogen	13-0081-86	1:2800
Biotin-B220	invitrogen	13-0452-85	1:1400
Biotin-CD127	invitrogen	13-1271-85	1:1400

Table 2.5: List of antibodies for the biotin-conjugated lineage antibody stain. Working dilutions are listed.

Antibody	Manufacturer	Cat #	Dilution
eFluor450-CD45.2	invitrogen	48-0454-82	1:100
FITC-CD34	invitrogen	11-0341-85	1:75
PE-Sca1	BioLegend	122508	1:300
PE/Cy7-CD45.1	BioLegend	110730	1:100
APC-c-Kit	invitrogen	17-1171-83	1:300
APC-eFluor780-streptavidin	invitrogen	47-4317-82	1:300

Table 2.6: List of antibodies for bone marrow engraftment chimerism analysis. Working dilutions are listed.

stain (Table 2.5) for 30 min at 4°C, washed once with PBS, and then stained with 300 μL of the bone marrow HSC antibody stain (Table 2.7) for 90 min at 4°C, and washed once in PBS. Subsequently, cells were resuspended in FACS buffer with propidium iodide (0.5 μg/mL, Sigma) and analyzed on the flow cytometer. The gating strategy is shown in Fig. 2.5.

2.5.4 HSPC culture

Flow cytometry was performed as described in [117]. Briefly, 50μL of resuspended HSPC cultures was collected, washed once in FACS buffer, and then stained with 50μL of the HSPC culture antibody stain (Table 2.8) for 30 min at 4°C. Cells were then washed once with FACS buffer,

Antibody	Manufacturer	Cat #	Dilution
BV421-CD48	BioLegend	103427	1:300
FITC-CD34	invitrogen	11-0341-85	1:75
PE-Sca1	BioLegend	122508	1:300
PE/Cy7-CD150	BioLegend	115914	1:300
APC-c-Kit	invitrogen	17-1171-83	1:300
APC-eFluor780-streptavidin	invitrogen	47-4317-82	1:300

Table 2.7: List of antibodies for bone marrow HSC analysis. Working dilutions are listed.

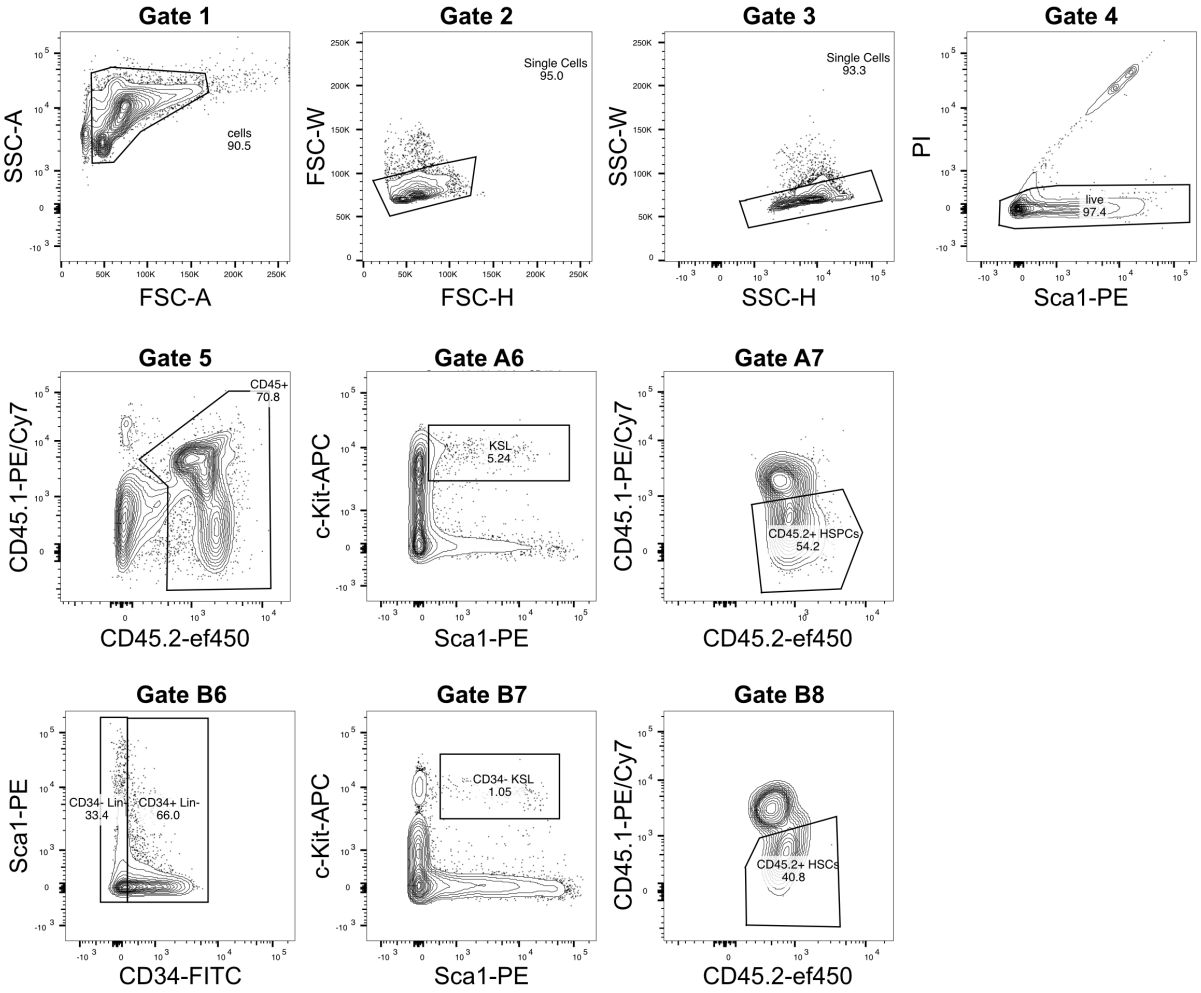


Figure 2.4: Representative gating scheme for flow cytometric analysis of bone marrow engraftment chimerism.

resuspended in FACS buffer with propidium iodide (0.5μg/mL, Sigma) and analyzed on the flow cytometer. The gating strategy is shown in Fig. 2.6.

2.5.5 B cell development - bone marrow

Mice were sacrificed and one femur and tibia were dissected out. Bone marrow was collected by flushing the bones with PBS using a 23-gauge needle. The cell suspension was then counted using a Pentra ES60 Hematology Analyzer (Horiba) and stained with 50μL of the antibody stain listed in Table 2.9 for 20 min at 4°C. Cells were washed twice in FACS buffer, then resuspended in FACS buffer and analyzed on the flow cytometer. The gating strategy is shown in Fig. 2.7.

Antibody	**Manufacturer**	**Cat #**	**Dilution**
BV421-c-Kit	BioLegend	105828	1:200
PE-Sca1	BioLegend	122508	1:200
PE/Cy7-CD150	BioLegend	115914	1:200
APC-CD201	invitrogen	17-2012-82	1:500
APC/Cy7-B220	BioLegend	103224	1:500
APC/Cy7-CD4	BioLegend	100526	1:500
APC/Cy7-CD8	BioLegend	100714	1:500
APC/Cy7-CD127	BioLegend	135040	1:500
APC/Cy7-Ly6G	BioLegend	108424	1:500
APC/Cy7-Ter119	BioLegend	116223	1:500

Table 2.8: List of antibodies for HSPC culture analysis. Working dilutions are listed.

Antibody	**Manufacturer**	**Cat #**	**Dilution**
ZombieAqua	BioLegend	423102	1:200
BV605-B220	BioLegend	103244	1:200
FITC-IgM	BioLegend	406506	1:200
PE-BP-1	BioLegend	108308	1:200
Percp/Cy5.5-CD43	BD Biosciences	562865	1:200
AlexaFluor700-CD24	BioLegend	101836	1:500
APC/Cy7-IgD	BioLegend	405716	1:500
FC-block	BD Biosciences	553141	1:500

Table 2.9: List of antibodies for B cell development analysis in the bone marrow. Working dilutions are listed.

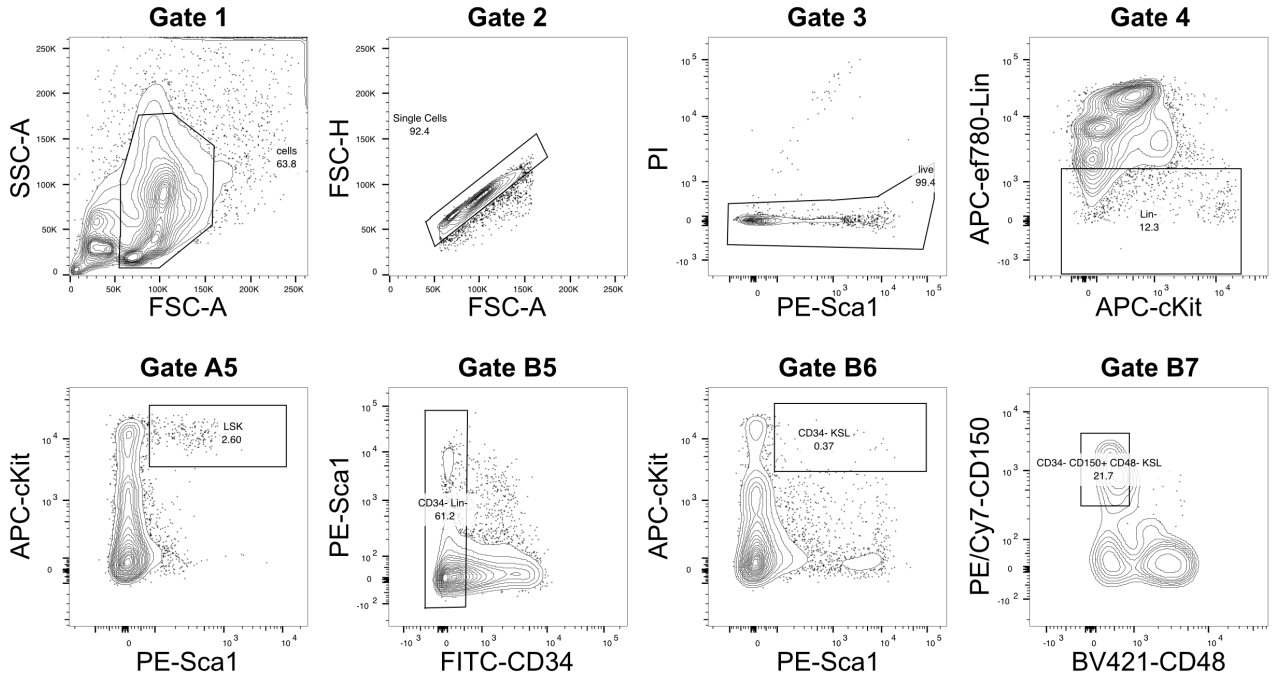


Figure 2.5: Representative gating scheme for flow cytometric analysis of HS(P)C populations in the bone marrow.

Antibody	**Manufacturer**	**Cat #**	**Dilution**
PacificBlue-IgM	invitrogen	48-5890-82	1:200
ZombieAqua	BioLegend	423102	1:200
BV605-B220	BioLegend	103244	1:200
FITC-IgD	BioLegend	406506	1:200
PE-CD23	BD Biosciences	553139	1:200
PerCP/Cy5.5-CD19	BioLegend	115534	1:500
APC-CD93	BioLegend	136510	1:100
APC/Cy7-CD21	BioLegend	123418	1:200
FC-block	BD Biosciences	553141	1:500

Table 2.10: List of antibodies for B cell development analysis in the spleen. Working dilutions are listed.

2.5.6 B cell development - spleen

Mice were sacrificed and the spleen was dissected out and mashed into a single-cell suspension. The cell suspension was then counted using a Pentra ES60 Hematology Analyzer (Horiba) and stained with 50μL of the antibody stain listed in Table 2.10 for 20 min at 4°C. Cells were washed twice in FACS buffer, then resuspended in FACS buffer and analyzed on the flow cytometer. The gating strategy is shown in Fig. 2.8.

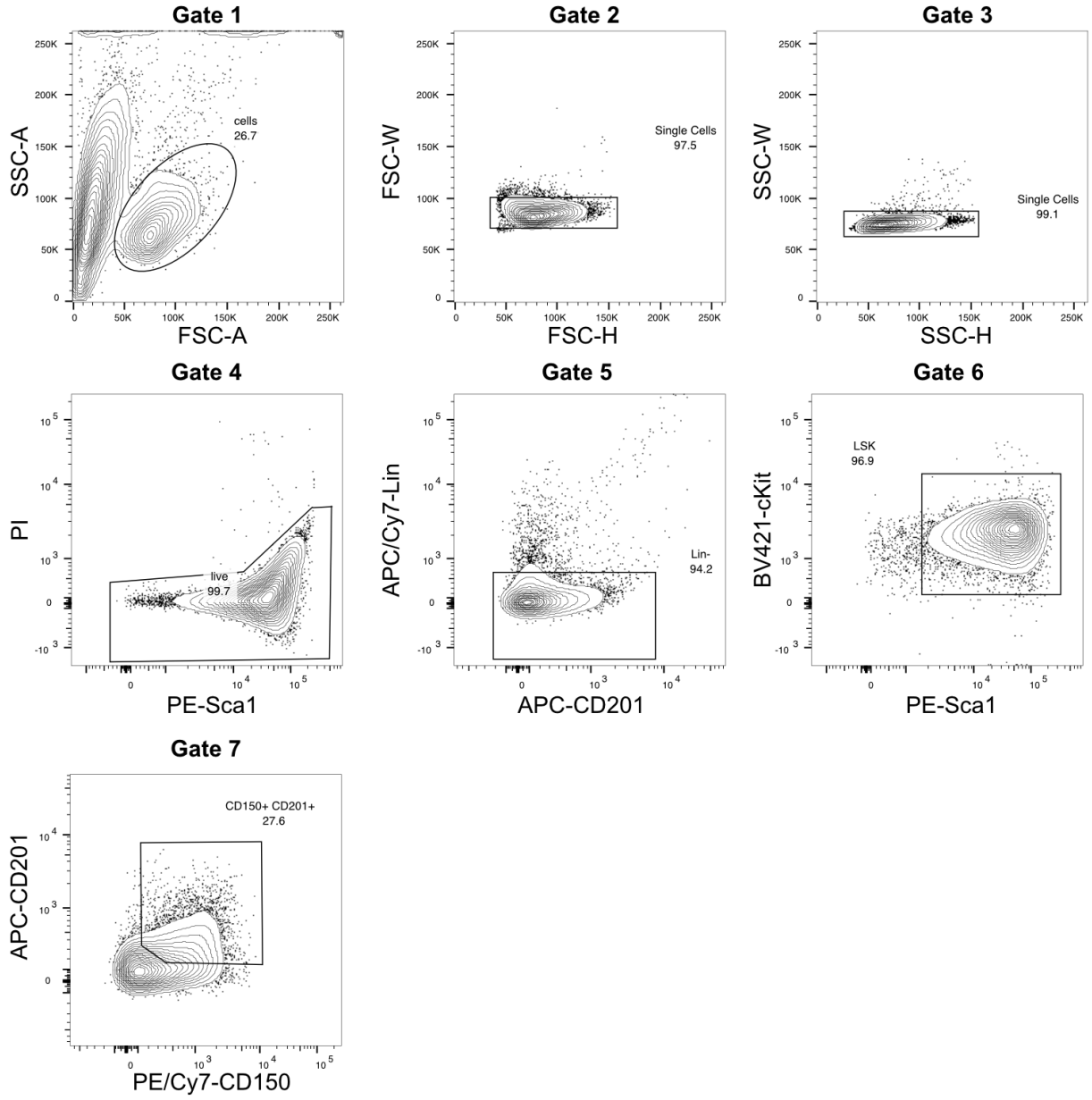


Figure 2.6: Representative gating scheme for flow cytometric analysis of HSPC culture purity.

2.5.7 T cell development

Mice were sacrificed and thymus was dissected out and mashed into a single-cell suspension. The cell suspension was then counted using a Pentra ES60 Hematology Analyzer (Horiba) and stained with 50 μL of the antibody stain listed in Table 2.11 for 20 min at 4°C. Cells were washed twice in FACS buffer, then resuspended in FACS buffer and analyzed on the flow cytometer. The gating strategy is shown in Fig. 2.9.

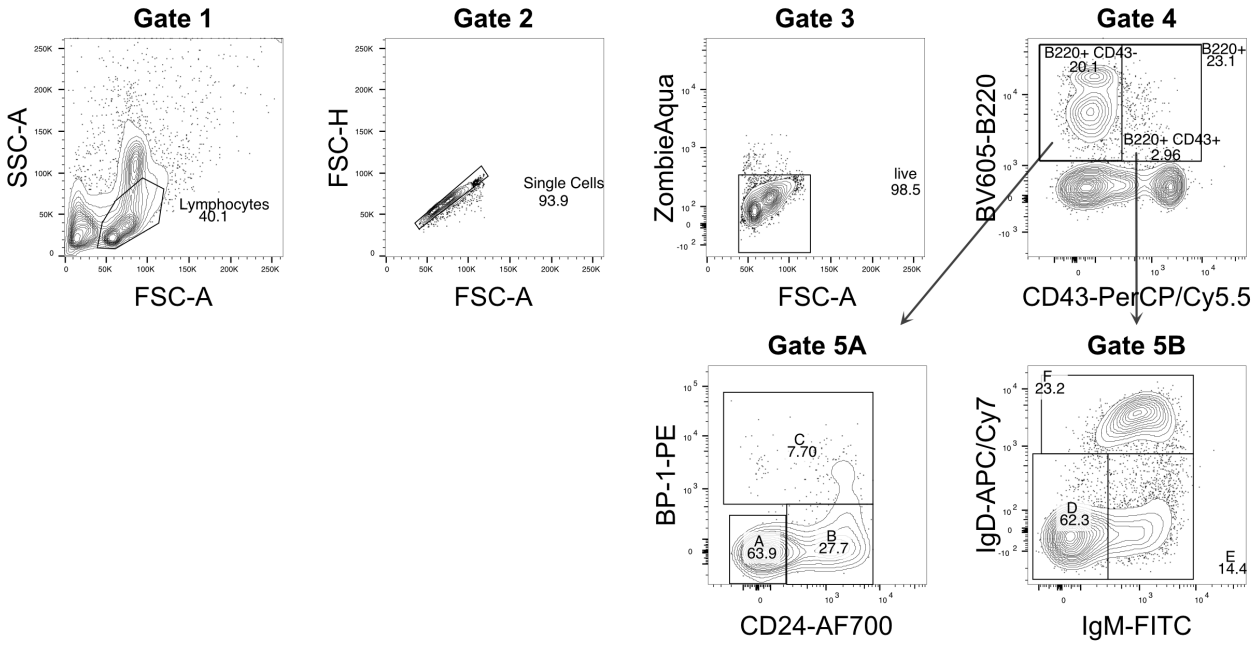


Figure 2.7: Representative gating scheme for flow cytometric analysis of B cell development in the bone marrow.

Antibody	**Manufacturer**	**Cat #**	**Dilution**
ZombieAqua	BioLegend	423102	1:200
BV605-TCR-β	BioLegend	109241	1:200
FITC-CD44	BioLegend	103022	1:200
PE-CD25	BioLegend	102008	1:500
PerCP-CD8a	BioLegend	100732	1:500
AlexaFluor700-CD4	BioLegend	100430	1:500
APC/Cy7-CD3	BioLegend	100248	1:100
FC-block	BD Biosciences	553141	1:500

Table 2.11: List of antibodies for T cell development analysis. Working dilutions are listed.

2.5.8 B cell isotype switching

B cells isolated from murine spleens and stimulated with cytokines for 96h (see Section 2.1.6) were resuspended in FACS buffer, blocked with Mouse BD FC Block (BD Biosciences), and stained with 50μL biotin-conjugated antibody cocktail listed in Table 2.12 for 20 min at 4°C. Cells were washed twice in FACS buffer, and then stained with 50μL secondary antibody cocktail listed in Table 2.13 for 20 min at 4°C. Cells were subsequently washed twice in FACS buffer, then resuspended in FACS buffer and analyzed on the flow cytometer. Cell proliferation was assessed using Cell Trace Violet (invitrogen) according to manufacturer's instructions. The gating strategy

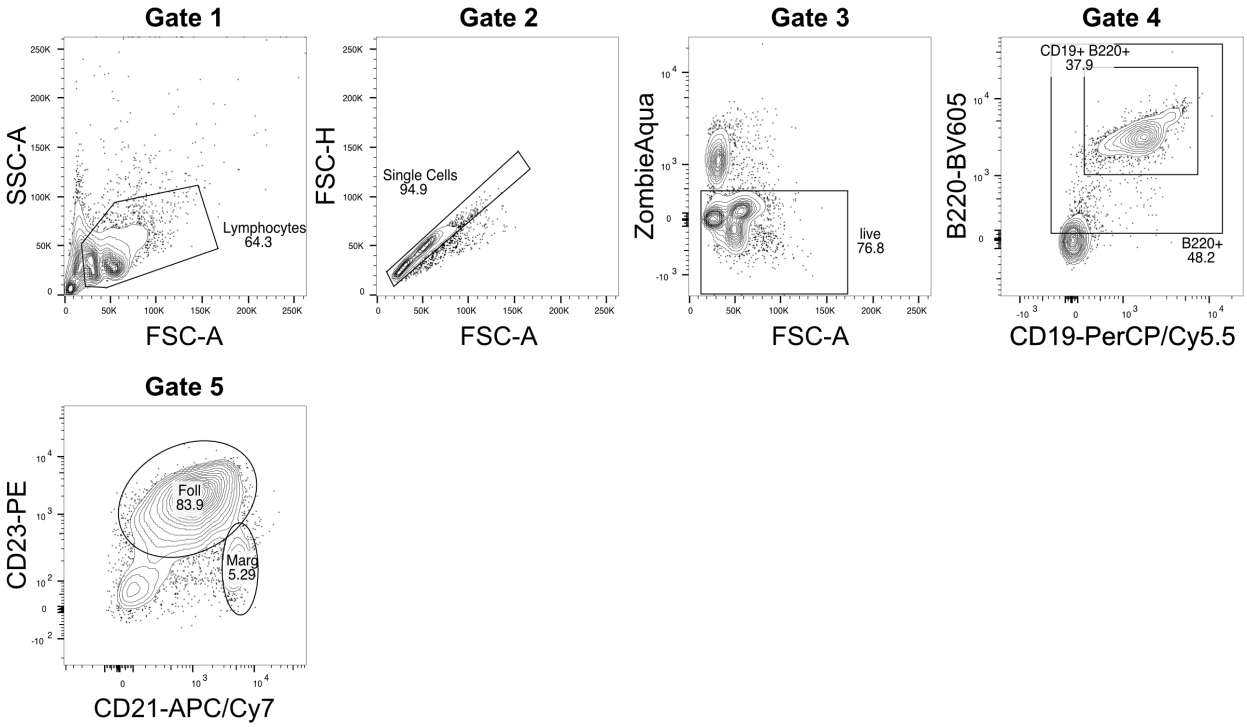


Figure 2.8: Representative gating scheme for flow cytometric analysis of B cell development in the spleen.

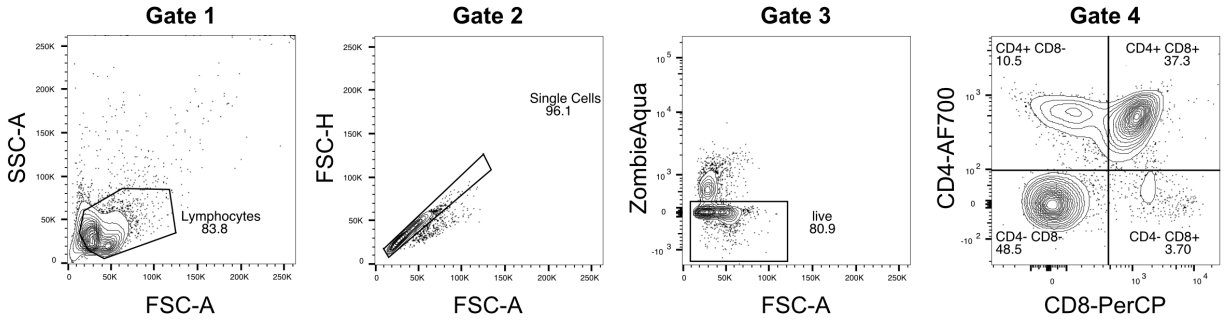


Figure 2.9: Representative gating scheme for flow cytometric analysis of T cell development in the thymus.

is shown in Fig. 2.10.

2.5.9 V(D)J recombination assay

v-Abl pre-B cells were seeded at a final concentration of 1×10^6 cells/mL in complete *v-Abl* media and treated with 3 μM imatinib (13139-25, Cambridge Bioscience) for 96h. Cells were then collected, washed once in FACS buffer, and stained with 50 μL of antibody stain containing hCD4-AF647 (1:50, 300520, BioLegend) and ZombieAqua (1:250, 423102, BioLegend) for 20 min at 4°C. Cells were twice washed with FACS buffer, then resuspended in FACS buffer and analysed

Antibody	Manufacturer	Cat #	Dilution
IgG1-biotin	BD Biosciences	553441	1:100
IgG2b-biotin	BD Biosciences	406704	1:100
IgG3-biotin	BD Biosciences	553401	1:100
FC-block	BD Biosciences	553141	1:500

Table 2.12: List of biotin-conjugated antibodies for B cell isotype switching analysis. Working dilutions are listed.

Antibody	Manufacturer	Cat #	Dilution
ZombieAqua	BioLegend	423102	1:200
Streptavidin-APC	invitrogen	17-4317-82	1:500
IgE-PE	BioLegend	406908	1:200

Table 2.13: List of secondary antibodies for B cell isotype switching analysis. Working dilutions are listed.

on the flow cytometer. The gating strategy is shown in Fig. 2.11

2.5.10 Micronucleus assay

Flow cytometry was performed as described in [119]. Briefly, 50μL of peripheral blood were collected from the tail vein into a heparin solution and immediately fixed using cold methanol. Samples were stored at -80°C until further processing. Cells were stained with FITC-CD71 (1:400, BioLegend, cat. no. 113806) for 60 min at 4°C and ultimately resuspended in FACS buffer with propidium iodide (0.5μg/mL, Sigma) and analysed on the flow cytometer. The gating strategy is shown in Fig. 2.12.

2.5.11 CSR assay in CH12-F3 cells

CH12-F3 cells were plated at 1.5×10^5 cells/mL in complete RPMI-1640 medium supplemented with anti-CD40 antibody (1μg/mL, Miltenyi Biotec), IL-4 (20ng/mL, PeproTech), and TGF-β (10ng/mL, PeproTech) to stimulate isotype switching from IgM to IgA. 48h after stimulation, cells were stained with IgA-PE (1040-09, SouthernBiotech, 1:250) and ultimately resuspended in FACS buffer and analyzed on the flow cytometer.The gating strategy is shown in Fig. 2.13.

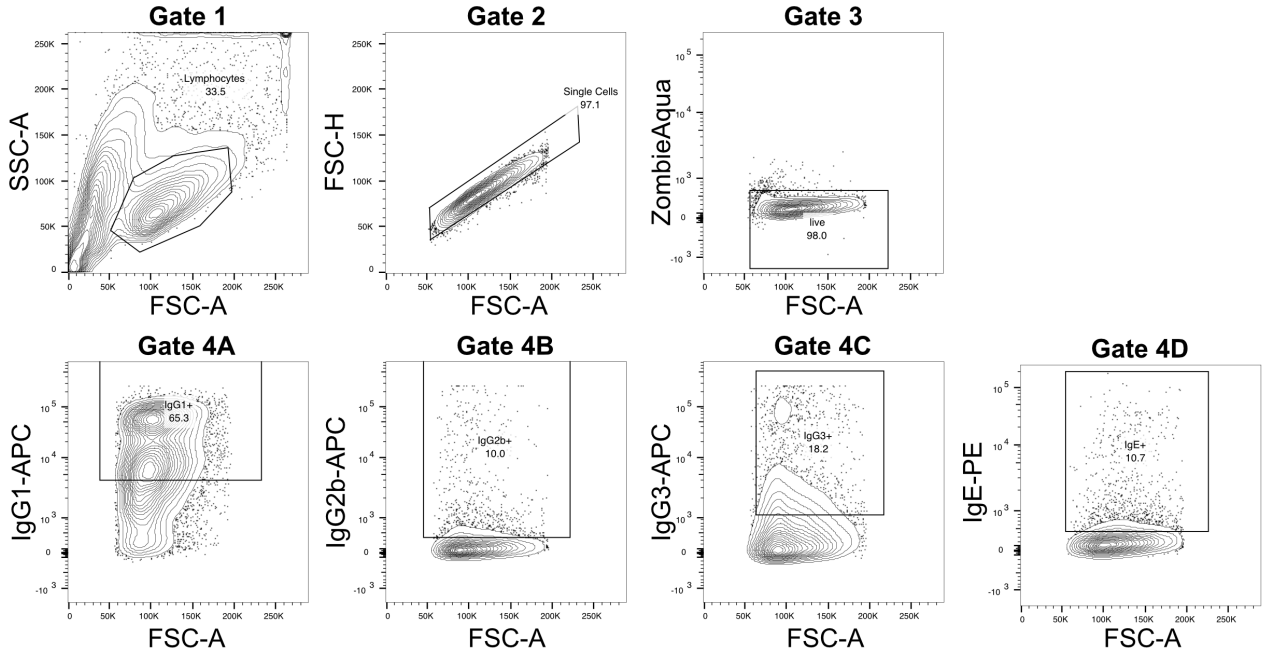


Figure 2.10: Representative gating scheme for flow cytometric analysis of B cell isotype switching.

2.6 Proteomics and Mass Spectrometry

2.6.1 Cell lysate preparation and complex immunoprecipiation

Cell pellets were collected from cultures of 1.4×10^8 CH12-F3 cells ±10 Gy irradiation and 60 min recovery. Cells were lysed in BLB (benzonase lysis buffer: 100 mM NEM, 40 mM KCl, 2 mM $MgCl_2$, 10% glycerol, 1% NP-40, 50 U/mL benzonase (Millipore), 1X cOmplete protease inhibitor cocktail (Roche), 0.5% (v/v) protease inhibitor cocktail 2 and 3 (Sigma-Aldrich) and incubated for 10 min on ice. The salt concentration was then adjusted to 450 mM KCl and lysates were further incubated on wheel for 30 min at 4°C. Subsequently, extracts were clarified by centrifugation at 17,000 G for 15 min and then diluted in NSB (no-salt buffer: 100 mM NEM, 20mM HEPES, pH 7.9, 0.5 mM DTT, 0.5 mM EDTA, 10% glycerol, 1X cOmplete protease inhibitor cocktail (Roche), and 0.5% (v/v) protease inhibitor cocktail 2 and 3 (Sigma-Aldrich) to reach a final salt concentration of 125 mM KCl. TwinStrep-HA-ERCC6L2 complexes were incubated with MagStrep-XT magnetic beads (2-5090-002, iba) for 4 h at 4°C, subsequently

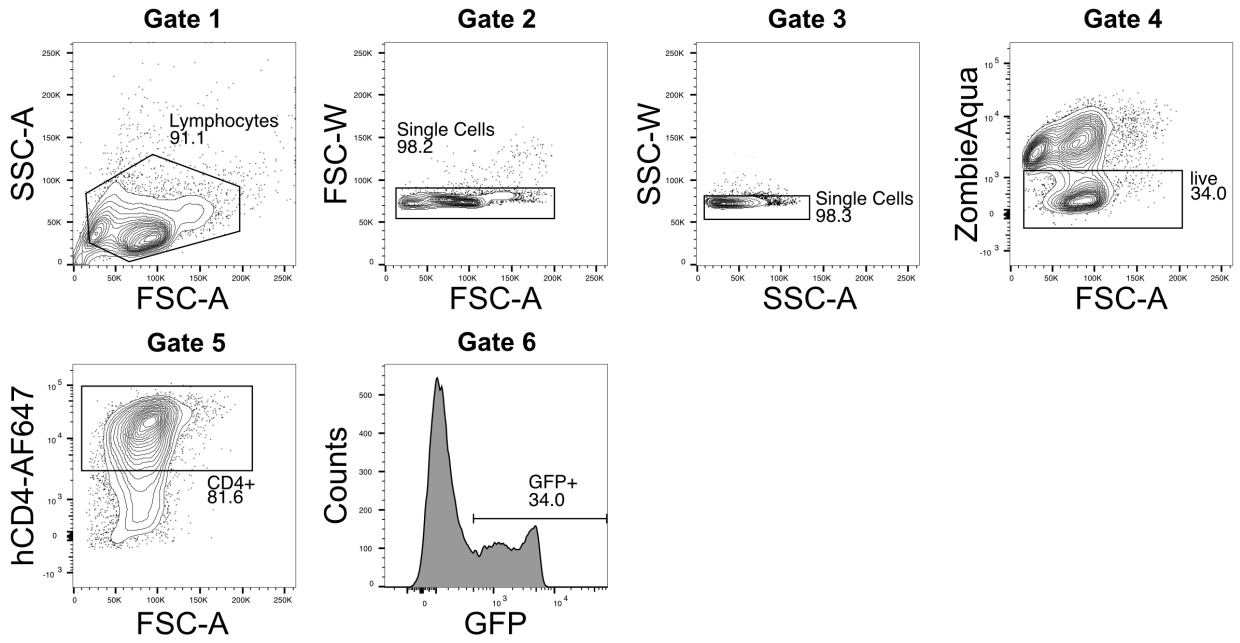


Figure 2.11: Representative gating scheme for flow cytometric analysis of V(D)J recombination in *v-Abl* pre-B cells.

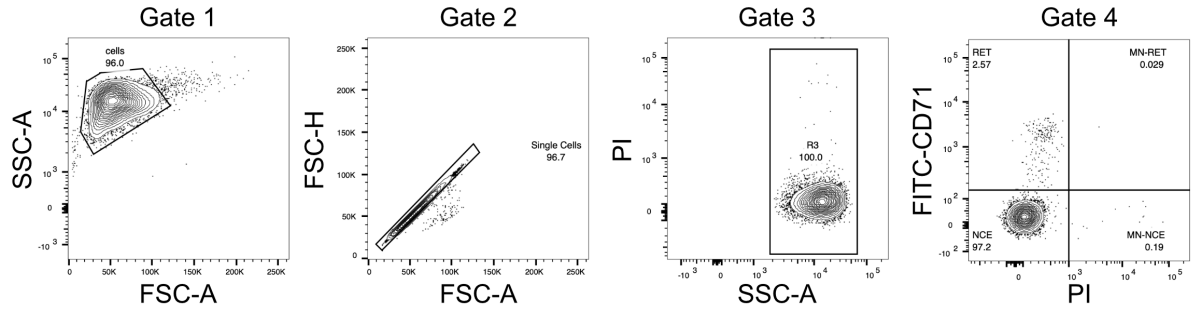


Figure 2.12: Representative gating scheme for flow cytometric analysis of micronucleus levels in mature erythrocytes.

washed in wash buffer (100 mM KCl, 20 mM HEPES, pH 7.9, 0.5 mM DTT, 10% glycerol, 0.05% NP-40, 1X cOmplete protease inhibitor cocktail (Roche), and finally eluted in 1X BXT buffer (iba) supplemented with 1X cOmplete protease inhibitor cocktail (Roche). An aliquot was mixed with Laemmli buffer and boiled at 95°C for 5 min for validation using Western blot. The remaining elutions were snap-frozen for further downstream processing for proteomic analysis.

3XFLAG-CYREN complexes were incubated with anti-FLAG-M2 magnetic beads (M8823, Merck) for 2 h at 4°C, subsequently washed in wash buffer (100 mM KCl, 20 mM HEPES, pH 7.9, 0.5 mM DTT, 10% glycerol, 0.05% NP-40, 1X cOmplete protease inhibitor cocktail (Roche), and

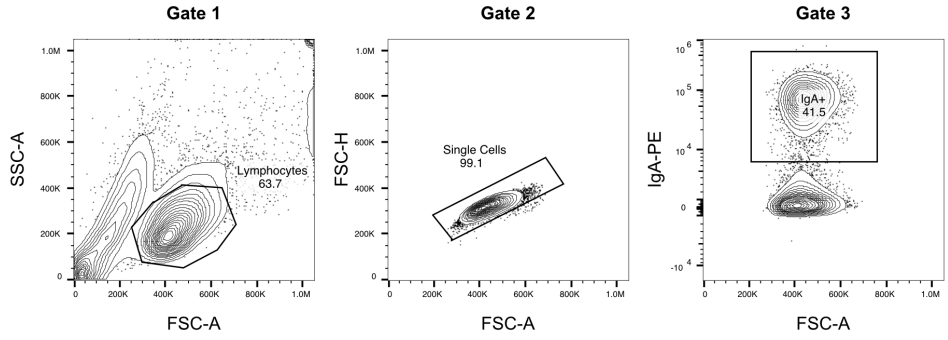


Figure 2.13: Representative gating scheme for flow cytometric analysis of isotype switching in CH12-F3 cells.

finally eluted in FLAG elution buffer (300ng/muL FLAG-peptide, 100 mM TrisHCl, pH 8, 150 mM NaCl, 1 mM EDTA) supplemented with 1X cOmplete protease inhibitor cocktail (Roche) for 30min at 4°C. The elution step was repeated once. Elutions were pooled, mixed with Laemmli buffer, and boiled at 95°C for 5 min subsequent analysis using Western blot.

2.6.2 sTRAP digestion and proteomic analysis

All subsequent steps were performed and described by the BSRC Mass Spectrometry and Proteomics Facility of the University of St Andrews. TwinStrep-HA-ERCC6L2 elution complexes were solubilized in 4% (w/v) SDS and 50 mM TEAB pH 8.5, reduced and alkylated using DTT and iodoacetamide and digested with trypsin on S-Trap microcolumns as per the manufacturer's instructions.

Peptides were subjected to LC-MS^2 using an Ultimate 3000 RSLC coupled to an Orbitrap Fusion Lumos Tribrid mass spectrometer with Easyspray source and FAIMS interface (Thermo Fisher Scientific). Peptides were injected onto a reverse-phase trap (Pepmap100 C18 100μm x 2cm) for pre-concentration with loading buffer (100% water, 0.05% TFA), at 5 μL/min for 10 min. The peptide trap was then switched in-line with the analytical column (Easy-spray Pepmap RSLC C18 2μm, 15cm x75μm ID). Peptides were eluted from the column using a linear solvent gradient at a flow rate of 300nL/min with the following gradient: linear 2-35% of buffer B (Mobile phase

B: 80% acetonitrile, 20% water, 0.1% formic acid, Mobile phase A: 100% water, 0.1% formic acid) over 44 min, sharp increase to 95% buffer B within 0.5 min, isocratic 95% of buffer B for 4.5 min, sharp decrease to 2% buffer B within 1.1 min and isocratic 2% buffer B for 5 min. The FAIMS interface was set to -45V and -65V at standard resolution. The mass spectrometer was operated in DDA positive ion mode with a cycle time of 1.5sec. The Orbitrap was selected as the MS1 detector at a resolution of 60000 with a scan range of m/z 375 to 1500. RF-lens was set to 30%. Peptides with charge states 2 to 5 at a minimum intensity threshold of 5.0×10^3 were selected for fragmentation in the ion trap using HCD as collision energy. Once selected, a dynamic exclusion for 45 sec with a 10 ppm mass tolerance was applied. A dependent scan was only performed on the single charge state per precursor.

The Ion Trap was selected for data dependent MS2 fragmentation with an isolation window of m/z 1.2 and no isolation offset. The peptides were fragmented with fixed HCD as activation type at 28% HCD Collision Energy. The Ion Trap scan rate was set to Rapid with a maximum injection time of 45 ms.

The raw data files were converted into mgf using MSconvert (ProteoWizard) and searched using Mascot (version 2.8.1, Matrix Science) with trypsin as the cleavage enzyme, carbamidomethylation on cysteines as fixed modification and oxidation of methionines as a variable modification against the Uniprot mouse database (downloaded on 28Oct2024, containing 54747 protein sequences). The mass accuracy for the MS scan was set to 20 ppm and for the fragment ion mass to 0.6Da.

Scaffold (version 5.3.3, Proteome Sofwarte Inc.) was used to validate MS^2-based peptide and protein identification. Peptide identifications were accepted if they could be established at > 95.0% probablility by the Peptide Prophet algorithm [120] with Scaffold delta-mass correction. Protein identifications were accepted if they could be established at > 99.0% probability and contained at least 1 identified peptide. Protein probabilities were assigned by the Protein Prophet algorithm [121].

Antibody	Manufacturer	Cat #	Dilution
α-TUBULIN	Sigma-Aldrich	00020911	1:10,000
CHK2	Sigma-Aldrich	05-649	1:1000
FLAG-M2	Sigma-Aldrich	F1804	1:2000
HA-tag	Cell Signaling	3724	1:1000
KU70	Cell Signaling	4588	1:1000

Table 2.14: List of primary antibodies used for Western blotting. Working dilutions are listed.

Antibody	Manufacturer	Cat #	Dilution
Goat anti-mouse IgG-HRP	invitrogen	31432	1:50,000
Goat anti-rabbit IgG-HRP	invitrogen	31462	1:50,000
Mouse TrueBlot ULTRA: rat anti-mouse Ig-HRP	Rockland	18-8817-30	1:5000

Table 2.15: List of secondary antibodies used for Western blotting. Working dilutions are listed.

2.7 SDS-PAGE-Western Blot

Protein lysates were mixed with Laemmli buffer and boiled at 95°C for 5 min before loading on NuPAGE 4-12% Bis-Tris 1.0 mm polyacrylamide gels (invitrogen). SDS-PAGE gels were run at 95V for 15min, followed by 150V for 60min, before transferring to 0.45-μm nitrocellulose membranes (GE Healthcare) at 50 mA overnight. After transfer, membranes were blocked for 30 min with 5% milk in PBS-T and then incubated with primary antibodies listed in Table 2.14 in 5% BSA in PBS-T with 0.025% NaN_3 overnight. On the next day, membranes were incubated with secondary antibodies conjugated to horseradish peroxidase (HRP) listed in Table 2.15 for at least 1 h at room temperature. Membranes were developed using Clarity Western enhanced chemiluminescence substrate (Bio-Rad) or SuperSignal West Pico Plus (Thermo Fisher) and imaged using a ChemiDoc MP (Bio-Rad). Images analyzed using Image Lab (Bio-Rad).

2.8 AlphaFold modelling

All AlphaFold modelling was performed by Dr. Yueru Sun. The structure of the ERCC6L2-MRI-KU complex was modelled using AlphaFold3 [122]. The sequences of *homo sapiens* ERCC6L2, MRI, KU70, annd KU80, along with DNA substrates consisting of two 30-nucleotide complementary strands with various length of single strand overhangs, were used as input into the AlphaFold 3 web server (`https://alphafoldserver.com/`) for protein complex prediction. The resulting structural models were assessed based on the per-residue confidence metric (pLDDT), and the model with the highest pLDDT score was selected for further analysis. Subsequent structural analysis was carried out at regions with high confidence (pLDDT>70). As for domain interface evaluation, only high confidence regions with accurate relative domain position (Expected Position Error <15) were considered.

2.9 Reverse-transcription qPCR

RNA was extracted from mature resting B cells isolated from murine spleens via a RNeasy Mini Kit (Qiagen) with an on-column DNase digestion step as per the manufacturer's protocol. RNA was then reverse-transcribed using 1μg of total RNA per sample via the iScript cDNA Synthesis Kit (Bio- Rad). Quantitative PCR reactions were then assembled in 96-well plates using QuantiFast SYBER Green Master Mix (Qiagen) and specific primer pairs (Table 2.16). qPCR was then performed and analyzed in a StepOnePlus Instrument (Applied Biosystems) using the *Quantitative - Comparative* C_T *(*$\Delta\Delta C_T$*)* protocol. Gene expression between samples was normalized using hypoxanthine-guanine phosphoribosyl-transferase 1 (*Hprt1*) as a housekeeping gene. All values were presented as fold-change compared to appropriate controls.

Target	Primer sequence
Ercc6l2 exon 2-3	AGGGGGCTCAGTTCCTCTAC GCAAAACTGCAGCCAGAAAT
Ercc6l2 exon 13-14	ACAGCCACAACATGGTTGAA CGCTGTAGAGCTCACAAACATC
Hprt1	CTGGTGAAAAGGACCTCTCG TGAAGTACTCATTATAGTCAAGGGCA

Table 2.16: Primer sequences for locus-specific amplification of cDNA for RT-qPCR.

2.10 Genome-wide CRISPR/Cas9 knockout screen in primary HSPCs

2.10.1 Titration of sgRNA library lentivirus

The genome-wide Mouse CRISPR Knockout Pooled Library (Brie; Addgene pooled library #73633) lentivirus was provided by the Vector Core of the University of Michigan and contains 4 gRNAs per gene. The virus was titered to determine the infectivity of the virus in primary murine HSPCs. Cells were plated at a cell density of 2×10^6 cells/mL in a fibronectin-coated 24-well plate (Corning). One well was not transduced. The virus was diluted 1:25, 1:50, 1:100, 1:200, 1:400, 1:800, 1:1600, and 1:3200 in complete HSPC media. Viral media was removed 6h post-transduction, cells were re-plated in a 24-well CellBind plate (Corning) at a density of 1×10^6 cells/mL and cells were collected 48h post-transduction. DNA was extracted using the PureLink Genomic DNA Mini Kit (Thermo Fisher) and ddPCR (BioRad) was performed as per manufacturer's recommendations using ddPCR Supermix for Probes (Bio-Rad). MOI was determined as $2 * \frac{copies[puromycin\,resistance]}{copies[Tfrc]}$. Primer sequences for ddPCR are listed in Table 2.17.

2.10.2 Cell culture

Wild-type and *Ercc6l2*$^{-/-}$ Rosa26$^{\text{CAG-Cas9/CAG-Cas9}}$ HSPCs were isolated as described in Section 2.1.7. The HSPCs of 4-5 mice per genotype were pooled and seeded at 3×10^5 cells/mL in a

Target	Primer sequence	Fluorophore
Tfrc	AGCAGCGGAAGTGGCTGGTCAGCT AGTTCCTCAGCTCAACCAAA TTGAACTGGTTCAGATCCTTCA	HEX
Puromycin resistance casette	CACCGGCCCAAGGAGCCCG AGCAACAGATGGAAGGCCTC GACGGTGGCCAGGAACCA	FAM

Table 2.17: Primer sequences for locus-specific amplification of DNA for ddPCR.

24-well CellBind plates (Corning). Cells were cultured as described in Section 2.2 and culture purity was assessed once a week as described in Section 2.5.4. Cells were grown for 14 days and then expanded by splitting 1:2 (i.e. one 24-well plate was split into 2 plates). 21 days after cell seeding, cells were transduced with the genome-wide Mouse CRISPR Knockout Pooled Library (Brie; Addgene pooled library #73633) lentivirus aiming for an MOI of 0.3 to avoid multiple viral insertions per cell. Cells were plated at a cell density of 2×10^6 cells/mL in fibronectin-coated 24-well plates (Corning) and virus was added at a ratio of 1:150 in complete HSPC media. Viral media was removed 6h post-transduction and cells were re-plated in 24-well CellBind plates (Corning) at a density of 1×10^6cells/mL. MOI was determined 48 h post-transduction as described in Section 2.10.1. 48h post-transduction, cells were also treated with 1μg/mL puromycin (Gibco). Puromycin selection was maintained for 48 h. 7 days post-transduction, cells were pooled per genotype and a cell pellet was frozen for subsequent genomic DNA extraction, taking care to leave enough cells growing to maintain a 250X coverage. gDNA for 3 more time points spaced 7 days (i.e. T0, T7, T14, T21) apart were collected as described above.

2.10.3 NGS sample preparation and sequencing

Genomic DNA (gDNA) was harvested using the QIAamp DNA Blood Maxi kit (Qiagen). Briefly, cell pellets were resuspended in PBS and lysed with Proteinase K and RNase A at 70°C for 15 min. Samples were further processed as per manufacturer's protocol. Finally, DNA was eluted in AE buffer, with the eluate reloaded on the column once and DNA concentrations were measured via NanoDrop.

Cycles	Temperature	Time
	95°C	1 min
14X	95°C 72°C (-1°C/cycle) 72°C	30 s 30 s 30 s
20X	95°C 58°C 72°C	30 s 30 s 30 s
	72°C	10 min
	4°C	∞

Table 2.18: Thermocycler conditions for amplification of lentiviral gRNA integrations.

Name	Primer sequence
P5	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTC TTCCGTCCGATCTTAAGTAGAGGCTTTATATATCTTGTGGAAAGGACGA AACACC
P7	CAAGCAGAAGACGGCATACGAGAT*NNNNNNNN*GTGACTGGAGTTCAGA CGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT

Table 2.19: Primer sequences for amplification of lentiviral gRNA integrations. $[N]_8$ is the unique barcode sequence for each individual sample.

To enrich the gRNA regions in the genome and amplify gRNAs with sequence adapters a one-step PCR was performed. Briefly, gRNAs were amplified using the NEBNext Ultra II Q5 master mix. The reactions were split up into individual 50 μL reactions with 2.5 μg DNA per reaction. Thermocycler conditions are listed in Table 2.18 and primers are listed in Table 2.19. The PCR reactions were then pooled for each sample and purified using a PCR purification kit (Qiagen). DNA concentrations were determined using the KAPA library quantification kit (Roche) as per manufacturer's recommendations and samples were pooled to achieve a coverage of at least 200X coverage per sample. Sequencing was performed by GENEWIZ (Azenta) on a NovaSeq 6000 (Illumina) using 150 bp paired-end reads with a 30% PhiX spike-in to increase sequence diversity on the flow cell. Sequencing data was then demultiplexed using bcl2fastq (Illumina) and only forward reads were used for further analysis.

2.10.4 Computational analysis

Sequencing data was analyzed using MAGeCK (v.0.5.9.5.) with default parameters and options

[123]. gRNAs counts were determined by matching reads to the Brie library reference for each sample. Essentiality scores (Bayes Factor (BF)) were then calculated on a gene-level-phenotype using the Bayesian Analysis of Gene Essentiality 2 (BAGEL2) algorithm using the list of core-essential (CEGv2) and non-essential (NEG) gene sets for calibration, and with each time point being compared to the gRNA counts of the plasmid used for lentivirus generation [124].

Chapter

DEFINING THE ROLE OF ERCC6L2 IN HAEMATOPOIESIS

3

3.1 Introduction

Whilst ERCC6L2 was originally discovered due to the inherited bone marrow failure syndrome caused by germline mutations, its role in haematopoiesis remains relatively unexplored. Previously generated mouse models focused entirely on the effects of ERCC6L2 on lymphocyte development, leaving the disease-causing haematopoietic stem and progenitor compartment uncharacterised [102].

In this chapter, I validated the *Ercc6l2*$^{-/-}$ mouse model generated for this thesis, including the characterization of its HS(P)C compartment, and its functionality in primary and secondary bone marrow transplants. Using the "two-tier model of aldehyde protection", I show that stem cell attrition and bone marrow failure in ERCC6L2 syndrome are unlikely a consequence of the action of endogenous aldehydes, suggesting that the disease has an aetiology unrelated to Fanconi anaemia. Finally, I performed a genome-wide CRISPR screen in primary murine HSPCs cultured

ex vivo, to investigate potential disease-relevant genetic modifiers of *Ercc6l2*-deficiency for the first time in a disease-relevant cell type.

3.2 Results

3.2.1 *Ercc6l2* knockout mice have normal numbers of HS(P)Cs

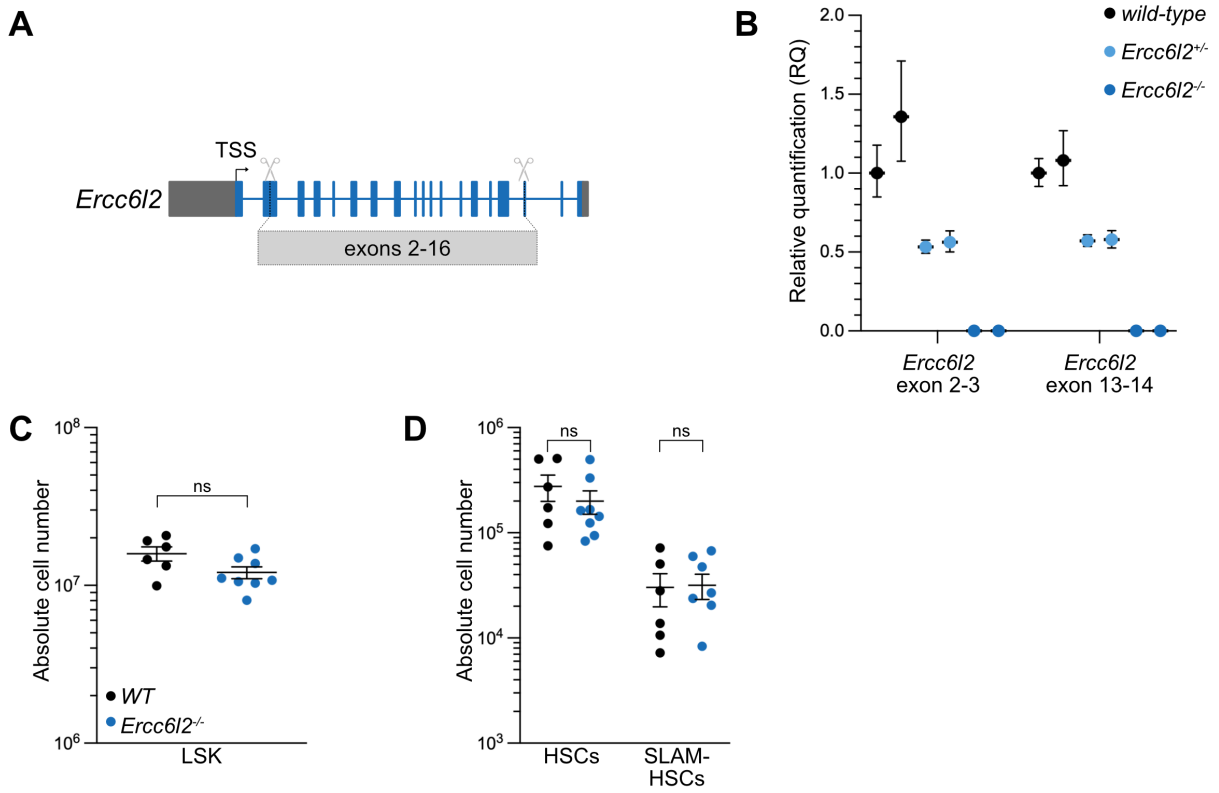


Figure 3.1: ***Ercc6l2*$^{-/-}$ mice have normal numbers of HS(P)Cs.** **A** Schematic representation of the *Ercc6l2* (knockout) allele generated for this thesis. TSS, transcription start site; grey, non-coding exons; blue, coding exons. **B** RT-qPCR result of Ercc6l2 transcript expression in splenic B cells isolated from mice of the indicated genotypes. Gene expression was normalized to *Hprt1* (n=2 per genotype, where each data point represents one mouse). **C** Absolute number of Lin$^-$ Sca1$^+$ c-Kit$^+$ (LSK) cells in the bone marrow (one femur and one tibia) (n= 6-8 mice per genotype, where each data point is a single mouse). Significance was determined by an unpaired two-tailed t-test (mean ± SEM). ns, not significant **D** Absolute number of CD34$^-$ LSK cells (HSCs) and CD34$^-$ CD48$^-$ CD150$^+$ LSKs (SLAM-HSCs) in the bone marrow (one femur and one tibia) (n= 6-8 mice per genotype, where each data point is a single mouse). Significance was determined by an unpaired two-tailed t-test (mean ± SEM).

In order to define the role of ERCC6L2 in haematopoiesis, an *Ercc6l2* knockout mouse model was generated by the laboratory of Dr. Ben Davies in an inbred C57BL/6J background (see

Section 2.1.2). Knockout mice were deleted for exons 2-16 of the *Ercc6l2* gene, which encompass sequences encoding its ATP-binding and ATPase domains (Fig. 3.1a). Due to lack of a functional antibody, knockout was validated via RT-qPCR of *Ercc6l2* transcripts in isolated splenic B cells (Fig. 3.1b). Reassuringly, no *Ercc6l2* transcripts could be detected in B cells of *Ercc6l2*$^{-/-}$ animals, whereas *Ercc6l2* transcripts were readily detectable in wild-type splenic B cells. This experiment also revealed that heterozygous animals have a 50% reduction of the level of *Ercc6l2* transcripts, suggesting that both alleles support optimal gene expression.

The production of blood depends on a small pool of long-term HSCs giving rise to a transient pool of progenitor cells (HSPCs) that then give rise to progenitors committed to specific lineages. In BMF syndromes, HS(P)C pools become depleted, ultimately leading to hypocellular bone marrow and multilineage cytopenias. I therefore quantified the frequency of HSPCs (i.e. LSK cells, Lin^{-} Sca1^{+} cKit^{+}) in *Ercc6l2*$^{-/-}$ mice. No significant reduction in the number of LSK cells compared to wild-type litter mates could be observed (Fig. 3.1c). Additionally, no significant reduction in HSC (CD34^{-} LSKs) or long-term HSC (CD34^{-}, CD150^{+}, CD48^{-} LSKs) subpopulations could be observed (Fig. 3.1d). Taken together, these results suggest that ERCC6L2 is dispensable for haematopoietic stem cell maintenance in mice. This is in clear contrast to the phenotype displayed by patients with *ERCC6L2* mutations, which typically present with hypocellular bone marrow and mild pancytopenias at time of referral to the clinic [88, 90].

3.2.2 ERCC6L2 is dispensable for bone marrow engraftment and blood repopulation

HSCs are defined by their ability to produce progeny ultimately giving rise to all blood lineages, whilst at the same time being able to self-renew long-term. Whilst the HSC pool of *Ercc6l2*$^{-/-}$ mice did not show any quantitative defects, I wondered if there were any underlying defects in their function as stem cells. To test this, I conducted a competitive bone marrow transplant experiment.

Namely, equal numbers of *Ercc6l2*$^{-/-}$ (CD45.2) and wild-type (CD45.1/2) bone marrow cells were transplanted into lethally irradiated recipient mice (CD45.1), with a wild-type (CD45.2) vs. wild-type (CD45.1/2) experiment run in parallel. The bone marrow's ability to contribute to peripheral blood production was assessed over several weeks via flow cytometry (Fig. 3.2a). Near-complete ablation of the recipient bone marrow was achieved via irradiation, allowing for unobstructed competition between the donor bone marrows (Fig. 3.2b). *Ercc6l2*$^{-/-}$ bone marrow was able to repopulate peripheral blood to near wild-type levels (Fig. 3.2c) and no clear difference in the lineage output could be observed (Fig. 3.2d).

After 16 weeks, recipient mice were sacrificed, and their bone marrow was assessed for engraftment of donor HS(P)Cs. Again, no significant difference in percentage of engrafted donor LSKs or HSCs compared to the wild-type experiment run in parallel could be observed (Fig. 3.2e). As whole bone marrow was transplanted, not only HSCs, but also progenitors at varying stages of differentiation were being used for the repopulation of blood in the primary transplant. This means that if the fitness defect is exclusive to HSCs, it may have been overlooked as HSCs are not yet required to contribute to to blood reconstitution during the 16 weeks being observed. I therefore conducted a secondary bone marrow transplant, where bone marrow from recipients of the primary transplant is being collected and injected into a new cohort of recipient mice. This puts additional proliferative stress on the system, forcing even long-term HSCs to contribute to blood production. Unfortunately, irradiation of recipients for the secondary transplant was sub-optimal, with ∼40-50% of peripheral blood being supplied by residual CD45.1^{+} cells (Fig. 3.3a). This severely limited the sensitivity of the assay and therefore the strength of the conclusions that could be draw from the experimental outcome. However once again, *Ercc6l2*$^{-/-}$ bone marrow was able to repopulate peripheral blood to near wild-type levels (Fig. 3.3b) and no clear difference in the lineage output could be observed (Fig. 3.3c). Engraftment of LSKs and HSCs in the bone marrow was difficult to interpret as a substantial portion of the bone marrow consisted of residual

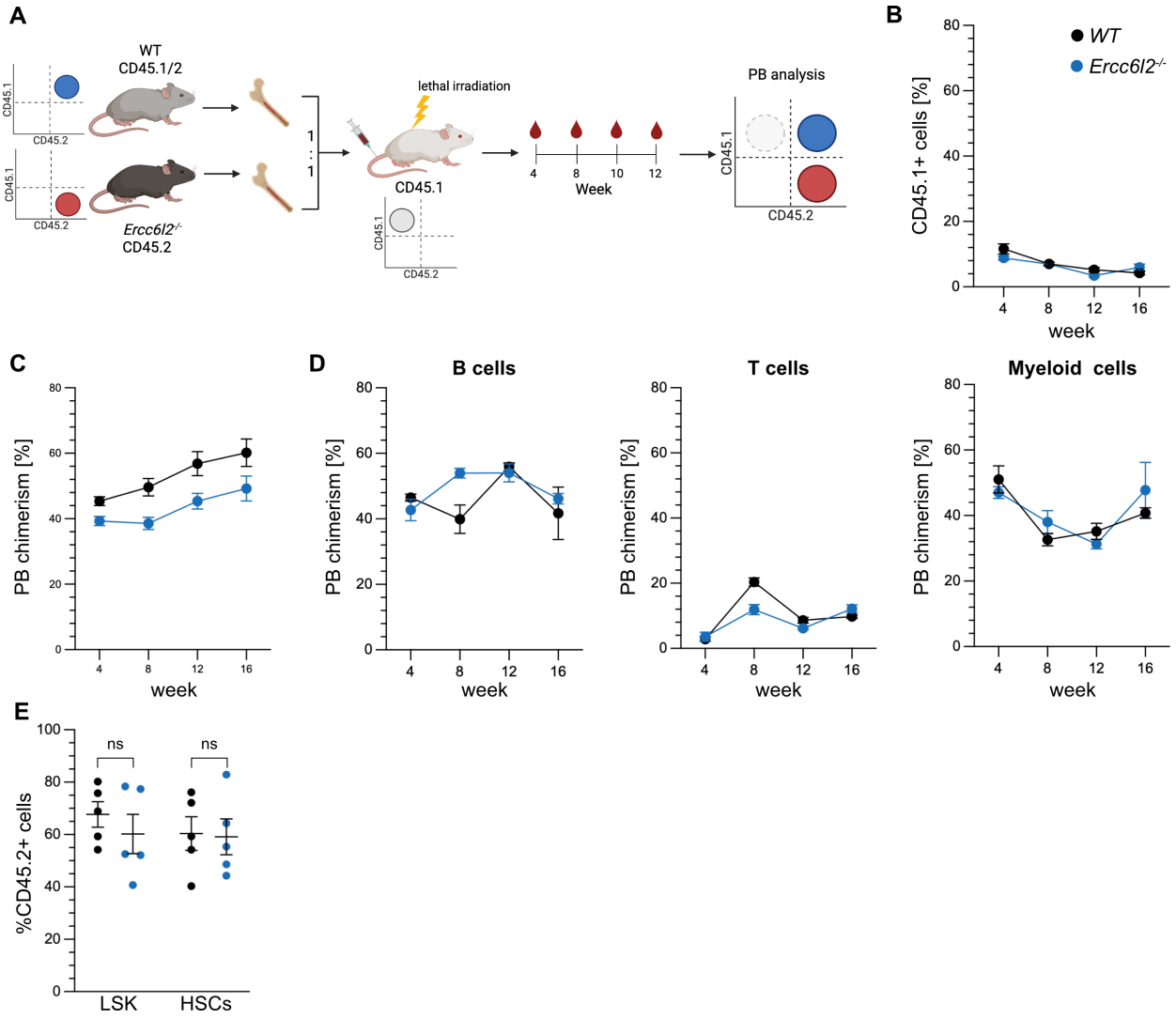


Figure 3.2: **Primary competitive bone marrow transplant.** **A** Schematic of the experimental set-up. Bone marrow from donor (CD45.2) and competitor (CD45.1/2) mice was mixed equally and injected into lethally irradiated recipient mice (CD45.1). Peripheral blood was collected over several weeks and analyzed using flow cytometry to track donor vs. competitor contribution to blood production. Schematic was created using BioRender. **B** Residual recipient (CD45.1) chimerism in nucleated peripheral blood (PB) **C** Percentage of donor chimerism in total nucleated peripheral blood. **D** Percentage of donor chimerism in peripheral B cells ($B220^+$), T cells ($CD4^+$ and/or $CD8^+$), or myeloid cells ($CD11b^+$ and/or $Ly6G^+$). **E** Percentage of donor chimerism in LSKs and HSCs in the bone marrow. Significance was determined by an unpaired two-tailed t-test (n= 4-5 mice per test condition for all plots, mean ± SEM).

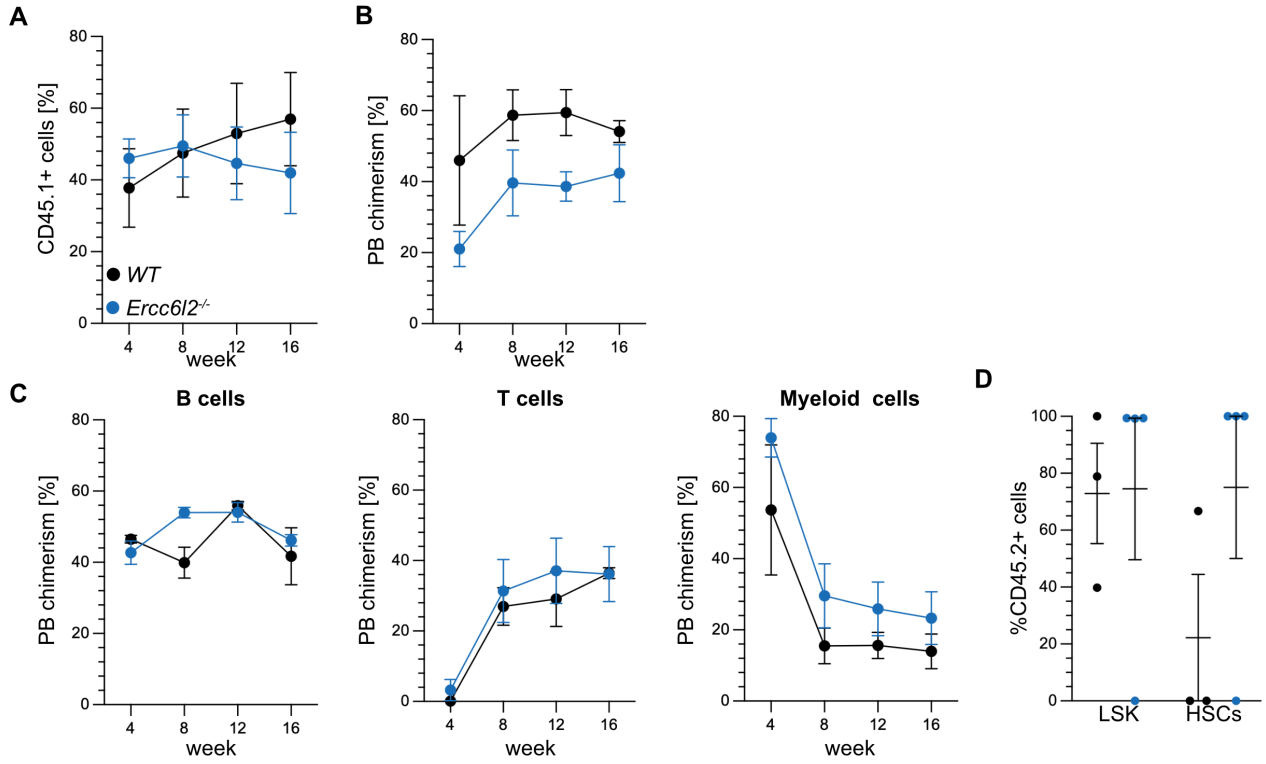


Figure 3.3: **Secondary competitive bone marrow transplant.** **A** Residual recipient (CD45.1) chimerism in nucleated peripheral blood. (PB) **B** Percentage of donor chimerism in total nucleated peripheral blood. **C** Percentage of donor chimerism in peripheral B cells ($B220^+$), T cells ($CD4^+$ and/or $CD8^+$), or myeloid cells ($CD11b^+$ and/or $Ly6G^+$). **D** Percentage of donor chimerism in LSKs and HSCs in the bone marrow(n= 3-4 mice per test condition for all plots, mean ± SEM).

cells from the recipient, with the percentage of $CD45.2^+$ LSKs and HSCs ranging from 0-100% (Fig. 3.3d).

Taken together, these results suggest that, in mice, ERCC6L2 is dispensable for both the maintenance of HSCs, and the production of mature blood cells. This is a counter-intuitive finding, which is in stark contrast to findings in humans were *ERCC6L2* mutations lead to an inherited bone marrow failure syndrome [88].

3.2.3 ERCC6L2 syndrome is not driven by aldehyde-induced DNA damage

The complete lack of a haematological phenotype in the *Ercc6l2* knockout mice was a puzzling result, which opened up the question if the stimulus needed for haematopoietic collapse was missing in the mouse model. Much in contrast to humans, mice lacking factors in the Fanconi anaemia

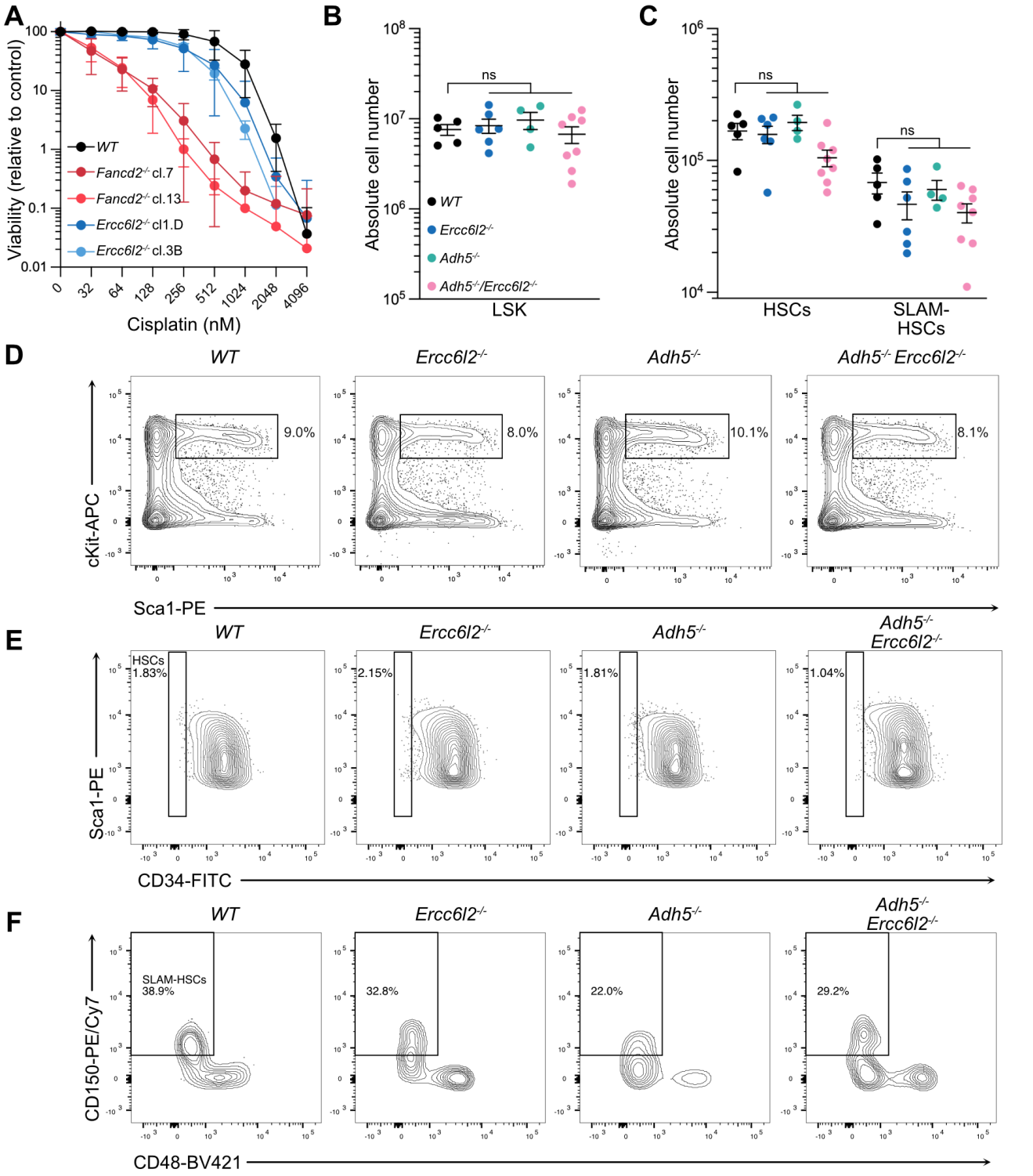


Figure 3.4: **Combined genetic loss of *Ercc6l2* and *Adh5* does not lead to synergistic loss of HS(P)Cs. A** Survival of CH12-F3 cells grown for 5 days in the presence of indicated doses of cisplatin (n= 3 biological replicates, with 3 technical replicates each, mean ± SD). **B** Absolute number of Lin$^-$ Sca1$^+$ c-Kit1$^+$ (LSK) cells in the bone marrow (one femur and one tibia) (n= 5-8 mice per genotype, where each data point is a single mouse). Significance was determined by an unpaired two-tailed t-test (mean ± SEM). **C** Absolute number of CD34$^-$ LSK cells (HSCs) and CD34$^-$ CD48$^-$ CD150$^+$ LSKs (SLAM-HSCs) in the bone marrow (one femur and one tibia) (n= 5-8 mice per genotype, where each data point is a single mouse). Significance was determined by an unpaired two-tailed t-test (mean ± SEM). **D** Representative flow cytometry plots showing Lin$^-$ cells used to quantify LSK frequencies **E** Representative flow cytometry plots showing LSK cells used to quantify HSC frequencies **F** Representative flow cytometry plots showing CD34$^-$ LSK cells used to quantify SLAM-HSC frequencies.

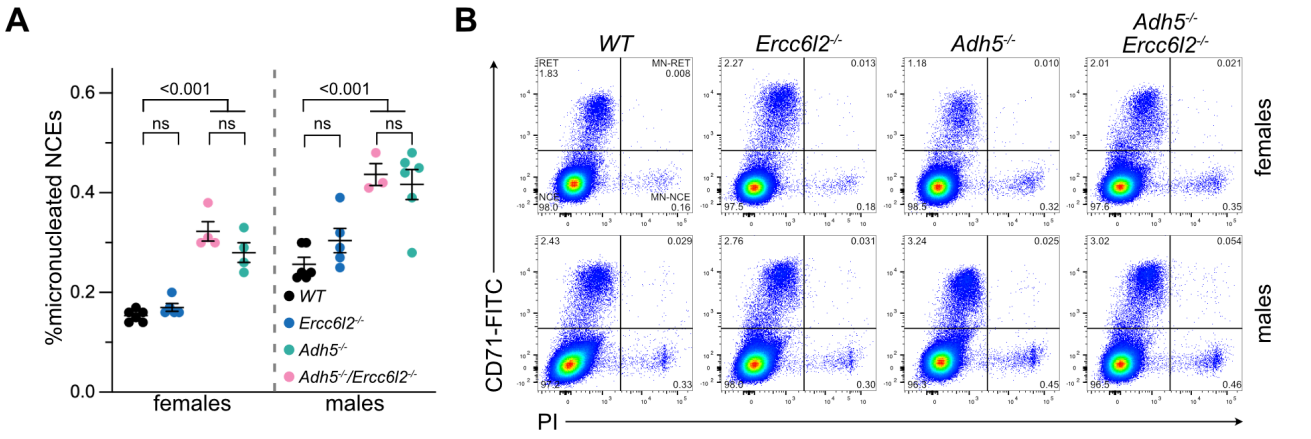


Figure 3.5: **Combined genetic loss of *Ercc6l2* and *Adh5* does not lead to synergistic genomic instability. A** Percentage of micronucleated normochromatic erythrocytes (NCEs) in the peripheral blood. Blood was taken from 8-12 week old mice via tail vein (n= 3-6 mice of each sex per genotype, where each data point is a single mouse). Significance was determined by an unpaired two-tailed t-test (mean ± SEM).**B** Representative flow cytometry plots showing peripheral blood used to quantify micronucleated erythrocytes. RET, reticulocytes; MN-RET, micronucleated reticulocytes; MN-NCE, micronucleated normochromatic erythrocytes; NCE, normochromatic erythrocytes

pathway (e.g. *Fancd2*) have superficially normal haematopoiesis, and do not spontaneously succumb to bone marrow failure or anaemia [125]. Only when ICL repair defects are genetically combined with defects in the detoxification of endogenous aldehydes (i.e. *Adh5* or *Aldh2*), do mouse models phenocopy the human disease, with mice rapidly developing aplastic anaemia and haematological malignancies [50, 51].

Fibroblasts collected from patients with ERCC6L2 syndrome have been reported to be sensitive to DNA interstrand crosslinking agents [90]. Using cell survival assays, I could recapitulate a mild sensitivity to cisplatin, another DNA crosslinker(Fig. 3.4a). I therefore hypothesized that ERCC6L2 - in addition to its role in NHEJ - might have a role in ICL repair and, subsequently, that ERCC6L2 syndrome could represent a milder form of Fanconi anaemia. To test this, I exploited the "two tier protection model" , intercrossing *Adh5*$^{-/-}$ mice with *Ercc6l2*$^{-/-}$ mice in order to derive *Adh5*$^{-/-}$*Ercc6l2*$^{-/-}$ progeny co-deficient in *Ercc6l2* and the metabolic clearance of endogenous aldehydes via ADH5. If mice had any underlying ICL repair defects, removal of protection from endogenous aldehydes via ADH5 would likely lead to loss of HSC function [51]. However, animals lacking both *Ercc6l2* and *Adh5* did not show a synergistic reduction in the

number of LSKs (Fig. 3.4b, d), or HSCs (Fig. 3.4c, e, f).

Moreover, genomic instability as defined by the level of micronuclei in mature erythrocytes was not increased in *Ercc6l2*$^{-/-}$ mice (Fig. 3.5a, b). During erythropoiesis, cells naturally enucleate, meaning any remaining DNA in peripherally circling erythrocytes is an indicator of global genomic instability [119]. As previously reported, *Adh5*$^{-/-}$ mice showed a significant increase in the number of micronucleated normochromatic erythrocytes (NCEs) [53]. However, unlike factors in ICL repair, combined loss of *Ercc6l2* and *Adh5* did not lead to a synergistic increase in the proportion of micronucleated erythrocytes [126].

These data suggest that ERCC6L2 does not have a role in ICL repair. Furthermore, endogenous aldehydes are unlikely to be the driver of haematopoietic collapse in ERCC6L2 syndrome. Taken together, this suggests that ERCC6L2 syndrome is a distinct disease unrelated to Fanconi anaemia.

3.2.4 A quest for answers - genome-wide CRISPR screen in *Ercc6l2*$^{-/-}$ HSPCs

Sample	Total sgRNAs	Zero counts	Gini-Index
plasmid	79633	437	0.05
WT_T0	79633	529	0.06
WT_T7	79633	1270	0.11
WT_T14	79633	2263	0.19
WT_T21	79633	2856	0.26
E6L2_T0	79633	518	0.07
E6L2_T7	79633	1332	0.13
E6L2_T14	79633	2405	0.19
E6L2_T21	79633	2549	0.24

Table 3.1: **Results from MAGeCK count algorithm.** Total sgRNAs expected, number of gRNAs undetected (Zero counts) and distribution of read counts (Gini-Index) are listed for each sample

Having excluded an ICL repair defect as the underlying cause for the phenotype of ERCC6L2-deficient patients, I decided to take a more agnostic approach to finding the stimulus required for HSC collapse by performing a genome-wide CRISPR screen in the disease-affected cell type

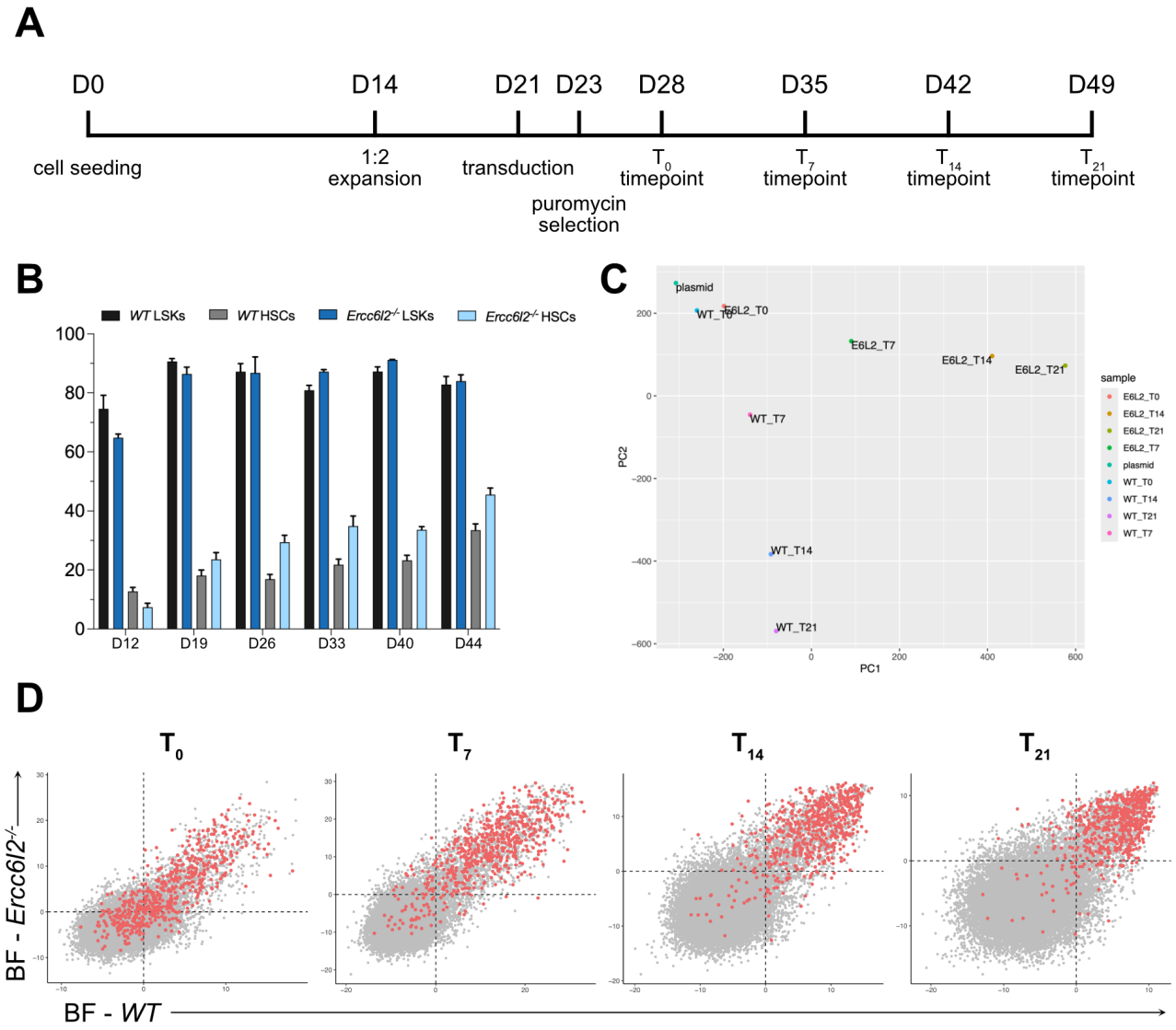


Figure 3.6: **Summary of set-up and quality control for genome-wide CRISPR screen A** Schematic of timeline from cell seeding to last timepoint. **B** Percentage of LSKs and HSCs ($CD150^+$ $CD201^+$ LSKs) in total cultured cells over time for both genotypes. One well per plate per genotype was analyzed (mean ± SD). **C** PCA analysis of sequencing results as part of the analysis with MAGeCK. The first two principle components (PCs) are shown. **D** Results of CRISPR screen in primary murine HSPCs. Bayes Factor (BF) for each gene in wild-type vs. *Ercc6l2*$^{-/-}$ backgrounds is plotted at each time point. Genes categorized as common essential by the BAGEL2 algorithm (CEG_mouse) are highlighted in red. The dashed line marks BF=0.

- primary HSPCs. To this end, mice ubiquitously expressing Cas9-eGFP in a wild-type and *Ercc6l2*$^{-/-}$ background were generated. HSPCs from these mice were isolated and cultured *in vitro* using a culture method allowing for large-scale *ex vivo* HSC expansion to achieve the cell numbers needed for a genome-wide CRISPR screen [118, 127]. After 3 weeks of expansion, a genome-wide CRISPR screen was performed with the mouse CRISPR Brie library containing sgRNAs targeting 19,674 genes (4 gRNAs/gene) and 1000 control gRNAs [128].

Transduced primary HSPC cultures were maintained for 5 weeks, with a fraction the cell pool collected for subsequent processing and sequencing each week (Fig. 3.6a). Flow cytometry confirmed good culture purity with cells consisting of $>$ 80% LSKs and 20-40% HSCs at the time of, and subsequently to transduction (Fig. 3.6b). gRNA sequencing revealed high library representation with 99.3% of gRNAs contained in the library detected at T_0 (Table3.1). Principle component analysis (PCA) of gRNA counts of each sample showed diverging trajectories for the two tested genotypes over time (Fig. 3.6c). Results were analyzed using BAGEL2, which scores genes based on their essentiality [129]. Reassuringly, common essential genes showed positive BF (Bayes Factor) scores (indicating essentiality) in both backgrounds and became gradually more enriched in the common essential quadrant (i.e. high BF_{WT} and $BF_{Ercc6l2}$) over time (Fig. 3.6d).

Having confirmed that the CRISPR screen worked well on a technical level, and with no obvious hits exclusively essential in either background, I next focused on exploring hits related to two area of special interest: DNA repair (GO:0006281, Fig. 3.7) and telomere maintenance (GO:0000723, Fig. 3.8). For the sake of simplicity, results are only shown for the timepoints T_7 and T_{21}, as these have the highest BF scores and longest time for selection of hits, respectively.

Encouragingly, *Trp53* (which encodes p53) scored as one of the top hits for genes that limit cell survival. Given that p53 plays a central role in regulating HSC quiescence and apoptosis, it is unsurprising that *Trp53* knockout emerges as one of the strongest hits enhancing HSC

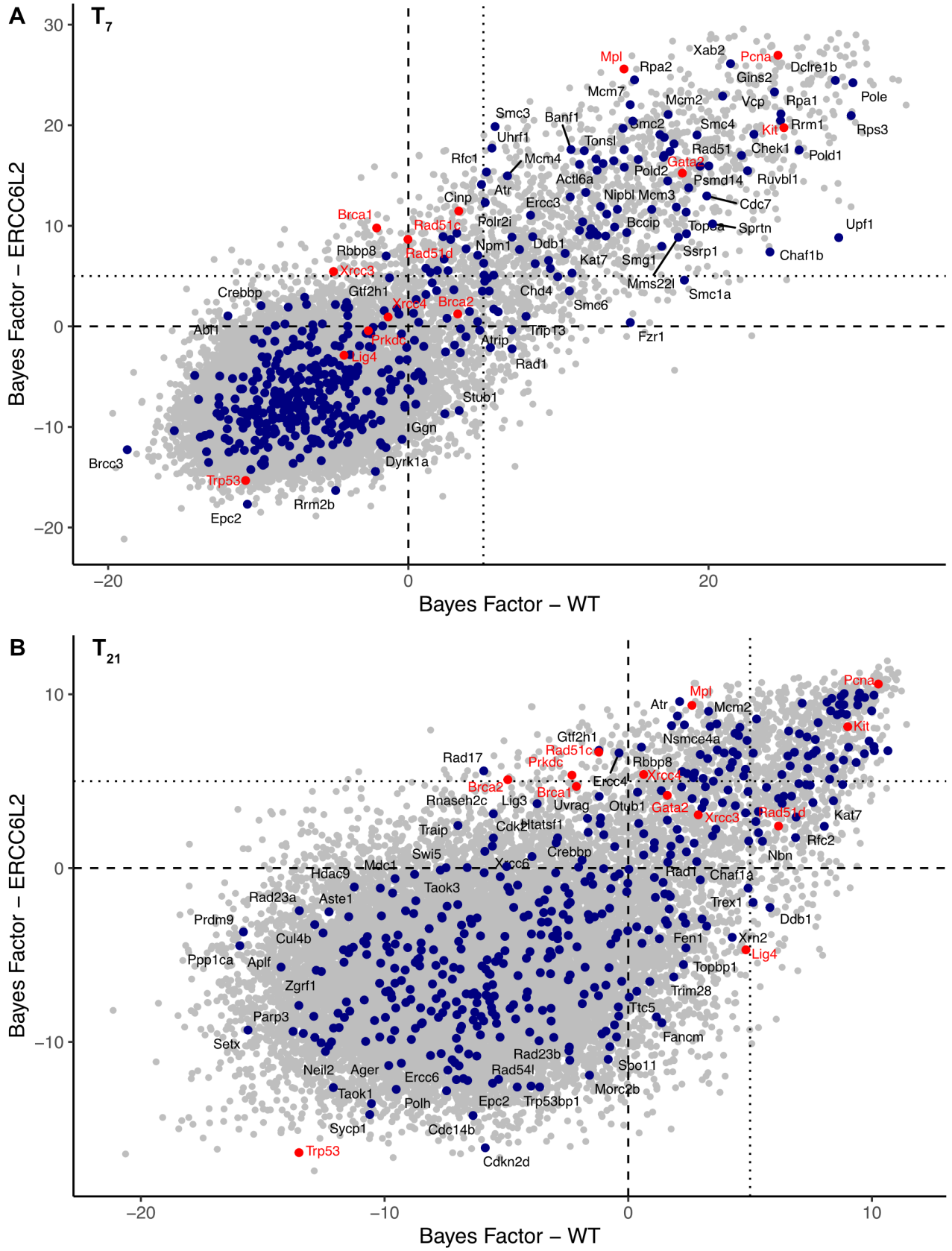


Figure 3.7: **CRISPR screen results - DDR hits.** Results of CRISPR screen with DNA repair (GO:0006281) hits highlighted in blue and hits discussed in the text highlighted in red. Bayes Factor (BF) for each gene is plotted in wild-type vs. *Ercc6l2*$^{-/-}$ backgrounds. The dashed line markes BF=0 and the dotted line marks BF=5 (minimum score for essentiality). **A** Results for timepoint T_7. **B** Results for timepoint T_{21}.

proliferation (Fig. 3.7) [130, 131]. On the other extreme, one of the top hits inhibiting cell survival is *Pcna* knockout. PCNA is a DNA sliding clamp that is essential for DNA replication, making its loss highly lethal [132]. Furthermore, genes well established to be crucial for HSC identity (i.e. *Gata2* - a key haematopoietic transcription factor [133], *Mpl* - the thrombopoietin receptor [134], and *Kit* - the SCF receptor [135]) were correctly identified as essential. These examples show that the experiment was able to accurately capture both general and HSC-specific genetic dependencies and validate the approach.

As the CRISPR screen aimed to uncover genetic interactions with *Ercc6l2*, I was particularly interested in hits where the gene showed a clear difference of essentiality in between genetic backgrounds. Specifically hits scoring BF>5 (signifying minimum confidence essentiality) in one background, but not the other (i.e. the upper left and lower right quadrant sectioned by the dotted lines) were examined. NHEJ factors (i.e. *Xrcc4*, *Prkdc* (encoding DNA-PKcs)) were showed significantly increased essentiality in an *Ercc6l2*-deficient background, likely due to repair intermediates relying more heavily on other factors in NHEJ when ERCC6L2 is not functional (Fig. 3.7). Curiously, *Lig4* showed the opposite trend, with it showing increased essentiality in wild-type cells. This could indicate that in *Ercc6l2*-deficient cells, DSB repair is already skewed towards pathways other than NHEJ, making loss of ligase IV less relevant, whereas in wild-type cells, loss of NHEJ is more detrimental. This would fit in well with the fact that loss of HR factors (such as *Brca1*, *Brca2*, *Rad51c*, *Rad51d*, *Xrcc3*) appeared as being selectively detrimental in *Ercc6l2*-deficient cells [136].

Due to a hypothesis that ERCC6L2 might have a function in the the protection of telomeres (discussed in detail in Section6.3), I was also interested in exploring genes involved in telomere maintenance. Telomere ends are normally not recognised by the DDR as the shelterin complex sequesters the DNA end in a T-loop, thereby "hiding" them from being recognised as a DSB [63]. Apollo (encoded by *Dclre1b)* is an exonuclease that is required for the resection of newly

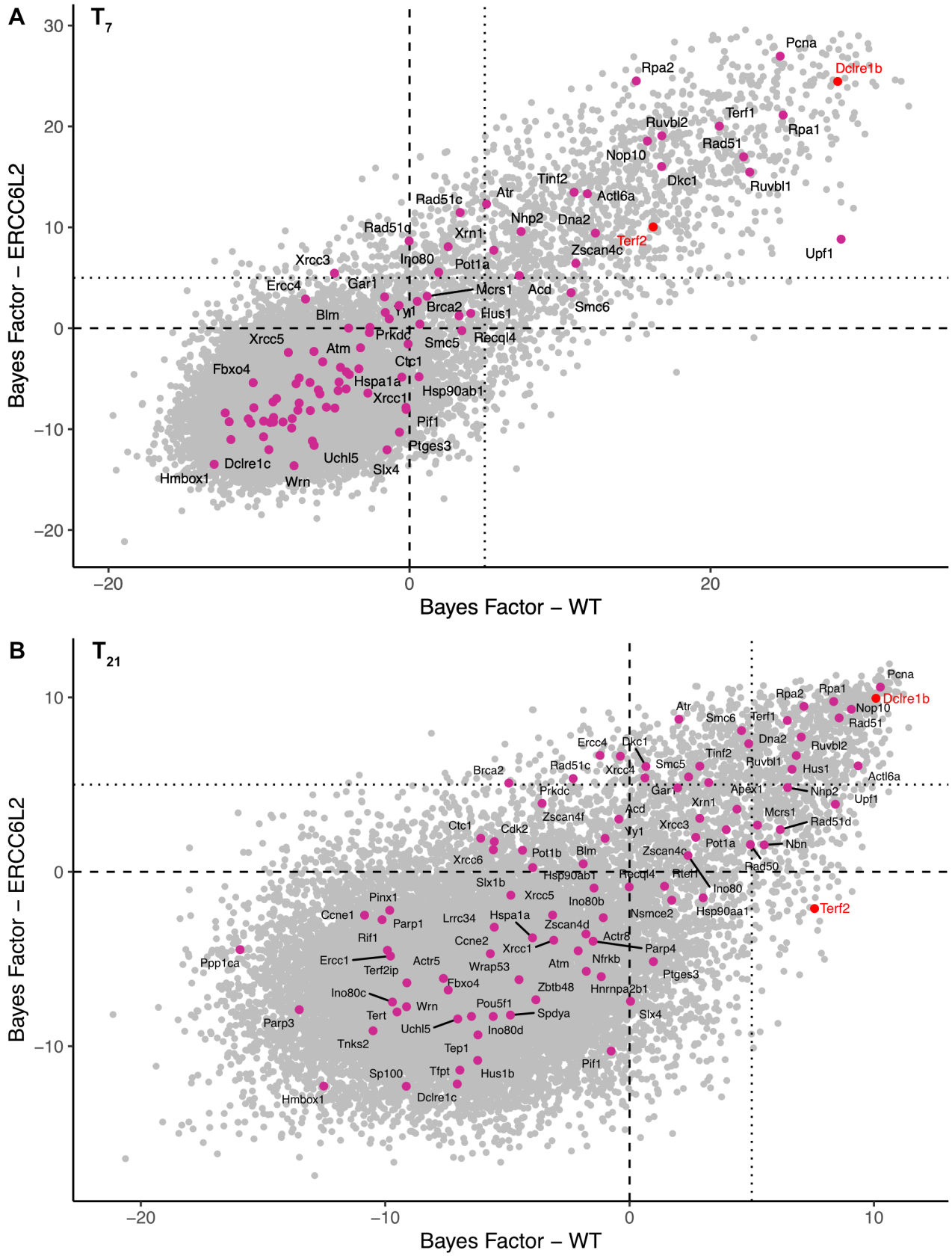


Figure 3.8: **CRISPR screen - telomere maintenance hits.** Results of CRISPR screen with telomere maintenance (GO:0006281) hits highlighted in pink and hits discussed in the text highlighted in red. Bayes Factor (BF) for each gene is plotted in wild-type vs. *Ercc6l2*$^{-/-}$ backgrounds. The dashed line marks BF=0 and the dotted line marks BF=5 (minimum score for essentiality). **A** Results for timepoint T_7. **B** Results for timepoint T_{21}.

synthesised telomeric ends, allowing for sequestration of the single-stranded DNA end in the T-loop [137, 138]. Without Apollo, the T-loop cannot be formed on newly synthesised telomeres and ATM gets activated by telomere ends during S phase. Accordingly, knockout *Dclre1b* is highly toxic and scores as highly essential in both genetic backgrounds (Fig. 3.8). Perhaps most interestingly, loss of *Terf2* (which encodes TRF2) was much more detrimental in wild-type cells compared to *Ercc6l2*$^{-/-}$ cells. Without TRF2, the T-loop cannot be formed, exposing DNA resembling a staggered DSB (i.e. a DSB with a single-strand overhang) [139]. As discussed in Section 5, this is the favoured substrate for ERCC6L2-supported NHEJ. It is therefore likely, that without ERCC6L2, telomeres are less efficiently fused in the absence of TRF2, conferring a selective advantage when both *Terf2* and *Ercc6l2* are knocked out compared to only *Terf2*.

Taken together, these results show that the CRISPR screen was able to identify well established factors inhibiting or promoting (haematopoietic stem) cell survival and proliferation and could also showcase more subtle differences of genetic interactions in between a wild-type and an *Ercc6l2*-deficient background. I was however unable to identify clear candidates for "first tier" genes encoding DSB protection activities that is selectively required in *Ercc6l2*-deficient cells. This suggests that (at least in the context of ERCC6L2), there is no two tiered protection against DSBs and that it is unlikely that bone marrow failure can be precipitated in an *Ercc6l2*$^{-/-}$ mouse model by the removal of DSB-protective factors.

Due to time constraints, and more promising lines of enquiry, the identified hits were not selected for functional experimental re-validation. Follow-up experiments are required to fully assess the functional relevance of identified genetic modifiers of *Ercc6l2*-deficiency in primary HSPCs.

3.3 Concluding remarks

Germline mutations in *ERCC6L2* lead to an inherited bone marrow failure syndrome - termed ERCC6L2 syndrome [88, 89]. Yet, how ERCC6L2 supports blood development and why its dysfunction leads to HSC attrition and leukaemic transformation remains poorly understood. In this chapter, I investigated how ERCC6L2 deficiency impacts blood stem cell functionality and genome stability, aiming to uncover the molecular basis of ERCC6L2 syndrome. To this end, an *Ercc6l2* knockout mouse model was generated. Having confirmed that the knockout mice do not express the *Ercc6l2* transcript, I characterised their HS(P)C compartment and found that they had normal numbers of both HSPCs and LT-HSCs. Stem cell functionality was tested in a primary and secondary competitive bone marrow transplant, with *Ercc6l2*-deficient cells engrafting into the bone marrow, and repopulating the blood to near wild-type levels. Together, these data show that ERCC6L2 is dispensable for the haematopoietic stem cell maintenance and blood production in the mouse.

Having recapitulated a mild sensitivity of *Ercc6l2*-deficient cells to cisplatin, I tested if a HSC phenotype could be precipitated by increasing the endogenous ICL burden through the removal of aldehyde detoxification by ADH5. Again, HS(P)C numbers were not decreased and genome stability was unaffected as seen by the level of micronucleated erythrocytes, excluding endogenous aldehydes and ICLs as the likely physiological cause of HSC attrition in absence of ERCC6L2. Nevertheless, these results exemplify the aetiological distinctions between ERCC6L2 syndrome and Fanconi anaemia, and exclude endogenous aldehydes as the source of DNA damage in the former disease.

Having ruled out endogenous aldehydes as a relevant source of causative DNA damage in *ERCC6L2*-disease, I undertook a genome-wide CRISPR screen in primary murine HSPCs to uncover selective genetic modifiers of *Ercc6l2*-deficiency that could help explain the disease phenotype. The screen

was able to replicate well known genetic dependencies relating to both general cell survival and haematopoietic stem cell biology, validating the approach. *Ercc6l2*-deficient cells showed an increased dependency on both NHEJ and HR factors, suggesting underlying DSB repair defects. However, I was unable to identify any synthetic interactions suggesting factors protecting cells against formation of DSBs. This suggests that there either is no non-redundant tier-1 protection activities suppressing the abundance of DSBs necessitating ERCC6L2 for DNA repair, and/or such a concept is irrelevant in the context of ERCC6L2. Considered together, it remains to be determined why the phenotypes of *ERCC6L2*-deficiency are so starkly different in between humans and mice. Several hypothesis that could explain the discrepancy are discussed in Section 6.

Chapter

DEFINING THE ROLE OF ERCC6L2 IN LYMPHOCYTE DEVELOPMENT

4

4.1 Introduction

With ERCC6L2 having been classified as a novel factor in NHEJ, investigating its function in lymphocyte development was a logical next step [90, 101–103]. In line with this, several groups also reported a significant reduction of isotype switching upon CSR using either the CH12-F3 B cell lymphoblastoid cell line model, and/or a mouse model deleted for *Ercc6l2* [101–103]. However, ERCC6L2's functional contribution to CSR remains ill-defined, and similarly its role in V(D)J recombination - and therefore lymphocyte differentiation - has been less well characterised. For example, Liu *et al.* reported a 50% reduction of total thymic cells in *Ercc6l2*$^{-/-}$ mice, yet distributions of T cell development stages were paradoxically comparable to wild-type controls [102]. Concurrently, total cell numbers in the bone marrow were described as being unaffected, and similarly, B cell development was reportedly superficially normal. However, a detailed characterisation of B cell precursor subsets using Hardy fractions was not performed.

This would suggest that ERCC6L2 is dispensable for V(D)J recombination. However, in a sensitised (*Xlf*-deficient) background, loss of *Ercc6l2* led to a near-total collapse of V(D)J recombination in *v-Abl* pre-B cells. Liu *et al.* therefore concluded that ERCC6L2 is an accessory cNHEJ factor, facilitating orientation-specific recombination in both V(D)J and CSR [102]. In a parallel publication, Francica *et al.* suggested ERCC6L2 might be part of 53BP1-RIF1-shieldin-CST pathway instead [103].

In this chapter, I characterise the lymphocyte development of *Ercc6l2*-deficient mice in detail. To address the conflicting literature on ERCC6L2's association with either cNHEJ or the 53BP1-RIF1-shieldin-CST pathway, I further characterise the lymphocyte development in mice lacking *Ercc6l2*, both in isolation, and in mice co-deficient in either *53bp1* or *Shld2*. Ultimately, my approach aims to use genetic epistasis (or lack thereof) in V(D)J recombination and CSR as a vehicle to better position ERCC6L2 mechanistically in DSB repair and advance the field's understanding of its functional significance in lymphocyte development.

4.2 Results

4.2.1 ERCC6L2 is dispensable for lymphocyte differentiation

The differentiation and maturation of lymphocytes is intrinsically linked with successful V(D)J recombination, which in turn is completely dependent on NHEJ. Lack of core NHEJ factors (such as KU80) leads to an inability to perform V(D)J recombination, therefore fully blocking B and T cell differentiation at early progenitor stages [140]. Curiously, whilst *XLF* mutations cause lymphocytopenia in humans, *Xlf*-deficient mice only show modest decreases in the number of lymphocytes [56, 57].

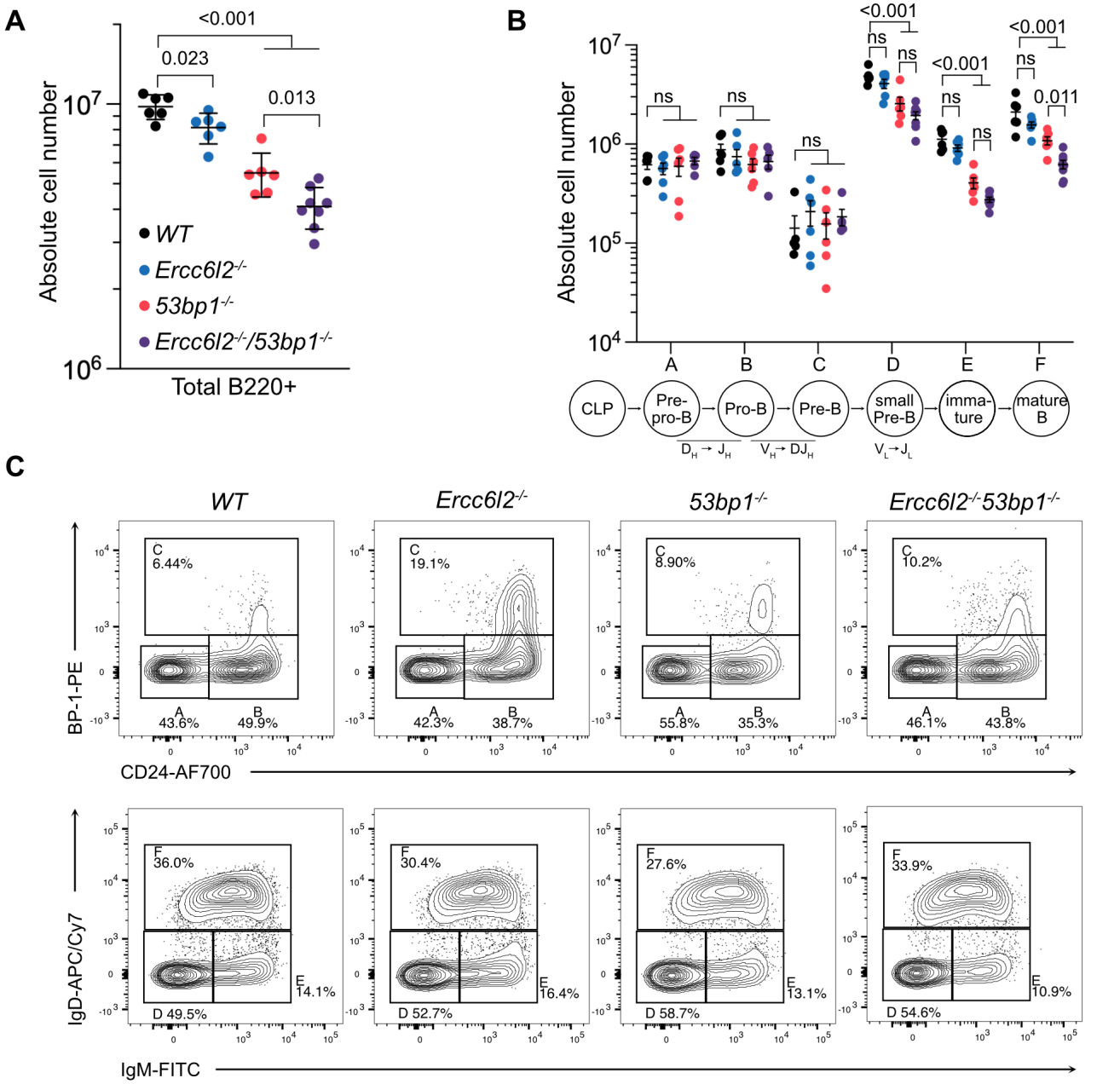


Figure 4.1: **ERCC6L2 is dispensable for B cell development - even in absence of 53BP1. A** Absolute number of B220$^+$ B cells in the bone marrow (one femur and one tibia). **B** Absolute number of B220$^+$ B cell precursors in the bone marrow (one femur and one tibia). **C** Representative flow cytometry plots used to quantify B cell development in the bone marrow, gating on B220$^+$ CD43$^+$ for Hardy fractions A, B, C, and on B220$^+$ CD43$^-$ for Hardy fractions D, E, F. For all quantitative graphs, n= 5-8 mice per genotype, where each data point is a single mouse and significance was determined by an unpaired two-tailed t-test (mean ± SEM). ns, not significant

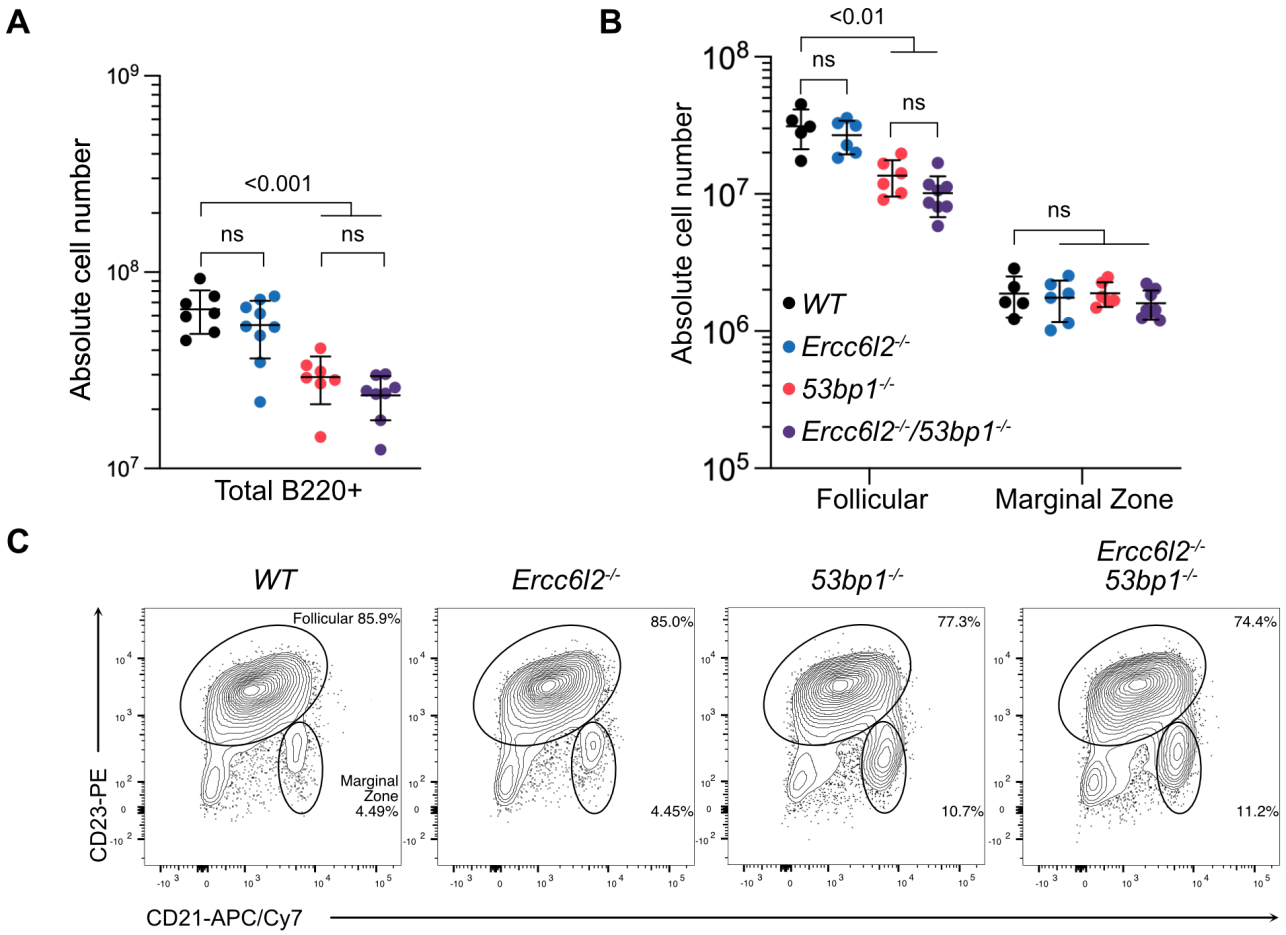


Figure 4.2: **ERCC6L2 is dispensable for mature B cells - even in absence of 53BP1. A** Absolute number of B220$^+$ B cells in the spleen. **B** Absolute number of follicular and marginal zone B cells in the spleen. **C** Representative flow cytometry plots showing B220$^+$ CD19$^+$ cells used to quantify mature B cell populations in the spleen. For all quantitative graphs, n= 5-8 mice per genotype, where each data point is a single mouse and significance was determined by an unpaired two-tailed t-test (mean ± SEM). ns, not significant

In contrast, deletion of DNA damage response proteins (such as ATM, H2AX, or 53BP1) causes more subtle effects on lymphocyte development [80, 141, 142]. For example, *53bp1*$^{-/-}$ mice show a 50-80% in the number of total lymphocytes due to impaired distal V(D)J recombination but can still produce mature B and T cells [143].

To test if ERCC6L2 is involved in V(D)J recombination, B and T lymphocyte development of *Ercc6l2*$^{-/-}$ mice were analysed using flow cytometry. At first, a marginal ≈ 15% reduction of total B cells in the bone marrow could be observed compared to wild-type controls (Fig. 4.1a). However, this did not hold true for later experiments, were *Ercc6l2*$^{-/-}$ mice showed no significant reduction of B cells (Fig. 4.6a). These later results also fit in with similar observations published

previously, where B cell numbers were unaffected [102].

Additionally, thorough immunophenotypic characterisation of individual B cell precursor stages using Hardy fractions showed no significant reduction of cell numbers at any particular B cell differentiation stage (Fig. 4.1b-c). In contrast, *53bp1*$^{-/-}$ mice had 50% fewer B cells than wild-type controls (Fig. 4.1a). In line with 53BP1's function in facilitating long-range DNA end-joining during V(D)J recombination, this reduction was restricted to the later stages of B cell development, when immunoglobulin locus should have been successfully rearranged and expressed (Fig. 4.1b-c) [143].

Importantly, combined loss of *Ercc6l2* and *53bp1* only led to a small additional decrease in total B cell numbers in bone marrow compared to *53bp1*$^{-/-}$ animals (Fig. 4.1a). This decrease was also not localized to any particular mature B cell differentiation stage, which would be expected if a this was due to a synergistic V(D)J recombination defect (Fig. 4.1b-c).

Similarly, loss of *Ercc6l2* did not impact on the total number of splenic B cells (Fig. 4.2a) or any B cell subpopulations in the spleen (Fig. 4.2b-c). In contrast, *53bp1*$^{-/-}$ animals had fewer splenic B cells overall and particularly follicular B cells numbers were reduced. Moreover, combined loss of *Ercc6l2* and *53bp1* had no synergistic effect on splenic B cells - either when looking at total B cell numbers or splenic B cell subpopulations (Fig. 4.2a-c).

In tandem with the characterisation of B cell development, T cell development was also analysed using flow cytometry. The total number of thymocytes was slightly elevated compared to wild-type (Fig. 4.3a). However, this did not correspond to a significant change in any T cell development change (Fig. 4.3b-c). As seen in the B cell populations, loss of *53bp1* lead to a reduction of total thymocytes (Fig. 4.3a), which is due to block of late T cell differentiation (Fig. 4.3b) caused by impaired V(D)J recombination [143]. Again, combined loss of *Ercc6l2* and *53bp1* did not cause a

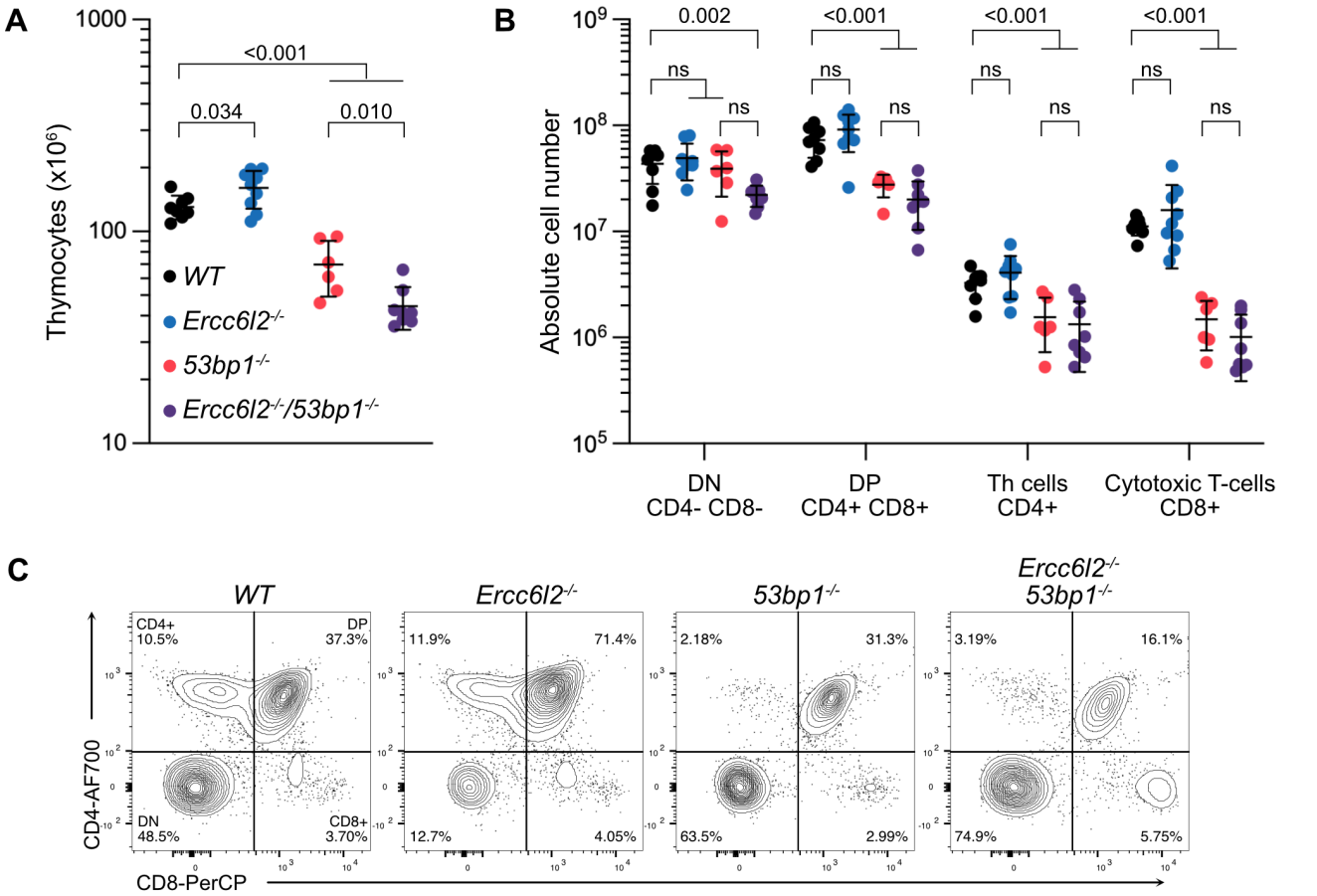


Figure 4.3: **ERCC6L2 is dispensable for T cell development - even in absence of 53BP1. A** Absolute number of nucleated cells in the thymus (n= 6-8 mice per genotype, where each data point is a single mouse) Significance was determined by an unpaired two-tailed t-test (mean ± SEM). **B** Absolute number of thymic T cells (n= 6-8 mice per genotype, where each data point is a single mouse). Significance was determined by an unpaired two-tailed t-test (mean ± SEM). ns, not significant; DN, double-negative; DP, double-positive **C** Representative flow cytometry plots showing thymocytes used to quantify T cells in each developmental stage.

total collapse of T cell differentiation - in fact, whilst total thymocyte counts decreased slightly, this was once again not localized to any particular differentiation stage.

To robustly quantify V(D)J recombination, I generated viral-Abelson kinase (*v-Abl*) transformed pre-B cell lines from *Ercc6l2*$^{-/-}$ and *Ercc6l2*$^{-/-}$/*53bp1*$^{-/-}$ mice and deleted *Ercc6l2* in *Xlf*$^{-/-}$ *vAbl*-pre B cells using CRISPR/Cas9. Treatment of *v-Abl* transformed pre-B cells with the *v-Abl* kinase inhibitor Imatinib causes G_1 cell cycle arrest, induction of RAG1/2 expression and rearrangement of the retrovirally-introduced V(D)J reporter substrate (pMX-INV) [116]. Successful rearrangement of the substrate leads to expression of GFP, which can subsequently be quantified using flow cytometry (Fig. 4.4a).

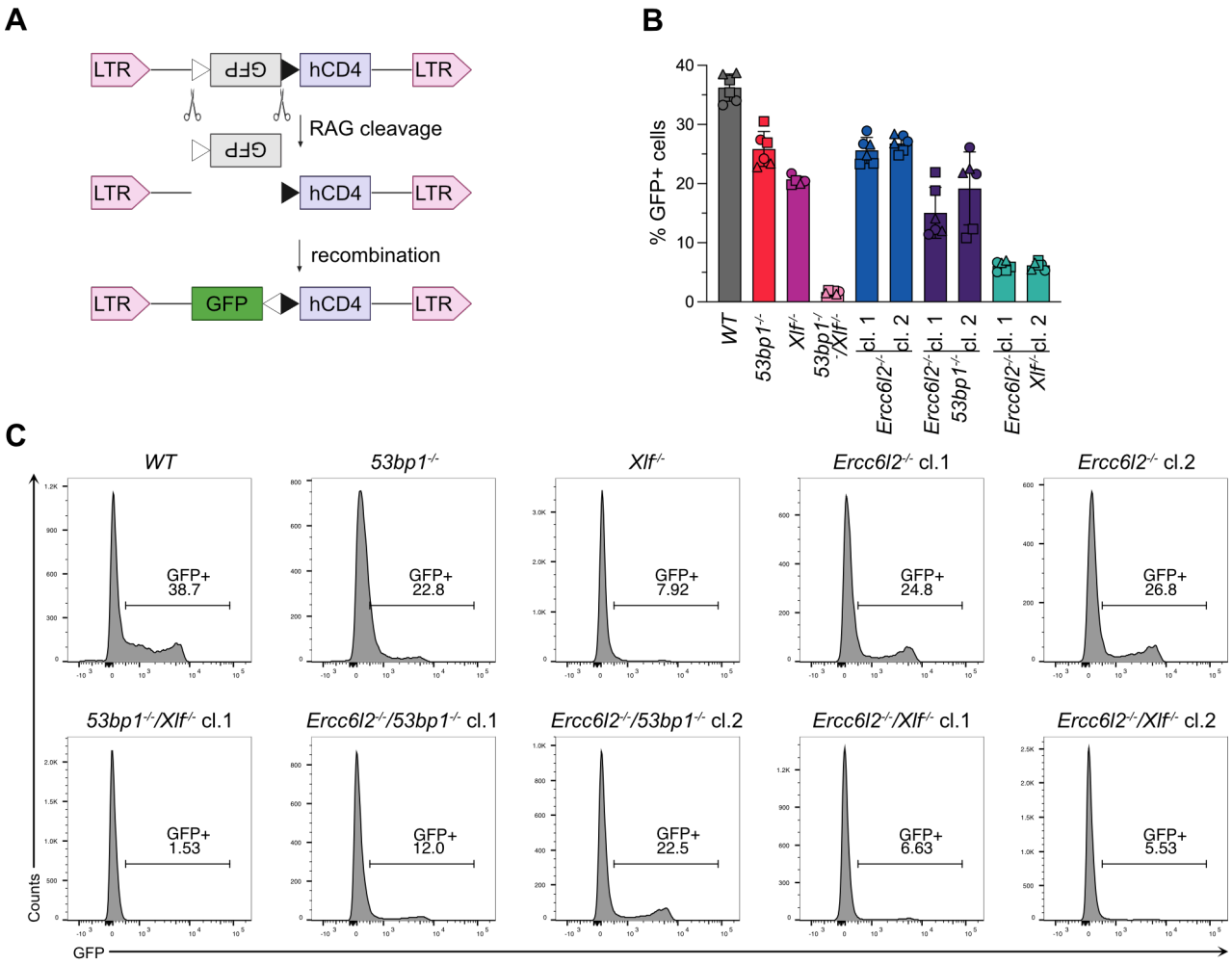


Figure 4.4: **ERCC6L2 participates in extra-chromosomal V(D)J recombination. A** Schematic of pMX-INV retroviral V(D)J substrate. The vector has two RSSs that flank an inverted GFP cassette. Functional V(D)J recombination flips the cassette, enabling GFP expression. **B** Bar plot showing V(D)J recombination efficiency of pMX-INV recombination substrate in *v-Abl* pre-B cell lines of the indicated genotypes. Cells were treated with Imatinib for 48 h and GFP expression was determined using flow cytometry. Bar represents mean ± SD, n=3 biological replicates, where each data point is one technical replicate. **C** Representative flow cytometry plots showing quantification of GFP expression.

Ercc6l2$^{-/-}$ *v-Abl* pre-B cells showed an equivalent V(D)J recombination defect to *53bp1*$^{-/-}$ cells, but the defect was slightly less pronounced than in *Xlf*$^{-/-}$ cells (Fig. 4.4b-c). Combined loss *Ercc6l2* with either *53bp1* or *Xlf* lead to a further decrease in GFP expression. However, neither combination was as severe as combined *53bp1/Xlf* deficiency, which yielded a near complete ablation of reporter recombination, as evidenced by the absence of GFP-expressing cells.

These data might suggest that ERCC6L2 has a concealed role in V(D)J recombination. However, due to the artificial nature of the assay, and the opposing results found *in-vitro*, I would interpret these results with caution. It seems more likely that ERCC6L2 has no role in physiological V(D)J

recombination, and only participates in more synthetic scenarios outside of the physiological Ig locus and context.

Taken together, these results suggest that ERCC6L2 is fully dispensable for lymphocyte development and that it does not have a role in physiological V(D)J recombination. Of special interest are the findings that combined loss of *Ercc6l2* and *53bp1* did not lead to a synergistic collapse of lymphocytes. These results are in stark contrast to similar experiments done with *53bp1*$^{-/-}$ *Xlf*$^{-/-}$ mice [144]. Not only were combined knockout mice significantly smaller than their littermates, lymphocyte development was completely blocked at early progenitor stages - phenotypes similar to mice deficient in core NHEJ factors.

In contrast, little to no synergy was found in *Ercc6l2*$^{-/-}$/*53bp1*$^{-/-}$ mice, showing that even in a "sensitised" background, the presence or abscence of ERCC6L2 was of little consequence for V(D)J recombination and lymphocyte development. Yet, it is essential to note that in similar experiments done with *Xlf*$^{-/-}$ and *Shld1*$^{-/-}$ *v-Abl* pre-B cells, combined loss of *Xlf* and *Shld1* was completely equivalent to *Xlf*-deficiency by itself, showing that shieldin does not participate in V(D)J recombination at all [145]. This would suggest that whilst ERCC6L2 does not participate in physiological V(D)J recombination, it also is not part of the 53BP1-RIF1-shieldin axis strictly required for CSR but fully dispensable for V(D)J recombination [25, 83].

4.2.2 ERCC6L2 supports CSR independently of the shieldin complex

Several groups have reported that ERCC6L2 participates in CSR [101–103]. To investigate ERCC6L2's functional relationship and epistasis with other proteins involved in the repair of AID-induced DSBs during CSR,*Ercc6l2*$^{-/-}$ mice were crossed with *53bp1*$^{-/-}$ or *Shld2*$^{-/-}$ mice.

Productive repair of AID-induced DSBs during CSR is entirely dependent on 53BP1, with loss of 53BP1 leading to even stronger defects in CSR than when core NHEJ factors are missing [80, 146].

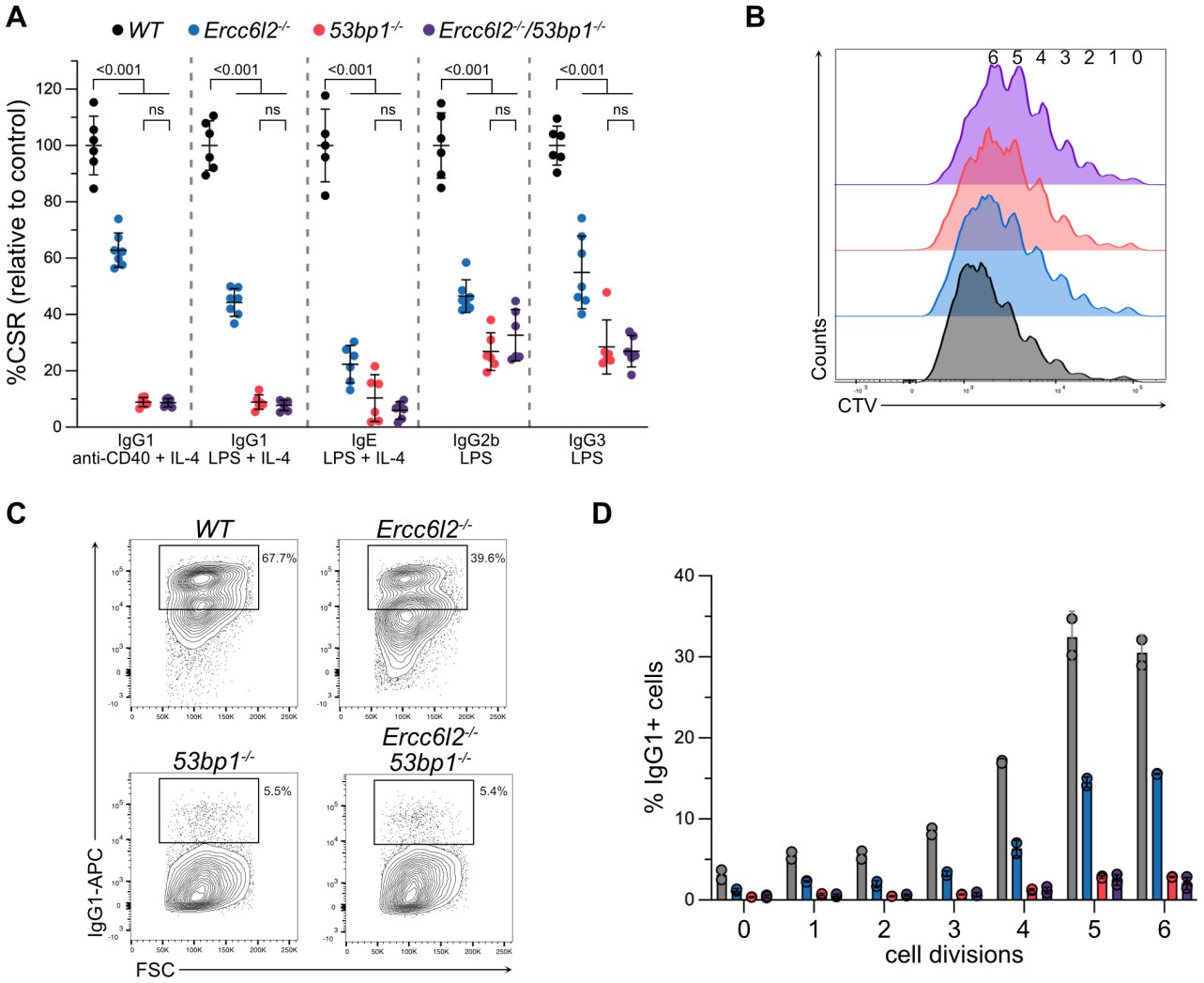


Figure 4.5: **ERCC6L2 supports NHEJ during CSR downstream of 53BP1.** **A** Splenic B cells cultured with indicated cytokines for 96 h and stained for IgG1, IgE, IgG2b, or IgG3. (n= 5-6 mice per genotype, where each data point represents a single mouse). Isotype switching frequency was normalized to two wild-type mice for each experiment. Significance was determined by a one-way ANOVA with Tukey's correction (mean ± SEM). ns, not significant **B** Representative flow cytometry plots of CTV dilution in purified splenic B cells stimulated with LPS and IL-4 for 96 h. **C** Representative flow cytometry plots showing splenic B cells cultured with anti-CD40 and IL-4 for 96 h and stained for IgG1. **D** Percentage of IgG1$^+$ B cells in each cell division as determined by proliferation-associated CTV dilution (n= 2 mice per genotype, where each data point is a single mouse).

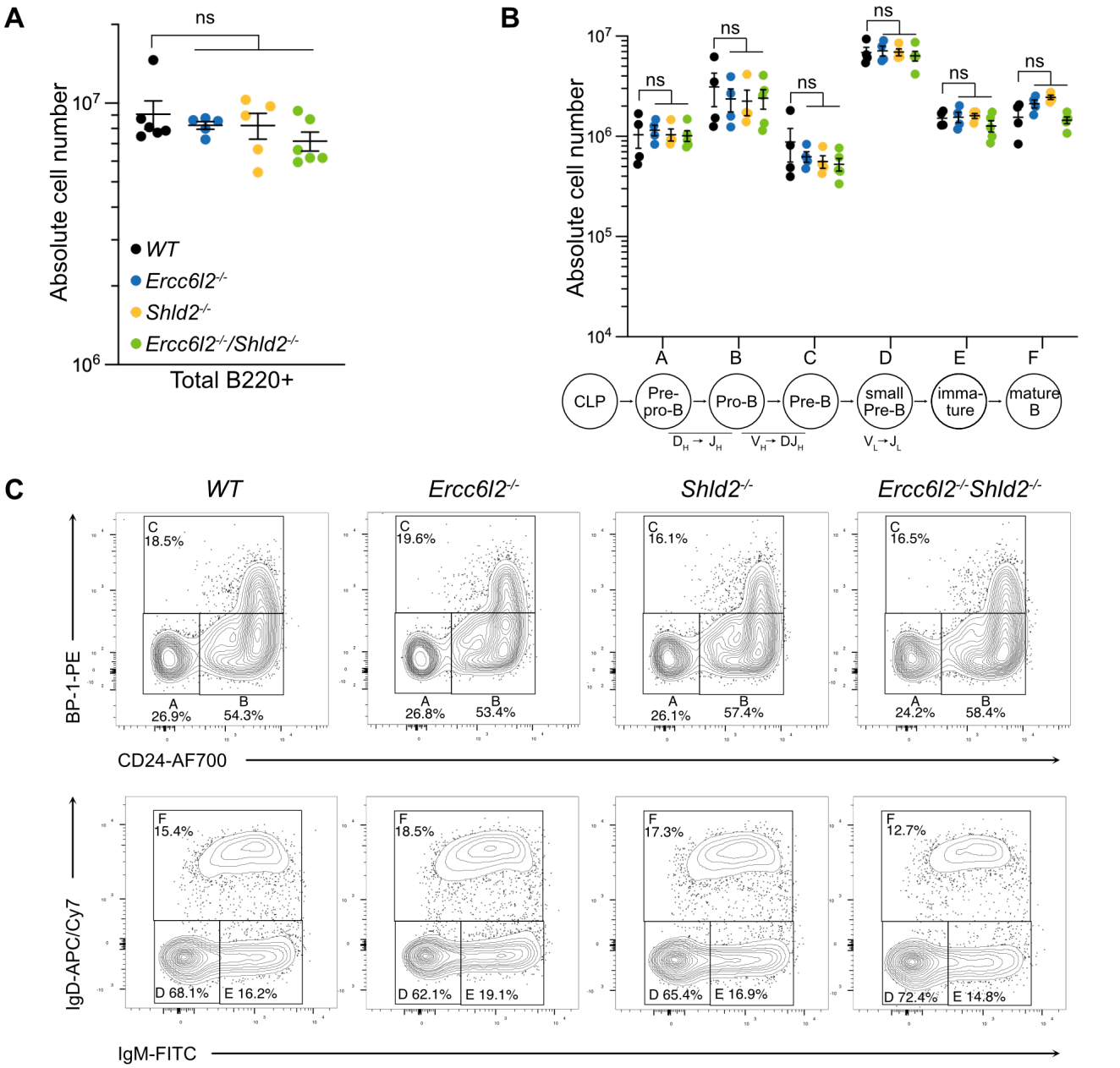


Figure 4.6: **Combined loss of *Ercc6l2* and *Shld2* does not affect B cell development. A** Absolute number of $B220^+$ B cells in the bone marrow (one femur and one tibia). **B** Absolute number of $B220^+$ B cell precursors in the bone marrow (one femur and one tibia). **C** Representative flow cytometry plots used to quantify B cell development in the bone marrow, gating on $B220^+$ $CD43^+$ for Hardy fractions A, B, C, and on $B220^+$ $CD43^-$ for Hardy fractions D, E, F. For all quantitative graphs, n= 5-6 mice per genotype, where each data point is a single mouse and significance was determined by an unpaired two-tailed t-test (mean ± SEM). ns, not significant

To better understand the interplay between ERCC6L2 and 53BP1 during CSR, mature splenic B cells from *Ercc6l2*$^{-/-}$, *53bp1*$^{-/-}$, and *Ercc6l2*$^{-/-}$/*53bp1*$^{-/-}$ mice were isolated and stimulated with cytokines *ex vivo* to analyse their ability to undergo antibody isotype switching.

In line with previously published work [102], *Ercc6l2*$^{-/-}$ B cells had a 2-fold reduction in CSR compared to wild-type controls (Fig. 4.5a, c). As expected, *53bp1*$^{-/-}$ mice had a severe CSR defect. Interestingly, *Ercc6l2*$^{-/-}$/*53bp1*$^{-/-}$ cells showed complete epistasis with 53bp1$^{-/-}$ cells, with isotype switching rates being fully equivalent. Cells were also stained with CellTrace Violet (CTV), which allows for tracking of cell proliferation through the dilution of the dye at every cell division. Proliferation profiles of B cells were similar for all genotypes, suggesting that any observed CSR defects were unlikely to be caused by differences in proliferation or cell survival (Fig. 4.5b, d). These findings confirm that ERCC6L2 is involved in NHEJ during CSR and that ERCC6L2's participation critically depends on 53BP1.

Next, I proceeded to interrogate the relationship of ERCC6L2 and the shieldin complex through experiments using *Ercc6l2*$^{-/-}$, *Shld2*$^{-/-}$, and *Ercc6l2*$^{-/-}$/*Shld2*$^{-/-}$ mice. As expected from two factors uninvolved in V(D)J recombination, combined loss of *Ercc6l2* and *Shld2* did not affect lymphocyte development. The total number of B cells in the bone marrow (Fig. 4.6a), as well as the individual B cell precursor populations (Fig. 4.6b-c) were completely unaffected. The total number of B cells in the spleen (Fig. 4.7a), as well as individual splenic B cell subpopulations were also comparable to wild-type controls (Fig. 4.7b-c).

The shieldin complex is a DSB-responsive adapter for the CST complex, facilitating DNA polymerase ζ-dependent fill-in of single-stranded overhangs, thus enabling DSB repair via NHEJ [25]. ERCC6L2's role in CSR is less well understood. However, it is thought to somehow promote directional end-joining to enable productive (i.e. deletional) repair of the immunoglobulin locus [102]

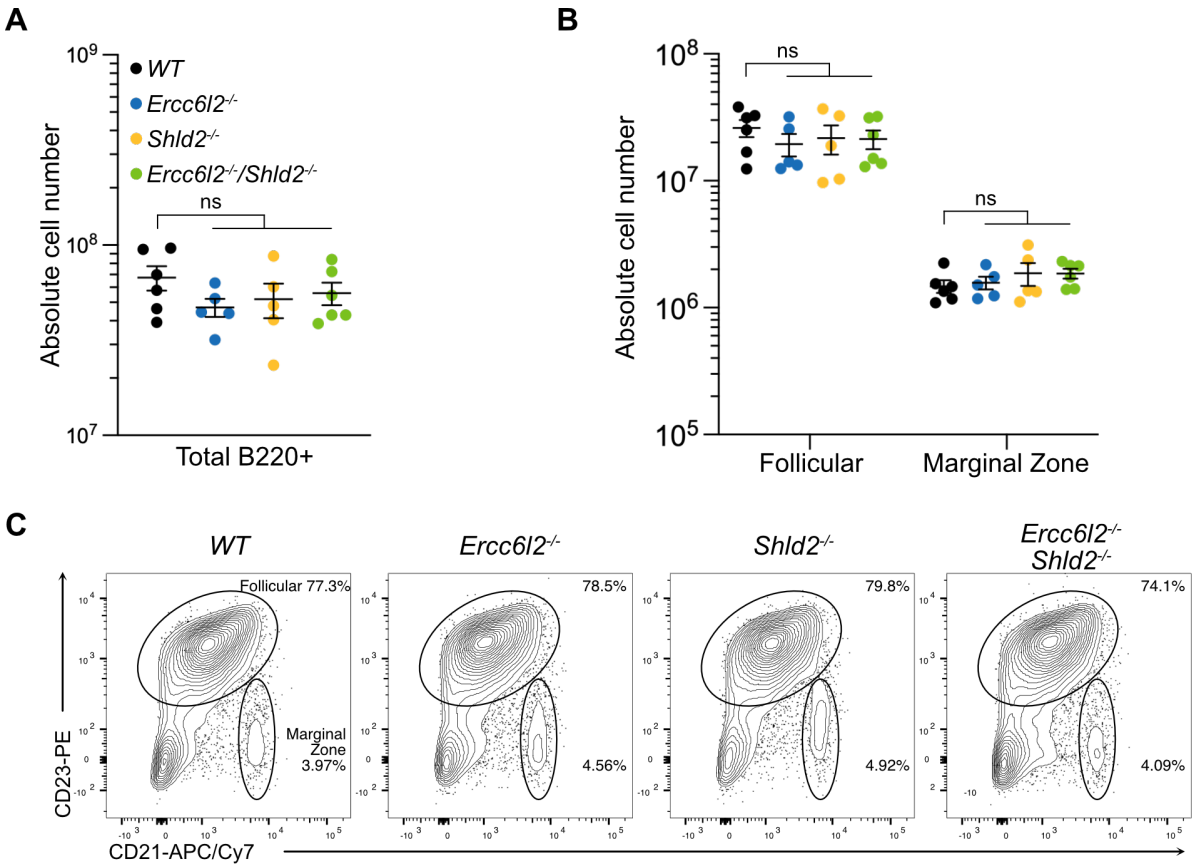


Figure 4.7: **Combined loss of *Ercc6l2* and *Shld2* does not affect mature splenic B cells. A** Absolute number of B220$^+$ B cells in the spleen. **B** Absolute number of follicular and marginal zone B cells in the spleen. C Representative flow cytometry plots showing B220$^+$ CD19$^+$ cells used to quantify mature B cell populations in the spleen. For all quantitative graphs, n= 5-6 mice per genotype, where each data point is a single mouse and significance was determined by an unpaired two-tailed t-test (mean ± SEM). ns, not significant

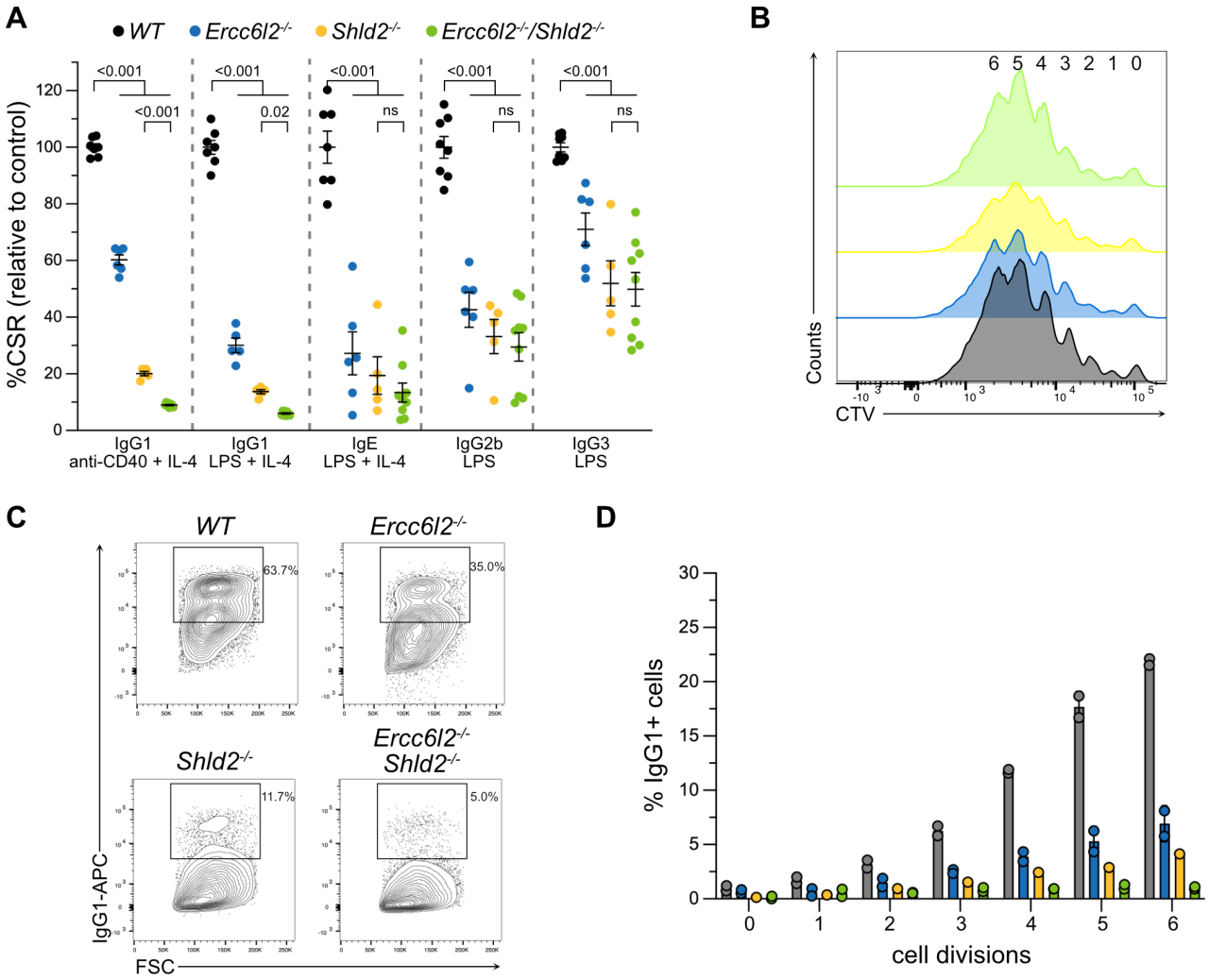


Figure 4.8: **ERCC6L2 supports NHEJ during CSR independently of the shieldin complex. A** Splenic B cells cultured with indicated cytokines for 96 h and stained for IgG1, IgE, IgG2b, or IgG3. (n= 5-6 mice per genotype, where each data point represents a single mouse). Isotype switching frequency was normalized to two wild-type mice for each experiment. Significance was determined by a one-way ANOVA with Tukey's correction (mean ± SEM). ns, not significant **B** Representative flow cytometry plots of CTV dilution in purified splenic B cells stimulated with LPS and IL-4 for 96 h. **C** Representative flow cytometry plots showing splenic B cells cultured with anti-CD40 and IL-4 for 96 h and stained for IgG1. **D** Percentage of $IgG1^+$ B cells in each cell division as determined by proliferation-associated CTV dilution (n= 2 mice per genotype, where each data point is a single mouse).

To investigate if ERCC6L2 exerts this function as part of the 53BP1-RIF1-shieldin-CST axis or if it operates independently, I again isolated mature splenic B cells from mice and stimulated them *ex vivo* to undergo antibody isotype switching. As previously published [25], *Shld2*$^{-/-}$ B cell had a 5-fold reduction of cells successfully switching their antibody isotype (Fig. 4.8a,c). Interestingly, combined loss of *Ercc6l2* and *Shld2* caused a near total collapse of CSR - equivalent to loss of *53bp1*. Again, this was not caused by any proliferation or survival defects (Fig. 4.8b, d).

This suggests that ERCC6L2 and the shieldin complex have independent, non-redundant functions in the repair of DSBs during CSR. Together with the complete epistasis seen with 53BP1, these results would therefore suggest that ERCC6L2 function in CSR critically depends on 53BP1, but that it operates in a distinct, parallel pathway to shieldin.

4.3 Concluding remarks

Since its classification as an accessory NHEJ factor, an involvement of ERCC6L2 in lymphocyte development seemed highly likely [90, 101–103]. To produce the vast repertoire of different antigen binding receptors our adaptive immune system depends on, programmed DSBs are utilized to generate diversity during V(D)J recombination and CSR. These programmed DSBs are subsequently repaired via NHEJ. A previous publication suggest that ERCC6L2 is involved in both V(D)J recombination and CSR [102].

In this chapter, I thoroughly examined the lymphocyte development of *Ercc6l2*-deficient mice. However, I did not observe any defects in their B and T cell development. Even when combined with *53bp1*-deficiency - which should exacerbate any underlying problems with V(D)J recombination - lymphocyte development remained unaffected. While recombination efficiency was reduced in the *in-vitro* extra-chromosomal V(D)J recombination assay in *Ercc6l2*$^{-/-}$ *v-Abl* pre-B

cells, these data should be interpreted cautiously due to the artificial nature of the assay and the contradictory findings *in vivo*. Taken together, these results suggest that ERCC6L2 is not required for V(D)J recombination or lymphocyte development under physiological conditions.

In contrast, during CSR, ERCC6L2 plays a integral part in the proper repair of DSBs facilitating efficient antibody isotype switching. This function is critically dependent of 53BP1, with combined loss of both factors showing complete epistasis. Intriguingly, combined loss of *Ercc6l2* and *Shld2* caused a near-complete collapse of CSR - a defect strikingly similar to loss of *53bp1*. ERCC6L2 must therefore support CSR independently of the shieldin-CST-mediated fill-in of single-stranded DNA overhangs. How ERCC6L2 may function molecularly, is investigated in Chapter 5.

However, much like the shieldin complex, the differential requirement for ERCC6L2 between V(D)J recombination and CSR is likely to be due to the different DNA end structures at the break. Whilst V(D)J recombination involves (almost) blunt DSBs that can be re-ligated directly using NHEJ, AID-dependent breaks contain single-strand DNA tails of varying lengths that require further processing before ligation [77]. Ultimately, these findings refine our understanding of the molecular mechanisms of DNA repair during CSR, suggesting ERCC6L2 acts in a 53BP1-dependent, but distinct, parallel pathway to shieldin to promote effective isotype switching.

Chapter

INVESTIGATING THE MOLECULAR FUNCTION OF ERCC6L2

5

5.1 Introduction

Since its discovery ERCC6L2 has been attributed to TC-NER [99], the FA pathway [88, 90], and - most convincingly - DSB repair in the form of NHEJ [90, 101–103]. However, how ERCC6L2 is supposed to contribute to these processes mechanistically, remains largely unexplored. Towards this goal, finding its interacting partners could shed some light on its mechanism of action. Previous co-immunoprecipitation-mass spectrometry (coIP-MS) experiments suggest that ERCC6L2 interacts with a wide array of proteins, with the top hit being DNA-PKcs [99]. However, this work has been done with ectopically overexpressed ERCC6L2, which carries the risk of detecting non-specific and low-affinity interactions. Other work has implicated CYREN/MRI and the KU heterodimer to interact with endogenously tagged ERCC6L2, but no unbiased approach has been used to explore its interactome [102].

In this chapter, I carry out cell survival assays to interrogate ERCC6L2's contribution the response

to several genotoxic agents. Additionally, the endogenous *Ercc6l2* locus is tagged in CH12-F3 cells and the cell line is subsequently used to perform coIP-MS analysis of the endogenous ERCC6L2 interactome. Using AlphaFold3 to predict the interactions with the identified interactors, I subsequently validate the interactions using an IP-immunoblotting approach. Finally, I show that loss of both ERCC6L2 and its interacting partner lead to an equivalent, epistatic CSR defect, suggesting that they work together to facilitate antibody isotype switching.

5.2 Results

5.2.1 ERCC6L2 is a NHEJ factor

To investigate ERCC6L2's contribution to the response to genotoxic agents, *Ercc6l2-/-* cells were generated in two different backgrounds - mouse embryonic fibroblasts (MEFs) and the murine B lymphoma cell line CH12-F3. In line with previous work, *Ercc6l2*-deficient cells were sensitive to etoposide, although not to the same extent as cells lacking the NHEJ core component XRCC4 (Fig. 5.1a-b) [101]. Normally , topoisomerase II (TOP2) removes topological stress on DNA (such as catenanes, supercoiling, or knots) by generating a transient DSB that subsequently gets re-ligated [147, 148]. However, etoposide "traps" TOP2 in a usually transient cleavage complex intermediate (TOP2cc), generating an irreversible TOP2-DNA-protein crosslink and a DSB, leading to etoposide hypersensitivity in cells defective in DSB repair.

Considering its putative ATPase/helicase function, I wondered if ERCC6L2 has a role in resolving DNA secondary structures to facilitate DNA repair. An example of DNA secondary structures are G-quadruplexes (G4). G4s are characterized by the formation of planar guanine quartets through Hoogsteen hydrogen bonds [149]. They generally form in G-rich regions of the genome, such as telomeres, promoters, and, notably, immunoglobulin switch regions [150]. G4s can be

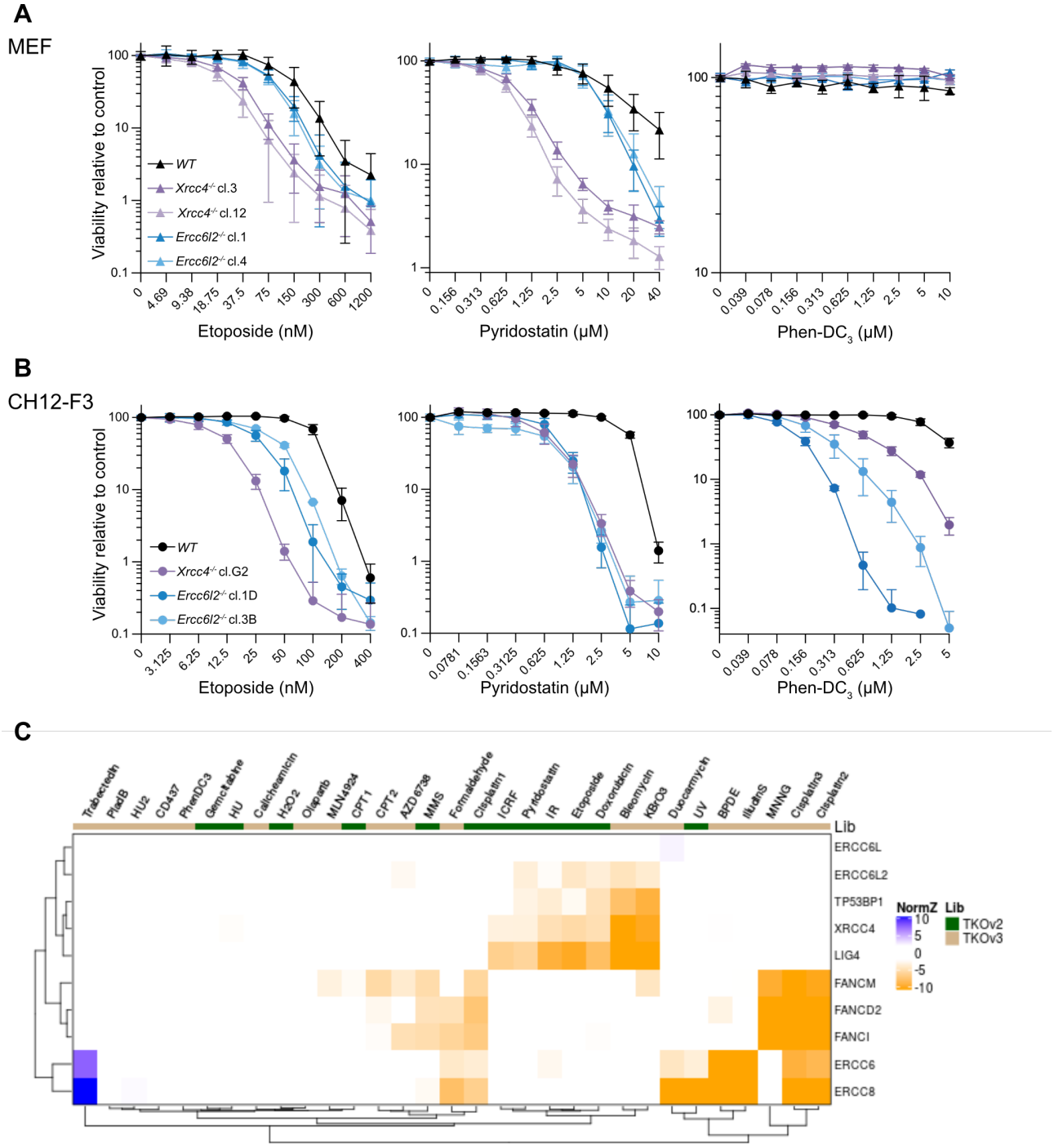


Figure 5.1: **ERCC6L2 is a NHEJ factor. A** Survival of the indicated MEF cell lines grown for 7 days in the presence of indicated doses of etoposide, pyridostatin, or Phen-DC$_3$ (n= 3 biological replicates for etoposide and pyridostatin, n= 1 for Phen-DC$_3$; with 3 technical replicates each, mean ± SD). **B** Survival of the indicated CH12-F3 cell lines grown for 5 days in the presence of indicated doses of etoposide, pyridostatin, or Phen-DC$_3$ (n= 3 biological replicates, with 3 technical replicates each, mean ± SD). **C** Heatmap showing NormZ scores of indicated genes in response to indicated genotoxic agents. Plot was obtained from `https://durocher.shinyapps.io/GenotoxicScreens`.

bound and stabilised by small-molecule G4 binders, such as pyridostatin or Phen-DC_3, leading to replication- and transcription-dependent DNA damage [151]. When cells were treated with pyridostatin and Phen-DC_3, a curious pattern emerged: *Ercc6l2*$^{-/-}$ MEFs showed an intermediate sensitivity compared to *Xrcc4*$^{-/-}$ MEFs (Fig. 5.1a). These results mirror the response to etoposide and therefore fit in well with recent insights showing that pyridostatin cytotoxicity involves TOP2 trapping [101]. Curiously, in CH12-F3 cells, *Ercc6l2*- and *Xrcc4*-deficiency lead to an equivalent sensitisation, suggesting that ERCC6L2 is as important as core NHEJ factors in response to G4-induced DSBs (Fig. 5.1b). A similar pattern could be seen in the cellular response to Phen-DC_3: MEF cell lines of all genotypes were completely resistant to Phen-DC_3 at the tested concentrations (Fig. 5.1a). In contrast, not only were *Ercc6l2*-deficient CH12-F3 cells hypersensitive to Phen-DC_3, their sensitivity exceeded that seen in *Xrcc4*-deficient cells (Fig. 5.1b). Whilst the complete lack of an effect after Phen-DC_3 in MEFs could be easily explained by - for example - a lack of drug uptake and/or efficient drug removal using efflux pumps, the different epistasis of ERCC6L2 and XRCC4 depending on cell line is confounding. Little is known about cell- and tissue-type specific DNA repair preferences and even less about about different effects of individual proteins depending on cell type.

However, it is clear from these data, that ERCC6L2 behaves like a NHEJ factor in response to genotoxic agents. This is further exemplified by the clustering of ERCC6L2 with other NHEJ factors, such as XRCC4, LIG4, or 53BP1, in the cellular response to genotoxins (Fig. 5.1c) [101]. Factors involved in TC-NER (e.g. ERCC6, ERCC8) and factors involved in the FA pathway (e.g. FANCD2, FANCI, FANCM) form their own clusters, discounting ERCC6L2's contribution to these pathways.

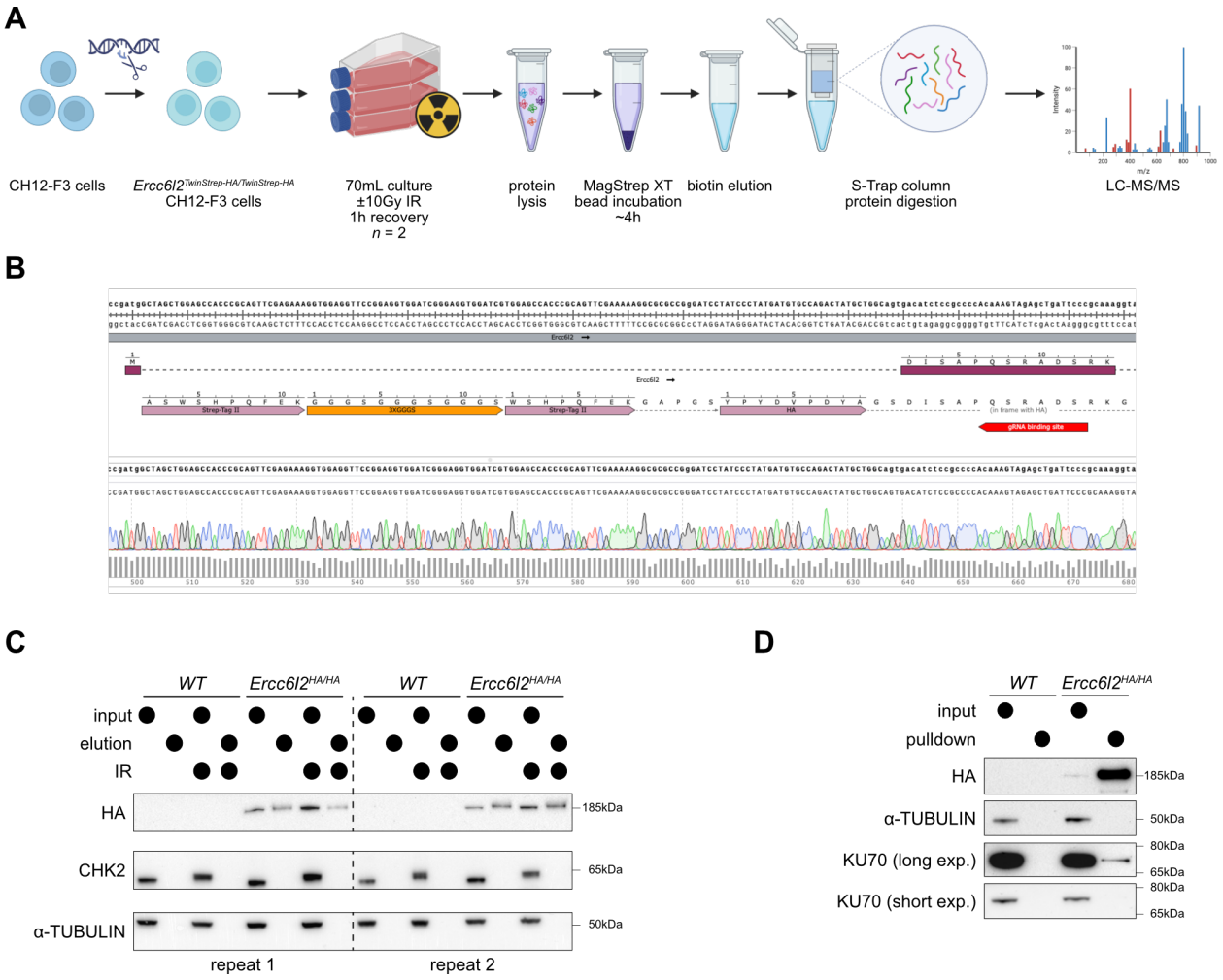


Figure 5.2: **Affinity purification-MS reveals the ERCC6L2 interactome. A** Schematic representation of TwinStrep-HA-ERCC6L2 purification and proteomic analysis strategy. Briefly, the endogenous Ercc6l2 locus was bi-allelically tagged with a TwinStrep-HA tag. 70 mL cell cultures were either mock-treated or irradiated (10 Gy) and lysates were collected following a 1 h recovery. ERCC6L2 was purified using MagStrep-XT beads and eluted in biotin. The elutions were then captured and tryptic-digested on S-Trap columns and the resulting peptides were analysed using LC-MS^2. This schematic was generated using BioRender. **B** SnapGene screenshot showing successful bi-allelic integration of the TwinStrep-HA tag in the endogenous *Ercc6l2* locus. **C** B Immunoblot analysis of CH12-F3 cell lysates from ERCC6L2 purification experiments used for subsequent analysis via LC-MS^2. **D** Immunoblot analysis of CH12-F3 cell lysates from an ERCC6L2 purification. Blot is representative of n= 2 biological replicates.

5.2.2 Proteomics reveal the ERCC6L2 interactome

To elucidate which proteins interact with ERCC6L2, an affinity purification-mass spectrometry approach was developed (Fig. 5.2a). Several lines of evidence indicate that ERCC6L2 is expressed at exceedingly low levels, chief among them the failure to produce an antibody capable of detecting endogenous ERCC6L2 via immunoblotting [102]. Because of this, it seemed essential to work with the endogenous protein instead of relying on ectopic overexpression that would most likely disturb normal stoichiometry and behaviour. To this end, I inserted a TwinStrep-HA tag on the endogenous *Ercc6l2* locus - which corresponds to the N-terminus of the protein - using CRISPR/Cas9 in CH12-F3 cells. Successful bi-allelic insertion in an isogenic clone was confirmed using Sanger sequencing (Fig. 5.2b). Due to ERCC6L2's extremely low abundance, successful purification required lysates isolated from $\approx 1.2 \times 10^8$ cells. ERCC6L2 was isolated from lysates using Strep-Tactin affinity purification and then subjected to liquid chromatography-tandem mass spectrometry (LC-MS^2), over two independent biological replicates. Successful affinity purification was confirmed by immunoblotting of a small aliquot of the elution sent for proteomic analysis (Fig. 5.2c). Considering the shieldin-CST interaction during CSR depends on ATM-dependent DNA damage signalling [25], I wondered if the ERCC6L2 interactome also changed when DNA damage is sensed. To stimulate a DNA-damage response, cells were X-ray irradiated. Irradiation causes a massive accumulation of DSB and subsequently triggers the DDR. The induction of a DNA-damage response was confirmed by the (ATM-dependent) phosphorylation of CHK2 via immunoblotting (Fig. 5.2c).

ERCC6L2 was readily detectable via mass spectrometry and the only other proteins exclusively and reproducibly detectable in the experimental samples were KU70, KU80, and MRI/CYREN (Table5.1. These results recapitulate previous work also showing that ERCC6L2 interacts with the KU heterodimer and MRI [102]. Additionally, this interaction seemed to be DNA damage-independent, as treatment of cells with ionizing radiation did not change the interactome. The

Protein symbol	Coverage [%]	Total unique spectra							
		R1_ TS-HA-E6L2	R1_ TS-HA-E6L2 +IR	R2_ TS-HA-E6L2	R2_ TS-HA-E6L2 +IR	R1_ CTR	R1_ CTR +IR	R2_ CTR	R2_ CTR +IR
ERCC6L2	13	14	21	12	16	0	0	0	0
KU80	6.3	3	4	3	3	0	0	0	0
KU70	5.9	3	4	3	1	0	0	0	0
MRI	14	1	2	2	2	0	0	0	0

Table 5.1: ERCC6L2 interactome detected via LC-MS^2. Interacting proteins were determined as in Fig. 5.2 (n=2 independent experiments per genotype and condition). Peptide identifications were accepted if they could be established at >95.0% probability, and protein identifications were accepted if they could be established at >99% probability. Proteins were accepted if they could not be detected in any of the control samples and at least 2 unique spectra were detected in the experimental samples.

interaction with KU was subsequently validated through immunoblotting of ERCC6L2 purification (Fig. 5.2d). Unfortunately, the MRI interaction could not be confirmed endogenously, as the only available MRI antibody cross-reacts with the Ig light chain expressed in CH12-F3 cells and the light chain and MRI have the same molecular weight.

5.2.3 ERCC6L2 interacts with KU via MRI

To study the molecular basis of the interaction of ERCC6L2, MRI, and KU on DNA, multiple rounds of structure modelling using AlphaFold 3 (AF3) were performed [122]. AF3 modelling predicted the formation of a protein complex with several high confidence (pLDDT score >90) domains and protein-protein interactions (Fig. 5.3a, c). Predicted Aligned Error (PAE) analysis predicts two high-confidence (expected position error < 15Å) protein-protein interactions for MRI (Fig. 5.3b):

Firstly, AF3 modelling predicted MRI to interact with KU80 by forming several hydrogen bonds and salt bridges between D106, D109, D116, Q119, and S143 on Ku80 and R11, P14, and W16 on MRI (Fig. 5.3e). This is consistent with a previously reported *Xenopus laevis* crystal structure, where the MRI and Ku70-80 interface is formed between MRI's KBM (Ku-Binding Motif) and

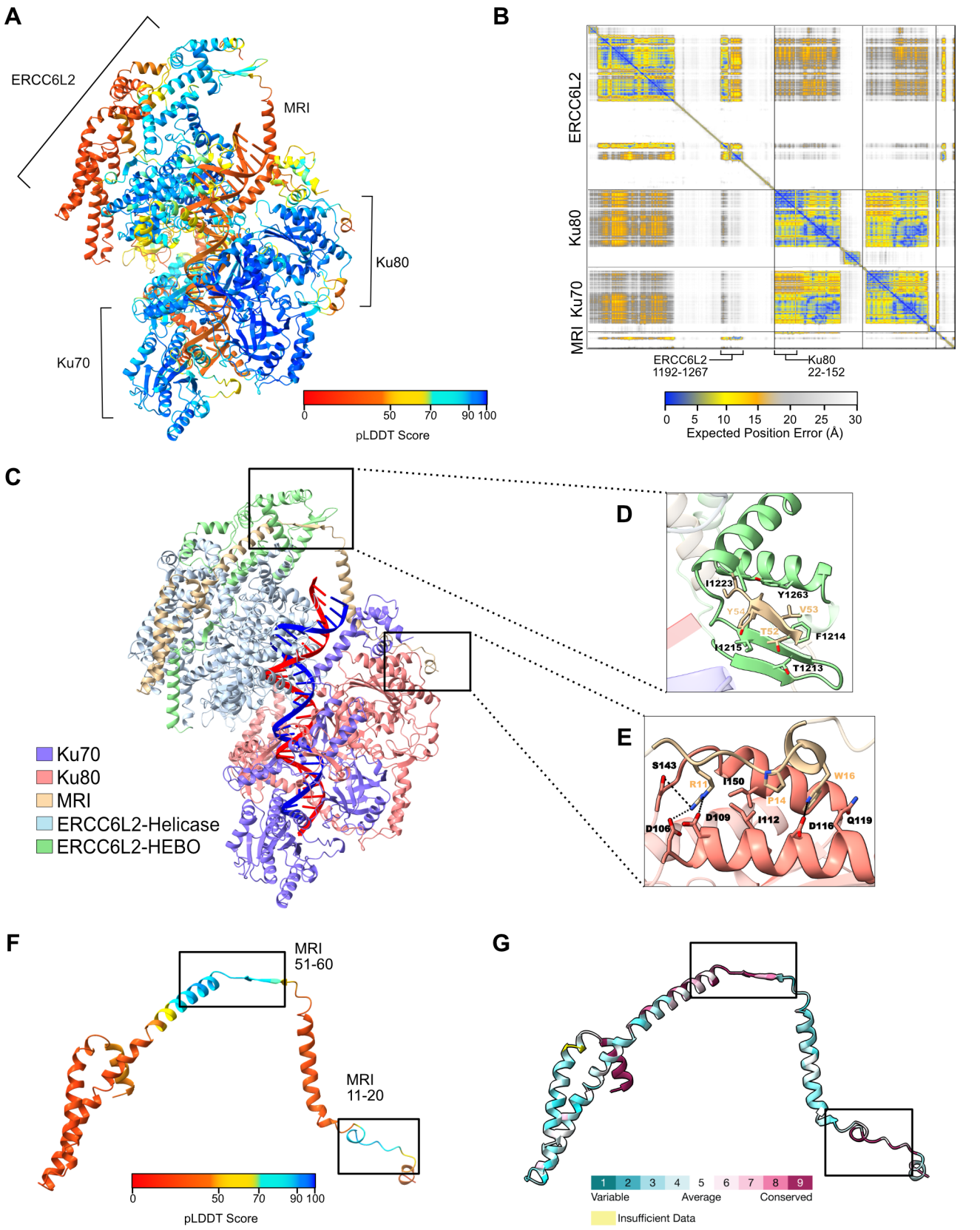


Figure 5.3: **AlphaFold3 model of the ERCC6L2-MRI-KU-complex.** **A** Top ranking ERCC6L2-MRI-KU-DNA complex predicted structure coloured according to pLDDT score. **B** PAE plot showing the expected position error of the highest-ranked predicted interactions of the ERCC6L2-MRI-KU-DNA complex. **C** Top ranking ERCC6L2-MRI-KU-DNA complex predicted structure coloured by individual subunits, in which MRI highlighted in pLDDT colour. DNA double-helix can be seen in blue and red. **D** Zoomed-in ERCC6L2-MRI binding site, interacting residues are labelled. **E** Zoomed-in Ku80-MRI binding site, interacting residues are labelled. **F** Predicted MRI strucure coloured according to pLDDT score and **G** ConSurf score. AF3 modelling was performed by Dr. Yueru Sun and Prof. Ross Chapman.

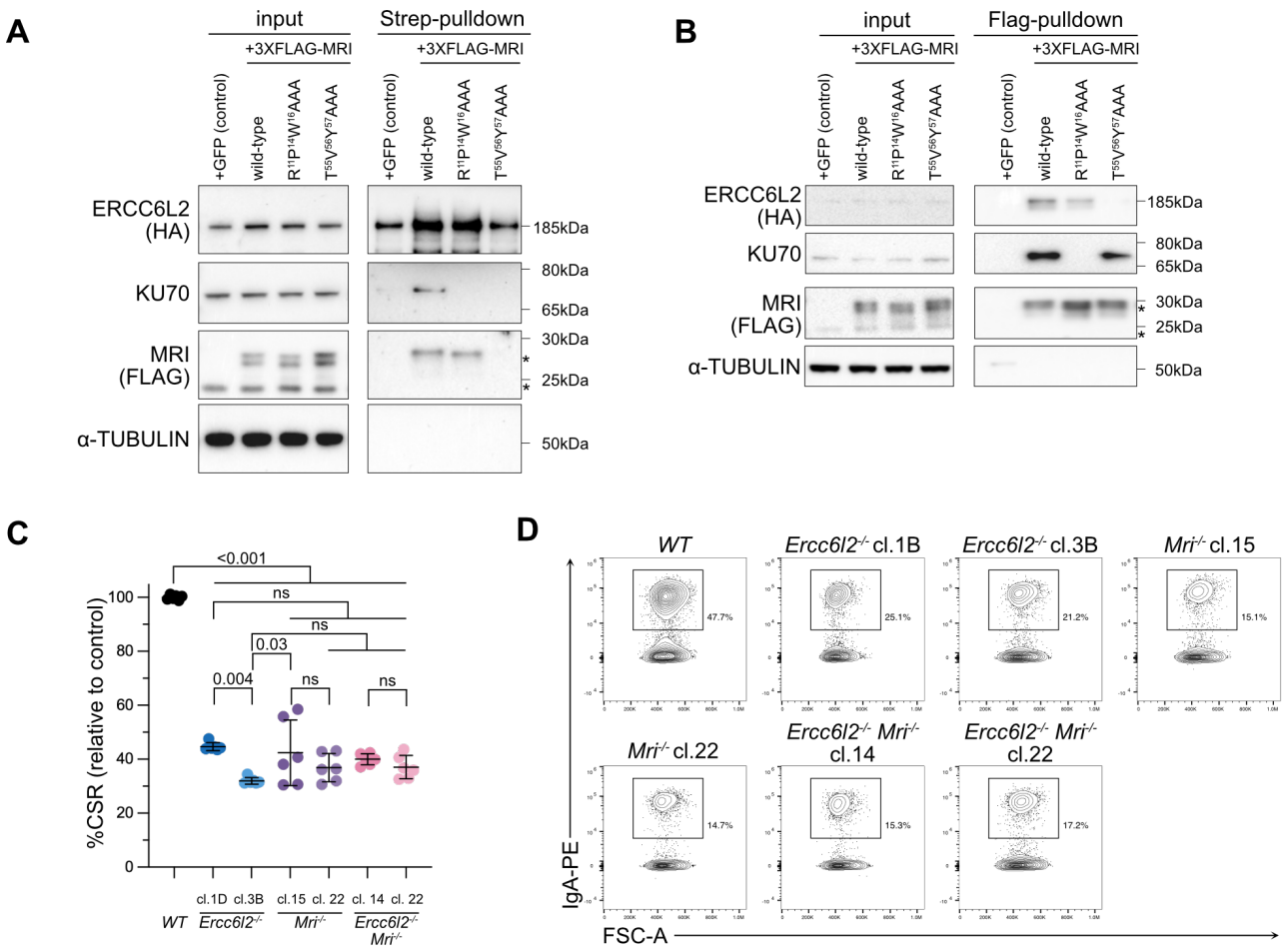


Figure 5.4: **ERCC6L2 interacts with KU via MRI to support NHEJ during CSR. A** Western blot analysis of CH12-F3 lysates from TwinStrep-ERCC6L2 pulldown experiments. Blot is representative of n= 2 biological replicates. **B** Western blot analysis of CH12-F3 lysates from FLAG-MRI pulldown experiments. Blot is representative of n= 2 biological replicates. **C** IgM-to-IgA CSR frequencies in *Ercc6l2* and *Mri* knockout CH12-F3 cells normalized to wild-type cells (n= 3 biological replicates, with two technical replicates per experiment). Significance was determined by a one-way ANOVA with Tukey's correction (mean ± SD). **D** Representative flow cytometry plots for IgM-to-IgA CSR.

Ku80's von Willebrand domain (vWD) [152]. Additionally, mutagenesis of these residues has been shown to abolish the interaction between KU and MRI [112].

Secondly, AF3 modelling predicted a striking novel interaction between MRI and ERCC6L2's HEBO domain of unknown function (Fig. 5.3d). Mediated by T52, V53, and Y54 on MRI and T1213, F1214, I1215, I1223, and Y1263 on ERCC6L2, MRI forms an intermolecular β-sheet with two β-strands of ERCC6L2. Further supporting the validity and structural relevance of these interaction is the high conservation score of the interacting residues on MRI according to ConSurf [153].

To test these *in silico*-predicted protein-protein interactions *in vitro*, I subsequently generated *Mri* $^{-/-}$ *Ercc6l2*$^{HA/HA}$ CH12-F3 cells to ablate endogenous MRI expression. Cells were then virally reconstituted with wild-type or mutant 3XFLAG-MRI, or a negative control (GFP). The tested mutants included: MRI with residues R11, P14, and W16 substituted to alanine - which are predicted to mediate the interaction with Ku80 - and MRI with residues T55, V56, and Y57 substituted to alanine - which are predicted to mediate the interaction with the ERCC6L2-HEBO domain.

Cell lysates were collected from the reconstituted cells and ERCC6L2 was subsequently isolated by affinity purification using StrepTactin-XT beads and complex formation was analysed by immununoblotting (Fig. 5.4a). This revealed a number of interesting insights: Firstly, the interaction between ERCC6L2, MRI, and KU could be confirmed. Secondly, without MRI (i.e. in the *Mri*$^{-/-}$ cells recomplemented with GFP), the interaction between KU and ERCC6L2 was completely abolished. Thirdly, the interaction between KU and ERCC6L2 fully depended on both the previously identified MRI/KU interaction and the newly discovered MRI/ERCC6L2-HEBO interaction. These observations were confirmed by isolation of MRI by affinity purification using FLAG-beads (Fig. 5.4b). Taken together, these data suggest that ERCC6L2 interacts with KU via MRI, with MRI anchoring the interaction between the other two proteins.

To test the functional relevance of this interaction, *Ercc6l2*$^{-/-}$, *Mri* $^{-/-}$, and *Ercc6l2*$^{-/-}$*/Mri* $^{-/-}$ CH12-F3 cells were generated and stimluated to undergo IgM-to-IgA class switching. Mirroring results obtained from *Mri* $^{-/-}$ mice [109], loss of MRI lead to a 2.5-fold reduction in CSR - an effect highly similar to what was seen in *Ercc6l2*$^{-/-}$ cells (Fig. 5.4c-d). Combined loss of MRI and ERCC6L2 was fully equivalent with loss of either protein, suggesting full epistasis between the two in the context NHEJ during CSR.

5.3 Concluding remarks

This chapter investigates the mechanistic role of ERCC6L2, a protein previously linked to multiple DNA repair pathways including transcription-coupled nucleotide excision repair (TC-NER), the Fanconi anemia (FA) pathway, and most strongly, non-homologous end joining (NHEJ). Despite these associations, its exact molecular function has remained unclear.

Here, I was able to confirm that knockout of *Ercc6l2* sensitises cells to etoposide, a chemotherapeutic that traps topoisomerase II and induces DSBs. Furthermore, I found that *Ercc6l2*-deficient cells were also sensitive to G-quadruplex (G4) stabilisers (i.e. pyridostatin and Phen-DC_3), mirroring results in *Xrcc4*-deficient cells, which further supports a role in DSB repair. Curiously, the results from experiments with G4 stabilisers are inconsistent between cell lines, with the results in MEFs supporting an accessory role in cNHEJ, whereas results in CH12-F3 cells could hint at a stronger requirement for ERCC6L2 in the presence of G4s. Further work is needed to untangle this discrepancy, e.g. by expanding the experiment to more cell lines or by confirming if ERCC6L2 can indeed unwind G4-forming DNA *in-vitro*. Notably, immunogloblin switch regions (and telomeres) have repetitive, G-rich sequences [154, 155]. This makes it tantalising to speculate that the ATPase/translocase ERCC6L2 might have a role in "melting" these regions to allow access for downstream repair factors. Finally, a genotoxic clustering analysis using publicly available data groups ERCC6L2 alongside cNHEJ factors (e.g. XRCC4, LIG4), and separates it from TC-NER and the FA pathway. Taken together, these results firmly support a role for ERCC6L2 in NHEJ, whilst discounting likely meaningful contributions to either TC-NER, or the FA pathway.

I also developed an unbiased affinity purification-mass spectrometry approach with endogenously tagged and expressed ERCC6L2. This identified KU70, KU80, and MRI/CYREN as interactors

of ERCC6L2. This was consistent across untreated and irradiated conditions, suggesting that these interactions were not dependent on DNA damage signalling. Due to antibody limitations, the interaction could not be confirmed for endogenous MRI. However, the interaction with KU70 (and by inference KU80) could be confirmed via immunoblotting.

Using AlphaFold3 modelling, a complex involving ERCC6L2, MRI, and the KU heterodimer was predicted with high confidence. AF3 predicted a known interaction between KU80 and MRI via its KU-binding motif. Additionally, a novel interaction between MRI and the ERCC6L2 via its HEBO domain was predicted, which was supported by high evolutionary conservation of the implicated residues. Functional mutagenesis experiments were able to confirm that loss of either the MRI/KU or the MRI/ERCC6L2 interface disrupted the complex - showing that MRI is essential as a bridging factor between ERCC6L2 and KU. Finally, using IgM-to-IgA class switching as a functional readout, I showed equivalent CSR defects in cells deleted for *Ercc6l2*, *Mri*, or both factors. This epistatic relationship supports the functional relevance of the identified complex in NHEJ during CSR.

In summary, my results confirm ERCC6L2's role as a bona fide NHEJ factor that forms a functionally important complex with KU70/80, supported by bridging interactions mediated by the adaptor protein MRI/CYREN. This interaction is critical for resolving DSBs during class switch recombination, shedding light on the molecular mechanism of ERCC6L2, and MRI, in adaptive immunity and genome maintenance. Yet, many open questions still remain.

Firstly, it is unclear how ERCC6L2 is recruited to DSBs - the KU heterodimer has a high affinity for broken DNA ends [34, 156], but ERCC6L2 purification experiments clearly show that only a subfraction of the available KU complex is bound to ERCC6L2 (Fig. 5.2d). Recruitment is also unlikely to be governed by DSB repair signalling (such as phosphorylation by ATM), as irradiation of cells with ionizing radiation did not change the ERCC6L2 interactome. Instead, I propose that

MRI is likely vital for its regulated recruitment. This is supported by the observation that the ERCC6L2/KU interactions are lost when MRI is absent or mutated at the its interface-relevant residues.

However, the function of MRI itself is contested. Arnoult *et al.* found that MRI *inhibits* cNHEJ specifically in S and G_2 phase of the cell cycle - hence their renaming of MRI to CYREN (cell cycle regulator of NHEJ) [112]. They also report that MRI preferentially inhibited NHEJ at breaks with overhangs, instead protecting them and favouring repair via HR. These results starkly contradict findings from this thesis, where MRI supports NHEJ during CSR. However, notably, CSR is thought to occur mainly in early G_1, not S/G_2 phase [157].

Meanwhile, Hung *et al.* found that MRI *promotes* cNHEJ. They also generated a *Mri* knockout mouse model that showed striking similarities with our *Ercc6l2* knockout mice - *Mri* $^{-/-}$ mice are viable, fertile, have normal lymphocyte development, but show an intermediate CSR defect and sensitivity to ionizing radiation. Mechanistically, they report that MRI forms multimeric complexes with diverse DDR factors, with the N-terminus of MRI interacting with cNHEJ factors, and the C-terminus with MRN and ATM. They also find that ERCC6L2 interacts with MRI, but do not discuss this further - notably, MRI interacts with ERCC6L2 through residues in the middle of the protein, not through either terminus. Data generated for this thesis certainly corresponds better to these results, than to the findings that MRI inhibits cNHEJ. It is possible that - much like KU - only a subfraction of MRI binds ERCC6L2 and MRI exerts different function depending on cellular contexts such as cell cycle phase and DSB end structure. It would be interesting to compare the stoichiometry of total MRI vs ERCC6L2-bound MRI using affinity purifications of endogenously expressed MRI either by generating a suitable antibody or adding a tag to the the endogenous *Mri* locus.

Secondly, ERCC6L2's exact molecular function remains to be clarified. ERCC6L2 is required for NHEJ during CSR - where DSBs with single-strand overhangs are generated - but not during V(D)J recombination - where DSBs are mostly blunt. This observation, together with ERCC6L2's likely enzymatic function as a DNA translocase, makes it tantalizing to speculate that ERCC6L2 functions by properly positioning KU at the break to facilitate NHEJ. For example, this could involve "sliding" KU to the double-strand-single strand (ds-ss) junction. Alternatively, as mentioned before, the sensitivity of *Ercc6l2*-deficient cells to G4 ligands could suggest that ERCC6L2 "melts" secondary structures to allow for efficient downstream repair. Ultimately, further investigation is needed to define the molecular function of ERCC6L2 in more detail.

Chapter

6 DISCUSSION

6.1 Why does the bone marrow not fail in *Ercc6l2*$^{-/-}$ mice?

ERCC6L2's one defining feature is the fact that bi-allelic mutations cause an inherited bone marrow failure syndrome, termed ERCC6L2 syndrome [89, 93]. ERCC6L2 syndrome is characterised by mild changes of peripheral blood although the bone marrow is severely hypocellular. There is however a high risk of progression to erythroid AML (AML M6), a exceedingly rare subtype of AML with a dismal prognosis. Astonishingly, *ERCC6L2* mutations seem to almost exclusively cause AML M6, making it the first time a germline mutations has been reported to cause a strictly specific form of acute leukaemia [95]. BMF is generally caused by the premature attrition of haematopoietic stem and progenitor cells, leading to hypocellular bone marrow and the collapse of blood production of some (or all) lineages. Curiously, *Ercc6l2*$^{-/-}$ mice do not share this phenotype. There are no changes of HS(P)C cell numbers, and neither do *Ercc6l2*-deficient HSPCs show any overt defects in bone marrow engraftment and blood repopulation in both a primary and secondary bone marrow transplant, meaning *Ercc6l2*-deficient HSPCs are fully functional, even under conditions of high physiological stress.

Initial reports suggested ERCC6L2 might have a role in DNA interstrand crosslink repair, proposing it could be a novel factor in the Fanconi anaemia pathway [88, 90]. Combined lack of aldehyde detoxification and the FA repair pathway is synthetically lethal in some cell lines and causes rapid bone marrow failure in early adulthood in mice [48, 50]. This has been coined the "two tier aldehyde protection model", showing that endogenously produced aldehydes are the source of DNA interstrand crosslinks ultimately causing BMF in Fanconi anaemia patients [158]. Having also confirmed a mild sensitivity of *Ercc6l2*$^{-/-}$ cells to cisplatin - which causes ICLs - I investigated if ERCC6L2 syndrome is a previously unidentified mild form of Fanconi anaemia. *Adh5*$^{-/-}$/*Ercc6l2*$^{-/-}$ mice were generated to test if increased aldehyde levels could precipitate bone marrow failure in absence of *Ercc6l2* - as it was shown to do in *Fancd2*$^{-/-}$ mice [51]. However, no synergy was found between *Adh5* and *Ercc6l2* and haematopoietic stem cell numbers were completely normal, discounting a physiological role of ERCC6L2 in the FA repair pathway. Additionally, this also suggests that whatever drives HSPC attrition in ERCC6L2, it is highly unlikely to be endogenous aldehydes.

Using the "two tier protection model" as an inspiration, I wondered if I could pinpoint the source of DNA damage that requires ERCC6L2 for its repair. To this end, I performed a genome-wide CRISPR screen in primary murine HSPCs cultured *ex vivo* in a wild-type and an *Ercc6l2*-deficient background looking for genetic targets that were selectively essential in either genotype.
So what could be the reason for this major discrepancy between the human and the mouse phenotype? Why does the bone marrow *not* fail in the mouse? There are several possible explanations for this phenomenon:

(1) Mice are an unsuitable model for mild human bone marrow failure syndromes.
The simplest, yet most unsatisfying explanation for this discrepancy is that mice might just be a poor model system for BMF syndromes. Despite their relatively close phylogenetic relationship to humans compared to other model organisms, mouse models do not necessarily replicate all human

disease. A prime example for this are *Atm*$^{-/-}$ mice. Ataxia-telangiectasia mutated (ATM) was identified and named according to the disease it causes: Ataxia-telangiectasia (AT) [159]. One of AT's most characteristic features is an unsteady gait (ataxia) caused by cerebellar atrophy. Yet, *Atm*$^{-/-}$ mice do not share the debilitating neurodegeneration phenotype, instead succumbing to rapidly developing thymic lymphomas [160]. Another example that is perhaps more relevant as it afflicts the same tissue, is *Fancd2*$^{-/-}$ mouse model, in which the bone marrow does not fail spontaneously [125]. Considering the stark discrepancy in the age at disease onset (7 years in FA vs. 18 years in ERCC6L2 syndrome), and in the general disease severity between Fanconi anaemia and ERCC6L2 syndrome, lack of ERCC6L2 might just not have enough impact on the murine haematopoietic system [89, 161]. Yet, closer analysis of *Fancd2*$^{-/-}$ mice does show defects in their HSPC number and function [162]. Meanwhile, ERCC6L2-status has no impact on murine blood development whatsoever - even during serial competitive bone marrow transplants when HS(P)Cs are challenged acutely. This could all be explained by a more robust haematopoietic system in mice compared to humans, or even just their difference in lifespan.

(2) Mice lack the correct stimulus to trigger BMF in absence of ERCC6L2.

Another possibility could be that - unlike humans - mice do not receive enough DNA damage requiring ERCC6L2 for repair to trigger BMF. Once again, the parallel to be drawn is the two-tier aldehyde protection model: *Fancd2*$^{-/-}$ mice don't spontaneously succumb to BMF. Only when the burden of ICLs is increased by additionally removing aldehyde detoxification, do mice succumb to BMF and leukaemia in a rapid manner, phenocopying Fanconi anaemia in humans [50, 51]. So perhaps mice also produce fewer DSBs compared to humans and the BMF phenotype could therefore be revealed by removing the first tier of protection from endogenous DSBs. Yet, the genome-wide CRISPR screen ±ERCC6L2 revealed few to no standout hits that were likely to be synthetically lethal/selectively detrimental in the absence of ERCC6L2. This would argue against a genetic/metabolic non-redundant protective tier against endogenous DSBs(at least in

the context of ERCC6L2). Considering that endogenous DSBs are mostly generated by processes strictly essential for cell survival (i.e. DNA replication, transcription, topological stress release via topoisomerases) and not by metabolites, this is not a very surprising result, as reducing the source of DSBs would also negatively affect cell viability. Whilst it is possible that the CRISPR screen missed hits protecting against DSBs (e.g. through inefficient knockout of the gene), this does seem unlikely.

(3) The DNA repair defect is a red herring.

Finally, it is conceivable that whilst ERCC6L2 does have a role in NHEJ, ERCC6L2 syndrome is not primarily caused by a DNA repair defect: With the exception of ERCC6L2 and LIG4, no other NHEJ factors have been associated with bone marrow failure syndromes [54, 88]. Generally, mutations in NHEJ factors (such as *XLF* or *DCLRE1C* (encoding Artemis)) cause lymphocytopenias and SCID due to the inability to properly carry out V(D)J recombination [56, 74]. Meanwhile, *ERCC6L2* mutations cause pancytopenia with a slight skew towards thrombocytopenia - whereas lymphocytes are not disproportionally affected [89]. These observations suggest a distinct disease mechanism for ERCC6L2 syndrome, separate from other factors in NHEJ. Instead, I propose that ERCC6L2 syndrome may actually be a previously overlooked telomeropathy. This hypothesis is discussed in detail in Section 6.3. Importantly, current mouse models struggle to replicate the phenotypes seen in telomeropathy, likely due to the much longer telomeres found in mice compared to humans, i.e. 20-150kb and 10-14kb respectively, as well as the differing expression patterns of telomerase in between species [63].

6.2 ERCC6L2 as an accessory cNHEJ factor

In this thesis, I was able to recapitulate previous findings that ERCC6L2 has a role in the repair of DSBs via NHEJ using cell survival assays [90, 101–103]. In contrast, I found no compelling

evidence supporting initial reports claiming a function of ERCC6L2 in the FA pathway and/or TC-NER [88, 90, 99]. Intriguingly, *Ercc6l2*-deficient cells were exquisitely sensitive to G4 stabilizers (i.e. pyridostatin and Phen-DC_3), possibly suggesting a role in resolving secondary structures hindering DSB repair via NHEJ. However, I was unable to recapitulate these results in a second cell line and further work is required to ascertain the link between ERCC6L2 and G4 quadruplexes.

A powerful tool to disentangle the contribution of NHEJ factors is the examination of the development and maturation of lymphocytes. This process is intrinsically linked to the successful rearrangement of antigen receptor loci - which itself is dependent on the repair of programmed DSBs via NHEJ [34]. Extensive characterisation of the B and T cell development of *Ercc6l2* knockout mice showed that ERCC6L2 is fully dispensable for lymphocyte development and therefore V(D)J recombination, even in a "sensitised" *53bp1*-deficient background. Meanwhile, lack of ERCC6L2 did cause a substantial CSR defect. This discrepancy between two processes that rely on repair of DSBs via NHEJ is likely due to the distinct structure of breaks being produced.

The opening of the hairpin by Artemis during V(D)J recombination generates a blunt break that requires minimal processing [34, 69]. Meanwhile, during CSR, the processing of uracil generated by AID produces DSBs with single-stranded overhangs of various lengths [77]. These "staggered" breaks require considerable processing before re-ligation is possible. In large parts, this is accomplished by the 53BP1-shieldin-CST pathway that allows for fill-in of the staggered end [25]. Using several genetic mouse crosses, I demonstrated that whilst ERCC6L2 functions downstream of 53BP1, it operates independently of the shieldin complex, suggesting a novel, parallel pathway to enable end ligation. Interestingly, the shieldin complex requires at least 60bp of ssDNA to bind to DNA [163], opening up the possibility that ERCC6L2 operates on staggered DSBs with overhangs shorter than 60bp (Fig. 6.1).

Mass-spectrometric analysis of the interactome of endogenously expressed ERCC6L2 protein revealed MRI and the KU70/80 heterodimer as interacting partners. Subsequent *in-silico* modelling using AlphaFold3 predicted these proteins to form a complex with high confidence and functional mutagenesis of the predicted interfaces confirmed the predictions. Further supporting the functional relevance of the discovered complex was the fully equivalent CSR defect observed in *Ercc6l2*, *Mri* and *Ercc6l2/Mri* knockout cells.

The interaction with the KU heterodimer suggests that ERCC6L2 performs its function as an early step during NHEJ before the ligase complex (LIG4, XRCC4, XLF/PAXX) is assembled. Whilst ERCC6L2's exact function remains to be explored, it is tantalising to speculate that it might assist in the proper positioning of KU on the break by using its enzymatic function as a ATPase/DNA translocase. This may become especially important at complex breaks such as staggered DSBs - possibly in combination with the removal of detrimental secondary DNA structures like G4 quadruplexes.

6.3 ERCC6L2 as a telomere maintenance factor?

Inherited bone marrow failure syndromes can be caused by defects in either ribosome biogenesis, DNA repair, or telomere maintenance [44]. However, as discussed previously, ERCC6L2's function as a NHEJ factor is unlikely to be the explanation for the bone marrow failure seen in ERCC6L2 syndrome. Yet ERCC6L2's participation in NHEJ only under specific circumstances may offer a clue towards the disease aetiology: Whilst ERCC6L2 is completely dispensable for the resolution of *blunt* DSBs generated during V(D)J, it has a crucial function in the resolution of *staggered* DSBs during CSR. Staggered DSBs (i.e. DSBs with single-stranded overhangs) are not only formed at *Igh* switch regions, but are also found at an integral part of all mammalian chromosomes: telomeres.

Mammalian chromosome ends are capped by telomeres - long stretches of TTAGGG repeats ending with a G-rich 3' overhang of 50-200 nucleotides of single-stranded DNA [164, 165]. Telomere ends are generally protected from being recognized as DSBs by the shelterin complex, which primarily functions by sequestering the 3'overhang in a lariat-like confirmation called the t-loop, thereby making the DSB inaccessible to DDR factors [166]. When telomeres are deprotected by the removal of shelterin components, NHEJ causes telomere fusions, causing cell death [167, 168]. So how would ERCC6L2 - a NHEJ factor - protect telomeres? The answer likely lies in its interacting partner: the KU70/80 heterodimer. Paradoxically, whilst KU promotes NHEJ at DNA breaks, it has also been reported to associate with functional telomeres both *in vitro* and *in vivo*, and that it functions to repress telomere fusions [169–171]. Notably, the KU heterodimer has also been reported to repress exchanges between sister telomeres (T-SCEs) - a form of HR - at deprotected telomeres [172]. This could be particularly important at "open" telomeres, i.e. during or immediately after replication when the t-loop is momentarily compromised. Importantly, T-SCEs can shorten or lengthen individual telomeres when the exchanged segments are not of equal length. It is tantalizing to speculate that the DNA-translocase ERCC6L2 has a function in correctly positioning KU to repress T-SCEs at telomeres, for example by repressing DNA resection. One prediction of this hypothesis would be that patients with *ERCC6L2* mutations might have an increased rate of T-SCEs, ultimately leading to increased telomere length heterogeneity.

Notably, a dual role in DSB repair via NHEJ and in the protection of telomeres is not exclusive to the KU heterodimer (and potentially ERCC6L2). A striking parallel can be drawn to another factor involved in the resolution of staggered DSBs during CSR: the CST complex. During CSR, the CTC1-STN1-TEN1 complex is recruited by shieldin to facilitate DNA fill-in synthesis [25]. Yet, CST's main function is not at switch regions, but at telomeres: During replication, lagging strand DNA synthesis is unable to completely duplicate the C-strand. CST-Polα-primase maintains the C-strand by mediating fill-in synthesis, thereby preventing premature shortening

of the telomere [64]. In line with its importance in telomere maintenance, mutations in *CTC1* and *STN1* have been linked to Coats plus, a pleiotropic disorder with shortened telomeres [173, 174]. With CST moonlighting at switch regions during CSR, but providing its main function at telomeres, it is therefore enticing to postulate a similar dual role for ERCC6L2.

The evidence supporting the classification of ERCC6L2 syndrome as a telomeropathy is currently inconclusive. Compared to healthy, age-matched controls, patients with telomeropathies have very short telomeres and display a characteristic triad of mucocutaneous symptoms: abnormal skin pigmentation, nail dystrophy, and oral leukoplakia [44, 175]. Patients with ERCC6L2 syndrome lack these symptoms and telomere length was determined to be variable in between patients, but not universally critically short. Taken together, this lead to the disease being dismissed as a telomeropathy [88–90]. Yet, the qPCR-based approach for quantifying telomere length used in these studies provides only an estimate of the average (bulk) telomere length and does not capture information on length heterogeneity or the presence of critically short telomeres at the individual level. Importantly, the shortest telomere, rather than the average length dictates cell viability and chromosome stability, meaning that telomere crisis can be triggered by a single critically short telomere [176, 177]. If *ERCC6L2* mutations lead to an increase in telomere length heterogeneity, rather than bulk shortening, this would be undetectable with previously used methods, which only report on average telomere length. It would therefore be possible that critically short telomeres remained undetected whilst still causing HSC attrition. Instead, high-resolution telomere sequencing (such as Telo-seq) of patients with ERCC6L2 syndrome should be able to shed light on the distribution of telomere length and could provide a straightforward approach to test this hypothesis [178]. Regardless, more work is needed to ascertain the molecular mechanisms underlying ERCC6L2 syndrome.

6.4 Conclusions and final model

- *Ercc6l2*$^{-/-}$ mice do not have any haematopoietic defects and cannot recapitulate the human ERCC6L2 syndrome.
- ERCC6L2 syndrome is unlikely to be precipitated by DNA damage caused by endogenous aldehydes, making it distinct from Fanconi anaemia.
- ERCC6L2 is a NHEJ factor selectively required at staggered DSBs.
- ERCC6L2 does not participate in physiological V(D)J recombination.
- ERCC6L2 is required for CSR in a 53BP1-dependent manner, but acts in a distinct, parallel pathway to the shieldin-CST axis.
- Structurally, ERCC6L2 forms a functionally relevant complex with MRI and the KU70/80 heterodimer.
- ERCC6L2 may have dual functions in cNHEJ during CSR and in the protection of telomeres, particularly by repressing T-SCEs.

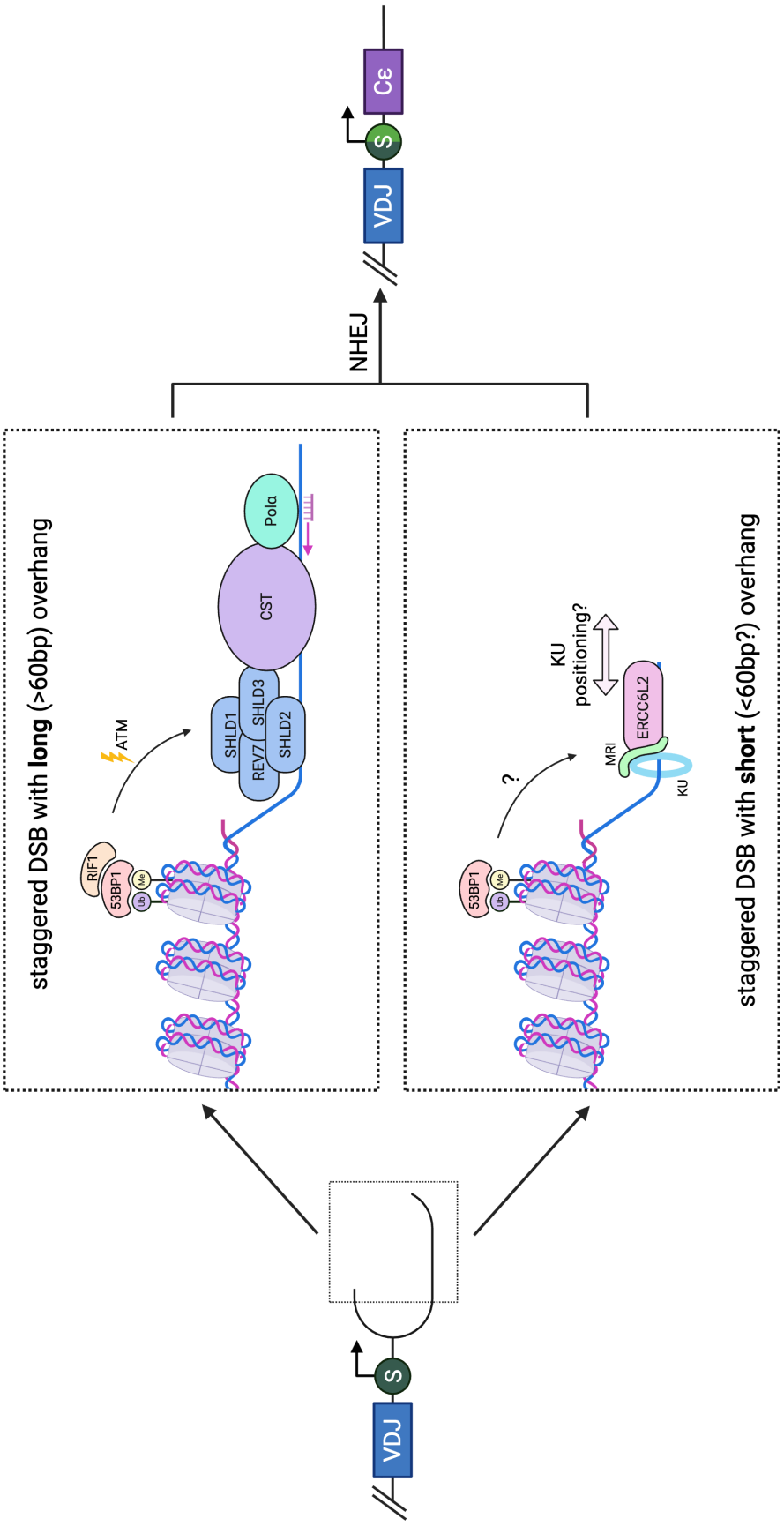


Figure 6.1: **Final model of the ERCC6L2's function during CSR.** During CSR, DSBs with staggered ends of varying lengths are generated. DSBs with long overhangs get processed for fill-in by shieldin-CST-Polα, whereas DSBs with short overhangs require ERCC6L2 - possibly to correctly position the KU heterodimer for subsequent ligation complex loading. Schematic was created using BioRender.

REFERENCES

1. S. P. Jackson and J. Bartek, "The DNA-damage response in human biology and disease," *Nature*, vol. 461, pp. 1071–1078, 10 2009.

2. D. Hanahan and R. Weinberg, "Hallmarks of cancer: The next generation," *Cell*, vol. 144, no. 5, pp. 646–674, 2011.

3. A. Tubbs and A. Nussenzweig, "Endogenous DNA damage as a source of genomic instability in cancer," *Cell*, vol. 168, no. 4, pp. 644–656, 2017.

4. R. Ceccaldi, B. Rondinelli, and A. D. D'Andrea, "Repair pathway choices and consequences at the double-strand break," *Trends in Cell Biology*, vol. 26, pp. 52–64, 2016.

5. R. Ceccaldi and P. Cejka, "Mechanisms and regulation of DNA end resection in the maintenance of genome stability," *Nature Reviews Molecular Cell Biology*, pp. 1–14, 2025.

6. A. Brambati, O. Sacco, S. Porcella, J. Heyza, M. Kareh, J. C. Schmidt, and A. Sfeir, "RHINO directs MMEJ to repair DNA breaks in mitosis," *Science*, p. eadh3694, 2023.

7. C. Gelot, M. T. Kovacs, S. Miron, E. Mylne, A. Haan, L. Boeffard-Dosierre, R. Ghouil, T. Popova, F. Dingli, D. Loew, J. Guirouilh-Barbat, E. D. Nery, S. Zinn-Justin, and R. Ceccaldi, "PolÎž is phosphorylated by PLK1 to repair double-strand breaks in mitosis," *Nature*, vol. 621, no. 7978, pp. 415–422, 2023.

8. J.-H. Lee and T. T. Paull, "ATM activation by DNA double-strand breaks through the mre11-rad50-nbs1 complex," *Science*, vol. 308, no. 5721, pp. 551–554, 2005.

9. E. P. Rogakou, D. R. Pilch, A. H. Orr, V. S. Ivanova, and W. M. Bonner, "DNA double-stranded breaks induce histone h2ax phosphorylation on serine 139*," *Journal of Biological Chemistry*, vol. 273, no. 10, pp. 5858–5868, 1998.

10. A. N. Blackford and S. P. Jackson, "ATM, ATR, and DNA-PK: The trinity at the heart of the DNA damage

response," *Molecular Cell*, vol. 66, pp. 801–817, 2017.

11. M. Stucki, J. A. Clapperton, D. Mohammad, M. B. Yaffe, S. J. Smerdon, and S. P. Jackson, "MDC1 directly binds phosphorylated histone h2ax to regulate cellular responses to DNA double-strand breaks," *Cell*, vol. 123, no. 7, pp. 1213–1226, 2005.

12. N. K. Kolas, J. R. Chapman, S. Nakada, J. Ylanko, R. Chahwan, F. D. Sweeney, S. Panier, M. Mendez, J. Wildenhain, T. M. Thomson, L. Pelletier, S. P. Jackson, and D. Durocher, "Orchestration of the DNA-damage response by the RNF8 ubiquitin ligase," *Science*, vol. 318, no. 5856, pp. 1637–1640, 2007.

13. N. Mailand, S. Bekker-Jensen, H. Faustrup, F. Melander, J. Bartek, C. Lukas, and J. Lukas, "RNF8 ubiquitylates histones at DNA double-strand breaks and promotes assembly of repair proteins," *Cell*, vol. 131, no. 5, pp. 887–900, 2007.

14. M. S. Huen, R. Grant, I. Manke, K. Minn, X. Yu, M. B. Yaffe, and J. Chen, "RNF8 transduces the DNA-damage signal via histone ubiquitylation and checkpoint protein assembly," *Cell*, vol. 131, no. 5, pp. 901–914, 2007.

15. T. Thorslund, A. Ripplinger, S. Hoffmann, T. Wild, M. Uckelmann, B. Villumsen, T. Narita, T. K. Sixma, C. Choudhary, S. Bekker-Jensen, and N. Mailand, "Histone h1 couples initiation and amplification of ubiquitin signalling after DNA damage," *Nature*, vol. 527, no. 7578, pp. 389–393, 2015.

16. F. Mattiroli, J. H. A. Vissers, W. J. V. Dijk, P. Ikpa, E. Citterio, W. Vermeulen, J. A. Marteijn, and T. K. Sixma, "RNF168 ubiquitinates k13-15 on h2a/h2ax to drive DNA damage signaling," *Cell*, vol. 150, pp. 1182–1195, 2012.

17. C. Doil, N. Mailand, S. Bekker-Jensen, P. Menard, D. H. Larsen, R. Pepperkok, J. Ellenberg, S. Panier, D. Durocher, J. Bartek, J. Lukas, and C. Lukas, "RNF168 binds and amplifies ubiquitin conjugates on damaged chromosomes to allow accumulation of repair proteins," *Cell*, vol. 136, no. 3, pp. 435–446, 2009.

18. S. Panier and S. J. Boulton, "Double-strand break repair: 53bp1 comes into focus," *Nature Reviews Molecular Cell Biology*, vol. 15, no. 1, pp. 7–18, 2014.

19. M. V. Botuyan, J. Lee, I. M. Ward, J.-E. Kim, J. R. Thompson, J. Chen, and G. Mer, "Structural basis for the methylation state-specific recognition of histone h4-k20 by 53bp1 and crb2 in DNA repair," *Cell*, vol. 127, no. 7, pp. 1361–1373, 2006.

20. A. Fradet-Turcotte, M. D. Canny, C. Escribano-Diaz, A. Orthwein, C. C. Y. Leung, H. Huang, M.-C. Landry, J. Kitevski-LeBlanc, S. M. Noordermeer, F. Sicheri, and D. Durocher, "53bp1 is a reader of the DNA-damage-induced h2a lys 15 ubiquitin mark," *Nature*, vol. 499, no. 7456, pp. 50–54, 2013.

21. S. Pellegrino, J. Michelena, F. Teloni, R. Imhof, and M. Altmeyer, "Replication-coupled dilution of h4k20me2 guides 53bp1 to pre-replicative chromatin," *Cell Reports*, vol. 19, no. 9, pp. 1819–1831, 2017.

22. S. F. Bunting, E. Callen, N. Wong, H.-T. Chen, F. Polato, A. Gunn, A. Bothmer, N. Feldhahn, O. Fernandez-

Capetillo, L. Cao, X. Xu, C.-X. Deng, T. Finkel, M. Nussenzweig, J. M. Stark, and A. Nussenzweig, "53bp1 inhibits homologous recombination in brca1-deficient cells by blocking resection of DNA breaks," *Cell*, vol. 141, no. 2, pp. 243–254, 2010.

23. M. D. Virgilio, E. Callen, A. Yamane, W. Zhang, M. Jankovic, A. D. Gitlin, N. Feldhahn, W. Resch, T. Y. Oliveira, B. T. Chait, A. Nussenzweig, R. Casellas, D. F. Robbiani, and M. C. Nussenzweig, "Rif1 prevents resection of DNA breaks and promotes immunoglobulin class switching," *Science*, vol. 339, pp. 711–715, 2 2013.

24. J. R. Chapman, 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, "RIF1 is essential for 53bp1-dependent nonhomologous end joining and suppression of DNA double-strand break resection," *Molecular Cell*, vol. 49, pp. 858–871, 3 2013.

25. A. King, P. I. Reichl, J. S. Metson, R. Parker, D. Munro, C. Oliveira, L. Sommerova, J. R. Becker, D. Biggs, C. Preece, B. Davies, and J. R. Chapman, "Shieldin and CST co-orchestrate DNA polymerase-dependent tailed-end joining reactions independently of 53BP1-governed repair pathway choice," *Nature Structural & Molecular Biology*, pp. 1–12, 2024.

26. J. R. Chapman, A. J. Sossick, S. J. Boulton, and S. P. Jackson, "BRCA1-associated exclusion of 53bp1 from DNA damage sites underlies temporal control of DNA repair," *Journal of Cell Science*, vol. 125, no. 15, pp. 3529–3534, 2012.

27. J. R. Becker, G. Clifford, C. Bonnet, A. Groth, M. D. Wilson, and J. R. Chapman, "BARD1 reads h2a lysine 15 ubiquitination to direct homologous recombination," *Nature*, vol. 596, pp. 433–437, 7 2021.

28. I. Ceppi, M. R. D. Stritto, M. Mutze, S. Braunshier, V. Mengoli, G. Reginato, H. M. P. Vo, S. Jimeno, A. Acharya, M. Roy, A. Sanchez, S. Halder, S. M. Howard, R. Guerois, P. Huertas, S. M. Noordermeer, R. Seidel, and P. Cejka, "Mechanism of BRCA1-BARD1 function in DNA end resection and DNA protection," *Nature*, pp. 1–9, 2024.

29. S. Salunkhe, J. M. Daley, H. Kaur, N. Tomimatsu, C. Xue, V. B. Raina, A. M. Jasper, C. M. Rogers, W. Li, S. Zhou, R. Mojidra, Y. Kwon, Q. Fang, J.-H. Ji, A. B. Shabestari, O. Fitzgerald, H. Dinh, B. Mukherjee, A. A. Habib, R. Hromas, A. V. Mazin, E. V. Wasmuth, S. K. Olsen, D. S. Libich, D. Zhou, W. Zhao, E. C. Greene, S. Burma, and P. Sung, "Promotion of DNA end resection by BRCA1-BARD1 in homologous recombination," *Nature*, pp. 1–10, 2024.

30. S. M. H. Sy, M. S. Y. Huen, and J. Chen, "PALB2 is an integral component of the BRCA complex required for homologous recombination repair," *Proceedings of the National Academy of Sciences*, vol. 106, no. 17, pp. 7155–7160, 2009.

31. M. Jasin and R. Rothstein, "Repair of strand breaks by homologous recombination," *Cold Spring Harbor Perspectives in Biology*, vol. 5, no. 11, p. a012740, 2013.

32. Z. Mirman and T. d. Lange, "53bp1: a DSB escort," *Genes & development*, vol. 34, pp. 7–23, 1 2020.

33. C. Escribano-Diaz, A. Orthwein, A. Fradet-Turcotte, M. Xing, J. Young, J. Tkac, M. Cook, A. Rosebrock, M. Munro, M. Canny, D. Xu, and D. Durocher, "A Cell Cycle-Dependent Regulatory Circuit Composed of 53BP1-RIF1 and BRCA1-CtIP Controls DNA Repair Pathway Choice," *Molecular Cell*, vol. 49, no. 5, pp. 872–883, 2013.

34. B. Zhao, E. Rothenberg, D. A. Ramsden, and M. R. Lieber, "The molecular basis and disease relevance of non-homologous DNA end joining," *Nature Reviews Molecular Cell Biology*, vol. 21, pp. 765–781, 12 2020.

35. T. Graham, J. Walter, and J. Loparo, "Two-stage synapsis of DNA ends during non-homologous end joining," *Molecular Cell*, vol. 61, no. 6, pp. 850–858, 2016.

36. H. H. Y. Chang, N. R. Pannunzio, N. Adachi, and M. R. Lieber, "Non-homologous DNA end joining and alternative pathways to double-strand break repair," *Nature Reviews Molecular Cell Biology*, vol. 18, pp. 495–506, 8 2017.

37. J. M. Pryor, C. A. Waters, A. Aza, K. Asagoshi, C. Strom, P. A. Mieczkowski, L. Blanco, and D. A. Ramsden, "Essential role for polymerase specialization in cellular nonhomologous end joining," *Proceedings of the National Academy of Sciences*, vol. 112, no. 33, pp. E4537–E4545, 2015.

38. S. Chen, L. Lee, T. Naila, S. Fishbain, A. Wang, A. E. Tomkinson, S. P. Lees-Miller, and Y. He, "Structural basis of long-range to short-range synaptic transition in NHEJ," *Nature*, vol. 593, no. 7858, pp. 294–298, 2021.

39. N. R. Pannunzio, G. Watanabe, and M. R. Lieber, "Nonhomologous DNA end-joining for repair of DNA double-strand breaks," *Journal of Biological Chemistry*, vol. 293, no. 27, pp. 10512–10523, 2018.

40. A. C. Wilkinson, K. J. Igarashi, and H. Nakauchi, "Haematopoietic stem cell self-renewal in vivo and ex vivo," *Nature Reviews Genetics*, vol. 21, no. 9, pp. 541–554, 2020.

41. S. H. Orkin and L. I. Zon, "Hematopoiesis: An evolving paradigm for stem cell biology," *Cell*, vol. 132, no. 4, pp. 631–644, 2008.

42. E. Laurenti and B. Goettgens, "From haematopoietic stem cells to complex differentiation landscapes," *Nature*, vol. 553, pp. 418–426, 2018.

43. R. Yamamoto, A. C. Wilkinson, and H. Nakauchi, "Changing concepts in hematopoietic stem cells," *Science*, vol. 362, pp. 895–896, 11 2018.

44. I. Dokal and T. Vulliamy, "Inherited bone marrow failure syndromes," *Haematologica*, vol. 95, no. 8, pp. 1236–1240, 2010.

45. M. Park, "Overview of inherited bone marrow failure syndromes," *Blood Research*, vol. 57, no. S1, pp. S49–S54, 2022.

46. R. Ceccaldi, P. Sarangi, and A. D. D'Andrea, "The Fanconi anaemia pathway: new players and new functions," *Nature Reviews Molecular Cell Biology*, vol. 17, pp. 337–349, 6 2016.

47. A. J. Deans and S. C. West, "DNA interstrand crosslink repair and cancer," *Nature Reviews Cancer*, vol. 11, no. 7, pp. 467–480, 2011.

48. I. V. Rosado, F. Langevin, G. P. Crossan, M. Takata, and K. J. Patel, "Formaldehyde catabolism is essential in cells deficient for the fanconi anemia DNA-repair pathway," *Nature Structural & Molecular Biology*, vol. 18, no. 12, pp. 1432–1434, 2011.

49. F. Langevin, G. P. Crossan, I. V. Rosado, M. J. Arends, and K. J. Patel, "Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice," *Nature*, vol. 475, no. 7354, pp. 53–58, 2011.

50. J. I. Garaycoechea, G. P. Crossan, F. Langevin, M. Daly, M. J. Arends, and K. J. Patel, "Genotoxic consequences of endogenous aldehydes on mouse haematopoietic stem cell function," *Nature*, vol. 489, pp. 571–575, 9 2012.

51. L. B. Pontel, I. V. Rosado, G. Burgos-Barragan, J. I. Garaycoechea, R. Yu, M. J. Arends, G. Chandrasekaran, V. Broecker, W. Wei, L. Liu, J. A. Swenberg, G. P. Crossan, and K. J. Patel, "Endogenous formaldehyde is a hematopoietic stem cell genotoxin and metabolic carcinogen," *Molecular Cell*, vol. 60, pp. 177–188, 10 2015.

52. A. Hira, H. Yabe, K. Yoshida, Y. Okuno, Y. Shiraishi, K. Chiba, H. Tanaka, S. Miyano, J. Nakamura, S. Kojima, S. Ogawa, K. Matsuo, M. Takata, and M. Yabe, "Variant ALDH2 is associated with accelerated progression of bone marrow failure in japanese fanconi anemia patients," *Blood*, vol. 122, no. 18, pp. 3206–3209, 2013.

53. F. A. Dingler, M. Wang, A. Mu, C. L. Millington, N. Oberbeck, S. Watcham, L. B. Pontel, A. N. Kamimae-Lanning, F. Langevin, C. Nadler, R. L. Cordell, P. S. Monks, R. Yu, N. K. Wilson, A. Hira, K. Yoshida, M. Mori, Y. Okamoto, Y. Okuno, H. Muramatsu, Y. Shiraishi, M. Kobayashi, T. Moriguchi, T. Osumi, M. Kato, S. Miyano, E. Ito, S. Kojima, H. Yabe, M. Yabe, K. Matsuo, S. Ogawa, B. Goettgens, M. R. Hodskinson, M. Takata, and K. J. Patel, "Two aldehyde clearance systems are essential to prevent lethal formaldehyde accumulation in mice and humans," *Molecular Cell*, vol. 80, pp. 996–1012.e9, 12 2020.

54. T. Altmann and A. R. Gennery, "DNA ligase IV syndrome; a review," *Orphanet Journal of Rare Diseases*, vol. 11, no. 1, p. 137, 2016.

55. A. Nijnik, L. Woodbine, C. Marchetti, S. Dawson, T. Lambe, C. Liu, N. P. Rodrigues, T. L. Crockford, E. Cabuy, A. Vindigni, T. Enver, J. I. Bell, P. Slijepcevic, C. C. Goodnow, P. A. Jeggo, and R. J. Cornall, "DNA repair is limiting for haematopoietic stem cells during ageing," *Nature*, vol. 447, pp. 686–690, 6 2007.

56. D. Buck, L. Malivert, R. d. Chasseval, A. Barraud, M.-C. Fondaneche, O. Sanal, A. Plebani, J.-L. Stephan, M. Hufnagel, F. l. Deist, A. Fischer, A. Durandy, J.-P. d. Villartay, and P. Revy, "Cernunnos, a novel nonhomologous end-joining factor, is mutated in human immunodeficiency with microcephaly," *Cell*, vol. 124, no. 2, pp. 287–299, 2006.

57. G. Li, F. W. Alt, H. L. Cheng, J. W. Brush, P. H. Goff, M. M. Murphy, S. Franco, Y. Zhang, and

S. Zha, "Lymphocyte-specific compensation for XLF/cernunnos end-joining functions in v(d)j recombination," *Molecular Cell*, vol. 31, pp. 631–640, 9 2008.

58. S. Avagyan, M. Churchill, K. Yamamoto, J. L. Crowe, C. Li, B. J. Lee, T. Zheng, S. Mukherjee, and S. Zha, "Hematopoietic stem cell dysfunction underlies the progressive lymphocytopenia in XLF/cernunnos deficiency," *Blood*, vol. 124, no. 10, pp. 1622–1625, 2014.

59. D. M. Townsley, B. Dumitriu, and N. S. Young, "Bone marrow failure and the telomeropathies," *Blood*, vol. 124, pp. 2775–2783, 10 2014.

60. J. D. Watson, "Origin of concatemeric t7dna," *Nature New Biology*, vol. 239, no. 94, pp. 197–201, 1972.

61. L. Hayflick and P. Moorhead, "The serial cultivation of human diploid cell strains," *Experimental Cell Research*, vol. 25, no. 3, pp. 585–621, 1961.

62. U. Herbig, W. A. Jobling, B. P. Chen, D. J. Chen, and J. M. Sedivy, "Telomere shortening triggers senescence of human cells through a pathway involving ATM, p53, and p21cip1, but not p16ink4a," *Molecular Cell*, vol. 14, no. 4, pp. 501–513, 2004.

63. C. Jones-Weinert, L. Mainz, and J. Karlseder, "Telomere function and regulation from mouse models to human ageing and disease," *Nature Reviews Molecular Cell Biology*, vol. 26, no. 4, pp. 297–313, 2025.

64. H. Takai, V. Aria, P. Borges, J. T. P. Yeeles, and T. d. Lange, "CST-polymerase a-primase solves a second telomere end-replication problem," *Nature*, vol. 627, no. 8004, pp. 664–670, 2024.

65. B. McClintock, "THE STABILITY OF BROKEN ENDS OF CHROMOSOMES IN ZEA MAYS," *Genetics*, vol. 26, no. 2, pp. 234–282, 1941.

66. W. Palm and T. d. Lange, "How shelterin protects mammalian telomeres," *Annual Review of Genetics*, vol. 42, no. 1, pp. 301–334, 2008.

67. B. Alberts, R. Heald, A. Johnson, D. Morgan, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell*. W. W. Norton Company, Incorporated, 7th ed. ed., 2022.

68. F. W. Alt, Y. Zhang, F. L. Meng, C. Guo, and B. Schwer, "Mechanisms of programmed DNA lesions and genomic instability in the immune system," *Cell*, vol. 152, pp. 417–429, 1 2013.

69. D. G. Schatz and Y. Ji, "Recombination centres and the orchestration of v(d)j recombination," *Nature Reviews Immunology*, vol. 11, pp. 251–263, 4 2011.

70. R. R. Hardy, C. E. Carmack, S. A. Shinton, J. D. Kemp, and K. Hayakawa, "Resolution and characterization of pro-b and pre-pro-b cell stages in normal mouse bone marrow.," *Journal of Experimental Medicine*, vol. 173, pp. 1213–1225, 5 1991.

71. R. R. Hardy and K. Hayakawa, "B CELL DEVELOPMENT PATHWAYS," *Annual Review of Immunology*, vol. 19, no. 1, pp. 595–621, 2001.

72. E. V. Rothenberg, J. E. Moore, and M. A. Yui, "Launching the t-cell-lineage developmental programme," *Nature Reviews Immunology*, vol. 8, no. 1, pp. 9–21, 2008.

73. K. Schwarz, G. H. Gauss, L. Ludwig, U. Pannicke, Z. Li, D. Lindner, W. Friedrich, R. A. Seger, T. E. Hansen-Hagge, S. Desiderio, M. R. Lieber, and C. R. Bartram, "RAG mutations in human b cell-negative SCID," *Science*, vol. 274, no. 5284, pp. 97–99, 1996.

74. D. Moshous, I. Callebaut, R. d. Chasseval, B. Corneo, M. Cavazzana-Calvo, F. L. Deist, I. Tezcan, O. Sanal, Y. Bertrand, N. Philippe, A. Fischer, and J.-P. d. Villartay, "Artemis, a novel DNA double-strand break repair/v(d)j recombination protein, is mutated in human severe combined immune deficiency," *Cell*, vol. 105, no. 2, pp. 177–186, 2001.

75. Z. Xu, H. Zan, E. J. Pone, T. Mai, and P. Casali, "Immunoglobulin class-switch DNA recombination: Induction, targeting and beyond," *Nature Reviews Immunology*, vol. 12, pp. 517–531, 7 2012.

76. J. J. Bednarski and B. P. Sleckman, "At the intersection of DNA damage and immune responses," *Nature Reviews Immunology*, vol. 19, pp. 231–242, 4 2019.

77. T. Saha, D. Sundaravinayagam, and M. D. Virgilio, "Charting a DNA repair roadmap for immunoglobulin class switch recombination," *Trends in Biochemical Sciences*, vol. 46, pp. 184–199, 3 2021.

78. C. T. Yan, C. Boboila, E. K. Souza, S. Franco, T. R. Hickernell, M. Murphy, S. Gumaste, M. Geyer, A. A. Zarrin, J. P. Manis, K. Rajewsky, and F. W. Alt, "IgH class switching and translocations use a robust non-classical end-joining pathway," *Nature*, vol. 449, pp. 478–482, 9 2007.

79. C. Boboila, C. Yan, D. R. Wesemann, M. Jankovic, J. H. Wang, J. Manis, A. Nussenzweig, M. Nussenzweig, and F. W. Alt, "Alternative end-joining catalyzes class switch recombination in the absence of both ku70 and DNA ligase 4," *Journal of Experimental Medicine*, vol. 207, no. 2, pp. 417–427, 2010.

80. J. P. Manis, J. C. Morales, Z. Xia, J. L. Kutok, F. W. Alt, and P. B. Carpenter, "53bp1 links DNA damage-response pathways to immunoglobulin heavy chain class-switch recombination," *Nature Immunology*, vol. 5, no. 5, pp. 481–487, 2004.

81. I. M. Ward, B. Reina-San-Martin, A. Olaru, K. Minn, K. Tamada, J. S. Lau, M. Cascalho, L. Chen, A. Nussenzweig, F. Livak, M. C. Nussenzweig, and J. Chen, "53bp1 is required for class switch recombination," *The Journal of Cell Biology*, vol. 165, no. 4, pp. 459–464, 2004.

82. R. Gupta, K. Somyajit, T. Narita, E. Maskey, A. Stanlie, M. Kremer, D. Typas, M. Lammers, N. Mailand, A. Nussenzweig, J. Lukas, and C. Choudhary, "DNA repair network analysis reveals shieldin as a key regulator of NHEJ and PARP inhibitor sensitivity," *Cell*, vol. 173, no. 4, pp. 972–988.e23, 2018.

83. H. Ghezraoui, C. Oliveira, J. R. Becker, K. Bilham, D. Moralli, C. Anzilotti, R. Fischer, M. Deobagkar-Lele, M. Sanchiz-Calvo, E. Fueyo-Marcos, S. Bonham, B. M. Kessler, S. Rottenberg, R. J. Cornall, C. M. Green, and J. R. Chapman, "53bp1 cooperation with the REV7âshieldin complex underpins DNA structure-specific

NHEJ," *Nature*, vol. 560, pp. 122–127, 8 2018.

84. S. M. Noordermeer, S. Adam, D. Setiaputra, M. Barazas, S. J. Pettitt, A. K. Ling, M. Olivieri, A. Alvarez-Quilon, N. Moatti, M. Zimmermann, S. Annunziato, D. B. Krastev, F. Song, I. Brandsma, J. Frankum, R. Brough, A. Sherker, S. Landry, R. K. Szilard, M. M. Munro, A. McEwan, T. G. d. Rugy, Z. Y. Lin, T. Hart, J. Moffat, A. C. Gingras, A. Martin, H. v. Attikum, J. Jonkers, C. J. Lord, S. Rottenberg, and D. Durocher, "The shieldin complex mediates 53bp1-dependent DNA repair," *Nature*, vol. 560, pp. 117–121, 7 2018.

85. H. Dev, T. W. W. Chiang, C. Lescale, I. d. Krijger, A. G. Martin, D. Pilger, J. Coates, M. Sczaniecka-Clift, W. Wei, M. Ostermaier, M. Herzog, J. Lam, A. Shea, M. Demir, Q. Wu, F. Yang, B. Fu, Z. Lai, G. Balmus, R. Belotserkovskaya, V. Serra, M. J. OâConnor, A. Bruna, P. Beli, L. Pellegrini, C. Caldas, L. Deriano, J. J. Jacobs, Y. Galanty, and S. P. Jackson, "Shieldin complex promotes DNA end-joining and counters homologous recombination in BRCA1-null cells," *Nature Cell Biology*, vol. 20, pp. 954–965, 8 2018.

86. Z. Mirman, F. Lottersberger, H. Takai, T. Kibe, Y. Gong, K. Takai, A. Bianchi, M. Zimmermann, D. Durocher, and T. d. Lange, "53bp1-RIF1-shieldin counteracts DSB resection through CST- and pola-dependent fill-in," *Nature*, vol. 560, pp. 112–116, 8 2018.

87. Z. Mirman, N. K. Sasi, A. King, J. R. Chapman, and T. d. Lange, "53bp1-shieldin-dependent DSB processing in BRCA1-deficient cells requires CST-pola-primase fill-in synthesis," *Nature Cell Biology*, vol. 24, pp. 51–61, 1 2022.

88. H. Tummala, M. Kirwan, A. J. Walne, U. Hossain, N. Jackson, C. Pondarre, V. Plagnol, T. Vulliamy, and I. Dokal, "ERCC6l2 mutations link a distinct bone-marrow-failure syndrome to DNA repair and mitochondrial function," *The American Journal of Human Genetics*, vol. 94, pp. 246–256, 2 2014.

89. M. Hakkarainen, I. Kaaja, S. P. M. Douglas, T. J. Vulliamy, I. Dokal, J. Soulier, L. Larcher, R. P. d. Latour, T. M. Leblanc, F. S. d. Fontbrune, T. Siitonen, O. Lohi, E. Hellstroem-Lindberg, G. Barbany, B. Tesi, A. Shimamura, F. Beier, S. R. Jackson, A. Kuperman, T. C. F. Zaccai, H. Tamary, C. Mecucci, I. Capolsini, K. Jahnukainen, U. Salmenniemi, R. Niinimaeki, T. Varilo, O. Kilpivaara, and U. Wartiovaara-Kautto, "The clinical picture of the ERCC6l2 disease - from bone marrow failure to acute leukemia," *Blood*, vol. 141, no. 23, pp. 2853–2866, 2023.

90. S. Zhang, 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. d. Villartay, "A nonsense mutation in the DNA repair factor hebo causes mild bone marrow failure and microcephaly," *Journal of Experimental Medicine*, vol. 213, pp. 1011–1028, 5 2016.

91. I. Shabanova, E. Cohen, M. Cada, A. Vincent, R. D. Cohn, and Y. Dror, "ERCC6l2-associated inherited bone marrow failure syndrome," *Molecular Genetics & Genomic Medicine*, vol. 6, pp. 463–468, 5 2018.

92. T. Jaerviaho, K. Halt, P. Hirvikoski, J. Moilanen, M. Moettoenen, and R. Niinimaeki, "Bone marrow failure

syndrome caused by homozygous frameshift mutation in the ERCC6l2 gene," *Clinical Genetics*, vol. 93, pp. 392–395, 2 2018.

93. M. W. Wlodarski, "ERCC6l2 syndrome: attack of the TP53 clones," *Blood*, vol. 141, no. 23, pp. 2788–2789, 2023.

94. F. Baccelli, D. Leardini, S. Cerasi, D. Messelodi, S. N. Bertuccio, and R. Masetti, "ERCC6l2-related disease: a novel entity of bone marrow failure disorder with high risk of clonal evolution," *Annals of Hematology*, vol. 1, pp. 1–7, 2 2023.

95. S. P. Douglas, P. Siipola, P. E. Kovanen, M. Pyoeraelae, S. Kakko, E. R. Savolainen, U. Salmenniemi, K. Orte, S. Kytoelae, E. Pitkaenen, K. Porkka, O. Kilpivaara, and U. Wartiovaara-Kautto, "ERCC6l2 defines a novel entity within inherited acute myeloid leukemia," *Blood*, vol. 133, pp. 2724–2728, 6 2019.

96. P. Boddu, C. B. Benton, W. Wang, G. Borthakur, J. D. Khoury, and N. Pemmaraju, "Erythroleukemia-historical perspectives and recent advances in diagnosis and management," *Blood Reviews*, vol. 32, no. 2, pp. 96–105, 2018.

97. D. Rose, T. Haferlach, S. Schnittger, K. Perglerova, W. Kern, and C. Haferlach, "Subtype-specific patterns of molecular mutations in acute myeloid leukemia," *Leukemia*, vol. 31, no. 1, pp. 11–17, 2017.

98. F. P. S. Santos, S. Faderl, G. Garcia-Manero, C. Koller, M. Beran, S. O'Brien, S. Pierce, E. J. Freireich, X. Huang, G. Borthakur, C. Bueso-Ramos, M. d. Lima, M. Keating, J. Cortes, H. Kantarjian, and F. Ravandi, "Adult acute erythroleukemia: an analysis of 91 patients treated at a single institution," *Leukemia*, vol. 23, no. 12, pp. 2275–2280, 2009.

99. H. Tummala, A. D. Dokal, A. Walne, A. Ellison, S. Cardoso, S. Amirthasigamanipillai, M. Kirwan, I. Browne, J. K. Sidhu, V. Rajeeve, A. Rio-Machin, A. A. Seraihi, A. S. Duncombe, M. Jenner, O. P. Smith, H. Enright, A. Norton, T. Aksu, N. Y. oezbek, N. Pontikos, P. Cutillas, I. Dokal, and T. Vulliamy, "Genome instability is a consequence of transcription deficiency in patients with bone marrow failure harboring biallelic ERCC6l2 variants," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 115, pp. 7777–7782, 7 2018.

100. W. Wang, J. Xu, J. Chong, and D. Wang, "Structural basis of DNA lesion recognition for eukaryotic transcription-coupled nucleotide excision repair," *DNA Repair*, vol. 71, pp. 43–55, 2018.

101. M. Olivieri, T. Cho, A. Alvarez-Quilon, K. Li, M. Schellenberg, M. Zimmermann, N. Hustedt, S. E. Rossi, S. Adam, H. Melo, A. M. Heijink, G. Sastre-Moreno, N. Moatti, R. Szilard, A. McEwan, A. Ling, A. Serrano-Benitez, T. Ubhi, I. Delgado-Sainz, M. Ferguson, G. Brown, F. Cortes-Ledesma, R. S. Williams, A. Martin, D. Xu, and D. Durocher, "A genetic map of the response to DNA damage in human cells," *Cell*, vol. 182, pp. 1–16, 7 2020.

102. X. Liu, T. Liu, Y. Shang, P. Dai, W. Zhang, B. J. Lee, M. Huang, D. Yang, Q. Wu, L. D. Liu, X. Zheng, B. O. Zhou, J. Dong, L.-S. Yeap, J. Hu, T. Xiao, S. Zha, R. Casellas, X. S. Liu, and F.-L. Meng, "ERCC6l2

promotes DNA orientation-specific recombination in mammalian cells," *Cell Research*, vol. 30, pp. 732–744, 9 2020.

103. P. Francica, M. Mutlu, V. A. Blomen, C. Oliveira, Z. Nowicka, A. Trenner, N. M. Gerhards, P. Bouwman, E. Stickel, M. L. Hekkelman, L. Lingg, I. Klebic, M. v. d. Ven, R. d. Korte-Grimmerink, D. Howald, J. Jonkers, A. A. Sartori, W. Fendler, J. R. Chapman, T. Brummelkamp, and S. Rottenberg, "Functional radiogenetic profiling implicates ERCC6l2 in non-homologous end joining," *Cell Reports*, vol. 32, p. 108068, 8 2020.

104. H. Armes, F. BewickeâCopley, A. RioâMachin, D. D. Bella, C. Philippe, A. Wozniak, H. Tummala, J. Wang, T. Ezponda, F. Prosper, I. Dokal, T. Vulliamy, O. Kilpivaara, U. Wartiovaara-Kautto, J. Fitzgibbon, and K. Rouault-Pierre, "Germline ERCC excision repair 6 like 2 (ERCC6l2) mutations lead to impaired erythropoiesis and reshaping of the bone marrow microenvironment," *British Journal of Haematology*, vol. 00, pp. 1–11, 9 2022.

105. A. Flaus, D. M. A. Martin, G. J. Barton, and T. Owen-Hughes, "Identification of multiple distinct snf2 subfamilies with conserved structural motifs," *Nucleic Acids Research*, vol. 34, no. 10, pp. 2887–2905, 2006.

106. S. Eustermann, A. B. Patel, K.-P. Hopfner, Y. He, and P. Korber, "Energy-driven genome regulation by ATP-dependent chromatin remodellers," *Nature Reviews Molecular Cell Biology*, vol. 25, no. 4, pp. 309–332, 2024.

107. M. Botuyan and G. Mer, "Chromatin signaling and diseases," *Section II: Histone Mark Readers*, pp. 149–165, 2016.

108. C. J. Carnie, L. Armstrong, M. Sebesta, A. Ariza, X. Wang, E. Graham, K. Zhu, and D. Ahel, "ERCC6l2 mitigates replication stress and promotes centromere stability," *Cell Reports*, vol. 42, no. 4, p. 112329, 2023.

109. P. J. Hung, B. Johnson, B.-R. Chen, A. K. Byrum, A. L. Bredemeyer, W. T. Yewdell, T. E. Johnson, B. J. Lee, S. Deivasigamani, I. Hindi, P. Amatya, M. L. Gross, T. T. Paull, D. J. Pisapia, J. Chaudhuri, J. J. Petrini, N. Mosammaparast, G. K. Amarasinghe, S. Zha, J. K. Tyler, and B. P. Sleckman, "MRI is a DNA damage response adaptor during classical non-homologous end joining," *Molecular Cell*, vol. 71, no. 2, pp. 332–342.e8, 2018.

110. T. Rolland, M. TaÅan, B. Charloteaux, S. Pevzner, Q. Zhong, N. Sahni, S. Yi, I. Lemmens, C. Fontanillo, R. Mosca, A. Kamburov, S. Ghiassian, X. Yang, L. Ghamsari, D. Balcha, B. Begg, P. Braun, M. Brehme, M. Broly, A.-R. Carvunis, D. Convery-Zupan, R. Corominas, J. Coulombe-Huntington, E. Dann, M. Dreze, A. Dricot, C. Fan, E. Franzosa, F. Gebreab, B. Gutierrez, M. Hardy, M. Jin, S. Kang, R. Kiros, G. Lin, K. Luck, A. MacWilliams, J. Menche, R. Murray, A. Palagi, M. Poulin, X. Rambout, J. Rasla, P. Reichert, V. Romero, E. Ruyssinck, J. Sahalie, A. Scholz, A. Shah, A. Sharma, Y. Shen, K. Spirohn, S. Tam, A. Tejeda, S. Trigg, J.-C. Twizere, K. Vega, J. Walsh, M. Cusick, Y. Xia, A.-L. Barabasi, L. Iakoucheva, P. Aloy, J. DeÂ LasÂ Rivas, J. Tavernier, M. Calderwood, D. Hill, T. Hao, F. Roth, and M. Vidal, "A proteome-scale map of the human interactome network," *Cell*, vol. 159, no. 5, pp. 1212–1226, 2014.

111. S. A. Slavoff, J. Heo, B. A. Budnik, L. A. Hanakahi, and A. Saghatelian, "A human short open reading frame (sORF)-encoded polypeptide that stimulates DNA end joining*," *Journal of Biological Chemistry*, vol. 289, no. 16, pp. 10950–10957, 2014.

112. N. Arnoult, A. Correia, J. Ma, A. Merlo, S. Garcia-Gomez, M. Maric, M. Tognetti, C. W. Benner, S. J. Boulton, A. Saghatelian, and J. Karlseder, "Regulation of DNA repair pathway choice in s and g2 phases by the NHEJ inhibitor CYREN," *Nature*, vol. 549, no. 7673, pp. 548–552, 2017.

113. I. M. Ward, K. Minn, J. v. Deursen, and J. Chen, "p53 Binding Protein 53BP1 Is Required for DNA Damage Responses and Tumor Suppression in Mice," *Molecular and Cellular Biology*, vol. 23, no. 7, pp. 2556–2563, 2003.

114. L. Liu, Y. Yan, M. Zeng, J. Zhang, M. A. Hanes, G. Ahearn, T. J. McMahon, T. Dickfeld, H. E. Marshall, L. G. Que, and J. S. Stamler, "Essential Roles of S-Nitrosothiols in Vascular Homeostasis and Endotoxic Shock," *Cell*, vol. 116, no. 4, pp. 617–628, 2004.

115. R. Platt, S. Chen, Y. Zhou, M. Yim, L. Swiech, H. Kempton, J. Dahlman, O. Parnas, T. Eisenhaure, M. Jovanovic, D. Graham, S. Jhunjhunwala, M. Heidenreich, R. Xavier, R. Langer, D. Anderson, N. Hacohen, A. Regev, G. Feng, P. Sharp, and F. Zhang, "CRISPR-Cas9 Knockin Mice for Genome Editing and Cancer Modeling," *Cell*, vol. 159, no. 2, pp. 440–455, 2014.

116. A. L. Bredemeyer, G. G. Sharma, C.-Y. Huang, B. A. Helmink, L. M. Walker, K. C. Khor, B. Nuskey, K. E. Sullivan, T. K. Pandita, C. H. Bassing, and B. P. Sleckman, "ATM stabilizes DNA double-strand-break complexes during v(d)j recombination," *Nature*, vol. 442, no. 7101, pp. 466–470, 2006.

117. H. M. Khoo, G. A. Meaker, and A. C. Wilkinson, "Ex Vivo Expansion and Genetic Manipulation of Mouse Hematopoietic Stem Cells in Polyvinyl Alcohol-Based Cultures," *Journal of Visualized Experiments*, no. 192, 2023.

118. A. C. Wilkinson, R. Ishida, H. Nakauchi, and S. Yamazaki, "Long-term ex vivo expansion of mouse hematopoietic stem cells," *Nature Protocols*, vol. 15, pp. 628–648, 1 2020.

119. G. Balmus, N. A. Karp, B. L. Ng, S. P. Jackson, D. J. Adams, and R. E. McIntyre, "A high-throughput in vivo micronucleus assay for genome instability screening in mice," *Nature Protocols*, vol. 10, pp. 205–215, 1 2015.

120. A. Keller, A. I. Nesvizhskii, E. Kolker, and R. Aebersold, "Empirical Statistical Model To Estimate the Accuracy of Peptide Identifications Made by MS/MS and Database Search," *Analytical Chemistry*, vol. 74, no. 20, pp. 5383–5392, 2002.

121. A. I. Nesvizhskii, A. Keller, E. Kolker, and R. Aebersold, "A Statistical Model for Identifying Proteins by Tandem Mass Spectrometry," *Analytical Chemistry*, vol. 75, no. 17, pp. 4646–4658, 2003.

122. J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard,

J. Bambrick, S. W. Bodenstein, D. A. Evans, C.-C. Hung, M. OâNeill, D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Zemgulyte, E. Arvaniti, C. Beattie, O. Bertolli, A. Bridgland, A. Cherepanov, M. Congreve, A. I. Cowen-Rivers, A. Cowie, M. Figurnov, F. B. Fuchs, H. Gladman, R. Jain, Y. A. Khan, C. M. R. Low, K. Perlin, A. Potapenko, P. Savy, S. Singh, A. Stecula, A. Thillaisundaram, C. Tong, S. Yakneen, E. D. Zhong, M. Zielinski, A. Zidek, V. Bapst, P. Kohli, M. Jaderberg, D. Hassabis, and J. M. Jumper, "Accurate structure prediction of biomolecular interactions with AlphaFold 3," *Nature*, vol. 630, no. 8016, pp. 493–500, 2024.

123. B. Wang, M. Wang, W. Zhang, T. Xiao, C. H. Chen, A. Wu, F. Wu, N. Traugh, X. Wang, Z. Li, S. Mei, Y. Cui, S. Shi, J. J. Lipp, M. Hinterndorfer, J. Zuber, M. Brown, W. Li, and X. S. Liu, "Integrative analysis of pooled CRISPR genetic screens using MAGeCKFlute," *Nature Protocols*, vol. 14, pp. 756–780, 2019.

124. E. Kim and T. Hart, "Improved analysis of CRISPR fitness screens and reduced off-target effects with the BAGEL2 gene essentiality classifier," *Genome Medicine*, vol. 13, no. 1, p. 2, 2021.

125. M. Whitney, G. Royle, M. Low, M. Kelly, M. Axthelm, C. Reifsteck, S. Olson, R. Braun, M. Heinrich, R. Rathbun, G. Bagby, and M. Grompe, "Germ cell defects and hematopoietic hypersensitivity to gamma-interferon in mice with a targeted disruption of the fanconi anemia c gene," *Blood*, vol. 88, no. 1, pp. 49–58, 1996.

126. J. I. Garaycoechea, 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, "Alcohol and endogenous aldehydes damage chromosomes and mutate stem cells," *Nature*, vol. 553, pp. 171–177, 1 2018.

127. A. C. Wilkinson, R. Ishida, M. Kikuchi, K. Sudo, M. Morita, R. V. Crisostomo, R. Yamamoto, K. M. Loh, Y. Nakamura, M. Watanabe, H. Nakauchi, and S. Yamazaki, "Long-term ex vivo haematopoietic-stem-cell expansion allows nonconditioned transplantation," *Nature*, vol. 571, no. 7763, pp. 117–121, 2019.

128. J. G. Doench, N. Fusi, M. Sullender, M. Hegde, E. W. Vaimberg, K. F. Donovan, I. Smith, Z. Tothova, C. Wilen, R. Orchard, H. W. Virgin, J. Listgarten, and D. E. Root, "Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-cas9," *Nature Biotechnology*, vol. 34, pp. 184–191, 2016.

129. E. Kim and T. Hart, "Improved analysis of CRISPR fitness screens and reduced off-target effects with the BAGEL2 gene essentiality classifier," *Genome Medicine*, vol. 13, no. 1, p. 2, 2021.

130. Y. Liu, S. E. Elf, Y. Miyata, G. Sashida, Y. Liu, G. Huang, S. D. Giandomenico, J. M. Lee, A. Deblasio, S. Menendez, J. Antipin, B. Reva, A. Koff, and S. D. Nimer, "p53 regulates hematopoietic stem cell quiescence," *Cell Stem Cell*, vol. 4, no. 1, pp. 37–48, 2009.

131. V. Pant, A. Quintas-Cardama, and G. Lozano, "The p53 pathway in hematopoiesis: lessons from mouse models, implications for humans," *Blood*, vol. 120, no. 26, pp. 5118–5127, 2012.

132. E. Boehm, M. Gildenberg, and M. Washington, "Chapter seven the many roles of PCNA in eukaryotic DNA replication," *The Enzymes*, vol. 39, pp. 231–254, 2016.

133. E. d. Pater, P. Kaimakis, C. S. Vink, T. Yokomizo, T. Yamada-Inagawa, R. v. d. Linden, P. S. Kartalaei, S. A. Camper, N. Speck, and E. Dzierzak, "Gata2 is required for HSC generation and survival," *Journal of Experimental Medicine*, vol. 210, no. 13, pp. 2843–2850, 2013.

134. H. Yoshihara, F. Arai, K. Hosokawa, T. Hagiwara, K. Takubo, Y. Nakamura, Y. Gomei, H. Iwasaki, S. Matsuoka, K. Miyamoto, H. Miyazaki, T. Takahashi, and T. Suda, "Thrombopoietin/MPL signaling regulates hematopoietic stem cell quiescence and interaction with the osteoblastic niche," *Cell Stem Cell*, vol. 1, no. 6, pp. 685–697, 2007.

135. L. A. Thoren, K. Liuba, D. Bryder, J. M. Nygren, C. T. Jensen, H. Qian, J. Antonchuk, and S.-E. W. Jacobsen, "Kit regulates maintenance of quiescent hematopoietic stem cells," *The Journal of Immunology*, vol. 180, no. 4, pp. 2045–2053, 2008.

136. N. Suwaki, K. Klare, and M. Tarsounas, "RAD51 paralogs: Roles in DNA damage signalling, recombinational repair and tumorigenesis," *Seminars in Cell & Developmental Biology*, vol. 22, no. 8, pp. 898–905, 2011.

137. P. Wu, M. v. Overbeek, S. Rooney, and T. d. Lange, "Apollo contributes to g overhang maintenance and protects leading-end telomeres," *Molecular Cell*, vol. 39, no. 4, pp. 606–617, 2010.

138. P. Wu, H. Takai, and T. D. Lange, "Telomeric 3' overhangs derive from resection by exo1 and apollo and fill-in by POT1b-associated CST," *Cell*, vol. 150, pp. 39–52, 7 2012.

139. L. A. Timashev and T. D. Lange, "Characterization of t-loop formation by TRF2," *Nucleus*, vol. 11, no. 1, pp. 164–177, 2020.

140. A. Nussenzweig, C. Chen, V. d. C. Soares, M. Sanchez, K. Sokol, M. C. Nussenzweig, and G. C. Li, "Requirement for ku80 in growth and immunoglobulin v(d)j recombination," *Nature*, vol. 382, no. 6591, pp. 551–555, 1996.

141. I. R. Matei, R. A. Gladdy, L. M. Nutter, A. Canty, C. J. Guidos, and J. S. Danska, "ATM deficiency disrupts tcra locus integrity and the maturation of CD4+CD8+ thymocytes," *Blood*, vol. 109, no. 5, pp. 1887–1896, 2007.

142. A. Celeste, S. Petersen, P. J. Romanienko, O. Fernandez-Capetillo, H. T. Chen, O. A. Sedelnikova, B. Reina-San-Martin, V. Coppola, E. Meffre, M. J. Difilippantonio, C. Redon, D. R. Pilch, A. Olaru, M. Eckhaus, R. D. Camerini-Otero, L. Tessarollo, F. Livak, K. Manova, W. M. Bonner, M. C. Nussenzweig, and A. Nussenzweig, "Genomic instability in mice lacking histone h2ax," *Science*, vol. 296, no. 5569, pp. 922–927, 2002.

143. S. Difilippantonio, E. Gapud, N. Wong, C.-Y. Y. Huang, G. Mahowald, H. T. Chen, M. J. Kruhlak, E. Callen, F. Livak, M. C. Nussenzweig, B. P. Sleckman, and A. Nussenzweig, "53BP1 facilitates long-range DNA end-joining during V(D)J recombination," *Nature*, vol. 456, pp. 529–533, 2008. http://www.nature.com/articles/nature07476.

144. X. Liu, W. Jiang, R. L. Dubois, K. Yamamoto, Z. Wolner, and S. Zha, "Overlapping functions between

XLF repair protein and 53bp1 DNA damage response factor in end joining and lymphocyte development," *Proceedings of the National Academy of Sciences*, vol. 109, no. 10, pp. 3903–3908, 2012.

145. E. Vincendeau, W. Wei, X. Zhang, C. Planchais, W. Yu, H. Lenden-Hasse, T. Cokelaer, J. P. d. Fonseca, H. Mouquet, D. J. Adams, F. W. Alt, S. P. Jackson, G. Balmus, C. Lescale, and L. Deriano, "SHLD1 is dispensable for 53bp1-dependent v(d)j recombination but critical for productive class switch recombination," *Nature Communications*, vol. 13, p. 3707, 6 2022.

146. C. Boboila, C. Yan, D. R. Wesemann, M. Jankovic, J. H. Wang, J. Manis, A. Nussenzweig, M. Nussenzweig, and F. W. Alt, "Alternative end-joining catalyzes class switch recombination in the absence of both ku70 and DNA ligase 4," *Journal of Experimental Medicine*, vol. 207, no. 2, pp. 417–427, 2010.

147. Y. Pommier, Y. Sun, S.-y. N. Huang, and J. L. Nitiss, "Roles of eukaryotic topoisomerases in transcription, replication and genomic stability," *Nature Reviews Molecular Cell Biology*, vol. 17, no. 11, pp. 703–721, 2016.

148. Y. Pommier, A. Nussenzweig, S. Takeda, and C. Austin, "Human topoisomerases and their roles in genome stability and organization," *Nature Reviews Molecular Cell Biology*, vol. 23, no. 6, pp. 407–427, 2022.

149. J. Spiegel, S. Adhikari, and S. Balasubramanian, "The structure and function of DNA g-quadruplexes," *Trends in Chemistry*, vol. 2, no. 2, pp. 123–136, 2020.

150. M. L. Duquette, P. Handa, J. A. Vincent, A. F. Taylor, and N. Maizels, "Intracellular transcription of g-rich DNAs induces formation of g-loops, novel structures containing g4 DNA," *Genes & Development*, vol. 18, no. 13, pp. 1618–1629, 2004.

151. R. Rodriguez, K. M. Miller, J. V. Forment, C. R. Bradshaw, M. Nikan, S. Britton, T. Oelschlaegel, B. Xhemalce, S. Balasubramanian, and S. P. Jackson, "Small-molecule-induced DNA damage identifies alternative DNA structures in human genes," *Nature Chemical Biology*, vol. 8, no. 3, pp. 301–310, 2012.

152. K. Kim, J. Min, T. W. Kirby, S. A. Gabel, L. C. Pedersen, and R. E. London, "Ligand binding characteristics of the ku80 von willebrand domain," *DNA Repair*, vol. 85, p. 102739, 2020.

153. B. Yariv, E. Yariv, A. Kessel, G. Masrati, A. B. Chorin, E. Martz, I. Mayrose, T. Pupko, and N. BenâTal, "Using evolutionary data to make sense of macromolecules with a âfaceâliftedâ ConSurf," *Protein Science*, vol. 32, no. 3, p. e4582, 2023.

154. T. M. Bryan, "G-quadruplexes at telomeres: Friend or foe?," *Molecules*, vol. 25, no. 16, p. 3686, 2020.

155. O. Deze, B. Laffleur, and M. Cogne, "Roles of g4-DNA and g4-RNA in class switch recombination and additional regulations in b-lymphocytes," *Molecules*, vol. 28, no. 3, p. 1159, 2023.

156. J. R. Walker, R. A. Corpina, and J. Goldberg, "Structure of the ku heterodimer bound to DNA and its implications for double-strand break repair," *Nature*, vol. 412, no. 6847, pp. 607–614, 2001.

157. S. Petersen, R. Casellas, B. Reina-San-Martin, H. T. Chen, M. J. Difilippantonio, P. C. Wilson, L. Hanitsch, A. Celeste, M. Muramatsu, D. R. Pilch, C. Redon, T. Ried, W. M. Bonner, T. Honjo, M. C. Nussenzweig,

and A. Nussenzweig, "AID is required to initiate nbs1/g-h2ax focus formation and mutations at sites of class switching," *Nature*, vol. 414, no. 6864, pp. 660–665, 2001.

158. J. I. Garaycoechea and K. J. Patel, "Why does the bone marrow fail in fanconi anemia?," *Blood*, vol. 123, pp. 26–34, 1 2014.

159. K. Savitsky, 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. J. Jaspers, A. M. R. Taylor, C. F. Arlett, T. Miki, S. M. Weissman, M. Lovett, F. S. Collins, and Y. Shiloh, "A single ataxia telangiectasia gene with a product similar to PI-3 kinase," *Science*, vol. 268, no. 5218, pp. 1749–1753, 1995.

160. M. F. Lavin, "The appropriateness of the mouse model for ataxia-telangiectasia: Neurological defects but no neurodegeneration," *DNA Repair*, vol. 12, no. 8, pp. 612–619, 2013.

161. A. Butturini, R. P. Gale, P. C. Verlander, B. Adler-Brecher, A. P. Gillio, and A. D. Auerbach, "Hematologic abnormalities in fanconi anemia: an international fanconi anemia registry study.," *Blood*, vol. 84, no. 5, pp. 1650–5, 1994.

162. K. Parmar, J. Kim, S. M. Sykes, A. Shimamura, P. Stuckert, K. Zhu, A. Hamilton, M. K. Deloach, J. L. Kutok, K. Akashi, D. G. Gilliland, and A. D'Andrea, "Hematopoietic stem cell defects in mice with deficiency of fancd2 or usp1," *STEM CELLS*, vol. 28, pp. 1186–1195, 7 2010.

163. S. Gao, S. Feng, S. Ning, J. Liu, H. Zhao, Y. Xu, J. Shang, K. Li, Q. Li, R. Guo, and D. Xu, "An OB-fold complex controls the repair pathways for DNA double-strand breaks," *Nature Communications*, vol. 9, no. 1, p. 3925, 2018.

164. R. McElligott and R. J. Wellinger, "The terminal DNA structure of mammalian chromosomes," *The EMBO Journal*, vol. 16, no. 12, pp. 3705–3714, 1997.

165. W. E. Wright, V. M. Tesmer, K. E. Huffman, S. D. Levene, and J. W. Shay, "Normal human chromosomes have long g-rich telomeric overhangs at one end," *Genes & Development*, vol. 11, no. 21, pp. 2801–2809, 1997.

166. T. d. Lange, "Shelterin-mediated telomere protection," *Annual Review of Genetics*, vol. 52, no. 1, pp. 223–247, 2018.

167. B. v. Steensel, A. Smogorzewska, and T. d. Lange, "TRF2 protects human telomeres from end-to-end fusions," *Cell*, vol. 92, no. 3, pp. 401–413, 1998.

168. A. Sfeir and T. d. Lange, "Removal of shelterin reveals the telomere end-protection problem," *Science*, vol. 336, no. 6081, pp. 593–597, 2012.

169. A. Bianchi and T. d. Lange, "Ku binds telomeric DNA in vitro *," *Journal of Biological Chemistry*, vol. 274, no. 30, pp. 21223–21227, 1999.

170. H.-L. Hsu, D. Gilley, E. H. Blackburn, and D. J. Chen, "Ku is associated with the telomere in mammals," *Proceedings of the National Academy of Sciences*, vol. 96, no. 22, pp. 12454–12458, 1999.

171. H.-L. Hsu, D. Gilley, S. A. Galande, M. P. Hande, B. Allen, S.-H. Kim, G. C. Li, J. Campisi, T. Kohwi-Shigematsu, and D. J. Chen, "Ku acts in a unique way at the mammalian telomere to prevent end joining," *Genes & Development*, vol. 14, no. 22, pp. 2807–2812, 2000.

172. G. B. Celli, E. L. Denchi, and T. d. Lange, "Ku70 stimulates fusion of dysfunctional telomeres yet protects chromosome ends from homologous recombination," *Nature Cell Biology*, vol. 8, no. 8, pp. 885–890, 2006.

173. B. H. Anderson, P. R. Kasher, J. Mayer, M. Szynkiewicz, E. M. Jenkinson, S. S. Bhaskar, J. E. Urquhart, S. B. Daly, J. E. Dickerson, J. O'Sullivan, E. O. Leibundgut, J. Muter, G. M. Abdel-Salem, R. Babul-Hirji, P. Baxter, A. Berger, L. Bonafe, J. E. Brunstom-Hernandez, J. A. Buckard, D. Chitayat, W. K. Chong, D. M. Cordelli, P. Ferreira, J. Fluss, E. H. Forrest, E. Franzoni, C. Garone, S. R. Hammans, G. Houge, I. Hughes, S. Jacquemont, P. Y. Jeannet, R. J. Jefferson, R. Kumar, G. Kutschke, S. Lundberg, C. M. Loureno, R. Mehta, S. Naidu, K. K. Nischal, L. Nunes, K. Ounap, M. Philippart, P. Prabhakar, S. R. Risen, R. Schiffmann, C. Soh, J. B. Stephenson, H. Stewart, J. Stone, J. L. Tolmie, M. S. V. D. Knaap, J. P. Vieira, C. N. Vilain, E. L. Wakeling, V. Wermenbol, A. Whitney, S. C. Lovell, S. Meyer, J. H. Livingston, G. M. Baerlocher, G. C. Black, G. I. Rice, and J. Yanick, "Mutations in CTC1, encoding conserved telomere maintenance component 1, cause coats plus," *Nature Genetics*, vol. 44, pp. 338–342, 2012.

174. A. J. Simon, A. Lev, Y. Zhang, B. Weiss, A. Rylova, E. Eyal, N. Kol, O. Barel, K. Cesarkas, M. Soudack, N. Greenberg-Kushnir, M. Rhodes, D. L. Wiest, G. Schiby, I. Barshack, S. Katz, E. Pras, H. Poran, H. Reznik-Wolf, E. Ribakovsky, C. Simon, W. Hazou, Y. Sidi, A. Lahad, H. Katzir, S. Sagie, H. A. Aqeilan, G. Glousker, N. Amariglio, Y. Tzfati, S. Selig, G. Rechavi, and R. Somech, "Mutations in STN1 cause coats plus syndrome and are associated with genomic and telomere defects," *Journal of Experimental Medicine*, vol. 213, no. 8, pp. 1429–1440, 2016.

175. A. J. Walne and I. Dokal, "Advances in the understanding of dyskeratosis congenita," *British Journal of Haematology*, vol. 145, no. 2, pp. 164–172, 2009.

176. M. T. Hemann, M. A. Strong, L.-Y. Hao, and C. W. Greider, "The shortest telomere, not average telomere length, is critical for cell viability and chromosome stability," *Cell*, vol. 107, no. 1, pp. 67–77, 2001.

177. S. W. A. Titen and K. G. Golic, "Telomere loss provokes multiple pathways to apoptosis and produces genomic instability in drosophila melanogaster," *Genetics*, vol. 180, no. 4, pp. 1821–1832, 2008.

178. T. T. Schmidt, C. Tyer, P. Rughani, C. Haggblom, J. R. Jones, X. Dai, K. A. Frazer, F. H. Gage, S. Juul, S. Hickey, and J. Karlseder, "High resolution long-read telomere sequencing reveals dynamic mechanisms in aging and cancer," *Nature Communications*, vol. 15, no. 1, p. 5149, 2024.