UNIVERSITY OF CALIFORNIA

Santa Barbara

Interstrand crosslinking of homologous repair template DNA enhances gene editing in human cells

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Molecular, Cellular, and Developmental Biology

by

Hannah Isabella Ghasemi

Committee in charge:

Professor Christopher D. Richardson, Chair

Professor Christopher S. Hayes

Professor David A. Low

Professor Denise J. Montell

September 2023

The dissertation of Hannah I. Ghasemi is approved.

_____

Christopher S. Hayes

_____

David A. Low

_____

Denise J. Montell

_____

Christopher D. Richardson, Committee Chair

September 2023

Interstrand crosslinking of homologous repair template DNA enhances gene editing in human cells

by

Hannah Isabella Ghasemi

Acknowledgments

I want to begin by acknowledging my mentor, Chris Richardson. Chris has empowered me to think like a scientist, to care and to be curious, to embrace my haphazard creative spirit, and to go after what I want unapologetically, something I have always grappled with as a woman of color. I aspire to be as innovative and versatile of a scientist as he and as compassionate of a human being.

To my parents, thank you for sacrificing so much for me to give me my life, the American Dream. You left behind a familiar world to start anew, your bravery and strength drives me. To my baba, thank you for introducing me to a range of music from Pink Floyd to roots reggae, the music you showed me gave me the stamina to push through tough times. Thank you for always reminding me that the sky is the limit and nothing is impossible. To my maman, thank you for pushing me to strengthen the right side of my brain, to exercise my creative muscles, and for showing me the allure of art and painting. Thank you, Mom and Dad, for embracing me as your only child, for taking me high when I felt low after days of getting knocked down, without your unconditional love and belief in me, I would not be here. To my late grandmother, thank you for raising me in my formative years, and for instilling our culture into my soul with your magical food, language, and warmth.

To my dear friends, family, and comrades, thank you for laughing and crying with me through some trying times, for continually reminding me that I was not alone, for giving me a reason to get out of bed every day, and for opening your arms to me, you have pushed and challenged me to grow.

To the brave women and men in Iran, thank you for your courage, you inspire me to be a woman, to be a fighter, and to never give up no matter how dark it gets, your fire burns eternally.

I would also like to take a moment to thank my thesis committee, Drs. Chris Hayes, David Low, and Denise Montell, for providing abundant guidance and support throughout my Ph.D.

Curriculum Vitae of Hannah Ghasemi

University of California, Santa Barbara
Department of Molecular, Cellular, and Developmental Biology

Education:

University of California, Santa Barbara: **2019-2023**
Doctor of Philosophy in Molecular, Cellular, and Developmental Biology

University of Colorado, Boulder: **2015-2019**
B.A. in Ecology & Evolutionary Biology | Minor in Spanish.

Awards and Honors:

Latin Honors – Magna cum laude | *University of Colorado, Boulder*
Undergraduate Honors Thesis | *University of Colorado, Boulder*
Biological Sciences Initiative Scholar | *University of Colorado, Boulder*
Cumulative GPA 4.0 | *University of California, Santa Barbara*
NSF GRFP Honorable Mention | *University of California, Santa Barbara*

Research Experience:

Doctoral Research:

University of California, Santa Barbara: 2019-2023

Research Advisor - Christopher Richardson, Ph.D.

- Pioneering and leading publication of a manuscript on the optimization of template DNA for CRISPR-Cas9 gene-editing.
- Deciphering the mechanism of crosslink-stimulated recombination using a combination of proteomic, genomic (pooled screens), and candidate-based approaches.

Undergraduate Research:

University of Colorado, Boulder: 2018-2019

Research Advisor - David Stock, Ph.D.

- Analyzed the extent to which an enhancer was co-opted in the evolution of morphologically novel bony armor plates in Stickleback fishes from the ancestral precursor of scales.

University of Colorado, Boulder: 2018-2019

Research Advisor – Julienne Ng, Robert Laport, Stacey Smith, Ph.D.

- Collected trait measurements from a variety of herbarium species using a Machine Learning program to assess trait differentiation amongst woody perennial species in Harvard Forest, as a part of the phylogenetic community project.

Publications & Patents:

1. Ghasemi, H.I., Bacal, J., Yoon, A.C. *et al.* Interstrand crosslinking of homologous repair template DNA enhances gene-editing in human cells. *Nat Biotechnol* (2023). https:// doi.org/10.1038/s41587-022-01654-y
2. U.S. Provisional Patent Application No. PCT/US2022/072020 Systems and Methods for Genetic Editing with Increased Efficiency Inventors - Hannah I. Ghasemi, Chris Richardson.
3. Ghasemi, HI, Stock, D. Evolution of gene regulation in association with the origin of armor in Stickleback fishes. Spring 2019.

Presentations:

Oral Presentations:

Hannah Ghasemi. "Interstrand Crosslinks Boost Homologous Recombination in Human Cells." | SoCal Genome Stability Symposium | University of California, Irvine – December 2022

Hannah Ghasemi. "Interstrand crosslinking of homologous repair template DNA enhances gene-editing in human cells." | Genome Engineering Seminar | Harvard Medical School – March 2023

Poster Presentations:

Hannah Ghasemi. "Interstrand Crosslinks in HDR Template DNA Boost Gene-editing in Human Cells." | MCDB Retreat | University of California, Santa Barbara – October 2022

Hannah Ghasemi. Interstrand Crosslinks in Donor DNA Boost Gene-editing in Human Cells | Precision Genome Engineering Keystone Symposia – April 2022

Teaching Experience:

Graduate Student Teaching Assistant

University of California, Santa Barbara – Introductory Bio Lab: Winter 2019; Molecular Genetics: Winter 2022.

- Held weekly discussion sections along with office hours.

Mentoring:

Undergraduate Students mentored:

Amanda Yoon – March 2021 – January 2023
Sofia Gottihold – February 2023 – July 2023

Graduate Students mentored:

Keile Hansen – September 2022 – December 2022

Abstract

Interstrand crosslinking of homologous repair template DNA enhances gene editing in human cells

by

Hannah Isabella Ghasemi

Non-viral gene editing strategies have been developed for basic science and therapeutic applications. Major gains in efficiency have come from the optimization of nuclease activity and delivery, with comparatively little focus on the role of the homology-directed repair template (HDRT) in this process.

Here, we will describe a strategy to boost the efficiency of gene editing via homology-directed repair (HDR) by covalently modifying the template DNA molecule with interstrand crosslinks. Crosslinked templates (xHDRTs) increase Cas9-mediated editing efficiencies by up to fivefold in K562, HEK293T, U2OS, iPS, and primary T cells. Increased editing from xHDRTs is driven by events on the template molecule and requires ataxia telangiectasia and Rad3-related (ATR) kinase and components of the Fanconi anemia (FA) pathway.

Chapter I

Introduction

General Introduction

The alteration of genetic sequences in eukaryotic cells has been made possible by the wide development of programmable nucleases. This has empowered the study of human genes and their precise functions, the correction of disease-causing mutations, and the construction of novel receptors on immune cells to target cancer.

The continued advancement of CRISPR-Cas9 technology has enabled scientists to effectively alter the genomes of human cells in a highly accessible and facile manner[1]. The means by which we can genetically alter human cells are now well-established, not to exclude the innovation of alternative editing strategies (i.e., base editing and prime editing)[2,3].

It is, however, essential to recognize that gene editing would not be possible without a cell's intrinsic DNA repair capabilities. At its core, gene-specific integrations and manipulations contend upon the introduction of a double-strand break (DSB), a severing of the two DNA strands coordinated by a targeted nuclease. The co-introduction of a DNA molecule then enables the replacement of the cleavage site by homologous recombination (HR), an endogenous means of genetic exchange. And yet, many things are not fully understood about how the

cell carries out repair processes on a molecular basis. Namely, how are exogenously supplied DNA templates sensed? Why does the cell preferentially utilize an exogenously supplied DNA template instead of its own sister chromosome? Moreover, how does this knowledge influence pathway choice or, in simpler terms, gene-editing efficiency?

Gene-editing tools were, in part, and continue to be developed out of the study of cellular DNA repair mechanisms, and in turn, researchers can effectively study DNA repair through the strategic introduction of DNA damage mediated by gene-editing tools. The resolution of this damage can then be monitored, giving us a glimpse into how the cell resolves DNA damage on a molecular level. The advancement of both fields, genetic engineering and DNA repair, is thus a mutually beneficial pursuit.

Gene-editing

Targeted Nucleases

The earliest DNA nucleases developed in the gene-editing space were meganucleases and zinc finger nucleases (ZFNs), followed by the development of transcription activator-like nucleases (TALENs)[4].

Meganucleases, though naturally occurring and abundant, are engineered restriction enzymes, limited in generalizability by their

unique 14-40 base pair (bp)-long recognition sites[5,6]. Later developed ZFNs and TALENs consist of the chimeric fusion of a nonsequence-specific Fok I endonuclease cleavage domain to a tunable DNA binding domain, either zinc-finger modules or TALE modules respectively, making them significantly more scalable[5–7].

In years to follow, born out of the bacterial adaptive immunity field, came the expansion of a new flexible gene-editing tool, clustered regularly interspaced palindromic repeat (CRISPR)-associated 9 (Cas9) nuclease. This programmable enzyme could be targeted to any genetic sequence in the human genome through its association with a short customizable RNA molecule.

CRISPR-Cas9

Two classes of CRISPR/Cas systems exist: class I (type I, III, and IV) and class II (type II, V, and VI). Multi-subunit Cas-protein complexes, containing multiple Cas proteins, occupy the class I systems, while class II systems employ a single multidomain Cas protein[8]. The type II CRISPR/Cas9 system has become the most expansive and well-developed CRISPR system to be used as a gene-editing tool, primarily due to the relatively uncomplicated structure of the Cas9 protein.

Composed of two lobes, the recognition (REC) and nuclease (NUC), the Cas9 protein can be escorted to any site in the human genome through its association with a two-RNA structure, a mature CRISPR RNA (crRNA) hybridized to a trans-activating CRISPR RNA (tracrRNA)[9,10]. The crRNA specifies complementary DNA sequences by physically binding them through Watson & Crick base pairing, while the tracrRNA, a long stretch of stem loops and anti-repeat sequences, acts as a molecular scaffold, a liaison between the Cas9 nuclease and the target DNA sequence[10].

The CRISPR/Cas9 system has been streamlined to be directed by a single guide RNA (sgRNA) chimera, a fusion of tracrRNA and crRNA[11]. The REC lobe of the Cas9 protein, composed of two domains, REC1 and REC2, binds the complementary region of the sgRNA, while the NUC lobe, composed of three domains, two nuclease domains: RuvC and HNH, and the Protospacer Adjacent Motif (PAM) interacting domain, coordinates the nuclease activities of the enzyme[10].

Upon RNA-DNA base pairing by the sgRNA, a targeted DSB can be subsequently introduced at a cut site 3bp upstream of the PAM, a 2-5bp conserved DNA sequence recognized by Cas9 and bound by its PAM interacting domain at an NGG consensus sequence, in which N can be any of the four nucleotides[9].

CRISPR-Cas9 as a gene-editing tool

It is well-recognized that the CRISPR-Cas9 system is remarkably effective at introducing site-specific double-strand breaks in the human genome. As such, the CRISPR-Cas9 system has been adapted into a highly successful gene-editing tool.

While DSBs are considered an alarming cellular insult and are highly detrimental if not effectively repaired in a host cell, DSBs are required for routine cellular maintenance and health. A DSB must occur for homologous recombination, the exchange of genetic information, to endogenously unfold. Accordingly, DSBs are intentionally and deliberately produced by a topoisomerase II-like enzyme, Spo11, in meiotic cells to promote the exchange of genetic material and chromosome segregation[12,13].

Introducing a DSB into the human genome prompts the rapid resolution of this cellular insult. This repair, though often error-prone when resolved through an end-joining process, is useful in disrupting or silencing genes of interest.

To reduce the frequency of end-joining repair processes, an exogenously supplied DNA template, flanked with sequence homology to the break site, can be co-introduced with the Cas9 complex. This co-introduced template increases the frequency at which the cell will resolve the nuclease-mediated break through homologous recombination, in

which the information encoded in the DNA template is copied over on a genomic level. Methodologies that encourage HR outcomes over end-joining outcomes thus remain heavily investigated in the field, as HR processes can support targeted integrations into the human genome. Understanding the mechanics of how HR and EJ processes unfold progresses novel methodologies for improved editing.

DNA Damage and Repair

Given the urgency to effectively resolve damaged DNA, mammalian cells have evolved rigorous repair networks that can productively repair an array of DNA damage, including DSBs. The two major constituents of the DNA damage response (DDR) are 1) recognition of the DNA damage or lesion, followed by 2) the tight coordination of a DNA repair response (DDR), mediated by signaling cascades, namely, protein phosphorylation, sustained by delays in cell-cycle progression[14].

Four semi-independent DSB sensors can effectively recognize and initiate a DNA repair response in mammalian cells: Ku70/Ku80 (referred to as Ku hereafter), Poly (ADP-ribose) polymerase (PARP1/2), MRE11-RAD50-NBS1 (MRN), and Replication Protein A (RPA)[15], all of which will be expanded upon in the text to ensue.

All eukaryotes express at least one of three members of a family of related kinases to orchestrate the activation of the DDR; ataxia-

telangiectasia mutated (ATM), ATM- and Rad3-related (ATR), and the DNA-dependent protein kinase (DNA-PK), all of which regulate the cell cycle through the inhibition and tight regulation of cyclin-dependent-kinase (CDK) activity[14].

A central role of ATR is to phosphorylate and activate the protein kinase CHK1 to promote the degradation of CDC25a. Degradation of CDC25a prevents the removal of inhibitory signals from CDKs, thus delaying the cell cycle and gracing the cell with more time to resolve any damage that may have activated ATR in the first place[14]. Unlike its siblings (ATM and DNA-PK), ATR recruitment to chromatin is not just limited to DSBs; ATR will also uniquely activate the Fanconi anemia pathway (among its 700 other substrates), a pathway known to resolve DNA interstrand crosslinks (ICLs)[14]. Simply put, ATM, ATR, and DNA-PK, activated by DSBs, phosphorylate a range of substrates to initiate the DDR; this activation then choreographs the recruitment of DNA repair factors and cell cycle arrest[16].

Double-strand breaks

DNA double-strand breaks (DSBs) are undesirable cellular lesions in which both strands of the double helix become severed, leaving a discontinuous gap within the double-stranded DNA (dsDNA) helix. DSBs can spawn from either endogenous or exogenous sources of DNA

damage, in which a replication fork may stall upon encountering a complex DNA lesion (introduced by DNA damage), triggering fork collapse and, consequently, DSB formation.

The major endogenous source of DNA damage that commonly contributes to DSBs is reactive oxygen species (ROS), a byproduct of cellular metabolism, while exogenous sources of DNA damage that contribute to DSBs include, but are not limited to, ionizing radiation, UV light exposure from sunlight, and crosslinks introduced by front-line chemotherapeutic agents[15,17,18]. An inability to resolve DSBs or the lesions that promote replication fork stalling can encourage genomic instability, abnormalities, or gross chromosomal rearrangements (GCRs), associated with aging, cancer, neurodegeneration, and cell death[18–21].

Double-strand break repair

To resolve a DSB, mammalian cells will most commonly execute one of two repair pathways: (1) The two ends of the break can be rapidly re-ligated back together in what is known as an end-joining process, (2) or the cell can utilize a separate homologous sequence of DNA, typically encoded within a sister chromatid, to copy the information spanning the break and resolve it in a process called homology-directed repair (HDR)[18]. Because this ladder activity, HDR, relies on a separate DNA molecule or template, an opportunity to incorporate targeted sequences at the site of

the break presents itself, making HDR extremely useful from a gene-editing perspective.

The genome can be precisely altered through the induction of a site-specific DSB, carried out by a targeted nuclease. This nuclease-induced DSB must then resolve through an HDR pathway to effectively incorporate target DNA sequences, supplied on an exogenously supplied template, flanked with sequence homology to the break site. Albeit, the cell can resolve the nuclease-induced DSB through either an end-joining or homology-directed repair process. In fact, classical non-homologous end-joining (cNHEJ) is the preferred pathway for DSB repair (DSBR) in adult mammalian cells[20].

Non-homologous end joining

Non-homologous end joining (NHEJ) is comprised of two subsets of repair: classical non-homologous end-joining (cNHEJ), the dominant NHEJ pathway, and microhomology-mediated end-joining (MMEJ), also known as alternative-NHEJ (alt-NHEJ). A considerable distinction between these two subsets is DSB end resection. cNHEJ does not require the nucleolytic processing of the free ends at the site of the break, whereas end resection is a prerequisite for both alt-NHEJ and HDR[15,22].

cNHEJ begins upon the rapid binding of the Ku heterodimer to the free ends of a double-strand break[15,17,18,23,24]. Bound Ku facilitates the

loading and activation of the catalytic subunit of DNA-dependent protein kinase (DNA-PKcs) to form an end-recognition complex (DNA-PK). This complex acts to stabilize DNA termini, thought to protect the DSB ends from short-range resection by the MRN complex[15]. The binding of Ku to DSB ends was previously thought to "compete" with HR by steering subsequent repair events towards NHEJ outcomes[25–27], while more recent models propose that the DNA-PK complex, together with Ku, will rapidly bind all DSB ends. This bound DNA-PK complex, a roadblock to HR, can then be excised by the endonucleolytic activity of the MRN complex[28]. Following DNA-PKcs loading, the downstream cNHEJ factor, DNA ligase IV (XRCC4/LIG4), is recruited to the break site to re-seal the broken DNA ends[17,18].

Poly (ADP-ribose) polymerase I (PARP1) promotes alt-NHEJ, an alternative to cNHEJ. Notably, PARP1 has been characterized to participate in the initial accumulation of the MRE11/RAD51/NBS1 (MRN) complex to DSBs, encouraging short-range resection events to be carried out[23,29]. The MRN complex, activated by the C-terminal binding protein interacting protein (CtIP), resects short stretches of dsDNA; this can be followed by more extensive end processing, observed in HR, facilitated by long-range nucleases such as exonuclease 1 (EXO1) and DNA replication helicase/nuclease 2 (DNA2)[15,21,28].

Homology-directed repair

Like alt-EJ, HDR is also initiated by the initial short-range resection of DSB ends by the MRN complex (stimulated by CDK-phosphorylated CtIP)[30], followed by an extensive long-range resection event coordinated by EXO1 or DNA2 and BLM1[16]. These subsequent resection events leave behind 3' single stranded DNA (ssDNA) overhangs, which become rapidly coated with Replication Protein A (RPA). RPA protects these ssDNA strands, susceptible to additional breakage, while acting as a landing platform for other downstream repair factors, such as RAD51, which displaces RPA[22,31].

RAD51 assembles into nucleoprotein filaments on the ssDNA overhangs, coordinating the RAD51-mediated strand invasion of homologous DNA sequences encoded within a template molecule, commonly a sister chromosome. This strand invasion forms triple-stranded D loop structures, in which two strands of the DNA helix become partially separated, held apart by a third strand of DNA. This D loop structure is later cleaved by the MUS81-EME1 endonuclease complex or is displaced by the regulator of telomere elongation helicase 1 (RTEL1)[15,22].

In short, HR resolves DSBs through an extensively resection-dependent process in which a homologous sequence of DNA is invaded

and used as a template to copy the information at the break site, restoring the original sequence identity relatively error-free.

DSB Pathway Choice

Many elements have been shown to influence DSB repair pathway preference, and many observations are still being developed in the field. The ability to promote HDR outcomes by influencing the cell's preference is advantageous from a gene-editing perspective.

As discussed above, resection of the free DNA ends at a double-stranded break is a prime step in modulating the decision between different DSB repair pathways. DNA end resection is considered a terminal commitment step in HR, necessary for MMEJ and HR to occur while antagonizing cNHEJ. This cNHEJ antagonization is likely made possible by the incapacity of the Ku heterodimer, a pivotal cNHEJ factor, to sufficiently bind the ssDNA overhangs left behind by MRN-mediated DSB end processing[32].

Another variable that influences DSB pathway choice is the cell cycle. While cNHEJ operates throughout every phase of the vertebrate cell cycle, HR is restricted to the S and G2 phases of the cell cycle[25]. More recently, it has been proposed that template accessibility is a major factor driving these cyclical fluctuations in pathway preference[19]. During

S phase up until anaphase, sister chromatids, the templates used to carry out HR, are closely associated and thus highly accessible if and when HR is to unfold. Additionally, cycle-dependent kinase (CDK) activity, up-regulated during the S phase of the cell cycle, activates resection machinery and downstream HR factors, encouraging HR outcomes during the S phase of the cell cycle, bleeding into G2[21,33–35].

While past and more recent studies have successfully established end resection, cell cycle, and homolog accessibility as considerable factors involved in cellular pathway preference post-DSB, several questions still beckon: How does the cell *sense* a proximal homologous template? How does it perform the "homology search"? Furthermore, why does the cell preferentially utilize an exogenously supplied, bacterially derived DNA temple instead of its own sister chromosome in pursuit of resolving a Cas9-induced DSB through homologous recombination?

Complex DNA lesions

Earlier in the text, we discussed how complex DNA lesions can contribute to DSB formation. Let us hone in on one such form of DNA damage, a DNA interstrand crosslink, or ICL for short. DNA interstrand crosslinks (ICLs) tether the two strands of the DNA backbone. These bulky lesions distort the double helix, impeding the two DNA strands

from separating, a requisite for basic cellular processes like DNA replication and transcription. We can thus classify ICLs as complex DNA lesions in that the covalent linkage formed by an interstrand crosslink precludes simple excision and repair synthesis pathways. Furthermore, in all replicating cells, the confrontation of an ICL during replication will ultimately result in a DSB[36,37].

Interstrand Crosslinks

The most common source of ICLs is endogenous aldehydes, a byproduct of lipid peroxidation, thought to be more prevalent in individuals with alcoholism or a high-fat diet[38,39]. ICLs can also be exogenously administered through interstrand crosslinking agents, developed as front-line chemotherapeutic agents in treating leukemia and solid tumors[40].

Interstrand crosslinking agents come in several different flavors. The four major classes of ICL-inducing drugs are alkylating agents (born out of WWII), platinums, mitomycin C (MMC) and furocoumarins, and psoralens[40]. These compounds bombard DNA with interstrand crosslinks, displaying a broadly similar mechanism of action in doing so, given that two chemically active leaving groups are required for an ICL to chemically form[40]. However, these compounds exhibit different base specificities and degrees of distortion and thus may elicit different cellular responses. For

example, psoralens will specifically generate covalent bonds to thymines on opposing DNA strands at TA sequences[41], while cisplatin-mediated ICLs distinctly occur at GC sequences, binding the N7 of guanine[42]. Much like DSBs, a systemic failure to resolve ICLs can have severe repercussions on a cellular and patient level, predisposing individuals to bone marrow failure and the development of cancer[43].

Interstrand Crosslink Repair

The Fanconi Anemia Pathway

The known interstrand crosslink (ICL) repair pathways in eukaryotic cells are resolved through either replication or transcription-coupled processes. This comes as no surprise, given that ICLs present a major roadblock to both processes. Replication-coupled ICL repair proceeds through the Fanconi anemia (FA) pathway, while transcription-coupled ICL repair occurs by Nucleotide Excision Repair (NER)[37,44]. Multiple lines of evidence suggest that ICL repair processes carried out by mammalian cells during DNA replication in S phase feature DSBs as repair intermediates and use HR machinery[45]. ICL repair pathways studied in Xenopus egg extracts have shown that replication-coupled ICL repair processes ensue when two replication forks converge at an ICL. This

replication-dependent fork collision prompts ATR checkpoint activation and the activation of the Fanconi anemia pathway[44].

The Fanconi anemia (FA) pathway was born out of the study of Fanconi anemia, a rare genetic disorder distinguished by early bone marrow failure, increased susceptibility to cancer, and decreased life expectancy[41]. FA presents in individuals with inactivating mutations in any of the 20 established FA genes[43]. Most reported cases of FA derive from mutations in *FANCA*, *FANCC,* and *FANCG*, all members of the FA core complex[40,46].

On a molecular level, the FA pathway has been best established to resolve interstrand crosslinks, the relevant lesion discussed in the previous section. Accordingly, FA-derived fibroblasts exhibit hypersensitivity to interstrand-crosslinking agents like mitomycin C; thus, traditional front-line chemotherapeutics prove to be useless in treating FA patients with cancer[36,40].

Each of the 19 FA genes encodes for FA/FANC proteins, which together orchestrate a three-step process in resolving interstrand crosslinks: recognition of the lesion, DNA incision, and, last, lesion bypass and repair[43]. The 19 FA proteins can be separated into three major functional groups: (1) there is the FA core complex, which

monoubiquitinates a FANCD2-FANCI heterodimer; (2) the FANCD2/FANCI heterodimer, which recruits (3) a series of effector proteins downstream of it to repair the ICL.

The FA core complex is composed of eight FANC proteins (FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL, and FANCM) along with three associated proteins (FAAP20, FAAP24, and FAAP100). The core complex is thought to play a pivotal role in lesion sensing and the activation of the FANCD2-FANCI heterodimer, the heart of the FA pathway.

The aforementioned core of the FA pathway, the FANCD2-FANCI heterodimer, is recruited to ICL lesions upon its monoubiquitination by the FA core complex, namely FANCL and FANCT (UBE2T), a ubiquitin-conjugating enzyme[40,41]. This serves as a crucial step in repairing the ICL. If this monoubiquitination event fails to occur, the ICL will remain unresolved. More recently, it was reported that these monoubiquitination events specifically stabilize the FANCD2-FANCI heterodimer onto dsDNA, allowing it to form filament-like arrays on long stretches of dsDNA[29]. It has also been proposed that the MRN complex, responsible for short-range resection in both alt-NHEJ and HR processes, regulates FANCD2 stability and function[45].

Downstream of FANCD2-FANCI recruitment, a series of effector proteins are recruited to repair the ICL in a complex mechanism that has yet to be fully understood.

The Fanconi Anemia Pathway and Homologous Recombination

The Fanconi anemia pathway engages in cross-talk between several classical ICL repair pathways, including the mutagenic translesion synthesis (TLS), homologous recombination (HR), and nucleotide excision repair (NER) to organize the detection and repair of genomic ICLs[41]. It has been proposed that the FA pathway can be separated into two branches, in which a FANCD2-I ubiquitination branch engages HR outcomes while a core complex-dependent branch encourages the activity of TLS; this model can be further corroborated by an upregulation of TLS activity in response to FANCD2-I deficiency[47].

Coming full circle, the resolution of ICLs through the Fanconi anemia pathway can require the co-option of canonical homologous recombination machinery. BRCA1, BRCA2, BRIP1, PALB2, RAD51, and RAD51C are among this class of proteins, playing a dual role in ICL repair through the FA pathway and HR[48]. Activated (monoubiquitinated) FANCD2 has been shown to colocalize with BRCA1, BRCA2, and RAD51[49].

The Fanconi anemia pathway, particularly the FA core complex and FANCD2-FANCI heterodimer, has also been characterized to play an indispensable role in homologous recombination, in addition to its roles in ICL repair and the protection of stalled replication forks[45,50,51].

Summary of Introduction

Altogether, the precise repair of an array of DNA damage, DSBs or interstrand crosslinks, can dictate life or death on a molecular and cellular basis. Accordingly, the mammalian cell is stocked with an arsenal of repair machinery, at the ready to combat lesions and obstacles scattered throughout the human genome.

We have discussed the significance of homology-directed repair, a DSB repair pathway, its usefulness in altering DNA sequences in the pursuit of treating mutation-derived diseases and disorders, as well as the Fanconi anemia pathway and its concurrent role in HDR and the resolution of interstrand crosslinks, a complex DNA lesion.

The following thesis will discuss a strategy to enhance or increase the frequency of HDR outcomes post Cas9-mediated DSB, providing a method by which non-viral gene-editing efficiency can be improved.

Current gene therapies available in the United States commonly involve viral vectors to deliver gene-editing reagents into human cells. While effective, viral vectors are costly to produce and have recently

presented a series of safety concerns. Thus, the expansion of strategies that enhance non-viral gene-editing efficiency will continue to behoove the entire gene-editing community, allowing scientists to make desired changes to a cell more rapidly, enabling faster scientific advancements and discoveries, and increasing the likelihood of success with impactful treatments.

The thesis to follow will describe the general observation made during my Ph.D. rotation in the Richardson lab that covalent modifications made to the DNA template, specifically interstrand-crosslinking of the template molecule used in Cas9-based gene-editing workflows, boost editing efficiencies by up to five-fold. We establish this observation as highly generalizable, being that it occurs in a number of cell types and that it is not a locus-specific or a topology-specific phenomenon. This generalizability compelled us to pursue the mechanism by which this enhancement in efficiency occurs.

Chapter II

Interstrand crosslinking of homologous repair template DNA enhances gene-editing in human cells

Abstract

Co-introduction of targeted nucleases and DNA/RNA templates encoding new genomic sequence is the basis for rapid, effective, and iterable gene-editing workflows for therapeutic and basic science applications. Extensive optimization of reagent delivery and nuclease activity have improved genome editing workflows, but comparatively few efforts have been made to alter the gene-editing activity of template molecules. Here, we report template DNA modified with interstrand crosslinks (ICLs) – xHDRTs - increases editing frequencies in Cas9-directed gene-editing workflows by up to five-fold. xHDRTs increase gene-editing frequencies independent of DNA template topology, amount of sequence added, or cell type. xHDRT-stimulated gene-editing boosts yields of edited primary T-cells by approximately three-fold. Gene-editing using xHDRTs requires the DNA repair kinase, ATR, and components of the Fanconi anemia pathway but is independent of other ICL-repair pathways. Covalent modification of donor DNA thus presents a compelling opportunity to improve nonviral gene-editing workflows.

Note: The present chapter incorporates data reprinted from the publication listed below of which I was first collaborating author: Ghasemi HI, Bacal J, Yoon AC, Tavasoli KU, Cruz C, Vu JT, Gardner BM,

Richardson CD. Interstrand crosslinking of homologous repair template DNA enhances gene-editing in human cells. *Nat Biotechnology* (2023).

Introduction

CRISPR/Cas9 enables gene-editing via DNA double-strand break (DSB) generation and subsequent activation of cellular DNA repair pathways. Depending on the repair pathway that is engaged, outcomes can include disruption of the targeted gene or replacement with new sequence that restores or introduces functionality[52]. These latter gene replacement events require the delivery of template DNA encoding new sequences to levels that support gene replacement but do not adversely affect cell viability. In translational applications, template molecules are often delivered by viral vectors. While effective, viral workflows are expensive, difficult to scale, and potentially toxic to cells. Use of non-viral template DNA is thus an appealing alternative, but the efficiency and acute toxicity of non-viral templates can be inferior to viral delivery[53]. Improved non-viral gene-editing would be a powerful approach to unravel DNA repair mechanisms, a useful laboratory technique, and a promising strategy for the treatment of a multitude of diseases[54].

One high efficiency non-viral gene-editing strategy co-delivers ribonucleoprotein (RNP) formulations comprising the targeted nuclease

Cas9, a single guide RNA (sgRNA), and a template molecule that contains homology to the region being edited as well as the sequence to be modified or inserted[55]. These RNPs introduce DSBs at targeted regions in the genome, which are then repaired by error prone end joining (EJ) processes that rejoin the ends of the break, or homology-directed repair (HDR) processes that resolve DSBs using sequence encoded in a separate template molecule[56] (**Fig. S1A**). Use of HDR to introduce new DNA sequence into targeted locations enables exciting gain-of-function applications[57]. Strategies to increase HDR frequency may therefore improve outcomes and decrease costs in laboratory and biomedical workflows.

Gains in non-viral HDR efficiency have been achieved through optimization of editing reagents, including protein engineering of Cas9 and related nucleases[6], improving delivery of reagents into cells[58], biophysical optimization of RNP parameters[59], optimization of size and orientation of the homology region of template DNA[60,61], and tethering template to editing reagents[62–64]. Parallel lines of research have focused on defining the cellular response to editing reagents with the goal of redirecting repair events through desired repair pathways[65,66]. These studies have developed key insights into DNA repair processes that underlie gene-editing, but with few exceptions[67,68], it has been hard to translate this understanding into treatments that bias DSB repair towards

desirable outcomes. One limitation may be an inability to upregulate DNA repair processes that contribute to DSB repair. For example, we and others demonstrated that non-viral gene-editing requires the Fanconi anemia (FA) pathway and that these FA proteins localize to DSBs[65,69,70]. However, overexpression of key FA genes failed to increase HDR beyond frequencies seen in control strains[65].

We reasoned that adding substrates for desired DNA repair pathways to template DNA would be an effective approach to activate desired DNA repair activities. Here we report that adding interstrand crosslinks (ICLs) – substrates for the FA DNA repair pathway – to template DNA stimulates HDR by approximately three-fold on a per mole basis in human cell lines, iPS cells, and stimulated T-cells, without increasing mutation frequencies or altering EJ repair outcomes.

Results

We adapted a nonviral gene-editing workflow to measure the effect of covalent modification of double-stranded HDR templates (HDRTs) on gene-editing efficiency. Interstrand crosslinks (ICLs) added to an HDRT – which we refer to as xHDRTs – dramatically improve editing rates in non-viral gene-editing workflows in a dose-dependent manner (**Figure 1A**). ICLs are perturbing DNA lesions, which covalently tether both DNA strands together, and are repaired in human cells by replication- and

transcription-coupled mechanisms[71–73]. Common crosslinking agents include psoralen, which crosslinks TA sites[74], and cisplatin, which crosslinks GC sites[75]. Both psoralen and cisplatin crosslinking reagents stimulate HDR when used to make xHDRTs, suggesting that the HDR stimulation is general to ICLs and not to a specific chemistry (**Figures 1A and S1B**). Psoralen crosslinking requires long-wave UV irradiation, thus unreacted psoralen cannot cause genomic ICLs in cells (where no UV exposure occurs), so we prioritized the development of psoralen-derived xHDRTs. Incubation of HDRTs with varying concentrations of psoralen and 365nm UV radiation creates xHDRTs that increase integration of GFP into the *HBB* locus of human cells ~three-fold (**Figure 1A**). This effect was not caused by transcription from the template molecule, as psoralen ICLs inhibit transcription from reporter genes expressed on the xHDRT (**Figure S1C**). Nor was this effect caused by nonspecific integration of donor sequence into the genome, as xHDRTs that attach GFP to the N-terminus of *LMNB1* produce signal consistent with the fusion protein, and side products indicative of frequent off-target insertion do not appear in the edited samples (**Figure S2A**). Addition of xHDRTs to cells causes a slight enrichment of cells in the G2 phase of the cell cycle over asynchronous controls, but this is indistinguishable from cells treated with uncrosslinked templates (**Figure S2B**). We note that HDRTs containing primarily thymidine dimers[76] caused by longwave UV radiation

did not support elevated levels of HDR (**Figure 1A, 0μM(UV)**), and so increased editing is specific to ICLs and not nonspecifically caused by damaged donor DNA. Overall, xHDRTs can be used in existing gene-editing workflows to boost HDR by approximately three-fold on a per-mole basis.

Psoralen crosslink density is a function of the TA content of the DNA, the psoralen concentration, and the UV dosage and may thus vary between HDRTs. To estimate the optimal number of ICLs per xHDRT, we developed a qPCR-based assay that approximates the number of crosslinks within a given DNA molecule (**Figure S2C**). Using primers that amplify a 94 base pair region of the HDRT plasmid backbone, we determined the probability that at least one crosslink has been introduced in this region. We calculated the ratio (expressed as ΔCt) of qPCR signal produced from xHDRTs generated with different psoralen concentrations or uncrosslinked templates. The editing activity of xHDRTs relative to uncrosslinked controls peaked at three-fold, which occurs at a mean ΔCt value of 4.5 (**Figure 1B**). This translates to an average crosslink density of ~60 crosslinks per xHDRT (**Figure S2C**). These parameters were consistent for xHDRTs homologous to the *HBB* and *RAB11A* loci.

To define the generalizability of our xHDRTs, we tested these constructs in the context of different donor DNA topologies and

sequences. xHDRTs boost gene-editing in the context of linear and circular double-stranded molecules, and for HDR payloads including three nucleotide SNPs (~five-fold), GFP-tag constructs (~two-fold), and promoter-reporter constructs (~three-fold) in K562 cells (**Figure 2A**). To validate our approach in other human cell lines, we confirmed that xHDRTs increase HDR by ~two-fold as compared to an uncrosslinked template in additional cell lines, including UMSCC1 and HEK293T cells (**Figure 2B**). We also validated that xHDRTs stimulate HDR in iPS cells (~three-fold) (**Figure 2C**), which are useful cells for regenerative medicine applications. Our overall conclusion is that xHDRTs boost gene-editing in multiple payloads and target cell types.

We subsequently tested xHDRTs in near-therapeutic T-cell editing workflows. xHDRTs increased the final edited cell yield (**Figure S3D**) by ~3 fold compared to uncrosslinked templates (**Figure 2D**). Edited cell yield measures the number of edited cells seven days after nucleofection and thus incorporates editing percentage as well as toxicity or transient cell cycle arrest caused by editing reagents. To optimize cell yield, we tested multiple doses of crosslinked or uncrosslinked linear template. Cell yield was greatest using 500ng of xHDRT per reaction, which yielded ~3.8-fold more edited T-cells than the same dose of uncrosslinked template. Higher doses of xHDRT further boosted editing percentages (**Figure S3A**), but viability deficits limited cell yield (**Figure S3B)**. We

observed stimulation of T-cell editing by crosslinked templates at multiple loci, with multiple payload sizes, and at sites edited with frequencies ranging from 10% to over 40% **(Figure S3C).** Overall, crosslinked templates are an effective strategy to boost cell yield in T-cell editing workflows. Our results further indicate that cell yield is limited by toxicity caused by electroporation and donor nucleic acid and approaches that limit this toxicity may boost cell yield further.

xHDRTs contain DNA lesions that are potentially mutagenic; however, we see no evidence that HDR using xHDRTs is more mutagenic than HDR using uncrosslinked templates. This is apparent during fluorescent tagging of endogenous genes, where we observe a ~3-fold increase in GFP cells rather than any decrease caused by frame- or codon-disrupting mutations in the GFP donor sequence (**Figure 2**). We further investigated mutation frequencies during SNP editing experiments and observed no increase in cumulative mutation frequencies in a window surrounding the Cas9 cut site relative to those observed during editing with RNP alone or with RNP and uncrosslinked template (**Figure S4A**). However, we note that the background mutation frequency (the noise) of our amplicon sequencing data is approximately $2x10^{-3}$ per nucleotide (**Figure S4A Unedited**). To boost the sensitivity of our assay, we focused on TA-sites, which are the substrates for psoralen crosslinks, and are present in the 50bp window surrounding the *BFP* (2) and *HBB* (1) cut

sites. We observe no increase in mutation frequency at these sites in xHDRTs relative to uncrosslinked controls (**Figure S4B**). Overall, we conclude that xHDRTs promote HDR without decreasing HDR fidelity.

xHDRTs could boost HDR through biophysical parameters, e.g. by altering delivery of editing reagents, or by altering the recognition of xHDRTs by cellular DNA repair pathways. To determine whether ICLs are detected in xHDRTs or trigger a cell-wide response that favors HDR, we tested if the ICL had to be present in *cis* on the homologous template molecule. We simultaneously transfected two plasmids, one containing homology to the break site, and one lacking homology, with ICLs present on the homologous, nonhomologous, or neither template DNA. Only ICLs on the homologous template, but not the nonhomologous template, boosted HDR at the *LMNB1* and *HBB* loci (**Figure 3A**). This suggests that the xHDRT mechanism acts through local activity on the template DNA molecule and not by globally altering DNA repair pathway preferences. Consistent with this model, we observe no change in EJ outcomes at the *HBB* or *RAB11A* loci for cells edited with crosslinked or uncrosslinked templates (**Figures S5A-B**). Both loci have preferred indel outcomes of -9nt (*HBB*) or -3nt (*RAB11A*) and preference for repair that produces these outcomes does not change in the presence of xHDRTs. We therefore conclude that xHDRTs specifically boost HR frequency rather than altering global DNA repair preferences.

We next tested if the xHDRT effect was caused by an increased nuclear abundance of our xHDRTs. We observed no change in nuclear abundance of xHDRTs relative to uncrosslinked controls 24 hours after nucleofection in U2OS (**Figure 3B**) or K562 (**Figure S6A**) cells. This indicates that ICLs do not increase the nuclear abundance of xHDRTs relative to uncrosslinked templates. It has been reported that biophysical alterations that change the size of RNP particles can improve editing outcomes[59]. We added anionic polymers (ssDNA) to editing reactions containing xHDRTs or uncrosslinked donors and observed robust increases in HDR in all contexts (**Figure S6B**), indicating that xHDRTs act independently from the anionic polymer effect. Together, these results indicate that higher levels of editing seen with xHDRTs requires recognition and processing of the template molecule.

To define these mechanisms, we recovered both linear (PCR-derived) and plasmid xHDRT-edited samples into media containing small molecule inhibitors of the apical DNA repair kinases ATM, ATR[78], and DNA-PK[79], which have previously been inhibited to alter the frequency and type of DSB repair outcomes[80]. We found that ATR inhibition profoundly reduces (up to 5-fold) the HDR frequencies of cells edited with linear or plasmid xHDRTs while modestly altering uncrosslinked HDR frequencies (**Figures 3C, S6C, S6D, and S6F**). ATM inhibition reduced xHDRT HDR frequency, increased linear HDRT HDR frequency,

but did not change plasmid HDRT HDR frequency (**Figures 3C and S6C**). Inhibition of DNA-PK caused slight increases in HDRT and xHDRT HDR (**Figure S6D**). ATM (5μM KU55933), ATR (400nM AZ20 or Ceralasertib), and DNA-PK (5μM NU7026) inhibition prevented the phosphorylation of downstream targets Chk2, Chk1, and DNA-PK, confirming that kinase inhibition was effective at these doses (**Figure S6E**). ATR inhibition also decreased xHDRT HR in primary T-cells (**Figure S6F**). These observations are most consistent with a model in which multiple DNA repair pathways can utilize uncrosslinked template DNA but xHDRTs are processed by ATR-dependent mechanisms.

Due to the local effect of the ICL, we hypothesized that DNA repair factors recruited to the ICL might prime the xHDRT for use as a template. Major pathways implicated in ICL-repair are the Fanconi anemia (FA) pathway, the nucleotide excision repair (NER) pathway, the base-excision repair (BER) pathway, and the NEIL3 glycosylase pathway[73]. We also tested the involvement of DSB-repair factors RAD51 and 53BP1[81]. We separately knocked down genes using stably integrated CRISPRi constructs or siRNA treatment (**Figures 3D, S7B and S7C**). Knockdown of FANCA significantly attenuated editing from xHDRT relative to uncrosslinked controls (**Figures S7B and S7D**). RAD51 inhibition reduced HDR from cells edited with uncrosslinked and crosslinked templates, indicating a role for this gene in both types of recombination (**Figure S7C**). CRISPRi and siRNA-

mediated knockdowns were effective in both K562 and U2OS cells (**Figures S7A, S7E, and S7F).**

To further define the involvement of the FA pathway, we individually tested knockdowns of FANCA, FANCF, FANCM, FANCJ, and FANCD2. FANCA, FANCF, FANCD2, and FANCM showed a significant reduction in xHDRT-stimulated HR, while FANCJ showed no significant reduction (**Figure 3D**). These results indicate that the FA core complex and the ID2 heterodimer are important for crosslink-stimulated HR while FANCJ helicase activities[82] are not. We therefore conclude that activation of the FANCD2-FANCI heterodimer contributes to increased HDR from xHDRTs.

Discussion

Our previous work showed that the FA pathway is required for HDR outcomes after Cas9-mediated genome editing, but overexpression of individual FA proteins did not boost HDR frequencies[65]. We therefore investigated if adding ICLs – a substrate of the FA pathway – to donor DNA increased the probability that these molecules would be used for HDR. Strikingly, we found that adding interstrand crosslinks (ICLs) to donor DNA in gene-editing reactions dramatically enhances the frequency with which the template is utilized in HDR. This enhancement occurred in many different cell types, and across a range of donors and editing

reactions. We also observed that xHDRTs can be used synergistically with other strategies to boost editing efficiency, suggesting a distinct mechanism of HDR enhancement.

We also uncover the outlines of this mechanism: xHDRT editing requires ATR signaling and is partially dependent on the Fanconi anemia pathway. The dependence on ATR, which is primarily activated through RPA[83], suggests that signaling from ATR-activating nuclear structures – and not the DSB – may play a key role in specifying HDR instead of EJ repair pathways. These ATR-activating structures are unlikely to be encoded on the xHDRT, as these xHDRT molecules do not act as an agonist of ATR (**Figure S6E, ATRi, lanes 1 and 5**), and may instead comprise sites of replication stress or resected DNA. Furthermore, the requirement for ATR activity during xHDRT editing indicates that this may be a mechanistically distinct form of recombination. Therefore, the choice between EJ and HDR may include more repair options than the binary EJ/HDR model (**Fig S1A**) specifies. Overall, we favor a model in which xHDRT ICLs are uncovered and repaired during HDR itself and the repair of these lesions, and the completion of HDR, requires ATR signaling.

While our genetic results suggest the Fanconi anemia pathway is involved in xHDRT processing, the precise mechanism of ICL recognition remains unclear. Proposed mechanisms for FA-mediated ICL repair

stipulate that DNA replication uncovers lesions, but degradation rates of HDRTs in cells are inconsistent with episomal replication of these elements **(Figure S8A)**. There are additional models for transcription-coupled repair of ICLs, but components of these ICL-repair pathways, for example XPF, are not required for xHDRT editing (**Figure S7B**). We also note that transcription itself is not required, as xHDRTs lacking any eukaryotic promoters support increased levels of HDR **(Figs 2 and S4A *LMNB1* and *BFP*)**. An intriguing possibility is therefore that xHDRT ICLs are uncovered during recombination between the DSB and the template. Validation of such a model in the context of our observation that crosslinks stimulate xHDRT recombination in *cis* would suggest that HDR is explored frequently during DSB repair and that detection of crosslinked DNA increases the likelihood that HDR will proceed. Future studies that more precisely control the location and number of crosslinks will determine if xHDRT repair occurs via known DNA repair pathways, or if a novel recognition mechanism is involved.

From a practical standpoint, xHDRTs support higher levels of HDR with multiple payloads and loci, and in multiple cell types. We thus introduce xHDRTs as a useful tool for laboratory gene-editing workflows. Using commercial reagents and the qPCR assay outlined in this manuscript to optimize crosslink density, milligram-scale xHDRT preparations can be completed in a day. Future developments of this

approach may enable faster and more effective *ex vivo* cell therapy manufacturing.

Materials and Methods

Cell lines and culture

HEK293T, K562, and U2OS cells were obtained from ATCC, UMSCC1 cells were obtained from the Fanconi anemia Research materials repository, held in partnership with the Oregon Health & Science University. K562 cells were cultured in RPMI medium supplemented with 10% fetal bovine serum, 1% sodium pyruvate and 100 μg ml^{-1} penicillin-streptomycin. HEK293T and UMSCC1 cells were cultured in DMEM media supplemented with 10% fetal bovine serum, 1% sodium pyruvate and 100 μg ml^{-1} penicillin-streptomycin. U2OS cells were cultured in DMEM supplemented with only 10% fetal bovine serum and 100 μg ml^{-1} penicillin-streptomycin. For routine passaging, cells were grown to ~70% confluency, washed with 1-3 ml DPBS, and subsequently treated with 1-2 ml 0.25% trypsin-EDTA (Gibco) for 3-5 minutes in a 37ºC incubator. Lifted cells were then quenched with their respective media. Cell lines were routinely tested for mycoplasma contamination using enzymatic (Lonza) and PCR-based assays (Bulldog Bio).

qPCR-quantification

Purified xHDRT or HDRT plasmids were diluted to 1e+09 and 1e+08 copies per μl based on measured concentration (Qubit BR kit, Thermo Fisher, Nanodrop, or Hoescht 33342). Diluted plasmids were analyzed by qPCR using primers annealing to the ampR gene (oCR3187: cagtgaggcacctatctcagc, oCR3188: taagccctcccgtatcgtagt). ΔCt values were calculated between the HDRT and xHDRT molecules after pooling of technical triplicates. ΔCts were averaged between two concentrations of input DNA. We based our quantification on the hypothesis that at least one crosslink on the amplicon will disrupt PCR amplification. Thus, the fraction of uncrosslinked xHDRT molecules at a given psoralen concentration is equivalent to $2^{-(Ct_{crosslinked}-Ct_{uncrosslinked})}$. We used the uncrosslinked fraction to approximate the probability mass function (code available upon request) generated by the Binomial distribution for n=8 AT sites and calculated the average number of crosslinks. Parameters calculated for the amplicon were scaled to obtain values for the whole template based on relative lengths.

Cas9, RNA, and HDRT preparation

S. pyogenes Cas9-NLS was obtained from the QB3 MacroLab at UC Berkeley. All sgRNAs were synthesized by Synthego as modified gRNAs with 2'-O-methyl analogs and 3' phosphorothioate internucleotide linkages at the first three 5' and 3' terminal RNA residues.

All dsDNA was derived from purified plasmid DNA from bacterial cultures containing the indicated plasmid (Qiagen Plasmid Plus) or by SPRI purification of amplified linear dsDNA.

Psoralen-mediated xHDRTs were generated by preparing double-strand DNA to a concentration of 100μg/ml in 1X TE buffer in a 1.5 ml micro-centrifuge tube. Psoralen (20mM in DMSO) was then added to the reaction tube to the desired final concentration. Each reaction mixture in an open microfuge tube, placed on ice, was then irradiated with long wavelength UV for 15 minutes in a Spectrolinker™ XL-1000 at 365nm. Non-reacted psoralen was removed by an isopropanol precipitation and crosslinked DNA was resuspended in 1x TE.

Cisplatin-mediated xHDRTs were generated by diluting double-strand DNA to a concentration of 100μg/ml in 1X TE buffer in a 1.5 ml micro-centrifuge tube. Cisplatin (3.3mM in 0.9% saline) was added to the reaction tube to the desired final concentration. The reaction was briefly vortexed and transferred to a 37ºC incubator for one hour. Non-reacted cisplatin was removed by isopropanol precipitation and crosslinked DNA was resuspended in 1x TE.

Cas9 RNP assembly and nucleofection

Per nucleofection, 0.50 µl of sgRNA (100µM) were added to 1 µl of 5x RNP buffer (100 mM HEPES, 750 mM KCl, 25 mM MgCl2, 25% glycerol, 5 mM TCEP) in a 1.5 ml microcentrifuge tube. 1µl of Cas9 protein (40µM) was added to the reaction mixture and then brought up to a volume of 4 µl with nuclease-free water. 1µg of dsDNA donor, prepared at 1 µg/µl, was then added to the RNP mixture. Each reaction mixture was then left to incubate for at least 5 minutes at room temperature to allow RNP formation. 2.5e+05 cells were collected and spun down at 500g for 3 minutes, washed once in 200 µl D-PBS, and resuspended in 15µl of nucleofection buffer (Lonza). RNP mixtures were then added to resuspended cell pellets. Reaction mixtures were electroporated in 4D Nucleocuvettes (Lonza), and later recovered to culture dish wells containing pre-warmed media.

Editing was measured at defined time points after electroporation by flow cytometry (standard times are 96 and 120 hours; 240 hours for *RAB11A* editing – due to transcription off the plasmid). Resuspension buffer and electroporation conditions are as follows for each cell line: K562 in SF with FF-120, UMSCC1 in P3 with DS-138, HEK293T in SF with DS-150, U2OS in SE with CM104, iPSC in P3 with CA-137, T-cell in P3 with EH-115.

Viability was measured at defined time points post electroporation by flow cytometry (standard times are 24 and 48 hours). Viable cells were size-gated using FSC, SSC size gating and PI stained.

Western Blot

~400,000 cells were lysed in 150µl of 2X Laemmli buffer *(20% glycerol, 120 mM 1M Tris-HCl pH 6.8, 4% SDS, 0.05% bromophenol blue) containing 100 mM DTT*. Samples were vortexed for 10 seconds at full speed, boiled for 8 minutes, and passed three times through a 25G needle. Whole cell extracts were separated via electrophoresis on Biorad TGX gels 4-20%. Prior to transfer, TGX chemistry was activated for 45 seconds and subsequently used as a loading control. Gels were transferred onto PVDF membranes and blocked for an hour in PBS with 0.1% Tween-20 and 5% milk. Membranes were incubated overnight in primary antibodies diluted in PBS with 0.1% Tween-20 and 3% BSA. Membranes were washed in PBS with 0.1% Tween-20 three times for 10 minutes and incubated for an hour at RT with HRP secondary antibodies (1:5000). Membranes were finally imaged on a Chemidoc (Image Lab™, BioRad). Phospho-Chk1 (1:1000) was detected using antibody #2348 from Cell Signaling. Phospho-Chk2 (1:1000) was detected using #2661 from Cell Signaling. GFP was detected

using #A11122 from ThermoFisher (1:2000). Phospho-DNA-PK was detected using #68716S
from Cell Signaling (1:1000). RAD51 was detected using #8875S from Cell Signaling (1:1000).

Dox-inducible transcription

K562 cells stably expressing the reverse tetracycline transactivator (RTTA - Addgene 26429) were nucleofected using a modified *LMNB1* donor expressing mCherry under a Tet promoter (PCR 2070). mCherry expression from the donor plasmid was monitored by flow cytometry upon doxycycline induction (1ug/ml).

T Cell isolation and culture

T cell isolation and culture were performed as previously described[84]. Peripheral blood mononuclear cells (PBMCs) were isolated from a fresh healthy donor (Donor A) blood by Ficoll centrifugation using SepMate tubes (STEMCELL, per manufacturer's instructions), or purchased as purified PBMCs (Donors B and C, STEMCELL). Donor A, B, and C T cells were further isolated from PBMCs via magnetic negative selection using an EasySep Human T Cell Isolation Kit (STEMCELL, per manufacturer's instructions). Isolated T cells were cultured at 1 million cells ml^{-1} in ImmunoCult medium (STEMCELL) with 5% fetal bovine serum (Bio

techné), 50 μM 2-mercaptoethanol (Sigma), and 10 mM N-Acetyl L-Cysteine (Sigma), and were stimulated for two days prior to electroporation with anti-human CD3/CD28 magnetic dynabeads (ThermoFisher) at a beads to cells concentration of 1:1, along with a cytokine cocktail of IL-2 at 200 U ml^{-1} (STEMCELL), IL-7 at 5 ng ml^{-1} (STEMCELL), and IL-15 at ng ml^{-1}(STEMCELL). T cells were harvested from their culture vessels and de-beaded on a magnetic rack for several minutes. Prior to nucleofection, de-beaded cells were centrifuged for 3 min at 500g, media was gently aspirated from the pellet, and cells were resuspended in buffer P3 (Lonza), in which 15 μL of buffer were used per one million T-cells.

T-Cell nucleofections

RNPs were made prior to electroporation as described above. One million stimulated T-cells were de-beaded for several minutes prior to nucleofection and pelleted at 500g for 3 minutes. The cell pellet was then washed with DPBS. DPBS was gently aspirated from the T cell pellet and then resuspended in 15 μL of buffer P3 (Lonza). The cell suspension was then transferred to the RNP mix and thoroughly triturated. Next, the cell suspension was transferred to the well of a 20 μL nucleocuvette and immediately nucleofected using the pulse code EH115. Post-nucleofection, cells were rapidly recovered in 1 ml of prewarmed media.

Recovery media was composed of ImmunoCult with 5% fetal bovine serum, 50 μM 2-mercaptoethanol, 10 mM N-Acetyl L-Cysteine, and 500 U mL^{-1} IL-2. Edited T-cells analyzed for viability and total cell yield were monitored daily and kept at a confluency of 1 million cells ml^{-1}.

iPSC culture

iPSCs (AICS-0090-391) were acquired from the Allen Institute and treated essentially as described[85]. Low-passage iPSCs were thawed and cultured in 10 ml sterile-filtered mTeSR1 (STEMCELL), without antibiotic, in a $10cm^2$ Matrigel-coated plate and grown to 70% confluency, five days post-thaw. For routine passaging, at 70% confluency, old media was aspirated and cells were washed with 5 ml room temp DPBS prior to dissociation. iPSCs were then treated with 3 ml pre-warmed Accutase (Innovative Cell Technologies) and the vessel was incubated at 37ºC for 5 min. Once cells began to detach, 3 ml DPBS were added to the Accutase-treated cells and dissociated cells were triturated. Cells were rinsed with an additional 7 ml of DPBS for a final wash, and the dissociated cell suspension was transferred to a 15 ml conical tube and centrifuged at 500g for 3 min at room temp. DPBS/Accutase supernatant was carefully aspirated and cells were resuspended in 10 ml fresh mTeSR1 containing ROCK inhibitor (ROCKi) and counted using a Countess slide. Cells were then seeded into a Matrigel-coated six-well dish at a density of 1.5e+05

per well in 3 ml mTeSR1 containing ROCKi. Old media containing ROCKi was aspirated from each well the next day and replaced with fresh mTeSR1 without ROCKi. mTeSR1 was changed daily, and ROCKi was used for each passaging event, and always removed 24 hours thereafter. All cell line and primary cell work was approved by UCSB BUA2019-15.

iPSC pre-assembly of Cas9 RNP

For each iPSC nucleofection, 1 μL of 5x RNP buffer (5x stock = 100 mM HEPES, 750 mM KCl, 25 mM MgCl2, 25% glycerol, 5 mM TCEP) and 2 μL of sgRNA (100uM) were mixed with 1.5 μL of 40μM Cas9 protein (QB3 MacroLab) in a microcentrifuge tube along with 1 μg of DNA and brought up to a volume of 6 μL with nuclease-free water. The RNP reaction was incubated at room temperature for 20 minutes.

iPSC Cas9 RNP Delivery

iPSC RNPs were made prior to electroporation as described above. Low-passage iPSCs, at 70% confluency, in the wells of a six-well Matrigel-coated plate were washed with 2ml DPBS. DPBS was aspirated and then 1 ml pre-warmed Accutase was added to each well. Accutase-treated cells were then incubated at 37ºC for 3-5 minutes. 2ml DPBS were added and lifted cells were triturated, followed by the addition of another 3 ml DPBS for a final wash. Lifted cells were then transferred to a 15 ml conical tube

and pelleted at 500g for 3 minutes. Cells were then resuspended in 10ml fresh mTeSR1 with ROCKi and counted using a Countess slide. 4e+05 cells were aliquoted per nucleofection and pelleted at 300 g for 5 minutes. Media was aspirated and cells were washed again with DPBS. DPBS was aspirated and cells were resuspended in 15 μL buffer P3 (Lonza). The cell suspension was then transferred to the RNP mix and thoroughly triturated in the RNP mix. 20 μL of the resulting cell suspension was carefully, (avoiding the introduction of bubbles), transferred into the well of a 20 μL nucleocuvette (Lonza;). Cells were immediately nucleofected using the 'Primary Cell P3' program and 'CA-137' pulse code. Post-nucleofection, cells were immediately recovered into the well of a pre-coated 12-well Matrigel plate containing 1 ml of mTeSR1 and ROCK inhibitor. Nucleofected cells were cold-shocked for two days post-nucleofection at 32ºC, transferred to the 37ºC incubator three days post-nucleofection. mTeSR1 media was changed the day after nucleofection, without ROCKi. Cells were grown to 80% confluency (typically three days post nucleofection), and passaged using Accutase and ROCKi. Cells were then flowed at 96 hours and 120 hours post-electroporation to measure editing.

Genomic DNA extraction (for amplicon sequencing)

Approximately 1e+06 cells were harvested two days post nucleofection and incubated in 200 μL of QuickExtract DNA Extraction Solution (Lucigen) at 65ºC for 15 minutes, 68ºC for 15 minutes, and 95ºC for 15 minutes. Extracts were diluted 1:4 with dH_2O and insoluble cell debris was removed by centrifugation. Supernatants were then transferred to a new tube for downstream analysis.

PCR amplification of edited regions

Edited loci were amplified using locus-specific primer pairs described in Supplementary Table 1 using GoTaq master mix (Promega) and 200ng of genomic DNA. The thermocycler was set for 1 cycle of 98˚C for 30s, 35 cycles of 98˚C for 10s, 62˚C for 10s and 72 ˚C for 30 s, respectively, and 1 cycle of 72 ˚C for 1 minute. PCR amplicons (PCR1) were purified using SPRI beads, run on a 1.0% agarose gel to validate size and quantified by Qubit. 100ng of purified PCR1 DNA was then reamplified with PCR2 primers as listed in Supplementary Table 1. PCR conditions are in order as follows, 95ºC for 2 minutes, 95ºC for 30s, 60ºC for 20 cycles, 72ºC for 30s, 72ºC for 2 minutes. PCR2 products were SPRI cleaned, quantified by Qubit, normalized, and pooled at equimolar amounts. PCR2 pools were sequenced using 2x300 chemistry on a Miseq.

Analysis of Amplicon Sequencing Data

Reads were adaptor and quality trimmed using trim_galore (version 0.6.6) and aligned to predicted amplicon sequences using bowtie2 (version 2.2.5, --very-sensitive-local mode). Nucleotide variants at each position of the aligned reads were quantified using bcftools mpileup and bcftools call (version 1.11-1-g87d355e, m -A flags passed to bcftools call). Nucleotide variants were extracted using bcftools query in two formats: all nucleotides in a 50bp window centered on the cut site, and all nucleotides in a 50bp window centered on the cut site with HDR nucleotides removed. These values were plotted on a per-nucleotide basis [**Fig S2D**] or summed to produce bar plots [**Fig S2C**].

PCR Amplification of PacBio Samples

Edited or unedited samples were amplified with primers described in Supplementary Table 1 (oCR3775-oCR3776 for *HBB*; oCR3807-oCR3808 for *RAB11A*) using GoTaq master mix (Promega) and 200ng of genomic DNA. The thermocycler was set for 1 cycle of 95˚C for 2 minutes, 35 cycles of 95˚C for 30s, 62˚C for 2:20 and 72 ˚C for 30 s, respectively, and 1 cycle of 72 ˚C for 2 minutes. PCR amplicons (PCR1) were purified using SPRI beads, run on a 1.0% agarose gel to validate size and quantified by Qubit. 50ng of purified PCR1 DNA was then reamplified with PCR2 primers as provided in the PacBio 96 Barcoded Universal Primers (BUP) plate. PCR2 conditions were 1 cycle of 98˚C for 30s, 20 cycles of 98˚C for

15s, 64˚C for 15s and 72 ˚C for 3 minutes, respectively, and 1 cycle of 72 ˚C for 7 minutes. PCR2 products were SPRI cleaned, quantified by Qubit, normalized, and pooled at equimolar amounts. Final preparation for sequencing was performed using the SMRTbell Express Template Prep Kit 2.0 (PacBio). Samples were sequenced on a Sequel II PacBio sequencer.

Processing and Analysis of PacBio Samples

Consensus sequence calling barcode demultiplexing were performed using the parameters listed (ccs --minLength 10 --maxLength 50000 --minPasses 3 --minSnr 2.5 –minPredictedAccuracy 0.99; lima --hifi-preset SYMMETRIC-ADAPTERS --min-score 80 --min-qv 20). Resulting FASTX files were subsampled using awk to include reads that could be clearly identified as EJ by filtering out reads greater than a specific length. Length filters applied were 1228bp for *HBB* and 1069bp for *RAB11A* (amplicon length + 100bp). Filtered FASTX files were analyzed using CRISPResso2 CRISPRessoBatch version 2.1.1. Insertion/deletion data as a function of nucleotide position (Deletion_histogram.txt) was reprocessed for display using Python (version 3+). Correlations between indel spectra for pairwise comparisons were calculated using Pearson correlations (seaborn v 0.12.0).

Nuclear localization experiments

HDRT and xHDRT DNA were Cy5-labeled using the *Label* IT® Nucleic Acid Labeling Reagents (Mirus) and used in a standard nucleofection protocol (see Cas9, RNP assembly and nucleofection, with about $1x10^6$ cells). At 2 and 20 hours, 5e+05 cells were collected and washed in PBS. 10% of the cells were analyzed by flow cytometry. The rest of the samples were processed for nuclei isolation as follows: cells were resuspended in 475 µl of hypotonic buffer (20mM Tris-HCl, pH 7.4, 10mM NaCl, 3mM $MgCl_2$) and incubated on ice for 15 minutes. 25µl of 10% NP40 were added and the samples were vortexed full speed for 20 seconds. Nuclei were spun for 5 minutes at 700g and resuspended in PBS. Nuclei were then assessed by flow cytometry. The quality of the nuclei was ascertained by analyzing the FCS/SSC channels (nuclei should be approximately one third of the size of the whole cell). For microscopic analysis of nuclear localization, U2OS cells were plated on a 96-well glass bottom plate (#1.5H) at a density of 1E4 cells/well. 20 hours later, cells were fixed for 10 minutes with 4% formaldehyde, permeabilized for 15 minutes with DPBS containing 0.25% Triton X100. Nuclei were then counterstained with DAPI and imaged on a spinning disc microscope. A DAPI mask was used to measure the Cy5 intensity in the nucleus.

Small molecule inhibition

After standard nucleofection, cells (K562s or T-cells) were recovered in media containing the indicated concentration of ATR inhibitor (AZ20 or Ceralasertib), ATM inhibitor (KU55933), or DNA-PK inhibitor (NU7026).

Lentiviral packaging

Lentiviral packaging was adapted from[86]. Lentivirus was produced by transfecting HEK293T cells with standard packaging vectors using the TransIT-LT1 Transfection Reagent (MIR 2306; Mirus Bio LLC). Viral supernatant was collected 48–72 h after transfection, snap-frozen and stored at −80 ˚C for future use.

CRISPRi knockdown

Lentiviral constructs encoding gRNAs targeting FANCA, FANCD2, FANCF, FANCJ, FANCM, 53BP1, NEIL3, XPC, XPF, CSA, CSB, DDB1, POLB, TRAIP, XRCC1, or a non-targeting sequence (Supplementary Table 1) were separately transduced into K562 cells containing dCas9-KRAB (clone K1e[65]). Resulting cell populations were selected to homogeneity using puromycin (1μg/mL). Pooled knockdown cell populations were tested as described in the manuscript, and knockdowns validated by qPCR.

qPCR for CRISPRi cell-lines

For qPCR, between 2.5e+05 and 1e+06 CRISPRi cells were harvested. RNA was extracted using RNeasy Mini Kits (Qiagen). RNA was quantified by nanodrop and cDNA was produced from 1 μg of purified RNA using the iScript Reverse Transcription Supermix for RT-qPCR (Bio-Rad Laboratories). qPCR reactions were performed using the SsoFast Universal SYBR Green Supermix (Bio-Rad Laboratories) in a total volume of 10 μl with primers at final concentrations of 500 nM. The thermocycler was set for 1 cycle of 95 ˚C for 2 min, and 40 cycles of 95 ˚C for 2 s and 55 ˚C for 8 s, respectively. Fold enrichment of the assayed genes over the housekeeping control *ACT1B* locus were calculated using the $2^{-\Delta\Delta C}$T method essentially as described.

siRNA Experiments

Between 1e+05 and 2e+05 U2OS cells were lipofectamine transfected with 50pmols of either RAD51 siRNA (ambion #s531930) or an NTC siRNA (ThermoFisher #4390843). Cells were siRNA treated for 48 hours, nucleofected, and an aliquot of cells were harvested for Western blot at the time of nucleofection. Cells were harvested for flow cytometry 96 hours post-nucleofection.

Cell Cycle Experiments

Cell cycle analysis was performed using Click-iT EdU Alexa Fluor 647 Flow Cytometry Assay Kit (#Thermofisher C10424) with the following modifications: Cells were pulse labelled with EdU at 10uM final for 30 minutes, fixed in 4% formaldehyde for 10 minutes, washed twice with PBS containing 1% BSA, permeabilized for 15min with PBS containing 0.5% triton X100. Click iT reaction was carried out following manufacturer instruction. After 3 washes with PBS containing 1% BSA, cells were treated for 30 minutes with RnaseA, and stained for 10 minutes with propidium iodide and run on the flow cytometer.

Pairwise comparisons between data

Statistical comparisons in Figures 1, 2, 3, S6, and S7 and elsewhere in the manuscript were made using unpaired two-tailed t-tests with equal variance or unpaired two-tailed t-tests with unequal variance, where specified by the *F*-test of equality of variances. Nucleofections in S6F were split into different drug treatment wells and so comparisons were made using paired two-tailed t-tests.

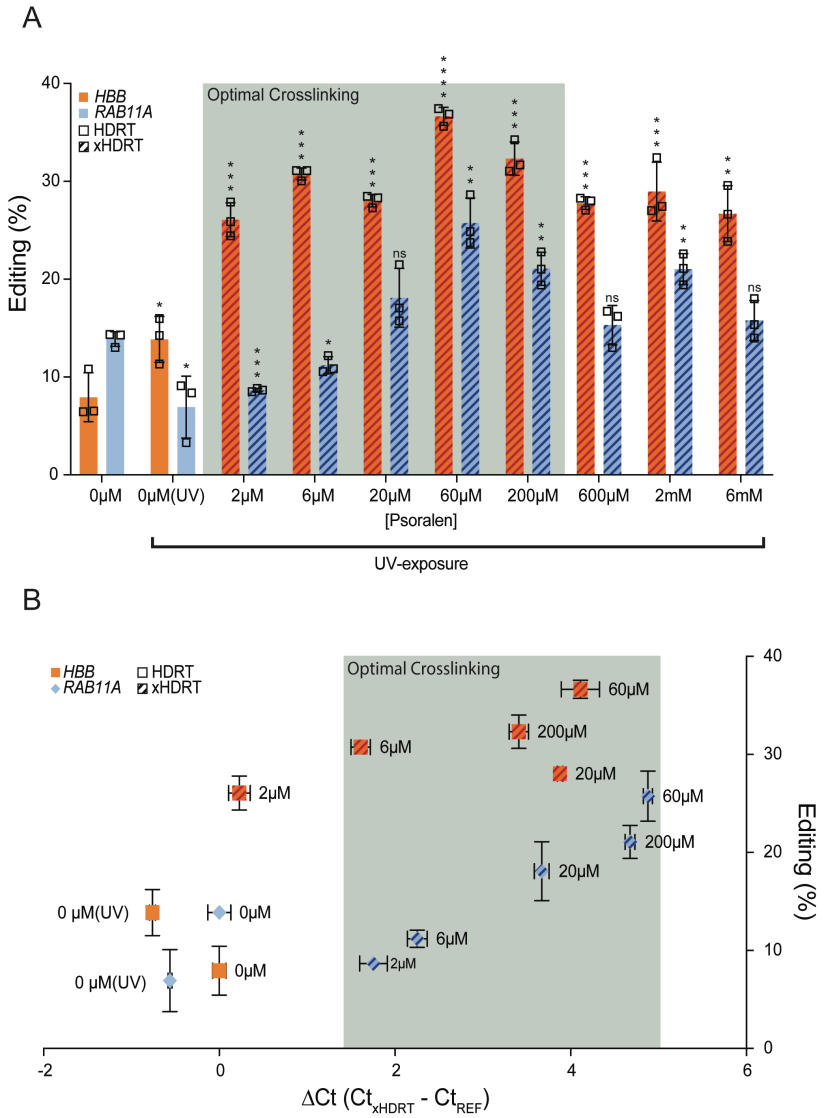


Figure 1. Modification of HDRTs with an optimal number of interstrand crosslinks increases HDR during gene-editing. (A) Percent of cells GFP positive after editing with pSFFV-GFP (*HBB*) or N-terminal GFP fusion (*RAB11A*) constructs in human K562 myeloid leukemia cells. xHDRTs were produced by treatment with the indicated amount of psoralen and UV exposure. 0μM – isopropanol precipitated plasmid HDRT (no UV, no psoralen). Significance of experimental conditions versus 0μM control is displayed above columns (* – p≤0.05, ** – p≤0.01, *** – p≤0.001, **** – p≤0.0001, ns – not significant). (B) Percent of cells GFP positive (y-axis) as a function of qPCR signal loss (x-axis), an approximation of crosslinks per unit length, for xHDRTs produced with the indicated psoralen concentration. Data displayed as the mean ± SD for n=3 biological replicates.

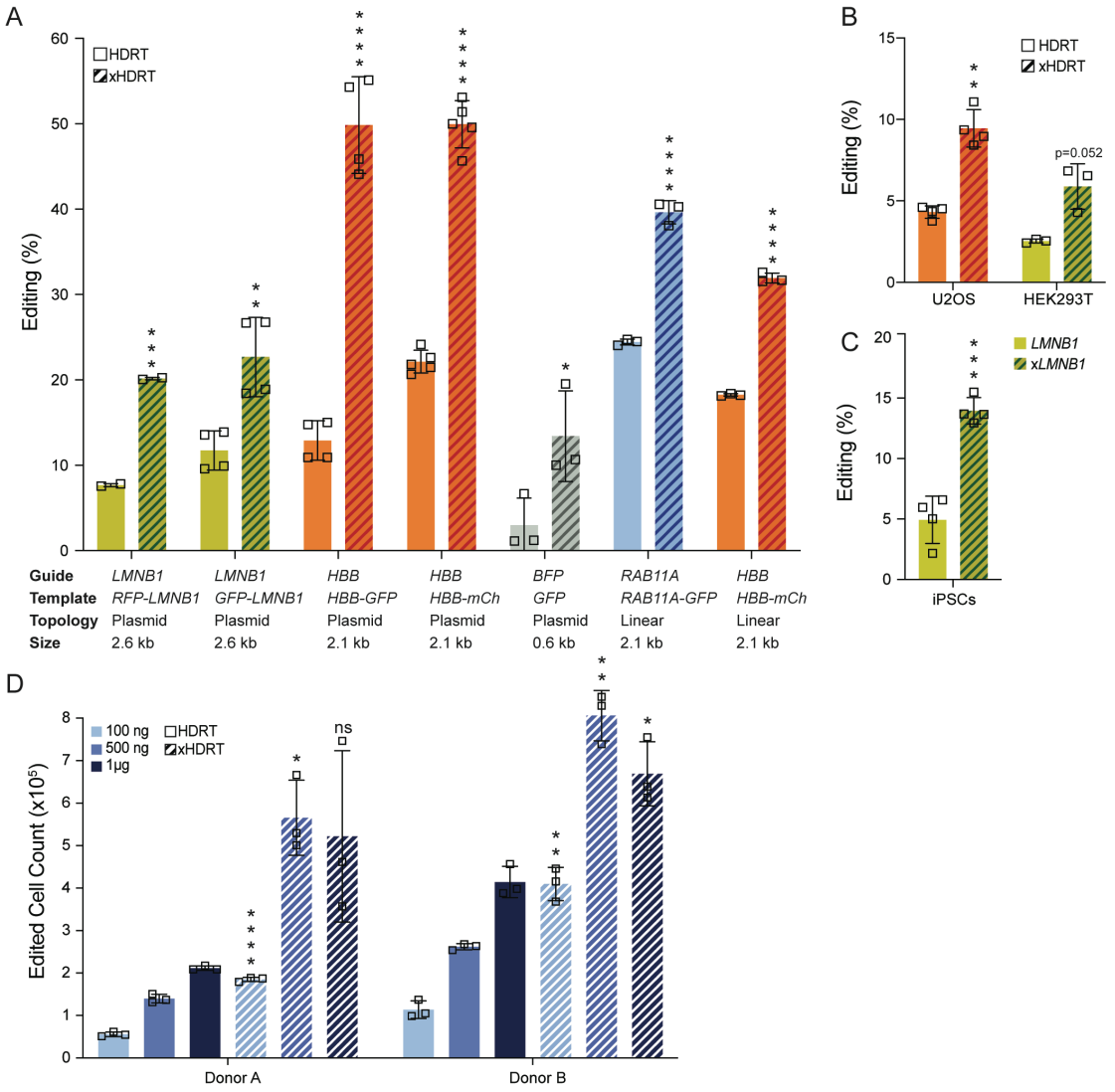


Figure 2. xHDRTs increase HDR in broad gene-editing applications. (A) Percent incorporation of GFP-tag (*LMNB1, RAB11A*), promoter-reporter (*HBB*), and SNP (*BFP*) sequences using plasmid or linear PCR-derived double-stranded DNA of indicated sizes (homology + payload) in K562 cells. Percent incorporation of a fluorophore at the *LMNB1* locus of (B) UMSCC1 and HEK293T cells or (C) iPS cells. (D) Absolute yield of *RAB11A*-GFP positive, viable T-cells from two blood donors 168 hours after editing with linear HDRT or xHDRT as gated in Fig. S3D. Data were obtained by flow cytometry and displayed as the mean ± SD of at least n=3 biological replicates; comparisons between xHDRT-edited samples versus HDRT-edited controls. Significance values are displayed above the experimental sample, * – p≤0.05, ** – p≤0.01, *** – p≤0.001, **** – p≤0.0001, ns – not significant.

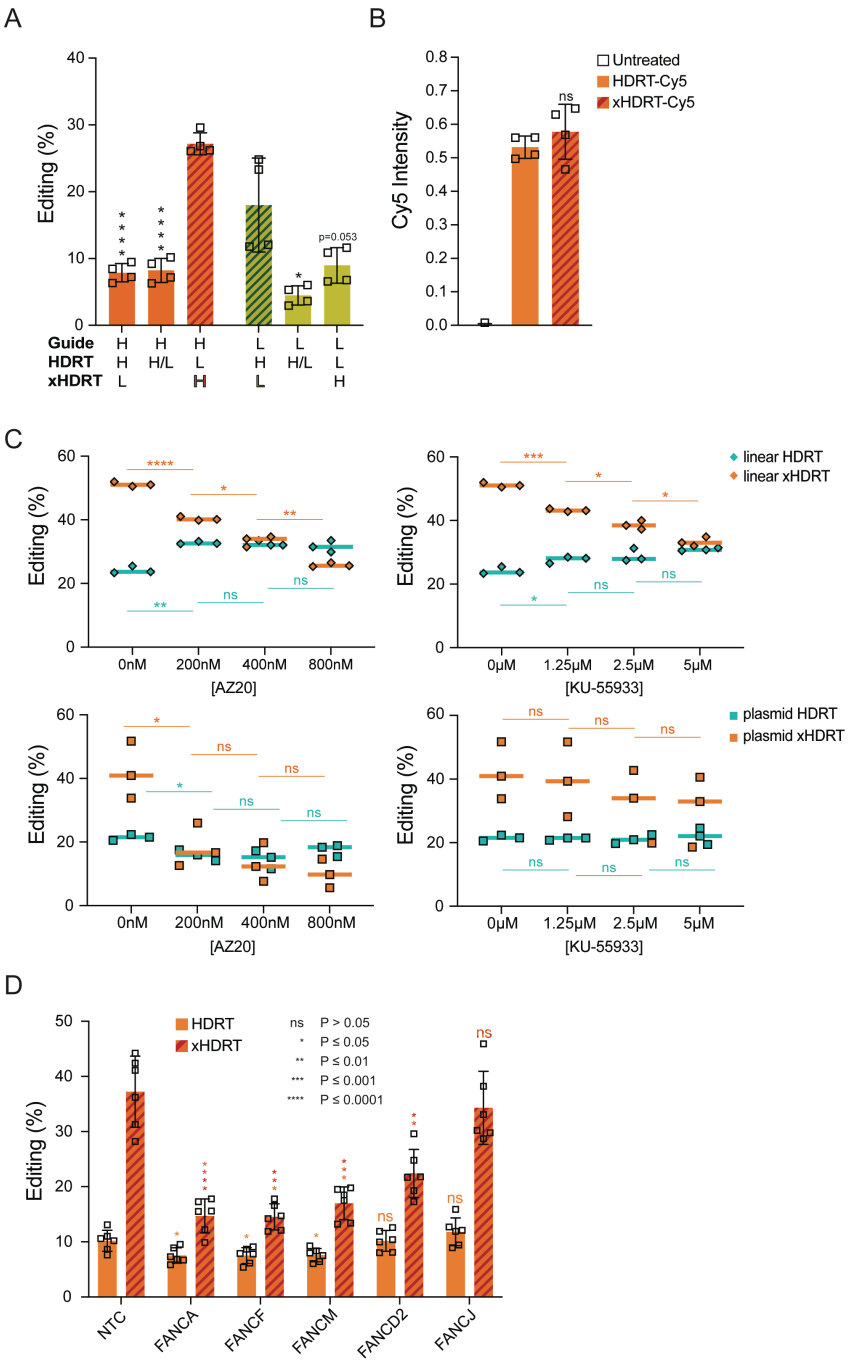


Figure 3. Enhanced editing from xHDRTs requires the activity of DNA repair pathways that are partially distinct from those that support HDR from uncrosslinked plasmids. (A) ICLs stimulate HDR in *cis*. Percent incorporation of a fluorophore encoded by crosslinked (xHDRT)

or uncrosslinked (HDRT) templates homologous to the *HBB* (H) and/or *LMNB1* (L) loci in K562 cells. DSBs are introduced at only one locus by providing sgRNA targeting *HBB* (H) or *LMNB1* (L). Maximal editing percentages (bold letters) occur when guide and xHDRT match the same locus. Data displayed as the mean ± SD of n=4 biological replicates; comparisons between xHDRT-edited samples versus HDRT-edited controls. (B) ICLs do not increase nuclear abundance of xHDRTs. Percent incorporation of an *HBB*-GFP construct (crosslinked and uncrosslinked) fused to Cy5 in U2OS cells (left). Nuclear Cy5 intensity of labeled xHDRTs as compared to uncrosslinked HDRTs and untreated U2OS cells (right). Data displayed as the mean ± SD of at least n=3 biological replicates; comparisons between xHDRT-treated samples versus HDRT-treated controls. (C) xHDRT activity is ATR and ATM dependent. Percent incorporation of *HBB*-mCherry encoded by linear PCR-derived (top) or plasmid (bottom) HDRT or xHDRT in K562 cells treated with titrated concentrations of AZ20 (ATR inhibitor), or KU-55933 (ATM inhibitor). Data shown as n=3 biological replicates; comparisons between edited untreated samples versus edited drug-treated controls. (D) xHDRT activity requires components of the Fanconi anemia pathway. Percent incorporation of *HBB*-GFP in cells edited with HDRTs (solid) or xHDRTs (striped). Data shown as n=3 of m=2 independent knockdown cell lines; comparisons between knockdown samples versus NTC controls. Knockdown efficiency with CRISPRi is shown in Figure S7F. Significance values are displayed above the experimental sample, * – $p \leq 0.05$, ** – $p \leq 0.01$, *** – $p \leq 0.001$, **** – $p \leq 0.0001$, ns – not significant.

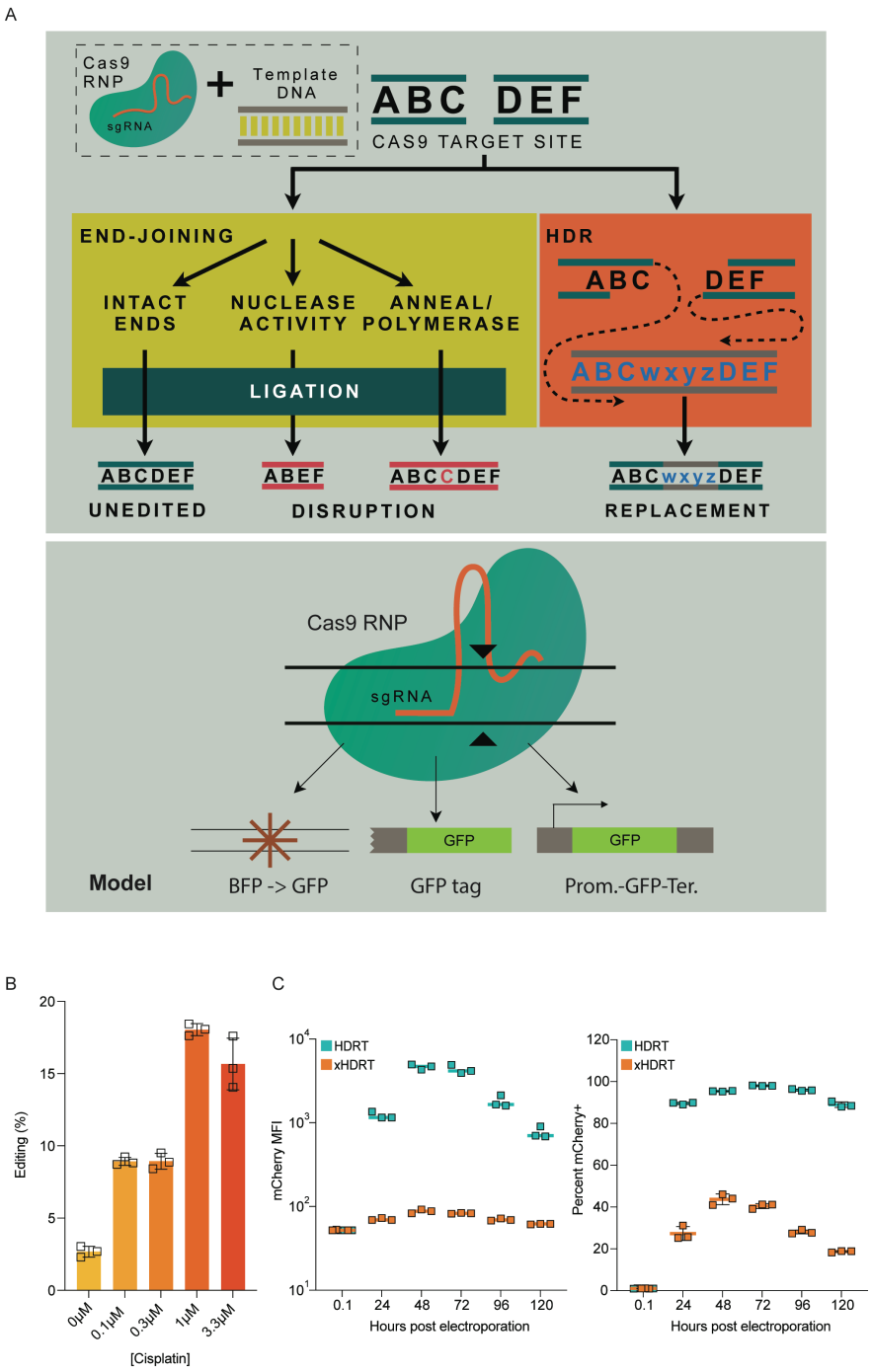


Figure S1. Modification of HDRTs with interstrand crosslinks increases HR during gene-editing. (A) Top panel: Cas9 RNPs introduce a double-strand DNA break (DSB) at a targeted region in the genome, which can be repaired by error prone end joining (EJ) processes that rejoin the ends of the break, or homology-directed repair (HDR) processes that resolve DSBs

using sequence encoded in a separate template molecule. Bottom panel: HDR gene-editing applications can be approximated using marker-based assays as diagrammed. These editing events are initiated by electroporation of Cas9, sgRNA, and HDRT into human cells, and monitored by flow cytometry or high throughput sequencing. B) Incorporation frequency of a pSFFV-GFP construct into the *HBB* locus of K562 cells using plasmid DNA treated with the indicated amount of cisplatin. (C) Transcription is inhibited from xHDRTs. Expression of dox-inducible mCherry presented both as mean fluorescence intensity (MFI) (left) and percent (right) of cells expressing mCherry encoded by uncrosslinked or xHDRT plasmid DNA. Data displayed as the mean ± SD of n=3 biological replicates.

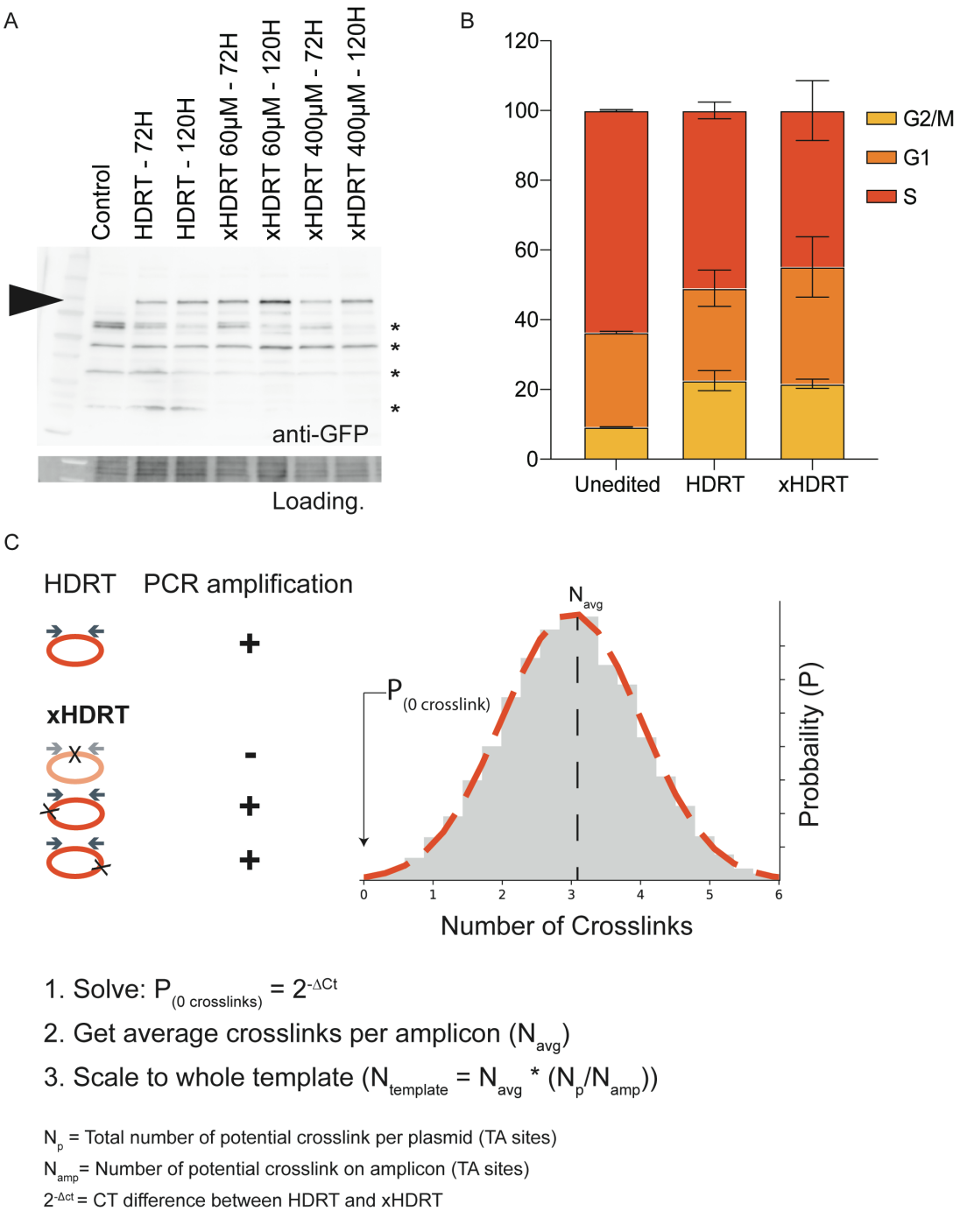


Figure S2. (A) Western blot for GFP in K562 cells edited with xHDRTs that insert GFP at the N-terminus of *LMNB1*. Cross-reacting bands (asterisks) are shown on the blot. Predicted size of GFP-*LMNB1* indicated by solid arrow. Blot is representative of n=3 biological replicates. (B) xHDRTs do not alter the cell cycle more than uncrosslinked templates. Percent of

asynchronous cells edited with uncrosslinked donors or xHDRTs at the indicated point in the cell cycle. Data displayed as the mean ± SD of n=3 biological replicates. Data were obtained by flow cytometry. (C) Schematic of crosslinking quantification by qPCR. Untreated (HDRT) or xHDRT molecules were amplified using PCR primers that produce a 94bp amplicon. Cycle thresholds (Cts) were calculated for each sample and subtracted from an uncrosslinked control to obtain ΔCt. ΔCt numbers were used to calculate a probability of no crosslinks ($P_{0\ crosslinks}$). $P_{0\ crosslinks}$ was used to calculate an average number of crosslinks per amplicon (N_{avg}). N_{avg} was scaled to the size of the xHDRT.

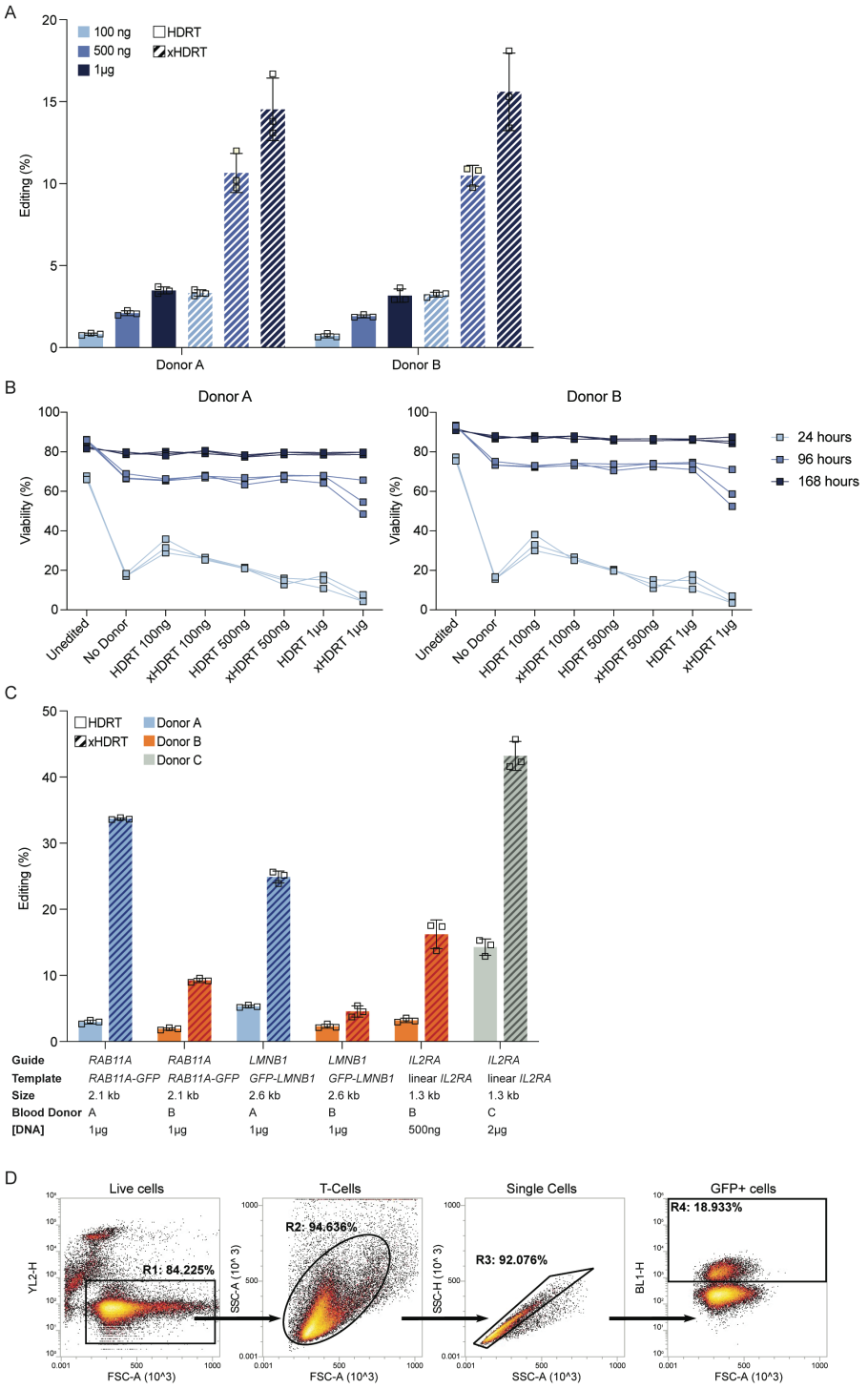


Figure S3. (A) Editing frequencies in two T-cell blood donors achieved using titrated concentrations of PCR-derived linear *RAB11A* template. (B) T-cell viability shown in two blood donors in response to titrated amounts of uncrosslinked and crosslinked PCR-derived *RAB11A*-GFP

template. (C) Crosslinked templates support higher editing efficiencies than uncrosslinked templates in primary T-cells from different blood donors and at different loci. Percent incorporation of a fluorophore at the *LMNB1*, *IL2RA,* or *RAB11A* loci in primary T-cells. Data were obtained by flow cytometry and displayed as the mean ± SD of at least n=3 biological replicates. (D) T-cell gating strategy. T-cells were stained with PI, gated for viable cells (R1), morphology by FSC and SSC (R2), single cells (R3), and GFP+ or "edited" cells (R4).

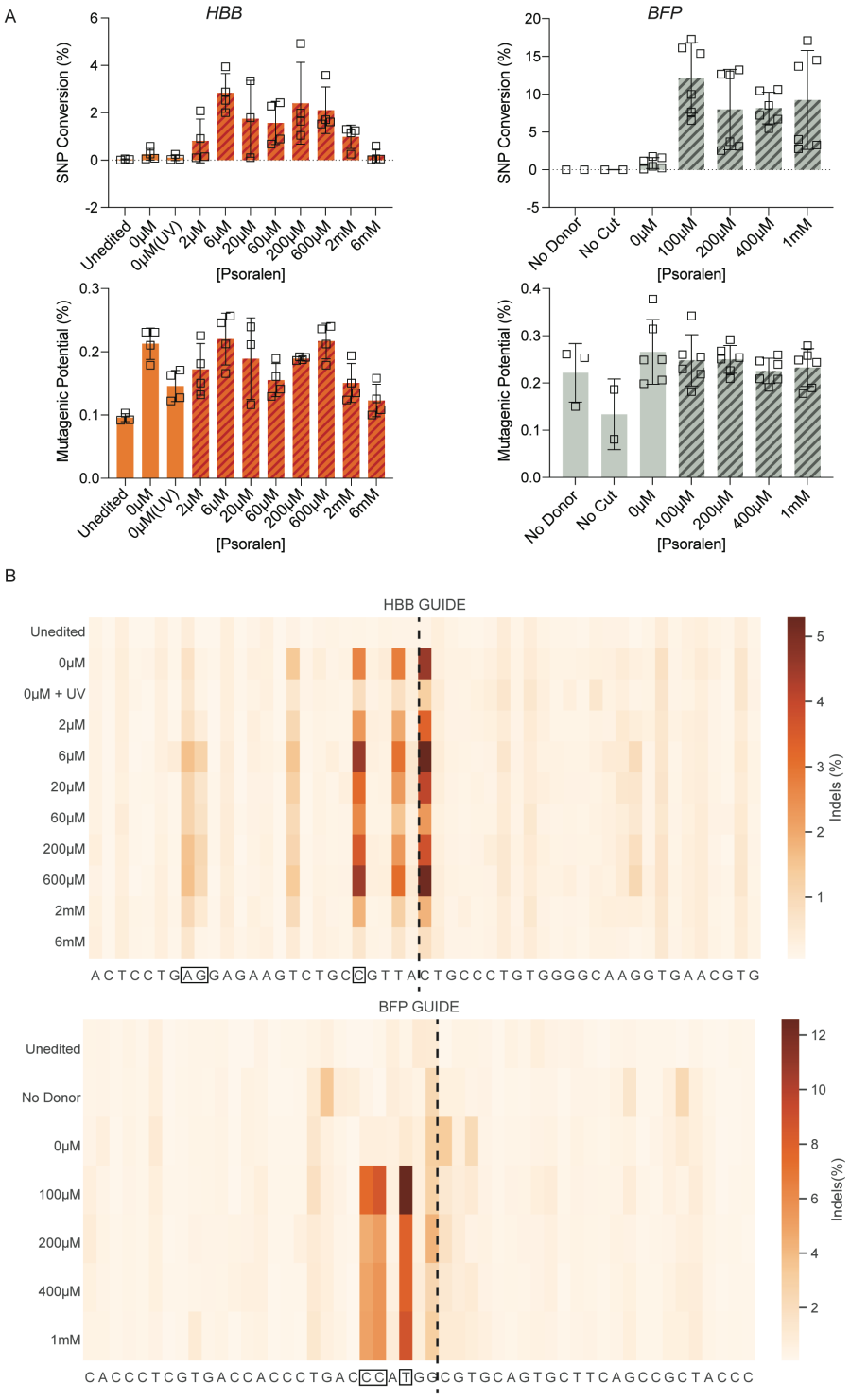


Figure S4. (A) xHDRTs boost SNP conversion without increasing total number of mutations in a window surrounding the cut site. SNP conversion as a function of crosslink frequency at the *HBB* E7V (left) or

BFP (right) loci in K562 cells (top panels). Cumulative probability of a non-HDR mutation (mutagenic potential) arising within a 50bp window surrounding the Cas9 cut site for samples edited with the indicated homology donors at *HBB* (left) or *BFP* (right) loci in K562 cells (bottom panels). Data generated from at least n=4 biological replicates. (B) xHDRTs do not increase the mutation frequency at non-SNP bases. Heatmap showing mutation frequency at each base within a window surrounding Cas9 cut site (black dashed line) for samples edited with the indicated homology donors at *HBB* (top) or *BFP* (bottom). Nucleotides altered by successful HDR are outlined with black squares.

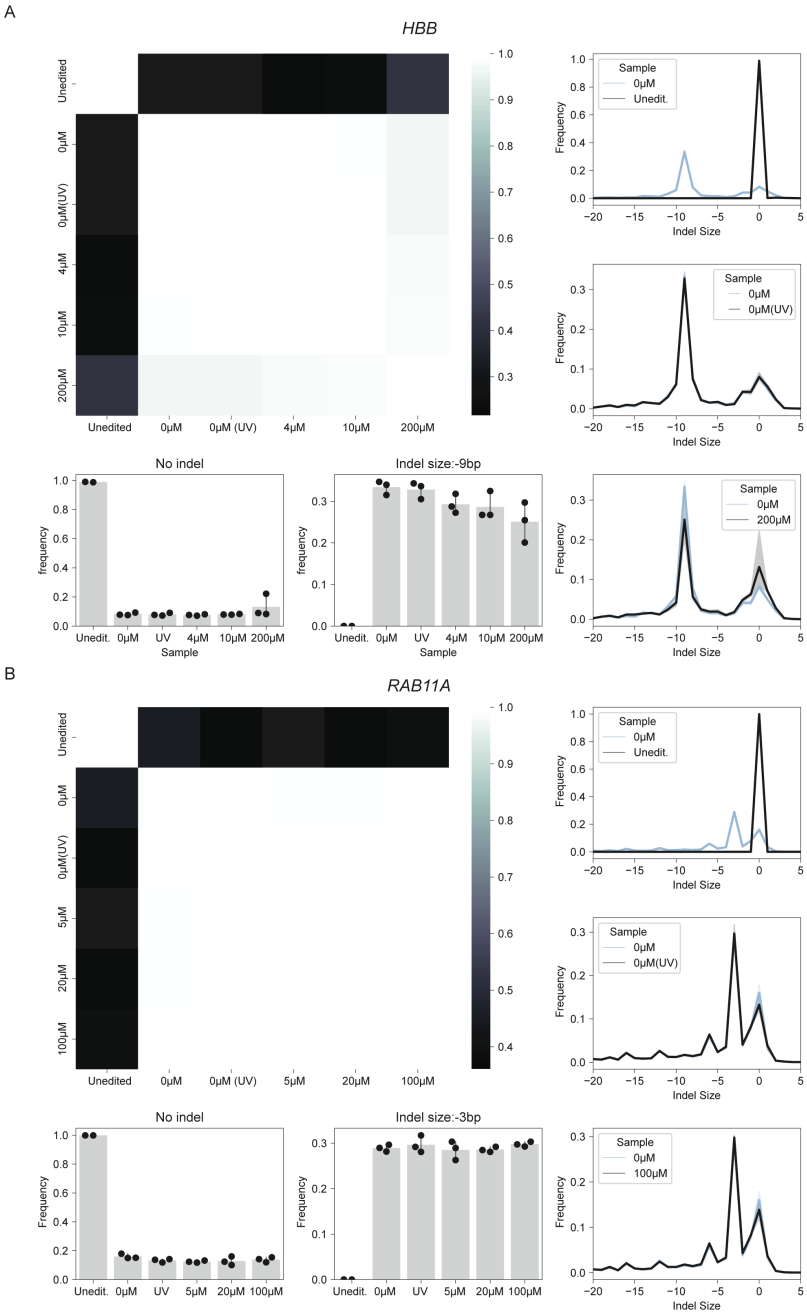


Figure S5. Crosslinked templates do not alter EJ outcomes at the *HBB* (A) or *RAB11A* (B) loci. Clockwise from top left: (i) Pearson correlations between indel spectra for unedited or cells edited with xHDRTs produced using the indicated psoralen concentrations, (ii) frequency of indels with the indicated sizes for unedited cells vs cells edited with mock treated template, (iii) frequency of indels with the indicated sizes for cells edited with mock versus UV treated template, (iv) frequency of indels with the

indicated sizes for cells edited with mock versus crosslinked template, (v) frequency of characteristic indel for cells edited using the indicated parameters, and (vi) frequency of unedited (no insertions of deletions) alleles for cells edited using the indicated parameters. All data were generated from n=3 biological replicates.

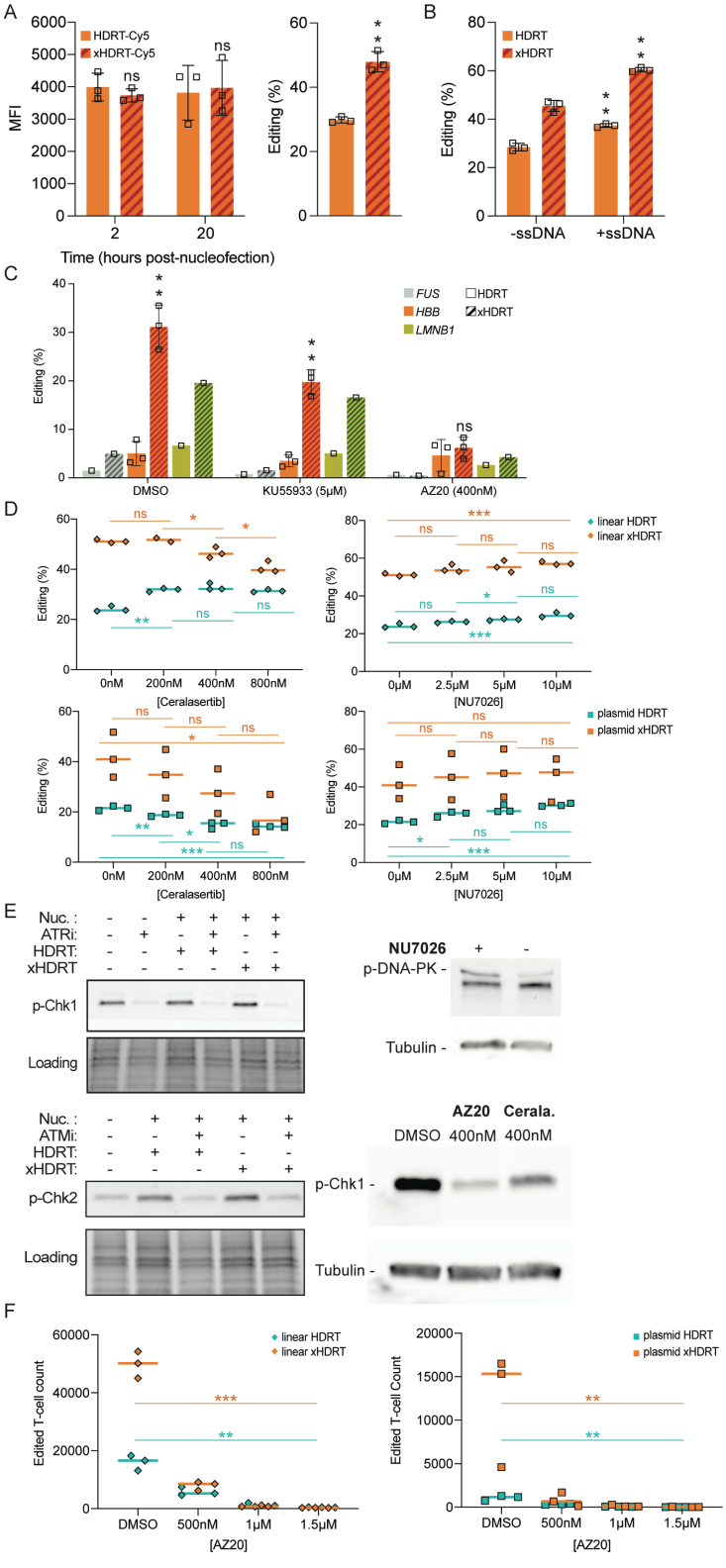
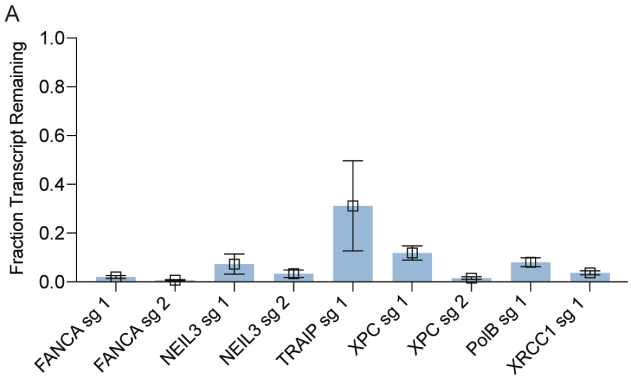
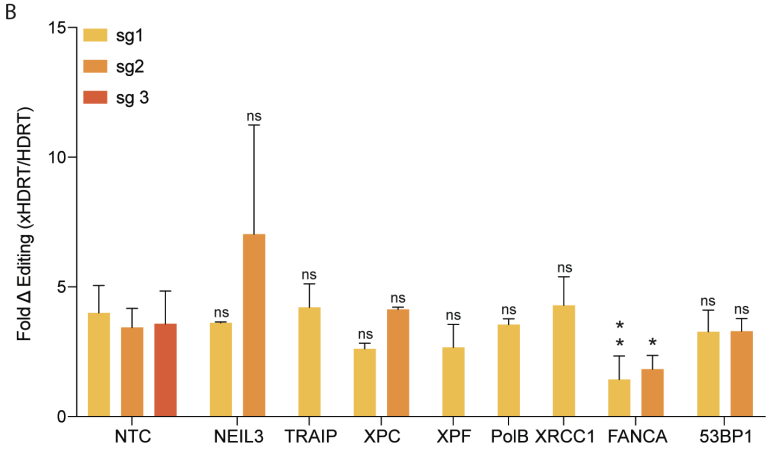
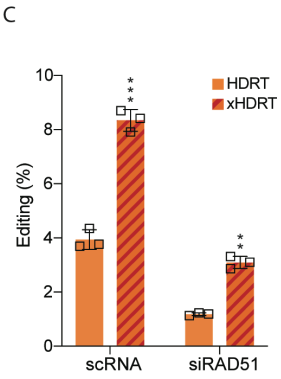
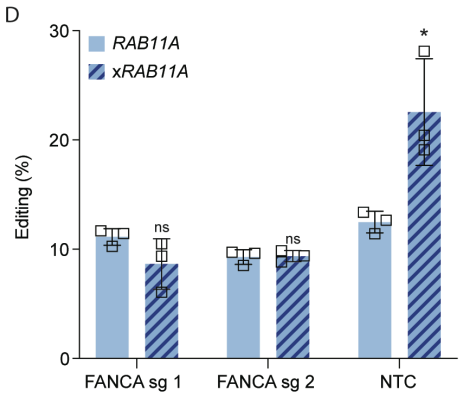

A
MFI
HDRT-Cy5
xHDRT-Cy5
ns
Editing (%)
Time (hours post-nucleofection)
B
HDRT
xHDRT
-ssDNA
+ssDNA
C
FUS
HBB
LMNB1
HDRT
xHDRT
DMSO
KU55933 (5μM)
AZ20 (400nM)
D
linear HDRT
linear xHDRT
plasmid HDRT
plasmid xHDRT
[Ceralasertib]
[NU7026]
E
Nuc. :
ATRi:
HDRT:
xHDRT
p-Chk1
Loading
NU7026
p-DNA-PK -
Tubulin -
ATMi:
p-Chk2
AZ20
Cerala.
DMSO
400nM
p-Chk1 -
F
Edited T-cell count
[AZ20]

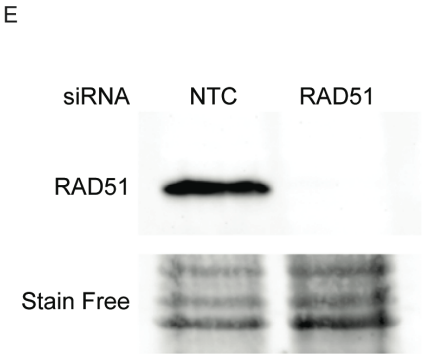
Figure S6. (A) ICLs do not increase nuclear abundance of xHDRTs. Mean fluorescence intensity (abundance) of Cy5 labeled HDRT or xHDRT DNA in isolated nuclei shown 2- and 20-hours post-electroporation (left). Percent incorporation of an *HBB*-GFP construct (crosslinked and uncrosslinked) fused to Cy5 in K562 cells (right). Data displayed is n=3 biological replicates; comparisons between xHDRT-treated samples versus HDRT-treated controls. (B) xHDRTs work additively with the anionic polymer effect. Percent incorporation of a multi-kilobase (*HBB*-mCherry) construct with or without 100 pmoles of nonhomologous ssDNA. Data shown are the mean ± SD of n=3 biological replicates; comparisons between ssDNA-treated samples versus untreated controls. (C) xHDRT activity is ATR-dependent in primary T-cells. Percent incorporation of *RAB11A*-GFP achieved using either linear PCR-derived or plasmid HDRT/xHDRT in primary T-cells treated with titrated concentrations of AZ20 (ATR inhibitor). Data represented was calculated from n=3 biological replicates; comparisons between xHDRT-edited samples versus HDRT-edited controls. (D) xHDRT activity is ATR dependent and DNA-PK independent. Percent incorporation of *HBB*-mCherry achieved using linear PCR-derived (top) and plasmid (bottom) HDRT or xHDRT in K562 cells treated with titrated concentrations of Ceralasertib (ATR inhibitor) and NU7026 (DNA PK inhibitor). Data represented was calculated from n=3 biological replicates; comparisons between treatments indicated by horizontal bars. (E) Phosphatidylinositol-3 kinase-related kinases (PIKK) inhibitors prevent substrate phosphorylation. Western blots for phospho-Chk1, and phospho-Chk2, and phospho-DNA-PK 24 hours after the indicated treatments. Data shown is representative of n=3 blots. (F) xHDRT activity is ATR-dependent. Percent incorporation for GFP-tag (*FUS, LMNB1*) or promoter-reporter (*HBB*) sequences in K562 cells treated with DMSO, KU55933 (ATM inhibitor), or AZ20 (ATR inhibitor). Data displayed was calculated from n=3 biological replicates (*HBB*) or n=1 sample (*LMNB1, FUS*); comparisons between AZ20-treated versus untreated conditions indicated by horizontal bars. Significance values are displayed above the experimental sample, * – $p \leq 0.05$, ** – $p \leq 0.01$, *** – $p \leq 0.001$, **** – $p \leq 0.0001$, ns – not significant.


A
Fraction Transcript Remaining
1.0
0.8
0.6
0.4
0.2
0.0
FANCA sg 1
FANCA sg 2
NEIL3 sg 1
NEIL3 sg 2
TRAIP sg 1
XPC sg 1
XPC sg 2
PolB sg 1
XRCC1 sg 1


B
sg1
sg2
sg 3
Fold Δ Editing (xHDRT/HDRT)
15
10
5
0
ns
NTC
NEIL3
TRAIP
XPC
XPF
PolB
XRCC1
FANCA
53BP1


C
HDRT
xHDRT
Editing (%)
10
8
6
4
2
0
scRNA
siRAD51


D
RAB11A
xRAB11A
Editing (%)
30
20
10
0
ns
FANCA sg 1
FANCA sg 2
NTC


E
siRNA
NTC
RAD51
RAD51
Stain Free

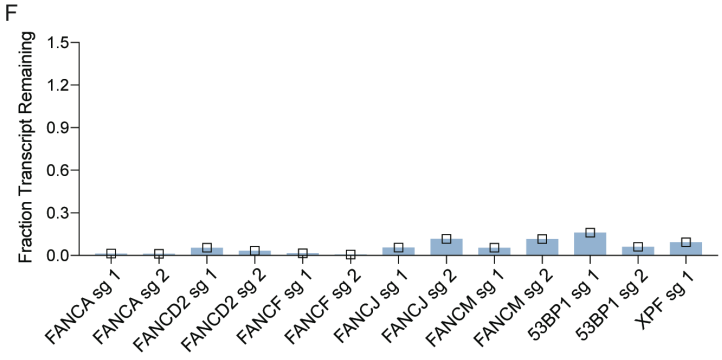

F
Fraction Transcript Remaining
1.5
1.2
0.9
0.6
0.3
0.0
FANCA sg 1
FANCA sg 2
FANCD2 sg 1
FANCD2 sg 2
FANCF sg 1
FANCF sg 2
FANCJ sg 1
FANCJ sg 2
FANCM sg 1
FANCM sg 2
53BP1 sg 1
53BP1 sg 2
XPF sg 1

Figure S7. (A) CRISPRi knockdown of ICL-repair genes was effective. Fraction transcript remaining for indicated genes in CRISPRi cell lines as measured by qPCR. Data displayed as the mean ± SD of n=3 biological replicates. (B) xHDRT activity is partially dependent on the Fanconi anemia pathway. Fold change in editing supported by xHDRTs normalized against HDRT editing for the indicated knockdowns. NTC = non-targeting knockdown. Data presented was calculated from at least n=3 biological replicates and multiple independent guides were shown where indicated; comparisons between knockdown samples versus NTC controls. Knockdown efficiency is shown in Figs. S7A and S7F. (C) RAD51 siRNA knockdown decreases editing from both HDRT and xHDRT. (D) xHDRT activity is partially dependent on FANCA. Percent incorporation of a GFP-tag construct (*RAB11A*-GFP) in two independent FANCA knockdown K562 cell lines. Data displayed as the mean ± SD of n=3 biological replicates; comparisons between knockdown samples versus NTC controls. (E) RAD51 siRNA-treated U2OS cells are effectively knocked down at the time of nucleofection. Western blot for RAD51 shown in U2OS cells that had been siRNA-treated for 48 hours. (F) CRISPRi knockdown of FA pathway genes was effective. Fraction knockdown of indicated transcripts in CRISPRi cell lines as measured by qPCR. Data displayed as the mean ± SD of n=3 biological replicates. Significance values are displayed above the experimental sample, * – $p \leq 0.05$, ** – $p \leq 0.01$, *** – $p \leq 0.001$, **** – $p \leq 0.0001$, ns – not significant.

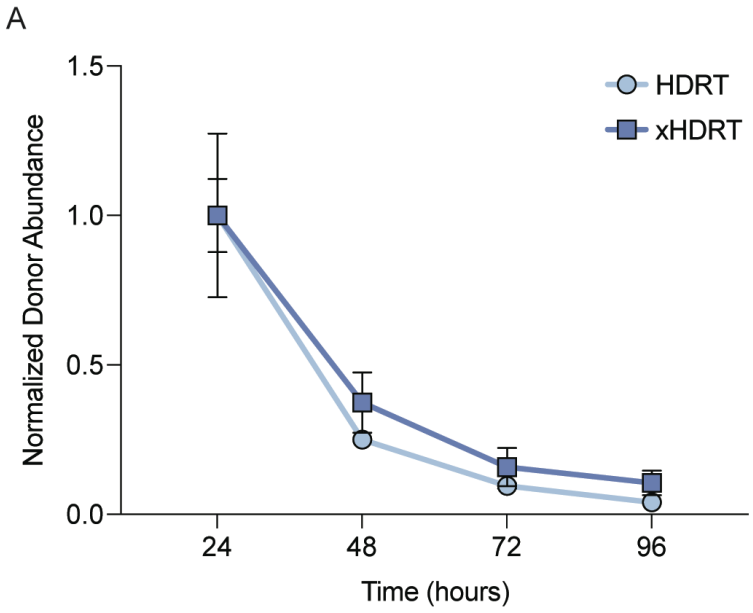


Figure S8. (A) Abundance of both xHDRTs and HDRTs decreases over time in cells. qPCR plasmid quantification (AU= $2^{(Ct^{plasmid}-Ct^{genome})_{tN}}/2^{(Ct^{plasmid}-Ct^{genome})_{t24}}$) at the indicated times after electroporation. Data normalized to plasmid abundance at t=24.

Chapter III

Discussion and Concluding Remarks

Summary of Research Findings

In Chapter II, we introduce the observation that interstrand crosslinks introduced into the backbone of a DNA template enhance gene editing efficiency in a multitude of human cell types. We establish that this editing boost proceeds through a localized mechanism given that the ICLs must be present on the template supplying information to the break site, the homology-directed repair template (HDRT). This localized phenomenon lends to the notion that the ICLs need to be introduced into the template's homology portion – not the template's non-homologous backbone.

We further observe that this enhancement in editing is supported by canonical HDR factors, such as RAD51, along with several members of the Fanconi anemia pathway - corroborating data in the field. We further observe that ICL-stimulated recombination has a distinct requirement for the checkpoint kinase ATR, upstream of the Fanconi anemia pathway, encouraging us to believe that the mechanism by which this editing boost occurs is separate from a canonical HR pathway.

In brief, when we provide the cell with a damaged DNA template riddled with ICLs, the frequency in which the cell resolves a Cas9-induced DSB through an HDR pathway increases. This phenomenon is contingent upon well-established HDR factors and has a special requirement for the checkpoint kinase ATR.

Interstrand-crosslinked template DNA is sensed differently by the cell

Contextualization of Findings

The discovery that interstrand crosslinks in template DNA improve non-viral gene-editing efficiencies is practical from a benchtop perspective, enabling scientists to improve gene-editing efficiencies for basic science applications. With further development, the discovery may prove to be applicable in clinical settings.

Even more fascinating about this observation is that it provides a glimpse into how the cell "thinks" or logically approaches complex problems. In the first chapter, we discussed several metrics thought to influence DSB pathway preference: DSB end resection, template abundance, and cell cycle position. We now observe that when we bombard a DNA template with interstrand crosslinks and blast it into a cell, the frequency at which the cell proceeds through a homology-directed repair pathway significantly increases, incorporating the sequences we provide on the template into its genome more frequently. Nevertheless, *what* about this damaged template is appetizing to the cell? The premise that a cell would *want*, if not *prefer,* to use damaged goods to repair a life-threatening form of DNA damage (a DSB) seems peculiar and counterintuitive, evoking many questions.

This preference for a crosslinked DNA template and change in editing activity suggests that interstrand-crosslinked template DNA is

sensed differently by the cell. To unpack this question (*how is an interstrand-crosslinked template sensed differently by the cell?),* let's first address several models for how the homology search occurs within a cell, a requisite for HDR to transpire in the first place.

Models for ICL-stimulated recombination

In one model of RAD51-mediated homologous recombination, it is thought that the structure formed by RAD51-bound ssDNA at DSB ends probes the entire genome, sifting through unrelated duplex DNA contents, to locate a homolog that contains, at minimum, 70 bases of homology[87]. Other models find it highly improbable that a genome-wide search for a homologous DNA sequence occurs[88]. If we accept this latter model, is it possible that the biggest hindrance to transpiring HR processes is the difficulty in locating a homolog deep inside the genome to resolve a time-sensitive dilemma? This would not stray far from the model that sister chromosome proximity is pivotal to the cell's decision between EJ and HR processes, introduced in Chapter I.

As a rule of thumb, the more DNA template you supply the cell with, the more editing is generally observed in gene-editing reactions (see Figure S3A, Chapter II), and perhaps this is a byproduct of statistics, the more of something you give the cell when it needs it, the more likely the cell is going to use it. Given this, at any concentration of DNA template

supplied, we do not observe an editing enhancement when we provide the cell with millions of copies of an untreated or unmodified repair template, which can be conveniently used to resolve the DSB within the genome; we *uniquely* observe an editing boost when ICLs are present on this homolog.

One model for how this might happen is that ICLs are the first thing sensed on the DNA template. ICLs are a substantial cellular insult, so it may be possible that the cell will strategically prioritize the resolution of this complex DNA lesion in what it may perceive as "self" DNA (*PCR-derived templates used in our gene-editing reactions are de-methylated*). In response, all of the relevant endogenous ICL-repair machinery is loaded onto the template, in which homology is conveniently revealed during the repair of the ICL within the DNA template. Once the homolog has been located and restored, the cell licenses and proceeds through an HDR pathway. In summary, homology on the DNA template may become more *discoverable* or accessible when interstrand crosslinks are present in it by coincidence.

Another possibility is that the ICL in the DNA template is discovered *after* a D-loop structure has formed, post RAD51-mediated strand exchange, during homology-directed repair. In this scenario, the increase in editing could perhaps be explained by an enrichment of factors at the intersection of ICL-repair and HDR, *like the Fanconi anemia*

pathway, capable of executing or following through with a homology-directed repair process, preventing the cell from backing out and pursuing any other options it may have to resolve the DSB, such as end-joining.

Additional Findings

Two significant questions remain unaddressed regarding the mechanism of ICL-stimulated recombination: 1. Does ICL-stimulated recombination require known HR factors or novel factors? When and how are the ICLs in xHDRTs recognized? And what I mean by this is, are the ICLs recognized on the template or at a recombination structure?

To address the first question, a series of proteomic, screening, and candidate-based approaches need to be undertaken, the preliminary results of which are described below.

Testing the Order of Events

Chapter II establishes that ATR and the Fanconi anemia pathway are required for ICL-stimulated recombination. Chapter I indicates that ATR uniquely activates the Fanconi anemia pathway, achieved through the direct phosphorylation of FANCA, FANCD2, and FANCM[41]. ATR is also thought to contribute to the physical recruitment of the FA core complex onto DNA. A model for this suggests that once ATR and its binding

partner ATRIP are recruited to RPA-bound ssDNA at stalled replication forks, ATR/ATRIP interact with HCLK2, a factor that compels the FANCM-FAAP24 heterodimer, predicted to play a significant role in the recruitment of the core complex to DNA[41]. Conversely, during M phase of the cell cycle, FANCM is hyperphosphorylated by ATM and ATR to promote the degradation and dissociation of the core complex from chromatin.

To examine if the Fanconi anemia pathway was acting canonically downstream of ATR, I treated the FA knockdown cell lines produced in Chapter II, with and without ATR inhibition. I expected to observe a decrease in editing supported by the ICL-crosslinked HDRT when FA factors (FANCA, FANCD2, FANCF, FANCM) were knocked down, as seen in Chapter II (Figure 3D), as well as a decline in editing supported by the xHDRT in non-targeting control (NTC) cells treated with ATR inhibition, corroborated by Figure 3C in Chapter II. However, if we were to observe an *additional* significant drop in editing upon the inhibition of ATR in the FA knockdown cells, we could postulate that ATR was not acting canonically upstream of the Fanconi anemia pathway. In the FA knockdown cells, the editing rates supported by the crosslinked homology template (represented by the striped bar plots) decrease as expected and do not significantly deviate upon treatment of ATR inhibition (Figure 1), indicating that the Fanconi anemia pathway is likely

acting canonically downstream of ATR. At the same time, NTC xHDRT-mediated editing rates significantly tanked upon treatment with AZ20, the small molecule inhibitor of ATR, as expected. Further experiments should be conducted to test if direct phosphorylation by ATR is required for ICL-stimulated recombination to unfold. Additionally, it would be interesting to explore the elevated editing rates supported by an uncrosslinked DNA template in ATR-inhibited cells.

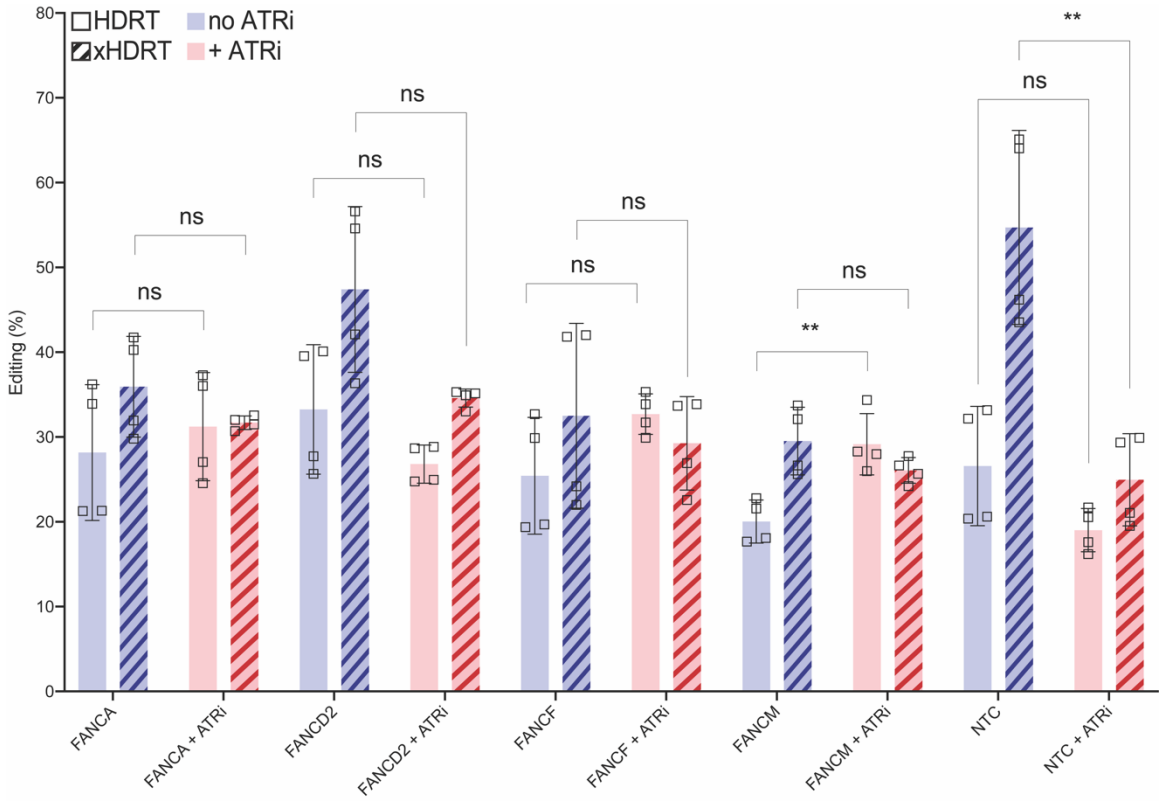


Figure 1. ATR acts canonically upstream of the Fanconi anemia pathway in ICL-stimulated recombination. Percent of GFP positive cells after editing using either uncrosslinked or ICL-crosslinked (striped) pSFFV-GFP (*HBB*) plasmid constructs in human K1e cells with or without treatment of AZ20 (ATR inhibitor), represented by red. Data displayed as n=4; comparisons between ATR-treated and untreated editing rates supported by either HDRT or xHDRT. Significance values are displayed above compared treatments, (* – p≤0.05, ** – p≤0.01, *** – p≤0.001, **** – p≤0.0001, ns – not significant).

Mass Spec Experiments

To characterize which proteins were distinctively required for ICL-stimulated HDR, we biotin-labeled crosslinked and uncrosslinked DNA templates to be used in gene-editing reactions. We edited cells with these biotinylated templates to later pull them out at 16 hours, using a Streptavidin-biotin conjugation. This approach would theoretically pull out the template DNA along with any bound protein. We hoped to be able to compare and contrast the proteomic landscape of the biotinylated ICL-crosslinked templates to biotin-labeled uncrosslinked templates.

We first found that almost every protein in the Mass Spec dataset was equally abundant on crosslinked and uncrosslinked DNA templates. The factors enriched on ICL-crosslinked templates relative to uncrosslinked HDRTs were only humbly enriched and, more importantly, later found to be unimportant in the scope of ICL-stimulated recombination when tested by knockdown. Knockdown of these enriched factors did not hinder or limit an editing boost in any capacity expected of a gene required for ICL-stimulated recombination. This suggests that the most enriched or physically abundant factors are not necessarily involved in the distinct mechanism of the editing boost.

Another explanation for this lack of significance is that samples were assessed at an irrelevant time point; perhaps if the cells had been

harvested sooner or later than 16 hours, a more relevant proteomic landscape could have been observed.

Pooled Screen Experiments

To salvage, if not unabandon, the bountiful mass spec dataset, we generated a pooled library containing knockdowns for genes acquired from the mass spec data set, selectively including proteins involved in DNA repair. Our logic here was that at least one of the bound factors, even if unenriched, would be involved in the mechanism of ICL-stimulated recombination.

We lentivirally transduced our pooled library into immortalized U2OS cells to later edit a bulk population of these cells at the *H2BJ* locus, using both crosslinked and uncrosslinked HDRTs containing a sequence encoding for the mouse surface receptor CD90. All of the edited U2OS cells would then express mouse CD90. This approach was taken to mitigate hours of cell sorting, a requirement if a fluorophore was incorporated into the genomes of edited cells as opposed to a surface receptor.

Conveniently, we could simply pull all edited cells out of the bulk population to sequence the guide sequences in these cells later. This allows us to identify the genes enriched or depleted in samples edited with a crosslinked template compared to cells edited with an

uncrosslinked template. Once sequencing is complete, this strategy will allow us to make more informative conclusions about which factors may or may not be involved in the mechanism of ICL-stimulated recombination without entirely relying on the physical proximity of proteins to tell us about ICL-stimulated recombination. Hits from the screen should be further validated by CRISPRi or siRNA knockdown.

Future Work

Resolving the mechanism of ICL-stimulated HDR

Some lingering questions I look forward to seeing answered by the Richardson lab are: *how are the interstrand crosslinks in the ICL-crosslinked templates sensed by cells? Do ICLs need to be introduced into the homology region of the template?* To follow, *are the ICLs resolved by the cell either prior to or during use in homology-directed repair?*

I have generated some preliminary proof-of-concept data suggesting that the cell resolves the interstrand crosslinks in DNA templates over time pertaining to transcription. We had previously discussed in Chapter II that ICLs in plasmid DNA inhibit transcription (Figure S1C), supporting the concept that ICLs hinder vital cellular processes like transcription and replication, as introduced in Chapter I.

To assess if ICLs in DNA templates were resolved by the cell over time, I blasted in 1µg of uncrosslinked and ICL-crosslinked *RAB11A*

plasmids, capable of self-transcribing GFP, without introducing a DSB or any RNP complex into K562 cells. Transcriptional frequencies could, therefore, be quantified by measuring the frequency of GFP-positive cells. At 24 hours post-transfection, we initially observe substantially low transcriptional frequencies supported by xHDRTs relative to the basal transcriptional levels expressed off uncrosslinked HDRTs (Figure 2). Over time, from 48 hours until 72 hours, transcription is progressively restored on the xHDRT, suggesting that the interstrand crosslinks are resolving. After 96 hours, transcription expectedly drops off in cells nucleofected with uncrosslinked and crosslinked templates.

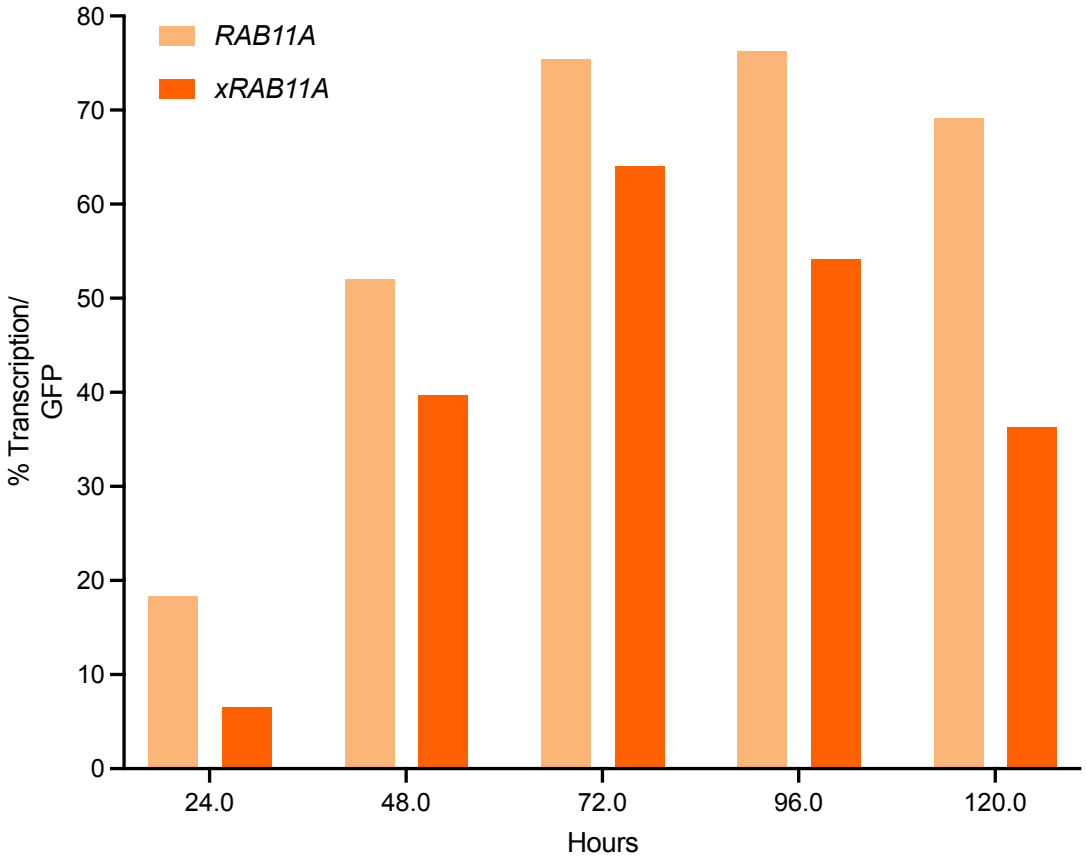


Figure 2. Transcriptional frequencies inhibited by the presence of ICLs are restored over time. Frequency of GFP positive cells after plasmid-only nucleofection using either uncrosslinked or ICL-crosslinked *RAB11A* plasmid constructs in human K562 cells. Data displayed as n=1.

Concluding Remarks

Interstrand-crosslinked DNA templates enhance non-viral gene editing frequencies. We understand that ATR and the Fanconi anemia pathway play a major role in this localized mechanism, amongst other canonical HR factors like RAD51.

We hope the recent screening approaches undertaken will illuminate the specific host pathways involved in processing xHDRTs, driving ICL-stimulated recombination.

This discovery paves the way for the continued expansion and exploration of alternative template modifications for improved non-viral gene editing. The advancement of non-viral gene editing strategies is paramount in addressing the limitations associated with current viral vector-based therapies, making room for safer, more affordable, and equally effective gene therapies.

I look forward to witnessing the follow-up to the lingering questions I have proposed by future and prospective members of the Richardson lab. I am curious to see what xHDRTs reveal about how the search for homology is carried out and the basic science of how cells resolve complex lesions and dilemmas.

Citations

1. Doudna, J. A. & Charpentier, E. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096 (2014).
2. Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420–424 (2016).
3. Anzalone, A. V. *et al.* Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149–157 (2019).
4. Li, H. *et al.* Applications of genome editing technology in the targeted therapy of human diseases: mechanisms, advances and prospects. *Signal Transduct. Target. Ther.* **5**, 1 (2020).
5. Porteus, M. H. & Carroll, D. Gene targeting using zinc finger nucleases. *Nat. Biotechnol.* **23**, 967–973 (2005).
6. Adli, M. The CRISPR tool kit for genome editing and beyond. *Nat. Commun.* **9**, 1911 (2018).
7. Joung, J. K. & Sander, J. D. TALENs: a widely applicable technology for targeted genome editing. *Nat. Rev. Mol. Cell Biol.* **14**, 49–55 (2013).
8. Makarova, K. S. *et al.* Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.* **18**, 67–83 (2020).
9. Jinek, M. *et al.* A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337**, 816–821 (2012).

10. Nishimasu, H. *et al.* Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. *Cell* **156**, 935–949 (2014).
11. Cong, L. *et al.* Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **339**, 819–823 (2013).
12. Murakami, H. & Keeney, S. Regulating the formation of DNA double-strand breaks in meiosis. *Genes Dev.* **22**, 286–292 (2008).
13. Zickler, D. & Kleckner, N. Meiotic Chromosomes: Integrating Structure and Function. *Annu. Rev. Genet.* **33**, 603–754 (1999).
14. Blackford, A. N. & Jackson, S. P. ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. *Mol. Cell* **66**, 801–817 (2017).
15. Ciccia, A. & Elledge, S. J. The DNA Damage Response: Making It Safe to Play with Knives. *Mol. Cell* **40**, 179–204 (2010).
16. Zhao, F., Kim, W., Kloeber, J. A. & Lou, Z. DNA end resection and its role in DNA replication and DSB repair choice in mammalian cells. *Exp. Mol. Med.* **52**, 1705–1714 (2020).
17. Chapman, J. R., Taylor, M. R. G. & Boulton, S. J. Playing the End Game: DNA Double-Strand Break Repair Pathway Choice. *Mol. Cell* **47**, 497–510 (2012).
18. Scully, R., Panday, A., Elango, R. & Willis, N. A. DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nat. Rev. Mol. Cell Biol.* **20**, 698–714 (2019).

19. Shrivastav, M., De Haro, L. P. & Nickoloff, J. A. Regulation of DNA double-strand break repair pathway choice. *Cell Res.* **18**, 134–147 (2008).
20. Chakraborty, A. *et al.* Classical non-homologous end-joining pathway utilizes nascent RNA for error-free double-strand break repair of transcribed genes. *Nat. Commun.* **7**, 13049 (2016).
21. Huertas, P. DNA resection in eukaryotes: deciding how to fix the break. *Nat. Struct. Mol. Biol.* **17**, 11–16 (2010).
22. Hartlerode, A. J. & Scully, R. Mechanisms of double-strand break repair in somatic mammalian cells. *Biochem. J.* **423**, 157–168 (2009).
23. Frit, P., Barboule, N., Yuan, Y., Gomez, D. & Calsou, P. Alternative end-joining pathway(s): Bricolage at DNA breaks. *DNA Repair* **17**, 81–97 (2014).
24. Mahaney, B. L., Meek, K. & Lees-Miller, S. P. Repair of ionizing radiation-induced DNA double-strand breaks by non-homologous end-joining. *Biochem. J.* **417**, 639–650 (2009).
25. Takata, M. *et al.* Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *EMBO J.* **17**, 5497–5508 (1998).

26. Pierce, A. J., Hu, P., Han, M., Ellis, N. & Jasin, M. Ku DNA end-binding protein modulates homologous repair of double-strand breaks in mammalian cells. *Genes Dev.* **15**, 3237–3242 (2001).
27. Frank-Vaillant, M. & Marcand, S. Transient Stability of DNA Ends Allows Nonhomologous End Joining to Precede Homologous Recombination. *Mol. Cell* **10**, 1189–1199 (2002).
28. DNA-dependent protein kinase promotes DNA end processing by MRN and CtIP. *Sci. Adv.* (2020).
29. Hochegger, H. *et al.* Parp-1 protects homologous recombination from interference by Ku and Ligase IV in vertebrate cells. *EMBO J.* **25**, 1305–1314 (2006).
30. Kimble, M. T., Johnson, M. J., Nester, M. R. & Symington, L. S. Long-range DNA end resection supports homologous recombination by checkpoint activation rather than extensive homology generation. *eLife* **12**, e84322 (2023).
31. Maréchal, A. & Zou, L. RPA-coated single-stranded DNA as a platform for post-translational modifications in the DNA damage response. *Cell Res.* **25**, 9–23 (2015).
32. Dynan, W. S. & Yoo, S. Interaction of Ku protein and DNA-dependent protein kinase catalytic subunit with nucleic acids. *Nucleic Acids Res.* **26**, 1551–1559 (1998).

33. Aylon, Y., Liefshitz, B. & Kupiec, M. The CDK regulates repair of double-strand breaks by homologous recombination during the cell cycle. *EMBO J.* **23**, 4868–4875 (2004).
34. Ira, G. *et al.* DNA end resection, homologous recombination and DNA damage checkpoint activation require CDK1. *Nature* **431**, 1011–1017 (2004).
35. Tomimatsu, N. *et al.* Phosphorylation of EXO1 by CDKs 1 and 2 regulates DNA end resection and repair pathway choice. *Nat. Commun.* **5**, 3561 (2014).
36. Niedernhofer, L. J., Lalai, A. S. & Hoeijmakers, J. H. J. Fanconi Anemia (Cross)linked to DNA Repair. *Cell* **123**, 1191–1198 (2005).
37. De Silva, I. U., McHugh, P. J., Clingen, P. H. & Hartley, J. A. Defining the Roles of Nucleotide Excision Repair and Recombination in the Repair of DNA Interstrand Cross-Links in Mammalian Cells. *Mol. Cell. Biol.* **20**, 7980–7990 (2000).
38. Brooks, P. J. & Theruvathu, J. A. DNA adducts from acetaldehyde: implications for alcohol-related carcinogenesis. *Alcohol* **35**, 187–193 (2005).
39. Folmer, V., Soares, J. C. M., Gabriel, D. & Rocha, J. B. T. A High Fat Diet Inhibits δ-Aminolevulinate Dehydratase and Increases Lipid Peroxidation in Mice (Mus musculus). *J. Nutr.* **133**, 2165–2170 (2003).

40. Deans, A. J. & West, S. C. DNA interstrand crosslink repair and cancer. *Nat. Rev. Cancer* **11**, 467–480 (2011).
41. Lopez-Martinez, D., Liang, C.-C. & Cohn, M. A. Cellular response to DNA interstrand crosslinks: the Fanconi anemia pathway. *Cell. Mol. Life Sci.* **73**, 3097–3114 (2016).
42. Malinge, J.-M., Giraud-Panis, M.-J. & Leng, M. Interstrand cross-links of cisplatin induce striking distortions in DNA. *J. Inorg. Biochem.* **77**, 23–29 (1999).
43. Ceccaldi, R., Sarangi, P. & D'Andrea, A. D. The Fanconi anaemia pathway: new players and new functions. *Nat. Rev. Mol. Cell Biol.* **17**, 337–349 (2016).
44. Räschle, M. *et al.* Mechanism of Replication-Coupled DNA Interstrand Crosslink Repair. *Cell* **134**, 969–980 (2008).
45. Roques, C. *et al.* MRE11–RAD50–NBS1 is a critical regulator of FANCD2 stability and function during DNA double-strand break repair. *EMBO J.* **28**, 2400–2413 (2009).
46. Kee, Y. & D'Andrea, A. D. Molecular pathogenesis and clinical management of Fanconi anemia. *J. Clin. Invest.* **122**, 3799–3806 (2012).
47. Moldovan, G.-L. & D'Andrea, A. D. How the Fanconi Anemia Pathway Guards the Genome. *Annu. Rev. Genet.* **43**, 223–249 (2009).
48. Li, X. & Heyer, W.-D. Homologous recombination in DNA repair and DNA damage tolerance. *Cell Res.* **18**, 99–113 (2008).

49. Taniguchi, T. *et al.* S-phase–specific interaction of the Fanconi anemia protein, FANCD2, with BRCA1 and RAD51. *Blood* **100**, 2414–2420 (2002).

50. Richardson, C. D. *et al.* CRISPR–Cas9 genome editing in human cells occurs via the Fanconi anemia pathway. *Nat. Genet.* **50**, 1132–1139 (2018).

51. Nakanishi, K. *et al.* Human Fanconi anemia monoubiquitination pathway promotes homologous DNA repair. *Proc. Natl. Acad. Sci.* **102**, 1110–1115 (2005).

52. Nambiar, T. S., Baudrier, L., Billon, P. & Ciccia, A. CRISPR-based genome editing through the lens of DNA repair. *Mol. Cell* **82**, 348–388 (2022).

53. Cornu, T. I., Mussolino, C. & Cathomen, T. Refining strategies to translate genome editing to the clinic. *Nat. Med.* **23**, 415–423 (2017).

54. Doudna, J. A. The promise and challenge of therapeutic genome editing. *Nature* **578**, 229–236 (2020).

55. Kim, S., Kim, D., Cho, S. W., Kim, J. & Kim, J.-S. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* **24**, 1012–1019 (2014).

56. Nambiar, T. S., Baudrier, L., Billon, P. & Ciccia, A. CRISPR-based genome editing through the lens of DNA repair. *Mol. Cell* **82**, 348–388 (2022).

57. Azimi, C. S., Tang, Q., Roybal, K. T. & Bluestone, J. A. NextGen cell-based immunotherapies in cancer and other immune disorders. *Curr. Opin. Immunol.* **59**, 79–87 (2019).

58. van Haasteren, J., Li, J., Scheideler, O. J., Murthy, N. & Schaffer, D. V. The delivery challenge: fulfilling the promise of therapeutic genome editing. *Nat. Biotechnol.* **38**, 845–855 (2020).

59. Nguyen, D. N. *et al.* Polymer-stabilized Cas9 nanoparticles and modified repair templates increase genome editing efficiency. *Nat. Biotechnol.* **38**, 44–49 (2020).

60. Chen, F. *et al.* High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. *Nat. Methods* **8**, 753 (2011).

61. Gallagher, D. N. & Haber, J. E. Repair of a Site-Specific DNA Cleavage: Old-School Lessons for Cas9-Mediated Gene Editing. *ACS Chem. Biol.* **13**, 397–405 (2018).

62. Carlson-Stevermer, J. *et al.* Assembly of CRISPR ribonucleoproteins with biotinylated oligonucleotides via an RNA aptamer for precise gene editing. *Nat. Commun.* **8**, 1711 (2017).

63. Lee, K. *et al.* Synthetically modified guide RNA and donor DNA are a versatile platform for CRISPR-Cas9 engineering. *eLife* **6**, e25312 (2017).

64. Ling, X. *et al.* Improving the efficiency of precise genome editing with site-specific Cas9-oligonucleotide conjugates. *Sci. Adv.* **6**, eaaz0051 (2020).
65. Richardson, C. D. *et al.* CRISPR–Cas9 genome editing in human cells occurs via the Fanconi anemia pathway. *Nat. Genet.* **50**, 1132–1139 (2018).
66. Hussmann, J. A. *et al.* Mapping the genetic landscape of DNA double-strand break repair. *Cell* **184**, 5653-5669.e25 (2021).
67. Lin, S., Staahl, B. T., Alla, R. K. & Doudna, J. A. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *eLife* **3**, e04766 (2014).
68. Canny, M. D. *et al.* Inhibition of 53BP1 favors homology-dependent DNA repair and increases CRISPR-Cas9 genome-editing efficiency. *Nat. Biotechnol.* **36**, 95–102 (2018).
69. Nakanishi, K. *et al.* Human Fanconi anemia monoubiquitination pathway promotes homologous DNA repair. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 1110 LP-- 1115 (2005).
70. Roques, C. *et al.* MRE11-RAD50-NBS1 is a critical regulator of FANCD2 stability and function during DNA double-strand break repair. *EMBO J.* **28**, 2400–2413 (2009).
71. Enoiu, M., Jiricny, J. & Schärer, O. D. Repair of cisplatin-induced DNA interstrand crosslinks by a replication-independent pathway involving

transcription-coupled repair and translesion synthesis. *Nucleic Acids Res.* **40**, 8953–8964 (2012).

72. Iyama, T. *et al.* CSB interacts with SNM1A and promotes DNA interstrand crosslink processing. *Nucleic Acids Res.* **43**, 247–258 (2015).

73. Semlow, D. R. & Walter, J. C. Mechanisms of Vertebrate DNA Interstrand Cross-Link Repair. *Annu. Rev. Biochem.* **90**, 107–135 (2021).

74. Hearst, J. E. Photochemistry of the psoralens. *Chem. Res. Toxicol.* **2**, 69–75 (1989).

75. Alderden, R. A., Hall, M. D. & Hambley, T. W. The Discovery and Development of Cisplatin. *J. Chem. Educ.* **83**, 728 (2006).

76. Rochette, P. J. UVA-induced cyclobutane pyrimidine dimers form predominantly at thymine-thymine dipyrimidines and correlate with the mutation spectrum in rodent cells. *Nucleic Acids Res.* **31**, 2786–2794 (2003).

77. Hickson, I. *et al.* Identification and Characterization of a Novel and Specific Inhibitor of the Ataxia-Telangiectasia Mutated Kinase ATM. *Cancer Res.* **64**, 9152–9159 (2004).

78. Foote, K. M. *et al.* Discovery of 4-{4-[(3 *R*)-3-Methylmorpholin-4-yl]-6-[1-(methylsulfonyl)cyclopropyl]pyrimidin-2-yl}-1 *H* -indole (AZ20): A Potent and Selective Inhibitor of ATR Protein Kinase with

Monotherapy in Vivo Antitumor Activity. *J. Med. Chem.* **56**, 2125–2138 (2013).

79. Robert, F., Barbeau, M., Éthier, S., Dostie, J. & Pelletier, J. Pharmacological inhibition of DNA-PK stimulates Cas9-mediated genome editing. *Genome Med.* **7**, 93 (2015).
80. Roidos, P. *et al.* A scalable CRISPR/Cas9-based fluorescent reporter assay to study DNA double-strand break repair choice. *Nat. Commun.* **11**, 4077 (2020).
81. Symington, L. S. & Gautier, J. Double-Strand Break End Resection and Repair Pathway Choice. *Annu. Rev. Genet.* **45**, 247–271 (2011).
82. Cantor, S. *et al.* The BRCA1-associated protein BACH1 is a DNA helicase targeted by clinically relevant inactivating mutations. *Proc. Natl. Acad. Sci.* **101**, 2357–2362 (2004).
83. Blackford, A. N. & Jackson, S. P. ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. *Mol. Cell* **66**, 801–817 (2017).
84. Wienert, B. *et al.* Timed inhibition of CDC7 increases CRISPR-Cas9 mediated templated repair. *Nat. Commun.* **11**, (2020).
85. Skarnes, W. C., Pellegrino, E. & McDonough, J. A. Improving homology-directed repair efficiency in human stem cells. *Methods* **164–165**, 18–28 (2019).
86. Gilbert, L. A. *et al.* Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159**, 647–661 (2014).

87. Ira, G. & Haber, J. E. Characterization of *RAD51* -Independent Break-Induced Replication That Acts Preferentially with Short Homologous Sequences. *Mol. Cell. Biol.* **22**, 6384–6392 (2002).

88. Weiner, A., Zauberman, N. & Minsky, A. Recombinational DNA repair in a cellular context: a search for the homology search. *Nat. Rev. Microbiol.* **7**, 748–755 (2009).