HHS Public Access

Author manuscript

DNA Repair (Amst). Author manuscript; available in PMC 2021 January 01.

Published in final edited form as:

DNA Repair (Amst). 2020 January ; 85: 102738. doi:10.1016/j.dnarep.2019.102738.

Absence of *XRCC4* and its paralogs in human cells reveal differences in outcomes for DNA repair and V(D)J recombination

Brian Ruis, **Amy Molan**[1], **Taylor Takasugi**[2], **Eric A. Hendrickson**[*]

Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota Medical School, Minneapolis, MN, 55455, United States

Abstract

The repair of DNA double-stranded breaks (DSBs) is an essential function performed by the Classical Non-Homologous End-Joining (C-NHEJ) pathway in higher eukaryotes. C-NHEJ, in fact, does double duty as it is also required for the repair of the intermediates formed during lymphoid B- and T-cell recombination. Consequently, the failure to properly repair DSBs leads to both genomic instability and immunodeficiency. A critical DSB protein required for C-NHEJ is the DNA Ligase IV (LIGIV) accessory factor, X-Ray Cross Complementing 4 (XRCC4). XRCC4 is believed to stabilize LIGIV, participate in LIGIV activation, and to help tether the broken DSB ends together. XRCC4′s role in these processes has been muddied by the identification of two additional XRCC4 paralogs, XRCC4-Like Factor (XLF), and Paralog of XRCC4 and XLF (PAXX). The roles that these paralogs play in C-NHEJ is partially understood, but, in turn, has itself been obscured by species-specific differences observed in the absence of one or the other paralogs. In order to investigate the role(s) that XRCC4 may play, with or without XLF and/or PAXX, in lymphoid variable(diversity)joining [V(D)J] recombination as well as in DNA DSB repair in human somatic cells, we utilized gene targeting to inactivate the *XRCC4* gene in both parental and *XLF*$^{-}$ HCT116 cells and then inactivated *PAXX* in those same cell lines. The loss of XRCC4 expression by itself led, as anticipated, to increased sensitivity to DNA damaging agents as well as an increased dependence on microhomology-mediated DNA repair whether in the context of DSB repair or during V(D)J recombination. The additional loss of XLF in these cell lines sensitized the cells even more whereas the presence or absence of PAXX was scarcely negligible. These studies demonstrate that, of the three LIG4 accessory factor paralogs, the absence of XRCC4 influences DNA repair and recombination the most in human cells.

[*]Corresponding author at: BMBB Department, University of Minnesota Medical School, 6-155 Jackson Hall, 321 Church St., SE, Minneapolis, MN, 55455, United States. hendr064@umn.edu (E.A. Hendrickson).

[1]Current address: Imanis Life Sciences Technology, 3605 US Highway 52N, Building 10, Rochester, MN 55901, United States.

[2]Current address: Department of Laboratory Medicine, Yale University School of Medicine, New Haven, CT 06510, United States.

Declaration of Competing Interest

E.A.H. declares that he is a member of the scientific advisory boards of Horizon Discovery, Ltd. and Intellia Therapeutics, companies that special in applying gene editing technology to basic research and therapeutics.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.dnarep.2019.102738.

Keywords

X-ray cross complementing 4 (XRCC4); XRCC4-like factor (XLF); Paralog of XRCC4 and XLF (PAXX); Classic non-homologous end joining (C-NHEJ); Alternative-EJ (aEJ); Variable(Diversity)Joining recombination [V(D)J recombination]

1. Introduction

The accurate repair of DNA DSBs is absolutely necessary for cellular survival and genomic stability. Albeit rarer than most other types of DNA damage, DSBs can be caused not only by exposure to exogenous agents such as ionizing radiation (IR) or chemotherapeutic compounds, but they are also spontaneously created by endogenous processes [1] including errors of DNA replication [2] as well as during V(D)J recombination and class switch recombination, the latter two of which are required for lymphogenesis [3].

DSBs are especially pathologic because of their ability to induce chromosomal translocations [4]. In order to protect the integrity of the genome from such occurrences, mammalian cells utilize two main pathways to repair DSBs: homology dependent recombination (HDR) and NHEJ. The repair of DSBs by the HDR pathway is predominately error-free, but requires a non-damaged template from which to enact repair and thus is usually only active in the late S or G2 phases of the cell cycle when an undamaged sister chromatid may be available as a donor template [5]. Because most of the cells in a human being are, however, non-cycling and thus never in S or G2 of the cell cycle, NHEJ is per force the main pathway used for the repair of DSBs [6]. NHEJ consists of at least two genetically and mechanistically distinct pathways: C-NHEJ, the predominant pathway, and Alternative-EJ (aEJ), which appears primarily to be a back-up pathway [7]. The key proteins involved in C-NHEJ include the Ku (Ku70 and Ku86) heterodimer, the catalytic subunit of the DNA-dependent protein kinase complex (DNA-PK_{cs}/PRKDC), Artemis, XRCC4, XLF/Cernunnos (hereafter XLF), and LIGIV. For simple DNA DSBs, the DNA ends can likely be rejoined by LIGIV alone [8,9]. However, if either one or both of the ends requires processing first, then the entire complement of C-NHEJ factors is utilized in a three-step process involving: *i*) recognition of the DSB, *ii*) processing of the DNA ends, and ultimately *iii*) DNA ligation [6]. The first step consists of detection of the DSB by the Ku (Ku70/Ku86) heterodimer. Once bound to the ends, Ku recruits DNA-PK_{cs} leading to the formation of the DNA-PK holoenzyme that, upon autophosphorylation [10], undergoes a conformational change that allows for the recruitment of the DNA end-processing enzymes including Artemis, polynucleotide kinase/phosphatase, and/or DNA polymerases. Finally the XRCC4:LIGIV complex ligates the processed broken ends back together.

In mammalian cells, in addition to DNA DSB repair, C-NHEJ is also required for V(D)J recombination. V(D)J recombination is a process used by the immune system to generate functional B-cell (immunoglobulin) and T-cell receptors [3,11]. Briefly, in lymphoid progenitor cells, variable (V), diversity (D) and joining (J) segments are excised from three loci on chromosomes 2, 14 and 22 and enzymatically recombined to generate genes capable of encoding functional immunoglobulin or T-cell receptor proteins. V(D)J recombination is a site-specific process that is facilitated by recombination signal sequences (RSS), which are

comprised of conserved heptamer and nonomer sequences separated by non-conserved spacer sequences of either 12 or 23 bp. During V(D)J recombination the proteins encoded by the recombination activating genes-1 and −2 (*RAG1* and *RAG2*) form the RAG complex that introduces nicks between individual V, D, or J coding sequences at the heptamer junction of the 12- and 23- signals. This RAG-mediated cleavage produces two blunt 5'-phosphorylated signal ends and two covalently-sealed (hairpin) coding ends. Subsequent joining of the signal ends and coding ends by C-NHEJ forms both signal and coding joints. Although C-NHEJ is required for V(D)J recombination, individual factors in the pathway can preferentially influence the formation of either coding or signal joints. For example, both DNA-PK_{cs} and Artemis are required significantly more for coding, rather than signal, joint formation due to their hairpin opening activities [12]. In addition, ablating XRCC4′s affinity for XLF results in a reduction of coding joint but not signal joint formation [13].

XRCC4 was first characterized almost 25 years ago after it was identified as the factor that complemented the DSB repair deficiencies in a hamster IR hypersensitive mutant cell line, XR-1 [14,15]. Although XRCC4 has no apparent enzymatic activity, it is required for the stabilization and activation of LIGIV [16,17]. A complete loss-of-function of XRCC4 appears to be exceedingly rare and has only been reported for a handful of patients afflicted with microcephalic primordial dwarfism [18,19]. Relevantly, these patients (in contrast to, say, LIGIV patients) do not have a recognizable immunological phenotype indicating that genetic redundancy for XRCC4′s V(D)J recombination activity must exist *in vivo*. Importantly, polymorphisms within the *XRCC4* gene have been identified as causing susceptibility to a bevy of cancers including bladder [20], breast [21], prostate, hepatocellular carcinoma, lymphoma and multiple myeloma [22]. XLF/Cernunous was identified through its association with patients exhibiting developmental anomalies, such as microcephaly and (unlike XRCC4) immunodeficiency, as well as through its interaction with XRCC4 [23,24]. XLF and XRCC4 share similar structural features including an N-terminal head domain and a C-terminal coiled-coiled domain that is required for homodimerization [25–27]. Importantly, cells lacking XLF exhibit impaired V(D)J recombination using either plasmid substrates [23,24,28] or chromosomal loci [29,30].

XRCC4 and XLF, besides homotypically interacting, can also interact with each other to form a filamentous complex that extends along DNA. These filaments are thought to bridge separate DNA molecules independently of LIGIV. It has been suggested that these filaments enhance the ligation of DSBs by forming a scaffold that assists in synapsis of the broken ends [31–36]. This modestly well-understood (albeit hypothetical) mechanism was significantly complicated by the discovery of a third XRCC4-like paralog, PAXX [37–39]. PAXX appears to interact more with Ku than with either of its paralogs [37,38], but the role of this protein in C-NHEJ and how functionally redundant it is with either XRCC4 or XLF — especially in human cells — is still unclear although it seems likely that it is not via filament formation [40].

In order to get a better understanding of how XRCC4 and its paralogs function in NHEJ, and to examine how their loss affects DSB repair in general, we used recombinant adeno-associated virus (rAAV)-mediated gene targeting to create both $XRCC4^{-/-}$ and $XRCC4^{-/-}$:$XLF^{-/-}$ human HCT116 cell lines. Here we show that the absence of XRCC4

leads to several pronounced phenotypes including increased sensitivity to the DNA-damaging agents etoposide and IR, a decrease in overall DSB repair with a concomitant increase in the use of aEJ. Moreover, we show that cells lacking XRCC4, XLF or both factors, are not only unable to complete V(D)J recombination efficiently, but result in disparate recombination configurations. Finally, we utilized clustered regularly interspersed palindromic repeats (CRISPR)/CRISPR-associated 9 (Cas9)-mediated gene editing to generate HCT116 cells lacking all three XRCC4 paralogs (*XRCC4*$^{-/-}$:*XLF*$^{-/-}$:*PAXX*$^{-/-}$). These cells were viable and demonstrated that, in human somatic cells, the presence or absence of PAXX seems to have no detectable effect for C-NHEJ.

2. Materials and methods

2.1. Targeting vector construction

Construction of the pAAV-XRCC4 exon 4 neomycin resistance (Neo) targeting vector was carried out by PCR followed by restriction enzyme digestion and subsequent DNA ligation. Briefly, HCT116 genomic DNA was used as template for PCR reactions to create homology arms flanking exon 4 of the *XRCC4* locus. Primers used to create either the left or right homology arms included XRCC4.3F1: 5'-ATACATACGCG GCCGCGTAATGACCCCCAGAAAGGCAACC-3′, XRCC4.3 SacIIR: 5'-TTATCCGCGGTGGAGCTCCAGCTTTTGTTCCCTTTAGAAAAGTAAATGACTACACA TGAG-3′, XRCC4.3KpnF: 5'-ATGGTACCCAATTCGCCCTATAGTGAGTCGTATTACTCCAAAATGTTA CATAGTAAAATG-3′, and XRCC4.3R1: 5'-ATACATACGCGGCCGCGTTTCTCTGCATTATTCCCTACAC-3′, XRCC4.3R: 5'-CTTGGGCCACAGGAAAGAACAC-3′. Fusion PCR was then performed using the left and right homology arms that had been generated by PCR along with a PvuI restriction fragment from a pNeDaKO vector to create a NotI-digestible vector fragment that was subsequently ligated into pAAV-MCS.

Alternatively, a pAAV-XRCC4-Exon4Fusion-Neo vector was constructed for the purpose of functionally inactivating XRCC4 in the *XLF*$^{-/-}$ HCT116 cell line. Primers used to create this targeting vector included fusXRCC4LarmF: 5'-ATCGGCGGCCGCAGCGTGCACCACCATGTCTGGCTCAT-3′, fusXRCC4LarmR: 5'-GCATCGATCGGGCTTGATTTTC TGCAATGGTGTCCAAGCAATAAC-3′, XRfusNeo Forward: 5'-ATCGATCGATAACTTCGTATAGCATACATTATACGAAGTTATGCATGGCTGAACAAG ATGG-3′, XRfusNeo Reverse: 5'-TACTCTCGAGCATACATATGCACAGTGGTAC-3′, fusXRCC4RarmF: 5'-GCTACTCGAGAAATCAAGCCAAAAATGAGCACCTGC-3′, fusXRCC4RarmR: 5'-ATCGGCGGCCGCCACCCAAGCTACAGATACAGGTCCAAG-3′. PCR products were digested with the restriction enzymes NotI, and PvuI, or XhoI and subsequently ligated into pAAV-MCS.

The targeted disruption of PAXX was accomplished using the CRISPR/Cas9 gene targeting system [41]. A CRISPR guide sequence (5'-TGACCGACGCCGCGGAGCTT-3′), that was complementary with exon 2 of the *PAXX* gene, was cloned into the PX458 CRISPR/

Cas9:green fluorescent protein (GFP) vector. Transfected cells were sorted for GFP (*i.e.*, Cas9) expression and clones were subsequently isolated by single cell sub-cloning.

2.2. Viral production

rAAV-XRCC4 Exon 4 Neo virus or rAAV-XRCC4 FusionExon 4 Neo was generated using a triple transfection strategy in which the targeting vector (8 μg) was mixed with pAAV-RC and pAAV-helper (8 μg each) and was then transfected onto 4×10^6 AAV-293 cells using Lipofectamine 2000 (Invitrogen). Virus was isolated from the AAV-293 cells 48 h later by scraping the cells into 1 ml of media followed by three rounds of freeze/thawing in liquid nitrogen.

2.3. Infections

HCT116 cells were grown to ~70 to 80% confluency on 6-well tissue culture plates. Fresh media (1 ml) was added at least 30 min prior to the addition of virus. At that time, the required amount of virus was added drop-wise to the plates. The cells and virus were allowed to incubate for 2 h before adding back more media (3 ml). The infected cells were then allowed to grow for 2 days before they were trypsinized and plated at 2000 cells per well onto 96-well plates.

2.4. Isolation of genomic DNA and PCR

Genomic DNA for PCR was isolated using a PureGene DNA Purification Kit (Qiagen). Cells were harvested from confluent wells of a 24-well tissue culture plate. DNA was resuspended in 50 μL of a hydration solution, 2 μL of which was used for each PCR reaction. For targeting events produced from the *XRCC4* Exon 4 Neo virus, a control PCR was performed for the 3′-side of the targeted locus using the primer set RArmF 5’-CGCCCTATAGTGAGTCGTATTAC-3′ and XRCC4.4RR 5’-ATACATACGCGGCCGCGTCTATACAGAGCAATCACAATGG-3′ while correct targeting was determined using RArmF and XRCC4.4ER 5’-ATGGTTTCCTGATACACAGTTTGG-3′. For the second round of targeting, a control PCR for the 5’-side of the targeted locus was accomplished using the primer pair XRCC4.4.LF 5’-ATACATACGCGGCCGCAGCATAGCATAGCAGAATCTTGAC-3′ and NeoR2 5’-AAAGCGCCTCCCCTACCCGGTAGG-3′ while correct targeting was determined by using XRCC4.4EF2 5’-CTCTTCTACGGACTCAGTGCTACCAGC-3′ and NeoR2. Alternatively, for targeting events produced from the XRCC4 FusExon 4 Neo virus, PCR was performed for the 5’-side of the targeted locus using the primer pair FusXRCC4-EF3 5’-GTCCACTTCTCTCAAAGCTACCTGAC-3′ and Neo-Rev1 5’-GAATAGCCTCTCCACCCAAGCGGCC-3′. For the second round of targeting, a PCR for the 3′-side of the targeted locus was performed using the primer pair Neo Screen Fwd Mid 5’-CCAAGCGAAACATCGCATCGAGC-3′ and FusXRCC4-ER3 5’-CACACACACTGTTACAGCACAGGGTG-3′. Final *XRCC4* exon 4 status for each allele of correctly targeted clones was determined using the primer pair XRCC4 5’F1 5’-GCTTAAAACCAGGCTTCTCAATCTTG-3′ and XRCC4 3′R1 5’-TAAAGAGGGCTCCGTATTTTTATACC-3′, each of which flank exon 4 of the *XRCC4* locus. Screening primers for *PAXX* CRISPR/Cas9-GFP targeting included PAXX ScrF 5’-TTCGTGTGCTACTGCGAAGG-3′ and PAXX ScrR 5’-

CACAGCTTGCTGCTCACAGG-3′ or PAXX ScrF2 5'-GAAGTCGTTCTTCCGCGCAG-3′ and PAXX ScrR2 5'-GGAGAGGTCAAAGGC CAGTG-3′.

2.5. Immunoblotting

Proteins were subjected to electrophoresis in a 4–20% gradient acrylamide gel, electroblotted onto a nylon membrane and detected using either ECL-Plus (GE Healthcare) or by Licor. Antibodies used included mouse anti-XRCC4 (Origene TA500650), rabbit anti-XLF, (Abcam ab126353), rabbit anti-LIGIV, mouse anti-β-catenin (Santa Cruz sc-7963), mouse anti-β-tublin, rabbit anti-PAXX (Abcam, ab126353) and rabbit anti-Ku70 (Santa Cruz sc-9033). Licor secondary antibodies included anti-rabbit IRDye 800CW 926–32213, anti-mouse IRDye 680LT 926 68022 and rabbit anti-XLF IRDye 680LT 926–68021.

2.6. Gene targeting strategies

During both *XRCC4*$^{-/-}$ cell line constructions, each round of targeting was performed using the same individual targeting vector. This fact necessitated that once a correctly targeted first round clone was identified, the Neo selection cassette needed to be removed before continuing with a second round of targeting. To this end, the inclusion of locus of crossover for phage P1 (LoxP) sites flanking the drug selection cassette facilitated its removal by the use of transient cyclization recombinase (Cre) expression. Briefly, the cells were transfected with a pML-Cre expression plasmid using Lipofectamine LTX or by infecting the cells with AdCre virus, after which they were plated at limited dilutions onto 10 cm dishes and allowed to form colonies.

All *PAXX*$^{-/-}$ cell lines were created using CRISPR/Cas9-mediated gene targeting. The cell lines targeted included, HCT116 WT, *XRCC4*$^{-/-}$, *XLF*$^{-/-}$, and the *XLF*$^{-/-}$*:XRCC4*$^{-/-}$ cell lines. In all four cases, the cell lines were transfected with the PX458 PAXX CRISPR/Cas9:GFP plasmid by electroporation using the Neon transfection system (Life Technologies). Transfected cells were then sorted based on GFP expression and subsequently plated for single cell isolation.

2.7. Etoposide sensitivity assay

For all cell lines including parental, *XRCC4*$^{+/-}$, *XRCC4*$^{-/-}$, and *XRCC4*$^{-/-}$ complemented cells, 3000 cells were plated in duplicate onto six-well plates containing 0, 0.025, 0.05, 0.1, 0.2, or 0.3 μM of etoposide. Cells were allowed to grow for ~ 10 days before the colonies were fixed, stained, counted and the cell survival percentage was calculated.

2.8. X-ray survival assay

For each cell line, either 300, in the case of the parental and *XRCC4*$^{+/-}$ cells, or 3000, for the *XRCC4*$^{-/-}$ and *XRCC4*$^{-/-}$ complemented cells, were plated in duplicate in T25 flasks 24 h before irradiation. Cells were then irradiated using an RS2000 X-ray irradiator at doses of 0, 0.2, 0.5, 1.0, 2.5, and 5.0 Gray (Gy). After irradiation, cells were allowed to grow for ~ 10 days before the colonies were fixed, stained, and counted and cell survival percentage was calculated.

2.9. Gene targeting efficiency in XRCC4$^{-/-}$ cells

To determine if the loss of XRCC4 confers higher relative gene targeting frequencies by rAAV, three different loci were chosen. The loci targeted in *XRCC4*$^{-/-}$ cells included exon 2 of the *Artemis* gene, exon 3 of the hypoxanthine-guanine phosphoribosyltransferase (*HPRT*) gene and exon 6 of the phosphatidylinositol N-acetylglucosaminyl-transferase subunit A (*PIGA*) gene [42].

For *Artemis* exon 2, a rAAV-Artemis-Neo virus with homology arms flanking exon 2 was used for viral infection as detailed above. 120 G418-resistant single colonies were isolated from 96-well plates and expanded to 24 well plates for isolation of genomic DNA. The harvested DNA was then subjected to PCR to determine correct targeting using the primer pair ZeoF2 5'-GCCGAGGAGCAGGACTGAATA-3′ and Art2ER1 5'-GTCACAGGTGACCAAAAAAAATTACTG-3′ as well as confirming the presence of the integrated vector using the primer pair Art2F-1 5'-GAGCCACCATGTCCAACTGGTTTAG-3′ and NeoR2 5'-AAAGCGCCTCCCCTACCCGGTAGG-3′. Of the 120 clones initially picked, only 94 clones eventually grew well enough to undergo PCR screening. Of these 94, 81 were confirmed to be positive for the vector by PCR. Five of these 81 were determined to be correctly targeted; resulting in a relative targeting frequency of 6.2%. WT HCT116 cells were treated similarly yielding a much lower targeting frequency; 3/176 correct targeting events or a 1.7% relative targeting frequency.

Similarly, for *HPRT* exon 3, a rAAV-HPRT-Neo virus with homology arms flanking exon 3 was used. 120 G418-resistant single colonies were isolated from 96 well plates and expanded to 24 well plates for isolation of genomic DNA which was then subjected to PCR to determine correct targeting using the primer pair ZeoF2 and HPRT.3 ER3 5'-GCTTATGGGTAGGAAGTAGTGTTATG-3′ as well as confirming the presence of the integrated vector using the primer pair ZeoF2 and HPRT.3 SbfIRR 5'-ATACATACGCGGCCGCTTAAATGGCTGCCCAATCACCTGCAGGATTGATG-3′. Of the 120 clones initially picked, only 98 clones grew well enough to undergo PCR screening. Of these, 98 were confirmed to be positive for the vector by PCR of which only 1 was determined to be correctly targeted; resulting in a relative targeting frequency of 1.0%. As a control for this experiment, WT HCT116 cells were subjected to the same treatment yielding similar results with 1/116 correct targeting events or a 0.8% relative targeting frequency.

Finally, for *PIGA*, 1×10^6 cells, in triplicate, were infected with rAAV-PIGA virus as described [42]. Infected cells were plated without counting onto one 10 cm plate per infection using 1 mg/ml G418 for selection. G418-resistant cells were allowed to form colonies over a period of 12 days after which they were collected and fixed with 4% paraformaldehyde. The fixed cells were then incubated with a FLAER antibody specific to PIGA and analyzed by flow cytometry. As *PIGA* is located on the X-chromosome, any correct gene-targeting events will result in the complete absence of PIGA protein on the cell surface (since HCT116 were derived from a male and only have one X chromosome), the absence of which corresponds to and can be quantitated by the decrease in FLAER signal.

2.10. In vivo plasmid end-joining assay

The *in vivo* plasmid end-joining assay utilizing the pEGFP-Pem1-Ad2 plasmid has been described [43–45]. Briefly, before transfection, the pEGFP-Pem1-Ad2 plasmid was digested with either HindIII, or I-SceI restriction enzymes in order to generate distinctly different reparable DNA ends. Parental HCT116, *XRCC4*$^{+/-}$, *XRCC4*$^{-/-}$, and complemented *XRCC4*$^{-/-}$ cells as well as *XLF*$^{-/-}$, *XLF*$^{-/-}$:*XRCC4*$^{-/Neo}$ cells were transfected with either a HindIII-linearized, or an I-SceI-linearized pEGFP-Pem1-Ad2 plasmid using Lipofectamine 2000 (Invitrogen) per the manufacturer's instructions. To control for transfection efficiency, a pCherry plasmid (Clontech) was co-transfected with the linearized pGFP-Pem1-Ad2 plasmid. Cells were allowed to repair the linearized plasmid for 24 h after which they were harvested and either assayed for total repair by fluorescence-activated cell sorting (FACS) analysis or used for recovery of the repaired plasmid. For FACS analysis, cells were harvested, washed with 1X phosphate buffered saline (PBS) and analyzed directly with a BD Biosciences FACSCalibur. As the total repair events is dependent on the number of cells successfully transfected with the pEGFP-Pem1-Ad2 plasmid, the ratio of cells that are both red and green divided by the total number of red cells was determined for each transfection. The final value for each mutant is reported as a percent repair compared to parental HCT116 cells. To assess the repair junctions produced during individual repair events, each cell line was similarly transfected with either the HindIII-, or I-SceI-linearized pEGFP-Pem1-Ad2 plasmid alone. 48 h later, the repaired plasmids were rescued from the human cells using a Qiagen plasmid mini-prep kit. The recovered plasmids were then used to transform Max efficiency DH5α cells (Invitrogen) and colonies were selected on LB plates containing 50 μg/ml kanamycin. To assess the overall use of microhomology mediated repair, HindIII restriction enzyme digestion on the HindIII-restricted/repaired plasmids or on colony PCR products produced using the primer pair GFP-F2 5'-GACGACGGCAACTACAAGACC-3′ and GFP-R2 5'-CGATGCCCTTCAGCTCGATG-3′ was performed. DNA sequencing using primers located upstream and downstream of the restriction enzyme recognition site (primer sequences used are available upon request), allowed for precise repair junction examination.

2.11. Microhomology mediated repair end-joining assay

The microhomology mediated repair assay was performed as described [45,46]. Briefly, 3.2 μg of EcoRV and AfeI restriction enzyme-digested pDVG94 was transfected into cells that had been plated at 1×10^6 cells per well of a 6 well plate using Lipofectamine 2000 (Invitrogen) according to manufacturer's directions. After transfection (48 h), repaired plasmids were recovered from cells using a Qiagen plasmid mini-prep kit. Repair junctions from recovered pDVG94 plasmids were PCR amplified using primers FM30 and DAR5 [46]. The resulting PCR products were restriction enzyme digested with BstXI and the fragments, as well as any undigested products, were separated by polyacrylamide gel electrophoresis using a 7.5% Ready gel (BioRad) in Tris/borate/EDTA (TBE) buffer. The gel was then bathed in TBE containing 1X SYBR Gold (Invitrogen) in order to visualize the products. The undigested (180 bp) and digested (120 bp and 60 bp) PCR products were quantified using Quantity One 1-D Analysis Software (BioRad).

2.12. V(D)J recombination repair assay

The V(D)J recombination assay was performed transiently using three V(D)J recombination substrate plasmids; pGG49 (SJ reporter), pGG51 (CJ reporter) and pGG52 (Inversion/Hybrid reporter), which were each co-transfected into target cells along with *RAG1* and *RAG2* expression plasmids. The V(D)J recombination substrate plasmids were kindly provided by Michael Lieber [47] while Jean-Pierre de Villartay provided the *RAG1* and *RAG2* co-expression plasmids [48]. Briefly, the *RAG1* and *RAG2* plasmids were co-transfected with each of the V(D)J reporter plasmids using Lipofectamine LTX into desired cells in a 6-well format. The cells were allowed to grow 48 h post-transfection to enable V(D)J recombination to occur. V(D)J recombined plasmids were recovered from each cell line after transfection using a plasmid mini-prep kit. The recovered plasmids were then subjected to DpnI restriction enzyme digestion to remove non-replicated plasmids. This DNA was then used to transform Max-efficiency DH5α *E. coli* that were then plated in parallel on ampicillin (100 μg/ml) and ampicillin-chloramphenicol (100 μg/ml, 34 μg/ml) plates. Colonies formed on the ampicillin plates measure the transformation efficiency while the ampicillin-chloramphenicol plates measure the V(D)J recombination events. In order to examine the nature of individual repair events, colony PCR was performed using the primer pair SCR21 and pGG51Reverse and the resulting products were sequenced using pGG51Reverse.

3. Results

3.1. Generation of a homozygous XRCC4$^{-/-}$ HCT116 cell line

A *XRCC4*$^{+/-}$ HCT116 cell line was constructed utilizing a rAAV gene targeting vector that was designed to replace exon 4 of the *XRCC4* locus (Fig. 1A) with a LoxP-flanked (floxed) Neo drug selection cassette (Fig. 1B). The *XRCC4* gene is expressed from 8 exons on chromosome 5 and the loss of exon 4, and subsequent splicing of exon 3 to exon 5, will create a truncated protein that terminates immediately after the first amino acid encoded from exon 5. Importantly, any residual protein that might be expressed will be lacking its LIGIV interaction domain [49–51] and should be non-functional. The vector was generated as described [52,53] with a 929 bp left homology arm and a 915 bp right homology arm with a central floxed Neo drug selection cassette (Fig. 1B). Correct targeting of the *XRCC4* locus replaced exon 4 with the floxed Neo cassette resulting in cells resistant to G418 selection (Fig. 1C and H). In order to use the same vector for the second round of targeting, it was necessary that the Neo cassette be removed before proceeding (Fig. 1D). A *XRCC4*$^{+/Neo}$ clone was therefore transiently exposed to the Cre recombinase via a pML-Cre transfection, which resulted in the isolation of a *XRCC4*$^{+/-}$ clone. This clone was then subjected to a second round of targeting (Fig. 1E and H). Two independent clones were confirmed to be null for *XRCC4* (i.e., *XRCC4*$^{Neo:-}$; Fig. 1F and H). One of the clones was further sub-cloned and then exposed to an adenoviral vector expressing Cre (AdCMV-Cre) in order to, once again, remove the drug selection cassette (Fig. 1G and H) to generate a viable *XRCC4*$^{-/-}$ cell line. PCR analysis of each clone along with WT HCT116 and a *XRCC4*$^{+/-}$ clone using primers that flank exon 4 confirmed that both alleles of *XRCC4* had been targeted (Fig. 1H).

3.2. Complementation of XRCC4$^{-/-}$ HCT116 cells

In order to show that the absence of XRCC4 was responsible for any subsequent phenotypes, one of the *XRCC4*$^{-/-}$ clones was complemented using an *XRCC4* cDNA expression vector constructed in Dr. Mauro Modesti's laboratory (INSERM, France) and kindly provided to us by Dr. Katheryn Meek (Michigan State University, USA). Whole cell extracts from two resulting clones, *XRCC4*$^{-/-:XRCC4-4}$ and *XRCC4*$^{-/-:XRCC4-5}$, along with WT, *XRCC4*$^{+/-}$, *XRCC4*$^{NEO/-}$, and *XRCC4*$^{-/-}$ clones were used for western blot analyses. LIGIV is required for the stability of XRCC4 (and *vice versa*), and therefore a whole cell extract from *LIGIV* $^{-/-}$ cells [54] was used as an additional positive control. Cells lacking one copy of *XRCC4* appeared to have only ~ 20% less XRCC4 compared to WT cells while, as expected, *LIGIV* $^{-/-}$ cells resulted in a ~ 50% reduction of XRCC4 (Fig. 1I). Importantly, both the *XRCC4*$^{Neo/-}$, and *XRCC4*$^{-/-}$ clones showed no detectable XRCC4 expression (Fig. 1I, J, and K) while the complemented clones #4 and #5 either expressed 2X more or less, respectively, protein than WT cells (Fig. 1I).

3.3. Generation of homozygous XLF$^{-/-}$:XRCC4$^{+/-}$, and XLF$^{-/-}$:XRCC4$^{-/-}$ HCT116 cell lines

In order to carry out epistasis experiments, we generated *XLF* mutant cell lines that were also mutant for *XRCC4*. To this end, a *XLF*$^{-/-}$ clone [55] was targeted with an rAAV-XRCC4-Exon4Fusion-Neo vector that introduced an in-frame Neo gene, which was also flanked by two Lox P recombination sites, into exon 4 of the *XRCC4* gene (Fig. 1E). In contrast to the vector used to make the *XRCC4*$^{-/-}$ cell line, which deleted Exon 4, correct targeting with this vector resulted in the production of three clones containing a severely truncated XRCC4:Neo fusion gene product that conferred G418 resistance. Successful AdCMV-Cre treatment of one of these clones resulted in a *XLF*$^{-/-}$:*XRCC4*$^{+/-}$ G418-sensitive clone that was subsequently used for a second round of targeting using the same XRCC4-Exon4Fusion-Neo vector. Two independent *XLF*$^{-/-}$:*XRCC4*$^{Neo/-}$ clones were obtained (Fig. 1H). To confirm that correct targeting of *XRCC4* in the *XLF*$^{-/-}$ cell line resulted in the loss of expression of XRCC4, western blot analysis was performed (Fig. 1J). In contrast to *XLF*$^{-/-}$ cells, which expressed XRCC4 at WT levels, and *XRCC4*$^{-/-}$ cells, which expressed XLF at near WT levels, each *XLF*$^{-/-}$:*XRCC4*$^{Neo/-}$ cell line exhibited the complete lack of expression of both proteins.

3.4. Generation of PAXX$^{-/-}$, PAXX$^{-/-}$:XRCC4$^{-/-}$, PAXX$^{-/-}$:XLF$^{-/-}$ and PAXX$^{-/-}$:XLF $^{-/-}$:XRCC4$^{-/-}$ HCT116 cell lines

In order to complete our set of epistasis reagents, we next generated *PAXX*$^{-/-}$ cell lines in four genetic backgrounds, including the parental HCT116, *XRCC4*$^{-/-}$, *XLF*$^{-/-}$, and *XLF* $^{-/-}$:*XRCC4*$^{-/-}$ cells utilizing the CRISPR/Cas9 gene targeting system. To this end, a CRISPR guide sequence (5'-TGACCGACGCCGCGGAGCTT-3′) that is specific to exon 2 of the *PAXX* gene was used to disrupt the gene. Successfully targeted HCT116 clones were identified by PCR. Candidate clones were selected based on the presence of a PCR product smaller than that produced from non-transfected HCT116 cells. Finally, the loss of PAXX expression in each cell line was confirmed by western blot analysis (Fig. 1K).

3.5. Biological endpoints for XRCC4-defective cell lines — DNA damage sensitivity

As XRCC4 is a key member of the C-NHEJ pathway it was expected and has been repeatedly demonstrated that lowered levels of XRCC4 should lead to sensitivity to DNA damaging agents [28,55–59]. In order to confirm this expectation in HCT116 cells, the cells were exposed to either increasing levels of etoposide, a topoisomerase II poison that induces DNA DSBs [60], or to increasing doses of IR. As expected, in contrast to *XRCC4*$^{+/-}$ cells, which showed sensitivity to etoposide at levels nearly equivalent to wild-type (WT) cells, *XRCC4*$^{-/-}$ cells exhibited an extreme sensitivity to etoposide at levels greater than an order of magnitude compared to WT cells (Fig. 2B). Interestingly, both low (#5) and high (#4) levels of XRCC4 expression in *XRCC4*$^{-/-}$:XRCC4 complemented cell lines resulted in complete suppression of the etoposide sensitivity. Similar to the etoposide sensitivity, *XRCC4*$^{-/-}$ cells displayed an extreme hypersensitivity to IR. The WT cell line had a D_{37} (the dose required to reduce survival to 37%) of ~ 2 Gy, the *XRCC4* $^{+/-}$ cell line was ~ 1.5 Gy, whereas the *XRCC4*$^{-/-}$ cell line was ~ 0.2 Gy (Fig. 2C). Both *XRCC4*$^{-/-}$:*XRCC4* complemented cell lines (#4 and #5) displayed significant complementation, with IR sensitivities intermediate between WT and *XRCC4*$^{+/-}$ levels (Fig. 2C).

3.6. NHEJ DNA repair activity in XRCC4-, XLF-, and PAXX-deficient cells

While the sensitivity of mammalian *XRCC4*-null cells to DNA damaging agents has been well documented [28,55–59], the impact of the absence of XRCC4 on DSB repair activity has not. Thus, to determine what effect the absence of XRCC4, XLF or PAXX have on NHEJ, an extra-chromosomal assay consisting of a GFP reporter plasmid (pEGFP-Pem1-Ad2) was used [43–45]. This plasmid allows for both the analysis of the relative repair frequency as well as the visualization of individual repair events by subsequent DNA sequencing. The plasmid contains an EGFP gene interrupted by a 2.4 kb intron derived from the rat phosphatidylethanolamine methyltransferase 1 (*Pem1*) gene that is interrupted by an adenoviral (Ad2) exon, which itself is flanked on both sides with HindIII and I-SceI restriction enzyme recognition sites (Fig. 3A). In its unmodified form, the presence of an adenoviral (Ad2) exon within the *Pem1* intron prevents the expression of EGFP due to the Ad2 exon being incorporated into the GFP mRNA (Fig. 3C). Upon digestion with either a HindIII or I-SceI restriction enzyme, linearized plasmids lacking the Ad2 exon with either compatible cohesive ends (HindIII) or incompatible ends (I-SceI) are generated. While the digestion of the HindIII recognition sites results in cohesive 4-bp overlapping ends, the fact that the I-SceI sites have been arranged in an inverted orientation requires the processing of the DNA ends before they can be rejoined (Fig. 3B). Importantly, both undigested and partially digested plasmids will not contribute to the GFP positive readout for this assay as any inclusion of the Ad2 exon will create mRNAs with an interrupted EGFP coding region (Fig. 3C). In addition, the introns spanning the Ad2 exon provide a substrate that can undergo long-range resection events before losing the capability of expressing EGFP. Therefore, the impact of the loss of XRCC4 expression on end joining can be assessed by FACS analysis. In order to determine the repair efficiency, cells are co-transfected with a pCherry expression vector that acts as a transfection control. The repair efficiency is reported as a percentage of cherry-positive cells that are also green (i.e., repaired). Finally, because the pEGFP-Pem1-Ad2 vector contains a *ColE1* bacterial origin of replication and a kanamycin antibiotic resistance gene, repaired plasmids recovered from human cells can be

efficiently rescued in *E. coli*. Subsequently, the resulting repair junctions can be determined by DNA sequencing, allowing for the elucidation of the possible repair mechanisms used during the repair event.

When parental (WT) HCT116 cells were transfected with either a HindIII- or I-SceI-linearized pEGFP-Pem1-Ad2 plasmid along with a pCherry transfection control plasmid, a baseline level of doubly EGFP- and Cherry-positive cells was established by flow cytometry (Fig. 3D). The loss of one copy of *XRCC4* reduced the repair efficiency only slightly, while in contrast, the complete loss of XRCC4 expression reduced the repair frequency by an order of magnitude. Importantly, as had been observed with the DNA damage sensitivity assays, the repair frequency for each *XRCC4*$^{-/-}$ complemented cell line was restored to near WT levels (Fig. 3D). In addition, the absence of XLF also led to a marked reduction in relative rejoining while the loss of both XLF and XRCC4 mirrored the results observed with the single mutant *XRCC4*$^{-/-}$ cells. Somewhat unexpectedly, there was no major difference in repair frequencies observed between the HindIII- or I-SceI-digested plasmids (Fig. 3E) suggesting that the configuration of the DSB ends had little effect on repair efficiency. Interestingly, the absence of PAXX in all possible genetic backgrounds: *i*) alone, *ii*) combined with *XLF*-null, *iii*) combined with *XRCC4*-null or *iv*) combined with *XRCC4*:*XLF* doubly-null cells had no significant additional impact on the relative repair efficiency of that cell line (Fig. S1). From these experiments we concluded that, in human somatic cells, XRCC4 appeared to be the most important factor for C-NHEJ repair activity. The absence of XLF also significantly negatively impacted C-NHEJ whereas the presence or absence of PAXX appeared irrelevant.

In addition to the repair frequencies acquired by FACS analysis, the repair events were characterized molecularly. For this analysis at least 30 plasmids were recovered from each cell line encompassing multiple independent experiments. The HindIII-linearized plasmid contains 4 bp compatible overhangs that essentially constitute microhomology. If this microhomology is utilized during the repair event, it will produce a plasmid that now contains one, instead of the former two, HindIII sites. To determine this frequency of "perfect joining", 20 colony PCR products generated from recovered pEGFP-Pem1-Ad2 plasmids for each cell line were subjected to HindIII digestion (Fig. 4A). While the WT and *XRCC4*$^{+/-}$ cell lines were able to perfectly rejoin the substrate 20% and 25% of the time, respectively, the *XRCC4*$^{-/-}$ cell line exhibited an increase to a frequency of 50% perfect rejoining that indicated an increased dependence on microhomology-mediated repair (Fig. 4A). Interestingly, the *XLF*$^{-/-}$ and *XLF*$^{-/-}$:*XRCC4*$^{\mathrm{Neo}/-}$ cell lines, appeared to be even more dependent on microhomology than cells lacking XRCC4 alone as they exhibited perfect joining frequencies of 70% and 75%, respectively.

Finally, for each cell line, at least 32 plasmids, recovered from non-digestible HindIII repair events as well as from I-SceI repair events, were sequenced (Tables S1 and S2, respectively). The degree of microhomology-mediated repair, using 3 or more bp as the cutoff for defining a microhomology-mediated end joining event [61], was strikingly increased in the *XRCC4*$^{-/-}$, *XLF*$^{-/-}$ as well as in the *XLF*$^{-/-}$:*XRCC4*$^{\mathrm{Neo}/-}$ cell lines (Fig. 4B, Tables S1 and S2). Indeed, while only 50% of the plasmid recovered from *XRCC4*-null cells showed perfect rejoining (Fig. 4A), 100% of the plasmids recovered from the same cells used some

degree of microhomology in their repair event and the length of microhomology was often quite large (an average of 6 nt; Tables S1 and S2). Some regions of microhomology seemed to be preferentially used for repair. For example, an eight bp sequence (GAGTGACC) was used 33% of the time during the repair of HindIII-linearized plasmids in *XRCC4*$^{-/-}$ cells and 25% of the time in the case of I-SceI-restricted plasmids repaired in *XLF*$^{-/-}$:*XRCC4*$^{Neo/-}$ cells (Tables S1 and S2). While theoretically some of these events could have been produced clonally while propagating the plasmids through *E. coli*, this possibility was experimentally excluded as the plasmids were recovered from three completely separate experiments, confirming their independent origins.

Unexpectedly, there appeared to be a bias in the symmetry with which the DSB ends were resected; with the 5′-end of the linearized plasmid being resected somewhat more than the 3′-end in WT, *XRCC4*$^{+/-}$ and complemented *XRCC4*-null cell lines (Fig. S2A and S2B). Interestingly, this asymmetry was by and large absent in the *XRCC4*-null, *XLF*-null and *XRCC4:XLF*-double-null cell lines. The mechanistic basis of this bias and its absence in the mutant cell lines is not understood, but clearly may have to do with differential access by nucleases to the DNA ends in the presence and absence of the XRCC4 paralogs. For each repair event, the amount of resection symmetry was determined by subtracting the total 3′ resection from the total 5' resection. Perfect symmetry was defined as having resection events that were equidistant from the point of the initial resection and thus yielded a value of zero. Repair events that were more asymmetrical on the 5'-side were given a positive value while those on the 3′-side a negative value. Importantly, to eliminate any bias that may arisen from the inclusion of clonally produced (either in HCT116 cells or in *E. coli*) repair events, we included only unique repair events for this analysis (*i.e.*, the four clones that utilized the GAGTGACC microhomology — Table S1 — to facilitate their repair in *XRCC-4* null cells were only counted as a single event). When the resection symmetry values for the repair of the HindIII-digested plasmid were thus assessed, it was clear that *XRCC4*$^{-/-}$ (20%), *XLF*$^{-/-}$ (25%) and *XLF*$^{-/-}$:*XRCC4*$^{Neo/-}$ (16%) cells had fewer asymmetric events than any of the control cell lines (Fig. S2C).

Lastly, the resection events appeared to be somewhat biphasic, with the majority of events falling within 300 bp of the DSB and the residual being quite large (> 1000 bp; note that the plasmids were randomly recovered and not selected for and thus there was no biological restraint on the extent of the deletions and as such some of the deletions removed the promoter or the EGFP coding sequences). As each of these plots (Fig. S2A and S2B) contained many repair events that lay close to zero, the total percent of clones that had a resection symmetry less than or equal to 300 was determined (Fig. S2C). In agreement with the previous results, both *XRCC4*$^{-/-}$, *XLF*$^{-/-}$ and *XLF*$^{-/-}$:*XRCC4*$^{Neo/-}$ cells displayed an elevated amount of symmetrical repair events compared to WT and *XRCC4*$^{+/-}$ cells. Interestingly, a similar biphasic nature has been observed for deletions associated with aberrant V(D)J recombination events in murine DNA-PK-defective cells [62]. Whether these *in vivo* observations are a reflection of the well-accepted biphasic mechanistic nature of DSB end resection [63] is currently not clear.

3.7. The loss of XRCC4 leads to an increase in aEJ microhomology-mediated repair events in human somatic cells

In order to further examine the extent at which microhomology-mediated repair plays a role in the repair of DSBs in *XRCC4*-null cells, we employed a second reporter plasmid (pDVG94). This plasmid predominately detects aEJ repair events and allows the overall efficiency of aEJ versus C-NHEJ to be determined [45,46]. pDVG94 contains both AfeI and EcoRV restriction enzyme sites that, when digested, results in a blunt-ended linear plasmid with 6-bp repeats at both ends (Fig. 5A). C-NHEJ can repair this plasmid in multiple ways, which result in the generation of many different repair junctions while in contrast aEJ almost exclusively produces one product, which includes the use of microhomology-mediated repair to combine the two repeats into one and in the process of so doing generates a unique BstXI restriction enzyme recognition site. PCR amplification of plasmids recovered after repair using primers that flank the repair junction yield a ~ 180 bp product that is digestible by BstXI if it was repaired using microhomology (Fig. 5B). Briefly, WT, $XRCC4^{+/-}$, $XRCC4^{-/-}$, and $XRCC4^{-/-}$ complemented cell lines were transfected with the linearized pDVG94 plasmid. The repaired plasmids were recovered from the cells 48 h later, purified and then used in a PCR reaction that generated the 180 bp product that included the repair junction. The PCR products were then subjected to BstXI restriction enzyme digestion after which they were electrophoresed along with the uncut product on a polyacrylamide gel and visualized using SYBR Gold (Fig. 5C). The quantitation of two independent experiments demonstrated that the WT HCT116 cell line as well as the $XRCC4^{+/-}$ and both $XRCC4^{-/-}$ complemented lines exhibited microhomology-mediated repair levels near 20%. In sharp contrast, the $XRCC4^{-/-}$ cell line showed ~90% microhomology use (Fig. 5C and 5D). These results agree with those observed using the pEGFP-Pem1-Ad2 vector (Fig. 3,4, Tables S1 and S2) and show that cells lacking XRCC4 are severely reduced in their ability to perform C-NHEJ, but instead utilize an alternative pathway (aEJ) for DSB repair.

3.8. V(D)J recombination is impaired in XRCC4-deficient cells

XRCC4 is clearly required for lymphoid V(D)J recombination, however a detailed analysis of the defects associated with the absence of XRCC4 in human cells has not been reported. To address this deficiency, we utilized an extrachromosomal vector-based assay. In brief, three plasmid vectors, pGG49, pGG51, and pGG52, which are able to detect signal joints, coding joints, and inversion or hybrid joints, respectively [47], were utilized (Fig. 6A). Importantly, for each vector, two V(D)J RSSs flank a premature prokaryotic transcription terminator (stop) that prevents the expression of chloramphenicol (Cam) unless it is removed by V(D)J recombination. Each RSS pair is oriented such that they will form the indicated joins upon successful V(D)J recombination. In addition to the chloramphenicol resistance gene, these plasmids also contain an ampicillin (Amp) resistance gene that allows for a determination of V(D)J recombination frequencies by transforming *E. coli* with reporter plasmids recovered from the human cells, and plating the bacteria onto both Amp and Amp + Cam selection plates. The V(D)J recombination frequency for each cell line is determined by dividing the total number of colonies formed on Amp + Cam plates by the number of colonies formed on the Amp plates. As HCT116 cells do not express the RAG proteins that are necessary for recognizing and cleaving the recombination signal sequences two additional plasmids, pRag1 and pRag2, were co-transfected into target cells along with each

of the individual V(D)J assay vectors. For each cell line tested, at least three independent transfections were performed.

When this assay was performed on parental HCT116 cells, a baseline average level of 0.19% SJ, 0.22% CJ, and 0.04% Inv/Hyb joint repair efficiency was observed (Fig. 6B). In sharp contrast, the *XRCC4*$^{-/-}$ cell line yielded dramatically reduced repair efficiencies exhibiting an average of 0.002% SJ, 0.02% CJ, and 0.006% Inv/Hyb joints or relative recombination frequencies of 1%, 9% and 13%, respectively (Fig. 6B). In addition to the marked decrease in repair efficiencies seen between WT and *XRCC4*$^{-/-}$ cells, when 30 random WT and 40 random *XRCC4*$^{-/-}$ Inv/Hyb repair events were sequenced, the overall ratio of Inversion/Hybrid joint formation was strikingly different with WT cells producing 76% inversion joints and 24% hybrid joints while *XRCC4*$^{-/-}$ cells produced only 10% inversion joints and 90% hybrid joints. As a control, *XLF*$^{-/-}$ cells were, as expected [55], highly V(D)J defective with average repair efficiencies of 0.009% SJ, 0.001% CJ, and 0.0% Inv/Hyb joint repair yielding relative repair efficiencies of 5%, 1%, and 0%, respectively (Fig. 6B). Impressively, even though the single knockout cell lines carried out only very low levels of V(D)J recombination, the *XLF*$^{-/-}$:*XRCC4*$^{Neo/-}$ double knockout cell line was reduced by yet at least another order of magnitude for both SJ (0.0002%) and CJ (0.0001%) formation resulting in relative repair efficiencies of 0.1% for both repair substrates (Fig. 6C).

A sequence analysis of V(D)J recombination repair junctions revealed striking differences between *XRCC4*-, *XLF*- and *XLF*:*XRCC4*- doubly-deficient cells. During V(D)J recombination RAG1 cleaves the RSS at the 5' end of the heptameric sequence, 5'-CACAGTG-3′. Sequenced repair junctions producing signal joints, coding joints, as well as both inversion and hybrid joints can therefore be aligned in comparison to the starting un-cleaved substrate to identify nts lost during the rearrangement process (Fig. S3, S4, S5 and S6). Signal joints are formed by the direct joining of these sequences leading to the production of a unique ApaLI restriction enzyme recognition site, 5'-GTGCAC-3′. Therefore, when examining signal joints, the PCR products were first assessed by digestion with ApaLI. The parental (WT) HCT116 cells displayed 100% ApaLI sensitivity and were thus inferred to represent perfect signal joining (40 samples, Fig. S3). Three of these were randomly sequenced and confirmed to be perfect signal joints. The *XRCC4*$^{-/-}$ cell line, in contrast, not only produced no perfect signal joints, but displayed some evidence of microhomology-based repair with significant resection seen in almost all of the 47 samples examined (Fig. S3). Surprisingly, the *XLF*$^{-/-}$ cell line more closely resembled the parental cell line with a perfect signal joint frequency of 63% and a predominance of shorter resection than seen in *XRCC4*$^{-/-}$ cells (Fig. S3). The frequency of V(D)J recombination was so low in the *XLF*$^{-/-}$:*XRCC4*$^{Neo/-}$ double-knockout cell line that only a single signal joint was recovered and this showed significant resection and the use of microhomology.

Coding joints form the functional basis of antibody specificity. Parental HCT116 cells displayed a wide variety of repair junctions with 21 (62%) unique junctions out of 34 plasmids sequenced (Fig. S4). Similar to the results seen when examining signal joints, the *XRCC4*$^{-/-}$ cell line displayed evidence for the use of microhomology-mediated repair with 11 (48%) of the 23 coding joints examined being the same microhomology-mediated repair event. Moreover, the number of nucleotides deleted during coding end formation was biased

toward larger deletions in the *XRCC4*$^{-/-}$ cells than in the WT cells (Fig. 6D). Similar results were observed with the limited amount of clones recovered from both the *XLF*$^{-/-}$ and *XLF* $^{-/-}$: *XRCC4*$^{Neo/-}$ cell lines.

When the RSSs are arranged such that neither signal nor coding joints can be formed, two outcomes are possible: either the intervening sequence can be inverted, resulting in an inversion joint, or the intervening sequence is removed and the 5' coding junction is ligated to the 3′ signal junction resulting in a hybrid joint (Fig. 6A). When V(D)J recombined inversion/hybrid joint reporter plasmids were recovered from WT cells, 23 of 30 were found to contain inversions while the remaining 7 were hybrid joints yielding an approximately 3:1 inversion/hybrid ratio (Fig. S5, S6). Conversely, when the same repair events were examined in *XRCC4*$^{-/-}$ cells only 4 of 40 plasmids could be verified to be inversions yielding a 1:10 inversion/hybrid ratio (Fig. S5, S6). Subsequent sequencing of the junctions formed in each cell line revealed sharp differences in their ability to repair these recombination events. WT HCT116 cells displayed a variety of different inversion joints with two out of 15 appearing to be produced using two base pairs of microhomology, one of which was recovered three independent times (Fig. S5). In contrast, only two different inversional repair events were recovered from the *XRCC4*$^{-/-}$ cells. Each contained either 2 or 6 bp of microhomology, the latter of which was recovered three times (Fig. S5). Interestingly, the signal junction in both of these clones was unexpectedly perfect. Similarly, the hybrid joints produced in WT cells revealed that very little resection occurred before the ends were ligated. However, the hybrid joints recovered from the *XRCC4*$^{-/-}$ cells were all resected well beyond the RSS that is normally ligated bluntly to the coding joint (Fig. S6). From these data, we concluded that the absence of XRCC4 is more deleterious than the absence of XLF and that virtually all of the V(D)J recombination products formed in *XRCC4*-null cells are aberrant with extensive deletions and a high use of microhomology.

3.9. The loss of XRCC4 leads to a variable impact on rAAV-mediated gene targeting

As the absence of efficient C-NHEJ activity due to an absence of Ku is associated with increased gene targeting [64], we thought that cells lacking XRCC4 may also exhibit a similar phenotype. To test this hypothesis, rAAV-mediated gene targeting vectors were used to target three different genes including exon 2 of the *Artemis* gene located on chromosome 10, exon 3 of the *HPRT* gene located on the X chromosome and finally exon 6 of *PIGA* also located on the X chromosome. For both the *Artemis* and *HPRT* genes, a standard method of gene targeting frequency determination was employed. Briefly, cells were infected with the rAAV vector of choice followed by drug selection that would allow only those cells infected by the rAAV targeting vector to grow into colonies. As rAAV-mediated gene targeting can result in random targeting as well as targeting at the desired locus, the number of correct targeting events can be determined by the use of a PCR-based screen using primers specific for the targeting vectors and genomic flanking primers. When the *Artemis* gene was targeted, *XRCC4*$^{-/-}$ HCT116 cells exhibited a 3.5-fold increase in gene targeting compared to WT cells (Fig. 7). In contrast, when the *HPRT* gene was targeted, there was no statistical increase in gene targeting observed in the *XRCC4*-deficient cells (Fig. 7). Finally, the gene targeting frequency was actually reduced ~ 2-fold in *XRCC4*-null cells when the *PIGA* gene was targeted (Fig. 7). Altogether, these results demonstrate that, unlike Ku-deficient cells,

where gene targeting is uniformly elevated [64], *XRCC4*-deficient cells do not show increased frequencies of rAAV-mediated gene targeting.

4. Discussion

4.1. A human somatic cell model for the absence of XRCC4

Human C-NHEJ has been rather exhaustively studied in either patient-derived or immortalized cancer cell lines genetically engineered to contain loss-of-function mutations. One exception to this is the C-NHEJ factor, XRCC4. One of the reasons for this is that *XRCC4* mutant patients present with dwarfism or encephalocardiomyopathy [18,19]. This is in stark contrast to almost all of the other described C-NHEJ mutant human patients [6], who generally present with some form of radiation sensitive-severe combined immune deficiency (RS-SCID). As a consequence, there have been almost no published reports utilizing *XRCC4* patient-derived cell lines. Similarly, there have been only a few descriptions of somatic human cancer cells engineered to lack the expression of XRCC4: these include HCT116 [28,45,56], 293 T [57] and HAP1 [58,59] cell lines. However, in all of these publications, virtually the only phenotype that has been characterized has been sensitivity to DNA damaging agents. The lack of the expected phenotypes in patient-derived cell lines and the paucity of somatic cell models prompted us to establish a *XRCC4* loss-of-function mutation in the HCT116 cell line.

4.2. XRCC4-null cells display defective C-NHEJ and a greatly increased dependence on microhomology-mediated repair

We have, to our knowledge, carried out the first extensive analysis of the impact of the absence of human XRCC4 on C-NHEJ (Fig. 3, 4B, S1, S2, Table S1, S2), aEJ (Fig. 4A,5) and V(D)J recombination (Fig. 6, S3–S6). These studies demonstrate that the absence of XRCC4 virtually abrogates C-NHEJ activity and V(D)J recombination and as such effectively phenocopies *LIGIV*-null cells [54,65], which had previously been the most useful experimental human C-NHEJ-defective cell line. Although many aspects of the residual *XRCC4*-null DSB repair phenotypes were interesting, we would like to emphasize two features here: 1) the reliance on microhomology, and 2) the completely inverted asymmetry between inversions and hybrid joint formation during V(D)J recombination.

The use of microhomology in residual DSB repair events in *XRCC4*-null cells was characterized using two experimental approaches: either by providing pre-formed microhomology at the ends of the DSBs (HindIII-restricted pEGFP-Pem1-Ad2 or linearized pDVG94) or by sequencing the resulting junctions when non-homologous ends (I-SceI-restricted pEGFP-Pem1-Ad2 and all V(D)J recombination products) were provided. Impressively, regardless of whether there existed pre-formed terminal microhomology or not, virtually every *XRCC4*-null DSB repair event utilized some amount of microhomology, especially when DSB repair was assessed. Importantly, the laboratory of Jeremy Stark has recently elegantly defined the use of 3 or more nts of microhomology as evidence of (albeit not proof of) aEJ usage whereas 1 or 2 nts of microhomology can still be consistent with C-NHEJ [61]. In virtually of the junctions formed in *XRCC4*-null cells at least 3 nt of microhomology was observed (Tables S1, S2, Fig. S3). Thus, it seems likely that the loss of

XRCC4, like that of Ku [45] and LIGIV [54,65], essentially ablates C-NHEJ and forces the cell to utilize aEJ to perform end joining. This abrogation of C-NHEJ stands in contrast to the loss of either DNA-PK_{cs} or XLF; where the loss of either of these factors drastically impairs C-NHEJ, but does not ablate it ([66], Fig. 3). The simplest interpretation of these data is that without the C-NHEJ end-binding factor (Ku) or the C-NHEJ ligation complex (XRCC4 and LIGIV), repair intermediates are ultimately repaired by the aEJ pathway. This "either/or" model is, however, likely overly simplistic. Thus, while it is true that the absence of human XRCC4 reduces the frequency of chromosomal translocations [67] and increases the use of microhomology found at the residual translocations [68] the absence of LIGIII and LIGIV (which, theoretically should ablate both aEJ and C-NHEJ) still permits translocations with microhomology-mediated junctions [69]. Thus, human cells may very well contain an additional pathway(s) of end joining besides the two well-recognized pathways of C-NHEJ and aEJ and this pathway(s) may also rely on microhomology.

Wild type cells presented with a V(D)J recombination substrate that can undergo inversion, will, as expected, invert those sequences the majority of the time. Occasionally, the ends can become mis-sorted and a hybrid junction (which is a biologically non-functional recombination product) results instead. When V(D)J inversion/hybrid joint reporter plasmids were recovered from WT cells, they showed a 3:1 inversion/hybrid ratio (Fig. S5, S6). In sharp contrast, when the same repair events were examined in *XRCC4*$^{-/-}$ cells the frequency of inversions to hybrid joints was heavily skewed in a 1:10 inversion/hybrid ratio (Fig. S5, S6). We interpret this data as some of the strongest evidence for a role of XRCC4 in bridging DNA molecules. Thus, a productive inversional rearrangement requires that all 4 DNA ends, the RAG proteins and the C-NHEJ repair proteins exist in a synaptic complex [11]. In this configuration, a coding end is joined to another coding end and the two RSS ends are joined together. If the synaptic complex is somehow disturbed, a coding end can aberrantly join toRSS end forming a hybrid junction [70]. Thus, the 30-fold increase in hybrid junction formation strongly suggests that the RAG:DNA synaptic complex cannot be properly formed in *XRCC4*-null cells. While other mechanisms are possible, we propose that XRCC4′s ability to form filaments (alone or in conjunction with XLF) is critical for the assembly or maintenance of the synaptic complex. Thus, in the absence of XRCC4, the V(D)J DNA ends within the aberrant complex can now incorrectly be resorted at a much higher frequency.

4.3. XRCC4-null cells do not display an increase in gene targeting frequency

Gene targeting (*aka*, precision genome engineering) requires the use of HDR to replace or alter genomic sequences with DNA sequences of the investigator's choosing [71]. It is intuitively appealing to believe, and has been repeatedly proposed, that the various cellular DSB repair pathways may compete for the ends of the donor and/or the recipient DNA and thus influence the frequency with which gene targeting can occur [72]. A corollary of this hypothesis is that if the competing pathways of C-NHEJ and/or aEJ were disrupted, then HDR factors should have unimpeded access to the DNA ends and so greatly augment the gene targeting process. In the case of the absence of Ku, this turns out to be true. Filamentous fungi, which normally carry out only low or modest levels of gene targeting, perform gene targeting at ~100% in the absence of Ku [73–77]. Moreover, human cells with

significantly reduced levels of Ku (the complete absence of Ku is lethal in humans [78];) also show highly elevated rates of gene targeting [64]. This attribute, however, is clearly not provided by the complete absence of XRCC4. While gene targeting was 4-fold elevated at the *Artemis* locus, it was unchanged at the *HPRT* locus and even reduced at the *PIGA* locus when *XRCC4*-null cells were compared to the parental HCT116 cell line (Fig. 7). This observation is consistent with the small (2-fold) increase in gene targeting observed with *LIGIV*-null cells [54] and a similar lack of a positive effect on gene targeting in *XLF*-heterozygous cells [55].

The most parsimonious interpretation of these results is that the fate of DNA ends in human cells is predominately determined by the proximal DNA end binding factor, Ku. When Ku binds a DNA end it shunts it into the C-NHEJ repair pathway. In Ku's absence, other DNA end binding factors — presumably HDR-relevant factors — can instead bind and force the ends into the HDR pathway [79–81]. In contrast, in the absence of other C-NHEJ factors (LIGIV, XLF and XRCC4), Ku is still present and it presumably (and futilely) still tries to shunt the DNA ends down the now inactivated C-NHEJ pathway. This non-functional sequestration of the ends does not, therefore, enhance the number of ends available for HDR and no increase in gene targeting is observed. Altogether, these studies strongly suggest that small molecule inhibitors to Ku would be the most efficacious way forward in enhancing gene targeting in human cells.

4.4. XRCC4 paralogs are not created equal

There are three XRCC4 paralogs (XRCC4, XLF and PAXX) that are structurally related [26,37–39]. Moreover, there is significant precedent in biology for paralogs providing functional redundancy for one another (*e.g.* in tubulin and actin gene families). This is certainly the case with the XRCC4 paralogs as well. Thus, for example, while *XLF*-and *PAXX*-null mice are viable, a double knockout mouse is not [59,82] strongly suggesting that the two proteins can, at some level, compensate for one another. With that said, the paralogs are also clearly not equivalent in terms of their contribution to cellular viability, DNA DSB repair activity nor even V(D)J recombination capacity in the mouse [29,30,57,59,82,83] and in human cells [28,57]. In our HCT116 human somatic cell system, the loss of XRCC4 and XLF nearly equivalently reduced the frequency of V(D)J recombination by 2 orders of magnitude over wild type cells (Fig. 6). Nonetheless, there were distinct differences between these cell lines. Thus, *XRCC4*-null cells seemed to be less capable of forming signal junctions whereas *XLF*-null cells were less capable of forming coding junctions (Fig. 6, S3 and S4). Perhaps not surprisingly then, the doubly mutant *XRCC4:XLF*-null cells were reduced by another order of magnitude over each single mutant for both reactions, demonstrating that there are at least some recombination events for which these factors are redundant. In terms of C-NHEJ DSB repair activity, there was a clearer hierarchy between the factors. Thus, the loss of XRCC4 was in all instances, the most deleterious and the residual repair demonstrated an almost complete reliance on microhomology (Fig. 3, S1, Tables S1, S2). The loss of XLF, however, while clearly detrimental, still allowed for a modest level of repair activity and showed a corresponding reduced reliance on microhomology. In contrast, the loss of PAXX had almost no impact on C-NHEJ activity; either by itself or in combination with the loss of either XRCC4, XLF or both (Fig. S1).

Thus, in human HCT116 cells the apparent importance of the XRCC4 paralogs is XRCC4 > XLF > > > PAXX.

Homodimers of XRCC4 and XLF can physically interact and have been proposed to form filaments that hold or bridge the ends of a DSB together to facilitate repair [28,31,32,34–36,49,84,85]. If this was the proteins' only function, however, then the individual knockouts should phenocopy each other — and they don't; neither in mouse models nor in human somatic cell models. Instead, we suggest that as the partner and stabilizer of LIGIV, XRCC4 plays a more significant role because there is at least a subset of repair and recombination events that need little to no processing [8,9,86] and thus have no requirement for bridging and therefore XLF. This hypothesis is consistent with recent *in vitro* work, which demonstrated an absolute requirement for XRCC4 and a more auxiliary role for XLF in synapsis [40]. For the majority of repair events, however, some form of processing is likely required at either one or both ends. For these repair events, we envision that XRCC4:XLF filaments/sleeves are required to hold the ends in close proximity while the processing is occurring and to subsequently facilitate the ligation reaction. This model posits no role for PAXX, which is consistent with its lack of contribution to any phenotypes that we investigated. There is no experimental evidence to suggest that PAXX interacts with or is required for LIGIV activity nor is there evidence for its ability to form filaments either with XRCC4, XLF or XRCC4:XLF [40]. Thus, it seems apparent that PAXX likely functions in a mechanistically distinct manner from XRCC4 and XLF. To date, the one biochemical role ascribed to PAXX is its ability to physically interact with Ku [29,37–39,57]. What the PAXX interaction does biochemically to Ku is unknown but it has been suggested that it promotes Ku's accumulation at DNA ends [87]. Superficially, this would seem unlikely since Ku is highly abundant and has an extremely high affinity for DNA ends and consequently likely needs no assistance in DNA end binding. However, it is certainly possible that there are a minority of DNA ends that have a unique structure such that they are no longer good DNA substrates for Ku binding. We posit that these rare "special" ends may require PAXX to facilitate Ku binding (which, in turn, subsequently permits the processing and repair of that end). This model would explain why the absence of PAXX seems to have a negligible effect in virtually every DNA repair assay where it has been tested, but seems to have an impact *in vivo*. *In vitro* assays invariably utilize canonical DNA repair substrates with "clean" pre-cleaved ends whereas *in vivo*, special ends may have a chance to be formed, even if only rarely.

In conclusion, these experiments have revealed similarities and some striking differences between *XRCC4* paralog mutant cell lines. These cell lines will be powerful tools for studying both *XRCC4* paralog patient mutations as well as determining how these paralogs differ in their contributions to cellular DNA repair and recombination pathways.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding

Funding for the Hendrickson laboratory was provided in part through grants from the National Institutes of Health (GM088351) and the National Cancer Institute (CA154461 and CA190492). These agencies had no involvement in the study design, the data collection, the analysis nor the interpretations presented here.

Abbreviations:

Ad2	adenovirus 2
aEJ	alternative end joining
Cas9	CRISPR-associated 9
C-NHEJ	classic non-homologous end joining
Cre	cyclization recombinase
CRISPR	clustered regularly interspersed palindromic repeats
DNA-PK$_{cs}$/PRKDC	catalytic subunit of the DNA-dependent protein kinase complex/protein kinase DNA-activated, catalytic subunit
DSBs	double-stranded breaks
FACS	fluorescence-activated cell sort
Floxed	LoxP-flanked
GFP	green fluorescent protein
Gy	gray
HDR	homology dependent recombination
HPRT	hypoxanthine-guanine phosphoribosyltransferase
IR	ionizing radiation
Ku	Ku70 and Ku86 heterodimer
LIGIV	DNA ligase IV
LoxP	locus of crossover for phage P1
Neo	neomycin phosphotransferase
P	palindromic
PAXX	paralog of XRCC4 and XLF
PBS	phosphate buffered saline
PEM1	phosphatidylethanolamine methyltransferase 1

PIGA	phosphatidylinositol N-acetylglucosaminyltransferase subunit A
rAAV	recombinant adeno-associated virus
RAG-1	recombination activating gene-1
RAG-2	recombination activating gene-2
RSS	recombination signal sequence
RS-SCID	radiation sensitive-severe combined immune deficient
TBE	tris/borate/EDTA
V(D)J	variable(diversity)joining recombination
XLF	XRCC4-like factor
XRCC4	X-ray cross complementing 4

References

[1]. Tubbs A, Nussenzweig A, Endogenous DNA damage as a source of genomic instability in cancer, Cell 168 (2017) 644–656. [PubMed: 28187286]

[2]. Ait Saada A, Lambert SAE, Carr AM, Preserving replication fork integrity and competence via the homologous recombination pathway, DNA Repair (Amst.) (2018).

[3]. Alt FW, Zhang Y, Meng FL, Guo C, Schwer B, Mechanisms of programmed DNA lesions and genomic instability in the immune system, Cell 152 (2013) 417–429. [PubMed: 23374339]

[4]. Baird DM, Hendrickson EA, Telomeres and chromosomal translocations : there's a ligase at the end of the translocation, Adv. Exp. Med. Biol 1044 (2018) 89–112. [PubMed: 29956293]

[5]. Jasin M, Rothstein R, Repair of strand breaks by homologous recombination, Cold Spring Harb. Perspect. Biol 5 (2013) a012740. [PubMed: 24097900]

[6]. Woodbine L, Gennery AR, Jeggo PA, The clinical impact of deficiency in DNA non-homologous end-joining, DNA Repair (Amst) 16 (2014) 84–96. [PubMed: 24629483]

[7]. Frit P, Barboule N, Yuan Y, Gomez D, Calsou P, Alternative end-joining pathway (s): bricolage at DNA breaks, DNA Repair (Amst.) 17 (2014) 81–97. [PubMed: 24613763]

[8]. Oh S, Harvey A, Zimbric J, Wang Y, Nguyen T, Jackson PJ, Hendrickson EA, DNA ligase III and DNA ligase IV carry out genetically distinct forms of end joining in human somatic cells, DNA Repair (Amst.) 21 (2014) 97–110. [PubMed: 24837021]

[9]. Waters CA, Strande NT, Pryor JM, Strom CN, Mieczkowski P, Burkhalter MD, Oh S, Qaqish BF, Moore DT, Hendrickson EA, Ramsden DA, The fidelity of the ligation step determines how ends are resolved during nonhomologous end joining, Nat. Commun 5 (2014) 4286. [PubMed: 24989324]

[10]. Jette N, Lees-Miller SP, The DNA-dependent protein kinase: a multifunctional protein kinase with roles in DNA double strand break repair and mitosis, Prog. Biophys. Mol. Biol 117 (2015) 194–205. [PubMed: 25550082]

[11]. Helmink BA, Sleckman BP, The response to and repair of RAG-mediated DNA double-strand breaks, Annu. Rev. Immunol 30 (2012) 175–202. [PubMed: 22224778]

[12]. Ma Y, Pannicke U, Schwarz K, Lieber MR, Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination, Cell 108 (2002) 781–794. [PubMed: 11955432]

[13]. Roy S, Andres SN, Vergnes A, Neal JA, Xu Y, Yu Y, Lees-Miller SP, Junop M, Modesti M, Meek K, XRCC4's interaction with XLF is required for coding (but not signal) end joining, Nucleic Acids Res. 40 (2012) 1684–1694. [PubMed: 22228831]

[14]. Li Z, Alt FW, Identification of the XRCC4 gene: complementation of the DSBR and V(D)J recombination defects of XR-1 cells, Curr. Top. Microbiol. Immunol 217 (1996) 143–150. [PubMed: 8787623]

[15]. Li Z, Otevrel T, Gao Y, Cheng HL, Seed B, Stamato TD, Taccioli GE, Alt FW, The XRCC4 gene encodes a novel protein involved in DNA double-strand break repair and V(D)J recombination, Cell 83 (1995) 1079–1089. [PubMed: 8548796]

[16]. Grawunder U, Wilm M, Wu X, Kulesza P, Wilson TE, Mann M, Lieber MR, Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells, Nature 388 (1997) 492–495. [PubMed: 9242410]

[17]. Bryans M, Valenzano MC, Stamato TD, Absence of DNA ligase IV protein inXR-1 cells: evidence for stabilization by XRCC4, Mutat. Res 433 (1999) 53–58. [PubMed: 10047779]

[18]. Rosin N, Elcioglu NH, Beleggia F, Isguven P, Altmuller J, Thiele H, Steindl K, Joset P, Rauch A, Nurnberg P, Wollnik B, Yigit G, Mutations in XRCC4 cause primary microcephaly, short stature and increased genomic instability, Hum. Mol. Genet 24 (2015) 3708–3717. [PubMed: 25839420]

[19]. Murray JE, van der Burg M, Carroll IJH,P, Wu Q, Ochi T, Leitch A, Miller ES, Kysela B, Jawad A, Bottani A, Brancati F, Cappa M, Cormier-Daire V, Deshpande C, Faqeih EA, Graham GE, Ranza E, Blundell TL, Jackson AP, Stewart GS, Bicknell LS, Mutations in the NHEJ component XRCC4 cause primordial dwarfism, Am. J. Hum. Genet 96 (2015) 412–424. [PubMed: 25728776]

[20]. Mittal RD, Gangwar R, Mandal RK, Srivastava P, Ahirwar DK, Gene variants of XRCC4 and XRCC3 and their association with risk for urothelial bladder cancer, Mol. Biol. Rep 39 (2012) 1667–1675. [PubMed: 21617942]

[21]. Zhou LP, Luan H, Dong XH, Jin GJ, Ma DL, Shang H, Association of functional polymorphisms of the XRCC4 gene with the risk of breast cancer: a meta-analysis, Asian Pac. J. Cancer Prev 13 (2012) 3431–3436. [PubMed: 22994773]

[22]. Cifci S, Yilmaz M, Pehlivan M, Sever T, Okan V, Pehlivan S, DNA repair genes polymorphisms in multiple myeloma: no association with XRCC1 (Arg399Gln) polymorphism, but the XRCC4 (VNTR in intron 3 and G-1394T) and XPD (Lys751Gln) polymorphisms is associated with the disease in Turkish patients, Hematology 16 (2011) 361–367. [PubMed: 22183071]

[23]. Ahnesorg P, Smith P, Jackson SP, XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining, Cell 124 (2006) 301–313. [PubMed: 16439205]

[24]. Buck D, Malivert L, de Chasseval R, Barraud A, Fondaneche MC, Sanal O, Plebani A, Stephan JL, Hufnagel M, le Deist F, Fischer A, Durandy A, de Villartay JP, Revy P, Cernunnos, a novel nonhomologous end-joining factor, is mutated in human immunodeficiency with microcephaly, Cell 124 (2006) 287–299. [PubMed: 16439204]

[25]. Junop MS, Modesti M, Guarne A, Ghirlando R, Gellert M, Yang W, Crystal structure of the Xrcc4 DNA repair protein and implications for end joining, EMBO J. 19 (2000) 5962–5970. [PubMed: 11080143]

[26]. Andres SN, Modesti M, Tsai CJ, Chu G, Junop MS, Crystal structure of human XLF: a twist in nonhomologous DNA end-joining, Mol. Cell 28 (2007) 1093–1101. [PubMed: 18158905]

[27]. Li Y, Chirgadze DY, Bolanos-Garcia VM, Sibanda BL, Davies OR, Ahnesorg P, Jackson SP, Blundell TL, Crystal structure of human XLF/Cernunnos reveals unexpected differences from XRCC4 with implications for NHEJ, EMBO J. 27 (2008) 290–300. [PubMed: 18046455]

[28]. Roy S, de Melo AJ, Xu Y, Tadi SK, Negrel A, Hendrickson E, Modesti M, Meek K, XRCC4/XLF interaction is variably required for DNA repair and is not required for ligase IV stimulation, Mol. Cell. Biol 35 (2015) 3017–3028. [PubMed: 26100018]

[29]. Lescale C, Lenden Hasse H, Blackford AN, Balmus G, Bianchi JJ, Yu W, Bacoccina L, Jarade A, Clouin C, Sivapalan R, Reina-San-Martin B, Jackson SP, Deriano L, Specific roles of XRCC4 paralogs PAXX and XLF during V(D)J recombination, Cell Rep. 16 (2016) 2967–2979. [PubMed: 27601299]

[30]. Kumar V, Alt FW, Frock RL, PAXX and XLF DNA repair factors are functionally redundant in joining DNA breaks in a G1-arrested progenitor B-cell line, Proc. Natl. Acad. Sci. U. S. A 113 (2016) 10619–10624. [PubMed: 27601633]

[31]. Andres SN, Vergnes A, Ristic D, Wyman C, Modesti M, Junop M, A human XRCC4-XLF complex bridges DNA, Nucleic Acids Res. 40 (2012) 1868–1878. [PubMed: 22287571]

[32]. Hammel M, Rey M, Yu Y, Mani RS, Classen S, Liu M, Pique ME, Fang S, Mahaney BL, Weinfeld M, Schriemer DC, Lees-Miller SP, Tainer JA, XRCC4 protein interactions with XRCC4-like factor (XLF) create an extended grooved scaffold for DNA ligation and double strand break repair, J. Biol. Chem 286 (2011) 32638–32650. [PubMed: 21775435]

[33]. Hammel M, Yu Y, Fang S, Lees-Miller SP, Tainer JA, XLF regulates filament architecture of the XRCC4.lIgase IV complex, Structure 18 (2010) 1431–1442. [PubMed: 21070942]

[34]. Reid DA, Keegan S, Leo-Macias A, Watanabe G, Strande NT, Chang HH, Oksuz BA, Fenyo D, Lieber MR, Ramsden DA, Rothenberg E, Organization and dynamics of the nonhomologous end-joining machinery during DNA double-strand break repair, Proc. Natl. Acad. Sci. U. S. A 112 (2015) E2575–2584. [PubMed: 25941401]

[35]. Brouwer I, Sitters G, Candelli A, Heerema SJ, Heller I, de Melo AJ, Zhang H, Normanno D, Modesti M, Peterman EJ, Wuite GJ, Sliding sleeves ofXRCC4-XLF bridge DNA and connect fragments of broken DNA, Nature 535 (2016) 566–569. [PubMed: 27437582]

[36]. Reid DA, Conlin MP, Yin Y, Chang HH, Watanabe G, Lieber MR, Ramsden DA, Rothenberg E, Bridging of double-stranded breaks by the non-homologous end-joining ligation complex is modulated by DNA end chemistry, Nucleic Acids Res. 45 (2017) 1872–1878. [PubMed: 27924007]

[37]. Xing M, Yang M, Huo W, Feng F, Wei L, Jiang W, Ning S, Yan Z, Li W, Wang Q, Hou M, Dong C, Guo R, Gao G, Ji J, Zha S, Lan L, Liang H, Xu D, Interactome analysis identifies a new paralogue of XRCC4 in non-homologous end joining DNA repair pathway, Nat. Commun 6 (2015) 6233. [PubMed: 25670504]

[38]. Ochi T, Blackford AN, Coates J, Jhujh S, Mehmood S, Tamura N, Travers J, Wu Q, Draviam VM, Robinson CV, Blundell TL, Jackson SP, DNA repair. PAXX, a paralog of XRCC4 and XLF, interacts with Ku to promote DNA double-strand break repair, Science 347 (2015) 185–188. [PubMed: 25574025]

[39]. Craxton A, Somers J, Munnur D, Jukes-Jones R, Cain K, Malewicz M, XLS (c9orf142) is a new component of mammalian DNA double-stranded break repair, Cell Death Differ. 22 (2015) 890–897. [PubMed: 25941166]

[40]. Zhao B, Watanabe G, Morten MJ, Reid DA, Rothenberg E, Lieber MR, The essential elements for the noncovalent association of two DNA ends during NHEJ synapsis, Nat. Commun 10 (2019) 3588. [PubMed: 31399561]

[41]. Wright AV, Nunez JK, Doudna JA, Biology and applications of CRISPR systems: harnessing nature's toolbox for genome engineering, Cell 164 (2016) 29–44. [PubMed: 26771484]

[42]. Karnan S, Konishi Y, Ota A, Takahashi M, Damdindorj L, Hosokawa Y, Konishi H, Simple monitoring of gene targeting efficiency in human somatic cell lines using the PIGA gene, PLoS One 7 (2012) e47389. [PubMed: 23056640]

[43]. Seluanov A, Mittelman D, Pereira-Smith OM, Wilson JH, Gorbunova V, DNA end joining becomes less efficient and more error-prone during cellular senescence, Proc. Natl. Acad. Sci. U. S. A 101 (2004) 7624–7629. [PubMed: 15123826]

[44]. Wang M, Wu W, Wu W, Rosidi B, Zhang L, Wang H, Iliakis G, PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways, Nucleic Acids Res. 34 (2006) 6170–6182. [PubMed: 17088286]

[45]. Fattah F, Lee EH, Weisensel N, Wang Y, Lichter N, Hendrickson EA, Ku regulates the non-homologous end joining pathway choice of DNA double-strand break repair inhuman somatic cells, PLoS Genet. 6 (2010) e1000855. [PubMed: 20195511]

[46]. Verkaik NS, Esveldt-van Lange RE, van Heemst D, Bruggenwirth HT, Hoeijmakers JH, Zdzienicka MZ, van Gent DC, Different types of V(D)J recombination and end-joining defects in DNA double-strand break repair mutant mammalian cells, Eur. J. Immunol 32 (2002) 701–709. [PubMed: 11870614]

[47]. Gauss GH, Lieber MR, Unequal signal and coding joint formation in human V(D)J recombination, Mol. Cell. Biol 13 (1993) 3900–3906. [PubMed: 8321197]

[48]. Nicolas N, Moshous D, Cavazzana-Calvo M, Papadopoulo D, de Chasseval R, Le Deist F, Fischer A, de Villartay JP, A human severe combined immunodeficiency (SCID) condition with increased sensitivity to ionizing radiations and impaired V(D)J rearrangements defines a new DNA recombination/repair deficiency, J. Exp. Med 188 (1998) 627–634. [PubMed: 9705945]

[49]. Wu PY, Frit P, Meesala S, Dauvillier S, Modesti M, Andres SN, Huang Y, Sekiguchi J, Calsou P, Salles B, Junop MS, Structural and functional interaction between the human DNA repair proteins DNA ligase IV and XRCC4, Mol. Cell. Biol 29 (2009) 3163–3172. [PubMed: 19332554]

[50]. Grawunder U, Zimmer D, Lieber MR, DNA ligase IV binds to XRCC4 via a motif located between rather than within its BRCT domains, Curr. Biol 8 (1998) 873–876. [PubMed: 9705934]

[51]. Sibanda BL, Critchlow SE, Begun J, Pei XY, Jackson SP, Blundell TL, Pellegrini L, Crystal structure of an Xrcc4-DNA ligase IV complex, Nat. Struct. Biol 8 (2001) 1015–1019. [PubMed: 11702069]

[52]. Khan IF, Hirata RK, Russell DW, AAV-mediated gene targeting methods for human cells, Nat. Protoc 6 (2011) 482–501. [PubMed: 21455185]

[53]. Kohli M, Rago C, Lengauer C, Kinzler KW, Vogelstein B, Facile methods for generating human somatic cell gene knockouts using recombinant adeno-associated viruses, Nucleic Acids Res. 32 (2004) e3. [PubMed: 14704360]

[54]. Oh S, Wang Y, Zimbric J, Hendrickson EA, Human LIGIV is synthetically lethal with the loss of Rad54B-dependent recombination and is required for certain chromosome fusion events induced by telomere dysfunction, Nucleic Acids Res. 41 (2013) 1734–1749. [PubMed: 23275564]

[55]. Fattah FJ, Kweon J, Wang Y, Lee EH, Kan Y, Lichter N, Weisensel N, Hendrickson EA, A role for XLF in DNA repair and recombination in human somatic cells, DNA Repair (Amst.) 15 (2014) 39–53. [PubMed: 24461734]

[56]. Katsube T, Mori M, Tsuji H, Shiomi T, Shiomi N, Onoda M, Differences in sensitivity to DNA-damaging agents between XRCC4- and Artemis-deficient human cells, J. Radiat. Res 52 (2011) 415–424. [PubMed: 21785230]

[57]. Tadi SK, Tellier-Lebegue C, Nemoz C, Drevet P, Audebert S, Roy S, Meek K, Charbonnier JB, Modesti M, PAXX is an accessory c-NHEJ factor that associates with Ku70 and has overlapping functions with XLF, Cell Rep. 17 (2016) 541–555. [PubMed: 27705800]

[58]. Dewan A, Xing M, Lundbaek MB, Gago-Fuentes R, Beck C, Aas PA, Liabakk NB, Saeterstad S, Chau KTP, Kavli BM, Oksenych V, Robust DNA repair in PAXX-deficient mammalian cells, FEBS Open Bio 8 (2018) 442–448.

[59]. Castaneda-Zegarra S, Xing M, Gago-Fuentes R, Saeterstad S, Oksenych V, Synthetic lethality between DNA repair factors Xlf and Paxx is rescued by inactivation of Trp53, DNA Repair (Amst.) 73 (2019) 164–169. [PubMed: 30579708]

[60]. Montecucco A, Biamonti G, Cellular response to etoposide treatment, Cancer Lett. 252 (2007) 9–18. [PubMed: 17166655]

[61]. Bhargava R, Sandhu M, Muk S, Lee G, Vaidehi N, Stark JM, C-NHEJ without indels is robust and requires synergistic function of distinct XLF domains, Nat. Commun 9 (2018) 2484. [PubMed: 29950655]

[62]. Hendrickson EA, Schlissel MS, Weaver DT, Wild-type V(D)J recombination in scid pre-B cells, Mol. Cell. Biol 10 (1990) 5397–5407. [PubMed: 2118996]

[63]. Mimitou EP, Symington LS, DNA end resection-unraveling the tail, DNA Repair (Amst.) 10 (2011) 344–348. [PubMed: 21227759]

[64]. Fattah FJ, Lichter NF, Fattah KR, Oh S, Hendrickson EA, Ku70, an essential gene, modulates the frequency of rAAV-mediated gene targeting in human somatic cells, Proc. Natl. Acad. Sci. U. S. A 105 (2008) 8703–8708. [PubMed: 18562296]

[65]. Grawunder U, Zimmer D, Fugmann S, Schwarz K, Lieber MR, DNA ligase IV is essential for V(D)J recombination and DNA double-strand break repair in human precursor lymphocytes, Mol. Cell 2 (1998) 477–484. [PubMed: 9809069]

[66]. Ruis BL, Fattah KR, Hendrickson EA, The catalytic subunit of DNA-dependent protein kinase regulates proliferation, telomere length, and genomic stability in human somatic cells, Mol. Cell. Biol 28 (2008) 6182–6195. [PubMed: 18710952]

[67]. Ghezraoui H, Piganeau M, Renouf B, Renaud JB, Sallmyr A, Ruis B, Oh S, Tomkinson AE, Hendrickson EA, Giovannangeli C, Jasin M, Brunet E, Chromosomal translocations in human cells are generated by canonical non-homologous end-joining, Mol. Cell 55 (2014) 829–842. [PubMed: 25201414]

[68]. Liddiard K, Ruis B, Takasugi T, Harvey A, Ashelford KE, Hendrickson EA, Baird DM, Sister chromatid telomere fusions, but not NHEJ-mediated inter-chromosomal telomere fusions, occur independently of DNA ligases 3 and 4, Genome Res. 26 (2016) 588–600. [PubMed: 26941250]

[69]. Liddiard K, Ruis B, Kan Y, Cleal K, Ashelford KE, Hendrickson EA, Baird DM, DNA Ligase 1 is an essential mediator of sister chromatid telomere fusions in G2 cell cycle phase, Nucleic Acids Res. 47 (2019) 2402–2424. [PubMed: 30590694]

[70]. Raghavan SC, Tong J, Lieber MR, Hybrid joint formation in human V(D)J recombination requires nonhomologous DNA end joining, DNA Repair (Amst.) 5 (2006) 278–285. [PubMed: 16275127]

[71]. Hendrickson EA, Gene targeting in human somatic cells, in: Conn PM (Ed.), Source Book of Models for Biomedical Research, Humana Press, Inc., Totowa, NJ, 2008, pp. 509–525.

[72]. Kass EM, Jasin M, Collaboration and competition between DNA double-strand break repair pathways, FEBS Lett. 584 (2010) 3703–3708. [PubMed: 20691183]

[73]. Kooistra R, Hooykaas PJ, Steensma HY, Efficient gene targeting in Kluyveromyces lactis, Yeast 21 (2004) 781–792. [PubMed: 15282801]

[74]. Ninomiya Y, Suzuki K, Ishii C, Inoue H, Highly efficient gene replacements in Neurospora strains deficient for nonhomologous end-joining, Proc. Natl. Acad. Sci. U. S. A 101 (2004) 12248–12253. [PubMed: 15299145]

[75]. Colot HV, Park G, Turner GE, Ringelberg C, Crew CM, Litvinkova L, Weiss RL, Borkovich KA, Dunlap JC, A high-throughput gene knockout procedure for Neurospora reveals functions for multiple transcription factors, Proc. Natl. Acad. Sci. U. S. A 103 (2006) 10352–10357. [PubMed: 16801547]

[76]. da Silva Ferreira ME, Kress MR, Savoldi M, Goldman MH, Hartl A, Heinekamp T, Brakhage AA, Goldman GH, The akuB(KU80) mutant deficient for nonhomologous end joining is a powerful tool for analyzing pathogenicity in Aspergillus fumigatus, Eukaryot. Cell 5 (2006) 207–211. [PubMed: 16400184]

[77]. Goins CL, Gerik KJ, Lodge JK, Improvements to gene deletion in the fungal pathogen Cryptococcus neoformans: absence of Ku proteins increases homologous recombination, and co-transformation of independent DNA molecules allows rapid complementation of deletion phenotypes, Fungal Genet. Biol 43 (2006) 531–544. [PubMed: 16714127]

[78]. Li G, Nelsen C, Hendrickson EA, Ku86 is essential in human somatic cells, Proc. Natl. Acad. Sci. U. S. A 99 (2002) 832–837. [PubMed: 11792868]

[79]. Zentilin L, Marcello A, Giacca M, Involvement of cellular double-stranded DNA break binding proteins in processing of the recombinant adeno-associated virus genome, J. Virol 75 (2001) 12279–12287. [PubMed: 11711618]

[80]. Di Primio C, Galli A, Cervelli T, Zoppe M, Rainaldi G, Potentiation of gene targeting in human cells by expression of Saccharomyces cerevisiae Rad52, Nucleic Acids Res. 33 (2005) 4639–4648. [PubMed: 16106043]

[81]. Ishibashi K, Suzuki K, Ando Y, Takakura C, Inoue H, Nonhomologous chromosomal integration of foreign DNA is completely dependent on MUS-53 (human Lig4 homolog) in Neurospora, Proc. Natl. Acad. Sci. U. S. A 103 (2006) 14871–14876. [PubMed: 17003123]

[82]. Balmus G, Barros AC, Wijnhoven PW, Lescale C, Hasse HL, Boroviak K, le Sage C, Doe B, Speak AO, Galli A, Jacobsen M, Deriano L, Adams DJ, Blackford AN, Jackson SP, Synthetic lethality between PAXX and XLF in mammalian development, Genes Dev. 30 (2016) 2152–2157. [PubMed: 27798842]

[83]. Hung PJ, Chen BR, George R, Liberman C, Morales AJ, Colon-Ortiz P, Tyler JK, Sleckman BP, Bredemeyer AL, Deficiency of XLF and PAXX prevents DNA double-strand break repair by

non-homologous end joining in lymphocytes, Cell Cycle 16 (2017) 286–295. [PubMed: 27830975]

[84]. Ropars V, Drevet P, Legrand P, Baconnais S, Amram J, Faure G, Marquez JA, Pietrement O, Guerois R, Callebaut I, Le Cam E, Revy P, de Villartay JP, Charbonnier JB, Structural characterization of filaments formed by human Xrcc4-Cernunnos/XLF complex involved in nonhomologous DNA end-joining, Proc. Natl. Acad. Sci. U. S. A 108 (2011) 12663–12668. [PubMed: 21768349]

[85]. Normanno D, Negrel A, de Melo AJ, Betzi S, Meek K, Modesti M, Mutational phospho-mimicry reveals a regulatory role for the XRCC4 and XLF C-terminal tails in modulating DNA bridging during classical non-homologous end joining, Elife 6 (2017).

[86]. Chang HH, Watanabe G, Gerodimos CA, Ochi T, Blundell TL, Jackson SP, Lieber MR, Different DNA end configurations dictate which NHEJ components are most important for joining efficiency, J. Biol. Chem 291 (2016) 24377–24389. [PubMed: 27703001]

[87]. Liu X, Shao Z, Jiang W, Lee BJ, Zha S, PAXX promotes KU accumulation at DNA breaks and is essential for end-joining in XLF-deficient mice, Nat. Commun 8 (2017) 13816. [PubMed: 28051062]

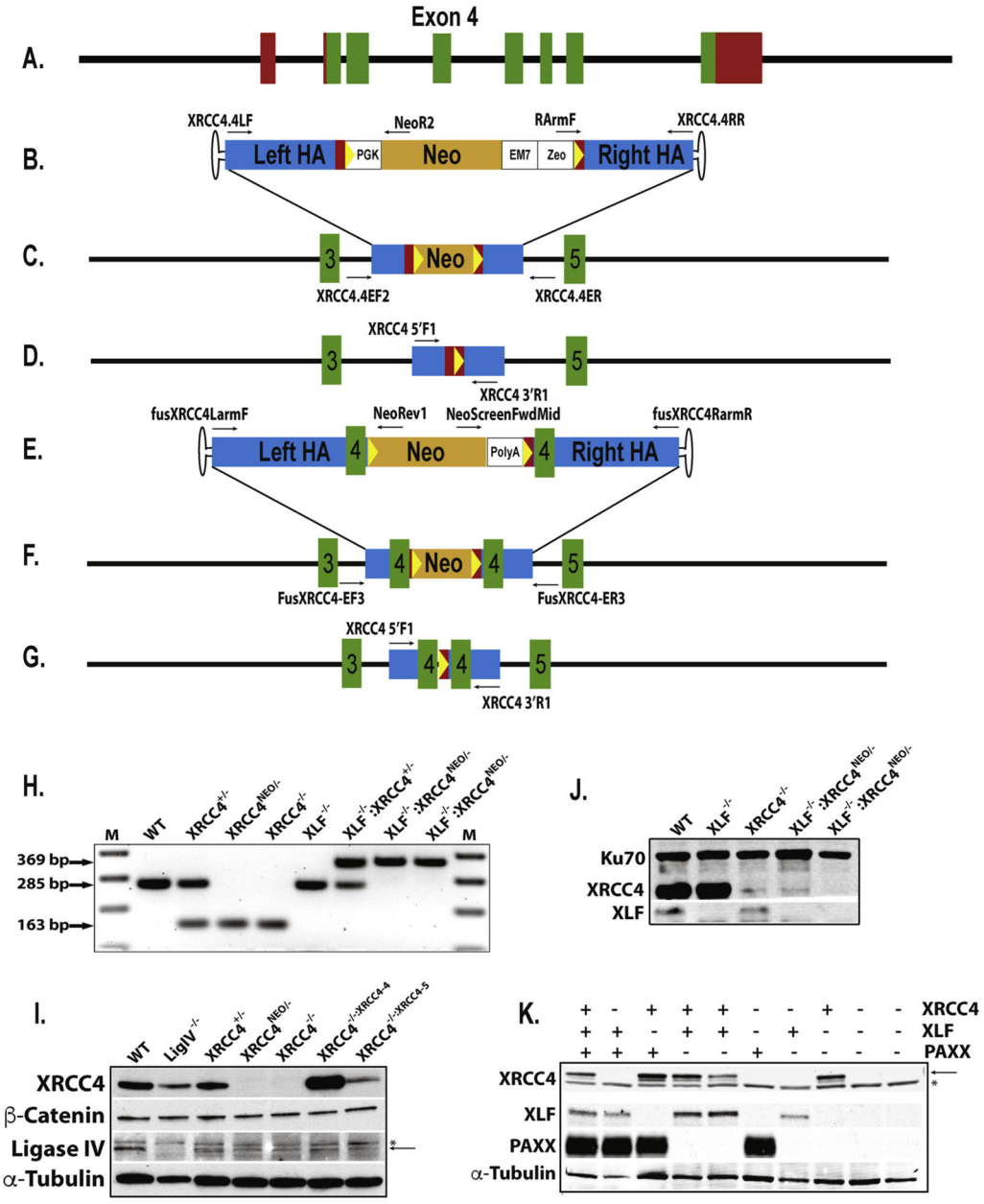


Fig. 1.

Generation of XRCC4 +/−, XRCC4−/−, and XLF−/−:XRCC4−/− HCT116 cell lines.(A) A diagram representing a partial XRCC4 genomic locus in the HCT116 cell line. Exons are shown (not to scale) as red (5’ and 3’ untranslated regions) or green (coding sequence) colored rectangles. (B) A diagram of the XRCC4 exon 4 Neo rAAV targeting vector. The left and right homology arms (Left HA and Right HA, respectively) used for facilitating HDR between the targeting vector and the genomic locus are shown as blue rectangles. Open and red rectangles depict vector sequences, while LoxP sites are shown as yellow triangles. The orange rectangle represents the Neo gene. The direction and approximate location of vector-specific PCR primers are shown as horizontal arrows. (C) A diagram of a first-round targeted allele using the rAAV-XRCC4 exon 4 Neo targeting vector. The

targeting vector has replaced exon 4 on one chromosome. All symbols are the same as in (A and B). Chromosome-specific primers used for determining targeting events are shown as horizontal arrows. (D) A diagram of a XRCC4 exon 4 first-round targeted locus following Cre-mediated LoxP recombination. PCR primers flanking the single LoxP site are shown as horizontal arrows. (E) A diagram of the rAAV-XRCC4 exon 4 fusion Neo targeting vector. All symbols are the same as (A, B and C). (F) A diagram of a first-round targeted allele using rAAV-XRCC4 exon 4 fusion Neo targeting vector. The targeting vector has integrated an in-frame (LoxP site-Neo-PolyA-LoxP) cassette within exon 4 on one chromosome. All symbols are the same as in (A and E). (G) A diagram of a XRCC4 exon 4 fusion first-round targeted allele following Cre-mediated LoxP recombination. All symbols are the same as in (A, E and F). (H) PCR confirmation of XRCC4+/−, XRCC4−/−, and XLF−/−:XRCC4−/− cell lines. A final confirmation PCR was performed using genomic DNA isolated from XRCC4+/−, XRCC4−/Neo, XRCC4−/−, XLF−/− :XRCC4 +/−, XLF−/−:XRCC4Neo/−, and XLF−/−:XRCC4Neo/− clones. Genomic DNA from WT and XLF−/− HCT116 cells was used as controls, lanes 1 and 5 respectively. (I) XRCC4−/− cells lack detectable XRCC4 protein. Whole cell extracts were prepared from parental (WT), LigIV−/−, XRCC4+/−, XRCC4Neo/−, XRCC4−/−, and two XRCC4−/− complemented clones (#4 and #5). The extracts were analyzed by immunoblot for XRCC4 as well as LigIV protein with b-catenin and a-tubulin as loading controls. In the LIGIV panel, a non-specific band is marked with an asterisk (*) and the correct band is marked with an arrow. (J) XLF−/− and XRCC4−/− cells lack detectable XLF and XRCC4 protein. Whole cell extracts were prepared from parental (WT), XLF−/−, XRCC4−/−, as well as two XLF−/−:XRCC4Neo/− cell lines. The extracts were analyzed by immunoblot for XRCC4 expression using Ku70 as a loading control and then subsequently probed with an XLF antibody. (K) PAXX−/− cells lack detectable PAXX protein. Whole cell extracts were prepared from parental (WT), XRCC4−/−, XLF−/−, two PAXX−/−, XRCC4−/−:XLF−/−, XRCC4−/−:PAXX−/−, XLF−/−:PAXX−/−, as well as two XRCC4−/−:XLF−/−:PAXX−/− cell lines. The extracts were analyzed by immunoblotting for XRCC4, XLF and PAXX using a-tubulin as a loading control. In the XRCC4 blot, a non-specific band is marked with an asterisk (*) and the correct band is marked with an arrowhead.

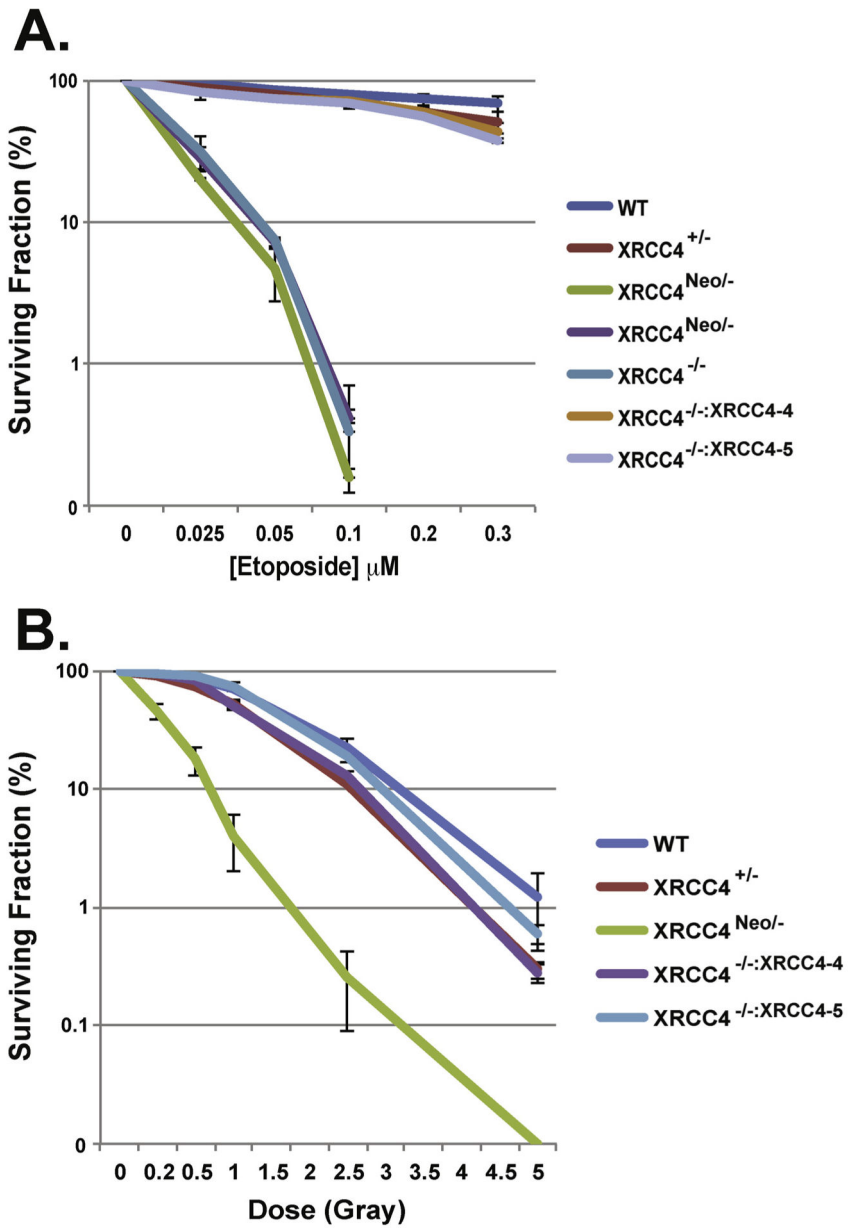


Fig. 2.

XRCC4-deficient HCT116 cells are sensitive to DNA-damaging agents.(A) Etoposide sensitivity. For each cell line, 3000 cells were plated in duplicate and exposed to the indicated levels of etoposide. Surviving cells that grew into colonies of at least 50 cells after ~10 days were scored. The averages (+/− standard deviations) of two experiments are shown. (B) IR sensitivity. For each cell line, either 300, in the case of the WT and XRCC4+/− cells, or 3000, for the XRCC4−/− and XRCC4−/− complemented cells, were plated in duplicate and X irradiated at the indicated doses. Surviving cells that grew into colonies of at least 50 cells after ~10 days were scored. The averages (+/− standard deviations) of two experiments are shown.

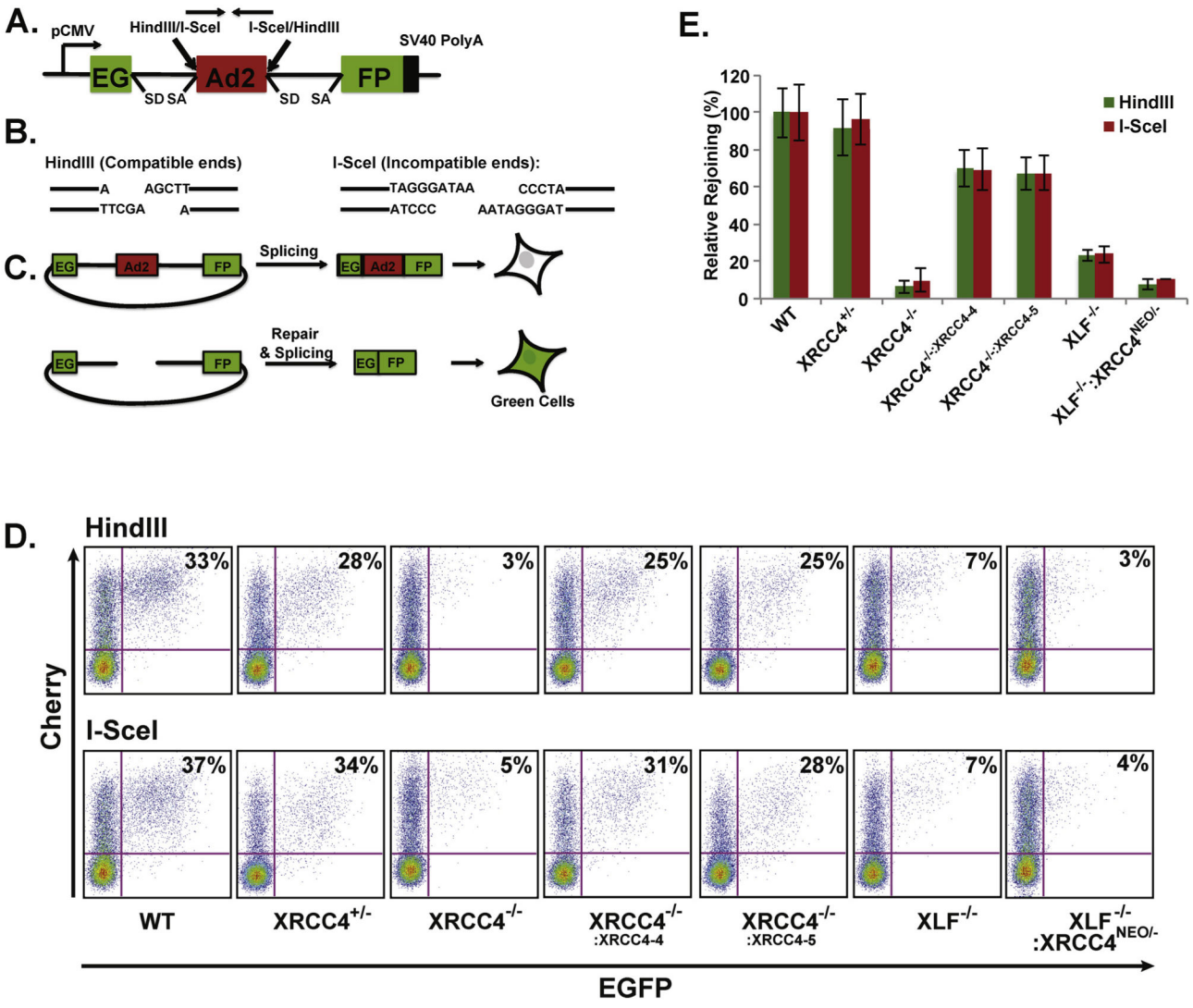


Fig. 3.

NHEJ activity in XRCC4−/−, XRCC4−/− complemented, and XLF−/−:XRCC4Neo/− cells. (A) A diagram of the pEGFP-Pem1-Ad2 NHEJ reporter substrate plasmid. The vector contains an EGFP expression cassette that is driven by a CMV promoter and terminated by the SV40 polyA sequence. The 5' (G) portion of GFP is separated from the 3' (FP) portion by a 2.4 kb long Pem1 intron interrupted by an Ad2 exon that is flanked by HindIII and I-SceI restriction enzyme recognition sites. Both splice donor (SD) and splice acceptor (SA) sites are also noted. (B) Restriction enzyme recognition sites used for the introduction of DSBs. Digestion of the two HindIII sites produces compatible cohesive ends while digestion of the inverted 18 bp non-palindromic I-SceI sites produces incompatible ends. (C) The Ad2-exon, within the Pem1 intron, is efficiently spliced into the middle of the GFP ORF resulting in an inactivated GFP gene product that leads to GFP-negative cells. HindIII/I-SceI restriction enzyme recognition sites flank each side of the Ad2 exon. The removal of the Ad2 exon by either of these endonucleases results in a linearized plasmid that, upon successful intracellular recircularization, leads to GFP expression that can be quantitated by

flow cytometry. (D) The loss of XRCC4, XLF, as well as both XRCC4 and XLF expression reduces end-joining. The indicated cell lines were transiently co-transfected with HindIII (top panels) or I-SceI (lower panels) linearized pEGFP-Pem1-Ad2 plasmid along with a pCherry plasmid. The cells were analyzed by FACS analysis 24 hr post transfection. The total number of both EGFP- and pCherry-positive (upper right-hand quadrant) cells versus all pCherry-positive (upper left and right-hand quadrant) cells was determined. The resulting percentage for each cell line is shown in the upper right-hand quadrant of each plot. (E) Quantitation of the results shown in (D), plotted as relative plasmid rejoining for HindIII (green bars) or I-SceI (red bars). The graph shown is the average (+/− the standard deviation) from five (WT, XRCC4+/−, XRCC4−/−, XRCC4−/− complemented) or two (XLF−/− and XLF−/−:XRCC4−/−) independent experiments.

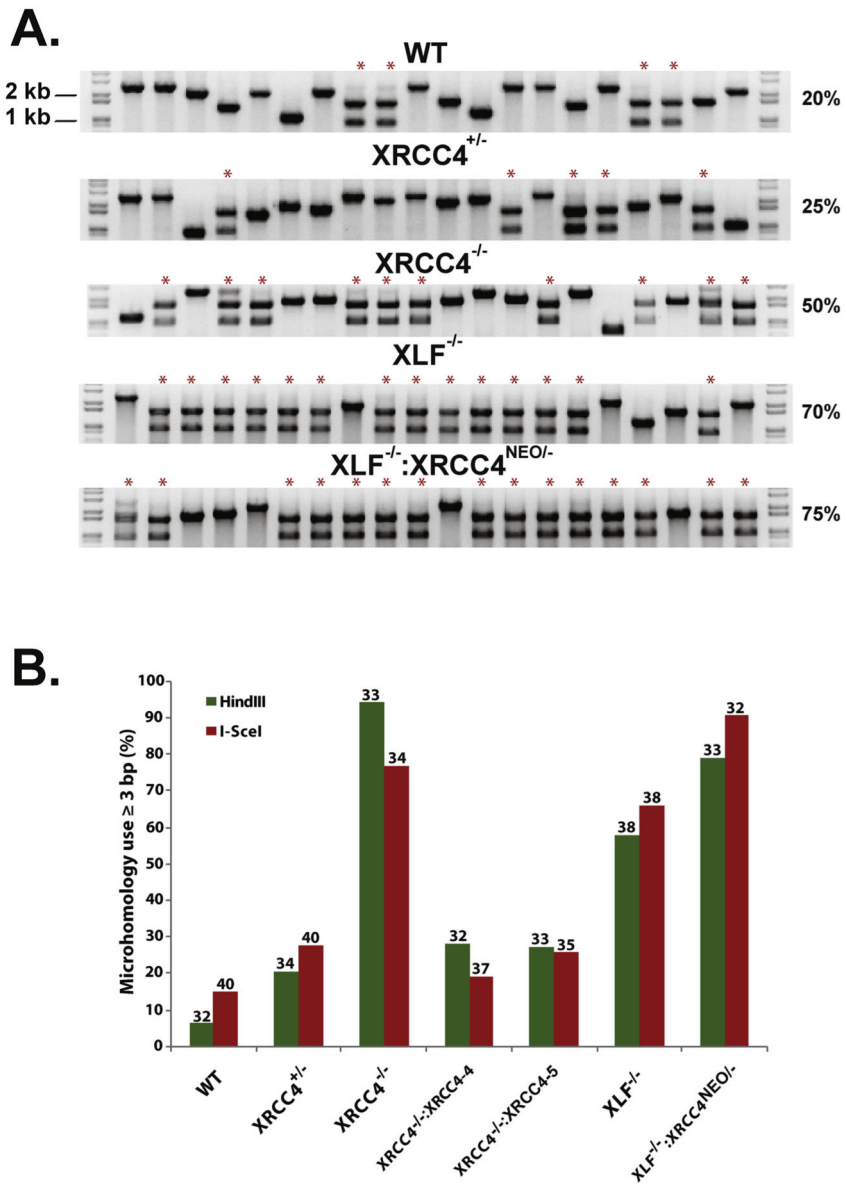


Fig. 4.

XRCC4−/−, XLF−/−, and XLF−/−:XRCC4−/− cells exhibit an increase in microhomology-mediated DNA repair.(A) Repaired HindIII-linearized plasmids were recovered from the indicated cell lines and propagated in E. coli for the purpose of determining the total ("perfect rejoining") frequency. For each cell line, twenty colony PCR products produced using GFP-specific primers were then subjected to HindIII restriction enzyme digestion and subjected to gel electrophoresis. Red asterisks indicate those PCR products where perfect joining occurred. (B) XRCC4−/−, XLF−/−, and XLF−/−:XRCC4−/− cells exhibit elevated use of microhomology-mediated repair. An additional ~30 plasmids that had not undergone perfect rejoining were analyzed by DNA sequencing. Those results (compiled from Tables S1 and S2) are displayed here with the percent microhomology used (defined by at least 3 nt of microhomology) for during the repair event. The total number of events (n) analyzed for each sample is indicated by the number above the bars.

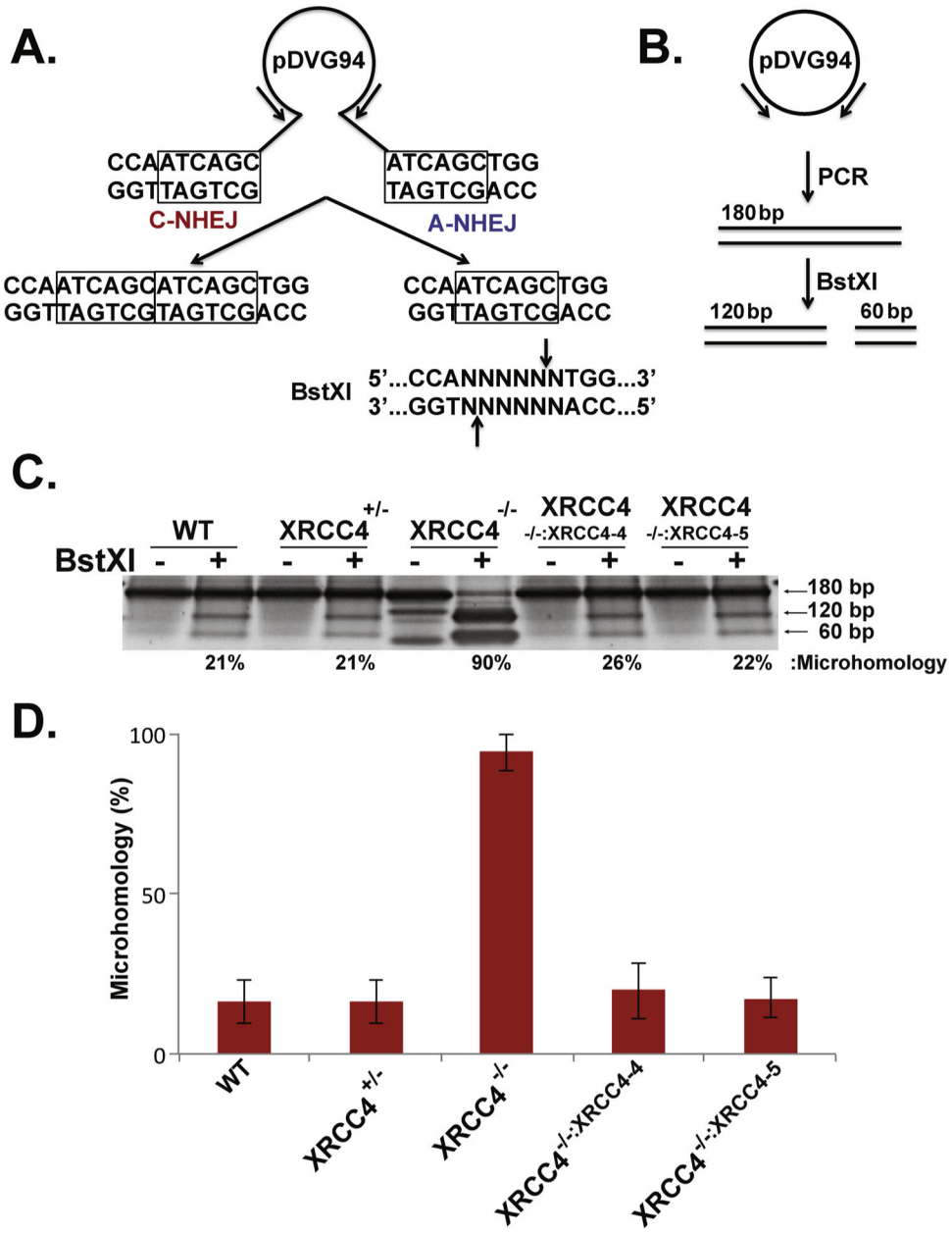
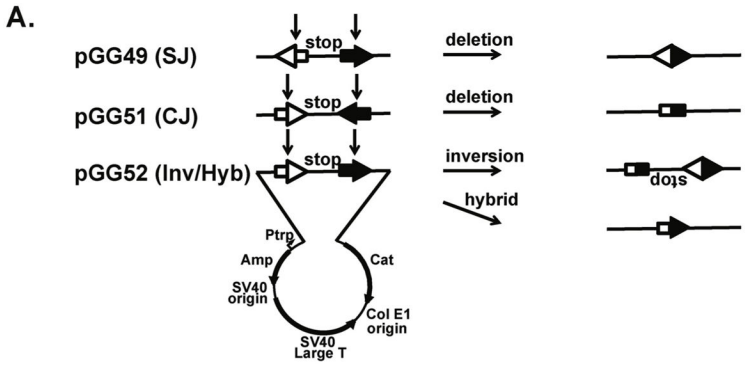


Fig. 5.

Confirmation of microhomology-mediated end-joining in XRCC4-deficient cells.(A) A microhomology-directed NHEJ (aEJ) biased reporter substrate plasmid has been engineered so that digestion with Eco47III and EcoRV yields a blunt-ended linear substrate with 6-bp repeats (boxes) at each end. Repair of this plasmid by C-NHEJ joining will result in the retention of both repeats while aEJ should yield a single repeat, which conforms to a BstXI restriction enzyme recognition site. (B) Experimental strategy for analysis of repair events produced in transfected cells. The recovered plasmids are subjected to PCR amplification using primers flanking the repair junction. These ~ 180 bp PCR products are then subjected to BstXI restriction enzyme digestion that will yield 120 bp and 60 bp products if successfully cleaved. The restriction fragments are visualized with SYBR Gold following

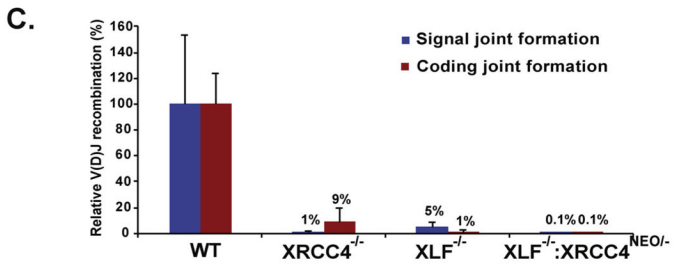
polyacrylamide gel electrophoresis. (C) The results of one such experiment. (D) Two independent experiments including the one shown in (C) were quantitated using ImageJ software and averaged.



B.

Cell Lines	pGG49 (SJ Substrate) DAC/DA	SJ (%)	Relative eff. (%)	pGG51 (CJ Substrate) DAC/DA	CJ (%)	Relative eff. (%)	pGG52 (Inversion/Hybrid Substrate) DAC/DA	Inv/Hyb (%)	Relative eff. (%)
WT HCT116	46/83020	0.06		200/78960	0.25		45/96420	0.047	
	290/226000	0.13		198/131750	0.15		77/227500	0.034	
	806/339830	0.24	100	1057/440665	0.24	100	230/474165	0.049	100
	528/194000	0.27		753/287000	0.26		114/294500	0.039	
	148/113500	0.13		303/179000	0.17		75/1940000	0.039	
XRCC4$^{-/-}$	2/44420	0.0045		14/84760	0.017		2/53620	0.004	
	3/131500	0.0023	1	58/84250	0.069	9	13/60000	0.022	13
	4/280330	0.0014		11/237630	0.005		14/410500	0.003	
XLF$^{-/-}$	39/464500	0.0084		7/318000	0.002		0/489850	0	
	3/59000	0.0051	5	0/81500	0.000	1	0/142000	0	0
	11/44550	0.0247		0/36750	0.000		0/87900	0	
XLF$^{-/-}$ XRCC4$^{-/-}$	1/230500	0.0004		1/364900	0.0003		0/150600	0	
	0/95750	0.0000	0.1	0/91500	0.0000	0.1	0/97000	0	0
	0/161250	0.0000		0/164100	0.0000		0/261450	0	

DAC, numbers of ampicillin-chloramphenicol-resistant E. coli transformants after DpnI treatment; DA, numbers of ampicillin-resistant tranformants after DpnI treatment. Transient V(D)J recombination assays were done in the presence of RAG1 and RAG2 expression vectors.



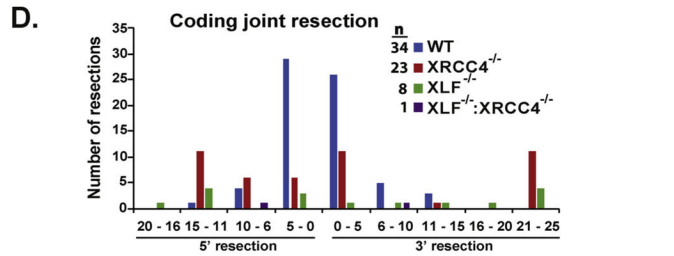


Fig. 6.

HCT116 cells lacking XRCC4, XLF, and both XRCC4 and XLF exhibit greatly reduced levels of V(D)J recombination.(A) A human V(D)J recombination substrate plasmid diagram extracted from Gauss and Lieber [46]. Three different plasmids are diagrammed. Each plasmid is a dual drug selection vector that includes an Amp (ampicillin resistance) gene as well as a Ptrp (prokaryotic tryptophan promoter) driving expression of a promoter-less chloramphenicol (Cat) resistance gene. Importantly, a phage transcription terminator is represented by “stop” in the three cassettes. This terminator is flanked by 12- and 23-bp V(D)J RSSs arranged in different orientations for each vector (open and closed arrowheads, respectively). Transfection of each V(D)J substrate plasmid along with RAG1 and RAG2 expression plasmids into human cells and subsequent V(D)J recombination will result in the

deletion or inversion of the terminator sequence. Plasmids are then recovered from the human cells and propagated in bacteria, where only recombined plasmids confer resistance to both chloramphenicol and ampicillin while nonrecombined plasmids confer resistance only to ampicillin. (B) V(D)J recombination activity table. The results of five separate experiments for parental (WT) HCT116 as well as three separate experiments for XRCC4-, XLF-, and XLF:XRCC4-deficient cells are shown. (C) The relative V(D)J recombination frequencies for each cell line were determined from the average of these three individual experiments. (D) Incidence of nucleotide deletion from the coding ends. Coding joints (compiled from Fig. S3) were examined for the number of bases lost from each coding end in both directions and binned into the indicated 5 nt intervals. The total number of events (n) analyzed for each genotype is indicated.

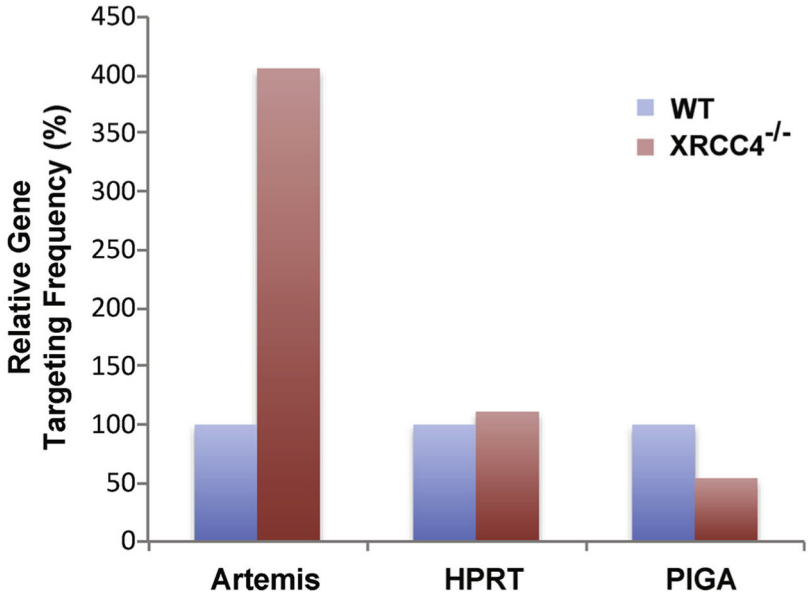


Fig. 7.

rAAV-mediated gene targeting frequency in XRCC4-deficient cells.A graph displaying the relative rAAV-mediated gene targeting frequencies for each locus observed for WT and XRCC4−/− cells. For Artemis and HPRT, the data are from single experiments where 81 to 116 drug-resistant colonies were analyzed for each cell line and for each gene. For PIGA, the data represent the average of two independent cell-sorting experiments that were combined together.