

# A Genetic Screen Identifies FAN1, a Fanconi Anemia-Associated Nuclease Necessary for DNA Interstrand Crosslink Repair

Agata Smogorzewska,<sup>1,2,3,\*</sup> Rohini Desetty,<sup>3</sup> Takamune T. Saito,<sup>4</sup> Michael Schlabach,<sup>1</sup> Francis P. Lach,<sup>3</sup> Mathew E. Sowa,<sup>5</sup> Alan B. Clark,<sup>6</sup> Thomas A. Kunkel,<sup>6</sup> J. Wade Harper,<sup>5</sup> Monica P. Colaiácovo,<sup>4</sup> and Stephen J. Elledge<sup>1,\*</sup>

<sup>1</sup>Howard Hughes Medical Institute, Department of Genetics, Harvard Medical School, Department of Medicine, Division of Genetics, Brigham and Women's Hospital, Boston, MA 02115, USA

<sup>2</sup>Department of Pathology, Massachusetts General Hospital, Boston MA 02114, USA

<sup>3</sup>Laboratory of Genome Maintenance, The Rockefeller University, New York, NY 10065, USA

<sup>4</sup>Department of Genetics

<sup>5</sup>Department of Pathology

Harvard Medical School, Boston MA 02115, USA

<sup>6</sup>Laboratory of Molecular Genetics and Laboratory of Structural Biology, National Institute of Environmental Health Sciences, National Institutes of Health, DHHS, Research Triangle Park, NC 27709, USA

\*Correspondence: [asmogorzewska@rockefeller.edu](mailto:asmogorzewska@rockefeller.edu) (A.S.), [selledge@genetics.med.harvard.edu](mailto:selledge@genetics.med.harvard.edu) (S.J.E.)

DOI 10.1016/j.molcel.2010.06.023

## SUMMARY

The Fanconi anemia (FA) pathway is responsible for interstrand crosslink repair. At the heart of this pathway is the FANCI-FANCD2 (ID) complex, which, upon ubiquitination by the FA core complex, travels to sites of damage to coordinate repair that includes nucleolytic modification of the DNA surrounding the lesion and translesion synthesis. How the ID complex regulates these events is unknown. Here we describe a shRNA screen that led to the identification of two nucleases necessary for crosslink repair, FAN1 (KIAA1018) and EXDL2. FAN1 colocalizes at sites of DNA damage with the ID complex in a manner dependent on FAN1's ubiquitin-binding domain (UBZ), the ID complex, and monoubiquitination of FANCD2. FAN1 possesses intrinsic 5'-3' exonuclease activity and endonuclease activity that cleaves nicked and branched structures. We propose that FAN1 is a repair nuclease that is recruited to sites of crosslink damage in part through binding the ubiquitinated ID complex through its UBZ domain.

## INTRODUCTION

Cells in all organisms experience massive amounts of spontaneous DNA damage each day. A failure to properly respond to this genotoxic stress can lead to both developmental abnormalities and tumorigenesis. Organisms have evolved a complex signal transduction pathway called the DNA damage response (DDR) that senses genotoxic stress and orchestrates a response by activating specific types of repair, arresting the cell cycle and altering transcription. At the core of this signal transduction pathway are two PI-3 kinase-like protein kinases, ATM and

ATR (Bakkenist and Kastan, 2004; Bartek et al., 2004; Harper and Elledge, 2007), which support the damage-induced phosphorylation of hundreds of substrates to coordinate DNA repair (Matsuoka et al., 2007; Stokes et al., 2007).

A life-threatening lesion is the DNA double-strand crosslink, which covalently connects the Watson and Crick strands of DNA to create a bidirectional polymerase block. A repair pathway known as the Fanconi anemia (FA) pathway has evolved to specifically deal with these types of lesions. FA is a recessive developmental and cancer predisposition syndrome whose patients display multiorgan defects, bone marrow failure in childhood (Fanconi, 1967; Schmid and Fanconi, 1978), and a high incidence of malignancies (Alter et al., 2003). Cells from FA patients are hypersensitive to DNA interstrand crosslinking (ICL) agents such as mitomycin C (MMC) (Auerbach and Wolman, 1976). To date, 13 proteins have been implicated in FA. At the center of this pathway is the FANCI/FANCD2 (ID) complex, which loads onto sites of crosslinks to direct DNA repair. The ID complex is chromatin bound, and when it encounters a DNA replication structure stalled due to a DNA crosslink, it becomes phosphorylated by the ATR/ATRIP kinase, which is localized through recognition of RPA at the lesion (Zou and Elledge, 2003). Phosphorylation of both I and D2 is required for ID function (Andreasen et al., 2004; Ho et al., 2006; Ishiai et al., 2008) and leads to the monoubiquitination of both subunits by a multisubunit E3 ligase formed by eight FA proteins (FANCA/B/C/E/F/G/L/M) and the E2-conjugating enzyme UBE2T (Cole et al., 2010; Machida et al., 2006; Meetei et al., 2004). Ubiquitinated ID then accumulates at the damage site and directs repair (Smogorzewska et al., 2007; Knipscheer et al., 2009; Raschle et al., 2008).

The repair of a crosslink is thought to involve two incision events on a single DNA strand flanking the lesion, followed by bypass synthesis over the lesion on the remaining intact strand using a translesion polymerase, possibly Rev1 (Niedzwiedz et al., 2004; Simpson and Sale, 2003) in combination with Rev3 and Rev7 (Lehmann et al., 2007; Raschle et al., 2008). A recently identified A family nuclear DNA polymerase, PolN, might also

play a role in this step (Moldovan et al., 2009; Zietlow et al., 2009). After bypass synthesis, two more incision events flanking the lesion occur, allowing it to be excised. The initially cleaved strand can then be repaired by gene conversion using homologous recombination (HR). Both the incision step and the bypass polymerase step are dependent upon ubiquitination of the ID complex (Knipscheer et al., 2009; Raschle et al., 2008). The nucleases responsible for the incision and excision events are not precisely known, although XPF-ERCC1 and Mus81 complexes have been implicated (Ciccica et al., 2008). Recently, SLX4, a scaffold for various DNA repair nucleases, has been identified to be necessary for resistance to crosslinking agents. SLX4 interacts with both Mus81 and XPF and together with SLX1 forms a Holliday junction resolvase, although it is unclear which SLX4 activity is responsible for conferring resistance to DNA crosslinks (Fekairi et al., 2009; Munoz et al., 2009; Saito et al., 2009; Svendsen et al., 2009). After bypass synthesis, the two strands liberated by the first two incision events constitute a double-strand break and are repaired by HR. Since HR requires 3' overhangs to initiate recombination, it is likely these strands are processed by a 5'-3' exonuclease to generate the recombination substrate that can gap repair off the strand from which the lesion was removed.

A number of outstanding issues remain unresolved in the FA pathway. First and foremost is how the ID complex orchestrates repair and the role of ubiquitination. Does it play a structural role in simply maintaining the complex at sites of damage, or does it recruit repair factors? In addition, the identities of the enzymatic factors involved in manipulation of the DNA at the site of the lesion are still unknown. Furthermore, it is unknown if there are additional backup pathways that can also function to repair crosslinks in the absence of FA. The FA pathway appears to be conserved in metazoans, but is missing from prokaryotes and unicellular eukaryotes such as *S. cerevisiae* and *S. pombe*. Thus, alternative repair pathways do exist in nature. To address these issues and identify new genes involved in crosslink repair, we performed an RNAi screen for genes required for resistance to the crosslinking agent MMC and discovered two nucleases required for crosslink repair, one of which, FAN1, possesses both endonuclease and exonuclease activities and is recruited to lesions by the monoubiquitinated ID complex.

## RESULTS

### Genome-wide shRNA Screen to Identify Proteins Necessary for Resistance to Interstrand Crosslink Damage

To identify proteins involved in resistance to crosslinking agents, we screened U2OS cells transduced with a library of 74,905 retroviral shRNAs targeting 32,293 unique human transcripts (Figure 1A). Using competitive hybridization of probes that detect the half-hairpins (HHs) of the shRNAs, we compared the relative abundance of shRNAs between untreated cells and cells treated with low levels of MMC. Cells bearing shRNAs that conferred sensitivity were depleted from the treated population. About 2173 hairpins targeting 2017 genes conferred sensitivity to MMC using the criteria of an average loss of 2-fold from

the treated population ( $\log_2 > 1$ ) (see Table S1 available online). Among these were previously known DDR proteins including BRCA1, TOPBP1, RAD18, RAD17, RAD51, RAD54, FANCE, and others. We employed the multicolor competition assay (MCA) (Smogorzewska et al., 2007) with 379 shRNAs against a selected group of genes that made the cutoff in the primary screen to retest for MMC sensitivity. Eighty-four shRNAs tested conferred MMC sensitivity (Table S2). Figure 1B shows the top 38 scoring hairpins in which resistance to MMC was below 80% of control shRNA. To further examine damage sensitivity, pools of siRNAs were also tested (Figure 1C). Based on the results of this assay and domain analysis, two genes, EXDL2 and KIAA1018, were chosen for further study.

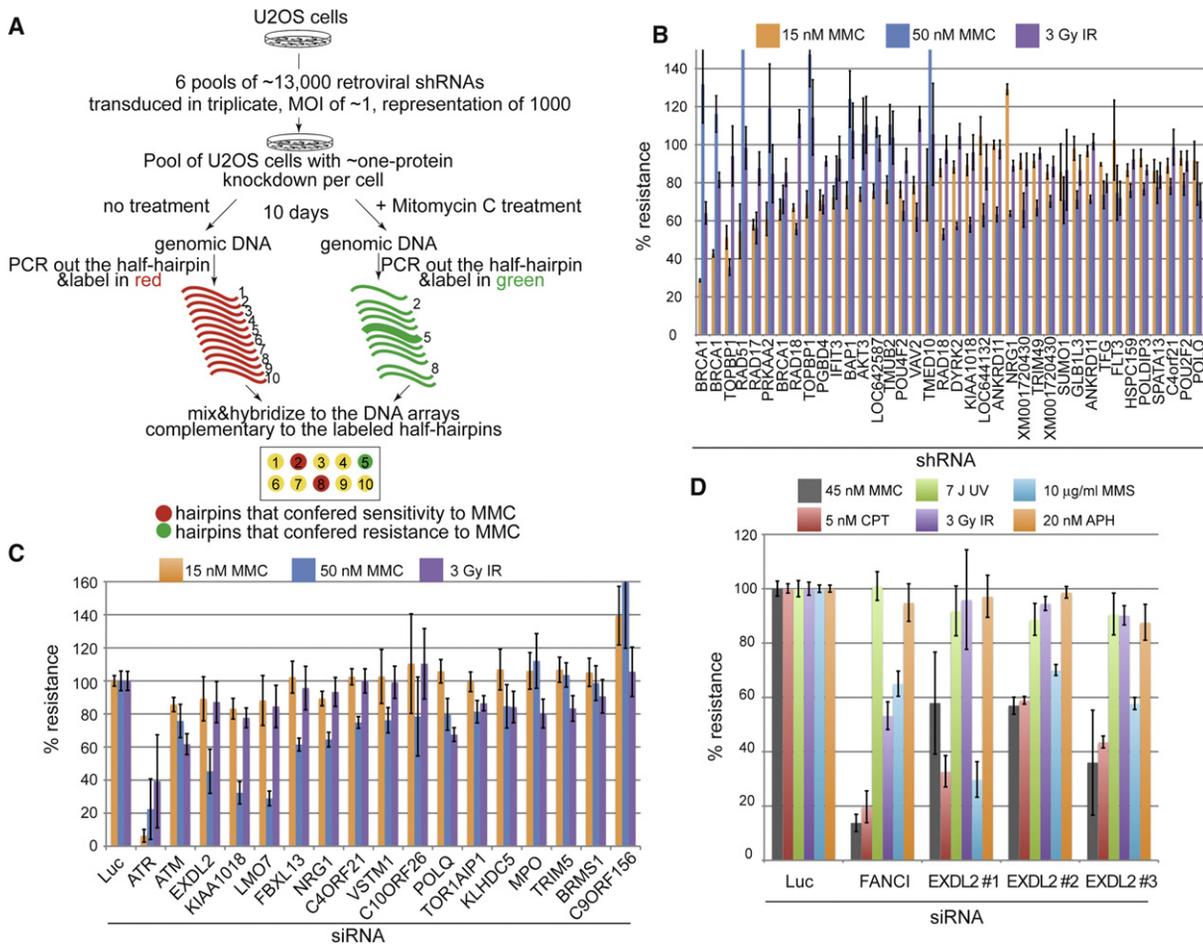
### EXDL2—A Putative 3'-5' Exonuclease Necessary for Resistance to MMC

EXDL2 is an uncharacterized 621 amino acid protein with a nuclease domain most similar to the WRN-exo domain in the WRN protein (Figures S1A and S1B) (Perry et al., 2006). Based on the conservation of the four key negatively charged residues (DEDD) that serve as ligands for the metal ions, as well as a tyrosine residue that has been shown to be important for the catalysis, EXDL2 is predicted to be a 3'-5' exonuclease. A mutation in a *D. melanogaster* ortholog of EXDL2, CG6744, displays a phenotype of hyperrecombination (Cox et al., 2007). In human cells, depletion of EXDL2 using three different siRNAs led to sensitivity to MMC, the Topo1 inhibitor camptothecin (CPT), and the alkylating agent MMS (Figure 1D and Figure S1C). This spectrum of sensitivities is similar to mutants in the FA pathway. Therefore, we tested if FANCD2 ubiquitination was affected in the EXDL2-depleted cells. Based on the normal ubiquitination of FANCD2 before and after damage (Figure S1D), we conclude that EXDL2 is either downstream of FANCD2 in the Fanconi pathway or in a parallel pathway of crosslink repair.

### FAN1 (KIAA1018) Is Required for the Resistance to Crosslinking Agents

A second protein with an interesting domain structure is KIAA1018, which we renamed FAN1 (Fanconi-associated nuclease 1) based on the data presented below. FAN1 has a nuclease-like fold called DUF994 (later renamed VRR-NUC) at its C terminus, potential DNA-binding (SAP) and protein-protein interaction (TPR) motifs in its midsection, and a Rad18-like ZnF domain at the N terminus (Kinch et al., 2005) (Figure 4A). Depletion of FAN1 with multiple siRNAs leads to sensitivity to crosslinking agents including MMC, chlorambucil, carboplatin, and oxaliplatin, as well as CPT and MMS (Figures 2A and 2B).

To examine the evolutionary conservation of FAN1's role in crosslink repair, we examined the crosslink sensitivity of a *C. elegans* FAN1 mutant (*tm423*) that carries a deletion of the SAP domain (Figure S2). *fan-1* mutants lay normal numbers of eggs and show normal larval development as well as no increase in either embryonic lethality or the percent of males among their progeny that would suggest a meiotic phenotype (Figure 2D). However, treatment with either MMC or nitrogen mustard (HN2) results in decreased embryonic viability compared to wild-type as judged by decreased hatching ( $p < 0.0001$ ,



**Figure 1. Whole-Genome shRNA Screen to Identify Genes Necessary for Crosslink Resistance**

(A) Schematic of the primary screen. Changes in hairpin abundance after transduction with the shRNA library and MMC treatment were followed by competitive DNA array hybridization.

(B) MCA in U2OS cells transduced with the indicated shRNAs. Of the tested hairpins, only those that showed less than 80% resistance to 15 or 50 nM MMC treatment are shown. Resistance of cells transduced with a hairpin against luciferase was set at 100% in all MCA experiments. All hairpins used in MCA are shown in Table S2.

(C) MCA in U2OS cells transfected with the indicated siRNAs (pools of four siGENOME siRNAs from Dharmacon). Cells transfected with siRNAs against ATM and ATR were used as a control.

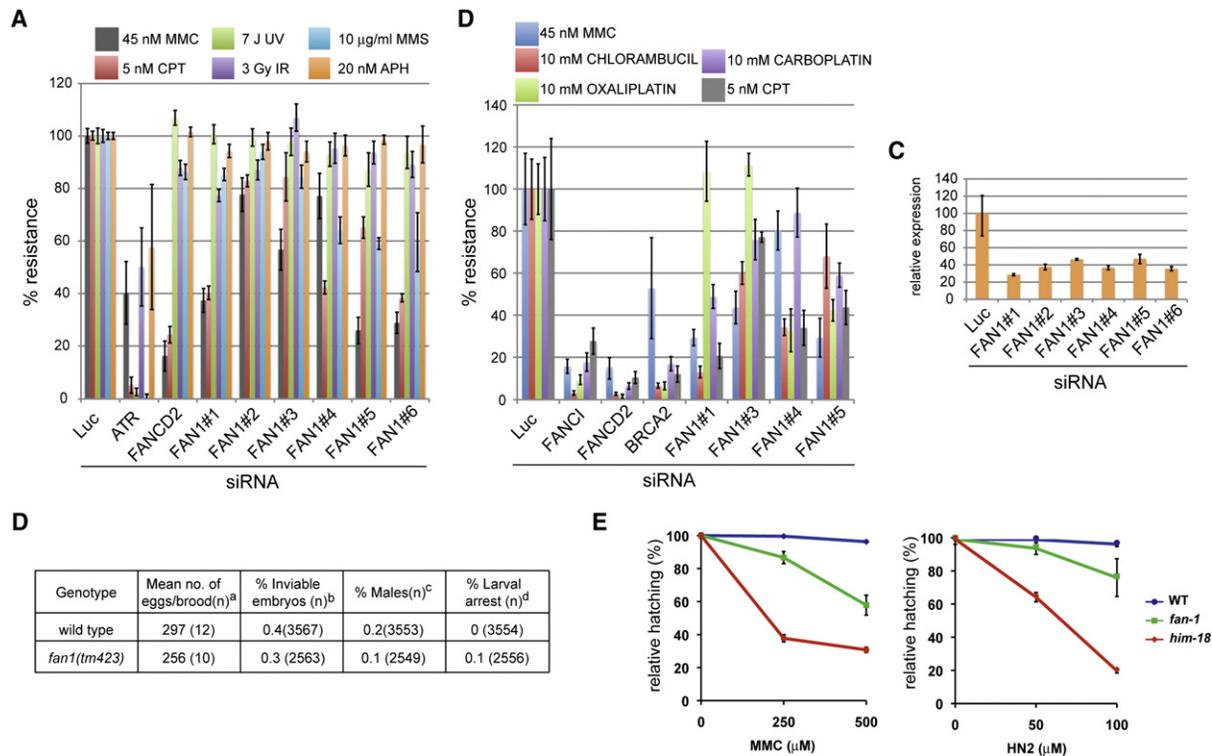
(D) MCA in U2OS cells transfected with three separate siRNAs against EXDL2. Cells transfected with siRNAs against FANCI were used as a control. Error bars represent standard deviation across three technical replicates.

respectively, chi-square test) (Figure 2E). Thus, FAN1 plays an evolutionarily conserved role in resistance to crosslink damage. Since the *tm423* mutant may be a hypomorph, a meiotic phenotype could emerge in null animals.

### FAN1 Associates with Mismatch Repair Proteins

To identify FAN1-associated proteins, a HA-tagged FAN1 was purified from 293TRES cells, and interacting proteins were identified by LC-MS/MS. Proteomic data were processed using the Comparative Proteomic Analysis Software Suite (CompPASS) (Sowa et al., 2009). The top-scoring interacting proteins were the mismatch repair proteins MLH1, MLH3, PMS1, and PMS2 (Figure 3A). KIAA1018 was previously detected in a proteomic analysis of MLH1-interacting proteins (Cannavo et al., 2007);

however, the interaction was not independently confirmed. To confirm these interactions, we performed immunoprecipitations. HA-FAN1 was identified in immunoprecipitates of MLH1, MLH3, and PMS2 (Figure 3B). Reciprocal immunoprecipitations identified all four proteins in precipitation reactions with anti-HA antibodies (Figure 3C). Endogenous FAN1 immunoprecipitations brought down both FAN1 and MLH1 in a FAN1-dependent manner (Figure 3D). The interaction with the mismatch repair machinery raised the possibility that FAN1's role in ICL resistance is dependent on mismatch repair. To test this, we depleted FAN1 from the HCT116 cells, which carries inactivating mutations in both alleles of MLH1 (Figure 3E). FAN1 depletion with four different siRNAs still increased sensitivity of these cells to crosslinking agents.



**Figure 2. Sensitivity to Crosslinking Agents and Camptotecin in Human Cells Depleted of FAN1 and in a FAN1 Mutant *C. elegans* Strain**

(A) MCA in U2OS cells transfected with six separate siRNAs against FAN1 and treated with different types of DNA damaging agents.

(B) MCA in U2OS cells transfected siRNAs against FAN1 and treated with different DNA crosslinking agents.

(C) RT-qPCR in U2OS cells transfected with the different siRNAs against FAN1.

(A–C) Error bars represent the standard deviation (SD) across three technical replicates.

(D) Plate phenotypes. Parentheses indicate the total number of: <sup>a</sup>singled hermaphrodites for which entire brood sizes were scored, <sup>b</sup>fertilized eggs scored, <sup>c</sup>adults scored, <sup>d</sup>L1–L4 worms. ND, not determined due to n = 0. Wild-type data are from Saito et al. (2009).

(E) Relative hatching of wild-type, *fan-1*, and *him-18* mutants after treatment with the indicated doses of MMC and nitrogen mustard (HNZ). Hatching is plotted as a fraction of the hatching observed in untreated animals. Error bars indicate standard error of the mean for at least 20 animals in each of three independent experiments.

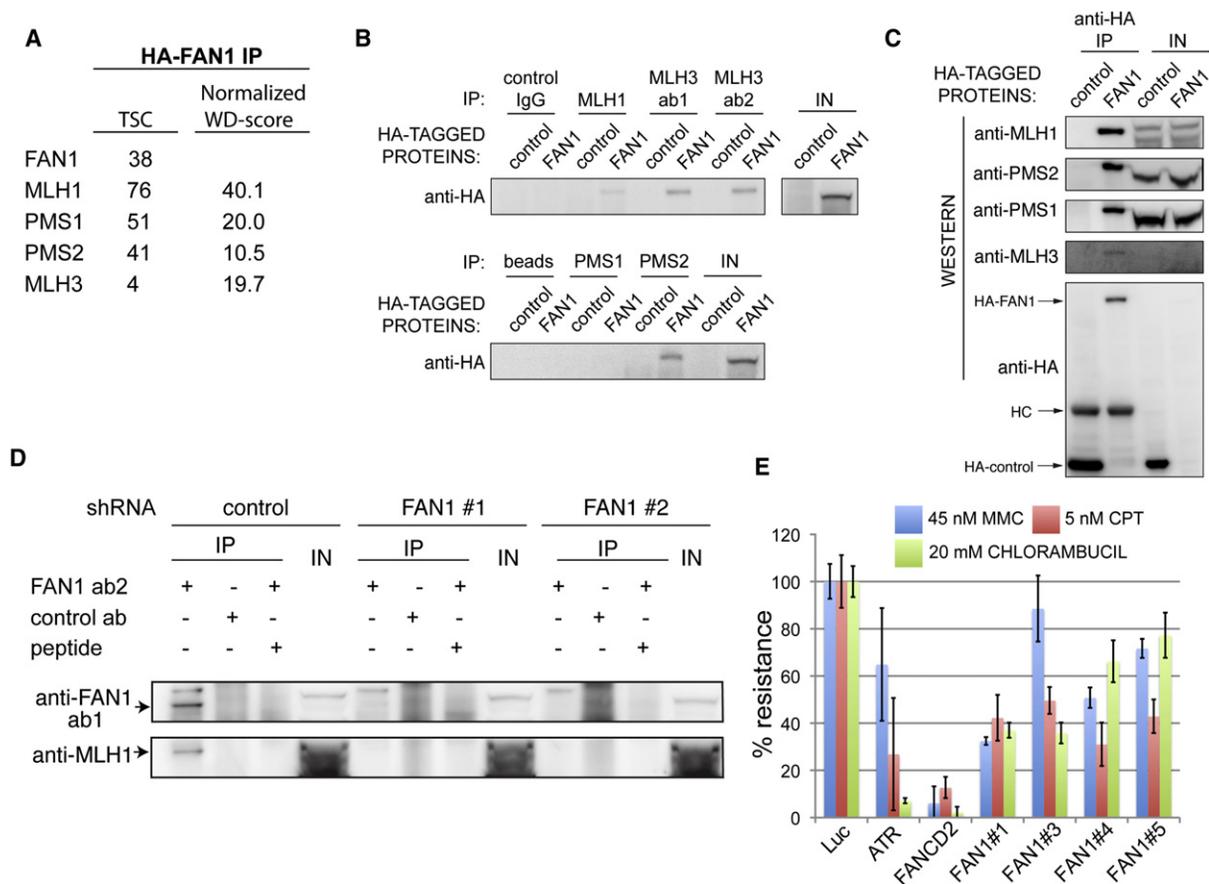
### FAN1 Colocalizes to Sites of Damage with FANCD2 via Its UBZ Domain

The domain structure of FAN1 with its well-conserved UBZ, SAP, and nuclease domains (Figures 4A and 4B) led us to test FAN1 localization to sites of DNA damage. We subjected GFP-tagged FAN1 cells to laser microirradiation, which results in localized DNA damage tracks (Bekker-Jensen et al., 2006). Within 15 min of microirradiation, GFP-FAN1 localized to sites of DNA damage along with  $\gamma$ -H2AX, a marker of DNA damage (Figure 4C) (Rogakou et al., 1999). The GFP tracks were also seen in the absence of  $\gamma$ -H2AX staining (data not shown). We next asked if FAN1 localized to sites of crosslinked DNA damage caused by MMC. Most MMC-induced foci contained both FAN1 and FANCD2 (Figure 4D). We next tested various FAN1 domain mutants for their ability to localize to sites of laser microirradiation (Figure 5A). Mutants in the nuclease domain behaved like wild-type. The N-terminal 373 amino acids lacking SAP or a protein with a SAP domain mutation (L477P) still localized to sites of damage, but the strength of the GFP-FAN1 signal was substantially diminished. Only the mutant with two conserved cysteine residues of the UBZ domain substituted by alanines (C44A,

C47A) completely failed to localize to microirradiation tracks (Figure 5B). The same UBZ mutant did not form foci in MMC-treated cells (Figure 5C) but was expressed at similar levels as the other alleles (Figure 5C, top panel “no triton”). Several other mutant alleles localized to sites of damage with reduced efficiency (Figure S3). Interestingly, the N-terminal 90 amino acids containing the UBZ domain were able to colocalize with FANCD2, although with diminished efficiency. Based on these experiments, we conclude that the UBZ domain is critical for localizing FAN1 to sites of damage. However, other parts of the protein also assist in localization. Careful kinetic analysis of the behavior of the different mutants will be necessary to fully understand the role of all domains in localization of FAN1 to sites of damage.

### FANCI and FANCD2 Are Required for FAN1 Localization to Sites of DNA Damage

To examine possible dependency on the ID complex, we assessed formation of GFP-FAN1 foci after MMC treatment in cells depleted of FANCD2 and FANCI. In FANCI- or FANCD2-depleted cells, FAN1 was no longer able to form foci (Figure 6A and Figure S4B). The percentage of cells with foci in cells



**Figure 3. FAN1 Associates with Mismatch Repair Proteins**

(A) HA-FAN1 was expressed in 293 TREX cells using an inducible retroviral system and cell extracts were subjected to immunoprecipitation and LC-MS/MS as described in the Experimental Procedures. High-confidence candidate-interacting proteins are shown as determined using CompPASS to derive normalized WD score for individual proteins in the immune complex. TSC, total spectral count. This experiment was done in duplicate, and average values are reported. In both experiments all listed proteins were identified.

(B) Cell extracts of 293 TREX cells expressing HA-FAN1 were subjected to immunoprecipitation using the indicated antibodies. HA-FAN1 was identified by immunoblotting with an anti-HA antibody. IN represents 5% input.

(C) Cell extracts of 293 TREX cells expressing HA-FAN1 were subjected to immunoprecipitation using an anti-HA antibody, and immunoprecipitates were probed with the indicated antibodies. HC, heavy chain of the antibody used in the immunoprecipitation.

(D) Endogenous FAN1 was immunoprecipitated ( $\pm$  antigenic peptide) from HeLa cells with or without FAN1 depletion using two separate shRNAs and immunoblotted for FAN1 and MLH1. shRNAs #1–739 and shRNA #2–600. Note that the FAN1 antibody does not recognize endogenous protein in straight western, only in IPs.

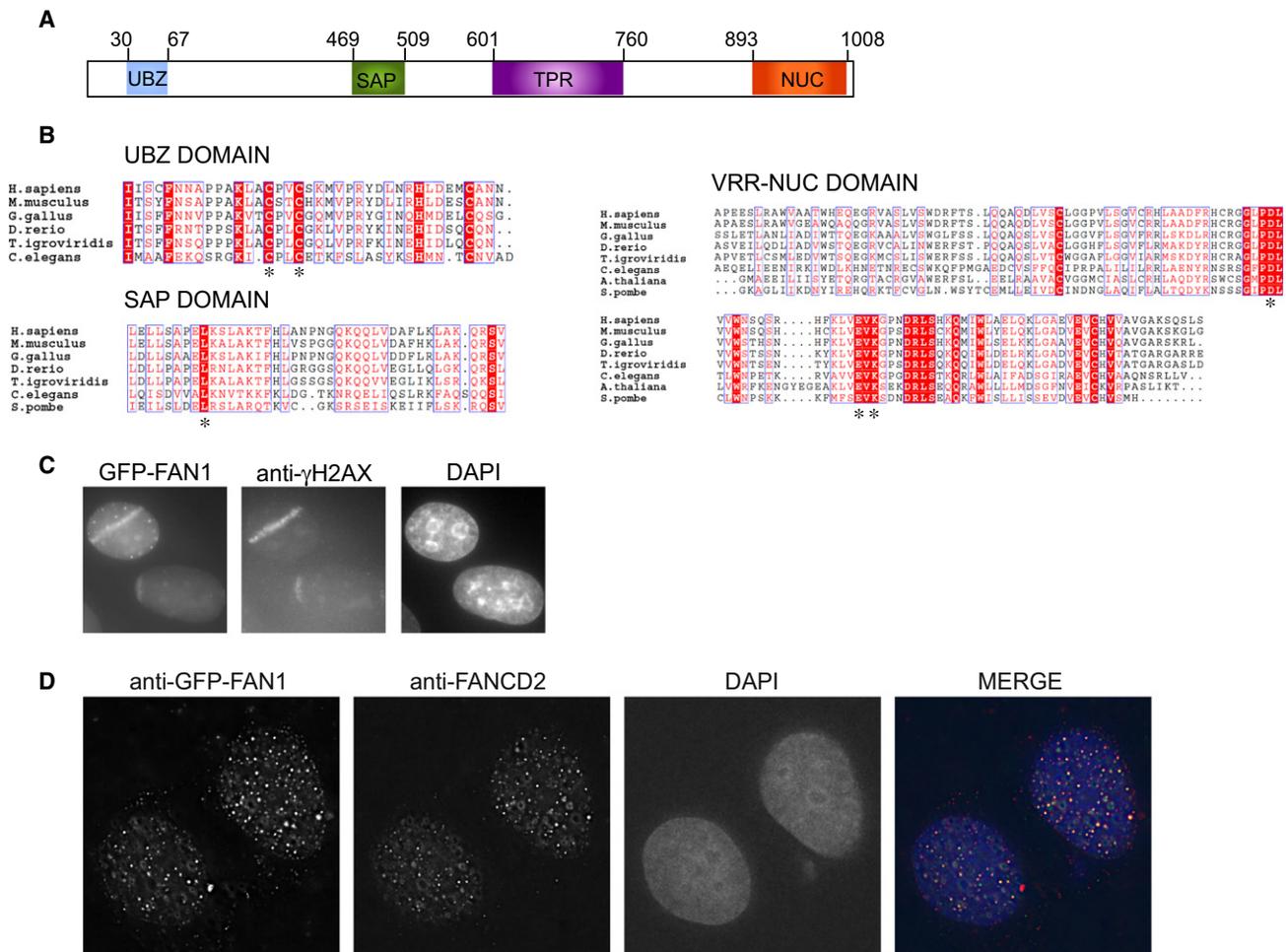
(E) MMC and camptotecin sensitivity caused by depletion of FAN1 are not dependent on MLH1 or MSH3. MCA in HCT116 cells transfected with the indicated siRNAs. Error bars represent the standard deviation (SD) across three technical replicates.

depleted of either FANCD1 or FANCD2 was 28% and 24%, respectively, while 93% of control depleted cells showed FAN1 foci. FAN1 also fails to form foci in PD20 cells, which lack FANCD2 protein (Figure 6B), but FAN1 foci reappear upon complementation with wild-type FANCD2. PD20 cells expressing the monoubiquitination-defective FANCD2 K561R mutant also fail to form FAN1 foci (Figure 6B). We next asked if FAN1 is required for FANCD2 monoubiquitination. Cells transfected with four different siRNAs against FAN1 were treated with MMC. FANCD2 in these cells was ubiquitinated to the same extent as in cells treated with siRNAs against luciferase (Figure 6C). We conclude that FA proteins and specifically FANCD2 monoubiquitination are

necessary for the recruitment of FAN1 to sites of damage and that the crosslink sensitivity of FA-defective cells is in part due to a failure to recruit FAN1 to sites of DNA damage.

### FAN1 Interacts with FANCD2

The dependence of FAN1 foci formation on the presence of monoubiquitinated FANCD2 raised a possibility that FANCD2 and FAN1 interact. The chromatin fraction from cells expressing GFP-FAN1 were subjected to immunoprecipitation with anti-FANCD2 antibodies. GFP-FAN1 strongly coimmunoprecipitated with FANCD2 (Figure 6D). Therefore, FAN1 and FANCD2 interact in vivo.



**Figure 4. FAN1 Is an Evolutionarily Conserved Protein that Localizes to Sites of DNA Damage**

(A) Schematic of the domain architecture of FAN1. Conserved domains are indicated: UBZ, ubiquitin-binding zinc finger; SAP, SAF-A/B, Acinus and PIAS; TPR, tetratricopeptide repeat; Nuc (VRR-NUC) virus type replication-repair nuclease.

(B) ClustalW2 alignment of UBZ, SAP, and VRR-NUC domains across different species. Note that the *S. pombe* carries the SAP and Nuc domain but no recognizable UBZ domain. Stars indicate residues mutated in subsequent experiments.

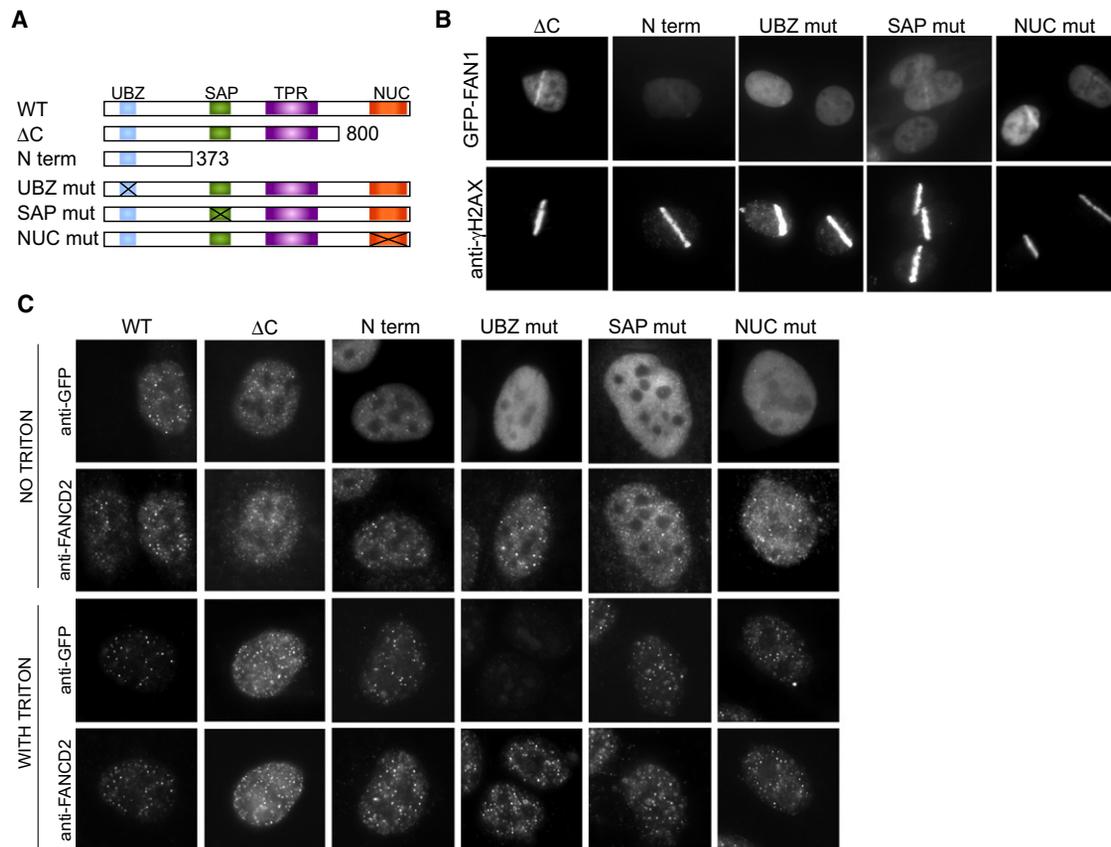
(C) U2OS cells expressing GFP-FAN1 were laser microirradiated and after 30 min were processed for imaging of GFP-FAN1 and  $\gamma$ -H2AX. GFP was visualized directly. Nuclei were stained with DAPI.

(D) U2OS cells expressing GFP-FAN1 were treated with 1  $\mu$ M MMC for 24 hr and processed for indirect immunofluorescence with antibodies against GFP and FANCD2. Nuclei are stained with DAPI. Images were captured and deconvolved using the DeltaVision Image Restoration Microscope.

### FAN1 Has an Endonuclease and 5' Exonuclease Activity

FAN1 contains a conserved nuclease domain. To assess its activity, we examined whether HA-FAN1 complexes immunoprecipitated from 293T cells could act on radiolabeled substrates including 3' and 5' flaps, replication forks, and nicked substrates. Two nuclease activities were observed. One was a 5' to 3' exonuclease especially active on 3'FLAP substrate (Figure 7A, lane 8) as well as a nicked substrate (Figure 7A, lane 17). The second activity was an endonuclease activity on a 5'FLAP substrate (Figure 7A, lane 2), replication fork substrate (Figure 7A, lane 11), and nicked substrate (Figure 7A, lane 20). The endonuclease activity on the 5'FLAP substrate was seen on the top strand at the junction of the two DNA duplexes, not on the single strand flap itself. Neither the exonuclease nor the endonuclease activities were

seen in the immunoprecipitates from cells expressing control HA protein or a mutant FAN1 with the key residues predicted to be necessary for the catalysis substituted to alanines (Q864A, D960A, E975A, K977A) (Figure 7A, lanes 3, 9, 12, 18, and 21), although the wild-type and mutant proteins were immunoprecipitated to the same extent (Figure S5A). To further show that this activity is intrinsic to FAN1 as opposed to a factor like MLH1/PMS2 associated with FAN1 in mammalian cells, we expressed the last 644 amino acids of human FAN1 (aa 373–1017) as a His<sub>6</sub> fusion protein in bacteria (Figure 7B, Figures S5B and S5C). Indeed, the purified protein had the activities seen with the mammalian FAN1. More robust activity of FAN1 purified from bacteria gave us an opportunity to define the substrate specificity. Using a 3'-labeled substrate, we have confirmed



**Figure 5. Localization of FAN1 to Damage Sites Depends on the UBZ Domain**

(A) Schematic of mutant proteins used in FAN1 localization experiments.

(B) U2OS cells expressing the indicated GFP-FAN1 mutants were laser microirradiated and after 30 min were processed for imaging of GFP-FAN1 and  $\gamma$ -H2AX.

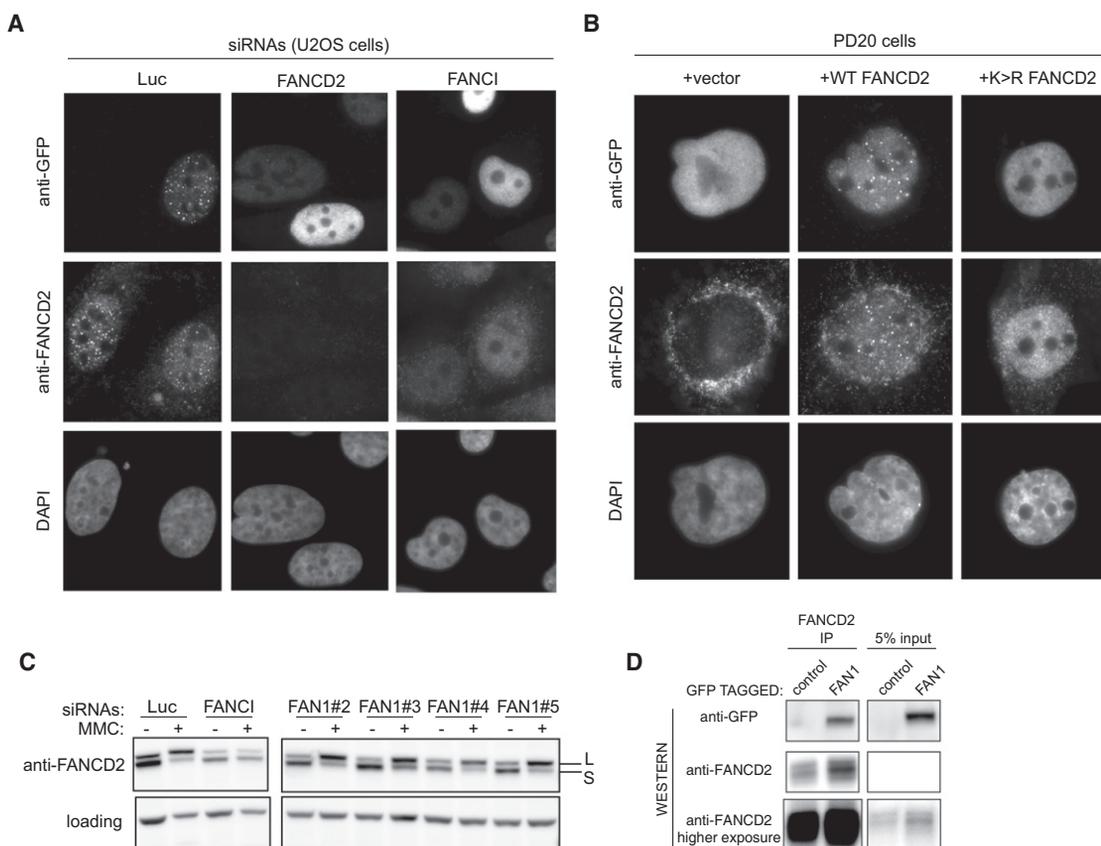
(C) U2OS expressing the indicated GFP-FAN1 mutants were treated with 1  $\mu$ M MMC for 24 hr and processed for indirect immunofluorescence with antibodies against GFP and FANCD2. Cells were either directly fixed with formaldehyde or pre-extracted with Triton X-100 before fixation. Exposures were adjusted to show the presence of foci even if foci were dim. For the comparison of the intensity of foci, see Figure S3.

that the exonuclease activity was a 5' to 3' activity (Figure 7B, lane 26). The major endonuclease activity of FAN1 was observed on the top strand of the 5'FLAP substrate at the junction of the DNA duplexes (Figure 7B, lane 8) and across from a nick (Figure 7B, lane 35). A weak 5'FLAP endonuclease activity was seen on the bottom strand (Figure 7B, lane 11). On a replication fork, the major activity was on a top strand (Figure 7B, lane 29), with some activity seen on the bottom strand (Figure 7B, lane 32). FAN1 activity was stimulated by ATP (Figure S5E). We noted a well-conserved motif in FAN1 (GFDQGIHGEGST, amino acids 826–837 of human protein) that may be able to bind to ATP. The significance of this motif remains to be determined. A mutant FAN1 with the key residues predicted to be necessary for the catalysis substituted to alanines (Q864A, D960A, E975A, K977A) (Figure 7B) or a mutant with just two mutations (E975A, K977A) (Figure S5F) lacked the 5'-3' exonuclease and the endonuclease activities. The bacterially purified proteins did have some contaminating 3' exonuclease activity (Figure 7B, lane 27). Based on these in vitro experiments, we conclude that FAN1 possesses an intrinsic nuclease activity that participates in DNA repair.

## DISCUSSION

### An shRNA Screen Identifies Many Putative Players in Crosslink Resistance

ICLs are among the most lethal lesions to cells, and their repair pathway(s) remain poorly understood. Given the difficulty that ICLs create during replication and transcription, it is critical to identify all the participants in the repair process. Therefore we performed an shRNA screen to identify new components of crosslink repair. Despite the limitations of RNAi, we were able to confirm a number of proteins involved in resistance to crosslinking agents using shRNAs and siRNAs. Among these were C4orf21, which has both a ZF-GRF, a presumed DNA-binding domain, and a domain with similarities to helicases and FLJ25006, which has a kinase domain with predicted serine/threonine activity. Among the known DNA repair proteins was POLQ, an error-prone translesion DNA polymerase (Arana et al., 2008; Seki et al., 2004). POLQ in chicken cells has been shown to be involved in repair of oxidative damage, but not in crosslink repair (Yoshimura et al., 2006). However, in *C. elegans*, POLQ has been implicated in ICL repair, in a pathway distinct



**Figure 6. Localization of FAN1 Depends on the Fanconi Anemia Proteins FANCI and FANCD2**

(A) U2OS cells expressing GFP-FAN1 and transfected with the indicated siRNAs were treated with 1  $\mu$ M MMC for 24 hr, fixed, and processed for indirect immunofluorescence with antibodies against GFP and FANCD2. The corresponding Triton X-100 extraction experiment is shown in Figure S4B.

(B) PD20 cells (FANCD2 negative) complemented with vector, wild-type FANCD2, and the FANCD2K561R mutant (which cannot be ubiquitinated) and expressing GFP-FAN1 were treated with 1  $\mu$ M MMC for 24 hr, fixed, and processed for indirect immunofluorescence with antibodies against GFP and FANCD2. The corresponding Triton X-100 extraction experiment is shown in Figure S4C.

(C) U2OS cells transfected with the indicated siRNAs were treated with 1  $\mu$ M MMC for 24 hr and collected for western blotting with anti-FANCD2 antibodies.

(D) Cell extracts of 293 cells expressing GFP-FAN1 were subjected to immunoprecipitation using FANCD2 antibody. GFP-FAN1 and FANCD2 were identified by immunoblotting.

from the FA pathway but in the same pathway as the *C. elegans* BRCA1 ortholog (Muzzini et al., 2008).

Among the validated genes with strong phenotypes were two putative nucleases. We validated involvement of one, EXDL2, in crosslink sensitivity using multiple siRNAs. EXDL2 has a WRN-like exonuclease domain and promises to shed light on the mechanism of crosslink resistance. The other protein FAN1, with an intriguing domain structure, has now been placed as a bona fide nuclease necessary for crosslink repair in the FA pathway.

#### Evolutionary Conservation of FAN1 Function

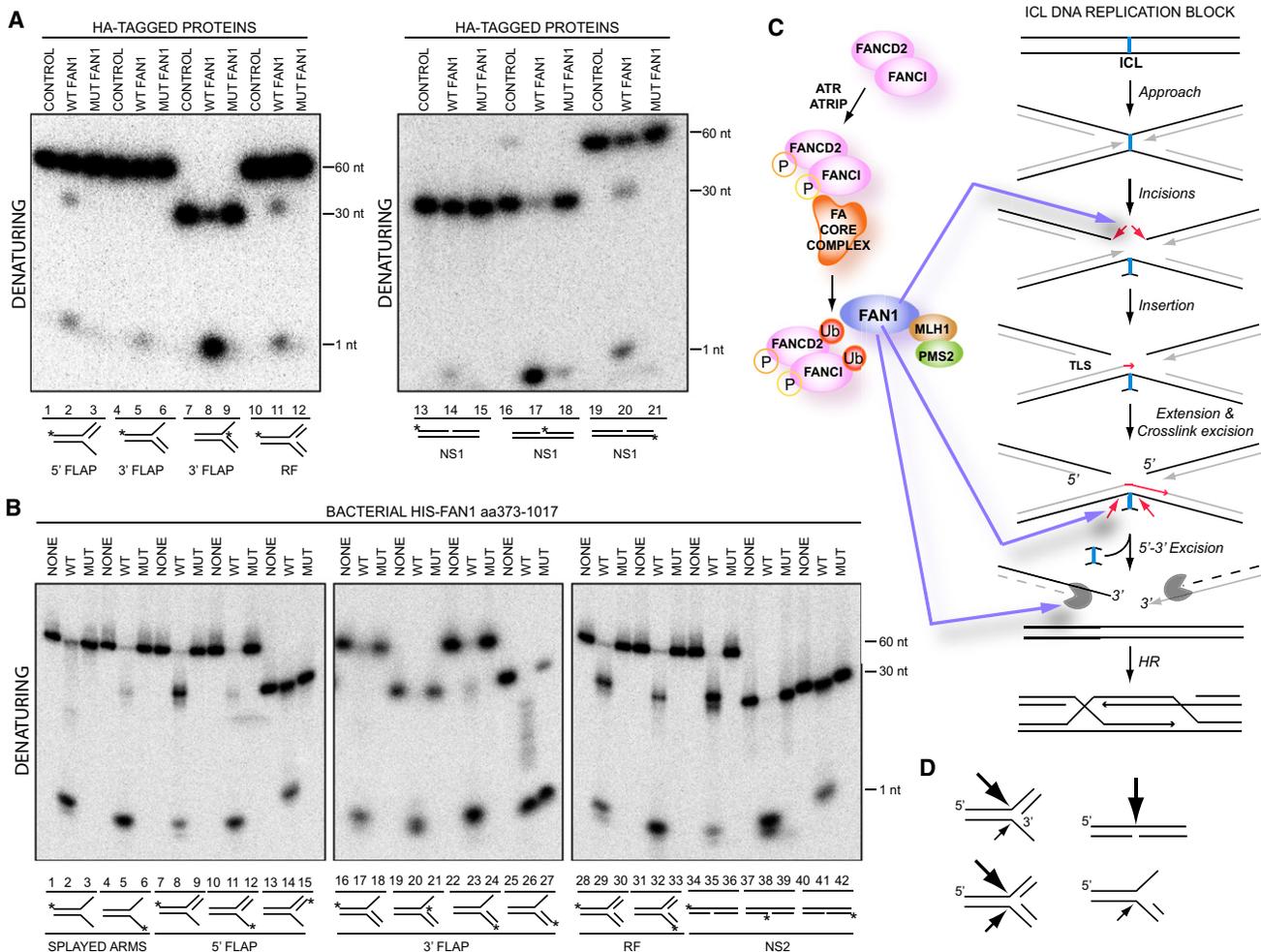
FAN1 is easily identifiable in *Dictyostelium discoideum*, which has orthologs of several FA proteins (Zhang et al., 2009), but no obvious orthologs are apparent in *Drosophila* or *Xenopus*. Interestingly, an *S. pombe* ortholog of FAN1 has a SAP domain, a TRP domain, and a nuclease domain, and thus it is predicted to also function in DNA repair transactions. Fission yeast does not possess the classical FA proteins except for an ortholog of

FANCM, Fml1, which promotes Rad51-dependent gene conversion at stalled replication forks and limits crossing over during mitotic double-strand break repair (Sun et al., 2008). It will be interesting to test if the *S. pombe* FAN1 mutant is sensitive to crosslinking agents and, if so, how it functions without the other proteins present in human cells. This could shed light on alternative pathways for crosslink repair in human cells.

#### Interaction of FAN1 with Mismatch Repair Proteins

We identified all human MutL proteins in nearly stoichiometric complexes with FAN1, suggesting a highly conserved function. Despite the strong interaction with MutLs, we have yet to identify a mismatch repair defect in extracts derived from cells depleted of FAN1 using shRNAs. This could be due to the limitations of RNAi to create the equivalent of a null mutation. Alternatively, the MutL complexes may play roles outside of mismatch repair.

During mismatch repair in bacteria, MutL complexes are recruited to sites of mismatches by MutS complexes that sense the mismatch lesions. MutL complexes then recruit the UvrD



**Figure 7. FAN1 Possesses an Intrinsic Endonuclease and Exonuclease Activity**

(A) Control HA empty vector (CONTROL), HA-FAN1, or HA-FAN1 nuclease mutant (MUT FAN1) (Q864A\_D960A\_E975A\_K977A) complexes were precipitated from 293 TRES cells and incubated with the indicated <sup>32</sup>P-end-labeled substrates prior to electrophoresis on denaturing gels. Asterisk (\*) indicates the position of label on the labeled strand.

(B) Bacterial His<sub>6</sub>-tagged FAN1 aa 373–1017 (wild-type or mutant Q864A\_D960A\_E975A\_K977A) were incubated with <sup>32</sup>P-end-labeled substrates prior to electrophoresis on denaturing gels.

(C) A model of FAN1 activity in the FA pathway. Right-hand side is based on Knipscheer et al. (2009), Raschle et al. (2008). See text for details.

(D) Summary of FAN1's endonucleolytic activity. The size of the arrows corresponds to the strength of the endonuclease.

helicase and the MutH endonuclease, which it also activates, to initiate repair. MutL complexes have also been shown to have an endogenous endonuclease activity that could participate in mismatch repair (Kadyrov et al., 2006). It is possible that MutL complexes could play a similar role in ICL repair. In place of MutS complexes, the monoubiquitinated ID complex would play the analogous role as an ICL lesion sensor to recruit the FAN1-MutL complex to introduce one or more of the four incisions needed for ICL repair. In addition, this complex could recruit additional factors to aid in repair. The MutL $\alpha$  (MLH1-PMS2 heterodimer) complex has been identified as an interactor of FANCD2 (BRIP1), a helicase involved in crosslink repair (Peng et al., 2007). A direct MLH1-FANCD2 interaction was required for FANCD2 to complement the crosslink sensitivity of a patient cell line with a mutation in the FANCD2 gene. Thus, it is possible

that FAN1 and MutL $\alpha$  might act as a bridge between the ID complex and FANCD2 allowing the nuclease activity of FAN1 to pair with FANCD2's helicase activity during the crosslink repair process. It will be important to determine if the interaction of MLH1 with FAN1 and FANCD2 is mutually exclusive and whether FAN1 or MLH1 is required to localize FANCD2 to the sites of crosslinks.

**FAN1 Localization to Sites of Crosslink Damage and Interaction with FANCD2**

FAN1 localizes to laser microirradiation sites as well as to crosslink-induced damage foci, where it colocalized with FANCD2. Localization of FAN1 to foci was dependent on the UBZ domain of FAN1, the presence of FANCI and FANCD2, and the monoubiquitination of FANCD2. Based on these findings, we

hypothesize that FAN1 is recruited to sites of damage by the ID complex using monoubiquitinated FANCD2 and FANCI as an interaction platform for FAN1's UBZ domain.

Since several mutant alleles of FAN1 examined, including the truncation allele lacking the SAP domain and the TPR domain, showed decreased GFP-FAN1 staining at the sites of DNA damage, it is very likely that in addition to the interaction between the ubiquitinated ID complex and the UBZ domain, FAN1 is stabilized in damage foci by interactions with the DNA or other proteins.

FAN1 was also identified by our group as a phosphoprotein in a proteomic analysis of DDR (C. Zhou and S.J.E., unpublished data). The identified phosphorylated SQ site, S210, in the human protein is evolutionarily conserved in mouse, chicken, and fish and is the only SQ/TQ site in FAN1 with such a high extent of conservation. Phosphorylation plays a pivotal role in activating FA pathway, and it is likely that the activity of enzymes participating in repair will be tightly regulated. Therefore, it will be important to establish the functional consequences of abolishing FAN1 phosphorylation.

### FAN1 Is a Nuclease in the Fanconi Anemia Pathway

Biochemical analysis of FAN1 isolated from human cells revealed that FAN1 acts both as an endonuclease and as an exonuclease. The major endonucleolytic activity of FAN1 appears to act opposite a nick and at branched structures with a 3' end at the branch point (Figure 7C). FAN1's minor activities are toward a 3'FLAP and toward replication forks across a 5' end at the branch point. The 5' to 3' exonuclease activity is active on most 5' ends, although its strength varies depending on the specific substrate, with those substrates processed efficiently by the endonuclease activity being poorer substrates for the exonuclease than those inefficiently processed by the endonuclease. The ultimate test of FAN1's activity will come from examining its function on crosslinked substrates in the setting of DNA replication, since this is where ATR is activated and where the interaction with the monoubiquitinated ID complex positions FAN1 at the sites of damage. However, the *in vitro* activities of FAN1 seen so far fit well with the repair activities hypothesized to be present at the crosslink and other lesions that necessitate replication restart (Figure 7D). FAN1 may be involved in the unhooking of the ICL in the setting of replication. The ID complex is required for this process, and since it is also necessary for FAN1 localization at the site of DNA damage, FAN1 is a good candidate for this activity. The exonuclease activity of FAN1 might also be important for the processing of the DNA strands to generate 3' overhangs for the homologous recombinational repair that is necessary to restart the replication fork after the crosslink is removed or to gap repair unreplicated regions when both forks reach the ICL. Lastly, the final steps of removal of the fully unhooked crosslink might also rely on FAN1 activity. This is among the first instances for which we have an understanding of how a nuclease important for crosslink resistance localizes to sites of damage. Although other nucleases including XPF-ERCC1, MUS81-EME1, and SLX1 have been implicated in crosslink repair (Ciccia et al., 2008; Munoz et al., 2009; Svendsen et al., 2009), we still do not understand how they are recruited

to the sites of DNA damage or the nature of their relationship to FA proteins.

### FAN1 Is a Candidate Tumor Suppressor

Based on FAN1's function in ICL resistance, colocalization, dependence on FANCI and FANCD2 function, and interaction with FANCD2, FAN1 is a candidate FA gene. Since three of the FA genes, FANCD1 (BRCA2), FANCN (PALB2), and FANCI (BRIP1), are also mutated in familial breast cancer predisposition syndrome (Rahman et al., 2007; Wooster et al., 1995), FAN1 should be sequenced in the appropriate cohort of patients who display familial predisposition to breast cancer but lack identified predisposing mutations.

The identification of FAN1 brings us closer to an understanding of the biochemical pathway involved in DNA crosslink repair and sets the stage for more precise examination of the repair process in reconstituted crosslink repair systems.

## EXPERIMENTAL PROCEDURES

### Cell Culture, Plasmids, Antibodies, RNAi, RT-qPCR

U2OS, DR-U2OS, HeLa, 293T, 293TRES, and HCT116 cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% (v/v) FBS (Invitrogen), 100 units of penicillin per ml, and 0.1 mg streptomycin per ml. PD20 cells were grown as above but with 15% FBS. Plasmids were constructed using recombinational cloning via the Gateway system (Invitrogen). KIAA1018 clone was obtained from Origene, and the wild-type or truncation mutants were amplified and recombined into pDONR223 (Lamesch et al., 2007). pDONR223 derivatives were recombined into appropriate recipient vectors using LR clonase (Invitrogen). Mutagenesis was performed using multisite mutagenesis kit (Agilent) (primers are listed in Table S4). Antibodies against FAN1 were raised in rabbits against peptide CGQSDSAKREVKQKIS (YenZym) and affinity purified using the antigenic peptide. The other antibodies were FANCD2 (Novus NB100-182), GFP (Roche 11814460001), HA (Covance MMS-101R), MLH1 (Santa Cruz sc-582), MLH3 (Bethyl A301-849A and A301-850A), PMS1 (Santa Cruz sc-615), and PMS2 (BD Pharmingen 556415). siRNA transfections were performed using Lipofectamine RNAiMAX as suggested by the manufacturer with the final siRNA concentration of 50 nM. siRNA sequences are listed in Table S5. For RT-qPCR, Superscript III reverse transcriptase followed by Platinum cybergreen super mix (Invitrogen) were used according to the instructions. GAPDH or actin was used as control.

### Whole-Genome shRNA Screen

The pool-based shRNA screen using HH barcode deconvolution was performed as described previously (Schlachbach et al., 2008).

### Protein Purification and Mass Spectrometry

The 293 TRES cells expressing FAN1 were lysed and immunoprecipitated using anti-HA antibodies (Sowa et al., 2009; Svendsen et al., 2009). Complexes were either used for DNA cleavage assays, subjected to immunoblotting, or eluted with HA peptide and trypsinized prior to mass spectrometry (Sowa et al., 2009; Svendsen et al., 2009). Processing of samples for mass spectrometry as well as analysis of proteomic data using CompPASS were as described (Sowa et al., 2009). For immunoprecipitations with FANCD2, chromatin fraction was prepared as described (Moldovan et al., 2009). For DNA cleavage assays, the immune complexes were washed three times in buffer containing 20 mM Tris HCl (pH 8.0), 5 mM MgCl<sub>2</sub>, 1 mM DTT. Bacterial proteins (FAN1 373–1017) were expressed using pDEST17 (Invitrogen), induced with L-arabinose, and purified on a Ni-NTA column. The proteins were dialyzed against a buffer containing 50 mM Tris HCl (pH 7.5), 100 mM NaCl, 0.01%NP40, 10% glycerol, 1 mM DTT, 0.5 mM EDTA and were stored at –80°C.

**Multicolor Competition Assay**

Experiments were done as described (Smogorzewska et al., 2007).

**Laser-Induced Damage and Immunofluorescence**

Microirradiation was performed as described previously (Bekker-Jensen et al., 2006). Immunofluorescence experiments were done as described (Smogorzewska et al., 2007).

**In Vitro Cleavage Assays**

In vitro cleavage of DNA substrates was performed using the FAN1 immune complexes or bacterially purified FAN1 in conjunction with previously described DNA substrates (Ciccina et al., 2003; Ip et al., 2008; Rass and West, 2006; Svendsen et al., 2009). DNA cleavage assays were performed using 5' <sup>32</sup>P- or 3' <sup>32</sup>P-end-labeled DNA substrates. Substrates were generated by annealing oligonucleotides and were purified by polyacrylamide gel electrophoresis as described previously (Ciccina et al., 2003; Ip et al., 2008; Rass and West, 2006; Svendsen et al., 2009). The sequences of substrates are provided in Table S4. Radiolabeled substrates were incubated with the indicated immune complexes or bacterially purified FAN1. DNA cleavage assays were performed in 20 mM Tris HCl (pH 8.0), 5 mM MgCl<sub>2</sub>, 1 mM ATP, and 1 mM DTT. For bacterially purified FAN1, 20 ng of wild-type protein or 40 ng of mutant protein was used with each substrate. After 30 min (for the bacterial protein) or 2 hr (for the immunoprecipitated proteins) at 37°C, reaction mixtures were treated with 1% Proteinase K in SDS prior to electrophoresis on either 12% polyacrylamide gels (native) or 16% polyacrylamide-urea gels (denaturing). Reaction products were visualized by autoradiography.

**C. elegans Genetics**

*C. elegans* strains were cultured at 20°C under standard conditions (Brenner, 1974). The N2 Bristol strain was used as the wild-type background. The following mutations and chromosome rearrangements were used in this study: LGIII, *him-18(tm2181)* (Saito et al., 2009), *qC1[dpy-19(e1259) glp-1(q339) qIs26]* (III); LGIV, *fan-1(tm423)*, *nT1[unc-?(n754) let-? qIs50]* (IV; V), *nT1[qIs51]* (IV; V).

The *fan-1(tm423)* mutant, obtained from the Japanese National Bioresource Project, carries a 411 bp in-frame deletion encompassing parts of exons 5–8. This deletion results in the loss of the predicted SAP motif.

**DNA Interstrand Crosslink Sensitivity Assay in C. elegans**

Young adult worms were treated with 0, 250, or 500 μM of MMC (Sigma) in M9 buffer containing *E. coli* OP50 with slow shaking in the dark for 19 hr. Treatment with nitrogen mustard (mechlorethamine hydrochloride; Sigma) was similar, but with doses of 0, 50, or 100 μM. Following treatment with MMC or HN2, animals were plated to allow recovery for 3 hr. Twenty animals were plated five per plate, and hatching was assessed for the time period 22–26 hr from the start of treatment. Each damage condition was replicated at least three times in independent experiments. *him-18/slx-4* mutants, shown previously to be extremely sensitive to ICL-inducing agents (Saito et al., 2009), were used as a control. Since untreated *him-18(tm2181)* mutants have reduced hatching, embryonic viability after DNA damage treatment was plotted as a percentage of the hatching after DNA damage normalized by that in untreated animals (relative hatching) (Saito et al., 2009).

**SUPPLEMENTAL INFORMATION**

Supplemental Information includes five figures, four tables, Supplemental Experimental Procedures, and Supplemental References and can be found with this article at doi:10.1016/j.molcel.2010.06.023.

**ACKNOWLEDGMENTS**

We thank Jen Svendsen and members of the Elledge lab for protocols and discussion. Special thanks to Alberto Ciccina for insightful suggestions and comments. This work is supported by grants from the National Institutes of Health (NIH) and CMCR grant #1U19A1067751-01 to S.J.E., NIH grant

R01GM072551 and a Giovanni Armenise-Harvard Foundation award to M.P.C., a NIH grant to J.W.H., and Project Z01 ES065089 to T.A.K., in the Division of Intramural Research of the National Institute of Environmental Health Sciences, NIH. A.S. was supported by T32CA09216 to the Pathology Department at the Massachusetts General Hospital and by Burroughs Wellcome Fund Career Award for Medical Scientists and is an Irma T. Hirsch scholar. S.J.E. is an investigator of the Howard Hughes Medical Institute.

Received: April 13, 2010

Revised: June 7, 2010

Accepted: June 9, 2010

Published: July 8, 2010

**REFERENCES**

- Alter, B.P., Greene, M.H., Velazquez, I., and Rosenberg, P.S. (2003). Cancer in Fanconi anemia. *Blood* 101, 2072.
- Andreassen, P.R., D'Andrea, A.D., and Taniguchi, T. (2004). ATR couples FANCD2 monoubiquitination to the DNA-damage response. *Genes Dev.* 18, 1958–1963.
- Arana, M.E., Seki, M., Wood, R.D., Rogozin, I.B., and Kunkel, T.A. (2008). Low-fidelity DNA synthesis by human DNA polymerase theta. *Nucleic Acids Res.* 36, 3847–3856.
- Auerbach, A.D., and Wolman, S.R. (1976). Susceptibility of Fanconi's anaemia fibroblasts to chromosome damage by carcinogens. *Nature* 261, 494–496.
- Bakkenist, C.J., and Kastan, M.B. (2004). Initiating cellular stress responses. *Cell* 118, 9–17.
- Bartek, J., Lukas, C., and Lukas, J. (2004). Checking on DNA damage in S phase. *Nat. Rev. Mol. Cell Biol.* 5, 792–804.
- Bekker-Jensen, S., Lukas, C., Kitagawa, R., Melander, F., Kastan, M.B., Bartek, J., and Lukas, J. (2006). Spatial organization of the Mamm Genome surveillance machinery in response to DNA strand breaks. *J. Cell Biol.* 173, 195–206.
- Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. *Genetics* 77, 71–94.
- Cannavo, E., Gerrits, B., Marra, G., Schlapbach, R., and Jiricny, J. (2007). Characterization of the interactome of the human MutL homologues MLH1, PMS1, and PMS2. *J. Biol. Chem.* 282, 2976–2986.
- Ciccina, A., Constantinou, A., and West, S.C. (2003). Identification and characterization of the human mus81-eme1 endonuclease. *J. Biol. Chem.* 278, 25172–25178.
- Ciccina, A., McDonald, N., and West, S.C. (2008). Structural and functional relationships of the XPF/MUS81 family of proteins. *Annu. Rev. Biochem.* 77, 259–287.
- Cole, A.R., Lewis, L.P., and Walden, H. (2010). The structure of the catalytic subunit FANCL of the Fanconi anemia core complex. *Nat. Struct. Mol. Biol.* 17, 294–298.
- Cox, L.S., Clancy, D.J., Boubriak, I., and Saunders, R.D. (2007). Modeling Werner Syndrome in *Drosophila melanogaster*: hyper-recombination in flies lacking WRN-like exonuclease. *Ann. N Y Acad. Sci.* 1119, 274–288.
- Fanconi, G. (1967). Familial constitutional panmyelocytopenia, Fanconi's anemia (F.A.). I. Clinical aspects. *Semin. Hematol.* 4, 233–240.
- Fekairi, S., Scaglione, S., Chahwan, C., Taylor, E.R., Tissier, A., Coulon, S., Dong, M.Q., Ruse, C., Yates, J.R., 3rd, Russell, P., et al. (2009). Human SLX4 is a Holliday junction resolvase subunit that binds multiple DNA repair/recombination endonucleases. *Cell* 138, 78–89.
- Harper, J.W., and Elledge, S.J. (2007). The DNA damage response: ten years after. *Mol. Cell* 28, 739–745.
- Ho, G.P., Margossian, S., Taniguchi, T., and D'Andrea, A.D. (2006). Phosphorylation of FANCD2 on two novel sites is required for mitomycin C resistance. *Mol. Cell Biol.* 26, 7005–7015.

- Ip, S.C., Rass, U., Blanco, M.G., Flynn, H.R., Skehel, J.M., and West, S.C. (2008). Identification of Holliday junction resolvases from humans and yeast. *Nature* **456**, 357–361.
- Ishiai, M., Kitao, H., Smogorzewska, A., Tomida, J., Kinomura, A., Uchida, E., Saberi, A., Kinoshita, E., Kinoshita-Kikuta, E., Koike, T., et al. (2008). FANCI phosphorylation functions as a molecular switch to turn on the Fanconi anemia pathway. *Nat. Struct. Mol. Biol.* **15**, 1138–1146.
- Kadyrov, F.A., Dzantiev, L., Constantin, N., and Modrich, P. (2006). Endonucleolytic function of MutLalpha in human mismatch repair. *Cell* **126**, 297–308.
- Kinch, L.N., Ginalski, K., Rychlewski, L., and Grishin, N.V. (2005). Identification of novel restriction endonuclease-like fold families among hypothetical proteins. *Nucleic Acids Res.* **33**, 3598–3605.
- Knipscheer, P., Raschle, M., Smogorzewska, A., Enoiu, M., Ho, T.V., Scharer, O.D., Elledge, S.J., and Walter, J.C. (2009). The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. *Science* **326**, 1698–1701.
- Lamesch, P., Li, N., Milstein, S., Fan, C., Hao, T., Szabo, G., Hu, Z., Venkatesan, K., Bethel, G., Martin, P., et al. (2007). hORFeome v3.1: a resource of human open reading frames representing over 10,000 human genes. *Genomics* **89**, 307–315.
- Lehmann, A.R., Niimi, A., Ogi, T., Brown, S., Sabbioneda, S., Wing, J.F., Kannouche, P.L., and Green, C.M. (2007). Translesion synthesis: Y-family polymerases and the polymerase switch. *DNA Repair (Amst.)* **6**, 891–899.
- Machida, Y.J., Machida, Y., Chen, Y., Gurtan, A.M., Kupfer, G.M., D'Andrea, A.D., and Dutta, A. (2006). UBE2T is the E2 in the Fanconi anemia pathway and undergoes negative autoregulation. *Mol. Cell* **23**, 589–596.
- Matsuoka, S., Ballif, B.A., Smogorzewska, A., McDonald, E.R., 3rd, Hurov, K.E., Luo, J., Bakalarski, C.E., Zhao, Z., Solimini, N., Lerenthal, Y., et al. (2007). ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* **316**, 1160–1166.
- Meetei, A.R., Yan, Z., and Wang, W. (2004). FANCL replaces BRCA1 as the likely ubiquitin ligase responsible for FANCD2 monoubiquitination. *Cell Cycle* **3**, 179–181.
- Moldovan, G.L., Madhavan, M.V., Mirchandani, K.D., McCaffrey, R.M., Vinciguerra, P., and D'Andrea, A.D. (2009). DNA polymerase POLN participates in crosslink repair and homologous recombination. *Mol. Cell. Biol.* **30**, 1088–1096.
- Munoz, I.M., Hain, K., Declais, A.C., Gardiner, M., Toh, G.W., Sanchez-Pulido, L., Heuckmann, J.M., Toth, R., Macartney, T., Eppink, B., et al. (2009). Coordination of structure-specific nucleases by human SLX4/BTBD12 is required for DNA repair. *Mol. Cell* **35**, 116–127.
- Muzzini, D.M., Plevani, P., Boulton, S.J., Cassata, G., and Marini, F. (2008). *Caenorhabditis elegans* POLQ-1 and HEL-308 function in two distinct DNA interstrand cross-link repair pathways. *DNA Repair (Amst.)* **7**, 941–950.
- Niedzwiedz, W., Mosedale, G., Johnson, M., Ong, C.Y., Pace, P., and Patel, K.J. (2004). The Fanconi anaemia gene FANCC promotes homologous recombination and error-prone DNA repair. *Mol. Cell* **15**, 607–620.
- Peng, M., Litman, R., Xie, J., Sharma, S., Brosh, R.M., Jr., and Cantor, S.B. (2007). The FANCI/MutLalpha interaction is required for correction of the cross-link response in FA-J cells. *EMBO J.* **26**, 3238–3249.
- Perry, J.J., Yannone, S.M., Holden, L.G., Hitomi, C., Asaithamby, A., Han, S., Cooper, P.K., Chen, D.J., and Tainer, J.A. (2006). WRN exonuclease structure and molecular mechanism imply an editing role in DNA end processing. *Nat. Struct. Mol. Biol.* **13**, 414–422.
- Rahman, N., Seal, S., Thompson, D., Kelly, P., Renwick, A., Elliott, A., Reid, S., Spanova, K., Barfoot, R., Chagtai, T., et al. (2007). PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nat. Genet.* **39**, 165–167.
- Raschle, M., Knipscheer, P., Enoiu, M., Angelov, T., Sun, J., Griffith, J.D., Ellenberger, T.E., Scharer, O.D., and Walter, J.C. (2008). Mechanism of replication-coupled DNA interstrand crosslink repair. *Cell* **134**, 969–980.
- Rass, U., and West, S.C. (2006). Synthetic junctions as tools to identify and characterize Holliday junction resolvases. *Methods Enzymol.* **408**, 485–501.
- Rogakou, E.P., Boon, C., Redon, C., and Bonner, W.M. (1999). Megabase chromatin domains involved in DNA double-strand breaks in vivo. *J. Cell Biol.* **146**, 905–916.
- Saito, T.T., Youds, J.L., Boulton, S.J., and Colaiacovo, M.P. (2009). *Caenorhabditis elegans* HIM-18/SLX-4 interacts with SLX-1 and XPF-1 and maintains genomic integrity in the germline by processing recombination intermediates. *PLoS Genet.* **5**, e1000735. 10.1371/journal.pgen.1000735.
- Schlabach, M.R., Luo, J., Solimini, N.L., Hu, G., Xu, Q., Li, M.Z., Zhao, Z., Smogorzewska, A., Sowa, M.E., Ang, X.L., et al. (2008). Cancer proliferation gene discovery through functional genomics. *Science* **319**, 620–624.
- Schmid, W., and Fanconi, G. (1978). Fragility and spiralization anomalies of the chromosomes in three cases, including fraternal twins, with Fanconi's anemia, type Estren-Dameshek. *Cytogenet. Cell Genet.* **20**, 141–149.
- Seki, M., Masutani, C., Yang, L.W., Schuffert, A., Iwai, S., Bahar, I., and Wood, R.D. (2004). High-efficiency bypass of DNA damage by human DNA polymerase Q. *EMBO J.* **23**, 4484–4494.
- Simpson, L.J., and Sale, J.E. (2003). Rev1 is essential for DNA damage tolerance and non-templated immunoglobulin gene mutation in a vertebrate cell line. *EMBO J.* **22**, 1654–1664.
- Smogorzewska, A., Matsuoka, S., Vinciguerra, P., McDonald, E.R., 3rd, Hurov, K.E., Luo, J., Ballif, B.A., Gygi, S.P., Hofmann, K., D'Andrea, A.D., et al. (2007). Identification of the FANCI protein, a monoubiquitinated FANCD2 paralog required for DNA repair. *Cell* **129**, 289–301.
- Sowa, M.E., Bennett, E.J., Gygi, S.P., and Harper, J.W. (2009). Defining the human deubiquitinating enzyme interaction landscape. *Cell* **138**, 389–403.
- Stokes, M.P., Rush, J., Macneill, J., Ren, J.M., Spratt, K., Nardone, J., Yang, V., Beausoleil, S.A., Gygi, S.P., Livingstone, M., et al. (2007). Profiling of UV-induced ATM/ATR signaling pathways. *Proc. Natl. Acad. Sci. USA* **104**, 19855–19860.
- Sun, W., Nandi, S., Osman, F., Ahn, J.S., Jakovleska, J., Lorenz, A., and Whitby, M.C. (2008). The FANCM ortholog Fml1 promotes recombination at stalled replication forks and limits crossing over during DNA double-strand break repair. *Mol. Cell* **32**, 118–128.
- Svendsen, J.M., Smogorzewska, A., Sowa, M.E., O'Connell, B.C., Gygi, S.P., Elledge, S.J., and Harper, J.W. (2009). Mammalian BTBD12/SLX4 assembles a Holliday junction resolvase and is required for DNA repair. *Cell* **138**, 63–77.
- Wooster, R., Bignell, G., Lancaster, J., Swift, S., Seal, S., Mangion, J., Collins, N., Gregory, S., Gumbs, C., and Micklem, G. (1995). Identification of the breast cancer susceptibility gene BRCA2. *Nature* **378**, 789–792.
- Yoshimura, M., Kohzaki, M., Nakamura, J., Asagoshi, K., Sonoda, E., Hou, E., Prasad, R., Wilson, S.H., Tano, K., Yasui, A., et al. (2006). Vertebrate POLQ and POLbeta cooperate in base excision repair of oxidative DNA damage. *Mol. Cell* **24**, 115–125.
- Zhang, X.Y., Langenick, J., Traynor, D., Babu, M.M., Kay, R.R., and Patel, K.J. (2009). Xpf and not the Fanconi anaemia proteins or Rev3 accounts for the extreme resistance to cisplatin in *Dictyostelium discoideum*. *PLoS Genet.* **5**, e1000645. 10.1371/journal.pgen.1000645.
- Zietlow, L., Smith, L.A., Bessho, M., and Bessho, T. (2009). Evidence for the involvement of human DNA polymerase N in the repair of DNA interstrand cross-links. *Biochemistry* **48**, 11817–11824.
- Zou, L., and Elledge, S.J. (2003). Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science* **300**, 1542–1548.

## **Supplemental Information**

### **A Genetic Screen Identifies FAN1, a Fanconi Anemia-Associated Nuclease Necessary for DNA Interstrand Crosslink Repair**

**Agata Smogorzewska, Rohini Desetty, Takamune T. Saito, Michael Schlabach,  
Francis P. Lach, Mathew E. Sowa, Alan B. Clark, Thomas A. Kunkel, J. Wade Harper,  
Monica P. Colaiácovo, and Stephen J. Elledge**

Inventory of Supplemental information:

**Figure S1**, related to Figure 1.

**Figure S2**, related to Figure 2.

**Figure S3**, related to Figure 5

**Figure S4**, related to Figure 6

**Figure S5**, related to Figure 7

**Table S1**, related to Figure 1

**Table S2**, related to Figure 1

**Table S3**, related to Figure 1

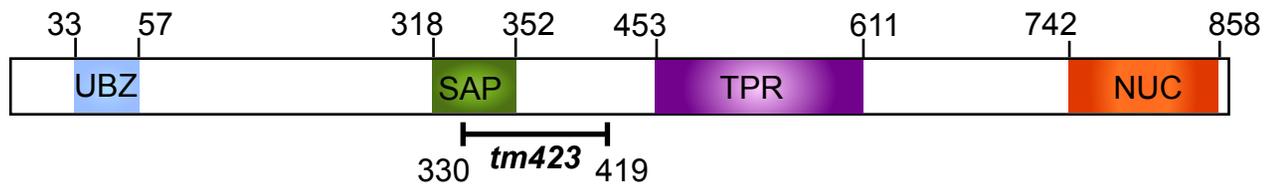
**Table S4**, related to Experimental Procedures

**Supplemental Experimental Procedures**

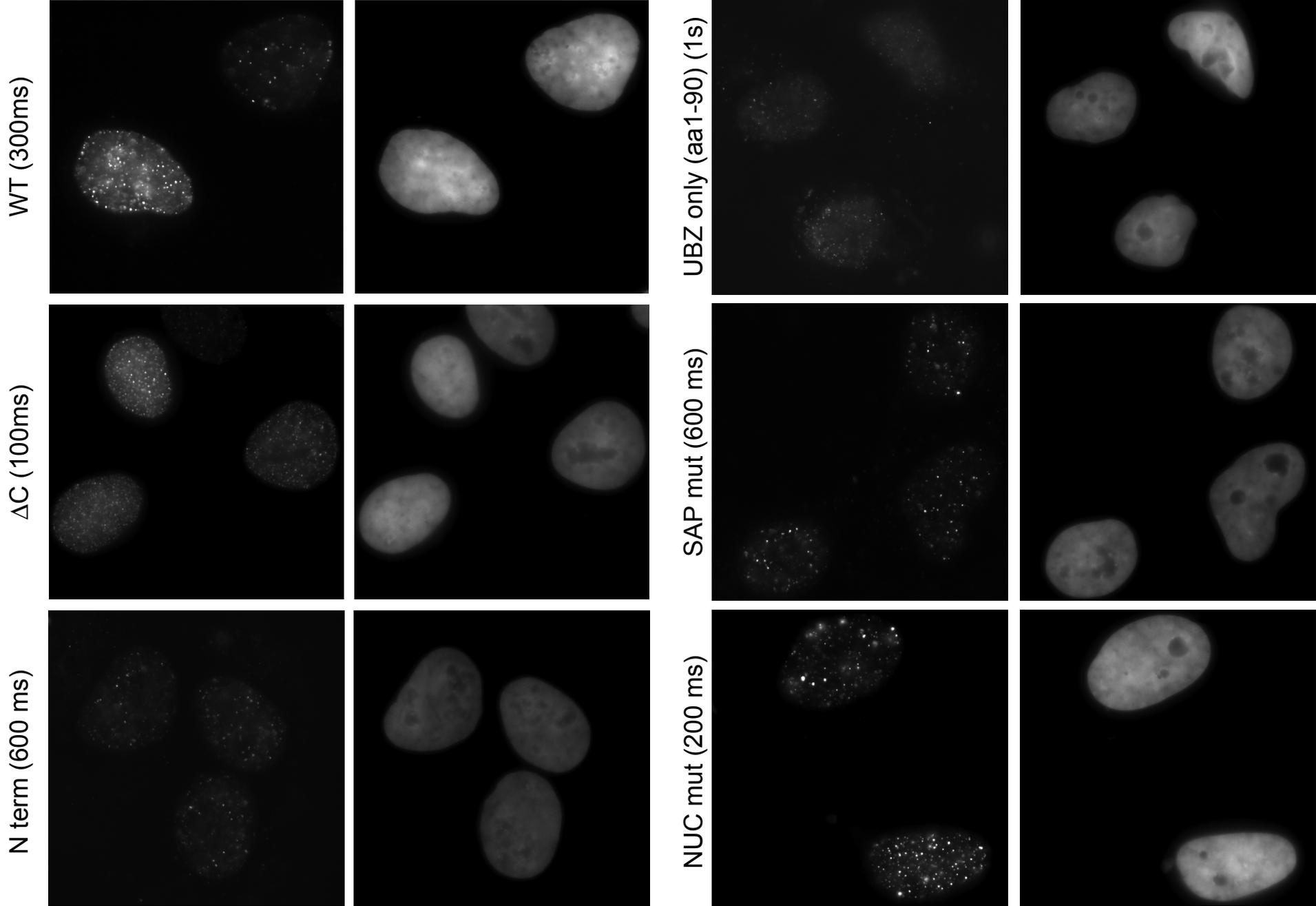
**Supplemental References**



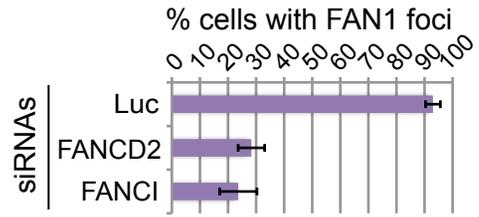
Supplementary Figure 2 Smogorzewska et al.



Supplementary Figure 3 Smogorzewska et al.

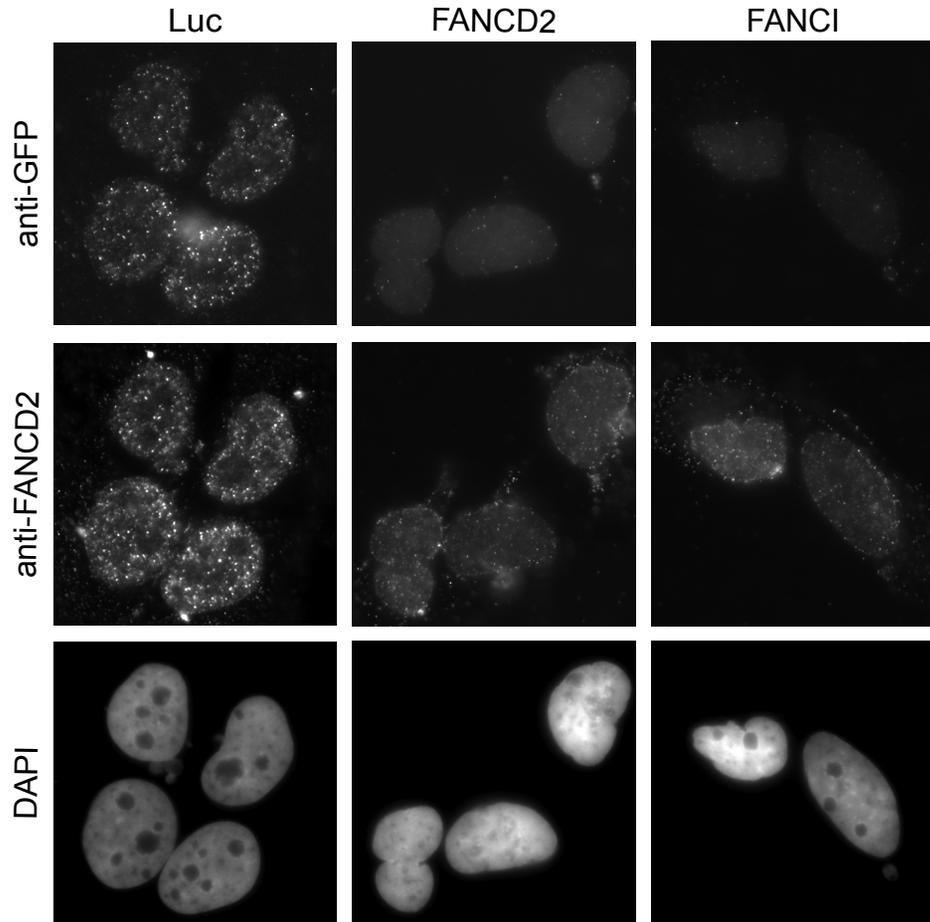


**A**



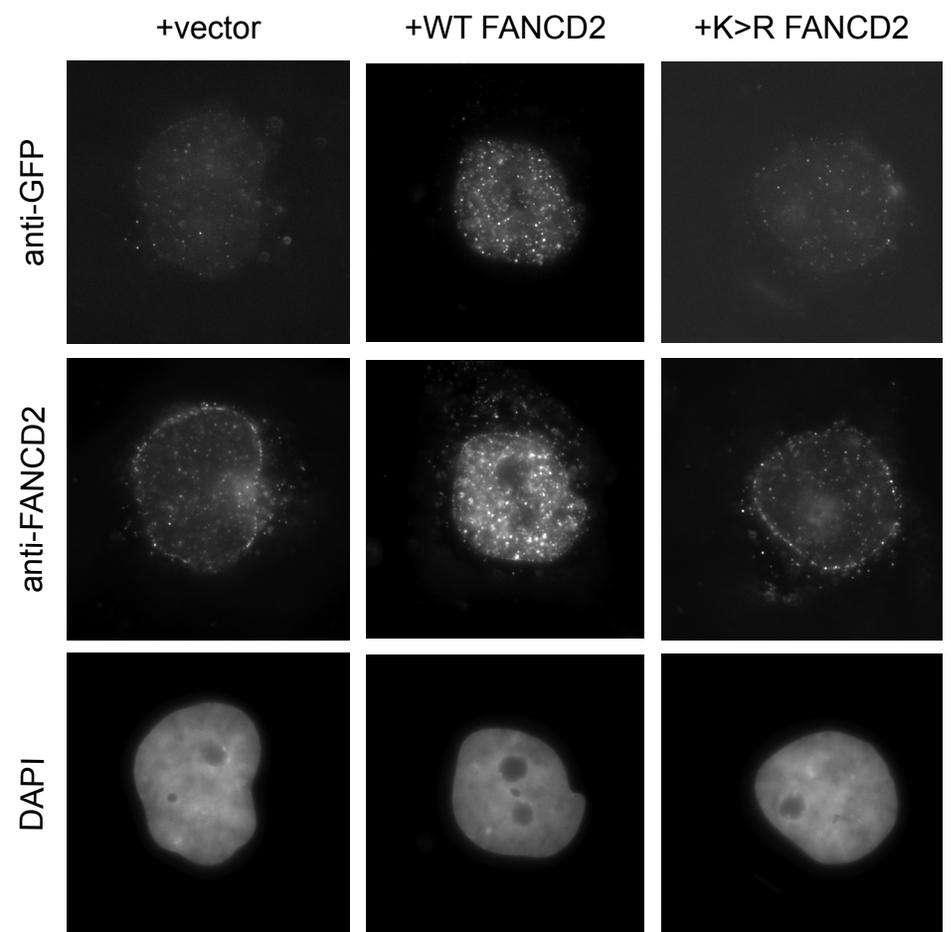
**B**

siRNAs (U2OS cells plus triton)

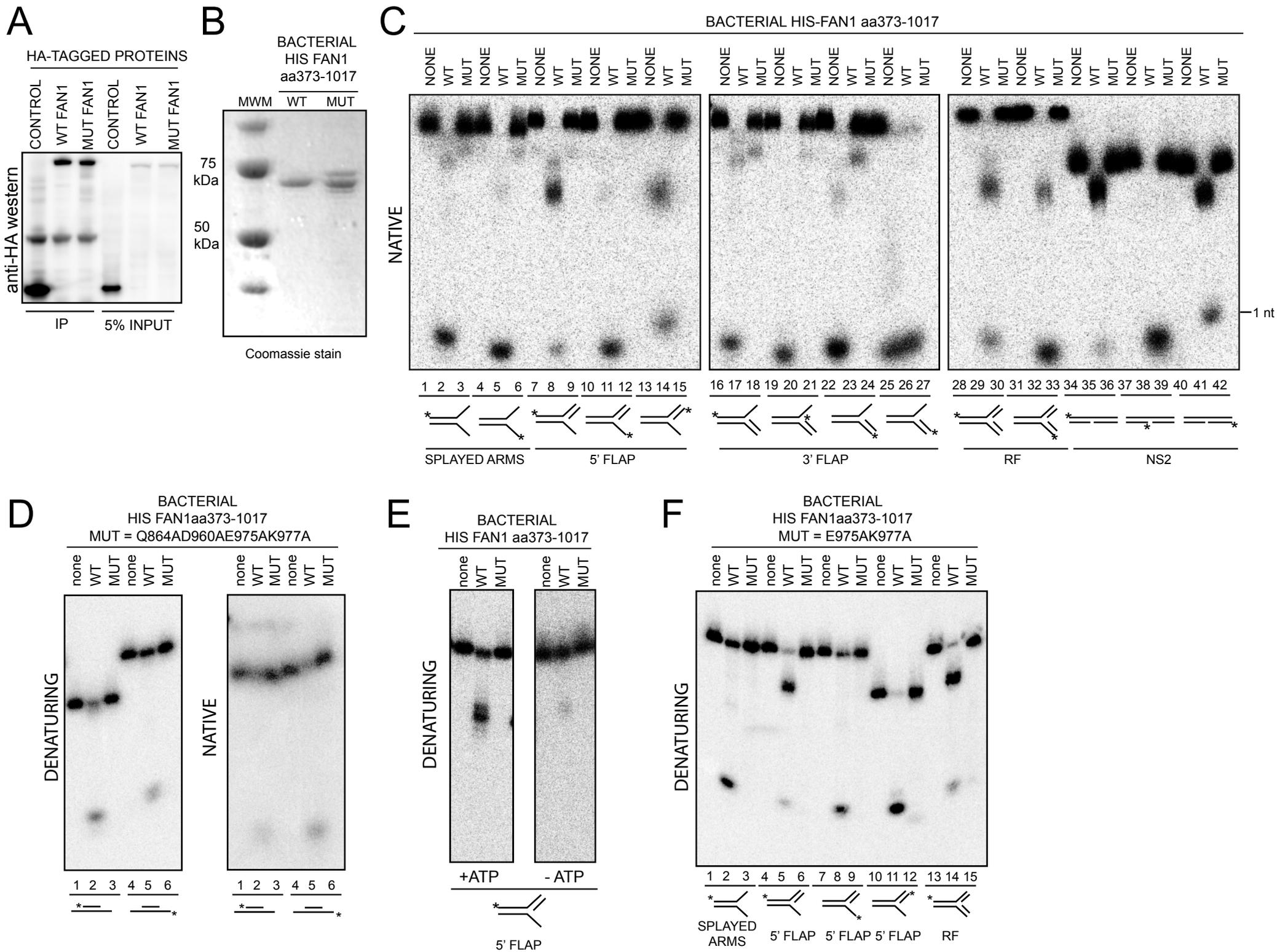


**C**

PD20 (plus triton)



**Supplementary Figure 5** Smogorzewska et al.



**Figure S1**, related to Figure 1.

**A.** Schematic representation of EXDL2. **B.** ClustalW2 alignment of the EXDL2 nuclease domain. Human WRN protein was used in the alignment. The residues colored in orange have been shown to be essential for the exonuclease activity of WRN protein.

**C.** RT-qPCR in U2OS cells transfected with the different siRNAs against EXDL2. **D.** U2OS cells transfected with the indicated siRNAs were treated with 1  $\mu$ M MMC for 24 hours and collected for western blotting with anti-FANCD2 antibody.

**Figure S2**, related to Figure 2.

Schematic representation of the predicted *C. elegans* FAN1 (open reading frame C01G5.8) protein structure. The region deleted in the *tm423* mutant allele is indicated. Domains are as in Figure 3.

**Figure S3**, related to Figure 5

U2OS expressing the indicated GFP-FAN1 mutants were treated with 1  $\mu$ M MMC for 24 hours and pre-extracted with TritonX-100 for 5 minutes and processed for indirect immunofluorescence with antibodies against GFP and FANCD2. Exposures are indicated to show relative intensity of the FAN1 foci.

**Figure S4**, related to Figure 6

**A.** Percentage of cells with FAN1 foci after transfection of indicated siRNAs in U2OS cells expressing GFP-FAN1. Two hundred cells were counted in triplicate and the mean with the standard deviations are indicated. **B.** U2OS cells expressing GFP-FAN1 and transfected with the indicated siRNAs were treated with 1  $\mu$ M MMC for 24 hours, pre-extracted with TritonX-100, fixed and processed for indirect immunofluorescence with antibodies against GFP and

FANCD2. **C.** PD20 cells (FANCD2 negative) complemented with vector, wild type FANCD2 and FANCD2K561R mutant (which cannot be ubiquitinated) and expressing GFP-FAN1 were treated with 1  $\mu$ M MMC for 24 hours, pre-extracted with TritonX-100, fixed and processed for indirect immunofluorescence with antibodies against GFP and FANCD2.

**Figure S5**, related to Figure 7

**A.** Western blot analysis with HA antibodies of immunoprecipitates used for experiments in Figure 7A. **B.** Coomassie stained gel of bacterially-purified His-tagged FAN1 used for experiments shown in Figure 7B. **C.** FAN1 activity on various substrates. Bacterial His<sub>6</sub>-tagged FAN1aa373-1017 (WT or mutant Q864A\_D960A\_E975A\_K977A) were incubated with <sup>32</sup>P-end labeled substrates prior to electrophoresis on native gels. The samples are identical to the ones shown on a denaturing gel in Figure 7C. After the incubations, the same samples were divided and run on denaturing and native gels. **D.** FAN1 activity on additional substrates. Bacterial His<sub>6</sub>-tagged FAN1 aa373-1017 (WT or mutant (Q864A\_D960A\_E975A\_K977A) were incubated with <sup>32</sup>P-end labeled substrates prior to electrophoresis on native or denaturing gels. **E.** Activity of the of bacterially-purified His-tagged FAN1 with and without the addition of ATP. **F.** Bacterial His<sub>6</sub>-tagged FAN1 aa373-1017 (WT or mutant E975A\_K977A) were incubated with <sup>32</sup>P-end labeled substrates prior to electrophoresis on a denaturing gel.

## Supplemental Experimental Procedures

### Whole genome shRNA screen

The pool-based shRNA screen using half-hairpin (HH) barcode deconvolution was performed as described before (Schlabach et al., 2008). shRNA library containing 74,905 retroviral shRNAs targeting 32,293 unique human gene transcripts (including 19,542 RefSeqs) were screened as 6 pools of ~13,000 shRNAs per pool in independent triplicates. The genome-wide mir30shRNA library was expressed using the retroviral vector MSCV-PM (Schlabach et al., 2008) and is available through Open Biosystems Inc. Retroviral pools were prepared by transfecting 293T cells using *TransIT*®-293 Transfection Reagent (Mirus) in the presence of Gag-Pol and VSVG-expressing plasmids. The supernatant was collected on day 2 and 3 after transfection and supplemented with Polybrene (8  $\mu$ g/ml). U2OS cells were transduced with the above pools of retroviral shRNA at a representation of ~1,000 and a multiplicity of infection (MOI) of 1-3. Following puromycin selection, the cells were divided into two experimental arms. Half of the cells were left untreated. The other half were treated with 10 nM MMC on days 1, 3, 6, and 8. Cells were collected on day 10. For each passage a minimal representation of 1000 was maintained. Following genomic DNA isolation (de Lange et al., 1990), shRNA HH barcode was PCR-recovered from untreated and MMC-treated cells and labeled with Cy5 and Cy3 dyes respectively. The labeled HH barcode amplicons were competitively hybridized to a microarray containing the corresponding probes. Custom microarrays with HH barcode probe sequences were from Roche Nimblegen. Array hybridization and scanning protocols were based on manufacturer's instructions. For the analysis, only the informative probes (i.e. those with raw signal 2-fold above negative control probes) were used and are listed in Supplementary table 1. shRNAs with the average  $\log_2 > 1$  were considered to confer sensitivity to MMC.

### **Multicolor competition assay**

Gfp U2OS cells were transfected/transduced with a control siRNA/shRNA (luciferase) and rfp cells with an siRNA/shRNA of interest. Gfp and rfp cells were counted and mixed in 1 to 1 ratio and were left untreated or were treated with IR or indicated drugs. After 7 days of culture, all cells were collected and analyzed using LSRII FACS analyzer (BD Bioscience). Relative survival of Luc siRNA-treated cells after damage was set to 100%.

### **Laser-induced damage**

Microirradiation was performed as described previously (Bekker-Jensen et al., 2006). Cells were pre-treated with 10  $\mu$ M BrdU for 24 hours to sensitize them to the UVA laser. Cells were microirradiated using PALM MicroBeam with fluorescence illumination (Zeiss). The power of the 355 nm laser was set between 40-45%, which resulted in localized damage as judged by  $\gamma$ -H2AX staining. DAPI staining was intact under these conditions. U2OS cells were fixed 15-30 min after irradiation and stained with DAPI. GFP-fusion proteins were visualized directly. For co-localization experiments, antibodies against  $\gamma$ -H2AX were used in indirect immunofluorescence staining.

### **Immunofluorescence**

Experiments were done as described (Smogorzewska et al., 2007). Cells grown on autoclaved cover slips were rinsed in PBS and fixed in 3.7% (w/v) formaldehyde (Sigma) diluted in PBS for 10 minutes at room temperature. Cells were washed once with PBS, permeabilized in 0.5% (v/v) NP40 in PBS for 10 minutes, washed again in PBS, and blocked with PBG (0.2% [w/v] cold fish gelatin, 0.5% [w/v] BSA in PBS) for 20 minutes. Coverslips were incubated for 2 hours at either room temperature or at 4°C overnight in a humidified chamber with a primary antibody, then washed 3 times for 5 minutes in PBG, and incubated with the appropriate secondary antibody. After three additional washes in PBG, the coverslips were embedded in Vectashield

(Vector Laboratories) supplemented with DAPI. Triton pre-extraction was performed by incubating cells for 5 minutes at room temperature with 0.5% TritonX-100 in PBS. Cells were fixed and processed as above. Images were captured with an Axioplan2 Zeiss microscope or Observer A2 with a AxioCam MRM Zeiss digital camera supported by Axiovision software. Any co-staining experiments included proper controls to exclude crossing of signal between channels. The co-localization images shown in figure 4D were captured and deconvolved using the DeltaVision Image Restoration Microscope.

### **Analysis of Protein Motifs**

Pfam (Sonnhammer et al., 1997) and HHpred (Soding et al., 2005) were applied to FAN1 for motif prediction. CDD domain analysis was performed on the EXDL2 (Marchler-Bauer et al., 2009). Domain alignment was performed using ClustalW2 (<http://www.ebi.ac.uk/Tools/clustalw2/>) and rendered using ESPript2.2 (<http://esript.ibcp.fr>) (Gouet et al., 1999)

## Supplemental References

Bekker-Jensen, S., Lukas, C., Kitagawa, R., Melander, F., Kastan, M.B., Bartek, J., and Lukas, J. (2006). Spatial organization of the mammalian genome surveillance machinery in response to DNA strand breaks. *J Cell Biol* 173, 195-206.

de Lange, T., Shiue, L., Myers, R.M., Cox, D.R., Naylor, S.L., Killery, A.M., and Varmus, H.E. (1990). Structure and variability of human chromosome ends. *Mol Cell Biol* 10, 518-527.

Gouet, P., Courcelle, E., Stuart, D.I., and Metz, F. (1999). ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15, 305-308.

Marchler-Bauer, A., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., *et al.* (2009). CDD: specific functional annotation with the Conserved Domain Database. *Nucleic Acids Res* 37, D205-210.

Schlabach, M.R., Luo, J., Solimini, N.L., Hu, G., Xu, Q., Li, M.Z., Zhao, Z., Smogorzewska, A., Sowa, M.E., Ang, X.L., *et al.* (2008). Cancer proliferation gene discovery through functional genomics. *Science* 319, 620-624.

Smogorzewska, A., Matsuoka, S., Vinciguerra, P., McDonald, E.R., 3rd, Hurov, K.E., Luo, J., Ballif, B.A., Gygi, S.P., Hofmann, K., D'Andrea, A.D., *et al.* (2007). Identification of the FANCI protein, a monoubiquitinated FANCD2 paralog required for DNA repair. *Cell* 129, 289-301.

Soding, J., Biegert, A., and Lupas, A.N. (2005). The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* 33, W244-248.

Sonnhammer, E.L., Eddy, S.R., and Durbin, R. (1997). Pfam: a comprehensive database of protein domain families based on seed alignments. *Proteins* 28, 405-420.