Binding of FANCI-FANCD2 Complex to RNA and R-Loops Stimulates Robust FANCD2 Monoubiquitination

Highlights

- FANCD2 colocalizes with co-transcriptional R-loops in human cells
- Human FANCI-FANCD2 robustly binds ssRNA, but not RNA:DNA hybrids
- Human FANCI-FANCD2 binds R-loops via the displaced ssDNA strand and ssRNA tail
- ssRNA and R-loop can stimulate robust FANCI-FANCD2 monoubiquitination

Authors

Zhuobin Liang, Fengshan Liang, Yaqun Teng, ..., Li Lan, Patrick Sung, Gary M. Kupfer

Correspondence
sungp@uthscsa.edu (P.S.), gary.kupfer@yale.edu (G.M.K.)

In Brief

Fanconi anemia pathway has a well-known role in the repair of DNA crosslinks, but its recently identified role in suppression of co-transcriptional R-loops remains elusive. Here, Liang et al. show that FANCI-FANCD2 has intrinsic RNA and R-loop binding activity and provide mechanistic insights into FA pathway activation upon transcription stress.
Binding of FANCI-FANCD2 Complex to RNA and R-Loops Stimulates Robust FANCD2 Monoubiquitination

Zhuobin Liang, Fengshan Liang, Yaqn Teng, Xiaoyong Chen, Jingchun Liu, Simonne Longerich, Timsi Rao, Allison M. Green, Natalie B. Collins, Yong Xiong, Li Lan, Patrick Sung, and Gary M. Kupfer

INTRODUCTION

Fanconi anemia (FA) is characterized by developmental abnormalities, bone marrow failure, and cancer predisposition. FA cells are hypersensitive to DNA replicative stress and accumulate co-transcriptional R-loops. Here, we use the Damage At RNA Transcription assay to reveal colocalization of FANCD2 with R-loops in a highly transcribed genomic locus upon DNA damage. We further demonstrate that highly purified human FANCI-FANCD2 (ID2) complex binds synthetic single-stranded RNA (ssRNA) and R-loop substrates with high affinity, preferring guanine-rich sequences. Importantly, we elucidate that human ID2 binds an R-loop structure via recognition of the displaced ssDNA and ssRNA but not the RNA:DNA hybrids. Finally, a series of RNA and R-loop substrates are found to strongly stimulate ID2 monoubiquitination, with activity corresponding to their binding affinity. In summary, our results support a mechanism whereby the ID2 complex suppresses the formation of pathogenic R-loops by binding ssRNA and ssDNA species, thereby activating the FA pathway.

SUMMARY

Fanconi anemia (FA) is characterized by developmental abnormalities, bone marrow failure, and cancer predisposition. FA cells are hypersensitive to DNA replicative stress and accumulate co-transcriptional R-loops. Here, we use the Damage At RNA Transcription assay to reveal colocalization of FANCD2 with R-loops in a highly transcribed genomic locus upon DNA damage. We further demonstrate that highly purified human FANCI-FANCD2 (ID2) complex binds synthetic single-stranded RNA (ssRNA) and R-loop substrates with high affinity, preferring guanine-rich sequences. Importantly, we elucidate that human ID2 binds an R-loop structure via recognition of the displaced ssDNA and ssRNA but not the RNA:DNA hybrids. Finally, a series of RNA and R-loop substrates are found to strongly stimulate ID2 monoubiquitination, with activity corresponding to their binding affinity. In summary, our results support a mechanism whereby the ID2 complex suppresses the formation of pathogenic R-loops by binding ssRNA and ssDNA species, thereby activating the FA pathway.

INTRODUCTION

Fanconi anemia (FA) is a recessive genetic disorder characterized by congenital abnormalities, bone marrow failure, and a strong predisposition to cancer. FA is a multigenic disease, with at least 22 complementation genes (FANCA–FANCQ) having been identified thus far, which, together with other partner proteins, constitute the FA pathway of DNA damage sensing and repair (Nalepa and Clapp, 2018). The FA core complex harbors the ubiquitin E3 ligase FANCL and catalyzes the monoubiquitination of the FANCI-FANCD2 (ID2) heterodimer complex during S phase and upon occurrence of interstrand DNA crosslinkings (ICLs) and replicative stress (Huang et al., 2014; Rajendra et al., 2014; van Twest et al., 2017). The monoubiquitinated ID2 complex stably associates with damaged DNA, where it is thought to nucleate the formation of repair centers that harbor a number of other FA proteins involved in ICL removal and subsequent DNA repair including SLX4/FANCP, BRCA1/FANCQ, BRCA2/FANCD1, BACH1/FANCJ, and PALB2/FANCN (Cantor et al., 2001; Godthelp et al., 2006; Kim et al., 2011; Xia et al., 2007). Importantly, mutation of the ubiquitin acceptor site in FANCD2 impairs the assembly of DNA repair foci comprising the aforementioned FA proteins (Chen et al., 2014; Wang et al., 2004; Yamamoto et al., 2011). Thus, ID2 monoubiquitination represents a pivotal event in the activation of the FA pathway.

Previous studies demonstrated that in vitro monoubiquitination of purified FANCI and ID2 by FANCL-UBE2T is greatly stimulated upon DNA binding (Longerich et al., 2009, 2014; Sato et al., 2012), which can cause a conformational change of ID2 revealing the concealed monouniquition sites (Joo et al., 2011; Liang et al., 2016). In addition, ID2 exhibits binding preference for DNA substrates that mimic intermediates of a stalled replication fork underlying the physiological relevance of ID2 DNA binding in ICL repair.

To further characterize the FA pathway in mitigating conflict of DNA replication and RNA transcription upon DNA damage, recent studies have revealed roles of several FA factors including FANCD2, FANCA, BRCA1/FANCQ, and BRCA2/FANCD1 in the metabolism of R-loops, which contain RNA:DNA hybrids that can interfere with DNA replication and can result in DNA double-strand breaks and genome instability (Bhatia et al., 2014; García-Rubio et al., 2015; Hatchi et al., 2015; Schwab et al., 2015).

In this study, we confirm the presence of increased R-loops in FA mutant cells, as previously reported (García-Rubio et al., 2015; Hatchi et al., 2015; Schwab et al., 2015).
2015; Schwab et al., 2015), show that FANCD2 monoubiquitination is required to prevent their accumulation and demonstrate that FANCD2 colocalizes with R-loops in an actively transcribed genomic region. We also demonstrate that purified recombinant human ID2 can directly bind RNA with a predilection for single-stranded and high-guanine-content RNA, as well as R-loop structures. Importantly, these same RNA and R-loop species are able to stimulate robust monoubiquitination of ID2 complex in a reconstituted system. Taken together, our results provide direct evidence for RNA-mediated ID2 monoubiquitination in the avoidance of genome instability that is induced by the accumulation of pathogenic R-loops.

RESULTS

**FANCD2 Colocalizes with R-Loops at a DNA Damaged, Actively Transcribed Genomic Site**

Recent work has demonstrated that the FA pathway has a role in preventing R-loop accumulation upon replication and transcription stress (García-Rubio et al., 2015; Schwab et al., 2015). In order to confirm these data, we compared FA-D2 mutant cells with absent FANCD2 expression transduced with different FANCD2 expression constructs and found that cells expressing FANCD2 monoubiquitination-dead FANCD2-K561R mutant (FA-D2-K561R) or vector alone had elevated background level of R-loops as visualized by S9.6 immunofluorescence microscopy, which further accumulated over the time-course of DNA cross-linking agent mitomycin C (MMC) treatment (Figures S1A and S1B). Conversely, expression of RNA:DNA endonuclease RNase H1 markedly reduced S9.6 signal and partially rescued the survival of FA-D2 mutant cells in MMC (Figures S1C–S1E). Nondenaturing bisulfite sequencing also demonstrated increased R-loops at the highly transcribed GAPDH gene in cells defective in FANCD2 monoubiquitination (Figure 1F). Collectively, our results demonstrate that ID2 and its monoubiquitination facilitate R-loop avoidance upon replication stress, confirming previously published results (García-Rubio et al., 2015; Schwab et al., 2015).

To understand how transcription influences recruitment and activation of ID2 to prevent co-transcriptional R-loops, we deployed the DNA Damage At RNA Transcription (DART) system to examine the association of FANCD2 (and by implication, ID2) with transcription upon DNA damage. DART utilizes the light-inducible chromophore-modified KillerRed (KR) to generate reactive oxygen species (ROS) at a genome-integrated tet response element (TRE) locus to induce focal DNA damage (Lan et al., 2014; Wei et al., 2015). The transcriptional status of the TRE locus can be modulated by fusion of KR with either transcription activator (TA) or repressor (tetR) (Figure 1A).

We first tested FANCD2 recruitment to the TRE locus in four combinations of transcription status and DNA damage by expressing different effector proteins. FANCD2 became highly enriched at the TRE locus when constitutively active transcription was stalled by ROS-induced DNA damage (TA-KR) but not in cells with only active transcription (TA-Cherry), only DNA damage (tetR-KR), or no transcription and DNA damage (tetR-Cherry) (Figures 1B and 1C). This result clearly supports the idea that both transcription and DNA damage are required to recruit FANCD2 (ID2) to chromatin for its subsequent monoubiquitination to form foci, consistent with a similar observation reported recently that endogenous FANCD2 colocalizes with large transcription units in response to aphidicolin-induced replication stress (Okamoto et al., 2018).

We hypothesized that perturbation of active transcription by ROS damage in TA-KR cells would lead to R-loop accumulation. Indeed, we observed that the TA-KR site was an R-loop hotspot as indicated by the formation of S9.6 foci in a transcription and DNA damage dependent fashion (Teng et al., 2018). Importantly, enriched foci of FANCD2 and S9.6 colocalized in ~75% of the TA-KR positive cells tested, and ~96% of the cells with distinct S9.6 foci at the TRE locus were positive for colocalized FANCD2 foci (Figures 1D and 1E). Furthermore, overexpression of wild-type RNase H1, but not the nuclease-dead RNase H1-D210N mutant, in TA-KR positive cells drastically diminished FANCD2 foci intensity at the TRE locus (Figures 1F and 1G), indicating that FANCD2 enriches at DNA-damaged, transcriptionally active genomic sites in a manner that depends on R-loops. In addition, small interfering RNA (siRNA) depletion of endogenous FANCD2 significantly increased S9.6 foci intensity at the TRE locus in TA-KR positive cells to a similar extent as the depletion of RNA helicase Aquarius (AQR) (Figures 1H and 1I), which is known to resolve R-loops (Sollier et al., 2014), demonstrating the utility of the DART system to study R-loop metabolism.

**Human FANCI-FANCD2 Complex Binds Single-Stranded RNA but Not RNA:DNA Hybrids**

Given that FANCD2 colocalizes with R-loops in cells, we asked whether ID2 complex possesses RNA binding activity. We first tested several biotinylated oligomers for their ability to pull down FANCD2 from FA-D2 mutant cells transduced with either wild-type FANCD2 or FANCD2-K561R. Both FANCD2 and FANCI from cell extracts associated with RNA similarly as with DNA, and FANCD2 monoubiquitination is dispensable for RNA interaction as the K561R mutant can be pulled down also (Figures S2A and S2B). The majority of FANCD2 that was pulled down is not ubiquitinated even after MMC treatment, confirming that ubiquitination is indeed not required for RNA interaction. To investigate whether FANCD2 association with RNA could have nucleotide sequence preference, we tested RNA homopolymers for pull-down of FANCD2 and found that poly-rG has the highest pull-down efficiency (Figure S2C), suggesting RNA binding of FANCD2 may have a preference for guanine-rich sequences, which are often present in co-transcriptional R-loops.

In order to directly test RNA interaction of ID2 with RNA, we purified the human ID2 complex from insect cells (Figure 2A) and tested for RNA binding activity using the electrophoretic mobility shift assay (EMSA). We first examined a set of DNA and RNA substrates with the same nucleotide sequence in both single-stranded (ss) and double-stranded (ds) forms (Table S1). We observed that ID2 had significantly higher affinity for the tested ssRNA over ssDNA (Figure 2B). Unexpectedly, despite the fact that human ID2 bound ssRNA avidly, it had the lowest binding affinity for duplex substrates containing at least one RNA strand, such as dsRNA and RNA:DNA hybrids, while it bound dsDNA with a similar affinity as ssDNA (Figure 2C). Taken together, the
order of human ID2 binding affinity from high to low in our experiments was ssRNA &gt; ssDNA &gt; dsDNA &gt; dsRNA &gt; RNA:DNA hybrids (Figure 2D).

We previously reported that FANCI DNA binding mutants in complex with wild-type FANCD2 can impede the DNA binding activity of ID2 (Longerich et al., 2009). To test whether ID2 shares the same dependence on FANCI in binding RNA, we tested the FANCI (KKEE)/D2 and FANCI (R1285X)/D2 mutants (Figure S2D), which are both impaired for DNA binding, for the ability to bind RNA. Our results showed that both ID2 mutant complexes were also defective in binding RNA (Figures S2E and S2F). To investigate the contribution of each subunit of the ID2 complex toward RNA binding activity, we first tested purified human FANCi and found it binds ssRNA, dsRNA, and RNA:DNA hybrids (Figures S3A and S3B). We were not able to test human FANCD2 protein alone, because, in the absence of FANCI, FANCD2 forms high-molecular-weight aggregates during purification. Instead, we took advantage of chicken FANCI and FANCD2 proteins, which can be purified in a non-aggregated form either alone or as a heterodimer, and found that ID2, FANCI, and FANCD2 all have strong binding activity toward ssRNA, dsRNA, and RNA:DNA hybrids (Figures S3C–S3H). These results confirm the findings of a recent study showing that purified chicken FANCi, FANCD2, and ID2 complex can bind RNA:DNA hybrids.
with high affinity (Okamoto et al., 2018). Taken together, our results suggest that while FANCI and FANCD2 both have strong affinity for RNA, the RNA binding activity of ID2 depends on the FANCI nucleic acid binding motif.

**Guanine Content of RNA Positively Regulates Binding by ID2**

We noted that FANCD2 preferentially interacted with RNA poly-

**Figure 2. Recombinant ID2 Complex Preferentially Binds Single-Stranded and Guanine-Rich RNA**

(A) Purified ID2 heterodimer proteins run on SDS-PAGE and stained with Coomassie Blue.

(B) EMSA showing binding of ssDNA and ssRNA by recombinant ID2.

(C) EMSA showing binding of dsDNA, dsRNA, and RNA:DNA hybrids by recombinant ID2.

(D) Quantification of shifted nucleic acid substrates in (B) and (C).

(E) EMSA showing binding of ssDNA substrates with different guanine content (8%, 28%, or 48%) by recombinant ID2.

(F) Quantification of shifted substrates in (E).

(G) EMSA showing binding of ssRNA substrates with different guanine content (8%, 28%, or 48%) by recombinant ID2.

(H) Quantification of shifted substrates in (G).

For (D), (F), and (H), means were calculated from 3 replicates; error bars represent SD.
the guanine content of nucleic acids on their binding by ID2, we compared three 60-mer ssDNA and ssRNA substrates with varying percentages of guanine content (see Table S1 for sequence information). These substrates are termed here as G8, G28, and G48 based on their guanine content. G28 is the same sequence used in Figure 2B, while G8 is derived from G28 by converting some G bases to A or T and G48 is derived from G28 by converting some A or T bases to G. As predicted, increasing the guanine content of both DNA and RNA enhanced their binding by ID2, and ID2 showed stronger affinity for ssRNA as compared to ssDNA of the same sequence among these tested substrates (Figures 2E–2H). In addition, we also observed similar effect of guanine content on human ID2 binding of dsDNA and dsRNA, despite dsRNA binding was weaker, as expected (Figures S3I–S3L).

**FANCI-FANCD2 Complex Binds R-Loops via the Displaced ssDNA Strand and ssRNA Tail**

To test ID2 for the ability to recognize R-loops, we first prepared a synthetic R-loop structure by hybridizing a guanine-rich RNA strand to a DNA bubble scaffold and compared ID2 binding to it versus a D-loop substrate of the same sequence. Our results showed that the R-loop substrate was bound by ID2 to the same extent as the D-loop (Figures S4A and S4B). Replacing the RNA:DNA hybrid portion of the R-loop to a A:T-rich sequence that has almost no binding affinity to ID2 based on our previous observations (Longerich et al., 2014) actually slightly increased the overall binding affinity (Figures S4C and S4D), suggesting that human ID2 R-loop binding occurs not via recognition of the RNA:DNA hybrids, consistent with our finding that human ID2 complex has poor binding affinity to RNA:DNA hybrids (Figure 2C).

R-loop binding of ID2 could be dependent on the displaced ssDNA strand and influenced by its guanine content and secondary structure. To address these questions, we constructed a series of R-loop substrates with different guanine contents in the displaced DNA strand and found that the R-loop binding affinity of ID2 positively correlates with the guanine content of the displaced ssDNA (Figures 3A and 3D). We also tested an R-loop substrate with the potential to form G-quadruplex, a frequent secondary structure found in the displaced ssDNA strand of R-loops and observed slight enhancement of binding affinity over the control R-loop with the same guanine content but without the G-quadruplex forming sequence (Figures 3B and 3D). Because ID2 robustly bound ssRNA, we hypothesized that another binding site of ID2 to an R-loop could be the ssRNA tail from a dysregulated mRNA transcript. To test this possibility, we synthesized additional R-loop substrates that harbor a 40-nt...
Figure 4. RNA and R-Loop Stimulate FANCD2 Monoubiquitination

(A) *In vitro* ubiquitination reaction of recombinant ID2 with either ssDNA or ssRNA substrates. FANCD2 was ubiquitinated maximally in presence of random sequence of ssRNA or ssDNA and ubiquitination was abrogated by T or U-rich nucleic acids.

(B) DNA and RNA binding mutant FANCI(KKEE)/D2 complex failed to be ubiquitinated.

(C) Comparison of ID2 ubiquitination efficiency using various nucleic acids substrates with different binding affinity. ssRNA conferred the most ubiquitination.

(D) Comparison of ID2 ubiquitination efficiency using nucleic acids substrates with different guanine content. Nucleic acid with the highest guanine content conferred the most ubiquitination.

(E) Comparison of ID2 ubiquitination efficiency using two configurations of D-loops and R-loops. D- and R-loops with guanine-containing single-strand DNA (ssDNA) supported ID2 ubiquitination similarly.

(legend continued on next page)
ssRNA tail at the 5’ end and observed that the additional ssRNA tail markedly increased R-loop ID2 binding affinity (Figures 3C and 3D). Furthermore, comparison of G-rich ssRNA and R-loop in a series of competition EMSA assays showed ID2 associating with both substrates robustly, with slightly higher affinity in ssRNA over R-loop for the substrates we tested (Figures S4E–S4H). Collectively, these results demonstrated that both the ssRNA and ssDNA moieties of a synthetic R-loop are critical components for recognition by ID2, thus implicating them in the recruitment and/or retention of ID2 at pathological R-loops in cells.

RNA and R-loop Support ID2 Monoubiquitination

Here, we asked whether RNA would similarly support FANCD2 monoubiquitination as DNA in a reconstituted in vitro ubiquitination system containing the E2-E3 enzymatic pair UBE2T and FANCL proteins. We found that ssRNA with a random sequence stimulated FANCD2 monoubiquitination to the same extent as the equivalent ssDNA counterpart, whereas poly-U (RNA) and poly-T (DNA) had little or no effect (Figure 4A), consistent with their weak binding affinity to ID2. As expected, ID2 complex containing FANCI DNA and RNA binding mutations failed to support ID2 monoubiquitination (Figure 4B). Indeed, the same set of DNA, RNA substrates used in EMSA all stimulated the monoubiquitination of FANCD2 in a manner consistent with their individual binding affinity (Figures 4C and 4D). As expected, R-loop substrates to which ID2 bound with high affinity stimulated FANCD2 monoubiquitination robustly (Figures 4E and 4F).

DISCUSSION

While the FA pathway has a well-known role in sensing and orchestrating the repair of DNA crosslinks, its recently described role in suppression or resolution of R-loops remains to be defined. In this study, we report findings of RNA and R-loop binding by ID2 and provide mechanistic insights of FA pathway activation in response to transcription stress. First, we show FANCD2 colocalizes with R-loops in cells and is able to bind ssRNA with G-rich sequence and R-loop structures with high affinity. Furthermore, our data provide experimental support that ID2 binds R-loops via recognition of the inherent single-stranded nucleic acid species, including the displaced DNA and unhybridized RNA. Importantly, we have shown that RNA and R-loop binding can stimulate robust ID2 monoubiquitination.

Based on these findings and current knowledge, we propose a model in which ID2 is recruited not only to sites of DNA damage such as the ICL, but also to transcription perturbed sites with accentuated R-loops (Figure 4G). We speculate that ID2 recruitment to R-loop, like to ICL, is dependent on the activity of the FA core complex proteins. This notion is supported by recent findings that showed assembly of the majority of FANCD2 foci requires R-loops (García-Rubio et al., 2015) and FANCD2 foci diminish upon depletion of FANCA (Okamoto et al., 2018). Furthermore, we suggest that ID2 monoubiquitination facilitates R-loop resolution via the recruitment of downstream factors. These factors could include nucleases, helicases, topoisomerases, RNA binding proteins, and DNA repair factors that have been implicated in either preventing the accumulation of pathological R-loops or in the resolution of these structures (Santos-Pereira and Aguilera, 2015). For example, the tumor suppressor BRCA1 has been shown to interact with the putative RNA helicase Senataxin (SETX), and the resulting complex localizes to R-loops (Hatchi et al., 2015). BRCA2 and its binding partner TREX also colocalize to R-loops (Bhatia et al., 2014). Finally, FANCM has been proposed to eliminate R-loops via its RNA:DNA helicase activity (Hodson et al., 2018; Schwab et al., 2015). We note that chromatin recruitment of BRCA1, BRCA2, and RAD51 upon replication stress depends monoubiquitination of ID2 (García-Higuera et al., 2001; Taniguchi et al., 2002; Tripathi et al., 2016), which emphasizes a likely pivotal role of ID2 monoubiquitination in the recruitment and/or activation of factors that function in R-loop metabolism.

One unexpected finding of our study is that human ID2 has only weak affinity for RNA:DNA hybrids without any single-stranded characteristics. We speculate this may avoid inappropriate activation of the FA pathway in human cells where R-loops serve an important function in gene expression regulation (Skouri-Stathaki and Proudfoot, 2014). While the exact mechanism of such substrate recognition of human ID2 warrants further investigation, differences in primary protein sequences of FANCI and FANCD2 among vertebrates (Joo et al., 2011) support the notion that human ID2 might have evolved additional layers of nucleic acid recognition modules that are absent in ID2 homologs of lower vertebrates.

The preference of ID2 for guanine-rich DNA and RNA species likely reflects the higher propensity for such sequences to form secondary structures, which hints at the possibility that ID2 has specific affinity for these structures in its nucleic acid targets. Indeed, FANCD2 has been shown to associate with fragile genomic regions with strong and large transcriptional units (Madireddy et al., 2016; Okamoto et al., 2018), where genomic secondary structures stemming from aberrant DNA replication and transcription intermediates accentuate (Wilson et al., 2015). We also speculate that both FANCI and FANCD2 subunits contribute to the recognition of guanine-rich sequences based on the observation of similar binding activity using chicken proteins (Figure S3).

Our data add credence to the premise that any failure to prevent the formation of pathological R-loops or in the resolution of these structures contributes to the genomic instability seen in FA cells. Based on the phenotypic similarities shared by FA and other genetic diseases marked by aberrant RNA metabolism such as dyskeratosis congenita (DC), Diamond Blackfan anemia (DBA), and Shwachman-Diamond syndrome (SDS) (Dokal and

(F) Comparison of ID2 ubiquitination efficiency using different R-loop substrates. R-loops with guanine-containing ssDNA and ssRNA tail supported ID2 ubiquitination with the highest efficiency. For (C)–(F), percentage of ubiquitinated FANCD2 is calculated by the intensity of the Ub-D2 band over total intensity of the Ub-D2 and D2 bands.

(G) Model of R-loop binding and its stimulated monoubiquitination of FANCI-FANCD2 in response to DNA damage-induced transcriptional perturbation.

570 Cell Reports 26, 564–572, January 15, 2019
Vulliamy, 2010), it is possible that R-loop formation leads to genomic instability seen in other cancer predisposition syndromes as well.

**STAR METHODS**

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **CONTACT FOR REAGENT AND RESOURCE SHARING**
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - Mammalian cell lines
  - Insect cell lines
- **BACTERIA STRAINS**
- **METHOD DETAILS**
  - Damage At RNA Transcription (DART) assay
  - Immunofluorescence staining of RNA:DNA hybrids
  - Cell survival assay
  - Nondenaturing bisulfite sequencing
  - Biotinylated RNA pull-down
  - Immunoblot analysis
  - Protein expression and purification
  - EMSA
  - ID2 In vitro ubiquitination assay
- **QUANTIFICATION AND STATISTICAL ANALYSIS**

**SUPPLEMENTAL INFORMATION**

Supplemental Information includes four figures and one table and can be found with this article online at https://doi.org/10.1016/j.celrep.2018.12.084.

**ACKNOWLEDGMENTS**

This study was supported by the NIH grants R01CA168635, R01CA20123, P30CA54174, and U54DK106857 to G.M.K. and P.S. and GM118833 to L.L. Z.L. was supported by the Damon Runyon Cancer Research Foundation (DRG-[2253-16]), Y.T. was supported by China Scholarship Council, and T.R. by DNA damage distinguishes all Fanconi anemia complementation groups from DNA damage repair. Mol. Cell 57, 636–647.

**AUTHOR CONTRIBUTIONS**

G.M.K., P.S., Z.L., F.L., Y.T., J.L., and L.L. conceived the experiments and analyzed the data. Z.L., F.L., Y.T., X.C., J.L., S.L., T.R., A.G., and N.B.C. performed the experiments. Z.L., G.M.K., and P.S. wrote the manuscript.

**DECLARATION OF INTERESTS**

The authors declare no competing interests.

Received: June 11, 2018
Revised: November 15, 2018
Accepted: December 18, 2018
Published: January 15, 2019

**REFERENCES**

Bhatia, V., Barroso, S.I., Garcia-Rubio, M.L., Tumini, E., Herrera-Moyano, E., and Aguilera, A. (2014). BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. Nature 517, 362–365.

Cantor, S.B., Bell, D.W., Ganesan, S., Kass, E.M., Drapkin, R., Grossman, S., Wahrer, D.C., Sgroi, D.C., Lane, W.S., Haber, D.A., and Livingston, D.M. (2001). BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function. Cell 105, 149–160.

Chen, X., Wilson, J.B., McChesney, P., Williams, S.A., Kwon, Y., Longerich, S., Marriott, A.S., Sung, P., Jones, N.J., and Kupfer, G.M. (2014). The Fanconi anemia proteins FANCd2 and FANCj interact and regulate each other’s chromatin localization. J. Biol. Chem. 289, 25774–25782.

Dokal, I., and Vulliamy, T. (2010). Inherited bone marrow failure syndromes. Haematologica 95, 1236–1240.

Garcia-Higuera, I., Taniguchi, T., Ganesan, S., Meyn, M.S., Timmers, C., Hejna, J., Grompe, M., and D’Andrea, A.D. (2001). Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway. Mol. Cell 7, 249–262.

Garcia-Rubio, M.L., Pérez-Calero, C., Barroso, S.I., Tumini, E., Herrera-Moyano, E., Rosado, I.V., and Aguilera, A. (2015). The Fanconi anemia pathway protects genome integrity from R-loops. PLoS Genet. 11, e1005674.

Godthelp, B.C., Wiegant, W.W., Waisfisz, Q., Medhurst, A.L., Anwert, F., Joenje, H., and Zdzienicka, M.Z. (2006). Inducibility of nuclear Rad51 foci after DNA damage distinguishes all Fanconi anemia complementation groups from D1/BRCA2. Mutat. Res. 594, 39–48.

Hatchi, E., Skourtis-Stathaki, K., Ventz, S., Pinello, L., Yen, A., Kamieniarz-Gdula, K., Dimitrov, S., Pathania, S., McKinney, K.M., Eaton, M.L., et al. (2015). BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair. Mol. Cell 57, 636–647.

Hodson, C., Rourke, J.J., van Tweest, S., Murphy, W.J., Dunn, E., and Deans, A.J. (2018). FANCM-family branchpoint translocases remove co-transcriptional R-loops. bioRxiv, arXiv:248161. https://doi.org/10.1101/248161.

Huang, Y., Leung, J.W., Lowery, M., Matsushita, N., Wang, Y., Shen, X., Huang, D., Takata, M., Chen, J., and Li, L. (2014). Modularized functions of the Fanconi anemia core complex. Cell 158, 849–859.

Joo, W., Xu, G., Persky, N.S., Smogorzewska, A., Rudge, D.G., Buzovetsky, O., Elledge, S.J., and Pavletich, N.P. (2011). Structure of the FANCI-FANCd2 complex: Insights into the Fanconi anemia DNA repair pathway. Science 333, 312–316.

Kim, Y., Y. Lach, F.P., Desetty, R., Hanenberg, H., Auerbach, A.D., and Smogorzewska, A. (2011). Mutations of the SLX4 gene in Fanconi anemia. Nat. Genet. 43, 142–146.

Lan, L., Nakajima, S., Wei, L., Sun, L., Hsieh, C.L., Sobol, R.W., Bruchez, M., Van Houten, B., Yasui, A., and Levine, A.S. (2014). Novel method for site-specific induction of oxidative DNA damage reveals differences in recruitment of repair proteins to heterochromatin and euchromatin. Nucleic Acids Res. 42, 2330–2345.

Liang, C.C., Li, Z., Lopez-Martinez, D., Nicholson, W.V., Vénien-Bryan, C., and Cohn, M.A. (2016). The FANCd2-FANCJ complex is recruited to DNA interstrand crosslinks before monoubiquitination of FANCd2. Nat. Comm. 7, 12142.

Longerich, S., San Filippo, J., Liu, D., and Sung, P. (2009). FANCJ binds branched DNA and is monoubiquitinated by UBE2T-FANCL. J. Biol. Chem. 284, 23182–23186.

Longerich, S., Kwon, Y., Tsai, M.S., Hlaing, A.S., Kupfer, G.M., and Sung, P. (2014). Regulation of FANCJ and FANCJ monoubiquitination by their interaction and by DNA. Nucleic Acids Res. 42, 5657–5670.

Madiredy, A., Kosiyatratkul, S.T., Boisvert, R.A., Herrera-Moyano, E., Garcia-Rubio, M.L., Gerhardt, J., Vuono, E.A., Owen, N., Yan, Z., Olson, S., et al. (2016). FANCd2 Facilitates Replication through Common Fragile Sites. Mol. Cell 64, 388–404.

Nalepa, G., and Clapp, D.W. (2018). Fanconi anaemia and cancer: An intricate relationship. Nat. Rev. Cancer 18, 168–185.

Nguyen, H.D., Yadav, T., Giri, S., Saez, B., Graubert, T.A., and Zou, L. (2017). Functions of replication protein A as a sensor of R loops and a regulator of RNAseH1. Mol. Cell 65, 832–847.

Okamoto, Y., Iwasaki, W.M., Kugou, K., Takahashi, K.K., Oda, A., Sato, K., Kobayashi, W., Kawai, H., Sakasai, R., Takaori-Kondo, A., et al. (2018). Replication stress induces accumulation of FANCd2 at central region of large fragile genes. Nucleic Acids Res. 46, 2932–2944.
Rajendra, E., Oestergaard, V.H., Langevin, F., Wang, M., Dorman, G.L., Patel, K.J., and Passmore, L.A. (2014). The genetic and biochemical basis of FANCD2 monoubiquitination. Mol. Cell 54, 858–869.

Santos-Pereira, J.M., and Aguilera, A. (2015). R loops: New modulators of genome dynamics and function. Nat. Rev. Genet. 16, 583–597.

Sato, K., Toda, K., Ishiai, M., Takata, M., and Kurumizaka, H. (2012). DNA robustly stimulates FANCID2 monoubiquitylation in the complex with FANCI. Nucleic Acids Res. 40, 4553–4561.

Schwab, R.A., Nieminuszczy, J., Shah, F., Langton, J., Lopez Martinez, D., Liang, C.C., Cohn, M.A., Gibbons, R.J., Deans, A.J., and Niedzwiedz, W. (2015). The Fanconi anemia pathway maintains genome stability by coordinating replication and transcription. Mol. Cell 60, 351–361.

Skourti-Stathaki, K., and Proudfoot, N.J. (2014). A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. Genes Dev. 28, 1384–1396.

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

Taniguchi, T., Garcia-Higuera, I., Andreassen, P.R., Gregory, R.C., Grompe, M., and D’Andrea, A.D. (2002). S-phase-specific interaction of the Fanconi anemia protein, FANCID2, with BRCA1 and RAD51. Blood 100, 2414–2420.

Teng, Y., Yadav, T., Duan, M., Tan, J., Xiang, Y., Gao, B., Xu, J., Liang, Z., Liu, Y., Nakajima, S., et al. (2018). ROS-induced R loops trigger a transcription-coupled but BRCA1/2-independent homologous recombination pathway through CSB. Nat. Commun. 9, 4115.

Tripathi, K., Mani, C., Clark, D.W., and Palle, K. (2016). Rad18 is required for functional interactions between FANCID2, BRCA2, and Rad51 to repair DNA topoisomerase 1-poisons induced lesions and promote fork recovery. Oncotarget 7, 12537–12553.

van Twist, S., Murphy, V.J., Hodson, C., Tan, W., Swuec, P., O’Rourke, J.J., Heierhorst, J., Crismani, W., and Deans, A.J. (2017). Mechanism of ubiquitination and deubiquitination in the Fanconi anemia pathway. Mol. Cell 65, 247–259.

Wang, X., Andreassen, P.R., and D’Andrea, A.D. (2004). Functional interaction of monoubiquitinated FANCID2 and BRCA2/FANCID1 in chromatin. Mol. Cell. Biol. 24, 5850–5862.

Wei, L., Nakajima, S., Böhm, S., Bernstein, K.A., Shen, Z., Tsang, M., Levine, A.S., and Lan, L. (2015). DNA damage during the G0/G1 phase triggers RNA-templated, Cockayne syndrome B-dependent homologous recombination. Proc. Natl. Acad. Sci. USA 112, E3495–E3504.

Wilson, T.E., Artt, M.F., Park, S.H., Rajendran, S., Paulsen, M., Ljungman, M., and Glover, T.W. (2015). Large transcription units unify copy number variants and common fragile sites arising under replication stress. Genome Res. 25, 189–200.

Xia, B., Dorsman, J.C., Ameziane, N., de Vries, Y., Roomans, M.A., Sheng, Q., Pals, G., Errami, A., Gluckman, E., Llera, J., et al. (2007). Fanconi anemia is associated with a defect in the BRCA2 partner PALB2. Nat. Genet. 39, 159–161.

Yamamoto, K.N., Kobayashi, S., Tsuda, M., Kurumizaka, H., Takata, M., Kono, K., Jiricny, J., Takeda, S., and Hirota, K. (2011). Involvement of SLX4 in inter-strand cross-link repair is regulated by the Fanconi anemia pathway. Proc. Natl. Acad. Sci. USA 108, 6492–6496.

Yu, K., Chedin, F., Hsieh, C.L., Wilson, T.E., and Lieber, M.R. (2003). R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. Nat. Immunol. 4, 442–451.
## STAR METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| **Antibodies**      |        |            |
| Rabbit polyclonal anti-FANCD2 | Abcam | Cat# ab2187; RRID:AB_302885 |
| Rabbit polyclonal anti-FANCD2 | Bethyl Laboratories | Cat# A302-174A; RRID:AB_1659803 |
| Mouse monoclonal anti-FANCD2 | Santa Cruz Biotechnology | Cat# sc-20022; RRID:AB_2278211 |
| Rabbit polyclonal anti-FANCI | Bethyl Laboratories | Cat# A301-254A; RRID:AB_890616 |
| Rabbit polyclonal anti-RNase H1 | Abcam | Cat# ab229078 |
| Mouse monoclonal anti-Ku86 | Santa Cruz Biotechnology | Cat# sc-5280; RRID:AB_672929 |
| Mouse monoclonal S9.6 | Kerafast or purify from mouse BALB/c cells | Cat# ENH001; RRID:AB_2687463 |
| mouse monoclonal anti-HA | Roche | Cat# 11583816001; RRID:AB_514505 |
| Goat anti-Mouse IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 488 | Thermo Fisher | Cat# A-11001; RRID:AB_2534069 |
| Goat anti-Mouse IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 405 | Thermo Fisher | Cat# A-31553; RRID:AB_221604 |
| Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 488 | Thermo Fisher | Cat# A-11008; RRID:AB_143165 |
| Goat anti-Mouse IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 555 | Thermo Fisher | Cat# A-21422; RRID:AB_141822 |
| ECL Rabbit IgG, HRP-linked whole Ab (from donkey) | Amersham Health | Cat# NA934; RRID:AB_772206 |
| ECL Mouse IgG, HRP-linked whole Ab (from sheep) | Amersham Health | Cat# NA931; RRID:AB_772210 |
| **Bacterial and Virus Strains** |        |            |
| Rosetta(DE3) pLysS | Novagen | Cat# 70956 |
| NEB 5-alpha Competent E. coli (High Efficiency) | NEB | Cat# C29871 |
| **Chemicals, Peptides, and Recombinant Proteins** |        |            |
| Mitomycin C (MMC) | Sigma | Cat# M4287 |
| Crystal Violet | Sigma | Cat# C0775 |
| Sodium bisulfite | Sigma | Cat# 243973 |
| High Capacity Streptavidin Agarose | Pierce | Cat# 20357 |
| 3x Flag peptide | APExBIO | Cat# A6001 |
| anti-FLAG M2 Affinity Agarose Gel | Sigma | Cat# A2220 |
| Ni-NTA Agarose | QIAGEN | Cat# 30250 |
| Imidazole | Sigma | Cat# I5513 |
| Recombinant human FANCI | This paper | N/A |
| Recombinant human FANCI-FANCD2 | This paper | N/A |
| Recombinant chicken FANCI | This paper | N/A |
| Recombinant chicken FANCD2 | This paper | N/A |
| Recombinant chicken FANCI-FANCD2 | This paper | N/A |
| Recombinant human FANCL | Longerich et al., 2014 | N/A |
| Recombinant human UBE2T | Longerich et al., 2014 | N/A |
| UBE1 | Boston Biochem | Cat# E-305 |
| HA-ubiquitin | Boston Biochem | Cat# U-110 |
| RNaseOUT Recombinant Ribonuclease Inhibitor | Thermo Fisher | Cat# 10777019 |
| **Critical Commercial Assays** |        |            |
| Bradford Protein Assay Kit | Pierce | Cat# 23200 |
| Supersignal West Pico Kit | Pierce | Cat# 34078 |

(Continued on next page)
CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Gary M Kupfer (gary.kupfer@yale.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mammalian cell lines

FANCD2 deficient PD20 cells (GM16756, Coriell Institute, male) was used as the initial FA-D2 cell line to construct derivative PD20+pMMP-vector (FA-D2), PD20+pMMP-Flag-K561R (FA-D2+K561R) ubiquitination mutant, and PD20+pMMP-Flag-FANCD2 (FA-D2+FANCD2) wild-type cells. Cells were maintained at 37°C in a 5% CO2 incubator and cultured in Dulbecco’s modified Eagle medium (DMEM, Thermo Fisher) containing 10% fetal bovine serum (FBS, Biowest) and 1% Pen-Strep (Thermo Fisher). Cells stably transfected with pEGFP-N1-RNASEH1 (gift of James Manley, Columbia University) and pEGFP-N1 empty vector were cultured in DMEM containing 10% FBS, 1% Pen-Strep, and 400 μg/ml G418 (Sigma). U2OS TRE cells (female) used in DART assay were cultured in dark at 37°C in a 5% CO2 incubator using DMEM (Lonza) with 10% FBS (GEMINI Bio-product). Mouse hybridoma cell line (HB-8730, ATCC, sex unknown) was used to purify S9.6 antibody.

Insect cell lines

Sf9 insect cells (Thermo Fisher) were used to produce initial baculovirus and High Five insect cells (Thermo Fisher) were used as the host for protein expression.

BACTERIA STRAINS

Rosetta(DE3) pLysS was used as the strain for protein expression and DH5a (NEB) was used as the strain for plasmid cloning.
METHOD DETAILS

Damage At RNA Transcription (DART) assay
The DART system has been described in previous publications (Lan et al., 2014; Wei et al., 2015). Briefly, the KillerRed (KR) or mCherry fusion protein expression vectors TA-KR, tetR-KR, TA-Cherry or tetR-Cherry in pBroad3 backbone were transfected into genetically engineered U2OS TRE cells in 35 mm glass-bottom dishes (MatTek). To introduce RNase H1 in the DART system, HA-RNase H1 (WT) or HA-RNase H1 (D210N) (Nguyen et al., 2017) was co-transfected with TA-KR. To induce DNA damage by KillerRed, cells that had been cultured in the dark for 36-48 hours were exposed to light (15 W Sylvania cool white fluorescent bulb) for 30 minutes and allowed to recover for 1 hour. For S9.6 foci staining, cells were then treated with 75 mM KCl (pre-warmed to 37°C) for 10 minutes and fixed by freshly made methanol-acetic acid (3:1) overnight, followed by 3 washes in PBS and steaming in TE buffer (10 mM Tris-HCl, 2 mM EDTA, pH 9.0) on a 95°C heating block for 20 minutes to expose the antigen. Next, cells were cooled, washed 3 times in PBS and blocked by 5% bovine serum albumin (BSA) (Sigma) in 0.1% PBS-Tween (PBST) for 1 hour at room temperature. S9.6 primary antibody (ENH001, Kerafast) and goat anti-mouse Alexa Fluor 488 or 405 secondary antibody (A-11001 or A-31553, Thermo Fisher) were used. For staining of endogenous FANCD2 foci in DART system, cells were rinsed in PBS and fixed in 4% paraformaldehyde (Affymetrix) for 15 minutes at room temperature. They were then washed 3 times by PBS, permeabilized by 0.2% Triton X-100 in PBS for 10 minutes, washed 3 times in PBS, and blocked by 5% BSA in 0.1% PBS-T for 30 minutes at room temperature. Primary FANCD2 antibody (sc-20022, Santa Cruz) was diluted in blocking buffer and incubated with cells overnight at 4°C. The cells were then washed 3 times in 0.05% PBS-T and incubated with goat anti-mouse Alexa Fluor 488 secondary antibody for 1 hour at room temperature. Finally, they were washed 3 times by 0.05% PBS-T and imaged in PBS. For co-staining with S9.6 and anti-FANCD2 in TA-KR cells, the method was the same as above for staining of S9.6 foci, and rabbit FANCD2 (A302-174A, Bethyl) antibody was added, followed by goat anti-rabbit Alexa Fluor 488 secondary antibody (A-11008, Thermo Fisher). For co-staining of FANCD2 and HA-RNase H1, the method was the same as staining of endogenous FANCD2 foci, and mouse HA antibody (11583816001, Roche) was added, followed by goat anti-mouse Alexa Fluor 405 secondary antibody. Images were acquired using the Olympus FV1000 confocal microscopy system. The foci intensity was directly measured by arbitrary unit from the ImageJ 1.50i software.

Immunofluorescence staining of RNA:DNA hybrids
For staining of RNA:DNA hybrids in FA-D2 cell lines, S9.6 antibody was purified from the mouse BALB/c hybridoma cells (gift of Tae Hoon Kim, Yale University) using NAb Protein A/G Spin Kit (Thermo Fisher) according to the manufacturer’s instructions. Cells were plated in 8 chamber slides, grown to 50% confluence, and then treated with 500 nM MMC (Sigma) for the indicated time. Slides were then rinsed in PBS, fixed in 4% paraformaldehyde in PBS for 5 minutes, rinsed in PBS again, permeabilized in 0.5% Triton X-100 in PBS for 5 minutes, rinsed in PBS once more, and then blocked in PBS with 0.5% BSA, 0.1% NP-40 and 0.5% BSA was added to the slides, followed by an overnight incubation at 4°C. Slides were then washed three times in 0.1% PBS-T, and goat anti-mouse Alexa Fluor 555 secondary antibody (A-21422, Thermo Fisher) diluted 1:1000 was added to slides for 2 hours at room temperature. After washing 3 times in 0.1% PBS-T and 2 more times in PBS, slides were mounted using DAPI Vectashield Hard-Set (Vector Laboratories) and images were captured in a TE2000-E Eclipse inverted fluorescent microscope (Nikon). Velocity software (Perkin Elmer) was used to quantify immunofluorescence.

Cell survival assay
Cells were seeded into 6-well plates at 3,000 cells per well and incubated at 37°C for 24 hours. After the addition of MMC (Sigma), cell were incubated at 37°C for 5 days. Surviving cells were fixed in 10% methanol/10% acetic acid and stained with crystal violet solution (1% in methanol) for 5 minutes at room temperature on a rocker. Plates were then rinsed in water and allowed to dry, and crystal violet dye was extracted with 0.1% SDS in methanol for 1 hour at room temperature. Dye concentration was measured in a microplate reader (BioTek) by absorbance at 595 nm.

Nondenaturing bisulfite sequencing
R-loops were detected by using nondenaturing bisulfite treatment as previously reported (Yu et al., 2003). Briefly, genomic DNA was extracted by QIAGEN DNA Easy Kit and then treated with sodium bisulfite (Sigma) overnight at 37°C and purified. For the PCR reactions, the constitutively active glyceraldehyde 3-phosphate dehydrogenase (GAPDH) gene was amplified using the following primer pair, forward: 5'-GTCAAGGCTGAGAACGGGAA-3' reverse: 5'-CTCCCCACATCACCCCTCTA-3'. The PCR products were then gel purified and Sanger-sequenced.

Biotinylated RNA pull-down
Cells were lysed in buffer containing 50 mM Tris pH 7.5, 150 mM NaCl, 1% Triton X-100, 1 mM MgCl2, and protease and phosphatase inhibitors (2 μg/ml aprotinin, 1 μg/ml pepstatin, 2 μg/ml leupeptin, 1 mM PMSF, 1 mM sodium pyrophosphate, and 1 mM Na3VO4). Cell suspensions were sonicated briefly and cleared by centrifugation, and protein concentration was determined by the Bradford assay (Pierce). Various amount of 60-mer biotinylated oligos (Keck Facility, Yale University) (5 μg in Figure S2A, 10 μg in Figure S2B).
and 2 ug in Figure S2C) were added to reaction buffer (10 mM Tris pH 7.4, 1 mM DTT, 50 mM NaCl, 1 mM MgCl₂, 0.5 mM EDTA, 5% glycerol, and 1% DEPC) containing ~2 mg lysis proteins. After 2-hour incubation at 4°C, streptavidin resin (Pierce) was added, followed by another 2-hour incubation at 4°C. The resin was collected by centrifugation and washed 4 times in the reaction buffer. SDS loading buffer was then added to resin and heated at 95°C for 5 minutes. Samples were subjected to standard immune blotting procedure described below.

Immunoblot analysis

Nitrocellulose membranes (Bio-Rad) were blocked in 5% milk in 0.1% PBS-T for 1 hour at room temperature, then incubated overnight at 4°C with diluted primary antibody. Membranes were washed three times in 0.1% PBS-T then incubated at room temperature for 1 hour with secondary antibody diluted 1:2000 in 0.5% milk in 0.1% PBS-T. Finally, membranes were washed five times in 0.1% PBS-T and developed by enzyme-linked chemiluminescence using SuperSignal West Pico Kit (Pierce). GelQuant.NET (http://biochemlabsolutions.com/GelQuantNET.html) software was used to quantify intensity of bands in immunoblots. For immunoblotting, rabbit polyclonal FANCD2 antibody was obtained from Abcam (ab2187), rabbit polyclonal FANC2 antibody was obtained from Bethyl (A301-254A), rabbit polyclonal RNase H1 antibody was obtained from Abcam (ab229078), and mouse monoclonal Ku86 antibody was obtained from Santa Cruz (sc-5280). ECL anti-rabbit (NA934) and anti-mouse (NA931) IgG HRP-linked secondary antibodies were purchased from Sigma.

Protein expression and purification

The human ID2 complex was expressed in insect cells and purified as previously described (Longerich et al., 2014; Longerich et al., 2009) with modifications. In brief, baculoviruses containing 6xHis-FANCI or 3xFlag-FANCD2 were generated by the Bac-to-Bac Baculovirus Expression System (Thermo Fisher) following the manufacturer’s manual. Table S1. Duplex DNA, R-loop and D-loop substrates were prepared by annealing of DNA and/or RNA oligos (IDT). The assay was conducted according to our published protocol (Longerich et al., 2009). Nuclei acid substrates used in this study are reported in Table S1. Duplex DNA, R-loop and D-loop substrates were prepared by annealing of DNA and/or RNA oligos (IDT). Following electrophoresis in a native 12% acrylamid gel at room temperature in 1X TBE buffer, substrates were eluted from gel slices by dialysis and concentrated by centrifugation using Micro Bio-Spin Chromatography Columns (Bio-Rad). All the reactions contained 10 mM 32P-labeled substrates. The RNase inhibitor RNaseOUT (Thermo Fisher) was added to reactions (0.16 units/ul, 1:250) containing DNA substrates. Reactions were incubated at 30°C for 5 minutes to allow the primer extension to proceed.

EMSA

The assay was conducted according to our published protocol (Longerich et al., 2009). Nuclei acid substrates used in this study are reported in Table S1. Duplex DNA, R-loop and D-loop substrates were prepared by annealing of DNA and/or RNA oligos (IDT). Following electrophoresis in a native 12% acrylamid gel at room temperature in 1X TBE buffer, substrates were eluted from gel slices by dialysis and concentrated by centrifugation using Micro Bio-Spin Chromatography Columns (Bio-Rad). All the reactions contained 10 mM 32P-labeled substrates. The RNase inhibitor RNaseOUT (Thermo Fisher) was added to reactions (0.16 units/ul, 1:250) containing DNA substrates. Reactions were incubated at 30°C for 5 minutes to allow the primer extension to proceed.
**ID2 in vitro ubiquitination assay**

Reactions containing ID2 complex, HA-ubiquitin, E1, UBE2T, FANCL, and the indicated nucleic acid ligand were carried out as previously described (Longerich et al., 2014), with RNaseOUT (Thermo Fisher) being included (0.16 units/ul, 1:250) to prevent RNA degradation. Reactions were incubated overnight (~16 hours) in Figure 4C and for 7 hours in Figures 4D–4F prior to loading on 4%–15% Mini-PROTEAN TGX Precast Protein Gels (Bio-Rad). The GelQuant.NET (http://biochemlabsolutions.com/GelQuantNET.html) software was used to quantify Ub-FANCD2 and unmodified FANCD2.

**QUANTIFICATION AND STATISTICAL ANALYSIS**

Quantification and statistical analysis were done using Prism 7.05 (GraphPad). For DART assay, 50 cells were analyzed, and statistical significance was determined by unpaired t test. Error bars represent SEM, *p < 0.05, **p < 0.01, ***p < 0.001. For staining of RNA:DNA hybrids in FA-D2 and derivative cells, Volocity software (Perkin Elmer) was used to quantify S9.6 immunofluorescence. S9.6 signal of each time point was normalized to the signal at 0 hour, and percentages of change were plotted. Approximately 10 cells were analyzed for each cell line, and statistical significance was determined by unpaired t test. Error bars represent SD. *p < 0.05, **p < 0.01, ***p < 0.001. For survival assay, crystal violet dye concentration was measured in a microplate reader (BioTek) by absorbance at 595 nm and normalized to untreated control to calculate the percentage of survival. Means were calculated from 3 replicates, error bars represent SD. For EMSA assay, signal of the non-shifted substrate in each lane was quantified using GelQuant.NET (http://biochemlabsolutions.com/GelQuantNET.html) software in order to calculate the percentage of the shifted substrates. Means were calculated from 3 replicates, error bars represent SD. For ID2 in vitro ubiquitination assay, signal of the ub-FANC2 and FANC2 bands were quantified using GelQuant.NET (http://biochemlabsolutions.com/GelQuantNET.html) software to calculate the percentage of ub-FANC2.
Supplemental Information

Binding of FANCI-FANCD2 Complex to RNA and R-Loops Stimulates Robust FANCD2 Monoubiquitination

Zhuobin Liang, Fengshan Liang, Yaqun Teng, Xiaoyong Chen, Jingchun Liu, Simone Longerich, Timsi Rao, Allison M. Green, Natalie B. Collins, Yong Xiong, Li Lan, Patrick Sung, and Gary M. Kupfer
Figure S1. FANCD2 monoubiquitination is required to prevent formation of pathogenic R-loops, related to Figure 1. (A) FA-D2+empty vector, +FANCD2-K561R monoubiquitination-dead mutant, and +wild-type FANCD2 complemented cells treated with MMC for the indicated times were subjected to immunofluorescence microscopy to visualize RNA:DNA hybrids using the S9.6 antibody. Increased S9.6 signal along MMC treatment was seen in the 2 mutant cell lines. (B) Quantification of S9.6 intensity in (A). S9.6 signal of each time point was normalized to the signal at 0 hour, and percentages of change were plotted. n=10 cells, unpaired t-test, error bars represent SD. *P < 0.05, **P < 0.01, ***P < 0.001. (C) Lysates from FA-D2 mutant cells+vector control, +wild-type FANCD2, or +GFP-RNase H1 were immunoblotted for FANCD2 and RNase H1 to confirm expression. (D) Overexpression of RNase H1 reduced RNA:DNA hybrids immunostained by the S9.6 antibody in cells. (E) Overexpression of RNase H1 partially rescued FA-D2 cell survival in MMC. Means were calculated from 3 replicates, error bars represent SD. (F) Representative result of bisulfite Sanger sequencing revealed C to T conversion in the highly expressed GAPDH gene in FA-D2 +vector and +FANCD2(K561R) mutant cells but not in cells corrected with wild type FANCD2.
Figure S2. FANCI-FANCD2 in cell extracts binds RNA and prefers guanine-rich sequence, related to Figure 2.

(A) Immunoblots were performed against FANCD2 and Ku86 after streptavidin pull-down of biotinylated single stranded DNA and RNA oligos with a random sequence. Oligos were incubated with extracts prepared from FA-D2 mutant cells transduced with either wild-type FANCD2 or FANCD2-K561R ubiquitination dead mutant, treated with or without MMC. MMC treatment of FA-D2+FANCD2 cells induced FANCD2 monoubiquitination. Monoubiquitinated FANCD2 (Ub-D2) can be separated from the non-ubiquitinated form (D2) in PAGE. (B) Immunoblot was performed against FANCI and Ku86 after streptavidin pull-down of the indicated biotinylated oligos incubated with extracts prepared from same cells as in (A). (C) Immunoblot was performed against FANCD2 after streptavidin pull-down of the indicated biotinylated RNA polymer species incubated with extracts prepared from FA-D2 mutant cells transduced with either vector control, wild-type FANCD2, or FANCD2-K561R treated with or without MMC. r: ribonucleotide; C: cytosine; U: uracil; G: guanine; A: adenine. (D) Purified FANCI(KKEE)/D2 and FANCI(R1285X)/D2 heterodimers were analyzed by SDS-PAGE and Coomassie Blue staining. (E) EMSA showing binding of ssDNA by recombinant wild-type ID2, FANCI(KKEE)/D2 or FANCI(R1285X)/D2. (F) EMSA showing binding of ssRNA by recombinant wild-type ID2, FANCI(KKEE)/D2 or FANCI(R1285X)/D2.
Figure S3. Binding of RNA substrates by recombinant human FANCI, chicken FANCI, FANCD2 and ID2, related to Figure 2.

(A) EMSA showing binding of ssRNA, dsRNA and RNA:DNA hybrids by recombinant human FANCI. (B) Quantification of shifted nucleic acid substrates in (A). (C) EMSA showing binding of ssRNA, dsRNA and RNA:DNA hybrids by recombinant chicken FANCI. (D) Quantification of shifted nucleic acid substrates in (C). (E) EMSA showing binding of ssRNA, dsRNA and RNA:DNA hybrids by recombinant chicken FANCD2. (F) Quantification of shifted nucleic acid substrates in (E). (G) EMSA showing binding of ssRNA, dsRNA and RNA:DNA hybrids by recombinant chicken ID2 complex. (H) Quantification of shifted nucleic acid substrates in (G). (I) EMSA showing binding of dsDNA with variable guanine content by recombinant human ID2 complex. (J) Quantification of shifted nucleic acid substrates in (I). (K) EMSA showing binding of dsRNA with variable guanine content by recombinant human ID2 complex. (L) Quantification of shifted nucleic acid substrates in (K). For (B), (D), (F), (H), (J) and (L), means were calculated from 3 replicates, error bars represent SD.
Figure S4. ID2 binds R-loop and D-loop structures with similar affinity, related to Figure 3.

(A) EMSA showing binding of D-loop(GGC) and R-loop(GGC) substrates by recombinant ID2. The “GGC” annotation represents the base composition of the bubble portion of the substrates, meaning annealing product of G-rich top DNA strand, G-rich middle RNA strand, and C-rich bottom DNA strand. (B) Quantification of shifted substrates in (A). (C) EMSA showing binding of D-loop(GAT) and R-loop(GAT) substrates by recombinant ID2. (D) Quantification of shifted substrates in (C). (E) Competitive EMSA to compare ID2 binding of guanine-rich ssRNA and R-loop using P32-labeled (hot) ssRNA(G28) and R-loop(G30). (F) Quantification of shifted substrates in (E). (G) Competitive ID2 EMSA using either P32-labeled (hot) R-loop(G30) or ssRNA(28) under titration of non-labeled (cold) competitor substrates. (H) Quantification of shifted substrates in (G). For (B), (D), (F) and (H), means were calculated from 3 replicates, error bars represent SD.
Table S1. Nucleic acid substrates used in this study and their oligonucleotide sequences, related to Figure 2-4.

| Nucleic acid substrate | Oligo type | Oligo sequence (5′-3′) | Length | Associated figures |
|------------------------|------------|------------------------|--------|-------------------|
| Poly-dC | RNA | CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC...