Supplemental Information

1. Seven Supplemental Figures

Figure S1. Further characterization of the patient cell lines. Related to Figure 1.
Figure S2. Homologous recombination in cells expressing RAD51 T131P. Related to Figure 2.
Figure S3. RPA activation in RA2630 cells following MMC treatment. Related to Figures 3 and 4.
Figure S4. DNA2 and WRN2 dependent activation of RPA in RA2630 cells after MMC treatment. Related to Figure 3.
Figure S5. Inhibition of MRE11 with Mirin is not equivalent to depletion of MRE11 with siRNAs. Related to Figure 3.
Figure S6. In vitro ATPase and DNA binding activity of T131P RAD51 mutant. Related to Figure 5.
Figure S7. In vitro recombinase activity of T131P RAD51 mutant. Related to Figure 5 and 6.

2. Supplemental Figure Legends

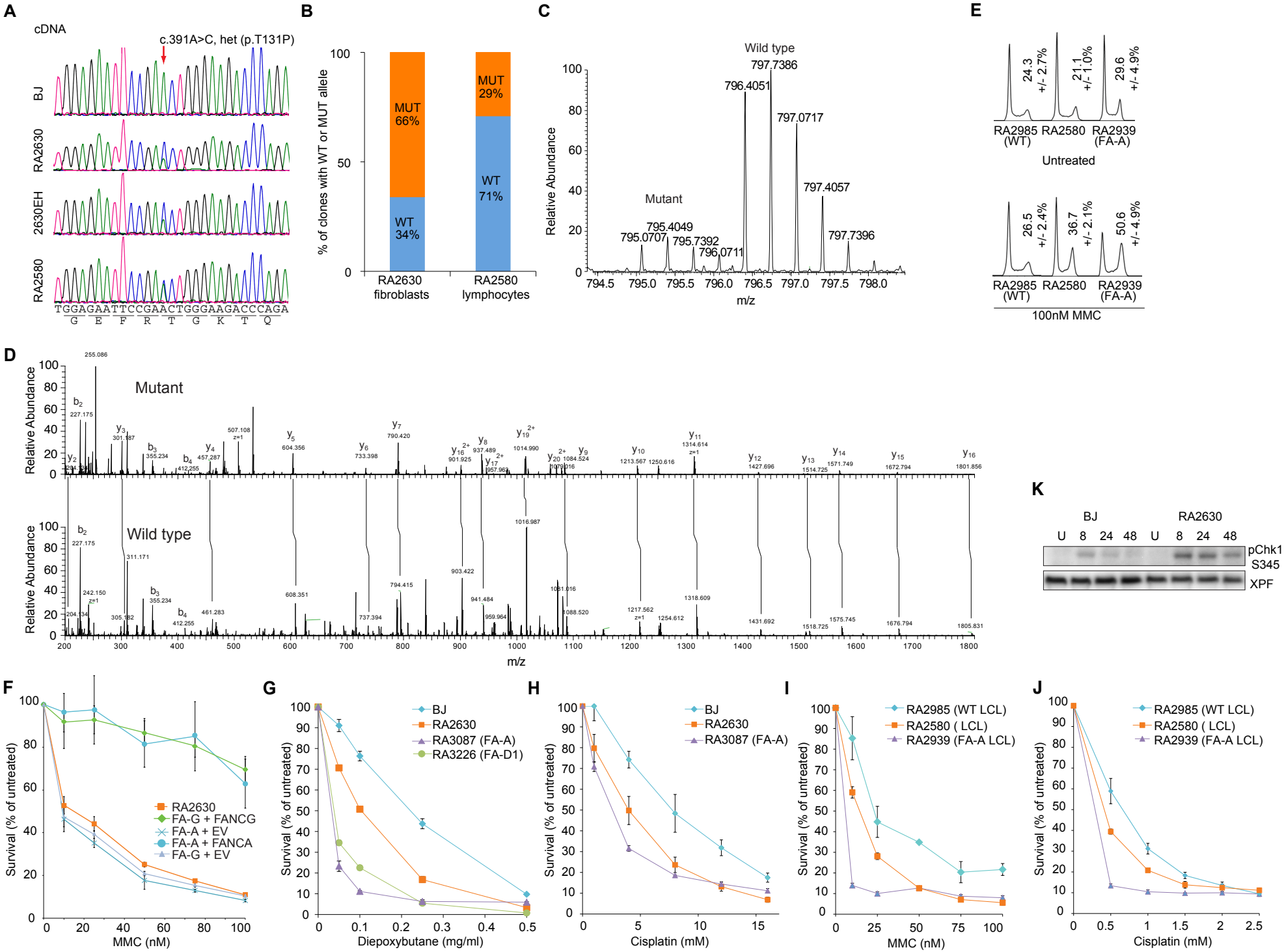
3. Four Supplemental Tables

Table S1. Patient phenotype. Related to Figure 1.
Table S2. Chromosome breakage analysis. Related to Figure 1.
Table S3. Chromosome breakage analysis in BJ cells overexpressing WT RAD51 or T131P RAD51 with and without diepoxybutane (DEB) or mitomycin C (MMC) treatment. Related to Figure 1.
Table S4. Sequences of siRNAs, RT-PCR, and mutagenesis primers. Related to Figure 1, 2, and 3.

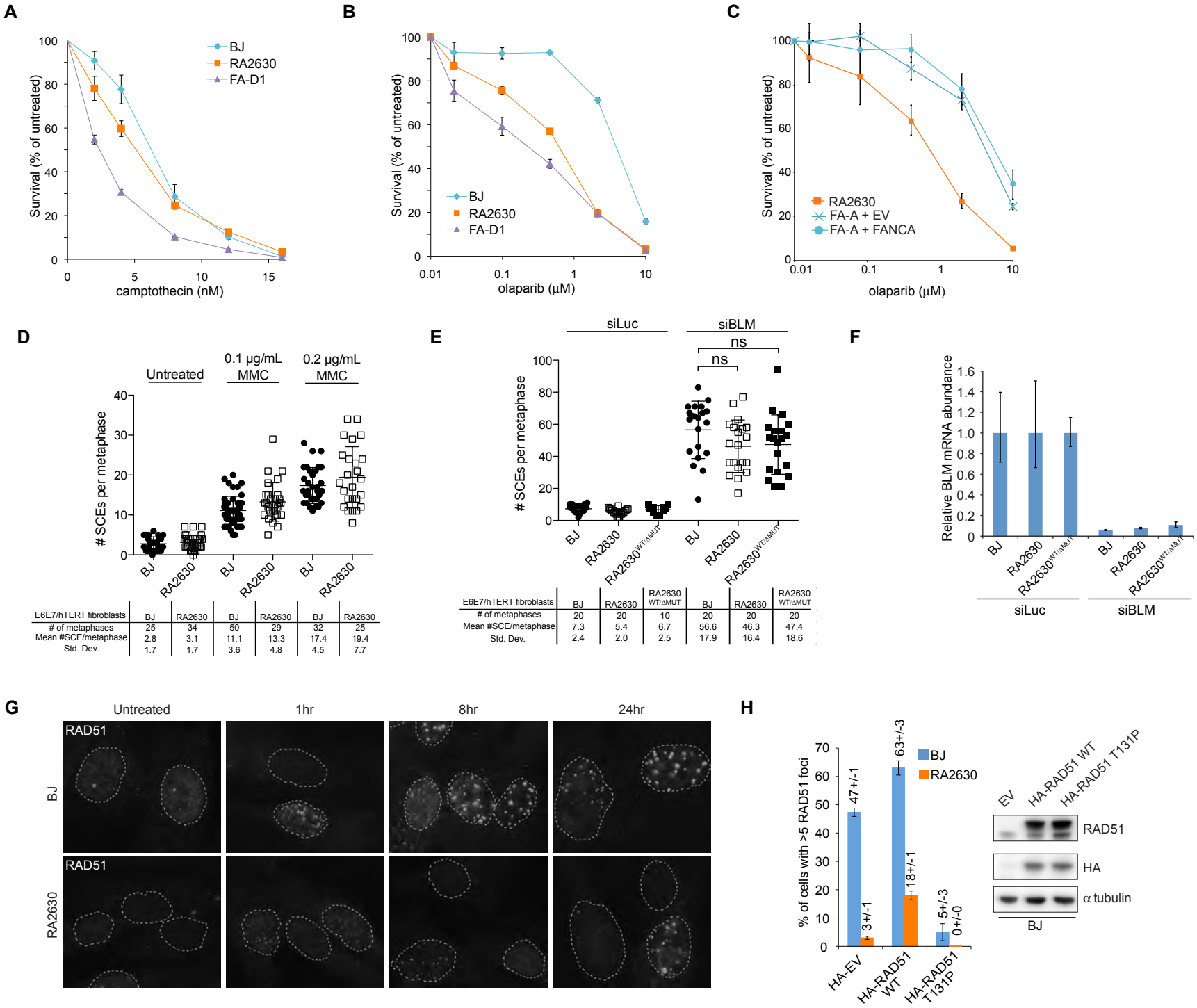
4. Supplemental Experimental Procedures

5. Supplemental References

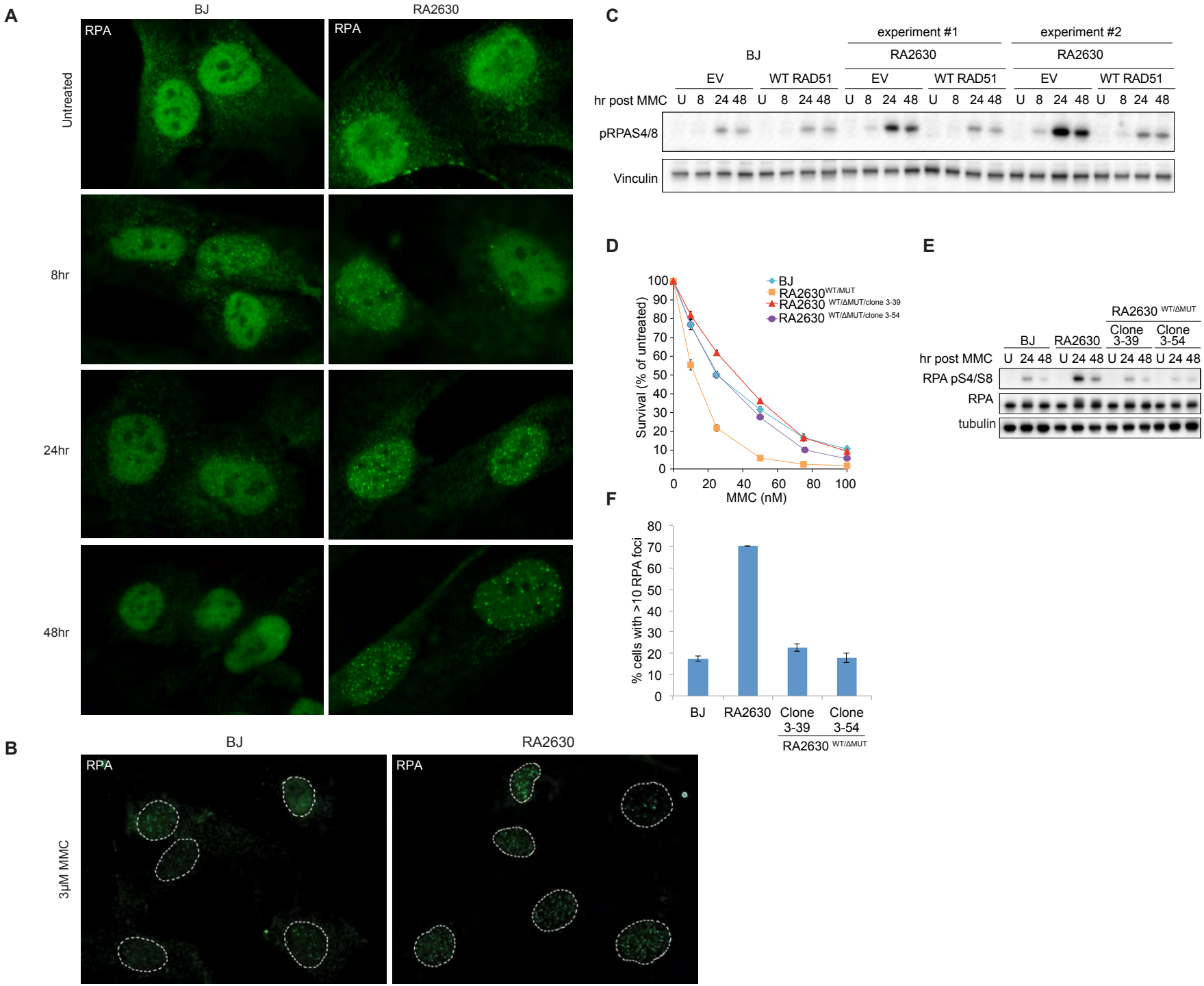
Wang et al. Supplemental Figure 1



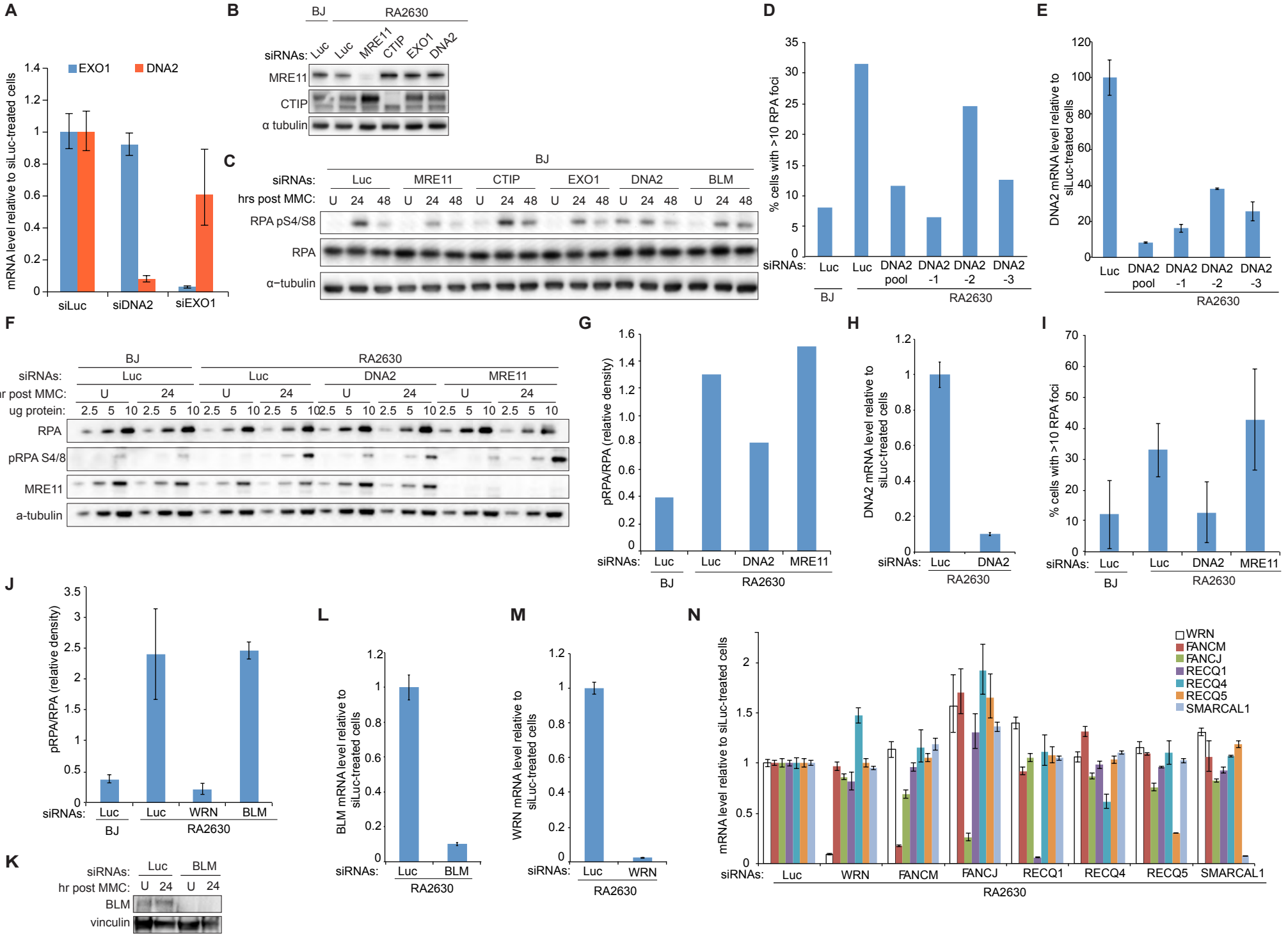
Wang et al. Supplemental Figure 2



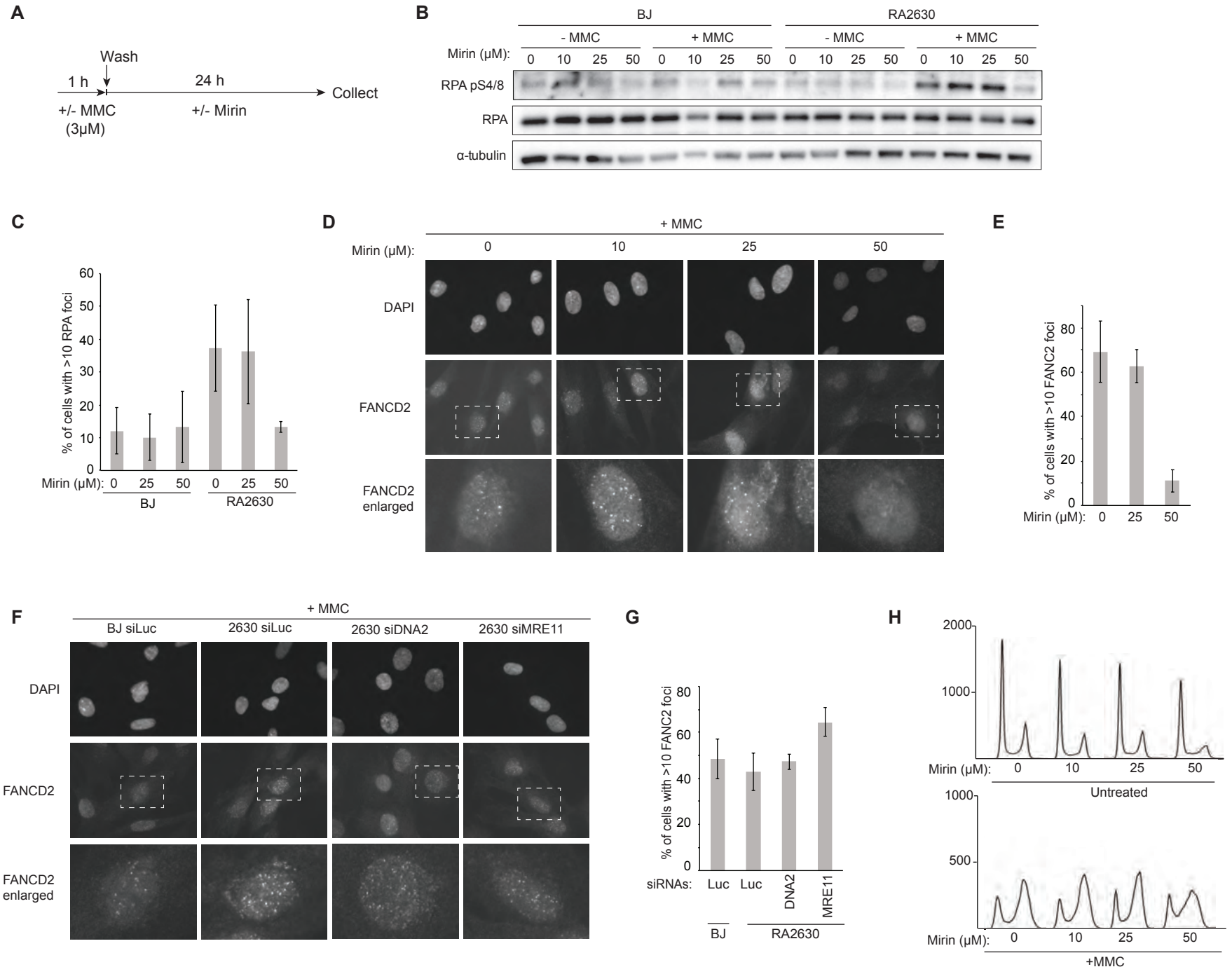
Wang et al. Supplemental Figure 3



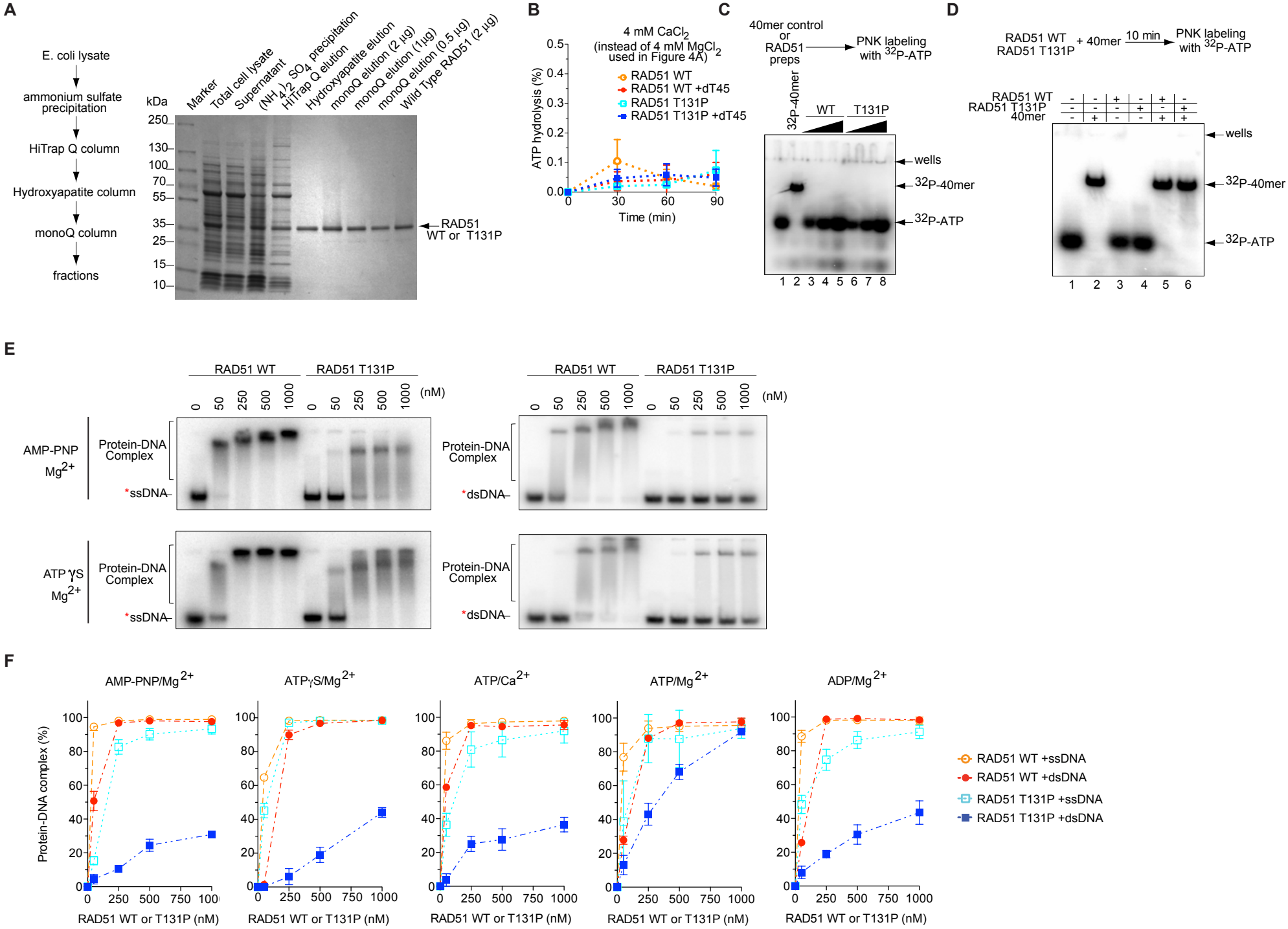
Wang et al. Supplemental Figure 4



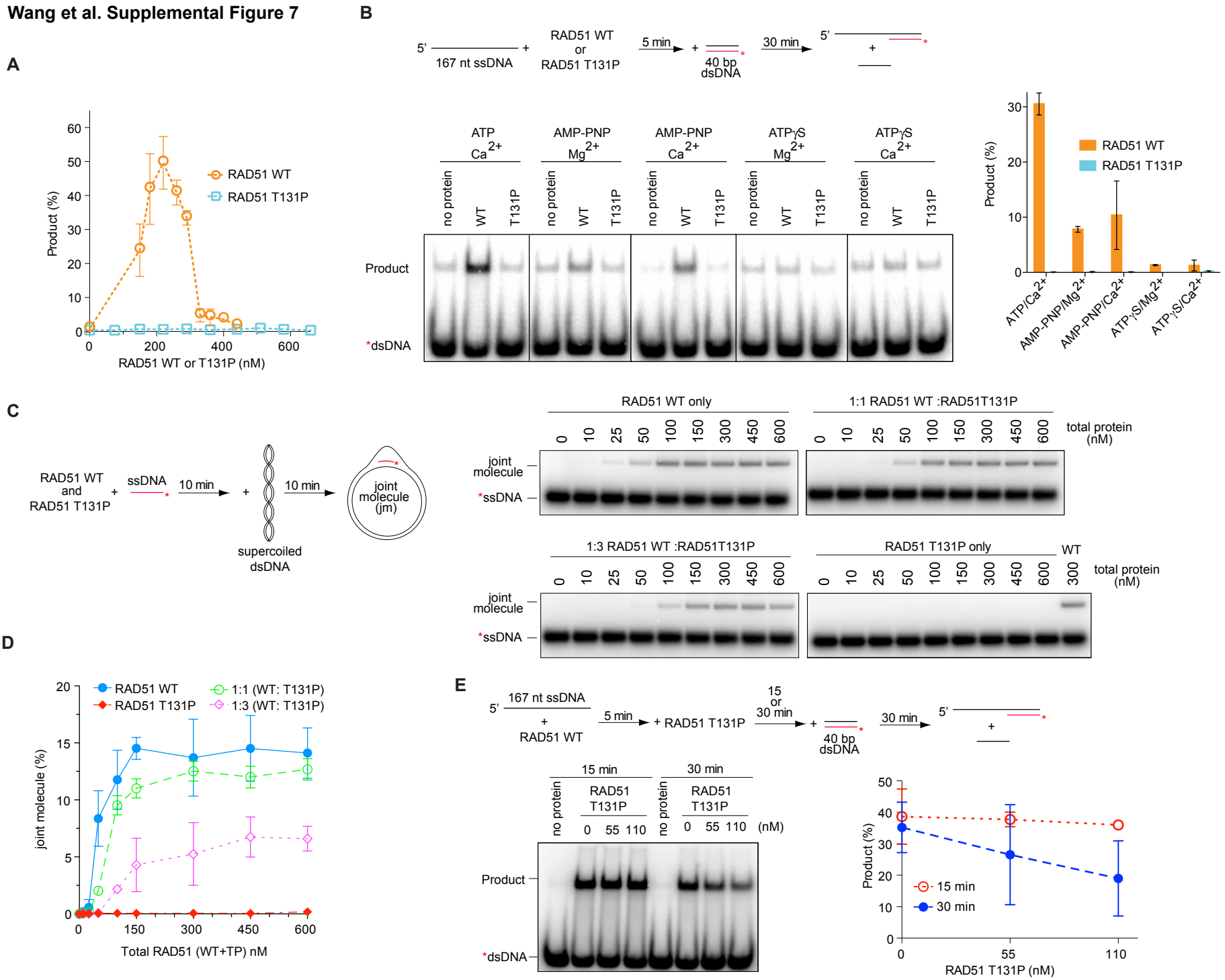
Wang et al. Supplemental Figure 5



Wang et al. Supplemental Figure 6



Wang et al. Supplemental Figure 7



Supplemental figure legends

Figure S1. Further characterization of the patient cell lines. **Related to Figure 1** (A) Sequencing traces of RAD51 cDNAs derived from BJ, primary patient fibroblasts (RA2630), immortalized and transformed patient fibroblasts (RA2630EH) and patient LCLs (RA2580). (B) Topo cloning analysis of the levels of the WT and mutant RAD51 mRNA transcripts in the patient fibroblasts and lymphocytes. (C) Summed MS spectra over the entire elution period for the mutant (LLQGGIETGSITEM(ox)FGEFRPGK) (m/z 795.0707) and the wild type (LLQGGIETGSITEM(ox)FGEFRTGK) (m/z 796.4051) triply charged peptides of RAD51 measured in fibroblast sample. (D) Annotated MS/MS spectra of the triply charged mutant and wild type peptides: LLQGGIETGSITEM(ox)FGEFRPGK (upper panel) and LLQGGIETGSITEM(ox)FGEFRTGK (lower panel), respectively. Fragment ions y1 through y20 was matched for both peptides. When comparing the tandem MS spectra of the mutant and the wild type peptides y-ion, fragment ions y3 to y20 displayed a median mass shift of 3.996 Da (s.d.: 0.068) compared to a theoretical fragment ion mass shift of 3.9949 Da corresponding to a threonine to proline mutation at peptide position 3. In accordance with the sequence, the fragment ions: y2, b2, b3 and b4 did not display this mass shift when comparing the two. (E) Cell cycle profiles of indicated LCLs following MMC treatment. (F-J) Survival of indicated fibroblasts (F to H) and LCLs (I and J) following treatment with increasing doses of indicated DNA cross-linking agents. Error bars: s.d. (n=3). (K) Immunoblotting analysis of CHK1 phosphorylation on Serine 345 residue (pCHK1 S345) at the indicated times following one-hour 3μM MMC treatment. U denotes untreated samples.

Figure S2. Homologous recombination in cells expressing RAD51 T131P. Related to Figure 2. (A-C) Survival of indicated fibroblasts following treatment with increasing doses of camptothecin (A) or olaparib (B and C). FA-A + EV are *FANCA* deficient patient cells (RA3087) expressing empty vector; FA-A + FANCA: *FANCA* deficient patient cells complemented with FANCA overexpression; FA-G + EV: *FANCG* (RA3291) patient cells expressing empty vector; FA-G + FANCG: *FANCG* patient cells complemented with FANCG overexpression. (D) Sister chromatid exchange assay in BJ and RA2630 after treatment with indicated doses of MMC. No statistically significant differences were identified between BJ and RA2630 cells in any of the examined conditions. (E) Sister chromatid exchange assay in BJ, RA2630, and RA2630$^{WT/\Delta MUT}$ (also see Fig. 4 in the main text) after depletion of BLM helicase. No statistically significant differences (ns) were identified between studied cell lines. (F) qRT-PCR analysis of BLM expression in control (siLuc) and BLM-depleted cells. Error bars: s.d. (n=3) (G) RAD51 foci formation following ionizing radiation (IR). Dashed lines mark outline of nuclei. Panel shows examples of immunofluorescence images quantified in Fig. 2F in the main text. (H) Quantification of RAD51 foci formation eight hours after 12Gy IR treatment in BJ cells stably transduced with empty vector (EV) or vector encoding HA-FLAG tagged WT or T131P RAD51 cDNA under CMV promoter. Error bars: s.d. (n=3). Levels of RAD51 overexpression were determined by anti-RAD51 and anti-HA immunoblotting.

Figure S3. **RPA activation in RA2630 cells following MMC treatment. Related to Figure 3 and 4.** (A) RPA foci formation following MMC treatment. BJ and RA2630 cells were treated with 3μM MMC for one hour then incubated in drug-free media for the indicated times before being fixed and analyzed for RPA foci formation. (B) Cells described in (A) were triton pre-extracted before being fixed. (C) Suppression of RPA phosphorylation by expression of RAD51 WT in RA2630 cells treated with MMC.

Two experiments are shown (#1 and #2). (D-F) Cell survival (D), reversion of RPA phosphorylation (E), and reversion of RPA foci formation (F) after MMC treatment in an additional clone (clone 3-54) obtained using CRISPR/Cas9 in RA2630 cells. This clone has a 47-nt deletion in the mutant allele of RAD51 predicted to lead to loss of function of that allele. Clone 3-39 is the clone shown in the main Figure 4.

Figure S4. DNA2 and WRN2 dependent activation of RPA in RA2630 cells after MMC treatment. Related to Figure 3. (A) qRT-PCR analysis of mRNA levels of EXO1 and DNA2 in RA2630 cells treated with indicated siRNA in experiment shown in Fig. 3E and F in the main text. Error bars: s.d. (n=3). (B) Immunoblotting analysis showing protein depletion level by indicated siRNAs in experiment shown in Fig. 3E and F in the main text. (C) RPA phosphorylation after MMC treatment in BJ cells transfected with indicated siRNAs was determined by immunoblotting. This is a control for the experiment in Figure 3E in the main text. (D) RPA foci formation after MMC treatment in cells transfected with a pool of siRNAs or individual siRNA against DNA2. (E) qRT-PCR analysis of mRNA levels in RA2630 cells used in the experiment shown in panel D. Error bars: s.e.m (n=3). (F) Immunoblotting analysis of RPA phosphorylation in cells treated with indicated siRNA. (G) Quantification of pRPA/RPA relative signal density of 5 μg extract from panel F. (H) qRT-PCR analysis of mRNA levels of DNA2 in RA2630 cells treated with indicated siRNA. Error bars: s.em. (n=3). (I) RPA foci formation after MMC treatment in cells transfected with indicated siRNAs. Blinded samples from three independent experiments were counted. Errors: s.d. (n=3). (J) Quantification of pRPA/RPA relative signal density of 5 μg shown in Fig 3G in the main text. (K) Immunoblotting analysis showing protein depletion level by indicated siRNAs in experiment shown in Fig. 3G and H in the main text. (L) qRT-PCR analysis of mRNA levels of BLM in RA2630 cells treated with indicated siRNA in

experiment shown in Fig. 3G and H in the main text. Error bars: s.em. (n=3). (M) qRT-PCR analysis of mRNA levels of WRN in RA2630 cells treated with indicated siRNA in experiment shown in Fig. 3G and H in the main text. Error bars: s.e.m. (n=3). (N) qRT-PCR analysis of mRNA levels in RA2630 cells treated with indicated siRNA in experiment shown in Fig. 3I in the main text. Error bars: s.d. (n=3).

Figure S5. Inhibition of MRE11 with Mirin is not equivalent to depletion of MRE11 with siRNAs. Related to Figure 3. (A) Schematic of MMC and mirin treatment in the experiments shown in this Figure. (B) Immunoblotting analysis of RPA phosphorylation following treatment performed as shown in panel A. (C) Quantification of RPA foci formation following treatment described in panel A. Blinded samples from two independent experiments were counted. Error bars: s.d. (n=3). (D) Representative images of FANCD2 foci following treatment performed as shown in panel A. (E) Quantification of FANCD2 foci from the experiment shown in panel D. Blinded samples from three independent experiments were counted. Errors: s.d. (n=3). (F) Representative images of FANCD2 at 24 hours following one-hour 3μM MMC treatment in cells transfected with indicated siRNAs. (G) Quantification of FANCD2 foci from the experiment shown in panel F. Blinded samples from three independent experiments were counted. Errors: s.d. (n=3). (H) Cell cycle profiles following treatment performed as shown in panel A.

Figure S6. In vitro ATPase and DNA binding activity of T131P RAD51 mutant. Related to Figure 5. (A) Purification of recombinant RAD51 proteins. RAD51 was expressed in *E. coli* and the soluble fraction of cell extract was ammonium sulfate precipitated. The precipitate containing RAD51 was then dissolved and further purified by using sequential HiTrap Q, hydroxyapatite and MonoQ columns. (B)

ATP hydrolysis by WT or RAD51 T131P with or without ssDNA, dT45, was examined in the presence of $CaCl_2$ in place of $MgCl_2$. (C) Purified RAD51 proteins at increasing concentrations were checked for the presence of DNA or RNA by radioactive labeling using ^{32}P-ATP and T4 polynucleotide kinase (PNK). A 40-mer DNA oligonucelotide was used as a control for the labeling reaction. (D) RAD51 WT and T131P were pre-incubated with 40-mer oligo at 37ºC for 10 minutes prior to starting PNK labeling reaction. Mock labeling (no DNA, no protein) and labeling reaction with 12.5 pmoles of 40-mer were used as controls. (E) EMSA showing WT or RAD51 T131P in the presence of ATP or its analogues with $MgCl_2$ or $CaCl_2$. (F) Quantification of EMSA shown in panel E and other EMSA experiments performed in the presence of ATP or ADP and $CaCl_2$ or $MgCl_2$. Error bars: s.d. (n=3).

Figure S7. **In vitro recombinase activity of T131P RAD51 mutant. Related to Figure 5 and 6.** (A) Quantification of DNA strand exchange product formation shown in Figure 5D in the main text. (B) Assessment of DNA strand exchange by RAD51 WT and RAD51 T131P in the presence of indicated cofactors. Quantification is shown on the graph. Error bars: s.d. (n=3). (C) Joint molecule formation by RAD51 WT and RAD51 T131P proteins alone or pre-mixed at 1:1 or 1:3 ratios. (D) Quantification of product formation in panel C. (E) Assessment of the ability of RAD51 T131P mutant to interfere with DNA strand exchange mediated by WT RAD51 DNA filament. Experimental setup is show in a schematic on the top. RAD51 WT protein (220 nM) was pre-incubated with ssDNA for 5 minutes to allow RAD51 filament formation before adding different concentrations of RAD51 T131P. The reaction mixture was further incubated for 15 or 30 minutes before dsDNA was added. Reaction products were deproteinized and analyzed by PAGE using a 6% gel and 1X TAE buffer. Product formation from gels was quantified by densitometry and plotted on the right. Error bars: s.d. (n=3).

Table S1. Patient phenotype. Related to Figure 1

Phenotype	The subject was a 7lb 11oz product of a 39-week gestation born to a gravida 1 para 1 27-year-old female delivered by caesarian section due to prior caesarian section. At birth, she was noted to have bilateral radial ray defects with absent right thumb and right radial dysplasia, flattened left thenar eminence and small left thumb, microcephaly, microophthalmia, displaced left kidney in the pelvis with normal right kidney, type I Chiari malformation and tethered spinal cord. At later ages, audiometric evaluation in childhood revealed pure tone thresholds in the normal range in the right ear with moderate to severe hearing loss in the left ear at 250-750 hz ; CT temporal bones revealed normal cochlea, vestibule and semicircular canals but small round window niche on right and ossified round window niche on left. Endocrine evaluation including hypothalamic-pituitary axis was normal. Cardiology evaluation by ECG and echocardiogram was normal. Dental evaluation revealed congenital missing teeth # 6, 20, and 29 with mixed tooth morphology. Skeletal survey revealed mild scoliosis in the upper thoracic spine and absence of L4, L5 spinous processes. Bone age was slightly advanced of chronological age. Skin evaluation revealed two café au lait patches, freckling on the upper cheeks and nose, and scattered benign appearing nevi. Developmental milestones were slightly delayed - walking at 16 months and toilet trained at 4 years. IQ was reported as above average and there are no identified learning disabilities at 13 years old.
Family history	Polycystic ovaries, epilepsy. No early cancers or marrow failure.
Clinical breakage studies	Excessive chromosomal breakage of PHA stimulated lymphoblasts and skin fibroblasts in response to diepoxybutane (0.1 ug/mL) and mitomycin C (20 nmol/mL) was demonstrated, although at the lower range of that observed in Fanconi anemia. Notably, results were variable between three laboratories with some testing not diagnostic for Fanconi anemia. Tests for Nijmegen breakage syndrome and mutations for BRCA2 were negative.
Current hematological status	The patient has been followed at the University of Minnesota at least annually since age of 9 months (currently 13 years). Most recent blood counts revealed a hemoglobin 14.7, MCV 82, platelet counts 328,000, white blood count 7.2 with normal differential (absolute neutrophil count 3.7 and absolute lymphocyte count 2.8). Marrow aspirate revealed trilineage hematopoiesis with normal differentiation and biopsy demonstrated 60-70% cellularity; cytogenetics revealed 46 XX.

Table S2. Chromosome breakage analysis in the indicated cell lines with and without diepoxybutane (DEB) or mitomycin C (MMC) treatment. BJ foreskin fibroblasts from a healthy donor (ATCC) were used as a normal fibroblast control. RA3087 and RA3157 cell lines are Fanconi anemia patient cell line with biallelic mutation in *FANCA* gene. Primary cells are non-immortalized. Related to Figure 1.

	Cell line	Agent	Concentration (μg/mL for DEB, nM for MMC)	# Metaphases	# Total breaks*	# Chromatid breaks	# Radials	# Metaphases with breaks	# Breaks per metaphase**
Fibroblasts	BJ primary	none	NA	30	1	1	0	0	0.03
	RA3087 primary (FA-A)	none	NA	30	3	3	0	3	0.1
	RA2630 primary (RAD51 mutant)	none	NA	27	11	11	0	10	0.4
	BJ E6E7/hTERT	none	NA	30	3	3	0	3	0.1
	RA3087 E6E7/hTERT (FA-A)	none	NA	30	10	8	1	5	0.3
	RA2630 E6E7/hTERT (RAD51 mutant)	none	NA	20	3	3	0	3	0.2
	BJ primary	DEB	0.1	34	1	1	0	1	0.03
	RA3087 primary (FA-A)	DEB	0.1	14	41	25	8	13	2.9
	RA2630 primary (RAD51 mutant)	DEB	0.1	31	20	8	6	9	0.7
	BJ E6E7/hTERT	DEB	0.1	33	8	2	3	3	0.2
	RA3087 E6E7/hTERT (FA-A)	DEB	0.1	11	173	73	50	11	15.7
	RA2630 E6E7/hTERT (RAD51 mutant)	DEB	0.1	23	143	59	42	10	6.2
	BJ primary	MMC	50	30	2	2	0	2	0.07
	RA3087 primary (FA-A)	MMC	50	18	63	33	15	16	3.5
	RA2630 primary (RAD51 mutant)	MMC	50	25	36	18	9	14	1.4
	BJ E6E7/hTERT	MMC	50	25	9	9	0	8	0.4
	RA3087 E6E7/hTERT (FA-A)	MMC	50	25	123	65	29	23	4.9
	RA2630 E6E7/hTERT (RAD51 mutant)	MMC	50	25	145	83	31	17	5.8
Lymphoblasts (EBV transformed)	RA2907 (nonFA)	none	NA	20	2	2	0	1	0.1
	RA3157 (FA-A)	none	NA	20	11	11	0	8	0.6
	RA2580 (RAD51 mutant)	none	NA	30	2	2	0	2	0.07
	RA2907 (nonFA)	DEB	0.1	22	1	1	0	1	0.05
	RA3157 (FA-A)	DEB	0.1	17	130	64	33	14	7.7
	RA2580 (RAD51 mutant)	DEB	0.1	28	32	30	1	12	1.1
	RA2907 (nonFA)	MMC	50	20	6	4	1	4	0.3
	RA3157 (FA-A)	MMC	50	20	55	25	15	16	2.8
	RA2580 (RAD51 mutant)	MMC	50	21	19	15	2	7	0.9

* Total number of breaks including chromatid breaks and radial chromosomes

** Fanconi anemia range: 0.68-1.10 breaks per metaphase after DEB treatment

[1] Auerbach, A.D., *Diagnosis of fanconi anemia by diepoxybutane analysis.* Curr Protoc Hum Genet, 2003. **Chapter 8**: Unit 8.7.

Table S3. Chromosome breakage analysis in BJ cells overexpressing WT RAD51 or T131P RAD51 with and without diepoxybutane (DEB) or mitomycin C (MMC) treatment. Related to Figure 1.

BJ E6E7/hTERT	Agent	Concentration (μg/mL for DEB, nM for MMC)	# Metaphases	# Total breaks*	# Chromatid breaks	# Radials	# Metaphases with breaks	# Breaks per metaphase
+ empty vector	none	NA	30	1	1	0	1	0.03
+ WT RAD51	none	NA	30	21	17	2	6	0.7
+ T131P RAD51	none	NA	30	50	4	5	14	1.7
+ empty vector	DEB	0.1	25	9	9	0	2	0.4
+ WT RAD51	DEB	0.1	25	36	34	1	6	1.4
+ T131P RAD51	DEB	0.1	18	142	76	33	7	7.9
+ empty vector	MMC	50	25	4	2	1	3	0.2
+ WT RAD51	MMC	50	25	16	13	3	12	0.6
+ T131P RAD51	MMC	50	25	203	151	26	17	8.1

* Total number of breaks including chromatid breaks and radial chromosomes

Table S4. Sequences of siRNAs, RT-PCR, and mutagenesis primers. Related to Figures 1, 2, and 3.

siRNAs:	**Sense sequence**	**Company**
siLuciferase	CGUACGCGGAAUACUUCGA	Sigma
siRAD51-1	GGUAGAAUCUAGGUAUGCAtt	Ambion
siRAD51-2	CAGUGGUAAUCACUAAUCAtt	Ambion
siRAD51-3	CCAGCUCCUUUAUCAAGCAtt	Ambion
siMRE11-1	GAUAGACAUUAGUCCGGUUtt	Ambion
siMRE11-2	CCCGAAAUGUCACUACUAAtt	Ambion
siMRE11-3	CGACUGCGAGUGGACUAUAtt	Ambion
siCTIP-1	GUACAAGGUUUACAAGUAAtt	Ambion
siCTIP-2	GGAUCUGUCUGAUCGAUUUtt	Ambion
siCTIP-3	GGGUCUGAAGUGAACAAGAtt	Ambion
siEXO1-1	GCCUGAGAAUAAUAUGUCUtt	Ambion
siEXO1-2	CUUUUGAACAGAUCGAUGAtt	Ambion
siEXO1-3	GGCUAGGAAUGUGCAGACAtt	Ambion
siDNA2-1	CAUCCAAUAUUUUCCCGUAtt	Ambion
siDNA2-2	CCGUACAGGCAGCAAUUAAtt	Ambion
siDNA2-3	GUAACUUGUUUAUUAGACAtt	Ambion
siBLM-1	CCCACTACTTTGCAAGTAA	Ambion
siBLM-2	GGATGTTCTTAGCACATCA	Ambion
siBLM-3	GATATCTTCCAAAACGAAA	Ambion
siWRN_1	GGAGGGUUUCUAUCUUACUtt	Ambion
siWRN_2	CUGUAGCAAUUGGAGUAAAtt	Ambion
siWRN_3	CGAUGCUAGUGAUUGCUCUtt	Ambion
RT-PCR primers:	**sequence**	
DNA2_Fwd	GCTGTCCTGAGTGAAACTTTTAGG	
DNA2_Rev	CCTCATGGAGAACCGTACCA	
EXO1_Fwd	CTTTCTCAGTGCTCTAGTAAGGACTCT	
EXO1_Rev	TGGAGGTCTGGTCACTTTGA	
BLM_Fwd	TTTATCCTGATGCCGACTGG	
BLM_Rev	ACCCCAGGAGAAACACAGG	
WRN_Fwd	GATGTTGCCAATAAAAAGCTGA	
WRN_Rev	GTTTACCTAAGAGGTGTTTAACCAGAC	
Actin_Fwd	GCTACGAGCTGCCTGACG	
Actin_Rev	GGCTGGAAGAGTGCCTCA	
Mutagenesis primers:	**sequence**	
Rad51 a391c sense	TGTTTGGAGAATTCCGACCTGGGAAGACCCAGATC	
Rad51 a391c antisense	GATCTGGGTCTTCCCAGGTCGGAATTCTCCAAACA	

Supplemental Experimental Procedures

Subjects

Research subjects were accrued in the International Fanconi Anemia Registry under institutional review board protocol. Cell lines and genomic DNA samples were derived from the patients and family members with informed consents.

Whole exome sequencing

Proband and one of the parents were sequenced at the Broad Institute. Libraries for whole exome sequencing were constructed and sequenced on either an Illumina HiSeq 2000 or Illumina GA-IIX using 76 bp paired-end reads. Details of whole exome library construction have been detailed elsewhere(Fisher et al., 2011). Standard quality control metrics, including error rates, percentage passing filter reads, and total Gb produced, were used to characterize process performance before downstream analysis. The Illumina pipeline generates data files (BAM files) that contain the reads together with quality parameters. Output from Illumina software was processed by the "Picard" data processing pipeline to yield BAM files containing aligned reads (via MAQ or BWA, to the NCBI Human Reference Genome Build hg18) with well-calibrated quality scores (DePristo et al., 2011).

The other parents to complete the trio was sequenced at the New York Genome Center. Exome capture was performed using Agilent SureSelect Human All Exon V4 capture kit. 100 bp paired-end sequencing was done on Illumina HiSeq 2500. Sequence was aligned to human genome build GRCh37 using BWA (Burrows-Wheeler Aligner) (Li and Durbin, 2009). Duplicate reads were marked using Picard [http://picard.sourceforge.net]. GATK (Genome Analysis Toolkit) (DePristo et al., 2011) was used for base quality score recalibration (BQSR) and local realignment around indels to refine alignment artifacts around putative insertions or deletions. Variant discovery was performed in two steps - variant calling with GATK HaplotypeCaller followed by joint genotyping using GATK GenotypeGVCFs. The resulting

variant call set was refined using Variant Quality Score Recalibration (VQSR). The VQSR scores were used to define low quality variants for downstream processing. Variant annotation was performed using SnpEff(Cingolani et al., 2012), VCFtools(Danecek et al., 2011) and in-house software. The annotations include predictions of the effect of nucleotide change on protein sequence using snpEFF(Cingolani et al., 2012); Variant frequencies in different populations from 1000 Genomes project(Genomes Project et al., 2010), NHLBI GO Exome Sequencing Project (ESP) [http://evs.gs.washington.edu/EVS/], dbSNP 138(Sherry et al., 2001); Cross species conservation scores from PhyloP(Cooper et al., 2005), GERP(Davydov et al., 2010), PhastCons(Hubisz et al., 2011); functional prediction scores from Polyphen2(Adzhubei et al., 2010), SIFT(Kumar et al., 2009) and CADD(Kircher et al., 2014); Variant disease association from OMIM [http://omim.org], ClinVar [www.ncbi.nlm.nih.gov/clinvar]; Regulatory annotations from ENCODE(Consortium, 2004), Regulome(Boyle et al., 2012), ORegAnno(Griffith et al., 2008); Gene Ontology annotations for biological process, cellular component, molecular function(Ashburner et al., 2000). For consistency, the proband and parent not sequenced at NYGC were reanalyzed with the NYGC sequence analysis pipeline.

Analysis of variant calls in the family trio

The identified variants were first filtered using variant quality score recalibration (VQSR) with a cutoff of 99.8 for SNPs and 92 for indels. Next, we filtered out all variants present in dbSNP database. Concentrating on moderate and high impact effect variants(Cingolani et al., 2012) we have identified 58 unique variants (43 and 15 indels) in one of the parents, 137 unique variants (74 SNPs and 63 indels) in the other parent. 118 variants (81 SNPs and 37 indels) were present in the proband sample and one or both of the parents. 9 variants (3 SNPs and 6 indels) were *de novo* variants present in the proband only but in neither of the parents. Only two genes had two significant variants identified in them: CTBP2 (p.Thr973Ser/c.2917A>T, p.Asn956Lys/c.2868C>A) and HLA-DRB1 (p.Asn956Lys/c.2868C>A,

p.Gly154Ala/c.461G>C). For CTBP2, manual inspection of reads revealed that the call might have been false positive. Thus the analysis did not identify any candidates consistent with recessive inheritance, which is the usual inheritance of Fanconi anemia. We turned to the *de novo* mutations identified in the proband and absent from either parent. All of the indels identified as *de novo* were false positive by manual inspection of individual reads. Three of the *de novo* SNPs appeared true variants. These were FAHD2A p.Pro101Ala/c.301C>G, LIMK2 p.Met642Arg/c.1925T>G, and RAD51 Thr131Pro/c.391A>C.

FAHD2A (fumarylacetoacetate hydrolase domain containing 2A) is an enzyme that is presumed to be involved in catabolism of amino acids. A similar enzyme, fumarylacetoacetate hydrolase (FAH), mutated in hereditary tyrosinemia type I (OMIM 276700) (Grompe and al-Dhalimy, 1993) which is an autosomal recessive disorder is the last enzyme of tyrosine degradation. Lack of FAH causes progressive liver disease and death from cirrhosis or hepatocellular carcinoma at a young age (Mitchell et al., 1994). Amino acid altered by the *de novo* variant is outside of the hydrolase domain and the second allele did not contain a potential causative mutation. LIMK2 (LIM domain kinase 2) is a kinase, which by phosphorylating cofilin contributes to Rho-induced reorganization of the actin cytoskeleton (Maekawa et al., 1999). LIMK2 affects only the longer isoform of the gene (isoform1). Since neither FAHD2A nor LIMK2 variants were likely to be responsible for the patient cellular phenotype of abnormal DNA damage response, we focused on the strongest candidate, RAD51 variant Thr131Pro/c.391A>C.

RAD51 cDNA, mutagenesis and plasmids

RAD51 cDNA in pAL vector was a kind gift from the Pavletich laboratory (Memorial Sloan Kettering Cancer Center, NY). Mutagenesis of RAD51 was performed using multisite mutagenesis kit (Agilent)

using mutagenesis primers listed in Supplementary data Table 4. For HA-FLAG-tagged RAD51 expression in human cells, wildtype and mutant RAD51 cDNAs were PCR-amplified and cloned into pDONR223 vector by GATEWAY cloning using BP clonase (Invitrogen). pDONR223 derivatives encoding RAD51 were recombined with pMSCV retroviral vector or pCMV lentiviral vector using LR clonase (Invitrogen). For non-tagged RAD51 expression in human cells, RAD51 was PCR amplified and subcloned into MSCV-Hygro vector (Clontech) by sequence-ligation-independent cloning(Li and Elledge, 2007) using HpaI. For bacterial expression of RAD51, WT and mutant RAD51 cDNAs were subcloned into pET16b vector (Novagen) using restriction enzymes NcoI and BamHI.

Cell culture, transfection and viral transduction

Patient cell lines were transformed by HPV16 E6E7 expression and/or immortalized by expression of catalytic subunit of human telomerase (hTERT) as indicated. Patient fibroblasts and BJ foreskin fibroblasts (ATCC) were grown in 3% oxygen and maintained in Dulbecco Modified Eagle Medium supplemented with 15% (v/v) fetal bovine serum, non-essential amino acids, glutamax and 100 units of penicillin per ml, and 0.1 mg streptomycin per ml (all from Invitrogen).

Transfection of fibroblasts with siRNAs was performed using Lipofectamine RNAiMax (Invitrogen) as according to manufacturer's instruction with final siRNA concentration of 25nM. Sequence information of siRNAs used in this study is provided in Supplementary data Table 4. Viruses for transducing expression constructs into fibroblasts were first packaged in HEK293T cells using TransIT reagent (MirusBio) as according to manufacturer's protocol. Viral supernatants collected were used to infect fibroblasts in the presence of 4μg/mL polybrene. Cells were selected in the appropriate antibiotics for at least one week post viral infection.

GFP HDR assay

BJ and RA2630 fibroblasts were transfected with pDRGFP(Pierce et al., 1999) (Addgene plasmid 26475, originally from Jasin lab, MSKCC) using Amaxa® Cell Line Nucleofector® Kit R (Lonza) as according to manufacturer's protocol. Cells were than selected in 1μg/mL puromycin before clones with single copy DR-GFP substrate stably integrated were isolated. To assess the ability of cells to repair the defective GFP gene in the DR-GFP substrate(Pierce et al., 1999), the isolated clones with DR-GFP substrates were infected with AdNgus24i adenovirus expressing I-SceI endonuclease as previously described(Smogorzewska et al., 2007). Cells were harvested for FACS analysis of GFP expression 60 hours post infection.

Cell cycle, breakage and cell survival analyses

Analyses of cell cycle and chromosomal breakage following treating patient fibroblasts with DNA damaging agents were performed as previously described(Kim et al., 2013). For cell survival assays, cells were seeded overnight before treated with drugs at indicated concentrations. Cells were allowed to grow to near confluency and passaged once at appropriate ratio. Once cells reached near confluency after passaging, cell number was determined by using Z2 Coulter counter (Beckman Coulter). In the case of olaparib treatment, fresh drug was replenished every 48 hours. For cell survival assay following ionizing radiation, cells were first irradiated in suspension before plated out to grow. In the case of NU7026 pre-treatment for inhibiting DNA-PK_{cs}, cells were first pretreated with 1μM NU7026 for three hours before being harvested for irradiation. Cells were then plated out in media containing 1μM NU7026 for the duration of the survival assay.

Analysis of Sister Chromatid Exchanges in BJ and RA2630 cells

Protocol used was as previously described(Garner et al., 2013). Cells were cultured for 24 hr in medium containing 10 μg/mL BrdU (BD Pharmingen). Cells were then treated with 0.1 μg/mL or 0.2 μg/mL

MMC for one hour (Sigma Aldrich), washed once with PBS, and then cultured for an additional 24 hours with fresh medium containing 10 μg/mL BrdU before harvest and spreading of metaphases. Spreads were stained with 20 μg/mL Hoechst 33342 (Invitrogen) for 30 min, rinsed and mounted in PBS, exposed to 254 nm UV for 3 hr, incubated in 65ºC 2 x SSC for 2 hr, rinsed in room temperature 1 x GURR buffer (Invitrogen) and then stained with 8% Giemsa Karyomax (Invitrogen) diluted in GURR buffer for three minutes. Metasystems Metafer Slide Scanning Platform was used to capture metaphases with differentially stained chromosomes. One-way Anova with multiple comparisons was used to test statistical significance.

For sister chromatid exchange analysis of siRNA depleted cells, cells were transfected twice with a pool of BLM siRNAs. Media was supplemented with 10ug/mL BrdU after the second transfection and cultured for 24 hours. The cells were fed fresh medium containing BrdU, cultured for an additional 24 hours, and harvested and prepared for analysis as described above.

Immunoblotting and immunofluorescence

Whole cell extracts were prepared by directly lysing cell pellets in Laemmli sample buffer (4% SDS, 20% glycerol, 125 mM Tris-HCl pH 6.8) and vortexing samples at highest speed for 20 seconds. Samples were boiled at 95°C for 5 minutes before proteins were separated on 4-12% or 3-8% gradient gels (Invitrogen) by SDS-PAGE. Immunoblotting analyses were performed using antibodies listed below. Immunofluorescence samples were performed as previously described(Wang et al., 2011). In brief, cells seeded on coverslips were washed once with PBS and fixed in 3.7% (w/v) formaldehyde in PBS for 10 minutes at room temperature before permeablized with 0.5% (v/v) Triton in PBS for 5 minutes at room temperature. Coverslips were then blocked overnight with 5% (v/v) fetal bovine serum in PBS at 4°C. Primary antibody incubation was performed for 2 hours at room temperature before

coverslips were washed 3 times in PBS. Coverslips were then stained with secondary antibodies conjugated with Alexafluor 488 or 594 (Invitrogen) and washed 3 times with PBS before mounted in Vectashield (Vector Laboratories) containing DAPI. Where indicated, triton pre-extraction was performed before fixing by incubating cells for 5 minutes at room temperature with 0.5% (v/v) Triton in PBS. Cells were fixed and processed as above. Image analyses of the stained slides were performed using the Axio Observer.A1 fluorescence microscope (Carl Zeiss), equipped with a Plan- Apochromat 63× NA-1.4 oil objective, the AxioCam CCD camera, and the AxioVision Rel Version 4.7 software.

Antibodies used in this study were anti-RAD51 (clone SWE47, a kind gift from Steve West), anti-RAD51 (clone H9) (Santa Cruz), anti-RAD51 (Abcam), anti-CTIP, anti-RPA and anti-RPA pS4/S8 (Bethyl laboratories), anti-alpha Tubulin (Sigma), anti-MRE11 (a kind gift from John Petrini), anti-FANCA (Rockland), anti-FANCD2 (Novus biologicals), anti-FANCI (in house), and anti-HA (Covance).

Reverse transcription and quantitative real-time PCR

Total messenger RNA was extracted by using RNeasy plus kit (Qiagen) and reverse transcribed to synthesize cDNA by using SuperScript III Reverse Transcriptase (Invitrogen). The relative transcript levels of genes of interest were determined by quantitative real-time PCR using Platinum SYBR Green qPCR SuperMix-UDG (Invitrogen) and normalized against actin. The primers for specific gene of interest are listed in Supplementary data Table 4.

CRISPR/Cas9-mediated genome-editing

RA2630-Cas9 cells were generated by lentiviral transduction of pCW-Cas9 vector encoding dox-inducible Cas9 expression cassette(Wang et al., 2014) (plasmid 50661, Addgene, originally from Lander and Sabatini labs, MIT). Following infection, cells were selected in the presence of 1μg/mL puromycin.

Stable RA2630-Cas9 cells were then infected with lentivirus expressing CRISPR single guide RNA (sgRNA)(Sanjana et al., 2014) (modified from plasmid 52963, Addgene, originally from Zhang lab, MIT) against target sequence (5'-AGGCAACAGCCTCCACAGTA-3') in exon 3 of RAD51. Following selection in 50 μg/mL hygromycin, RA2630-Cas9 cells with sgRNA stably integrated were treated with 500 ng/mL doxycyclin for 48 hours before single cells were plated out in 96 well plates. Isolated clones were screened for CRISPR-mediated genome editing at target region by sequencing of genomic DNA. TOPO cloning of RAD51 cDNA PCR products was performed using Zero Blunt TOPO PCR cloning kit (Invitrogen) to identify RAD51 transcripts expressed in clones with genome editing.

Protein purification

Wild type (WT) RAD51 and T131P mutant proteins (Hilario et al., 2009), BRCA2 (Jensen et al., 2010) and RPA(Walther et al., 1999) were purified as previously described. The concentrations of both RAD51 proteins and of BRCA2 were determined using extinction coefficients at 280 nm of 12,800 $M^{-1}cm^{-1}$ and 365,160 $M^{-1}cm^{-1}$, respectively. The concentration of RPA was determined by Bradford assay using bovine serum albumin as a standard.

ATPase assay

ATPase activity was measured by release of $^{32}P_i$ from [γ-^{32}P] ATP. The assay mixture (10 μL) contained 20 mM TrisHCl (pH7.5), 0.5 mM ATP, 200 nCi [γ-^{32}P] ATP, 1 mM dithiothreitol (DTT), 4 mM $MgCl_2$, 3 μM (nucleotides (nts)) ssDNA (dT_{45}), and 1 μM RAD51 WT or T131P protein. The reaction mixture was incubated at 37 °C and, at indicated time points, 1 μL of each reaction was spotted onto polyethylene amine thin layer chromatography plates (EMD Chemicals). The separation of $^{32}P_i$ from [γ-^{32}P] ATP was performed in buffer containing 0.8 M LiCl and 1 M formic acid. Plates were exposed on PhosphorImager screens, and the percentage release of $^{32}P_i$ was quantified using ImageQuant software (Molecular Dynamics).

^{32}P-labeling of nucleic acids in protein preparations

To detect the possible presence of contaminating nucleic acids in the RAD51 protein preparations, wild type and T131P mutant proteins (2 μg) were treated with T4 polynucleotide kinase and [γ-^{32}P] ATP. The reaction mixture was incubated for 60 min at 37° C. For the control involving preformed RAD51-ssDNA complexes, WT and T131P were pre-incubated with TK-1 (40-mer; 0.5 μM molecules) at 37° C for 10 min prior to the labeling reaction. The products were analyzed by polyacrylamide gel electrophoresis (PAGE) using a 12% gel and 1X TAE (40 mM TrisOAc (pH 8.0), and 1 mM EDTA) buffer, run at 60 V for 60 minutes at room temperature. The gel was then dried and exposed to a PhosphorImager screen. The screen was scanned on a Molecular Dynamics Storm 840 PhosphorImager.

Electrophoretic mobility shift assay

Oligonucleotide substrates were obtained from IDT (Ultramers) and were purified by polyacrylamide gel electrophoresis. The following oligonucleotides were utilized: TK-167-mer (5'-CTG CTT TAT CAA GAT AAT TTT TCG ACT CAT CAG AAA TAT CCG TTT CCT ATA TTT ATT CCT ATT ATG TTT TAT TCA TTT ACT TAT TCT TTA TGT TCA TTT TTT ATA TCC TTT ACT TTA TTT TCT CTG TTT ATT CAT TTA CTT ATT TTG TAT TA TCC TTA TCT TAT TTA-3'); TK-Oligo1 (5'-TAA TAC AAA ATA AGT AAA TGA ATA AAC AGA GAA AAT AAA G-3'); TK-Oligo2 (5'-CTT TAT TTT CTC TGT TTA TTC ATT TAC TTA TTT TGT ATT A-3'). The ssDNA substrate was TK-167-mer that was radiolabeled with ^{32}P at the 5'-end. The dsDNA was generated by first radiolabeling TK-Oligo1 with ^{32}P at the 5'-end and annealing it to TK-Oligo2. The assay mixture (10 μL) contained 25 mM TrisOAc (pH 7.5), 1 mM $MgCl_2$, 2 mM $CaCl_2$, 0.1 mg/mL BSA, 2 mM ATP, 1 mM DTT, 1 nM (molecules) radio-labeled DNA substrate and either RAD51 wild type or T131P

mutant protein, at the indicated concentrations. When ATP analogues were used, 2 mM of ATPγS or AMP-PNP and 3 mM $MgCl_2$ or $CaCl_2$ were used. An ATP regeneration system (1 mM PEP (phosphoenolpyruvate) and 80 U/ml pyruvate kinase) was used for the reactions with ATP and $MgCl_2$. The reaction mixture was incubated for 10 min at 37° C and mixed with 1 μL loading buffer (2.5% Ficoll-400, 10 mM Tris-HCl (pH 7.5), and 0.0025% Xylene cyanol). The reactions were resolved by PAGE using a 5% gel and 1X TAE buffer, and run for 120 minutes at 60 V in the cold room (4° C). The gel was then dried and exposed to a PhosphorImager screen. The screen was scanned on a Molecular Dynamics Storm 840 PhosphorImager and bands quantified using ImageQuant software. The percentage of protein-DNA complexes was calculated as the free radio-labeled DNA remaining in a given lane relative to the protein-free lane, which defined the value of 0% complex (100% free DNA).

DNA strand exchange assays

TK-167-mer and ^{32}P labeled dsDNA (annealed TK-Oligo1 and TK-Oligo2), which were described in above EMSA analysis, were used as the ssDNA substrate and homologous dsDNA, respectively. The buffer contained 25 mM TrisOAc (pH 7.5), 1 mM $MgCl_2$, 2 mM $CaCl_2$, 0.1 mg/mL BSA, 2 mM ATP, and 1 mM DTT. When ATP analogues were used, 2 mM of the indicated nucleotides and 3 mM of indicated divalent ions were used. All pre-incubations and reactions were at 37 °C. The DNA and proteins were at the following concentrations unless otherwise indicated: RAD51 proteins (0.22 μM), RPA (25 nM), BRCA2 (80 nM), ssDNA (4 nM molecules), and dsDNA (4 nM molecules). Where proteins were omitted, storage buffer was substituted. The reactions were terminated by adding SDS to 0.25% and proteinase K to 0.5 mg/mL and further incubation for 10 minutes. The reactions were analyzed by PAGE using a 6% gel and 1X TAE buffer, run at 60 V for 70 minutes at room temperature.

The gel was then dried and exposed to PhosphorImager screen. The percentage of DNA strand exchange product was calculated as labeled product divided by total labeled input DNA in each lane.

Joint molecule formation assay

The ssDNA, JM1, a PAGE-purified 100-mer (5'-GGA GTC GCT GAG GTA ACG CTG AGG TTA AGC TGA GGT ACA GCT GAG GAC TAG TGC TGA GGT TAA GCT GAG GAC TAG CTG AGG AAG TGC TGA GGA TAT GCT G-3') and pSE1 supercoil DNA were gifts from Naofumi Handa (S.C.K. laboratory). JM1 was labeled with ^{32}P at the 5'-end. To form filaments, the indicated concentrations of RAD51 WT or T131P protein were incubated with ^{32}P-labeled ssDNA (900 nM nts; 9 nM molecules) in buffer containing 25 mM TrisOAc (pH 7.5), 2 mM ATP, 1 mM $MgCl_2$, 2 mM $CaCl_2$, 0.1 mg/mL BSA, and 1 mM DTT for 10 min at 37°C. Joint molecule (D-loop) formation was initiated by addition of pSE1 supercoil DNA (27 μM; 9 nM molecules). Reactions were deproteinized by adding SDS to 1% and proteinase K to 1 mg/mL by incubation for 15 min at 37°C. Products were analyzed by electrophoresis using 1% agarose and 1X TAE, run at 50 V for 45 minutes at room temperature. After electrophoresis, the gel was dried and exposed to a PhosphorImager screen. The screen was scanned on a Molecular Dynamics Storm 840 PhosphorImager and bands quantified using ImageQuant software. The formation of joint molecules was calculated as labeled product divided by total labeled input DNA in each lane, and expressed as a percentage relative to the amount of ssDNA.

Mass Spectrometry

Proteins bands were digested with either Lysyl Endopeptidease C (Lys-C) (Wako Chemicals USA, Inc, Richmond, VA) or trypsin (Promega, Madison, WI, USA). The generated peptides were measured by nano LC-MS/MS using a Q-Exactive mass spectrometer (Thermo, Bremen, Germany) coupled to a

Dionex NCP3200RS HPLC setups (Thermo, Sunnyvale, CA, USA). Samples were desalted and concentrated on a trap column prior to separation on a packed-in-emitter C_{18} column, 75 µm by 12 cm, 3 µm particles (Nikkyo Technos Co., Ltd. Japan). The analytical gradient were generated at 200 nL/min increasing from 10% Buffer B (0.1% formic acid in acetonitrile) / 90% Buffer A (0.1% formic acid) to 45% Buffer B / 55% Buffer A in 40 minutes followed by a wash step with 90% Buffer B / 10% Buffer A for 10 minutes and conditioning for 15 minutes with 1% Buffer B / 99% Buffer A Buffer B. Data Dependent Acquisition experiment: MS survey scans were scanned from m/z 300 to m/z 1400 at resolution of 70,000@200 Th (AGC: 1e6 and Maximum IT: 60 ms). Up to the 10 most abundant ions were subjected to MS/MS and measured at a resolution of 17,500 (AGC: 1e6 and Maximum IT: 120 ms) with m/z 100 as lowest mass. For the tandem MS experiments, an isolation window of 2.0 Th was used. Targeted experiment: cycling between isolating and fragmenting m/z 795.4 using a 4 m/z window (res: 35,000, AGC : 1e5, Max IT: 110ms, NCE: 25) and a full scan from m/z to 760-816 (res: 70,000, AGC : 1e6, Max IT: 100ms). Both wild type (LLQGGIETGSITEM(ox)FGEFRTGK) and mutant (LLQGGIETGSITEM(ox)FGEFRPGK) RAD51 peptides were consistently detected in patient fibroblast samples (Supplementary data Fig 1c-e). The non-oxidized form of these peptides were also observed but were always less abundant. Tandem MS spectra of the two methionine oxidized peptides, differing solely by a T->P mutation (theoretical fragment ion mass shift of 3.9949 Da, measured median mass shift of 3.996 Da, s.d. 0.068) (Supplementary data Fig. 1d), elutes less than half-a-minute apart (Supplementary data Fig. 1c). A chromatographic duplet associated with diastereomers of oxidized methionine is observed for both mutant and wild type peptides (Bunkenborg et al., 2010).

To test if differences in digestion efficiency of the two RAD51 proteins or/and differences in ionization efficiency of the two peptides can account for the measured ~5-fold difference, and for additional validation, a control LC-MS/MS experiment was designed using purified wild type and mutant RAD51.

Equal amounts of mutant and wild type RAD51 was: 1) loaded in separate lanes and 2) mixed 1:1 and loaded in one lane of a SDS-PAGE gel. Excised RAD51 protein bands were digested and subjected to standard data dependent acquisition (DDA) as well as a targeted LC-MS/MS. DDA generated data was queried against wild type and mutant RAD51 sequences inserted into a custom database (5,225 sequence entries including common contaminants(Bunkenborg et al., 2010). Using extracted ion chromatograms from six RAD51 peptides not differing between mutant and wild type we tested for loading (mutant vs. wild type: of 1.3, s.d. 0.35). In the same two samples we measured a fold differences of 2.8 (mutant: wild type) between the peptides differing in sequence. For the 1:1 mixed sample a fold difference of 1.6 (n=5, s.d. 0.30) was measured. These measurements suggests that our strategy may be slightly over-estimating amounts of the T>P mutated RAD51.

Supplemental References

Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., and Sunyaev, S.R. (2010). A method and server for predicting damaging missense mutations. Nature methods *7*, 248-249.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T.*, et al.* (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet *25*, 25-29.

Boyle, A.P., Hong, E.L., Hariharan, M., Cheng, Y., Schaub, M.A., Kasowski, M., Karczewski, K.J., Park, J., Hitz, B.C., Weng, S.*, et al.* (2012). Annotation of functional variation in personal genomes using RegulomeDB. Genome research *22*, 1790-1797.

Bunkenborg, J., Garcia, G.E., Paz, M.I., Andersen, J.S., and Molina, H. (2010). The minotaur proteome: avoiding cross-species identifications deriving from bovine serum in cell culture models. Proteomics *10*, 3040-3044.

Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly *6*, 80-92.

Consortium, E.P. (2004). The ENCODE (ENCyclopedia Of DNA Elements) Project. Science *306*, 636-640.

Cooper, G.M., Stone, E.A., Asimenos, G., Program, N.C.S., Green, E.D., Batzoglou, S., and Sidow, A. (2005). Distribution and intensity of constraint in mammalian genomic sequence. Genome research *15*, 901-913.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., *et al.* (2011). The variant call format and VCFtools. Bioinformatics *27*, 2156-2158.

Davydov, E.V., Goode, D.L., Sirota, M., Cooper, G.M., Sidow, A., and Batzoglou, S. (2010). Identifying a high fraction of the human genome to be under selective constraint using GERP++. PLoS computational biology *6*, e1001025.

DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., *et al.* (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet *43*, 491-498.

Fisher, S., Barry, A., Abreu, J., Minie, B., Nolan, J., Delorey, T.M., Young, G., Fennell, T.J., Allen, A., Ambrogio, L., *et al.* (2011). A scalable, fully automated process for construction of sequence-ready human exome targeted capture libraries. Genome Biol *12*, R1.

Garner, E., Kim, Y., Lach, F.P., Kottemann, M.C., and Smogorzewska, A. (2013). Human GEN1 and the SLX4-Associated Nucleases MUS81 and SLX1 Are Essential for the Resolution of Replication-Induced Holliday Junctions. Cell Rep *5*, 207-215.

Genomes Project, C., Abecasis, G.R., Altshuler, D., Auton, A., Brooks, L.D., Durbin, R.M., Gibbs, R.A., Hurles, M.E., and McVean, G.A. (2010). A map of human genome variation from population-scale sequencing. Nature *467*, 1061-1073.

Griffith, O.L., Montgomery, S.B., Bernier, B., Chu, B., Kasaian, K., Aerts, S., Mahony, S., Sleumer, M.C., Bilenky, M., Haeussler, M., *et al.* (2008). ORegAnno: an open-access community-driven resource for regulatory annotation. Nucleic Acids Res *36*, D107-113.

Grompe, M., and al-Dhalimy, M. (1993). Mutations of the fumarylacetoacetate hydrolase gene in four patients with tyrosinemia, type I. Human mutation *2*, 85-93.

Hilario, J., Amitani, I., Baskin, R.J., and Kowalczykowski, S.C. (2009). Direct imaging of human Rad51 nucleoprotein dynamics on individual DNA molecules. Proc Natl Acad Sci U S A *106*, 361-368.

Hubisz, M.J., Pollard, K.S., and Siepel, A. (2011). PHAST and RPHAST: phylogenetic analysis with space/time models. Briefings in bioinformatics *12*, 41-51.

Jensen, R.B., Carreira, A., and Kowalczykowski, S.C. (2010). Purified human BRCA2 stimulates RAD51-mediated recombination. Nature *467*, 678-683.

Kim, Y., Spitz, G.S., Veturi, U., Lach, F.P., Auerbach, A.D., and Smogorzewska, A. (2013). Regulation of multiple DNA repair pathways by the Fanconi anemia protein SLX4. Blood *121*, 54-63.

Kircher, M., Witten, D.M., Jain, P., O'Roak, B.J., Cooper, G.M., and Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet *46*, 310-315.

Kumar, P., Henikoff, S., and Ng, P.C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc *4*, 1073-1081.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754-1760.

Li, M.Z., and Elledge, S.J. (2007). Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. Nature methods *4*, 251-256.

Maekawa, M., Ishizaki, T., Boku, S., Watanabe, N., Fujita, A., Iwamatsu, A., Obinata, T., Ohashi, K., Mizuno, K., and Narumiya, S. (1999). Signaling from Rho to the actin cytoskeleton through protein kinases ROCK and LIM-kinase. Science *285*, 895-898.

Mitchell, G.A., Lambert, M., and Tanguay, R.M. (1994). Hypertyrosinemi. In The Metabolic and Molecular Basis of Inherited Disease, J. Scriver, A.L. Beaudet, W. Sly, and D. Valle, eds. (New York: MacGraw-Hill), pp. pp1077-1106.

Pierce, A.J., Johnson, R.D., Thompson, L.H., and Jasin, M. (1999). XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. Genes Dev *13*, 2633-2638.

Sanjana, N.E., Shalem, O., and Zhang, F. (2014). Improved vectors and genome-wide libraries for CRISPR screening. Nature methods *11*, 783-784.

Sherry, S.T., Ward, M.H., Kholodov, M., Baker, J., Phan, L., Smigielski, E.M., and Sirotkin, K. (2001). dbSNP: the NCBI database of genetic variation. Nucleic Acids Res *29*, 308-311.

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.

Walther, A.P., Gomes, X.V., Lao, Y., Lee, C.G., and Wold, M.S. (1999). Replication protein A interactions with DNA. 1. Functions of the DNA-binding and zinc-finger domains of the 70-kDa subunit. Biochemistry *38*, 3963-3973.

Wang, A.T., Sengerova, B., Cattell, E., Inagawa, T., Hartley, J.M., Kiakos, K., Burgess-Brown, N.A., Swift, L.P., Enzlin, J.H., Schofield, C.J.*, et al.* (2011). Human SNM1A and XPF-ERCC1 collaborate to initiate DNA interstrand cross-link repair. Genes & Development *25*, 1859-1870.

Wang, T., Wei, J.J., Sabatini, D.M., and Lander, E.S. (2014). Genetic screens in human cells using the CRISPR-Cas9 system. Science *343*, 80-84.