Cell Reports, Volume *30*

Supplemental Information

The ARK Assay Is a Sensitive and Versatile Method for the Global Detection of DNA-Protein Crosslinks

Qianghua Hu, Naeh Klages-Mundt, Rui Wang, Erica Lynn, Liton Kuma Saha, Huimin Zhang, Mrinal Srivastava, Xi Shen, Yanyan Tian, Hyeung Kim, Yin Ye, Tanya Paull, Shunichi Takeda, Junjie Chen, and Lei Li

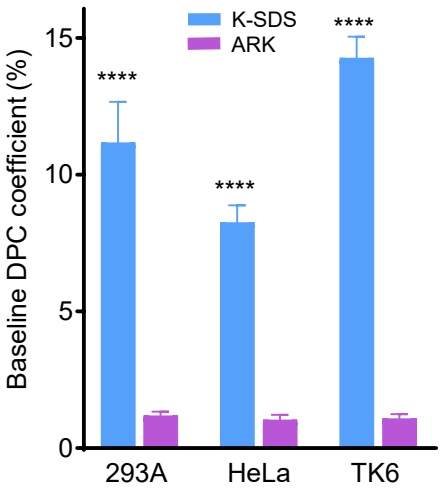

Figure S1. Parallel comparison of background DPC level using the conventional K-SDS and the ARK assay in three cell lines as indicated. Biological repeats with triplication: n = 3. Related to Figure 2.

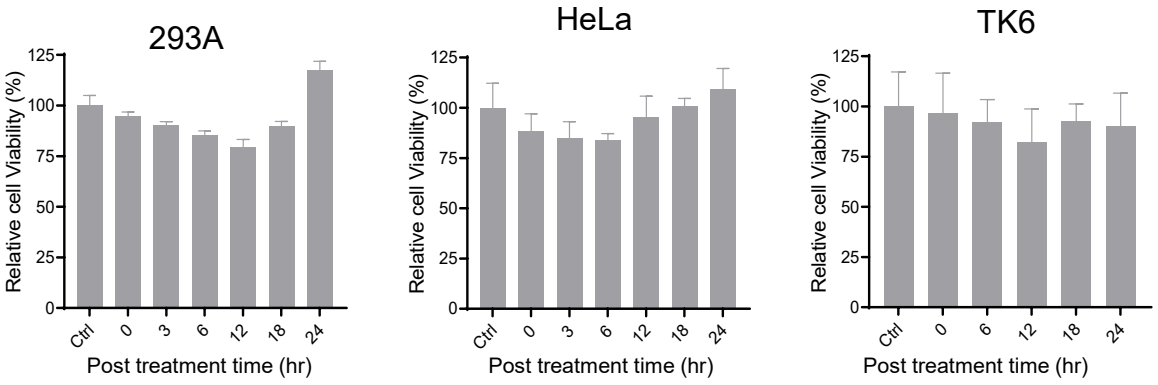

Figure S2. Parallel cell viability assessment in 293A, HeLa, and TK6 cells used in Figure 3A**.** Related to Figure 3A.

After 400 mM formaldehyde treatment for 2 hours, cells were collected and viability determined using automatic Trypan Blue exclusion counter at indicated time points. Viability of Mock-treated (Ctrl) cells was set as 100%. Biological repeats with duplication: n = 3 for each cell line.

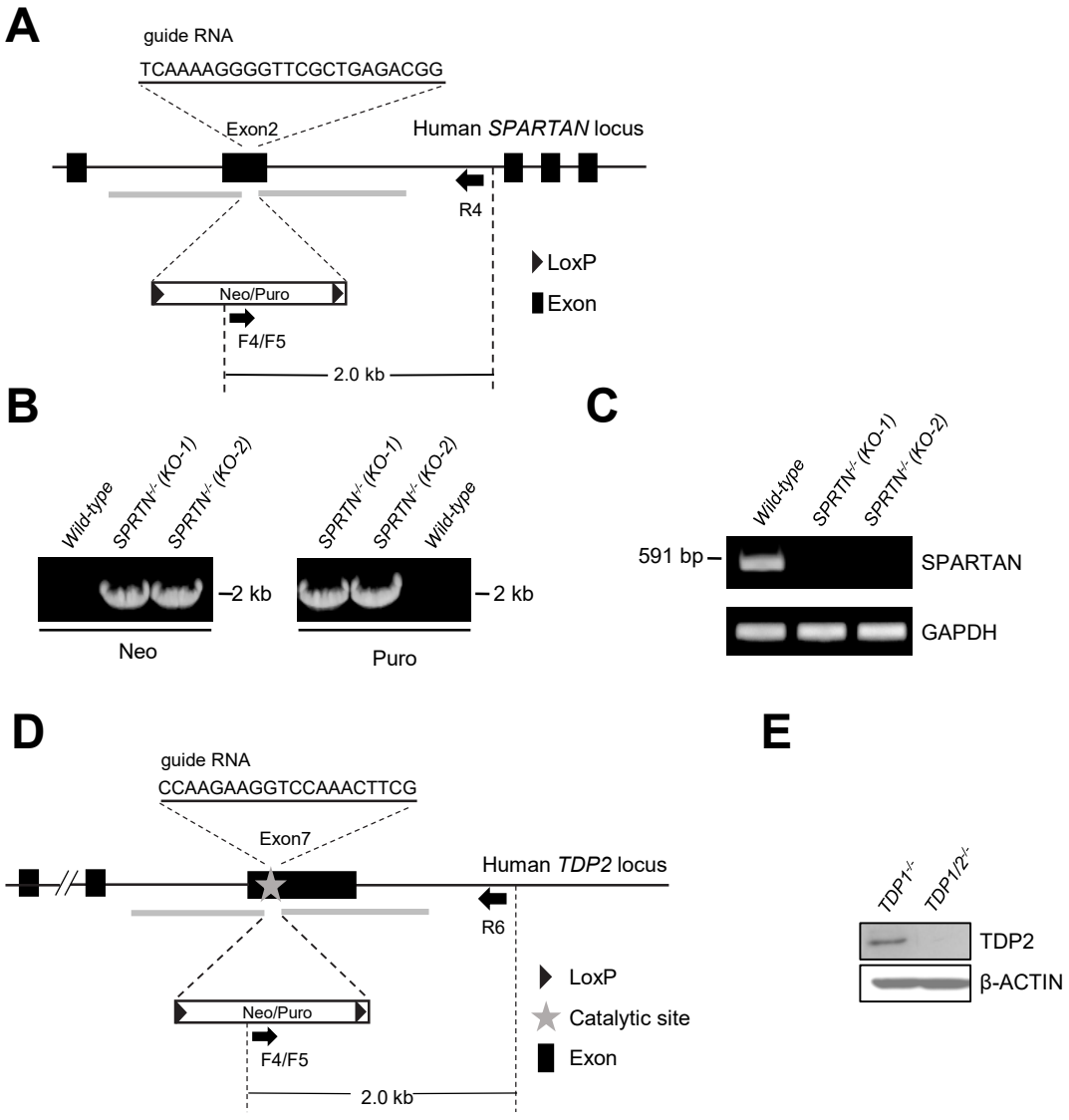

Figure S3. Generation and validation of SPRTN$^{-/-}$ and TDP1/2$^{-/-}$ TK6 cells. Related to Figure 5. (**A**) The SPRTN gene targeting strategy at the human SPARTAN locus. The black boxes represent the SPRTN exons and the SPRTN homology regions flanking the Neo/Puro resistance genes in the targeting vectors. (**B**) PCR genotyping of SPRTN knock-out clones. The forward primer F4, F5, and reverse R4 were used for genotyping PCR to verify precise target events. (**C**) Reverse transcription (RT)-PCR analysis to determine transcript levels of SPRTN . The primer F6 and R5 (see Methods for sequence information) were used to amplify a 591 bp cDNA fragment. (**D**) Schematic representation of the TDP2-genedisruption construct and the site recognized by guide RNA. PCR primers F4, F5, and R6 were used for genotyping PCR to verify precise target events. (**E**) Western blotting using anti-TDP2 antibody in *TDP1*$^{-/-}$ and *TDP1/2*$^{-/-}$ TK6 mutants. β-actin is a loading control.

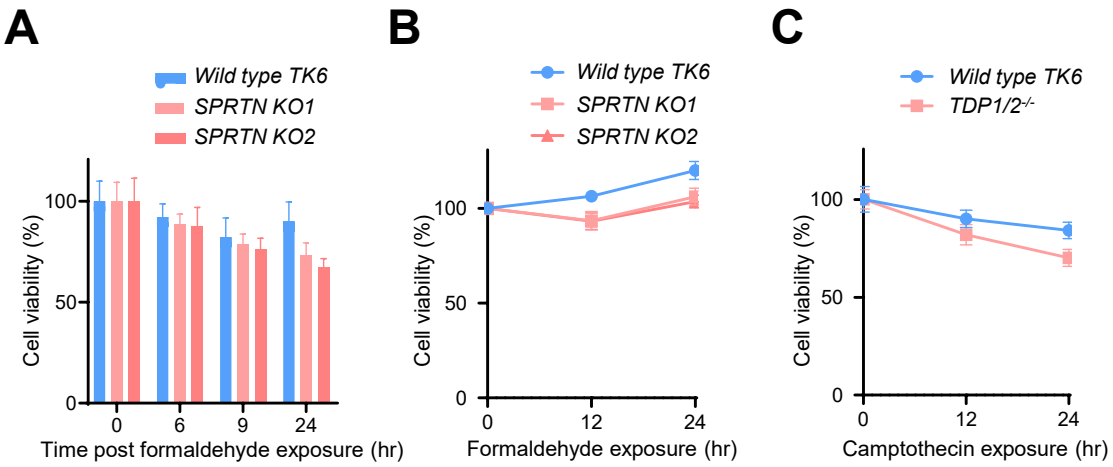

Figure S4. Parallel assessments of cell viability in parental and knockout cells used in Figure 5. Related to Figure 5

(**A**) TK6 and the *SPRTN* knockout derivatives (*KO1* and *KO2*) were treated with 400 μM formaldehyde for 2 hr. Cells were collected at indicated time points and viability determined using automatic Trypan Blue exclusion counter. Cell viability immediately after formaldehyde exposure was set as 100%. (**B**) TK6 and the *SPRTN* knockout derivatives (*KO1* and *KO2*) were treated with 50 μM formaldehyde continuously for 12 and 24 hours. Cells were collected at indicated time points and viability determined using automatic Trypan Blue exclusion counter. Cell viability of mock-treated cells was set as 100%. (**C**) TK6 and the *TDP1/2*$^{-/-}$ knockout mutant were treated with 75 nM of camptothecin continuously for 12 and 24 hours. Cells were collected at indicated time points and viability determined using automatic Trypan Blue exclusion counter. Cell viability of mock-treated cells was set as 100%. Each data point was generated from ≥ 4 biological repeats with duplication. Error bars depict standard variations.

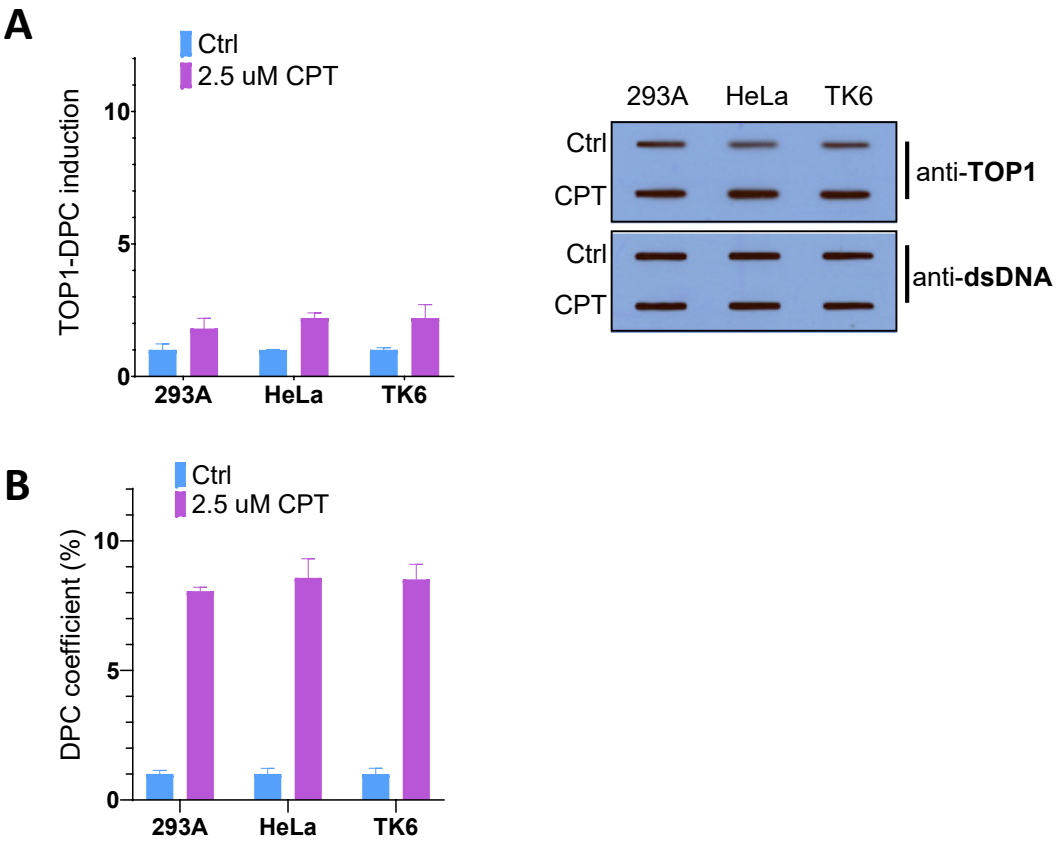

Figure S5: Comparison of assay readouts between the ARK and RADAR assays at a lower dose of CPT**.** Related to Figure 6.
(**A**) RADAR assay detection of DPCs from 3 indicated cell lines treated with 2.5 μM of CPT for 1 hour. Left panel: Relative Top1-DPC induction. The data were derived by normalizing the Top1-DPC band intensity against that of its corresponding dsDNA loading control. Right panel: a representative slot blot showing the TOP1 and dsDNA loading control bands. Background level of mock-treated sample (Ctrl) for each cell line was set to 1. **B**) ARK assay of identical cell samples used in panel A. Number of biological repeats with duplication: n = 4 for panel (A), n = 3 for panel (B). Error bars depict standard deviations.

A HeLa CRISPR/Cas9 Knockout Genotypes

XPA$^{+/+}$ (exon 1)
ATGGCGGCGGCCGACGGGGCTTTGCCGGAGGCGGCGGCTTTAGAGCAACCCGCGGAGCTGCCTGCCTCGGTGCGGGC
GAGTATCGAGCGGAAGCGGCAGCGGGCACTGATGCTGCGCCAGGCCCGGCTGGCTGCCCGGCCCTACTCGGCGACGG
CGGCTGCGGCTACTGGAG

XPA$^{-/-}$ Allele 1
-17 bp -5 bp
ATGGCGGCGGC_AGGCGGCGGCTTTAGAGCAAC_GAGCTGCCTGCCTCGGTGCGGGCGAGTATCGAGCGGAAGCGGC
AGCGGGCACTGATGCTGCGCCAGGCCCGGCTGGCTGCCCGGCCCTACTCGGCGACGGCGGCTGCGGCTACTGGAG

XPA$^{-/-}$ Allele 2
-1 bp -7 bp
ATGGCGGCGGCCGACGGGGCTTTGCC_GAGGCGGCGGCTTTAG_CGCGGAGCTGCCTGCCTCGGTGCGGGCGAGTAT
CGAGCGGAAGCGGCAGCGGGCACTGATGCTGCGCCAGGCCCGGCTGGCTGCCCGGCCCTACTCGGCGACGGCGGCTG
CGGCTACTGGAG

XPA$^{-/-}$ */FANCL*$^{-/-}$ Allele 1
-106 bp
ATGGCGGCGGCCGACGGGG_GCTGGCTGCCCGGCCCTACTCGGCGACGGCGGCTGCGGCTACTGGAG

XPA$^{-/-}$ */FANCL*$^{-/-}$ Allele 2
-34 bp -1 bp
ATGGCGGC_GAGCAA_CCGCGGAGCTGCCTGCCTCGGTGCGGGCGAGTATCGAGCGGAAGCGGCAGCGGGCACTGAT
GCTGCGCCAGGCCCGGCTGGCTGCCCGGCCCTACTCGGCGACGGCGGCTGCGGCTACTGGAG

B

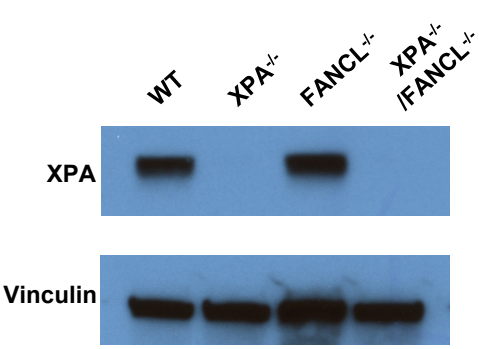

Figure S6. Generation of CRISPR/Cas9-mediated single- and double-knockout mutants in HeLa cells. Related to Figure 7.
(**A**) Genotypes of *XPA*$^{-/-}$, *FANCL*$^{-/-}$, and *XPA*$^{-/-}$*/FANCL*$^{-/-}$ mutants generated in HeLa cells. Each mutants harbors bi-allelic frame disrupting mutations in Exon 1 of the targeted loci. (**B**) Western blotting of the target protein for each mutant.

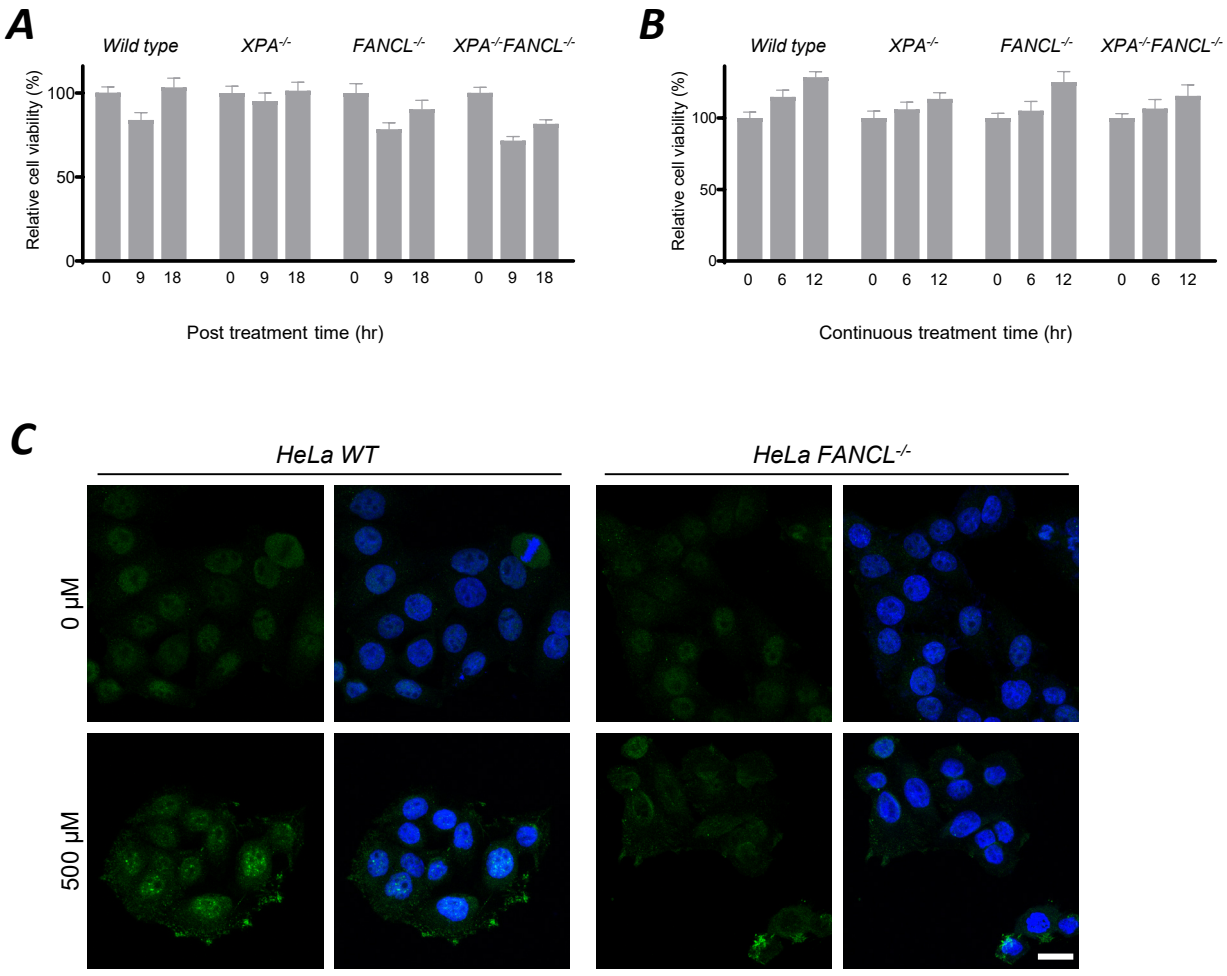

Figure S7. Cell viability and FANCD2 foci formation in wild type and mutant HeLa cells exposed to formaldehyde. Related to Figure 7**.**
After 500 μM formaldehyde treatment for 2 hours (**A**) or 75 μM formaldehyde treatment continuously for up to 12 hours (**B**), cells were collected and viability determined using automatic Trypan Blue exclusion counter at indicated time points. Viability of Mock-treated (Ctrl) cells was set as 100%. Number of biological repeats with duplication: n = 2 for each genotype. Error bars depict standard deviations. (**C**) Immunostaining of FANCD2 foci in wild type and *FANCL*$^{-/-}$ HeLa cells treated with indicated concentrations of formaldehyde for 6 hours. Scale bar = 10 μm

Table S1. Primer information for generation of mutant in TK6 cells. Related to Figure 7.

Primer Name	Primer Sequences
F1	5’- TTAAATGAGGGGAGTCTGGTTTTGGACCTG -3’
R1	5’- CATCTCCAAGCATGTTTTAACATGCTTCCC -3’
F2	5’- GCGAATTGGGTACCGGGCCTTAAATGAGGGGAGTCTGGT -3’
R2	5’- CTGGGCTCGAGGGGGGGCCCTAGAAAACAATCACCATAA -3’
F3	5’- TGGGAAGCTTGTCGACTTAAGTATTTTTCGTGTAAACTTG -3’
R3	5’- CACTAGTAGGCGCGCCTTAACATCTCCAAGCATGTTTTAA -3’
F4	5’- AACCTGCGTGCAATCCATCTTGTTCAATGG -3’
F5	5’- GTGAGGAAGAGTTCTTGCAGCTCGGTGA -3’
R4	5’- TACACCTACCTTTCTCTTTATCATCTGAGC -3’
F6	5’- CCGGACTTGCAGGCACTGTTTGTTCAGTTT -3’
R5	5’-TTTATTCTCTGCGGCCAATACTGGTTCCTT-3’
R6	5’- CCACAGAAATGTACTATTGTATTACCTCTA -3’
F7	5’- GCGAATTGGGTACCGGGCCAATGGTGAATTGGTGTTTAATGGGTAC -3’
R7	5’- CTGGGCTCGAGGGGGGGCCCCTCTTCTGCTGCTGCTCTGAAAAATA -3’
F8	5’- TGGGAAGCTTGTCGACTTAACTGTGCAACTTAGATATAATATTGTAA -3’
R8	5’- CACTAGTAGGCGCGCCTTAATGGAGTGAGAAGCAAATGGAAATCT -3’