# Phosphorylation of PNKP mediated by CDKs promotes end-processing of Okazaki fragments during DNA replication

**Kaima Tsukada[1)], Rikiya Imamura[1)], Kotaro Saikawa[1)], Mizuki Saito[1)], Naoya Kase[1)], Tomoko Miyake[2)], Masamichi Ishiai[3)], Yoshihisa Matsumoto[1)], and Mikio Shimada[1) 4)*]**

[1)]Laboratory for Zero-Carbon Energy, Institute of Innovative Research, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo, 152-8550,

[2)]Cosmetic R&D Department, Takara Belmont Corp, 7-1-19 Akasaka, Minato-ku, Tokyo 107-0052, Japan

[3)]Japan National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo, 104-0045, Japan

[4)]Lead contact

*Correspondence: mshimada@zc.iir.titech.ac.jp

Phone: +81-3-5734-3703

Fax: +81-3-5734-3703

Running title: Phosphorylation of PNKP by CDK for DNA replication

Figure X7

Supplementary Figure X8

## Abstract

Polynucleotide kinase phosphatase (PNKP) has enzymatic activities as 3′ phosphatase and 5′ kinase of DNA ends to promote DNA ligation. Here, we show that PNKP is involved in progression of DNA replication through end-processing of Okazaki fragments (OFs). Cyclin-dependent kinases (CDKs) regulate phosphorylation on threonine 118 (T118) of PNKP, and which phosphorylation allows it to be recruited to OFs. Loss of PNKP and T118 phosphorylation significantly increased unligated OFs and high-speed DNA synthesis in replication forks, suggesting that PNKP T118 phosphorylation is required for OFs ligation for its maturation. Furthermore, phosphatase-dead PNKP also exhibited an accumulation of unligated OFs and high-speed DNA synthesis. Overall, our data suggested that CDK-mediated PNKP phosphorylation at T118 is important for its recruitment to OFs and PNKP subsequently promotes end-processing for OFs maturation for stable cell proliferation.

## Keyword

PNKP, DNA replication, Okazaki fragments, DNA repair, CDKs, poly ADP-ribosylation

## Introduction

DNA replication integrity is essential to ensure genome stability and accurate cell proliferation (O'Donnell et al., 2013; Siddiqui et al., 2013). DNA replication is initiated at the origin of replication in the leading strand and at the short RNA-primed DNA fragments, known as Okazaki fragments (OFs), in the lagging strand (Leonard and Mechali, 2013; Okazaki et al., 1968). The DNA replication machinery is precisely controlled by cyclin-dependent kinases (CDKs), which phosphorylate several replication factors to allow them to enter S phase and promote DNA synthesis (Swaffer et al., 2016). During replication fork progression, single-stranded DNA is fragile and protected by replication protein A (RPA) (Wold, 1997). When replication forks stall, ataxia telangiectasia mutated- and Rad3-related protein (ATR) is activated and phosphorylates CHK1 and RPA to resolve or remove it, promoting replication fork recovery (Cortez et al., 2001; Zou and Elledge, 2003). Furthermore, DNA damages, e.g., base damage, single-strand breaks (SSBs), and double-strand breaks (DSBs), occur upon genotoxic stresses and replication errors, which lead to DNA replication stress, genomic instability, genetic mutation, and tumorigenesis. Poly (ADP-ribose) polymerase 1/2 (PARP1/2) are involved in the repair of SSBs and DSBs, and multiple DNA replication processes (Ray Chaudhuri et al., 2012; Ray Chaudhuri and Nussenzweig, 2017). PARP1 activity is required for not only efficient SSBs repair and filling OF gaps during DNA replication (Hanzlikova et al., 2018).

Polynucleotide kinase phosphatase (PNKP) is a key enzyme with dual roles. It has 3′-phosphatase and 5′-kinase activity (Chappell et al., 2002; Jilani et al., 1999; Karimi-Busheri et al., 1999). PNKP is recruited to the sites of SSBs and DSBs depending on the interactions with XRCC1 and XRCC4, respectively, and is involved in base excision repair (BER), SSB repair, and non-homologous end joining (NHEJ) for DSB repair (Breslin and Caldecott, 2009; Mani et al., 2019; Mani et al., 2010; Tsukada et al., 2020a; Tsukada et al., 2020b). Human PNKP consists of 521 amino acids, including the forkhead-associated (FHA) domain (amino acid residue 1–110) in the amino-terminal region and phosphatase (146–337) and kinase (341–516) domains in the carboxy-terminal region, which are

connected by a linker region (111–145). Phosphorylation of PNKP at serine 114 by ataxia telangiectasia mutated (ATM) is required for protein stability and efficient DNA repair for cellular survival (Parsons et al., 2012; Segal-Raz et al., 2011). Meanwhile, PNKP mutations are associated with the human inherited disease microcephaly and seizures (MCSZ) (Shen et al., 2010), ataxia oculomotor apraxia 4 (AOA4) (Bras et al., 2015) and Charcot-Marie-Tooth disease (CMT2B2) (Pedroso et al., 2015). These mutations are mostly sited in phosphatase or kinase domains and attenuated phosphatase and kinase activity (Bermudez-Guzman et al., 2020; Kalasova et al., 2020; Reynolds et al., 2012).

In this study, we revealed that PNKP is required for DNA replication process. Defects in PNKP induced over-accelerated DNA synthesis and genome instability. We also found that CDKs phosphorylate PNKP at threonine 118 (T118), mainly in S phase, and the CDK-mediated PNKP phosphorylation allows it to be recruited to unligated OFs via the interaction with replication protein A 2 (RPA2). Moreover, PNKP enzymatic activity, especially those of phosphatase activity, is subsequently required for end-processing of OFs during OFs maturation. Taken together, our data suggested that phosphorylation-mediated PNKP recruitment to OFs and end-processing activity of PNKP are critically important for preventing genome instability through the regulation of DNA replication.

## Results

### *PNKP is important for the DNA replication integrity*

In a previous study, depletion of PNKP resulted in increasing formations of micronuclei and chromosome bridges even without any treatment (Tsukada et al., 2020b). Since formations of micronuclei and chromosome bridges are associated with DNA replication, followed by genome instability (Mackenzie et al., 2017), we considered the possibility that PNKP is involved in DNA replication integrity. To investigate this possibility, we depleted with PNKP using siRNA in the human

osteosarcoma cells U2OS and analyzed the formation of micronuclei and chromosome bridges (Fig. 1A, B, and C). PNKP-depleted cells showed significant increases of both micronuclei and chromosome bridges compared to XRCC1- or XRCC4-depleted cells, suggesting that PNKP might be involved in the suppression of genome instability via a pathway independent of XRCC1 and XRCC4. Since insufficient loading of RPA to the single-stranded DNA triggers the activation of ATR and DNA replication stress (Zou and Elledge, 2003), we assessed the RPA-loading at single-stranded DNA after induction of DNA replication stress by CPT in PNKP-depleted cells. PNKP-depleted cells showed decreased RPA2 intensity compared to control cells (Fig. 1D and E). These results suggested that PNKP is required for the DNA replication integrity.

***Generation of PNKP$^{-/-}$ U2OS cells by CRISPR/Cas9 genome editing***

We next generated PNKP knockout U2OS cells (PNKP$^{-/-}$ cells) using CRIPSR/Cas9 Nickase (D10A) targeting exon 4 of *PNKP* coding region (Chiang et al., 2016). We obtained two clones (C1 and C2) which showed complete loss of PNKP, confirmed by western blotting analysis and DNA sequencing (Fig. 2A and S1). To test whether the PNKP$^{-/-}$ cells had deficiencies in DNA repair, we analyzed cellular sensitivity to the DNA damage treatment. A combination of poly ADP-ribose glycohydrolase inhibitor (PARGi) and SSBs induction, such as IR, can detect SSBs stained by anti-pan-ADP-ribose binding reagents by immunofluorescence (Hoch et al., 2017). We assessed SSB repair activity in PNKP$^{-/-}$ cells. PNKP$^{-/-}$ cells showed a marked increase in ADP-ribose intensity, indicating that PNKP$^{-/-}$ cells impaired the repair of SSBs (Fig. 2C). PNKP$^{-/-}$ cells showed a delay in diminishing the phosphorylation of KAP1, DSBs marker (Fig. 2B). Furthermore, PNKP$^{-/-}$ cells showed increased sensitivity to the hydrogen peroxide ($H_2O_2$), a powerful inducer of SSBs, as well as IR (Fig. 2D and E). These results confirmed that PNKP$^{-/-}$ cells defect in SSB and DSB repair and which is consistent with a previous report (Chalasani et al., 2018), indicating that PNKP$^{-/-}$ cells were successfully established.

***PNKP is important for DNA replication and cell proliferation***

To address PNKP function during DNA replication, we analyzed the cellular sensitivity to the DNA replication stress. PNKP$^{-/-}$ cells showed increased sensitivity to the HU and CPT treatment (Fig. 2F and G). Further, flowcytometric analysis revealed that cells in S phase are significantly increased in PNKP$^{-/-}$ cells (Fig. 3A, S1C). Next, to investigate whether PNKP is involved in replication fork progression, we used DNA fiber assays in PNKP$^{-/-}$ cells. PNKP$^{-/-}$ cells had faster progression of DNA replication fork than WT cells (Fig. 3B). A similar phenotype has been reported in PCNA polyubiquitination mutant (K164R) cells, which were unable to reduce the replication fork's speed (Thakar et al., 2020). This result suggested that PNKP might be involved in the progression of the DNA replication fork. Since DNA replication is closely associated with cell proliferation, we analyzed cell growth by cell counting. Despite the replication fork's high-speed progression, PNKP$^{-/-}$ cells showed slower cell growth than WT cells (Fig. 3C). However, PCNA KR mutant cells also showed slow cell proliferation because the accelerated replication speed failed to protect nascent DNA degradation, leading to slow cell proliferation, a phenotype is reminiscent of PNKP$^{-/-}$ cells. To identify the region of PNKP involving in fork progression during DNA replication, we measured the growth rate and the speed of fork progression in PNKP$^{-/-}$ cells transiently expressing PNKP deletion mutants (D1: the FHA domain, D2: linker region, D3: phosphatase domain and D4: kinase domain) (Fig. 3D, E and S2A) (Tsukada et al., 2020b). The D2 mutant expressing cells showed slower cell proliferation than cells expressing PNKP WT and other mutants. Furthermore, D2 mutant expressing cells also showed an increased speed of replication fork compared with WT and D1 mutant expressing cells, although D3 and D4 showed mildly high-speed fork progression (Fig. 3F). For further analysis, we attempted to identify essential amino acids for DNA replication in the linker region and identified five potential phosphorylation sites: serine 114, threonine118, threonine 122, serine 126 and serine 143, using PhosphoSitePlus (Fig. 3G) (Hornbeck et al., 2015). The amino acids S114, T118, T122 and S143 are highly conserved among mammalian species. S114 and S126 form a typical SQ/TQ motif, and it is

reported that S114 is phosphorylated by ATM and S126 is phosphorylated by ATM and DNA-PKcs (Segal-Raz et al., 2011; Zolner et al., 2011). T118, T122, and S143 are novel sites for phosphorylation, and T118 and T122 are predicted to be a CDK phosphorylation substrate motif (S/TP). To elucidate the importance of these amino acids in DNA replication, we constructed point-mutant expression vectors of five predicted phosphorylation amino acids substituted with alanine and measured growth rates of these transfectants (Fig. 3H). T118A mutant expressing cells exhibited marked delay in cell growth, which was not in S114A, although T122A, S126A, and S143A were slightly delayed. These results suggested that the linker region of PNKP, especially the phosphorylation of PNKP T118, is required for cell proliferation and normal DNA replication.

***CDK-dependent phosphorylation of PNKP at T118 is required for interaction with RPA***

Since PNKP-depleted cells showed decreased RPA2 loading at single-stranded DNA against DNA replication (Fig. 1D and E), we examine the interaction between PNKP and RPA2 using GFP-pulldown immunoprecipitation assays. GFP-tagged PNKP (GFP-PNKP) was expressed in HEK293 cells, and cell extract was immunoprecipitated with GFP-trap magnetic beads. As we expected, PNKP was co-immunoprecipitated with RPA2 (Fig. 4A and S4A). ATM and XRCC4 were used as interaction controls. We next attempted to confirm whether T118 is required for interaction with RPA. We expressed GFP-PNKP WT, R35A/R48A, which is mutated at the protein interaction center of the FHA domain, S114A, which is mutated at the most common phosphorylation site for DNA repair, and T118A mutants in HEK293 cells and analyzed RPA2 interactions by GFP-pulldown immunoprecipitation assays (Fig. 4B). Although PNKP WT and S114A mutant interacted with interaction controls, XRCC1 and XRCC4, and RPA2, PNKP R35A/R48A did not interact with any of these proteins, confirming that the FHA domain is essential for all protein interactions. Interestingly, T118A mutant diminished the interaction with RPA2, although it interacted with XRCC1 and XRCC4, suggesting that T118 is specifically important for the interaction with RPA2. Next, to assess the

importance of T118 phosphorylation in DNA replication, we generated antibodies that recognize phosphorylated S114 (pS114), a control to confirm pT118 specificity in DNA replication, and phosphorylated T118 (pT118) peptide. Sufficient titer and specificity of pS114 and pT118 antibodies were confirmed using ELISA assays (Fig. S5 and S6). pT118 PNKP antibody was used for western blotting for lysate from U2OS WT, GFP-PNKP WT, or T118A expressing cells. Unfortunately, this antibody was found to cross-react with proteins around 55kDa, which is close to the apparent molecular mass of endogenous PNKP expression (Fig. S5C). Nevertheless, this antibody clearly recognized GFP-PNKP. Therefore, hereafter, we used GFP-PNKP expression to examine the phosphorylation of T118. To determine whether phosphorylation on T118 is DNA replication-specific, we synchronized HCT116 cells transiently expressing GFP-PNKP by double thymidine block and released them at specific times (Fig. S4B). After synchronization at the indicated cell cycle phase, cells were extracted and used for western blotting. Since Cyclin A2 peaks at mid to late S/G2 phase, and Cyclin E1 peaks at early S phase, we used these proteins for cell cycle markers (Fig. 4C) (Fung et al., 2007; Honda et al., 2005; Pagano et al., 1992). Surprisingly, PNKP pT118 was detected in asynchronized cells, particularly in the S phase, similar with Cyclin A2. However, it was very weak in mitosis, suggesting that the T118 phosphorylation has a specific role in the S phase. We then attempted to identify the kinase of PNKP at T118. After CDKs inhibitor, roscovitine (CDKi) (Bukanov et al., 2006; Meijer et al., 1997), treatment for 1h, GFP-PNKP expressing PNKP$^{-/-}$ cells showed decreased phosphorylation on T118. In contrast, pS114 was not affected by CDKi treatment, indicating that CDKs specifically phosphorylate PNKP on T118. To identify the type of CDK which phosphorylates PNKP on T118, we investigated the phosphorylation level of PNKP on T118 under co-overexpression of Cyclin A2 and CDK1, 2, which are targets of roscovitine and main binding partners of Cyclin A2, or CDK4, as a control (Fig. 4E). CDK2/Cyclin A2 overexpression showed an increased phosphorylation level on T118, although CDK1/Cyclin A2 and CDK4/Cyclin A2 overexpression also increased the PNKP phosphorylation level, albeit to a lesser extent than the CDK2/Cyclin A2. Taken together, these

results suggested that CDK2/Cyclin A2 complex potentially regulates the phosphorylation level of PNKP on T118 in S phase to interact with RPA2 and proceed DNA replication.

***Phosphorylation of PNKP at T118 is required for Okazaki fragment maturation***

To elucidate the further role of PNKP in DNA replication, we used HU and Emetine (EME), which is a DNA replication inhibitor blocked OFs formation (Burhans et al., 1991). We analyzed the extent of SSBs by poly ADP-ribosylation in PNKP$^{-/-}$ cells (Fig. 5A). U2OS HU treatment slightly increased the amount of poly ADP-ribosylation, and EME prevented poly ADP-ribosylation by inhibiting OFs formation in the S phase in U2OS WT cells. These results are consistent with those of a previous study (Hanzlikova et al., 2018). Although PNKP$^{-/-}$ cells showed a high level of ADP-ribose intensity spontaneously, HU and EME treatment reduced the ADP-ribose intensity, but slightly higher than WT cells, indicating that PNKP is required for SSBs repair in both S phase and outside of S phase. Especially, EME treatment dramatically reduced poly ADP-ribosylation in PNKP$^{-/-}$ cells, which suggested that the OFs formation might causes a difference in spontaneous SSBs between the presence and absence of PNKP. Therefore, we next attempted to clarify the role of PNKP pT118 during the OFs maturation. We transiently expressed PNKP WT or T118A in PNKP$^{-/-}$ cells and incubated them with EdU-containing medium to separately assess ADP-ribose intensity in S phase (EdU-positive) or outside of S phase (EdU-negative) (Fig. 5B and C). In outside of S phase, T118A expressing cells showed high ADP ribose intensity, similar with PNKP$^{-/-}$ cells, and PNKP WT rescued this effect. The ADP ribose intensity increased significantly in the S phase, suggesting that PNKP function and T118 phosphorylation are required for SSBs-less DNA replication. For further analysis, we used flap endonuclease 1 inhibitor (FEN1i) which prevents resection of overhanging nucleotides from the end of OFs leading to unligated OFs (Exell et al., 2016; Ward et al., 2017; Zheng and Shen, 2011). We assessed the ADP-ribose intensity of PNKP WT or T118A expressing PNKP$^{-/-}$ cells treated with either DMSO or FEN1i only in EdU-positive cells (Fig. 5D and E). PNKP$^{-/-}$ cells and T118A expressing

cells strongly increased the intensity of ADP ribose compared with U2OS and PNKP WT expressing cells in the S phase even after DMSO treatment. FEN1i treated U2OS and PNKP WT expressing cells showed a higher ADP-ribose intensity than did DMSO treated cells. In contrast, FEN1i treated PNKP$^{-/-}$ or T118A expressing cells disappeared the difference of the ADP-ribose intensity, suggesting that PNKP may suppress the OFs-mediated SSBs in an epistatic pathway with FEN1. We then assessed the implication of the OFs maturation in the replication fork progression by DNA fiber assay (Fig. 5F). FEN1i treatment induced high-speed fork progression, although HU treatment led to stalled forks, suggesting that the inhibition of OFs maturation by FEN1i leads to a defect in replication fork protection. Next, we analyzed this phenotype in PNKP$^{-/-}$ cells and T118A expressing cells (Fig. 5G). Surprisingly, PNKP$^{-/-}$ cells and T118A expressing cells showed spontaneous fast fork progression, and FEN1i treatment did not accelerate the DNA replication speed additionally. Moreover, T118 phospho-mimetic mutant (T118D) expressing cells showed normal DNA replication speed as similar as PNKP WT expressing cells, suggesting that PNKP and its phosphorylation at T118 play essential roles in OFs maturation and accurate DNA replication (Fig. 5H). Taken together, we here concluded that the phosphorylation of PNKP at T118 is required for processing unligated OFs for preventing OFs-mediated SSBs formation and high-speed DNA replication.

***Phosphatase activity of PNKP is important for the end-processing of Okazaki fragments***

The main function of PNKP is to catalyze 5’- phosphorylation and 3’-dephosphorylation of the DNA ends. To clarify the role of these PNKP enzymatic activities for DNA end-processing during OFs maturation in the lagging strand, we constructed phosphatase-dead (D171A) and kinase-dead (K378A) PNKP mutants (Jilani et al., 1999; Kalasova et al., 2019; Reynolds et al., 2012). We first tried to obtain stable clones expressing both mutants, but we could not obtain phosphatase-dead (D171A) stable clones. This is consistent with previous observations, in which PNKP-mutated MCSZ patient cells mutated in phosphatase domain showed unstable expression of PNKP (Shen et al., 2010), and

recombinant phosphatase-dead PNKP is unstable (Kalasova et al., 2020). Therefore, these mutants were expressed transiently in PNKP$^{-/-}$ cells and detected protein expression by western blotting (Fig. 6A). Next, we analyzed the speed of fork progression in cells expressing these mutants using DNA fiber assay (Fig. 6B). Interestingly, D171A expressing cells showed spontaneously higher speed of fork progression than PNKP WT expressing cells, and FEN1i treatment did not accelerate the DNA replication speed additionally, although K378A expressing cells slightly showed spontaneous faster fork progression than PNKP WT expressing cells, suggesting that PNKP enzymatic activities, especially those of phosphatase, are required for the accurate fork progression and protection. Next, we elucidated whether these end-processing activities of PNKP are important for preventing OFs-mediated SSBs formation. These PNKP mutants expressing PNKP$^{-/-}$ cells were treated with FEN1i, and then we measured the ADP-ribose intensity of EdU-positive cells (Fig. 6 C and D). D171A showed a high level of ADP-ribose intensity, and K378A also showed a relatively high level of ADP-ribose intensity without FEN1i treatment. Moreover, FEN1i treated cells showed a high level of ADP-ribose intensity in all cells. Altogether, these results suggested that PNKP phosphatase and kinase activities, especially those of phosphatase, play an important role in the end-processing of unligated OFs ends for the suppression of the OFs-mediated SSBs formation and accurate DNA replication.

***Phosphorylation of PNKP at T118 is required for the genome stability***

Finally, since PNKP-depleted cells showed micronuclei and chromosome bridges (Fig. 1A, B, and C), we addressed whether the phosphorylation of PNKP at T118 is required for preventing genome stability. PNKP T118A expressing cells showed increasing frequency of micronuclei and chromosome bridges compared with PNKP WT expressing cells (Fig. 7 A and B). Furthermore, we assessed the importance of the phosphorylation of PNKP at T118 in IR-induced SSB and DSB repair abilities by poly ADP-ribosylation and γH2AX foci formation, respectively (Fig. S7 A, B, and C). PNKP T118A

expressing cells showed high level of ADP-ribose intensity and increasing γH2AX foci positive cells, suggesting that phosphorylation of T118 is essential for both SSB and DSB repairs. To address weather phosphorylation of PNKP on T118 is required for recruitment to the gapped DNA structure, we assessed gapped DNA binding ability using cellular extracts from T118A and T118D mutants expressing HCT116 cells (Fig. 7C). PNKP T118D (phosphor mimic) mutant showed relatively higher binding ability than PNKP T118A mutant suggesting that phosphorylation of PNKP on T118 is required for recruitment to the gapped DNA structure including DNA SSBs and unligated OFs. Overall, recruitment of PNKP to the OFs through phosphorylation on T118 is important for end-processing of OFs ends and genome stability.

## Discussion

In this study, we uncovered a novel role for PNKP in DNA replication. PNKP has dual functions as dephosphorylation of 3' end and phosphorylation of 5' end of DNA. Here, we showed that these end-processing activities of PNKP are required for OFs maturation during DNA replication. Furthermore, we also identified the phosphorylation of PNKP at T118 catalyzed by CDKs to be important for the recruitment to the OFs-mediated SSBs through the interaction with RPA. Defects of PNKP phosphorylation at T118 resulted in unmatured OFs during DNA replication.

PNKP consists of four regions: the FHA domain, linker region, phosphatase domain, and kinase domains. Compared with the other domains, the role of linker region is still poorly understood, although the linker region includes several residues that have possibility to be post-translationally modified. DNA damage signaling and DNA replication progressing are often regulated by protein modifications such as phosphorylation and ubiquitination (Blackford and Jackson, 2017; Cortez et al., 1999; Kolas et al., 2007). Lysine is the main target of E3 ubiquitin ligase (Laney and Hochstrasser, 1999), and there is a clustered lysine region (137–142) in the PNKP linger region. However, we reported that this region includes a nuclear localization signal (NLS) and the alanine's substitution led

to it not being transported to the nucleus (Tsukada et al., 2020a). Therefore, we focused on the other region from amino acids 137–142. In this study, we identified five predicted phosphorylation sites (S114, T118, T122, S126, and S143) in the linker region. Using laser micro-irradiation analysis, we assessed the accumulation of these PNKP mutants to the DNA damage sites (Fig. S3). Although T122A, S126A, and S143A mutants showed no significant difference from PNKP WT, these mutants expressing cells slightly reduced cell proliferation (Fig. 3H), suggesting that these residues might be involved in the other cellular functions. In contrast, S114A and T118A mutants significantly attenuated the accumulation to the DNA damage sites. In addition, phosphorylation mimic mutants, S114D and T118D recovered accumulation to the DNA damage sites (Fig. S3A, B, C, and D). S114A mutant expressing cells showed the presence of SSBs repair and normal cell growth (Fig. 3G and S8), suggesting that the phosphorylation of PNKP at S114 is likely to be important for DSBs repair. On the other hand, T118A mutants showed the absence of both SSBs and DSBs repair (Fig. 3G, and S7). However, residual SSBs often cause DNA replication-coupled DSBs formation, the reason of impaired DSB repair ability of T118A mutant might be come from the combination of impaired SSB repair ability and DNA replication-coupled DSBs formation.

Regulating the speed of DNA synthesis is important for accurate DNA replication (Genois et al., 2021; Kunkel, 2004). PNKP$^{-/-}$ and T118A expressing cells showed high-speed of DNA synthesis, resulting in a slow cell proliferation and genome instability (Fig. 3, 5 and 7). These observations are consistent with those from PCNA KR mutant cells, and PARPi treated cells (Maya-Mendoza et al., 2018; Thakar et al., 2020). We also found that the enzymatic activity, especially those of phosphatase activity, of PNKP is required for end-processing of OFs during DNA replication. PNKP$^{-/-}$ cells and mutant expressing cells strongly increased the ADP-ribose intensity in the S phase even without FEN1i treatment, although FEN1i treatment in these cells leads to attenuated or diminished increase in the ADP-ribose intensity compared with PNKP WT expressing cells, suggesting that PNKP act as an epistatic pathway in FEN1 (Fig. 5E and 6D). FEN1 catalyzes the removal of overhanging RNA/DNA

fragments from OFs on the lagging strand. Our data also suggests that both end-processing activity of PNKP and exonuclease activity of FEN1 are required for the ligation of unligated OFs (Fig. 5G). This observation is consistent with that HU and EME treated PNKP $^{-/-}$ cells showed decrease in the ADP-ribose intensity (Fig. 5A). Since EME selectively inhibits the OFs synthesis (Burhans et al., 1991) and HU treatment increases the amount of OF-like DNA fragments (Laipis and Levine, 1973; Magnusson, 1973a, b), they inhibit matured OFs structure as PNKP appropriate substrates. We also showed that phosphorylation of PNKP at T118 is regulated in a cell cycle-dependent and CDKs-associated manner (Fig. 4C, D, and E). We concluded that phosphorylation on T118 of PNKP allows it to be recruited to OFs on lagging strand in S phase and processes OFs ends for OFs maturation. Furthermore, in parallel, PNKP is required for PARP-XRCC1-dependent OFs gap-filling pathway when OFs maturation is stacked (Fig. 5A, 7D and S7A).

In summary, since the role of DNA replication is central to life, the function of PNKP in OFs maturation is important for biological development. In fact, although PNKP mutations cause several inherited diseases (MCSZ, AOA4 and CMT2B2), almost all of the mutations are found in the FHA, phosphatase or kinase domains and not in the linker region. This fact might be explained that mutations of phosphorylation residue in linker region is embryonic lethal. We show here a novel role of PNKP in OFs maturation during DNA replication, which might contribute to the elucidation of the biological basis of DNA replication and increase our understanding of the mechanisms of the onset of inherited diseases.

## Materials and Methods

### *Cell culture*

The human embryonic kidney cell line HEK293, the human colon cancer cell line HCT116 and the human osteosarcoma cell line U2OS were obtained from the American Type Culture Collection (ATCC) and U2OS PNKP$^{-/-}$ cell lines were established in this study. All cell lines were maintained in

Dulbecco's modified Eagle's medium (DMEM; Nacalai Tesque Inc.) supplemented with 10% v/v fetal bovine serum (FBS; Hyclone, GE Healthcare) and penicillin/streptomycin (Nacalai Tesque Inc.) at 37°C in humidified atmosphere containing 5% $CO_2$ conditions. All cell lines were tested for mycoplasma contamination using e-Myco™ mycoplasma detection PCR kit (iNtRON Biotechnology, Inc., cat# 25235).

***Construction of plasmid DNA and mutagenesis***

pEGFP-C1 was purchased from Clontech. Full-length human PNKP cDNA was obtained by PCR from the cDNA pool of U2OS cells and inserted into pEGFP-C1. Mutations were introduced using PrimeSTAR mutagenesis basal kit (Takara Bio, cat# R046A), according to the manufacturers' instruction. All DNA constructs were verified by DNA sequence analysis. All primers for mutagenesis were designed using the Agilent QuikChange primer design program and are shown in Supplementary table 1.

***cDNA and siRNA transfection***

For cDNA transfections, PEI-MAX (Polysciences, Inc. cat# 24765) or Lipofectamine 3000 (Invitrogen in Thermo Fisher Scientific) were used according to the manufacturers' instruction. For siRNA transfections (treatment time is typically for 48-72 h), Lipofectamine RNAiMAX (Invitrogen in Thermo Fisher Scientific) was used according to the manufacturer's instruction. All siRNAs were used at a final concentration of 50 nM. The targeting sites and the sequence of the siRNA oligonucleotides are shown in Supplementary table 2.

***Establishment of GFP-PNKP stably expressing cell lines***

To establish cell lines stably expressing GFP-PNKP, pEGFP-C1-PNKP was transfected into U2OS cells using Lipofectamine 3000, according to the manufacturer's instruction. On the subsequent day of

the transfection, the media was replaced by fresh growth media containing 400 μg/ml Geneticin G418 (Nacalai Tesque Inc.). After 10-14 days of incubation, single colonies were isolated using cloning rings, and the cells were cultured for expansion. GFP-PNKP expression in the clones was examined by western blotting and fluorescence microscopy.

***Genome editing by CRISPR/Cas9 system and establishment of PNKP$^{-/-}$ cell lines***

pSpCas9n(BB)-2A-Puro (PX462) vector was purchased from Addgene. For the establishment of PNKP$^{-/-}$ cells using CRISPR/Cas9 system, the sgRNA target sequences were cloned into the pSpCas9n(BB)-2A-Puro (PX462) vector and verified by DNA sequencing. U2OS cells were transfected with the targeting vectors and incubated for 2 days before addition of selective medium containing 1.5 μg/ml puromycin (Invivogen, cat# ant-pr-1). After 5 days, clonal cells were isolated by limiting dilution in 96-well plates. PNKP expression of single clones were analyzed by western blotting. Genomic DNA mutations in both PNKP alleles present in U2OS cells were verified by DNA sequencing of a PCR-amplified genomic fragment that was cloned into the pEGFP-C1 vector. sgRNA target sequences are shown in Supplementary figure 2.

***Laser micro-irradiation and live-cell imaging***

U2OS cells were grown on glass-bottom 35 mm dishes (Matsunami Glass Ind., Ltd.) and the indicated GFP-tagged cDNA were transfected into the cells at 2 days before the observation using the method as described above. On the subsequent day of transfection, culture media containing the transfection reagents were replaced with normal culture media without transfection reagents. At the day of observation, culture media were replaced with phenol red-free DMEM (Nacalai Tesque Inc.) supplemented with 10% v/v FBS and 0.5 μg/ml Hoechst 33258 (Sigma-Aldrich) was added for enhancement of the DNA damage at 30 min prior to the observation. Leica TCS SP8 LIGHTNING Confocal Microscope (Leica microsystems) with a 63x/1.40 oil immersion objective lens was used for

induction of localized DNA damages by laser micro-irradiation, live-cell observation and capturing pictures (Inano et al., 2017; Mochizuki et al., 2017). For the quantification, 20 cells were irradiated with the laser and analyzed the green fluorescent intensity in each experiment. Time-lapse images were started before the laser micro-irradiation. A 100% intensity scan (0.25 sec) from a 405 nm laser was used to induce localized DNA damages in this study. For observation of green fluorescence, 488 nm light was irradiated to the cells. The green fluorescence intensity at the irradiated sites were converted into numerical values by using SP8 software (Leica microsystems). The relative green fluorescence intensity was acquired by subtraction of the background intensity in the cells and division by the intensity at the irradiated area.

***SDS-PAGE and western blotting***

Cells were lysed in a radioimmunoprecipitation assay (RIPA) buffer (50 mM Tris HCl, pH 8.0, 250 mM NaCl, 25 mM ethylenediaminetetraacetic acid (EDTA), 0.5% v/v Triton X-100, 0.5% w/v sodium dodecyl sulfate (SDS), 0.5% w/v sodium deoxycholate) containing protease inhibitor cocktail (Nacalai Tesque Inc., cat# 25955-11) and phosphatase inhibitor cocktail (Nacalai Tesque Inc., cat# 07575-51), and the protein concentration was measured by a bicinchoninic acid (BCA) assay kit (Takara Bio) using bovine serum albumin as the standard. In all experiment, 20 μg of protein was loaded onto SDS polyacrylamide gel electrophoresis (SDS-PAGE) plates. The proteins were electrophoresed at 30 mA/gel plate for 1-1.5 h, and transferred onto a polyvinylidene fluoride (PVDF) membrane at 100 V for 1.5 h. Next, the PVDF membrane was blocked with either 2% w/v BSA/TBS-T (tris-buffered saline and Tween 20) or 5% w/v skim milk/TBS-T for 1 h at room temperature on a shaker. For primary antibody reactions, the following primary antibodies were used for 1- 4 h at room temperature: PNKP (rabbit, 1:1000, Novus, cat# NBP1-87257), pT118-PNKP (rabbit, 1:1000, generated in this paper), pS114-PNKP (rabbit, 1:1000, generated in this paper), XRCC1 (mouse, 1:1000, Invitrogen in Thermo Fisher Scientific, cat# MA5-13412), XRCC4 (rabbit, 1:1000, generated in our laboratory (Kamdar and

Matsumoto, 2010)), PCNA (rabbit, 1:500, Santa Cruz Biotechnology, cat# sc-7907), KAP1 (rabbit, 1:1000, abcam, cat# ab10483), pS824-KAP1 (rabbit, 1:1000, BETHYL, cat# A300-767A-2), Cyclin-A2 (mouse, 1:1000, Cell signaling, cat# BF683), Cyclin-E1 (rabbit, 1:1000, Sigma-Aldrich, cat# C4976), GFP (mouse, 1:3000, Nacali Tesque Inc., cat# GF200), ATM (mouse, 1:2000, Sigma Aldrich, cat# A1106), ATR (rabbit, 1:1000, BETHYL, cat# A300-137A), pT1989-ATR (rabbit, 1:1000, GeneTex, cat# GTX128145), DNA-PKcs (rabbit, 1:2000, abcam, cat# Y393), p53 (mouse, 1:5000, Santa Cruz Biotechnology, cat# sc-126), pS15-p53 (mouse, 1:500, abcam, cat# ab223868), RPA2 (mouse, 1:3000, abcam, cat# ab2175), pS4/8-RPA2 (rabbit, 1:1000, BETHYL, cat# A300-245A), CHK1 (mouse, 1:1000, Santa Cruz Biotechnology, cat# sc-8408), pS317-CHK1 (rabbit, 1:1000, Cell signaling, cat# D12H3), GAPDH (mouse, 1:10000, EMD Millipore, cat# MAB374). The PVDF membrane was washed three times with TBS-T. For secondary antibody reactions, horseradish peroxidase (HRP)-conjugated rabbit or mouse antibodies (Dako, cat# P0399 or P0447, respectively) were used for 1 h at room temperature. After washing five times with TBS-T, the membrane was developed by enhanced chemiluminescence (LI-COR, Biosciences) and detected by C-digit (LI-COR, Biosciences).

***Immunoprecipitation (IP)***

For sample preparation for immunoprecipitation (IP), HEK293 cells were grown on 10 cm dish and washed twice in phosphate-buffered saline (PBS: Nacalai Tesque Inc.), and lysed in lysis buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 0.2% NP-40, 1 mM $MgCl_2$, 10% glycerol), supplemented with cocktails of protease inhibitors and phosphatase inhibitors. After incubation for 30 min on the rotator at 4 °C, lysates were cleared by the centrifugation 20,000 x g for 20 mins at 4 °C. Next, lysates were incubated with 10 μl of GFP-Trap magnetic agarose beads (ChromoTek, GmbH) for 4 h with mixing on a rotator at 4 °C. The beads were then washed five times with lysis buffer and proteins were eluted in 2X SDS sample buffer (125 mM Tris-HCl, pH 6.8, 4% w/v SDS, 20% v/v glycerol, 0.01% w/v

bromophenol blue, 5% v/v 2-mercaptoethanol).

***Immunofluorescence***

Cells were grown on glass coverslips and fixed with 4% formaldehyde for 15 min at 4°C. Cells were subsequently permeabilized with PBS containing 0.2% Triton X-100 for 5 min at 4°C. Following 30 min of blocking in PBS supplemented with 2% w/v BSA, primary antibody reactions were performed in PBS-T supplemented with 1% w/v BSA for 2 h at room temperature. Cells were washed three times with PBS, and secondary antibody reactions were performed in PBS-T supplemented with 1% BSA for 1 h at room temperature in the dark. After washing five times with PBS, coverslips were mounted in mounting medium (Dako) containing the nuclear staining dye 4',6-diamidino-2-phenylindole dihydrochloride (DAPI), and allowed to dry for 2 h at room temperature in the dark. For a primary antibody, anti-γ-H2AX mouse antibody (Merck Millipore, JBW301) was used at 1:1000 dilution. Alexa Fluor 594-conjugated mouse secondary antibody (Invitrogen in Thermo Fisher Scientific, cat# A32742) was used at 1:2000 dilution.

For the quantification of γ-H2AX foci formation, nuclei and foci-positive cells were counted using ImageJ software. Foci-positive cells were defined as those containing more than 10 foci and at least 100 cells were counted. The representative pictures are shown in the figures.

For the measurement of SSB repair efficiency, cells were pre-treated with 10 μM of Poly (ADP-ribose) glycohydrolase inhibitor (PARGi, TOCRIS Bio-Techne, PDD 00017273) for 30 mins prior to IR exposure. As the primary reaction to detect SSBs, a PAN ADP-ribose binding reagent (rabbit, Merck, cat# 9QQ12P) was used at 1:1000 dilution in PBS-T supplemented with 1% w/v BSA. Alexa Fluor 488- or 594-conjugated rabbit secondary antibody (Invitrogen, cat# A32731, A32740, respectively) was used at 1:2000 dilution. The mean intensity of ADP-ribose in the nucleus stained with DAPI was measured by ImageJ software. At least 100 cells were counted, and the average value of ADP-ribose intensity were also calculated by Graphpad Prism 8 (GraphPad Software Inc.).

For measurement of genome instability, cells with micronuclei and chromosome bridge were counted by DAPI staining. At least 300 cells were counted by ImageJ software.

***γ-ray irradiation***

To examine the sensitivity and response to ionizing radiation (IR), cells were irradiated using $^{60}$Co γ-ray source in Tokyo Institute of Technology. The dose rate was measured using an ionizing chamber-type exposure dosimeter C-110 (Oyo Giken) and corrected for decay by calculation.

***Colony formation assay***

Survival fraction was determined by the colony formation assay. Cells were plated on 100-mm dishes. The number of the cells plated was adjusted for higher doses of indicated DNA damaging agents so as to obtain appropriate number of colonies. After incubation for 12-14 h at 37°C under 5% $CO_2$ conditions, cells were exposed to grading doses of γ-ray (1, 3 and 5 Gy), $H_2O_2$ (100 and 200 μM, for 2 h), HU (1, 2 and 4 mM, for 24 h) and CPT (5, 10 and 20 nM, for 24 h). Cells were further incubated for 10-14 days to form colonies. After washing with PBS, cells were fixed with 99.5% ethanol and stained with staining solution (0.02% w/v crystal violet; 2.5% v/v methanol). After washing the plates twice with water and drying them overnight, colonies consisting of more than 50 cells were counted manually. Plating efficiency was calculated as the number of colonies divided by the number of plated cells. Surviving fraction was calculated as the plating efficiency of irradiated cells divided by the plating efficiency of unirradiated cells. Experiments were repeated at least three times independently.

***Cell growth assay***

Cell growth and growth rate were analyzed with the number of cells at several time points (In Fig. 3B: Day 0, 1, 2, 3, 4; Fig. 3D and 3G: Day 0, 1, 4). Cells were prepared at 70–80% confluency on 100-mm dishes or 60-mm dishes and spread into 6-well plates ($2\times10^5$ cells/well). The cells were cultured at

37°C in humidified atmosphere containing 5% $CO_2$ conditions. After indicated duration of incubation, cells were harvested by trypsinization and cell numbers were counted using Coulter counter (Beckman Coulter) in all experiments.

***Cell cycle distribution analysis by flowcytometry***

The procedure for cell cycle distribution analysis was described in our recent publication (Tsuchiya et al., 2021) and appropriately modified for this study. In brief, the nascently synthesized DNA was labeled with 5-ethynyl-2′-deoxyuridine (EdU) and Alexa Fluor 488 azide through Click reaction using a Click-iT EdU Imaging kit (Life Technologies, cat# C10337) and then cells were stained with propidium iodide (PI) using a Cell Cycle Phase Determination Kit (Cayman Chemicals, cat# 10009349) according to the manufacturer's instruction. Cells grown on 6-well plates or 60-mm dish at 70-90% confluency were treated with 10 μM of EdU for 1 h and harvested by trypsinization. Harvested cells were washed with 0.1% w/v BSA/PBS and fixed/permeabilized by BD cytofix /cytoperm buffer (BD biosciences, cat# 554714) according to the manufacturer's instruction. Subsequently, cells were washed with 1x BD perm/wash buffer and resuspended in click-it reaction cocktail for 1 h at room temperature in the dark. After click-it reaction, cells were washed with 1x BD perm/wash buffer and resuspend cells in PBS containing 0.02% w/v sodium azide, 0.02% w/v RNaseA and 0.01% w/v PI for 1 h at room temperature in the dark. Cell suspension was supplemented with 500 μl of 0.1% w/v BSA/PBS for adjustment of the volume prior to the analysis and subjected to flowcytometric analysis using Cell Lab Quanta SC (Beckman Coulter).

***Measurement of RPA-loading at damaged single-stranded DNA***

RPA-loading assay was referred from a previous report (Shorrocks et al., 2021) and appropriately modified for this study. U2OS cells treated with indicated siRNA were incubated with 1 μM CPT for 1.5 h and harvested by trypsinization. Harvested cells were pre-extracted in PBS containing 0.2% v/v

Triton X-100 for 10 min on ice, and fixed using BD Cytofix/Cytoperm solution (BD Biosciences) according to the manufacturer's instructions. RPA2 was detected by a RPA2 antibody (mouse, 1:1000, abcam, cat# ab2175) diluted in BD Perm/Wash buffer (BD Biosciences) for 1 h at room temperature, followed by incubation with Alexa Fluor 488-conjugated mouse secondary antibody (1/1000, Invitrogen, cat# A32731) diluted in BD Perm/Wash buffer (BD Biosciences) for 1 h at room temperature in the dark. After washing in BD Perm/Wash buffer, cells were resuspended in PBS containing 0.02% w/v RNaseA, 0.02% w/v sodium azide and 0.01% w/v PI for 30 min at room temperature in the dark. Samples were resuspended in BD perm/wash buffer for subjecting to flowcytometric analysis using Cell Lab Quanta SC (Beckman Coulter). Cell cycle was determined by DNA contents. RPA2-loading at single-stranded DNA was quantified by subtracting the RPA2 intensity of untreated S phase cells from the RPA2 intensity of CPT-treated S phase cells in each condition.

***DNA fiber analysis***

DNA fiber assay was basically performed according to previous protocol paper (Schwab and Niedzwiedz, 2011) and appropriately modified for this study. Cells grown on 60-mm dish at 70-90% confluency were firstly labeled with 50 μM of 5-Iodo-2'-deoxyuridine (IdU) for 15 min and subsequently labeled with 250 μM of 5-Chloro-2'-deoxy-uridine (CldU) for 15 min in the humidified $CO_2$ incubator. Labeled cells were harvested by trypsinization and resuspended in ice-cold PBS at 1 x $10^6$ - 1 x $10^7$ cells/ml concentration. Two μl of cell suspension was spotted on one end of the glass slides (Matsunami glass, cat# S8215) and air-dried for 5 min. Seven μl of DNA fiber lysis buffer (200 mM Tris-HCl, pH 7.5, 50 mM EDTA, 0.5% w/v SDS,) were added on the cell suspension, gently stirred with the pipette tip, and incubated for 2 min. Glass slides were tilted to 15° to allow the fibers to spread along the slide, and air-dried once the fiber solution has been reached at the end of the glass slide. Glass slide were immersed in fixative (75% v/v methanol and 25% v/v acetic acid) and incubated

for 10 min. After two times washing with distilled water, glass slides were immersed in 2.5 M HCl for 80 min followed by three times wash with PBS and blocking with 5% w/v BSA/PBS for 30 min. For primary antibody reaction, anti-BrdU (mouse, 1/100, BD biosciences, cat# 347580, react to IdU) and anti-BrdU (rat, 1/400, abcam, cat# ab6326, react to CIdU) diluted in 5% w/v BSA/PBS were used and incubated in a humidified case for 2 h, After three times wash with PBS, goat anti-rat Alexa Fluor 488 (Invitrogen, 1/1000, cat# A110060) and goat anti-mouse Alexa Fluor 594 (Invitrogen, 1/1000, cat# A11005) were put onto the glass slides for secondary antibody reaction and incubated for 1 h in the dark. Glass slides were subsequently washed with PBS-T three times and mounted with mounting medium (Dako). For the observation of DNA fiber, OLYMPUS IX71 (OLYMPUS) or Zeiss LSM880 (Carl Zeiss) fluorescence microscope were used and at least 50 fibers were measured in each experiment. Tract length of DNA fibers were measured using ImageJ software and analyzed using GraphPad Prism 8 (GraphPad Software Inc.).

***Protein-DNA binding assay***

EpiQuik Colorimetric General Protein-DNA Binding Assay Kit (Epigentek Inc., cat# P-2004-96) was used for measuring DNA binding ability of PNKP WT, T118A and T118D, according to the manufacturers' instruction. Nuclear extracts were harvested from U2OS PNKP$^{-/-}$ cells transfected with GFP-PNKP WT, T118A or T118D expression vectors in non-denaturing lysis buffer (150 mM KCl, 50 mM Tris HCl, pH 8.3, 1 mM EDTA and 1 mM DTT) supplemented with protease inhibitors. To measure the DNA binding ability to the gapped DNA, a biotinylated oligonucleotide (BioF20: 5'-Biotin/TAGCACCTACCGATTGTATG/Phos-3') and a non-biotinylated oligonucleotide (F15: 5'/TACGTTTTTGTGTCG/3') were annealed to a complementary strand oligonucleotide (R36: 5'-Phos/CGACACAAAAACGTATCATACAATCGGTAGGTGCTA/3') in annealing buffer (10 mM Tris, pH 7.5, 50 mM NaCl and 1 mM EDTA). Oligonucleotides in annealing buffer were incubated at 95 for 5min and cooled down slowly. Twenty ng of biotinylated double-stranded oligonucleotides

and 10 μg of nuclear extracts were applied for the DNA binding reaction in the streptavidin-coated tubes. One μg/ml GFP antibody (Nacali Tesque Inc., cat# GF200) and 0.5 μg/ml HRP-conjugated mouse antibody (Dako, cat# P0447) were used for detecting DNA-binding proteins. The absorbance at 450 nm wavelength was measured by iMark Microplate Absorbance Reader (Bio-Rad Laboratories).

***Statistical analysis***

Statistical analysis was performed using either GraphPad Prism 8 (GraphPad Software Inc.) or Microsoft Excel. Unpaired (two-tailed) t-test was applied to analyze the statistical significance of difference between two experimental groups. Sample scales are indicated in figure legends. All experiments were independently performed at least two times, with similar results. In all experiments: not significant (ns) is defined as $p > 0.05$, *: $0.01 < p \leqq 0.05$; **: $0.005 < p \leqq 0.01$; ***: $0.001 < p \leqq 0.005$; ****: $0.0005 < p \leqq 0.001$.

## Acknowledgements

Authors thank Isao Yoda at $Co^{60}$ radiation center, Hiromi Yanagihara, Hirofumi Nakano, and Daisuke Morishita for technical assistance, and Matsumoto laboratory member for critical discussion. This work was supported by The Uehara Memorial Foundation [to MS], Takeda Science Foundation [to MS], Kato Memorial Bioscience Foundation [to MS], Japan Atomic Energy Agency [to MS], and Chubu Electric Power [to MS], Tokyo Tech Academy for Co-creative Education of Environment and Energy Science [to KT], Tokyo Tech Academy for Leadership [to KT], Grant-in-Aid for Scientific Research from Japan Society for the Promotion of Science [Grant Numbers JP15H02817, JP17K20042, JP20H04334 to YM and JP18K11642 to MI], Grant-in-Aid for Japan Society for the Promotion of Science Fellows [Grant Number JP20J13601 to KT] and Radiation Effects Association

[to MI].

## Author contributions

K. T. and M. S. designed this study. K.T. constructed plasmid DNA vectors. K. T., R. I., K. S., M. S., N. K., T. M., and M. S. performed cell biology experiments. K.T. and M.I. carried out laser micro-irradiation assays. K.T., Y.M. and M.S. purified PNKP phosphorylation-specific antibodies. M. I., Y, M., and M. S. supervised and provided advice. K. T. and M. S. wrote the manuscript with comments from the authors.

## Competing interests

The authors declare no competing interests.

## References

Bermudez-Guzman, L., Jimenez-Huezo, G., Arguedas, A., and Leal, A. (2020). Mutational survivorship bias: The case of PNKP. PLoS One *15*, e0237682.

Blackford, A.N., and Jackson, S.P. (2017). ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. Mol Cell *66*, 801-817.

Bras, J., Alonso, I., Barbot, C., Costa, M.M., Darwent, L., Orme, T., Sequeiros, J., Hardy, J., Coutinho, P., and Guerreiro, R. (2015). Mutations in PNKP cause recessive ataxia with oculomotor apraxia type 4. Am J Hum Genet *96*, 474-479.

Breslin, C., and Caldecott, K.W. (2009). DNA 3'-phosphatase activity is critical for rapid global rates of single-strand break repair following oxidative stress. Mol Cell Biol *29*, 4653-4662.

Bukanov, N.O., Smith, L.A., Klinger, K.W., Ledbetter, S.R., and Ibraghimov-Beskrovnaya, O. (2006). Long-lasting arrest of murine polycystic kidney

disease with CDK inhibitor roscovitine. Nature *444*, 949-952.

Burhans, W.C., Vassilev, L.T., Wu, J., Sogo, J.M., Nallaseth, F.S., and DePamphilis, M.L. (1991). Emetine allows identification of origins of mammalian DNA replication by imbalanced DNA synthesis, not through conservative nucleosome segregation. EMBO J *10*, 4351-4360.

Chalasani, S.L., Kawale, A.S., Akopiants, K., Yu, Y., Fanta, M., Weinfeld, M., and Povirk, L.F. (2018). Persistent 3'-phosphate termini and increased cytotoxicity of radiomimetic DNA double-strand breaks in cells lacking polynucleotide kinase/phosphatase despite presence of an alternative 3'-phosphatase. DNA Repair (Amst) *68*, 12-24.

Chappell, C., Hanakahi, L.A., Karimi-Busheri, F., Weinfeld, M., and West, S.C. (2002). Involvement of human polynucleotide kinase in double-strand break repair by non-homologous end joining. EMBO J *21*, 2827-2832.

Chiang, T.W., le Sage, C., Larrieu, D., Demir, M., and Jackson, S.P. (2016). CRISPR-Cas9(D10A) nickase-based genotypic and phenotypic screening to enhance genome editing. Sci Rep *6*, 24356.

Cortez, D., Guntuku, S., Qin, J., and Elledge, S.J. (2001). ATR and ATRIP: partners in checkpoint signaling. Science *294*, 1713-1716.

Cortez, D., Wang, Y., Qin, J., and Elledge, S.J. (1999). Requirement of ATM-dependent phosphorylation of brca1 in the DNA damage response to double-strand breaks. Science *286*, 1162-1166.

Exell, J.C., Thompson, M.J., Finger, L.D., Shaw, S.J., Debreczeni, J., Ward, T.A., McWhirter, C., Sioberg, C.L., Martinez Molina, D., Abbott, W.M., *et al.* (2016). Cellularly active N-hydroxyurea FEN1 inhibitors block substrate entry to the active site. Nat Chem Biol *12*, 815-821.

Fung, T.K., Ma, H.T., and Poon, R.Y. (2007). Specialized roles of the two mitotic cyclins in somatic cells: cyclin A as an activator of M phase-promoting factor. Mol Biol Cell *18*, 1861-1873.

Genois, M.M., Gagne, J.P., Yasuhara, T., Jackson, J., Saxena, S., Langelier, M.F., Ahel, I., Bedford, M.T., Pascal, J.M., Vindigni, A., *et al.* (2021). CARM1 regulates replication fork speed and stress response by stimulating PARP1. Mol Cell *81*, 784-800 e788.

Hanzlikova, H., Kalasova, I., Demin, A.A., Pennicott, L.E., Cihlarova, Z., and Caldecott, K.W. (2018). The Importance of Poly(ADP-Ribose) Polymerase as a Sensor of Unligated Okazaki Fragments during DNA Replication. Mol Cell *71*, 319-331 e313.

Hoch, N.C., Hanzlikova, H., Rulten, S.L., Tetreault, M., Komulainen, E., Ju, L., Hornyak, P., Zeng, Z., Gittens, W., Rey, S.A., *et al.* (2017). XRCC1 mutation is associated with PARP1 hyperactivation and cerebellar ataxia. Nature *541*, 87-91.

Honda, R., Lowe, E.D., Dubinina, E., Skamnaki, V., Cook, A., Brown, N.R., and Johnson, L.N. (2005). The structure of cyclin E1/CDK2: implications for CDK2 activation and CDK2-independent roles. EMBO J *24*, 452-463.

Hornbeck, P.V., Zhang, B., Murray, B., Kornhauser, J.M., Latham, V., and Skrzypek, E. (2015). PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. Nucleic Acids Res *43*, D512-520.

Inano, S., Sato, K., Katsuki, Y., Kobayashi, W., Tanaka, H., Nakajima, K., Nakada, S., Miyoshi, H., Knies, K., Takaori-Kondo, A., *et al.* (2017). RFWD3-Mediated Ubiquitination Promotes Timely Removal of Both RPA and RAD51 from DNA Damage Sites to Facilitate Homologous Recombination. Mol Cell *66*, 622-634 e628.

Jilani, A., Ramotar, D., Slack, C., Ong, C., Yang, X.M., Scherer, S.W., and Lasko, D.D. (1999). Molecular cloning of the human gene, PNKP, encoding a polynucleotide

kinase 3'-phosphatase and evidence for its role in repair of DNA strand breaks caused by oxidative damage. J Biol Chem *274*, 24176-24186.

Kalasova, I., Hailstone, R., Bublitz, J., Bogantes, J., Hofmann, W., Leal, A., Hanzlikova, H., and Caldecott, K.W. (2020). Pathological mutations in PNKP trigger defects in DNA single-strand break repair but not DNA double-strand break repair. Nucleic Acids Res.

Kalasova, I., Hanzlikova, H., Gupta, N., Li, Y., Altmuller, J., Reynolds, J.J., Stewart, G.S., Wollnik, B., Yigit, G., and Caldecott, K.W. (2019). Novel PNKP mutations causing defective DNA strand break repair and PARP1 hyperactivity in MCSZ. Neurol Genet *5*, e320.

Kamdar, R.P., and Matsumoto, Y. (2010). Radiation-induced XRCC4 association with chromatin DNA analyzed by biochemical fractionation. J Radiat Res *51*, 303-313.

Karimi-Busheri, F., Daly, G., Robins, P., Canas, B., Pappin, D.J., Sgouros, J., Miller, G.G., Fakhrai, H., Davis, E.M., Le Beau, M.M., *et al.* (1999). Molecular characterization of a human DNA kinase. J Biol Chem *274*, 24187-24194.

Kolas, N.K., Chapman, J.R., Nakada, S., Ylanko, J., Chahwan, R., Sweeney, F.D., Panier, S., Mendez, M., Wildenhain, J., Thomson, T.M., *et al.* (2007). Orchestration of the DNA-damage response by the RNF8 ubiquitin ligase. Science *318*, 1637-1640.

Kunkel, T.A. (2004). DNA replication fidelity. J Biol Chem *279*, 16895-16898.

Laipis, P.J., and Levine, A.J. (1973). DNA replication in SV40-infected cells. IX. The inhibition of a gap-filling step during discontinuous synthesis of SV40 DNA. Virology *56*, 580-594.

Laney, J.D., and Hochstrasser, M. (1999). Substrate targeting in the ubiquitin system. Cell *97*, 427-430.

Leonard, A.C., and Mechali, M. (2013). DNA replication origins. Cold Spring Harb

Perspect Biol *5*, a010116.

Mackenzie, K.J., Carroll, P., Martin, C.A., Murina, O., Fluteau, A., Simpson, D.J., Olova, N., Sutcliffe, H., Rainger, J.K., Leitch, A., *et al.* (2017). cGAS surveillance of micronuclei links genome instability to innate immunity. Nature *548*, 461-465.

Magnusson, G. (1973a). Hydroxyurea-induced accumulation of short fragments during polyoma DNA replication. I. Characterization of fragments. J Virol *12*, 600-608.

Magnusson, G. (1973b). Hydroxyurea-induced accumulation of short fragments during polyoma DNA replication. II. Behavior during incubation of isolated nuclei. J Virol *12*, 609-615.

Mani, R.S., Mermershtain, I., Abdou, I., Fanta, M., Hendzel, M.J., Glover, J.N.M., and Weinfeld, M. (2019). Domain analysis of PNKP-XRCC1 interactions: Influence of genetic variants of XRCC1. J Biol Chem *294*, 520-530.

Mani, R.S., Yu, Y., Fang, S., Lu, M., Fanta, M., Zolner, A.E., Tahbaz, N., Ramsden, D.A., Litchfield, D.W., Lees-Miller, S.P., *et al.* (2010). Dual modes of interaction between XRCC4 and polynucleotide kinase/phosphatase: implications for nonhomologous end joining. J Biol Chem *285*, 37619-37629.

Maya-Mendoza, A., Moudry, P., Merchut-Maya, J.M., Lee, M., Strauss, R., and Bartek, J. (2018). High speed of fork progression induces DNA replication stress and genomic instability. Nature *559*, 279-284.

Meijer, L., Borgne, A., Mulner, O., Chong, J.P., Blow, J.J., Inagaki, N., Inagaki, M., Delcros, J.G., and Moulinoux, J.P. (1997). Biochemical and cellular effects of roscovitine, a potent and selective inhibitor of the cyclin-dependent kinases cdc2, cdk2 and cdk5. Eur J Biochem *243*, 527-536.

Mochizuki, A.L., Katanaya, A., Hayashi, E., Hosokawa, M., Moribe, E., Motegi, A., Ishiai, M., Takata, M., Kondoh, G., Watanabe, H., *et al.* (2017). PARI Regulates Stalled

Replication Fork Processing To Maintain Genome Stability upon Replication Stress in Mice. Mol Cell Biol *37*.

O'Donnell, M., Langston, L., and Stillman, B. (2013). Principles and concepts of DNA replication in bacteria, archaea, and eukarya. Cold Spring Harb Perspect Biol *5*.

Okazaki, R., Okazaki, T., Sakabe, K., Sugimoto, K., and Sugino, A. (1968). Mechanism of DNA chain growth. I. Possible discontinuity and unusual secondary structure of newly synthesized chains. Proc Natl Acad Sci U S A *59*, 598-605.

Pagano, M., Pepperkok, R., Verde, F., Ansorge, W., and Draetta, G. (1992). Cyclin A is required at two points in the human cell cycle. EMBO J *11*, 961-971.

Parsons, J.L., Khoronenkova, S.V., Dianova, II, Ternette, N., Kessler, B.M., Datta, P.K., and Dianov, G.L. (2012). Phosphorylation of PNKP by ATM prevents its proteasomal degradation and enhances resistance to oxidative stress. Nucleic Acids Res *40*, 11404-11415.

Pedroso, J.L., Rocha, C.R., Macedo-Souza, L.I., De Mario, V., Marques, W., Jr., Barsottini, O.G., Bulle Oliveira, A.S., Menck, C.F., and Kok, F. (2015). Mutation in PNKP presenting initially as axonal Charcot-Marie-Tooth disease. Neurol Genet *1*, e30.

Ray Chaudhuri, A., Hashimoto, Y., Herrador, R., Neelsen, K.J., Fachinetti, D., Bermejo, R., Cocito, A., Costanzo, V., and Lopes, M. (2012). Topoisomerase I poisoning results in PARP-mediated replication fork reversal. Nat Struct Mol Biol *19*, 417-423.

Ray Chaudhuri, A., and Nussenzweig, A. (2017). The multifaceted roles of PARP1 in DNA repair and chromatin remodelling. Nat Rev Mol Cell Biol *18*, 610-621.

Reynolds, J.J., Walker, A.K., Gilmore, E.C., Walsh, C.A., and Caldecott, K.W. (2012). Impact of PNKP mutations associated with microcephaly, seizures and developmental delay on enzyme activity and DNA strand break repair. Nucleic Acids Res *40*, 6608-6619.

Schwab, R.A., and Niedzwiedz, W. (2011). Visualization of DNA replication in the

vertebrate model system DT40 using the DNA fiber technique. J Vis Exp, e3255.

Segal-Raz, H., Mass, G., Baranes-Bachar, K., Lerenthal, Y., Wang, S.Y., Chung, Y.M., Ziv-Lehrman, S., Strom, C.E., Helleday, T., Hu, M.C., *et al.* (2011). ATM-mediated phosphorylation of polynucleotide kinase/phosphatase is required for effective DNA double-strand break repair. EMBO Rep *12*, 713-719.

Shen, J., Gilmore, E.C., Marshall, C.A., Haddadin, M., Reynolds, J.J., Eyaid, W., Bodell, A., Barry, B., Gleason, D., Allen, K., *et al.* (2010). Mutations in PNKP cause microcephaly, seizures and defects in DNA repair. Nat Genet *42*, 245-249.

Shorrocks, A.K., Jones, S.E., Tsukada, K., Morrow, C.A., Belblidia, Z., Shen, J., Vendrell, I., Fischer, R., Kessler, B.M., and Blackford, A.N. (2021). The Bloom syndrome complex senses RPA-coated single-stranded DNA to restart stalled replication forks. Nat Commun *12*, 585.

Siddiqui, K., On, K.F., and Diffley, J.F. (2013). Regulating DNA replication in eukarya. Cold Spring Harb Perspect Biol *5*.

Swaffer, M.P., Jones, A.W., Flynn, H.R., Snijders, A.P., and Nurse, P. (2016). CDK Substrate Phosphorylation and Ordering the Cell Cycle. Cell *167*, 1750-1761 e1716.

Thakar, T., Leung, W., Nicolae, C.M., Clements, K.E., Shen, B., Bielinsky, A.K., and Moldovan, G.L. (2020). Ubiquitinated-PCNA protects replication forks from DNA2-mediated degradation by regulating Okazaki fragment maturation and chromatin assembly. Nat Commun *11*, 2147.

Tsuchiya, H., Shimada, M., Tsukada, K., Meng, Q., Kobayashi, J., and Matsumoto, Y. (2021). Diminished or inversed dose-rate effect on clonogenic ability in Ku-deficient rodent cells. J Radiat Res *62*, 198-205.

Tsukada, K., Matsumoto, Y., and Shimada, M. (2020a). Linker region is required for efficient nuclear localization of polynucleotide kinase phosphatase. PLoS One *15*,

e0239404.

Tsukada, K., Shimada, M., Imamura, R., Saikawa, K., Ishiai, M., and Matsumoto, Y. (2020b). The FHA domain of PNKP is essential for its recruitment to DNA damage sites and maintenance of genome stability. Mutat Res *822*, 111727.

Ward, T.A., McHugh, P.J., and Durant, S.T. (2017). Small molecule inhibitors uncover synthetic genetic interactions of human flap endonuclease 1 (FEN1) with DNA damage response genes. PLoS One *12*, e0179278.

Wold, M.S. (1997). Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. Annu Rev Biochem *66*, 61-92.

Zheng, L., and Shen, B. (2011). Okazaki fragment maturation: nucleases take centre stage. J Mol Cell Biol *3*, 23-30.

Zolner, A.E., Abdou, I., Ye, R., Mani, R.S., Fanta, M., Yu, Y., Douglas, P., Tahbaz, N., Fang, S., Dobbs, T., *et al.* (2011). Phosphorylation of polynucleotide kinase/phosphatase by DNA-dependent protein kinase and ataxia-telangiectasia mutated regulates its association with sites of DNA damage. Nucleic Acids Res *39*, 9224-9237.

Zou, L., and Elledge, S.J. (2003). Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. Science *300*, 1542-1548.

## Figure legends

**Fig. 1. PNKP is important for the genome stability**

A: Representative images of normal nucleus, a micronucleus and chromosome bridge. The micronucleus and chromosome bridge are indicated by white arrows.

B and C: Quantified results of Fig. 1A. Luciferase- (as control), PNKP-, XRCC1- and

XRCC4-depleted U2OS cells. DNA were stained by DAPI at 24 h after each treatment. Cells with micronucleus (B) and chromosome bridge (C) were counted and graphed. At least 200 cells were counted.

D and E: Flowcytometric analysis for measurement of RPA loading on single-stranded DNA Luciferase- (con) and PNKP-depleted U2OS cells were used. The cells were pre-extracted and stained with Propidium Iodide (PI) and RPA2 antibody. Cell cycle and RPA intensity were analyzed by flowcytometry. RPA intensity in S phase was graphed. At least 20,000 counts were measured.

In all panels, error bars represent SEM.

**Fig. 2. PNKP knockout cells showed high sensitivity to the genotoxic stress**

A: Protein expression analysis of PNKP knockout U2OS cells ($PNKP^{-/-}$ cells) generated by CRISPR/Cas9 genome editing. Protein expression of PNKP in $PNKP^{-/-}$ clone 1 (C1) and clone 2 (C2) were confirmed by western blotting.

B: DNA double-strand break repair abilities in $PNKP^{-/-}$ C1 cells measured by western blotting. DNA double-strand breaks were analyzed by pS824-KAP1 antibody after indicated time points recovered from IR 5Gy exposure.

C: Measurement of DNA single-strand break repair abilities of $PNKP^{-/-}$ cells. Single-strand breaks were analyzed by immunofluorescence using PAN-ADP-ribose binding reagents at 30min after IR 2Gy exposure in U2OS WT and $PNKP^{-/-}$ (C1 and C2) cells treated with 10 μM PARGi for 30min prior to IR exposure.

D and E: Cellular sensitivity of $PNKP^{-/-}$ cells to the DNA damages, especially DNA strand breaks, was measured by colony formation assay. Cells were treated by hydrogen peroxide ($H_2O_2$: D) and IR (E) exposure at indicated dose.

F and G: Cellular sensitivity of $PNKP^{-/-}$ cells to the DNA replication stress was measured by colony formation assay. Cells were treated by HU (G) and CPT (H) at indicated dose.

In all panels, scale bar indicates 10 μm and error bars represent SEM.

**Fig. 3. Linker region of PNKP is required for cell proliferation and DNA replication**

A: Flowcytometric analysis for cell cycle distribution of U2OS WT and PNKP$^{-/-}$ C1 and C2 cells. The cells were stained with PI, Synthesized DNA was labeled by EdU. Percentage of each cell cycle is shown in vertical axis, cell types are shown in horizontal axis.

B: Representative images and quantified results of DNA fiber length in U2OS WT and PNKP$^{-/-}$ C1 and C2 cells. Nascent synthesized DNA was labeled by CIdU.

C: Measurement of growth rate of U2OS WT and of PNKP$^{-/-}$ C1 and C2 cells. Cell numbers (shown in vertical axis) were counted at indicated time points (shown in horizontal axis).

D: Schematic diagrams of the structure of GFP -tagged human PNKP WT and deletion mutants (D1-D4).

E: Cell growth rate in U2OS WT and PNKP deletion mutant expressing PNKP$^{-/-}$ C1 cells. Cell growth rate was normalized by GFP expressing U2OS cells.

F: Representative images and quantified results of DNA fiber length in U2OS WT and PNKP$^{-/-}$ C1 cells expressing indicated PNKP deletion mutants. Nascent synthesized DNA was labeled by CIdU.

G: Alignment of amino acid sequences in linker region among mammalian species. SQ/TQ motif (blue) is PI3 kinase substrate motif. TP (red) is CDK substrate motif.

H: Cell growth rate in U2OS WT and PNKP$^{-/-}$ C1 cells expressing indicated point mutants. Cell growth rate was normalized by GFP expressing U2OS cells.

In all panels, error bars represent SEM.

**Fig. 4. CDKs mediated PNKP phosphorylation at T118 is required for its RPA binding**

A: Analysis of immunoprecipitation (IP) between GFP-trap magnetic agarose beads and GFP or

GFP-PNKP in HEK293 cells. DNA-PKcs, ATM, XRCC4, and RPA2 antibodies were used for western blot analysis. 1% of whole cell lysate used in IP experiment represented as input 1% as a loading control.

B: Interaction between indicated PNKP mutants and XRCC1, XRCC4 and RPA2 was assessed by GFP-pulldown immunoprecipitation assays. Lysates from GFP, GFP-WT, R35/48A (the FHA domain), S114A, and T118A PNKP were expressed cells were immunoprecipitated with GFP-trap magnetic beads. 1% of whole cell lysate used in IP experiment represented as input 1% as a loading control.

C: Scheme and protein expression levels of GFP-PNKP expressing HCT116 cells after release from double thymidine block. After released, cells were collected at indicated time points, and used for western blotting. pT118 PNKP antibody was generated for this study. Cyclin A2 and E1 were used for cell cycle markers. GAPDH was used as a loading control.

D: T118-PNKP phosphorylation levels of GFP-PNKP expressing PNKP$^{-/-}$ C1 cells after treatment of 10mM CDKs inhibitor, roscovitine, for 1h. Cells were lysed and applied for western blotting. GFP and PNKP antibodies were used for confirming GFP-PNKP expression levels. Cyclin A2 antibody was used as a cell cycle marker. pS15-p53 antibody was used for confirming CDKs inhibition. pS114-PNKP antibody, generated in this study, was used for confirming pT118 specificity against CDKs inhibition. PCNA antibody was used as a loading control.

E: HCT116 cells were lysed at 5 days and 3 days after transfection with GFP-PNKP and co-transfection with mCherry2-Cyclin A2, and 3XFLAG-CDK1, CDK2 or CDK4, respectively, and applied for western blotting. pT118 PNKP level was analyzed by pT118 PNKP specific antibody. PNKP expression was analyzed with PNKP or GFP antibodies. Cyclin A2 expression was analyzed with RFP antibody. Cyclin E antibody was used as cell cycle marker. KAP1 antibody was used as loading control.

**Fig. 5. Phosphorylation of PNKP at T118 is required for Okazaki fragments maturation**

A: Measurement of SSBs, detected by ADP-ribose intensity, in U2OS WT and PNKP$^{-/-}$ C1 cells under HU or EME treatment. Cells were pre-treated with PARGi for 30 min prior to the treatment of DNA replication inhibitors. ADP-ribose intensities were normalized by the intensity of U2OS NT.

B and C: Representative images (B) and quantified results (C) of measurement of ADP-ribose intensity in U2OS WT and PNKP WT, PNKP T118 expressing PNKP$^{-/-}$ C1 cells. Synthesized DNA was labelled by EdU and EdU-positive cells were defined as S phase (grey) and the other cells were defined as G1+G2/M phase (white).

D and E: Scheme, representative images (D) and quantified results (E) of the experiments for measurement of ADP-ribose intensity in U2OS WT and PNKP$^{-/-}$ C1 cells expressing GFP, PNKP WT, and PNKP T118A mutant under DMSO (negative control) and FEN1i treatment only in EdU-positive (S phase) cells.

F: Measurement of DNA synthesis speed in U2OS WT cells under non-treatment (NT), 2 mM HU and 10 μM FEN1 inhibitor treatment analyzed by DNA fiber assay. At least 100 DNA fibers were measured.

G and H: Measurement of DNA synthesis speed in U2OS WT and PNKP$^{-/-}$ C1 cells expressing GFP (con), PNKP, T118A (G), or T118D (H) mutants under 10 μM FEN1 inhibitor treatment analyzed by DNA fiber assay. At least 100 DNA fibers were measured.

In all panels, scale bar indicates 10 μm and error bars represent SEM.

**Fig. 6. Phosphatase activity of PNKP is essential for end-processing of Okazaki fragments**

A: Protein expression analysis for PNKP phosphatase-dead (D171A) and kinase-dead (K378A) mutants in U2OS WT and PNKP$^{-/-}$ C1 cells expressing GFP, PNKP WT, D171A and K378A mutants confirmed by western blotting. KAP1 antibody was used as a loading control.

B: Measurement of the speed of DNA synthesis in U2OS WT and PNKP$^{-/-}$ C1 cells expressing GFP (con), PNKP, D171A and K378A mutants under 10 μM FEN1 inhibitor treatment analyzed by DNA

fiber assay. At least 100 DNA fibers were measured.

C and D: Representative images (C) and quantified results (D) of the experiments for measurement of ADP-ribose intensity in U2OS WT and PNKP$^{-/-}$ C1 cells expressing GFP, PNKP WT, PNKP D171A and PNKP K378A mutants under DMSO (negative control) and FEN1i treatment only in EdU-positive (S phase) cells.

In all panels, scale bar indicates 10 μm and error bars represent SEM.

**Fig. 7. Phosphorylation of PNKP at T118 is required for genome stability**

A and B: Measurement of the formations of micronuclei and chromosome bridges in U2OS WT and PNKP$^{-/-}$ C1 cells expressing GFP, PNKP WT, and PNKP T118A mutant. DNA were stained by DAPI at 24 h after each treatment. Cells with micronucleus (A) and chromosome bridge (B) were counted and graphed. At least 200 cells were counted. Error bars represent SEM.

C: Measurement of the gapped DNA binding ability of PNKP WT, T118A, and T118D mutants. Nuclear extracts were harvested from U2OS PNKP$^{-/-}$ C1 cells at 2 days after transfection with GFP-PNKP WT, T118A or T118D mutants. GFP antibody was used to detect GFP-PNKPs bound to the gapped double-stranded DNA.

D: Schematic mechanisms of the involvement of PNKP in DNA replication, especially in Okazaki fragments (OFs) maturation. T118 of PNKP is phosphorylated by CDKs for the recruitment to the OFs gaps via the interaction with RPA2. OFs ends are processed by PNKP and become the ligatable ends prior to the OFs maturation. In PNKP-deficient, OFs gaps become SSBs, and the SSBs are sensed by PARP for proceeding PARP-dependent gap-filling pathway which required SSB repair proteins, XRCC1, LIG3 and PNKP. Impaired PARP-dependent gap filling leads to the high-speed DNA synthesis and the genome instability.

## Supplementary figure legends

**Fig. S1. Related to Fig. 2. PNKP knockout cells showed high sensitivity to the genotoxic stress**

A: Targeting gene locus of CRISPR/Cas9 genome editing for generation of PNKP$^{-/-}$ cells. PNKP L105 and P119 located on exon 4 were targeted by Cas9 D10A.

B: DNA sequencing results in PNKP exon 4 of U2OS WT, PNKP$^{-/-}$ C1 and C2. DNA sequences of all allele were aligned and PAM and sgRNA sequences were indicated. C1 and C2 have 2 and 3alleles, respectively, and all reading frames were frame-shifted. “Purple”, “Light blue” and “Red” characters indicated mutations, protospacer adjacent motif (PAM), single-guide RNA (sgRNA) sequences, respectively. “-” indicates a deletion of nucleotide.

C: Flowcytometric analysis of cell cycle distribution in U2OS WT, PNKP$^{-/-}$ C1 and C2 cells. Nascent synthesized DNA was detected by EdU (vertical axis) and whole DNA was stained by PI (horizontal axis). Each cell cycle was determined by gating of broken lines (Blue: G1 phase, Light green: S phase, Red: G2/M phase). Quantified result was shown in Fig. 2F.

**Fig. S2. Related to Fig. 3. Linker region of PNKP is required for cell proliferation and DNA replication**

A: Protein expression analysis of U2OS WT and PNKP$^{-/-}$ C1 cells transiently expressing GFP or indicated PNKP deletion mutants. GFP antibody was used for confirming the expression of exogenous PNKP deletion mutants. PNKP antibody was used for comparing expression levels of endogenous and exogenous PNKP. KAP1 antibody was used as a loading control. “*” indicates endogenous PNKP.

B: Protein expression analysis of U2OS WT and PNKP$^{-/-}$ C1 cells transiently expressing GFP or indicated PNKP point mutants. GFP antibody was used for confirming the expression of exogenous PNKP point mutants. PNKP antibody was used for comparing expression levels of endogenous and exogenous PNKP. KAP1 antibody was used as a loading control. “*” indicates endogenous PNKP.

**Fig. S3. Related to Fig. 3. Linker region of PNKP is required for cell proliferation and DNA replication**

A: Representative live-cell images of the GFP-tagged PNKP S114 substituted mutants (S114A and S114D) after laser micro-irradiation. Micro-laser was irradiated on the line between red arrows. Cells were sensitized by Hoechst 33258 at 30 min prior to the laser micro-irradiation. Scale bar indicates as 10 μm.

B: Relative green fluorescence intensity of indicated PNKP S114 substituted mutants after laser micro-irradiation. Green fluorescence intensity was scored every 0.25 s up to 60 s. 20 cells were scored at each point. Error bars represent SEM.

C: Representative live-cell images of the GFP-tagged PNKP T118 substituted mutants (T118A and T118D) after laser micro-irradiation. Micro-laser was irradiated on the line between red arrows. Scale bar indicates as 10 μm.

D: Relative green fluorescence intensity of indicated PNKP T118 substituted mutants (T118A and T118D) after laser micro-irradiation. 20 cells were scored at each point. Error bars represent SEM.

E, F and G: Relative green fluorescence intensity of indicated PNKP point substituted mutants (T122A (E), S126A (F) and S143A (G)) after laser micro-irradiation. 20 cells were scored at each point. Error bars represent SEM.

**Fig. S4. Related to Fig. 4. CDK-dependent PNKP phosphorylation at T118 is required for its RPA binding**

A: Analysis of immunoprecipitation between GFP-trap magnetic agarose beads and GFP or GFP-PNKP in HEK293 cells. Immunoprecipitated results were shown in Fig. 4B. GFP antibody was used for western blotting. 1% of whole cell lysate used in IP experiment represented as input 1% as a loading control.

B: Flowcytometric analysis of cell cycle distribution of HCT116 cells transiently expressing

GFP-PNKP at indicated time after release from double thymidine block. Cell cycle was determined by DNA contents measured by PI staining (horizontal axis).

**Fig. S5. Related to Fig. 4. CDK-dependent PNKP phosphorylation at T118 is required for its RPA binding**

A: Representative image of ELISA assay for pT118-PNKP antibody. Dilutions were shown in vertical axis, and ligands and antibodies were shown in horizontal axis.

B: Quantified results of Fig. S6A for confirming the titer and specificity of pT118-PNKP antibody. Absorbance was shown in vertical axis, and dilutions were shown in horizontal axis.

C: Specificity of pT118-PNKP antibody for cell lysate from U2OS WT and PNKP$^{-/-}$ C1 cells transiently expressing GFP-PNKP or GFP-PNKP T118A mutant. Asterisks indicate non-specific detections. GFP and PNKP antibodies were used for confirming the expression of GFP-PNKP and GFP-PNKP T118A mutant. GAPDH antibody was used as a loading control.

**Fig. S6. Related to Fig. 4. CDK-dependent PNKP phosphorylation at T118 is required for its RPA binding**

A: Representative image of ELISA assay for pS114-PNKP antibody. Dilutions were shown in vertical axis, and ligands and antibodies were shown in horizontal axis.

B: Quantified results of Fig. S7A for confirming the titer and specificity of pS114-PNKP antibody. Absorbance was shown in vertical axis, and dilutions were shown in horizontal axis.

**Fig. S7. Related to Fig. 7. Phosphorylation of PNKP at T118 is required for genome stability**

A: Representative images of immunofluorescence using PAN ADP-ribose binding reagents at 30 min after 2Gy IR exposure in U2OS WT and PNKP$^{-/-}$ cells transiently expressing GFP or indicated PNKP mutants. PARGi was added 30 min prior to IR exposure. Images detected by ADP-ribose, DAPI and

DIC were shown. Scale bar indicates as 10 μm.

B: Quantified result of Fig. S8A. Relative ADP-ribose intensity is shown in vertical axis and cell types were shown in horizontal axis. At least 100 cells were analyzed for the quantification. Statistical significance was indicated as not significant (ns) and ****: $0.0005 < p \leqq 0.001$.

C: Measurement of DNA DSB repair ability of indicated cells after 2Gy IR exposure. Cells were harvested at 30 min and 6 h after IR exposure. Percentage of γH2AX positive cells is shown in vertical axis and conditions were shown in horizontal axis. NT indicates non-treatment. At least 100 cells were analyzed for the quantification. Statistical significance was indicated as not significant (ns) and ****: $0.0005 < p \leqq 0.001$.

Figure 1

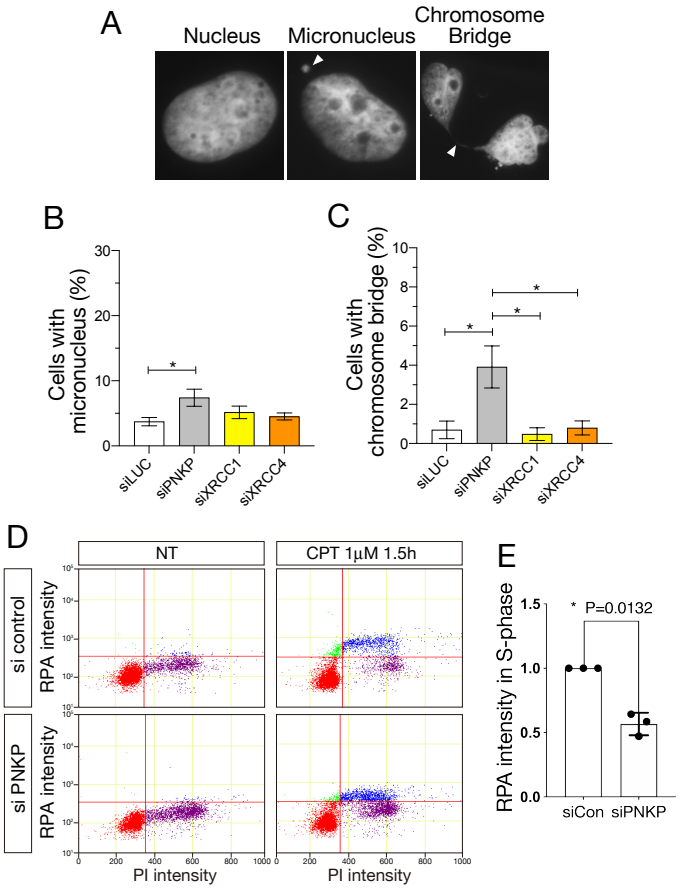

A
Nucleus
Micronucleus
Chromosome Bridge
B
Cells with micronucleus (%)
30
20
10
0
*
siLUC
siPNKP
siXRCC1
siXRCC4
C
Cells with chromosome bridge (%)
10
8
6
4
2
0
*
*
*
siLUC
siPNKP
siXRCC1
siXRCC4
D
NT
CPT 1μM 1.5h
si control
si PNKP
RPA intensity
PI intensity
E
RPA intensity in S-phase
1.5
1.0
0.5
0.0
* P=0.0132
siCon
siPNKP

Figure 2

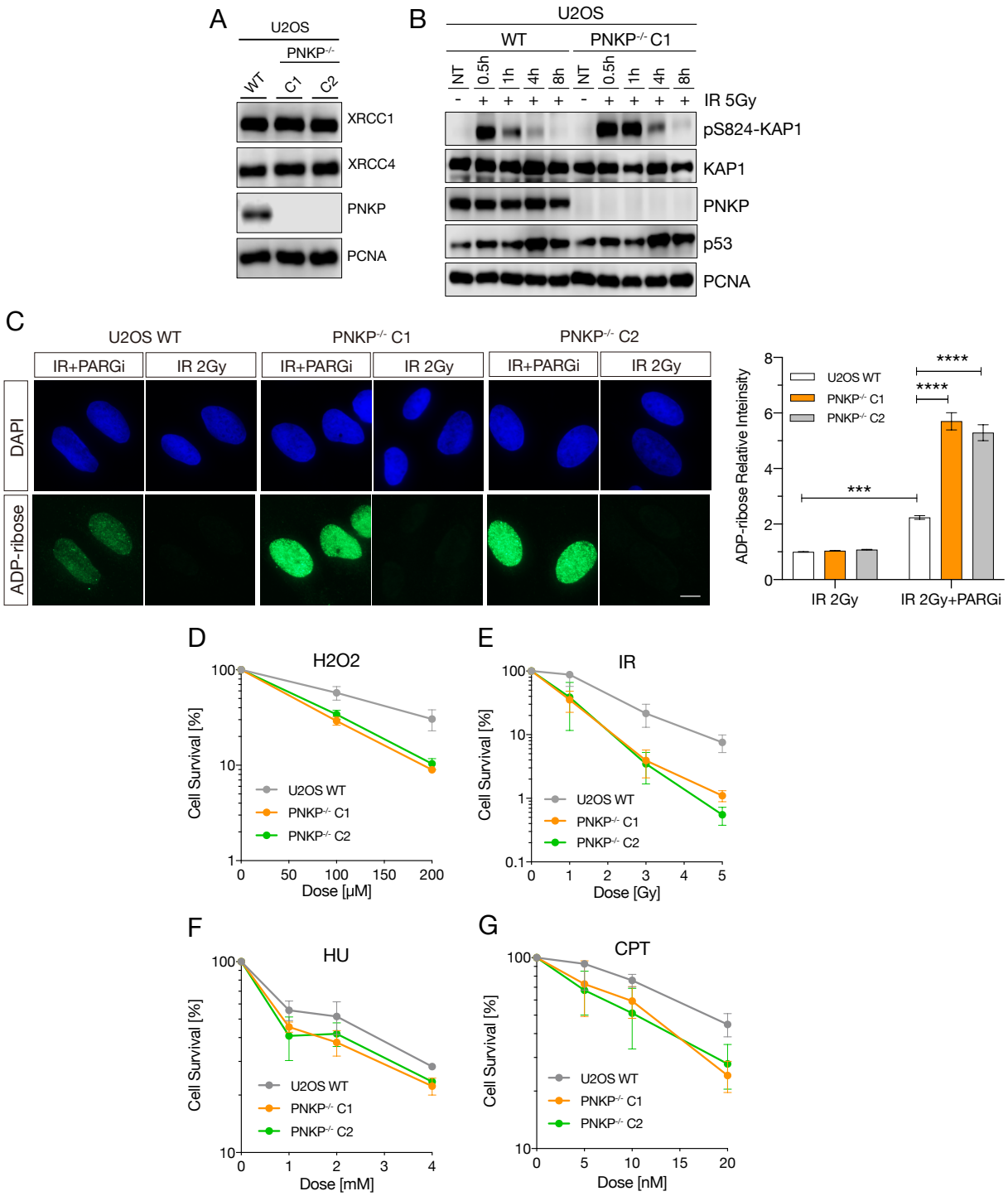

A
U2OS
PNKP-/-
WT
C1
C2
XRCC1
XRCC4
PNKP
PCNA
B
U2OS
WT
PNKP-/- C1
NT
0.5h
1h
4h
8h
IR 5Gy
pS824-KAP1
KAP1
PNKP
p53
PCNA
C
U2OS WT
PNKP-/- C1
PNKP-/- C2
IR+PARGi
IR 2Gy
DAPI
ADP-ribose
ADP-ribose Relative Inteinsity
U2OS WT
PNKP-/- C1
PNKP-/- C2
****
***
IR 2Gy
IR 2Gy+PARGi
D
H2O2
Cell Survival [%]
Dose [µM]
E
IR
Dose [Gy]
F
HU
Dose [mM]
G
CPT
Dose [nM]

Figure 3

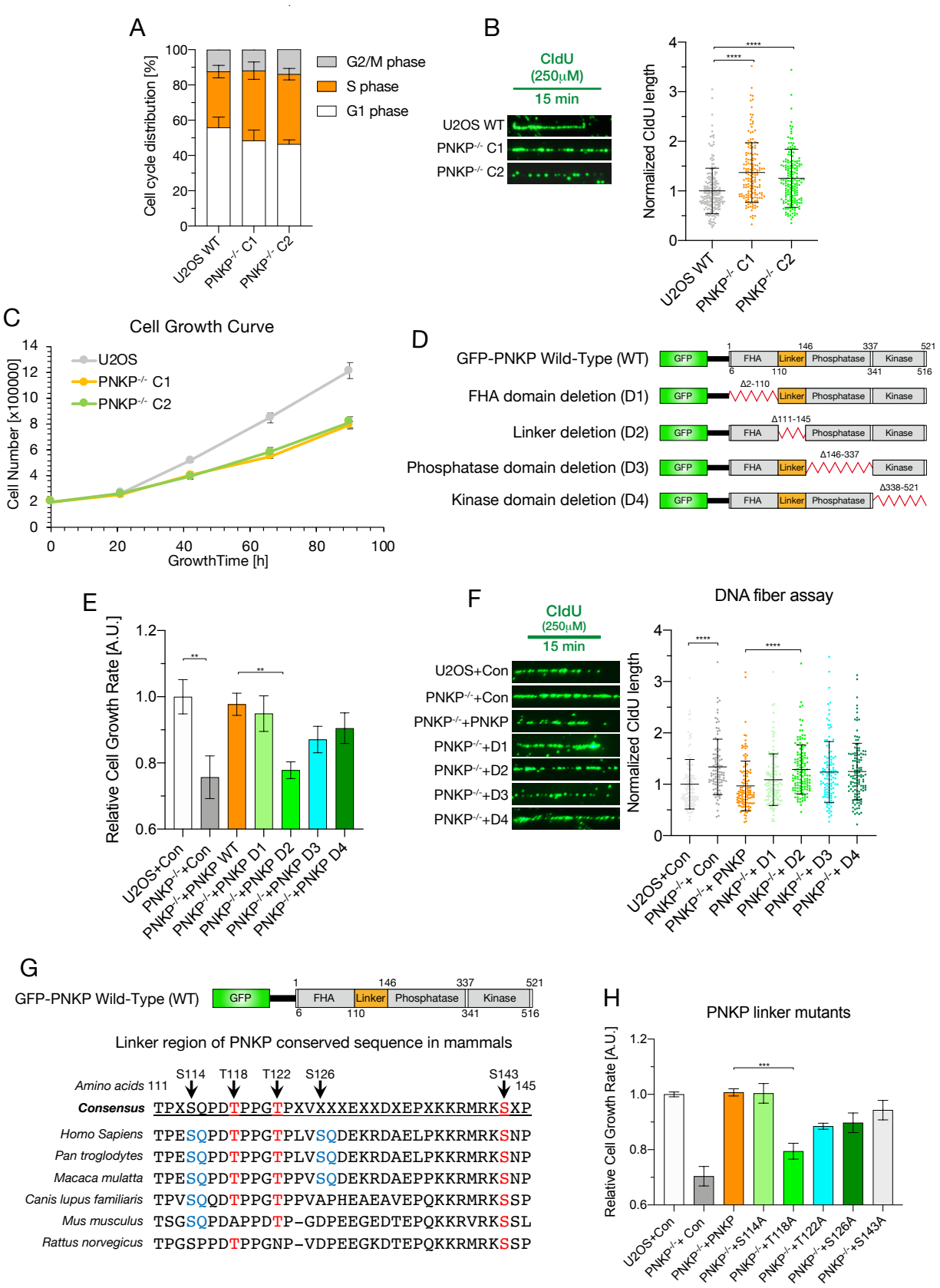

A
Cell cycle distribution [%]
G2/M phase
S phase
G1 phase
U2OS WT
PNKP-/- C1
PNKP-/- C2
B
CldU (250μM)
15 min
U2OS WT
PNKP-/- C1
PNKP-/- C2
Normalized CldU length
C
Cell Growth Curve
U2OS
PNKP-/- C1
PNKP-/- C2
Cell Number [x100000]
GrowthTime [h]
D
GFP-PNKP Wild-Type (WT)
FHA domain deletion (D1)
Linker deletion (D2)
Phosphatase domain deletion (D3)
Kinase domain deletion (D4)
GFP
FHA
Linker
Phosphatase
Kinase
Δ2-110
Δ111-145
Δ146-337
Δ338-521
E
Relative Cell Growth Rate [A.U.]
U2OS+Con
PNKP-/-+Con
PNKP-/-+PNKP WT
PNKP-/-+PNKP D1
PNKP-/-+PNKP D2
PNKP-/-+PNKP D3
PNKP-/-+PNKP D4
F
CldU (250μM)
15 min
U2OS+Con
PNKP-/-+Con
PNKP-/-+PNKP
PNKP-/-+D1
PNKP-/-+D2
PNKP-/-+D3
PNKP-/-+D4
DNA fiber assay
Normalized CldU length
G
GFP-PNKP Wild-Type (WT)
Linker region of PNKP conserved sequence in mammals
Amino acids 111
S114 T118 T122 S126 S143
145
Consensus TPXSQPDTPPGTPXVXXXEXXDXEPXKKRMRKSXP
Homo Sapiens TPESQPDTPPGTPLVSQDEKRDAELPKKRMRKSNP
Pan troglodytes TPESQPDTPPGTPLVSQDEKRDAELPKKRMRKSNP
Macaca mulatta TPESQPDTPPGTPPVSQDEKRDAEPPKKRMRKSNP
Canis lupus familiaris TPVSQQDTPPGTPPVAPHEAEAVEPQKKRMRKSSP
Mus musculus TSGSQPDAPPDTP-GDPEEGEDTEPQKKRVRKSSL
Rattus norvegicus TPGSPPDTPPGNP-VDPEEGKDTEPQKKRMRKSSP
H
PNKP linker mutants
Relative Cell Growth Rate [A.U.]
U2OS+Con
PNKP-/-+Con
PNKP-/-+PNKP
PNKP-/-+S114A
PNKP-/-+T118A
PNKP-/-+T122A
PNKP-/-+S126A
PNKP-/-+S143A

Figure 4

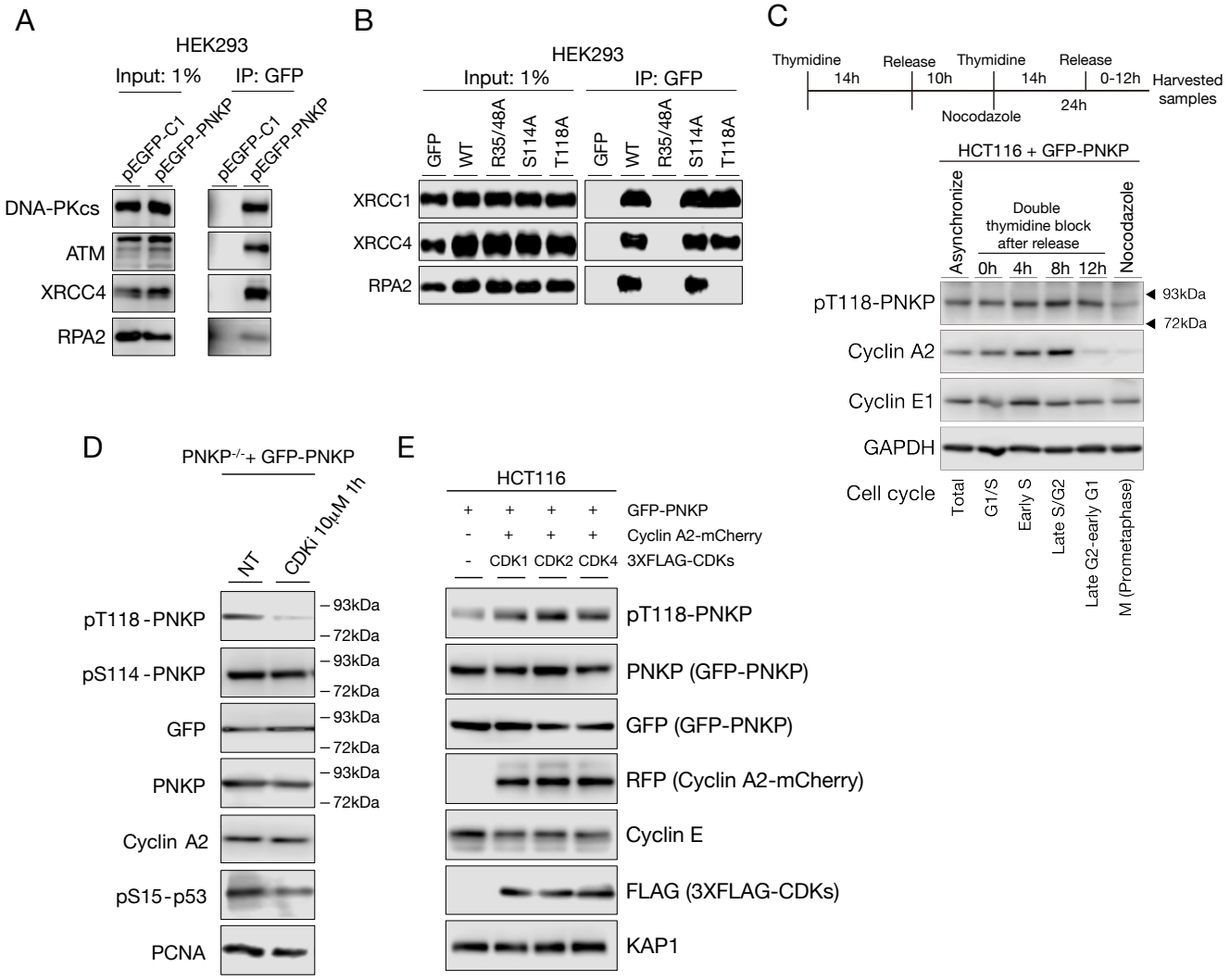

A
HEK293
Input: 1%
IP: GFP
pEGFP-C1
pEGFP-PNKP
DNA-PKcs
ATM
XRCC4
RPA2
B
HEK293
Input: 1%
IP: GFP
GFP
WT
R35/48A
S114A
T118A
XRCC1
XRCC4
RPA2
C
Thymidine
14h
Release
10h
Thymidine
14h
Release
0-12h
Harvested samples
Nocodazole
24h
HCT116 + GFP-PNKP
Asynchronize
Double thymidine block after release
0h
4h
8h
12h
Nocodazole
pT118-PNKP
93kDa
72kDa
Cyclin A2
Cyclin E1
GAPDH
Cell cycle
Total
G1/S
Early S
Late S/G2
Late G2-early G1
M (Prometaphase)
D
PNKP-/-+ GFP-PNKP
NT
CDKi 10μM 1h
pT118-PNKP
pS114-PNKP
GFP
PNKP
Cyclin A2
pS15-p53
PCNA
93kDa
72kDa
E
HCT116
GFP-PNKP
Cyclin A2-mCherry
3XFLAG-CDKs
CDK1
CDK2
CDK4
pT118-PNKP
PNKP (GFP-PNKP)
GFP (GFP-PNKP)
RFP (Cyclin A2-mCherry)
Cyclin E
FLAG (3XFLAG-CDKs)
KAP1

Figure 5

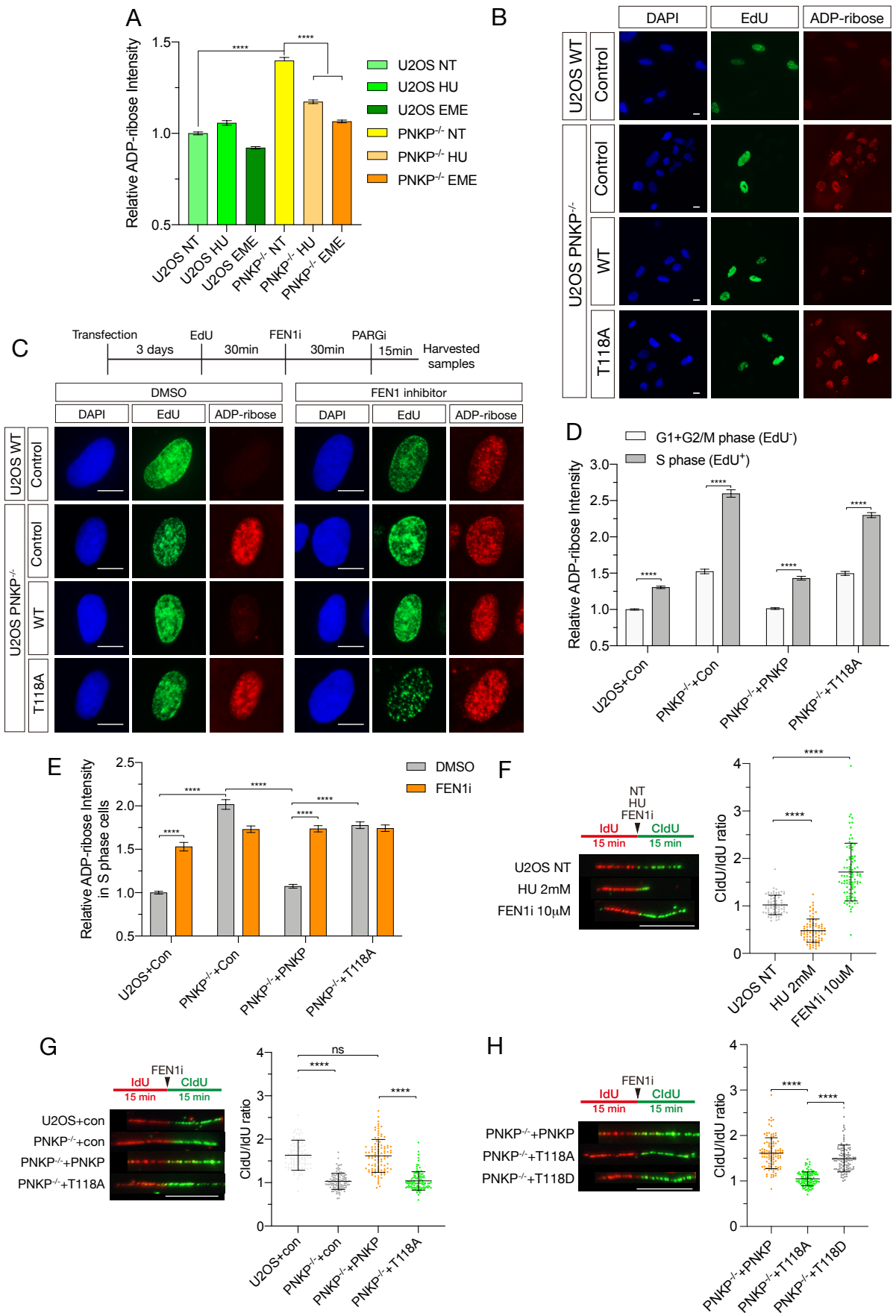

A
Relative ADP-ribose Intensity
1.5
1.0
0.5
****
****
U2OS NT
U2OS HU
U2OS EME
PNKP-/- NT
PNKP-/- HU
PNKP-/- EME
B
DAPI
EdU
ADP-ribose
U2OS WT
Control
U2OS PNKP-/-
Control
WT
T118A
C
Transfection
3 days
EdU
30min
FEN1i
30min
PARGi
15min
Harvested samples
DMSO
FEN1 inhibitor
DAPI
EdU
ADP-ribose
U2OS WT
Control
U2OS PNKP-/-
Control
WT
T118A
D
G1+G2/M phase (EdU-)
S phase (EdU+)
Relative ADP-ribose Intensity
3.0
2.5
2.0
1.5
1.0
0.5
U2OS+Con
PNKP-/-+Con
PNKP-/-+PNKP
PNKP-/-+T118A
E
Relative ADP-ribose Intensity in S phase cells
2.5
2.0
1.5
1.0
0.5
DMSO
FEN1i
U2OS+Con
PNKP-/-+Con
PNKP-/-+PNKP
PNKP-/-+T118A
F
NT
HU
FEN1i
IdU
CldU
15 min
15 min
U2OS NT
HU 2mM
FEN1i 10μM
CldU/IdU ratio
4
3
2
1
0
U2OS NT
HU 2mM
FEN1i 10μM
G
FEN1i
IdU
CldU
15 min
15 min
U2OS+con
PNKP-/-+con
PNKP-/-+PNKP
PNKP-/-+T118A
CldU/IdU ratio
ns
U2OS+con
PNKP-/-+con
PNKP-/-+PNKP
PNKP-/-+T118A
H
FEN1i
IdU
CldU
15 min
15 min
PNKP-/-+PNKP
PNKP-/-+T118A
PNKP-/-+T118D
CldU/IdU ratio
PNKP-/-+PNKP
PNKP-/-+T118A
PNKP-/-+T118D

Figure 6

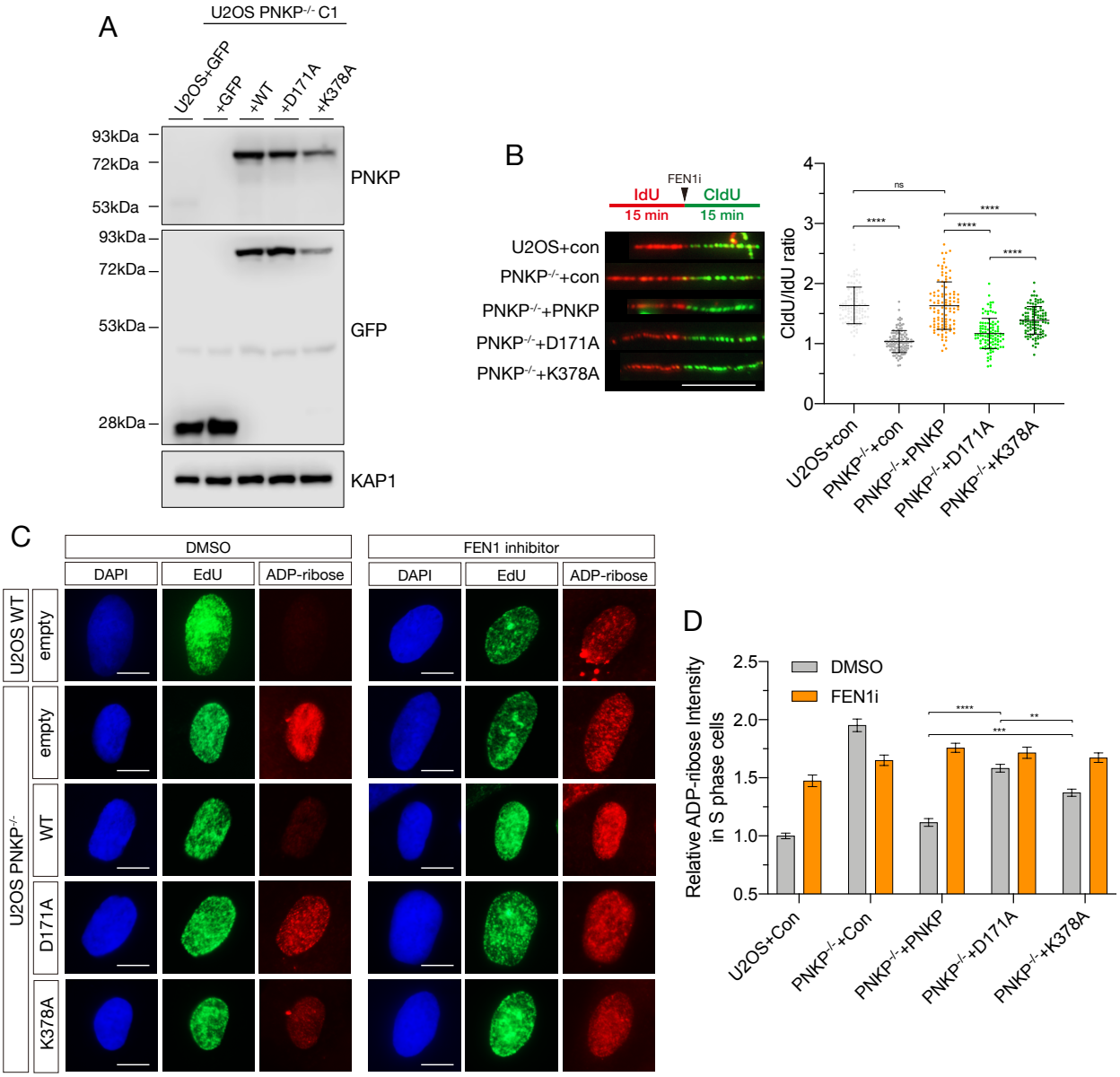
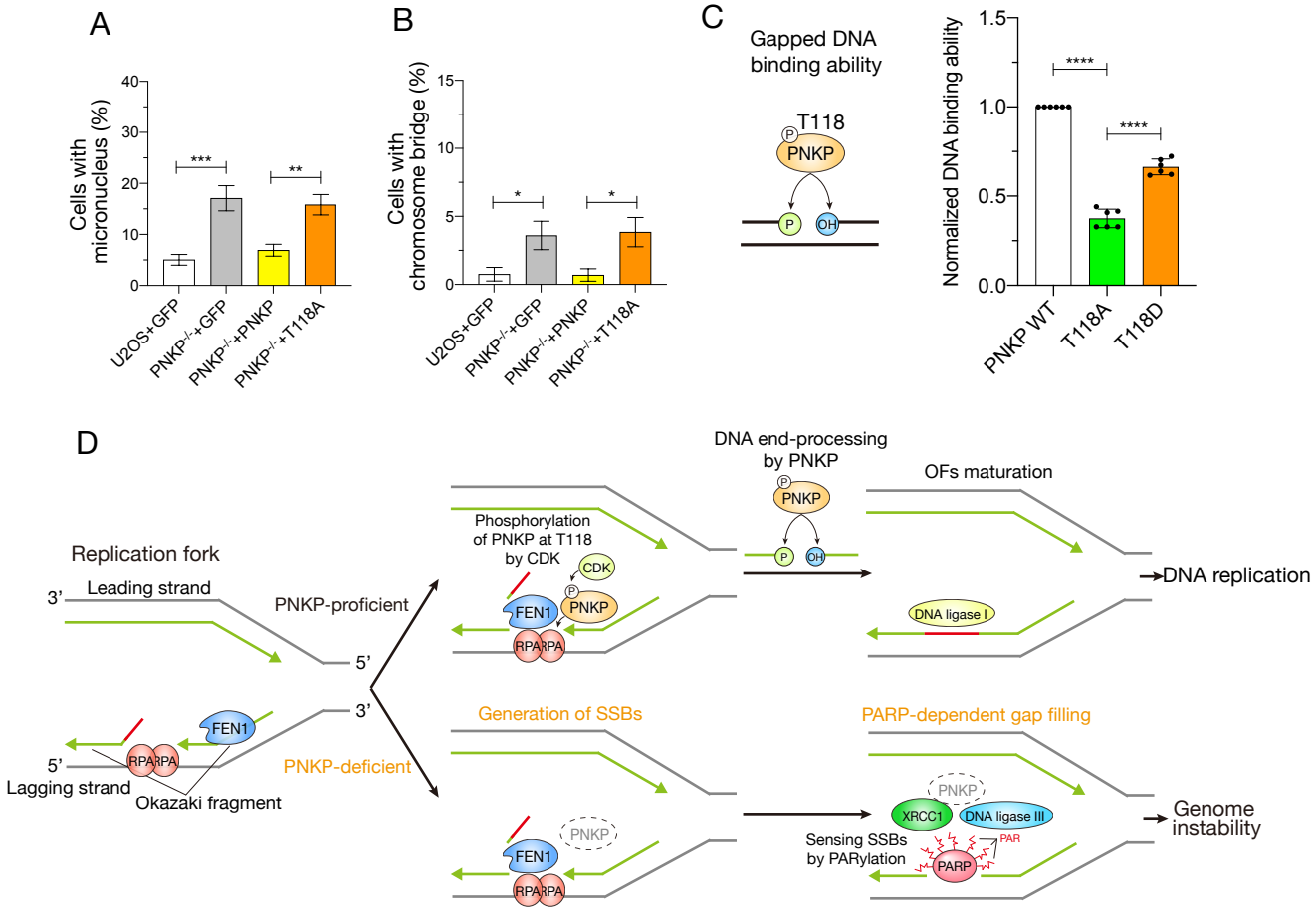

A
U2OS PNKP-/- C1
U2OS+GFP
+GFP
+WT
+D171A
+K378A
93kDa
72kDa
53kDa
PNKP
93kDa
72kDa
53kDa
28kDa
GFP
KAP1
B
FEN1i
IdU
CldU
15 min
15 min
U2OS+con
PNKP-/-+con
PNKP-/-+PNKP
PNKP-/-+D171A
PNKP-/-+K378A
CldU/IdU ratio
ns
****
U2OS+con
PNKP-/-+con
PNKP-/-+PNKP
PNKP-/-+D171A
PNKP-/-+K378A
C
DMSO
FEN1 inhibitor
DAPI
EdU
ADP-ribose
U2OS WT
U2OS PNKP-/-
empty
empty
WT
D171A
K378A
D
DMSO
FEN1i
Relative ADP-ribose Intensity in S phase cells
****
**
***
U2OS+Con
PNKP-/-+Con
PNKP-/-+PNKP
PNKP-/-+D171A
PNKP-/-+K378A

Figure 7


A
Cells with micronucleus (%)
40
30
20
10
0
***
**
U2OS+GFP
PNKP-/-+GFP
PNKP-/-+PNKP
PNKP-/-+T118A
B
Cells with chromosome bridge (%)
15
10
5
0
*
*
U2OS+GFP
PNKP-/-+GFP
PNKP-/-+PNKP
PNKP-/-+T118A
C
Gapped DNA binding ability
T118
P
PNKP
P
OH
Normalized DNA binding ability
1.5
1.0
0.5
0.0
****
****
PNKP WT
T118A
T118D
D
Replication fork
3'
Leading strand
5'
PNKP-proficient
3'
FEN1
5'
RPA
RPA
Lagging strand
Okazaki fragment
PNKP-deficient
Phosphorylation of PNKP at T118 by CDK
CDK
P
PNKP
FEN1
RPA
RPA
DNA end-processing by PNKP
P
PNKP
P
OH
OFs maturation
DNA ligase I
DNA replication
Generation of SSBs
FEN1
PNKP
RPA
RPA
PARP-dependent gap filling
PNKP
XRCC1
DNA ligase III
PAR
PARP
Sensing SSBs by PARylation
Genome instability