PNAS RESEARCH ARTICLE | DEVELOPMENTAL BIOLOGY

DPPA5A suppresses the mutagenic TLS and MMEJ pathways by modulating the cryptic splicing of *Rev1* and *Polq* in mouse embryonic stem cells

Fangjie Jiang[a,b,c,1], Lin Wang[a,d,1], Yuping Dong[a,b,d], Wenhui Nie[a], Hu Zhou[e], Jing Gao[e], and Ping Zheng[a,d,f,2]

Edited by Janet Rossant, Gairdner Foundation, Toronto, Canada; received March 30, 2023; accepted June 13, 2023

Genetic alterations are often acquired during prolonged propagation of pluripotent stem cells (PSCs). This ruins the stem cell quality and hampers their full applications. Understanding how PSCs maintain genomic integrity would provide the clues to overcome the hurdle. It has been known that embryonic stem cells (ESCs) utilize high-fidelity pathways to ensure genomic stability, but the underlying mechanisms remain largely elusive. Here, we show that many DNA damage response and repair genes display differential alternative splicing in mouse ESCs compared to differentiated cells. Particularly, *Rev1* and *Polq*, two key genes for mutagenic translesion DNA synthesis (TLS) and microhomology-mediated end joining (MMEJ) repair pathways, respectively, display a significantly higher rate of cryptic exon (CE) inclusion in ESCs. The frequent CE inclusion disrupts the normal protein expressions of REV1 and POLθ, thereby suppressing the mutagenic TLS and MMEJ. Further, we identify an ESC-specific RNA binding protein DPPA5A which stimulates the CE inclusion in *Rev1* and *Polq*. Depletion of DPPA5A in mouse ESCs decreased the CE inclusion of *Rev1* and *Polq*, induced the protein expression, and stimulated the TLS and MMEJ activity. Enforced expression of DPPA5A in NIH3T3 cells displayed reverse effects. Mechanistically, we found that DPPA5A directly regulated CE splicing of *Rev1*. DPPA5A associates with U2 small nuclear ribonucleoprotein of the spliceosome and binds to the GA-rich motif in the CE of *Rev1* to promote CE inclusion. Thus, our study uncovers a mechanism to suppress mutagenic TLS and MMEJ pathways in ESCs.

embryonic stem cells | cryptic exon | DPPA5A | *Rev1* | *Polq*

Significance

Understanding how ESCs maintain genomic integrity would pave the way to produce safe PSCs. Although it is known that ESCs employ high-fidelity pathways to preserve genomic stability, the underlying mechanisms remain largely elusive. Here, we found that mouse ESCs utilized cryptic exon (CE) inclusion to decrease the protein expression of REV1 and POLθ, thereby suppressing the mutagenic translesion DNA synthesis (TLS) and microhomology-mediated end joining (MMEJ) repair. Importantly, we identified an ESC-specific RNA binding protein DPPA5A, which regulates the CE inclusion in *Rev1* and *Polq* and suppresses the mutagenic DNA replication and repair pathways in ESCs.

Embryonic stem cells (ESCs) are derived from epiblast cells of blastocysts. Owing to the fundamental role in organism development, maintaining a stable genome is of critical importance to ESCs and their in vivo counterparts (1). Concordantly, ESCs and their in vivo counterparts contain a more stable genome than the differentiated somatic cells. For example, mouse ESCs display a 100-times lower genome mutation rate than their isogenic mouse embryonic fibroblasts (MEFs) (2). Perturbation in genome stability of in vivo pluripotent cells can result in developmental failure or defects (3, 4). Although the mutation rate is lower in ESCs, many recurrent genetic aberrations, including chromosomal abnormalities, copy number variations (CNVs), and nucleotide point mutations, as well as epigenetic changes, can be acquired in human ESCs during prolonged culture (5–11). These recurrent aberrations are considered to provide a proliferative advantage to human ESCs under increased replication stress. For example, gain of chromosome 17q in human ESCs provides selective advantage under replicative stress conditions (7). Recurrent CNVs within the 20q11.21 region increase cell survival after passaging (12). Point mutations in cancer-associated gene *BCOR* promote malignancies and appear in 26.9% of human embyonic stem (ES) cell lines (13). DNA methylation patterns similar to tumors can also be acquired in human ESCs after long-term culture (11). These genomic mutations in human ESCs or induced pluripotent stem cells (iPSCs) would compromise their clinical applications by affecting the differentiation potential as well as safety (6, 13, 14). Therefore, elucidating how ESCs preserve genomic stability would shed light on the etiology of developmental failure or defeats and facilitate developing strategies to overcome the genomic instability acquired during generation and/or in vitro expansion of human ESCs and iPSCs (6).

Except for expressing a relatively higher level of genes involved in DNA damage response (DDR) and repair (15), ESCs are capable of employing unique regulators and pathways to better protect their genome when compared to differentiated cells (16). These ESC-specific mechanisms play determinant roles in various processes of genomic stability regulation, including DDR, DNA double-strand break (DSB) repair (17–19), DNA

Author contributions: P.Z. designed research; F.J., L.W., Y.D., W.N., H.Z., and J.G. performed research; F.J. and L.W. analyzed data; and L.W. and P.Z. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2023 the Author(s). Published by PNAS. This article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

[1]F.J. and L.W. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: zhengp@mail.kiz.ac.cn.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2305187120/-/DCSupplemental.

Published July 17, 2023.

replication stress response (3, 20), telomere maintenance (21–24), and cell fate determination after DNA damage (25–29). For example, PARP-catalyzed PARylation is essential for many DDR processes. An ESC-specific protein FILIA can physically interact with PARP1 and stimulate PARP1 activity to increase DDR and damage repair efficiency (18). SALL4, an ESC-stemness regulator, is able to translocate onto DSB sites to enhance the ATM activation (17). DNA replication stress is a major source of endogenous DNA damage. ESCs possess greater ability to resolve replication stress and avoid replication-associated DNA damage (20, 30). ESC-specific factors, e.g., FILIA-FLOPED protein complex and long noncoding RNA *Discn* and *Lnc956*, have been found to play critical roles in replication stress response (3, 20, 31). Unlike somatic cells, ESCs lengthen the telomeres using a telomerase-independent mechanism known as alternative lengthening of telomeres (ALT), which relies on telomere recombination and requires the ESC-specific protein ZSCAN4 (21). Mechanisms of telomere protection are also distinct between ESCs and somatic cells. Telomere protection in ESCs does not require TRF2, which, however, is a central regulator in somatic cells (23, 24). Thus, to comprehensively understand how ESCs efficiently preserve genomic stability, it is essential to focus on the ESC-specific players and mechanisms.

Alternative splicing (AS) occurs in >90% of genes and acts as a versatile mechanism to generate protein diversity and to regulate development and tissue identity (32–34). For example, spliceosomal repression in mouse ESCs drives a pluripotent-to-totipotent state transition (35). ESC-specific exon usage in *Foxp1* transcripts plays a pivotal role in the regulation of pluripotency (36). Whether ESCs differentially splice DDR and repair genes to ensure a highly stable genome is unexplored. If so, what ESC-prevalent RNA binding proteins (RBPs) regulate these splicing events remains elucidated. In this study, we reported the essential role of ESC-prevalent RNA splicing events in safeguarding the genomic stability of mouse ESCs. Moreover, we identified an ESC-specific RBP DPPA5A (also called DPPA5, ECAT2, ESG1), which ensures genome stability by regulating the ESC-prevalent cryptic exon (CE) splicing of key DNA replication and repair genes *Rev1* and *Polq*.

Results

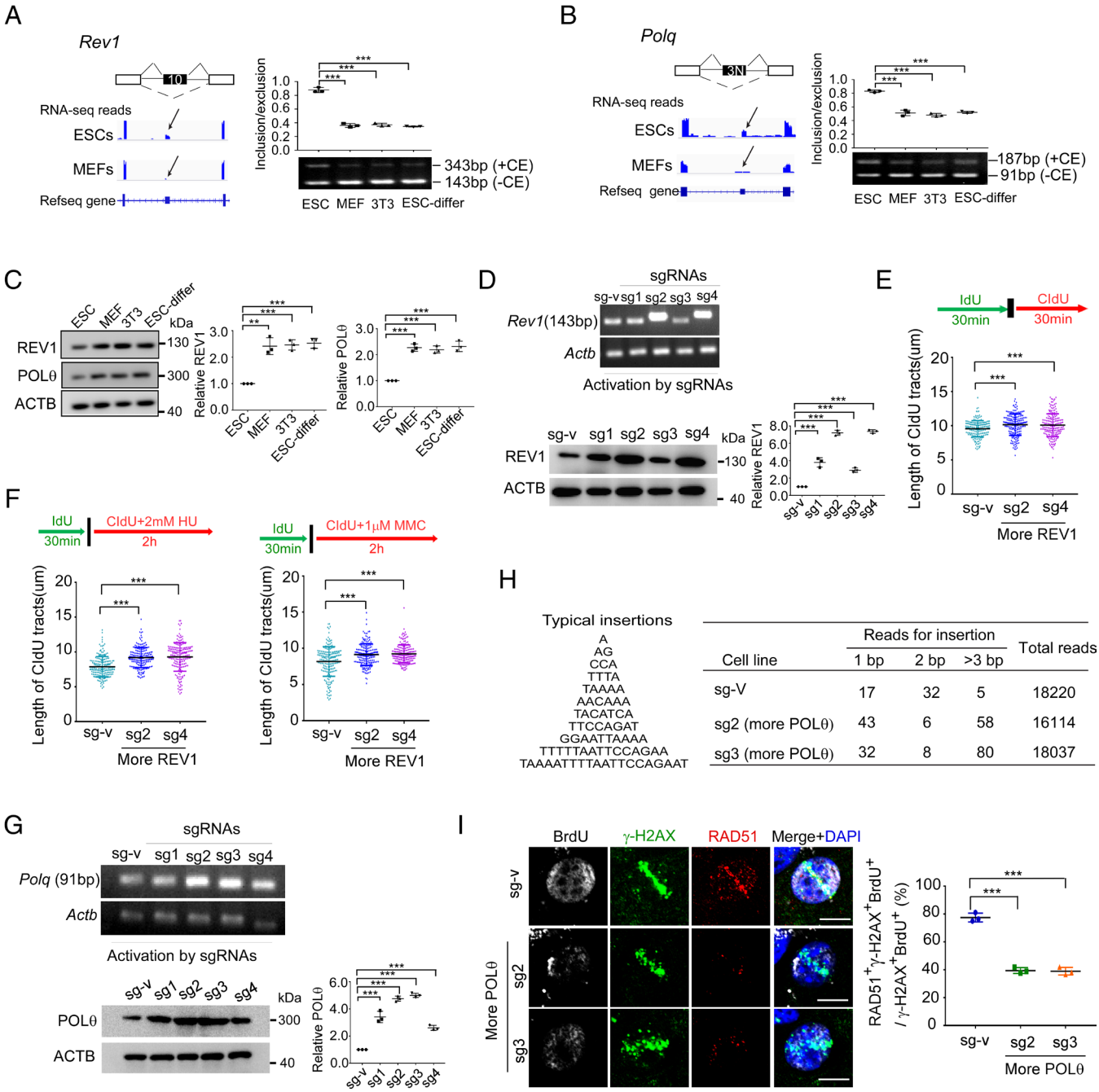
ESCs Have Alternative Splicing Signature of DDR and Repair Genes. We performed RNA sequencing analysis to compare the expression of a set of DDR regulators (805 genes) from the AmiGO 2 database (http://amigo.geneontology.org/amigo) between mouse ESCs and MEFs. Among the 805 DDR genes, 298 genes (37%) showed differential expression levels between ESCs and MEFs (fold change ≥ 2) (Dataset S1). We then examined whether these DDR genes displayed differential splicing events between ESCs and somatic cells. Notably, 162 DDR genes (20.1%), whose expression levels were comparable between ESCs and MEFs, exhibited differential AS patterns (225 AS events). These AS events fell into variable categories, with the cassette inclusion or exclusion being the most frequent (*SI Appendix*, Fig. S1*A* and Dataset S2).

Among the AS events, splicing of CE (also called pseudoexons) has a definite impact on protein expression. CE is normally an intronic sequence. When aberrantly getting spliced into mature mRNA transcripts, CE often introduces frame shifts and/or premature stop codons. In most cases, CE inclusion halts translation and promotes nonsense-mediated RNA decay, causing the decrease of protein expression (37). Among the 162 DDR genes showing differential splicing patterns between ESCs and MEFs, 15 genes undergo CE inclusion, and the inclusion frequency was distinct between ESCs and MEFs (*SI Appendix*, Fig. S1*B*). Of them, nine genes (*Rev1*, *Polq*, *Casp2*, *Ercc4, Rev3l, Taok2, Alkbh3, Dclre1c,* and *Pidd1*) exhibited a higher rate of CE inclusion in ESCs than in MEFs, whereas the others (*Hdac10*, *Casp9*, *Ogg1*, *Smc5*, *Iffo1,* and *Eif2ak4*) displayed more frequent CE splicing in MEFs. We validated the differential CE inclusion rate of several DDR genes (*Rev1*, *Polq*, *Casp2,* and *Alkbh3*) by RT-PCR between ESCs and several types of somatic cells including MEFs, NIH3T3 cells, and isogenic ESC-differentiated cells (Fig. 1 *A* and *B* and *SI Appendix*, Fig. S2 *A–C*).

A High Rate of CE Inclusion in *Rev1* and *Polq* Suppresses Mutagenic DNA Replication and Repair Pathways in Mouse ESCs. CEs in the above genes introduce premature stop codons and are anticipated to cause a protein expression decrease. Taking key DDR regulators REV1 and POLθ (encoded by the *Polq* gene) as examples, we validated that mouse ESCs expressed less REV1 and POLθ proteins when compared to the MEFs, 3T3 cells, and isogenic ESC-differentiated cells (Fig. 1*C*). REV1 is a Y-family DNA polymerase which plays an essential role in structurally organizing the protein complex for translesion DNA synthesis (TLS). TLS can bypass the DNA lesions during DNA replication and render the cells tolerance to DNA damage. However, it is highly mutagenic, and the activity is under stringent control (38, 39). POLθ is an A-family DNA polymerase and plays a key role in the microhomology-mediated end-joining (MMEJ) repair pathway for DNA DSBs. MMEJ competes with high-fidelity homologous recombination (HR)-mediated repair pathway and is mutagenesis due to the template DNA insertion at the sites of DSB joining (40–42). The reduced expression of REV1 and POLθ proteins in mouse ESCs could decrease the mutagenic TLS- and MMEJ-mediated DNA replication and repair.

To validate whether the protein expression changes in REV1 and POLθ can modulate the TLS and MMEJ activity, respectively, in mouse ESCs, we simply increased REV1 and POLθ expression in ESCs via CRISPR/Cas9-mediated gene activation strategy and investigated the influence. We utilized the DNA fiber assay to analyze the DNA replication under a normal or replication stress condition (43). Cells were sequentially pulse-labeled with 5-iodo-2′-deoxyuridine (IdU) and 5-chloro-2′-deoxyuridine (CIdU) to track the active forks. Fork progression is monitored by measuring the length of CIdU track. Under a normal culture condition, elevating REV1 expression in ESCs (Fig. 1*D*) significantly increased the CIdU track length when compared to control ESCs (Fig. 1*E*). We also examined the fork behavior under the replication stress condition. We used a low dose of hydroxyurea (HU, which blocks DNA synthesis through inhibition of ribonucleotide reductase) or mitomycin C (MMC, which blocks DNA replication by cross-linking guanine and cytosine) to induce mild replication stress. Under the mild stress condition, replication forks move with a reduced rate (*SI Appendix*, Fig. S3 *A* and *B*), allowing us to investigate the fork progression changes. Similarly, rising REV1 expression in ESCs increased CIdU track length under the replication stress condition (Fig. 1*F*). Thus, the expression level of REV1 is positively correlated with the TLS activity.

We also measured MMEJ repair according to the strategy described previously (44). MMEJ usually leaves behind a recognizable genomic scar such as templated insertions when repairing chromosome DSB ends. Templated insertion is a specific footprint of MMEJ activity and consequently has been used as a biomarker of MMEJ activity (45). A DNA DSB was generated at the 3′ end of the *Discn* locus using the CRISPR-Cas9 system (3). The genome scars were then evaluated by deep sequencing of the breakpoint junction. POLθ-mediated MMEJ usually generates insertions which



Cell line	Reads for insertion 1 bp	2 bp	>3 bp	Total reads
sg-V	17	32	5	18220
sg2 (more POLθ)	43	6	58	16114
sg3 (more POLθ)	32	8	80	18037

Fig. 1. A high rate of CE inclusion in *Rev1* and *Polq* suppresses mutagenic TLS and MMEJ pathways in mouse ESCs. (*A* and *B*) Alternative splicing (AS) of the exon in *Rev1* (*A*) and *Polq* (*B*) genes. The left panel shows the RNA-seq reads. The black box indicates a CE; white boxes indicate flanking exons. AS of CE was validated by RT-PCR, and the ratio of CE inclusion/exclusion was used to quantify the CE inclusion frequency. (*C*) ESCs expressed a lower level of REV1 and POLθ proteins when compared to several types of differentiated cells. The relative protein levels were normalized with ACTB. (*D*) CRISPR/Cas9-mediated gene activation increased the mRNA (*Upper*) and protein expression (*Lower*) of *Rev1*. (*E* and *F*) The DNA fiber assay showed that replication forks progressed faster when REV1 expression was elevated under the normal (*E*) or replication stress condition induced by HU or MMC treatment (*F*). (*G*) CRISPR/Cas9-mediated gene activation increased the mRNA (*Upper*) and protein expression (*Lower*) of *Polq*. (*H*) Increasing POLθ protein expression elevated the MMEJ activity. (*I*) Conversely, increasing POLθ protein expression decreased the HR activity. The relative protein levels in (*D* and *G*) were normalized by ACTB. Results were from three independent experiments. Data are shown as mean ± SEM. $^{**}P < 0.01$ and $^{***}P < 0.001$, two-tailed Student's *t* test.

are typically 2 to 30 bp long in mammalian cells (46). After elevating POLθ expression in ESCs (Fig. 1*G*), the frequencies of insertions ≥2 bp long were robustly increased compared to the control (Fig. 1*H*), indicating that the MMEJ repair pathway was significantly activated. MMEJ competes with HR repair. Conversely, increasing POLθ level in ESCs suppressed the HR repair pathway (Fig. 1*I*). Thus, a high rate of CE inclusion in *Rev1* and *Polq* reduced the protein expression and suppressed mutagenic DNA replication and repair pathways in mouse ESCs.

ESC-Specific RNA-Binding Protein DPPA5A Regulates AS Events of DDR Genes and Safeguards Genomic Stability. AS events are regulated by RBPs (47). ESCs have highly expressed RBP repertoire (48, 49). The trans-acting regulators of differential AS of DDR genes in ESCs are not known. Identifying these RBPs is crucial to understand the mechanism safeguarding the stem cell genomic stability and biology. We were particularly interested in elucidating the regulators of CE splicing of two critical DDR genes *Rev1* and *Polq*, whose protein expression levels fine tune the DNA replication

and repair fidelity. RBP regulating CE splicing can bind to the CE of target pre-RNAs (50) and could be identified through RNA pull down combined with mass spectrometry analysis. We then utilized this strategy to identify the potential RBP regulating CE splicing of *Rev1*. To this end, the CE of the *Rev1* gene was transcribed in vitro with biotin labeling. RNA pull down combined with mass spectrometry analyses identified a list of interaction protein candidates (Fig. 2*A* and Dataset S3). Among these candidates, DPPA5A is the only RBP showing an ESC-specific expression pattern (51). DPPA5A contains an atypical K-homology (KH) RNA-binding domain, and the gene is abundantly transcribed in mouse and human ESCs (51–53). However, DPPA5A depletion did not influence ESC self-renewal (52), and its function in ESCs remained unexplored. We therefore decided to focus on DPPA5A to investigate its possible role in regulating AS events of DDR genes and genomic stability in mouse ESCs.

We first validated the in vitro association of DPPA5A with the CE of *Rev1* by immunoblotting analysis of RNA pull-down samples (Fig. 2*B*). *Dppa5a* was then efficiently knocked down (KD) in mouse ESCs via two independent doxycycline (Dox)-inducible short hairpin RNAs (shRNAs) (*SI Appendix*, Fig. S4 *A* and *B*). Consistent with the previous report (52), *Dppa5a* KD ESCs displayed normal colony

Proteins	Numbers of peptides: Negative control	*Rev1*-CE pull down
PTRHD1	0	2
MED11	0	2
DPPA5A	0	1
RWDD4	0	1
HNRNPA1	0	1
LSM2	0	1
DDX46	0	1
RPL21	0	1
TAF12	0	1

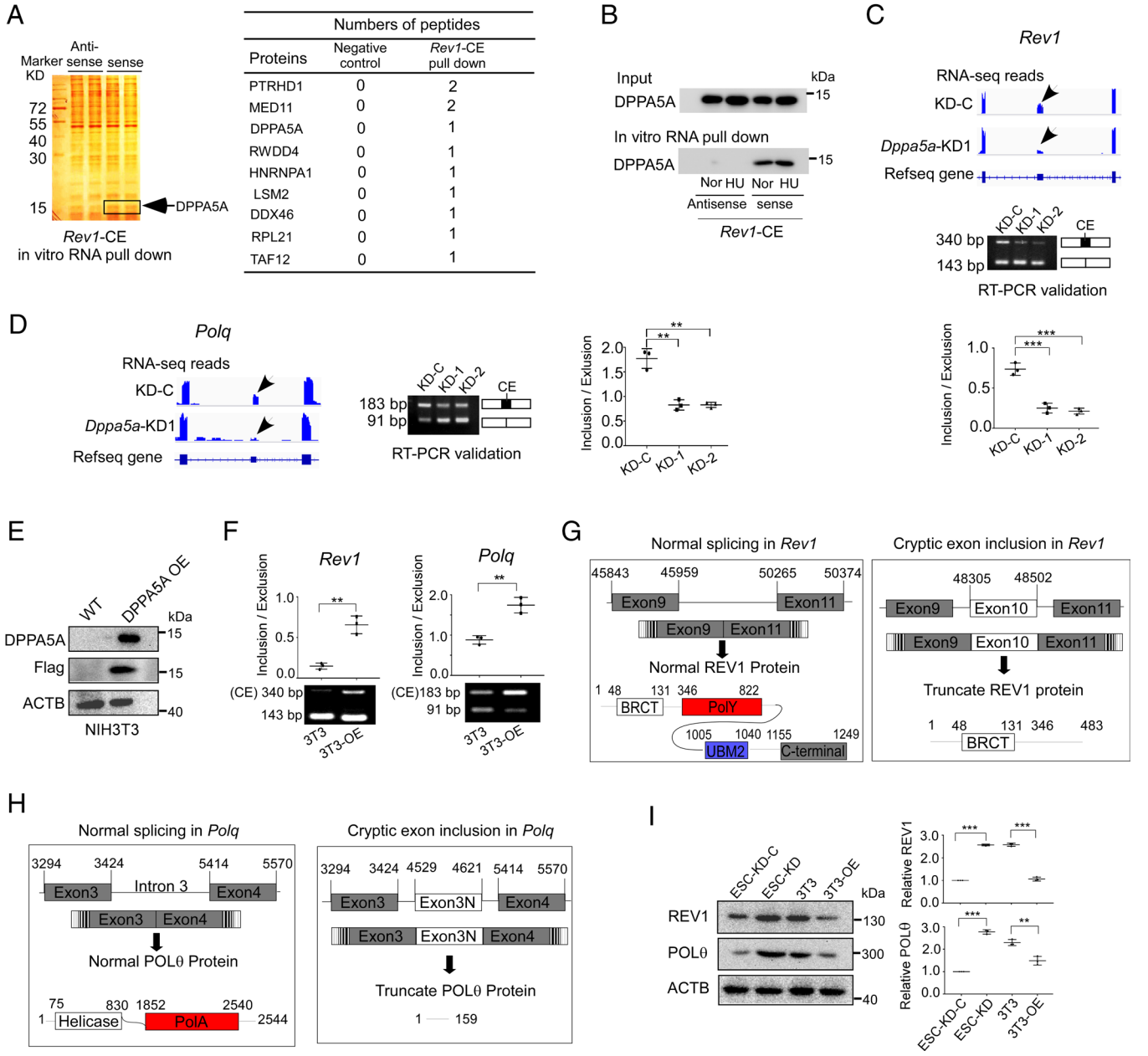


Fig. 2. ESC-specific RNA-binding protein DPPA5A regulates CE inclusion in *Rev1* and *Polq*. (*A*) The CE of *Rev1* was labeled with biotin and subject to in vitro RNA pull-down and mass spectrometry analysis. The left panel shows the silver staining of pull-down samples, and the right panel shows the protein candidates identified by mass spectrometry. (*B*) Immunoblotting analysis of in vitro RNA pull-down samples validated the direct interaction between DPPA5A and the CE of *Rev1*. (*C* and *D*) RNA-seq analysis revealed that CE inclusion in *Rev1* (*C*) and *Polq* (*D*), respectively, was reduced in *Dppa5a* knockdown (KD) ESCs when compared to KD control (KD-C) ESCs. RT-PCR analysis further validated the reduced CE inclusion after *Dppa5a* KD. (*E*) Immunoblotting confirmed the overexpression (OE) of Flag-tagged DPPA5A in NIH3T3 cells. (*F*) Overexpression of DPPA5A in NIH3T3 cells (3T3-OE) significantly stimulated the CE inclusion rates in *Rev1* and *Polq*. (*G*) Schematic diagram showing that the inclusion of exon 10 (CE) in *Rev1* leads to the generation of truncated protein. (*H*) Schematic diagram showing that the inclusion of a new exon 3 (exon 3N, CE) leads to the generation of truncated POLθ protein. (*I*) Immunoblotting showed that KD of *Dppa5a* in mouse ESCs increased the REV1 and POLθ protein expressions. Conversely, overexpression of DPPA5A in NIH3T3 cells (3T3-OE) decreased the protein expressions. The results were from three independent experiments. Data are shown as mean ± SEM. Two-tailed Student's *t* test, $^{**}P < 0.01$ and $^{***}P < 0.001$.

morphology *SI Appendix*, Fig. S4*C*) and had unperturbed cell cycle and proliferation rate (*SI Appendix*, Fig. S4 *D* and *E*). Moreover, the protein expressions of several key pluripotency regulators were not significantly altered by *Dppa5a* KD (*SI Appendix*, Fig. S4*F*).

We then examined the influence of *Dppa5a* KD on overall AS events. Compared to shControl, knocking down of *Dppa5a* in mouse ESCs affected the AS of 2,421 genes and of 3,192 events (Dataset S4). Gene ontology (GO) enrichment analysis showed that the affected genes were highly enriched in the processes of DNA repair, chromatin organization, cellular response to DNA damage stimuli, regulation of transcription, and RNA splicing (*SI Appendix*, Fig. S5*A*). Because DPPA5A is an ESC-specific RBP, we examined whether it regulated the ESC-prevalent AS events of DDR genes. Focusing on the 162 DDR genes and their 225 AS events showing differential AS patterns between ESCs and MEFs, we found that 29 genes (17.9%) and 31 AS events (13.7%) were affected by *Dppa5a* depletion (Dataset S5). Of the 15 DDR genes displaying differential CE inclusion between ESCs and MEFs, five genes (*Rev1*, *Polq*, *Casp2, Taok2,* and *Alkbh3*) showed affected CE splicing in ESCs after DPPA5A knockdown (Dataset S5). This observation suggested that DPPA5A could modulate a set of DDR genes at a posttranscriptional level and regulate genomic stability in ESCs. Indeed, KD of *Dppa5a* in mouse ESCs increased DNA DSBs, as indicated by the elevated level of γH2AX (*SI Appendix*, Fig. S5 *B* and *C*) and increased comet tail length in the neutral comet assay (*SI Appendix*, Fig. S5*D*). *Dppa5a* KD ESCs also displayed chromosome instability (CIN) phenotypes including aneuploidy (*SI Appendix*, Fig. S5*E*) and micronuclei formation (*SI Appendix*, Fig. S5*F*).

DPPA5A Regulates CE Inclusion of *Rev1* and *Polq* Genes. DPPA5A may function in different ways to regulate the AS of its target DDR genes. We focused on CE splicing and examined the influence of *Dppa5a* KD on CE inclusion of two key DDR genes *Rev1* and *Polq*. Notably, *Dppa5a* KD in mouse ESCs decreased the CE splicing of *Rev1* (Fig. 2*C*) and *Polq* (Fig. 2*D*). To further verify the regulation of DPPA5A on CE inclusion in the two genes, we ectopically expressed Flag-tagged DPPA5A in NIH3T3 cells (Fig. 2*E*) and examined the influence. Consistently, DPPA5A expression stimulated the CE inclusion of *Rev1* and *Polq* in NIH3T3 cells (Fig. 2*F*). We next validated the sequences of CEs as well as the alternative splicing sites in *Rev1* and *Polq* by Sanger sequencing. CE inclusion introduced a premature stop codon at the PolY domain of REV1 protein, leading to the generation of REV1 truncate containing only 483 amino acids (Fig. 2*G*). Similarly, CE inclusion in *Polq* generated a truncate containing 159 amino acids (Fig. 2*H*). In line with the regulation of DPPA5A on CE inclusion in *Rev1* and *Polq*, *Dppa5a* KD in mouse ESCs increased the protein expression of REV1 and POLθ. Conversely, ectopic expression of DPPA5A in NIH3T3 cells decreased the REV1 and POLθ protein expression (Fig. 2*I*).

DPPA5A Suppresses the Mutagenic TLS and Ensures DNA Replication Stress Response in Mouse ESCs. REV1 regulates mutagenic TLS. We speculated that ESC-specific protein DPPA5A may suppress TLS to ensure genomic stability. The DNA fiber assay showed that under the normal culture condition, *Dppa5a* KD in mouse ESCs significantly increased the CIdU track length when compared to KD control ESCs. Blocking REV1 activity with inhibitor JH-RE-06 (54, 55) in *Dppa5a* KD cells reversed the phenotype (Fig. 3*A*). Consistently, KD of *Dppa5a* in ESCs drastically increased the CIdU track length under HU (Fig. 3*B*) or MMC treatment condition (Fig. 3*C*). The phenotype was fully reversed by blocking TLS activity via JH-RE-06.

Continuous fork progression under DNA replication stress condition can also be induced by DNA replication via repriming or new origin firing, which generates single-stranded DNA (ssDNA) gaps not easily visible in the standard DNA fiber assay. To rule out the contribution of repriming or new origin firing to fork progression in *Dppa5a* KD ESCs, we performed the DNA fiber assay in the presence of S1 nuclease, which degrades DNA with an ssDNA gap. Treatment of S1 nuclease had no influence on the CIdU tract length change induced by *Dppa5a* depletion, excluding the involvement of repriming or new origin firing in continuity of fork progression (*SI Appendix*, Fig. S6*A*). Thus, depletion of *Dppa5a* significantly induced the TLS activity. In line with the observations in *Dppa5a* KD ESCs, forced expression of DPPA5A in NIH3T3 cells decreased the CIdU length under the normal culture or mild replication stress condition (Fig. 3 *D*–*F*). These results altogether support that DPPA5A suppresses the mutagenic TLS by decreasing REV1 expression. Notably, inhibition of REV1 activity in *Dppa5a*-proficient mouse ESCs had no influence on fork progression under the normal culture condition (Fig. 3*A*) or exhibited very mild influence on fork progression under the replication stress condition (Fig. 3 *B* and *C*), suggesting that TLS activity was very marginal in wild-type (WT) ESCs under either normal or replication stress condition.

TLS bypasses the DNA lesions to enable replication to continue. This avoids the formation of ssDNA gaps and suppresses the replication stress response, leading to the cellular tolerance to replication stress (56). Given that TLS was enhanced in *Dppa5a* KD ESCs, we speculated that *Dppa5a* KD ESCs would display the above phenotypes. We then examined the ssDNA gaps by nondenaturing immunofluorescent staining following CIdU incorporation. Under the MMC or HU-induced replication stress condition, *Dppa5a* KD significantly reduced the ssDNA formation (Fig. 4*A* and *SI Appendix*, Fig. S6*B*). The ssDNA initiates the ATR-CHK1 replication stress signaling pathway (57). Concordantly, ATR-CHK1 signaling was robustly compromised in *Dppa5a* KD ESCs when compared to *Dppa5a*-proficient cells. The downstream p53 activation was also suppressed by *Dppa5a* KD (Fig. 4*B* and *SI Appendix*, Fig. S6*C*). ATR-CHK1 signaling activates the intra S-phase checkpoint which prevents the cell cycle progression (58). We treated the ESCs with MMC or HU and measured the kinetics of cell cycle progression from the S to G2/M phase at different time points after release from treatment. In line with the S-phase checkpoint deficiency, *Dppa5a* KD ESCs displayed faster dynamics than *Dppa5a*-proficient ESCs to exit the S phase (Fig. 4*C* and *SI Appendix*, Fig. S6*D*) and to progress into the G2/M phase (Fig. 4*D* and *SI Appendix*, Fig. S6*E*). Moreover, *Dppa5a* KD increased the tolerance of ESCs to DNA replication stress. After MMC or HU treatment, *Dppa5a* KD ESCs were resistant to undergo apoptosis, as indicated by the lower levels of CASPASE 3 activation and PARP1 cleavage (Fig. 4*E* and *SI Appendix*, Fig. S6*F*). The classical colony formation assay also showed that *Dppa5a* KD ESCs formed more alkaline phosphate (AP)-positive colonies than *Dppa5a*-proficient cells following MMC or HU treatment (Fig. 4*F* and *SI Appendix*, Fig. S6*G*). Thus, *Dppa5a* depletion in ESCs compromised the activation of DNA replication stress response and rendered cells tolerant to replication stress.

DPPA5A Suppresses POLθ-Mediated Mutagenic MMEJ and Promotes HR Repair of DNA DSBs. POLθ performs a key role in MMEJ repair of DNA DSBs (40). MMEJ can compete with the HR pathway and is highly error prone (42). Compared to WT ESCs, the protein level of POLθ in *Dppa5a* KD ESCs was elevated more than two folds (Fig. 2*I*), suggesting that the HR repair pathway might be overridden by MMEJ. To test

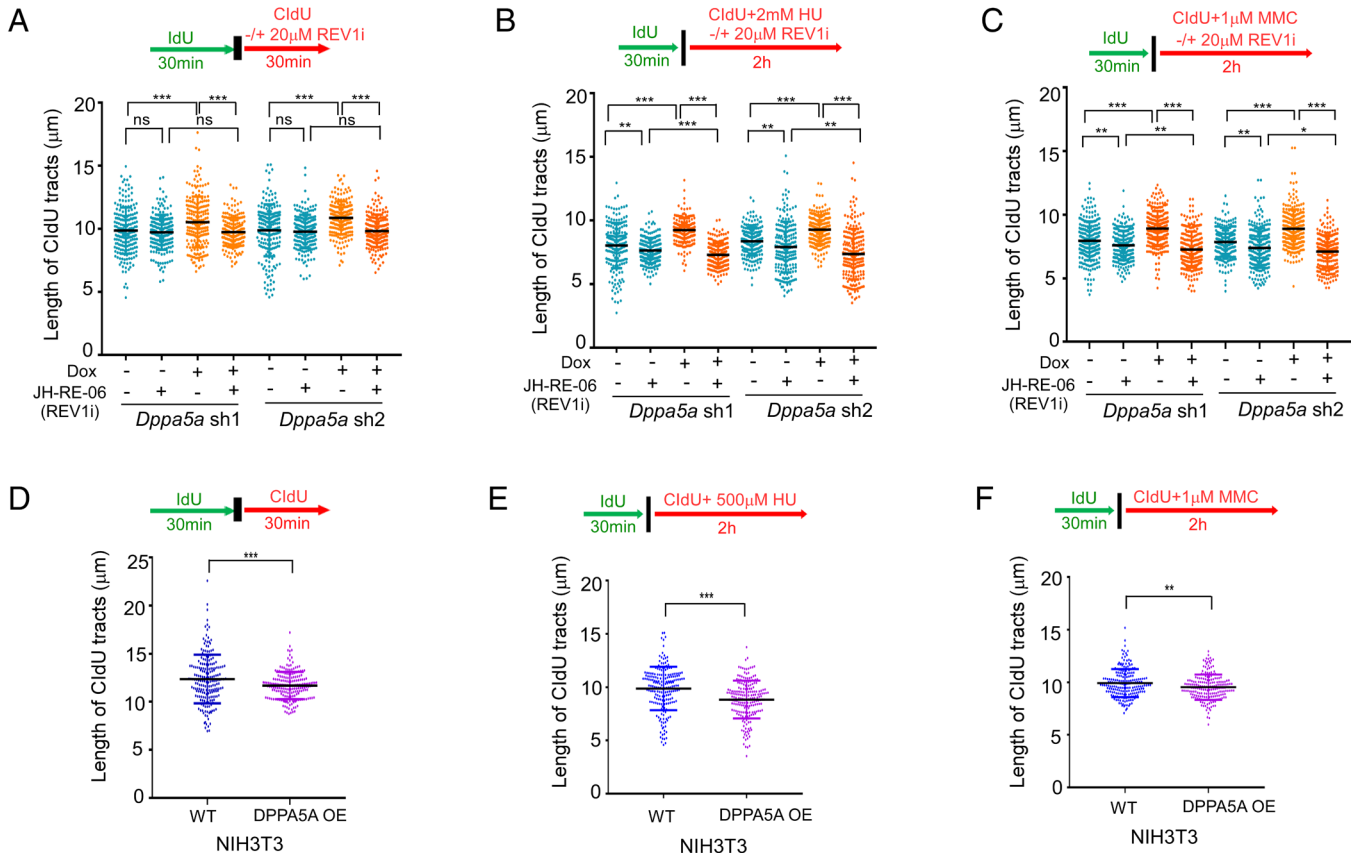


Fig. 3. DPPA5A suppresses the mutagenic TLS. (*A*) The DNA fiber assay showed that under the normal culture condition, KD of *Dppa5a* in mouse ESCs increased the CldU tract length. Blocking REV1 activity with the inhibitor (JH-RE-06) counteracted the effect, indicating that *Dppa5a* KD increased the TLS activity. (*B* and *C*) ESCs were treated with HU (*B*) or MMC (*C*) to induce replication stress. Under stress conditions, *Dppa5a* KD significantly increased the TLS activity in ESCs. Blocking REV1 activity in WT ESCs slightly reduced the CldU tract length, suggesting that ESCs marginally employ TLS to cope with replication stress. (*D*) Under the normal culture condition, overexpression of DPPA5A (DPPA5A OE) in NIH3T3 cells reduced the CldU tract length. (*E* and *F*) Similarly, DPPA5A OE in NIH3T3 cells reduced the CldU tract length under HU (*E*) or MMC (*F*) treatment condition. At least 200 fibers from three independent experiments were analyzed. Data are shown as mean ± SEM. Two-tailed Student's *t* test, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, and ns, no significance.

this idea, we treated the ESCs with etoposide to induce DNA DSBs and examined the influence of *Dppa5a* KD on the three DSB repair pathways, including nonhomologous end joining (NHEJ), HR, and MMEJ. Classical NHEJ repair of DSBs was monitored by the green fluorescent protein (GFP)-reporter system (59). Notably, DPPA5A depletion had no influence on NHEJ (*SI Appendix*, Fig. S7*A*). However, HR repair was significantly repressed in *Dppa5a* KD ESCs, as indicated by the reduced proportions of ESCs capable of forming RAD51 foci at DSB sites following laser microirradiation at the S-phase (Fig. 5*A*). We also measured MMEJ repair. Compared to the DPPA5A-proficient ESCs, the frequencies of insertions ≥2 bp long were robustly increased in *Dppa5a* KD ESCs (Fig. 5*B*), indicating that the MMEJ repair pathway was significantly enhanced when DPPA5A was depleted. Taken together, these observations suggested that DPPA5A regulated the DNA DSB repair pathway choice between HR and MMEJ, an event which occurs after the initiation of end resection. By suppressing the expression of POLθ, DPPA5A favored the DNA DSB repair via the HR pathway. MMEJ repair contributes to the occurrence of chromosomal translocation (45). We found that *Dppa5a* KD ESCs had increased telomere fusion (Fig. 5*C*), which manifests the elevated chromosome translocation at telomeric regions.

Because TLS and MMEJ are highly mutagenic, we wondered whether *Dppa5a* KD in ESCs increased tumorigenicity. To this end, different numbers of *Dppa5a*-proficient and -deficient ESCs were subcutaneously injected to BALB/c nude mice and teratoma formation was examined 1 mo later. The teratoma formation rates were significantly higher in *Dppa5a* KD ESCs than in *Dppa5a*-proficient ESCs (Fig. 5*D*). Thus, DPPA5A plays critical roles in preventing ESCs from transformation.

DPPA5A Associates with the U2 snRNA Complex. We next investigated how DPPA5A regulated CE splicing of the two genes. We first analyzed the interacting proteins of DPPA5A. To this end, we expressed Flag-tagged DPPA5A in WT ESCs and established a stable cell line (*SI Appendix*, Fig. S7*B*). Using Flag antibody for coimmunoprecipitation followed by mass spectrometry analysis, we identified the potential interactome of DPPA5A under the HU-induced replication stress condition. Among the candidates, the core components of U2 small nuclear ribonucleoprotein (snRNP) of the spliceosome, for instance, SF3A3/SF3B3 and U2AF2 (60), were significantly represented (Fig. 6*A* and Dataset S6). GO term analysis of the interaction proteins also enriched the term of "RNA splicing" (Fig. 6*B*). The association of DPPA5A with U2 snRNP components including SF3B1, SF3B3, SF3A3, and U2AF2 was validated by reciprocal immunoprecipitation under normal and HU treatment conditions (Fig. 6*C*). To further verify the association of DPPA5A with the U2 complex, we individually knocked down U2AF2 or SF3B1 in ESCs by shRNAs (*SI Appendix*, Fig. S7 *C* and *D*) and reexamined the protein interaction. Depletion of U2AF2 disrupted the U2 complex assembly, as indicated by the failure of SF3B3 to pull down other protein components of U2 (Fig. 6*D*).

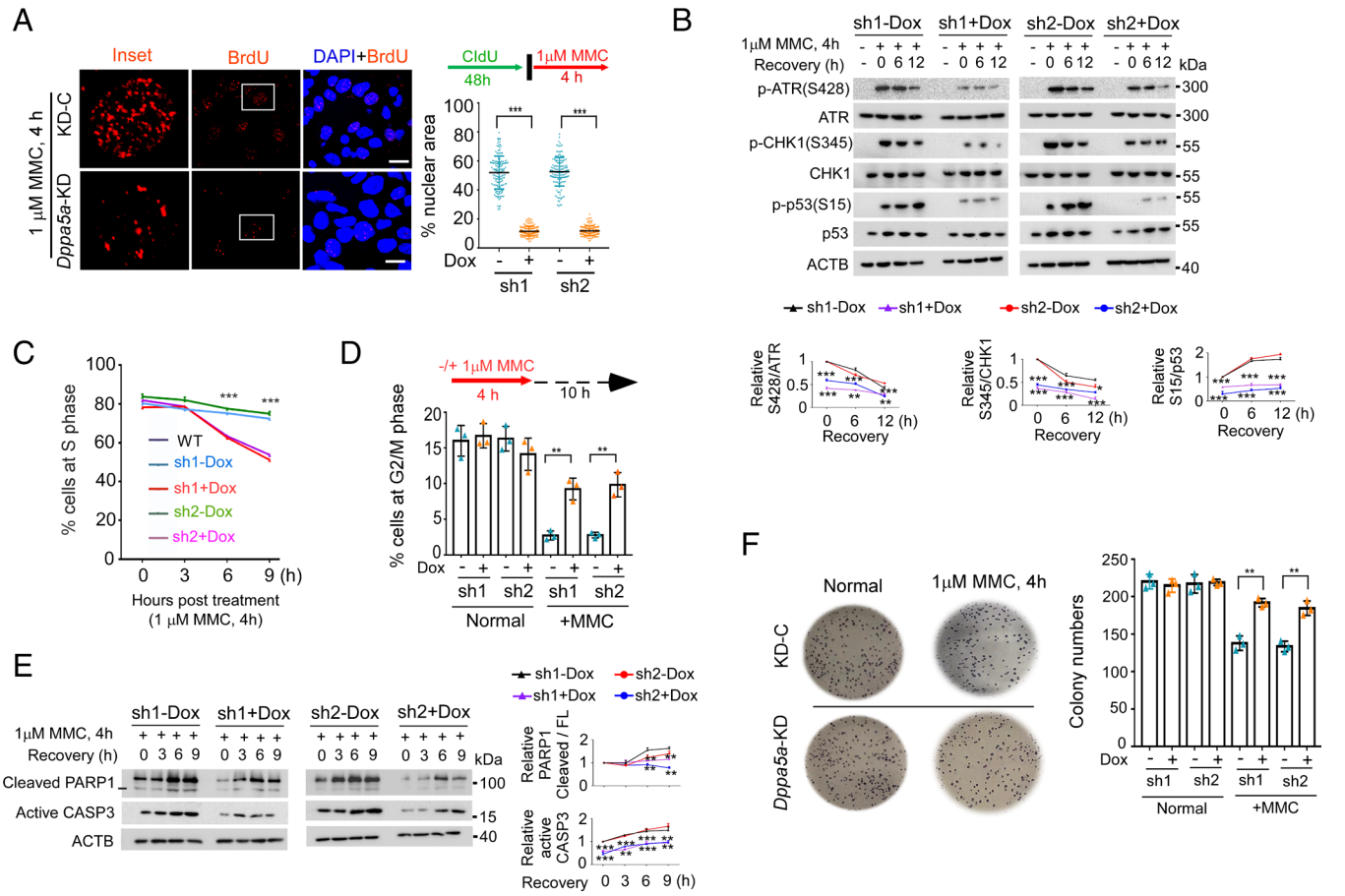


Fig. 4. DPPA5A KD ESCs are tolerant to DNA replication stress. (*A*) *Dppa5a* KD in ESCs compromised the generation of ssDNA induced by MMC treatment. ssDNA was labeled and detected under the nondenaturing condition. The left panel shows the representative images, and the right panel shows the quantification. At least 150 cells from three independent experiments were calculated. (Scale bar, 10 μm.) (*B*) Immunoblotting showed that the ATR-CHK1-p53 signaling was compromised in *Dppa5a* KD ESCs under MMC treatment. (*C*) After MMC treatment, the intra-S-phase checkpoint prevented the S-phase progression in DPPA5A-proficient ESCs. However, less S-phase cells were detected in *Dppa5a*-KD ESCs, indicating the deficiency in intra-S-phase checkpoint. (*D*) Consistently, significantly higher proportions of cells progressed to the G2/M phase after MMC treatment in *Dppa5a*-KD ESCs than in DPPA5A-proficient ESCs. (*E*) Compared to DPPA5A-proficient ESCs, *Dppa5a* KD ESCs were more tolerant to MMC treatment-induced apoptosis, as monitored by CASP3 activation and PARP1 cleavage. (*F*) The classical colony formation assay showed that *Dppa5a* KD ESCs formed more alkaline phosphatase–positive (AP^+) colonies than DPPA5A-proficient ESCs after MMC treatment. The relative protein levels in (*B* and *E*) were normalized by ACTB. All experiments were repeated three times with similar results. Data are shown as mean ± SEM. Two-tailed Student's *t* test, $^*P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$.

In U2AF2 KD ESCs, DPPA5A failed to coimmunoprecipitate with the U2 components (SF3B1, SF3B3, SF3A3, and U2AF2) (Fig. 6*E*). Similarly, depletion of SF3B1 compromised the U2 assembly (Fig. 6*F*) and diminished the association of DPPA5A with SF3B3, SF3A3, and U2AF2 (Fig. 6*G*). Of note, the interaction of DPPA5A with U2 components was reproducibly detected by reciprocal immunoprecipitation in NIH3T3 cells with overexpression of DPPA5A (Fig. 6*H*). These observations suggested that DPPA5A was a U2 complex–associated RBP in mouse ESCs. Importantly, KD of *Dppa5a* did not affect U2 components association (Fig. 6*I*), indicating that DPPA5A was not essential for the U2 complex assembly in ESCs.

DPPA5A Directly Binds to the GA-Rich Motif in the CE of *Rev1* and Regulates CE Inclusion. To further understand how DPPA5A regulated the CE splicing of *Rev1* and *Polq* genes, we went on to identify the binding motif of DPPA5A on its target RNAs. To this end, we performed the RNA immunoprecipitation (RIP) combined with sequencing to find the potential DPPA5A-bound RNAs under normal and HU treatment conditions. We obtained a list of genes whose RNA transcripts were associated with DPPA5A (Dataset S7). In line with the observation that *Dppa5a* KD affected the AS events of many DDR genes, GO term analysis of these genes enriched the processes of DDR and repair (Dataset S7 and *SI Appendix*, Fig. S7 *E* and *F*). Notably, among the five DDR genes (*Rev1*, *Polq*, *Casp2, Taok2,* and *Alkbh3*) whose CE splicing was affected by *Dppa5a* KD, only *Rev1* transcripts were detected to interact with DPPA5A, as validated by RT-PCR analysis of RIP pull-down samples (Fig. 7*A*). This result suggested that the regulation of DPPA5A on *Rev1* was direct, whereas its influences on *Polq, Casp2, Taok2,* and *Alkbh3* were indirect. Based on the DPPA5A-associated RNA species, we carried out motif enrichment analysis and predicted a potential motif that might engage in DPPA5A binding (Fig. 7*B*). This RNA motif was characterized by the GA-rich element and was detected in the CE of the *Rev1* gene and in other DPPA5A target genes (Dataset S8), but not in the CE of *Polq, Casp2, Taok2,* and *Alkbh3* genes. To verify the binding motif, we synthesized the biotin-labeled RNA fragments within the CE of *Rev1* carrying the GA-rich element. We also mutated GA elements by converting G to C and synthesized the biotin-labeled RNA fragments (Fig. 7*C*). The in vitro RNA pull-down assay showed that the partial mutation of GA elements decreased the binding of DPPA5A to CE fragment, and the mutation of all GA elements further reduced the CE-DPPA5A association (Fig. 7*C*). Thus, the GA-rich element is essential for binding to DPPA5A.

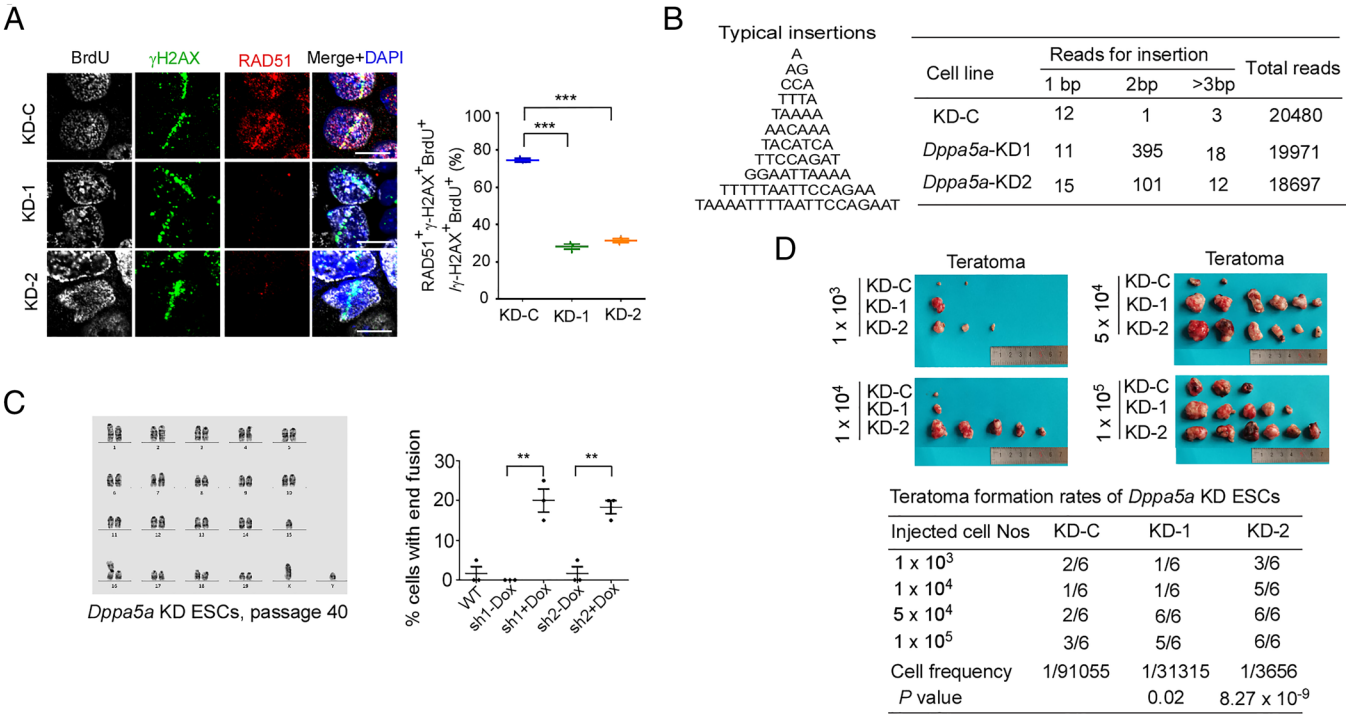


Fig. 5. DPPA5A promotes HR repair of DNA DSBs by suppressing the MMEJ pathway. (*A*) ESCs were subject to laser microirradiation. Compared to *Dppa5a* KD control (KD-C), *Dppa5a* KD significantly impaired HR repair of DSBs at the S phase ($BrdU^+$), as indicated by the failure to recruit RAD51 to DSB sites labeled with γ-H2AX. The left panel shows the representative images, and the right panel shows the quantification. At least 50 cells from three independent experiments were analyzed in each group. (Scale bar, 10 μm.) (*B*) *Dppa5a* KD significantly increased the frequencies of insertions ≥2 bp long, indicating that the MMEJ repair pathway was significantly enhanced by *Dppa5a* KD. Examples of insertions are shown in the left panel, and the reads for each type of insertion are summarized in the right panel. The experiments were repeated twice with similar results. (*C*) G-band karyotyping revealed the higher frequency of chromosomal fusion in *Dppa5a* KD ESCs than in DPPA5A-proficient ESCs. The left panel shows the G-band karyotype. The right panel shows the statistical analysis. At least 60 metaphase spreads from three independent experiments were analyzed in each group. (*D*) The teratoma formation rates were significantly higher in *Dppa5a* KD ESCs than in *Dppa5a* KD control (KD-C) ESCs. The upper panel shows the teratomas, and the lower panel shows the rates. Six BALB/c nude mice were used in each group. Data are presented as mean ± SEM. Two-tailed Student's *t* test, $^{**}P < 0.01$ and $^{***}P < 0.001$.

Next, we investigated whether the binding of DPPA5A to this GA-rich motif was required for regulating CE inclusion in *Rev1*. To this end, we performed the minigene assay (61). The minigene was composed of the intron 1, exon 2 (5′ exon), intron 2, and exon 3 fused with exon 4 (3′ exon) of the human *SERPING1/CINH* gene. The expression of the minigene was driven by the elongation factor 1alpha short (EFS) promoter (62). The CE variants of the *Rev1* gene (WT and mutants containing G to C conversion) were inserted into intron 2. The plasmids containing the minigene were transfected into WT ESCs. The RNA splicing pattern of minigene transcripts was verified by RT-PCR using primers located on the 5′ exon and 3′ exon (Fig. 7*D*). In line with the decreased binding of mutant CE to DPPA5A, the mutations compromised the CE inclusion rate (Fig. 7*D*). Thus, DPPA5A binding to the GA-rich motif was essential for regulation of CE inclusion in *Rev1*. Taken together, we proposed a working model in which DPPA5A associates with the U2 complex and binds to the GA-rich motif in the CE of *Rev1* to promote its splicing (Fig. 7*E*).

Discussion

Avoiding mutagenic DNA replication and repair pathways is crucial to maintain genomic stability. TLS-mediated DNA replication and MMEJ-mediated DSB repair are highly mutagenic. In this study, we show that mouse ESCs, which require a very stable genome to maintain stem cell identity, possess low TLS and MMEJ activity due to the low expression of REV1 and POLθ proteins. Notably, the protein expression of REV1 and POLθ is regulated at a posttranscriptional level via CE inclusion. Compared to differentiated cells, mouse ESCs had a significantly higher rate of CE inclusion in *Rev1* and *Polq* genes. CE inclusion in the two genes disrupted the normal protein expression and attenuated TLS and MMEJ activity. Moreover, we identified an ESC-specific RBP DPPA5A as the regulator of CE inclusion in *Rev1* and *Polq*. Thus, the DPPA5A-regulated CE inclusion in *Rev1* and *Polq* represents an important mechanism safeguarding the genomic stability of ESCs. DPPA5A directly regulated the CE inclusion of *Rev1*, whereas its regulation on CE splicing of *Polq* was indirect. In the future, identifying the RBP with direct role on CE inclusion of *Polq* warrants further investigation.

Due to the multiple functions of DPPA5A in regulating DNA replication and DSB repair pathway choice, depletion of DPPA5A in ESCs caused a wide range of defects in DNA replication stress response and DSB repair. *Dppa5a* KD ESCs prefer to utilize the TLS pathway to bypass DNA lesions and continue DNA replication. This renders the cells' resistance to replication stress. Indeed, after treatment with MMC or HU, the ATR-CHK1 replication stress signaling activation was compromised, and the KD ESCs were tolerant to damage treatment. In addition, the DSB repair pathway shifted from the high-fidelity HR pathway to the highly mutagenic MMEJ pathway in *Dppa5a* KD ESCs. As a result, *Dppa5a* KD ESCs were tumorigenic as evidenced by the high tendency to form teratomas. DPPA5A is highly expressed in undifferentiated ESCs and is used as a marker of pluripotency. Our study indicates that DPPA5A can be used as a marker for genomic stability and stem cell quality. DPPA5A is also highly expressed in oocytes and is conserved in human and nonhuman primates. A study on monkey oocyte aging showed that the RNA expression level of *DPPA5A* in

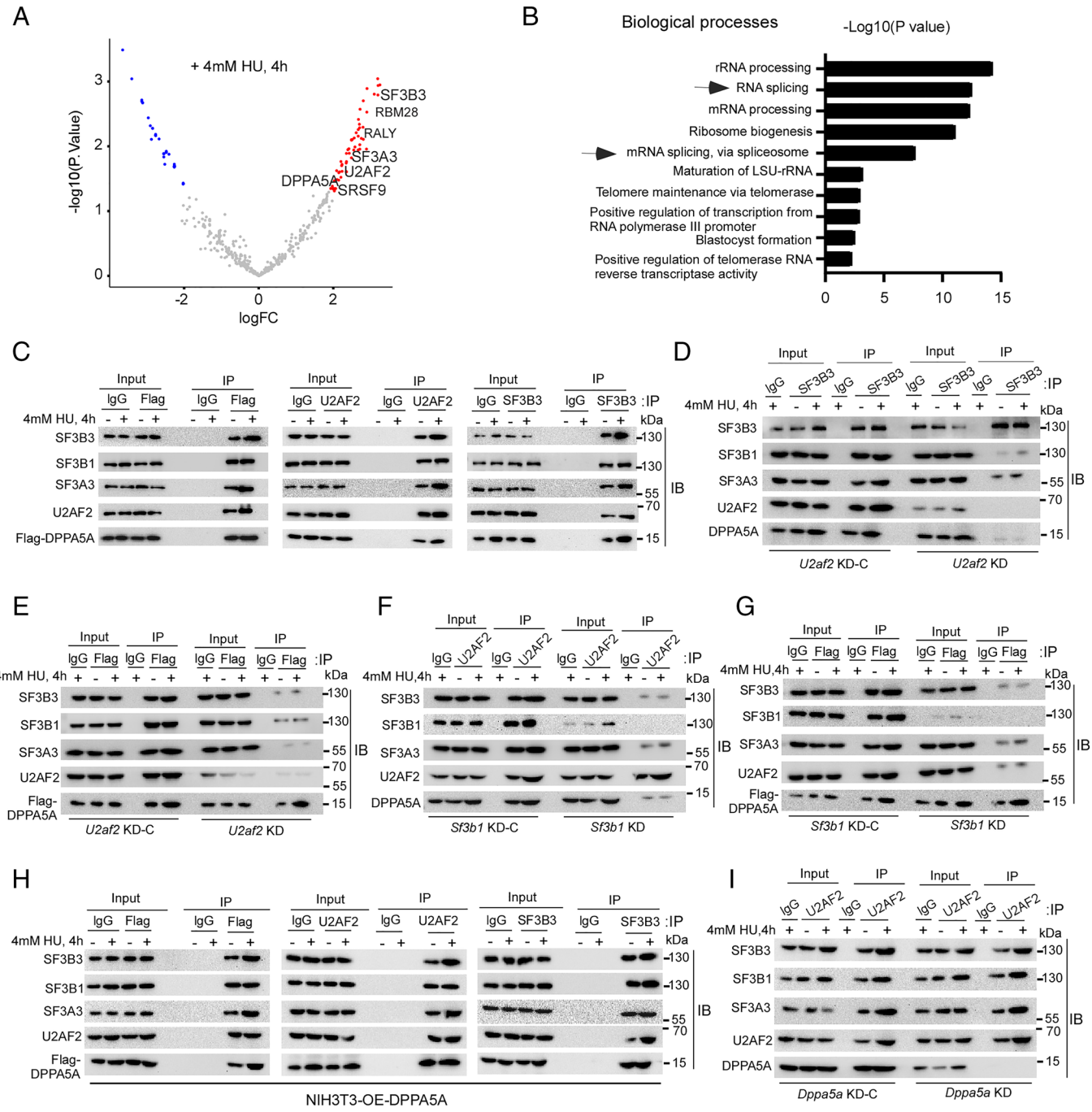


Fig. 6. DPPA5A associates with the U2 snRNA complex, but is not required for the U2 complex assembly. (*A*) Volcano plot analysis of the DPPA5A-interacting proteins identified by immunoprecipitation followed by mass spectrometry analysis under the HU treatment condition. IP with IgG was used as control. Blue dots indicate proteins enriched by IgG IP ($P < 0.05$, fold change ≥ 2). Red dots indicate proteins enriched by DPPA5A IP ($P < 0.05$, fold change ≥ 2). (*B*) Gene ontology analysis of proteins enriched by DPPA5A IP. Top 10 biological process terms are shown. (*C*) Reciprocal immunoprecipitation under normal and HU treatment conditions showed that DPPA5A interacted with the core components of U2 RNP of spliceosome SF3B3, SF3B1, SF3A3, and U2AF2 in mouse ESCs. (*D*) KD of *U2af2* in ESCs disrupted the U2 complex assembly, as indicated by the failure to co-immunoprecipitate SF3B3 with other protein components of U2. (*E*) In *U2af2* KD ESCs, DPPA5A failed to coimmunoprecipitate with the U2 components. (*F*) Similarly, KD of *Sf3b1* compromised the U2 assembly. (*G*) KD of *Sf3b1* diminished the association of DPPA5A with the U2 components SF3B3, SF3B1, SF3A3, and U2AF2. (*H*) In NIH3T3 cells with overexpression (OE) of Flag-tagged DPPA5A, the interaction of DPPA5A with U2 components was reproducibly detected by reciprocal immunoprecipitation. (*I*) KD of *Dppa5a* did not affect the U2 complex assembly in ESCs. Experiments were repeated three times with similar results.

oocytes decreased significantly in old monkeys (63), suggesting that DPPA5A in oocytes may play a similarly important function in regulating the oocyte quality. Although this study reported the important role of DPPA5A in regulating the cryptic splicing of *Rev1* and *Polq* to suppress the mutagenic TLS and MMEJ pathways in mouse ESCs, the functions of DPPA5A in human ESCs were not assessed and should be thoroughly characterized in the future.

Materials and Methods

Cell Culture and ESC Differentiation. Mouse ESCs were derived from the inner cell mass of embryonic day 3.5 (E3.5) blastocysts of C57BL/6J mouse strain in our laboratory and maintained on the mitomycin-treated MEFs (18). Mouse ESCs were differentiated in ESC culture medium without leukemia inhibitory factor (LIF). More details can be found in *SI Appendix*, Materials and Methods.

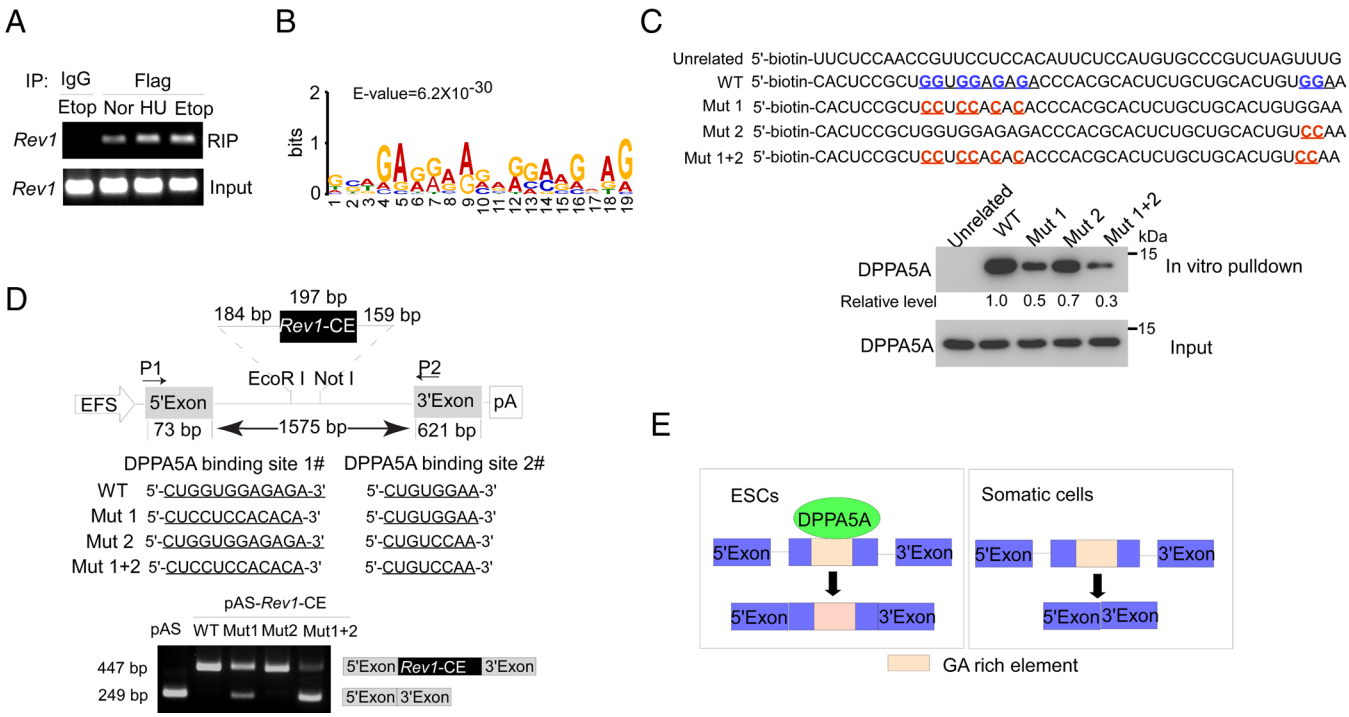


Fig. 7. DPPA5A binds to the GA-rich motif in the CE of *Rev1* and regulates CE inclusion. (*A*) RT-PCR analysis of RIP samples confirmed the interaction of DPPA5A with *Rev1* transcripts under normal culture and replication stress conditions. (*B*) The potential GA-rich RNA motif implicated in binding to DPPA5A. (*C*) Binding affinity assay between DPPA5A and *Rev1* RNA fragments containing WT and mutants' motifs using in vitro RNA pull down. (*D*) *Rev1* CE variants containing WT motif or mutant motif were inserted into the splicing reporter. Minigene assays showed that the mutation of the GA-rich element compromised the CE inclusion in the minigene. (*E*) A working model showing that DPPA5A associates with the U2 complex and binds to the GA-rich motif in the CE of the *Rev1* gene to promote the CE splicing. Experiments were repeated three times with similar results.

RNA-seq and Alternative Splicing Analysis. RNA-seq libraries were constructed and sequenced in Novogene company using an Illumina NovaSeq 6000 platform. RNA-seq reads were mapped to mm10 genome using Tophat2 software (version 2.0.9) (64), and alternative splicing events were identified using ASD software (65). More details can be found in *SI Appendix, Materials and Methods*.

In Vitro RNA Pull-Down and Mass Spectrometry Analysis. RNA fragments were synthesized or transcribed with T7 RNA polymerase and labeled with biotin. In vitro RNA pull down was conducted as previously described (3). The proteins were recovered and subject to SDS-PAGE for mass spectrometry (MS) and western blot analysis. More details can be found in *SI Appendix, Materials and Methods*.

RNA Immunoprecipitation Sequencing (RIP-seq). RNA immunoprecipitation was conducted as previously described (3). Detailed methods can be found in *SI Appendix, Materials and Methods*.

Lentivirus Package and Cell Line Generation. The expression vectors were cotransfected with psPAX2 and pMD2.G packaging plasmids in 293 T cells to package lentivirus. ESCs were transducted with virus to establish stable cell lines. Detailed methods were shown in *SI Appendix, Materials and Methods*.

Immunoprecipitation, Immunoblotting, and Immunofluorescent Staining. Immunoprecipitation, immunoblotting, and immunofluorescent staining were performed following standard procedures. Detailed methods can be found in *SI Appendix, Materials and Methods*.

qRT-PCR. qRT-PCR was conducted following standard procedures using the SYBR Green Premix Ex Taq II (Takara Bio, RR820A). Detailed methods can be found in *SI Appendix, Materials and Methods*.

DNA Fiber Assay. Replicating DNA labeled by IdU and CldU was treated or untreated with S1 nuclease and subject to DNA fiber assays as described (3). Detailed methods can be found in *SI Appendix, Materials and Methods*.

DNA Damage and Genome Stability Assay. The DNA damage and genome stability assay including the ssDNA assay, neutral comet assay, karyotyping, and micronucleus analysis were performed as described previously (3, 18). Detailed methods can be found in *SI Appendix, Materials and Methods*.

S Phase and G2/M Cell Cycle Checkpoints. Cells were labeled with BrdU for S phase checkpoint analysis and stained with p-H3(S10) for G2/M checkpoint analysis. Detailed methods can be found in *SI Appendix, Materials and Methods*.

Colony Formation Assay. Cells were seeded at 450 cells/well and grown in a 12-well plate, treated with or without HU or MMC, and cultured for 9 d. The colonies were stained for alkaline phosphatase (AP) and counted. More details can be found in *SI Appendix, Materials and Methods*.

HR and NHEJ DNA Repair Pathway Analysis. The HR DNA repair pathway was analyzed using the laser microirradiation assay. The NHEJ DNA repair pathway was analyzed using a mouse NHEJ-GFP reporter ESC line. More details can be found in *SI Appendix, Materials and Methods*.

MMEJ-Mediated DSB Repair Assay. The Cas9-gRNA system was used to induce DNA DSBs at a defined genetic locus. Cells were lysed, and genomic scars were amplified using the standard PCR protocol. Amplified products were added adapters and sequenced on an Illumina HiSeq platform. The template insertions were measured. More details can be found in *SI Appendix, Materials and Methods*.

Teratoma Formation. Mouse ESCs were subcutaneously injected into the BALB/c nude mice. After a month, teratoma formation ability and stem cell frequency were analyzed (https://bioinf.wehi.edu.au/software/elda/). More details can be found in *SI Appendix, Materials and Methods*. The studies on mice received ethical approval of the Animal Care and Use Committee of the Kunming Institute of Zoology, Chinese Academy of Sciences.

Minigene Assay. The minigene assay was performed using a splicing reporter as previously described with a minor modification (62). Detailed methods can be found in *SI Appendix, Materials and Methods*.

Statistical Analysis. Statistical tests with the sample size were listed in the text and figure legends. Statistical analyses were performed using GraphPad Prism 7.0 (GraphPad Software) by two-tailed Student's *t* test. $P < 0.05$ was considered significant. All data were provided as the mean ± SEM.

Data, Materials, and Software Availability. All other data are included in the manuscript and/or supporting information. RNA-Seq and RIP-seq data were deposited in Gene Expression Omnibus (GEO) under accession number

GSE221654 (66). The raw data of mass spectrometry in Fig. 6*A* were deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the iProX partner repository with the dataset identifier PXD042052 (67).

ACKNOWLEDGMENTS. We thank Prof. An-yong Xie in Zhejiang University School of Medicine for providing Non-homologous end joining-Green fluorescent protein (NHEJ-GFP) reporter embryonic stem cells and the I-SceI expression plasmid. This work was supported by the National Key Research & Developmental Program of China (2021YFA1102000), the National Natural Science Foundation of China (31930027), and Yunnan Fundamental Research Projects (Grant no. 202301AS070062).

Author affiliations: [a]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China; [b]University of Chinese Academy of Sciences, Beijing 101408, China; [c]Department of Reproductive Medicine, The Second Affiliated Hospital of Kunming Medical University, Kunming 650101, China; [d]Key Laboratory of Animal Models and Human Disease Mechanisms of Yunnan Province, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China; [e]Department of Analytical Chemistry and Key Laboratory of Receptor Research of Chinese Academy of Sciences, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China; and [f]The Chinese University of Hong Kong and Kunming Institute of Zoology Joint Laboratory of Bioresources and Molecular Research in Common Diseases, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China

1. S. P. Wyles, E. B. Brandt, T. J. Nelson, Stem cells: The pursuit of genomic stability. *Int. J. Mol. Sci.* **15**, 20948–20967 (2014).
2. E. D. Tichy, P. J. Stambrook, DNA repair in murine embryonic stem cells and differentiated cells. *Exp. Cell Res.* **314**, 1929–1936 (2008).
3. L. Wang *et al.*, A novel lncRNA Discn fine-tunes replication protein A (RPA) availability to promote genomic stability. *Nat. Commun.* **12**, 5572 (2021).
4. B. S. Heyer, A. MacAuley, O. Behrendtsen, Z. Werb, Hypersensitivity to DNA damage leads to increased apoptosis during early mouse development. *Genes Dev.* **14**, 2072–2084 (2000).
5. O. Thompson *et al.*, Low rates of mutation in clinical grade human pluripotent stem cells under different culture conditions. *Nat. Commun.* **11**, 1528 (2020).
6. P. W. Andrews *et al.*, The consequences of recurrent genetic and epigenetic variants in human pluripotent stem cells. *Cell Stem Cell* **29**, 1624–1636 (2022).
7. L. M. Mus *et al.*, Recurrent chromosomal imbalances provide selective advantage to human embryonic stem cells under enhanced replicative stress conditions. *Genes Chromosomes Cancer* **60**, 272–281 (2021).
8. J. Halliwell, I. Barbaric, P. W. Andrews, Acquired genetic changes in human pluripotent stem cells: Origins and consequences. *Nat. Rev. Mol. Cell Biol.* **21**, 715–728 (2020).
9. F. T. Merkle *et al.*, Whole-genome analysis of human embryonic stem cells enables rational line selection based on genetic variation. *Cell Stem Cell* **29**, 472–486 (2022).
10. Y. Avior, E. Lezmi, K. Eggan, N. Benvenisty, Cancer-related mutations identified in primed human pluripotent stem cells. *Cell Stem Cell* **28**, 10–11 (2021).
11. S. Bar, N. Benvenisty, Epigenetic aberrations in human pluripotent stem cells. *EMBO J.* **38**, e101033 (2019).
12. H. T. Nguyen *et al.*, Gain of 20q11.21 in human embryonic stem cells improves cell survival by increased expression of Bcl-xL. *Mol. Hum. Reprod.* **20**, 168–177 (2014).
13. F. J. Rouhani *et al.*, Substantial somatic genomic variation and selection for BCOR mutations in human induced pluripotent stem cells. *Nat. Genet.* **54**, 1406–1416 (2022).
14. F. T. Merkle *et al.*, Human pluripotent stem cells recurrently acquire and expand dominant negative P53 mutations. *Nature* **545**, 229–233 (2017).
15. P. J. Stambrook, E. D. Tichy, Preservation of genomic integrity in mouse embryonic stem cells. *Adv. Exp. Med. Biol.* **695**, 59–75 (2010).
16. P. Zheng, Current understanding of genomic stability maintenance in pluripotent stem cells. *Acta. Biochim. Biophys. Sin. (Shanghai)* **54**, 858–863 (2022).
17. J. Xiong *et al.*, Stemness factor Sall4 is required for DNA damage response in embryonic stem cells. *J. Cell Biol.* **208**, 513–520 (2015).
18. B. Zhao *et al.*, Filia is an ESE-specific regulator of DNA damage response and safeguards genomic stability. *Cell Stem Cell* **16**, 684–698 (2015).
19. E. D. Tichy *et al.*, Mouse embryonic stem cells, but not somatic cells, predominantly use homologous recombination to repair double-strand DNA breaks. *Stem Cells Dev.* **19**, 1699–1711 (2010).
20. B. Zhao *et al.*, Mouse embryonic stem cells have increased capacity for replication fork restart driven by the specific Filia-Floped protein complex. *Cell Res.* **28**, 69–89 (2018).
21. M. Zalzman *et al.*, Zscan4 regulates telomere elongation and genomic stability in ES cells. *Nature* **464**, 858–863 (2010).
22. R. Le *et al.*, Dcaf11 activates Zscan4-mediated alternative telomere lengthening in early embryos and embryonic stem cells. *Cell Stem Cell* **28**, 732–747 (2021).
23. M. Markiewicz-Potoczny *et al.*, TRF2-mediated telomere protection is dispensable in pluripotent stem cells. *Nature* **589**, 110–115 (2021).
24. P. Ruis *et al.*, TRF2-independent chromosome end protection during pluripotency. *Nature* **589**, 103–109 (2021).
25. T. Lin *et al.*, p53 induces differentiation of mouse embryonic stem cells by suppressing Nanog expression. *Nat. Cell Biol.* **7**, 165–171 (2005).
26. R. Dumitru *et al.*, Human embryonic stem cells have constitutively active Bax at the Golgi and are primed to undergo rapid apoptosis. *Mol. Cell* **46**, 573–583 (2012).
27. M. Li *et al.*, Distinct regulatory mechanisms and functions for p53-activated and p53-repressed DNA damage response genes in embryonic stem cells. *Mol. Cell* **46**, 30–42 (2012).
28. J. C. Liu *et al.*, High mitochondrial priming sensitizes hESCs to DNA-damage-induced apoptosis. *Cell Stem Cell* **13**, 483–491 (2013).
29. H. Ma, Y. Ning, L. Wang, W. Zhang, P. Zheng, Lnc956 regulates mouse embryonic stem cell differentiation in response to DNA damage in a p53-independent pathway. *Sci. Adv.* **9**, eade9742 (2023).
30. A. K. Ahuja *et al.*, A short G1 phase imposes constitutive replication stress and fork remodelling in mouse embryonic stem cells. *Nat. Commun.* **7**, 10660 (2016).
31. W. Zhang *et al.*, Lnc956-TRIM28-HSP90B1 complex on replication forks promotes CMG helicase retention to ensure stem cell genomic stability and embryogenesis. *Sci. Adv.* **9**, eadf6277 (2023).
32. Q. Pan, O. Shai, L. J. Lee, B. J. Frey, B. J. Blencowe, Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* **40**, 1413–1415 (2008).
33. E. T. Wang *et al.*, Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**, 470–476 (2008).
34. F. E. Baralle, J. Giudice, Alternative splicing as a regulator of development and tissue identity. *Nat. Rev. Mol. Cell Biol.* **18**, 437–451 (2017).
35. H. Shen *et al.*, Mouse totipotent stem cells captured and maintained through spliceosomal repression. *Cell* **184**, 2843–2859 (2021).
36. M. Gabut *et al.*, An alternative splicing switch regulates embryonic stem cell pluripotency and reprogramming. *Cell* **147**, 132–146 (2011).
37. A. T. Bademosi, A. K. Walker, Cryptic inclusions UNCover losses driving neurodegeneration. *Trends Genet.* **38**, 889–891 (2022).
38. A. Vaisman, R. Woodgate, Translesion DNA polymerases in eukaryotes: What makes them tick? *Crit. Rev. Biochem. Mol. Biol.* **52**, 274–303 (2017).
39. J. G. Jansen, A. Tsaalbi-Shtylik, N. de Wind, Roles of mutagenic translesion synthesis in mammalian genome stability, health and disease. *DNA Repair (Amst)* **29**, 56–64 (2015).
40. D. A. Ramsden, J. Carvajal-Garcia, G. P. Gupta, Mechanism, cellular functions and cancer roles of polymerase-theta-mediated DNA end joining. *Nat. Rev. Mol. Cell Biol.* **23**, 125–140 (2022).
41. J. H. Seol, E. Y. Shim, S. E. Lee, Microhomology-mediated end joining: Good, bad and ugly. *Mutat. Res.* **809**, 81–87 (2018).
42. P. A. Mateos-Gomez *et al.*, Mammalian polymerase theta promotes alternative NHEJ and suppresses recombination. *Nature* **518**, 254–257 (2015).
43. H. Techer *et al.*, Replication dynamics: Biases and robustness of DNA fiber analysis. *J. Mol. Biol.* **425**, 4845–4855 (2013).
44. B. Song, S. Yang, G. H. Hwang, J. Yu, S. Bae, Analysis of NHEJ-Based DNA Repair after CRISPR-Mediated DNA cleavage. *Int. J. Mol. Sci.* **22**, 6397 (2021).
45. J. Schimmel, R. van Schendel, J. T. den Dunnen, M. Tijsterman, Templated Insertions: A Smoking Gun for Polymerase Theta-Mediated End Joining. *Trends Genet.* **35**, 632–644 (2019).
46. M. J. Yousefzadeh *et al.*, Mechanism of suppression of chromosomal instability by DNA polymerase POLQ. *PLoS Genet.* **10**, e1004654 (2014).
47. L. Seufert, T. Benzing, M. Ignarski, R. U. Muller, RNA-binding proteins and their role in kidney disease. *Nat. Rev. Nephrol.* **18**, 153–170 (2022).
48. S. C. Kwon *et al.*, The RNA-binding protein repertoire of embryonic stem cells. *Nat. Struct. Mol. Biol.* **20**, 1122–1130 (2013).
49. C. He *et al.*, High-resolution mapping of RNA-binding regions in the nuclear proteome of embryonic stem cells. *Mol. Cell* **64**, 416–430 (2016).
50. X. R. Ma *et al.*, TDP-43 represses cryptic exon inclusion in the FTD-ALS gene UNC13A. *Nature* **603**, 124–130 (2022).
51. T. S. Tanaka *et al.*, Esg1, expressed exclusively in preimplantation embryos, germline, and embryonic stem cells, is a putative RNA-binding protein with broad RNA targets. *Dev. Growth Differ.* **48**, 381–390 (2006).
52. H. Amano *et al.*, Identification and targeted disruption of the mouse gene encoding ESG1 (PH34/ECAT2/DPPA5). *BMC Dev. Biol.* **6**, 11 (2006).
53. S. K. Kim *et al.*, Identification of developmental pluripotency associated 5 expression in human pluripotent stem cells. *Stem Cells* **23**, 458–462 (2005).
54. N. Chatterjee *et al.*, REV1 inhibitor JH-RE-06 enhances tumor cell response to chemotherapy by triggering senescence hallmarks. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 28918–28921 (2020).
55. J. L. Wojtaszek *et al.*, A small molecule targeting mutagenic translesion synthesis improves chemotherapy. *Cell* **178**, 152–159 (2019).
56. B. Pilzecker, O. A. Buoninfante, H. Jacobs, DNA damage tolerance in stem cells, ageing, mutagenesis, disease and cancer therapy. *Nucleic Acids Res.* **47**, 7163–7181 (2019).
57. A. Simoneau, L. Zou, An extending ATR-CHK1 circuitry: The replication stress response and beyond. *Curr. Opin. Genet. Dev.* **71**, 92–98 (2021).
58. D. R. Iyer, N. Rhind, The intra-S checkpoint responses to DNA damage. *Genes (Basel)* **8**, 74 (2017).
59. Y. L. Feng *et al.*, H2AX facilitates classical non-homologous end joining at the expense of limited nucleotide loss at repair junctions. *Nucleic Acids Res.* **45**, 10614–10633 (2017).
60. C. van der Feltz, A. A. Hoskins, Structural and functional modularity of the U2 snRNP in pre-mRNA splicing. *Crit. Rev. Biochem. Mol. Biol.* **54**, 443–465 (2019).
61. P. Gaildrat *et al.*, Use of splicing reporter minigene assay to evaluate the effect on splicing of unclassified genetic variants. *Methods Mol. Biol.* **653**, 249–257 (2010).
62. I. Tournier *et al.*, A large fraction of unclassified variants of the mismatch repair genes MLH1 and MSH2 is associated with splicing defects. *Hum. Mutat.* **29**, 1412–1424 (2008).
63. S. Wang *et al.*, Single-cell transcriptomic atlas of primate ovarian aging. *Cell* **180**, 585–600 (2020).
64. C. Trapnell *et al.*, Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010).
65. X. Zhou *et al.*, Transcriptome analysis of alternative splicing events regulated by SRSF10 reveals position-dependent splicing modulation. *Nucleic Acids Res.* **42**, 4019–4030 (2014).
66. F. Jiang *et al.*, Data for "Next generation sequencing discovers alternative splicing pattern among mouse embryonic stem cells, mouse embryonic fibroblasts and Dppa5a knockdown mouse embryonic stem cells and Dppa5a target RNAs". *NCBI Gene Expression Omnibus*. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE221654. Deposited 23 December 2022.
67. F. Jiang *et al.*, Data for "Identification of the interaction proteins of DPPA5A in mouse embryonic stem cells". iProX, an official member of ProteomeXchange Consortium. https://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD042052. Deposited 8 May 2023.