Supplementary Information

Simian-specific SAGE1 enhances germline genome stability by promoting homologous recombination

The PDF file includes:

Supplementary Figs S1 - S15

Supplementary Tables S4 - S5

Other Supplementary Materials for this manuscript include: Supplementary Tables S1(.doc), S2, S3 (.xlsx)

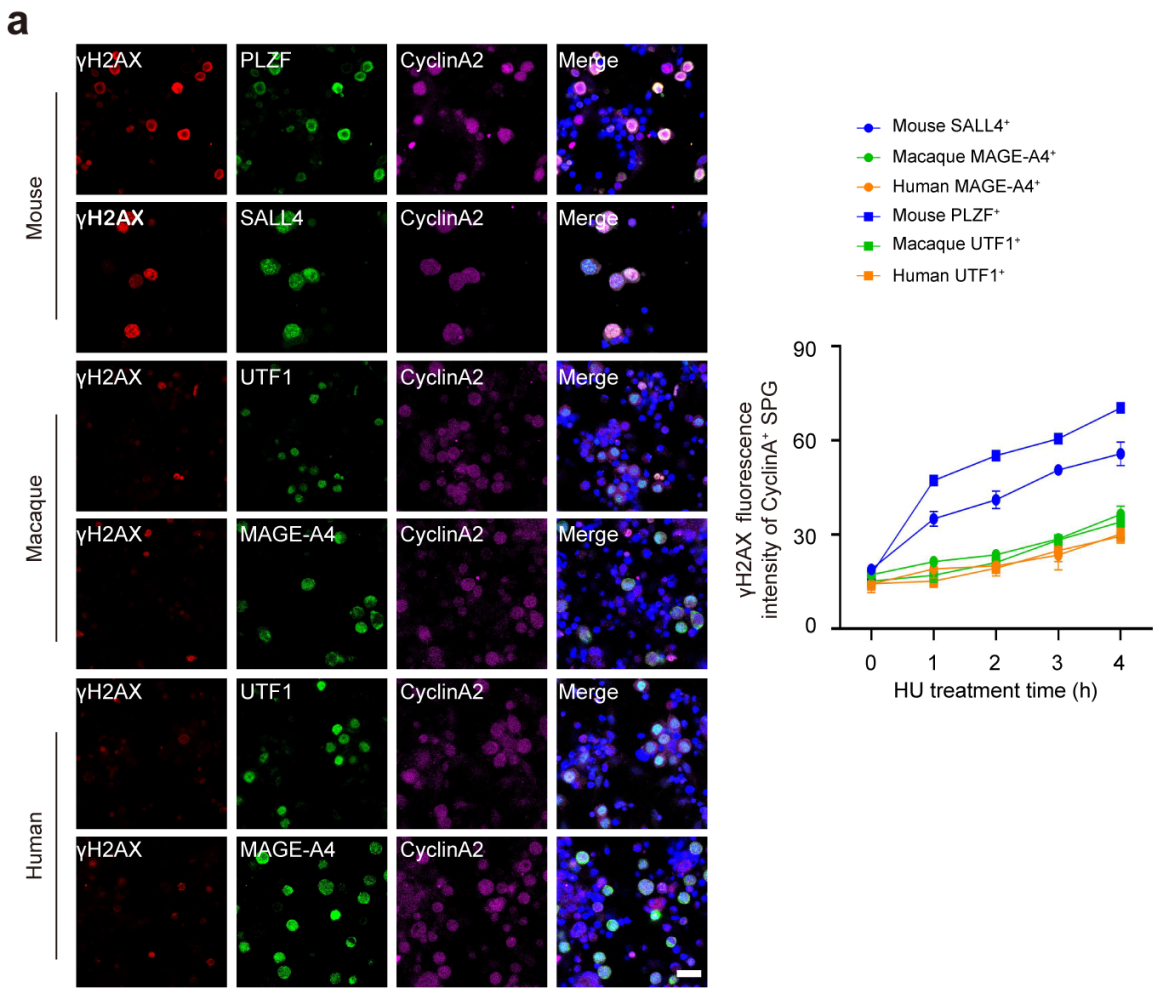

Fig. S1. Comparison of DNA damage sensitivity in cycling SPG across species. a Representative IF images of primary germ cells from human, crab-eating macaque, and mouse testicular tissues stained with antibodies against SPG markers (*Mm*SPG: PLZF, SALL4; *Mf*SPG and *Hs*SPG: UTF1, MAGE-A4), Cyclin A2 (S/G2 phase marker) and γH2AX after 2 h HU (2 mM) treatment. Scale bar, 20 μm. $n = 3$ independent replicates with at least 100 cells per condition were analyzed. For all experiments, data were shown as mean ± SD.

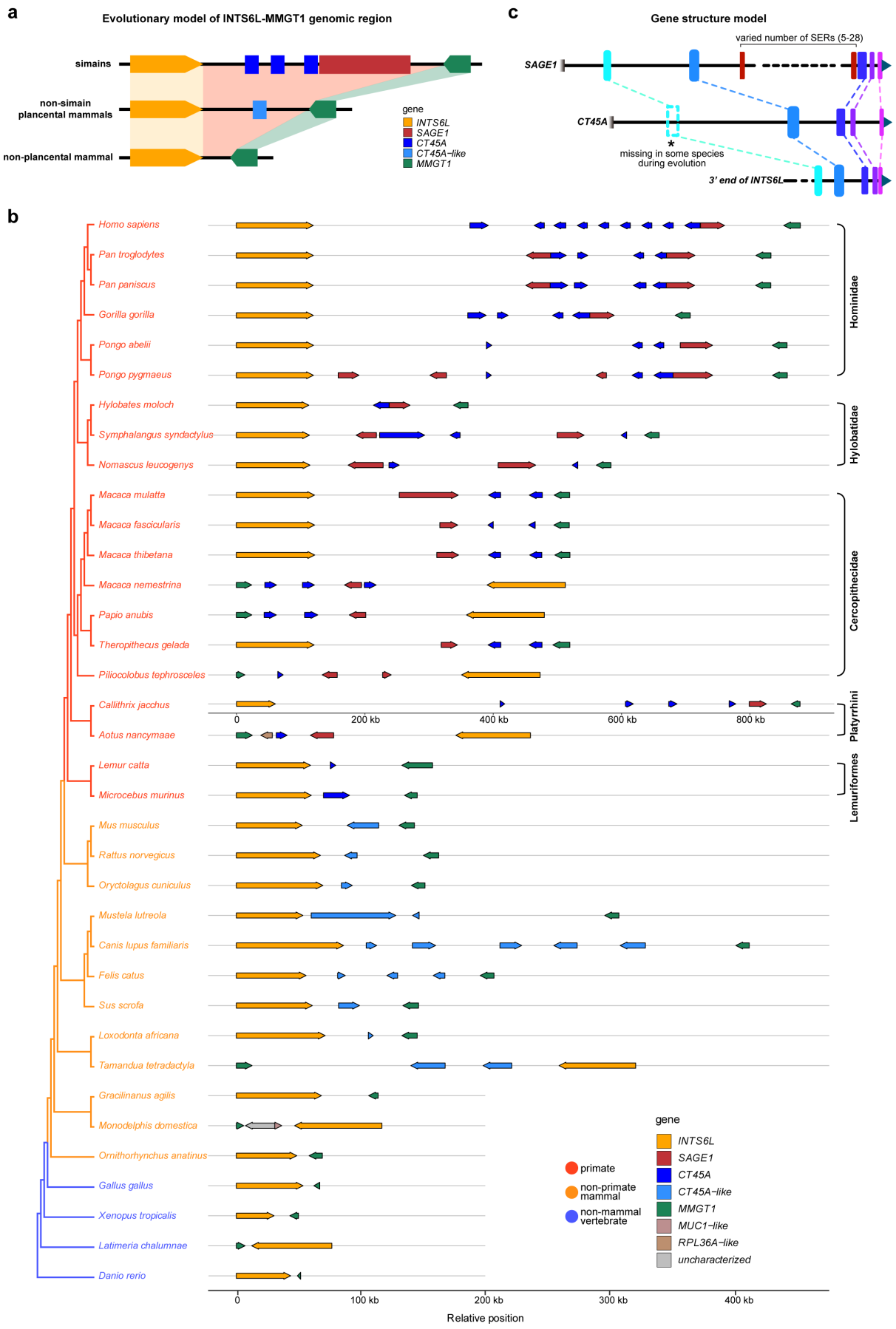
a
Evolutionary model of INTS6L-MMGT1 genomic region
simains
non-simain plancental mammals
non-plancental mammal
gene
INTS6L
SAGE1
CT45A
CT45A−like
MMGT1
c
Gene structure model
varied number of SERs (5-28)
SAGE1
CT45A
missing in some species during evolution
3' end of INTS6L
b
Homo sapiens
Pan troglodytes
Pan paniscus
Gorilla gorilla
Pongo abelii
Pongo pygmaeus
Hylobates moloch
Symphalangus syndactylus
Nomascus leucogenys
Macaca mulatta
Macaca fascicularis
Macaca thibetana
Macaca nemestrina
Papio anubis
Theropithecus gelada
Piliocolobus tephrosceles
Callithrix jacchus
Aotus nancymaae
Lemur catta
Microcebus murinus
Mus musculus
Rattus norvegicus
Oryctolagus cuniculus
Mustela lutreola
Canis lupus familiaris
Felis catus
Sus scrofa
Loxodonta africana
Tamandua tetradactyla
Gracilinanus agilis
Monodelphis domestica
Ornithorhynchus anatinus
Gallus gallus
Xenopus tropicalis
Latimeria chalumnae
Danio rerio
Hominidae
Hylobatidae
Cercopithecidae
Platyrrhini
Lemuriformes
0
200 kb
400 kb
600 kb
800 kb
primate
non-primate mammal
non-mammal vertebrate
gene
INTS6L
SAGE1
CT45A
CT45A−like
MMGT1
MUC1−like
RPL36A−like
uncharacterized
0
100 kb
200 kb
300 kb
400 kb
Relative position

Fig. S2. Phylogenetic synteny analysis of the *SAGE1* genomic locus. a Schematic model depicting the stepwise expansion of the *INTS6L-MMGT1* genomic region during vertebrate evolution and the emergence of *CT45A* and *SAGE1* families. **b** Gene composition and relative orientation within *INTS6L-MMGT1* genomic region across representative vertebrate species, from zebrafish to human. Genes whose protein sequences were incomplete were excluded. **c** Schematic gene structures of *SAGE1*, *CT45A* and the 3' end of *INTS6L*. Exons sharing high nucleotide identity are indicated by matching colors.

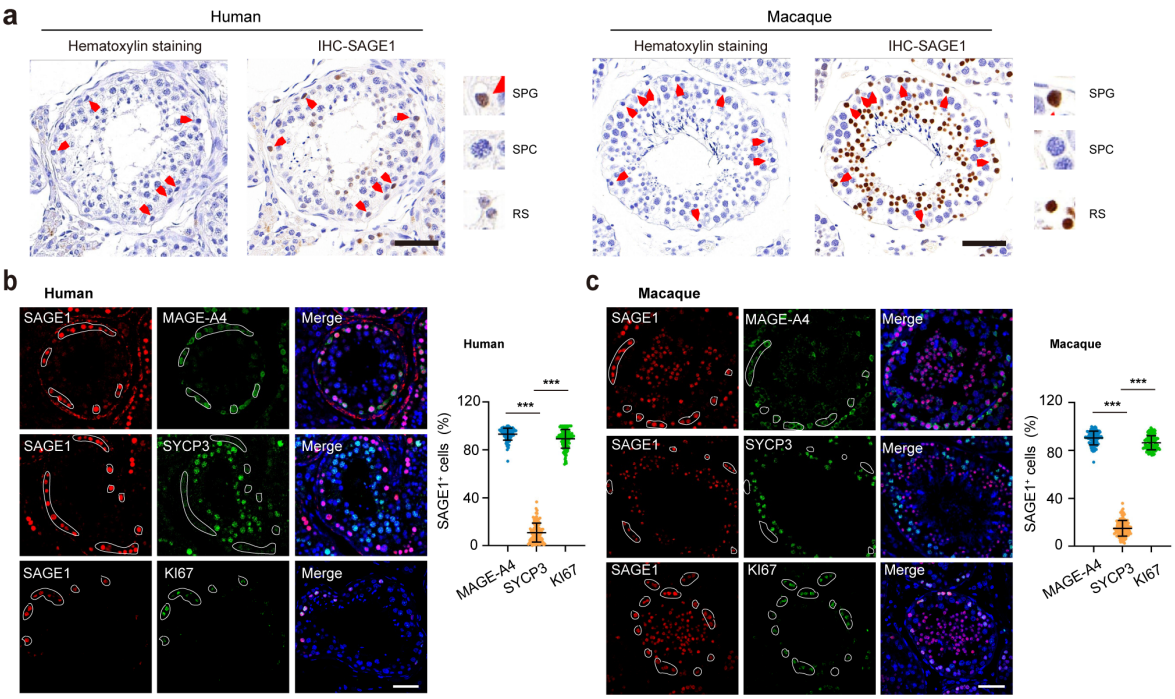

Fig. S3. SAGE1 is expressed in human and macaque spermatogonia. a Human and macaque adult testis sections were stained with hematoxylin to characterize SPG morphology, and then stained with anti-SAGE1 antibody. Representative images of tubular cross-sections are shown. Red arrows denote SPG. Scale bar, 50 μm. **b** Representative IF images for SAGE1 co-staining with MAGE-A4 (marker for *Hs*SPG), SYCP3 (marker for human spermatocytes) and KI67 in adult human testes (left). The quantification of SAGE1 proportion (right). Scale bar, 50 μm. **c** Representative IF images for SAGE1 co-staining with MAGE-A4 (marker for *Mf*SPG), SYCP3 (marker for macaque spermatocytes) and KI67 in adult macaque testes (left). The quantification of SAGE1 proportion (right). Scale bar, 50 μm. $n = 100$ seminiferous tubules from two macaque testes. For all experiments, data were shown as mean ± SD. Significance was determined with two-tailed Student's t-test. ***$P < 0.001$.

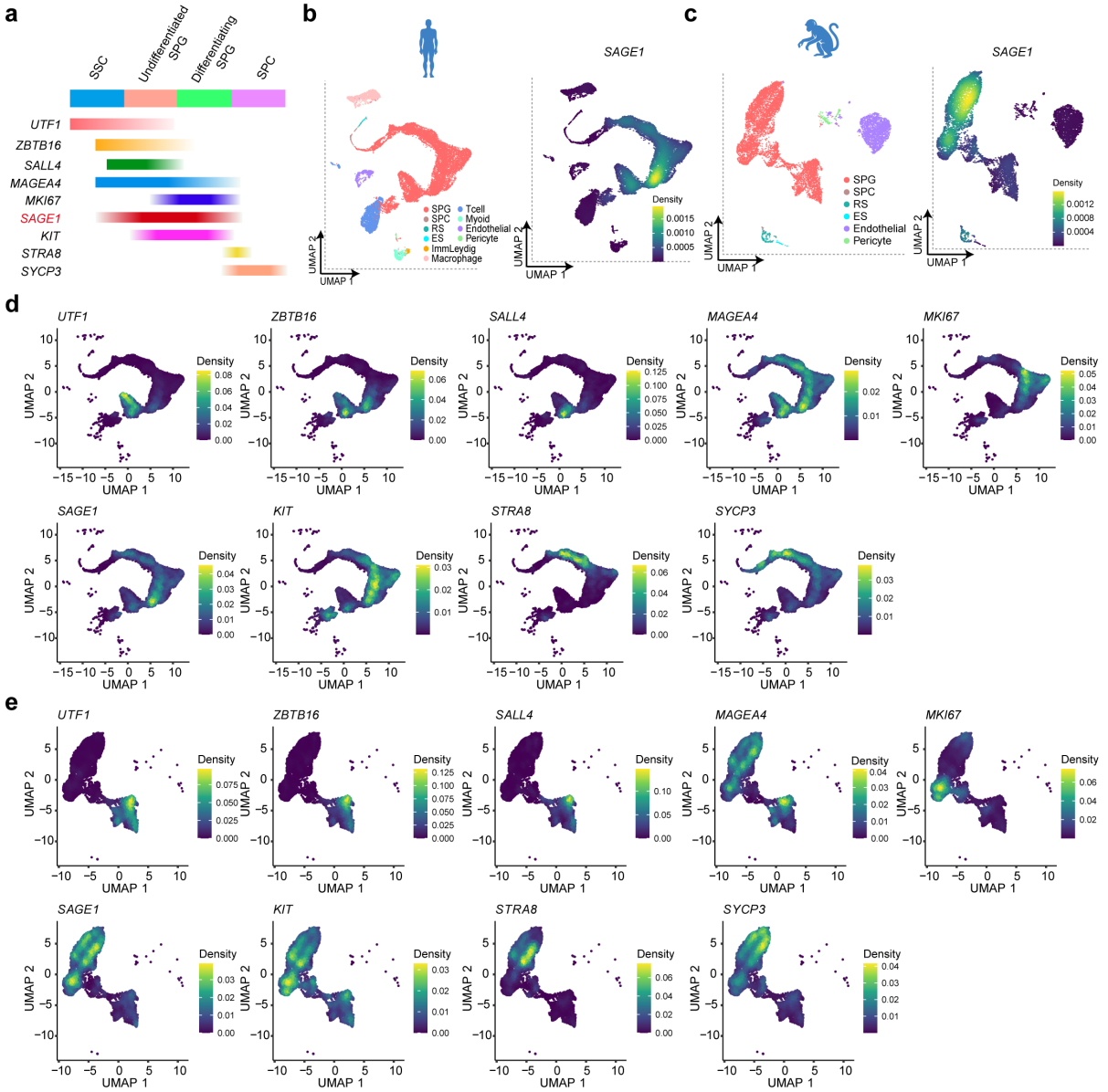

Fig. S4. Stage-specific expression of *SAGE1* during spermatogonial development. a Gene expression dynamics of key markers during human spermatogenesis. **b,c** scRNA-seq analysis of *SAGE1* expression in human and macaque SPG: UMAP representations showing testicular cell types (left panels) and the gene-weighted density of *SAGE1* expression (right panels) in sorted human (**b**) and crab-eating macaque (**c**) testicular cells. **d,e** UMAP representations of the gene-weighted density of *SAGE1* and established markers for spermatogonia (SPG) and spermatocytes (SPC) in human (**d**) and macaque (**e**) datasets.

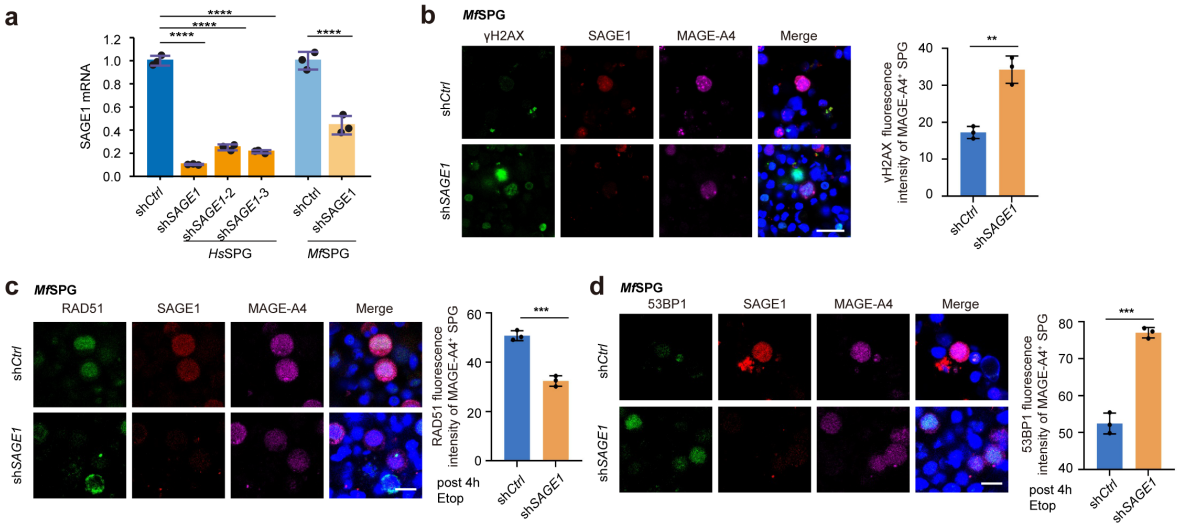

Fig. S5. SAGE1 promotes primate spermatogonia genome stability. a Quantitative RT-PCR confirmed SAGE1 knockdown efficiency by shRNA in *Hs*SPG and *Mf*SPG. **b** Representative IF images of sh*Ctrl* and sh*SAGE1 Mf*SPG stained with antibodies against MAGE-A4, SAGE1 and γH2AX, with the quantification shown on the right. Scale bar, 20 μm. $n = 3$ independent replicates with at least 100 cells per condition were analyzed. **c,d** Representative IF images of sh*Ctrl* and sh*SAGE1 Mf*SPG stained with antibodies against MAGE-A4, SAGE1 and RAD51(**c**) / 53BP1(**d**) after 4 h etoposide (10 μM) treatment, with the quantification shown on the right. Scale bar, 10 μm. $n = 3$ independent replicates with at least 100 cells per condition were analyzed. For all experiments, data were shown as mean ± SD. Significance was determined with two-tailed Student's t-test. $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.

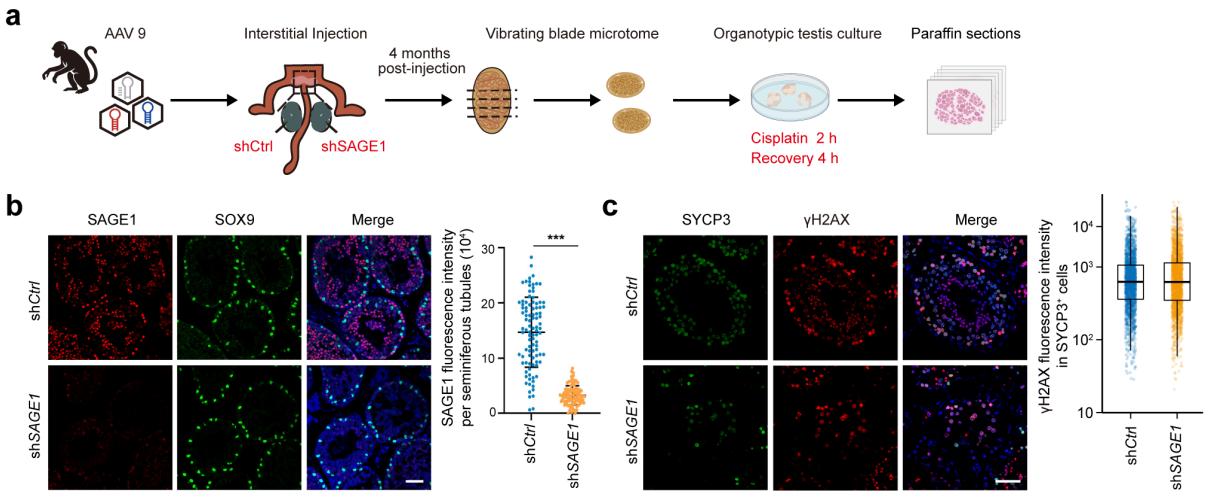

Fig. S6. SAGE1 promotes macaque spermatogonia DSB repair *in vivo*. **a** Schematic representation of the experimental workflow for the SAGE1 knockdown macaque testis model. AAV vectors expressing shRNA are used to knock down SAGE1 expression in macaque testis, followed by *ex vivo* culture, cisplatin treatment, and sectioning of the tissue for further analysis. **b** Immunostaining of SAGE1 and SOX9 (marker for macaque sertoli) in control and SAGE1 knockdown macaque testes 4 months after AAV9 injected. In sh*SAGE1* testes, SAGE1 signal is completely absent in germ cells. The quantification is shown on the right. Scale bar, 50 μm. **c** Representative IF images of the testes with γH2AX and SYCP3 staining in sh*Ctrl* and sh*SAGE1* macaque at 4 h after 2 h cisplatin treatment, with the quantification shown on the right. Scale bar, 50 μm. For all experiments, data were shown as mean ± SD from two macaque each group. Significance was determined with two-tailed Student's t-test. ***$P < 0.001$.

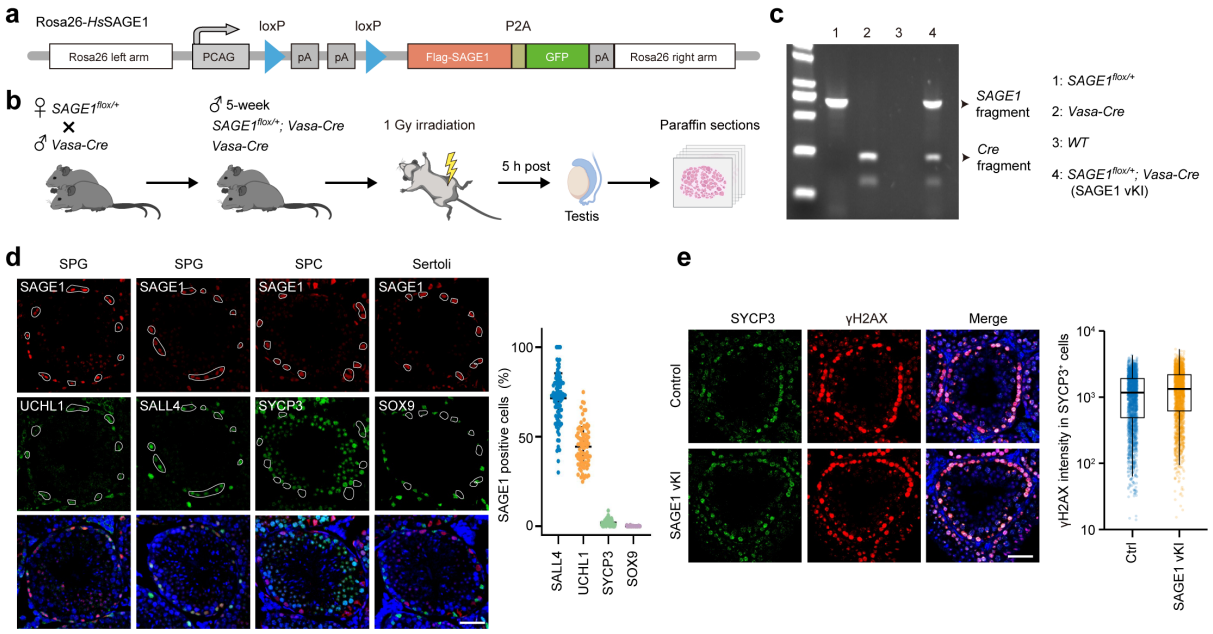

Fig. S7. Ectopically expressed human SAGE1 promotes mouse spermatogonia DSB repair *in vivo*. **a** Schematic diagram of the *Rosa26-SAGE1* mouse generation. The donor vector consists of a Flag-tagged human SAGE1, linked via a self-cleaving P2A peptide to EGFP. Expression is driven by the ubiquitous CAG promoter and is controlled by a loxP-stop-loxP cassette, allowing inducible expression of SAGE1 upon Cre recombinase activation. **b** Schematic diagram of germ cell specific *SAGE1* conditional knock-in mouse mating strategies and the following ionizing radiation induced DDR analysis. *Vasa-Cre* mouse line: Cre recombinase expression is driven by the *Vasa* (*Ddx4*) gene promoter, specifically in germ cells. **c** Diagram of genotype identification. PCR analysis was performed to identify the genotype of the mice, targeting the exogenous gene *SAGE1* and the *Cre* recombinase gene. PCR results with the corresponding genotype are shown. **d** Immunostaining of SGAE1 and Sall4, Sycp3, Sox9 in SAGE1 vKI testes at age of 5 weeks showing SPG-specific SAGE1 signals (Sall4: markers for mouse SPG; Sycp3: marker for mouse spermatocytes; Sox9: marker for mouse Sertoli). The quantification is shown on the right. Scale bar, 50 μm. **e** Representative IF images of the testis with γH2AX and Sycp3 staining in Control and SAGE1 vKI mouse at 5 h after 1

Gy irradiation, with the quantification showed on the right. Scale bar, 50 μm. For all experiments, data were shown as mean ± SD from at least three mice each group. Significance was determined with two-tailed Student's t-test.

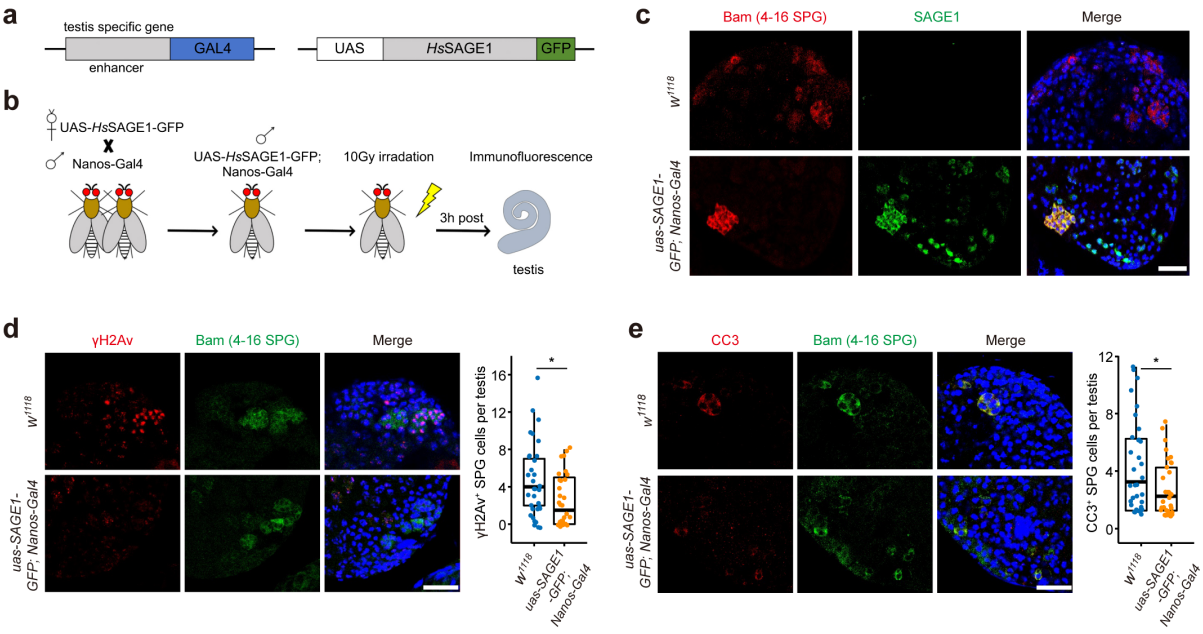

Fig. S8. Ectopically expressed human SAGE1 promotes *Drosophila* spermatogonia DSB repair *in vivo*. **a** Schematic diagram of the *Hs*SAGE1 *Drosophila* generation. **b** Schematic diagram of germ cell specific SAGE1 conditional knock-in *Drosophila* mating strategies and the following ionizing radiation induced DDR analysis. The expression of a SAGE1::GFP fusion protein was ectopically induced in *Drosophila* SPG with Nanos-GAL4. **c** IF images of the testis apical tip with Bam (4-16 SPG marker) and SAGE1 staining in w^{1118} and *UAS-SAGE1-GFP; Nanos-Gal4* files. Scale bar, 20 μm. **d** Representative IF images of the testis apical tip with γH2Av and Bam staining in w^{1118} and *UAS-SAGE1-GFP;Nanos-Gal4* files at 3 h after 10 Gy irradiation. The quantification is shown on the right. Scale bar, 20 μm. **e** Representative IF images of the testis apical tip with CC3 and Bam staining in w^{1118} and *UAS-SAGE1-GFP;Nanos-Gal4* files at 6 h after 10 Gy irradiation. The quantification is shown on the right. Scale bar, 20 μm. For all experiments, data were shown as mean ± SD. Significance was determined with two-tailed Student's t-test. *$P < 0.05$.

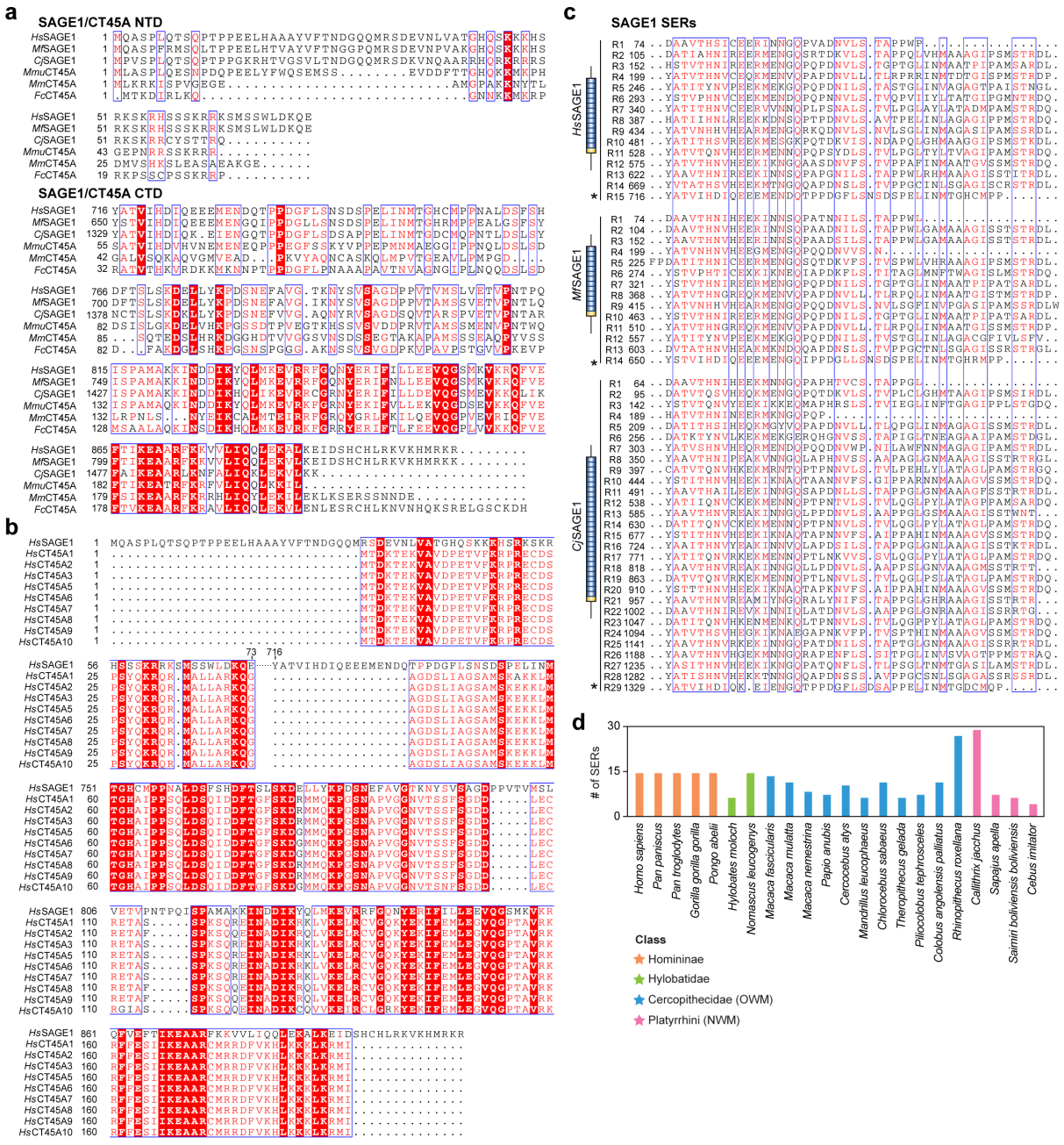

Fig. S9. Sequence analysis of SAGE1. a Amino acid sequence alignment of the NTDs and CTDs of SAGE1 or CT45A from various mammalian species: human (*Homo sapiens*) (*Hs*SAGE), crab-eating macaque (*Macaca fascicularis*) (*Mf*SAGE), marmoset (*Callithrix jacchus*) (*Cj*SAGE1), mouse lemur (*Microcebus murinus*) (*Mmu*CT45A), mouse (*Mus musculus*) (*Mm*CT45A), and cat (*Felis catus*) (*Fc*CT45A). **b** Amino acid sequence alignment of the NTD and CTD of human SAGE1 with the full-length sequences of human CT45A1-

CT45A10. **c** Alignment of SAGE1 SERs of *Hs*SAGE, *Mf*SAGE, and *Cj*SAGE1. The C-terminal SER repeat that lacks six amino acids is denoted by an asterisk (*). **d** The number of SERs in SAGE1 gene across various simian species.

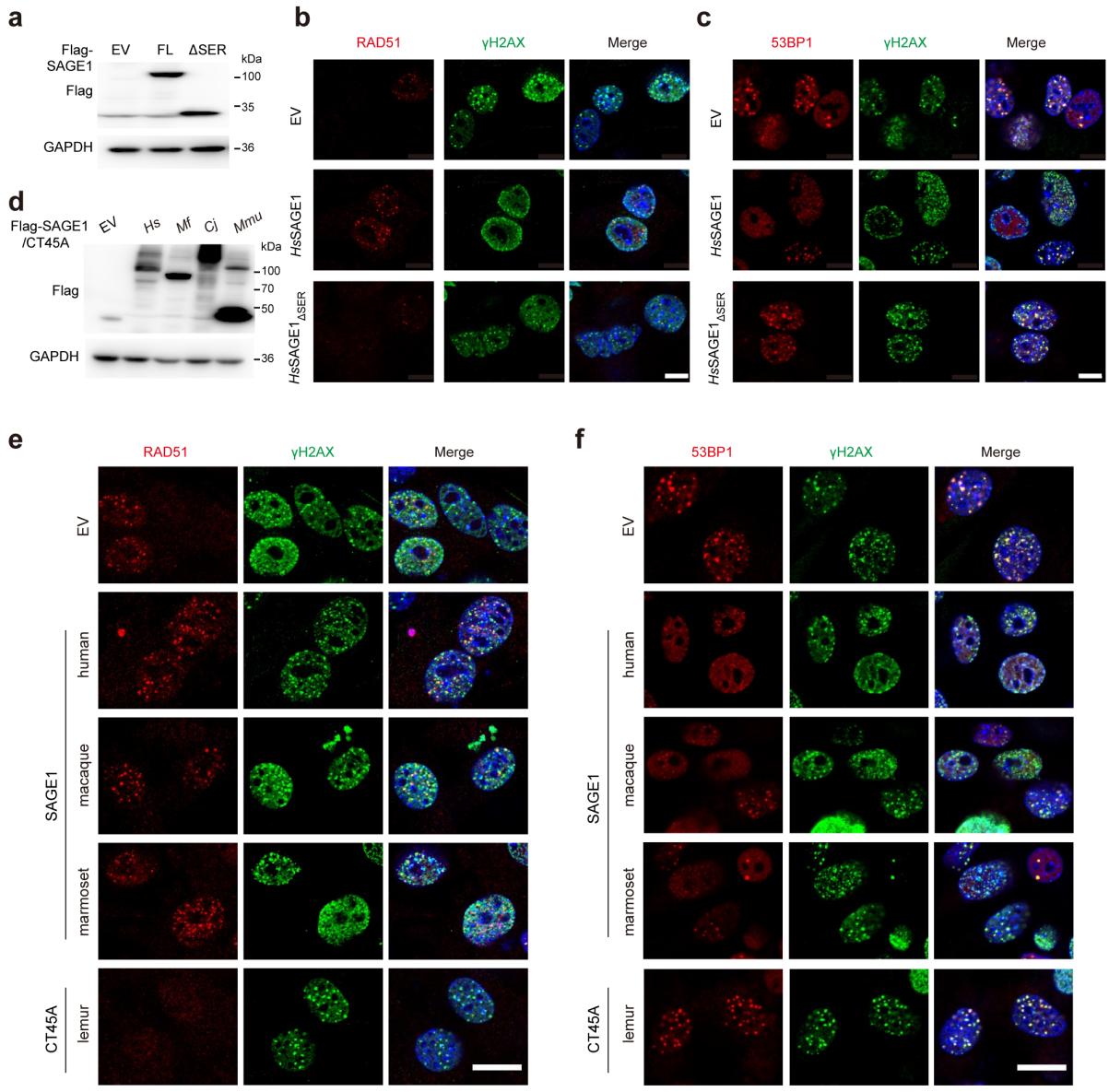

Fig. S10. SAGE1 tandem SER repeats promote genome stability. a Ectopic expression of Flag-EV (empty vector), Flag-*Hs*SAGE1, or Flag-*Hs*SAGE1$_{\Delta SER}$ in *Mm*SPG was validated by WB. **b** Representative IF images of RAD51 and γH2AX foci in *Mm*SPG expressing EV, *Hs*SAGE1, or *Hs*SAGE1$_{\Delta SER}$ after etoposide treatment. Scale bar, 5 μm. **c** Representative IF images of 53BP1 and γH2AX foci in *Mm*SPG expressing EV, *Hs*SAGE1, or *Hs*SAGE1$_{\Delta SER}$ after etoposide treatment. Scale bar, 5 μm. **d** Ectopic expression of Flag-tagged EV, SAGE1 or CT45A from different species (*Hs*SAGE1, *Mf*SAGE1, *Cj*SAGE1, or *Mmu*CT45A) in MEF cells was validated by WB. **e** Representative IF images of RAD51 and γH2AX foci in MEF

expressing EV, SAGE1 or CT45A from different species after etoposide treatment. Scale bar, 20 μm. **f** Representative IF images of 53BP1 and γH2AX foci in MEF expressing EV, SAGE1 or CT45A from different species after etoposide treatment. Scale bar, 20 μm.

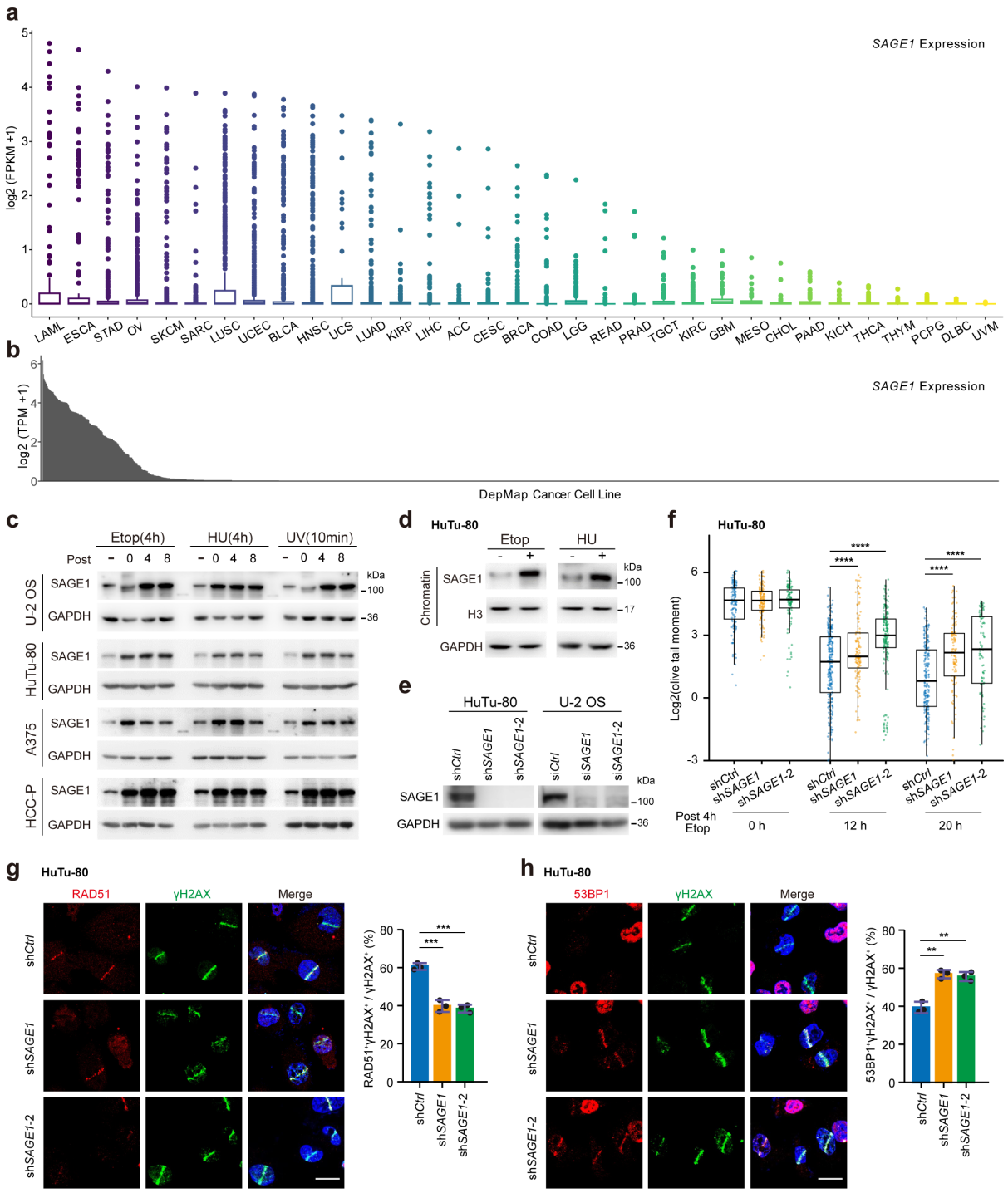

Fig. S11. SAGE1 promotes cancer cell genome stability. a Expression profiles of *SAGE1* in major TCGA cancer types. **b** Expression profiles of *SAGE1* in DepMap cancer cell lines. **c** Human cancer cells (U-2 OS, HuTu-80, A375, and HCC-P) were treated with HU (2 mM), etoposide (10 μM), or UV light (130 mJ) for indicated periods. SAGE1 expression was

analyzed by WB. **d** HuTu-80 cells were treated with etoposide or HU for 4 h. Chromatin-bound proteins were fractionated and immunoblotted with anti-SAGE1 antibody. **e** WB confirmed efficient SAGE1 knockdown in HuTu-80 and U-2 OS cells by shRNA or siRNA. **f** sh*Ctrl* and sh*SAGE1* HuTu-80 cells were treated with etoposide for indicated periods. The relative comet tail moment was quantified by the neutral comet assay. At least 100 cells were analyzed. **g** Representative IF images of laser-microirradiated sh*Ctrl* and sh*SAGE1* HuTu-80 cells stained with antibodies against RAD51 and γH2AX respectively, with the quantification shown on the right. Scale bar, 20 μm. $n = 3$ independent replicates with at least 100 cells per condition were analyzed. **h** Representative IF images of laser-microirradiated sh*Ctrl* and sh*SAGE1* HuTu-80 cells stained with antibodies against 53BP1 and γH2AX respectively, with the quantification shown on the right. Scale bar, 20 μm. $n = 3$ independent replicates with at least 100 cells per condition were analyzed. For all experiments, data were shown as mean ± SD. Significance was determined with two-tailed Student's t-test. $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.

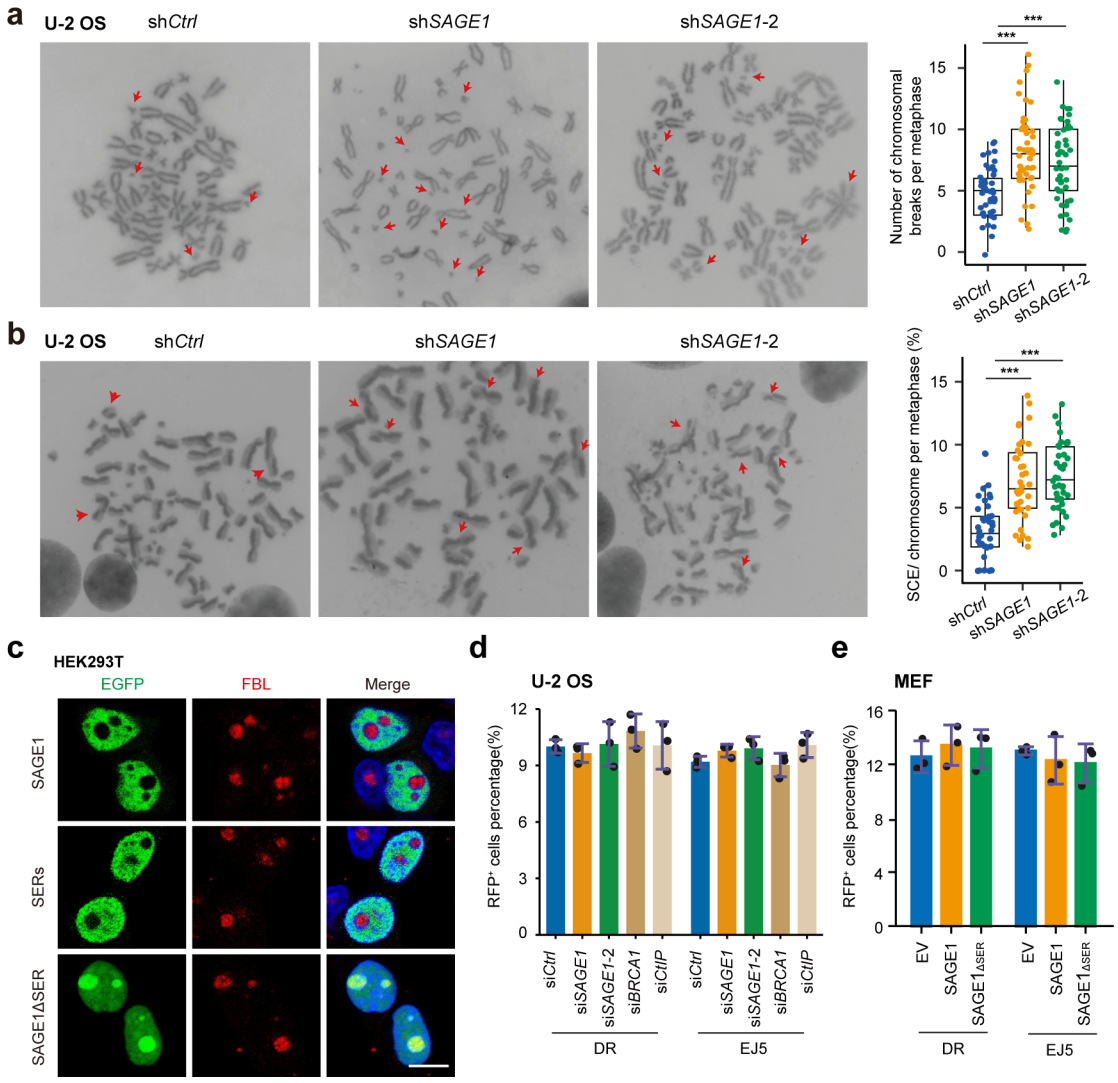

Fig. S12. SAGE1 promotes cancer cell genome stability. a Quantification of chromosome abnormality in sh*Ctrl* and sh*SAGE1* U-2 OS cells, $n = 40$ cells per condition were analyzed. **b** Quantification of sister chromatid exchange (SCE) rates in sh*Ctrl* and sh*SAGE1* U-2 OS cells, $n = 40$ cells per condition were analyzed. **c** IF analysis of EGFP and nucleolus marker fibrillarin (FBL) co-localization in HEK293T cells expressing EGFP-tagged SAGE1, $SAGE1_{SER}$ or $SAGE1_{\Delta SER}$. Scale bar, 10 μm. **d,e** Plasmids expressing RFP were co-transfected into cells with other constructions as internal controls to demonstrate the identical transfection efficiency. For all experiments, data were shown as mean ± SD. Significance was determined with a two-tailed Student's t-test. ***$P < 0.001$.

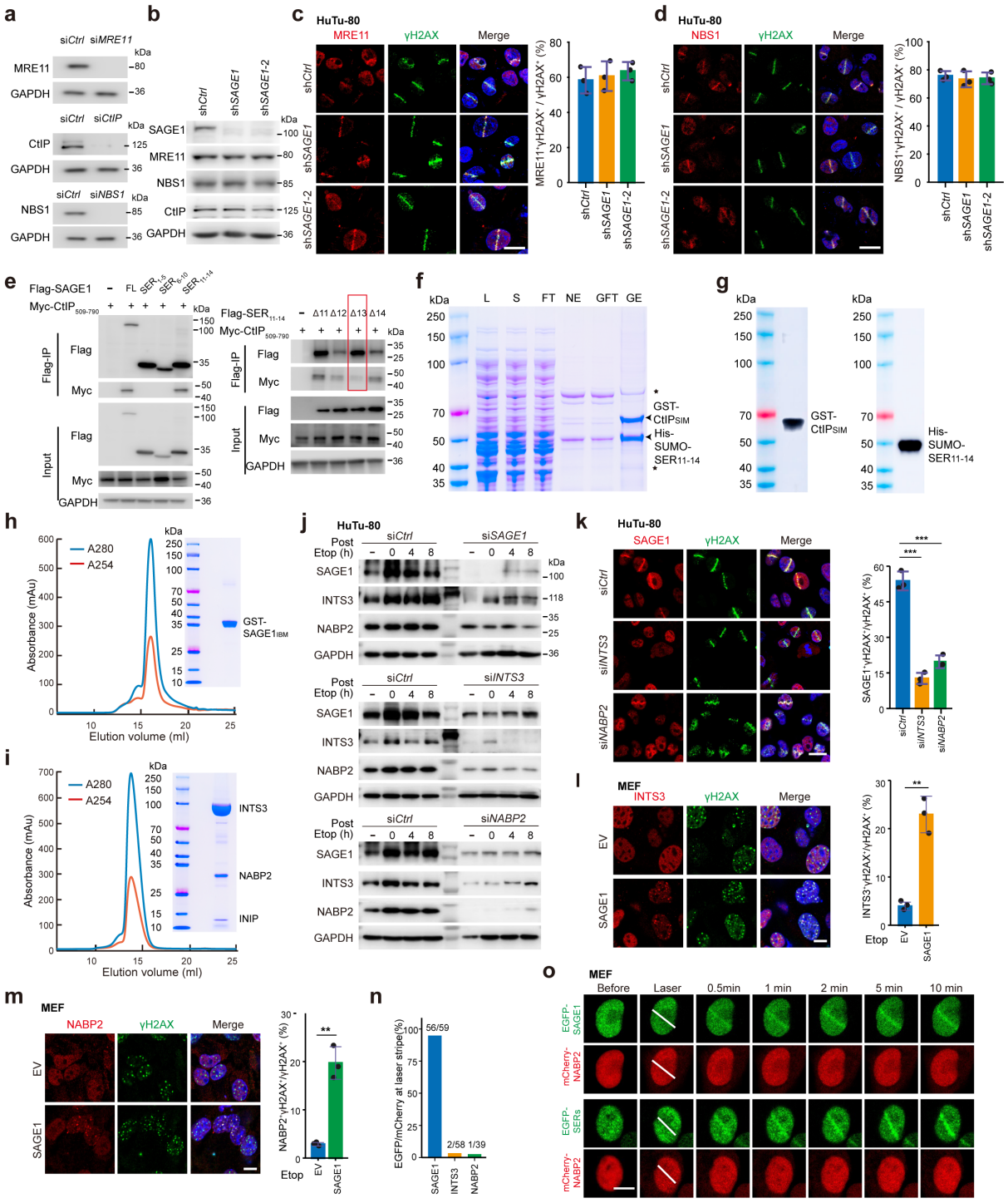

Fig. S13. Tandem SER repeats mediate the SAGE1 interaction with CtIP and SAGE1 recruits the SOSS complex to DSBs. a WB analysis showed the knockdown effects of MRE11, CtIP, and NBS1 by siRNA in HuTu-80 cells. **b** Protein levels of MRE11, CtIP, and NBS1 in sh*Ctrl* and sh*SAGE1* HuTu-80 cells were analyzed by WB. **c** Representative IF images of

MRE11 and γH2AX signals in laser-microirradiated sh*Ctrl* and sh*SAGE1* HuTu-80 cells, with co-localization quantified on the right. Scale bar, 20 μm. *n* = 3 independent replicates with at least 100 cells per condition were analyzed. **d** Representative IF images of NBS1 and γH2AX in laser-microirradiated sh*Ctrl* and sh*SAGE1* HuTu-80 cells, with co-localization quantified on the right. Scale bar, 20 μm. *n* = 3 independent replicates with at least 100 cells per condition were analyzed. **e** Co-IP experiments in HEK293T cells co-transfected with plasmids expressing Myc-CtIP$_{509\text{-}790}$ and various Flag-SAGE1$_{SER}$ truncations. FL: full-length. **f,g** Two-step affinity purification of the SAGE1$_{SER11\text{-}14}$-CtIP$_{SIM}$ complex. Coomassie-stained SDS-PAGE (**f**) and WB (**g**) of the purified SAGE1$_{SER11\text{-}14}$-CtIP$_{SIM}$ complex. Corresponding bands are labeled, and contaminant proteins are denoted by stars. L: lysate, S: supernatant, FT: flow-through, NE: nickel-bead eluant, GFT: flow-through after GST beads binding. GE: GST-bead eluant. **h** Purification of GST-SAGE1$_{IBM}$. Size exclusion chromatographic profile of GST-SAGE1$_{IBM}$ (left) and coomassie-stained SDS-PAGE of the corresponding peak in the profile (right). **i** Purification of the SOSS complex. Size exclusion chromatographic profile of the SOSS complex (left) and coomassie-stained SDS-PAGE of the corresponding peaks in the profile (right). **j** SAGE1, INTS3, and NABP2 were analyzed by WB in HuTu-80 cells transfected with indicated siRNAs, after etoposide treatment for indicated periods. **k** Representative IF images of SAGE1 and γH2AX signals in laser-microirradiated si*Ctrl*, si*NABP2*, and si*INTS3* HuTu-80 cells, with co-localization quantified on the right. Scale bar, 20 μm. *n* = 3 independent replicates with at least 100 cells per condition were analyzed. **l** Representative IF images of INTS3 and γH2AX foci in MEF expressing EV or SAGE1 after etoposide treatment, with the quantification shown on the right. Scale bar, 10 μm. *n* = 3 independent replicates with at least 100 cells per condition were analyzed. **m** Representative IF images of NABP2 and γH2AX foci in MEF expressing EV or SAGE1 after etoposide treatment, with the quantification shown on the right. Scale bar, 10 μm. *n* = 3 independent replicates with at least 100 cells per condition

were analyzed. **n** The percentage of EGFP-SAGE1, mCherry-INTS3, or mCherry-NABP2 signals at laser stripes in MEF expressing EGFP-SAGE1, mCherry-INTS3, or mCherry-NABP2 respectively, with the number of examined cells indicated in the figure. **o** MEFs were co-transfected with mCherry-NABP2 and EGFP-SAGE1 or EGFP-SAGE1$_{SERs}$, and subjected to laser stripping, and then the real-time EGFP recruitment was monitored by confocal microscopy. White lines indicate laser paths. Scale bar, 10 μm. For all experiments, data were shown as mean ± SD. Statistical significance was determined with two-tailed Student's t-test. $^{**}P < 0.01$, $^{***}P < 0.001$.

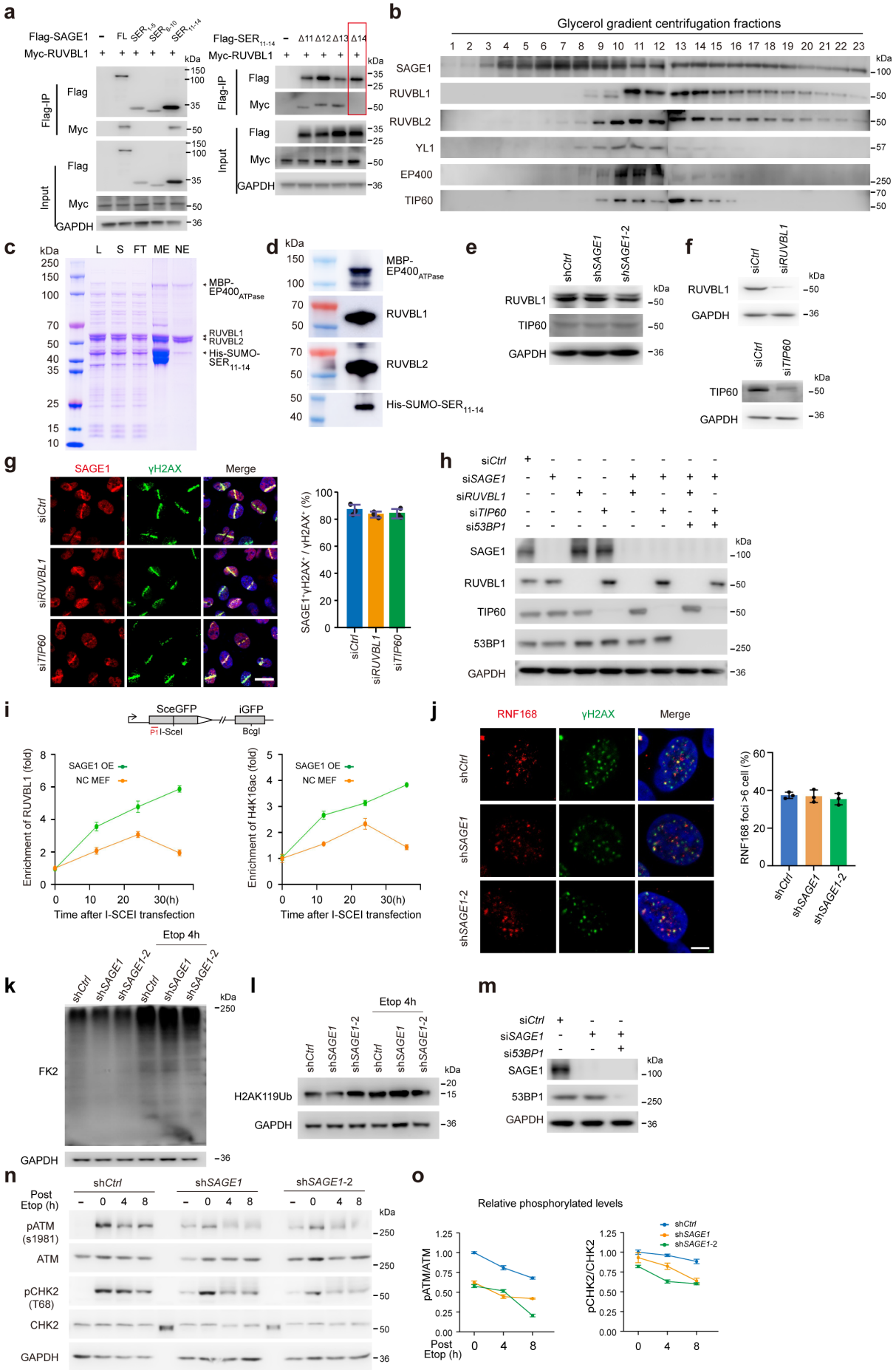
a
Flag-SAGE1
Myc-RUVBL1
FL
Flag-SER11-14
Δ11 Δ12 Δ13 Δ14
Flag-IP
Input
Flag
Myc
GAPDH
kDa
b
Glycerol gradient centrifugation fractions
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
SAGE1
RUVBL1
RUVBL2
YL1
EP400
TIP60
c
L S FT ME NE
MBP-EP400ATPase
RUVBL1
RUVBL2
His-SUMO-SER11-14
d
e
shCtrl
shSAGE1
shSAGE1-2
f
siCtrl
siRUVBL1
siTIP60
g
SAGE1
γH2AX
Merge
siCtrl
siRUVBL1
siTIP60
SAGE1+γH2AX+ / γH2AX+ (%)
h
siCtrl
siSAGE1
siRUVBL1
siTIP60
si53BP1
SAGE1
RUVBL1
TIP60
53BP1
GAPDH
i
SceGFP
iGFP
PI I-SceI
BcgI
SAGE1 OE
NC MEF
Enrichment of RUVBL1 (fold)
Enrichment of H4K16ac (fold)
Time after I-SCEI transfection
j
RNF168
γH2AX
Merge
shCtrl
shSAGE1
shSAGE1-2
RNF168 foci >6 cell (%)
k
Etop 4h
FK2
GAPDH
l
H2AK119Ub
GAPDH
m
siCtrl
siSAGE1
si53BP1
SAGE1
53BP1
GAPDH
n
Post Etop (h)
pATM (s1981)
ATM
pCHK2 (T68)
CHK2
GAPDH
o
Relative phosphorylated levels
pATM/ATM
pCHK2/CHK2
Post Etop (h)

Fig. S14. SAGE1 SER repeats recruit the NuA4 complex to DSBs and activate ATM. a Co-IP experiments in HEK293T cells co-transfected with plasmids expressing Myc-RUVBL1 and various Flag-SAGE1$_{SER}$ truncations. FL: full-length. **b** *In vivo* NuA4-SAGE1 association was analyzed by glycerol density gradient ultracentrifugation. SAGE1 and NuA4 subunits were detected with indicated antibodies. **c** Two-step affinity purification of the SAGE1$_{SER11-14}$-NuA4 submodule complex. Coomassie-stained SDS-PAGE of the purified SAGE1$_{SER11-14}$ and NuA4 submodule components. Corresponding bands are labeled. L: lysate, S: supernatant, FT: flowthrough, ME: MBP-bead eluant, NE: nickel-bead eluant. **d** Purified SAGE1$_{SER11-14}$-NuA4 submodule complex was confirmed by WB. **e** Protein levels of RUVBL1 and TIP60 in sh*Ctrl* and sh*SAGE1* HuTu-80 were analyzed by WB. **f** WB confirmed RUVBL1 and TIP60 knockdown efficiencies in HuTu-80 cells by siRNA. **g** Representative IF images of SAGE1 and γH2AX in laser-microirradiated si*Ctrl*, si*RUVBL1*, and si*TIP60* HuTu-80 cells, with co-localization quantified on the right. Scale bar, 20 μm. $n = 3$ independent replicates with at least 100 cells per condition were analyzed. **h** WB confirmed protein knockdown efficiencies in HuTu-80 cells by siRNAs. **i** DR-GFP plasmid and I-SceI plasmid were co-transfected in MEF cells ectopically expressing EV and HsSAGE1. ChIP-qPCR performed with RUVBL1 and H4K16ac antibodies in the presence of I-SceI induced DSBs in control and SAGE1-overexpressing MEFs. **j** IF analysis of RNF168 and γH2AX co-localization at DSB sites in sh*Ctrl*, sh*SAGE1*, sh*SAGE1*-2 U2OS cells, with the quantification shown on the right. Scale bar, 5 μm. $n = 3$ independent replicates with at least 100 cells per condition were analyzed. **k** WB analysis of total ubiquitin level in sh*Ctrl* and sh*SAGE1* U2OS cells were treated with 10 μM etoposide for 4 h. **l** WB analysis of H2AK119 ubiquitin level in sh*Ctrl* and sh*SAGE1* U2OS cells were treated with 10 μM etoposide for 4 h. **m** WB confirmed SAGE1 and 53BP1 knockdown efficiencies in HuTu-80 cells by siRNAs. **n** WB analysis of ATM-CHK2 activation. sh*Ctrl* and sh*SAGE1* HuTu-80 were treated with etoposide for 4 h and released for indicated

time periods. Total cell lysates were analyzed by immunoblotting with indicated antibodies. **o** Quantification of phosphorylated protein levels shown in panel (**n**). For all experiments, data were shown as mean ± SD. Significance was determined with two-tailed Student's t-test.

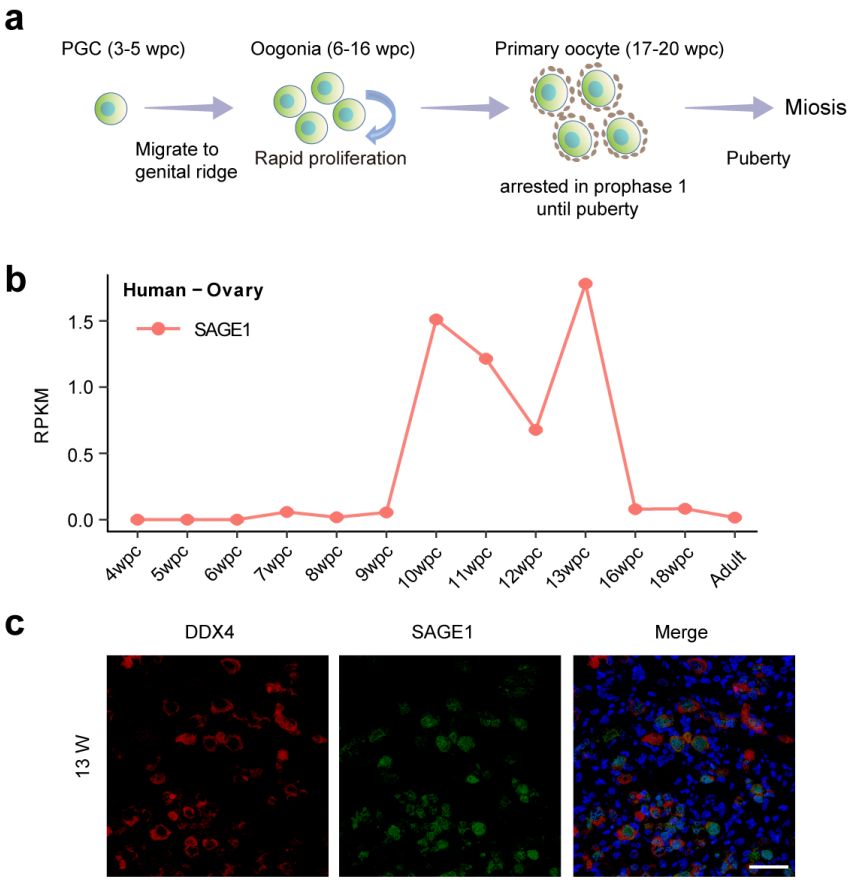

Fig. S15. SAGE1 is expressed in human germline precursors in fetal ovary. a Schematic diagram of human female germ cell development[64]. **b** Transcriptomic expression profiles of the human ovary from 4 weeks post-conception (WPC) to adulthood[65]. **c** IF co-staining of SAGE1 and germ cell marker DDX4 in 13-WPC fetal ovary. Scale bar, 50 μm.

Supplementary Data S1. Amino acid sequences of SAGE1 /CT45A across different species.

See attached Doc file

Supplementary Table S2. IP-MASS data information.

See attached Excel file

Supplementary Table S3. Evolutionary analysis of proteins involved in DSB repair.

See attached Excel file

Supplementary Table S4. Oligonucleotide templates, real-time quantitative PCR primers, and PCR primers used in this study.

Gene	Oligonucleotides	
Mouse genotype PCR primers		
SAGE1-KI	Forward	5'-AGCATACGTGTTCACCAACGAC-3'
	Reverse	5'-ATGTGATGGTGGCGTAAGGATC-3'
Vasa-cre	Forward	5'-CACGTGCAGCCGTTTAAGCCGCGT-3'
	Reverse	5'-TTCCCATTCTAAACAACACCCTGAA-3'
shRNA sequences		
sh*Ctrl*	Forward	5'-ccggCAACAAGATGAAGAGCACCAActcgagTTGGTGCTCTTCATCTTG TTGttttttg-3'
	Reverse	5'-aattcaaaaaaCAACAAGATGAAGAGCACCAActcgagTTGGTGCTCTTCAT CTTGTTG-3'
Human-sh*SAGE1*	Forward	5'-ccggCCAAGGAGAAACAAGGACATActcgagTATGTCCTTGTTTCTCCT TGGttttttg -3'
	Reverse	5′-aattcaaaaaaCCAAGGAGAAACAAGGACATActcgagTATGTCCTTGTTTC TCCTTGG-3'
Human-sh*SAGE1*-2	Forward	5'-ccggCCCAAACTGATAAGGTCATATctcgagATATGACCTTATCAGTTTGG Gttttttg -3'
	Reverse	5′-aattcaaaaaaCCCAAACTGATAAGGTCATATctcgagATATGACCTTATCAGT TTGGG -3'
Human-sh*SAGE1*-3	Forward	5'-ccggCTGATAAGGTCATATCAAATGctcgagCATTTGATATGACCTTAT CAGttttttg -3'
	Reverse	5'-aattcaaaaaaCTGATAAGGTCATATCAAATGctcgagCATTTGATATGACC TTATCAG-3'
Macaque-sh*SAGE1*	Forward	5′-ccggCCAAGGAGAAACAAGAACATActcgagTATGTTCTTGTTTCTCCT TGGttttttg-3'
	Reverse	5′-aattcaaaaaaCCAAGGAGAAACAAGAACATActcgagTATGTTCTTGTTTC TCCTTGG-3'
siRNA sequences		
si*HsSAGE1*-1	Forward	5′-CUGUUCAACCAGUGAUUAUdTdT-3'
	Reverse	5′-AUAAUCACUGGUUGAACAGdTdT-3'
si*HsSAGE1*-2	Forward	5′-CUGAUAAGGUCAUAUCAAAdTdT-3'
	Reverse	5′-UUUGAUAUGACCUUAUCAGdTdT-3'
si*MRE11*-1	Forward	5′-GGAGGUACGUCGUUUCAGAdTdT-3'

	Reverse	5′-UCUGAAACGACGUACCUCCdTdT-3'
si*MRE11*-2	Forward	5′-UGGCUAAUGACUCUGAUGAdTdT-3'
	Reverse	5′-UCAUCAGAGUCAUUAGCCAdTdT-3'
si*NBS1*-1	Forward	5′-CCAACUAAAUUGCCAAGUAdTdT-3'
	Reverse	5′-UACUUGGCAAUUUAGUUGGdTdT-3'
si*NBS1*-2	Forward	5′-CAUACGUAGCUGACACAGAdTdT-3'
	Reverse	5′-UCUGUGUCAGCUACGUAUGdTdT-3'
si*CtIP*-1	Forward	5′-UCCACAACAUAAUCCUAAUdTdT-3'
	Reverse	5′-AUUAGGAUUAUGUUGUGGAdTdT-3'
si*CtIP*-2	Forward	5′-GCUAAAACAGGAACGAAUCdTdT-3'
	Reverse	5′-GAUUCGUUCCUGUUUUAGCdTdT-3'
si*INTS3*	Forward	5′-CGUGAUGGCAUGAAUAUUGdTdT-3'
	Reverse	5′-CAAUAUUCAUGCCAUCACGdTdT-3'
si*NABP2*	Forward	5′-CGACGGA GACCUUUGUGAAdTdT-3'
	Reverse	5′-UUCACAAAGGUCUCCGUCGdTdT-3'
si*RUVBL1*	Forward	5′-GUUUACUCAACUGAGAUCAdTdT-3'
	Reverse	5′-UGAUCUCAGUUGAGUAAACdTdT-3'
si*BRCA1*	Forward	5′-GGAACCUGUCUCCACAAAGdTdT-3'
	Reverse	5′-CUUUGUGGAGACAGGUUCCdTdT-3'
si*53BP1*	Forward	5′-GAAGGACGGAGUACUAAUAdTdT-3'
	Reverse	5′-UAUUAGUACUCCGUCCUUCdTdT-3'
si*TIP60*	Forward	5′-CCUCAAUCUCAUCAACUACdTdT-3'
	Reverse	5′-GUAGUUGAUGAGAUUGAGGdTdT-3'
RT-PCR primers		

DR-Primer-1	Forward	5′-TCCCCTTCTCCATCTCCAG-3'
	Reverse	5′-GGCATGAACATGGTTAGCAG-3'
DR-Primer-2	Forward	5′-TGAGAATGAAGTGTATGTGGAACAG-3'
	Reverse	5′-TGTGGTTTCCAAATGTGTCAG-3'
EJ5-Primer-1	Forward	5′-ATGCCCTGGCTCACAAATAC-3'
	Reverse	5′-CCCATATGTCCTTCCGAGTG-3'
EJ5-Primer-2	Forward	5′-CTGGTAGGGACCCGAGTGTA-3'
	Reverse	5′-CACTTGGCCCTTGATGATCT-3'
*Hs*SAGE1	Forward	5′- ACTTCAAACGAGTCAACCAACT-3'
	Reverse	5′- TCTAACCACGAGGACATACTCTT-3'
*Mf*SAGE1	Forward	5′- TCCACCAGGGTGTACTAATC-3'
	Reverse	5′- CTGGATATCGTGAATGACAG-3'
*Hs*ACTB	Forward	5′- GCACAGAGCCTCGCCTT-3'
	Reverse	5′- GTTGTCGACGACGAGCG-3'
Vector construction primers		
*Hs*SAGE1-puast-attb-GFP	Forward	5′- CCGGAATTCCGGATGCAGGCATCCCCTCTGCAG-3'
	Reverse	5′- CGGGGTACCCCGCCTCTTCCTCATGTGCTTCAC-3'
Flag-*Hs*SAGE1-PBTRE	Forward	5′- GCCGAATTCGCCACCATGGACTACAAG-3'
	Reverse	5′- CCGACGCGTTTACCTCTTCCTCATGTGCTTCACC-3'
Flag-*Mf*SAGE1-PBTRE	Forward	5′- GCCGAATTCGCCACCATGGACTACAAG-3'
	Reverse	5′-CCGACGCGTTTATTTTTTTCTCATGTGCTTAACTTTTCTG-3'
Flag-*Cj*SAGE1-PBTRE	Forward	5′- GCCGAATTCGCCACCATGGACTACAAG-3'
	Reverse	5′-CCGACGCGTCTACTTTTTTAGCACCTTCTCGAGTTTC-3′
Flag-	Forward	5′-GCCGAATTCGCCACCATGGACTACAAG-3′

*Mmu*CT45A-PBTRE	Reverse	5′-CCGACGCGTTTAAAGGATCTTCTTAAGTTGCTGTATTAAAAC-3′
SAGE1-plvx-EGFP-NLS	Forward	5′-CCGCTCGAGGCCACCATGCAGGCATCCCCTCTGCAGAC-3′
	Reverse	5′-CGCGGATCCCCTCTTCCTCATGTGCTTCACC-3′
SER-plvx-EGFP-NLS	Forward	5′-CCGCTCGAGGCCACCATGGCCGTGACCCACTCTATCTGCG-3′
	Reverse	5′-CGCGGATCCGTACAGGTCTCTGGTGCTCCGACAAG-3′
△SER-plvx-EGFP-NLS	Forward	5′-GGCTGGACAAGCAGGAGGATGCCACAGTGATCCACGATATCC-3′
	Reverse	5′-GGATATCGTGGATCACTGTGGCATCCTCCTGCTTGTCCAGC -3′
INTS3-plvx-mCherry-NLS	Forward	5′-CCGCTCGAGGCCACCATGGAGTTGCAGAAGGGAAAAGGG-3′
	Reverse	5′-CGCGGATCCAACTTTGCGTTTCTTTTTCGGAGATCCGGAACTGTCA CTGTCAGAGCCCACTG-3′
NABP2-plvx-mCherry-NLS	Forward	5′-CTCGTCCTTAGACAGGGAGGTGAAGGGCATCCTCCTGCTTGTCCAG-3′
	Reverse	5′-CGCGGATCCAACTTTGCGTTTCTTTTTCGGAGATCCGGAACTTCTCT TGCTGCTCCTCCGG-3′
Flag-*Hs*SAGE1-PCDH	Forward	5′- AGAGCTAGCGAATTCATGCAGGCATCCCCTCTGCA-3'
	Reverse	5′- GATCCTTGCGGCCGCTCACCTCTTCCTCATGTGCTTCACC-3'
Flag-*Hs*SAGE1$_{\Delta SER}$-PCDH	Forward	5′-GGCTGGACAAGCAGGAGGATGCCACAGTGATCCACGATATCC-3'
	Reverse	5′-GGATATCGTGGATCACTGTGGCATCCTCCTGCTTGTCCAGCC-3'
pEGFP-3Flag-SAGE1	Forward	5′-CCGCTCGAGATGCAGGCATCCCCTCTGCAGAC-3'
	Reverse	5′-CGCGGATCCTTACCTCTTCCTCATGTGCTTCACC-3'
pEGFP-3Flag-SER	Forward	5′-CCGCTCGAGGCCGTGACCCACTCTATCTGCG-3'
	Reverse	5′-CGCGGATCCTTAGTACAGGTCTCTGGTGCTCCGACAAG-3'
pEGFP-3Flag-△SER	Forward	5′-GGCTGGACAAGCAGGAGGATGCCACAGTGATCCACGATATCC-3'
	Reverse	5′-GGATATCGTGGATCACTGTGGCATCCTCCTGCTTGTCCAGCC-3'
pEGFP-3MYC-CtIP	Forward	5′-CGCGTCGACATGAACATCTCGGGAAGCAGC-3'
	Reverse	5′-CGCGGATCCCTATGTCTTCTGCTCCTTGCC-3'
pEGFP-3MYC-	Forward	5′-CGCGTCGACATGAACATCTCGGGAAGCAGC-3'

CtIP(1-145)	Reverse	5′-CGCGGATCCTTAATCATTCTCAATTTTCTGCTGGAGTTG-3'
pEGFP-3MYC-CtIP(145-509)	Forward	5′-CGCGTCGACGATCAACAGCATCAAGCAGCTGAG-3'
	Reverse	5′-CGCGGATCCTTAACTTCCTTGGCTTTTCTCTTGACG-3'
pEGFP-3MYC-CtIP(509-790)	Forward	5′-CGCGTCGACAGTGAGACTTCTAAAAACAAATTTAGGC-3'
	Reverse	5′-CGCGGATCCTTACAAGCTAGTCTCTCTTTCATCACC-3'
pEGFP-3MYC-CtIP(790-897)	Forward	5′-CGCGTCGACTTGCAAAATTTTCCTCATATTGAGGTGG-3'
	Reverse	5′-CGCGGATCCCTATGTCTTCTGCTCCTTGCC-3'
pEGFP-3MYC-RUVBL1	Forward	5′-CCGCTCGAGATGAAGATTGAGGAGGTGAAGAGCAC-3'
	Reverse	5′-CGCGGATCCTCACTTCATGTACTTATCCTGCTGGTCAG-3'
pET28-6His-SUMO-SAGE1(525-716)	Forward	5′-CGCGGATCCATGTCTACCAAGGACCTGTAC-3'
	Reverse	5′-CCGCTCGAGTTAGTACAGGTCTCTGGTGCTCCG-3'
pGEX-6P1-CtIP(509-790)	Forward	5′-CGCGGATCCAGTGAGACTTCTAAAAACAAATTTAGGCAAGTG-3'
	Reverse	5′-CGCGTCGACTTACAAGCTAGTCTCTCTTTCATCACC-3'
pFastBac-GST-INTS3	Forward	5′-CGCGGATCCATGGAGTTGCAGAAGGGAAAAGGG-3'
	Reverse	5′-CCGCTCGAGTCAGTCACTGTCAGAGCCCACTGC-3'
pFastBac-6His-NABP2	Forward	5′-CGCGGATCCATGACGACGGAGACCTTTGTGAAG-3′
	Reverse	5’-CCGCTCGAGCTATCTCTTGCTGCTCCTCCGGGTTT -3′
pFastBac-6His-INIP	Forward	5′-CGCGGATCCATGGCAGCAAACTCTTCAGGAC-3′
	Reverse	5′-CGCGTCGACTCATTCTGGGTCAAGGCGAGG-3′
pETDuet-RUVBL1-RUVBL2	Forward (R1)	5′-CTCGGATCCATGAAGATTGAGGAGGTGAA-3′
	Reverse (R1)	5′-ATTTGCGGCCGCTCACTTCATGTACTTATCCTGCTGGTCAG-3′
	Forward (R2)	5′-CTCGGATCCATGGCAACCGTTACAGCCAC-3′
	Reverse (R2)	5′-CCGCTCGAGTCAGGAGGTGTCCATGGTCTCG-3′

Supplementary Table S5. Antibodies used in this study.

Name	Company	Catalog#	Clone#	Dilution
CtIP Mouse monoclonal antibody	Santa Cruz Biotechnology	sc-271339	D-4	IF 1:50 WB 1:100
NBS1 Rabbit polyclonal antibody	Novus	NB100-143		IF 1:200 WB 1:1000
ATM Mouse monoclonal antibody	Novus	NB100-309	2C1	WB 1:1000
anti-Chk2(phospho T68) Rabbit monoclonal antibody	Abcam	ab32148	Y171	WB 1:1000
anti-Chk1(phospho S345) Rabbit monoclonal antibody	Cell Signaling Technology	2348T	133D3	WB 1:1000
Chk1 Rabbit polyclonal antibody	ABclonal	A7653		WB 1:1000
Bam Mouse monoclonal antibody	DSHB	AB_1057032 7		IF 1:10
Anti-gamma H2Av (phospho S137) Rabbit monoclonal antibody	Rockland	600-401-914		IF 1:100
Cleaved Caspase-3 (Asp175) Rabbit polyclonal antibody	Cell Signaling Technology	9661T		IF 1:100
ATR Rabbit monoclonal antibody	ABclonal	A21253	ARC51749	WB 1:1000
Chk2 Rabbit monoclonal antibody	ABclonal	A19543	ARC57076	WB 1:1000
EP400 Rabbit polyclonal antibody	ABclonal	A12151		WB 1:1000
RUVBL2 Rabbit polyclonal antibody	ABclonal	A1905		WB 1:1000
RUVBL1 Rabbit polyclonal antibody	Novus	NBP2-20245		IF 1:100 WB 1:1000
TIP60 Rabbit polyclonal antibody	Novus	NBP2-20647		IF 1:500 WB 1:1000
Anti-Mre11 Mouse monoclonal antibody-BSA and Azide free	Abcam	ab214	12D7	IF 1:200 WB 1:1000
SAGE1 Rabbit polyclonal antibody	Novus	NBP1-84355		IF 1:300 WB 1:1000
Anti-gamma H2A.X(phospho S139) Rabbit monoclonal antibody	Abcam	ab81299	EP854(2)Y	IF 1:500
Anti-BrdU Rat monoclonal antibody	Abcam	ab6326	BU1/75 (ICR1)	IF 1:1000
Anti-gamma H2A.X(phospho S139) Mouse monoclonal antibody	Abcam	ab22551	3F2	IF 1:500
Anti-BrdU Mouse monoclonal antibody	BD	347580	B44	IF 1:500
Anti-ATR(phospho S428) Rabbit monoclonal antibody	Abcam	ab178407	EPR2184	WB 1:1000

Anti-ATM(phospho S1981) Rabbit monoclonal antibody	Abcam	ab81292	EP1890Y	WB 1:1000
Anti-53BP1 Rabbit monoclonal antibody	Abcam	ab175933	EPR2172	IF 1:500
Anti-53BP1 Mouse monoclonal antibody	Santa Cruz Biotechnology	sc-517281		IF 1:50
Anti-BRCA1 Mouse monoclonal antibody	Santa Cruz Biotechnology	sc-6954	D-9	IF 1:100
GAPDH Mouse monoclonal antibody	Proteintech	60004-1-Ig	1E6D9	WB 1:5000
HRP-conjugated Affinipure Goat Anti-Mouse IgG(H+L)	Proteintech	SA00001-1		WB 1:5000
HRP-conjugated Affinipure Goat Anti-Rabbit IgG(H+L)	Proteintech	SA00001-1		WB 1:5000
HRP-conjugated Affinipure Mouse Anti-Rabbit IgG (Light-Chain Specific)	Cell Signaling Technology	93702S	D4W3E	WB 1:1000
MYC-tag Rabbit polyclonal antibody	Proteintech	16286-1-AP		WB 1:2000
DYKDDDDK tag Rabbit polyclonal antibody	Proteintech	20543-1-AP		WB 1:2000
His tag Mouse monoclonal antibody	GenScript	A02003	C157	WB 1:1000
MBP tag Mouse monoclonal antibody	Proteintech	66003-1-Ig	4C6H4	WB 1:2000
Anti-OBFC2B Rabbit polyclonal antibody	Proteintech	14809-1-AP		IF 1:100 WB 1:1000
Anti-VPS72 Rabbit polyclonal antibody	Proteintech	15143-1-AP		WB 1:1000
Anti-Rad51 Mouse polyclonal antibody	Abcam	ab88572		IF 1:500
Anti-H4K16ac Rabbit monoclonal antibody	Cell Signaling Technology	13534S		ChIP
Anti-Pan Acetyl-Lysine Rabbit polyclonal antibody	ABclonal	A2391		IP
Anti-RPA32/RPA2 Rat monoclonal antibody	Cell Signaling Technology	2208T	4E4	IF 1:200 WB 1:1000
Anti-pRPA32(S4/S8) Rabbit polyclonal antibody	Bethyl Laboratories	A300-245A-T		WB 1:1000
Anti-INTS3 Rabbit polyclonal antibody	Proteintech	16620-1-AP		IF 1:300 WB 1:1000
Anti-H3 Rabbit monoclonal antibody	Cell Signaling Technology	4499T	D1H2	WB 1:1000
anti-Fibrillarin Rabbit monoclonal antibody	Abcam	ab166630	EPR10823(B)	IF 1:1000
Anti-SOX9 Rabbit polyclonal antibody	Abcam	ab185966		IF 1:1000

Anti-SYCP3 Rabbit polyclonal antibody	Abcam	ab15093		IF 1:3000
Anti- UTF1 Mouse monoclonal antibody	Sigma	MAB4337	5G10.2	IF 1:50
Anti- PLZF Mouse monoclonal antibody	Santa Cruz Biotechnology	sc-28319		IF 1:50
Anti- CyclinA2 Mouse monoclonal antibody	Proteintech	66391-1-Ig		IF 1:100
Anti-MAGEA4 Rabbit polyclonal antibody	Sigma	HPA021942		IF 1:200
Anti-SALL4 Mouse polyclonal antibody	MXB	MAB0691		IF 1:1
Anti-SALL4 Rabbit polyclonal antibody	Abcam	ab29112		IF 1:100
Anti-cKit Rabbit monoclonal antibody	Cell Signaling Technology	37805		IF 1:100
Anti-RNF168 Rabbit polyclonal antibody	Proteintech	21393-1-AP		IF 1:500
Anti- H2AK119ub Rabbit monoclonal antibody	Abcam	ab193203		WB 1:1000
FlexAble 2.0 CoraLite® Plus 555 Antibody Labeling Kit for Rabbit IgG	Proteintech	KFA502		
FlexAble 2.0 CoraLite® Plus 647 Antibody Labeling Kit for Mouse IgG2a	Proteintech	KFA543		
Goat anti-Rabbit IgG Secondary Antibody, Alexa Fluor 488	Thermo Fisher Scientific	A11034		IF 1:500
Goat anti-Rabbit IgG Secondary Antibody, Alexa Fluor 555	Thermo Fisher Scientific	A21429		IF 1:500
Goat anti-Mouse IgG Secondary Antibody, Alexa Fluor 488	Thermo Fisher Scientific	A11001		IF 1:500
Goat anti-Mouse IgG Secondary Antibody, Alexa Fluor 555	Thermo Fisher Scientific	A21424		IF 1:500
Donkey anti-Rat IgG Secondary Antibody, Alexa Fluor 488	Thermo Fisher Scientific	A21208		IF 1:500
Goat anti-Rabbit IgG Secondary Antibody, Alexa Fluor 647	Thermo Fisher Scientific	A21245		IF 1:500
Goat anti-Mouse IgG Secondary Antibody, Alexa Fluor 647	Thermo Fisher Scientific	A21236		IF 1:500

IF: immunofluorescent, IP: immunoprecipitation, WB: western blot