Differential regulation of S region hypermutation and class switch recombination by non-canonical functions of uracil DNA glycosylase

Ashraf S. Yousif, Andre Stanlie, Samiran Mondal, Tasuku Honjo, and Nasim A. Begum

Department of Immunology and Genomic Medicine, Kyoto University Graduate School of Medicine, Yoshida, Sakyo-ku, Kyoto 606-8501, Japan

Corresponding author: Tasuku Honjo; Phone +81-75-753-4371; Fax +81-75-753-4388; E-mail honjo@mfour.med.kyoto-u.ac.jp

Running title

Non-canonical function of UNG in SHM and CSR

Summary

Activation-induced cytidine deaminase (AID) is essential to class switch recombination (CSR) and somatic hypermutation (SHM) in both V region (v-SHM) and S region (s-SHM). Uracil DNA glycosylase (UNG), a member of the base excision repair (BER) complex, is required for CSR. Strikingly, however, UNG deficiency causes augmentation of SHM, suggesting involvement of distinct functions of UNG in SHM and CSR. Here we show that non-canonical scaffold functions of UNG regulate s-SHM negatively and CSR positively. The s-SHM suppressive function of UNG is attributed to the recruitment of faithful BER components at the cleaved DNA locus with competition against error-prone polymerases. By contrast, CSR promoting function of UNG enhances AID-dependent S-S synapse formation by recruiting 53BP1 and DNA PKcs. Several loss-of-catalysis mutants of UNG discriminated CSR promoting activity from s-SHM suppressive activity. Taken together, the non-canonical function of UNG regulates the steps after AID-induced DNA cleavage: error-prone repair suppression in s-SHM and end-joining promotion in CSR.

Significance Statement

UNG has been known as a critical BER protein required for CSR and SHM. On the other hand, its precise function in both CSR and SHM is extremely debatable and elusive. Here we showed that UNG suppresses s-SHM by recruiting the faithful DNA repair complex and in the absence of UNG, the error-prone repair complex that induces s-SHM overrides. Moreover, UNG promotes AID induced CSR by regulating S-S synapse and DNA end-repair. Interestingly enzymatic activity of UNG is dispensable for s-SHM suppression and CSR promotion.

/body

Introduction

Activation-induced cytidine deaminase (AID) is essential for somatic hypermutation (SHM) and class switch recombination (CSR), the genetic alterations that engrave antigen memory in the immunoglobulin gene (Ig) locus (1, 2). Both events are dependent on transcription of the target loci and take place during the G1 phase of the cell cycle (3). AID appears to be required for two distinct functions during CSR, namely DNA cleavage and recombination associated with its N-terminal and C-terminal regions, respectively (4-7).

AID introduces single-strand DNA break (SSB) on DNA to initiate both SHM and CSR (8, 9). Currently, the molecular mechanism for DNA cleavage by AID is extensively debated (2, 10). One hypothesis proposes that AID directly deaminates cytosine in DNA and the other considers the possibility that AID deaminates RNA to cause DNA cleavage.

SHM has been considered to depend on two events: a) SSB generation and b) repair by error-prone DNA synthesis using translesion polymerases (TLP). Recent studies clearly demonstrated that Polη and Polζ are the major TLP to introduce SHM (11, 12). Although CSR is also initiated by AID dependent SSB, CSR has several distinct features from SHM such as SSB processing to double strand break (DSB) and joining of appropriate pairs of DSB ends (13, 14).

Uracil DNA glycosylase (UNG) is known to be a key enzyme of the base excision repair (BER) system that carries out faithful repair. UNG removes damaged bases or mis-incorporated uracil on DNA to generate a basic site (15). The reaction is followed by a series of BER enzymes, including AP endonuclease, PARP1, XRCC1, Polβ,

FEN1, ligase1/3 (16, 17). The BER pathway is highly conserved from *E. coli* to mammalian systems. UNG is known to form a large complex with members of the BER system. Not only their physical interactions are demonstrated but also their genetic interactions are well documented (18, 19) Genetic defects of all these enzymes have been shown to inhibit the correct repair of DNA damage and enhance error-prone repair (20, 21).

UNG is also known to have a non-canonical function. The HIV-1 accessory protein Vpr recruits mammalian UNG to its integrase-complex required for DNA synthesis and recombination between viral and host genomes(22). Curiously, this function of UNG is independent of its catalytic activity. Similarly, vaccinia virus UNG protein but not its catalytic activity is essential for the viral replicative cycle by the formation of a large complex to support processive DNA synthesis (23, 24). UNG is also known to be recruited to DSB foci in association with γH2AX (25). Curiously, this recruitment is independent of its catalytic activity but dependent on the WxxF motif of UNG, which mediates interaction with the HIV-1 accessory protein Vpr. Interestingly, the WxxF motif was also found to be critical site for CSR (26). Requirement of a non-canonical activity of UNG was proposed in AID induced CSR (27, 28), especially because no correlation could be found between CSR efficiency and catalytic activity of UNG. More than 10 mutants, with wide ranges of enzymatic activity, were used in this study. A WxxF site mutant that retains 20% of the WT catalytic activity fails to support CSR whereas over 300 fold catalytically crippled mutants like D145N/H268L supports CSR as efficiently as WT. Shroyer *et al*., (29) also reported the non-enzymatic activity function of UNG in damage base repair. Authors showed that the *E. coli* UNG can bind a basic sites or base gaps independently of its enzymatic activity, suggesting an alternative mechanism of lesion

processing by UNG downstream of the damage base removal.

UNG has been proposed to be involved in AID-dependent DNA cleavage by DNA deamination hypothesis, in which AID generates uracil from cytosine on DNA, providing the substrate for UNG in the Ig locus (30, 31). Indeed, UNG deficiency drastically reduces CSR efficiency (26, 28, 30, 32). If UNG is involved AID dependent DNA cleavage, UNG deficiency is supposed to reduce not only CSR but also SHM because SHM is not restricted C/G and depends on error-prone repair of DNA breakage (12). Surprisingly, however, careful data analyses indicate that UNG deficiency rather augments SHM not only in Ig loci but also other target loci such as c-myc and bcl-6 (30, 32-34).

To solve this paradox, we examined the role of UNG in s-SHM and found that UNG suppresses s-SHM by recruiting BER enzymes. The UNG-BER complex competes against TLP for binding to the DNA damaged sites, by which correct and error-prone repairs appear to be balanced. We have further shown that this function of UNG does not depend on the catalytic activity of UNG. On the other hand, UNG deficiency inhibits AID-induced long-range interaction between S regions. UNG is also involved in recruiting synapse forming factors, such as 53BP1 and DNA PKcs to facilitate ligation of correct end pairs for CSR. This function is also independent of its catalytic activity. We thus conclude that the non-canonical function of UNG is involved in s-SHM and CSR by distinct mechanisms after DNA cleavage.

Results and Discussion

UNG suppresses s-SHM and promotes CSR

Since the UNG deficiency causes apparently opposite effects on CSR and SHM, we

speculated that UNG may have two distinct functions: a suppressor of SHM and a positive regulator of CSR. In order to understand the differential roles of UNG in SHM and CSR systematically, we investigated the effect of UNG expression on s-SHM and CSR in three different types of B cells (Wild type, $UNG^{-/-}$ and $AID^{-/-}$ $UNG^{-/-}$).

We examined whether UNG possesses the mutation suppressive activity by UNG overexpression in wild type B cells. Indeed, the mutation frequency at the 5' sequence of the core Sμ region, which is frequently targeted by AID induced mutations, was drastically reduced (~8 fold) without alteration of the mutation base profile (Fig. 1A, 1B; Table S2 and S3). Similarly, expression of UNG suppressed AID-induced mutations in UNG deficient B cells. Interestingly, however, the GC biased mutation profile, which is typical of UNG deficiency, remained uncorrected even after UNG overexpression (Fig. 1A, 1B and Table S3). We assumed that the failure of restoring the mutation base bias could be due to a technical limitation; GC biased mutations quickly accumulated before UNG expression using the retroviral vector. Retroviral transduction of UNG requires pre-activation of the splenic B cells, which induces AID prior to UNG expression and causes the mutation base bias. In order to test this possibility we co-expressed AID and UNG in $AID^{-/-}$ $UNG^{-/-}$ B cells and observed the mutation suppression, as well as the complete restoration of the GC/AT mutation ratio (Fig. 1A, 1B, and Table S3).

By contrast, similar overexpression of UNG hardly affected the CSR efficiency in wild type B cells (Fig. 1C and S1D), indicating distinct roles of UNG in s-SHM and CSR regulation. We confirmed that UNG expression restores CSR activity in $UNG^{-/-}$ as well as $AID^{-/-}$ $UNG^{-/-}$ B cells when AID is expressed (Fig. 1C and S1D). We also

confirmed that a UNG mutant with the N-terminal 90 residues deletion (Δ90 UNG) showed similar or even better CSR rescue, albeit slightly less s-SHM suppression, indicating that the core domain of UNG plays a major role in mutation suppression as well as CSR promotion. Consistent to our finding, recent observation by Zhan *et al*., (33) revealed that UNG deficient mice have even stronger SHM in the V region than WT. N-terminal of UNG is unique in that it possesses Proliferating cell nuclear antigen (PCNA) and Replication protein A (RPA) interacting sites (Fig.S3), but the the region is not necessary for CSR (26, 28). However, Guenzel *et al*, (22) found UNG-RPA interaction is necessary to suppress the mutation frequency in HIV-1, which is in contrast to what we observed in the case of AID induced s-SHM.

Distinct regulation of s-SHM and CSR by UNG mutants

Since we reported dispensability of UNG catalytic activity in CSR (26-28), we investigated the contribution of UNG catalytic activity in the mutation suppression process. Unexpectedly, two loss-of-catalysis mutations located in the core domain, H268L and D145N, both of which are CSR proficient, acted differentially for mutation suppression; H268L mutant suppressed the mutation frequency, whereas D145N mutant did not inhibit mutations at all (Fig. 1D, Table S1 and S2). As both mutants are severely defective in U-removal activity (Table S1) (35), differential s-SHM suppressive effects cannot be explained by the presence or absence of the catalytic activity. The N-terminal deletion of both UNG mutants affected s-SHM suppression marginally, although the N-terminal domain is critical for CSR restoration function of both mutants in agreement with previous report (28) (Fig. 1D and Table S1).

Similarly, we tested the mutation suppression activity of the WxxF motif mutant with

the N-terminal truncation (ΔN W231A), which is unable to support CSR (28). The W231A mutant with or without the N-terminal 90 residues was as active as H268L mutant for the mutation suppression, again indicating that UNG plays distinct roles in s-SHM and CSR (Fig. 1D, Table S1 and S2). Clear functional dissociation is also evident for other loss-of-catalysis mutants D145E and N204V, which are CSR proficient but much less efficient in mutation suppression (Fig. 1D, Table S1 and S2). When we plot s-SHM suppression activities versus enzymatic activities of individual mutants, we found the absence of correlation between mutation suppression function of UNG and its U removal activity (Fig. 1E). The phenomenon is equally applicable for both types of mutants, catalytic activity less than 1% or over. Most strikingly, when we plot the s-SHM suppression activity vs CSR promotion activity, almost all mutants are mapped off the proportional line (Fig. 1F). Thus we conclude UNG non-catalytically regulates CSR and SHM differentially.

UNG recruits BER enzymes to suppress s-SHM

Since the catalytic activity of UNG is not required for its s-SHM suppression activity, it is possible that UNG functions as a scaffold protein to enhance a faithful DNA repair cascade by which error-prone repair at AID-induced SSB can be suppressed. It has been reported that the core domain of UNG interacts with FEN1 and XRCC1 as well, which is known to associate with a series of BER proteins including Polβ, APE1, Ligase III and PARP1 (18). If so, UNG deficiency may reduce the amount of the BER protein complex loaded at the break site. In order to test such possibility, we first examined whether UNG is recruited to the target S-region after AID activation.

Recruitment of UNG to the target loci (Sμ and Sγ1) was clearly augmented in wild type B cells compared to $AID^{-/-}$ B cells under the stimulated condition (Fig. 2A),

suggesting that increased UNG binding occurs upon AID induced damage response in agreement with previous reports (32, 36). Next, we examined the effect of UNG deficiency on the recruitment of Polβ, XRCC1, FEN1 and PARP1, after confirming their normal expression in $UNG^{-/-}$ cells (Fig. S2C). In $UNG^{-/-}$ B cells, ChIP signals for all these BER proteins were drastically reduced (Fig. 2A), indicating that UNG is required for their loading after DNA damage induction. Indeed, their signals in $UNG^{-/-}$ cells were as low as those in the $AID^{-/-}$ condition. Low levels of Polβ and XRCC1 proteins observed in the Sγ1 region of $AID^{-/-}$ cells are probably due to the presence of a basal UNG level, which is involved in surveillance of DNA damage during other cellular functions such as DNA replication or CSR independent general DNA repair. UNG deficiency also reduced the relative nuclear distribution of Polβ and XRCC1 (Fig. S2D). Thus the enhanced mutation frequency in the absence of UNG is likely due to the failure of recruitment of faithful repair components to the damaged loci. In fact, genetic defect of Polβ, XRCC1, FEN1 and PARP1, all are reported to enhance the AID-induced mutation frequency but have no effects on CSR (37-40).

UNG deficiency enhances TLP recruitment

Strikingly, in the absence of UNG and faithful repair components, AID target loci showed elevated association with REV1, which is another large scaffold protein that recruits multiple error-prone TLP polymerases (Fig. 2A) (41). Consistently, we observed elevated recruitment the catalytic subunit of Polζ (REV3) in both Sμ and Sγ1 but much less in Cμ in UNG deficiency (Figure 2A). Polζ , known as an efficient damage bypass enzyme, also helps the extension of patch synthesis by TLP Polη (12). Several studies, including yeast and mammals, indicate that GC base biased mutation during SHM can also be initiated by REV1 and REV3 (42-44). Massive reduction of

GC as well as AT mutations was observed upon conditional inactivation of REV3 in mature B cells (12). Consistently, highly frequent mutations with altered base bias were evident in a knock-in mouse model of hyper-mutagenic mutant of REV3 (Rev3L2610F) (12).

We also detected increased recruitment of MSH2 (Fig. 2A), whose deficiency in mice leads to a moderate decrease in the mutation frequency (45, 46), suggesting occupancy of UNG in the S-region is repellent to MSH2, and thereby UNG blocks mutagenic effects exerted jointly by MSH2 and TLPs.

Contrary to our finding, Zan *et al*., (32) reported that REV1 recruits UNG through its WxxF motif and subsequently promotes CSR and SHM in a similar manner. However, as observed in their own study, neither s-SHM frequency nor CSR efficiency is comparable between REV1 and UNG deficiency. Here we show that both W231A and Δ90 W231A can suppress s-SHM. If the WxxF motif is indeed required by REV1 for the recruitment of UNG and functions similarly in the s-SHM pathway as proposed above, the mutation frequency in W231A expressing cells is expected to be as high as $UNG^{-/-}$ cells, which is not the case (Fig. 1D). In addition, we have shown that W231A, but not Δ90 W231A, is CSR proficient (28).

s-SHM suppression function of UNG linked with FEN1

Although the inteaction between faithfull repair factors and UNG has been well documented, precise interaction sites were not studied in many instances. We were particularly interested to address whether the interaction with any of the faithful repair factors is defective in the case of the loss-of-catalysis mutant D145N, which failed to suppress the enhanced mutations in UNG deficiency. We examined the interaction of

UNG with several known BER candidates by co-immunoprecipitation assay and found that FEN1, the Flap-end processing enzyme interacted with wild type UNG and the H268L mutant but not with D145N, D145E and N204V (Fig. 2B and S2B). It seems that more than one active site residues support UNG's interaction or association with FEN1 directly, or indirectly through other proteins. In any case, these UNG mutants irrespective of their N-terminal truncation or in combination with other mutation are inefficient in s-SHM suppression (Fig. 1D). On the other hand, wild type UNG, Δ90UNG, H268L and WxxF site mutants are proficient in mutation suppression and FEN1 association.

In order to confirm the biological relevance of loss of interaction between UNG and FEN1 we selected D145N mutant for further study. We evaluated recruitment of FEN1, Polβ, REV1 and REV3 to the S-region in stimulated $UNG^{-/-}$ B cells expressing the D145N mutant. We found that indeed FEN1 loading to cleavage target loci was dramatically dropped in the presence of the D145N mutant compared with UNG and H268L mutant (Fig. 2C). Similarly Polβ recruitment is reduced by D145N mutation compared with UNG and H268L mutation. In addition, expression of the D145N mutant caused elevated deposition of REV1 and REV3. These results supports the idea that the defect of S region mutation suppression by D145N is due to the failure to form a complex with other BER enzymes including FEN1 and Polβ. In the presence of the D145N mutant, REV1, REV3 and probably other TLP polymerases may be more efficiently recruited to AID targets to introduce mutations. Therefore, it possible that FEN1 and /or its associated complex are recruited in AID induced DNA damage sites via UNG, which is critical to error free DNA repair and mutation suppression.

Defective FEN1 was reported to increase mutation load and aberrant genomic

rearrangements (40). We therefore conclude that a part of the catalytic domain of UNG is also utilized for the non-canonical function of UNG, which plays a critical role in suppression of AID induced mutation. Taken together, UNG mediated balance of recruitment between error-free and error-prone repair machineries may well explain why loss of UNG augments mutations while its presence leads to mutation suppression.

UNG deficiency inhibits S-S synapse formation during CSR

Unlike SHM, CSR requires at least three different steps: (i) the S-S synapse that brings a correct pair of the cleaved-ends in donor and acceptor S regions to proximity, (ii) end processing of SSB to DSB and (iii) cleaved end repair and ligation. Therefore, defect in either of the three steps can give rise to severe blockade of CSR. Although the latter two steps are likely to be common to many DNA damage-induced recombination, the first step may be unique to CSR to secure its efficient and correct recombination. We thus examined the possibility that the S-S synapse formation (47) may be defective in UNG deficiency. Although it is not well understood how distantly located donor Sμ and acceptor S regions come to proximity and form the S-S synapse during end joining, long-range interactions are known to take place between specific S regions. Repair factors like ATM, ATR, 53BP1, γH2AX, and DNA-PKcs are known to be involved in early stages of DSB repair but not in SHM (48-51).

In order to examine if UNG plays a role in the formation of the recombination associated mega-complex between two S-regions, we employed the chromosome conformation capture (3C) technique, which allows us to detect the long range interaction between Eu and 3' Ea as well as between Sμ and Sγ1 (52). As shown in Fig. 3A-3C, UNG deficiency caused 2-7 fold reduction of Eμ-Sγ1, Eu-3'Ea and Sμ-

Sγ1 associations. In contrast, the conformations between the other S-S regions that do not recombine under the stimulation conditions showed no significant difference between the absence and presence of UNG. The observation is highly reproducible in independent experiments. We further validated the data by rescuing the S-S synapse defect in $UNG^{-/-}$ deficiency by wild type UNG complementation (Fig. 3D). Consistent with the defect of synapse formation in UNG deficiency, we observed loss of 53BP1 and DNA-PKcs from the acceptor and donor S regions in the absence of UNG. We confirmed that the deposition defect of 53BP1 and DNA-PKcs is not due to the reduced protein expression in UNG deficiency (Fig. 3E and S2E). Thus, UNG appears to be involved in stabilization of the long-range conformation of the IgH locus, which holds the DSB ends in proximity and plays a critical role to execute CSR efficiently.

Altered balance of repair and synapse factors leads to end-joining defect

Another step where UNG may play a role is processing of SSB to DSB and its protection. We examined the effect of UNG deficiency on the recruitment of the factors involved in end joining repair at the target loci. All the ChIP experiments were conducted under DSB inducible condition and $AID^{-/-}$ cells were used as control. End-processing and end-protecting factors such as Ku80, NBS1 and XRCC4 were expressed normally in $UNG^{-/-}$ cells, but showed elevated deposition in UNG deficiency compared with UNG proficiency (Fig. 3E and S2E).

It is also evident that Ku80, NBS1 and XRCC4 deposition is AID dependent, which further emphasizes the increase in unrepaired DSB in the absence of UNG. Excess Ku80 and XRCC4 deposition at the DSB ends may delay their joining and preventing them from end-ligation.

We reasoned the cumulative effects, S-S synapse defect and enhanced recruitment of NHEJ factors, would be reflected on the S-S junctional signature. We thus compared Sμ-Sγ1 junctions in CSR induced B cells between WT and $UNG^{-/-}$ mice. IgG1 positive cells were isolated from both groups and S-S junction sequences were analyzed; junctions with blunt end were 3 fold higher in $UNG^{-/-}$ compared to WT, and the average microhomology length at the junctions was concomitantly reduced (Fig. 4A, 4B and Table S4). Thus, the classical NHEJ (C-NHEJ) was promoted over alternative end joining (A-EJ) in the absence of UNG, which is consistent with the elevated level of C-NHEJ factors like Ku80 and XRCC4 (Fig. 3E) and decrease of A-EJ factor like PARP1 at the breakage loci (Figure 2A). As Ku80 and PARP1 are known to compete (53), the data is in good agreement with the reciprocal recruitment status and choice of C-NHEJ over A-EJ repair pathway. Similarly, a clear reciprocal recruitment trend was observed between UNG and MSH2 loading at the S-region, which may also contribute to the choice of DSB end joining pathways (54, 55).

CSR-defective UNG mutants fail to recruit the synapse factors

As we observed UNG deficiency causes recruitment alterations of synapse and repair factors like 53BP1, DNA PKcs, and Ku80, we examined the loading of these factors in the presence of the Δ90 D145N and Δ90 W231A mutants that fail to complement CSR (Fig. 1D, S2F and S2G). Strikingly, we observed that the recruitments status of the three proteins followed the same profile (decrement of 53BP1 and DNA-PKcs, and enhancement of Ku80) as observed in UNG deficiency (Fig. 4C and 4D). It is also interesting to note that 53BP1 and DNA-PKcs accumulation in the presence of Δ90 D145N or Δ90 W231A dropped to the same level as in the absence of AID, emphasizing the fact that UNG is essential to mount AID dependent 53BP1 and

DNA-PKcs accumulation in the S region and the Δ90 D145N and Δ90 W231A mutants are defective to elicit such a response.

Under the identical condition, expression of UNG, Δ90 UNG, H268L, D145N, and W231A mutants, which are CSR proficient, did not show any perturbation of synapse/repair protein recruitment (Fig. 4C and 4D). Clearly the relative accumulation of synapse factors at S regions well correlates with CSR induction (Fig.4E). These data well explain why the full length but not the N-terminally truncated UNG, either with loss-of-catalysis mutation or with WxxF motif mutation, are fully capable of CSR complementation. These results suggest that the N-terminal domain of UNG may serve as an accessory site, to provide a structural support to the core domain when mutations are introduced in the catalytic or WxxF motif residues to carry out a scaffold function required for CSR.

The fact that D145N, D145E and N204V mutants affect specifically s-SHM suppression but not CSR promoting activity, most convincingly indicates that UNG regulates the two genetic events differently. Since loss-of-catalysis mutations of UNG do not necessarily reduce s-SHM or CSR, it is likely that CSR and s-SHM are regulated by the scaffold function of different surfaces of the UNG structure. The study provides compelling evidence that the recruitment of critical DNA repair factors for s-SHM and CSR are orchestrated by UNG, and the entire events are dependent on AID (Fig. 5C, proposed model). Intriguingly, active site residues of UNG were found to be critical for interacting with DNA endonuclease FEN1. Therefore, the catalytic pocket of UNG is not solely for the enzymatic function, rather adapted to non-canonical function in DNA repair (25). It has been recently shown that loss-of-catalysis MUG (mismatch uracil DNA glycosylase) mutants bind damaged single

strand DNA as a dimeric complex (56). Interestingly, we also observed WT and loss-of-catalysis UNG mutants can form a dimer in the cell (Fig. 5A and B), which raises a possibility that the dimeric structure of UNG could favor a more dynamic scaffolding function as observed in the case of vaccinia virus UNG (23, 24). Future study may reveal how the structure of UNG promotes protein-protein interaction during SHM and CSR at the physiological milieu.

Experimental procedures

Mice

UNG-deficient mice were kindly provided by Dr. Rudolf Jaenisch (Department of Biology, Massachusetts Institute of Technology) and CSR defect was analyzed previously (27). AID-deficient mice were generated in our laboratory (1) and were crossed with UNG-deficient mice to obtain $AID^{-/-}$ $UNG^{-/-}$ double deficient mice.

Retroviral constructs

Nuclear form of human UNG (hUNG2) was amplified by RT-PCR and cloned into EcoRI and SalI sites of the retroviral expression vector pFB-IRES-GFP. Appropriate primer pairs were designed (Table S5) to generate individual catalytic and WxxF site mutants following the procedure of Quick Change II Site-Directed Mutagenesis system (Agilent Technologies). Single and double point mutations were initially generated in wild-type cDNA in TOPO-Blunt vector. After sequence verification, mutated cDNAs were transferred into pFB-IRES-GFP vector (Fig. S1A). All the constructs retained their natural Kozak consensus sequence for the initiation of translation. Mutated amino acid positions were shown in the sequence alignment of nuclear isoform of mouse and human UNG (Fig. S3). For the purpose of clarity, the name of the human UNG mutants were kept same as described (27, 35). In order to

express AID and UNG from a single retroviral vector, hAID was first cloned into EcoRI and SalI1 sites of the retroviral expression vector pFB-IRES-GFP. Later GFP portion was replaced by UNG fused with GFP-Flag at the C terminus using BstX1 and Not1 sites (Fig. S1A).

In vitro culture and CSR assay

B lymphocytes were isolated from 8-12 week-old mouse spleens using BD IMag™ B Lymphocyte Enrichment Set - DM (557792) and cultured at a concentration of 1.0 x 10^6 cells/ml in complete RPMI medium containing 25 μg/ml LPS and 7.5 ng/ml IL-4 to undergo class switching to IgG1. During retroviral transduction cells were pre-activated prior to infection by culturing in presence of LPS and IL4 for 48h. Standard protocol was followed to prepare the retroviral supernatants and for the infection of WT, $UNG^{-/-}$ and $AID^{-/-}UNG^{-/-}$ spleen cells. Flow cytometric analysis of IgG1 expression was performed using Biotinylated-anti IgG1(Pharmingen) and APC conjugated Streptavidin (eBioscience) on day 3 (Fig. S1B). And IgG1 switch efficiency was calculated from infected GFP positive cells in the live gate.

Mutation analysis of 5' Sμ

Genomic DNA was isolated from either $IgG1^+$ or IgG^+GFP^+ sorted cells under different conditions applied. A 565 bp region located 5' of core Sμ was PCR amplified by 5'-AATGGATACCTCAGTGGTTTTTAATGGTGG (Forward primer) and 5'-GCGGCCCGGCTCATTCCAGTTCATTACAG (Reverse primer) using high fidelity Pyrobest DNA polymerase (TAKARA). The PCR product was cloned in ZeroBlunt-Topo vector for sequencing and mutation analysis. The sequences were determined using an ABI 3130xl Genetic Analyzer (Applied Biosystems) and sequence analysis was performed by Sequencher DNA analysis software. For each set

96 clones or more were sequenced bidirectionally and only the unique mutations were counted.

Analysis of switch recombination junction

Genomic DNA was isolated from $IgG1^+$ B cells derived from WT and $UNG^{-/-}$ or $UNG^{-/-}$ complemented by either UNG or Δ90 UNG activated for 3 days (Fig. S1B). Approximately 400 ng of DNA was used to amplify Sμ-Sγ1 junction from each sample by nested PCR using Pyrobest DNA polymerase. The primers for the first round (Sμ1 and Sγ1.1) and second round (Sμ2 and Sγ1.1) PCR are listed in Table S5. The PCR condition (1^{st} round) was 6 cycles of 93°C (40 s), 64°C (40 s), and 72°C (120 s), followed by a further 24 cycles with the annealing performed at 55 °C. The 2^{nd} round PCR was 24 cycles of 93 °C (40 s), 55 °C (40 s), and 72 °C (120 s). Sequence of each clone was subjected to BLAST analysis against reference sequences (Sμ-J00440.1, Sγ1-D78344.1) to identify the junctions.

Immunoprecipitation of UNG

WT UNG and loss of catalysis function mutants H268L and D145N were expressed as GFP-Flag fused UNG in UNG deficient B cells through retroviral expression vector. Vector expressing GFP only used as mock (negative control). After stimulation for 3 days, cells were crosslinked by 1% formaldehyde and cell lysates were prepared as described by Akbari *et al*., (18). Immunoprecipitated complex was recovered by ani-Flag-M2 agarose beads and it was directly lysed in SDS sample buffer for gel electrophoresis and subsequent Western blot analysis. IP efficiency was examined by comparing 0.5–1 × 10^6 cell equivalent of total cell lysate and the IP product. Unbound fractions were also analyzed to confirm that more than 90% of UNG was bound to the bead. Western blot analysis was performed according to standard protocols and the antibodies used were listed in Table S6.

Chromatin immunoprecipitation (ChIP) Analysis

The ChIP assay was performed using Active Motif ChIP-IT Express Kit according to the manufacturer's instructions. In brief, 5×10^6 cells were fixed in the presence of 1% formaldehyde for 5 min at room temperature. The reaction was stopped by the addition of glycine to a final concentration of 0.125 M. A soluble chromatin fraction containing fragmented DNA of 500–200 bp was obtained after cell lysis and sonication. Immunoprecipitation was performed by incubating the lysate with 2–3 μg antibody (Table S6). The pulled-down DNA was subjected to detection by real-time PCR normalized to the amount of input followed by the maximum value in each data set.

Protein expression and BiFC assay

Expression of the DNA repair factors in WT and $UNG^{-/-}$ was done by standard western blot. WT and $UNG^{-/-}$ splenic B cells stimulated for 3 days. Cells lysate were prepared using RIPA buffer supplemented with protease inhibitor. BiFC constructs of UNG were performed by fusing UNG with either N- or C-terminal fragment of Kusabira Geen Fluorescent protein (mKG) as described by Ueyama *et al*., (57). Co-transfection of the constructs were performed in appropriate combinations in 293T cells and cells were harvested either 24h or 48h for FACS analysis.

Chromosome Conformation Capture (3C) Assay

A modified 3C assay was adopted based on a previously described assay by Wuerffel *et al* (52). In brief, $1x10^7$ activated B cells from WT and $UNG^{-/-}$ mice were washed by PBS and subjected to 1% formaldehyde cross linking for 5 min at RT, the reaction was stopped by adding 1.25M glycine followed by 1x wash with PBS. Nuclear lyaste was prepared using 500 μl buffer following the instruction of Active motif ChIP Kit manual. Crosslinked chromatin was digested with Hind III, and ligation of the

digested crosslinked chromatin was performed by T4 DNA ligase (Takara #2011A). Ligated chromatin was treated with proteinase K, reverse crosslinked and DNA was purified by phenol chloroform extraction. The details of the protocol can be available upon request. PCR conditions and the primers (Table S5) used were same as described previously.

Acknowledgments

We thank Ms. Y. Shiraki for her kind assistance in manuscript preparation and Ms. K. Yurimoto for her excellent technical assistance. This work was supported by Grant-in-Aid for Specially Promoted Research 17002015 (to T.H.), Grant-in-Aid for Scientific Research (C) 24590352 (to N.A.B.). ASY is supported by MEXT scholarship and SM is a CD-HFSP fellow.

References:

1. Muramatsu M, *et al.* (2000) Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* 102(5):553-563.
2. Muramatsu M, Nagaoka H, Shinkura R, Begum NA, & Honjo T (2007) Discovery of activation-induced cytidine deaminase, the engraver of antibody memory. *Adv Immunol* 94:1-36.
3. Schrader CE, Guikema JE, Linehan EK, Selsing E, & Stavnezer J (2007) Activation-induced cytidine deaminase-dependent DNA breaks in class switch recombination occur during G1 phase of the cell cycle and depend upon mismatch repair. *Journal of immunology* 179(9):6064-6071.
4. Doi T, *et al.* (2009) The C-terminal region of activation-induced cytidine deaminase is responsible for a recombination function other than DNA cleavage in class switch recombination. *Proc Natl Acad Sci U S A* 106(8):2758-2763.
5. Shinkura R, *et al.* (2004) Separate domains of AID are required for somatic hypermutation and class-switch recombination. *Nat Immunol* 5(7):707-712.
6. Barreto V, Reina-San-Martin B, Ramiro AR, McBride KM, & Nussenzweig MC (2003) C-terminal deletion of AID uncouples class switch recombination from somatic hypermutation and gene conversion. *Mol Cell* 12(2):501-508.
7. Ta VT, *et al.* (2003) AID mutant analyses indicate requirement for class-switch-specific cofactors. *Nat Immunol* 4(9):843-848.

8. Faili A, *et al.* (2002) AID-dependent somatic hypermutation occurs as a DNA single-strand event in the BL2 cell line. *Nat Immunol* 3(9):815-821.
9. Kinoshita K, Tashiro J, Tomita S, Lee CG, & Honjo T (1998) Target specificity of immunoglobulin class switch recombination is not determined by nucleotide sequences of S regions. *Immunity* 9(6):849-858.
10. Conticello SG, Langlois MA, Yang Z, & Neuberger MS (2007) DNA deamination in immunity: AID in the context of its APOBEC relatives. *Advances in immunology* 94:37-73.
11. Saribasak H, *et al.* (2012) DNA polymerase zeta generates tandem mutations in immunoglobulin variable regions. *The Journal of experimental medicine* 209(6):1075-1081.
12. Daly J, *et al.* (2012) Altered Ig hypermutation pattern and frequency in complementary mouse models of DNA polymerase zeta activity. *J Immunol* 188(11):5528-5537.
13. Honjo T, Kinoshita K, & Muramatsu M (2002) Molecular mechanism of class switch recombination: linkage with somatic hypermutation. *Annu Rev Immunol* 20:165-196.
14. Chaudhuri J, *et al.* (2007) Evolution of the immunoglobulin heavy chain class switch recombination mechanism. *Adv Immunol* 94:157-214.
15. Lindahl T, Ljungquist S, Siegert W, Nyberg B, & Sperens B (1977) DNA N-glycosidases: properties of uracil-DNA glycosidase from Escherichia coli. *J Biol Chem* 252(10):3286-3294.
16. Caldecott KW (2007) Mammalian single-strand break repair: mechanisms and links with chromatin. *DNA Repair (Amst)* 6(4):443-453.
17. Wilson SH & Kunkel TA (2000) Passing the baton in base excision repair. *Nature structural biology* 7(3):176-178.
18. Akbari M, *et al.* (2010) Direct interaction between XRCC1 and UNG2 facilitates rapid repair of uracil in DNA by XRCC1 complexes. *DNA Repair (Amst)* 9(7):785-795.
19. Akbari M, *et al.* (2004) Repair of U/G and U/A in DNA by UNG2-associated repair complexes takes place predominantly by short-patch repair both in proliferating and growth-arrested cells. *Nucleic Acids Res* 32(18):5486-5498.
20. Parsons JL & Dianov GL (2013) Co-ordination of base excision repair and genome stability. *DNA Repair (Amst)* 12(5):326-333.
21. Dianov GL & Hubscher U (2013) Mammalian base excision repair: the forgotten archangel. *Nucleic Acids Res* 41(6):3483-3490.
22. Guenzel CA, *et al.* (2012) Recruitment of the nuclear form of uracil DNA glycosylase into virus particles participates in the full infectivity of HIV-1. *J Virol* 86(5):2533-2544.
23. De Silva FS & Moss B (2003) Vaccinia Virus Uracil DNA Glycosylase Has an Essential Role in DNA Synthesis That Is Independent of Its Glycosylase Activity: Catalytic Site Mutations Reduce Virulence but Not Virus Replication in Cultured Cells. *Journal of Virology* 77(1):159-166.
24. Stanitsa ES, Arps L, & Traktman P (2006) Vaccinia virus uracil DNA glycosylase interacts with the A20 protein to form a heterodimeric processivity factor for the viral DNA polymerase. *J Biol Chem* 281(6):3439-3451.
25. Zeitlin SG, *et al.* (2011) Uracil DNA N-glycosylase promotes assembly of human centromere protein A. *PLoS One* 6(3):e17151.

26. Begum NA, *et al.* (2007) Requirement of non-canonical activity of uracil DNA glycosylase for class switch recombination. *J Biol Chem* 282(1):731-742.
27. Begum NA, *et al.* (2004) Uracil DNA glycosylase activity is dispensable for immunoglobulin class switch. *Science* 305(5687):1160-1163.
28. Begum NA, *et al.* (2009) Further evidence for involvement of a noncanonical function of uracil DNA glycosylase in class switch recombination. *Proc Natl Acad Sci U S A* 106(8):2752-2757.
29. Shroyer MJ, Bennett SE, Putnam CD, Tainer JA, & Mosbaugh DW (1999) Mutation of an active site residue in Escherichia coli uracil-DNA glycosylase: effect on DNA binding, uracil inhibition and catalysis. *Biochemistry* 38(15):4834-4845.
30. Rada C, *et al.* (2002) Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. *Curr Biol* 12(20):1748-1755.
31. Chaudhuri J & Alt FW (2004) Class-switch recombination: interplay of transcription, DNA deamination and DNA repair. *Nat Rev Immunol* 4(7):541-552.
32. Zan H, *et al.* (2012) Rev1 recruits ung to switch regions and enhances du glycosylation for immunoglobulin class switch DNA recombination. *Cell reports* 2(5):1220-1232.
33. Zahn A, *et al.* (2013) Separation of Function between Isotype Switching and Affinity Maturation In Vivo during Acute Immune Responses and Circulating Autoantibodies in UNG-Deficient Mice. *J Immunol* 190(12):5949-5960.
34. Liu M, *et al.* (2008) Two levels of protection for the B cell genome during somatic hypermutation. *Nature* 451(7180):841-845.
35. Mol CD, *et al.* (1995) Crystal structure and mutational analysis of human uracil-DNA glycosylase: structural basis for specificity and catalysis. *Cell* 80(6):869-878.
36. Ranjit S, *et al.* (2011) AID binds cooperatively with UNG and Msh2-Msh6 to Ig switch regions dependent upon the AID C terminus. *J Immunol* 187(5):2464-2475.
37. Saribasak H, *et al.* (2011) XRCC1 suppresses somatic hypermutation and promotes alternative nonhomologous end joining in Igh genes. *J Exp Med* 208(11):2209-2216.
38. Poltoratsky V, Prasad R, Horton JK, & Wilson SH (2007) Down-regulation of DNA polymerase beta accompanies somatic hypermutation in human BL2 cell lines. *DNA Repair (Amst)* 6(2):244-253.
39. Wu X & Stavnezer J (2007) DNA polymerase beta is able to repair breaks in switch regions and plays an inhibitory role during immunoglobulin class switch recombination. *J Exp Med* 204(7):1677-1689.
40. Larsen E, *et al.* (2008) Early-onset lymphoma and extensive embryonic apoptosis in two domain-specific Fen1 mice mutants. *Cancer Res* 68(12):4571-4579.
41. Guo C, *et al.* (2003) Mouse Rev1 protein interacts with multiple DNA polymerases involved in translesion DNA synthesis. *The EMBO journal* 22(24):6621-6630.
42. Masuda K, *et al.* (2009) A critical role for REV1 in regulating the induction of C:G transitions and A:T mutations during Ig gene hypermutation. *J Immunol* 183(3):1846-1850.

43. Auerbach P, Bennett RA, Bailey EA, Krokan HE, & Demple B (2005) Mutagenic specificity of endogenously generated abasic sites in Saccharomyces cerevisiae chromosomal DNA. *Proc Natl Acad Sci U S A* 102(49):17711-17716.
44. Roche H, Gietz RD, & Kunz BA (1995) Specificities of the Saccharomyces cerevisiae rad6, rad18, and rad52 mutators exhibit different degrees of dependence on the REV3 gene product, a putative nonessential DNA polymerase. *Genetics* 140(2):443-456.
45. Rada C, Ehrenstein MR, Neuberger MS, & Milstein C (1998) Hot spot focusing of somatic hypermutation in MSH2-deficient mice suggests two stages of mutational targeting. *Immunity* 9(1):135-141.
46. Bardwell PD, *et al.* (2004) Altered somatic hypermutation and reduced class-switch recombination in exonuclease 1-mutant mice. *Nat Immunol* 5(2):224-229.
47. Adams MM & Carpenter PB (2006) Tying the loose ends together in DNA double strand break repair with 53BP1. *Cell division* 1:19.
48. Reina-San-Martin B, Chen HT, Nussenzweig A, & Nussenzweig MC (2004) ATM is required for efficient recombination between immunoglobulin switch regions. *J Exp Med* 200(9):1103-1110.
49. Reina-San-Martin B, *et al.* (2003) H2AX is required for recombination between immunoglobulin switch regions but not for intra-switch region recombination or somatic hypermutation. *J Exp Med* 197(12):1767-1778.
50. Reina-San-Martin B, Chen J, Nussenzweig A, & Nussenzweig MC (2007) Enhanced intra-switch region recombination during immunoglobulin class switch recombination in 53BP1-/- B cells. *Eur J Immunol* 37(1):235-239.
51. Callen E, *et al.* (2009) Essential role for DNA-PKcs in DNA double-strand break repair and apoptosis in ATM-deficient lymphocytes. *Mol Cell* 34(3):285-297.
52. Wuerffel R, *et al.* (2007) S-S synapsis during class switch recombination is promoted by distantly located transcriptional elements and activation-induced deaminase. *Immunity* 27(5):711-722.
53. Wang M, *et al.* (2006) PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways. *Nucleic Acids Res* 34(21):6170-6182.
54. Kracker S, *et al.* (2010) Impaired induction of DNA lesions during immunoglobulin class-switch recombination in humans influences end-joining repair. *Proc Natl Acad Sci U S A* 107(51):22225-22230.
55. Schrader CE, Vardo J, & Stavnezer J (2002) Role for mismatch repair proteins Msh2, Mlh1, and Pms2 in immunoglobulin class switching shown by sequence analysis of recombination junctions. *J Exp Med* 195(3):367-373.
56. Grippon S, *et al.* (2011) Differential modes of DNA binding by mismatch uracil DNA glycosylase from Escherichia coli: implications for abasic lesion processing and enzyme communication in the base excision repair pathway. *Nucleic Acids Res* 39(7):2593-2603.
57. Ueyama T KT, Karasawa S, Kawasaki T, Shimizu A, Son J, Leto TL, Miyawaki A, Saito N (2008) Sequential binding of cytosolic Phox complex to phagosomes through regulated adaptor proteins: evaluation using the novel monomeric Kusabira-Green System and live imaging of phagocytosis. *Journal of immunology* 181(1):629-640.

Figure legends

Figure 1. s-SHM suppression and CSR promotion by UNG is catalytic activity independent. (A-C) Comparison of mutation frequency, base bias and CSR efficiency in WT, $UNG^{-/-}$ and $AID^{-/-}UNG^{-/-}$ B cells with and without complementation by UNG and AID. In the case of $UNG^{-/-}$ B cells, CSR was rescued either by UNG or Δ90 UNG. $AID^{-/-}UNG^{-/-}$ B cells were complemented by co-expressing AID and UNG from a single expression vector. Individual retroviral constructs and vector control introduced are indicated below each plot. (A and B) Upper and middle panel show analysis of hypermutation at 5' Sμ in $IgG1^{+}$ cells in each case and GC or AT base substitution percentage, respectively. (C) The bottom panel shows IgG1 switch efficiency as calculated from $IgG1^{+}GFP^{+}$ population, which was compiled from three independent experiments, UNG always consider as 100%. (D) Dispensability of catalytic activity of UNG and functional dissociation of SHM and CSR, Upper graph and lower graph represent the mutation frequency and CSR efficiency, respectively. The data was generated using $IgG1^{+}GFP^{+}$ cells obtained from $AID^{-/-}UNG^{-/-}$ B cells supplemented with indicated expression constructs of AID and UNG. D145N, H268L, D145E and N204V are the catalytic mutants of UNG and W231A is one of the WxxF site mutants (Begum et al., 2009). (E) Correlation between mutation suppression and U removal activity of UNG, the large graph illustrates mutants with U removal activity >1% and the inset graph represent mutants with <1% U removal activity. (F) Differential regulation of s-SHM and CSR by UNG. *P* value *< 0.05 and **<0.001).

Figure 2. AID-dependent UNG binding in the S region recruits BER proteins but inhibits MSH2, REV1 and REV3 recruitment. (A) WT and $AID^{-/-}$ splenic B cells were subjected to ChIP analysis of UNG and various repair factors indicated above

each plot. ChIP primer positions were designed to monitor Sμ, Sγ1, and Cμ. The ChIP signal at Cμ that serves as a control is considered to be DNA-break or -repair independent event. Mock represents the background or no antibody ChIP signal. The data was normalized to the DNA input signals and presented relative to the maximum value (fixed as 1) in Sμ and Sγ1 for each set. All experiments are done in triplicate. (B) Immunoprecipitation and Western blot analysis of GFP-Flag tagged UNG showing defective interaction between UNG mutant D145N and FEN1. Upper panel shows comparable expression of UNG, H268L and D145N in transduced $UNG^{-/-}$ splenic B cells. In the lower panel anti-Flag IPed products of UNG and indicated UNG mutants were analyzed; FEN1 can be detected with UNG and H268L but not with D145N. (C) ChIP assay showing FEN1, Polβ, REV1 and REV3 binding to Sμ and Sγ1 in $UNG^{-/-}$ B cells expressing Vector, UNG and two catalytic H268L and D145N. H268L expression but not D145N supports FEN1 recruitment in the S region. ChIP values were normalized to the DNA input signals as described. Three independent experiments showed similar result.

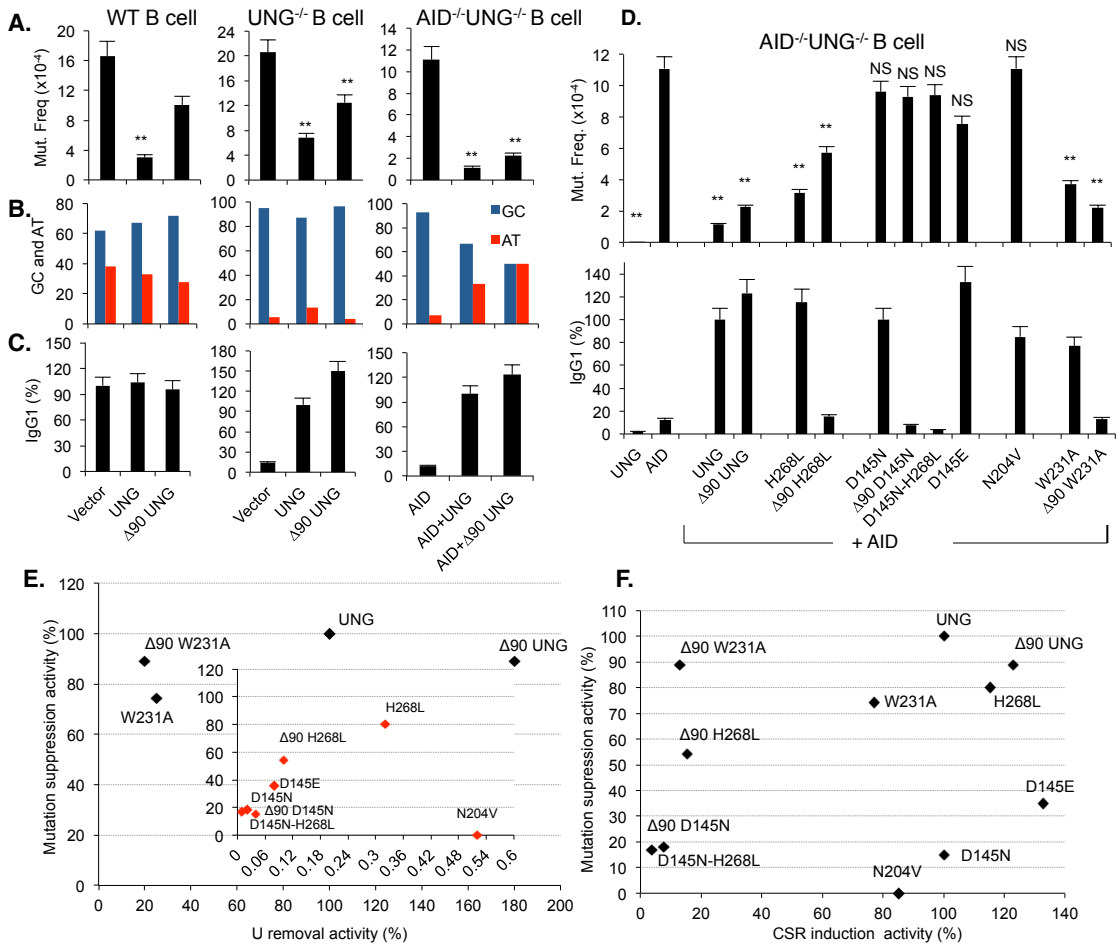
Figure 3. UNG deficiency perturbs S-S synapse formation required for CSR

(A) Chromosome conformation capture (3C) assay detecting long-range interaction (depicted by arches) of various acceptor S regions with Eμ and Sμ as anchors. (B) Interaction specific PCR products were detected by designed paired primer combination and indicated above the panel in spleen B cells with indicated genotypes. (C) Normalized image quantitation data represent the crosslinking frequency between the interacted regions, error bar represent mean of three independent experiments. (D) 3C assay in $UNG^{-/-}$ B cell expressing vector control and wild type UNG. Red rectangular box signifies the signal decrement upon UNG deficiency. Accordingly, the red arches in scheme-A highlight the defective interaction. A representative data

set was shown; the entire experiments were performed more than three times and similar results were obtained. (E) ChIP assay showing UNG deficiency reduces loading of 53BP1 and DNA PKcs in the S regions. Enhanced binding of Ku80, XRCC4 and NBS1 to Sμ and Sγ1 in $UNG^{-/-}$ compared to WT and $AID^{-/-}$ splenic B cells. ChIP data were normalized to the DNA input signals, followed by the maximum value in each data set as described in Fig.1.

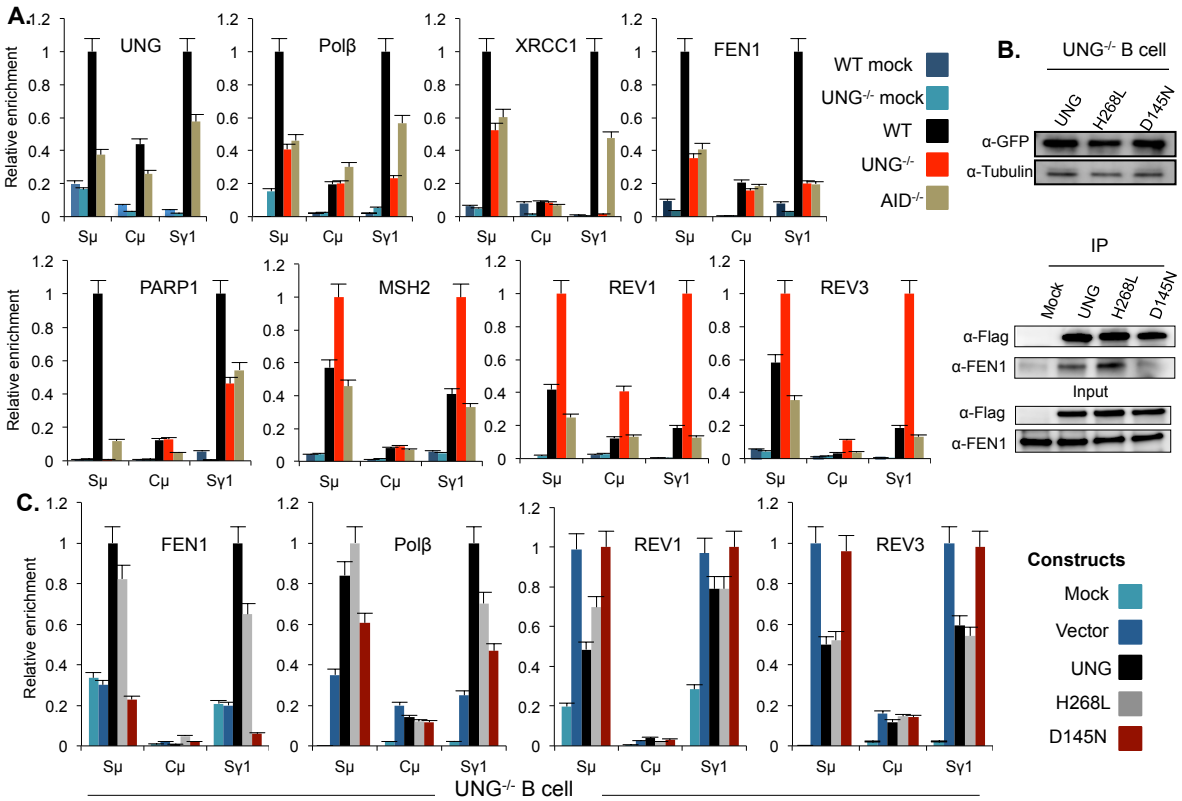
Figure 4. Altered occupancy of NHEJ repair factor skews CSR end-joining in UNG deficiency. (A) Sμ-Sγ1 switch junction analysis in $IgG1^{+}$ cells derived from WT and $UNG^{-/-}$ mice. First plots show the percentage of sequences with nucleotide overlap for Sμ and Sγ1. Junctions were grouped either as C-NHEJ or AEJ (middle panel) and the average microhomology length was calculated from the total junctions (rightmost plot). *P* value *< 0.05 and **<0.001. (B) Ectopic expression of UNG restores the altered end-joining defect in UNG deficiency. Sμ-Sγ1 switch junction analysis in $IgG1^{+}$ cells derived from $UNG^{-/-}$ splenic B cells expressing UNG, Δ90 UNG and vector control. Details of the junction analysis are summarized in Table S4. (C) S-region ChIP assay showing restoration of 53BP1 and DNA PKcs loading defect of $UNG^{-/-}$cells by expressing UNG but not Δ90 UNG mutants. ChIP assay was performed in $AID^{-/-}UNG^{-/-}$ splenic B cells and the various UNG constructs co-expressed with AID were indicated. (D) Ku80 ChIP assay in $UNG^{-/-}AID^{-/-}$ splenic B cells expressing either UNG or UNG mutant in conjunction with AID. Expression Δ90 D145N, which failed to rescue CSR, caused maximum or aberrant deposition of Ku80 in the S-region. (E) A scatterplot showing the positive correlation between CSR induction and the accumulation of 53BP1 or DNA-PKcs in various mutant background as indicated. The plot is derived from Figure 4(C).

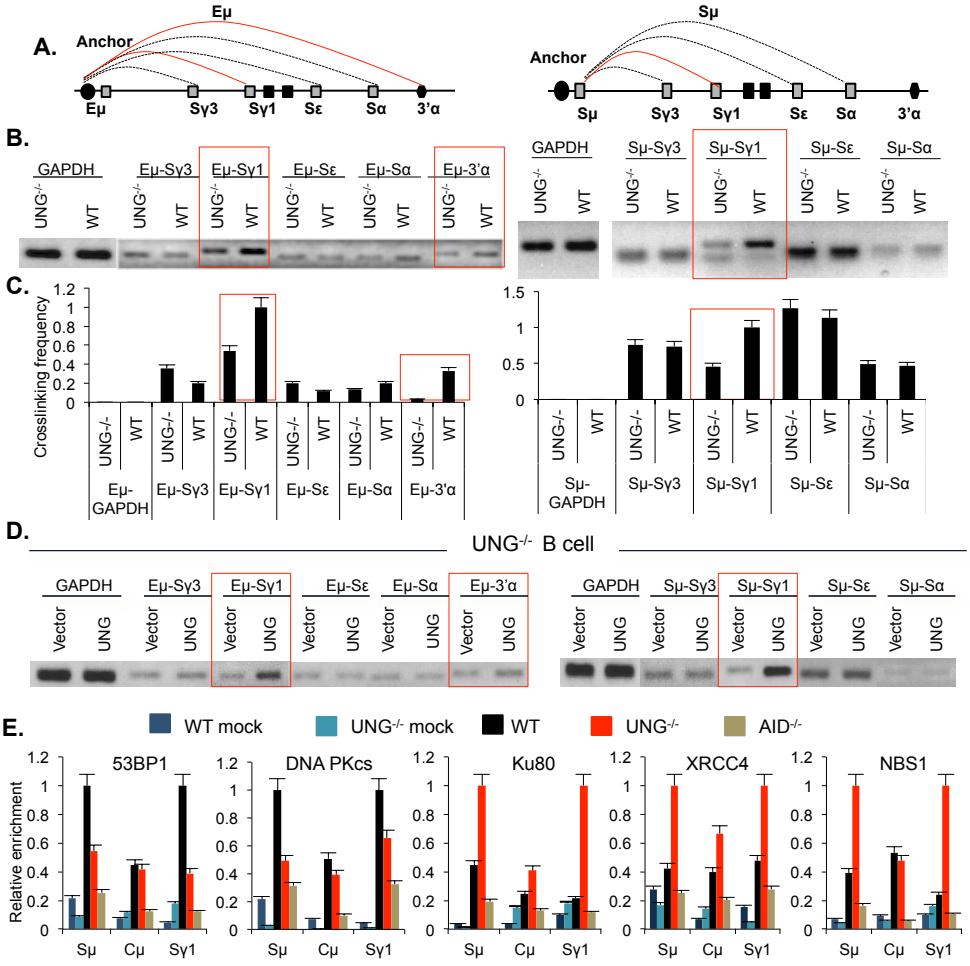
Figure 5. Bimolecular fluorescence complementation (BiFC) analysis shows UNG dimer formation in 293T cells and proposed model. (A) Schematic representation of the BiFC assay principle using a Kusabira Green Fluorescent protein. (B) Four types of constructs were prepared as shown in (A) for WT and D145N and the constructs were transfected in 4 possible pair combinations (indicated above each FACS profile). Percentages of fluorescent positive cells were indicated inside each FACS plot. Both the wild type UNG and its D145N catalytic mutant efficiently dimerizes in living cells. (C) Model of s-SHM and CSR regulation by UNG. s-SHM relies on the introduction of AID-induced SSBs at the S regions that are repaired by either error-free repair pathway mediated through UNG-BER complex, or error-prone repair by TLP. The two repair mechanisms compete with each other and the frequency of SHM is determined by the balance of the two: correct or error. In the case of CSR, AID-induced SSBs in the S region are processed to DSBs. UNG is required to recruit several DNA synapse and recombination factors that facilitate the end joing-process essential to efficient recombination to complete CSR. The dimer formation of UNG is important to bring two DSBs together.

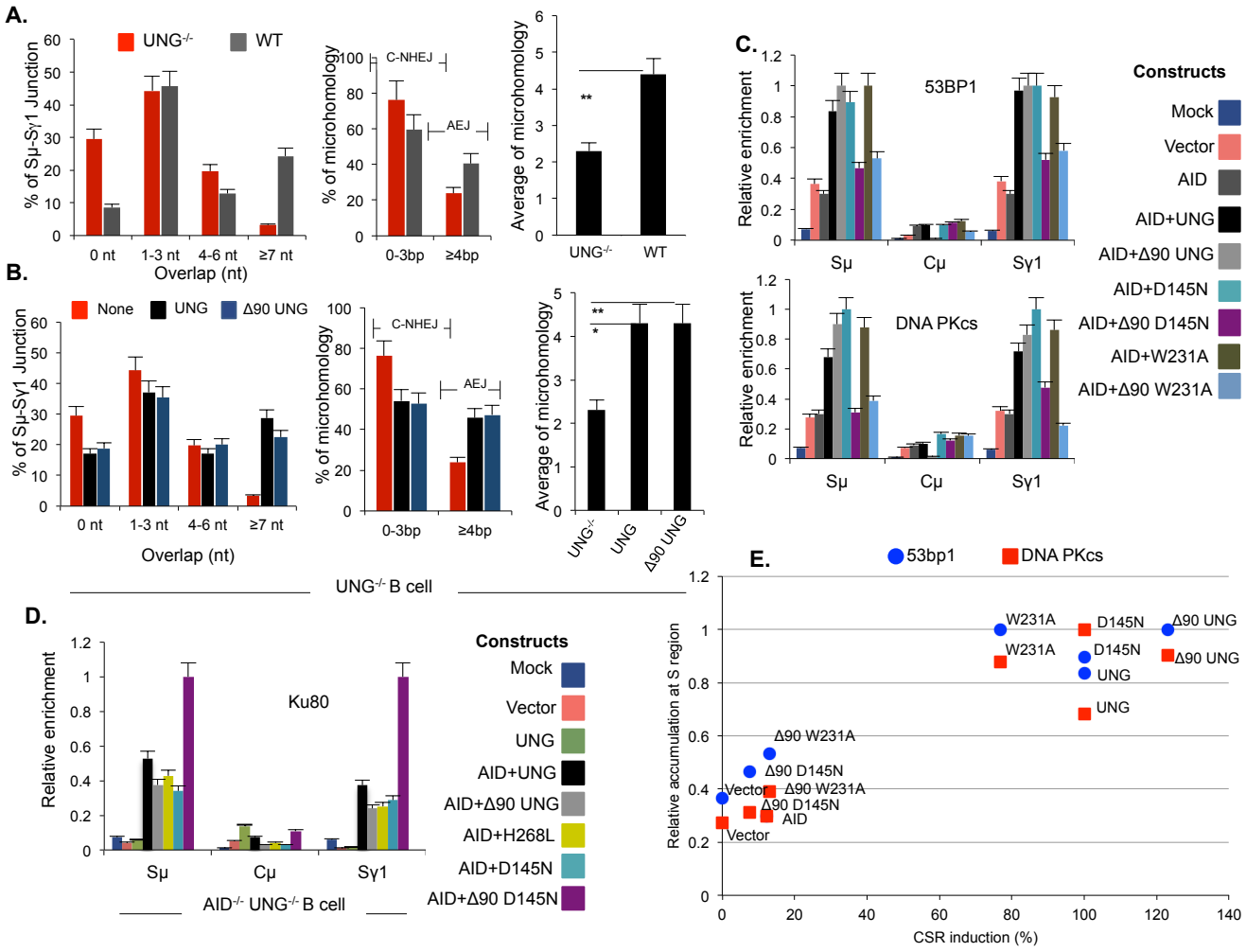

A.
WT B cell
UNG-/- B cell
AID-/-UNG-/- B cell
Mut. Freq (x10-4)
B.
GC and AT
GC
AT
C.
IgG1 (%)
Vector
UNG
Δ90 UNG
AID
AID+UNG
AID+Δ90 UNG
D.
AID-/-UNG-/- B cell
Mut. Freq. (x10-4)
NS
IgG1 (%)
UNG
AID
UNG Δ90 UNG
H268L Δ90 H268L
D145N Δ90 D145N D145N-H268L
D145E
N204V
W231A Δ90 W231A
+ AID

E.
Mutation suppression activity (%)
U removal activity (%)
UNG
Δ90 W231A
Δ90 UNG
W231A
H268L
Δ90 H268L
D145E
D145N
Δ90 D145N
D145N-H268L
N204V

F.
Mutation supression activity (%)
CSR induction activity (%)
UNG
Δ90 W231A
Δ90 UNG
W231A
H268L
Δ90 H268L
D145E
Δ90 D145N
D145N-H268L
D145N
N204V


A.
UNG
Polβ
XRCC1
FEN1
PARP1
MSH2
REV1
REV3
Relative enrichment
Sμ
Cμ
Sγ1
WT mock
UNG-/- mock
WT
UNG-/-
AID-/-
B.
UNG-/- B cell
UNG
H268L
D145N
α-GFP
α-Tubulin
IP
Mock
α-Flag
α-FEN1
Input
C.
FEN1
Polβ
REV1
REV3
Constructs
Mock
Vector
UNG
H268L
D145N
UNG-/- B cell


A.
Anchor
Eμ
Sγ3
Sγ1
Sε
Sα
3'α
Sμ
B.
GAPDH
Eμ-Sγ3
Eμ-Sγ1
Eμ-Sε
Eμ-Sα
Eμ-3'α
Sμ-Sγ3
Sμ-Sγ1
Sμ-Sε
Sμ-Sα
UNG-/-
WT
C.
Crosslinking frequency
Eμ-GAPDH
Sμ-GAPDH
D.
UNG-/- B cell
Vector
UNG
E.
WT mock
UNG-/- mock
WT
UNG-/-
AID-/-
Relative enrichment
53BP1
DNA PKcs
Ku80
XRCC4
NBS1
Sμ
Cμ
Sγ1


A.
UNG-/-
WT
% of Sμ-Sγ1 Junction
0 nt
1-3 nt
4-6 nt
≥7 nt
Overlap (nt)
% of microhomology
C-NHEJ
AEJ
0-3bp
≥4bp
Average of microhomology
**
UNG-/-
WT
B.
None
UNG
Δ90 UNG
% of Sμ-Sγ1 Junction
0 nt
1-3 nt
4-6 nt
≥7 nt
Overlap (nt)
% of microhomology
C-NHEJ
AEJ
0-3bp
≥4bp
Average of microhomology
**
*
UNG-/-
UNG
Δ90 UNG
UNG-/- B cell
C.
53BP1
DNA PKcs
Relative enrichment
Sμ
Cμ
Sγ1
Constructs
Mock
Vector
AID
AID+UNG
AID+Δ90 UNG
AID+D145N
AID+Δ90 D145N
AID+W231A
AID+Δ90 W231A
D.
Ku80
Relative enrichment
Sμ
Cμ
Sγ1
AID-/- UNG-/- B cell
Constructs
Mock
Vector
UNG
AID+UNG
AID+Δ90 UNG
AID+H268L
AID+D145N
AID+Δ90 D145N
E.
53bp1
DNA PKcs
Relative accumulation at S region
CSR induction (%)
W231A
D145N
Δ90 UNG
W231A
D145N
Δ90 UNG
UNG
UNG
Δ90 W231A
Δ90 D145N
Vector
Δ90 W231A
Δ90 D145N
AID
Vector

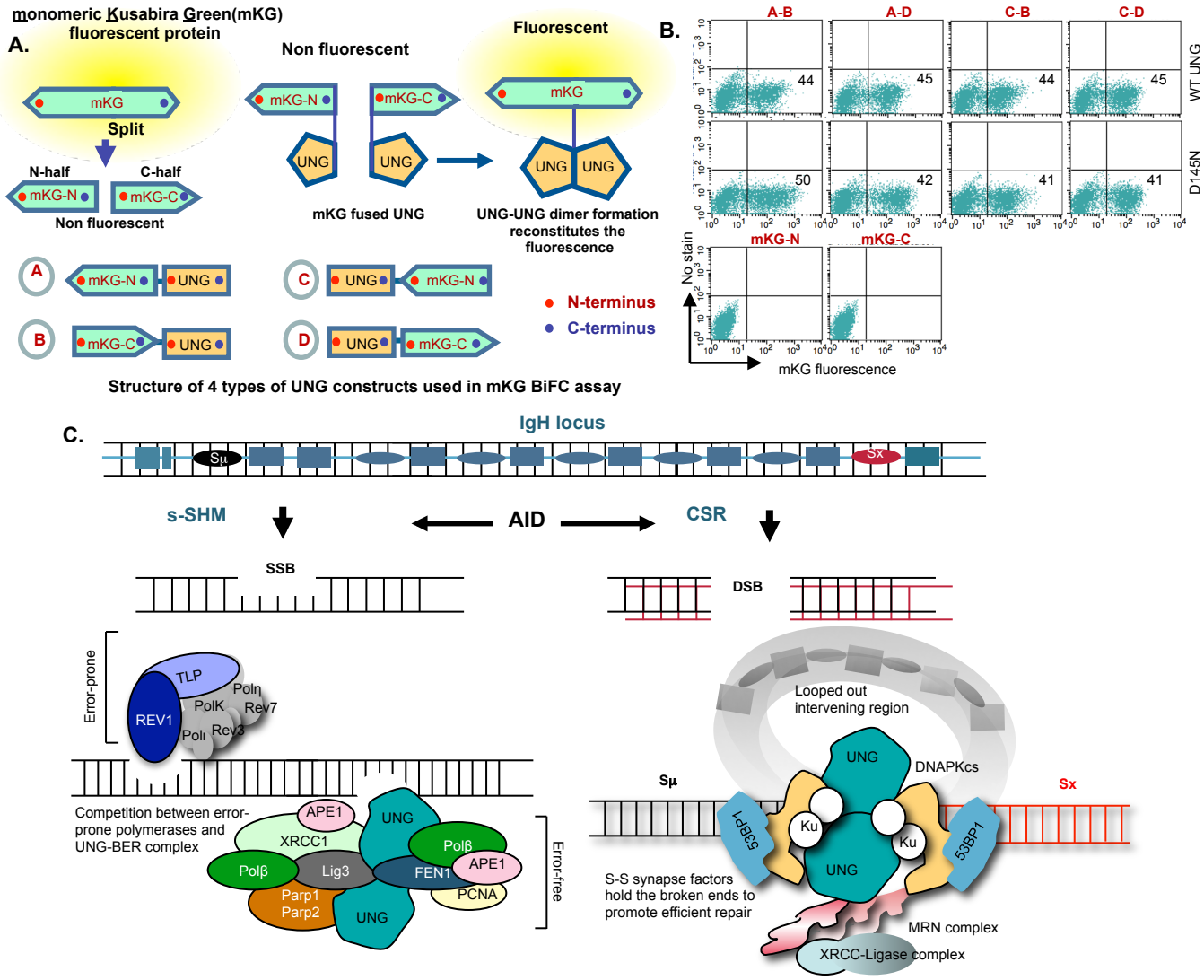

monomeric Kusabira Green(mKG)
fluorescent protein
A.
Non fluorescent
Fluorescent
mKG
Split
N-half
C-half
mKG-N
mKG-C
Non fluorescent
UNG
mKG fused UNG
UNG-UNG dimer formation
reconstitutes the
fluorescence
A
B
C
D
N-terminus
C-terminus
Structure of 4 types of UNG constructs used in mKG BiFC assay
B.
A-B
A-D
C-B
C-D
WT UNG
D145N
44
45
50
42
41
No stain
mKG fluorescence
C.
IgH locus
Sμ
Sx
s-SHM
AID
CSR
SSB
DSB
Error-prone
TLP
REV1
PolK
Polη
Rev7
Polι
Rev3
Competition between error-
prone polymerases and
UNG-BER complex
APE1
XRCC1
UNG
Polβ
Lig3
FEN1
PCNA
Parp1
Parp2
Error-free
Looped out
intervening region
DNAPKcs
Ku
53BP1
S-S synapse factors
hold the broken ends to
promote efficient repair
MRN complex
XRCC-Ligase complex

Fig. S1. Retroviral constructs, FACs profile and protein expression of UNG in WT, UNG$^{-/-}$ and AID$^{-/-}$UNG$^{-/-}$ B cell. (A) Schematic representation of retroviral constructs of UNG or AID and UNG used in the present study. (B) The diagram depicts the position (not in scale) of 5' Sμ selected for the mutation analysis. (C) Expression of UNG and Δ90 UNG after viral transduction. (D) A representative data showing IgG1 switching of UNG and UNG mutants. Numbers indicate percentages of IgG1+ cells in GFP+ gated cells. (E) Expression of UNG and different UNG mutants in AID$^{-/-}$UNG$^{-/-}$ B cell.

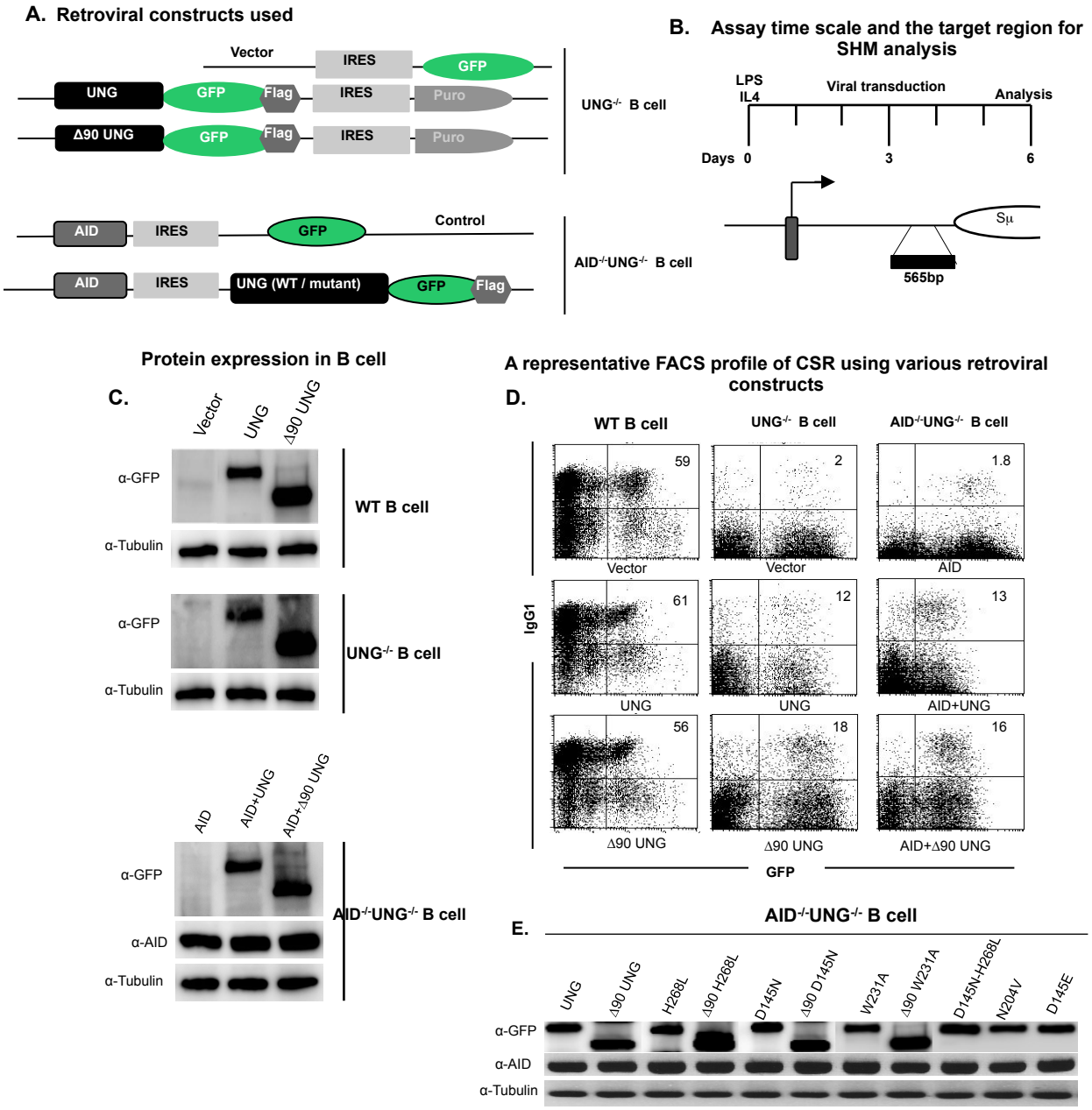
Fig. S2. Expression of BER/TLP and DNA repair factors in UNG$^{-/-}$ B cell.

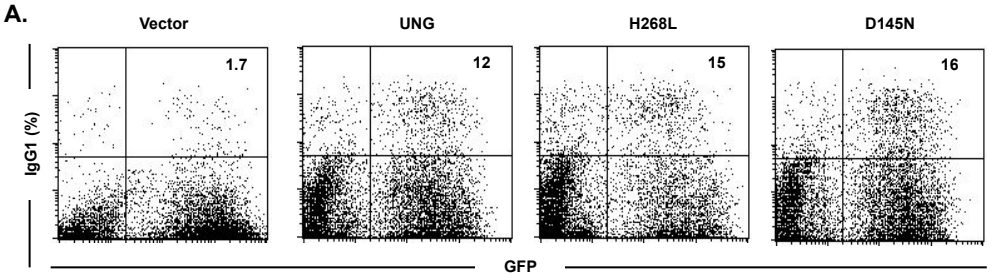
(A) CSR rescue FACs profile of UNG$^{-/-}$ B cells following expression of WT UNG and catalytic mutants of UNG. Numbers indicate percentages of IgG1+ cells in GFP+ gated cells. Activated cells were harvested on day 3 after retroviral expression and subjected to CSR assay and IP analysis as shown in Fig.2B. (B) CoIP analysis using anti-Flag (C and D) Protein expression of several repair factors analyzed by ChIP in this study (E) UNG is required for Polβ nuclear distribution and XRCC1 stabilization. Western blot of nuclear and cytoplasmic extracts of WT and UNG$^{-/-}$ splenic B cells cultured in presence of LPS + IL-4 for 3 days. Tubulin and TBP (TATA binding protein) were used as loading control for cytoplasm and nuclear fractions, respectively. (F) CSR efficiency following co-expression of AID and UNG in AID$^{-/-}$ UNG$^{-/-}$ B cells. The plot shows the percentages of IgG1+ cells in GFP+ gate and the total percentage of GFP^{+} cells. GFP expression is an indicator of the relative expression of various UNG constructs that are fused with GFP. (G) A representative Western blot showing AID and UNG expressions using anti-AID and anti-GFP antibodies, respectively.

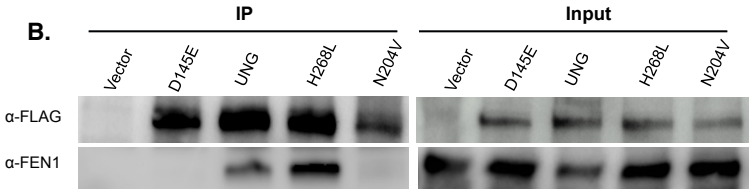
Fig. S3. UNG amino acid mutated position

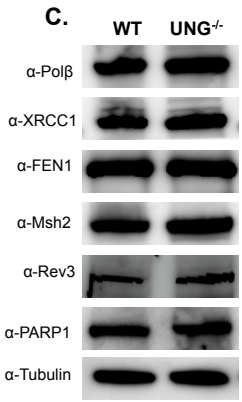
Amino acid sequence alignment of nuclear UNG of human (h) and mouse (m) origin**.** Catalytic (red) sites and WxxF motif (green) are shown in respect of a reference sequence (35) and mutants are designated accordingly. Rectangles indicate the human UNG mutants generated in the present work. The position of N-terminal truncation (arrowhead) and PCNA and RPA binding sites are also shown. Genbank acc. no. hUNG (NP_550433.1), mUNG (NP_001035781.1)

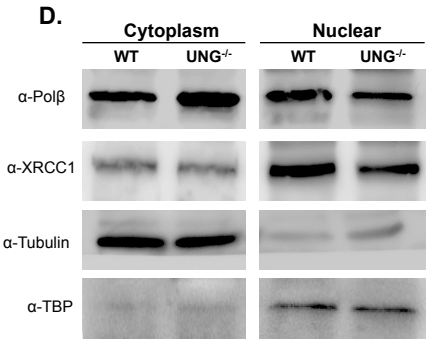
Table S1. Catalytic activity versus CSR and SHM efficiency of UNG
Table S2. Mutation analysis of 5' Sμ
Table S3. Mutation base
Table S4. Analysis of Sμ-Sγ1 recombination junctions
Table S5. List of primers set used in this study
Table S6. List of antibodies used in this study

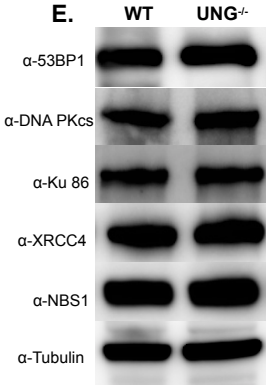

A. Retroviral constructs used
Vector
IRES
GFP
UNG
GFP
Flag
IRES
Puro
Δ90 UNG
GFP
Flag
IRES
Puro
UNG-/- B cell
AID
IRES
GFP
Control
AID
IRES
UNG (WT / mutant)
GFP
Flag
AID-/-UNG-/- B cell
B. Assay time scale and the target region for SHM analysis
LPS
IL4
Viral transduction
Analysis
Days 0
3
6
Sμ
565bp
Protein expression in B cell
C.
Vector
UNG
Δ90 UNG
α-GFP
α-Tubulin
WT B cell
α-GFP
α-Tubulin
UNG-/- B cell
AID
AID+UNG
AID+Δ90 UNG
α-GFP
α-AID
α-Tubulin
AID-/-UNG-/- B cell
A representative FACS profile of CSR using various retroviral constructs
D.
WT B cell
UNG-/- B cell
AID-/-UNG-/- B cell
IgG1
59
2
1.8
Vector
Vector
AID
61
12
13
UNG
UNG
AID+UNG
56
18
16
Δ90 UNG
Δ90 UNG
AID+Δ90 UNG
GFP
E.
AID-/-UNG-/- B cell
UNG
Δ90 UNG
H268L
Δ90 H268L
D145N
Δ90 D145N
W231A
Δ90 W231A
D145N-H268L
N204V
D145E
α-GFP
α-AID
α-Tubulin

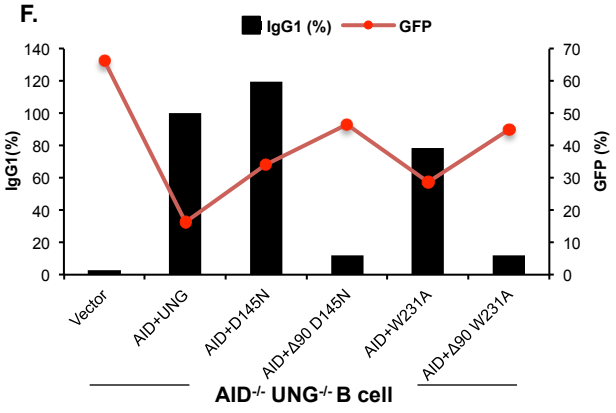

A.
Vector
UNG
H268L
D145N
1.7
12
15
16
IgG1 (%)
GFP

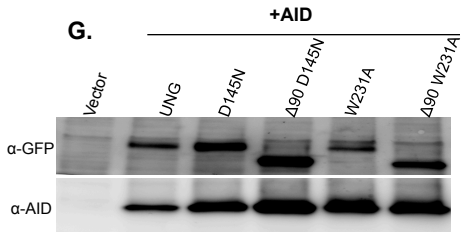

B.
IP
Input
Vector
D145E
UNG
H268L
N204V
α-FLAG
α-FEN1

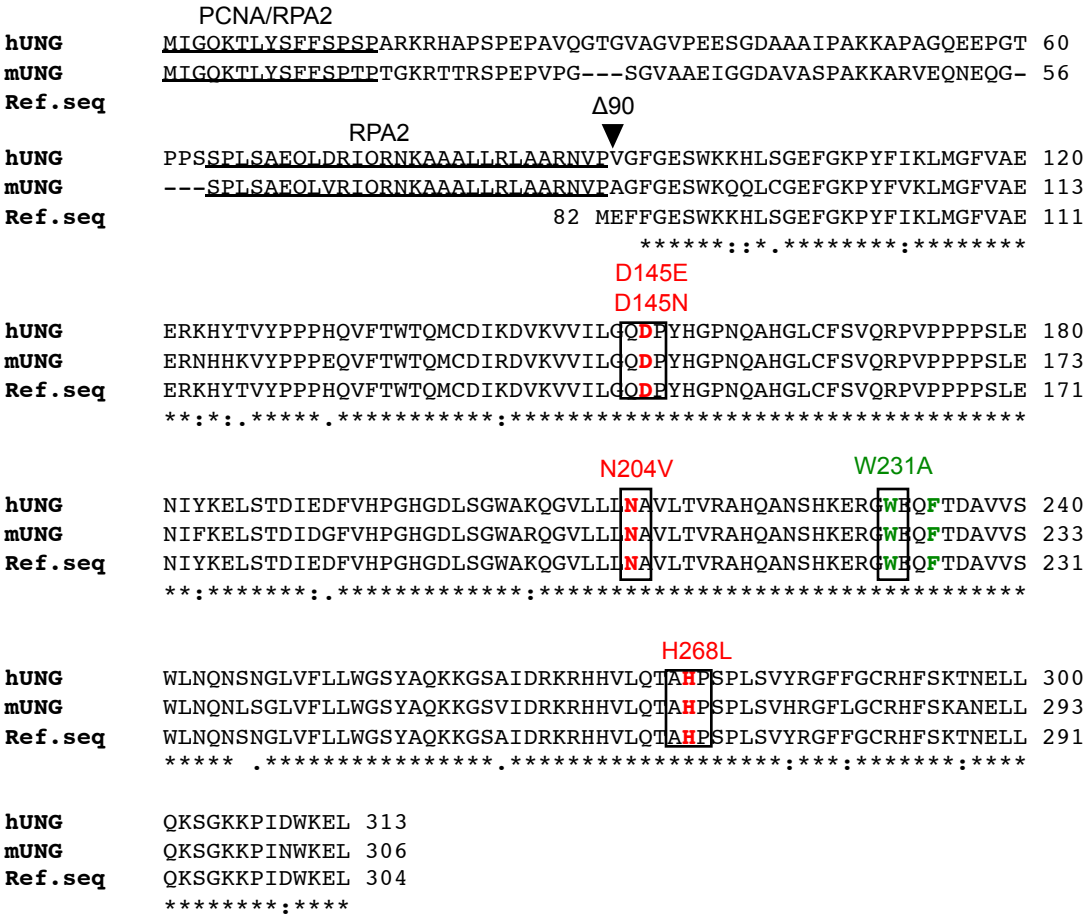

C.
WT
UNG-/-
α-Polβ
α-XRCC1
α-FEN1
α-Msh2
α-Rev3
α-PARP1
α-Tubulin


D.
Cytoplasm
Nuclear
WT
UNG-/-
α-Polβ
α-XRCC1
α-Tubulin
α-TBP


E.
WT
UNG-/-
α-53BP1
α-DNA PKcs
α-Ku 86
α-XRCC4
α-NBS1
α-Tubulin


F.
IgG1 (%)
GFP
IgG1(%)
GFP (%)
Vector
AID+UNG
AID+D145N
AID+Δ90 D145N
AID+W231A
AID+Δ90 W231A
AID-/- UNG-/- B cell


G.
+AID
Vector
UNG
D145N
Δ90 D145N
W231A
Δ90 W231A
α-GFP
α-AID


PCNA/RPA2
RPA2
Δ90
D145E
D145N
N204V
W231A
H268L
hUNG MIGQKTLYSFFSPSPARKRHAPSPEPAVQGTGVAGVPEESGDAAAIPAKKAPAGQEEPGT 60
mUNG MIGQKTLYSFFSPTPTGKRTTRSPEPVPG---SGVAAEIGGDAVASPAKKARVEQNEQG- 56
hUNG PPSSPLSAEQLDRIQRNKAAALLRLAARNVPVGFGESWKKHLSGEFGKPYFIKLMGFVAE 120
mUNG ---SPLSAEQLVRIQRNKAAALLRLAARNVPAGFGESWKQQLCGEFGKPYFVKLMGFVAE 113
Ref.seq 82 MEFFGESWKKHLSGEFGKPYFIKLMGFVAE 111
hUNG ERKHYTVYPPPHQVFTWTQMCDIKDVKVVILGQDPYHGPNQAHGLCFSVQRPVPPPPSLE 180
mUNG ERNHHKVYPPPEQVFTWTQMCDIRDVKVVILGQDPYHGPNQAHGLCFSVQRPVPPPPSLE 173
Ref.seq ERKHYTVYPPPHQVFTWTQMCDIKDVKVVILGQDPYHGPNQAHGLCFSVQRPVPPPPSLE 171
hUNG NIYKELSTDIEDFVHPGHGDLSGWAKQGVLLLNAVLTVRAHQANSHKERGWEQFTDAVVS 240
mUNG NIFKELSTDIDGFVHPGHGDLSGWARQGVLLLNAVLTVRAHQANSHKERGWEQFTDAVVS 233
Ref.seq NIYKELSTDIEDFVHPGHGDLSGWAKQGVLLLNAVLTVRAHQANSHKERGWEQFTDAVVS 231
hUNG WLNQNSNGLVFLLWGSYAQKKGSAIDRKRHHVLQTAHPSPLSVYRGFFGCRHFSKTNELL 300
mUNG WLNQNLSGLVFLLWGSYAQKKGSVIDRKRHHVLQTAHPSPLSVHRGFLGCRHFSKANELL 293
Ref.seq WLNQNSNGLVFLLWGSYAQKKGSAIDRKRHHVLQTAHPSPLSVYRGFFGCRHFSKTNELL 291
hUNG QKSGKKPIDWKEL 313
mUNG QKSGKKPINWKEL 306
Ref.seq QKSGKKPIDWKEL 304

Constructs	Enzymatic activity (%)[a]	CSR induction (%)[b]	Mutation suppression (%)[c]
WT UNG	100	100	100
Δ90 UNG	180	123	89
H268L	0.32	115.4	80
Δ90 H268L	0.1	15.4	54.2
D145N	0.04	100	15
Δ90 D145N	0.02	7.7	18
D145N-H268L	0.01	3.8	17
D145E	0.08	133	35
N204V	0.52	85	0
W231A	25	77	74.3
Δ90 W231A	20	13	89

a, Taken from previous data (28 and 35)
b and c, Taken from the present work

Splenic B cell	Expression constructs	Mutated clone/Total	Mutation (bp)	Total Sequenced	Mutation Frequency	Del/(ins)(bp)	Del(ins)/Total clone
WT	Vector	12/46	45	27120	16.6E-04	0	0/46
	UNG	4/35	6	19775	3.03E-04	11(1)	1/35
	Δ90 UNG	10/45	25	25425	10.0E-04	22	2/45
$UNG^{-/-}$	Vector	43/139	162	78535	20.6E-04	(1)	1/139
	UNG	26/137	53	77405	6.80E-04	0	0/137
	Δ90 UNG	48/136	96	76840	12.5E-04	4	1/136
$AID^{-/-}UNG^{-/-}$	AID	31/139	87	78535	11.1E-04	0	0/139
	AID+UNG	5/140	9	79100	1.14E-04	0	0/140
	AID+Δ90 UNG	4/142	18	80230	2.24E-04	3(1)	1/142
	AID+H268L	20/174	31	98310	3.15E-04	3	2/174
	AID+Δ90 H268L	25/157	51	88705	5.70E-04	0	0/157
	AID+D145N	38/166	90	93790	9.60E-04	5(1)	4/166
	AID+Δ90 D145N	34/133	70	75145	9.30E-04	0	0/133
	AID+H268L-D145N	26/94	50	53110	9.40E-04	0	0/94
	AID+D145E	19/181	67	102265	7.6E-04	10(1)	2/181
	AID+N204V	51/187	118	106220	11.1E-04	0	0/187
	AID+W231A	23/182	38	102830	3.70E-04	2	1/182
	AID+Δ90 W231A	8/96	12	54240	2.20E-04	1	1/96

Genomic DNA was isolated for mutation analysis from $IgG1^{+}GFP^{+}$ switched cells, with or without retroviral expression constructs indicated, at day 3 of post infection. A 565 bp fragment located 5' of core Sμwas subjected to mutation analysis, which corresponds to reference sequence J00440.1 (4596- 5161 bp).

WT B cell

From \ To	G	C	A	T	Total	%
G		3	8	4	15	62
C	3		4	6	13	
A	4	3		2	9	38
T	3	1	4		8	
Vector					45	

From \ To	G	C	A	T	Total	%
G		0	1	1	2	67
C	0		0	2	2	
A	2	0		0	2	33
T	0	0	0		0	
UNG					6	

From \ To	G	C	A	T	Total	%
G		0	8	2	10	72
C	3		0	5	8	
A	2	2		0	4	28
T	2	0	1		3	
Δ90 UNG					25	

$UNG^{-/-}$ B cell

From \ To	G	C	A	T	Total	%
G		0	67	3	70	94
C	1		0	82	83	
A	1	1		2	4	6
T	4	1	0		5	
Vector					162	

From \ To	G	C	A	T	Total	%
G		1	31	2	34	87
C	0		0	12	12	
A	1	1		0	2	13
T	1	3	1		5	
UNG					53	

From \ To	G	C	A	T	Total	%
G		1	56	0	57	96
C	0		1	34	35	
A	0	0		1	1	4
T	0	0	3		3	
Δ90 UNG					96	

$AID^{-/-}UNG^{-/-}$ B cell

From \ To	G	C	A	T	Total	%
G		0	53	0	53	93
C	0		0	28	28	
A	3	1		1	5	7
T	0	1	0		1	
AID					87	

From \ To	G	C	A	T	Total	%
G		0	3	1	4	67
C	1		0	1	2	
A	2	0		0	2	33
T	0	1	0		1	
AID+UNG					9	

From \ To	G	C	A	T	Total	%
G		0	5	0	5	50
C	1		0	3	4	
A	4	0		3	7	50
T	1	1	0		2	
AID+Δ90 UNG					18	

Splenic B cells	Expression constructs	Microhomologies (%)										Insertions (%)			
		Blunt	1bp	2bp	3bp	4bp	5bp	6bp	8bp	9bp	≥10bp	1bp	2-3bp	≥4bp	Total
$UNG^{-/-}$	None	18(29.5)	12(19. 7)	11(18)	4(6.6)	3(4.9)	4(6.6)	5(8.2)	0(0.0)	0(0.0)	2(3.3)	1(1.6)	0(0.0)	1(1.6)	61
WT		6(8.6)	16(22.9)	1(15.7)	5(7.1)	4(5.7)	2(2.9)	3(4.3)	3(4.3)	2(2.9)	7(10)	3(4.3)	2(2.9)	1(1.4)	70
$UNG^{-/-}$	UNG	6(17)	7(20)	6(17)	0(0)	0(0)	2(5.7)	4(11.4)	0(0)	6(17.1)	2(5.7)	0(0.0	0(0)	0(0)	35
	Δ90 UNG	16(18.8)	11(12.9)	6(7.1)	7(8.2)	6(7.1)	7(8.2)	4(4.7)	1(1.2)	2(2.4)	6(7.1)	0(0.0	2(2.4)	7(8.2)	85

Purified splenic mature B cells from WT and $UNG^{-/-}$ mice were stimulated for 3 days with LPS and IL-4, and then Sμ-Sγ1 junctions were amplified by PCR and sequenced as described in Methods. Microhomology was determined by identifying the longest region of perfect donor/acceptor identity.

Mutagenesis primers of UNG mutants

W231A (Forward)	5'-TCTCATAAGGAGCGAGGCGCGGAGCAGTTCACTGATGCA
W231A (Reverse)	5'-TGCATCAGTGAACTGCTCCGCGCCTCGCTCCTTATGAGA
D145N (Forward)	5'-GTCATCCTGGGACAGAATCCATATCATGGACCT
D145N (Reverse)	5'-AGGTCCATGATATGGATTCTGTCCCAGGATGAC
H268L (Forward)	5'-GTACTACAGACGGCTCTTCCCTCCCCTTTGTCA
H268L (Reverse)	5'-TGACAAAGGGGAGGGAAGAGCCGTCTGTAGTAC
D145E (Forward)	5'-GTCATCCTGGGACAGGAACCATATCATGGACCT
D145E (Reverse)	5'-AGGTCCATGATATGGTTCCTGTCCCAGGATGAC
N204V (Forward)	5'-CAAGGTGTTCTCCTTCTCGTCGCTGTCCTCACGGTTCGT
N204V (Reverse)	5'-ACGAACCGTGAGGACAGCGACGAGAAGGAGAACACCTTG

5' Sμ Mutaion region primers

Forward	5’-AATGGATACCTCAGTGGTTTTTAATGGTGG
Reverse	5’-GCGGCCCGGCTCATTCCAGTTCATTACAG

ChIP primers

Sμ (Forward)	5’-GTATCAAAGGACAGTGCTTAGATCCAAGGT
Sμ (Reverse)	5’-TTTCTCAATTCTGTACAGCTGTGGCCTTCC
Cμ (Forward)	5’-CAGCACCATTTCCTTCACCTGGAACTACCA
Cμ (Reverse))	5’-GGCTAGGTACTTGCCCCCTGTCCTCAGTGT
Sγ1 (Forward)	5’-AGTGTGGGAACCCAGTCAAA
Sγ1 (Reverse)	5’-GTACTCTCACCGGGATCAGC

Junction analysis primers

Sμ1 (Forward)	5'-TAGTAAGCGAGGCTCTAAAAAGCAT
Sμ2 (Forward)	5'-ATCGAATTCGCTTGAGCCAAAATGAAGTAGACT
Sγ1.1 (Reverse)	5'-CTGTAACCTACCCAGGAGACC
Sγ1.2 (Reverse)	5'-GTCGAATTCCCCCATCCTGTCACCTATA

3C assay

Eμ	3’-GGAACAATTCCACACAAAGACTC
Eα	3’-CAAGGTGTTAAGGAAAACTTGCTC
Sμ	3’-GCTGACATGGATTATGTGAGG
Sγ1	5’-CGACACTGGGCAGTTCATTTTG
Sγ3	3’-AGAGGAACCAAGTAGATAGGAC
Sε	3’-TGTGATTACCTACCTGATCCC
Sα	5’ -GCCTAGCCCAGACCATGCCA
GAPDH (Forward)	5’-AGTAGTGCGTTCTGTAGATTCC
GAPDH (Reverse)	3’-CAGTAGACTCCACGACATAC

Antibody	Catalog #	Company
UDG (FL-313)	sc:28719	Santa cruz
Pol β (18S)	ab3181	Abcam
XRCC1	ab9147	Abcam
FEN1	ab17993	Abcam
PARP1	ab6079	Abcam
REV1 (H-300)X	sc-48806 X	Santa cruz
Pol ζ (REV3L) (H-220)	sc-48814	Santa cruz
Ku86 (M-20)	sc-1485	Santa cruz
XRCC4 (G-10)	sc-365118	Santa cruz
Anti-p95 NBS1	ab32074	Abcam
53BP1	NB100-304	Novous
DNA PKcs (H-163)	sc-9051	Santa cruz
Msh2	sc-494	Santa cruz
Tubulin	CP06	Calbiochem
TBP	ab818	Abcam
GFP	A11122	Invitrogen
Flag	F-9291	Sigma-Aldrich
Rabbit IgG HRP	18-8816-33	eBioscience
Mouse IgG HRP	18-8817-33	eBioscience
AID	MAID-2(K4698)	eBioscience
Anti-Rat IgG (H+L)	712-035-153	Jakson