RNA editing of hepatitis B virus transcripts by activation-induced cytidine deaminase

Guoxin Liang[*,†], Kouichi Kitamura[†], Zhe Wang[†], Guangyan Liu[†], Sajeda Chowdhury[†], Weixin Fu[†,‡], Miki Koura[†], Kousho Wakae[†], Tasuku Honjo[*] and Masamichi Muramatsu[†]

[*]Department of Immunology and Genomic Medicine, Graduate School of Medicine, Kyoto University, Kyoto 606-8501, Japan; [†]Department of Molecular Genetics, Kanazawa University Graduate School of Medical Science, Kanazawa 920-8640, Japan, and [‡]Center of Laboratory Technology and Experimental Medicine, China Medical University, Shenyang, 11000, China

Classification: Biological Science, Immunology

To whom correspondence should be addressed.

Email: muramatu@med.kanazawa-u.ac.jp and Tel: +81-76-265-2175**,** and honjo@mfour.med.kyoto-u.ac.jp and Tel: +81-75-753-4371

Abstract

Activation-induced cytidine deaminase (AID) is essential for the somatic hypermutation (SHM) and class switch recombination (CSR) of immunoglobulin genes. The mechanism by which AID triggers SHM and CSR has been explained by two distinct models. In the DNA deamination model, AID converts cytidine bases in DNA into uridine. The uridine is recognized by the DNA repair system, which produces DNA strand breakages and point mutations. In the alternative model, RNA edited by AID is responsible for triggering CSR and SHM. However, RNA deamination by AID has not been demonstrated. Here we found that C-to-T and G-to-A mutations accumulated in the hepatitis B virus (HBV) nucleocapsid DNA when AID was expressed in HBV-replicating hepatic cell lines. AID expression caused C-to-T mutations in the nucleocapsid DNA of RNase H-defective HBV, which does not produce plus-strand viral DNA. Furthermore, the RT-PCR products of nucleocapsid viral RNA from AID-expressing cells exhibited significant C-to-T mutations, whereas viral RNAs outside the nucleocapsid did not accumulate C-to-U mutations. Moreover, AID was packaged within the nucleocapsid by forming a ribonucleoprotein complex with HBV RNA and the HBV polymerase protein. The encapsidation of the AID protein with viral RNA and DNA provides an efficient environment for evaluating AID's RNA and DNA deamination activities. A *bona fide* RNA-editing enzyme, APOBEC1, induced a similar level of C-to-U mutations in nucleocapsid RNA as AID. Taken together, the results indicate that AID can deaminate the nucleocapsid RNA of HBV.

¥body

Introduction

Antigen-induced immunoglobulin diversification is the basis of antibody memory, which is critical for effective vaccination. Activation-induced cytidine deaminase (AID) is required for somatic hypermutation (SHM) and class switch recombination (CSR), both of which are initiated by DNA cleavage induced by AID (1-4). The molecular mechanism by which AID triggers DNA cleavage specifically in immunoglobulin loci is a key question for understanding acquired immunity. There are two hypotheses for explaining how AID initiates DNA cleavage at the immunoglobulin locus: the DNA deamination model and the RNA editing hypothesis.

The DNA deamination model is based on the observation that AID induces mutations in the genomic DNA of *E. coli* and deaminates dC in single-stranded DNA *in vitro* (5-8). The resulting dU/dG mismatches are proposed to be recognized by enzymes in the base excision repair pathway, which cleave the DNA phosphodiester bond. However, it has not been directly demonstrated that AID generates dU specifically in the immunoglobulin locus. By contrast, in the RNA editing hypothesis, AID deaminates RNA, and the edited RNA is involved in DNA cleavage at the immunoglobulin genes (4, 9). This model was initially based on the structural similarity of AID with apolipoprotein B mRNA editing catalytic polypeptide 1 (APOBEC1), which is a *bona fide* RNA-editing enzyme (3, 4). Subsequently, various AID mutants were shown to have distinct defects in either CSR or SHM, suggesting that AID has at least two functions: one (DNA cleaving activity) shared by SHM and CSR and the other (DNA end-repairing activity) specific to CSR (9). The latter activity is dependent on the translation of a new protein (10). In addition, AID's

C-terminal region interacts with poly A-containing RNA (11). However, neither RNA deamination activity nor a target RNA have been demonstrated for AID.

HBV is a small DNA virus whose replication depends on reverse transcription (Fig. S1). To study the deaminase activity of AID against HBV, we used an *in vitro* model of HBV viral replication, in which an HBV replicon plasmid is transfected into a human hepatocyte cell line such as HepG2 or Huh7. The HBV replicon plasmid carries the full viral genomic sequence with an additional epsilon (ε) sequence (Fig. S2). After transfection, the replicon plasmid transcribes all the viral genes necessary for its replication, including pregenomic (pg) RNA and the mRNAs for viral proteins (P, core, X, and S) (Fig. S1A, left). The core protein encapsidates pgRNA to form a nucleocapsid, where the P protein reverse transcribes pgRNA to produce minus-strand DNA (Fig. S1B and C). The P protein digests the pgRNA in the RNA/DNA hybrid through its RNase H activity (Fig. S1D), and synthesizes plus-strand DNA to generate relaxed circular (RC) DNA in the nucleocapsid (Fig. S1E). Finally, the nucleocapsid is secreted as a virion after acquiring surface proteins. A small fraction of the nucleocapsid is transferred to the nucleus, where RC-DNA is released from the nucleocapsid and converted to covalently closed circular (ccc) DNA. In HBV infections *in vivo*, the cccDNA accumulates and remains in the nucleus of hepatocytes, where it serves as a transcription template for all the viral RNAs. The X gene encodes a protein that is not essential for viral replication, but may play a role in the development of hepatocellular carcinomas (12, 13).

AID belongs to the APOBEC (A) protein family. In humans, this family has at least 11 members: AID, A1, A2, A3A, A3B, A3C, A3DE, A3F, A3G, A3H, and A4 (14, 15). Recently, A3G was reported to induce hypermutation in the HBV nucleocapsid DNA. A3G is encapsidated with and hypermutates viral DNA (16). Rosler *et al*.

reported that the overexpression of A1 results in hypermutation not only in the viral DNA but also in the viral RNA of HBV (17). Bishop *et al*. found that C-to-U mutations accumulated in HIV-1 viral RNAs produced in the presence of A1 (18). More recently, Vartanian *et al*. suggested that AID has HBV hypermutation activity at a comparable level to that of A3G *in vitro* (19). Since A1 deaminates both viral RNA and DNA, we were interested in whether AID also deaminates HBV RNA.

In this study, using the *in vitro* HBV replication model, we found that the AID protein associates physically with the P protein and HBV RNA. Furthermore, AID deaminates C in viral RNA and causes C-to-U mutations in nucleocapsid RNA. These data demonstrate that AID can directly deaminate C in viral RNA.

Results

AID caused G-to-A and C-to-T hypermutation in RC-DNA

To examine AID-induced hypermutation in the HBV viral genome, a human AID expression vector and an HBV replicon plasmid (pHBV1.5 (20)) (Fig. S2A) were transfected into a human hepatic cell line (HepG2). Transfection with the HBV replicon plasmid initiated HBV viral replication and reproduced the intracellular viral life cycle in the HepG2 cells. Three days after transfection, the cells were lysed, treated with DNase I to eliminate residual transfected plasmids, and the nucleocapsids were then purified from the cytoplasmic lysates. The nucleocapsids were digested by proteinase K in the presence of SDS to isolate RC-DNA. The purified RC-DNA was subjected to a 3D-PCR assay of the X gene (21) because the X gene region is a good target of APOBEC deaminases (22).

In 3D-PCR, hypermutated DNA can be selectively amplified by lowering the denaturation temperature in a PCR protocol, because hypermutated DNA has a high content of A/T bases and melts at a lower temperature. Therefore, amplification of PCR target at lower denaturation temperatures indicates the presence of hypermutated DNA in a given sample. As shown in Fig. 1A, the RC-DNA samples isolated from AID-expressing cells exhibited clear PCR amplification bands at lower denaturation temperatures than the GFP-expressing control samples, consistent with a previous report (19).

To quantify the mutation frequency, samples of RC-DNA were sequenced. The X gene fragments in the RC-DNA were amplified from AID- or GFP-expressing samples using a standard PCR protocol (denaturation = 94°C) and cloned into the T vector. From each group, 55 clones were randomly selected and sequenced. The

RC-DNAs from the AID transfectants contained 15 G-to-A and 10 C-to-T mutations, whereas the RC-DNA samples from the GFP transfectants contained only a single G-to-A mutation (Fig. 1B). Similar results were obtained in two independent experiments using a hepatic cell line (Huh7) stably expressing HBV (Fig. S3) and HepG2 cell lines stably expressing AID-estrogen receptor (ER) fusion proteins (Fig. 2). Transfectants of the human hepatic cell line (HepG2) expressing AID-ER were established by retrovirus-mediated gene transduction, followed by drug selection. The AID-ER fusion protein was expressed as an inactive form that was activated by the addition of tamoxifen (4-OHT) (10). 3D-PCR of the purified RC-DNA showed amplified bands at lower temperatures (85.3° and 83.9°C) 6 hours after 4-OHT stimulation in human and mouse AID-ER transfectants (Fig. 2C). In contrast, 3D-PCR of the mock-ER transfectants did not produce signals at 85.3° and 83.9°C, indicating that stimulated AID-ER transfectants specifically induced hypermutation. To verify that AID-inducing hypermutation was dependent on its deaminase activity, we used a mutant AID (P19) that lacks deaminase activity due to a mutation in its catalytic domain (23), and found that it did not induce hypermutation (Fig. 2C).

The X gene segments were cloned from the band amplified at 83.9°C (Fig. 2C, indicated with an open box), which should contain AT-rich DNA due to deamination. Interestingly, in each of six randomly picked-up clones, the mutations were exclusively G-to-A or C-to-T, but not a mixture of both (Figs. 2D and 2E). Considering the mutation frequency (3.6×10^{-3}) in non-enriched clones (Fig. 1B), the massive accumulation of G-to-A and C-to-T mutations in the single clone must be rare events which were detected only after enrichment by 3D-PCR. Exclusive G-to-A or C-to-T hypermutation pattern suggests that the deamination reaction probably took place processively and only once, at either the RNA, the minus-strand

DNA, or the plus-strand DNA (Fig. S1). However, the efficiency of deamination of plus-strand DNA is considered to be much lower than that of the others, because the region of plus-strand DNA in RC-DNA is normally double-stranded (12, 13). Previous *in vitro* experiments showed that single-stranded DNA but not a DNA/RNA hybrid or double-stranded DNA can serve as a substrate for the AID, A1, or A3 deaminase (24, 25).

Absence of AID-induced deamination in replicon plasmid and non-encapsidated viral RNA

The frequent C-to-T mutations in AID-expressing RC-DNA samples (Fig. 1B) suggested that pgRNA may be deaminated by AID. We therefore carried out a series of experiments to examine whether pgRNA could be deaminated by AID. First, we determined the mutation load of non-encapsidated HBV transcripts. The X gene segment was amplified from HBV cDNA using the standard PCR protocol (denaturation = 94°C). If transfected replicon plasmid was directly deaminated by AID, transcripts from the plasmid should also show sign of hypermutation. However, as shown in Fig. 1C, sequencing of the cDNA derived from the AID-expressing HepG2 cells revealed no C-to-U mutations. Therefore, the transfected replicon plasmid and non-encapsidated viral RNA probably were not deaminated by AID. It is important to note that the encapsidated pgRNA was excluded by the organic RNA extraction method employed here, because it was packaged tightly in the nucleocapsid and removed with the protein fraction during the organic extraction.

AID deaminated HBV transcripts in the nucleocapsid

To examine whether the AID-induced hypermutation was due to deamination of the RC-DNA plus strand as suggested for A3G (26), an RNase H mutant replicon

(pHBV-ΔRNase H) (Fig. S2B) was used. A single amino acid replacement in the RNase H domain of the P protein causes the loss of its RNase H activity while maintaining its DNA polymerase activity (27). Thus, the viral life cycle of HBV-ΔRNase H transfectants stalls at the RNA/DNA hybrid stage in the nucleocapsid, and it cannot efficiently generate the RC-DNA plus strand (Fig. S1C) (27, 28).

Southern blot analysis of the wild-type HBV after RNase H digestion revealed two bands, corresponding to the RC-DNA and the single-stranded (minus) DNA (Fig. S4, lane 2). By contrast, the RNase H-defective virus showed only a single band corresponding to the single-stranded DNA after RNase H digestion (Fig. S4, lane 4), indicating that most of the nucleocapsid viral genome in the mutant virus stalled at the DNA/RNA hybrid step, as shown before (27, 28).

Using the HBV-ΔRNase H replicon plasmid, the mutation load of the X gene fragment in the nucleocapsid DNA was determined as described above. The control GFP expression showed no C-to-T hypermutation per 9185 bases, while the X gene fragment in the AID transfectants contained 4 C-to-T mutations per 9185 bases (Fig. 3A). The relative frequency of G-to-A hypermutation was lower than that in wild-type HBV (Fig. 1B). RNase H activity may be required for efficient hypermutation in minus-strand RC-DNA, because minus-strand DNA exists as a DNA/RNA hybrid in the RNase H-defective HBV.

The detection of C-to-T mutations in the nucleocapsid DNA from pHBV-ΔRNase H transfectants (Fig. 3A) and the absence of C-to-U mutations in HBV transcripts outside the nucleocapsid (Fig. 1C) suggested that AID may deaminate viral transcripts after encapsidation. Thus, we directly analyzed the encapsidated viral RNA. The pHBV-ΔRNase H replicon plasmid was transfected

into Huh7 cells with the AID (or GFP) expression vector, and the cells were treated with an HBV polymerase inhibitor (lamivudine) to completely inhibit leaky DNA synthesis of plus and minus-strands (Fig. S1B). Blocking the viral DNA synthesis by lamivudine is also effective for purifying the encapsidated RNA, because DNase I treatment may not completely digest the DNA strand of the RNA/DNA hybrid (29). The inhibition of RC-DNA synthesis by lamivudine was demonstrated by Southern blotting using native agarose gel electrophoresis (NAGE) with alkaline treatment (Fig. 3B).

The purified nucleocapsid RNA was reverse transcribed, and its cDNA was subjected to the 3D-PCR assay to detect the presence of C-to-U mutations in the X gene fragment from AID-, A1-, and A3G-expressing cells. The results demonstrated that the X gene fragment was amplified at a slightly lower denaturation temperature in the AID- and A1-expressing samples (Fig. 3C, right) compared with the GFP and A3G controls. To identify the mutations, the PCR fragments amplified at the lower denaturation temperature (87.8°C) were cloned and sequenced. As expected, C-to-U mutations were dominant in the A1- and AID-expressing samples (Fig. 3D), indicating that A1 and AID deaminated viral RNA. A3G did not induce 3D-PCR bands at lower temperatures than GFP, indicating the absence of viral RNA deamination.

We further confirmed that C-to-U mutations were derived from encapsidated RNA by removing the free viral RNAs outside the nucleocapsid, by immunoprecipitation (IP) using an anti-core antibody. The nucleocapsid RNA in the IP fraction and viral RNAs in the flow-through fraction were extracted after proteinase K digestion in the presence of SDS. RC-DNA and plasmid DNA, if present, were digested with DNase I before reverse transcription. The

reverse-transcribed viral RNAs were subjected to the 3D-PCR assay. The results showed that the X gene fragment was amplified at lower denaturation temperatures in the AID-expressing IP fraction but not in the GFP-expressing IP or flow-through fractions (Fig. 4A). DNA sequencing of the 87.3°C-amplified DNA band revealed C-to-U mutations almost exclusively (Figs. 4B and 4C), indicating that the nucleocapsid RNA was an RNA editing target of AID. The positions of the C-to-U mutations caused by AID were not highly specific but clustered near the tandem array of C (Fig. 4C). Taken together, these data indicate that AID deaminates nucleocapsid RNA.

AID incorporation into HBV nucleocapsids

To understand how AID interacts with nucleocapsid RNA, we examined the encapsidation of the AID protein. Since A3G is packaged into the HBV nucleocapsid in a P protein-dependent manner (30), we used A3G as a control for encapsidation. FLAG-tagged A3G, GFP, or non-tagged AID vectors were co-transfected into 293T cells together with an HBV replicon plasmid (pPB). The pgRNA's transcription in pPB was controlled by the CMV promoter (Fig. S2C) (31). Proteins associated with the nucleocapsid were then immunoprecipitated using an anti-core antibody. As shown in Fig. 5A, AID and A3G were co-immunoprecipitated with the nucleocapsid core protein. In contrast, control GFP was not immunoprecipitated, indicating that the physical association between AID and the nucleocapsid proteins was specific. Similar results were obtained in independent experiments using two hepatic cell lines (Huh7 and HepG2) (Figs 5B and 5C). We also tested whether AID physically associates with the P protein. 293T cells were co-transfected with expression vectors for the AID and FLAG-tagged P proteins. As shown in Fig. 5D, AID was co-immunoprecipitated with

the FLAG-tagged P protein, indicating that AID interacts with the P protein in the absence of the core protein and pgRNA.

Physical association of AID with HBV RNA

We previously demonstrated that the AID protein binds RNA (11). Therefore, we tested whether AID associates with viral RNA. The AID expression vector and HBV replicon plasmid were transfected into Huh7 cells. The protein-RNA complex was immunoprecipitated with an anti-AID antibody, and the IP fraction was divided into two aliquots for western blotting and RT-PCR analysis. Western blotting confirmed the successful IP of the AID protein (Fig. 5E, right). RT-PCR of the IP fraction with the AID antibody demonstrated that the HBV RNA was associated with AID, whereas HPRT mRNA was not precipitated with the anti-AID antibody (Fig. 5E, left). The IP fraction produced using control IgG did not contain HBV transcripts or HPRT mRNA (lane 3 in Fig. 5E, left). These results indicated that there was a physical association between the AID protein and the HBV RNAs. The physical association of the AID protein with the core and P proteins, as well as with the HBV RNAs, suggested that AID forms a ribonucleoprotein (RNP) complex with HBV viral proteins and RNA during nucleocapsid assembly (Fig. S1A). This study did not determine whether any of these interactions were direct or indirect, or which HBV transcripts were components of the RNP complex. However, it is likely that AID associates with viral RNA such as pgRNA through an interaction with the P protein, since the P protein binds pgRNA through the packaging signal, i.e., epsilon (12).

Endogenous AID induced HBV hypermutation in B cells

To determine whether the endogenous AID protein could induce HBV hypermutation, the HBV replicon plasmid (pPB) or a replication-defective replicon plasmid (pPB-ΔP)

(Fig. S2D) was introduced into a human EBV-transformed B cell line (BL2 cells), in which somatic hypermutation in Ig genes can be continuously induced (32). pPB-ΔP has a frameshift mutation in the spacer domain of the P protein, which cannot support viral replication (30). The GFP expression vector was transfected at the same time to monitor the transfection efficiency. Southern and Western blotting analyses using the NAGE assay demonstrated that BL2 cells could support HBV nucleocapsid formation (Fig. 6A). The production of RC-DNA in the nucleocapsid was completely dependent on the polymerase activity of the P protein (Fig. 6A, left). The transfection of pPB into *AICDA*-proficient and -deficient BL2 cells resulted in the production of RC-DNA in both, although the *AICDA*-proficient cells had a lower signal intensity for nucleocapsid DNA (Fig. 6A, right).

The RC-DNAs were then purified from both the *AICDA*-proficient and -deficient BL2 cells, and 3D-PCR analysis was performed. The X gene fragments from the *AICDA*-proficient cells were amplified at a lower denaturation temperature (83°C) than those from the *AICDA*-deficient ones (87.2°C) (Fig. 6B). The X gene segments obtained using the standard PCR protocol (94°C) were then cloned and sequenced to estimate the mutation frequency (Fig. 6C). The mutations were predominantly targeted at G and C, and C-to-T and G-to-A mutations were dominant. C-to-T mutations accumulated in the RC-DNA from the *AICDA*-proficient cells. The mutation preferences in the BL2 cells were consistent with those of the AID-overexpressing hepatic cell lines (Fig. 1B). The *AICDA*-deficient cells showed background levels of G-to-A and C-to-T mutations, which could be due to other endogenous APOBEC deaminases or to rTaq error. Semi-quantitative RT-PCR showed that BL2 cells expressed most of the AID/APOBEC deaminases (Fig. S5).

These results indicated that endogenous AID also induced HBV hypermutation in B cells.

Discussion

In this study, we expressed AID in HBV-replicating cells to evaluate the target of the AID deaminase activity. We found that AID forms an RNP complex with viral RNA as well as the core and P proteins. The X gene segment of RC-DNA accumulated C-to-T mutations, which were still observed when the plus-strand synthesis of RC-DNA was blocked. AID and the *bona fide* RNA-editing enzyme A1 caused C-to-U mutations in the viral RNA (Figs. 3 and 4). Since AID induced G-to-A mutations in the RC-DNA of HBV, it also deaminates minus-strand DNA (Figs. 1B and 2D). These results indicate that AID can deaminate both the viral RNA and the single-stranded DNA of HBV, suggesting the following series of events (Fig. S6). When AID is expressed in HBV-replicating cells, it may associate with an RNP complex containing pgRNA and the P protein. The RNP complex with AID may also attract the core protein to form a nucleocapsid, in which AID deaminates C of the viral RNA including pgRNA (Fig. S6C). The P protein reverse transcribes the edited pgRNA so that the C-to-T mutation is fixed in the minus strand of the RC-DNA. During synthesis of the minus-strand RC-DNA, AID may introduce further DNA deamination, giving rise to dG-to-dA mutations in the RC-DNA (Fig. S6D). Theoretically, the plus strand of RC-DNA can be deaminated, but AID is known to be inefficient in deaminating double-stranded DNA, which is the form in which the plus-strand DNA exists.

With current experimental approaches, it is difficult for us to demonstrate the endogenous AID activity in our hepatic cell lines (Huh7 and HepG2), possibly because the AID expression level does not reach the level to cause measurable hypermutation. Previous studies using clinical samples have reported that AID expression is induced in the liver of hepatitis patients, suggesting that AID could be

induced by chronic inflammation and may inhibit HBV replication *in vivo* (19, 33). Further study is required to determine the role of AID in the liver of hepatitis patients. It was reported that mononuclear cells such as B lymphocytes can serve as a reservoir for HBV (34-36). We demonstrated that a human B cell line, BL2, can support HBV replication after its transfection with the HBV replicon plasmid, and we found that endogenous AID expression introduces hypermutation in HBV (Fig. 6). *AICDA*-proficient BL2 cells produced less nucleocapsid DNA than did *AICDA*-deficient cells (Fig. 6A, right), suggesting that endogenous AID had an antiviral activity against HBV in the B cells, in agreement with previous studies in which AID was overexpressed in HBV-replicating hepatic cell lines (19, 37).

Materials and Methods

Hypermutation assay

The 3D-PCR was performed as previously described, with minor modifications (21). A fragment of the X gene was amplified using a nested procedure. The 3D-PCR was performed using a gradient program that generated 3°C–12°C gradients in the denaturation temperature using a thermal cycler, i.e., a MyCycler (Bio-Rad) or a MasterCycler Pro (Eppendorf). The initial PCR conditions were: 95°C for 5 min and 35 cycles of 95°C for 30 s, 50°C for 30 s, and 72°C for 30 s. The second round of 3D-PCR used 1/25 of the products from the first round as templates. The amplification conditions were as follows: 83°C–94 °C for 5 min, then 35 cycles of 83°C–94°C for 1 min, 45°C for 30 s, and 72°C for 30 s, followed by 72°C for 10 min. The C gene fragment used for sequencing was amplified by standard PCR (94°C denaturation). The amplification conditions were as follows: 94 °C for 5 min, then 35 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 40 s, followed by 72°C for 10 min.

The background mutation load due to the low fidelity of rTaq and endogenous APOBEC deaminases was estimated by sequencing the X gene fragment of RC-DNA from GFP transfectants (Figs, 1B and 3A), which showed less than 2 mutations out of 9185 nt sequenced. The mutation load of GFP transfectants was used as a negative control to determine the AID activity. rTaq error predominantly produces T-to-C and A-to-G mutations (38). For sequencing analysis, PCR fragments from 3D-PCR or standard (94°C) PCR were cloned into a T vector (Promega), and the indicated number of successful recombinant clones was selected randomly and sequenced using an ABI PRISM 3130 (Applied Biosystems). Plasmids used in this study were

described in Supplementary Table 1. The primer sequences are shown in Supplementary Table 2.

Additional Materials and Methods information provided in the Supplementary Materials and Methods.

ACKNOWLEDGMENTS

We thank Dr C.A. Reynaud, Dr H.Y. Kim, and Dr A. Takaori for providing the AID-deficient BL2 cells, pPB, and APOBEC3G vector, respectively. We also thank Drs. K. Kinoshita, N.A. Begum, M. Kobayashi, and M. Aida for critical comments, and Ms Imayasu and Ms Shimadzu for technical support. This study was supported by the Founding Program for Next Generation World-Leading Researchers, a Grant-in-Aid for Scientific Research on Priority Areas “Cancer,” and a Grant-in-Aid for Young Scientists (B) from the Ministry of Education, Science, Sports, and Culture of Japan.

The authors declare no conflicts of interest.

References

1. Muramatsu M, *et al.* (1999) Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *J Biol Chem* 274:18470-18476.
2. Muramatsu M, *et al.* (2000) Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* 102:553-563.
3. 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.

4. Honjo T, *et al.* (2012) The AID dilemma: infection, or cancer? *Adv Cancer Res* 113:1-44.
5. Di Noia JM, Neuberger MS (2007) Molecular mechanisms of antibody somatic hypermutation. *Annu Rev Biochem* 76:1-22.
6. Petersen-Mahrt S (2005) DNA deamination in immunity. *Immunol Rev* 203:80-97.
7. Ganesh K, Neuberger M (2011) The relationship between hypothesis and experiment in unveiling the mechanisms of antibody gene diversification. *FASEB J* 25:1123-1132.
8. Pavri R, Nussenzweig M (2011) AID Targeting in Antibody Diversity. *Advances in Immunology* 110:1-26.
9. Kobayashi M, *et al.* (2009) AID-induced decrease in topoisomerase 1 induces DNA structural alteration and DNA cleavage for class switch recombination. *Proc Natl Acad Sci U S A* 106:22375-22380.
10. Doi T, Kinoshita K, Ikegawa M, Muramatsu M, Honjo T (2003) De novo protein synthesis is required for the activation-induced cytidine deaminase function in class-switch recombination. *Proc Natl Acad Sci U S A* 100:2634-2638.
11. Nonaka T, *et al.* (2009) Carboxy-terminal domain of AID required for its mRNA complex formation in vivo. *Proc Natl Acad Sci U S A* 106:2747-2751.
12. Nassal M (2008) Hepatitis B viruses: reverse transcription a different way. *Virus Res* 134:235-249.
13. Ngui SL, Hallet R, Teo CG (1999) Natural and iatrogenic variation in hepatitis B virus. *Rev Med Virol* 9:183-209.
14. Harris R, Liddament M (2004) Retroviral restriction by APOBEC proteins. *Nat Rev Immunol.* 4:868-877.
15. Hamilton CE, Papavasiliou FN, Rosenberg BR (2010) Diverse functions for DNA and RNA editing in the immune system. *RNA Biol* 7.
16. Prochnow C, Bransteitter R, Chen XS (2009) APOBEC deaminases-mutases with defensive roles for immunity. *Sci China C Life Sci* 52:893-902.
17. Rosler C, *et al.* (2005) APOBEC-mediated interference with hepadnavirus production. *Hepatology* 42:301-309.
18. Bishop KN, Holmes RK, Sheehy AM, Malim MH (2004) APOBEC-mediated editing of viral RNA. *Science* 305:645.
19. Vartanian JP, *et al.* (2010) Massive APOBEC3 Editing of Hepatitis B Viral DNA in Cirrhosis. *PLoS Pathog* 6:e1000928.
20. Bruss V, Ganem D (1991) The role of envelope proteins in hepatitis B virus assembly. *Proc Natl Acad Sci U S A* 88:1059-1063.
21. Bonvin M, *et al.* (2006) Interferon-inducible expression of APOBEC3 editing enzymes in human hepatocytes and inhibition of hepatitis B virus replication. *Hepatology* 43:1364-1374.
22. Xu R, *et al.* (2007) Association of human APOBEC3 cytidine deaminases with the generation of hepatitis virus B x antigen mutants and hepatocellular carcinoma. *Hepatology* 46:1810-1820.
23. Ta VT, *et al.* (2003) AID mutant analyses indicate requirement for class-switch-specific cofactors. *Nat Immunol* 4:843-848.
24. Yu Q, *et al.* (2004) Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. *Nat Struct Mol Biol* 11:435-442.

25. Pham P, Bransteitter R, Goodman MF (2005) Reward versus risk: DNA cytidine deaminases triggering immunity and disease. *Biochemistry* 44:2703-2715.
26. Suspene R, *et al.* (2005) Extensive editing of both hepatitis B virus DNA strands by APOBEC3 cytidine deaminases in vitro and in vivo. *Proc Natl Acad Sci U S A* 102:8321-8326.
27. Radziwill G, Tucker W, Schaller H (1990) Mutational analysis of the hepatitis B virus P gene product: domain structure and RNase H activity. *J Virol* 64:613-620.
28. Gerelsaikhan T, Tavis JE, Bruss V (1996) Hepatitis B virus nucleocapsid envelopment does not occur without genomic DNA synthesis. *J Virol* 70:4269-4274.
29. Sutton D, Conn G, Brown T, Lane A (1997) The dependence of DNase I activity on the conformation of oligodeoxynucleotides. *Biochem J.* 15:481-486.
30. Nguyen DH, Hu J (2008) Reverse transcriptase- and RNA packaging signal-dependent incorporation of APOBEC3G into hepatitis B virus nucleocapsids. *J Virol* 82:6852-6861.
31. Kim HY, *et al.* (2004) Oligomer synthesis by priming deficient polymerase in hepatitis B virus core particle. *Virology* 322:22-30.
32. Faili A, *et al.* (2002) AID-dependent somatic hypermutation occurs as a DNA single-strand event in the BL2 cell line. *Nat Immunol* 3:815-821.
33. Kou T, *et al.* (2007) Expression of activation-induced cytidine deaminase in human hepatocytes during hepatocarcinogenesis. *Int J Cancer* 120:469-476.
34. Pontisso P, Vidalino L, Quarta S, Gatta A (2008) Biological and clinical implications of HBV infection in peripheral blood mononuclear cells. *Autoimmun Rev* 8:13-17.
35. Michalak T (2000) Occult persistence and lymphotropism of hepadnaviral infection: insights from the woodchuck viral hepatitis model. *Immunol Rev.* 174:98-111.
36. Coffin CS, *et al.* (2011) Molecular characterization of intrahepatic and extrahepatic hepatitis B virus (HBV) reservoirs in patients on suppressive antiviral therapy. *J Viral Hepat.* 18:415-423.
37. Jost S, Turelli P, Mangeat B, Protzer U, Trono D (2007) Induction of antiviral cytidine deaminases does not explain the inhibition of hepatitis B virus replication by interferons. *J Virol* 81:10588-10596.
38. Keohavong P, Thilly WG (1989) Fidelity of DNA polymerases in DNA amplification. *Proc. Natl. Acad. Sci. USA* 86: 9253-9257.

Figure legends

Fig. 1. AID induced frequent C-to-T mutations in nucleocapsid DNA. HepG2 cells were co-transfected with an AID or GFP expression vector and pHBV1.5 (Fig. S2A). Three days post-transfection, the nucleocapsid DNAs were purified and subjected to the 3D-PCR assay (A) and DNA sequencing (B). The total RNA was also extracted for the mutation analysis in C. (A) The 3D-PCR assay was performed as described in Materials and Methods. (B) Mutation frequency of the X gene segments. The X gene fragments were amplified using the standard PCR protocol (94°C denaturation) and cloned into T vectors. Fifty-five clone sequences were determined. C-to-T and G-to-A mutations were significantly increased by AID expression compared with GFP expression (χ^2 test). (C) Mutation frequency of the X gene fragments amplified from reverse transcribed total RNA.

Fig. 2. Stable transfection with AID also induced HBV hypermutation. (A) Human (h) AID-ER, mouse (m) AID-ER, P19-ER, and mock-ER-HepG2 transfectants were established and transiently transfected with the pHBV1.5 replicon plasmid (Fig. S2A). Six hours after transfection, 1 μM tamoxifen (4-OHT) was added to the medium to activate the ER fusion proteins. Cells were harvested 72 h after transfection. (B) The expression of ER fusion proteins was confirmed by Western blotting. (C) The 3D-PCR assay of the purified nucleocapsid DNA obtained from A. (D) Alignment of hypermutated HBV sequences. A PCR fragment derived from the 83.9 °C denaturation temperature reaction shown in C (indicated by an open box) was excised from the agarose gel and cloned into the T vector, then six *E. coli* clones were selected randomly and their X gene segments were sequenced. The sequence

(GenBank accession number: X02763) from pHBV1.5 is shown at the top as a reference sequence. Dots in the alignment indicate identity with the reference sequence. (E) Mutation matrix of the X gene segments of six clones in D.

Fig. 3. AID and A1 induced C-to-U mutations in the viral RNA of RNase H-defective HBV. (A) Mutation matrices of the X gene segments. AID or GFP expression vectors with pHBV-ΔRNase H (Fig. S2B) were transfected into Huh7 cells, which were then cultured for 3 days. The X gene segments from the nucleocapsid DNA were amplified using the standard PCR protocol (94°C denaturation), and the mutation frequency was determined as shown in Fig. 1B. C-to-T mutations were significantly increased by AID expression compared with GFP (χ^2 test). (B) Huh7 cells were co-transfected with an AID or GFP expression vector and pHBV-ΔRNase H, then cultured for 3 days with or without lamivudine, a reverse-transcriptase inhibitor. The native agarose gel electrophoresis (NAGE) and Southern blotting were used to detect nucleocapsid DNA in the cell lysates. Protein loading for the NAGE assay was determined by Western blot for the GAPDH protein. Lamivudine treatment completely blocked the viral nucleocapsid DNA synthesis. (C) Huh7 cells were transfected with AID, A1, A3G, or GFP expression vectors, and pPB-ΔRNase H (Fig. S2E), then cultured with lamivudine for three days. After proteinase K and SDS treatment to release encapsidated RNA, the total RNA was extracted. Reverse transcription was done with (+) or without (−) reverse transcriptase (RT), and PCR was performed to amplify the X gene segments (left). The RT-PCR products served as templates for the subsequent 3D-PCR (right). The 3D-PCR showed PCR amplification at lower temperatures with cDNAs from AID- and A1-expressing samples than from GFP- and A3G-expressing samples. (D) The PCR fragments

(87.8°C) in C were excised and cloned into the T vector, and 64 and 66 independent clones were sequenced. C-to-U mutations increased more than other mutations (χ^2 test).

Fig. 4. AID deaminated the nucleocapsid RNA. (A) Huh7 cells were co-transfected with an AID or GFP expression vector and pHBV-ΔRNase H (Fig. S2B), then cultured for 3 days with lamivudine. Cytoplasmic lysates were subjected to IP with an anti-core antibody. The flow-through (FT) and immunoprecipitated (IP) fractions were harvested separately. Cytoplasmic RNA was purified from the FT and treated with DNase I. The IPs were pretreated with proteinase K and SDS, and the nucleocapsid RNA was purified. (A) RT-PCR was performed to amplify the X gene segments (left) with (+) or without (−) reverse transcriptase (RT). The RT-PCR products were subjected to 3D-PCR (right). (B) A PCR fragment from the 87.3 °C denaturation temperature reaction in A, was excised and cloned into the T vector, and 50 independent clones were sequenced. The C-to-U mutation significantly increased compared with other mutations (χ^2 test). (C) Alignment of the clones in B. The sequence (GenBank accession number: X02763) from pHBV1.5 is shown at the top as a reference sequence. Dots in the alignment indicate identity shared with the reference sequence.

Fig. 5. AID was physically associated with a ribonucleoprotein complex of HBV RNA and P protein. (A) 293T cells were transfected using FLAG-tagged A3G, GFP, or AID and an HBV replicon plasmid (pPB, Fig. S2C), and then cultured for 3 days. An anti-core antibody was used to immunoprecipitate proteins from the

cytoplasmic lysates. The immunoprecipitated proteins (IP) were blotted using the antibodies indicated to detect their physical association. The whole cell lysates (input) were also blotted in parallel. (B) Huh7 cells were co-transfected with an AID or GFP expression vector and pPB, then cultured for 3 days. Co-IP using the anti-core antibody and immunoblotting were performed as described in A. The crude cytoplasmic lysate before IP was used as the input control. (C) Double-stable cell lines expressing HBV and AID (or GFP) were established in HepG2 (see Materials and Methods). IP was performed as described in A. The crude cytoplasmic lysate before IP was used as the input control. (D) 293T cells were co-transfected with each expression vector as indicated and cultured for 3 days. Mock: pcDNA3tag1A. Anti-FLAG agarose beads were used to immunoprecipitate the FLAG-P protein. Whole-cell lysates were blotted in parallel (input). (E) Huh7 cells were co-transfected with an AID expression vector and pPB, then cultured for 3 days. Cytoplasmic lysate was prepared, and IPs (with anti-AID antibody or control IgG) were performed. The IP fractions were analyzed by RT-PCR (left) and Western blotting (right). The crude cytoplasmic lysate before IP was used as the input control. RNAs were extracted from the IP fraction and crude lysate, and RT-PCR was performed (left). RT-PCR detected the specific amplification of HBV RNA in the anti-AID IP fraction (RT+).

Fig. 6. Endogenous AID induced HBV hypermutation in B cells. BL2 (human B cell line) cells were transfected with the replicon plasmid (pPB), and nucleocapsid formation and hypermutation were determined at 3 days post-transfection. The GFP expression vector was also transfected at the same time. (A) BL2 cells were transfected using pPB or the pPB-ΔP replicon plasmid (Fig. S2C and S2D), and nucleocapsid formation in the cytoplasmic lysates was assessed using the NAGE

assay. Western blotting of GFP and GAPDH was also conducted to estimate the transfection efficiency and protein loading. P-protein-dependent nucleocapsid DNA production was detected (left) by the NAGE assay. At right, *AICDA* +/+ and −/− BL2 cells were transfected with pPB, while nucleocapsid formation and AID and GFP expression were determined the same way as on the left. (B) The nucleocapsid DNAs (in A, right side) were purified and subjected to the 3D-PCR assay. (C) Mutation matrices of the X gene segments in the nucleocapsid DNA. The PCR fragments amplified by the standard PCR protocol (94 °C denaturation) were sequenced. C-to-T and G-to-A mutations were significantly increased in *AICDA* +/+ compared with *AICDA* -/- BL2 cells (χ^2 test).

Fig. 1

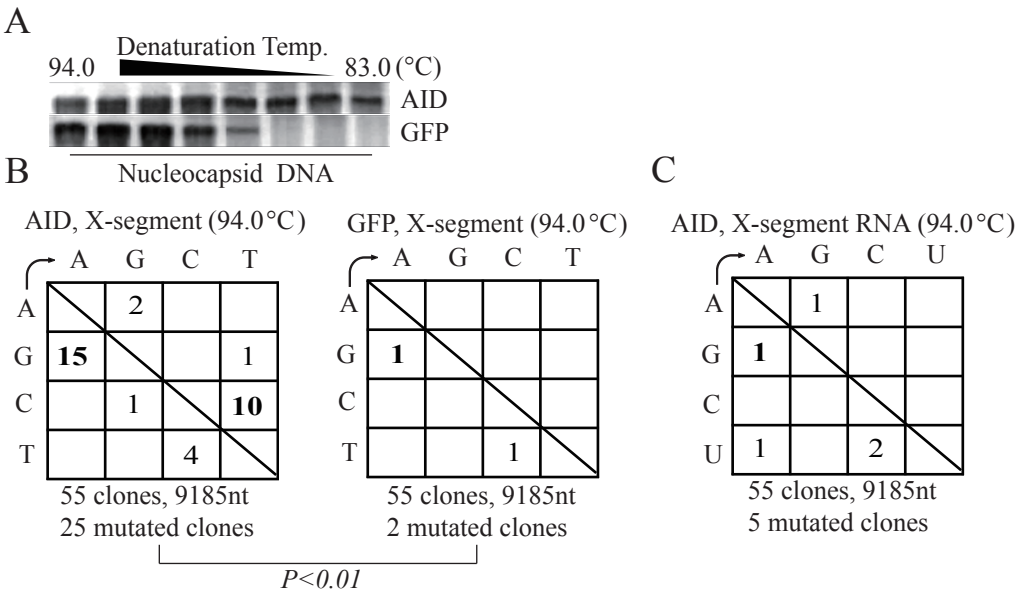

A
Denaturation Temp.
94.0
83.0 (°C)
AID
GFP
Nucleocapsid DNA
B
AID, X-segment (94.0°C)
A G C T
A 2
G 15 1
C 1 10
T 4
55 clones, 9185nt
25 mutated clones
GFP, X-segment (94.0°C)
A G C T
G 1
T 1
55 clones, 9185nt
2 mutated clones
P<0.01
C
AID, X-segment RNA (94.0°C)
A G C U
A 1
G 1
U 1 2
55 clones, 9185nt
5 mutated clones

Fig. 2

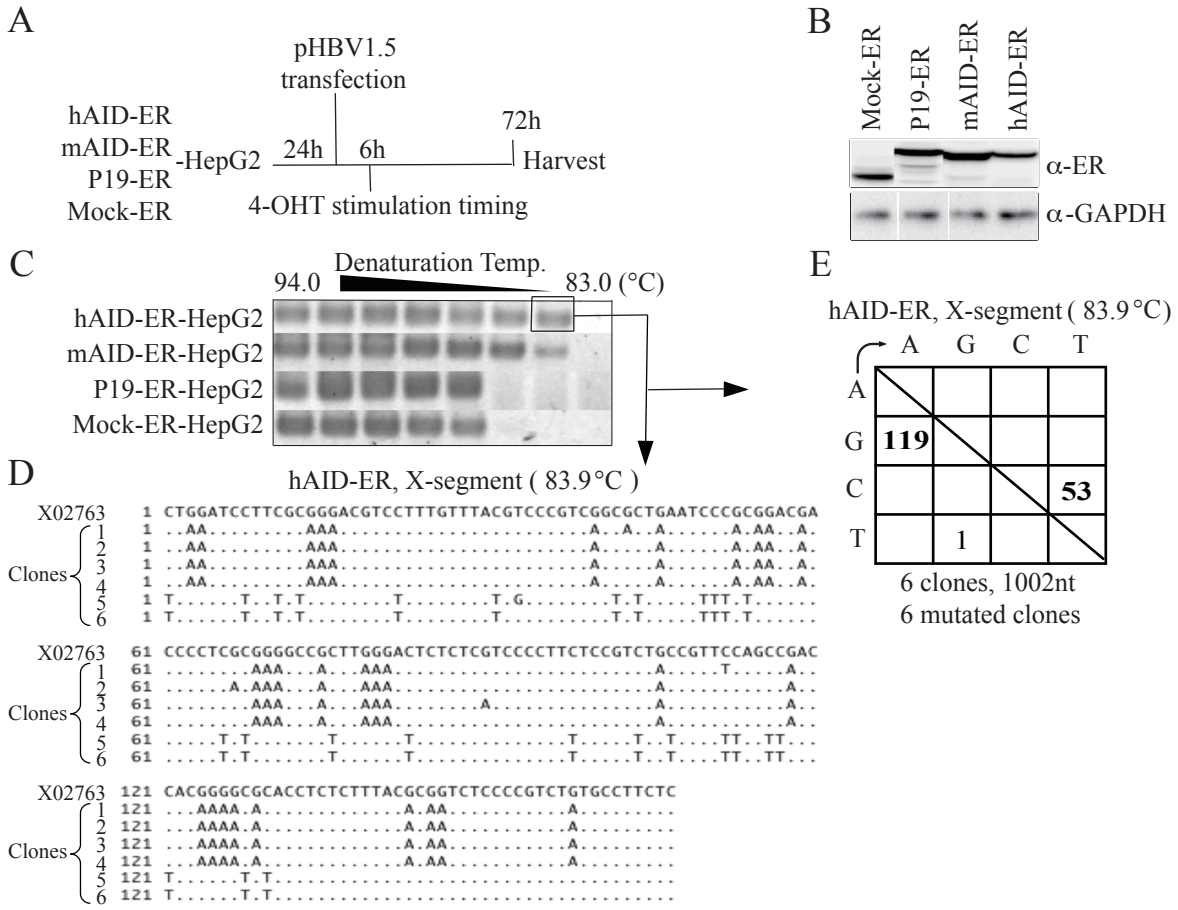

A
pHBV1.5
transfection
hAID-ER
mAID-ER
P19-ER
Mock-ER
-HepG2
24h
6h
72h
Harvest
4-OHT stimulation timing
B
Mock-ER
P19-ER
mAID-ER
hAID-ER
α-ER
α-GAPDH
C
Denaturation Temp.
94.0
83.0 (°C)
hAID-ER-HepG2
mAID-ER-HepG2
P19-ER-HepG2
Mock-ER-HepG2
D
hAID-ER, X-segment (83.9 °C)
X02763
Clones
E
hAID-ER, X-segment (83.9 °C)
A G C T
119
53
1
6 clones, 1002nt
6 mutated clones

Fig. 3

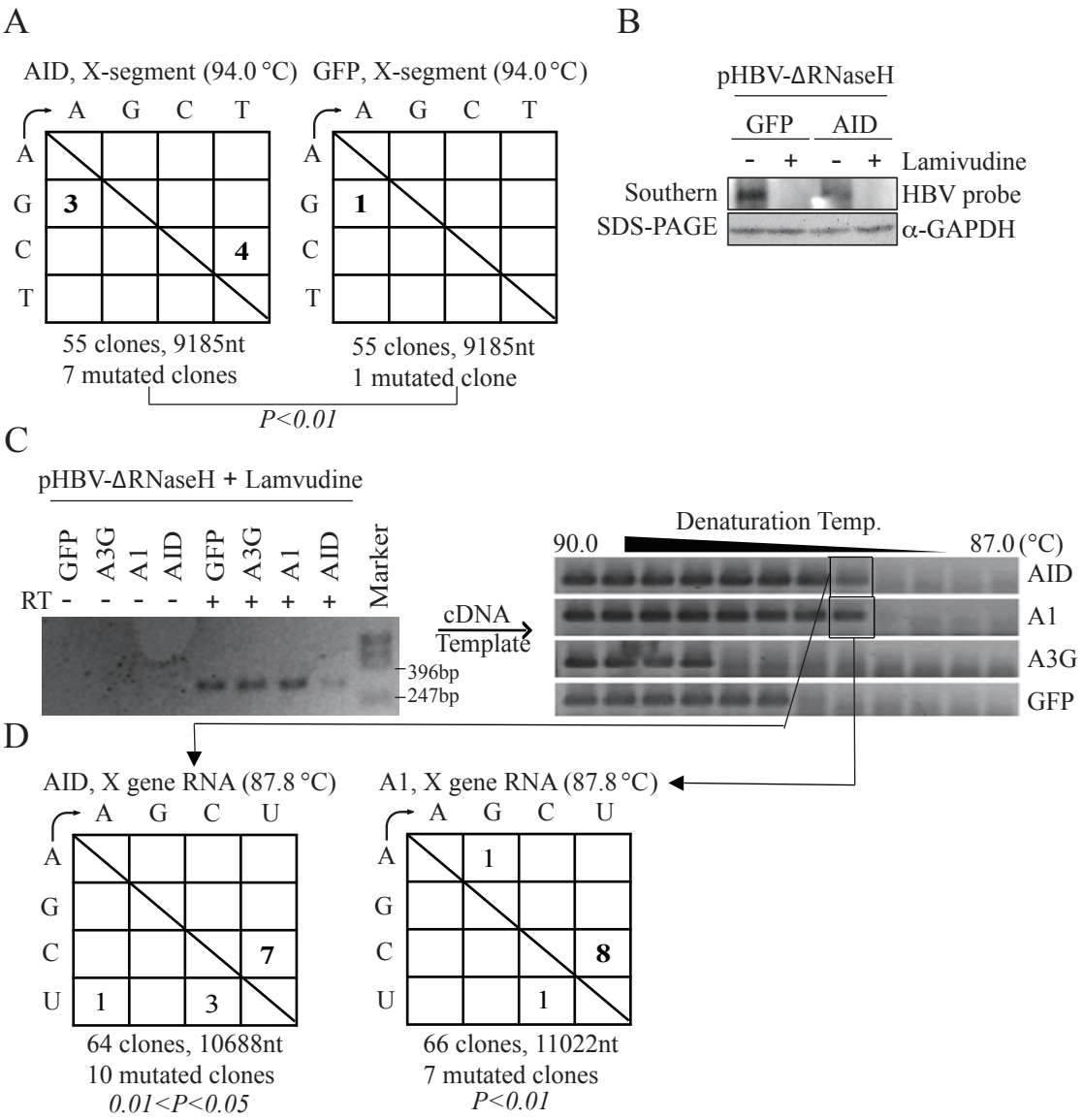

A
AID, X-segment (94.0 °C)
GFP, X-segment (94.0 °C)
A G C T
3
4
1
55 clones, 9185nt
7 mutated clones
55 clones, 9185nt
1 mutated clone
P<0.01
B
pHBV-ΔRNaseH
GFP AID
- + - + Lamivudine
Southern HBV probe
SDS-PAGE α-GAPDH
C
pHBV-ΔRNaseH + Lamvudine
GFP A3G A1 AID GFP A3G A1 AID Marker
RT - - - - + + + +
396bp
247bp
cDNA Template
Denaturation Temp.
90.0 87.0 (°C)
AID
A1
A3G
GFP
D
AID, X gene RNA (87.8 °C)
A G C U
7
1 3
64 clones, 10688nt
10 mutated clones
0.01<P<0.05
A1, X gene RNA (87.8 °C)
1
8
1
66 clones, 11022nt
7 mutated clones
P<0.01

Fig. 4

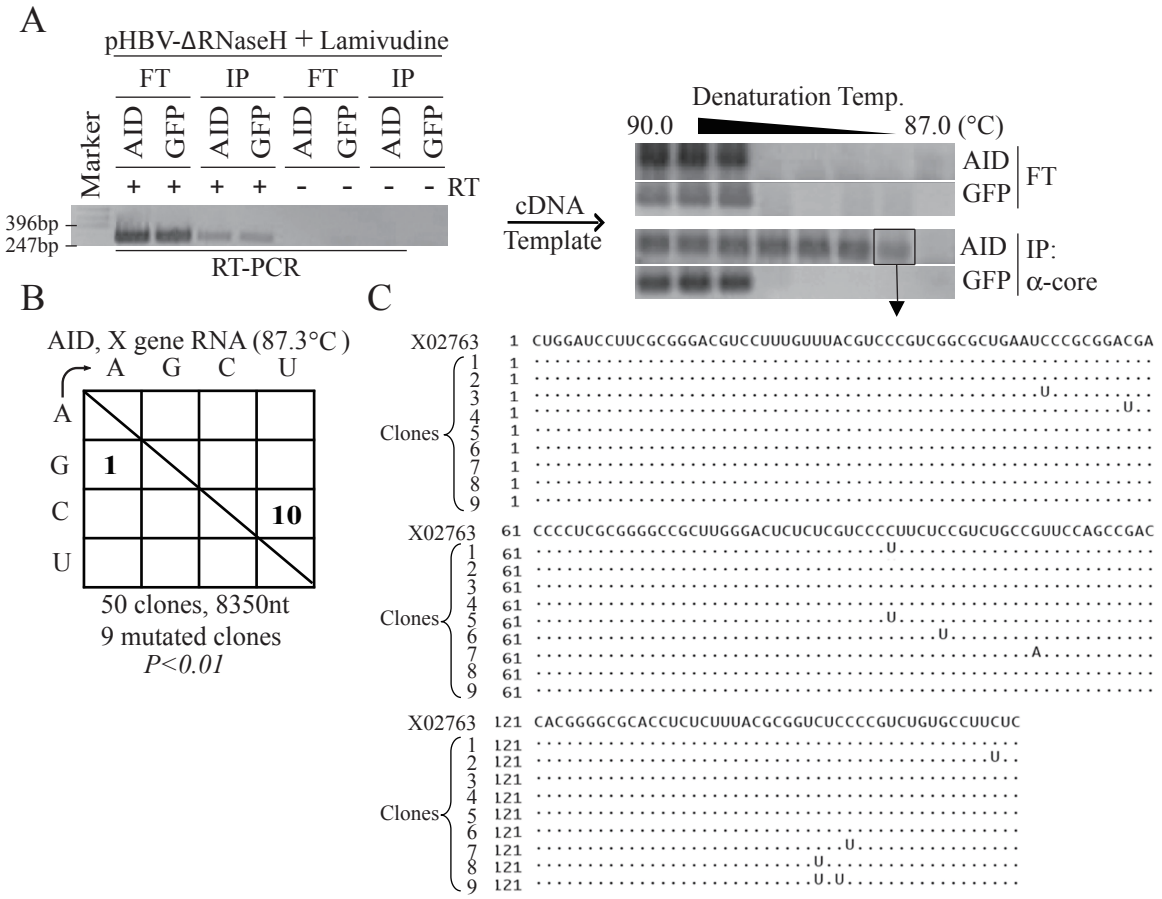

A
pHBV-ΔRNaseH + Lamivudine
FT
IP
FT
IP
Marker
AID
GFP
AID
GFP
AID
GFP
AID
GFP
+ + + + - - - - RT
396bp
247bp
RT-PCR
cDNA
Template
Denaturation Temp.
90.0
87.0 (°C)
AID
GFP
FT
AID
GFP
IP:
α-core
B
AID, X gene RNA (87.3°C)
A G C U
A
G
C
U
1
10
50 clones, 8350nt
9 mutated clones
P<0.01
C
X02763
Clones
CUGGAUCCUUCGCGGGACGUCCUUUGUUUACGUCCCGUCGGCGCUGAAUCCCGCGGACGA
CCCCUCGCGGGGCCGCUUGGGACUCUCUCGUCCCCUUCUCCGUCUGCCGUUCCAGCCGAC
CACGGGGCGCACCUCUCUUUACGCGGUCUCCCCGUCUGUGCCUUCUC

Fig. 5

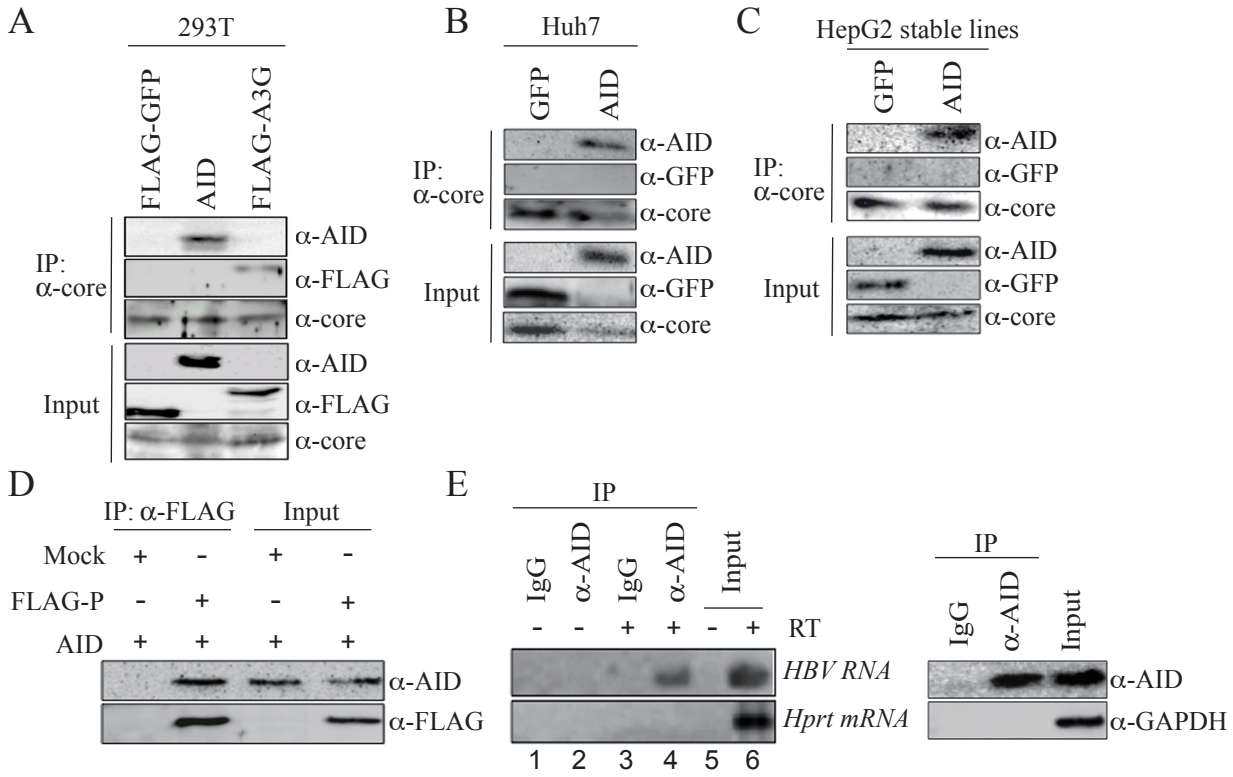
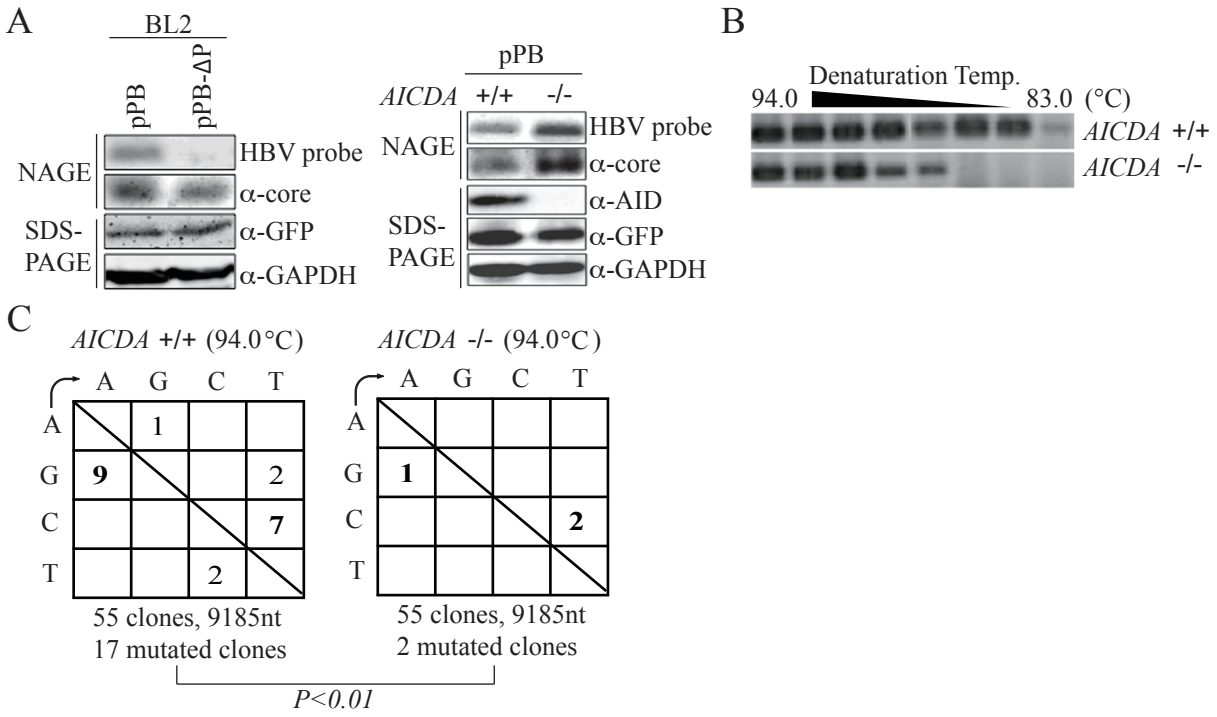

A
293T
FLAG-GFP
AID
FLAG-A3G
IP: α-core
α-AID
α-FLAG
α-core
Input
α-AID
α-FLAG
α-core
B
Huh7
GFP
AID
IP: α-core
α-AID
α-GFP
α-core
Input
α-AID
α-GFP
α-core
C
HepG2 stable lines
GFP
AID
IP: α-core
α-AID
α-GFP
α-core
Input
α-AID
α-GFP
α-core
D
IP: α-FLAG
Input
Mock + - + -
FLAG-P - + - +
AID + + + +
α-AID
α-FLAG
E
IP
IgG
α-AID
IgG
α-AID
Input
- - + + - +
RT
HBV RNA
Hprt mRNA
1 2 3 4 5 6
IP
IgG
α-AID
Input
α-AID
α-GAPDH

Fig. 6


A
BL2
pPB
pPB-ΔP
NAGE
SDS-PAGE
HBV probe
α-core
α-GFP
α-GAPDH
pPB
AICDA +/+ -/-
NAGE
SDS-PAGE
HBV probe
α-core
α-AID
α-GFP
α-GAPDH
B
Denaturation Temp.
94.0
83.0 (°C)
AICDA +/+
AICDA -/-
C
AICDA +/+ (94.0°C)
AICDA -/- (94.0°C)
A G C T
A 1
G 9 2
C 7
T 2
55 clones, 9185nt
17 mutated clones
A G C T
A
G 1
C 2
T
55 clones, 9185nt
2 mutated clones
P<0.01

Supplementary Materials and Methods

Cell culture and transfection

Cells (HepG2, Huh7, and 293T) and their transfectants were grown and maintained in Dulbecco's modified Eagle's medium (DMEM; Sigma) containing 10% fetal bovine serum, 100 U/mL penicillin, and 100 mg/mL streptomycin. Plasmids were transfected into Huh7, HepG2, and 293T cells using CalPhos (Clontech) or Fugene 6 (Roche), according to the manufacturer's instructions. *AICDA*-proficient and -deficient BL2 cells were cultured and transfected by electroporation using a Nucleofector II (Lonza), as previously described (11) .

Expression vectors

The HBV replicon plasmid (pPB) contains 1.04 copies of the HBV genomic DNA and expresses pgRNA under control of the CMV promoter (Fig. S2C) (31). The pPB-ΔP replicon plasmid (Fig. S2D) was produced through a four-base insertion at the EcoRI site in the P gene of pPB. Another HBV replicon plasmid (pHBV1.5, Fig. S2A) contains 1.5 copies of the HBV genomic DNA and expresses pgRNA under the control of a pre-core promoter (20). The pHBV-ΔRNase H replicon plasmid (Fig. S2B) was made by introducing a D737V mutation into the P gene from pHBV1.5 using a GeneTailor site-directed mutagenesis kit (Invitrogen) (27). For detailed information about the vectors, see Supplementary Table 1. Probes were labeled and Southern blotting signals were developed using an AlkPhos direct labeling system (Amersham). Further information on the vectors and primers used in this study are provided in the Supplementary Tables.

Immunoprecipitation and Western blotting

Cells were lysed using a buffer containing 50 mM Tris-HCl (pH 7.0), 20 mM NaCl, 1% NP-40, 1 mM EDTA, 2% glycerol, and a proteinase inhibitor cocktail (Roche). After centrifugation, the supernatants were incubated with the antibodies indicated and protein G sepharose (GE Healthcare). Anti-FLAG agarose beads (M2, Sigma) were used to immunoprecipitate FLAG-tagged proteins. Western blotting was performed using a standard method, and signals were detected with an LAS1000 Imager System (Fuji Film). The antibodies used in this study were as follows: rabbit anti-GAPDH (G9545, Sigma), mouse anti-FLAG (M2, Sigma), rabbit anti-core (B0586, Dako), rabbit anti-ER (sc-534, Santa Cruz), rabbit anti-GFP (Clontech, 632376), rat monoclonal anti-AID (eBioscience, 14-5959), anti-rat Igs HRP (Jackson ImmunoResearch, 712-035-153), anti-rabbit Igs-horseradish peroxidase (HRP) (ALI3404, eBiosource), and rabbit and mouse IgG Trueblot (eBioscience).

Establishment of stable transfectants

A stable line of HBV-expressing HepG2 cells was established using a standard method (2). Briefly, HepG2 cells were transfected with linearized pPB, and G418 drug selection was performed with limiting dilution. Of the resulting transfectants, a cell line with a high level of nucleocapsid DNA production was chosen from the G418-resistant clones. Human AID-ER-, mouse AID-ER-, P19-ER-, or Mock-ER-expressing HepG2 cells (Fig. 2) were produced by retrovirus-mediated gene transduction. In Fig. 5C, AID and GFP were introduced into the HBV-expressing HepG2 cells by retrovirus-mediated gene transduction. The inducible activation of the AID-ER fusion protein by tamoxifen (4-OHT, 1 μM) addition and retroviral gene transduction have been described (10, 23).

Establishment of a stable line of HBV-expressing Huh7 cells (7T7-8)

Stable transfectants were established as described previously (2). Briefly, Huh7 cells were transfected with pTETt-TAKzeo (a tetracycline activator expression vector, TET-off system) and subjected to limiting dilution with a zeocin selective medium. One clone (named 7T7) was chosen because of its good tetracycline activator expression. The 7T7 cells were transfected with an HBV replicon vector (pHBV1.2dsTETbsr), which expresses 1.2 copies of HBV genomic RNA under the control of a tetracycline promoter, and cloning was achieved by limiting dilution with blasticidin selection. pHBV1.2dsTETbsr has a stop codon in the S gene that does not affect the P gene reading frame, so Huh7 cells did not secrete virions, but accumulated HBV transcripts and nucleocapsids.

RT-PCR

Total RNA was extracted using a phenol-based RNA extraction reagent, Trisure (Bioline). The purified total RNA was digested further using DNase I (Takara) to eliminate the transfected plasmids, and purified again with Trisure. In the RT-PCR analysis, 1 μg of total RNA was treated with amplification grade DNase I (Invitrogen), then reverse transcribed with oligo(dT) using SuperScript III (Invitrogen).

Semi-quantitative RT-PCR

Total RNA was extracted using Trisure (Bioline), treated with DNase I (Takara) to eliminate transfected plasmids, and re-purified with Trisure. In the qRT-PCR analysis, 1 μg of the total RNA was treated with amplification grade DNase I (Invitrogen) and reverse transcribed with an oligo-dT or random primer using

SuperScript III (Invitrogen). The cDNA was amplified using specific primers for A3A, A3B, A3C, A3DE, A3F, A3G, A3H, ADAR1, A1, A2, and β-actin.

RNA immunoprecipitation

Cells were lysed with IP buffer (50 mM Tris-HCl [pH 7.0], 20 mM NaCl, 1% NP-40, 1 mM EDTA, 2% glycerol, and protease inhibitor cocktail). Immunoprecipitation was performed with protein G sepharose (Amersham) for 4 h at 4 °C in a chromatography column (BioRad) using anti-AID or control IgG antibodies. Protein G was washed twice each with cooled IP buffer and PBS. The immunoprecipitated RNA and proteins were eluted from the protein G sepharose using elution buffer (100 mM Tris-HCl [pH 8.0], 10 mM EDTA, and 1% SDS), and the RNA was purified further with Trisure. The immunoprecipitated RNA was treated with DNase I and re-purified with Trisure. The obtained RNAs and proteins were analyzed by RT-PCR and immunoblotting, respectively.

Native agarose gel electrophoresis (NAGE) assay

The NAGE assay has been described previously (30). Briefly, crude extracts containing HBV nucleocapsid particles were loaded onto an agarose gel and subjected to electrophoresis to separate intact nucleocapsid particles. The nucleocapsid particles in the gel were then denatured using an alkaline condition and transferred onto a nitrocellulose membrane (Roche). The nucleocapsid DNA and core protein were detected by Southern and Western blotting using a double-stranded HBV DNA probe that spanned the entire viral genome, and the anti-core antibody, respectively. The RC-DNA was purified from the cytoplasmic lysate by the following series: DNase I digestion to eliminate transfected plasmids, proteinase K and SDS digestion, phenol-chloroform extraction, and isopropanol precipitation.

Nucleocapsid pgRNA purification

Cells were treated with lamivudine (50-100μM) to inhibit HBV DNA synthesis and 3 days later were lysed with IP buffer (50 mM Tris-HCl [pH 7.0], 20 mM NaCl, 1% NP-40, 1 mM EDTA, 2% glycerol, and protease inhibitor cocktail). To isolate HBV nucleocapsids, immunoprecipitation was performed using protein G sepharose with an anti-core antibody for 4–6 h at 4°C in a chromatography column. The unbound RNA was recovered from the flow-through by Trisure purification and was treated further with DNase I. The immunoprecipitates were washed twice each with cooled IP buffer and PBS. The immunoprecipitates were eluted from protein G using elution buffer (100 mM Tris-HCl [pH 8.0], 10 mM EDTA, and 1% SDS) and treated further with proteinase K and SDS for 5–10 h at 37°C. Trisure was used to terminate the reaction directly and to purify the nucleocapsid RNA. The obtained RNAs were pretreated with DNase I and analyzed by RT-PCR or 3D-PCR.

Supplementary figure legends

Fig. S1. Schematic diagram of the HBV viral life cycle and outcomes of deamination in the RC-DNA plus strand. The HBV viral life cycle, experimental approaches to stopping the viral life cycle, possible targets for AID/APOBEC deaminases, and expected outcomes of deamination are depicted. (A) The P protein binds epsilon in pgRNA and forms an RNP complex in HBV-replicating cells. (B) The HBV RNP complex is encapsidated to form a nucleocapsid. (C) The P protein reverse transcribes nucleocapsid pgRNA to produce minus-strand DNA. This step is

blocked by lamivudine. (D) The P protein digests the pgRNA by its RNase H activity. In RNase H-defective mutant HBV, the viral life cycle stalls in the RNA/DNA-hybrid phase. Deamination in the minus-strand DNA is detected as a G-to-A mutation by DNA sequencing. (E) The P protein synthesizes plus-strand DNA to complete the production of RC-DNA. RNA deamination (in A and B) and plus-strand DNA deamination result in the accumulation of C-to-T mutations in the nucleocapsid DNA. (+): plus-strand DNA. (-): minus-strand DNA.

Fig. S2. Schematic diagram of the wild-type and mutant replicon plasmids. The partially redundant HBV genomic DNA is depicted as a black box, while the positions of the 5'- and 3'-epsilon are shown as pinheads. The corresponding open reading frames of the core (C), P, S, and X genes are shown as open boxes. The domain structure (terminal protein: TP, spacer, polymerase, RNase H) of the P protein is indicated within the P gene, separated by dotted lines. The positions of introduced point mutations are shown as an open triangle. The pPB (pPB, pPB-ΔRNase H, and pPB-ΔP) replicon plasmids transcribed viral genomic RNA (pgRNA) through the CMV promoter, whereas pHBV (pHBV1.5 and pHBV-ΔRNase H) transcribed pgRNA through the viral pre-core promoter; the positions of these promoters are indicated by arrows.

Fig. S3. AID induced HBV hypermutation in a stable HBV line. (A) A schematic representation of 7T7-8 cells (see Supplementary Materials and Methods). A stable HBV transfectant of Huh7 was established and designated 7T7-8. The induced expression of nucleocapsid DNA in 7T7-8 cells in the absence of tetracycline was demonstrated by Southern blotting of the NAGE assay (bottom). Zeo R: zeocin-resistance gene; bsd R: blasticidin-resistance gene. (B) 7T7-8 cells were

transfected with AID-GFP, HA-A3G, or GFP vectors and cultivated in the absence of tetracycline for 3 days. Nucleocapsid DNAs were purified and used as templates in the 3D-PCR assay. AID and A3G expression induced hypermutation in the RC-DNA.

Fig. S4. RNase H-defective HBV did not produce the plus-strand DNA of RC-DNA. Three days after the transfection of 293T cells with HBV replicon plasmids (pPB and pPB-ΔRNase H), the nucleic acids were purified from nucleocapsids, incubated with or without RNase H, and subjected to Southern blotting. The pPB plasmid was loaded in the right-hand lane as a control. The nucleocapsid nucleic acids from wild-type HBV produced a band corresponding to the relaxed circle (lane 1). RNase H digestion splits the signal of the relaxed circle in lane 1 into two bands (upper: RC-DNA; lower: single-stranded DNA) (lane 2). In contrast, the nucleic acids extracted from RNase H-defective HBV (lane 3) produced a signal similar to lane 1, although the signal intensity was lower than that of wild-type HBV. After digestion with RNase H *in vitro*, only the signal corresponding to single-stranded DNA was observed, indicating that plus-strand DNA synthesis was suppressed in the ΔRNase H mutant HBV. The interpretation is depicted below. Dotted lines: pgRNA. Solid lines: DNA. (+): plus-strand DNA. (-): minus-strand DNA.

Fig. S5. Expression profile of the APOBEC family in BL2 cells. Total RNA was extracted from *AICDA*+/+ and -/- BL2 cells, and semi-quantitative RT-PCR was performed to assess the expression of APOBEC deaminases. The cDNAs were serially diluted, and the expression of APOBEC deaminases was compared between the *AICDA*+/+ and -/- BL2 cells. ADAR1 and β-actin expression were used as controls.

Fig. S6. A proposed model for AID-mediated hypermutation of the HBV genome. (A) HBV RNP complex formation. (B) AID can associate with P protein and pgRNA. (C) The RNP complex containing AID is packaged with the core protein and the encapsidated AID protein deaminates pgRNA. (D) The P protein reverse transcribes the edited pgRNA to produce minus-strand DNA. During minus-strand DNA synthesis, AID further deaminates the minus-strand DNA. (E) AID also induces C-to-U conversion in the plus-strand DNA. Thus, AID causes C-to-T and G-to-A mutations in RC-DNA by acting on three forms of the viral genome. (+): plus-strand DNA. (-): minus-strand DNA.

Fig. S1

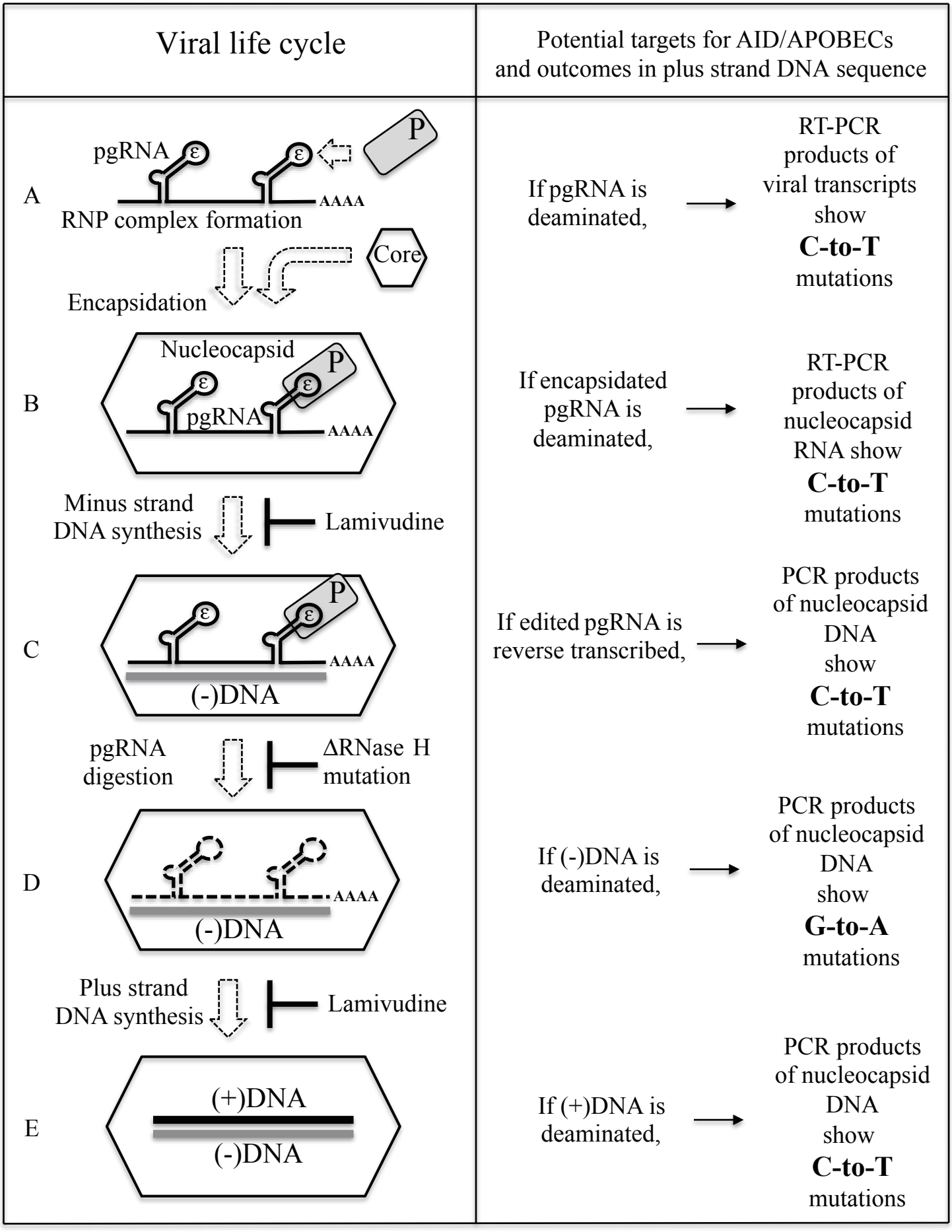

Viral life cycle
Potential targets for AID/APOBECs and outcomes in plus strand DNA sequence
A
pgRNA
ε
P
AAAA
RNP complex formation
Core
Encapsidation
If pgRNA is deaminated,
RT-PCR products of viral transcripts show C-to-T mutations
B
Nucleocapsid
pgRNA
Minus strand DNA synthesis
Lamivudine
If encapsidated pgRNA is deaminated,
RT-PCR products of nucleocapsid RNA show C-to-T mutations
C
(-)DNA
pgRNA digestion
ΔRNase H mutation
If edited pgRNA is reverse transcribed,
PCR products of nucleocapsid DNA show C-to-T mutations
D
Plus strand DNA synthesis
Lamivudine
If (-)DNA is deaminated,
PCR products of nucleocapsid DNA show G-to-A mutations
E
(+)DNA
If (+)DNA is deaminated,
PCR products of nucleocapsid DNA show C-to-T mutations

Fig. S2

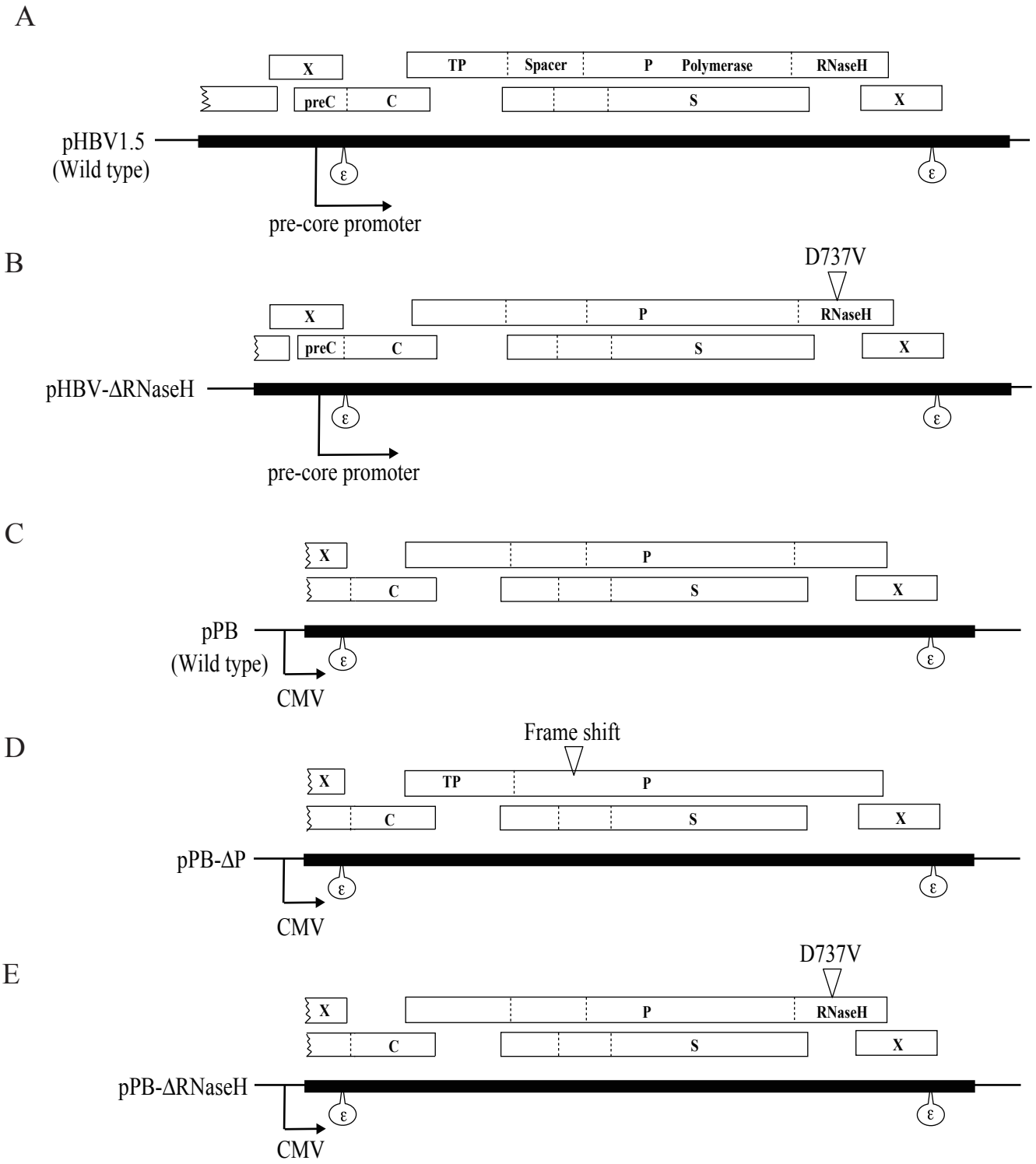

A
X
TP
Spacer
P
Polymerase
RNaseH
preC
C
S
X
pHBV1.5
(Wild type)
ε
ε
pre-core promoter
B
D737V
X
P
RNaseH
preC
C
S
X
pHBV-ΔRNaseH
ε
ε
pre-core promoter
C
X
P
C
S
X
pPB
(Wild type)
ε
ε
CMV
D
Frame shift
X
TP
P
C
S
X
pPB-ΔP
ε
ε
CMV
E
D737V
X
P
RNaseH
C
S
X
pPB-ΔRNaseH
ε
ε
CMV

Fig. S3

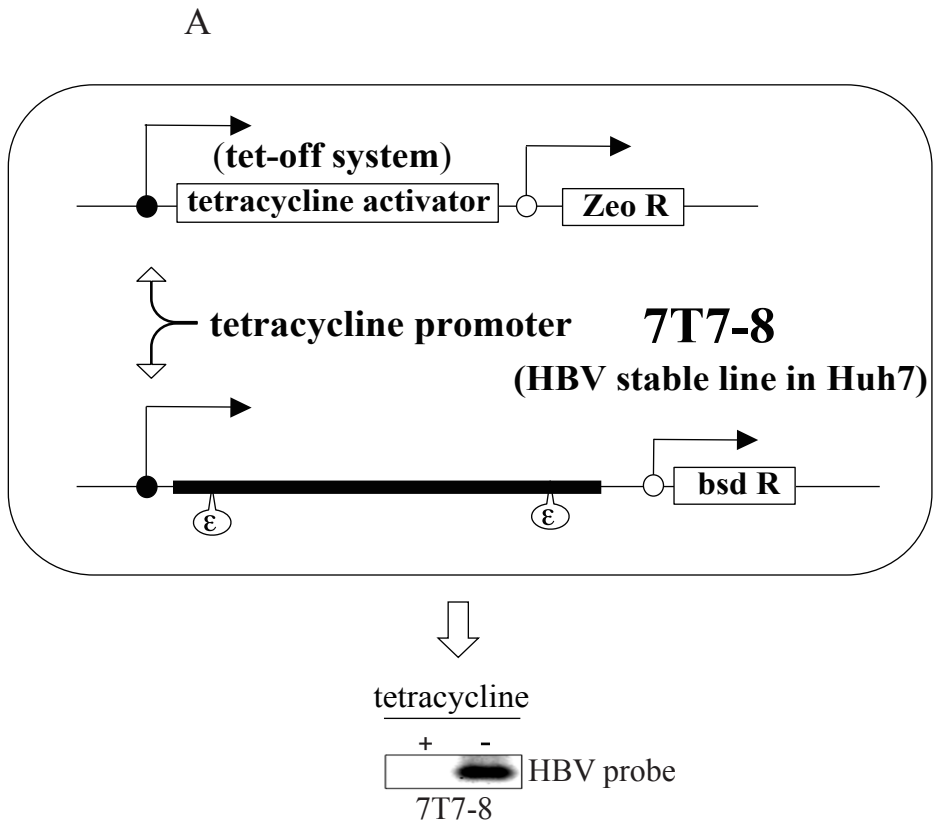

A
(tet-off system)
tetracycline activator
Zeo R
tetracycline promoter
7T7-8
(HBV stable line in Huh7)
bsd R
ε
ε
tetracycline
+
-
HBV probe
7T7-8
B
Denaturation Temp.
94.0
83.0(°C)
AID-GFP
HA-A3G
GFP
Nucleocapsid DNA

Fig. S4

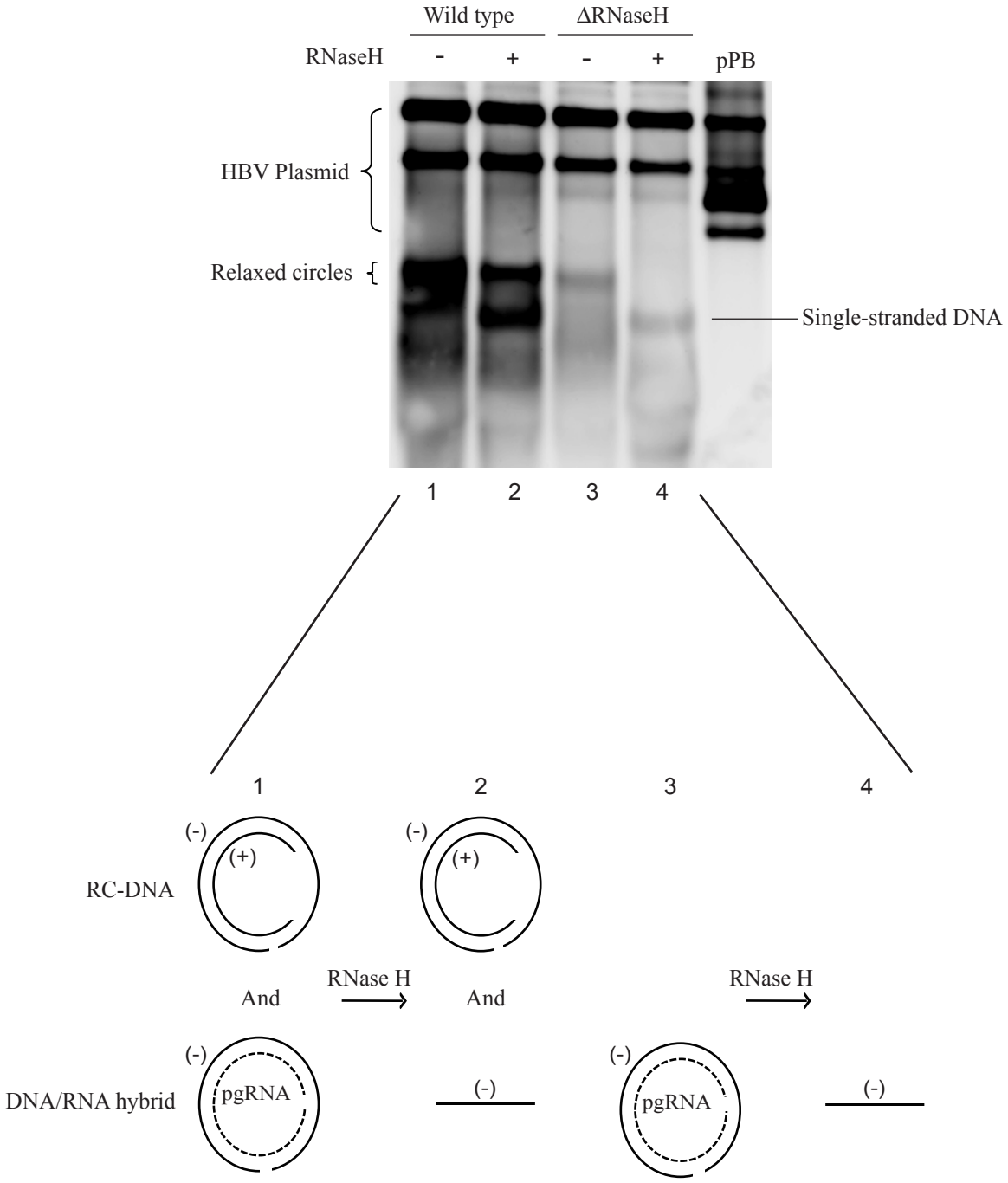

Wild type
ΔRNaseH
RNaseH
-
+
-
+
pPB
HBV Plasmid
Relaxed circles
Single-stranded DNA
1
2
3
4
1
2
3
4
RC-DNA
(-)
(+)
(-)
(+)
RNase H
And
And
RNase H
DNA/RNA hybrid
(-)
pgRNA
(-)
(-)
pgRNA
(-)

Fig. S5

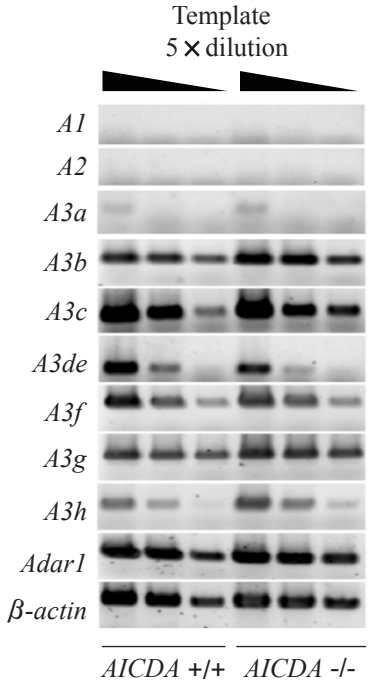

Template
5 × dilution
A1
A2
A3a
A3b
A3c
A3de
A3f
A3g
A3h
Adar1
β-actin
AICDA +/+
AICDA -/-

Fig. S6

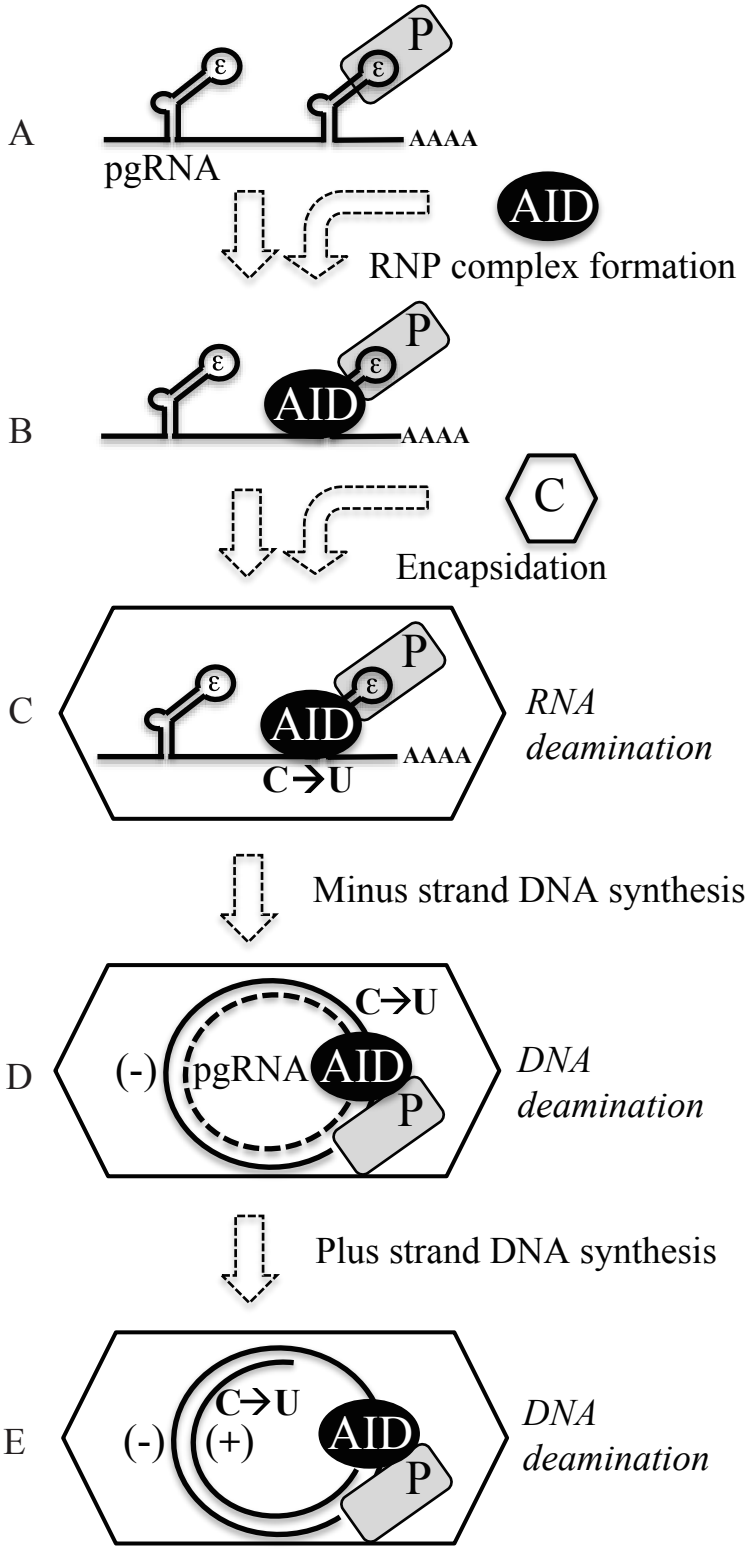

A
P
ε
ε
AAAA
pgRNA
AID
RNP complex formation
B
P
ε
ε
AID
AAAA
C
Encapsidation
C
P
ε
ε
AID
AAAA
C→U
RNA deamination
Minus strand DNA synthesis
D
C→U
(-)
pgRNA
AID
P
DNA deamination
Plus strand DNA synthesis
E
C→U
(-)
(+)
AID
P
DNA deamination

Table S1. Vector list

Name	Description
pHBV1.5	A replication-competent HBV replicon plasmid, in which the HBV pgRNA is driven by viral pre-core promoter (20).
pHBV-ΔRNaseH	A D737V mutation in the P gene has been shown to specifically delete RNaseH activity from the P protein (27). Thus a point mutation to make the D737V was introduced in pHBV1.5 using the GeneTailor site-directed mutagenesis kit (Invitrogen).
pPB	A replication-competent HBV replicon plasmid, in which the HBV pgRNA is driven by CMV promoter (31).
pPB-ΔP	An EcoRI site that locates in the middle of P gene in pPB was digested by EcoRI, treated with Klenow enzyme and ligated to creat 4 bases insertion in EcoRI site. This insertion results in a frame shift mutation in the spacer region of P gene.
pPB-ΔRNaseH	An EcoRV/SacII P gene fragment containing D737V mutation was replaced with corresponding fragment in pPB.
pCMV-AID	A DNA fragment of AID from pFBhAID (10) was replaced with one for GFP gene in pGFP2C3 (Biosignal Packard).
pCMV-hAID-GFP	previously mentioned (10).
pCMV-A1	IRES-EGFP gene in pIRES2-EGFP (Clontech) was deleted by SmaI/Bsp1407I digestion to make a CMV expression vector (designated pCMVneo). Human APOBEC1 ORF was inserted in multiple cloning site of pCMVneo.
pFLAG-P	HBV P protein ORF in pHBV1.5 was amplified by PCR and cloned into pCMV-3Tag-1A (Invitrogen) to make the vector that can express the FLAG-tagged P-protein.
hAID-pFB-ires-puro	Human AID ORF was inserted into multicloining site of retrovirus vector pFB-ires-puro. This was used for establishment of AID expressing HepG2 cells.
GFP-pFB-ires-puro	Human EGFP ORF was inserted into multicloining site of retrovirus vector pFB-ires-puro. This was used for establishment of GFP expressing HepG2 cells.
pFERp	pFB-AIDER-ires-puro retrovirus vector was described previousely (10). AID gene in pFB-AIDER-ires-puro was replaced with a synthetic DNA linker containing mutlicloning site to make a mock retrovirus vector and designated pFERp.
mock-pFERp	Two oligonucleotides (5'-AATTCTAGCACCATGGCTAAG-3', 5'-GATCCTTAGCCATGGTGCTAG-3') containing an in-frame ATG were hybridized and inserted in an EcoRI/BamHI site of the pFERp retroviral vector. After puromysin selection, this retrovirus vector stably expresses flag-ER chimerica protein. This is used as a mock control expressing flag-ER protein alone.
pFB-mAIDER-ires-puro	previously mentioned (10).
hAID-pFERp	hAID was inserted in pFERp to make a retrovirus vector expressing hAIDER-ires-puro mRNA.
p19-pFERp	p19-AID was inserted in pFERp to make a retrovirus vector exprssing p19ER-ires-puro mRNA.
pFLAG-GFP	A GFP2 ORF fragment in pGFP2N3 (Biosignal Packard) is inserted into pCMV3Tag2B (Invitrogen).
pFLAG-A3G	Human A3G ORF was cloned into pcDNA3Tag1A (Invitrogen).
pcDNA3/HA-A3G	kindly gifted by Dr. Takaori in Kyoto university .
pTETtTAK zeo	previously mentioned (2, 10)
pTet-splice-bsr	Blasticidin resistant gene expression cassette was inserted in BstXI site in pTET-splice (GIBCO).
pHBV1.2dsTet-bsr2	1.2 copy of HBV genomic DNA from pHBV1.5 which contains a stop codon in S open reading frame was inserted into multiple cloning (SalI/ClaI) of pTet-Splice-bsr.

Table S2. Primer list

Name	Sequence (5'→3')
HBV X segement 1st fwd	CGGAAATATACATCGTTTCCAT
HBV X segment 1st rev	AAGAGTYYTYTTATGTAAGACYTT
HBV X segment 2nd fwd	ATGGCTGCTARGCTGTACTGCCAA
HBV X segment 2nd rev	AAGTGCACACGGAYYGGCAGAT
HBV C segment fwd	CATGTCCCACTGTTCAAGCCT
HBV C segment rev	GAAACCACAATAGTTGCCTGATC
A3A fwd	ATGGCATTGGAAGGCATAAG
A3A rev	CAAAGAAGGAACCAGGTCCA
A3B fwd	TTCGAGGCCAGGTGTATTTCA
A3B rev	CAGAGATGGTCAGGGTGACA
A3C fwd	CAACGATCGGAACGAAACTT
A3C rev	TATGTCGTCGCAGAACCAAG
A3D fwd	ACCCAAACGTCAGTCGAATC
A3D rev	GCTCAGCCAAGAATTTGGTC
A3F fwd	GAAACACAGTGGAGCGAATG
A3F rev	GAAATGGGGCTCTGATGAAAG
A3G fwd	GGTCAGAGGACGGCATGAGA
A3G rev	GCAGGACCCAGGTGTCATTG
A3H fwd	CCCGCCTGTACTACCACTGG
A3H rev	GGGTTGAAGGAAAGCGGTTT
ADAR1 fwd	CGTGCTATTTGCTGTCGTGT
ADAR1 rev	TAGTCACGGGCAGCTTTCTT
A1 fwd	GAAGAATCGAACCCTGGGAGTTT
A1 rev	CCCAGGTGGGTAGTTGACAAA
A2 fwd	GGTGGGTCGACTCTTCATGTGG
A2 rev	GGGCCTTGTCCAGTATGATTGGT
β-Actin fwd	GACCTGACTGACTACCTCATGAAGA
β-Actin rev	GGGGCCGGACTCGTCATACTCCTGC
HPRT fwd	GCCCTGGCGTCGTGATTAGT
HPRT rev	CGAGCAAGACGTTCAGTCCTGTC

Y represents C and T.

R represents G and A.