APOBEC3G hypermutates genomic DNA and inhibits Ty1 retrotransposition in yeast

April J. Schumacher*,†, Dwight V. Nissley*, and Reuben S. Harris*†§

*Department of Biochemistry, Molecular Biology, and Biophysics and †Institute for Molecular Virology, University of Minnesota, Minneapolis, MN 55455; and ‡Basic Research Program, SAIC-Frederick, and Gene Regulation and Chromosome Biology Laboratory, National Cancer Institute, Frederick, MD 21702

Edited by Richard D. Kolodner, University of California at San Diego, La Jolla, CA, and approved May 16, 2005 (received for review March 1, 2005)

Human cells harbor a variety of factors that function to block the proliferation of foreign nucleic acid. The APOBEC3 enzyme inhibits the replication of retroviruses by deaminating nascent retroviral cDNA cytosines to uracils, lesions that can result in lethal levels of hypermutation. Here, we demonstrate that APOBEC3G is capable of deaminating genomic cytosines in Saccharomyces cerevisiae. APOBEC3G expression caused a 20-fold increase in frequency of mutation to canavanine-resistance, which was further elevated in a uracil DNA glycosylase-deficient background. All APOBEC3G-induced base substitution mutations mapped to the nuclear CAN1 gene and were exclusively C/G -> T/A transition mutations within a 5'-CC consensus. The APOBEC3G preferred sites were found on both strands of the DNA duplex, but were otherwise located in hotspots nearly identical to those found previously in retroviral cDNA. This unique genetic system further enabled us to show that expression of APOBEC3G or its homolog APOBEC3F was able to inhibit the mobility of the retrotransposon Ty1 by a mechanism that involves the deamination of cDNA cytosines. Thus, these data expand the range of likely APOBEC3 targets to include nuclear DNA and endogenous retroelements, which have pathological and physiological implications, respectively. We postulate that the APOBEC3-dependent innate cellular defense constitutes a tightly regulated arm of a conserved mobile nucleic acid restriction mechanism that is poised to limit internal as well as external assaults.

APOBEC3F | endogenous retrotransposon Ty1 | hypermutation | cytoedamination | Saccharomyces cerevisiae

The human apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3G (APOBEC3G) protein uses cytosine to uracil deamination to inhibit the replication of a variety of retroviruses, including HIV-1 (refs. 1–6; reviewed by in refs. 7–10). APOBEC3G localizes predominantly to the cytoplasm of mammalian cells (4, 11). In a retrovirus-infected cell, this localization may facilitate the incorporation of APOBEC3G into viral particles, which are released from the plasma membrane. APOBEC3G is also specifically incorporated into virions through an association with the viral Gag protein and/or viral genomic RNA (12–18). Once a retrovirus enters a cell, its genomic RNA is reverse transcribed, and during this process, APOBEC3G is capable of deaminating cDNA cytosines to uracils (C -> U). These lesions occur at such a high frequency that they ultimately inactivate the virus (causing G -> A hypermutation, as read-out on the genomic strand of the virus). APOBEC3G has several human homologs, and at least one, APOBEC3F, restricts HIV-1 infection by a similar mechanism (19–22). However, APOBEC3F and -G deaminate cytosines within different local contexts, preferring 5'-TC and 5'-CC, respectively (19). Both of the resulting mutational signatures are apparent in patient-derived HIV-1 sequences (e.g., ref. 23), suggesting that the APOBEC3-dependent restriction mechanism has a critical role in vivo.

To counteract the APOBEC3-dependent cellular defense, HIV-1 and most other related retroviruses encode a small protein, virion infectivity factor (Vif), which mediates the degradation of APOBEC3F and -G and thereby facilitates retroviral replication (reviewed in refs. 9 and 10). Vif binds APOBEC3G and mediates its polyubiquitination by a cellular E3 ligase complex containing Cullin 5 and Elongins B/C (24–26). Thus, the relative levels of Vif and APOBEC3 proteins appear to be key determinants of retroviral infection.

Although the APOBEC3-Vif interaction is likely important in primates, this relationship does not appear to be ubiquitous. First, APOBEC3 proteins appear to have evolved before Vif-encoding retroviruses (e.g., refs. 8, 27, and 28). All known mammals have APOBEC3G or at least one APOBEC3G-like protein, yet many of their retroviruses appear to lack an effective counterdefense protein such as HIV-1 Vif. For instance, rodents encode a single APOBEC3G homolog and host a variety of retroviruses, none of which appear to encode Vif or a Vif-like counterdefense. Second, many of the APOBEC family member proteins, including APOBEC3F and -G, show an extraordinarily high number of amino acid replacement mutations, indicating that they have been under a strong positive selection (27, 29). Evidence for positive selection predates by millions of years the time that HIV is predicted to have entered the human population. Thus, these observations have fueled speculation that the human APOBEC3 proteins possess another, possibly primary, cellular function (e.g., refs. 8, and 27–30).

Retroelements comprise a large proportion of the murine and human genomes (reviewed in refs. 31 and 32). Although the majority of these elements are apparently nonfunctional, their cis or trans-mobilization can contribute to a variety of genome destabilizing events including recombination and insertional mutagenesis. There are two major classes of endogenous retroelements. The first and largest is exemplified by long interspersed nucleotide elements (LINEs), which replicate upon integration via an LTR-independent mode. The second major class forms intracellular virus-like particles and replicates by using LTR sequences that flank the coding genes, much like HIV and related retroviruses. Examples of this class include Saccharomyces retrotransposons Ty1 and Ty3, Drosophila copia and gypsy elements and human endogenous retroviruses (HERVs). Although human LINE replication appears uninhibited by APOBEC3G expression (30), a role in controlling the spread of LTR-dependent retroelements remained a distinct possibility.

Here, we test this possibility by using Saccharomyces cerevisiae as a model system. We first ask whether human APOBEC3G is capable of functioning in yeast and, second, we apply this system to assess whether APOBEC3G is capable of inhibiting Ty1 retroposition. We show that APOBEC3G expression in yeast caused a mutator phenotype fully attributable to the deamination of genomic cytosines. This surprising functionality enabled us to further show that expression of APOBEC3G or -F inhibited replication of the yeast LTR-dependent retrotransposon Ty1 by a mechanism that involves cDNA cytosine deamination. These data indicate that the APOBEC3-dependent mechanism of retroelement restriction is...
highly conserved and that the range of APOBEC3 substrates may be far broader than originally anticipated.

Materials and Methods

Yeast Strains and Plasmids. See Supporting Text, which is published as supporting information on the PNAS web site.

Yeast Mutation Assays. pHybLex-Zeo, pJG4–5, pYES3-CT, and their derivatives were transformed into L40 and selected by using a synthetic complete medium containing zeacon and lacking tryptophan (SC + ZEO-TRP) (33). Several thousand viable cells from independent colonies were used to inoculate 2.5 ml SC + GAL + RAF + ZEO-TRP. Cultures were grown at 30°C for 3–4 days and concentrated 5-fold, and a fraction was plated to SC + CAN-ARG to obtain CanR mutants. Viable cell counts were obtained by plating a dilution to rich medium (33). Viable cells were counted after 2 days, and CanR colonies were counted after 3–4 days of incubation at 30°C. The CAN1 gene of CanR colonies was amplified by PCR and sequenced as reported (34). Accurate values for the mutation frequencies were obtained by using multiple independent cultures (n = 6–8) for each strain in each experiment and by repeating each experiment at least twice and as many as seven times. Sequencer (GenesCodes) was used for mutational analyses.

Immunoblotting. Cell pellets from a 10-ml log phase culture were washed with 1 ml of 20% trichloroacetic acid (TCA), resuspended in 50 μl of 20% TCA, and then lysed by vortexing with an equal volume of glass beads at 4°C. The supernatant was centrifuged to pellet the proteins. Pelleted proteins were resuspended in 100 μl of SDS-gel loading buffer, separated by SDS-PAGE, transferred to a poly(vinylidene difluoride) membrane, and probed with antibodies to APOBEC3G (35), LexA (Invitrogen), or Vif (36–38).

Ty1 Retrotransposition Assays. Ty-his3AI, TyHRT-lucAI, or TyHRT-lucAI plasmids were cotransformed with pJG4–5, pJG4–5-APOBEC3G, or pJG4–5-APOBEC3F into DG1251 or GRY1990 (39) and selected by using SC-URA-TRP + GLC (33). his3AI transformants were grown in SC-URA-TRP + GLC to saturation. 106 cells were subcultured in 1 ml of SC-URA-TRP + GAL for 12 h, and an aliquot was plated to SC-HIS. Cell viability was determined by plating a dilution to rich medium. Retrotransposition was quantified by determining the frequency of His* colonies. lucAI transformants were grown 1 day in SC-URA-TRP + GLC. Cells were transferred to SC-URA-TRP + GAL and grown for an additional 2 days at 30°C to induce retroelement expression and reverse transduction. Retrotransposition was quantified by measuring the relative active levels of luciferase to β-galactosidase (D.V.N. and J. Strathern, unpublished data). All incubations for plasmid-based Ty1 assays were at 30°C.

For endogenous retrotransposition assays, DG1141 was transformed with pJG4–5, pJG4–5-APOBEC3G, or pJG4–5-APOBEC3F. Single colonies were resuspended in water, and 10,000–50,000 cells were transferred to 2 ml of SC-TRP + GAL and grown at 20°C for 7–10 days until the cultures reached saturation. Dilutions of the starting and ending cultures were plated to rich media to determine the number of viable cells, and the equivalent of 1 ml of the saturated culture was plated to SC-HIS to score retrotransposition events.

Ty1 DNA Sequencing. Retrotransposed Ty1 and TyHRT cDNAs were isolated by growing His* colonies overnight in 10 ml of SC-HIS at 30°C and preparing DNA with a standard glass bead/phenol extraction method. The resulting DNA was used to amplify a 1,026-bp (Ty) or 971-bp (TyHRT) region spanning the RT gene and HIS3 by using 5′-TTC ATG TGG GAC ACT AGA GAT (TyRT) or 5′-CCT GAG TGG GAG TGT TTA (TyHRT) and 5′-TAT GAT ACA TGC TCT GGC CAA (HIS3). PCR products were purified (Qiagen) and sequenced with 5′-GT CTG CGA GGC AAG AAT GAT. GFP-negative retrotransposition events were obtained from pools of His* colony genomic DNA by transformation into Escherichia coli. GFP-negative colonies were identified with fluorescent light, and the resident plasmid DNA was amplified (as above except the product was 2.1 kb for Ty RT) and sequenced by using 5′-C GTT ATC CGG ATC ATA TGA and 5′-G TAG TTT CCC TCA TCT TGA.

Results

APOBEC3G Stimulates Mutation in S. cerevisiae via the Uracil Excision Pathway. To test whether human APOBEC3G could elicit its hallmark mutator activity in yeast, we expressed a LexA-APOBEC3G fusion protein in the haploid strain L40 and monitored the accumulation of mutations that conferred resistance to the toxic amino acid canavanine. Liquid cultures were grown from individual colonies expressing APOBEC3G or a control vector and then plated onto a solid medium containing canavanine. The numbers of canavanine-resistant (CanR) colonies were determined after 3- to 4-days growth. In contrast to cells expressing a control vector, those expressing LexA-APOBEC3G showed a 20-fold increase in the median frequency of CanR mutation, suggesting that APOBEC3G was capable of deaminating cytosines within yeast genomic DNA (Fig. 1A and Fig. 6, which is published as supporting information on the PNAS web site).

To begin to determine whether the LexA-APOBEC3G-induced mutator phenotype occurred by a C → U deamination mechanism, we asked whether a uracil DNA glycosylase deficiency would exacerbate this phenotype. Because most DNA-based organisms use uracil DNA glycosylase to rid their genomes of uracil (40), it was likely that, were this the mechanism, many of the APOBEC3G-induced uracil lesions would have been repaired and that the observed mutation frequency would be an underestimate of APOBEC3G activity. Indeed, yeast expressing both APOBEC3G and a uracil DNA glycosylase inhibitor (Ugi) protein showed a 320-fold increase in the median frequency of mutation to CanR (Fig. 1B). This stimulation was 6- and 26-fold higher than that observed in LexA-APOBEC3G-expressing and in Ugi-expressing yeast cells, respectively, indicating that many of the APOBEC3G-dependent uracils were repaired by a uracil excision mechanism.

In yeast, the major uracil DNA glycosylase is Ung1p (uracil DNA
N-glycosylase 1 protein). UngIp and most other Ung proteins from bacteria to humans are strongly inhibited by Ugi (41). However, Ugi-resistant uracil excision activities occur in mammalian cells, such as those elicited by the SMUG1 and TDG1 proteins (40). To eliminate the possibility some of the APOBEC3G-induced uracils might be repaired by auxiliary systems in yeast, we used homologous recombination to construct an UngIp deletion strain, L40 ung1::kanMX4. This strain showed levels of CanR mutation virtually indistinguishable from UngIp-expressing cells in the presence or absence of APOBEC3G (Fig. 7, which is published as supporting information on the PNAS website). Thus, the majority of APOBEC3G-induced lesions in yeast were repaired by an UngIp-dependent mechanism. Together with the exquisite specificity that UngIp has for uracil, these data indicated that the APOBEC3G-dependent mutator phenotype was attributable to a DNA cytosine deamination mechanism.

APOBEC3G is localized predominantly to the cytoplasm of mammalian cells (4, 11). Therefore, we were surprised by the high mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of theLexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the LexA tag. Therefore, we monitored the mutation frequencies caused by its expression in yeast and concerned that this might have been attributable to the DNA-binding properties of the Le
previously in model HIV and murine leukemia virus retroviral systems as 5′-YCCA (Y = C or T, ref. 19). Interestingly, APOBEC3G exhibited a strikingly similar 5′-CCCA preference in yeast (Fig. 3), indicating that its preference, as observed in other systems, was intact.

It is further notable that in addition to a large number of C/G → T/A transition mutations, four deletions and a single insertion were detected in the CAN1 gene of APOBEC3G-expressing yeast cells (Figs. 2 and 8; combined data including the Ugi experiments). Three of five of these alterations occurred either in or immediately adjacent to a preferred or potential APOBEC3G hotspot, 5′-CCC. In contrast, only 1 of 12 of the deletions and insertions found in control vector containing cells occurred at similar sites. The remainder (11 of 12) were distributed throughout the CAN1 gene and were presumably caused by a variety of mechanisms. The presence of deletions and insertions associated with APOBEC3G hotspots suggested that C → U deamination events are able to precipitate gross genomic instability. This finding is further supported by our observation that a small (<5%) proportion of Can+ mutants failed to yield a CAN1 gene-specific PCR product, potentially representing larger-scale lesions.

**Affect of HIV-1 Vif on APOBEC3G-Induced Yeast Hypermutation.** In primates such as humans and chimpanzees, Vif counteracts the antiretroviral activity of APOBEC3G by targeting it for proteasomal degradation (reviewed in refs. 9 and 10). Vif accomplishes this by binding to APOBEC3G. Some data suggest that this association alone may directly impair APOBEC3G function (44). Therefore, we wished to assess whether the interaction between Vif and APOBEC3G could be detected with this yeast assay system.

HIV-1 Vif, derived from the YU2 or the IIIB provirus, was expressed alongside APOBEC3G by using yeast two-hybrid bait or prey vectors. All possible pairwise combinations were tested for the ability to drive the yeast two-hybrid reporter genes lucZ or HIS3. No significant β-galactosidase activity or histidine prototrophy was observed despite repeated attempts (data not shown). This result was not attributable to an expression failure, as both proteins could be detected in cell lysates by immunoblotting (Fig. 9, which is published as supporting information on the PNAS web site).

However, because some weak or transient interactions may escape detection by the yeast two-hybrid assay, we reasoned that the sensitive CAN1 mutation assay might provide a more robust method for monitoring this interaction. To examine whether HIV-1 Vif could affect APOBEC3G-mediated hypermutation in yeast, we compared the Can+ mutation frequencies of cells coexpressing Vif and APOBEC3G with those of cells expressing either protein alone. The robust hypermutability of APOBEC3G was not significantly affected by HIV-1 Vif coexpression (Fig. 9). Therefore, we have been unable to detect a Vif-APOBEC3G interaction in yeast.

**APOBEC3F and APOBEC3G Inhibit Ty1 Retrotransposition.** To explore the possibility that APOBEC3 proteins function to impede the mobility of endogenous retroelements that replicate by using LTR sequences, we assayed the ability of the yeast retrotransposon Ty1 to replicate in the presence of APOBEC3G or its homolog APOBEC3F. Ty1 activity was monitored by using an intron-containing decay assay (Fig. 4A), Ty1 RNA expression, splicing, reverse transcription, and integration yield functional reporter gene cDNA copies, encoding either histidine prototrophy (45, 46) or luciferase activity.

We first measured the ability of Ty1-his3AI to retrotranspose in the presence of human APOBEC3F or -G (Fig. 4B). In comparison to cells containing a control vector, an average of 51% or 70% fewer His+ colonies were detected in the presence of APOBEC3F or -G, respectively. Slightly larger APOBEC3-dependent declines in Ty1-lucAI retrotransposition were observed, as monitored by the relative levels of luciferase present in liquid cultures (Fig. 4C). However, an almost total inhibition (94–98%) was observed when retrotransposition of a genomic Ty1-his3AI element was assayed in the presence of APOBEC3F or -G, suggesting that the ratio of APOBEC3 protein to retrotransposition intermediate (and/or Ty host factors) is a key determinant of this inhibitory mechanism (Fig. 10, which is published as supporting information on the PNAS web site). Together, these data clearly demonstrated that APOBEC3F or -G can inhibit Ty1 retrotransposition.

To assess whether the APOBEC3-dependent inhibition of retrotransposition in yeast could be influenced by the reverse transcriptase or the integration pathway, similar assays were performed with Ty1 constructs in which the normal reverse transcriptase was replaced with that from HIV-1 (TyHRT; refs. 45 and 46). TyHRT integration occurs predominantly by homologous recombination, whereas Ty1 integration mostly uses its own integrase. Retrotransposition of both TyHRT-his3AI and TyHRT-lucAI (i.e., the accumulation of HIV-1 reverse transcriptase products) was also inhibited by APOBEC3F or APOBEC3G expression (Fig. 4D and E). Levels of inhibition were roughly similar to those observed with Ty1
reverse transcriptase, indicating that neither the reverse transcriptase nor the integration pathway were key effectors of the APOBEC3-imposed retrotransposition block. These data further highlight the utility of the yeast Ty1 system for studying aspects of both APOBEC3 and HIV-1 biology.

Ty1 Restriction by APOBEC3F and APOBEC3G Involves a cDNA Cytosine Deamination Mechanism. As cDNA C → U deamination is a hallmark antiretroviral activity of APOBEC3F and -G, we asked whether this could account for the observed Ty1 retrotransposition block. If so, we expected to find an inordinate number of retrotransposon minus strand C → T transition mutations amongst the His+ integrants (equivalent to plus strand G → A transitions). We sequenced >26 and 47 kbp of TyRT-HIS3 template generated in the presence of APOBEC3F and -G, respectively, and found only two C → T transitions among the APOBEC3G-exposed templates. One transition occurred within a dinucleotide consensus 5'-YCC that is rarely preferred by this protein, and it therefore likely represents a reverse transcription or PCR error. The second transition occurred within the trinucleotide 5'-YCCC, which is the most common APOBEC3G-preferred site. However, this meager number of base substitutions may have been in part due to the fact that functional His+ (and not His-) integrants were analyzed. It is further possible that uracil residues within the retrotransposon cDNA triggered its degradation, as hypothesized originally for retroviruses (2).

Therefore, to address the former possibility and to enrich for mutations, we used a modified version of the Ty-his3AI system in which a GFP cassette was placed upstream of his3AI (Fig. 11, which is published as supporting information on the PNAS web site). This system enabled the selection of His+ integrants and a subsequent screen for unselected GFP-negative variants. Twenty independent GFP mutants were recovered from retrotransposition experiments in which APOBEC3G was expressed. Each sequence contained at least one mutation and as many as 15 mutations. In total, 56 base substitution mutations were identified, and 47 of these were minus strand C → T transitions (Figs. 5A and 11). Almost all of the APOBEC3G-dependent transitions occurred within the consensus 5'-YCC, very similar to the preferred cytosine deamination consensus site in the CAN1 gene and in a variety of other systems (e.g., compare Figs. 5B and 3). Moreover, many of the C → T transitions occurred at positions that were identical to those observed previously in GFP-encoding HIV or MLV (2, 19). A similar strand-specific transition bias and sequences with multiple mutations were found in GFP-negative templates produced in the presence of APOBEC3F (Figs. 5C and 11). However, in contrast to APOBEC3G, the APOBEC3F-dependent mutations occurred within a distinct 5'-TTGCC consensus (Fig. 5D; observed previously with an HIV substrate; ref. 19). We conclude that Ty1 retrotransposition can be inhibited by APOBEC3F and -G and that much (and possibly all) of this effect can be attributed to a cDNA cytosine deamination mechanism.

Discussion

The studies reported here demonstrate that APOBEC3G expression in S. cerevisiae results in a genomic mutator phenotype entirely attributable to a DNA cytosine deamination mechanism. The 5'-CC → CT transitions occurred within a broader local preference that could be found on either strand of the DNA duplex. In addition, the fact that APOBEC3G was functionally expressed in this system enabled us to show that both it and its homolog APOBEC3F could inhibit the mobility of the yeast retrotransposon Ty1 by a mechanism also involving DNA cytosine deamination. However, all of the Ty1 C → T transitions occurred preferentially on the cDNA (minus) strand of the retroelement. Thus, strand specificity distinguishes the mechanisms of genomic hypermutation and retroelement restriction.

A second distinguishing feature of these mechanisms is the fact that they manifest at dramatically different frequencies. CAN1 mutation occurred at a low frequency, with <1% of 50,000 of the APOBEC3G-expressing cells eliciting CanR. Because CAN1 is 1,771 bp, this frequency translates to ~0.000012 mutations per kb. In contrast, Ty1 restriction by APOBEC3G was highly efficient, because the vast majority of potential retrotransposition events could be inhibited (70–98%) or more depending on whether Ty1 transcripts were present in relatively high copy (plasmid-based, GAL promoter) or low copy (chromosome-based, Ty promoter), respectively (Figs. 4 and 10). The frequency of mutation in this system can be gauged by calculating the number of additional (hitch-hiking) mutations associated with those that inactivate the Ty1 GFP cassette. For APOBEC3G, 36 additional minus strand base substitutions (including 33 C → T transitions) were found among 20 independent 1,488-bp Ty1 cDNA sequences, translating to ~1.2 mutations per kb. This 5-log potency differential suggested that the APOBEC3G-dependent genomic lesions are occurring nonspecifically (possibly genome-wide), whereas Ty1 restriction may be due to the specific targeting of APOBEC3G (or -F) to the replicating retroelement. This observation is supported by the recent finding that APOBEC3G could associate with Ty1 Gag in an RNA-dependent manner (47). A similar specific association appears to be required for the restriction of HIV-1 by APOBEC3G (reviewed in refs. 7 and 8). Thus, taken together with our studies here, which uniquely revealed a major role for cDNA deamination in Ty1 restriction, it is likely that APOBEC3G and -F exploit an ancient and perhaps a structurally conserved pathway to accomplish the destruction of replicating retroelements.

Schwartz and colleagues (48) recently reported that APOBEC3G was able to block the transposition of two types of murine retroelements in HeLa cells. As for Ty1 and several retroviruses, much of the inhibitory effect was attributable to the deamination of cDNA cytosines. However, because the APOBEC3 proteins do not exist outside of mammals (e.g., refs. 8 and 28), our data showing that APOBEC3F or -G could inhibit yeast Ty1 retrotransposition were unexpected. Therefore, the Ty1 data not only demonstrate the remarkable conservation of this mechanism but, importantly, they also show that mammalian factors (in addition to APOBEC3F or -G) are not required for retroelement restriction.

This is the first report that a two-domain DNA cytosine deaminase such as APOBEC3G can mutate a nuclear gene within a eukaryotic cell. This was particularly surprising because APOBEC3G had shown a clear cytoplasmic localization in a variety of mammalian cell lines (4, 11). Therefore, the simple presence of a nuclear membrane or chromatinized DNA in eukaryotic cells is not sufficient to prevent APOBEC3G from deaminating genomic cytosines. Thus, our data imply the existence of regulatory mech-
anisms that serve to ensure the proper cytoplasmic compartmentalization of APOBEC3G in mammalian cells. A further implication is that the loss of such regulatory mechanisms might lead to genomic hypermutation, larger-scale lesions, and eventual carcinogenesis. The likelihood of such a pathological role is supported by data indicating that APOBEC3G is overexpressed in some breast cancer tissues (49).

The APOBEC3G-dependent yeast CAN1 mutations were found on both DNA strands, indicating for the first time that both the nontranscribed (displaced) and transcribed DNA strands can be substrates for its deamination activity. Previously, APOBEC3G had shown only a marked preference for single-strand DNA or for DNA that may have been displaced by transcription (e.g., refs. 2 and 49–51). Such an association with transcription was supported by our data, where a 6-fold bias in displaced strand C → T mutation was observed in APOBEC3G-expressing cells lacking Ung1p (Fig. 2). However, the fact that APOBEC3G-dependent C → T transitions were distributed equally over both DNA strands in the repair proficient parent strain indicated that the preferential repair of uracils in the displaced strand might also be occurring. Additional experiments will be required to further delineate this possibility.

APOBEC3F and -G are related to activation-induced deaminase (AID), a single-domain DNA cytosine deaminase with a prominent role in generating Dicer-mediated mutations (exogenous or endogenous). These proteins are around twice the size of AID, and they do not appear to possess nuclear localization signals. Nevertheless, both studies underscore the tremendous potential of the yeast system for studying these molecules. For instance, when our data are considered with the fact that in Cul5–Elongin B/C ubiquitin ligase complex is required for APOBEC3G degradation in human cells (24, 54), it becomes reasonable to use yeast to reconstitute the minimal components required for this interaction and thus open the door to alternative therapeutic screening strategies.

We propose that limiting the mobility of endogenous retroelements is the predominant physiological function of some of the mammalian APOBEC proteins, including the human APOBEC3F and -G DNA cytosine deaminases. A balance between limited mobility and maintained genome integrity is likely to be an important determinant of speciation, and it would certainly account for the fact that many of the APOBEC proteins are positively selected (27, 29). The advantage of having a potent retroelement restrictor clearly outweighs the disadvantage of encoding and regulating potential cellular genome mutants such as APOBEC3F and -G.

We thank A. Bielinsky, W. Brown, D. Kirkpatrick, D. Livingston, D. MacDuff, A. Rattray, and R. Ricke for helpful comments; K. Chris- tierman and F. Dobber for technical assistance; D. Garetinkel, J. Lingappa, M. Malim, and R. Wright for reagents; J. Strathern for unpublished reagents; and the University of Minnesota AGAC for sequencing. The HIV-1 Vif antibody (no. 319) was obtained from M. Malim through the AIDS Research and Reference Reagent Program, Division of AIDS, National Institute of Allergy and Infectious Diseases, National Institutes of Health. This work was supported by the University of Minnesota, a Scare Scholar, and a Burroughs–Wellcome Fund Hitchings–Elion Fellowship (to R.S.H.), and it was funded in part with federal funds from the National Cancer Institute, National Institutes of Health, under Contract NO1-CO-12400 (to D.V.N.).
Frequency His$^+$ (10$^{-8}$)

- Ty + Control
- Ty + A3G
- Ty + A3F
A

Reverse Transcription & Integration

GFP+

Sequence

B

AAG AAA CCA ATT ATT AAA GCC TTA CTT ACT GAT AGT AGA TCA ACG ATT AGT ATT ATT AAA CAA GCC TTA CAA AAT GAA GAG AAA TTT AGA AAT GAA TTT 372

TTT GCC ACA AAG GCA ATG AGA CTT AGA GAT GTA TCA GAT AAT AAT TTA AAC ATC ATC AGT ATA ATT AAG TCT ACA AAT GAA GAG AAA TTT AGA AAC AGA TTT 469

ATG ACA AAA CCT CTT CGG ATA AAA ACA TTT AAA CTA TTA ACT ACT AAT AAA TCG ATT CAT ATC GCG CCC CGA TCC CCC GAT TCA TTA ATG CAG 558

CTG GCA GCA CAG GCC GTT CCC CGA CTG GAA AGC TGA AAT TGA GAC TTA TAT GAA TCA GCT TCA TTA GCC ACC CCC CGA GCC TTT ACA 651

CTT TAT GCT TCC GCC TCG TAT GTC GAT GAA GTA TCA GAT GAA GTA TCA TTA GCA GCC GAT TCA TTA CCG CAA GAT TGC TTA TTA GAC TGA 744

ATG CCT GCA GCT CAA CCT TCG TAG AGG AGT CCC GCC CGG TAG AAA AAA TTA GTA AAG GAG AAG AAT TTT ATC CTT AGA TTG TCG TCC CAA TCC TGG 837

TTG AAT TAG ATG GTC ATG TTA ACT GCC ACA AAT TTT CTG TCA GTG GAG AGG TTT TAG ATG GAG ATG CAC TCA TTG AAC CAA TTT ACT TTA AAT 930

TTT GCA CTA CTG GAA AAC TAC CTC TCC CAT GCC CAA CAC TGG TCA CTA CTT TCC ATG TCA ACA TTT CCC CAA GCC TTA GCC TAC AAA 1023

TCA AAG GCC ATG ACT TTT TCA ACG GCT TAC CGG AAG GCC TTA ACG ATG AGA AAC GCA GCT TAT TTA TTA AAG TTA TTA GAA TGG GAA GCC TCC 1116

GTT CAG AGC TAT GTC GAA AGT ATG GAT AAG TTA ACG TAT GAT ATG TTA AAG ATG AAT TAC AAT GAT TTA AAG TTT ATC CAA TGC AAG 1209

AAC TGC AGT ACA ACT ACA AAT ACA ATG GAT CCG TTC AAC TAG CAG ACC ATT TCA AAG AGT GCA TGG TCC TTA GAT TGA TGC TCA GTG GGA GCG 1295

CTG AAT GTG ATG GAT CCG TTC ATG GTC TCA AAG TCA ATC AAT GAA TCA CTG ATG TGG TAT GGG TCT TCC TCC TAT GAG CAC ACA ATG GCA GCT 1388

GGA ATG AGC TCT ATC AGG TCT AGG CAA ATC AAG GGC GGT CCC TGC TCT CTA AAT TGG TCA CTT GCC ATC ATG AGG AGGC ATC ACC TTT TCC 1581

GGC TGC AGA TCC ATC ATC AGC CTT GAA TTA AAG TTA TTA AGG ATG GCC ATC ATG AAC GCT TTT GGT GCA CAG ATC TAA TAA GGAC TCA GCT 1674

GTC TCC TAT GCC ACG GC ACG CCT GAA GGC ACT CTC ACT ATG GTG ATG ATG ATT TCT GCC TGG CAG ACA ATC AAC ATG GAT GGT AAT 1767

Mutation to: ACGT

A

C

G

T