

Orthologous Mammalian APOBEC3A Cytidine Deaminases Hypermutate Nuclear DNA

Vincent Caval,¹ Rodolphe Suspène,¹ Jean-Pierre Vartanian,¹ and Simon Wain-Hobson^{*1}

¹Molecular Retrovirology Unit, Institut Pasteur, Paris, France

***Corresponding author:** E-mail: simon.wain-hobson@pasteur.fr.

Associate editor: Meredith Yeager

Abstract

The human *APOBEC3* gene cluster locus encodes polynucleotide cytidine deaminases. Although many act as viral restriction factors through mutation of single-stranded DNA, recent reports have shown that human *APOBEC3A* was capable of efficiently hypermutating nuclear DNA and inducing DNA breaks in genomic DNA. In addition, the enzyme was unique in efficiently deaminating 5-methylcytidine in single-stranded DNA. To appreciate the evolutionary relevance of these activities, we analyzed A3A-related enzymes from the rhesus and tamarin monkey, horse, sheep, dog, and panda. All proved to be orthologous to the human enzyme in all these activities revealing strong conservation more than 148 My. Hence, their singular role in DNA catabolism is a well-established mechanism probably outweighing any deleterious or pathological roles such as genomic instability and cancer formation.

Key words: APOBEC3, double-strand breaks, cytidine deaminase, hypermutation, nuclear DNA, methylcytidine.

Introduction

The *APOBEC3* (A3) locus, situated between the two conserved vertebrate genes chromobox 6 and 7 (*CBX6* and *CBX7*), is common to most placental mammals and arose from duplication and subsequent expansion of an ancestral paralogous gene, probably activation-induced deaminase (AID) (Jarmuz et al. 2002; Conticello et al. 2005). The human locus comprises seven *APOBEC3* genes (*APOBEC3A-H*) (supplementary fig. S1, Supplementary Material online), all being single-stranded DNA (ssDNA) cytidine deaminases (CDAs) that distinguish them from cytosine deaminase and CDAs involved in pyrimidine and nucleic acid precursor metabolism. Accordingly, they constitute a new class of deaminases referred to as polynucleotide CDAs (PCDs). *APOBEC3A*, -3C, and -3H (A3A, A3C, A3H) encode a single zinc finger domain while four (A3B, A3DE, A3F, A3G) encode two domains each with a zinc finger. The zinc finger motif, His-X-Glu-X₂₃₋₂₈-Pro-Cys-X₂₋₄-Cys, is singular to these deaminases, whereby the glutamic acid potentiates a water molecule that acts as the catalytic nucleophile in the oxidation of cytidine amino group, resulting in the conversion to uridine (Jarmuz et al. 2002).

A3 CDA function was first tied to innate immunity as antiviral restriction factors (Sheehy et al. 2002). Indeed, when deamination occurs in the confines of a virus particle away from DNA repair, cytidine deamination to uridine results in hypermutation of the viral genome pushing the mutation rate way beyond the error threshold (Harris et al. 2003; Lecossier et al. 2003; Mangeat et al. 2003; Mariani et al. 2003). A3F and A3G posed such a problem for lentiviruses so that their macrophage-specific virus precursors evolved the Vif protein that ultimately shunts the Vif:A3 complex to the proteasome (Conticello et al. 2003; Marin et al. 2003; Sheehy

et al. 2003). Hepatitis B virus (HBV) is particularly vulnerable to A3 editing as it has not evolved a countermeasure (Suspène, Guétard, et al. 2005; Vartanian et al. 2010).

Moreover, the nucleocytoplasmic localization of single domain A3A, A3C, and A3H along with the nuclear localization of the double domain A3B, through nuclear import mechanisms, also allows restriction of viruses replicating in the nucleus. Thus, human papillomavirus (HPV) (Vartanian et al. 2008) and human herpes simplex virus type 1 (HSV1) genomes are susceptible to A3 editing (Suspène, Aynaud, Koch, et al. 2011). Additionally, some retroelements and retrotransposons are vulnerable to A3A (Bogerd et al. 2006; Chen et al. 2006; Muckenfuss et al. 2006).

More recently, a nonviral role for human A3 enzymes was described. Although at least four enzymes can initiate the catabolism of cytosolic mitochondrial DNA (mtDNA), A3A is capable of efficiently editing nuclear DNA (nuDNA) (Suspène, Aynaud, Guétard, et al. 2011). It is possible that, to a lesser extent, A3B may be able to edit nuDNA (Shinohara et al. 2012; Burns et al. 2013). Cytidine deamination to uridine in nuDNA has two consequences. The uracil base is first rapidly excised by uracil N-glycosidase (UNG), thereby initiating DNA repair, which can involve the formation of double-stranded DNA breaks (DSBs) (Landry et al. 2011; Mussil et al. 2013). However, if cytidine deamination overwhelms DNA repair, it leads to apoptosis (Kaina 2003). Although hypermutation almost certainly overwhelms DNA repair, leading to unrepaired DSBs and apoptosis, lower levels of deamination, compatible with repair, can lead to the fixation of numerous mutations. This follows from the finding of small regions in some cancer genomes carrying large numbers of C->T transitions in the TpC dinucleotide, the "signature" context of A3A (Nik-Zainal et al. 2012). This hypothesis is

more interesting when it is realized that A3A is the only PCD capable of efficiently deaminating 5-methylated cytidine residues that have long been noted as mutation hotspots in cancer genomes (Carpenter et al. 2012; Wijesinghe and Bhagwat 2012; Suspène et al. 2013). As the A3A gene may be induced by interferon α and γ (Bonvin, et al. 2006; Argyris et al. 2007; Koning et al. 2009; Refsland et al. 2010; Stenglein et al. 2010), it is possible that chronic inflammation may maintain A3A expression with commensurate damage to nuDNA.

Phylogenetic analysis revealed a clustering among A3 zinc finger motifs, further classified as A3Z1, A3Z2, or A3Z3 (Conticello et al. 2005; Larue, Andresdottir, et al. 2008; LaRue, Jonsson, et al. 2008). This diversification probably preceded the split between placental mammals and marsupials (125–150 Ma) (Bininda-Emonds et al. 2007; Munk et al. 2012). There is considerable evidence for gene conversion within the A3 locus, the number of A3 genes varying from one to seven, expansion of the A3Z2 progenitor being the most frequent, while A3Z3 always remains a single copy (Larue, Andresdottir, et al. 2008). By contrast, the A3Z1 precursor was either expanded, or occasionally lost, notably in rodents and the pig. Given such extensive gene conversion, it is possible that many of these genes are no longer orthologs, both among themselves in the same A3 cluster and between mammals.

As we have shown that human A3A is involved in DNA hypermutation and catabolism of nuDNA (Suspène, Aynaud, Guétard, et al. 2011) and may well be an important determinant in the onset of cancer, we sought to determine whether the known mammalian A3A deaminases were orthologs of the human enzyme. A3A deaminases from the rhesus macaque, tamarin, cattle, sheep, horse, dog, and panda were studied. It turns out that all the mammalian A3A constructs tested retained the ability to edit nuDNA, 5-methylated cytidine residues and induce DSBs emphasizing the maintenance of DNA catabolism over 148 My of evolution.

Results

Mammalian APOBEC3A Sequences

From the *Primate* order, we selected a rhesus monkey A3A known to be active against lentivirus and referred to as rhesus a (Hultquist et al. 2011; Schmitt et al. 2011), along with a second allele reconstituted from a genomic DNA sequence (Henry et al. 2012), called rhesus b (fig. 1A). Although rhesus b exhibits a classical HAE motif in its active site, rhesus a protein contains HVE residues, reminiscent of AID sequence (fig. 1A). In the same way, the New World Monkey tamarin (*Saguinus oedipus*) A3A cDNA was inferred from genomic DNA sequences (Bulliard et al. 2011; Henry et al. 2012). For ruminant artiodactyls, A3Z1-like sequences were retrieved from genomic and expressed sequence tag (EST) data for both sheep (*Ovis aries*) and cattle (*Bos taurus*) (Larue, Andresdottir, et al. 2008). For the horse (*Equus caballus*), only the functional A3Z1b gene was synthesized (Zielonka et al. 2009). For *Carnivora*, A3Z1 PCD cDNA from the dog (*Canis lupus*) was identified during automated genome annotation. Blast-like alignment tool (BLAT) probing of recently available

panda (*Ailuropoda melanoleuca*) genome with the dog A3A sequence allowed us to reconstruct the A3Z1 sequence. As the precise localization of the splicing acceptor for the fifth panda exon was not clear, a cDNA comprising the first four exons was synthesized. An alignment of the A3A protein sequences (fig. 1A) revealed N- and C-terminal length variation with up to 44% amino acid sequence divergence (fig. 1A, supplementary fig S2A, Supplementary Material online). A concise phylogenetic analysis of those sequences, using neighbor-joining method (fig. 1B), reveals clustering among all A3A proteins with human A3C (A3Z2) and A3H (A33) as outliers, meaning that they are all bona fide A3Z1 proteins. A similar clustering was obtained with an evolutionary tree constructed from active site regions (supplementary fig S2C, Supplementary Material online). Interestingly, a two amino acids specific insertion was present in human and rhesus lineage within the active site. However, deleting those residues does not impair A3A editing properties (Carpenter et al. 2012; and data not shown).

Moreover, the previously identified Z1 specific amino acid I128 (fig. 1A) is strictly conserved (Larue, Andresdottir, et al. 2008). Interestingly, sequence variation and indels map to the surface and surface loops (fig. 1C), suggesting strong conservation of folding.

Synthesis and Expression of Mammalian APOBEC3As

Human A3A is expressed in two forms, p1 and p2, initiated from two ATGs at codons 1 and 13, both with “adequate” Kozak contexts (Thielen et al. 2010). The shorter sheep and cow amino termini correspond to initiation at the equivalent of p2 (fig. 1A). For the tamarin, panda, dog, and horse A3A proteins, a single protein form is expected given the strong Kozak context (supplementary fig S2B, Supplementary Material online). To compare A3As proteins functionality, A3A cDNAs were cloned into pcDNA3.1D/V5-His-TOPO with the same strong Kozak motif (ACCATG). Consistent with the literature, bovine A3A DNA was very difficult to clone (LaRue, Jonsson, et al. 2008), presumably due to protein toxicity in *Escherichia coli* resulting from a leaky T7 promoter expression. This problem was resolved by resynthesizing the cDNA with the complete 521 bp intron 4 from the bovine A3A gene. A3A expression was assessed in various cell lines. Western blot analysis of both human 293T-UGI cells, expressing Uracil-DNA Glycolase (UDG) inhibitor (UGI), (fig. 2) and quail QT6 (fig. 2B) transfected cells revealed strong and specific expression of V5-tagged A3A proteins, with the exception of tamarin A3A that was consistently weaker. The subcellular localization of overexpressed A3As was assessed in HeLa cells. Anti-V5 staining revealed that all A3As exhibit the classical nuclear and cytoplasmic distribution described for human A3A protein.

All APOBEC3As Are Potent PCDs

To test for A3A proteins activity, A3A expression plasmids were transfected into 293T cells, and the resulting cellular lysates were probed for deamination activity in a fluorescence resonance energy transfer (FRET)-based in vitro

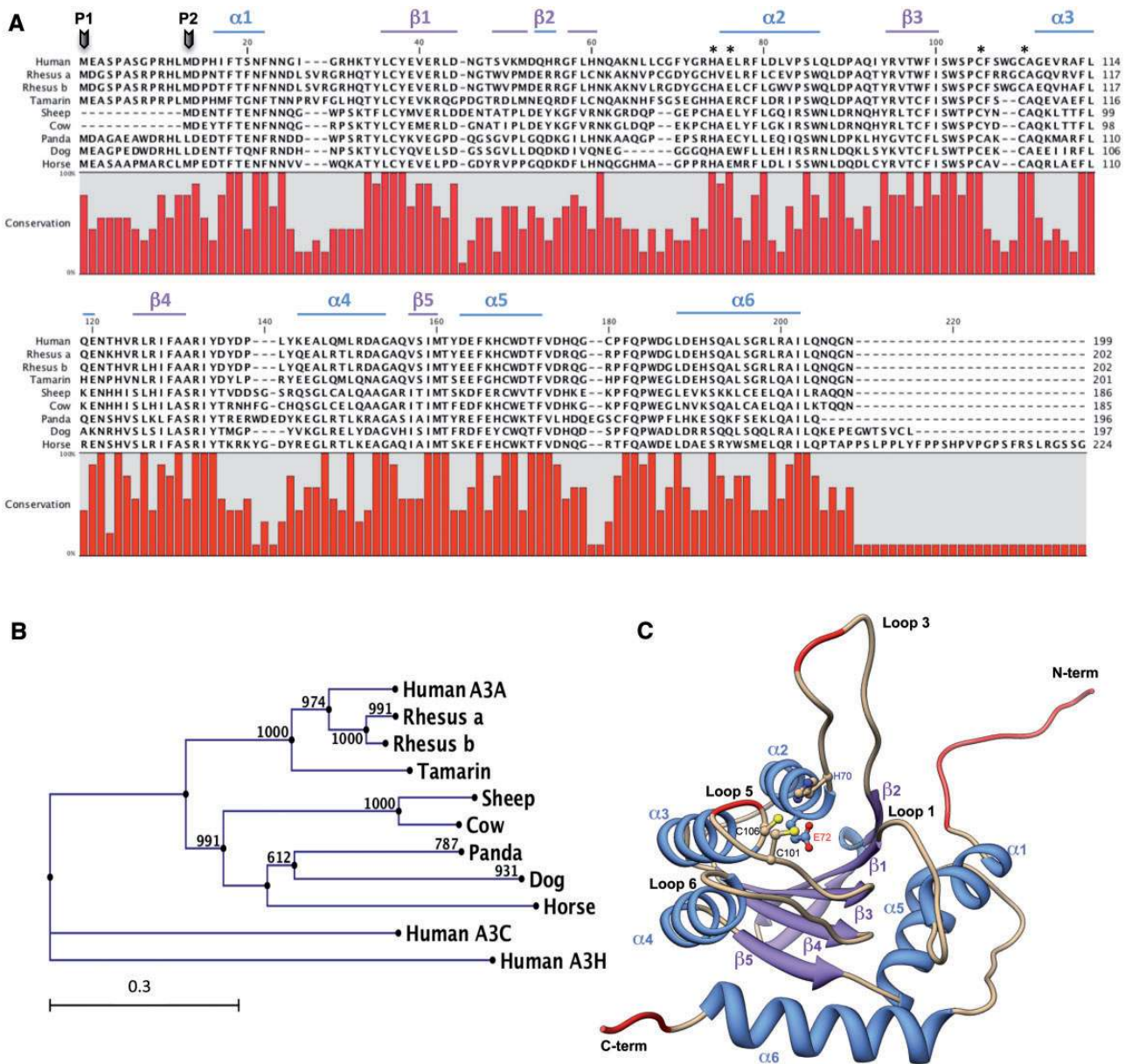


Fig. 1. Comparison of A3A CDA orthologs. (A) CLUSTALW alignment of A3A proteins. Sequence conservation is depicted in red for each residue, and human A3A described secondary structures are represented. Asterisks represent residues involved in zinc coordination responsible for enzymatic activity. (B) Evolutionary history of A3A CDAs inferred using the neighbor-joining method on protein sequences presented in (A). Numbers correspond to bootstrap values inferred from 1,000 replicates. Evolutionary distances were computed using Poisson correction method and are expressed in numbers of substitution per site. (C) Mapping of the variable regions on A3A putative structural model.

deamination assay (Stenglein et al. 2010). Briefly, cytidine deaminase expression causes the C to U conversion in a TAM-FAM labeled DNA oligo that will subsequently be cleaved by UDG activity, separating the quencher from the fluorophore, thus enabling fluorescence detection. All A3A transfected 293T whole cell lysates proved to exert C to U activity, albeit at different levels, while A3A catalytic mutant A3AC97S resulted in background levels of fluorescence, indicating that all the A3A proteins exhibit deamination activity (fig. 2C). A3A CDA function was further investigated in cellular context. Quail QT6 cells devoid of any APOBEC background activity (Henry et al. 2009) were transfected with A3A coding plasmids along with pCayw, an infectious molecular

clone HBV. As HBV viral replication occurs within capsid structures with an obligate reverse transcription step, it provides a good ssDNA substrate for APOBEC specific deamination (Suspène, Guétard, et al. 2005; Henry et al. 2009). Total cellular DNA from transfected cells was extracted after 48 h, and HBV editing analyzed by 3DPCR, a technique that allows recovery of AT-rich DNA (Suspène, Henry, et al. 2005).

For all A3A constructs, HBV DNA was recovered at temperatures far below the restrictive temperature of 89.0 °C obtained with mock-transfected cells or nontransfected cells (fig. 3A). If almost all A3A transfections allowed the recovery of amplified HBV DNA around 82.5 °C, for the tamarin A3A, 3DPCR recovered edited DNA at a higher temperature

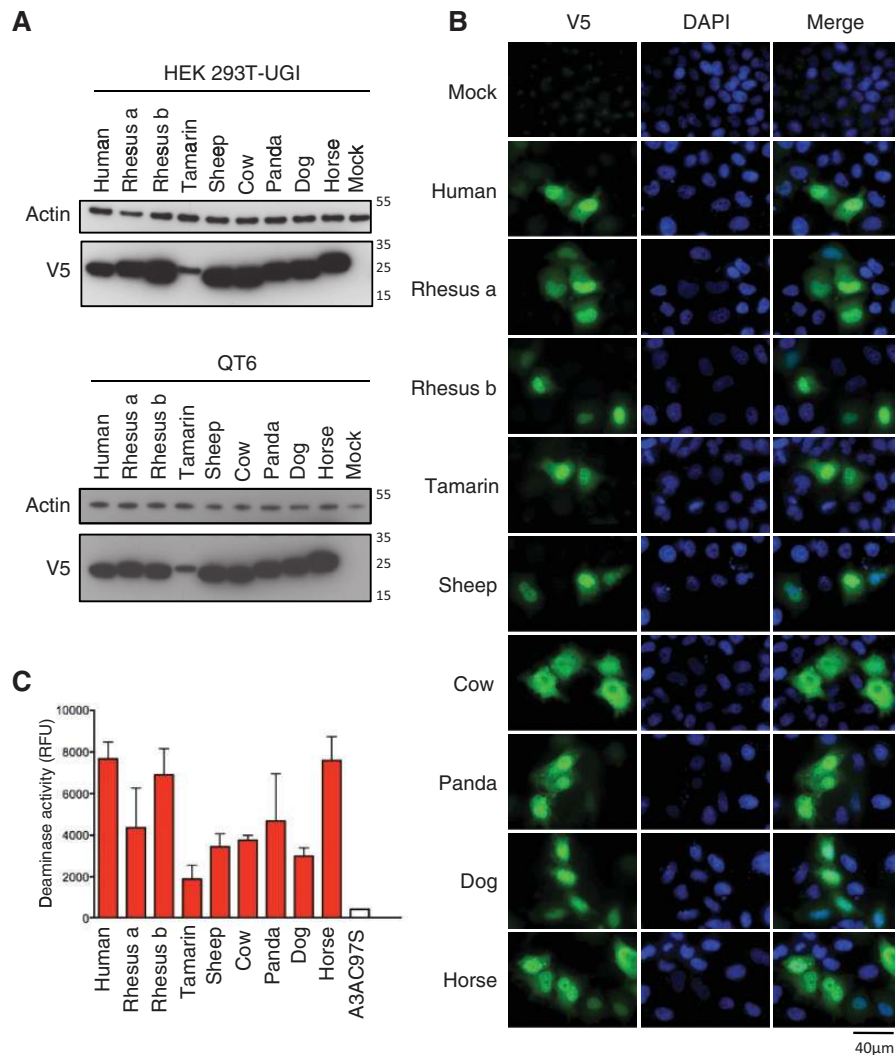


Fig. 2. Expression analysis of A3A proteins. (A) Western blot detection of V5-tagged A3A orthologs in both human HEK 293T-UGI and quail QT6 cells. β -Actin probing was used as loading control. (B) Cellular localization of A3A proteins. Confocal microscopy of V5-tagged A3A proteins performed in HeLa cells 24 h posttransfection. Nuclei are stained using DAPI. (C) FRET-based *in vitro* deamination assay performed in A3A 293T transfected cells. C to U activity was assessed by measuring fluorescence emission after A3A and UDG treatment on TAM-FAM DNA oligonucleotide containing Tp target sequence. Background fluorescence obtained with mock-transfected cells was subtracted, and A3A catalytic mutant A3AC97S was used as a negative control. Presented results correspond to two experiments performed in duplicate.

86.7 °C, suggesting either weaker editing activity and/or low protein levels in transfected cells. Molecular cloning and sequencing of 3DPCR products obtained at restrictive temperatures revealed extensively substituted sequences with a strong preference for C→U hypermutation on the HBV minus strand (fig. S3A and B, Supplementary Material online). The large number of sequences and mutations analyzed allowed comparison of the relative sequence context associated with deamination. In every case, there was a preference for 5'TpC and, to a lesser extent, 5'CpC, in keeping with previous data from human A3A (Henry et al. 2009; Suspène, Aynaud, Guétard, et al. 2011).

Cytosolic mtDNA too is sensitive to A3A deamination (Suspène, Aynaud, Guétard, et al. 2011), and quail mtDNA was amplified by 3DPCR from the same samples. Cytochrome c sequences were recovered below the restrictive temperature of 87.2 °C for all A3A transfections (fig. 3C). Again, the

tamarin A3A transfection only allowed amplification of edited DNA down to 84.8 °C, consistently with lower expression levels in QT6-transfected cells (fig. 2A), while the other mammalian A3A DNA was recovered at lower temperatures as far as 81.1 °C. Sequence analysis of cloned PCR products confirmed the hallmark TpC and CpC dinucleotide editing preference evidenced with HBV DNA (fig. 3D, and supplementary fig. S3C and D, Supplementary Material online). Taken together, those experiments revealed that mammalian A3A proteins are potent PCDs manifesting the mutational signature previously described for human A3A (Henry et al. 2009; Suspène, Aynaud, Guétard, et al. 2011).

APOBEC3A Editing of nuDNA Is Conserved

Singularly among human PCDs, A3A is able to edit nuDNA (Suspène, Aynaud, Guétard, et al. 2011). UNG excision of uracil bases in nuDNA is very rapid so that this A3A activity

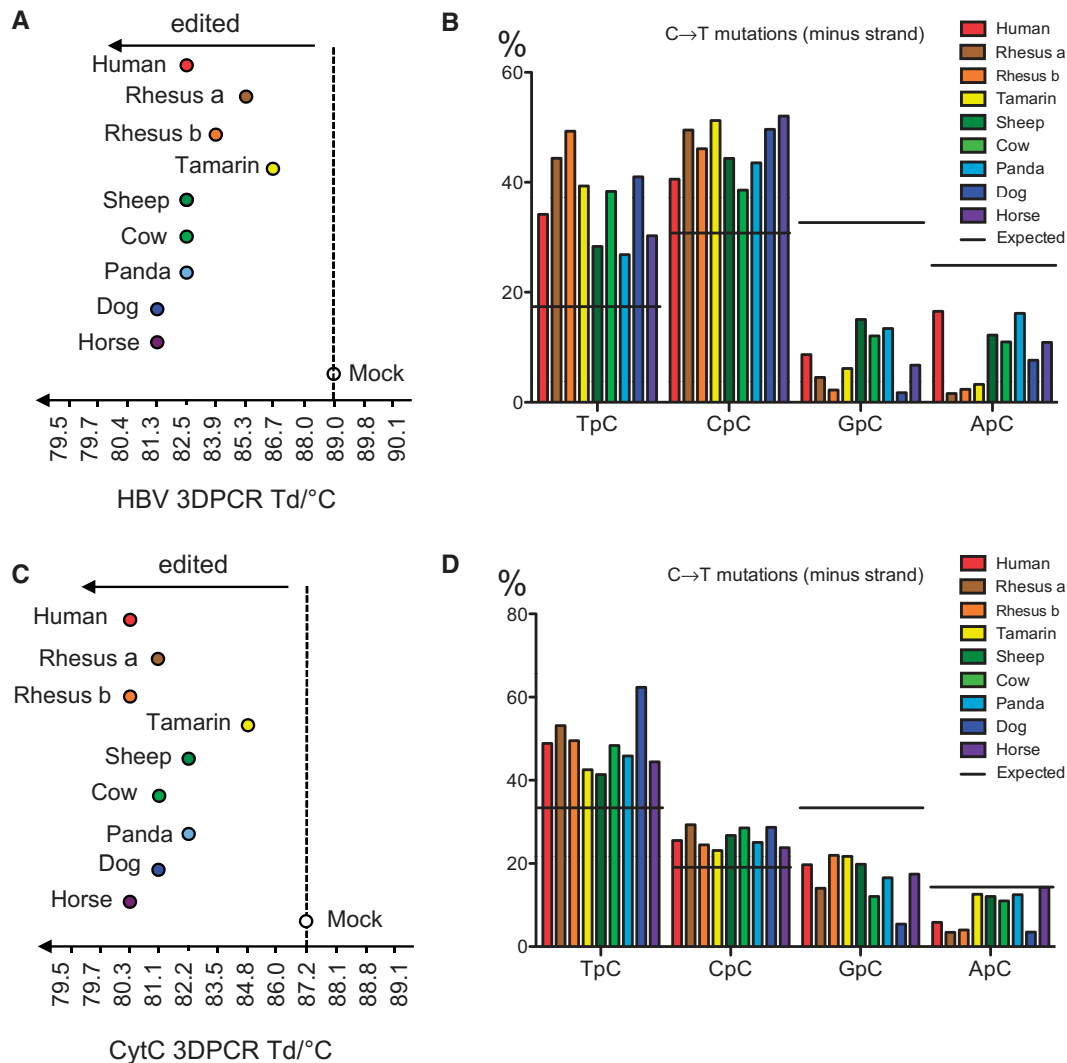


Fig. 3. A3A proteins orthologs and editing of cytosolic DNA. (A) Graphical representation of HBV DNA editing by A3A proteins. The last retrieved band by 3DPCR amplification is represented on the gradient. (B) 5' Dinucleotide analysis of the deamination context performed on HBV DNA minus strand for PCR products retrieved at 86.7°C. (C) Cytochrome C mtDNA editing by A3A proteins. The last retrieved band by 3DPCR amplification is represented on the gradient. (D) 5' Dinucleotide analysis of the deamination context performed on mtDNA minus strand for PCR products retrieved at 84.8°C. Dinucleotide context expected values, based on the dinucleotide composition of the minus strand, are given by horizontal lines.

can only be seen in UNG^{-/-} cell lines or else in 293T-UGI cells stably expressing the UNG inhibitor UGI (Suspène, Aynaud, Guétard, et al. 2011). nuDNA editing of *TP53* DNA was assessed by 3DPCR 72 h after transfection of 293T-UGI. Consistent with previous reports (Suspène, Aynaud, Guétard, et al. 2011; Aynaud et al. 2012), human A3A transfection allowed the recovery of hyperedited *TP53* DNA down to 86.6°C, the restrictive T_d being 89.7°C for the negative control. For all nonhuman A3A transfections, 3DPCR recovered DNA at temperatures lower than 89.7°C even if the shift was modest for the tamarin A3A. The other A3As proved to be as efficient as the human A3A reference if not more so (fig. 4A). Indeed, both the canine and equine A3A transfections resulted in *TP53* DNA recovery at even lower temperatures, 85.1°C and 84.6°C, respectively. Sequence analysis of cloned PCR products revealed hypermutated DNAs, with sequences peppered with C→T or G→A transitions, reflecting a nuDNA deamination process (fig. S3E and F). Dinucleotide

analysis of minus strand hypermutants showed a strong TpC preference for cytidine deamination, a trait of the human A3A enzyme (fig. 4B). Taken together, those results demonstrate that all tested mammalian A3A are proteins that are potent nuDNA mutators and are therefore orthologs.

Mammalian APOBEC3As Deaminate 5-Methylcytidine

Cytidine methylation is the most common DNA modification, playing a major role in epigenetic mechanisms of gene regulation and development. 5-Methyl CpG mutation hotspots are associated with cancer-related genes (Jones and Baylin 2007) while human A3A activity on modified cytidine has recently been reported (Carpenter et al. 2012; Wijesinghe and Bhagwat 2012; Suspène et al. 2013). A totally 5MeC substituted fragment of HIV-1 DNA was generated by PCR (Suspène et al. 2013) and QT6 cells transfected along with A3A expression plasmids. After 24 h, total DNA was extracted and edited DNA recovered by 3DPCR. As previously

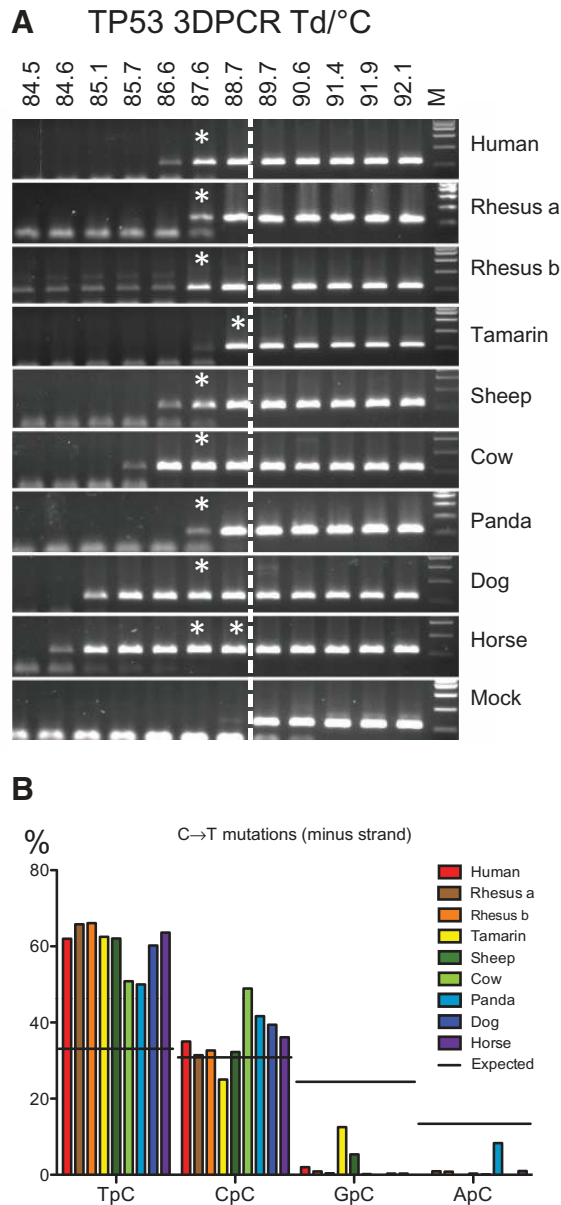


FIG. 4. A3A proteins orthologs and editing of nuDNA. (A) TP53-specific 3DPCR gels after 293T-UGI transfections with A3A proteins. Asterisks denote the samples that were molecularly cloned and sequenced. (B) 5' Dinucleotide analysis of the deamination context performed on DNA minus strand for cloned PCR products. 5' Dinucleotide analysis was performed at a lower temperature for horse PCR product (88.7°C) as sequences from 87.6°C amplification were more edited and homogeneous for context analysis.

described, human A3A could indistinctly edit 5MeC or classical dC-containing matrices, with 3DPCR products recovered down to 79.5°C. Control transfection with human A3C was only able to deaminate a classical matrix (fig. 5A) (Suspène et al. 2013). All the mammalian A3A deaminases were capable of deaminating 5-MeC substituted ssDNA (fig. 5A). When sequenced, cloned PCR products proved to be hypermutated (mean 41%, range 31–53% methylcytidines edited, supplementary fig. S3G and H, Supplementary Material online). Again, sequence analyses reveal a strong TpC preference for

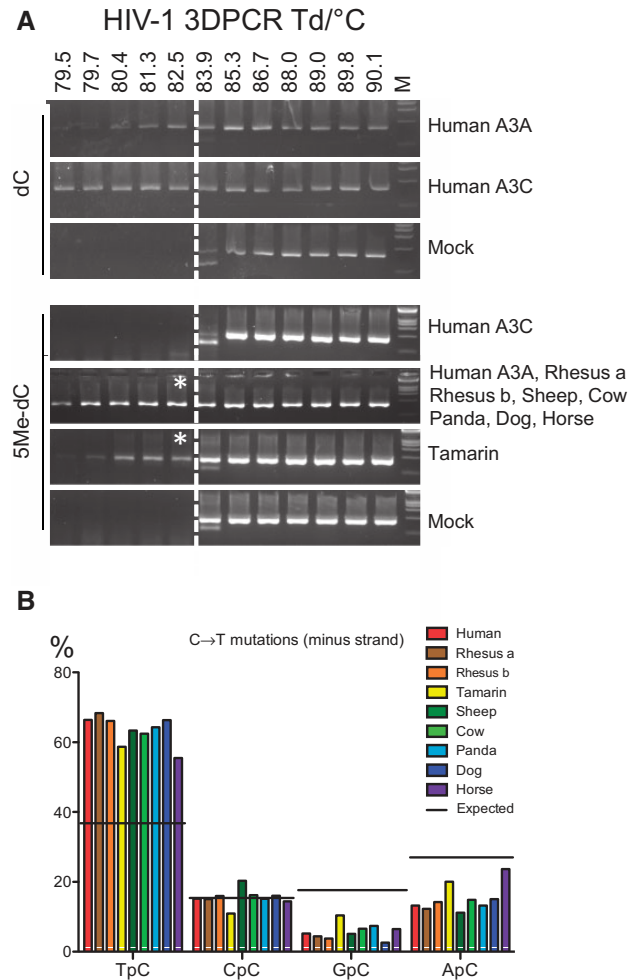


FIG. 5. A3A proteins can deaminate 5-methylcytidine in ssDNA. (A) HIV-1 specific 3DPCR gels after QT6 transfections with A3A proteins and human A3C protein along with classical dC or 5Me-dC substituted HIV-1 matrix. Asterisks denote the samples that were molecularly cloned and sequenced. (B) 5' Dinucleotide analysis of the deamination context performed on DNA minus strand for cloned PCR products.

editing, which was not altered by the 5-methyl group (fig. 5B). Once again, the tamarin A3A construct proved to be less efficient than the other mammalian enzymes.

Mammalian APOBEC3As Induce Double DNA Breaks

Following cytidine deamination of nuDNA, UNG excision of uracil bases and subsequent APE1 cleavage of the DNA backbone at abasic sites can result in DSBs characterized by γ H2AX phosphorylation (Rogakou et al. 1998). Although well established for AID, it has recently been shown to occur following human A3A deamination (Landry et al. 2011; Aynaud et al. 2012). As most mammalian A3A enzymes can deaminate, nuDNA induction of γ H2AX positive DSBs was assessed for all A3A proteins. As can be seen, A3A expression resulted in numerous γ H2AX foci, detected in 30–60% of V5 positive cells (fig. 6A and B) with the exception of the tamarin A3A transfection, which only resulted in background levels of γ H2AX positive cells. This observation can be attributed to the low level of expression of this enzyme,

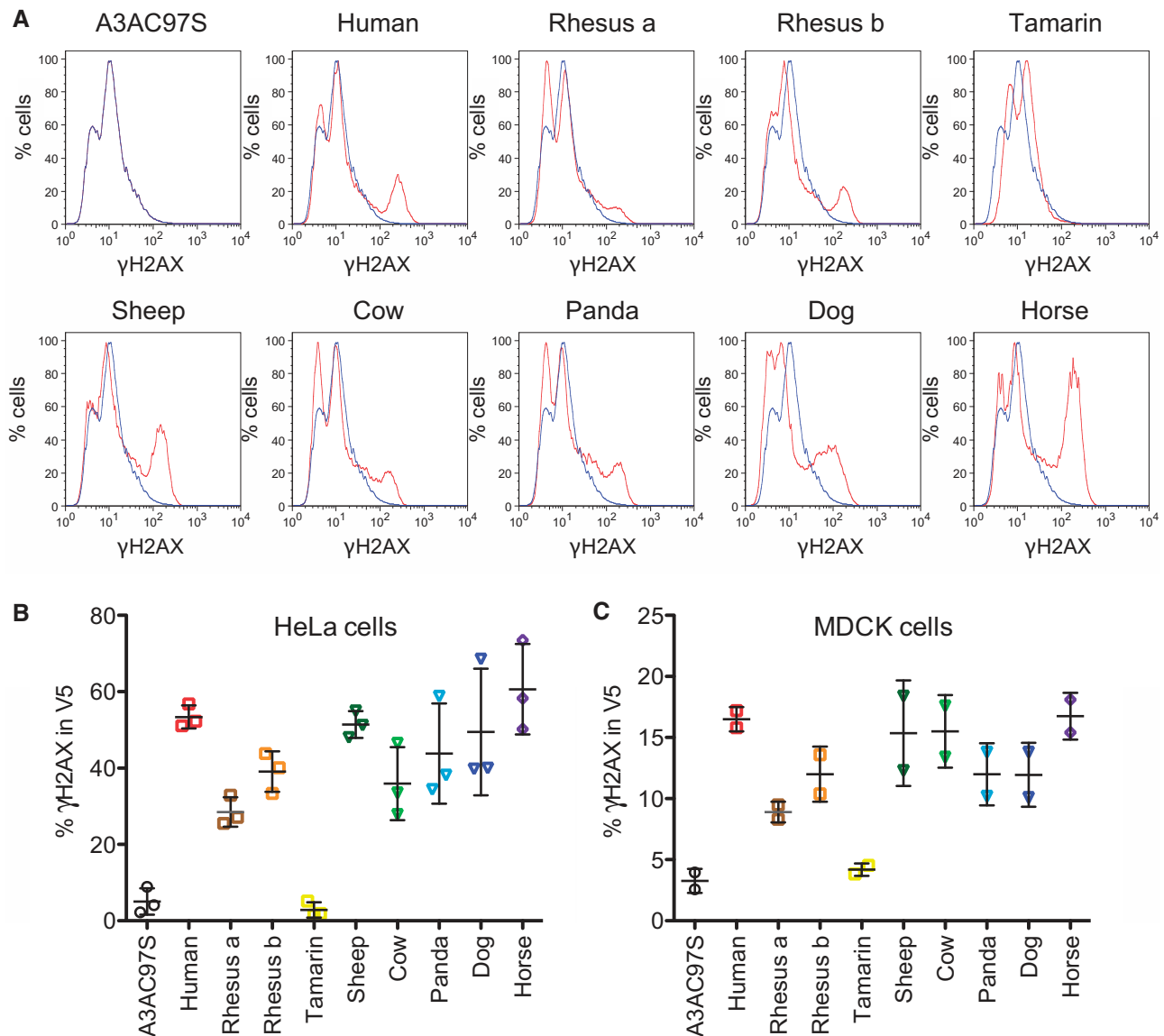


Fig. 6. A3A proteins can induce DSBs. (A) γ H2AX staining of A3A-transfected HeLa cells. Inactive human A3AC97S catalytic mutant is used as negative control. (B) FACS analysis of γ H2AX-positive HeLa cells gated on the V5-tagged A3A. (C) FACS analysis of γ H2AX-positive dog MDCK cells gated on the V5-tagged A3A.

responsible for modest nuDNA editing. If 3DPCR picks up rare events, γ H2AX staining and gating requires a substantial amount of DSB formation that may not be achieved after tamarin plasmid transfection, as DSB formation occurs during repair of numerous mutations in clusters. Accordingly, tamarin A3A's ability to induce DSB formation may well be underestimated, as H2AX staining was observed in few tamarin A3A positive cells by immunofluorescence imaging (fig. S4A). Similar evidence of A3A-sponsored DNA damage were obtained probing for p53-binding protein 1 (53BP1), a mediator protein recruited at DNA damage site, with the accumulation of numerous foci at DSB location after A3A transfection (fig. S4B) (Landry et al. 2011). To avoid any species or cell line effect, having so far only investigated nuDNA damage in human cells, we measured γ H2AX-positive DSBs in the canine Dog Madin Darby canine kidney (MDCK) cells

after A3A transfection (fig. 6C). The results were strongly correlated to those obtained with HeLa cells ($R^2 = 0.85$, Pearson correlation test $P = 0.0001$), the reduced levels probably resulting from a much reduced transfection efficiency.

The Carnivora APOBEC3 Locus

While analyzing the dog and panda A3 loci, we found the vestige of a cat A3A gene with complete exons 1, 2, 4, and 5 but with several indels in exon 3, including deletion of sequences encoding the crucial HAEX₂₃PCXXC zinc finger binding motif as well as splice acceptor mutations 5' to exons 3 and 4. As there were some unscored bases in intron 3, it was possible that the region was poorly sequenced or assembled. Accordingly, we amplified and sequenced exon 3, intron 3, and exon 4 from DNA of a primary cat tumor (*Felis domesticus*). Sequence alignment revealed numerous clustered

differences with the genomic contig (ISGSC *Felis_catus* 6.2/felCat5 assembly), suggesting errors rather than single nucleotide polymorphisms (supplementary fig. S5A, Supplementary Material online). However, the multiple indels in exon 3 were confirmed (supplementary fig. S5A and B, Supplementary Material online) as were the splice acceptor mutations pointing to a nonfunctional gene over a considerable period of time.

To date, the mammalian A3 locus is bounded by *CBX6* and *CBX7* with all A3 genes arranged in a head-to-tail manner (LaRue, Jonsson, et al. 2008). This is the case for the panda locus (*Ursinae* suborder, assembly BGI-Shenzhen 1.0/aiMel1) that encodes three A3 genes, *A3Z1* (i.e., *A3A*), *A3Z2*, and *A3Z3*, although the *A3Z1* gene is uniquely in the opposite orientation (supplementary fig. S1, Supplementary Material online). The dog A3 locus on chromosome 10 was composed of *A3Z1* and *A3Z3* genes (*Canis* suborder, assembly Broad CanFam3.1/canFam3), again with the *A3Z1* gene in an inverted orientation compared with *CBX6* and *CBX7*. However, there could well be assembly problems for there is a clearly recognizable *A3Z2* gene present in an isolated contig (UnJH373534). Reassembled, the dog locus is syntenic with the panda. Münk et al. (2008) have cogently argued for an expanded cat A3 locus with three *A3Z2* and a single *A3Z3* gene. As the defunct cat *A3Z1* gene is inverted compared with *CBX6* and *CBX7*, it appears that the three loci are highly syntenic with the cat locus showing further signs of gene conversion, the presumed organization of the loci being shown in supplementary figure S1, Supplementary Material online. However, it would be best to carefully reassemble the reads from the primary data, particularly for the dog locus.

Discussion

By a variety of different assays, all eight *A3Z1* enzymes are functional and orthologous to the human *A3A*, any differences being those of degree. Most importantly, the nuDNA editing function is conserved including deamination of 5-methylcytidine. As all are capable of hypermutation—up to 80% of target residues being edited—they must be seen as being involved in cell death and DNA catabolism, for the degree of mutation must overwhelm DNA damage responses (Landry et al. 2011; Mussil et al. 2013). Based on phylogenetic analyses, it has been cogently argued that the mammalian A3 precursor locus encoded a single copy of *A3Z1*, *A3Z2*, and *A3Z3*, indicating that the *A3Z1* nuDNA editing, DSB formation, and 5-methylcytidine deamination function goes back some 148 My (Munk et al. 2012). In turn, this demonstrates that the immediate benefits of catastrophic damage of somatic cell DNA—mutation and DSBs—far outweigh deleterious or pathological aspects over a longer time frame, notably genome instability and cancer.

As 3DPCR is not a quantitative technique, differences should not be overinterpreted, although the tamarin enzyme generally underperformed those of the others. It could be that the tamarin sequence represented a rare allele or contained a polymorphism or sequence error affecting its stability or function. However, as it could qualitatively perform all the functions of the human enzyme apart from

DSB formation, it is also probably an ortholog. As it can clearly edit nuDNA, DSBs should ensue. There are many levels by which genetic editing of nuDNA could vary among species. The Kozak context of the initiator methionines of natural *A3As* was “strong” in all sequences apart from the human and rhesus that were “adequate.” To permit comparisons, all genes were initiated from the same strong Kozak motif. The upstream promoter sequences can be reliably aligned for primate genomes but not across the mammals analyzed here. Accordingly, they could respond to different stimuli, although the sheep and rhesus *A3A* have a 5′ interferon response element just like its human counterpart.

Given the precursor *A3Z1* gene common to placental animals, it is logical that the mouse/rat, pig, and cat genomes have lost their *A3Z1* gene at some stage. Despite this, they still present with cancer, which begs the question as to whether there are other PCDs capable of editing nuDNA. The A3 locus has undergone considerable expansion and rearrangement with *A3A* vestiges in the human, horse, and cat genomes as well as duplication into the functional carboxy-terminal domains of human *A3B* and *A3G*, despite the loss of nuclear editing function. However, recent reports suggest that *A3B* can impact nuDNA, prompting base substitution and perhaps cancer (Shinohara et al. 2012; Burns et al. 2013). These results remain to be clarified as two other studied concomitantly reported that individuals with a homozygous *A3B* deletion had a higher odds ratio for developing breast or liver cancer (Zhang et al. 2012; Long et al. 2013).

Similarly, the human *A3DE* enzyme is nonfunctional while that from the macaque is functional, the difference residing in a single amino acid substitution, C320Y (Dang et al. 2011). In the same vein, the mouse *APOBEC1* (*A1*) deaminase is capable of hypermutating both viral RNA and ssDNA while the human enzyme can only hypermutate ssDNA (Petit et al. 2009). Accordingly, the reciprocal is perhaps possible, that is, gain of nuDNA editing function by some other mammalian PCD. Thus, while the mammalian *A3Z1* enzymes are clearly orthologous, the same may not be true for other *A3Z2* or *A3Z3* enzymes. *AID* and *A1* are PCDs both common to these species, with *AID* being the older of the two in evolutionary terms. Ectopic expression of *AID* is certainly linked to human cancers of the haematopoietic lineage (Matsumoto et al. 2007) while *AID* or *A1* transgenic mice generate cancers dictated by the promoter used in constructing the lineage (Yamanaka et al. 1995; Okazaki et al. 2003).

For the human locus *A3A* is the only PCD that efficiently deaminates nuDNA (Suspène, Aynaud, Guétard, et al. 2011), which does not rule out other PCDs being operative under certain circumstances, with human *AID* being the case in point. What is clear is that the physiological relevance of *A3A*-induced somatic hypermutation of host cell mtDNA and nuDNA far outweighs any pathological side effects.

Materials and Methods

Plasmids

Mammalian *A3A* cDNAs, from rhesus monkeys, tamarin, sheep, cow, dog, panda, and horse, were synthesized

(GeneCust) and subsequently cloned into pcDNA3.1D/V5-His-TOPO vector (Invitrogen) (supplementary table S1, Supplementary Material online). All plasmids were verified by sequencing.

Cells

The quail QT6 embryo fibroblast cell line was maintained in HAM's F40 medium (Eurobio), supplemented with 1% chicken serum, 10% fetal calf serum (FCS), 5% tryptose phosphate, 2 mM L-glutamine, 100 U/ml penicillin, and 100 mg/ml streptomycin. Human HeLa cells, 293T and 293T-UGI's cells stably expressing *Bacillus subtilis* phage UGI were maintained in DMEM glutamax medium (Invitrogen) supplemented with 10% FCS, 100 U/ml penicillin, and 100 mg/ml streptomycin. MDCK cells were maintained in DMEM glutamax medium (Invitrogen) supplemented with 10% FCS, 50 U/ml penicillin, and 50 mg/ml streptomycin.

Transfections

One million QT6 cells were cotransfected with 0.75 µg of pCayw HBV coding plasmid and 1.25 µg A3A expression plasmids using JetPrime (Polyplus) following manufacturer's recommendations and harvested 72 h posttransfection. For single plasmid transfections, 8×10^5 of HeLa, 293T-UGI, or MDCK cells were transfected using 2 µg A3A expression plasmids using JetPrime (Polyplus) following manufacturer's recommendations and harvested 72 h posttransfection. For the 5MeC deamination assay, a 679 bp fragment of HIV-1 pNL4.3 env gene was amplified using total substitution of dCTP by 5Me-dCTP (Trilink) using the primer pair 5'-TTGATGATCTGTAGTGCTACAGCA-3' and 5'-GCCTAATCCATGTGTACATGTGA-3'. The 5MeC containing DNA was heat denatured and chilled on ice, and 200 ng of synthesized DNA was transfected using JetPrime 24 h following initial transfection of A3 coding plasmids in QT6 cells as described earlier. For immunofluorescence labeling, coverslip-grown HeLa cells were transfected with 1 µg A3A expression plasmids using Fugene HD (Roche) following manufacturer's recommendations.

Western Blot Analysis

Western blot analysis was carried out according to standard procedures. After blocking, membranes were probed with 1/5,000 diluted mouse monoclonal antibody specific for the V5 epitope (Invitrogen) in phosphate buffer saline (PBS)-Tween 0.01%, and 5% dry milk was applied overnight. After PBS-Tween 0.1% washings and incubation with an anti-mouse IgG horseradish peroxidase-coupled secondary antibody (Amersham), the membrane was subjected to detection by enhanced chemiluminescence (Pierce). β -Actin was used as a loading control using 1/2,000 diluted mouse monoclonal antibody specific for β -Actin (Sigma).

Immunofluorescence

After PBS washing, coverslip-grown transfected HeLa cells were fixed with 4% paraformaldehyde (PFA) for 15 min. After PBS washing, cells were permeabilized with 0.1% Triton X-100 for 10 min, and the permeabilized cells were

incubated for 1 h with 1% bovine serum albumin (BSA) containing PBS. Mouse monoclonal anti-V5 antibody (Invitrogen) was then incubated at 1/200 for 1 h at room temperature, followed by incubation with a mouse-specific fluorescein isothiocyanate (FITC)-conjugated goat antibody for 1 h at room temperature. After washing, coverslips were stained with 4',6'-diamidino-2-phénylindole (DAPI) and mounted with Vectashield imaging medium (Vector Laboratories). Imaging was performed using Zeiss Wide-field inverted ApoTome microscope. A similar protocol was applied for DNA damage probing, using rabbit antibodies raised against phospho-histone H2AX (20E3, 1/400) and 53BP1 (1/100) (Cell Signaling) with rabbit-specific Cy3-conjugated goat antibody.

In Vitro Deamination Assay

At 48 h after transfection, APOBEC3-transfected 293T cells were extensively washed with PBS and mechanically harvested. Total proteins were extracted using specific lysis buffer (25 mM HEPES [pH 7.4], 10% glycerol, 150 mM NaCl, 0.5% Triton X-100, 1 mM EDTA, 1 mM $MgCl_2$, 1 mM $ZnCl_2$) supplemented with protease inhibitors. Deaminase activity was assessed by incubating whole cell lysates with 1 pmole DNA oligonucleotide 5'-(6-FAM)-AAATTCTAATAGATAATG TGA-(TAMRA)-3' in the presence of 0.4 U UDG (New England Biolabs) in a 20 mM Tris-HCl, 1 mM dithiothreitol, and 1 mM EDTA reaction buffer. After 2 h incubation at 37 °C, the generated abasic sites were cleaved by heating 2 min at 95 °C, and endpoint fluorescence were measured using realplex² mastercycler (Biorad) with FAM setting.

DNA Extraction and 3DPCR Amplification

Total DNA from transfected cells was extracted using the MasterPureTM complete DNA and RNA purification kit (Epicentre) and resuspended in 30 µl sterile water. All amplifications were performed using first-round standard PCR followed by nested 3DPCR (supplementary table S2, Supplementary Material online [Suspène, Henry, et al. 2005]). PCR was performed with 1 U Taq DNA polymerase (Eurobio) per reaction. After purification, PCR products were cloned onto TOPO 2.1 vector (Invitrogen) and sequencing was outsourced to GATC.

FACS Analysis of Double-Strand Breaks

At 48 h after transfection, cells were washed with PBS, fixed in 2–4% ice-cold paraformaldehyde (Electron Microscopy Sciences) for 10 min and permeabilized in 90% ice-cold methanol (Sigma) for 30 min. After washing with PBS, cells were incubated with 1:200 diluted mouse anti-V5 antibody (Invitrogen) in PBS-BSA 0.5% for 1 h. After PBS washings, incubation with 1:500 diluted Alexa Fluor 633 F(ab')₂ fragment of goat anti-mouse IgG (H+L) (Invitrogen) was performed for 45 min. DNA double-strand breaks were analyzed by staining for 1 h with 1:50 diluted Alexa Fluor 488-conjugated rabbit monoclonal anti- γ H2AX (20E3) antibody (Cell Signaling). All incubation steps were performed on ice. Stained samples were analyzed on a FACSCalibur using CellQuest Pro (BD

Biosciences, version 5.2) and data were analyzed with FlowJo software (Tree Star Inc. version 8.7.1).

Supplementary Material

Supplementary figures S1–S5 and tables S1–S2 are available at *Molecular Biology and Evolution* online (<http://www.mbe.oxfordjournals.org/>).

Acknowledgments

This work was supported by grants from the Institut Pasteur and Oséo (grant number FUI AAP 12). V.C. was supported by a postdoctoral grant from OSEO.

References

- Argyris EG, Acheampong E, Wang F, Huang J, Chen K, Mukhtar M, Zhang H. 2007. The interferon-induced expression of APOBEC3G in human blood-brain barrier exerts a potent intrinsic immunity to block HIV-1 entry to central nervous system. *Virology* 367:440–451.
- Aynaud MM, Suspène R, Vidalain PO, Mussil B, Guetard D, Tangy F, Wain-Hobson S, Vartanian JP. 2012. Human Tribbles 3 protects nuclear DNA from cytidine deamination by APOBEC3A. *J Biol Chem*. 287:39182–39192.
- Bininda-Emonds OR, Cardillo M, Jones KE, MacPhee RD, Beck RM, Grenyer R, Price SA, Vos RA, Gittleman JL, Purvis A. 2007. The delayed rise of present-day mammals. *Nature* 446:507–512.
- Bogerd HP, Wiegand HL, Hulme AE, Garcia-Perez JL, O'Shea KS, Moran JV, Cullen BR. 2006. Cellular inhibitors of long interspersed element 1 and Alu retrotransposition. *Proc Natl Acad Sci U S A*. 103:8780–8785.
- Bonvin M, Achermann F, Greeve I, et al. (11 co-authors). 2006. Interferon-inducible expression of APOBEC3 editing enzymes in human hepatocytes and inhibition of hepatitis B virus replication. *Hepatology* 43:1364–1374.
- Bulliard Y, Narvaiza I, Bertero A, et al. (13 co-authors). 2011. Structure-function analyses point to a polynucleotide-accommodating groove essential for APOBEC3A restriction activities. *J Virol*. 85:1765–1776.
- Burns MB, Lackey L, Carpenter MA, et al. (17 co-authors). 2013. APOBEC3B is an enzymatic source of mutation in breast cancer. *Nature* 494:366–370.
- Carpenter MA, Li M, Rathore A, et al. (12 co-authors). 2012. Methylcytosine and normal cytosine deamination by the foreign DNA restriction enzyme APOBEC3A. *J Biol Chem*. 287:34801–34808.
- Chen H, Lilley CE, Yu Q, Lee DV, Chou J, Narvaiza I, Landau NR, Weitzman MD. 2006. APOBEC3A is a potent inhibitor of adeno-associated virus and retrotransposons. *Curr Biol*. 16:480–485.
- Coticello SG, Harris RS, Neuberger MS. 2003. The Vif protein of HIV triggers degradation of the human antiretroviral DNA deaminase APOBEC3G. *Curr Biol*. 13:2009–2013.
- Coticello SG, Thomas CJ, Petersen-Mahrt SK, Neuberger MS. 2005. Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. *Mol Biol Evol*. 22:367–377.
- Dang Y, Abudu A, Son S, Harjes E, Spearman P, Matsuo H, Zheng YH. 2011. Identification of a single amino acid required for APOBEC3 antiretroviral cytidine deaminase activity. *J Virol*. 85:5691–5695.
- Harris RS, Bishop KN, Sheehy AM, Craig HM, Petersen-Mahrt SK, Watt IN, Neuberger MS, Malim MH. 2003. DNA deamination mediates innate immunity to retroviral infection. *Cell* 113:803–809.
- Henry M, Guetard D, Suspène R, Rusniok C, Wain-Hobson S, Vartanian JP. 2009. Genetic editing of HBV DNA by monodomain human APOBEC3 cytidine deaminases and the recombinant nature of APOBEC3G. *PLoS One* 4:e4277.
- Henry M, Terzian C, Peeters M, Wain-Hobson S, Vartanian JP. 2012. Evolution of the primate APOBEC3A cytidine deaminase gene and identification of related coding regions. *PLoS One* 7:e30036.
- Hultquist JF, Lengyel JA, Refsland EW, Larue RS, Lackey L, Brown WL, Harris RS. 2011. Human and rhesus APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H demonstrate a conserved capacity to restrict Vif-deficient HIV-1. *J Virol*. 84:11220–112234.
- Jarmuz A, Chester A, Bayliss J, Gisbourne J, Dunham I, Scott J, Navaratnam N. 2002. An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. *Genomics* 79: 285–296.
- Jones PA, Baylin SB. 2007. The epigenomics of cancer. *Cell* 128:683–692.
- Kaina B. 2003. DNA damage-triggered apoptosis: critical role of DNA repair, double-strand breaks, cell proliferation and signaling. *Biochem Pharmacol*. 66:1547–1554.
- Koning FA, Newman EN, Kim EY, Kunstman KJ, Wolinsky SM, Malim MH. 2009. Defining APOBEC3 expression patterns in human tissues and hematopoietic cell subsets. *J Virol*. 83:9474–9485.
- Landry S, Narvaiza I, Linfesty DC, Weitzman MD. 2011. APOBEC3A can activate the DNA damage response and cause cell-cycle arrest. *EMBO Rep*. 12:444–450.
- Larue RS, Andresdottir V, Blanchard Y, et al. (20 co-authors). 2008. Guidelines for naming non-primate APOBEC3 genes and proteins. *J Virol*. 83:494–497.
- LaRue RS, Jonsson SR, Silverstein KA, Lajoie M, Bertrand D, El-Mabrouk N, Hotzel I, Andresdottir V, Smith TP, Harris RS. 2008. The artiodactyl APOBEC3 innate immune repertoire shows evidence for a multifunctional domain organization that existed in the ancestor of placental mammals. *BMC Mol Biol*. 9:104.
- Lecossier D, Bouchonnet F, Clavel F, Hance AJ. 2003. Hypermutation of HIV-1 DNA in the absence of the Vif protein. *Science* 300:1112.
- Long J, Delahanty RJ, Li G, et al. (13 co-authors). 2013. A common deletion in the APOBEC3 genes and breast cancer risk. *J Natl Cancer Inst*. 105:573–579.
- Mangeat B, Turelli P, Caron G, Friedli M, Perrin L, Trono D. 2003. Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature* 424:99–103.
- Mariani R, Chen D, Schrofelbauer B, Navarro F, König R, Bollman B, Munk C, Nymark-McMahon H, Landau NR. 2003. Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif. *Cell* 114:21–31.
- Marin M, Rose KM, Kozak SL, Kabat D. 2003. HIV-1 Vif protein binds the editing enzyme APOBEC3G and induces its degradation. *Nat Med*. 9: 1398–1403.
- Matsumoto Y, Marusawa H, Kinoshita K, Endo Y, Kou T, Morisawa T, Azuma T, Okazaki IM, Honjo T, Chiba T. 2007. *Helicobacter pylori* infection triggers aberrant expression of activation-induced cytidine deaminase in gastric epithelium. *Nat Med*. 13:470–476.
- Muckenfuss H, Hamdorf M, Held U, Perkovic M, Lower J, Cichutek K, Flory E, Schumann GG, Munk C. 2006. APOBEC3 proteins inhibit human LINE-1 retrotransposition. *J Biol Chem*. 281:22161–22172.
- Munk C, Beck T, Zielonka J, et al. (13 co-authors). 2008. Functions, structure, and read-through alternative splicing of feline APOBEC3 genes. *Genome Biol*. 9:R48.
- Munk C, Willemsen A, Bravo IG. 2012. An ancient history of gene duplications, fusions and losses in the evolution of APOBEC3 mutators in mammals. *BMC Evol Biol*. 12:71.
- Mussil B, Suspène R, Aynaud M-M, Gauvrit A, Vartanian J-P, Wain-Hobson S. 2013. Human APOBEC3A isoforms translocate to the nucleus and induce DNA double strand breaks leading to cell stress and death. *PLoS One* 8:e73641.
- Nik-Zainal S, Alexandrov LB, Wedge DC, et al. (53 co-authors). 2012. Mutational processes molding the genomes of 21 breast cancers. *Cell* 149:979–993.
- Okazaki IM, Hiai H, Kakazu N, Yamada S, Muramatsu M, Kinoshita K, Honjo T. 2003. Constitutive expression of AID leads to tumorigenesis. *J Exp Med*. 197:1173–1181.
- Petit V, Guétard D, Renard M, Kerié A, Sitbon M, Wain-Hobson S, Vartanian JP. 2009. Murine APOBEC1 is a powerful mutator of retroviral and cellular RNA in vitro and in vivo. *J Mol Biol*. 385: 65–78.
- Refsland EW, Stenglein MD, Shindo K, Albin JS, Brown WL, Harris RS. 2010. Quantitative profiling of the full APOBEC3 mRNA repertoire in lymphocytes and tissues: implications for HIV-1 restriction. *Nucleic Acids Res*. 38:4274–4284.

- Rogakou EP, Pilch DR, Orr AH, Ivanova VS, Bonner WM. 1998. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J Biol Chem.* 273:5858–5868.
- Schmitt K, Guo K, Algaier M, Ruiz A, Cheng F, Qiu J, Wissing S, Santiago ML, Stephens EB. 2011. Differential virus restriction patterns of rhesus macaque and human APOBEC3A: implications for lentivirus evolution. *Virology* 419:24–42.
- Sheehy AM, Gaddis NC, Choi JD, Malim MH. 2002. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* 418:646–650.
- Sheehy AM, Gaddis NC, Malim MH. 2003. The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. *Nat Med.* 9:1404–1407.
- Shinohara M, Ito K, Shindo K, Matsui M, Sakamoto T, Tada K, Kobayashi M, Kadowaki N, Takaori-Kondo A. 2012. APOBEC3B can impair genomic stability by inducing base substitutions in genomic DNA in human cells. *Sci Rep.* 2:806.
- Stenglein MD, Burns MB, Li M, Lengyel J, Harris RS. 2010. APOBEC3 proteins mediate the clearance of foreign DNA from human cells. *Nat Struct Mol Biol.* 17:222–229.
- Suspène R, Aynaud MM, Guétard D, Henry M, Eckhoff G, Marchio A, Pineau P, Dejean A, Vartanian JP, Wain-Hobson S. 2011. Somatic hypermutation of human mitochondrial and nuclear DNA by APOBEC3 cytidine deaminases, a pathway for DNA catabolism. *Proc Natl Acad Sci U S A.* 108:4858–4863.
- Suspène R, Aynaud MM, Koch S, Padeloup D, Labetoulle M, Gaertner B, Vartanian JP, Meyerhans A, Wain-Hobson S. 2011. Genetic editing of herpes simplex virus 1 and Epstein-Barr herpesvirus genomes by human APOBEC3 cytidine deaminases in culture and in vivo. *J Virol.* 85:7594–7602.
- Suspène R, Aynaud M-M, Vartanian J-P, Wain-Hobson S. 2013. Efficient deamination of 5-methylcytidine and 5-substituted cytidine residues in DNA by human APOBEC3A cytidine deaminase. *PLoS One* 8: e63461.
- Suspène R, Guétard D, Henry M, Sommer P, Wain-Hobson S, Vartanian JP. 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.
- Suspène R, Henry M, Guillot S, Wain-Hobson S, Vartanian JP. 2005. Recovery of APOBEC3-edited human immunodeficiency virus G->A hypermutants by differential DNA denaturation PCR. *J Gen Virol.* 86:125–129.
- Thielen BK, McNevin JP, McElrath MJ, Hunt BV, Klein KC, Lingappa JR. 2010. Innate immune signaling induces high levels of TC-specific deaminase activity in primary monocyte-derived cells through expression of APOBEC3A isoforms. *J Biol Chem.* 285:27753–27766.
- Vartanian JP, Guétard D, Henry M, Wain-Hobson S. 2008. Evidence for editing of human papillomavirus DNA by APOBEC3 in benign and precancerous lesions. *Science* 320: 230–233.
- Vartanian JP, Henry M, Marchio A, et al. (12 co-authors). 2010. Massive APOBEC3 editing of hepatitis B viral DNA in cirrhosis. *PLoS Pathog.* 6:e1000928.
- Wijesinghe P, Bhagwat AS. 2012. Efficient deamination of 5-methylcytosines in DNA by human APOBEC3A, but not by AID or APOBEC3G. *Nucleic Acids Res.* 40:9206–9217.
- Yamanaka S, Balestra ME, Ferrell LD, Fan J, Arnold KS, Taylor S, Taylor JM, Innerarity TL. 1995. Apolipoprotein B mRNA-editing protein induces hepatocellular carcinoma and dysplasia in transgenic animals. *Proc Natl Acad Sci U S A.* 92:8483–8487.
- Zhang T, Cai J, Chang J, et al. (13 co-authors). 2012. Evidence of associations of APOBEC3B gene deletion with susceptibility to persistent HBV infection and hepatocellular carcinoma. *Hum Mol Genet.* 22: 1262–1269.
- Zielonka J, Bravo IG, Marino D, Conrad E, Perkovic M, Battenberg M, Cichutek K, Munk C. 2009. Restriction of equine infectious anemia virus by equine APOBEC3 cytidine deaminases. *J Virol.* 83: 7547–7559.