Powerful mutators lurking in the genome

Vincent Petit, Jean-Pierre Vartanian and Simon Wain-Hobson*

*Molecular Retrovirology Unit, CNRS URA 3015, Institut Pasteur, 28 rue du Docteur Roux,
75724 Paris cedex 15, France*

The human genome encodes numerous enzymes capable of deaminating polynucleotides. While they are capable of exquisite specificity, occasionally they result in hypermutation where up to 90 per cent of cytidine or adenosine residues may be edited. As such, they constitute a formidable anti-viral barrier, for no virus can survive such a high mutation rate. As the APOBEC3 group of cytidine deaminases edit single-stranded viral DNA, the crucial question is can they hyperedit chromosomal DNA? Everything points to a positive answer. Nonetheless, hypermutants *per se* have not yet been described, probably being countered by highly efficient mismatch repair. For the APOBEC3 genes, not only is their physiological function unknown, but also their role in the induction of cancer remains to be determined. Yet given the pace of research, all this is certain to change in the next few years.

Keywords: APOBEC3; activation-induced deaminase; cytidine deaminase; hypermutation; cancer

1. INTRODUCTION

Amine groups are vulnerable to modification such as alkylation, carboxylation and oxidation. They can be readily protonated and so provide excellent leaving groups. Of the five bases in nucleic acids, three have primary amino groups all of which are vulnerable to deamination. While this may form part of the metabolism of purines and pyrimidines as well as DNA and RNA precursors, once incorporated into polynucleotides, problems arise because these amino groups are involved in base pairing. Substitution of an amino for a keto group alters Watson–Crick base pairing. Thus, cytidine is changed to uridine, which base-pairs as thymidine, while adenosine is modified to yield inosine, which base-pairs mainly as guanosine. Guanosine deamination is not a major issue in nucleic acid biochemistry and will not be discussed further. In short, deamination can be mutagenic.

2. HOST ENZYMES WITH NUCLEIC ACIDS AS SUBSTRATES

Despite this, numerous cellular enzymes can deaminate RNAs with extraordinary specificity, such as the wobble position in the tRNA anticodon whereby an adenosine residue in the first base of the tRNA anticodon is edited to inosine by a series of genes collectively known as ADATs, adenosine deaminases acting on tRNA (Maas *et al.* 1999). As inosine is known to be able to base-pair with T, C and less readily A, it is able to decode codons with U or C in the third position. ADAR-1, -2 and -3 (ADAR=adenosine deaminase acting on RNA) can bind to double-stranded RNA (dsRNA) regions within cellular mRNAs and, by a base-flipping mechanism, specifically edit an adenosine residue (Valente & Nishikura

2005). Cytidine deamination too can be remarkably specific. For example, the human APOBEC1 enzyme specifically edits residue 6666 within the 14 121 base human apolipoprotein B (apoB) mRNA within intestinal tissue (Chen *et al.* 1987; Powell *et al.* 1987; Innerarity *et al.* 1996). In so doing, it generates a novel form of the protein that impacts triglyceride metabolism. In this particular case, specificity is conferred by another protein, APOBEC1 complementation factor. Specific editing is understandable, the evolutionary advantage is clear.

For the virologist, high mutation rates are part and parcel of RNA viruses where the genomic mutation rate can be as high as one or two per genome per round of replication (Drake 1993). While the mutation rate can be increased by a factor of 2 or so by the use of drugs (Holland *et al.* 1990), beyond that limit, the virus is mutated to death. This limit is referred to as the error threshold beyond which the informational integrity of the viral genome is irretrievably compromised (Biebricher & Eigen 2006). Among the 300 or more genes that are upregulated by type I interferons, is ADAR-1L (Samuel 2001). It is a variant of the constitutively expressed ADAR-1 gene. ADAR-1L binds to dsRNA and can edit as many as 90 per cent of adenosines within the same viral RNA. The prototypic example has always been the measles virus in cases of subacute sclerosing panencephalitis (Schmid *et al.* 1992). Subsequently, ‘blitzkrieg’ adenosine editing of cellular mRNA was also reported, particularly within dsRNA structures or mRNAs containing inverted Alu elements embedded within the 3′ untranslated region of mRNAs, or within introns of pre-mRNAs (Athanasiadis *et al.* 2004; Blow *et al.* 2004; Kim *et al.* 2004; Eisenberg *et al.* 2005; Levanon *et al.* 2005). The degree of editing ranges from a few bases to more than 40 per Alu.

Not surprisingly, such phenomena are referred to as hypermutants and the process hypermutation. The

* Author for correspondence (simon@pasteur.fr).

One contribution of 17 to a Discussion Meeting Issue ‘DNA deamination in immunity, virology and cancer’.

destruction of viral RNA via hypermutation, particularly by an interferon-induced gene, can be readily understood as part of an anti-viral response. The *raison d'être* of hyperediting mRNA is less clear and, although it could well be coupled to mRNA turnover, there is no clear picture for the moment (Scadden 2005). These observations were generally made by massive mRNA sequencing and computer comparisons with the human genome (Athanasiadis *et al.* 2004; Kim *et al.* 2004; Levanon *et al.* 2004).

Hyperediting of single-stranded DNA (ssDNA) has become *de rigueur* since the identification of activation-induced deaminase (AID) by Honjo's group back in 2000 (Muramatsu *et al.* 1999). There is no need to belabour the point in this volume. This remained a signal example although the homology to APOBEC1, with its specificity for ssRNA, was intriguing. Earlier work on transgenic mice bearing the human, mouse or rabbit APOBEC1 gene under the control of a liver-specific promoter clearly showed the development of hepatocellular carcinomas (HCC; Yamanaka *et al.* 1995) suggesting, but not proving, that the APOBEC1 enzyme might edit DNA (Harris *et al.* 2002; Petersen-Mahrt & Neuberger 2003). Interestingly, the same mice showed evidence of hyperediting of apoB mRNA, that is up to 10 edited C bases in the region surrounding residue 6666. Hence, the mechanism by which APOBEC1 expression resulted in these HCCs remained unclear. Direct evidence that APOBEC1 could edit DNA came from a sequence analysis of rifampicin-resistant clones generated by expressing APOBEC1 in *Escherichia coli* (Harris *et al.* 2002).

A second human APOBEC gene (APOBEC2) was identified in 1999, yet to date its physiological role remains unknown despite a beautiful crystal structure and knockout mice (Mikl *et al.* 2005; Prochnow *et al.* 2007). Navaratnam's group first identified the human APOBEC3 cluster on chromosome 22 in a remarkably complete paper without ascribing any function to the most studied member APOBEC3G (Jarmuz *et al.* 2002). They identified all but APOBEC3H (OhAinle *et al.* 2006). Given the remarkable sequence identity to APOBEC1, it was safe to conclude that APOBEC3s were cytidine deaminases but direct proof only came later (Harris *et al.* 2003; Lecossier *et al.* 2003; Mangeat *et al.* 2003; Mariani *et al.* 2003; Zhang *et al.* 2003; Suspène *et al.* 2004; Yu *et al.* 2004). However, APOBEC2, which shows considerable sequence homology to other cytidine deaminases with demonstrable activity, has yet to show positive in any deamination assay despite considerable study, which pleads in favour of a little 'hedge sitting'. In the meantime, the seminal paper from Mike Malim's laboratory highlighted the impact of APOBEC3G on the organization of the HIV genome (Sheehy *et al.* 2002). Although they noted that it belonged to the cytidine deaminase family, it took another crop of papers 2 years later to make the link with the long known phenomenon of retroviral G→A hypermutation (Harris *et al.* 2003; Lecossier *et al.* 2003; Mangeat *et al.* 2003; Mariani *et al.* 2003; Zhang *et al.* 2003). These are viral sequences in which between 10 and 60 per cent of G residues are monotonously substituted by A. They were so named to distinguish them from the A→I (or A→G) viral

hypermutants, notably the measles virus hypermutants mentioned above, which arose from ADAR deamination of RNA viral genomes (Suspène *et al.* 2008).

It transpires that over the course of time, two of these APOBEC3 genes—APOBEC3F and 3G—have been so devastating for the lentivirus subset of retroviruses that they evolved the unique *vif* gene to counter them. All but one of the lentiviruses encodes a *vif* gene. To a good first approximation, HIV-1 is resistant to all APOBEC3 cytidine deaminases although a recent report has suggested that wild-type virus is somewhat sensitive to a rare allele of APOBEC3H (OhAinle *et al.* 2008). APOBEC3F and A3G heavily impact HIV when the *vif* gene is either down or out (Harris *et al.* 2003; Lecossier *et al.* 2003; Mangeat *et al.* 2003; Mariani *et al.* 2003; Zhang *et al.* 2003; Liddament *et al.* 2004; Wiegand *et al.* 2004; Zheng *et al.* 2004). This helps explain the relative rarity of these sequences in HIV databases. This beautiful virological story is extensively developed in Michael Malim's accompanying paper. The HIV-1/A3F/A3G example parallels perfectly that of ADAR-1L editing of viral RNA. With the word parallel in mind, it is now clear that the A3G gene is strongly upregulated by type I and type II interferons, both *in vitro* and *in vivo* (Bonvin *et al.* 2006; Peng *et al.* 2006; Tanaka *et al.* 2006). Viral hypermutation indicates that nature, as opposed to Nature, has wised up to the fact that blitzkrieg mutation is yet another efficient way of fighting viruses.

3. SELECTIVE AMPLIFICATION OF CYTIDINE- AND ADENOSINE-EDITED HYPERMUTANTS

We realized that PCR could be simply adapted to select for G→A hypermutants. In standard PCR, the denaturation temperature (T_d) is invariably 94–95°C and held constant so that DNA is uniformly denatured. Yet it is widely known that AT-rich DNA melts at lower temperatures compared with GC-rich DNA. This is due to the lack of a third hydrogen bond (approx. one-third) and decreased stacking energy (approx. two-thirds). We reasoned that by modulating the T_d , conditions could be found allowing selective amplification of AT-rich APOBEC3 edited DNA to the detriment of unedited DNA. This turned out to be the case (figure 1a). Indeed, edited genomes at proportions as low as 10^{-4} could be selectively recovered (Suspène *et al.* 2005b). We have extensively used this method, referred to as 3DPCR, in our analyses of retroviral hypermutation (Mahieux *et al.* 2005; Suspène *et al.* 2005a,b; Bonvin *et al.* 2006; Delebecque *et al.* 2006; Suspène *et al.* 2008).

Inevitably, this made us wonder how the opposite could be achieved, that is the selective amplification of GC-rich DNA. Given the emerging importance of ADAR editing of a wide variety of RNAs, it would be useful to have a PCR-based method to allow selective amplification of GC-rich alleles. However, in view of the 3:2 hydrogen bonding rule for GC and AT base pairs, differential denaturation of target DNA would appear to be out of the question (Wain-Hobson 2006). Yet the beginnings to a solution lie in ADAR editing itself. Inosine base-pairs with cytidine through two hydrogen bonds rather than the three typical of a GC base pair (figure 2a).

Figure 1. Selective amplification of AT- and GC-rich DNA. (a) PCR amplification of molecular clones of HIV-1 differing by between 0 and 18 G→A transitions at the non-restrictive and restrictive temperatures of 95 and 83°C, respectively. At 83°C, only clones containing G→A transitions could be amplified. A sample of a HIVΔ*vif* virus from a molecular clone transfected on to 293T cells and subsequently used to infect PBMCs (peripheral blood mononuclear cells) was also amplified at 83°C. Cloning and sequencing showed the genomes to be G→A hypermutants resulting from APOBEC3G editing of viral cDNA. (b) 3DIPCR amplification of measles virus RNA produced in an interferon type I (i) insensitive (Vero) and (ii) sensitive (MRC5) cell line. (iv) For the MRC5 culture, DNA products could be recovered using a Td as low as 65°C, compared with (iii) the Vero control (67.4°C). Cloning and sequencing showed that the lower temperature products represent measles virus genomes edited by ADAR-1L.

Modified bases are often encountered in DNA bacteriophage genomes, usually as a means to avoid host restriction enzymes (Gommers-Ampt & Borst 1995). Invariably, modifications involve cytidine or thymidine, e.g. 5-hydroxymethyl cytidine in phage T4 DNA. There is, however, just one example of a modified purine, 2,6-diaminopurine, or 'D'. It is found in the cyanophage S-2L DNA genome where it totally substitutes for adenosine (Kirnos *et al.* 1977) and has the singular feature of base-pairing with thymidine (T) via three hydrogen bonds (figure 2b). As deoxyinosine

and deoxydiaminopurine triphosphates (dITP and dDTP) are commercially available, the outlines of a PCR-based method allowing selective amplification of GC-rich alleles becomes clear—a combination of differential denaturation PCR using the modified bases dITP and dDTP. Does it work? Indeed it does and the method is referred to as 3DIPCR, the inverse of 3DPCR (figure 1*b*). Now we have a pair of methods allowing amplification of AT- or GC-rich DNA.

These methods were applied to the problem of IgV hypermutation by AID (Suspène *et al.* 2008).

Figure 2. Chemistry of nucleotide deamination. (a) The products of APOBEC and ADAR deamination are uridine and inosine: (i) cytosine deamination C→U and (ii) adenine deamination A→I. (b) To enable PCR-based amplification of GC-rich DNA, two non-natural dNTPs are used. (i) Diaminopurine is an adenosine analogue and base-pairs with thymidine via three hydrogen bonds. (ii) Inosine pairs mainly with cytosine via two hydrogen bonds. Exploration of the PCR denaturation temperature allows selective amplification of GC-rich DNA, which in the form of TCID DNA, as opposed to TCGA DNA, melts at lower temperatures.

(a)

CDR1

V	T	I	T	C	R	A	S	Q	S	I	S	S	Y	L	N	W	Y	Q	Q	K	P
GTC	ACC	ATC	ACT	TGC	CGG	GCA	AGT	CAG	AGC	ATT	AGC	AGC	TAT	TTA	AAT	TGG	TAT	CAG	CAG	AAA	CCA
...	...	...	...	...	...	...	.A.	...	..T	...	..T	..T	.T.	...	...	...	...	...	...	...	...

CDR2

G	K	A	P	K	L	L	I	Y	A	A	S	S	L	Q	S	G	V	P	S	R	F
GGG	AAA	GCC	CCT	AAG	CTC	CTG	ATC	TAT	GCT	GCA	TCC	AGT	TTG	CAA	AGT	GGG	GTC	CCA	TCA	AGG	TTC
...	...	...	...	...	...	...	...	...	AT.	...	...	...	...	...	...	...	...	...	...	...	...
S	G	S	G	S	G	T	D	F	T	L	T	I									
AGT	GGC	AGT	GGA	TCT	GGG	ACA	GAT	TTC	ACT	CTC	ACC	ATC									
...	...	...	...	...	...	...	...	...	...	...	.T.	...									

(b)

	FR1	CDR1	FR2	CDR2	FR3	+	.	n.s.	s
A05	0	4	1	0	1	1	5	3	3
A2	0	4	0	2	1	5	2	4	3
C06	0	4	0	0	0	1	3	2	2
C07	1	0	2	0	0	1	2	2	1
E07	0	4	2	0	1	3	4	5	2
H10	1	0	0	4	0	2	3	3	2
C07	0	2	1	2	1	2	4	4	2
Σ	2	18	6	8	4	15	23	23	15

(c)

	0.2	0.1	C	+1
T	13	4	.	19
C	3	4	38	6
G	4	24	.	1
A	18	6	.	12
Con.	A	G	C	T
	T			A

Figure 3. AID ‘only’ edited transcribed human Vk1 sequences. (a) A single representative example is given. (b) Statistics of seven sequences bearing GC→AT transitions only. (c) The nucleotide context of the 38 GC→AT transitions resembles closely that of AID, which is WRCW, where W=A,T.

Complementary DNA corresponding to rearranged Vk1 regions from human CD14+ splenocytes was selectively amplified by 3DPCR and 3DIPCR. Although the latter yielded nothing in particular, 3DPCR recovered a series of exclusively AT-rich V region sequences (figure 3). Both strands were edited while the majority of mutations (68%) mapped to CDR1 and CDR2 (figure 3b). The local context for editing was WGCW, highly indicative of AID editing (figure 3c). These are among the few examples where

there is a trace of AID editing alone in IgV genes. What is particularly interesting is that they were derived from cDNA meaning the AID ‘only’ edited loci that were transcribed before DNA repair. As mentioned earlier, 3DPCR is acutely sensitive, down to 10⁻⁴, so these sequences could well reflect the exception. Although APOBEC3-associated hypermutation has been described for many retroviruses in a myriad of co-transfection experiments *in vitro*, for the moment, G→A hypermutants have only been identified *in vivo*

for HIV, other lentiviruses and hepatitis B virus (Meyerhans *et al.* 1989; Günther *et al.* 1997; Vartanian *et al.* 2002; Beale *et al.* 2004; Liddament *et al.* 2004; Kieffer *et al.* 2005; Noguchi *et al.* 2005; Suspène *et al.* 2005a,b, 2006; Kijak *et al.* 2008). A single hyperedited HTLV-1 sequence *in vivo* represents a rather singular case, it being derived from an experimentally infected New World squirrel monkey, which is not its natural host (Vartanian *et al.* 1997). As species-specific effects have been described *in vitro* concerning the restriction of lentiviruses, it is probably wise to put this example on hold. Although the primate spumaviruses appear to be quite sensitive to restriction by some APOBEC3 members *in vitro*, not a single hypermutant has been described *in vivo* or in virus cultures to date (Delebecque *et al.* 2006). A recent report of hyperedited human papillomavirus 1 and 16 DNA (HPV1 and HPV16) *in vivo* will be addressed below (Vartanian *et al.* 2008).

The database for APOBEC3-edited retroviral DNA is dominated by the primate lentiviruses, with HIV Δ vif at the top. In the absence of Vif, the proportion of edited genomes is approximately 100 per cent. Certainly, a SIV Δ vif virus fails to grow *in vivo* (Desrosiers *et al.* 1998). This is an interesting example. SIVs are found only among African monkeys and primates. The sooty mangabey virus (SIVsm) grows to high titres in its natural host without pathology. However, when transposed to the Asian Rhesus macaque, the virus was quickly adapted and resulted in overt pathology. This suggests that despite the many millions of years separating the two monkey lineages, the inter-host differences within the APOBEC3F and 3G genes were small fare for a rapidly evolving lentivirus such as SIV. For hepatitis B virus *in vivo*, the frequencies were very low, approximately 10^{-4} , while similar frequencies pertain to human papillomavirus (HPV). For the moment and from a virological standpoint, APOBEC3 cytidine deaminases are only efficient against the lentivirus subgroup of retroviruses.

4. APOBEC EDITING OF ENDOGENOUS RETROVIRUSES

Much has been written about the possible role of APOBEC3 genes in controlling endogenous retroviral transposition. It is an attractive hypothesis with much *in vitro* data in its favour (Esnault *et al.* 2005; Schumacher *et al.* 2005, 2008; Bogerd *et al.* 2006; Muckenfuss *et al.* 2006; Jern *et al.* 2007). The problem is the relative dearth of examples *in vivo*. Molecular biology is so powerful these days that hypothesis making can be a relatively easy affair. Hypothesis breaking, making sure that the observations are physiologically relevant, takes time and is frequently fraught with difficulties. The most straightforward rider is the fact that the avian lineage does not encode an APOBEC3 orthologue despite an abundance of endogenous retroviruses (ERVs) and a particularly famous exogenous retrovirus, Rous sarcoma virus. Reptile and fish genomes encode ERVs (Andersen *et al.* 1979; Huder *et al.* 2002; Shen & Steiner 2004); indeed some of the piscine ERVs are expressed *in vivo* (Paul *et al.* 2006). APOBEC3 cytidine

deaminases appear by the rodent lineage, where there is a single gene, moving up to six in equines (Bogerd *et al.* in press).

By contrast, *Macaca mulatta*, *Pan troglodytes* and *Homo sapiens* encode an expanded seven-gene APOBEC3 locus (A3A, A3B, A3C, A3DE, A3F, A3G and A3H). As *Papio*, *Cercopithecus* and some other monkey lineages encode clear homologues to A3G and A3H, it is a reasonable hypothesis for the moment that most monkey and Great Ape species encode seven APOBEC3 genes. An interesting 'problem' is afforded by baboon endogenous retrovirus (BaEV, *Papio* sp.), which has not only been identified in the peripheral blood of baboons but is also able to grow on many cell lines (Benveniste & Todaro 1974), suggesting that the virus is not restricted by baboon A3 deaminases. The explanation for this is the xenotropic nature of the surface envelope protein—the virus is unable to infect baboon cells yet is able to infect many non-baboon or foreign cells.

In an experimental setting, endogenous mouse intracisternal A type particles (IAPs) and MusD elements are vulnerable to hyperdeamination by a variety of APOBEC3 genes. However, the experimental system was human because the IAPs do not transpose well in murine cells. Showing a GC $\rightarrow$ AT mutation bias with an excess of G $\rightarrow$ A over C $\rightarrow$ T within IAPs compared with a consensus sequence is not *prima facie* evidence because similar biases can be found for pseudogenes compared with their cognate genes having corrected for CpG sites (Gojobori *et al.* 1982), mutated p53 alleles and disease alleles (Krawczak *et al.* 1995; Vartanian *et al.* 1997). As genes are generally GC rich, while transitions are the most frequent of mutations, a GC $\rightarrow$ AT bias is arguably the null hypothesis for any gene degradation process. Non-coding DNA is generally AT rich, which although proving nothing *per se*, is consistent with the notion that in mammals loss of information tends to end up as AT-rich DNA, if not deleted before.

What is perhaps a little surprising about extant analyses of ERVs is the relative dearth of hypermutants (Esnault *et al.* 2005; Jern *et al.* 2007; Jonsson *et al.* 2007; Armitage *et al.* 2008; Lee *et al.* 2008). The majority of ERVs known and studied to date are 'Moloney like' or 'mouse mammary tumour virus (MMTV) like'. Research on Moloney murine leukaemia virus (MoMLV) and MMTV go back to the pre-HIV days when retrovirology was effectively onco-retrovirology. When the first mammalian ERVs were described, it was noted that they had clear homology to either of these two murine retroviruses presumably reflecting past invasion of germinal cells. The effect of murine APOBEC3 on MMTV is slight with no deamination reported so far (Doehle *et al.* 2005; Abudu *et al.* 2006; Jern *et al.* 2007; Okeoma *et al.* 2007; Rulli *et al.* 2008). The APOBEC/MLV literature is highly variable, with a clear lack of consensus. Recent papers have suggested that one of the intrinsic resistance factors for Friend-MLV, *Rfv3*, is murine APOBEC3 (Santiago *et al.* 2008; Takeda *et al.* 2008). However, as there are multiple restricting factors for this retrovirus, their effects can only be dissected out in highly inbred mice. Whether or not mAPOBEC3 is

a dominant restriction factor in feral mice would seem to be an open question for the moment. In Balb/c Friend-sensitive mice, hyperdeamination of Fr-MLV is a rare event (Petit *et al.* 2009).

One glorious study showed that the rabbit genome harbours ERV remnants akin to a lentivirus, delightfully called RELIKs for rabbit endogenous lentivirus type K (Katzourakis *et al.* 2007). From a consideration of mutation rates, the authors concluded that RELIK had invaded the rabbit genome more than 7 Myr ago. From the present perspective, it must be noted *inter alia* that RELIK harboured a decayed *vif* gene meaning that the 'solution' to cytoplasmic APOBEC3 deaminases was solved a long time ago. It is not possible to know in which species this occurred for it could be argued that RELIK emerged in another species and crossed over to the rabbit.

5. APOBEC3 AND DNA VIRUSES

Not surprisingly, the current literature is dominated by work on retroviruses for that is where 'Hurricane APOBEC3' started. Double-stranded DNA comes alive through transcription and replication *via* ssDNA. As some APOBEC3 proteins can be found in the nucleus (Bogerd *et al.* 2006; Kinomoto *et al.* 2007), unlike A3F and A3G that are strongly cytoplasmic albeit with the occasional exceptions (Mangeat *et al.* 2003; Wiegand *et al.* 2004; Gallois-Montbrun *et al.* 2007, 2008), the question arises as to whether single-stranded cellular DNA can be edited by APOBEC3 deaminases. AID clearly can be specifically targeted to the IgV gene (Muramatsu *et al.* 2000). However, there is an ample literature demonstrating ectopic expression of AID in non-lymphoid tissues such as stomach (Matsumoto *et al.* 2007; Komori *et al.* 2008). Indeed, transgenic mice with AID under the control of a hepatotropic promoter give rise to HCCs, paralleling the finding for APOBEC1 transgenic mice (Yamanaka *et al.* 1995).

Albeit not within the APOBEC3 cluster, AID is to all intents and purposes an APOBEC3, or vice versa, in terms of exon/intron structure and phylogenetic relationships (figure 4). The most noticeable difference is the mutation fixation rate, which is extraordinarily low for AID compared with any of the APOBEC3 genes (Rogozin *et al.* 2005). So why could not some of the mainly nuclear human APOBEC3s restrict DNA viruses?

We were intrigued by the fact that a number of APOBEC3 genes were either upregulated or expressed in cutaneous keratinocytes and skin. For example, human A3A and hA3B are expressed in psoriatic keratinocytes, while hA3A is upregulated in acne lesions and can be induced by phorbol 12-myristate 13-acetate (Madsen *et al.* 1999; Trivedi *et al.* 2006). Interestingly, hA3H is also expressed in normal skin (Dang *et al.* 2006; OhAinle *et al.* 2006). From keratinocytes, it is only a short hop to HPVs. To a reasonable first approximation, HPVs can be generally divided into those that infect cutaneous or mucosal keratinocytes, a feature reflected in their phylogenetic clustering (de Villiers *et al.* 2004). Among the latter group are found the HPVs strongly associated with

Figure 4. Phylogenetic relationships among human APOBEC3 cytidine deamination domains (CDDs). The SplitsTree was made using protein sequences with APOBEC2 (hA2) and AID given as outliers. The CDD domains of the double-domain hA3s have been split and analysed as hA3Bn, hA3Bc, etc. where n and c indicate the N- and C-terminal domains, respectively. Those that edit HPV DNA have been highlighted.

cervical cancer. In a large European study, approximately 65 per cent and approximately 6 per cent of cervical cancers were associated with HPV16 and HPV18, respectively (Arbyn & Dillner 2007). In addition, for HPV31, 33, 35, 39, 45, 51, 52, 56, 58, 59 and 66, there is significant evidence that they can cause cervical cancer (Cogliano *et al.* 2005). To place ourselves in the best position to detect APOBEC3-edited HPV DNA, we selected the overlapping promoter/origin of replication region, which is located in the long non-coding region of all HPVs, based on the logic that this region would remain single-stranded for a longer period of time compared with other regions of the genome.

3DPCR was used to select for edited HPV DNA. We screened six HPV1a planter warts and found hypermutated DNA in one. Averaging across both strands, the mean cytidine editing frequency was approximately 11 per cent (range 4–61% per clone; figure 5a), which is highly comparable with the degree of editing seen for retroviruses. The dinucleotide context associated with editing was TpC and CpC (figure 5b). Such a preference is typical of APOBEC3 deaminases and antinomy for AID. To reduce the list of APOBEC3 candidates, transfection experiments were carried out with a full-length cloned HPV1a plasmid. Although there is no *in vitro* culture system for HPV, the promoter region can open up and support transcription.

It transpired that three out of seven human APOBEC3 deaminases, hA3A, hA3C and hA3H, were capable of editing HPV1a DNA. The mean cytidine editing frequencies *in vitro*, 25–29% (range 13–60% per clone), were approximately twofold higher than those *in vivo*, which may reflect stronger APOBEC3 gene expression driven by the powerful cytomegalovirus immediate-early promoter. Importantly, the editing context was almost identical to that found *in vivo*, supporting the hypothesis that editing arose as a result of an APOBEC3 enzyme (figure 5b). Yet we were unable to find traits that could help eliminate one or other of the enzymes. Indeed, as hA3A

Figure 5. APOBEC3 editing of HPV DNA *in vivo*. (a) Sample HPV1a.1. A small selection of HPV1a-edited DNA from a planter wart. Only sequence differences are given with respect to the unedited reference sequence. As only a fraction of the sequences are shown, the numbers to the right indicate the total number of mutations per sequence. (b) Dinucleotide context of HPV1a-edited sites where the dot indicates the edited site. The data for A3A, A3C and A3H were derived from transfection experiments. (c) Sample HPV16.29. A small selection of HPV16-edited DNA from a pre-cancerous cervical biopsy. Legend same as [figure 5a](#). The TATA box within the promoter region is highlighted. (d) Dinucleotide context of HPV16-edited sites where the dot indicates the edited site. Black bar, experimental; grey bar, *in vivo*.

and hA3H are expressed in keratinocytes, while hA3C is probably the most widely expressed of all the human A3 deaminases, there is no reason to exclude one or the other.

Although there is no data on APOBEC3 expression in mucosal keratinocytes, the finding of APOBEC3 editing of HPV1a led us to examine total DNA from nine HPV16-positive pre-cancerous cervical biopsies for evidence of hyperediting as opposed to cervical carcinoma tissue. The carcinoma harbours little HPV DNA and what there is is generally integrated. Furthermore, as the carcinoma is mono- or oligoclonal, sampling would be highly restricted. Of two out of nine samples, two showed evidence of hypermutation in the ori/promoter region (figure 5c). The mean editing frequency was approximately 9 per cent (range 5–24%). Once again, the dinucleotide editing context was TpC and CpC (figure 5d) indicating an APOBEC3 deaminase, possibly A3A, A3C and/or A3H.

Additional transfection experiments were performed using human APOBEC1 and 2 (hA1, hA2) as well as the murine deaminases mA1, mA2 and mA3. 3DPCR failed to identify any editing of HPV1a down to a frequency of 10^{-4} . Given the negative result for hA3B, and the fact that hA1 and mA1 traffic between the cytoplasm and nucleus, a nuclear localization *per se* is insufficient for deamination of HPV1a DNA.

These three human APOBEC3 enzymes encode a single cytidine deamination domain (CDD) as opposed

to hA3B, hA3DE, hA3F and hA3G that encode two CDDs. All three can be found within the cytoplasm and nucleus as can hA3B. Human A3DE, hA3F and hA3G are essentially cytoplasmic in location. Furthermore, A3A is phylogenically closely related to the active CDD2 domain of A3B while A3C is related to the active CDD2 domains of A3F and A3G, indicating that rather subtle differences must explain the HPV editing phenomenon by APOBEC3 enzymes (figure 4). It is curious to note that, similar to hA3A/C/H, AID too has a single CDD, and as mentioned earlier, AID is slightly closer to the APOBEC3 cluster of deaminases than APOBEC1, APOBEC2 or APOBEC4 (figure 4; Vartanian *et al.* 2008). Can AID edit HPV DNA *in vitro*? It can indeed, and the dinucleotide context is GpC and ApC just as it is for rearranged immunoglobulin V regions. In view of this, it is probable that AID is not involved in HPV deamination *in vivo*.

To resume, HPV genomic DNA can be edited by at least four human cytidine deaminases *in vitro* and possibly one to three *in vivo* and extends the range of viral targets for APOBEC3 enzymes to include double-stranded DNA viruses. It is a little surprising that the degree of editing can be as high as 61 per cent of potential targets when the whole panoply of DNA mismatch repair enzymes are omnipresent in the nucleus. For example, an edited IgV region with so many edited C residues has never been reported. These contrasting findings might be explained by the fact that

HPV genomes are packaged within the nucleus. It could be that the HPV hypermutants originate from neo-assembled viral cores where they would be sequestered from the mismatch repair machinery. Within the nucleus, G:U mismatch repair, which predominantly yields G:C, will mitigate against the effects of APOBEC3 editing. It is possible that a vestigial C → T transition reflects the opposing forces of APOBEC3 editing and G:U mismatch repair. Accordingly, the real frequency of hA3A/C/H editing should be considerably greater.

6. APOBEC3s AND CANCER

Transcriptionally active HPV genomes are in the form of mini-chromosomes complete with nucleosome. If they are vulnerable to cytidine deamination, say in a subset of cells with overexpressed hA3A/C/H, then it raises the intriguing question as to how chromosomal DNA is protected from these three nuclear APOBEC3 deaminases, which are, when given the chance, hugely efficient at editing ssDNA *in vivo* and *in vitro* (figure 5). Given the considerable literature linking ectopic expression of AID in a multitude of tissues to cancer (Kou *et al.* 2007; Matsumoto *et al.* 2007; Komori *et al.* 2008; Marusawa 2008; Morisawa *et al.* 2008), as well as AID editing outside of rearranged IgV regions, it appears that hA3A, hA3C and hA3H might well be implicated in a similar manner. As the three laws of thermodynamics tell us that no machine is perfect, it is plausible that they too, as AID, may occasionally turn against the human genome. Indeed this would seem to be the null hypothesis. While extremely efficient mutators, their effects are surely countered by highly efficient DNA repair machinery and so it may be extremely difficult to identify hyperedited human genes. As mentioned earlier, HPV hypermutants could well have been sequestered from the repair machinery by the packaging of the deaminated DNA. From this perspective, hyperedited dsDNA genomes could serve as surrogate markers for these powerful mutators. It seems necessary to investigate their promoters and identify the transcription factors involved as well as to establish their interactomes.

It is interesting that edited HPV genomes, whether *in vitro* or *in vivo*, represent low-frequency phenomena, and squares with the notion that these events arise from rare situations where control breaks down. That the degree of editing of some HPV genomes *in vivo* was strictly comparable with those *in vitro*, where the powerful cytomegalovirus immediate-early gene promoter drives gene expression, suggests that there must be a few cells *in vivo* where the expression of A3A/C/H must be very elevated indeed. In this context, it has been noted that numerous APOBEC3 genes are upregulated in tumour cell lines (Jarmuz *et al.* 2002).

It has not escaped the attention of many in the cancer genomics field that there are far more mutations in a cancer genome than hitherto suspected. Furthermore, the majority of mutations are C → T transitions (Greenman *et al.* 2007). Given the power of mismatch repair, a single C → T transition could reflect the vestige of local AID/APOBEC3 editing. Hyperediting of genomic DNA in sites of chronic inflammation would

set up a broad mutant spectrum from which a cancer genome could emerge. Yet hypermutation is unlike the sequential accumulation of point mutations, which is akin to the series WORD → WORE → GORE → GONE → GENE (Maynard Smith 1970). Hypermutation, even if severely corrected, allows jumps through sequence space that could propel the cell faster along the road to cancer.

We would like to thank past and present members of the laboratory for their support and company. This work was supported by grants from the Institut Pasteur, the ANRS and the CNRS.

REFERENCES

- Abudu, A., Takaori-Kondo, A., Izumi, T., Shirakawa, K., Kobayashi, M., Sasada, A., Fukunaga, K. & Uchiyama, T. 2006 Murine retrovirus escapes from murine APOBEC3 via two distinct novel mechanisms. *Curr. Biol.* **16**, 1565–1570. (doi:10.1016/j.cub.2006.06.055)
- Andersen, P. R., Barbacid, M., Tronick, S. R., Clark, H. F. & Aaronson, S. A. 1979 Evolutionary relatedness of viper and primate endogenous retroviruses. *Science* **204**, 318–321. (doi:10.1126/science.219480)
- Arbyn, M. & Dillner, J. 2007 Review of current knowledge on HPV vaccination: an appendix to the European Guidelines for Quality Assurance in Cervical Cancer Screening. *J. Clin. Virol.* **38**, 189–197. (doi:10.1016/j.jcv.2006.12.009)
- Armitage, A. E. *et al.* 2008 Conserved footprints of APOBEC3G on hypermutated human immunodeficiency virus type 1 and human endogenous retrovirus HERV-K(HML2) sequences. *J. Virol.* **82**, 8743–8761. (doi:10.1128/JVI.00584-08)
- Athanasiadis, A., Rich, A. & Maas, S. 2004 Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome. *PLoS Biol.* **2**, e391. (doi:10.1371/journal.pbio.0020391)
- Beale, R. C., Petersen-Mahrt, S. K., Watt, I. N., Harris, R. S., Rada, C. & Neuberger, M. S. 2004 Comparison of the differential context-dependence of DNA deamination by APOBEC enzymes: correlation with mutation spectra *in vivo*. *J. Mol. Biol.* **337**, 585–596. (doi:10.1016/j.jmb.2004.01.046)
- Benveniste, R. E. & Todaro, G. J. 1974 Evolution of type C viral genes: I. Nucleic acid from baboon type C virus as a measure of divergence among primate species. *Proc. Natl Acad. Sci. USA* **71**, 4513–4518. (doi:10.1073/pnas.71.11.4513)
- Biebricher, C. K. & Eigen, M. 2006 What is a quasispecies? *Curr. Top. Microbiol. Immunol.* **299**, 1–31. (doi:10.1007/3-540-26397-7_1)
- Blow, M., Futreal, P. A., Wooster, R. & Stratton, M. R. 2004 A survey of RNA editing in human brain. *Genome Res.* **14**, 2379–2387. (doi:10.1101/gr.2951204)
- Bogerd, H. P., Wiegand, H. L., Hulme, A. E., Garcia-Perez, J. L., O'Shea, K. S., Moran, J. V. & Cullen, B. R. 2006 Cellular inhibitors of long interspersed element 1 and Alu retrotransposition. *Proc. Natl Acad. Sci. USA* **103**, 8780–8785. (doi:10.1073/pnas.0603313103)
- Bogerd, H. P., Tallmadge, R. L., Oaks, J. L., Carpenter, S. & Cullen, B. R. In press. Equine infectious anemia virus resists the anti-retroviral activity of equine APOBEC3 proteins through a packaging-independent mechanism. *J. Virol.*
- 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. (doi:10.1002/hep.21187)

- Chen, S. H. *et al.* 1987 Apolipoprotein B-48 is the product of a messenger RNA with an organ-specific in-frame stop codon. *Science* **238**, 363–366. (doi:10.1126/science.3659919)
- Cogliano, V., Baan, R., Straif, K., Grosse, Y., Secretan, B., El Ghissassi, F. & WHO International Agency for Research on Cancer 2005 Carcinogenicity of human papillomaviruses. *Lancet Oncol.* **6**, 204.
- Dang, Y., Wang, X., Esselman, W. J. & Zheng, Y. H. 2006 Identification of APOBEC3DE as another antiretroviral factor from the human APOBEC family. *J. Virol.* **80**, 10 522–10 533. (doi:10.1128/JVI.01123-06)
- Delebecque, F., Suspène, R., Calattini, S., Casartelli, N., Saib, A., Froment, A., Wain-Hobson, S., Gessain, A., Vartanian, J. P. & Schwartz, O. 2006 Restriction of foamy viruses by APOBEC cytidine deaminases. *J. Virol.* **80**, 605–614. (doi:10.1128/JVI.80.2.605-614.2006)
- Desrosiers, R. C., Lifson, J. D., Gibbs, J. S., Czajak, S. C., Howe, A. Y., Arthur, L. O. & Johnson, R. P. 1998 Identification of highly attenuated mutants of simian immunodeficiency virus. *J. Virol.* **72**, 1431–1437.
- de Villiers, E. M., Fauquet, C., Broker, T. R., Bernard, H. U. & zur Hausen, H. 2004 Classification of papillomaviruses. *Virology* **324**, 17–27. (doi:10.1016/j.virol.2004.03.033)
- Doehle, B. P., Schafer, A., Wiegand, H. L., Bogerd, H. P. & Cullen, B. R. 2005 Differential sensitivity of murine leukemia virus to APOBEC3-mediated inhibition is governed by virion exclusion. *J. Virol.* **79**, 8201–8207. (doi:10.1128/JVI.79.13.8201-8207.2005)
- Drake, J. W. 1993 Rates of spontaneous mutation among RNA viruses. *Proc. Natl Acad. Sci. USA* **90**, 4171–4175. (doi:10.1073/pnas.90.9.4171)
- Eisenberg, E., Nemzer, S., Kinar, Y., Sorek, R., Rechavi, G. & Levanon, E. Y. 2005 Is abundant A-to-I RNA editing primate-specific? *Trends Genet.* **21**, 77–81. (doi:10.1016/j.tig.2004.12.005)
- Esnault, C., Heidmann, O., Delebecque, F., Dewannieux, M., Ribet, D., Hance, A. J., Heidmann, T. & Schwartz, O. 2005 APOBEC3G cytidine deaminase inhibits retrotransposition of endogenous retroviruses. *Nature* **433**, 430–433. (doi:10.1038/nature03238)
- Gallois-Montbrun, S., Kramer, B., Swanson, C. M., Byers, H., Lynham, S., Ward, M. & Malim, M. H. 2007 Antiviral protein APOBEC3G localizes to ribonucleoprotein complexes found in P bodies and stress granules. *J. Virol.* **81**, 2165–2178. (doi:10.1128/JVI.02287-06)
- Gallois-Montbrun, S., Holmes, R. K., Swanson, C. M., Fernandez-Ocana, M., Byers, H. L., Ward, M. A. & Malim, M. H. 2008 Comparison of cellular ribonucleoprotein complexes associated with the APOBEC3F and APOBEC3G antiviral proteins. *J. Virol.* **82**, 5636–5642. (doi:10.1128/JVI.00287-08)
- Gojobori, T., Li, W. H. & Graur, D. 1982 Patterns of nucleotide substitution in pseudogenes and functional genes. *J. Mol. Evol.* **18**, 360–369. (doi:10.1007/BF01733904)
- Gommers-Ampt, J. H. & Borst, P. 1995 Hypermodified bases in DNA. *FASEB J.* **9**, 1034–1042.
- Greenman, C. *et al.* 2007 Patterns of somatic mutation in human cancer genomes. *Nature* **446**, 153–158. (doi:10.1038/nature05610)
- Günther, S., Sommer, G., Plikat, U., Iwanska, A., Wain-Hobson, S., Will, H. & Meyerhans, A. 1997 Naturally occurring hepatitis B virus genomes bearing the hallmarks of retroviral G→A hypermutation. *Virology* **235**, 104–108. (doi:10.1006/viro.1997.8676)
- Harris, R. S., Petersen-Mahrt, S. K. & Neuberger, M. S. 2002 RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Mol. Cell* **10**, 1247–1253. (doi:10.1016/S1097-2765(02)00742-6)
- Harris, R. S., Bishop, K. N., Sheehy, A. M., Craig, H. M., Petersen-Mahrt, S. K., Watt, I. N., Neuberger, M. S. & Malim, M. H. 2003 DNA deamination mediates innate immunity to retroviral infection. *Cell* **113**, 803–809. (doi:10.1016/S0092-8674(03)00423-9)
- Holland, J. J., Domingo, E., de la Torre, J. C. & Steinhauer, D. A. 1990 Mutation frequencies at defined single codon sites in vesicular stomatitis virus and poliovirus can be increased only slightly by chemical mutagenesis. *J. Virol.* **64**, 3960–3962.
- Huder, J. B., Boni, J., Hatt, J. M., Soldati, G., Lutz, H. & Schupbach, J. 2002 Identification and characterization of two closely related unclassifiable endogenous retroviruses in pythons (*Python molurus* and *Python curtus*). *J. Virol.* **76**, 7607–7615. (doi:10.1128/JVI.76.15.7607-7615.2002)
- Innerarity, T. L., Boren, J., Yamanaka, S. & Olofsson, S. O. 1996 Biosynthesis of apolipoprotein B48-containing lipoproteins. Regulation by novel post-transcriptional mechanisms. *J. Biol. Chem.* **271**, 2353–2356. (doi:10.1074/jbc.271.5.2353)
- 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. (doi:10.1006/geno.2002.6718)
- Jern, P., Stoye, J. P. & Coffin, J. M. 2007 Role of APOBEC3 in genetic diversity among endogenous murine leukemia viruses. *PLoS Genet.* **3**, 2014–2022. (doi:10.1371/journal.pgen.0030183)
- Jonsson, S. R., LaRue, R. S., Stenglein, M. D., Fahrenkrug, S. C., Andresdottir, V. & Harris, R. S. 2007 The restriction of zoonotic PERV transmission by human APOBEC3G. *PLoS ONE* **2**, e893. (doi:10.1371/journal.pone.0000893)
- Katzourakis, A., Tristem, M., Pybus, O. G. & Gifford, R. J. 2007 Discovery and analysis of the first endogenous lentivirus. *Proc. Natl Acad. Sci. USA* **104**, 6261–6265. (doi:10.1073/pnas.0700471104)
- Kieffer, T. L., Kwon, P., Nettles, R. E., Han, Y., Ray, S. C. & Siliciano, R. F. 2005 G→A hypermutation in protease and reverse transcriptase regions of human immunodeficiency virus type 1 residing in resting CD4+T cells *in vivo*. *J. Virol.* **79**, 1975–1980. (doi:10.1128/JVI.79.3.1975-1980.2005)
- Kijak, G. H., Janini, L. M., Tovanabutra, S., Sanders-Buell, E., Arroyo, M. A., Robb, M. L., Michael, N. L., Bix, D. L. & McCutchan, F. E. 2008 Variable contexts and levels of hypermutation in HIV-1 proviral genomes recovered from primary peripheral blood mononuclear cells. *Virology* **376**, 101–111. (doi:10.1016/j.virol.2008.03.017)
- Kim, D. D., Kim, T. T., Walsh, T., Kobayashi, Y., Matise, T. C., Buyske, S. & Gabriel, A. 2004 Widespread RNA editing of embedded alu elements in the human transcriptome. *Genome Res.* **14**, 1719–1725. (doi:10.1101/gr.2855504)
- Kinamoto, M., Kanno, T., Shimura, M., Ishizaka, Y., Kojima, A., Kurata, T., Sata, T. & Tokunaga, K. 2007 All APOBEC3 family proteins differentially inhibit LINE-1 retrotransposition. *Nucleic Acids Res.* **35**, 2955–2964. (doi:10.1093/nar/gkm181)
- Kirnos, M. D., Khudyakov, I. Y., Alexandrushkina, N. I. & Vanyushin, B. F. 1977 2-Amino adenine is an adenine substituting for a base in S-2L cyanophage DNA. *Nature* **270**, 369–370. (doi:10.1038/270369a0)
- Komori, J. *et al.* 2008 Activation-induced cytidine deaminase links bile duct inflammation to human cholangiocarcinoma. *Hepatology* **47**, 888–896. (doi:10.1002/hep.22125)

- Kou, T. *et al.* 2007 Expression of activation-induced cytidine deaminase in human hepatocytes during hepatocarcinogenesis. *Int. J. Cancer* **120**, 469–476. (doi:10.1002/ijc.22292)
- Krawczak, M., Smith-Sorensen, B., Schmidtke, J., Kakkar, V. V., Cooper, D. N. & Hovig, E. 1995 Somatic spectrum of cancer-associated single basepair substitutions in the TP53 gene is determined mainly by endogenous mechanisms of mutation and by selection. *Hum. Mutat.* **5**, 48–57. (doi:10.1002/humu.1380050107)
- Lecossier, D., Bouchonnet, F., Clavel, F. & Hance, A. J. 2003 Hypermutation of HIV-1 DNA in the absence of the Vif protein. *Science* **300**, 1112. (doi:10.1126/science.1083338)
- Lee, Y. N., Malim, M. H. & Bieniasz, P. D. 2008 Hypermutation of an ancient human retrovirus by APOBEC3G. *J. Virol.* **82**, 8762–8770. (doi:10.1128/JVI.00751-08)
- Levanon, E. Y. *et al.* 2004 Systematic identification of abundant A-to-I editing sites in the human transcriptome. *Nat. Biotechnol.* **22**, 1001–1005. (doi:10.1038/nbt996)
- Levanon, E. Y., Hallegger, M., Kinar, Y., Shemesh, R., Djinic-Carugo, K., Rechavi, G., Jantsch, M. F. & Eisenberg, E. 2005 Evolutionarily conserved human targets of adenosine to inosine RNA editing. *Nucleic Acids Res.* **33**, 1162–1168. (doi:10.1093/nar/gki239)
- Liddament, M. T., Brown, W. L., Schumacher, A. J. & Harris, R. S. 2004 APOBEC3F properties and hypermutation preferences indicate activity against HIV-1 *in vivo*. *Curr. Biol.* **14**, 1385–1391. (doi:10.1016/j.cub.2004.06.050)
- Maas, S., Gerber, A. P. & Rich, A. 1999 Identification and characterization of a human tRNA-specific adenosine deaminase related to the ADAR family of pre-mRNA editing enzymes. *Proc. Natl Acad. Sci. USA* **96**, 8895–8900. (doi:10.1073/pnas.96.16.8895)
- Madsen, P. *et al.* 1999 Psoriasis upregulated phorbol-1 shares structural but not functional similarity to the mRNA-editing protein APOBEC-1. *J. Invest. Dermatol.* **113**, 162–169. (doi:10.1046/j.1523-1747.1999.00682.x)
- Mahieux, R., Suspène, R., Delebecque, F., Henry, M., Schwartz, O., Wain-Hobson, S. & Vartanian, J. P. 2005 Extensive editing of a small fraction of human T-cell leukemia virus type 1 genomes by four APOBEC3 cytidine deaminases. *J. Gen. Virol.* **86**, 2489–2494. (doi:10.1099/vir.0.80973-0)
- 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. (doi:10.1038/nature01709)
- Mariani, R., Chen, D., Schrofelbauer, B., Navarro, F., König, R., Bollman, B., Munk, C., Nymark-McMahon, H. & Landau, N. R. 2003 Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif. *Cell* **114**, 21–31. (doi:10.1016/S0092-8674(03)00515-4)
- Marusawa, H. 2008 Aberrant AID expression and human cancer development. *Int. J. Biochem. Cell Biol.* **40**, 1399–1402. (doi:10.1016/j.biocel.2008.01.018)
- Matsumoto, Y. *et al.* 2007 *Helicobacter pylori* infection triggers aberrant expression of activation-induced cytidine deaminase in gastric epithelium. *Nat. Med.* **13**, 470–476. (doi:10.1038/nm1566)
- Maynard Smith, J. 1970 Natural selection and the concept of a protein space. *Nature* **225**, 563–564. (doi:10.1038/225563a0)
- Meyerhans, A., Cheyner, R., Albert, J., Seth, M., Kwok, S., Sninsky, J., Morfeldt-Manson, L., Asjo, B. & Wain-Hobson, S. 1989 Temporal fluctuations in HIV quasispecies *in vivo* are not reflected by sequential HIV isolations. *Cell* **58**, 901–910. (doi:10.1016/0092-8674(89)90942-2)
- Mikl, M. C., Watt, I. N., Lu, M., Reik, W., Davies, S. L., Neuberger, M. S. & Rada, C. 2005 Mice deficient in APOBEC2 and APOBEC3. *Mol. Cell Biol.* **16**, 7270–7277. (doi:10.1128/MCB.25.16.7270-7277.2005)
- Morisawa, T., Marusawa, H., Ueda, Y., Iwai, A., Okazaki, I. M., Honjo, T. & Chiba, T. 2008 Organ-specific profiles of genetic changes in cancers caused by activation-induced cytidine deaminase expression. *Int. J. Cancer* **123**, 2735–2740. (doi:10.1002/ijc.23853)
- Muckenfuss, H., Hamdorf, M., Held, U., Perkovic, M., Lower, J., Cichutek, K., Flory, E., Schumann, G. G. & Munk, C. 2006 APOBEC3 proteins inhibit human LINE-1 retrotransposition. *J. Biol. Chem.* **281**, 22 161–22 172. (doi:10.1074/jbc.M601716200)
- Muramatsu, M., Sankaranand, V. S., Anant, S., Sugai, M., Kinoshita, K., Davidson, N. O. & Honjo, T. 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**, 18 470–18 476. (doi:10.1074/jbc.274.26.18470)
- Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y. & Honjo, T. 2000 Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* **102**, 553–563. (doi:10.1016/S0092-8674(00)00078-7)
- Noguchi, C., Ishino, H., Tsuge, M., Fujimoto, Y., Imamura, M., Takahashi, S. & Chayama, K. 2005 G to A hypermutation of hepatitis B virus. *Hepatology* **41**, 626–633. (doi:10.1002/hep.20580)
- OhAinle, M., Kerns, J. A., Malik, H. S. & Emerman, M. 2006 Adaptive evolution and antiviral activity of the conserved mammalian cytidine deaminase APOBEC3H. *J. Virol.* **80**, 3853–3862. (doi:10.1128/JVI.80.8.3853-3862.2006)
- OhAinle, M., Kerns, J. A., Li, M. M., Malik, H. S. & Emerman, M. 2008 Antiretroelement activity of APOBEC3H was lost twice in recent human evolution. *Cell Host Microbe* **4**, 249–259. (doi:10.1016/j.chom.2008.07.005)
- Okeoma, C. M., Lovsin, N., Peterlin, B. M. & Ross, S. R. 2007 APOBEC3 inhibits mouse mammary tumour virus replication *in vivo*. *Nature* **445**, 927–930. (doi:10.1038/nature05540)
- Paul, T. A., Quackenbush, S. L., Sutton, C., Casey, R. N., Bowser, P. R. & Casey, J. W. 2006 Identification and characterization of an exogenous retrovirus from atlantic salmon swim bladder sarcomas. *J. Virol.* **80**, 2941–2948. (doi:10.1128/JVI.80.6.2941-2948.2006)
- Peng, G., Lei, K. J., Jin, W., Greenwell-Wild, T. & Wahl, S. M. 2006 Induction of APOBEC3 family proteins, a defensive maneuver underlying interferon-induced anti-HIV-1 activity. *J. Exp. Med.* **203**, 41–46. (doi:10.1084/jem.20051512)
- Petersen-Mahrt, S. K. & Neuberger, M. S. 2003 *In vitro* deamination of cytosine to uracil in single-stranded DNA by apolipoprotein B editing complex catalytic subunit 1 (APOBEC1). *J. Biol. Chem.* **278**, 19 583–19 586. (doi:10.1074/jbc.C300114200)
- Petit, V., Guétard, D., Renard, M., Kerié, A., Sitbon, M., Wain-Hobson, S. & Vartanian, J. P. 2009 Murine APOBEC1 is a powerful mutator of retroviral and cellular RNA *in vitro* and *in vivo*. *J. Mol. Biol.* **385**, 65–78. (doi:10.1016/j.jmb.2008.10.043)
- Powell, L. M., Wallis, S. C., Pease, R. J., Edwards, Y. H., Knott, T. J. & Scott, J. 1987 A novel form of tissue-specific

- RNA processing produces apolipoprotein-B48 in intestine. *Cell* **50**, 831–840. (doi:10.1016/0092-8674(87)90510-1)
- Prochnow, C., Bransteitter, R., Klein, M. G., Goodman, M. F. & Chen, X. S. 2007 The APOBEC-2 crystal structure and functional implications for the deaminase AID. *Nature* **445**, 447–451. (doi:10.1038/nature05492)
- Rogozin, I. B., Basu, M. K., Jordan, I. K., Pavlov, Y. I. & Koonin, E. V. 2005 APOBEC4, a new member of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases predicted by computational analysis. *Cell Cycle* **4**, 1281–1285.
- Rulli Jr, S. J., Mirro, J., Hill, S. A., Lloyd, P., Gorelick, R. J., Coffin, J. M., Derse, D. & Rein, A. 2008 Interactions of murine APOBEC3 and human APOBEC3G with murine leukemia viruses. *J. Virol.* **82**, 6566–6575. (doi:10.1128/JVI.01357-07)
- Samuel, C. E. 2001 Antiviral actions of interferons. *Clin. Microbiol. Rev.* **14**, 778–809. (doi:10.1128/CMR.14.4.778-809.2001)
- Santiago, M. L., Montano, M., Benitez, R., Messer, R. J., Yonemoto, W., Chesebro, B., Hasenkrug, K. J. & Greene, W. C. 2008 APOBEC3 encodes Rfv3, a gene influencing neutralizing antibody control of retrovirus infection. *Science* **321**, 1343–1346. (doi:10.1126/science.1161121)
- Scadden, A. D. 2005 The RISC subunit Tudor-SN binds to hyper-edited double-stranded RNA and promotes its cleavage. *Nat. Struct. Mol. Biol.* **12**, 489–496. (doi:10.1038/nsmb936)
- Schmid, A., Spielhofer, P., Cattaneo, R., Baczko, K., ter Meulen, V. & Billeter, M. A. 1992 Subacute sclerosing panencephalitis is typically characterized by alterations in the fusion protein cytoplasmic domain of the persisting measles virus. *Virology* **188**, 910–915. (doi:10.1016/0042-6822(92)90552-Z)
- Schumacher, A. J., Nissley, D. V. & Harris, R. S. 2005 APOBEC3G hypermutates genomic DNA and inhibits Ty1 retrotransposition in yeast. *Proc. Natl Acad. Sci. USA* **102**, 9854–9859. (doi:10.1073/pnas.0501694102)
- Schumacher, A. J., Hache, G., Macduff, D. A., Brown, W. L. & Harris, R. S. 2008 The DNA deaminase activity of human APOBEC3G is required for Ty1, MusD, and human immunodeficiency virus type 1 restriction. *J. Virol.* **82**, 2652–2660. (doi:10.1128/JVI.02391-07)
- Sheehy, A. M., Gaddis, N. C., Choi, J. D. & Malim, M. H. 2002 Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* **418**, 646–650. (doi:10.1038/nature00939)
- Shen, C. H. & Steiner, L. A. 2004 Genome structure and thymic expression of an endogenous retrovirus in zebrafish. *J. Virol.* **78**, 899–911. (doi:10.1128/JVI.78.2.899-911.2004)
- Suspène, R. et al. 2004 APOBEC3G is a single-stranded DNA cytidine deaminase and functions independently of HIV reverse transcriptase. *Nucleic Acids Res.* **32**, 2421–2429. (doi:10.1093/nar/gkh554)
- Suspène, R., Guétard, D., Henry, M., Sommer, P., Wain-Hobson, S. & Vartanian, J. P. 2005a Extensive editing of both hepatitis B virus DNA strands by APOBEC3 cytidine deaminases *in vitro* and *in vivo*. *Proc. Natl Acad. Sci. USA* **102**, 8321–8326. (doi:10.1073/pnas.0408223102)
- Suspène, R., Henry, M., Guillot, S., Wain-Hobson, S. & Vartanian, J. P. 2005b Recovery of APOBEC3-edited human immunodeficiency virus G→A hypermutants by differential DNA denaturation PCR. *J. Gen. Virol.* **86**, 125–129. (doi:10.1099/vir.0.80426-0)
- Suspène, R., Rusniok, C., Vartanian, J. P. & Wain-Hobson, S. 2006 Twin gradients in APOBEC3 edited HIV-1 DNA reflect the dynamics of lentiviral replication. *Nucleic Acids Res.* **34**, 4677–4684. (doi:10.1093/nar/gkl555)
- Suspène, R. et al. 2008 Inverting the natural hydrogen bonding rule to selectively amplify GC-rich ADAR-edited RNAs. *Nucleic Acids Res.* **36**, e72. (doi:10.1093/nar/gkn295)
- Takeda, E., Tsuji-Kawahara, S., Sakamoto, M., Langlois, M. A., Neuberger, M. S., Rada, C. & Miyazawa, M. 2008 Mouse APOBEC3 restricts Friend leukemia virus infection and pathogenesis *in vivo*. *J. Virol.* **82**, 10 998. (doi:10.1128/JVI.01311-08)
- Tanaka, Y. et al. 2006 Anti-viral protein APOBEC3G is induced by interferon-alpha stimulation in human hepatocytes. *Biochem. Biophys. Res. Commun.* **341**, 314–319. (doi:10.1016/j.bbrc.2005.12.192)
- Trivedi, N. R., Gilliland, K. L., Zhao, W., Liu, W. & Thiboutot, D. M. 2006 Gene array expression profiling in acne lesions reveals marked upregulation of genes involved in inflammation and matrix remodeling. *J. Invest. Dermatol.* **126**, 1071–1079. (doi:10.1038/sj.jid.5700213)
- Valente, L. & Nishikura, K. 2005 ADAR gene family and A-to-I RNA editing: diverse roles in posttranscriptional gene regulation. *Prog. Nucleic Acid Res. Mol. Biol.* **79**, 299–338.
- Vartanian, J. P., Plikat, U., Henry, M., Mahieux, R., Guillemot, L., Meyerhans, A. & Wain-Hobson, S. 1997 HIV genetic variation is directed and restricted by DNA precursor availability. *J. Mol. Biol.* **270**, 139–151. (doi:10.1006/jmbi.1997.1104)
- Vartanian, J. P., Henry, M. & Wain-Hobson, S. 2002 Sustained G→A hypermutation during reverse transcription of an entire human immunodeficiency virus type 1 strain Vau group O genome. *J. Gen. Virol.* **83**, 801–805.
- Vartanian, J. P., 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. (doi:10.1126/science.1153201)
- Wain-Hobson, S. 2006 The third bond. *Nature* **439**, 539. (doi:10.1038/439539a)
- Wiegand, H. L., Doehle, B. P., Bogerd, H. P. & Cullen, B. R. 2004 A second human antiretroviral factor, APOBEC3F, is suppressed by the HIV-1 and HIV-2 Vif proteins. *EMBO J.* **23**, 2451–2458. (doi:10.1038/sj.emboj.7600246)
- Yamanaka, S., Balestra, M. E., Ferrell, L. D., Fan, J., Arnold, K. S., Taylor, S., Taylor, J. M. & Innerarity, T. L. 1995 Apolipoprotein B mRNA-editing protein induces hepatocellular carcinoma and dysplasia in transgenic animals. *Proc. Natl Acad. Sci. USA* **92**, 8483–8487. (doi:10.1073/pnas.92.18.8483)
- Yu, Q., Konig, R., Pillai, S., Chiles, K., Kearney, M., Palmer, S., Richman, D., Coffin, J. M. & Landau, N. R. 2004 Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. *Nat. Struct. Mol. Biol.* **11**, 435–442. (doi:10.1038/nsmb758)
- Zhang, H., Yang, B., Pomerantz, R. J., Zhang, C., Arunachalam, S. C. & Gao, L. 2003 The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* **424**, 94–98. (doi:10.1038/nature01707)
- Zheng, Y. H., Irwin, D., Kurosu, T., Tokunaga, K., Sata, T. & Peterlin, B. M. 2004 Human APOBEC3F is another host factor that blocks human immunodeficiency virus type 1 replication. *J. Virol.* **78**, 6073–6076. (doi:10.1128/JVI.78.11.6073-6076.2004)