Chapter 3

Genome Instability in DNA Viruses

R. Sanjuán, M. Pereira-Gómez, J. Risso
Universitat de València, Paterna, Spain

Chapter Outline

1. OVERVIEW

DNA viruses comprise important pathogens such as herpesviruses, smallpox viruses, adenoviruses, and papillomaviruses, among many others. DNA viruses are divided into three major categories: double-stranded DNA viruses (eg, poxviruses), single-stranded DNA viruses (eg, parvoviruses), and pararetroviruses (eg, hepadnaviruses) which replicate their genome through an RNA intermediate. Large DNA viruses (>10 kb) have double-stranded DNA, whereas small DNA viruses have circular single- or double-stranded DNA. These broad viral groups differ in their rates of spontaneous mutation, defined as the probability that an unrepaired genetic change is passed on to the viral progeny in each cell infection cycle [1,2]. For instance, single-stranded DNA microviruses such as bacteriophage ϕX174 and innoviruses produce about 10^{-6} spontaneous mutations per nucleotide per cell infection cycle (m/n/c), a mutation rate which is close to those of some RNA viruses. In contrast, the double-stranded DNA herpes simplex virus (HSV) and bacteriophage T4, both of which have genome sizes exceeding 150 kb, show clearly lower mutation rates (10^{-8}–10^{-7} m/n/c). As a result, there is an inverse correlation between genome size and per-base mutation rate in DNA viruses, while the per-genome mutation rate stays approximately constant. This correlation extends to unicellular organisms and is known as Drake's rule [3,4] (Fig. 3.1). RNA viruses also exhibit an inverse relationship between genome size and mutation rate, albeit with a different slope [5]. The main feature that distinguishes RNA viruses from DNA viruses in terms of genome stability is probably the absence of 3′-exonuclease proofreading activity from most RNA virus-encoded polymerases, which makes them particularly error prone [6]. The 3′-exonuclease activity leads to roughly 10-fold to 100-fold increase in replication fidelity [7,8]. In turn, differences in mutation rate among DNA viruses should be determined by their ability to access postreplicative repair. For instance, bacteriophage ϕX174 lacks sequence motifs required for methyl-directed mismatch repair (MMR) in *Escherichia coli*, therefore excluding the phage DNA from this major repair system [9]. In contrast, the interplay between viral replication and host postreplicative repair pathways is far more complex and less well understood in eukaryotic DNA viruses. Molecular evolution studies indicate that the classical dichotomy between fast-evolving RNA and slow-evolving DNA viruses becomes blurred when full-genome datasets are considered [10], suggesting that DNA viruses probably have other mechanisms for promoting genetic diversity. Some of these mechanisms have already been characterized and include gene amplification [11] and diversity-generating retro-elements (DGRs) [12], both of which act on specific genome regions. Additionally, retroviruses

Genome Stability. http://dx.doi.org/10.1016/B978-0-12-803309-8.00003-3

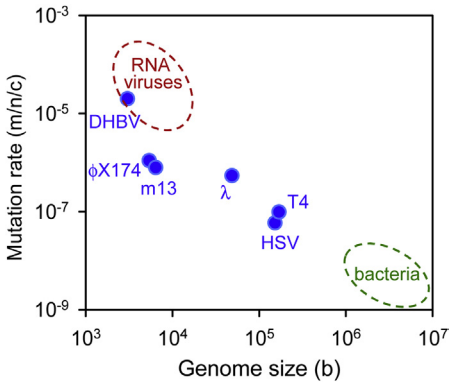

FIGURE 3.1 The relationship between genome size and the rate of spontaneous mutation in DNA viruses. *Dots* correspond to bacteriophages ϕX174, m13, λ, and T4, duck hepatitis B virus (DHBV), and herpes simplex virus (HSV). DHBV is a pararetrovirus, ϕX174 and m13 are single-stranded DNA viruses, and λ T4 and HSV are double-stranded DNA viruses. The approximate location of RNA viruses and bacteria is shown. See text for references from which these estimates are taken.

and several DNA viruses are subject to host-encoded DNA editing by enzymes of the APOBEC3 family, which can produce hypermutated viral genomes. Like all other biological systems, DNA viruses have to keep a balance between the avoidance of deleterious mutations and the production of diversity, and genome instability mechanisms probably play a central role in the maintenance of this balance.

2. RATES OF SPONTANEOUS MUTATION AND GENETIC DIVERSITY OF DNA VIRUSES

Although DNA viruses were traditionally believed to evolve slowly, analysis of sequences from field isolates with known sampling dates spanning years to decades suggested fast molecular evolution rates for several single-stranded DNA viruses, including emerging canine parvovirus strains [13], human parvovirus B19 [14], tomato yellow leaf curl geminivirus [15], and beak-and-feather disease circovirus [16]. During 2010s, it was further suggested that large double-stranded DNA viruses can also evolve fast. For instance, analysis of samples of the African swine fever virus (ASFV) spanning 70 years yielded estimated evolution rates in the order of 10^{-4} substitutions per nucleotide per year (s/n/y) [17], a value that falls within the typical range exhibited by many RNA viruses. In HSV, frameshift mutations, insertion/deletions, and large complex rearrangements are major sources of genetic diversity as well [18]. Next-generation sequencing (NGS) of HSV laboratory samples has detected the appearance of new mutations after few transfers, further questioning the long-believed genetic stability of DNA viruses [19]. It has also been found that serial plaque-to-plaque transfers can rapidly reduce the fitness of this virus by promoting the accumulation of deleterious mutations, thus echoing the results obtained with RNA viruses in 1990s [20,21]. Plaque-to-plaque transfers reduce viral effective population sizes dramatically, thus favoring the action of random genetic drift, but to set these processes into motion, spontaneous mutations need to occur at a relatively high rate. However, our current knowledge of DNA virus mutation rates is far more limited than for RNA viruses. These rates have been measured directly for a handful of DNA viruses, including HSV and bacteriophages ϕX174, m13, λ, and T4 [2,3,22–25]. A better understanding of DNA virus mutation rates is thus needed to evaluate baseline levels of genome instability. Probably, the main reason for this scarcity of data is of a technical nature. While, in many RNA viruses, mutation rates have been estimated based on sequence analysis, this has not been possible so far for DNA viruses in which mutations are less frequent. Classical Sanger sequencing does not provide deep-enough information to sample low-frequency, new spontaneous mutations. In turn, NGS platforms have the capacity to yield hundreds of billions of nucleotides of DNA sequences in a single experiment, but they are limited by their high per-read error rates which can be orders of magnitude higher than the mutation rate to be measured [26]. However, high-fidelity NGS techniques, developed during early 2010s, such as duplex sequencing [27,28] or circular sequencing [29,30] offer a promising solution for these limitations and should enable a much deeper understanding of spontaneous mutation rates and genome instability in DNA viruses (Fig. 3.2).

3. MUTATOR PHENOTYPES PRODUCED BY LOW-FIDELITY DNA VIRUS POLYMERASES

The most extreme form of genomic instability is achieved by mutators in which genome-wide rates of spontaneous mutation are elevated by orders of magnitude. Mutator strains and their evolutionary and clinical implications have been extensively

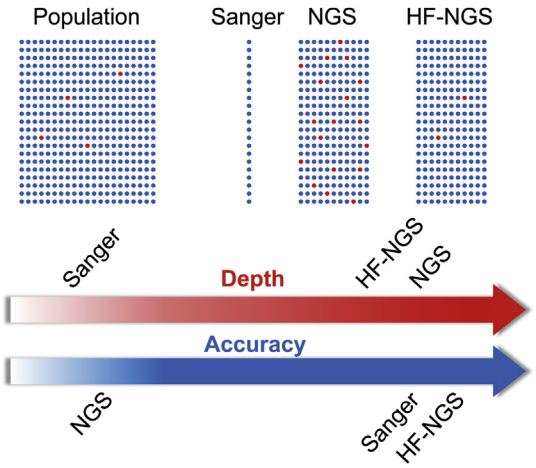

FIGURE 3.2 **Benefits of high-fidelity NGS.** Sequencing has been used for estimating the genetic diversity of viral populations and for characterizing hypermutation and other genome instability processes. Classical Sanger sequencing of PCR molecular clones is a reliable approach, but its coverage is typically limited to 10–100 reads per site, thus preventing sampling of low-frequency mutations. Using NGS platforms, sequencing coverage can be easily increased to >1000 reads per site, but the per-read per-base technical error rate is relatively high (0.1–1.0%). In contrast, the recently developed high-fidelity (HF) NGS technologies, such as duplex sequencing or circular sequencing, can achieve high coverage with an extremely low error rate (<0.001%). See text for references.

studied in bacteria [31–34] and constitute a highly active topic in cancer research [35,36]. DNA viruses can also adopt a mutator phenotype, as shown decades ago using the model bacteriophage T4. Although T4 has a low per-base spontaneous mutation rate compared to other DNA viruses, this rate is still about 30-fold higher than that of its host and is mainly determined by the fidelity of the viral polymerase [37,38]. Most T4 mutators are produced by replacements in the N-terminal domain of the polymerase where the 3′-exonuclease activity resides and can reduce replication fidelity by up to 400-fold [39]. This contrasts with RNA viruses including HIV-1, hepatitis C virus, and influenza virus in which natural mutators have not been described, probably because the wild-type mutation rate is already in the order of 0.1–1.0 new mutations per genome copying, a value that is presumably very close to the theoretical maximum level compatible with virus survival, also termed error threshold [40–42]. Changes in other genes involved in replication, including single-stranded DNA-binding proteins and helicase and clamp proteins, can also produce a mutator phenotype in T4, albeit typically more modestly than low-fidelity polymerases [43]. T4 antimutators showing 100-fold increase in replication fidelity have also been described and often map to the central exonuclease and palm subdomains and the carboxyl-terminal thumb subdomain of the viral polymerase [39]. It has also been noted that T4 antimutator polymerases tend to replicate DNA more slowly than wild-type polymerases, therefore negatively impacting viral fitness. This cost suggests that there is an upper limit for replication fidelity which is determined by the need to replicate fast. On the other hand, mutator phenotypes are also costly because mutations falling at essential genes inflate the genetic load of the population [44]. However, in theory, mutators may still be favored in populations that are maladapted or subject to rapidly changing environments because they boost the production of genetic diversity. However, their rise should be transient, particularly in recombining populations where the mutator locus rapidly unlinks from loci where positively selected mutations are found [45,46]. While in bacteria, these predictions have been largely confirmed, less is known about the evolutionary dynamics of mutators in DNA viruses. In addition to the well-studied T4 system, low-fidelity polymerases may play a central role in the production of diversity in other large DNA viruses such as, for instance, ASFV. Besides the replicative DNA polymerase, ASFV encodes a simple DNA repair system consisting of an endonuclease, a repair polymerase termed pol X and an ATP-dependent DNA ligase. Pol X, which belongs to the same family as the mammalian base-excision repair pol β, exhibits a high error rate, which is determined by the lack of 3′-exonuclease activity and a poor base discrimination capacity [47]. It has been suggested that the relatively high diversity found among ASFV isolates may in part have originated during mutagenic repair involving the highly error-prone DNA pol X [48].

4. DNA COLIPHAGES AND THE MMR SYSTEM

Compared with proofreading, relatively little attention has been paid to the role played by postreplicative repair in determining the genomic stability of DNA viruses. Inasmuch as the lack of 3′-exonuclease proofreading is believed to be a major determinant of RNA virus error-prone replication, access to postreplicative repair may dictate to a large extent the rate of spontaneous mutation of DNA viruses. An excellent model for addressing this question is the *E. coli* MMR

system which affords up to 1000-fold reduction in the rate of spontaneous mutations [49]. MMR is carried out by the MutHLS proteins which perform a strand-specific bidirectional repair [50]. Base mismatches or small insertion/deletion loops are recognized by MutS which interacts with MutL and leads to the activation of the MutH endonuclease which excises the daughter strand. DNA resynthesis from the parental strand is then carried out by DNA pol III followed by ligation of the nicked DNA. For this process to operate, the daughter and parental strands need to be distinguished. This is made possible because the parental strand has a methyl group in the adenosine of GATC sequence motifs which is added by *Dam* methylase. Therefore, MMR requires the presence of GATC motifs in the genome, which are normally found at a high frequency in the bacterial chromosome (about 1 in every 250 bases are expected by chance). However, GATC motifs are strikingly absent from the 5.4 kb genome of bacteriophage ϕX174. This strong GATC avoidance necessarily impairs MMR in the phage and should produce a major effect on the ϕX174 mutation rate. Supporting this, the mutation rate of bacteriophage ϕX174 (c. 10^{-6} m/n/c) is three orders of magnitude higher than that of *E. coli* [22]. However, the introduction of 20 GATC motifs in the ϕX174 genome using site-directed mutagenesis reduced the phage mutation rate only by eight fold, with varying effects of these motifs depending on their genome location, the lower-than-expected effect of GATC motifs on phage mutation rate being probably due to an inefficient methylation of the phage DNA [9]. Fast replication or the transient nature of double-stranded replicative intermediates may offer fewer chances for *Dam* methylation in the phage DNA compared to the bacterial chromosome. Although less marked, GATC depletion extends to other coliphages and plasmids, but it is still unclear whether this is a consequence of selection acting on either mutation rates or unrelated traits [51].

5. THE INTERACTION BETWEEN DNA VIRUSES AND THE EUKARYOTIC DNA DAMAGE RESPONSE

The DNA damage response (DDR) comprises a set of signaling pathways for the detection and repair of DNA damage and includes the MMR system for mispaired bases, the base excision repair system for small base modifications, the nucleotide excision repair for intrastrand crosslinks and pyrimidine dimers, the single-strand break repair and double-strand break (DSB) repair pathways involving homologous recombination and nonhomologous end joining [52]. The DDR is primarily controlled by two protein kinases, ataxia-telangiectasia mutated (ATM) and ATM/Rad3-related (ATR) protein kinases. ATM is mainly implicated in the repair of DSBs sensed by the protein complex MRN [53], whereas ATR responds to various types of DNA damage that have in common the presence of single-stranded DNA [54]. Numerous studies have demonstrated that viruses interact with DDR pathways and that, whereas many viruses evade DDR, others appear to benefit from it [55,56]. DNA viruses have developed different strategies to modulate DDR by altering the localization or promoting the degradation of DDR components. For instance, the adenovirus E4orf6 protein recruits a ubiquitin ligase and promotes the proteasomal degradation of TOPBP1, an activator of ATR [57]. Defects in the adenoviral E4 gene lead to the formation of genome concatemers constituted by ligated viral DNA with heterogeneous junctions [58], underscoring the importance of DDR evasion for adenoviruses (Fig. 3.3).

Similarly, HSV proteins such as the regulatory factor ICP0, antagonize DDR by promoting the mislocalization of ATR-interacting protein (ATRIP) [59]. As a result, mutants with ICP0 defects show very poor growth. DDR induction can produce undesirable effects for the virus, such as premature entry into apoptosis. Hence, the inhibition of downstream DDR pathways that stimulate apoptosis is also a common feature of DNA viruses. For instance, the HSV latency-associated protein M2 induces ATM activation, which results in p53 phosphorylation, the inhibition of DNA repair, the blockade of DNA damage-induced apoptosis, and the induction of G1 cell-cycle arrest [60]. Although DNA viruses tend to produce genomic instability in the infected cell, it is still poorly understood how DDR impairment affects DNA virus genomic stability and mutation rates. Highlighting the complexity of virus–DDR interactions, some DNA viruses also use DDR for their own benefit. For example, polyomaviruses induce and exploit the ATM signaling pathway [61]. The T antigen protein expressed by the SV40 polyomavirus activates ATM kinase and downstream targets that are required to obtain unit length viral replication products [62]. Other small DNA viruses such as papillomaviruses [63] and parvoviruses [64] also need to activate the DDR pathways for an efficient replication. These viruses share the common property of having small circular DNA genomes which do not encode their own polymerases and, therefore, they depend strictly on cellular polymerases for replication, as opposed to larger DNA viruses such as adenoviruses, herpesviruses, and poxviruses which encode autonomous replication complexes. Therefore, a possible explanation for why some small viruses promote DDR is that they need to prolong the S cell-cycle phase to create a more favorable environment for replication. By adopting circular genomes, these viruses would also avoid the formation of DDR-associated concatemers as those found in adenoviruses. Given the effects of repair avoidance in the mutation rates of prokaryotic viruses, changes in the expression, and localization of DDR repair-associated proteins might also have major effects on the genomic stability of eukaryotic DNA viruses.

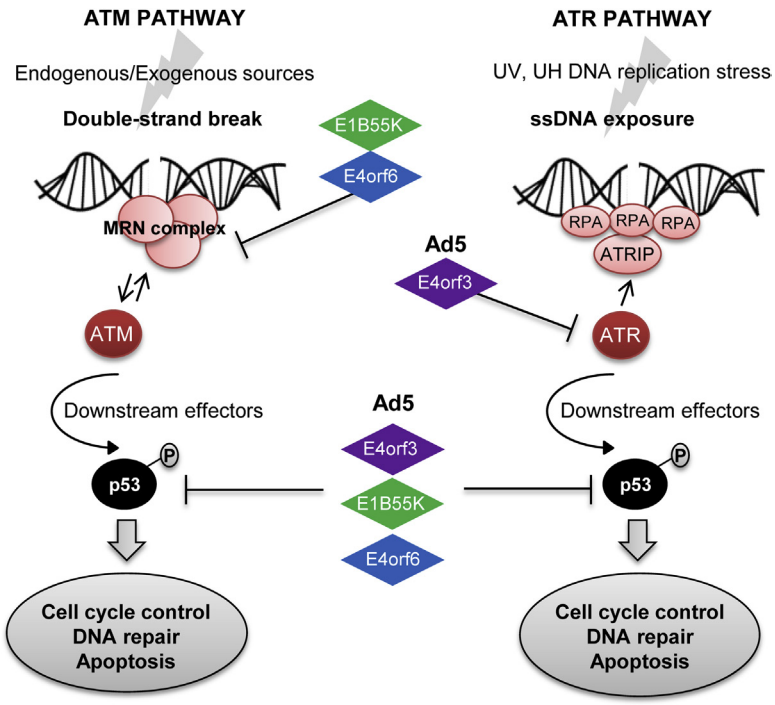

FIGURE 3.3 **Inhibition of ATM and ATR pathways by adenoviruses.** E1b55K–E4orf6 complexes promote the degradation of the MRN complex by recruiting cellular ubiquitin ligases, which prevents ATM signaling. These complexes also recruit ubiquitin ligases to promote the degradation of p53, avoiding apoptosis. The Ad5 E4orf3 protein abolishes the MRN-dependent activation of ATR, resulting in the inhibition of the ATR pathway, and it also inhibits p53. The Ad12 E1B55K protein recruits an E3 ubiquitin ligase to promote TOPB1 degradation, leading to ATR pathway suppression.

6. DIVERSITY-GENERATING RETRO-ELEMENTS IN BACTERIOPHAGES

Some DNA viruses have evolved the ability to target mutations to specific genome regions, thus avoiding the cost of genome-wide hypermutation. A unique and fascinating mechanism of mutation targeting is provided by DGRs. These elements are located in genome regions involved in host attachment and tropism, a trait which is frequently subject to rapidly changing selective pressures dictated by host availability. The first and best-studied DGR was found in the *Bordetella* BPP-1 bacteriophage [65]. This DGR consists of two sequence repeats of about 150 bp each and two ORFs (Fig. 3.4). The first repeat is called the variable repeat (VR) and is located in the 3′-end of the *mtd* gene (major tropism determinant) which encodes a tail fiber protein. Downstream of the VR is located the template repeat (TR) which, contrarily to VR, has a highly conserved sequence. A second ORF (*brt*) encodes a reverse transcriptase which synthesizes cDNA from the VR transcript. During this process, extensive mutagenesis occurs whereby adenines are systematically substituted for random bases by an as yet unknown mechanism. VR cDNA is then transferred to TR, thus producing a large number of variants of the *mtd* gene potentially capable of interacting with new ligands [12]. For this transfer to occur, several *cis*-acting elements are required, including an IMH (initiation of mutagenic homing) region which contains a 15-bp GC-only sequence identical to a portion of TR and a 21-bp sequence similar but not identical to another TR fragment followed by inverted repeats that can adopt a cruciform secondary structure [66]. Using a metagenomics approach, DNA viruses present in the human lower gastrointestinal tract were found to harbor hot spots of hypervariation in genes showing homology to BPP-1 DGR, along with other loci encoding the Ig-superfamily proteins, most of which were linked to genes encoding reverse transcriptases [67]. DGRs have also been found in plasmids, bacterial chromosomes, archaea, and archaeal viruses [68–70]. Although their absolute abundance is low, their powerful mutagenic effect may have a significant impact on the adaptability of prokaryotic viruses.

7. RECOMBINATION-DRIVEN GENOME INSTABILITY IN DNA VIRUSES

DGRs have not been described in eukaryotes or their viruses. The latter may thus use different mechanisms of targeted hypermutation. One such possible mechanism has been demonstrated in poxviruses and is based on recombination-mediated gene amplification. For instance, the inverted terminal repeats of the vaccinia virus genome are known to experience rapid changes in size [71]. This region contains abundant repeats of 10–100 bp sequence motifs that undergo frequent unequal crossover events [72]. While other regions of poxvirus genomes are believed to exhibit greater genome stability, diversity

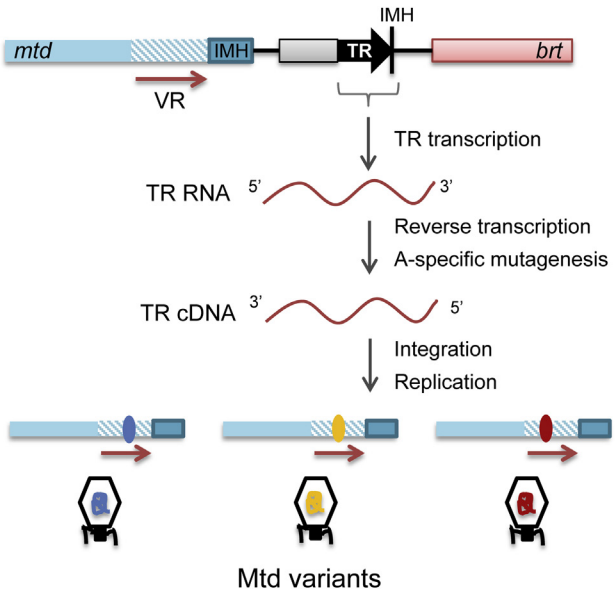

FIGURE 3.4 **Organization and function of DGRs.** The prototypic DGR of phage BPP-1 is linked to the major tropism determinant (*mtd*) gene which encodes a tail fiber protein. The TR transcript is converted to cDNA by the *brt*-encoded reverse transcriptase, and in this process, extensive mutagenesis of adenosines takes place. The cDNA then displaces the VR of the *mtd* gene, a process that is dependent on the IMH motif. As a result, a large number of Mtd variants are produced, allowing for rapid changes in host tropism at the viral population level.

is nevertheless required in these regions for immune escape and for the colonization of novel hosts. This requirement originates from the species-specific selective pressure exerted by host immunity. A central component of innate immunity is protein kinase R (PKR), which induces translational shutoff by phosphorylating the eukaryotic translation initiation factor 2 and mediates additional antiviral responses through its effects on protein phosphorylation status, mRNA stability, and apoptosis [73]. The host–pathogen evolutionary arms race has led to the coevolution of PKR and poxvirus proteins such as K3L and E3L which counteract PKR and contribute to virus-host specificity [74,75]. To investigate the plasticity of these genes, experimental evolution of a vaccinia virus deleted for E3L was carried out to impose a strong selection pressure favoring gain-of-function mutations in the other PKR suppressor, K3L [11]. The virus became adapted to this deletion by increasing the copy number of the K3L gene, inflating its total genome size by up to 10%. Low-frequency variants in the viral population carrying recombination breakpoints were identified as the most likely founders of these genomic expansions. The beneficial effect of gene amplification was two pronged. First, it increased K3L levels, thus providing a direct fitness advantage. Second, it also increased the number of targets available for the appearance of spontaneous gain-of-function mutations. Once these mutations were positively selected and became fixed in the viral population, K3L copy numbers were again reduced, a process which was probably driven by the cost of increased genome size. This thus led to accordion-like evolutionary dynamics whereby copy numbers expand and contract through time. Genomic accordions may also be relevant to the evolution of other poxviruses, such as for instance adaptive gene duplications found in myxomavirus [76]. More broadly, recombination plays a central role in DNA virus biology, including replication, the production of genetic diversity, and the preservation of genome integrity, and it has been associated with host range expansion, the emergence of new viruses, modifications of transmission vector specificity, pathogenesis, and host immunity evasion [77–79]. Early work suggested a nonhomologous recombination hot spot in the replication origin of phage m13 [80]. In phage λ, recombination can occur independently of DNA replication and is active even in cells deficient for the RecA protein (the main protein involved in *E. coli* recombination), which allowed for the identification of a phage-encoded homologous recombination system termed Red [81]. Herpesviruses also have their own recombination machinery used both for replication and DNA repair [81]. Sources of genome instability including DSBs and single-strand DNA breaks are sensed by DNA virus proteins and repaired using different recombination pathways depending of the type of DNA damage. Since these repair pathways are generally error prone, recombination hot spots may drive targeted genomic instability.

8. APOBEC3 PROTEINS AND DNA VIRUS GENOME INSTABILITY

The induction of viral genome instability by host-encoded factors is best illustrated by the action of the apolipoprotein-B mRNA-editing catalytic polypeptide-like 3 (APOBEC3) family of cytidine deaminases which constitutes an innate cellular

defense mechanism against retroviruses, endogenous retro-elements, and some DNA viruses [82]. APOBEC3 proteins produce mutations on viral DNA by deaminating cytidine to uracil [83,84]. The first studies showing the antiviral effect of APOBEC3 proteins were carried out in HIV-1 more than 10 years ago [85,86]. However, several subsequent studies have shown that APOBEC3 members can also edit the genomes of hepatitis B virus (HBV) and other DNA viruses that do not undergo a reverse transcription step. HBV has a partially double-stranded DNA genome of 3.2 kb which contains four highly overlapped open reading frames. The HBV genomic DNA is synthesized by the reverse transcription of a pregenomic RNA inside the nucleocapsid, and since reverse transcriptases are highly error prone, this is believed to be the main source of diversity in HBV. APOBEC3 proteins were first shown to inhibit HBV in a mutagenesis-independent manner when the amount of viral RNA associated with core particles was found to be reduced in the presence of APOBEC3 due to the inhibition of the pregenomic RNA encapsidation [87]. However, transfection experiments have shown that different APOBEC3 forms edit both plus and minus DNA strands [88] and have been found to produce hypermutated viral genomes in vivo [89], particularly in cirrhotic patients [90]. As opposed to HIV-1, HBV hypermutated genomes cannot usually be detected by molecular clone sequencing of conventional PCR products and require an ad hoc modified PCR protocol in which a lower denaturation temperature is used to favor the selective amplification of APOBEC3-edited A/T-rich sequences [88]. Interestingly, APOBEC3 footprints have also been detected in non-reverse transcribing DNA viruses. Human papillomavirus, a circular double-stranded DNA virus, has been found to be subject to APOBEC3 editing of both DNA strands in cotransfection experiments and in vivo, producing hypermutated viruses in benign and precancerous lesions [91]. Single-stranded DNA parvoviruses have also been found to be inhibited by APOBEC3, although in this case, this was not accompanied by hypermutation [92]. In contrast, transfusion-transmitted virus, another single-stranded DNA virus with no known homology to previously described viral families, has been found to be susceptible to hypermutation caused by APOBEC3 proteins [93]. Finally, APOBEC3-mediated editing has also been described in large double-stranded DNA viruses such as HSV and Epstein–Barr virus [94]. The primary effect of hypermutation is antiviral because a large number of deleterious missense or nonsense mutations are produced. For instance, one of the preferred APOBEC3G targets is the TGG trinucleotide which, after editing, can lead to TAG premature stop codons, most of which are lethal to the virus. In HIV-1, APOBEC3 expression levels can determine disease progression, with higher APOBEC3 activity associated with higher CD4 counts and slower progression [95,96]. However, a fraction of the edited genomes might be viable and could contribute to immune escape or drug resistance. In HIV-1, it has been shown that many drug-resistance mutations [97] and CTL-escape mutants [98] are located within typical APOBEC3G targets. Similarly, the 3TC-resistance M184I replacement in the HIV-1 reverse transcriptase arose faster in APOBEC3-expressing cells [99]. However, the role played by APOBEC3 proteins in the genetic diversity and virulence of DNA viruses still remains poorly characterized.

9. CONCLUSIONS AND FUTURE DIRECTIONS

The long-accepted genetic stability and slow evolution of DNA viruses have been challenged by multiple reports showing that DNA viruses can exhibit levels of genetic diversity approaching those of some RNA viruses. While for some small DNA viruses, this could be explained by a relatively high rate of spontaneous mutation, the few available estimates for large DNA viruses support a lower average mutation rate. High-fidelity NGS techniques should provide a powerful tool for the study of DNA virus mutation rates and genomic instability in the near future. Despite current uncertainties, DNA virus mutation rates appear to be higher than those of their hosts, probably because the former undergo less efficient DNA repair. The depletion of GATC motifs found in some coliphages provides evidence for repair avoidance, but further work is needed to clarify the evolutionary forces driving such avoidance. The relationship between DNA viruses and cellular repair pathways is much more complex in eukaryotes. It is well established that DNA virus infections modify DDR pathways, but the cause–effect relationships of these changes remain poorly understood. From the virus perspective, it appears that DDR is sometimes an undesired yet unavoidable cellular response to infection, whereas in other cases, DDR is a beneficial or even necessary cellular resource for the virus. Furthermore, the effects of DDR activation/inhibition on viral genomic stability as well as the implications for DNA virus genetic diversity remain obscure. Similar dualities apply to APOBEC3-mediated hypermutation of DNA virus genomes which, despite being primarily an antiviral response, can promote the appearance of immune escape or drug-resistance mutations. Another important realization is that although large DNA viruses show a higher average genomic stability than small DNA viruses and RNA viruses, mutational hot spots can be found at specific genome regions involved in dynamic virus–host interactions, and transient boosts of diversity may also be afforded by short-lived genome-wide mutators in DNA viruses. While the selective pressures acting on bacterial mutators have been well studied, much less is known about the fate of DNA virus mutators, particularly for eukaryotic viruses. DGRs provide a clear mechanistic basis for the ability of some DNA bacteriophages to target mutations to specific genome regions, and their in-depth characterization has both basic and practical implications for directed evolution purposes. Different mechanisms appear to be used for targeted hypermutation in large eukaryotic DNA viruses, in which recombination-driven genomic instability appears to play a central role.

GLOSSARY

Error threshold The theoretical maximal mutation rate tolerated by a given population. Above this threshold, natural selection fails to preserve the sequence of the fittest variants and other, less fit variants reach high population frequencies and may become fixed. This is expected to favor population extinction, although extinction is not a necessary consequence of error threshold crossing.

Genomic accordion An evolutionary expansion/contraction of a gene's copy number, typically associated with strong selection acting on this specific gene.

Hypermutation Strong elevation of the rate of spontaneous mutation which, in viruses, is typically associated with host-mediated edition of the viral genome and tends to be specific to some bases or sequence contexts.

Mutational hot spot Elevation of the spontaneous mutation rate at a specific genome region.

Mutator phenotype A highly increased rate of spontaneous mutation affecting the entire genome and caused by loss of fidelity mechanisms, including proofreading activity and/or postreplicative repair.

Plaque-to-plaque transfer A virus culture technique whereby a single viral plaque is picked and used to seed a fresh culture, in which new plaques develop, and so on. By passaging a virus in this manner, the effective population size is strongly reduced, thereby allowing for the accumulation of mutations under random genetic drift.

Rate of spontaneous mutation The probability that new genetic changes appear and are passed to the next generation. In viruses, a generation is typically defined as one cell infection cycle.

LIST OF ACRONYMS AND ABBREVIATIONS

APOBEC3 Apolipoprotein-B mRNA-editing catalytic polypeptide-like 3 (protein)
ASFV African swine fever virus
ATM Ataxia–telangiectasia mutated (protein)
ATR ATM/Rad3-related (protein)
DDR DNA damage response
DGR Diversity-generating retro-element
DSB Double-strand break
HBV Hepatitis B virus
HSV Herpes simplex virus
IMH Initiation of mutagenic homing (DGR element)
MMR Methyl-directed mismatch repair
NGS Next-generation sequencing
PKR Protein kinase R
TR Template repeat (DGR element)
VR Variable repeat

REFERENCES

[1] Duffy S, Shackelton LA, Holmes EC. Rates of evolutionary change in viruses: patterns and determinants. Nat Rev Genet 2008;9:267–76.
[2] Sanjuán R, Nebot MR, Chirico N, Mansky LM, Belshaw R. Viral mutation rates. J Virol 2010;84:9733–48.
[3] Drake JW. A constant rate of spontaneous mutation in DNA-based microbes. Proc Natl Acad Sci USA 1991;88:7160–4.
[4] Drake JW, Charlesworth B, Charlesworth D, Crow JF. Rates of spontaneous mutation. Genetics 1998;148:1667–86.
[5] Bradwell K, Combe M, Domingo-Calap P, Sanjuán R. Correlation between mutation rate and genome size in riboviruses: mutation rate of bacteriophage Qβ. Genetics 2013;195:243–51.
[6] Steinhauer DA, Domingo E, Holland JJ. Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase. Gene 1992;122:281–8.
[7] Kunkel TA. DNA replication fidelity. J Biol Chem 2004;279:16895–8.
[8] Schaaper RM. Antimutator mutants in bacteriophage T4 and *Escherichia coli*. Genetics 1998;148:1579–85.
[9] Pereira-Gómez M, Sanjuán R. Effect of mismatch repair on the mutation rate of bacteriophage ϕX174. Virus Evol 2015;1:vev010.
[10] Biek R, Pybus OG, Lloyd-Smith JO, Didelot X. Measurably evolving pathogens in the genomic era. Trends Ecol Evol 2015;30:306–13.
[11] Elde NC, Child SJ, Eickbush MT, Kitzman JO, Rogers KS, Shendure J, et al. Poxviruses deploy genomic accordions to adapt rapidly against host antiviral defenses. Cell 2012;150:831–41.
[12] Medhekar B, Miller JF. Diversity-generating retroelements. Curr Opin Microbiol 2007;10:388–95.
[13] Shackelton LA, Parrish CR, Truyen U, Holmes EC. High rate of viral evolution associated with the emergence of carnivore parvovirus. Proc Natl Acad Sci USA 2005;102:379–84.
[14] Shackelton LA, Holmes EC. Phylogenetic evidence for the rapid evolution of human B19 erythrovirus. J Virol 2006;80:3666–9.
[15] Duffy S, Holmes EC. Phylogenetic evidence for rapid rates of molecular evolution in the single-stranded DNA begomovirus tomato yellow leaf curl virus. J Virol 2008;82:957–65.

[16] Sarker S, Patterson EI, Peters A, Baker GB, Forwood JK, Ghorashi SA, et al. Mutability dynamics of an emergent single stranded DNA virus in a naive host. PLoS One 2014;9:e85370.
[17] Michaud V, Randriamparany T, Albina E. Comprehensive phylogenetic reconstructions of African swine fever virus: proposal for a new classification and molecular dating of the virus. PLoS One 2013;8:e69662.
[18] Szpara ML, Gatherer D, Ochoa A, Greenbaum B, Dolan A, Bowden RJ, et al. Evolution and diversity in human herpes simplex virus genomes. J Virol 2014;88:1209–27.
[19] Szpara ML, Tafuri YR, Parsons L, Shamim SR, Verstrepen KJ, Legendre M, et al. A wide extent of inter-strain diversity in virulent and vaccine strains of alphaherpesviruses. PLoS Pathog 2011;7:e1002282.
[20] Chao L. Fitness of RNA virus decreased by Muller's ratchet. Nature 1990;348:454–5.
[21] Duarte E, Clarke D, Moya A, Domingo E, Holland J. Rapid fitness losses in mammalian RNA virus clones due to Muller's ratchet. Proc Natl Acad Sci USA 1992;89:6015–9.
[22] Cuevas JM, Duffy S, Sanjuán R. Point mutation rate of bacteriophage FX174. Genetics 2009;183:747–9.
[23] Drake JW, Hwang CB. On the mutation rate of herpes simplex virus type 1. Genetics 2005;170:969–70.
[24] Kunkel TA. The mutational specificity of DNA polymerase-beta during in vitro DNA synthesis. Production of frameshift, base substitution, and deletion mutations. J Biol Chem 1985;260:5787–96.
[25] Luria SE. The frequency distribution of spontaneous bacteriophage mutants as evidence for the exponential rate of phage reproduction. Cold Spring Harb Symp Quant Biol 1951;16:463–70.
[26] Quinones-Mateu ME, Avila S, Reyes-Teran G, Martinez MA. Deep sequencing: becoming a critical tool in clinical virology. J Clin Virol 2014;61:9–19.
[27] Kennedy SR, Schmitt MW, Fox EJ, Kohrn BF, Salk JJ, Ahn EH, et al. Detecting ultralow-frequency mutations by Duplex Sequencing. Nat Protoc 2014;9:2586–606.
[28] Schmitt MW, Kennedy SR, Salk JJ, Fox EJ, Hiatt JB, Loeb LA. Detection of ultra-rare mutations by next-generation sequencing. Proc Natl Acad Sci USA 2012;109:14508–13.
[29] Acevedo A, Brodsky L, Andino R. Mutational and fitness landscapes of an RNA virus revealed through population sequencing. Nature 2014;505:686–90.
[30] Acevedo A, Andino R. Library preparation for highly accurate population sequencing of RNA viruses. Nat Protoc 2014;9:1760–9.
[31] Gonzalez C, Hadany L, Ponder RG, Price M, Hastings PJ, Rosenberg SM. Mutability and importance of a hypermutable cell subpopulation that produces stress-induced mutants in *Escherichia coli*. PLoS Genet 2008;4:e1000208.
[32] Jolivet-Gougeon A, Kovacs B, Le Gall-David S, Le BH, Bousarghin L, Bonnaure-Mallet M, et al. Bacterial hypermutation: clinical implications. J Med Microbiol 2011;60:563–73.
[33] LeClerc JE, Li B, Payne WL, Cebula TA. High mutation frequencies among *Escherichia coli* and *Salmonella* pathogens. Science 1996;274:1208–11.
[34] Sniegowski PD, Gerrish PJ, Lenski RE. Evolution of high mutation rates in experimental populations of *E. coli*. Nature 1997;387:703–5.
[35] Negrini S, Gorgoulis VG, Halazonetis TD. Genomic instability–an evolving hallmark of cancer. Nat Rev Mol Cell Biol 2010;11:220–8.
[36] Roberts SA, Gordenin DA. Hypermutation in human cancer genomes: footprints and mechanisms. Nat Rev Cancer 2014;14:786–800.
[37] Drake JW, Allen EF, Forsberg SA, Preparata RM, Greening EO. Genetic control of mutation rates in bacteriophageT4. Nature 1969;221:1128–32.
[38] Santos ME, Drake JW. Rates of spontaneous mutation in bacteriophage T4 are independent of host fidelity determinants. Genetics 1994;138:553–64.
[39] Mansky LM, Cunningham KS. Virus mutators and antimutators: roles in evolution, pathogenesis and emergence. Trends Genet 2000;16:512–7.
[40] Bull JJ, Sanjuán R, Wilke CO. Theory of lethal mutagenesis for viruses. J Virol 2007;81:2930–9.
[41] Crotty S, Cameron CE, Andino R. RNA virus error catastrophe: direct molecular test by using ribavirin. Proc Natl Acad Sci USA 2001;98:6895–900.
[42] Andino R, Domingo E. Viral quasispecies. Virology 2015;479–480C:46–51.
[43] Mufti S. Mutator effects of alleles of phage T4 genes 32, 41, 44, and 45 in the presence of an antimutator polymerase. Virology 1979;94:1–9.
[44] Sniegowski PD, Gerrish PJ, Johnson T, Shaver A. The evolution of mutation rates: separating causes from consequences. BioEssays 2000;22:1057–66.
[45] Taddei F, Radman M, Maynard-Smith J, Toupance B, Gouyon PH, Godelle B. Role of mutator alleles in adaptive evolution. Nature 1997;387:700–2.
[46] Raynes Y, Sniegowski PD. Experimental evolution and the dynamics of genomic mutation rate modifiers. Heredity 2014;113:375–80.
[47] Oliveros M, Yanez RJ, Salas ML, Salas J, Vinuela E, Blanco L. Characterization of an African swine fever virus 20-kDa DNA polymerase involved in DNA repair. J Biol Chem 1997;272:30899–910.
[48] Lamarche BJ, Kumar S, Tsai MD. ASFV DNA polymerase X is extremely error-prone under diverse assay conditions and within multiple DNA sequence contexts. Biochemistry 2006;45:14826–33.
[49] Fijalkowska IJ, Schaaper RM, Jonczyk P. DNA replication fidelity in *Escherichia coli*: a multi-DNA polymerase affair. FEMS Microbiol Rev 2012;36:1105–21.
[50] Jiricny J. Postreplicative mismatch repair. Cold Spring Harb Perspect Biol 2013;5:a012633.
[51] McClelland M. Selection against dam methylation sites in the genomes of DNA of enterobacteriophages. J Mol Evol 1984;21:317–22.
[52] Ciccia A, Elledge SJ. The DNA damage response: making it safe to play with knives. Mol Cell 2010;40:179–204.
[53] Lee JH, Paull TT. ATM activation by DNA double-strand breaks through the Mre11-Rad50-Nbs1 complex. Science 2005;308:551–4.
[54] Cimprich KA, Cortez D. ATR: an essential regulator of genome integrity. Nat Rev Mol Cell Biol 2008;9:616–27.

[55] Luftig MA. Viruses and the DNA damage response: activation and antagonism. Annu Rev Virol 2014;1:605–25.
[56] Weitzman MD, Lilley CE, Chaurushiya MS. Genomes in conflict: maintaining genome integrity during virus infection. Annu Rev Microbiol 2010;64:61–81.
[57] Blackford AN, Patel RN, Forrester NA, Theil K, Groitl P, Stewart GS, et al. Adenovirus 12 E4orf6 inhibits ATR activation by promoting TOPBP1 degradation. Proc Natl Acad Sci USA 2010;107:12251–6.
[58] Weiden MD, Ginsberg HS. Deletion of the E4 region of the genome produces adenovirus DNA concatemers. Proc Natl Acad Sci USA 1994;91: 153–7.
[59] Wilkinson DE, Weller SK. Herpes simplex virus type I disrupts the ATR-dependent DNA-damage response during lytic infection. J Cell Sci 2006;119:2695–703.
[60] Liang X, Pickering MT, Cho NH, Chang H, Volkert MR, Kowalik TF, et al. Deregulation of DNA damage signal transduction by herpesvirus latency-associated M2. J Virol 2006;80:5862–74.
[61] Dahl J, You J, Benjamin TL. Induction and utilization of an ATM signaling pathway by polyomavirus. J Virol 2005;79:13007–17.
[62] Sowd GA, Li NY, Fanning E. ATM and ATR activities maintain replication fork integrity during SV40 chromatin replication. PLoS Pathog 2013;9:e1003283.
[63] Gillespie KA, Mehta KP, Laimins LA, Moody CA. Human papillomaviruses recruit cellular DNA repair and homologous recombination factors to viral replication centers. J Virol 2012;86:9520–6.
[64] Luo Y, Qiu J. Parvovirus infection-induced DNA damage response. Future Virol 2013;8:245–57.
[65] Liu M, Deora R, Doulatov SR, Gingery M, Eiserling FA, Preston A, et al. Reverse transcriptase-mediated tropism switching in *Bordetella* bacteriophage. Science 2002;295:2091–4.
[66] Guo H, Tse LV, Nieh AW, Czornyj E, Williams S, Oukil S, et al. Target site recognition by a diversity-generating retroelement. PLoS Genet 2011;7:e1002414.
[67] Minot S, Grunberg S, Wu GD, Lewis JD, Bushman FD. Hypervariable loci in the human gut virome. Proc Natl Acad Sci USA 2012;109:3962–6.
[68] Arambula D, Wong W, Medhekar BA, Guo H, Gingery M, Czornyj E, et al. Surface display of a massively variable lipoprotein by a *Legionella* diversity-generating retroelement. Proc Natl Acad Sci USA 2013;110:8212–7.
[69] Doulatov S, Hodes A, Dai L, Mandhana N, Liu M, Deora R, et al. Tropism switching in *Bordetella* bacteriophage defines a family of diversity-generating retroelements. Nature 2004;431:476–81.
[70] Paul BG, Bagby SC, Czornyj E, Arambula D, Handa S, Sczyrba A, et al. Targeted diversity generation by intraterrestrial archaea and archaeal viruses. Nat Commun 2015;6:6585.
[71] Moss B, Winters E, Cooper N. Instability and reiteration of DNA sequences within the vaccinia virus genome. Proc Natl Acad Sci USA 1981;78:1614–8.
[72] Baroudy BM, Moss B. Sequence homologies of diverse length tandem repetitions near ends of vaccinia virus genome suggest unequal crossing over. Nucleic Acids Res 1982;10:5673–9.
[73] Munir M, Berg M. The multiple faces of proteinkinase R in antiviral defense. Virulence 2013;4:85–9.
[74] Elde NC, Child SJ, Geballe AP, Malik HS. Protein kinase R reveals an evolutionary model for defeating viral mimicry. Nature 2009;457:485–9.
[75] Rothenburg S, Seo EJ, Gibbs JS, Dever TE, Dittmar K. Rapid evolution of protein kinase PKR alters sensitivity to viral inhibitors. Nat Struct Mol Biol 2009;16:63–70.
[76] Kerr PJ, Hone J, Perrin L, French N, Williams CK. Molecular and serological analysis of the epidemiology of myxoma virus in rabbits. Vet Microbiol 2010;143:167–78.
[77] Martin DP, Biagini P, Lefeuvre P, Golden M, Roumagnac P, Varsani A. Recombination in eukaryotic single stranded DNA viruses. Viruses 2011;3:1699–738.
[78] Simon-Loriere E, Holmes EC. Why do RNA viruses recombine? Nat Rev Microbiol 2011;9:617–26.
[79] Thiry E, Meurens F, Muylkens B, McVoy M, Gogev S, Thiry J, et al. Recombination in alphaherpesviruses. Rev Med Virol 2005;15:89–103.
[80] Ehrlich SD, Bierne H, d'Alencon E, Vilette D, Petranovic M, Noirot P, et al. Mechanisms of illegitimate recombination. Gene 1993;135:161–6.
[81] Weller SK, Sawitzke JA. Recombination promoted by DNA viruses: phage lambda to herpes simplex virus. Annu Rev Microbiol 2014;68:237–58.
[82] Harris RS, Dudley JP. APOBECs and virus restriction. Virology 2015;479–480C:131–45.
[83] Armitage AE, Katzourakis A, de OT, Welch JJ, Belshaw R, Bishop KN, et al. Conserved footprints of APOBEC3G on Hypermutated human immunodeficiency virus type 1 and human endogenous retrovirus HERV-K(HML2) sequences. J Virol 2008;82:8743–61.
[84] Vieira VC, Soares MA. The role of cytidine deaminases on innate immune responses against human viral infections. Biomed Res Int 2013;2013:683095.
[85] Harris RS, Bishop KN, Sheehy AM, Craig HM, Petersen-Mahrt SK, Watt IN, et al. DNA deamination mediates innate immunity to retroviral infection. Cell 2003;113:803–9.
[86] Mangeat B, Turelli P, Caron G, Friedli M, Perrin L, Trono D. Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. Nature 2003;424:99–103.
[87] Turelli P, Mangeat B, Jost S, Vianin S, Trono D. Inhibition of hepatitis B virus replication by APOBEC3G. Science 2004;303:1829.
[88] Suspène R, Guetard D, Henry M, Sommer P, Wain-Hobson S, Vartanian JP. Extensive editing of both hepatitis B virus DNA strands by APOBEC3 cytidine deaminases in vitro and in vivo. Proc Natl Acad Sci USA 2005;102:8321–6.
[89] Baumert TF, Rosler C, Malim MH, von WF. Hepatitis B virus DNA is subject to extensive editing by the human deaminase APOBEC3C. Hepatology 2007;46:682–9.
[90] Vartanian JP, Henry M, Marchio A, Suspène R, Aynaud MM, Guetard D, et al. Massive APOBEC3 editing of hepatitis B viral DNA in cirrhosis. PLoS Pathog 2010;6:e1000928.

[91] Vartanian JP, Guetard D, Henry M, Wain-Hobson S. Evidence for editing of human papillomavirus DNA by APOBEC3 in benign and precancerous lesions. Science 2008;320:230–3.
[92] Chen H, Lilley CE, Yu Q, Lee DV, Chou J, Narvaiza I, et al. APOBEC3A is a potent inhibitor of adeno-associated virus and retrotransposons. Curr Biol 2006;16:480–5.
[93] Tsuge M, Noguchi C, Akiyama R, Matsushita M, Kunihiro K, Tanaka S, et al. G to A hypermutation of TT virus. Virus Res 2010;149:211–6.
[94] Suspène R, Aynaud MM, Koch S, Pasdeloup D, Labetoulle M, Gaertner B, et al. Genetic editing of herpes simplex virus 1 and Epstein-Barr herpesvirus genomes by human APOBEC3 cytidine deaminases in culture and in vivo. J Virol 2011;85:7594–602.
[95] Land AM, Ball TB, Luo M, Pilon R, Sandstrom P, Embree JE, et al. Human immunodeficiency virus (HIV) type 1 proviral hypermutation correlates with CD4 count in HIV-infected women from Kenya. J Virol 2008;82:8172–82.
[96] Cuevas JM, Geller R, Garijo R, López-Aldeguer J, Sanjuán R. Extremely high mutation rate of HIV-1 in vivo. PLoS Biol 2015;13:e1002251.
[97] Berkhout B, de Ronde A. APOBEC3G versus reverse transcriptase in the generation of HIV-1 drug-resistance mutations. AIDS 2004;18:1861–3.
[98] Wood N, Bhattacharya T, Keele BF, Giorgi E, Liu M, Gaschen B, et al. HIV evolution in early infection: selection pressures, patterns of insertion and deletion, and the impact of APOBEC. PLoS Pathog 2009;5:e1000414.
[99] Kim EY, Bhattacharya T, Kunstman K, Swantek P, Koning FA, Malim MH, et al. Human APOBEC3G-mediated editing can promote HIV-1 sequence diversification and accelerate adaptation to selective pressure. J Virol 2010;84:10402–5.