Published in final edited form as:

ACS Infect Dis. 2018 June 08; 4(6): 871–880. doi:10.1021/acsinfecdis.7b00273.

CRISPR/Cas9 and Genome Editing for Viral Disease - Is Resistance Futile?

Harshana S De Silva Feelixge¹, Daniel Stone¹, Pavitra Roychoudhury², Martine Aubert¹, and Keith R Jerome^{1,2,3,*}

¹Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, Seattle 98109, WA, USA

²Department of Laboratory Medicine, University of Washington, 1959 NE Pacific St, Seattle 98195, WA, USA

³Department of Microbiology, University of Washington, 1959 NE Pacific St, Seattle 98195, WA, USA

Abstract

Chronic viral infections remain a major public health issue affecting millions of people worldwide. Highly active antiviral treatments have significantly improved prognosis and infection-related morbidity and mortality, but have failed to eliminate persistent viral forms. Therefore, new strategies to either eradicate or control these viral reservoirs are paramount to allow patients to stop antiretroviral therapy and realize a cure. Viral genome disruption based on gene editing by programmable endonucleases is one promising curative gene therapy approach. Recent findings on RNA-guided Human immunodeficiency virus (HIV)-1 genome cleavage by Cas9 and other gene-editing enzymes in latently infected cells have shown high levels of site-specific genome disruption and potent inhibition of virus replication. However, HIV-1 can readily develop resistance to genome editing at a single antiviral target site. Current data suggests that cellular repair associated with DNA double stranded breaks (DSBs) can accelerate emergence of resistance. On the other hand, a combination antiviral target strategy can exploit the same repair mechanism to functionally cure HIV-1 infection *in vitro* while avoiding the development of resistance. This perspective summarizes recent findings on the biology of resistance to genome editing, and discusses the significance of viral genetic diversity on application of gene-editing strategies towards cure.

Graphical abstract

*Corresponding author information, kjerome@fredhutch.org.

Conflict of Interest

The authors declare no competing financial interest.

Keywords

CRISPR/Cas9; CRISPR Resistance; Genetic Diversity; HIV; Hepatitis B virus; Herpes simplex virus

While modern antiviral therapies can have remarkable clinical value, for some persistent infections such therapy can only suppress viral replication, and once therapy is stopped the virus re-emerges. For example, current combination anti-retroviral therapy is highly successful in suppressing HIV plasma viremia below detectable levels, but it is not able to cure HIV infections. A long-lived latent viral reservoir that harbors replication-competent proviruses enables HIV to persist, and facilitates new cycles of infection as soon as antiviral therapy is interrupted. Similarly, antivirals for hepatitis B virus (HBV) or herpes simplex virus (HSV) can effectively suppress viremia or mucosal shedding, but when treatment is stopped both processes resume. An alternative and promising curative gene therapy approach that capitalizes on major advances in targeted endonucleases and gene editing has been gaining popularity in recent years. In this perspective, we will discuss genome editing using clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9 (CRISPR/Cas9) or other enzymes against viral targets, and explore the emergence of resistance to such therapies. Particular emphasis will be paid to HIV, both because the preponderance of the literature to date has focused on it, and because the biology of this virus favors the emergence of resistance. Finally, although gene editing of viral coreceptors such as C-C chemokine receptor type 5 (CCR5) or other host dependency factors has shown substantial promise, the issues around potential resistance are substantially different for these targets, and as such they will not be addressed here.

The idea that DNA-modifying enzymes could be utilized to selectively eliminate persistent viral DNA forms was first experimentally demonstrated by Buchholz and colleagues in 2007¹. They utilized an evolved Cre recombinase to recognize a sequence within the long terminal repeat (LTR) of an HIV strain to efficiently excise the HIV provirus from the human genome, essentially curing cells of infection. However, LTR sequences from clinically-relevant HIV strains are divergent from the pseudo loxP site found in the LTR recognized by this evolved Cre recombinase, thus limiting its broad therapeutic applicability. Further advances in programmable gene-editing enzymes such as homing endonucleases (HEs), zinc finger nucleases (ZFNs), Transcription activator-like effector nucleases (TALENs), and CRISPR/Cas9 have allowed the field to adapt and expand this approach towards other viral genes and viral elements^{2–5}.

Endonucleases can be engineered to recognize a specific target DNA sequence and induce DNA double strand breaks (DSBs), thereby activating the cellular DNA DSB repair machinery. In the absence of a homologous DNA repair template, non-homologous end joining (NHEJ) occurs, leading to mutations at the cleavage site. Since NHEJ simply ligates broken DNA ends with little to no homology, it often introduces small insertions, deletions and/or substitutions. This trait can be manipulated to specifically introduce lethal mutations within viral sequences that are essential for virus replication, and therefore prevent productive viral replication. The first example of gene editing in the context of HIV was demonstrated by Aubert and colleagues⁶. They showed that an engineered HE can be directed towards an integrated proviral sequence and can induce mutations within the integrated proviral DNA. Furthermore, they showed that mutations derived through HES could lead to successful gene disruption that eliminates functional protein expression; suggesting that a functional cure for chronic viral infections may be achieved by sufficiently mutating essential viral genes to prevent viral replication and continued pathogenesis.

The CRISPR/Cas9 system has recently gained popularity as a gene-editing platform. CRISPR/Cas9 is a RNA-guided DNA endonuclease system that uses RNA-DNA complementarity to induce DSBs. By simply interchanging a short 20 base pair guide RNA (gRNA) sequence that associates a gene target with Cas9, it is possible to target and edit virtually any region of DNA. Cas9 binding and cleavage requires recognition of a short conserved sequence known as the protospacer adjacent motif (PAM) immediately after the 3' end of the gRNA complementary DNA sequence, and the presence of both gRNA complementary sequence and the PAM domain is necessary for Cas9 catalytic function⁷⁻⁸. The CRISPR/Cas9 system has driven recent interest in HIV genome editing because it is relatively simple, quick, and affordable to engineer. Consequently, it has been studied both *in vitro* and *in vivo*⁹⁻¹². A recent report took an important step forward by specifically excising the HIV provirus *in vivo*, using the BLT mouse model with an established chronic infection¹³. Yin et.al were able to deliver an AAV vector containing multiple different gRNAs along with Cas9 to latently infected cells, and detected HIV provirus excision/ mutations in multiple tissues and organs. These data support the potential application of genome editing therapies towards an HIV cure. However, the potential of this strategy as a curative therapy for HIV/AIDS or other viral infections would rely on the ability to avoid emergence of genetic resistance to CRISPR/Cas9.

Resistance to Genome Editing: a single gRNA is not sufficient to prevent viral escape

It is not surprising that HIV is able to escape genome editing, since the virus has been demonstrated to develop resistance to antiretroviral drugs, a multitude of immune pressures, and RNA-based shRNA antiviral therapeutics¹⁴⁻¹⁵. Emergence of resistance to antiretrovirals are directly attributable to the genetic diversity generated during error-prone reverse transcription, provirus transcription by RNA polymerase II, or direct apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3G (APOBEC3G)-derived hypermutation of viral RNAs¹⁶⁻¹⁹. For genome editing, escape mutations can occur through these mechanisms, and can also be facilitated by insertions, deletions and substitutions

acquired through error-prone NHEJ (Figure 1). While most mutations acquired through genome editing have deleterious effects on viral replication, some retain their coding capacity and are refractory to recognition by the same gene editing enzyme, thus providing a new template for viral resistance. The first evidence implicating DSB repair as a means of viral escape was shown by De Silva Felixge and colleagues²⁰. While assessing the infectivity profile of individual mutations derived after HIV POL-targeted ZFN treatment, they observed that despite successful genome editing, some mutants remained replication competent. Importantly, one fully replication-competent mutant was identified after therapy that had an insertion of 3 nucleotides within the ZFN target site, along with significantly diminished ZFN binding and susceptibility to re-cleavage, suggesting that this mutant was resistant to endonuclease therapy. Subsequently, several additional studies demonstrated that despite strong inhibition of virus replication, resistance to CRISPR/Cas9 gRNA could also eventually emerge. Multiple studies have shown that regardless of which regions a Cas9 single guide RNA (sgRNA) targets, and whether the target overlaps open reading frames of essential viral genes such as gag/pol and env/rev, or repetitive viral elements like the U3 region of the LTR, HIV is able to rapidly escape Cas9 inhibition. Importantly, almost all HIV escape variants arising after CRISPR/Cas9 treatment contained mutations clustered around the PAM recognition domain or at the Cas9 cleavage site (3bp away from the PAM). Sequence analysis of these escape variants confirm that resistance to CRISPR/Cas9 is specific to mutations acquired at the gRNA target sequence, which cause diminished or ablated interaction between the sgRNA and its cognate target sequence^{21–24}.

Resistance to almost all known antiviral drugs emerges predominantly through reverse transcriptase (RT)-derived base substitutions. Many of the resistant strains that emerge after Cas9 genome editing contain indels that are uncharacteristic of typical HIV evolution. However, this mutation spectrum is the hallmark of NHEJ, suggesting a clear role for error-prone DNA repair in generating viral resistance. Three recent studies evaluating Cas9 escape variants showed that mutations induced at Cas9 cleavage sites are often single base insertions. However, a significant number of gRNA sites in coding regions are altered by ± 3 base pairs, preserving open reading frames, and therefore potentially enabling the provirus to maintain replication fitness. Certain regions, such as the RRE RNA stem-loop or RRE bulge, were able to tolerate even larger multiple base indels as long as the proviruses retained correct reading frames^{21, 23–24}. In contrast, RT-derived errors largely consist of single base substitutions with transitions predominating over transversions. G-to-A transitions are the most common form of substitutions for RT-derived errors, but in the presence of Cas9, mutations with A-to-C transversions were predominant^{25–28}. RT-derived insertions, deletions, or deletions with insertions (indels) are particularly rare. These indels, regardless of their size, are generally accompanied by a short direct repeat of 1 to 3 nucleotides and/or a preceding run of identical nucleotides in the direction of the minus strand, and often yield replication-deficient progeny²⁹.

Several features of endonuclease design may determine how quickly resistance to genome editing can emerge. In a recent study, Wang et al observed that sgRNA target sites that are highly conserved across HIV isolates retained anti-viral activity longer compared to those that were poorly conserved. Specifically, regions that are highly conserved across essential protein coding regions of HIV, with a Shannon entropy of less than 0.2, had significantly

delayed viral escape. Not surprisingly, the escape variants that eventually emerged contained only minimal substitutions of 1 or 2 nucleotides, most of which were silent codon changes. Furthermore, indels were largely restricted to 3 nt insertions, such that they maintained the open reading frame and coding capacity²². This delayed escape is likely due to fewer potential escape variants in conserved regions. For example, highly-conserved regions such as structural proteins, integrase, and protease are all critical for transcription and may not tolerate large deletions or insertions, because there is pressure to maintain their coding potential. Similarly, critical transcriptional regulation motifs like TATA box in the U3 region of the HIV LTR may not tolerate even minor nucleotide changes, as it may disrupt transcriptional activity and therefore overall viral fitness³⁰. Nonetheless, even against the most-conserved sgRNAs that can potently suppress virus production, virus replication eventually breaks through.

Another factor that is not immediately apparent is the contribution of endonuclease potency to the rapid emergence of viral escape. In one study, Lebbink et al evaluated a panel of sgRNAs with various HIV inhibitory profiles and defined the mutational landscape of the viruses that could replicate in the presence of these sgRNAs³¹. Deep sequencing showed that in the presence of highly potent sgRNAs that are efficient DSB inducers, viral replication quickly occurs from multiple escape mutants. In contrast, replication is driven by the wildtype or founder viruses in the presence of low-potency sgRNAs³¹. Taken together, these data highlight that, in addition to errors introduced during the normal course of virus evolution, NHEJ repair associated with genome editing functions as a site-specific mutagen enhancing HIV diversity, and thus accelerates treatment resistance.

Combination therapy as a means to circumvent resistance

Two recent studies have provided promising data that a dual gRNA system targeting genes involved in different steps of the virus replication cycle can halt HIV replication and prevent viral escape^{22, 31}. However, complete ablation of viral escape was dependent upon both the therapeutic potency of the gRNA combination and the conservation of the target sites. Both studies showed that a combination of at least two potent sgRNAs, defined by their ability to strongly inhibit both viral gene expression and replication, is required to overcome the plasticity of HIV-1. It may be that potent sgRNAs can simultaneously induce efficient DSBs, thereby introducing large deletions that render the provirus inactive. In contrast, sub-optimal combinations may drive resistance similar to single gRNA therapy, as inefficient cleavage of one of the target sites will only result in partially edited genomes at either target site leading to incomplete control of HIV infection. Variants escaping one gRNA may subsequently be targeted at the second gRNA target site during subsequent rounds of infection, ultimately resulting in viral escape from combination therapy. Indeed, Lebbink et al showed that nearly all virus variants replicating in the presence of sub-optimal gRNA combinations contained escape mutations at all CRISPR/Cas9 target regions³¹. Additionally, deep sequencing analysis of proviruses treated with a strong and a weak gRNA simultaneously demonstrated the presence of escape mutations at both target sites in the same viral genome, further validating that sub-optimal gRNA pairing can readily yield resistance. Combination gRNA therapy that exclusively targets essential structural and regulatory proteins has also been shown to significantly delay/prevent the emergence of resistant variants³¹.

A recent study showed that there is a consistent and predictable pattern of mutagenic DNA repair following Cas9 cleavage *in vitro*³². Particularly, there is a chronological shift in the distribution of indels. Single-base pair insertions and small deletions (1–2 nucleotides) are dominant at early time points, whereas larger deletions appear at later time points, suggesting that minimally-edited target site variants may be vulnerable to recognition and re-cleavage by Cas9. Similarly, in the context of HIV dual gRNA therapy, Wang *et al* also observed continuous recognition of mildly mutated targets by Cas9, albeit with reduced efficiency. Long-term exposure to dual gRNA therapy (over 100 days) resulted in gradual emergence of larger deletions and insertions, and progressive loss of minor mutations as well as low-frequency wild type sequences. Phenotypic analysis of viral fitness corroborated this data, with complete loss of detectable replication competent virus over time. Overall, this data suggests that one way to address HIV resistance may be to prolong exposure to combination therapy, to allow favorable accumulation of hypermutations and/or larger indels over time. In the clinic, escape from CRISPR/Cas9 would need to occur from within the integrated proviral reservoir, in which virus presumably would not replicate as long as patients remain on ART during CRISPR/Cas9 therapy. If used in combination the CRISPR/Cas9 machinery could be exploited to create a viral graveyard of defective proviruses and thereby increase the likelihood of a cure³³. Whether such an effect will be observed with newer high-fidelity Cas9 variants remains to be investigated.

Quasispecies

One of the major obstacles for genome editing towards an HIV cure involves effective targeting of all the genetic viral variants existing within an infected individual. Within-host diversity, also known as viral quasispecies, has important implications for emergence of resistance. Inadequate targeting of certain genetic variants could lead to rebound and reseeded of the reservoir upon treatment³⁴. Thus far, studies used to demonstrate the viability of genome editing strategies have primarily focused on single molecular clones that provide ideal endonuclease target site recognition. However, within HIV-infected individuals, any given target site may be heterogeneous, with variants present at differing frequencies. In specific anatomical reservoir sites, certain target site variants may even be completely absent. The HIV quasispecies are dynamic. In the absence of cART the average rate of HIV evolution is estimated to be 5.3 mutations/kilobase (kb)/year. This rate gradually decreases to an average 1.02 mutations/kb/year in the total PBMC compartment, and 0.01–0.15 nucleotides/kb/year in latent resting memory CD4+ T cells during long-term cART suppressive therapy^{35–36}. Various selective pressures such as immune responses, therapeutic interventions, a person's general health, or simply homeostatic proliferation are thought to affect the distribution of pre-existing viral variants³⁵.

A recent modeling study predicts that as many as 10 HIV gRNAs targeting the LTR would be necessary to completely excise all known viral quasispecies present within the cellular PBMC compartment³⁷. Furthermore, a complete gRNA regimen could not be designed for a subset of infected individuals, in spite of their well-controlled infections (viral loads <50 copies/mL). Longitudinal next generation sequencing (NGS) sequencing analysis of one of the patients for whom a complete gRNA regimen could not be designed showed that an additional year of ART made such a regimen feasible. This corroborated previous reports

that long-term suppressive therapy can stabilize the reservoir and the amount of variation present within an individual^{36–37}. One outstanding issue is the number of gRNAs that could potentially be delivered as a single dose to achieve an effective therapeutic outcome. Even with customized target design catered toward a specific patient, NSG data will not provide information on the quasispecies present within all anatomical sites. Thus, synchronous delivery of a complete gRNA regimen to each target cell may be required. However, expression of multiple gRNAs within a single target cell may also augment off-target effects. Recently, significant efforts have been made to substantially reduce Cas9 off target cleavage. Newer high-fidelity Cas9 variants such as Cas9-HF, eSpCas9, and HypaCas9 have been shown to exhibit high on target activity and low off target activity^{38–39}. Nonetheless, an exclusive focus on excision of proviral DNA, by targeting the LTRs alone, may make therapeutic application of genome editing complex and ultimately ineffective.

A more effective solution to deal with target sequence diversity may be to target essential viral protein-coding regions that are highly conserved across various HIV isolates (Figure 2). A recent study revealed that across different HIV isolates, the average amino acid diversity is generally higher in HIV proteins that do not multimerize. In comparison, proteins that multimerize, such as capsid, nucleocapsid, protease, integrase, reverse transcriptase, Vpr, and the GP41 N terminus, were highly conserved, suggesting that these regions may have low tolerance for non-synonymous changes, thereby constraining evolution of resistance⁴⁰. In addition, multiplexed therapy with at least 3 gRNAs targeting distinct and highly conserved regions of the proviral genome would lead to either multiple mutations at different sites and/or deletion of large regions of the provirus¹³.

Global genetic diversity

Based on nucleotide diversity, HIV is divided into groups, of which group M accounts for over 90% of infections. Group M has been further classified into 9 different subtypes. The average divergence between HIV groups and subtypes are 25–35% and 15–20% respectively⁴¹. Furthermore, inter-subtype recombinants or circulating recombinant forms (CRFs) are also assumed to make an epidemiologically relevant contribution to the HIV epidemic. Currently there are over 50 CRFs, with a global prevalence estimated to be around 20%, which adds significant complexity to HIV genetic diversity⁴². Thus, one outstanding issue for genome editing is the development and application of a broadly-applicable and resistance-proof gRNA regimen that could account for global HIV genetic diversity. One way to overcome HIV diversity may be the development of individualized gRNA regimens that are quasispecies specific. However, such an approach would not be economical or practical in resource-limited settings.

Within essential viral protein-coding regions, nucleotide and/or amino acid diversity may be limited across different subtypes and CRFs. This may allow the design of a combination gRNA regimen to account for all group M viruses. Such an off the shelf therapeutic approach could be widely implementable. However, the most effective way forward may be a subtype specific gRNA regimen. In addition to considerably narrowing the genetic diversity, such an approach may also increase the genetic barrier for endonuclease resistance. It has been shown that there are subtype-specific propensities towards resistance

to different classes of antiretroviral drugs. Factors such as intrinsic differences in codon usages, amino acid differences associated with minor structural changes, and differences in sequence motifs have been attributed to this differential resistance spectrum⁴¹. Such intrinsic properties exist within HIV subtypes, and may also influence resistance profiles to CRISPR/Cas9 or other endonuclease-mediated genome editing therapy.

The HIV Reservoir

Currently, the lack of an efficient method to deliver therapeutic enzymes to latently-infected cells *in vivo* has hindered the application of proviral gene editing efforts towards a cure. The HIV reservoir is composed of infected cells that harbor transcriptionally-silent but replication-competent forms of viruses, and these infected cells are long lived and enable future reseeding of the reservoir. Quiescent CD4⁺ T lymphocytes with a memory phenotype are the best-characterized cellular reservoirs of HIV infection, though other cell types may contribute^{43–44}. Our current understanding of the reservoir suggests that latently infected cells harboring replication-competent virus are rare, with a frequency of around 1 per 10⁶ resting CD4⁺ T cells, although recent evidence suggests that this may be higher than previously estimated, at approximately 12–37 per 10⁶ resting CD4⁺T cells^{45–47}. Any meaningful effect on the reservoir will therefore rely on efficient and high-level gene delivery to these rare cells. Genome editing efforts will likely be attempted on individuals on long-term ART suppressive therapy. The ideal goal of endonuclease therapy is complete inactivation of provirus in all latently infected cells. However, a reasonable interim goal for endonuclease therapy is to achieve a reduction in replication-competent virus to a level that results in long-term virologic remission without the need for ongoing antiviral therapy. Mathematical models based on case reports of patients who achieved a reduction in the latent reservoir suggest that 2 to 3 logs of reservoir reduction are required to delay virus rebound for a few months, while > 4 logs reduction may prevent viral rebound altogether^{48–51}. Deeper understanding of the relevant contributions of different anatomical reservoir sites on residual viremia and viral rebound following treatment interruption would also help define characteristics of optimal delivery candidates. It may be that physiologically significant levels of reservoir reduction would require delivery of genome editing enzymes to only a selected number of reservoir sites. For instance, lymphoid tissue comprises a major reservoir site. A substantial fraction of the total lymphocytes in the body, approximately 40–60%, reside in the lymphatic system or in specialized sites along the gastrointestinal tract (Gut-Associated Lymphoid Tissue)^{52–54}. It is possible that reduction/elimination of proviruses only in such pathophysiologically-relevant latent tissue reservoirs might be sufficient to achieve a functional cure.

To fully tackle the issue of resistance, other secondary challenges also need to be addressed. One difficult problem is the dynamic range of assays to measure reservoir reduction. Assays that can accurately predict the time to viral rebound following genome editing are lacking. With a proviral gene editing strategy, patients who retain mutated viral sequences could be considered cured, but standard PCR-based assays are not able to differentiate between replication-competent proviruses and proviruses that have been either lethally mutated through gene editing enzymes or those that are permanently mutated but confer resistance to therapy. On the other hand, viral outgrowth assays underestimate the latent reservoir size,

because not all replication-competent proviruses can be induced by a single round of T cell activation, and thus these assay provide only a minimal estimate of the reservoir size. Depending on the log reduction of the reservoir, viral outgrowth assays may not even be feasible, as low levels of latently infected cells would require impractically large amounts of PBMCs to detect replication-competent viruses. Additionally, these types of assays may not be able to accurately detect/distinguish minor and rare subsets of proviruses that are escape variants^{55–57}. Thus, new ultra-sensitive assays that can accurately quantify the reservoir will be essential to determine the role of resistance to proviral genome editing therapy.

Other Viruses - Hepatitis B virus

Although antiviral treatments that effectively inhibit virus replication are available for hepatitis B virus (HBV) infections, they are not able to eradicate HBV covalently closed circular DNA (cccDNA), which has been the most challenging obstacle to developing a cure. The HBV viral genome is a partially double-stranded relaxed circular DNA (rcDNA), which is released into the nucleus upon infection and repaired to form cccDNA, which in turn acts as the template for all transcription and production of sub-genomic RNAs. cccDNA is highly stable, persists in the infected cell for its lifetime, and is resistant to antiviral therapy⁵⁸. Genome editing strategies that can either lethally mutate or degrade cccDNA represent a promising future treatment for HBV cure^{5, 59–67}. Similar to HIV, emergence of escape mutants following cccDNA genome editing needs to be addressed prior to its clinical application. HBV is a unique DNA virus that utilizes reverse transcription to create an RNA intermediate in order to replicate its genome. As a result, it has a very high rate of mutation for a DNA virus, approximately 10-fold higher than other DNA viruses. The high rate of mutations, approximately 10^{-5} nucleotide substitutions per site per year, and large number of virus progeny, $10^{11} - 10^{12}$ virions/day, allow HBV to generate tremendous genetic variation similar to that of HIV^{68–69}. There are 10 HBV genotypes (A–J) worldwide, with an approximate nucleotide divergence of >8%^{70–71}. However, unlike HIV, HBV evolution is constrained due to the extremely compact arrangement of the ORFs and regulatory elements. The four ORFs that encode for 7 essential viral proteins all at least partially overlap, while promoters, enhancers, and other regulatory elements are embedded within these ORFs. The multifunctional nature of its genome limits the plasticity of HBV and significantly increases its barrier to resistance. Therefore, drug resistance emerges much more slowly for HBV than for HIV. Based on this data one could predict that evolution of HBV escape following cccDNA editing and repair may also arise much more slowly compared to HIV⁷². Schiffer *et al.* previously developed mathematical models to predict evolution of viral mutants following cccDNA genome editing. The models predict that similar to HIV, a combination approach targeting different regions of cccDNA simultaneously should prevent emergence of endonuclease resistance; however, this prediction must be experimentally confirmed⁷³. cccDNA eradication/inactivation efforts are also hindered by secondary challenges. *In vivo* models that effectively mimic natural HBV infection are few, complex, and expensive. Consequently, accurate assessment of the efficacy of cccDNA genome editing *in vivo* has proved challenging. cccDNA is present at low concentrations in HBV-infected cells, and is estimated to be present at 1–10 copies per cell even in the absence of antiretrovirals. Furthermore, assays to measure cccDNA are complex and poorly standardized. Without

reliable methods to accurately evaluate cccDNA *in vivo*, assessing outcomes of resistance to genome editing will be extremely difficult⁷⁴.

Other Viruses - Herpes Simplex Virus

Herpes simplex virus (HSV) infections are highly prevalent and can result in serious disease outcomes, particularly in newborns and immunocompromised individuals^{75–76}. The most common complications of infection include oral and anogenital ulcers, corneal blindness, and encephalitis⁷⁷. In the United States alone, prevalence of neutralizing antibodies against HSV-1 and HSV-2 exceeds 50% and 16% respectively^{78–79}. Similar to HIV and HBV, replicating HSV is highly responsive to antiviral therapy, but antiviral therapy does not eliminate persistent infection and does not lead to a cure⁸⁰. HSV establishes latency in sensory neuronal cell bodies in either the trigeminal ganglia or dorsal root ganglia, where latent viral genomes exist as episomes, (2–50 genomes/neuron), and reactivate recurrently to cause lesions and/or episodes of viral shedding at the mucosa^{81–82}. HSV is an enveloped virus with a large linear double-stranded DNA genomes, and requires a large number of virally encoded proteins for replication^{83–84}. This makes it especially amenable to lethal mutagenesis through genome editing^{3, 5, 64, 67, 85}. For instance, HSV-1 encodes for 84 viral proteins, 31 of which are essential for virus replication in culture⁸⁶. Emergence of resistance following HSV genome editing could presumably pose a barrier toward an HSV cure. However, despite its high prevalence, the genetic diversity of HSV is low compared to HIV and HBV, owing to its high-fidelity DNA polymerase (3.5×10^{-8} mutations/site/year)⁸⁷. Phylogenetic analysis of worldwide clinical HSV-1 and HSV-2 isolates also show high nucleotide conservation among viral genes, with mutation rates ranging from 0.6% to 3.1% depending on the gene^{88–89}. These characteristics, along with its reactivation pattern, strongly favor a high genetic threshold for evolution of viral escape. One study in mice showed that HSV viral burden appears to be correlated with recurrent rates of reactivation, suggesting that it may not be necessary to eliminate all copies of the latent genome within all infected neurons to achieve a functional cure⁹⁰. Curative therapies that rely on overall reduction in viral load may be less affected by the presence of escape mutations. Despite the presence of resistant episomal genomic DNA, if the therapy is able to establish physiologically-relevant levels of viral inactivation, these resistance variants may never reactivate. Additionally, resistance variants generated during genome editing are likely to have lower viral fitness, and therefore may be easily controlled by the immune system. Nonetheless, comprehensive preclinical evaluation is essential to determine the parameters for a cure.

Conclusions

Chronic viral infections are responsible for millions of infections and deaths each year, causing immense disease and financial burdens. Genome editing strategies that eliminate and inactivate latent viral genomes are exciting approaches providing hope for cures. The development of treatment resistance to genome editing strategies is an important concern that needs to be addressed prior to their clinical application. However, excessive concern regarding resistance may be putting the cart before the horse, as no study has yet shown meaningful antiviral efficacy *in vivo*. Emerging data suggest that rational design of

multiplexed target sites to encompass highly conserved essential viral protein coding regions may prevent resistance. Furthermore, escape variants may have reduced fitness, and therefore be more easily controlled by the immune system or have lower pathogenesis. Most importantly, future antiviral gene editing therapies will likely not be used alone, but instead will be combined with traditional antivirals, which should also help limit the emergence of treatment resistance.

Acknowledgments

This work was supported by NIH grant UM1 AI126623 to KR Jerome and HP Kiem.

Abbreviations used

DSB	DNA double stranded breaks
CRISPR/Cas9	clustered regularly interspaced short palindromic repeats/ CRISPR-associated protein 9
CCR5	C-C chemokine receptor type 5
LTR	long terminal repeat
HE	homing endonuclease
ZFN	zinc finger nuclease
TALEN	Transcription activator-like effector nuclease
NHEJ	non-homologous end joining
gRNA	guide RNA
sgRNA	single guide RNA
PAM	protospacer adjacent motif
RT	reverse transcriptase
kb	kilobase
NGS	next generation sequencing
CRF	circulating recombinant form
cccDNA	covalently closed circular DNA
rcDNA	relaxed circular DNA

References

1. Sarkar I, Hauber I, Hauber J, Buchholz F. HIV-1 proviral DNA excision using an evolved recombinase. *Science*. 2007; 316(5833):1912–1915. DOI: 10.1126/science.1141453 [PubMed: 17600219]

2. Buchholz F, Hauber J. Antiviral therapy of persistent viral infection using genome editing. *Curr Opin Virol.* 2016; 20:85–91. DOI: 10.1016/j.coviro.2016.09.012 [PubMed: 27723558]
3. Schiffer JT, Aubert M, Weber ND, Mintzer E, Stone D, Jerome KR. Targeted DNA mutagenesis for the cure of chronic viral infections. *J Virol.* 2012; 86(17):8920–8936. DOI: 10.1128/JVI.00052-12 [PubMed: 22718830]
4. Stone D, Kiem HP, Jerome KR. Targeted gene disruption to cure HIV. *Curr Opin HIV AIDS.* 2013; 8(3):217–223. DOI: 10.1097/COH.0b013e32835f736c [PubMed: 23478911]
5. Stone D, Niyonzima N, Jerome KR. Genome editing and the next generation of antiviral therapy. *Hum Genet.* 2016; 135(9):1071–1082. DOI: 10.1007/s00439-016-1686-2 [PubMed: 27272125]
6. Aubert M, Ryu BY, Banks L, Rawlings DJ, Scharenberg AM, Jerome KR. Successful targeting and disruption of an integrated reporter lentivirus using the engineered homing endonuclease Y2 I-Anil. *PLoS One.* 2011; 6(2):e16825.doi: 10.1371/journal.pone.0016825 [PubMed: 21399673]
7. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F. Multiplex genome engineering using CRISPR/Cas systems. *Science.* 2013; 339(6121):819–823. DOI: 10.1126/science.1231143 [PubMed: 23287718]
8. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. RNA-guided human genome engineering via Cas9. *Science.* 2013; 339(6121):823–826. DOI: 10.1126/science.1232033 [PubMed: 23287722]
9. Ebina H, Misawa N, Kanemura Y, Koyanagi Y. Harnessing the CRISPR/Cas9 system to disrupt latent HIV-1 provirus. *Sci Rep.* 2013; 3:2510.doi: 10.1038/srep02510 [PubMed: 23974631]
10. Lisowski L, Dane AP, Chu K, Zhang Y, Cunningham SC, Wilson EM, Nygaard S, Grompe M, Alexander IE, Kay MA. Selection and evaluation of clinically relevant AAV variants in a xenograft liver model. *Nature.* 2014; 506(7488):382–386. DOI: 10.1038/nature12875 [PubMed: 24390344]
11. Kaminski R, Chen Y, Fischer T, Tedaldi E, Napoli A, Zhang Y, Karn J, Hu W, Khalili K. Elimination of HIV-1 Genomes from Human T-lymphoid Cells by CRISPR/Cas9 Gene Editing. *Sci Rep.* 2016; 6:22555.doi: 10.1038/srep22555 [PubMed: 26939770]
12. Zhu W, Lei R, Le Duff Y, Li J, Guo F, Wainberg MA, Liang C. The CRISPR/Cas9 system inactivates latent HIV-1 proviral DNA. *Retrovirology.* 2015; 12:22.doi: 10.1186/s12977-015-0150-z [PubMed: 25808449]
13. Yin C, Zhang T, Qu X, Zhang Y, Putatunda R, Xiao X, Li F, Xiao W, Zhao H, Dai S, Qin X, Mo X, Young WB, Khalili K, Hu W. In Vivo Excision of HIV-1 Provirus by saCas9 and Multiplex Single-Guide RNAs in Animal Models. *Mol Ther.* 2017; 25(5):1168–1186. DOI: 10.1016/j.ymthe.2017.03.012 [PubMed: 28366764]
14. Senserrich J, Pauls E, Armand-Ugon M, Clotet-Codina I, Moncunill G, Clotet B, Este JA. HIV-1 resistance to the anti-HIV activity of a shRNA targeting a dual-coding region. *Virology.* 2008; 372(2):421–429. DOI: 10.1016/j.virol.2007.10.045 [PubMed: 18068205]
15. Shah PS, Pham NP, Schaffer DV. HIV develops indirect cross-resistance to combinatorial RNAi targeting two distinct and spatially distant sites. *Mol Ther.* 2012; 20(4):840–848. DOI: 10.1038/mt.2012.3 [PubMed: 22294151]
16. Jern P, Russell RA, Pathak VK, Coffin JM. Likely role of APOBEC3G-mediated G-to-A mutations in HIV-1 evolution and drug resistance. *PLoS Pathog.* 2009; 5(4):e1000367.doi: 10.1371/journal.ppat.1000367 [PubMed: 19343218]
17. Kim EY, Bhattacharya T, Kunstman K, Swantek P, Koning FA, Malim MH, Wolinsky SM. Human APOBEC3G-mediated editing can promote HIV-1 sequence diversification and accelerate adaptation to selective pressure. *J Virol.* 2010; 84(19):10402–10405. DOI: 10.1128/JVI.01223-10 [PubMed: 20660203]
18. Neogi U, Shet A, Sahoo PN, Bontell I, Ekstrand ML, Banerjee AC, Sonnerborg A. Human APOBEC3G-mediated hypermutation is associated with antiretroviral therapy failure in HIV-1 subtype C-infected individuals. *J Int AIDS Soc.* 2013; 16:18472.doi: 10.7448/IAS.16.1.18472 [PubMed: 23443042]
19. Sadler HA, Stenglein MD, Harris RS, Mansky LM. APOBEC3G contributes to HIV-1 variation through sublethal mutagenesis. *J Virol.* 2010; 84(14):7396–7404. DOI: 10.1128/JVI.00056-10 [PubMed: 20463080]

20. De Silva Felixge HS, Stone D, Pietz HL, Roychoudhury P, Greninger AL, Schiffer JT, Aubert M, Jerome KR. Detection of treatment-resistant infectious HIV after genome-directed antiviral endonuclease therapy. *Antiviral Res.* 2016; 126:90–98. DOI: 10.1016/j.antiviral.2015.12.007 [PubMed: 26718067]
21. Ueda S, Ebina H, Kanemura Y, Misawa N, Koyanagi Y. Anti-HIV-1 potency of the CRISPR/Cas9 system insufficient to fully inhibit viral replication. *Microbiol Immunol.* 2016; 60(7):483–496. DOI: 10.1111/1348-0421.12395 [PubMed: 27278725]
22. Wang G, Zhao N, Berkhout B, Das AT. CRISPR-Cas9 Can Inhibit HIV-1 Replication but NHEJ Repair Facilitates Virus Escape. *Mol Ther.* 2016; 24(3):522–526. DOI: 10.1038/mt.2016.24 [PubMed: 26796669]
23. Wang Z, Pan Q, Gendron P, Zhu W, Guo F, Cen S, Wainberg MA, Liang C. CRISPR/Cas9-Derived Mutations Both Inhibit HIV-1 Replication and Accelerate Viral Escape. *Cell Rep.* 2016; 15(3): 481–489. DOI: 10.1016/j.celrep.2016.03.042 [PubMed: 27068471]
24. Yoder KE, Bundschuh R. Host Double Strand Break Repair Generates HIV-1 Strains Resistant to CRISPR/Cas9. *Sci Rep.* 2016; 6:29530.doi: 10.1038/srep29530 [PubMed: 27404981]
25. Muller V, Bonhoeffer S. Guanine-adenine bias: a general property of retroviral genomes that is unrelated to host-induced hypermutation. *Trends Genet.* 2005; 21(5):264–268. DOI: 10.1016/j.tig.2005.03.004 [PubMed: 15851060]
26. Berkhout B, de Ronde A. APOBEC3G versus reverse transcriptase in the generation of HIV-1 drug-resistance mutations. *AIDS.* 2004; 18(13):1861–1863. [PubMed: 15316354]
27. Berkhout B, Das AT, Beerens N. HIV-1 RNA editing, hypermutation, and error-prone reverse transcription. *Science.* 2001; 292(5514):7.doi: 10.1126/science.292.5514.7a [PubMed: 11292864]
28. Abram ME, Ferris AL, Das K, Quinones O, Shao W, Tuske S, Alvord WG, Arnold E, Hughes SH. Mutations in HIV-1 reverse transcriptase affect the errors made in a single cycle of viral replication. *J Virol.* 2014; 88(13):7589–7601. DOI: 10.1128/JVI.00302-14 [PubMed: 24760888]
29. Abram ME, Ferris AL, Shao W, Alvord WG, Hughes SH. Nature, position, and frequency of mutations made in a single cycle of HIV-1 replication. *J Virol.* 2010; 84(19):9864–9878. DOI: 10.1128/JVI.00915-10 [PubMed: 20660205]
30. van Opijnen T, Kamoschinski J, Jeeninga RE, Berkhout B. The human immunodeficiency virus type 1 promoter contains a CATA box instead of a TATA box for optimal transcription and replication. *J Virol.* 2004; 78(13):6883–6890. DOI: 10.1128/JVI.78.13.6883-6890.2004 [PubMed: 15194764]
31. Lebbink RJ, de Jong DC, Wolters F, Kruse EM, van Ham PM, Wiertz EJ, Nijhuis M. A combinatorial CRISPR/Cas9 gene-editing approach can halt HIV replication and prevent viral escape. *Sci Rep.* 2017; 7:41968.doi: 10.1038/srep41968 [PubMed: 28176813]
32. van Overbeek M, Capurso D, Carter MM, Thompson MS, Frias E, Russ C, Reece-Hoyes JS, Nye C, Gradia S, Vidal B, Zheng J, Hoffman GR, Fuller CK, May AP. DNA Repair Profiling Reveals Nonrandom Outcomes at Cas9-Mediated Breaks. *Mol Cell.* 2016; 63(4):633–646. DOI: 10.1016/j.molcel.2016.06.037 [PubMed: 27499295]
33. Wang G, Zhao N, Berkhout B, Das AT. A Combinatorial CRISPR-Cas9 Attack on HIV-1 DNA Extinguishes All Infectious Provirus in Infected T Cell Cultures. *Cell Rep.* 2016; 17(11):2819–2826. DOI: 10.1016/j.celrep.2016.11.057 [PubMed: 27974196]
34. Domingo E, Sheldon J, Perales C. Viral quasispecies evolution. *Microbiol Mol Biol Rev.* 2012; 76(2):159–216. DOI: 10.1128/MMBR.05023-11 [PubMed: 22688811]
35. Dampier W, Nonnemacher MR, Mell J, Earl J, Ehrlich GD, Pirrone V, Aiamkitsumrit B, Zhong W, Kercher K, Passic S, Williams JW, Jacobson JM, Wigdahl B. HIV-1 Genetic Variation Resulting in the Development of New Quasispecies Continues to Be Encountered in the Peripheral Blood of Well-Suppressed Patients. *PLoS One.* 2016; 11(5):e0155382.doi: 10.1371/journal.pone.0155382 [PubMed: 27195985]
36. Josefsson L, von Stockenstrom S, Faria NR, Sinclair E, Bacchetti P, Killian M, Epling L, Tan A, Ho T, Lemey P, Shao W, Hunt PW, Somsouk M, Wylie W, Douek DC, Loeb L, Custer J, Hoh R, Poole L, Deeks SG, Hecht F, Palmer S. The HIV-1 reservoir in eight patients on long-term suppressive antiretroviral therapy is stable with few genetic changes over time. *Proc Natl Acad Sci U S A.* 2013; 110(51):E4987–4996. DOI: 10.1073/pnas.1308313110 [PubMed: 24277811]

37. Dampier W, Nonnemacher MR, Sullivan NT, Jacobson JM, Wigdahl B. HIV Excision Utilizing CRISPR/Cas9 Technology: Attacking the Proviral Quasispecies in Reservoirs to Achieve a Cure. *MOJ Immunol.* 2014; 1(4):doi: 10.15406/moji.2014.01.00022
38. Kulcsar PI, Talas A, Huszar K, Ligeti Z, Toth E, Weinhardt N, Fodor E, Welker E. Crossing enhanced and high fidelity SpCas9 nucleases to optimize specificity and cleavage. *Genome Biol.* 2017; 18(1):190.doi: 10.1186/s13059-017-1318-8 [PubMed: 28985763]
39. Chen JS, Dagdas YS, Kleinstiver BP, Welch MM, Sousa AA, Harrington LB, Sternberg SH, Joung JK, Yildiz A, Doudna JA. Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature.* 2017; 550(7676):407–410. DOI: 10.1038/nature24268 [PubMed: 28931002]
40. Li G, Piampongsant S, Faria NR, Voet A, Pineda-Pena AC, Khouri R, Lemey P, Vandamme AM, Theys K. An integrated map of HIV genome-wide variation from a population perspective. *Retrovirology.* 2015; 12:18.doi: 10.1186/s12977-015-0148-6 [PubMed: 25808207]
41. Santoro MM, Perno CF. HIV-1 Genetic Variability and Clinical Implications. *ISRN Microbiol.* 2013; 2013:481314.doi: 10.1155/2013/481314 [PubMed: 23844315]
42. Hemelaar J, Gouws E, Ghys PD, Osmanov S, Isolation W-UNfH. Characterisation, Global trends in molecular epidemiology of HIV-1 during 2000–2007. *AIDS.* 2011; 25(5):679–689. DOI: 10.1097/QAD.0b013e328342ff93 [PubMed: 21297424]
43. Chomont N, El-Far M, Ancuta P, Trautmann L, Procopio FA, Yassine-Diab B, Boucher G, Boulassel MR, Ghattas G, Brenchley JM, Schacker TW, Hill BJ, Douek DC, Routy JP, Haddad EK, Sekaly RP. HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nat Med.* 2009; 15(8):893–900. DOI: 10.1038/nm.1972 [PubMed: 19543283]
44. Soriano-Sarabia N, Archin NM, Bateson R, Dahl NP, Crooks AM, Kuruc JD, Garrido C, Margolis DM. Peripheral Vgamma9Vdelta2 T Cells Are a Novel Reservoir of Latent HIV Infection. *PLoS Pathog.* 2015; 11(10):e1005201.doi: 10.1371/journal.ppat.1005201 [PubMed: 26473478]
45. Bruner KM, Murray AJ, Pollack RA, Soliman MG, Laskey SB, Capoferri AA, Lai J, Strain MC, Lada SM, Hoh R, Ho YC, Richman DD, Deeks SG, Siliciano JD, Siliciano RF. Defective proviruses rapidly accumulate during acute HIV-1 infection. *Nat Med.* 2016; 22(9):1043–1049. DOI: 10.1038/nm.4156 [PubMed: 27500724]
46. Chun TW, Stuyver L, Mizell SB, Ehler LA, Mican JA, Baseler M, Lloyd AL, Nowak MA, Fauci AS. Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. *Proc Natl Acad Sci U S A.* 1997; 94(24):13193–13197. [PubMed: 9371822]
47. Finzi D, Hermankova M, Pierson T, Carruth LM, Buck C, Chaisson RE, Quinn TC, Chadwick K, Margolick J, Brookmeyer R, Gallant J, Markowitz M, Ho DD, Richman DD, Siliciano RF. Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. *Science.* 1997; 278(5341):1295–1300. [PubMed: 9360927]
48. Saez-Cirion A, Bacchus C, Hocqueloux L, Avettand-Fenoel V, Girault I, Lecuroux C, Potard V, Versmisse P, Melard A, Prazuck T, Descours B, Guernon J, Viard JP, Boufassa F, Lambotte O, Goujard C, Meyer L, Costagliola D, Venet A, Pancino G, Autran B, Rouzioux C, Group AVS. Post-treatment HIV-1 controllers with a long-term virological remission after the interruption of early initiated antiretroviral therapy ANRS VISCONTI Study. *PLoS Pathog.* 2013; 9(3):e1003211.doi: 10.1371/journal.ppat.1003211 [PubMed: 23516360]
49. Persaud D, Gay H, Ziemniak C, Chen YH, Piatak M Jr, Chun TW, Strain M, Richman D, Luzuriaga K. Absence of detectable HIV-1 viremia after treatment cessation in an infant. *N Engl J Med.* 2013; 369(19):1828–1835. DOI: 10.1056/NEJMoa1302976 [PubMed: 24152233]
50. Hocqueloux L, Prazuck T, Avettand-Fenoel V, Lefeuvre A, Cardon B, Viard JP, Rouzioux C. Long-term immunovirologic control following antiretroviral therapy interruption in patients treated at the time of primary HIV-1 infection. *AIDS.* 2010; 24(10):1598–1601. [PubMed: 20549847]
51. Hill AL, Rosenbloom DI, Goldstein E, Hanhauser E, Kuritzkes DR, Siliciano RF, Henrich TJ. Real-Time Predictions of Reservoir Size and Rebound Time during Antiretroviral Therapy Interruption Trials for HIV. *PLoS Pathog.* 2016; 12(4):e1005535.doi: 10.1371/journal.ppat.1005535 [PubMed: 27119536]
52. Chun TW, Nickle DC, Justement JS, Meyers JH, Roby G, Hallahan CW, Kottlil S, Moir S, Mican JM, Mullins JJ, Ward DJ, Kovacs JA, Mannon PJ, Fauci AS. Persistence of HIV in gut-associated

- lymphoid tissue despite long-term antiretroviral therapy. *J Infect Dis.* 2008; 197(5):714–720. DOI: 10.1086/527324 [PubMed: 18260759]
53. Hatano H, Somsouk M, Sinclair E, Harvill K, Gilman L, Cohen M, Hoh R, Hunt PW, Martin JN, Wong JK, Deeks SG, Yukl SA. Comparison of HIV DNA and RNA in gut-associated lymphoid tissue of HIV-infected controllers and noncontrollers. *AIDS.* 2013; 27(14):2255–2260. DOI: 10.1097/QAD.0b013e328362692f [PubMed: 24157906]
 54. Yukl SA, Gianella S, Sinclair E, Epling L, Li Q, Duan L, Choi AL, Girling V, Ho T, Li P, Fujimoto K, Lampiris H, Hare CB, Pandori M, Haase AT, Gunthard HF, Fischer M, Shergill AK, McQuaid K, Havlir DV, Wong JK. Differences in HIV burden and immune activation within the gut of HIV-positive patients receiving suppressive antiretroviral therapy. *J Infect Dis.* 2010; 202(10):1553–1561. DOI: 10.1086/656722 [PubMed: 20939732]
 55. Bruner KM, Hosmane NN, Siliciano RF. Towards an HIV-1 cure: measuring the latent reservoir. *Trends Microbiol.* 2015; 23(4):192–203. DOI: 10.1016/j.tim.2015.01.013 [PubMed: 25747663]
 56. Eriksson S, Graf EH, Dahl V, Strain MC, Yukl SA, Lysenko ES, Bosch RJ, Lai J, Chioma S, Emad F, Abdel-Mohsen M, Hoh R, Hecht F, Hunt P, Somsouk M, Wong J, Johnston R, Siliciano RF, Richman DD, O'Doherty U, Palmer S, Deeks SG, Siliciano JD. Comparative analysis of measures of viral reservoirs in HIV-1 eradication studies. *PLoS Pathog.* 2013; 9(2):e1003174.doi: 10.1371/journal.ppat.1003174 [PubMed: 23459007]
 57. Siliciano JD, Siliciano RF. Assays to Measure Latency, Reservoirs, and Reactivation. *Curr Top Microbiol Immunol.* 2017; doi: 10.1007/82_2017_75
 58. Allweiss L, Dandri M. The Role of cccDNA in HBV Maintenance. *Viruses.* 2017; 9(6)doi: 10.3390/v9060156
 59. Bloom K, Ely A, Mussolino C, Cathomen T, Arbuthnot P. Inactivation of hepatitis B virus replication in cultured cells and in vivo with engineered transcription activator-like effector nucleases. *Mol Ther.* 2013; 21(10):1889–1897. DOI: 10.1038/mt.2013.170 [PubMed: 23883864]
 60. Cradick TJ, Keck K, Bradshaw S, Jamieson AC, McCaffrey AP. Zinc-finger nucleases as a novel therapeutic strategy for targeting hepatitis B virus DNAs. *Mol Ther.* 2010; 18(5):947–954. DOI: 10.1038/mt.2010.20 [PubMed: 20160705]
 61. Dong C, Qu L, Wang H, Wei L, Dong Y, Xiong S. Targeting hepatitis B virus cccDNA by CRISPR/Cas9 nuclease efficiently inhibits viral replication. *Antiviral Res.* 2015; 118:110–117. DOI: 10.1016/j.antiviral.2015.03.015 [PubMed: 25843425]
 62. Ely A, Moyo B, Arbuthnot P. Progress With Developing Use of Gene Editing To Cure Chronic Infection With Hepatitis B Virus. *Mol Ther.* 2016; 24(4):671–677. DOI: 10.1038/mt.2016.43 [PubMed: 26916283]
 63. Kennedy EM, Bassit LC, Mueller H, Kornepati AV, Bogerd HP, Nie T, Chatterjee P, Javanbakht H, Schinazi RF, Cullen BR. Suppression of hepatitis B virus DNA accumulation in chronically infected cells using a bacterial CRISPR/Cas RNA-guided DNA endonuclease. *Virology.* 2015; 476:196–205. DOI: 10.1016/j.virol.2014.12.001 [PubMed: 25553515]
 64. Aubert M, Madden EA, Loprieno M, DeSilva Felixge HS, Stensland L, Huang ML, Greninger AL, Roychoudhury P, Niyonzima N, Nguyen T, Magaret A, Galletto R, Stone D, Jerome KR. In vivo disruption of latent HSV by designer endonuclease therapy. *JCI Insight.* 2016; 1(14)doi: 10.1172/jci.insight.88468
 65. Scott T, Moyo B, Nicholson S, Maepa MB, Watashi K, Ely A, Weinberg MS, Arbuthnot P. ssAAVs containing cassettes encoding SaCas9 and guides targeting hepatitis B virus inactivate replication of the virus in cultured cells. *Sci Rep.* 2017; 7(1):7401.doi: 10.1038/s41598-017-07642-6 [PubMed: 28785016]
 66. Seeger C, Sohn JA. Complete Spectrum of CRISPR/Cas9-induced Mutations on HBV cccDNA. *Mol Ther.* 2016; 24(7):1258–1266. DOI: 10.1038/mt.2016.94 [PubMed: 27203444]
 67. Weber ND, Stone D, Sedlak RH, De Silva Felixge HS, Roychoudhury P, Schiffer JT, Aubert M, Jerome KR. AAV-mediated delivery of zinc finger nucleases targeting hepatitis B virus inhibits active replication. *PLoS One.* 2014; 9(5):e97579.doi: 10.1371/journal.pone.0097579 [PubMed: 24827459]
 68. Locarnini S, Zoulim F. Molecular genetics of HBV infection. *Antivir Ther.* 2010; 15(Suppl 3):3–14. DOI: 10.3851/IMP1619

69. Menendez-Arias L. Mutation rates and intrinsic fidelity of retroviral reverse transcriptases. *Viruses*. 2009; 1(3):1137–1165. DOI: 10.3390/v1031137 [PubMed: 21994586]
70. Gao S, Duan ZP, Coffin CS. Clinical relevance of hepatitis B virus variants. *World J Hepatol*. 2015; 7(8):1086–1096. DOI: 10.4254/wjh.v7.i8.1086 [PubMed: 26052397]
71. Li S, Wang Z, Li Y, Ding G. Adaptive evolution of proteins in hepatitis B virus during divergence of genotypes. *Sci Rep*. 2017; 7(1):1990.doi: 10.1038/s41598-017-02012-8 [PubMed: 28512348]
72. Margeridon-Thermet S, Shafer RW. Comparison of the Mechanisms of Drug Resistance among HIV, Hepatitis B, and Hepatitis C. *Viruses*. 2010; 2(12):2696–2739. DOI: 10.3390/v2122696 [PubMed: 21243082]
73. Schiffer JT, Swan DA, Stone D, Jerome KR. Predictors of hepatitis B cure using gene therapy to deliver DNA cleavage enzymes: a mathematical modeling approach. *PLoS computational biology*. 2013; 9(7):e1003131.doi: 10.1371/journal.pcbi.1003131 [PubMed: 23861664]
74. Li X, Zhao J, Yuan Q, Xia N. Detection of HBV Covalently Closed Circular DNA. *Viruses*. 2017; 9(6)doi: 10.3390/v9060139
75. Berrington WR, Jerome KR, Cook L, Wald A, Corey L, Casper C. Clinical correlates of herpes simplex virus viremia among hospitalized adults. *Clin Infect Dis*. 2009; 49(9):1295–1301. DOI: 10.1086/606053 [PubMed: 19807272]
76. Brown EL, Gardella C, Malm G, Prober CG, Forsgren M, Krantz EM, Arvin AM, Yasukawa LL, Mohan K, Brown Z, Corey L, Wald A. Effect of maternal herpes simplex virus (HSV) serostatus and HSV type on risk of neonatal herpes. *Acta Obstet Gynecol Scand*. 2007; 86(5):523–529. DOI: 10.1080/00016340601151949 [PubMed: 17464578]
77. Johnston C, Corey L. Current Concepts for Genital Herpes Simplex Virus Infection: Diagnostics and Pathogenesis of Genital Tract Shedding. *Clin Microbiol Rev*. 2016; 29(1):149–161. DOI: 10.1128/CMR.00043-15 [PubMed: 26561565]
78. Schiffer JT, Corey L. New concepts in understanding genital herpes. *Curr Infect Dis Rep*. 2009; 11(6):457–464. [PubMed: 19857385]
79. Xu F, Sternberg MR, Kottiri BJ, McQuillan GM, Lee FK, Nahmias AJ, Berman SM, Markowitz LE. Trends in herpes simplex virus type 1 and type 2 seroprevalence in the United States. *JAMA*. 2006; 296(8):964–973. DOI: 10.1001/jama.296.8.964 [PubMed: 16926356]
80. Corey L, Wald A, Patel R, Sacks SL, Tyring SK, Warren T, Douglas JM Jr, Paavonen J, Morrow RA, Beutner KR, Stratchounsky LS, Mertz G, Keene ON, Watson HA, Tait D, Vargas-Cortes M, Valacyclovir HSV-TSG. Once-daily valacyclovir to reduce the risk of transmission of genital herpes. *N Engl J Med*. 2004; 350(1):11–20. DOI: 10.1056/NEJMoa035144 [PubMed: 14702423]
81. Wang K, Lau TY, Morales M, Mont EK, Straus SE. Laser-capture microdissection: refining estimates of the quantity and distribution of latent herpes simplex virus 1 and varicella-zoster virus DNA in human trigeminal Ganglia at the single-cell level. *J Virol*. 2005; 79(22):14079–87. DOI: 10.1128/JVI.79.22.14079-14087.2005 [PubMed: 16254342]
82. Wang K, Mahalingam G, Hoover SE, Mont EK, Holland SM, Cohen JI, Straus SE. Diverse herpes simplex virus type 1 thymidine kinase mutants in individual human neurons and Ganglia. *J Virol*. 2007; 81(13):6817–6826. DOI: 10.1128/JVI.00166-07 [PubMed: 17459924]
83. Dolan A, Jamieson FE, Cunningham C, Barnett BC, McGeoch DJ. The genome sequence of herpes simplex virus type 2. *J Virol*. 1998; 72(3):2010–2021. [PubMed: 9499055]
84. McGeoch DJ, Dalrymple MA, Davison AJ, Dolan A, Frame MC, McNab D, Perry LJ, Scott JE, Taylor P. The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. *J Gen Virol*. 1988; 69(Pt 7):1531–1574. DOI: 10.1099/0022-1317-69-7-1531 [PubMed: 2839594]
85. Aubert M, Boyle NM, Stone D, Stensland L, Huang ML, Magaret AS, Galetto R, Rawlings DJ, Scharenberg AM, Jerome KR. In vitro Inactivation of Latent HSV by Targeted Mutagenesis Using an HSV-specific Homing Endonuclease. *Mol Ther Nucleic Acids*. 2014; 3:e146.doi: 10.1038/mtna.2013.75 [PubMed: 24496438]
86. Roizman, B., Knipe, D., Whitley, R. *Fields Virology*. 6. Lippincott Williams and Wilkins; Philadelphia, PA: 2013. Herpes Simplex Viruses; p. 1823-1897.
87. Baker RO, Hall JD. Impaired mismatch extension by a herpes simplex DNA polymerase mutant with an editing nuclease defect. *J Biol Chem*. 1998; 273(37):24075–24082. [PubMed: 9727026]

88. Norberg P, Bergstrom T, Rekabdar E, Lindh M, Liljeqvist JA. Phylogenetic analysis of clinical herpes simplex virus type 1 isolates identified three genetic groups and recombinant viruses. *J Virol.* 2004; 78(19):10755–10764. DOI: 10.1128/JVI.78.19.10755-10764.2004 [PubMed: 15367642]
89. Koelle DM, Norberg P, Fitzgibbon MP, Russell RM, Greninger AL, Huang ML, Stensland L, Jing L, Magaret AS, Diem K, Selke S, Xie H, Celum C, Lingappa JR, Jerome KR, Wald A, Johnston C. Worldwide circulation of HSV-2 $\times$ HSV-1 recombinant strains. *Sci Rep.* 2017; 7:44084.doi: 10.1038/srep44084 [PubMed: 28287142]
90. Hoshino Y, Pesnicak L, Cohen JI, Straus SE. Rates of reactivation of latent herpes simplex virus from mouse trigeminal ganglia ex vivo correlate directly with viral load and inversely with number of infiltrating CD8+ T cells. *J Virol.* 2007; 81(15):8157–8164. DOI: 10.1128/JVI.00474-07 [PubMed: 17522198]
91. Schneider TD, Stormo GD, Gold L, Ehrenfeucht A. Information content of binding sites on nucleotide sequences. *J Mol Biol.* 1986; 188(3):415–431. [PubMed: 3525846]

Figure 2. HIV sequence diversity

a. Phylogenetic trees representing HIV diversity were made using the Los Alamos National Laboratory (LANL) 2016 HIV complete genome alignments for all HIV sequences (left) or M group A–K genomes without recombinants (right). Trees were generated with Geneious using the Neighbor-Joining method and a Jukes-Cantor genetic distance model. Scale represents nucleotide substitutions per site. **b.** nucleotide diversity plots for HIV complete genomes, LTR, *gag*, *pol* and *env*, generated using data from the LANL 2016 M group

alignment. Nucleotide conservation is represented in terms of information content (range = 0–2 bits) calculated with a moving window width of 50 nt⁹¹.