



# Inactivation of Latent HIV-1 Proviral DNA Using Clustered Regularly Interspaced Short Palindromic Repeats/Cas9 Treatment and the Assessment of Off-Target Effects

Yufan Xu, Xiaorong Peng, Yanghao Zheng, Changzhong Jin, Xiangyun Lu, Dating Han, Haijing Fu, Chaoyu Chen and Nanping Wu\*

#### **OPEN ACCESS**

#### Edited by:

Nejat Duzgunes, University of the Pacific, United States

#### Reviewed by:

Vijai Singh, Indrashil University, India Ben Berkhout, University of Amsterdam, Netherlands

> \*Correspondence: Nanping Wu flwnp@zju.edu.cn

#### Specialty section:

This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: 13 November 2020 Accepted: 28 April 2021 Published: 26 May 2021

#### Citation:

Xu Y, Peng X, Zheng Y, Jin C, Lu X, Han D, Fu H, Chen C and Wu N (2021) Inactivation of Latent HIV-1 Proviral DNA Using Clustered Regularly Interspaced Short Palindromic Repeats/Cas9 Treatment and the Assessment of Off-Target Effects. Front. Microbiol. 12:629153. doi: 10.3389/fmicb.2021.629153 State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, National Clinical Research Center for Infectious Diseases, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, The First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, China

Viral DNA integrated in host cells is a major barrier to completely curing HIV-1. However, genome editing using the recently developed technique of clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 has the potential to eradicate HIV-1. The present study aimed to use a lentiviral vector-based CRISPR/Cas9 system combined with dual-small/single guide RNAs (sgRNAs) to attack HIV-1 DNA in the latency reactivation model J-Lat 10.6 cell line and to assess off-target effects using whole-genome sequencing (WGS). We designed 12 sgRNAs targeting HIV-1 DNA, and selected high-efficiency sgRNAs for further pairwise combinations after a preliminary evaluation of the editing efficiency. Three combinations of dual-sgRNAs/Cas9 with high editing efficiency were screened successfully from multiple combinations. Among these combinations, the incidences of insertions and deletions in the sgRNA-targeted regions reached 76% and above, and no credible off-target sites were detected using WGS. The results provided comprehensive basic experimental evidence and methodological recommendations for future personalized HIV-1 treatment using CRISPR/Cas9 genome editing technology.

Keywords: CRISPR/Cas9, HIV-1, dual-sgRNAs, off-target effect, genome editing

# INTRODUCTION

By the end of 2019, worldwide, 38 million people were estimated to be living with HIV. According to a WHO report, there were still about 1.7 million new HIV infections in 2019, and the HIV epidemic poses a major threat to global human health. A variety of drugs, such as reverse transcriptase inhibitors, protease inhibitors, and integrase strand transfer inhibitors, which have been widely used clinically, can play an antiviral role at different stages of the

1

HIV-1 life cycle. The majority of patients with HIV/AIDS treated with these antiviral drugs were able to achieve the virological inhibition criterion that the viral load in plasma was below the detection limit (Laskey and Siliciano, 2014). However, current antiviral therapy cannot completely remove the virus from the human body because of the stable integration of proviral DNA into the cellular genome. Genome editing technology might be one of the solutions to remove proviral DNA from HIV-1 infected cells.

The clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 system, derived from the immune system of prokaryotes, which could resist the invasion of exogenous genetic material, such as phages and viruses, has become the third generation of genome editing technology (Jinek et al., 2012; Cong et al., 2013). One of the most widely used applications of the CRISPR/Cas9 system is to knockout a target gene. Its principle is that small/single guide RNA (sgRNA) targeted to the gene of interest guides Cas9 protein-mediated cleavage of an exon in the gene to break the double strand, which prompts the cells to commence non-homologous end-joining (NHEJ) repair (Mali et al., 2013). In this process, several bases are randomly added or deleted between the breakpoints. If a frame shift mutation occurs, the target gene would be inactivated.

Multiple strategies, which could be divided into two categories according to whether they act on the HIV DNA or host factors, have been proposed for HIV/AIDS treatment using the CRISPR/Cas9 system. The first category included the design of sgRNAs based on the viral genome that prompt the sgRNA-Cas9 system to attack HIV-1 DNA (Panfil et al., 2018; Das et al., 2019), and also included the strategy of "shock and kill," which targets the HIV long-terminal repeat sequence to activate latent HIV via the CRISPR/dCas9 activator complex (Ji et al., 2016; Limsirichai et al., 2016; Saayman et al., 2016). The latter mainly targets the HIV co-receptor genes such as C-X-C motif chemokine receptor 4 (CXCR4) and C-C motif chemokine receptor 5 (CCR5; Liu et al., 2017, 2018; Xu et al., 2017; Yu et al., 2018). However, previous studies have voiced concerns regarding targeting to activate endogenous viral restriction factors to inhibit viral replication using CRISPR/ dCas9-synergistic activation mediators, such as apolipoprotein В mRNA editing enzyme, catalytic polypeptide-like (APOBEC)3G/3B (Bogerd et al., 2015), and bone marrow stromal cell antigen 2 (BST-2; Zhang et al., 2019).

Although HIV-1 replication is inhibited by CRISPR/Cas9induced mutations, viral escape could occur in sgRNA protected cells (Wang et al., 2016a). Another study constructed SpuT1 T cells stably expressing the Cas9 protein and dual sgRNAs, and observed the protective effect from HIV-1 LAI virus infection. The authors found that dual-sgRNAs CRISPR/Cas9 treatment could delay viral escape, block virus replication more efficiently, and even had the potential to extinguish all proviral DNAs in infected T cell cultures (Wang et al., 2016b). However, the efficiency and off-target effects of dual-sgRNAs combinations require further confirmatory studies.

Our study focused on attacking HIV-1 DNA using CRISPR therapy, for which we selected the latency reactivation model

J-Lat 10.6 cell line, which contains a full-length HIV-1 genome in which the Nef gene is replaced by the green fluorescent protein (GFP) gene and has a frame-shifted, non-functional Env gene. J-Lat 10.6 cells are suitable for the *in vitro* assessment of CRISPR/Cas9-mediated targeted mutagenesis of virus genes that encode vital proteins. Using the J-Lat 10.6 HIV-1 genomic DNA as a model, we designed and verified the mutation efficiency of several dual-sgRNA combinations. In addition, to study any off-target effects of the CRISPR gene-editing system, whole-genome sequencing (WGS) was performed.

# MATERIALS AND METHODS

#### Cells

Human embryonic kidney (HEK) 293 T cells and J-Lat 10.6 cells were grown in Dulbecco's modified Eagle's medium (DMEM, Gibco, Grand Island, NY, United States) and Roswell Park Memorial Institute (RPMI)-1640 medium (Gibco), respectively. Both media contained an additional 10% fetal bovine serum (Gibco), penicillin-streptomycin (Gibco), and non-essential amino acids (Gibco). The cells were cultured at  $37^{\circ}$ C in a humidified atmosphere of 5% CO<sub>2</sub>.

### **Plasmids**

The lentiviral vector lentiGuide-Puro (plasmid #52963, Addgene, Watertown, MA, United States) and lentiCas9-Blast (Addgene; plasmid #52962) were gifts from Feng Zhang. Control sgRNA encoding oligonucleotides and those encoding sgRNAs targeting HIV-1 were ligated into the LentiGuide-Puro vector *via* its BsmB1 site, as previously described (Sanjana et al., 2014).

# Lentiviral Vector Production and Transduction

The lentiviral vector was produced and titrated as previously described (ter Brake et al., 2006). Briefly, HEK 293 T cells were transfected with a packaging plasmid (psPAX2) and the lentiviral vector plasmid (pMD2.G) using Lipofectamine 2000 (Invitrogen, Waltham, MA, United States). At 6 h after transfection, the medium was replaced and the cell culture was continued for 2 days. Then, 0.45-µm filter membranes (Millipore, Billerica, MA, United States) were used to filter the lentivirus supernatants, which were divided into aliquots and placed at  $-80^{\circ}$ C. J-Lat 10.6 cells (3  $\times$  10<sup>5</sup> cells) in culture medium were transduced with LentiCas9-Blast virus particles (30 ng of viral capsid protein p24) and cultured in the presence of blasticidin (Thermo Fisher, Waltham, MA, United States) at 4 ng/ml for 10 days to select J-Lat-cas9 cells. Subsequently, these cells  $(3 \times 10^5 \text{ cells})$  were transduced with one or two LentiGuide-Puro virus particles (30 ng of viral capsid protein p24) and cultured in the presence of puromycin (Thermo Fisher) at 0.5 ng/ml for 10 days to produce J-Lat 10.6 cells expressing Cas9 and the sgRNAs. The dosage of two LentiGuide-Puro virus particles was prepared according to a 1:1 ratio of viral capsid protein p24 concentration.

## **Genome Editing Analysis**

The cellular genomic DNA was extracted using a MiniBEST Universal Genomic DNA Extraction Kit (Takara, Shiga, Japan) according to the manufacturer's instructions. PrimeSTAR® Max DNA Polymerase (Takara) and HIV-1 specific primers (**Supplementary Table S1**) were used to amplify the genomic DNA using PCR. The PCR products were purified using a MiniBEST Agarose Gel DNA Extraction Kit (Takara) and were detected for T7 Endonuclease I (T7EI)-based mutations using an EnGen® Mutation Detection Kit (NEB, Ipswich, MA, United States) following the manufacturer's protocol. The purified PCR products were cloned into a TA cloning vector and sequenced for subsequent analysis.

## Activation of the Expression of Latent HIV-1

J-Lat 10.6 cells expressing Cas9 and sgRNAs were treated with 0.1  $\mu$ M of phorbol 12-myristate 13-acetate (PMA, Sigma, St. Louis, MO, United States). After 16 h of treatment, the cells were harvested and suspended in phosphate-buffered saline (PBS, HyClone, Logan, UT, United States), and GFP expression was measured using flow cytometry.

# Whole-Genome Sequencing for Off-Target Assays

Whole-genome sequencing was performed using the cellular genomic DNA to assess the potential off-target effects. GENEWIZ (South Plainfield, NJ, United States) performed the WGS and analyzed the data.

# RESULTS

### Design of the sgRNA and the T7EI Assay

Based on the latent HIV-1 genome sequence of the J-Lat 10.6 cell line, a total of 12 sgRNAs with a high on-target

score, which targeted different protein-coding sequences, such as gag, pol, tat, and env genes, respectively, were designed using the online CRISPR design tool (Hsu et al., 2013; Figure 1; Table 1). Of these 12 target sites, five sgRNAs were located in the gag gene, four sgRNAs were in the pol gene, one sgRNA was in the tat gene, and two were in the env gene. Then, 12 cell lines stably expressing Cas9 and one sgRNA were constructed using lentiviral vectors. The mutation frequency of NHEJ caused by the CRISPR/Cas9 nuclease cleavage was first detected using the T7E1 assay to select effective sgRNAs. The annealed PCR product could be cleaved with T7EI. T7EI, a structure-specific enzyme, effectively identifies gene mismatches greater than 1 base. When a mismatch occurs, the two strands of the DNA molecule are cut into smaller fragments. Analyzing fragment maps could evaluate the efficiency of gene editing. The estimated mutation frequency ranged from 9 to 54% (Figure 2).

# Detecting the Efficiency of Combined sgRNAs Using TA Cloning

According to the results of the T7EI assay, sgRNAs with relatively high efficiency were selected and nine dual-sgRNA combinations, which included those targeting the same protein-coding gene and those targeting different protein-coding genes, were constructed in the lentiviral vector. The sgRNA target sites in these J-Lat 10.6 cells stably expressing Cas9 dual-guide RNAs (gRNAs) were sequenced to analyze the types and frequency of mutations. Nucleotide sequences analysis indicated that almost all of the mutations were clustered near the Cas9 cleavage site, at 3–4 nt upstream of the protospacer adjacent motif (PAM) sequence, confirming that NHEJ repair caused the mutations. The combinations with a high proportion of insertions and deletions of the HIV-1 genome were sg-Gag1-J + sg-Gag2-J (Insertions: 44%; Deletions: 32%), sg-Gag2-J + sg-Pol2-J (Insertions: 19%; Deletions: 63%), and sg-Pol2-J + sg-Tat1-J



#### TABLE 1 | Selected sgRNAs targeting HIV-1 DNA.

Name	Position/strand <sup>a</sup>	Guide sequence + PAM <sup>b</sup>	Specificity score <sup>c</sup>	Genomic region
sg-Gag1-J	1,185-1,207/R	CCTGGATGTTCTGCACTATA <b>GGG</b>	89	Gag (p24)
sg-Gag2-J	1,345-1,367/F	ACCATGCTAAACACAGTGGG <b>GGG</b>	74	Gag (p24)
sg-Gag3-J	1,664-1,686/R	CCGGTCTACATAGTCTCTAAAGG	88	Gag (p24)
sg-Gag4-J	1804-1826/R	ATTTCTTCTAGTGTAGCCGC <b>TGG</b>	92	Gag (p24)
sg-Gag5-J	1857-1879/F	CGGCCATAAGGCAAGAGTTT <b>TGG</b>	71	Gag (p24)
sg-Pol1-J	2,277-2,299/R	CCTATCTTTATTGTGACGAG <b>GGG</b>	93	Pol (Protease)
sg-Pol2-J	2,448-2,470/F	ATCTGTGGACATAAAGCTAT <b>AGG</b>	72	Pol (Protease)
sg-Pol3-J	3,018-3,040/R	ATGCTACTTTGGAATATTGC <b>TGG</b>	72	Pol (RT)
sg-Pol4-J	3,255-3,277/R	GGCTGTACTGTCCATTTATCAGG	88	Pol (RT)
sg-Tat1-J	5,967-5,989/R	TCTCCGCTTCTTCCTGCCATAGG	63	Tat/Rev
sg-Env1-J	6,981-7,003/R	AGCAGTTGAGTTGATACTAC <b>TGG</b>	87	Env (gp120)
sg-Env2-J	7,794-7,816/F	GGAGCAGCAGGAAGCACTAT <b>GGG</b>	74	Env (gp41)
sg-CTRL	-	GGAGACGGGATACCGTCTCT		-

<sup>a</sup>The position corresponds to the reference sequence HXB2. F, forward; R, reverse.

<sup>b</sup>The PAM sequence is in bold.

°The specificity score was calculated using the CRISPR design web tool from crispor.tefor.net.



FIGURE 2 | The efficiency of clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 treatment-induced cleavage of HIV-1 DNA. (A–E) Mutational outcomes of 12 sgRNA/Cas9 treatments were determined using a T7 endonuclease I (T7EI) assay. Cas9, cells only expressing the Cas9 protein; CTRL, cells before transfection.



(Insertions: 54%; Deletions: 25%), respectively (**Figure 3**). The mutated sequences of the sgRNA-targeted regions in the three most efficient dual-sgRNA combinations are displayed in **Figure 4**. Mutation sequences of other combinations are shown in **Supplementary Figures S1**, **S2**. The incidence of nucleotide substitutions (NSs) was low, only being detected in four out of nine combinations, and at less than 10% (**Figures 3C,E,H,I**). Some interesting results were found in the sequence analysis of HIV-1 mutation sites. In the combination sg-Gag1-J + sg-Gag2-J, we found deletion of the entire viral DNA between the two sgRNA-targeted sites in two cases, with 169 and 177 bp being deleted, respectively (**Supplementary Figure S3**). A chromosomal inversion occurred between two sgRNA-targeted sites in the combination sg-Gag2-J + sg-Gag5-J (**Supplementary Figure S4**).

### **Blocking of Latent HIV-1 Expression**

In addition, we performed a test to activate latent HIV-1 in the J-Lat 10.6 cells stably expressing Cas9 and dual-sgRNAs. After 16 h of treatment with 0.1  $\mu$ M PMA, GFP expression was monitored using flow cytometry. Except for the sg-Pol1-J + sg-Pol3-J combination, for which the inhibition effect was not obvious because of the low incidence of insertions and

deletions in the HIV-1 genome (**Figure 3G**), the GFP expression of the other eight combinations was significantly inhibited, among which the sg-Pol2-J + sg-Tat1-J combination showed the most significant inhibitory effect (**Figure 5**).

# Assessment of Off-Target Effects Using WGS

We selected three combinations with relatively high editing efficiency (sg-Gag1-J + sg-Gag2-J, sg-Gag2-J + sg-Pol2-J, and sg-Pol2-J + sg-Tat1-J), with the untreated J-Lat 10.6 cell as controls, to perform WGS (30x coverage) and detect single nucleotide variants (SNVs) and insertions and deletions (Indels). First, the SNV/Indel of a single treatment group was compared with the control to identify unique the mutation sites of the treatment group, and the upstream and downstream 100 bp sequences of these sites were obtained from the reference genome (GRCh37/hg19) as candidate potential off-target regions. Next, we aligned the sgRNA sequence with the candidate potential off-target regions, and filtered out the regions with more than six mismatched bases and those not containing the PAM (NRG) sequence. Finally, the potential off-target regions of three treatment groups after screening were identified, and all of these regions were located in intergenic or intronic

Ref WT WT WT -390 -71 -18 -8 -6 +1 +1 +2 +3 +5 +6 -160 -175

> Ref WT WT -394 -74 -20 -20 -11 -9 -7 -6 -6 +4 +76

Ref WT -9 -4 -1 +1 +2 +3 +5 +24

sg-Gag1-J target region (WT: 4/17)		sg-Gag2-J target region (WT: 4/17)
118         Part         Data Guide         124           CAGGTGAGCLANATTACCTA	Ref WT WT -280 -10 +1 +1 +1 +1 +1 +1 +1 +1 +2 +11 -169 -177	1313         2014 Guide         PAM         1388           NAGGAGCCCCCCACAGAGATTTAAACACCATGCTTAAACACAGTG         GGGGGGACATCAAGCAGCCATGCAAATGTTAAAAGAG           NAGGAGCCACCCCCACAGAGATTTAAACCACTGCTAAACACAGTG         GGGGGGACATCAAGCAGCCATGCAAATGTTAAAAGAG           NAGGAGCCACCCCCACAGAGATTTAAACCACTGCTAAACACAGTG         GGGGGGACATCAAGCAGCCATGCAAATGTTAAAAGAG           NAGGAGCCACCCCCACAGAGATTTAAACCATGCTAAACACAGTG         GGGGGGACATCAAGCAGCCATGCAAATGTTAAAAGAG           NAGGAGCCACCCCCACAGGATTTAAACCACTGCTAAACACAGTG         GGGGGGACATCAAGCAGCAGCAATGTAAAGAG           NAGGAGCCACCCCCACAGGATTTAAACCACTGCTAAACACAGTG         GGGGGGACATCAAGCAGCAGCAATGTAAAGAGAAAGAGAGCATGCAAGCAGCATGCAAGCAGCATGCAAGCATTAAAGAGG           NAGGAGCCACCCCCACAGGATTTAAACCACTGCTAAACCACAG         GGGGGGGACATCAAGCAGCAGCAGCATGCAAGCATTAAAGGAGAGAGCAGCATGCAAGCCATGCAATGTTAAAAGAGG           NAGGAGCCACCCCCACAAGGATTTAAACCATGCTAAACCACAGG         GGGGGGGACATCCAAGCAGCATGCAATGTTAAAGAGGAGAGCATGTAAGCAGCATGCAAGCCATGCAAGGCAGCATGGAAGCATGTAAGAGAGGAGGAGCATGCAAGGAGCATGCAAGGAGCATGCAAGGCATGTAAGAAGGAGGAGCACCCCCCAAGGATTTAAACCACTGCTAAACCACAGGCAGCATGCAAGGCAATGCAAGGCATGCAAGGAGCATGCAAGGAGCATGCAAGGAGCATGCAAGGAGCATGCAAGGAGGAGGAGCATGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGCCATGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA
B sg-Gag2-J + sg-Pol2-J so-Gad2-J target region (WT: 3/17)		sa-Pol2-J taraet region (WT: 3/15)
135     201 Outo     PM     139       CCACCCCACAGAGATTAAACACCATGTAAACACAG		201         201
sg-Pol2-J target region (WT: 4/16)		sg-Tat1-J target region (WT: 1/12)
2427         2001 Cuide         PAM         2402           TATCATCAGATACTCATAGAAATCTGEGACATAAAGCTATAGGTACAGTATAGTAGGACCTACACCTGTCAACA         Fef           TATCATCAGATACTCATAGAAATCTGEGACATAAAGCTATAGGTACAGTATAGTAGGACCTACACCTGTCAACA         FW           TATCATCAGATACTCATAGAAATCTGEGACATAAAGCTATAGGTACAGTATAGTAGGACCTACACCTGTCAACA         FW           TATCATCAGATACTCATAGAAATCTGEGACATAAAGCTATAGGTACAGTATATAGTAGGACCTACACCTGTCAACA         FW           TATCATCCAGATACTCATAGAAATCTGEGACATAAAGCTATAGGTACAGTATATAGTAGGACCTACACCTGTCAACA         FW           TATCATCCAGATACTCATAGAAATCTGEGACATAAAGCTATAGGTACAGTATATAGTAGGACCTACACCTGTCAACA         W           TATCATCCAGATACTCATAGGAAATCTGEGACATAAAGCT		545         PAM         20n Guide         900           ATGACAAAAGCCTTAGGCATCTCCTATG        CAGGAACAACGGAGCAAGCGACAAGAAGCCTAAGAACACCGAAGAAGCCT        CAGGAACAACGGAACACGGAAGAAGCCT         ATGACAAAAGCCTTAGGCATCTCCTATG        CAGGAACAACGGAACACGGACAAGCGACAGAAGACGCAAGAAG

**FIGURE 4** | Sequence analysis of HIV-1 mutation sites in dual-sgRNAs/Cas9-treated J-Lat 10.6 cells. Features of the mutations in the three most efficient dual-sgRNA combinations: (A) sg-Gag1-J + sg-Gag2-J. (B) sg-Gag2-J + sg-Pol2-J. (C) sg-Pol2-J + sg-Tat1-J. All sequences were aligned with the reference strain HXB2. The 20 nt guide sequence and the protospacer adjacent motif (PAM) sequence are underlined. The sequences of nucleotide substitutions, insertions, and deletions are indicated using red letters. Ref, reference sequence; -x/+x, x nt deleted/inserted; NS, nucleotide substitution; and WT, wild-type sequence.

regions (**Figure 6**). By determining the relative position between the mutation site and the 20 nt complementary region of the sgRNA in all potential off-target regions, we found no mutation sites located in the 20 nt complementary regions (the relative position between +1 and +20).

# DISCUSSION

In this study, among the 12 sgRNA sequences designed, sg-Tat1-J and sg-Env2-J refer to previously reported sequences because of their high activity in targeting HIV-1 DNA (Wang et al., 2016b), and the viral transcriptional repressive effect was significant, especially when targeting the sg-Tat1-J site. According to the results of the T7EI assay, sgRNAs with a low mutation rate were not included in the combination treatments. We selected three combinations with a high incidence of insertions and

deletions from nine dual-sgRNA combinations, and then performed WGS to detect any off-target effects. A variety of studies have confirmed the effectiveness of CRISPR/Cas9 system in eliminating HIV-1 DNA (Ebina et al., 2013; Hu et al., 2014; Liao et al., 2015; Zhu et al., 2015; Bella et al., 2018; Mefferd et al., 2018; Ophinni et al., 2018; Wang et al., 2018; Yin et al., 2018); however, few of them detected off-target effects using deep sequencing to assess the safety of the method. Targeting HIV-1 DNA was expected to have the advantage of inducing a low frequency of off-target mutations, because theoretically, an editing system targeting viral DNA would not affect the host genes. The effectiveness and safety of a treatment are equally important. Therefore, evaluation of off-target effects is essential for any gene editing therapy. Compared with detection methods involving directly amplifying the predicted off-target sites and sequencing by Sanger sequencing or next-generation sequencing, WGS is more comprehensive to assess off-target



effects because it can simultaneously detect coding and non-coding regions. Several studies have used WGS to detect off-target effects edited by the CRISPR/Cas9 system (Feng et al., 2014; Ghorbal et al., 2014; Li et al., 2018; Lyu et al., 2018); however, higher sequencing depth is required and the cost is relatively high. In the present study, we used WGS to assess off-target effects on three combined samples with high efficiency. Although many mutations were detected as potential off-target regions in the experimental groups, these mutation sites were not located within the 20 nt complementary region of the sgRNA. Therefore, these mutations are not credible off-target sites.

In the last few years, studies targeting HIV-1 co-receptor CCR5 using CRISPR-Cas9 have made significant progress, such as CCR5 disruption in induced pluripotent stem cells (Kang et al., 2015), inducing CCR5  $\Delta 32/\Delta 32$  homozygotes in CD4+ Jurkat cells and primary CD4+ cells (Qi et al., 2018), and even CCR5-Edited stem cells in a patient with HIV-1 and acute lymphocytic leukemia (Xu et al., 2019). However, there is a limitation in editing the *CCR5* gene, in that it can only prevent CCR5-tropic HIV-1 virus infection, but not CXCR4-tropic HIV-1 virus. In addition, in terms of safety, compared with directly attacking viral DNA using CRISPR/Cas9, the short and long-term effects on normal immune function caused by inactivation of *CCR5* by CRISPR/Cas9 genome editing technology require further research.

HIV-1 can be divided into many subtypes and recombinant subtypes, and the distribution of these subtypes varies in different countries and regions (Maartens et al., 2014). Genetically diverse viral populations still exist in the same subtype and even within each infected individual (Hemelaar, 2012). Therefore, it is a challenge to design general and conserved sgRNA sequences for all subtypes of HIV-1, which are both effective and safe (Roychoudhury et al., 2018). With the development of CRISPR/Cas9 genome editing technology, targeting HIV-1 DNA could be individualized *via* sgRNA design in the future (Dampier et al., 2018). The research object in this study was a cell model of latency; therefore, we did not need to consider the genetic diversity of HIV-1. If the object of further research is primary CD4+ cells in HIV-1 infected patients, the viral DNA sequence should be amplified and sequenced using deep sequencing. Multiple individualized high score sgRNAs should be designed, which should avoid the sites with a high mutation rate as much as possible (Dampier et al., 2018). Then, the efficacy and safety of the sgRNAs combinations could be verified *in vitro*, and the most appropriate combinations for individuals would be selected for CRISPR treatment. However, how to locate all the latently infected cells present in long-lived HIV-infected individuals and how to deliver the gene editing tools to these latently-infected cells remain a challenge.

In the present study, we found deletion of entire viral DNA and a chromosomal inversion between the two sgRNAs-targeted sites. These mutations seriously disrupted the integrity of the viral genome and led to loss of viral protein function. Thus, the dual-sgRNA combinations, consisting of two efficient and safe sgRNAs located between hundreds of bases in a proteincoding gene of HIV-1, could become a desirable and popular strategy to destroy the viral genome. Nonetheless, the low incidence of fragment excision and fragment inversion was an important issue that requires further attention (Binda et al., 2020). In addition to the dual-sgRNAs CRISPR treatment strategy, according to a recent study (Dash et al., 2019), a combination therapy comprising long-acting antiviral drugs and CRISPR treatment targeting HIV-1 DNA could lead to the safe elimination of HIV-1 infection, possibly representing a future cure for HIV-1. By contrast, CRISPR treatment targeting HIV-1 DNA combined with RNA interference can lead to cross-resistance (Zhao et al., 2017). Therefore, more studies are required to confirm whether CRISPR treatment combined with other treatments is appropriate.

<ul> <li>Apple Apple Apple</li></ul>	<b>a</b>	000003000000000000000000000000000000000	De del	D. C.	<b>T</b>	Control I	Providence 1	0
OutputOOImagePTTMUGPNo	g-Gag1-J	CUTGGATGTTUTGCAUTATANGG CUTGGATGTTUTGCAUTATANGG	Position	G	Treatment Hete (T/G)	Control	Functional region	Gene RNF168
Graph of the second of the		GTGGGATGTTCTGCACTAGAAGG	-68	CA	Hete (CA/C-)	CA	intronic	PITPNM1
<ul> <li>Martin Martin Martan Martin Martin Martin Martin Martin Martin Martin Martin Mar</li></ul>	-Gao2-I	ACCATGCTAAACACAGTGGG <b>NGG</b>	Position	Reference	Treatment	Control	Functional region	Gene
PM       CUMPUTATION       PM       CUMPUTATION       DCM       DCM <td>-Ougz-v</td> <td>GACCAACTAAACACAGTGTGCAG</td> <td>+91</td> <td>TAA</td> <td>Hete (TA/T-)</td> <td>Homo (TA/TA)</td> <td>ncRNA exonic</td> <td>LINC01883</td>	-Ougz-v	GACCAACTAAACACAGTGTGCAG	+91	TAA	Hete (TA/T-)	Homo (TA/TA)	ncRNA exonic	LINC01883
No.         No. <td></td> <td>TTCTAGTTGAACACAGTGGGAGG</td> <td>+70</td> <td>CTTTTTTTTT</td> <td>Hete(CTT-/C)</td> <td>Hete(CTTT/C)</td> <td>intergenic</td> <td>LOC339902;FHIT</td>		TTCTAGTTGAACACAGTGGGAGG	+70	CTTTTTTTTT	Hete(CTT-/C)	Hete(CTTT/C)	intergenic	LOC339902;FHIT
Participation         Partiteation         Participation         Participa		ACCATGCTAAACTCTATCATCAG	+55	G	Hete(GT/G-)	Homo (GT/GT)	intronic	ARHGAP24
Part of the second se		CACATGCTATACACAGTTTTTGG	+83	ATGTATTGTAT	Hete (ATGTAT/A)	Hete (ATGTATTGTAT/A)	intergenic	SLC25A51P1;LOC102723883
Mark         Mark <td< td=""><td></td><td>TATGTCCTCAACACAGTGGGTAG</td><td>-32</td><td>CAAAAAA</td><td>Homo (C-/C)</td><td>Hete(CA/C)</td><td>intergenic</td><td>RB1CC1;NPBWR1</td></td<>		TATGTCCTCAACACAGTGGGTAG	-32	CAAAAAA	Homo (C-/C)	Hete(CA/C)	intergenic	RB1CC1;NPBWR1
Math         Math <th< td=""><td></td><td>CCCATGCTAAACAGTGCAGGGAG</td><td>-53</td><td>GAAA</td><td>Homo (GA/GA)</td><td>Hete(GA/G)</td><td>intergenic</td><td>LINC01568;LOC101928035</td></th<>		CCCATGCTAAACAGTGCAGGGAG	-53	GAAA	Homo (GA/GA)	Hete(GA/G)	intergenic	LINC01568;LOC101928035
Partial       Partin       Partial       Partial		ACAGCACTAAACACAGTGCTGGG ACTATGCTAAACACTCTATCAAG	-92 -28	AAAGAAAGAAAGAAAG CAA	Hete (A/AAAAG) Homo (C-/C-)	Hete (AAAGAAAGAAAGAAAG/AAAAG) Homo (CA/CA)	ncRNA_intronic intergenic	MIR1268A SH3BGRL:POU3F4
Gamma Distribution         Name         Name         Transmin         Constr         Procision legin         Gene           GGGGG         GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG							intergente	
Gamma of the second o					sg-Gag2-J + sg-Pol2-J			
BACCAMCTANACLOSSING         ***         TAA         Bete (77/-7)         Henn (77/73)         exp(*)         exp(*)           ACCAMCTANACLOSSING         ************************************	Gag2-J	ACCATGCTAAACACAGTGGG <b>NGG</b>	Position	Reference	Treatment	Control	Functional region	Gene
TUTUNUTUALIZALULUSUNDO         0         CTITITITI         Hete (CT-/C)         Hete (CT-//C)         Hete (CT-		GACCAACTAAACACAGTGTG <b>CAG</b>	+91	TAA	Hete (TA/T-)	Homo (TA/TA)	ncRNA_exonic	LINC01883
No.L. N		TTCTAGTTGAACACAGTGGGAGG	+70	CTTTTTTTTT	Hete (CTT-/C)	Hete (CTTT/C)	intergenic	LOC339902;FHIT
$ \begin{array}{cccc} \mbod label{eq:linear_line$		AUCATGUTAAACTCTATCATCAG	+55	G AUCUAUUCUAU	Hete (GT/G=)	HOMO (GT/GT)	intronic	AKHGAP24
NEMA       CUSTORGACTABAGETANGE       Pain       Reference       Teament       Control       Bee (TAUATABACTAMA CONTORGA       Pain ALLANABARAANA         ACCUMMARCATABAGETANGE       -37       CITT       Bet (TAUATABACTAMA CONTORGA       -37       CITT       Bet (TAUATABACTAMA CONTORGA       -37       CITT       Bet (TAUATABACTAMA CONTORGA       EPAIA       Bet (CAUABACATAMA CONTORGA       EPAIA       Bet (CAUABACATAMA CONTORGA       Bet (CAUABACATAMA BET (CAUABACATAMA BET (CAUACATAMA BET (CAUACATAMA CONTORGA       Bet (CAUABACATAMA CONTORGA       BET (CAUABACATAMA BET (CAUACATAMA BET (CAUACATAMA BET (CAUACATAMA BET (CAUACATAMA BET (CAUACATAMA BET (CAUACATAMA BET (CAUABACATAMA BET (CAUACATAMA BET (		ACTATGCTATACACAGTTTT <b>TGG</b>	-28	CA	CA/CA	Homo (C/C)	intergenic	SH3BGRL;POU3F4
Match 1996         1 <td>Pol2-I</td> <td>ATCTGTGGACATAAAGCTATNGG</td> <td>Position</td> <td>Reference</td> <td>Treatment</td> <td>Control</td> <td>Functional region</td> <td>Gene</td>	Pol2-I	ATCTGTGGACATAAAGCTATNGG	Position	Reference	Treatment	Control	Functional region	Gene
Publy         ACCUMPNENDEGRAPHAGE         2:2         CTT         Hete (TAC)         Hete (TAC)         Hete (TAC)         Hete (TAC)           ACCUMPNENDEGRAPHAGE/TACTURGE         -2:3         CDAAAAA         Hete (CAAACA)         Hete (CAAACAACA)         Hete (CAAACAACA)         Hete (CAAACAACA)         Hete (CAAACAACA)         Hete (CAAACAACA)         Hete (CAACAACA)         Hete (CAAACAACA)         Hete (CAACAACA)         Hete (CAAACAACA)         Hete (CAACAACA)         Hete (CAACAACA)         Hete (CAACAACA)         Hete (CAACAACA)         Hete (CAACAACA)         Hete (CAACACA)         Hete (CAACACA)         Hete (CAACACA)         Hete (CAACACA)         Hete (CAACACA)<	. 018-3	AGCCATAGACATAAAGCTAGCAG	-37	ТАААААААА	Hete (TAAAA/T)	Hete (TAAAAAA/T)	intronic	EPHA5
NACCONSIGNATION CONTINUE       424       ACCONSIGNATION CONTINUE       Bite (CANADAC)       Bite (CANAC)       Bite (C		CTTTCTTGACATAAAGCTCCAGG	+32	CTTT	Hete(CTT/C)	Hete(CT/C)	intronic	ARHGAP26
Pul2         CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC		ACATAAAGACATAAAGCTAG <b>GAG</b>	+24	ATGTG	Hete (ATGTGTGTGTGTGTGTGTG/A)	Hete (ATGTGTGTGTGTGTGTGTG/A)	intronic	SH3BGRL2
PMP       CONTRACTOR 4-64       CANA       Reference       Reference       Control 1002001         TICLAMACARTANACCTATION       -76       CANAAAAAAAA       Reference       Reference       COCI00320012         TICLAMACARTANACCTATION       -76       CANAAAAAAAA       Reference       Reference       COCI00320012         TICLAMACARTANACCTATION       -76       CANA       Reference       Reference       COCI00320012         TICLAMACARTANACCTATION       -76       CANA       Reference       TICLAMACARTANACCTATION       PACE       Reference       Reference       REFERENCE       REFERENCE       REFERENCE       COCI00320012         TICLAGACARTANACCTATIONE       PSE       TAA       Reference       TISTIGE       REFERENCE       RE		CTGGGTGGACATAAACTTTTGAG	-85	CAAAAAAA	Hete (CAAAAAAA/CAA)	Hete (CAAAAAAA/C)	intergenic	NPTX2;TMEM130
PM2.1         CONDANABACCTATOGC -56         CANAAAAAAAA         Hete (TANAAACAAA/C)         Hete (TANAAAAAAA)         Hete (TANAAAAAAA)         Hete (TANAAT)         Immonic         CANAAAAAAAAA           ACCATAGACCTATOGC -53         TANAAAAAA         Hete (TANA-7)         Hete (TAAAAA)         Hete (TAAAAA)         Hete (TAAAAAA)         Hete (TAAAAAA)         Hete (TAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA		ATCTGTGGACAAGACTATCT <b>TGG</b>	+48	CAAAA	Hete(CAA/C)	Hete(CA/C)	ncRNA_intronic	LOC101928012
PdC2T07060ACT30ACTCVATAGC         -52         CAA         Hete (TAAA/TA)         Hete (TAAA/TA)         Hete (TAAA/TA)         Heto (TAA/TA)         He		TTGAAAAGACATAAAGCTAT <b>CAG</b>	-76	CAAAAAAAAAA	Hete(CAAAAAAAAAA/C-)	Hete (CAAAAAAAAAA(CA)	intergenic	LOC105376453;OTUD1
PackAAAAGCAATAAACCTAYCAG         C         CAA         Hete (CAAA/C)         Hete (CAAA/C)         intonice         DZANKI           TTATGECIACATAAACCTAYAAC         970         C         Hete (C/A)         G         intonice         BAGE2JAACE3BAAGE3BAA           ATCTGTGGCAATAAACCTAYAC         955         G         Hete (C/A)         G         intonice         TAGE2JAACE3BAAGE3BAA           ATCTGTGGCAATAAACCTAYCG         955         G         Bete (TAA/TA)         G         intergenic         LOC100421746,MIR435G-1           PULJ         ATCTGTGGCAATAAACCTAYCGA         955         G         Bete (TAAAAA-T)         Hete (TAAAAAA/T)         intergenic         DC100421746,MIR435G-1           PULJ         ATCTGTGGCAATAAACTAYCGA         977         TAAAAAAAAA         Hete (C/A)         Hete (TAAAAAA/T)         intergenic         EPHA/S           ATCTGTGGCAATAAACTAYCGA         974         TATGG         Hete (C/A)         Hete (C/A)         Hete (C/A)         Hete         MICONTGGCAATAAACTAYCGA         977         TAAAAAAA         Hete (C/A)         Hete (C/A)         Hete         MICONTGGCAATAAACTAYCGA         977         TAAAAAAA         Hete (C/A)         Hete (C/A)         Hete         MICONTGGCAATAAACTAYCGA         977         TAAAAAAA         Hete (C/A)         Hete (C/A)         Hete		ATCTGTGGACATTGGCTGAT <b>TGG</b>	-35	TAAAAAAA	Hete(TAAA/TA)	Hete(TAAA/T)	intronic	CACNA1A
Introduct NAMACTAINAGE 196         T         Hete (7/G)         T         Introduct NAMACTAINAGE 197         G         Hete (7/G)         G         Introduct NAMACTAINAGE 197         G         Hete (7/G)         G         Introduct NAMACTAINAGE 197         G         Hete (7/G)         Introduct NAMACTAINAGE 197         Introduct NAMAACTAINAGE 197         Introdu		AAGAAAAGACATAAAGCTAT <b>CAG</b>	-62	CAA	Hete(CAAA-/C)	Hete (CAAAA/C)	intronic	DZANK1
TATCSCAACATAAACCTAAUM       #9       G       Hete (TAA/TA)       G       Intronic       RAGE_BAGE_BAGE_BAGE_BAGE_BAGE_BAGE_BAGE_B		TTATGGCAACATAAAGCTAA <b>AAG</b>	+96	Т	Hete(T/G)	Т	intronic	BAGE2;BAGE3;BAGE4;BAG
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$		TTATGGCAACATAAAGCTAAAAG	+97	G	Hete (G/A)	G	intronic	BAGE2;BAGE3;BAGE4;BAG
Pol2J       ATCTGTGGACATIANGC TATING       Position       Reference       Treatment       Control       Functional region       Gene         ACCCATAGACATIANAGCTATINGG       -37       TAAAAAAAAA       Hete (TAAAAA-/T)       Hete (TAAAAA/T)       Hete (TAAAAAA/T)       Hete (TAAAAAAAAAA       Hete (TAAAAAAAAAAAAAA       Hete (TAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA		ATCTGTTGACCTAAAGCTAGAGG	+55	TAA	Hete (TAA/TA)	Hete (TAA/T)	intronic	TRIOBP
Ph24         ATCTGTGGGATAMAGCTAMOG         Moin         Refere         Temm         Contol         Bendination         Bendination         Gene           ACCATAGGATAMAGCTAMOG         -37         TAMAAMAA         Het (TAMAA-7)         Het (TAMAA-7)         Het (TAMAAAA)         Het (TAMAAAA)         Het (TAMAAAA)         Het (TAMAAAAA)         Het (TAMAAAAAA)         Het (TAMAAAAAAA)         Het (TAMAAAAAAA)         Het (TAMAAAAAAA)         Het (TAMAAAAAAA)         Het (TAMAAAAAAAA)         Het (TAMAAAAAAAAAAAAA)         Het (TAMAAAAAAAA)         Het (TAMAAAAAAAAA)         Het (TAMAAAAAAAA)         Het (TAMAAAAAAAAAA)         Het (TAMAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA			100	5		5	intergenie	LOC100421740, MIR4550-1
Phol2ATCTOTOGACATABAGCTATAGGPositioReferenceTeatmentControlHete (TABABAAT)IntronicEPHA 5ACCTATABAGCTATABAGCTACAGG-37TABAABAABAHete (TABABAAT)Hete (TABABAAT)IntronicSH3GCRL2ACTABAGCATABAGCTATAGAG+24ATGTGHomo (AHete (CA/CABAABAA)Hete (CA/CABAABACA)Hete (CA/CA)Hete (CA/C)Hete (CA/					sg-Pol2-J + sg-Tat1-J			
ACCCTTGGCATTAAACCTACGGG       -37       TAAAAAAAAA       Hete (TAAAAAAA/T)       Hete (TAAAAAAA/T)       Intronic       EPHAS         ACATAAAGCATAAAGCATAAAGCATAAAGCAAAACTTTGGG       -24       ATGG       Hete (C/CAAAAAAA)       Hete (C/CAAAAAAA)       Intronic       SHBGRL2         CGGGTGGGCATAAAACTTTGGG       -55       CAAAAAA       Hete (C/CAAAAAAA)       Hete (C/CAAAAAAA)       Intronic       LOCI01928012         ACGCATGGGCATAAAACTTACTGG       -47       ACT       Hete (C/CAAAAAAA)       Hete (C/CAAAAAAA)       Intronic       LOCI01928012         ACGCATGGGCATAAAACTTACTGG       -47       ACT       Hete (C/CAAAAAAA)       Hete (C/CAAAAAACTTACGGACGATAAAGCTACGACGACGACGACGACTTGCCTGATTGG       -53       TAAAAAAA       Hete (C/CACAAAACCA)       Hete (TAAA/T)       Intronic       CACNAIA         ATGCTGGGGCATAAAAGCATAAAGCATAAAGCACA       Hete (C/CACAAAACC-)       Hete (CAAAA/CA)       Intronic       CACNAIA         ATGCTGGGGCATAAAAACACAAAAGCATAAAACCA       48       Hete (C/C/CAACACAAAACCA)       Hete (C/C/CAACACAAAACCA)       Intronic       CACNAIA         ATACGCGGGCATAAAAACGAGGG       +55       6       Hono (C/C/TAA)       Hete (C/C/CAACACAAAACCA)       Intronic       ENAKI         TATGCGGCATAAAAACGGCA       +55       6       Hono (C/C)       Hete (C/C/G)       Intronic       NAKAAAAAAAAAAAAAAAAAAA	-Pol2-J	ATCTGTGGACATAAAGCTAT <b>NGG</b>	Position	Reference	Treatment	Control	Functional region	Gene
ACATABAAGCATCHACKC         24         ATCTC         Homo (A         Hete (AATCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG		AGCCATAGACATAAAGCTAG <b>CAG</b>	-37	TAAAAAAAA	Hete (TAAAAA-/T)	Hete (TAAAAAA/T)	intronic	EPHA5
CTGCGTGGACATTAACTTTTGG-55CAAAAAAHete (CA/CAAAAAAA)Hete (CA/CAAAAAAAAHete (CA/CAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA		ACATAAAGACATAAAGCTAG <b>GAG</b>	+24	ATGTG	Homo (A/A)	Hete (ATGTGTGTGTGTGTGTGTGTG/A)	intronic	SH3BGRL2
A ATCTTGTGGACABAGCLTATCTGG+48CAAAAHete (CA/C)Hete (CA/C)ncRNA_intronicLOC10928012GCCTTGGGACATAAAGATAGTAG+37THete (CA/C)TintergenicKDM4C/DMAC1TTGGACATAAAGCTATAAGCTATCGG-47ACTHete (A/ACT)ACTintergenicLOC10376453,0TUD1ATCTGTGGACATAAAGCTATAGG-47ACTHete (TAA-/TA)Hete (TAA/T)intronicCANAIAATCTGTGGACATTAGCTATCGG-55TAAAAAAAHete (CAAAAA/C-)Hete (CAAAA/CA)intronicDZANKIATCTGTGGACATAAAGCTATAAGC-56CGHomo (C/C)Hete (CAAAA/CA)intronicDAGE2BAGE3BAGE3BAGE4BAATCTGTGGGCATAAAATAGCGG+55GHomo (T/T)GintergenicLOC100421746,MIR4536-1CCTGGGGCATAAAATAGCGG+55GHomo (C/C)GintergenicLOC100421746,MIR4536-1CCTGCGCTCTCTCGCCCTGGG+48TAAAHomo (C/C)GintergenicMNR200A12P;LOC102722CCTCCGCTTCTTCCTCGCCATGG-47GHete (CA/C)intronicDNM3GCTCCCTTCTTCTCGCCATGG-48Hete (AAAAAAAAAAAAAGACA		CTGGGTGGACATAAACTTTTGAG	-85	CAAAAAA	Hete (CA/CAAAAAA)	Hete (C/CAAAAAA)	intergenic	NPTX2;TMEM130
GCCTTTGGACATAAAGATACHAG +37THete (C/T)TintergenicKDM4C;DMAC1TTGAAAAGACATAAAGCTATCAG -47ACTHete (A-/ACT)ACTintergenicLOC105376453;OTUD1ATCTGTGGACATGGCTGATTGG -35TAAAAAAAHete (TAA-/TA)Hete (TAAA/T)intronicCACNA1AAAGAAAAGCATAAAGCTATCAG -62CAAAHete (CAAAAA/C-)Hete (TAAA/T)intronicDAGE2;BAGE3;BAGE4;BAATCTGTGGCAATAAAGCTATAAG +26THomo (C/C)Hete (T/C)intronicDAGE2;BAGE3;BAGE4;BAATCTGTGGCCAATAAAGCTATTGG +55GHomo (TA'TA)Homo (T/T)intergenicLOC1042146;MR453-61CCTTGTGGACATGAAAGCTTTTGG +48TAAAHomo (TA/TA)Homo (T/T)intergenicANKRD20A12P;LOC102723TatI-JTCTCCCCTTCTTCCTGCCATGG -47GHomo (C/C)Hete (C/G)intronicDNM3GCTTCCCTTCTTCCTGCCACTGG -47GHomo (C/C)Hete (C/G)intronicDNM3GCTTCCCTTTCTTCCTGCCACTGG -41CHete (CTT/C)Hete (C/G)intronicDNM3CCTCCCCTTCTTCTTCCTGCCACTGG -41CAAAAHomo (C/CA)CintronicDNM3CCTCCCCTTCTTCCTGCCACTGG -51AAAAAAAAAAAAGAGA(A		ATCTGTGGACAAGACTATCT <b>TGG</b>	+48	CAAAA	Hete(CAA/C)	Hete (CA/C)	ncRNA_intronic	LOC101928012
International regions introduced at Long       -1/       ACT       Intergenic       LOCI0337643,07UD1         ATCTEGGACATTEGGACATTEGAC       -52       TANAANA       Hete (TAA)       Hete (TAA)       Intronic       CACNIA         AAGAAAAGCATAAAGCTATAGG       -52       CAAA       Hete (CAAAAA/C-)       Hete (TAA)       intronic       DZANKI         ATCTEGGACATTEGGCATTAGG       -52       CAAA       Hete (CAAAAA/C-)       Hete (TAC)       intronic       DZANKI         ATCTEGGACATAAAGCTATAG       -52       CAAA       Hete (CAAAA/C-)       Hete (TAC)       intronic       DZANKI         ATCTEGGACATAAAGCTATAG       +26       T       Homo (C/C)       Hete (TAC)       intronic       DZANKI         ATCTEGGACATAAAGCTATAGCTAGG       +48       TAA       Homo (TA/TA)       Homo (T/T)       intergenic       LOCI0421746;MIR4536-1         CCTCCCTCTCTCTCTCCTGCCATGGG       -47       G       Homo (C/C)       Hete (C/G)       intergenic       ANKRD20A12P;LOCI02723         CCTCCCCTCTCTCTCTCTCTCCCGCCATGG       -47       G       Hete (CTT/C)       C       intronic       DNM3         CCTCCCCTCTCTCTCTCTCTCCCGCCATGG       -47       G       Hete (CTT/C-)       Hete (CAC)       intronic       GRE2         CCTCCCCTCTCTCTCTCTCCCCGCCATGG		GGCTTTGGACATAAAGATAGTAG	+37	T	Hete (C/T)	T	intergenic	KDM4C;DMAC1
AACCASCIGNATION CONCENTING       -53       IADMADDAM       RECE (LAR/T)       RECE (LAR/T)       INTONIC       CACNAIA         AACAAAACACATAAAACCTATCAA       -62       CAA       Hete (CAAAAA/C)       Hete (T/C)       intronic       BAGE2;BAGE3;BAGE4;BA         ATCTGFGGCCATAAAAGCTATAAAC       +26       T       Homo (C/C)       Hete (T/C)       intronic       DAGE2;BAGE3;BAGE4;BA         ATCTGFGGCATAAAAACGCTATAGA       +26       T       Homo (C/C)       Hete (T/C)       intronic       LC100421746;MIR4536-1         CCTTGTGGGAACATGAAGCTTTGG       +55       G       Homo (T/T)       Homo (T/T)       Intergenic       TTY12;LINC00279         Tatl-J       TCTCGGCTTCTTGTCGCCATGG       Point       Reference       Treatment       Control       Functional region       Gene         GCTTCCTTGTCTGCCACTGG       -71       GAAA       Homo (C/C)       Bete (C/G)       intronic       DNM3         TCTCGCGCTTGCTTTTGTCGCACTGG       -71       CAAAA       Homo (C/C)       C       intronic       DNM3         TCTCGCGCTTGCTGTCTGTCGCATGG       -71       CAAAA       Homo (C/C)       Hete (C/A)       intronic       DNM3         TCTCGCGCTGTGTGTGTGCACAGG       +80       C       Hete (C/T/C)       Bete (C/C)       intronic       RH212 <td></td> <td>TTGAAAAGACATAAAGCTATCAG</td> <td>-4/</td> <td>ACT</td> <td>Hoto (TA) (TA)</td> <td></td> <td>intergenic</td> <td>LUC105376453;OTUD1</td>		TTGAAAAGACATAAAGCTATCAG	-4/	ACT	Hoto (TA) (TA)		intergenic	LUC105376453;OTUD1
TATAGECARACATABAGV2OWNINCOMPANYINCOMPANYINCOMEDZANNITTATGECARACATABAGV2THomo (C/C)Hete (T/C)intronicBAGE2;BAGE3;BAGE4;BAATCTGTGGGCATABAATAGGAGG+55GHomo (GT/GT)GintergenicLOC100421746;MIR4536-1CCTTGTGGACATGAAGCTTTTGG+48TAAAHomo (TA/TA)Homo (T/T)intergenicTTTY12;LINC00279Tatl-JTCTCCGCTTCTTCCTGCCCATGG-47GHomo (C/C)Hete (C/G)intergenicANKRD20A12P;LOC102723CCTCCGCTTCTTCTTCTCTGCCCAGG-47GHomo (C/C)Hete (C/G)intronicDNM3CCTCCGCTTCTTCTTCTCTGCCCAGG-47GHomo (C/C)Hete (C/C)intronicDNM3CCTCCGCTTCTTCTTCTCTGCCACTGG-71CAAAAHomo (C/CA)Hete (C/C)intronicTRNM2CCTCCGCTTCTGCTCTCCACTGG-71CAAAAHomo (C/CA)Hete (C/C/C)intronicGRHL2CCTCCGCTTCTGCCACTGG-58AAAAAAAAAAAGAGHete (AAAAAAAAAGAG/A)AAAAAAAAAAAGAGintronicGRHL2CCTCCCACTTCTCCGCCACGG-58AAAAAAAAAAAGAGHete (CTT/C-)Hete (CTT/CT)intronicILA1217GATCACTTCTCTCGCCACTGG+82CTTTTTHete (G/T)Homo (A/A)ncRNA_intronicDX12PGATCACTCTCTCCGCCACGG+85GHete (G/A)Homo (A/A)ncRNA_intronicDX12PACCCCCCTCTTCCCCCCCTGGA+85GHete (G/A)AncRNA_intronicRNF213-AS1ACCCCCCCTCTTCCCCCCTGTAA+45 </td <td></td> <td>AAGAAAAGACATAAAGCTATCAC</td> <td>-53</td> <td>CAAA</td> <td>Hete (CAAAAA/C-)</td> <td>Hete (CAAAAA/CA)</td> <td>intronic</td> <td>DZANKI</td>		AAGAAAAGACATAAAGCTATCAC	-53	CAAA	Hete (CAAAAA/C-)	Hete (CAAAAA/CA)	intronic	DZANKI
ATCTGTGGGCATAGAAATAGGAG       -       Intercent       Intercent       DAGE_DAGE_DAGE_DAGE_DAGE_DAGE_DAGE_DAGE_		TTATGGCAACATAAAGCTAAAAC	+26	T	Homo (C/C)	Hete (T/C)	intronic	BAGE2-BAGE3-BAGE4-BAG
CCTTGTGGGACATGAGGTTTTGG44TAAHomo (TA/TA)Homo (T/T)IntegenicITTY12;LINC0042/140;MIN45041Tatl-JTCCCGGCTTCTCCCGCGCATGGPositionReferenceTreatmentControlFunctional regionGeneGCTTCCCTTCTCTCGCGCCTGGG-47GHomo (C/C)Hete (C/G)intergenicANKRD20A12P;LOC102723CCTCGCGTTCTTCTCTCGCCACGG-47GHomo (C/C)Hete (C/G)intergenicANKRD20A12P;LOC102723CCTCGCGTTCTTCTTCTCGCCACGG-47GHete (CTT/C)CintronicDNM3CTTCCGCTTCTTCTTCTCGCCACGG-47GHete (CTT/C)CintronicDNM3CTTCTCTCTCTCCGCCACGG-58AAAAAAAAAAAGGGHete (CA/C)intronicTENM2CCTGTCCACGCTTCGCCACGGG-58AAAAAAAAAAAAGGGHete (CA/C)intronicGRHL2CCTGTCCACGCTTCGCCACGGG-58AAAAAAAAAAAAGGGHete (CTT/CT)intergenicIERSL,LOC101929331GATCACTTCTCTCGCCACGGG+82CTTTTTHete (CTT/C-)Hete (CTTTTT/C)intronicKIAA1217TTTCTCTTCTTCCTGCCCCTGGG+85GHete (G/A)Homo (T/T)ncRNa, intronicDX12PACCCCGCTTCTTCCCCCTGTAAG+26CHete (G/A)AncRNA_intronicRNF213-AS1ACCCCGCTTCTTCCCCCTGTAAG+29GHete (G/A)AncRNA_intronicRNF213-AS1ACCCCGCTTCTCCCCCTGTAAG+45AHete (G/A)AncRNA_intronicRNF213-AS1ACCCCGCTTCTCCCCCTGTAAG+45AHete (G/A)A<		ATCTGTGGGCATAAAATAGGAGG	+55	G	Homo (GT/GT)	G	intergenic	LOC100421746 MIR4536-1
Tatl-JTCTCCCCTTCTCCTCCCCCCCCGPositioReferenceTreatmentControlFunctional regionGeneGCTTCCCTTCTTCTCTCCCCCCCCG-47GHomo (C/C)Hete (C/G)intergenicANKRD20A12P;LOC10272:CCTCCCCTTCTTCTTCTCCCGCCACG-40CHete (CTT/C)CintronicDNM3CCTCCCTTTCTTCCTCCCCCCACG-50AAAAAAAAAAAGAGHete (AAAAAAAAGAG/AHete (CA/C)intronicTENM2CCTCCCTTCTTCTCTCCGCCACGG-50AAAAAAAAAAAGAGHete (AAAAAAAAAGAG/AAAAAAAAAAGAGIntronicIRSL;LOC101929331CCTCCCTCTCTCTCCTCCCCCCAGG+82CTTTTTHete (CTT/C-)Hete (CTTTT/C)intronicIAAAAAAAAGAGGATCACTTCTCTCTCCTCCCCCTGGG+86GHete (G/A)Homo (A/A)ncRNa_intonicDX12PTTTCTTCTTCTCCTCCCCCTGTAG+86GHete (A/C)CncRNa_intonicRNF213-AS1ACCCCCCTTTCCCCCCTGTAGA+50AHete (G/A)AncRNa_intonicRNF213-AS1ACCCCCCTTTCCCCCCTGTAGA+50AHete (G/A)AncRNa_intonicRNF213-AS1ACCCCCCTTTCCCCCCTGTAGA+50AHete (G/A)AncRNa_intonicRNF213-AS1ACCCCCCTTCTCCCCCCTGTAGA+50AHete (G/A)AncRNa_intonicRNF213-AS1ACCCCCCTTCTCCCCCTGTAGA+50AHete (G/A)AncRNa_intonicRNF213-AS1ACCCCCCTTCTCCCCCTGTAGA+50AHete (G/A)AncRNa_intonicRNF213-AS1ACCCCCCCTTCTCCCCCTGTAGA+50 <t< td=""><td></td><td>CCTTGTGGACATGAAGCTTT<b>TGG</b></td><td>+48</td><td>TAAA</td><td>Homo (TA/TA)</td><td>Homo (T/T)</td><td>intergenic</td><td>TTTY12;LINC00279</td></t<>		CCTTGTGGACATGAAGCTTT <b>TGG</b>	+48	TAAA	Homo (TA/TA)	Homo (T/T)	intergenic	TTTY12;LINC00279
GCTTCCCTTCTTCCTCGCCTCGG-47GHomo (C/C)Hete (C/G)IntergenicANKRD20A12P;LOC10272:CCTCCCTTTCTTTCTTCCCGCACTGG+40CHete (CTT/C)CintronicDNM3CTACCTTTTCTTCCTCCCCCCTGG-71CAAAHomo (CA/CA)Hete (CA/C)intronicTENM2CTACCTTTTCTCCTCCCCCCTTGG-58AAAAAAAAAAAGAGHete (AAAAAAAAGAG/A)AAAAAAAAAGAGIntronicGRHL2CCTCCCACTTCTCCCCCCTTGG+82CTTTTTHete (CTT/C-)Hete (CTTTTC)intronicIER5L;LOC101929331GATCACTTCTTCCTCCTCCTCGCATGG+82CTTTTTHete (CTTT-/C)Hete (CTTTTC)intronicIKIA1217GATCACTTCTTCCTCCTCCTCGG+86GHete (G/A)Homo (A/A)ncRNA_intronicDX12PTTTCTTCTTCTCCTCCCCCTCTAG+85GHete (G/A)Homo (A/A)ncRNA_intronicRNF213-AS1ACCCCGCTTCTTCCCCCCTCTAG+29GHete (G/A)AncRNA_intronicRNF213-AS1ACCCGCCTTCTTCCCCCCTCTAG+45AHete (G/A)AncRNA_intronicRNF213-AS1ACCCGCCTTCTTCCCCCCTCTAG+45AHete (G/A)AncRNA_intronicRNF213-AS1ACCCGCCTTCTTCCCCCCTCTAG+55AHete (G/A)AncRNA_intronicRNF213-AS1GCTTCCCTTCTCCCCCTCTAG+56AHete (G/A)AncRNA_intronicRNF213-AS1GCTTCCCTTCTCCCCCTCTAG+55AHete (G/A)AncRNA_intronicNNE;MIR3648-1	Tat1-J	TCTCCGCTTCTTCCTGCCAT <b>NGG</b>	Position	Reference	Treatment	Control	Functional region	Gene
CCTCCCCTTCTTCTTCCTCCCCCCG 40CHete(CTT/C)CintronicDNM3CTACCTTTCTTCCTCCCCCCGG 70CAAAHom (CA/CA)Hete (AAAAAAAACAGintronicTENM2TCTCCCCTTTCTCCACATGG 75AAAAAAAAAAAAAAGAGHete (AAAAAAAAAAGAG/A		GCTTCCCTTCTTCCTGCCCTCGG	-47	G	Homo (C/C)	Hete (C/G)	intergenic	ANKRD20A12P;LOC102723
CTACCTTTCTCCTCCCCCTGG-71CAAAAHomo(CA/CA)Hete(CA/CA)Hete(A/CA)Hete(A/CA)Hete(A/CA)Hete(A/CA)Hete(A/CA)Hete(A/CA)Hete(CT/CTA)IntronicCRHL2CCTGTCCACGTCTCGCCATGG-58AAAAAAAAGAGHete(AAAAAAAAAGAGA/A)AAAAAAAAGAGAIntronicCRHL2CCTGTCCACGTCTCGCCATGG+26CTTTTTTHete(CTT/C-)Hete(CTT/CT)IntergenicIERSL;LOC101929331GATCACTCTCTCCTCGCCATGG+78CTTTTHete(CTTT-/C)Hete(CTT/T/C)intronicKIAA1217TTTCTTCTTCTCTCCTCCTCCTGG+55GHete(G/A)Homo(A/A)mcRNA_intronicDX12PACCCCCCCTTCTCCCCCCTGTAGG+55GHete(A/C)CncRNA_intronicRNF213-AS1ACCCCCCCTTCTCCCCCCTGTAGG+59GHete(G/A)AncRNA_intronicRNF213-AS1ACCCCCCTTCTCCCCCCTGTAGG+50AHete(G/A)AncRNA_intronicRNF213-AS1ACCCCCCTTCTCCCCCCTGTAGG+55AHete(G/A)AncRNA_intronicRNF213-AS1ACCCCCCTTCTCCCCCCTGTAGG+55AHete(G/A)AncRNA_intronicRNF213-AS1ACCCCCCTTCTCCCCCTGTAGG-55AHete(G/A)AntergenicNONE;MIR3648-1		CCTCCGCTTCTTCTTCTCCAG	+40	С	Hete(CTTT/C)	с	intronic	DNM3
TCTCCCCCTCTCTCCCCCCTCTAGG-58AAAAAAAAAAAGAGHete (AAAAAAAAAAGAG/A)AAAAAAAAAAAGAGintronicGRHL2CCTGTCCCAGCTTCTGCCATGG+82CTTTTTHete (CTT/C-)Hete (CT/CT)intronicIESL_LOCI01929331GATTCACTTCTTCCTGCCAAGGG+78CTTTTHete (CTT/C-)Hete (CTT/CT)intronicIESL_LOCI01929331TTTCTTCTTCTTCCTGCCCACGGG+36GHete (G/T)Hoto (T/T)ncRNA_intronicDDX12PTTTCTTCTTCTTCTCCCCCCTGGG+85GHete (G/A)Homo (A/A)ncRNA_intronicDDX12PACCCCGCTTCTTCCCCCCTGTAGG+29GHete (A/C)CncRNA_intronicRNF213-AS1ACCCCGCTTCTTCCCCCCTGTAGG+45AHete (G/A)AncRNA_intronicRNF213-AS1ACCCCGCTTCTTCCCCCCTGTAGG+45AHete (G/A)AncRNA_intronicRNF213-AS1ACCCCGCTTCTTCCCCCCTGTAGG+55AHete (G/A)AncRNA_intronicRNF213-AS1ACCCCGCTTCTTCCCCCCTGTAGG+55AHete (G/A)AncRNA_intronicNNF2iMI3648-1GCTTCCCTTCTCCCCCCTGTAGG-55AHete (G/A)AncRNA_intronicNNF2iMI3648-1		CTACCTTTTTCTTCCTGCCACTGG	-71	CAAAA	Homo (CA/CA)	Hete (CA/C)	intronic	TENM2
CCTGTCCACCTTCCTGCCATAGG+82CTTTTTHete (CTT/C-)Hete (CTT/C1)intergenicIERSI,LOC101929331GATCACTTCCTGCCATAGGA+78CTTTTHete (CTT/C-)Hete (CTTTC/C)intronicKIAA1217TTTCTCTCTTCTCCTGCCTCTGG+36GHete (CTT/C-)Hom (T/T)ncRNA_intronicDDX12PTTTCTTCTTCTCCTCCCCCCTCTGG+85GHete (G/A)Hom (A/A)ncRNA_intronicDDX12PACCCCCCTCTTCCCCCCCTCTAGG+26CHete (A/C)CncRNA_intronicRNF213-AS1ACCCCCCCTCTTCCCCCCCTCTAGG+29GHete (T/G)GncRNA_intronicRNF213-AS1ACCCCCCTCTTCCCCCCCTCTAGG+45AHete (G/A)AncRNA_intronicRNF213-AS1GCTTCCCTTCTCCCCCCCCCCCCCG-65AHete (G/A)AintergenicNONE;MIR3648-1		TCTCCGCTTCTGTCTTCACATGG	-58	AAAAAAAAAAGAG	Hete (AAAAAAAAAAGAG/A)	AAAAAAAAAGAG	intronic	GRHL2
GATICACITICTTCCTGCCARGG +78CTTTTHete (CTTT-/C)Hete (CTTTT/C)intronicKIAA1217TTCTCTCTTCTTCCTGCCTCTGG +36GHete (CTTT-/C)Homo (T/T)ncRNA_intronicDDX12PTTTCTTCTTCTTCCTCCCCCTCTAGG +85GHete (G/A)Homo (T/T)ncRNA_intronicDDX12PACCCCGCTTCTTCCCCCCTCTAGG +26CHete (A/C)CncRNA_intronicRNF213-AS1ACCCCGCTTCTTCCCCCCTCTAGG +45AHete (T/G)GncRNA_intronicRNF213-AS1ACCCCGCTTCTTCCCCCCTCTAGG +45AHete (G/A)AncRNA_intronicRNF213-AS1GCTTCCCTTCTCCCCCTCTGGC -65AHete (G/A)AintergenicNONE;MIR3648-1		CCTGTCCAGCTTCCTGCCAT <b>TGG</b>	+82	CTTTTTTTT	Hete(CTT/C-)	Hete (CTT/CT)	intergenic	IER5L;LOC101929331
THEELECTORG436GHete (G/A)Homo (T/T)ncRNA_intronicDDX12PTTCTTCTTCTTCCTCCCCCTCTAG+25GHete (G/A)Homo (A/A)ncRNA_intronicDDX12PACCCCCCTTCTTCCCCCCCTCTAG+26CHete (A/C)CncRNA_intronicRNF213-ASIACCCCCCTTCTTCCCCCCCTCTAG+29GHete (T/G)GncRNA_intronicRNF213-ASIACCCCCCTTCTTCCCCCCCTCTAG+45AHete (G/A)AncRNA_intronicRNF213-ASIACCCCCCTTCTTCCCCCCCTCGC-55AHete (G/A)AintergenicNONE;MIR3648-1		GATTCACTTCTTCCTGCCAAGAG	+78	CTTTTT	Hete (CTTTT-/C)	Hete (CTTTTT/C)	intronic	KIAA1217
ACCCCGCTTCTTCCCCCCTGTAAG+25GHete(A/C)CnckNA_intronicDDX12PACCCCGCTTCTTCCCCCCTGTAAG+29GHete(A/C)CncRNA_intronicRNF213-AS1ACCCCGCTTCTTCCCCCCTGTAAG+29GHete(G/A)AncRNA_intronicRNF213-AS1ACCCCGCTTCTTCCCCCCTGTAAG+45AHete(G/A)AncRNA_intronicRNF213-AS1GCTTCCCTTCTTCCTCCCCCTGCG-65AHete(G/A)AintergenicNONE;MIR3648-1		TTTCTTCTTCTTCCTGCCTC <b>TGG</b>	+30	G	Hoto (G/A)	Home (A/A)	ncKNA_intronic	DDX12P
ACCCCGCTTCTTCCCCCCCTCTARG     +29     G     Hete (G/A)     G     mcRNA_intronic     RNF213-AS1       ACCCCGCTTCTTCCCCCCCTCTARG     +45     A     Hete (G/A)     A     ncRNA_intronic     RNF213-AS1       GCTTCCCTTCTTCCCCCCCTCGG     -65     A     Hete (G/A)     A     intergenic     NONE;MIR3648-1		ACCCCCCCTTCTTCCCCCCCTCTGG	+26	c	Hete (B/A)	nonio (A/A)	ncRNA_intronic	DUA12P DNE212_AS1
ACCCCCCTTTTCCCCCCCTTARG     +45     A     Hete(G/A)     A     ncRNA_introic     RNF13-AS1       GCTTCCCTTCTCCCCCCCTGGG     -65     A     Hete(G/A)     A     intergenic     NONE;MIR3648-1		ACCCCCCTTTTTCCCCCTCTAAG	+29	Ğ	Hete (T/G)	Ğ	ncRNA_intronic	RNF213-AS1
GCTTCCCTTCTTCCTGCCCTCGG -65 A Hete (G/A) A intergenic NONE;MIR3648-1		ACCCCGCTTCTTCCCCCTGTAAG	+45	Ā	Hete (G/A)	A	ncRNA intronic	RNF213-AS1
			-65	A	Hete (G/A)	А	intergenic	NONE;MIR3648-1
		GCTTCCCTTCTTCCTGCCCTCGG						

A novel Adeno-associated virus (AAV)-based vector (AAV9P1) with a synthetic surface peptide can deliver CRISPR/ Cas9 system to astrocytes and might facilitate inactivation of persisting HIV-1 proviruses (Kunze et al., 2018), which is of great significance to eliminate viral reservoirs in the central nervous system. AAV is also a promising tool for gene delivery *in vivo*. Excision of HIV-1 proviral DNA in solid tissues and organs in humanized mice with chronic HIV-1 infection can be achieved *via* AAV-delivered CRISPR/Cas9 (Yin et al., 2017). According to a recent study, AAV9-CRISPR/Cas9 gene editing was designed to eliminate simian immunodeficiency virus (SIV) proviral DNA in ART-treated

non-human primates (Mancuso et al., 2020). These findings offer a significant step toward the elimination of HIV-1 reservoirs in the clinic. Notably, CRISPR-Cas13a could inhibit HIV-1 infection by directly targeting viral RNA and diminishing viral gene expression (Yin et al., 2020). RNA editing technology has the potential to treat HIV-1 infection; however, the off-target effects cannot be ignored (Mao et al., 2019).

#### CONCLUSION

We designed and verified multiple sgRNAs and combinations comprising dual-sgRNAs that targeted HIV-1 genomic DNA in a cell line model of latency. WGS was used to assess off-target effects in the three efficient combinations selected, and no reliable off-target sites were detected. The study provided a comprehensive experimental basis and methodological suggestions for CRISPR/Cas9 individualized treatment of HIV-1 in the future, from the design of sgRNAs to the selection of efficient combinations and verification of safety. On the premise of effectiveness and safety, the dual-sgRNAs CRISPR/ Cas9 treatment targeting HIV-1 protein-coding genes could serve as a therapeutic procedure to achieve a cure and epidemic control of HIV/AIDS.

### DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

## AUTHOR CONTRIBUTIONS

YX and XP performed the research, analyzed the data, and wrote the manuscript. NW designed the research and revised the manuscript. YZ, CJ, XL, DH, HF, and CC performed

### REFERENCES

- Bella, R., Kaminski, R., Mancuso, P., Young, W. B., Chen, C., Sariyer, R., et al. (2018). Removal of HIV DNA by CRISPR from patient blood engrafts in humanized mice. *Mol. Ther. Nucleic Acids* 12, 275–282. doi: 10.1016/j. omtn.2018.05.021
- Binda, C. S., Klaver, B., Berkhout, B., and Das, A. T. (2020). CRISPR-Cas9 dual-gRNA attack causes mutation, excision and inversion of the HIV-1 proviral DNA. *Viruses* 12:330. doi: 10.3390/v12030330
- Bogerd, H. P., Kornepati, A. V., Marshall, J. B., Kennedy, E. M., and Cullen, B. R. (2015). Specific induction of endogenous viral restriction factors using CRISPR/Cas-derived transcriptional activators. *Proc. Natl. Acad. Sci. U. S.* A. 112, E7249–E7256. doi: 10.1073/pnas.1516305112
- Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823. doi: 10.1126/science.1231143
- Dampier, W., Sullivan, N. T., Mell, J. C., Pirrone, V., Ehrlich, G. D., Chung, C. H., et al. (2018). Broad-spectrum and personalized guide RNAs for CRISPR/ Cas9 HIV-1 therapeutics. *AIDS Res. Hum. Retrovir.* 34, 950–960. doi: 10.1089/ aid.2017.0274

the research. All authors read and approved the final version of the manuscript.

## FUNDING

This study was supported financially by the National Science and Technology Major Project during the 13th 5-year Plan of China (grant no. 2017ZX10202101-004-004) and the National High Technology Research and Development Program of China (863 Program, grant no. SS2015AA020310).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.629153/ full#supplementary-material

Supplementary Figure S1 | Analysis of the sequences of HIV-1 mutation sites in dual-sgRNAs/Cas9-treated J-Lat 10.6 cells. Feature of the mutations in the three most efficient dual-sgRNA combinations: (A) sg-Gag1-J + sg-Gag4-J
(B) sg-Gag1-J + sg-Pol1-J (C) sg-Gag2-J + sg-Pol1-J. All sequences were aligned with the reference strain HXB2. The 20 nt guide sequence and the PAM sequence are underlined. The sequences of nucleotide substitutions, insertions, and deletions are indicated using red letters. Ref, reference sequence; -x/+x, x nt deleted/inserted; NS, nucleotide substitution; WT, wild-type sequence.

**Supplementary Figure S2** | Sequence analysis of HIV-1 mutation sites in J-Lat 10.6 cells treated with dual-sgRNAs/Cas9. Features of the mutations in the three most efficient dual-sgRNA combinations: (A) sg-Gag2-J + sg-Gag5-J (B) sg-Pol1-J + sg-Pol3-J (C) sg-Tat1-J + sg-Env2-J. All sequences were aligned with the reference strain HXB2. The 20 nt guide sequence and the PAM sequence are underlined. The sequences of nucleotide substitutions, insertions, and deletions are indicated using red letters. Ref, reference sequence; -x/+x, x nt deleted/ inserted; NS, nucleotide substitution; WT, wild-type sequence.

**Supplementary Figure S3** | Deletion of the entire viral DNA between two sgRNA-targeted sites in the combination sg-Gag1-J + sg-Gag2-J (A) 169 bp deletion (B) 177 bp deletion with two single-base substitutions.

**Supplementary Figure S4** | Chromosomal inversion between two sgRNAtargeted sites in the combination sg-Gag1-J + sg-Gag5-J.

- Das, A. T., Binda, C. S., and Berkhout, B. (2019). Elimination of infectious HIV DNA by CRISPR-Cas9. Curr. Opin. Virol. 38, 81–88. doi: 10.1016/j. coviro.2019.07.001
- Dash, P. K., Kaminski, R., Bella, R., Su, H., Mathews, S., Ahooyi, T. M., et al. (2019). Sequential LASER ART and CRISPR treatments eliminate HIV-1 in a subset of infected humanized mice. *Nat. Commun.* 10:2753. doi: 10.1038/ s41467-019-10366-y
- Ebina, H., Misawa, N., Kanemura, Y., and Koyanagi, Y. (2013). Harnessing the CRISPR/Cas9 system to disrupt latent HIV-1 provirus. *Sci. Rep.* 3:2510. doi: 10.1038/srep02510
- Feng, Z., Mao, Y., Xu, N., Zhang, B., Wei, P., Yang, D. L., et al. (2014). Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.* 111, 4632–4637. doi: 10.1073/pnas.1400822111
- Ghorbal, M., Gorman, M., MacPherson, C. R., Martins, R. M., Scherf, A., and Lopez-Rubio, J. J. (2014). Genome editing in the human malaria parasite *Plasmodium falciparum* using the CRISPR-Cas9 system. *Nat. Biotechnol.* 32, 819–821. doi: 10.1038/nbt.2925
- Hemelaar, J. (2012). The origin and diversity of the HIV-1 pandemic. Trends Mol. Med. 18, 182–192. doi: 10.1016/j.molmed.2011.12.001

- Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* 31, 827–832. doi: 10.1038/nbt.2647
- Hu, W., Kaminski, R., Yang, F., Zhang, Y., Cosentino, L., Li, F., et al. (2014). RNA-directed gene editing specifically eradicates latent and prevents new HIV-1 infection. *Proc. Natl. Acad. Sci. U. S. A.* 111, 11461–11466. doi: 10.1073/pnas.1405186111
- Ji, H., Jiang, Z., Lu, P., Ma, L., Li, C., Pan, H., et al. (2016). Specific reactivation of latent HIV-1 by dCas9-SunTag-VP64-mediated guide RNA targeting the HIV-1 promoter. *Mol. Ther.* 24, 508–521. doi: 10.1038/mt.2016.7
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821. doi: 10.1126/science.1225829
- Kang, H., Minder, P., Park, M. A., Mesquitta, W. T., Torbett, B. E., and Slukvin, I. I. (2015). CCR5 disruption in induced pluripotent stem cells using CRISPR/Cas9 provides selective resistance of immune cells to CCR5tropic HIV-1 virus. *Mol. Ther. Nucleic Acids* 4:e268. doi: 10.1038/mtna.2015.42
- Kunze, C., Börner, K., Kienle, E., Orschmann, T., Rusha, E., Schneider, M., et al. (2018). Synthetic AAV/CRISPR vectors for blocking HIV-1 expression in persistently infected astrocytes. *Glia* 66, 413–427. doi: 10.1002/glia.23254
- Laskey, S. B., and Siliciano, R. F. (2014). A mechanistic theory to explain the efficacy of antiretroviral therapy. *Nat. Rev. Microbiol.* 12, 772–780. doi: 10.1038/nrmicro3351
- Li, J., Manghwar, H., Sun, L., Wang, P., Wang, G., Sheng, H., et al. (2018). Whole genome sequencing reveals rare off-target mutations and considerable inherent genetic or/and somaclonal variations in CRISPR/Cas9-edited cotton plants. *Plant Biotechnol. J.* 17, 858–868. doi: 10.1111/pbi.13020
- Liao, H. K., Gu, Y., Diaz, A., Marlett, J., Takahashi, Y., Li, M., et al. (2015). Use of the CRISPR/Cas9 system as an intracellular defense against HIV-1 infection in human cells. *Nat. Commun.* 6:6413. doi: 10.1038/ncomms7413
- Limsirichai, P., Gaj, T., and Schaffer, D. V. (2016). CRISPR-mediated activation of latent HIV-1 expression. Mol. Ther. 24, 499–507. doi: 10.1038/mt.2015.213
- Liu, Z., Chen, S., Jin, X., Wang, Q., Yang, K., Li, C., et al. (2017). Genome editing of the HIV co-receptors CCR5 and CXCR4 by CRISPR-Cas9 protects CD4 + T cells from HIV-1 infection. *Cell Biosci.* 7:47. doi: 10.1186/ s13578-017-0174-2
- Liu, S., Wang, Q., Yu, X., Li, Y., Guo, Y., Liu, Z., et al. (2018). HIV-1 inhibition in cells with CXCR4 mutant genome created by CRISPR-Cas9 and piggyBac recombinant technologies. *Sci. Rep.* 8:8573. doi: 10.1038/s41598-018-26894-4
- Lyu, C., Shen, J., Wang, R., Gu, H., Zhang, J., Xue, F., et al. (2018). Targeted genome engineering in human induced pluripotent stem cells from patients with hemophilia B using the CRISPR-Cas9 system. *Stem Cell Res Ther* 9:92. doi: 10.1186/s13287-018-0839-8
- Maartens, G., Celum, C., and Lewin, S. R. (2014). HIV infection: epidemiology, pathogenesis, treatment, and prevention. *Lancet* 384, 258–271. doi: 10.1016/ S0140-6736(14)60164-1
- Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., et al. (2013). RNA-guided human genome engineering via Cas9. *Science* 339, 823–826. doi: 10.1126/science.1232033
- Mancuso, P., Chen, C., Kaminski, R., Gordon, J., Liao, S., Robinson, J. A., et al. (2020). CRISPR based editing of SIV proviral DNA in ART treated non-human primates. *Nat. Commun.* 11:6065. doi: 10.1038/s41467-020-19821-7
- Mao, S., Liu, Y., Huang, S., Huang, X., and Chi, T. (2019). Site-directed RNA editing (SDRE): off-target effects and their countermeasures. J. Genet. Genom. 46, 531–535. doi: 10.1016/j.jgg.2019.11.005
- Mefferd, A. L., Bogerd, H. P., Irwan, I. D., and Cullen, B. R. (2018). Insights into the mechanisms underlying the inactivation of HIV-1 proviruses by CRISPR/Cas. Virology 520, 116–126. doi: 10.1016/j.virol.2018.05.016
- Ophinni, Y., Inoue, M., Kotaki, T., and Kameoka, M. (2018). CRISPR/Cas9 system targeting regulatory genes of HIV-1 inhibits viral replication in infected T-cell cultures. *Sci. Rep.* 8:7784. doi: 10.1038/s41598-018-26190-1
- Panfil, A. R., London, J. A., Green, P. L., and Yoder, K. E. (2018). CRISPR/ Cas9 genome editing to disable the latent HIV-1 provirus. *Front. Microbiol.* 9:3107. doi: 10.3389/fmicb.2018.03107
- Qi, C., Li, D., Jiang, X., Jia, X., Lu, L., Wang, Y., et al. (2018). Inducing CCR5Δ32/Δ32 homozygotes in the human jurkat CD4+ cell line and primary CD4+ cells by CRISPR-Cas9 genome-editing technology. *Mol. Ther. Nucleic Acids* 12, 267–274. doi: 10.1016/j.omtn.2018.05.012

- Roychoudhury, P., De Silva Feelixge, H., Reeves, D., Mayer, B. T., Stone, D., Schiffer, J. T., et al. (2018). Viral diversity is an obligate consideration in CRISPR/Cas9 designs for targeting the HIV reservoir. *BMC Biol.* 16:75. doi: 10.1186/s12915-018-0544-1
- Saayman, S. M., Lazar, D. C., Scott, T. A., Hart, J. R., Takahashi, M., Burnett, J. C., et al. (2016). Potent and targeted activation of latent HIV-1 using the CRISPR/dCas9 activator complex. *Mol. Ther.* 24, 488–498. doi: 10.1038/ mt.2015.202
- Sanjana, N. E., Shalem, O., and Zhang, F. (2014). Improved vectors and genomewide libraries for CRISPR screening. *Nat. Methods* 11, 783–784. doi: 10.1038/ nmeth.3047
- ter Brake, O., Konstantinova, P., Ceylan, M., and Berkhout, B. (2006). Silencing of HIV-1 with RNA interference: a multiple shRNA approach. *Mol. Ther.* 14, 883–892. doi: 10.1016/j.ymthe.2006.07.007
- Wang, Q., Liu, S., Liu, Z., Ke, Z., Li, C., Yu, X., et al. (2018). Genome scale screening identification of SaCas9/gRNAs for targeting HIV-1 provirus and suppression of HIV-1 infection. *Virus Res.* 250, 21–30. doi: 10.1016/j. virusres.2018.04.002
- Wang, Z., Pan, Q., Gendron, P., Zhu, W., Guo, F., Cen, S., et al. (2016a). CRISPR/Cas9-derived mutations both inhibit HIV-1 replication and accelerate viral escape. *Cell Rep.* 15, 481–489. doi: 10.1016/j.celrep.2016.03.042
- Wang, G., Zhao, N., Berkhout, B., and Das, A. T. (2016b). A combinatorial CRISPR-Cas9 attack on HIV-1 DNA extinguishes all infectious provirus in infected T cell cultures. *Cell Rep.* 17, 2819–2826. doi: 10.1016/j.celrep.2016.11.057
- Xu, L., Wang, J., Liu, Y., Xie, L., Su, B., Mou, D., et al. (2019). CRISPR-edited stem cells in a patient with HIV and acute lymphocytic leukemia. N. Engl. J. Med. 381, 1240–1247. doi: 10.1056/NEJMoa1817426
- Xu, L., Yang, H., Gao, Y., Chen, Z., Xie, L., Liu, Y., et al. (2017). CRISPR/ Cas9-mediated CCR5 ablation in human hematopoietic stem/progenitor cells confers HIV-1 resistance in vivo. *Mol. Ther.* 25, 1782–1789. doi: 10.1016/j. ymthe.2017.04.027
- Yin, L., Hu, S., Mei, S., Sun, H., Xu, F., Li, J., et al. (2018). CRISPR/Cas9 inhibits multiple steps of HIV-1 infection. *Hum. Gene Ther.* 29, 1264–1276. doi: 10.1089/hum.2018.018
- Yin, C., Zhang, T., Qu, X., Zhang, Y., Putatunda, R., Xiao, X., et al. (2017). In vivo excision of HIV-1 provirus by saCas9 and multiplex single-guide RNAs in animal models. *Mol. Ther.* 25, 1168–1186. doi: 10.1016/j. ymthe.2017.03.012
- Yin, L., Zhao, F., Sun, H., Wang, Z., Huang, Y., Zhu, W., et al. (2020). CRISPR-Cas13a inhibits HIV-1 infection. *Mol. Ther. Nucleic Acids* 21, 147–155. doi: 10.1016/j.omtn.2020.05.030
- Yu, S., Yao, Y., Xiao, H., Li, J., Liu, Q., Yang, Y., et al. (2018). Simultaneous knockout of CXCR4 and CCR5 genes in CD4+ T cells via CRISPR/Cas9 confers resistance to both X4- and R5-tropic human immunodeficiency virus type 1 infection. *Hum. Gene Ther.* 29, 51–67. doi: 10.1089/hum.2017.032
- Zhang, Y., Ozono, S., Yao, W., Tobiume, M., Yamaoka, S., Kishigami, S., et al. (2019). CRISPR-mediated activation of endogenous BST-2/tetherin expression inhibits wild-type HIV-1 production. *Sci. Rep.* 9:3134. doi: 10.1038/ s41598-019-40003-z
- Zhao, N., Wang, G., Das, A. T., and Berkhout, B. (2017). Combinatorial CRISPR-Cas9 and RNA interference attack on HIV-1 DNA and RNA can lead to cross-resistance. *Antimicrob. Agents Chemother.* 61, e01486–e01487. doi: 10.1128/AAC.01486-17
- Zhu, W., Lei, R., Le Duff, Y., Li, J., Guo, F., Wainberg, M. A., et al. (2015). The CRISPR/Cas9 system inactivates latent HIV-1 proviral DNA. *Retrovirology* 12:22. doi: 10.1186/s12977-015-0150-z

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Xu, Peng, Zheng, Jin, Lu, Han, Fu, Chen and Wu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.