

SCIENTIFIC REPORTS



OPEN

The variances of Sp1 and NF- κ B elements correlate with the greater capacity of Chinese HIV-1 B'-LTR for driving gene expression

Di Qu¹, Chuan Li¹, Feng Sang², Qiang Li², Zhi-Qiang Jiang², Li-Ran Xu², Hui-Jun Guo², Chiyu Zhang¹ & Jian-Hua Wang¹

Received: 07 April 2016
Accepted: 15 September 2016
Published: 04 October 2016

The 5' end of HIV-1 long terminal repeat (LTR) serves as a promoter that plays an essential role in driving viral gene transcription. Manipulation of HIV-1 LTR provides a potential therapeutic strategy for suppressing viral gene expression or excising integrated provirus. Subtype-specific genetic diversity in the LTR region has been observed. The minor variance of LTR, particularly in the transcription factor binding sites, can have a profound impact on its activity. However, the LTR profiles from major endemic Chinese subtypes are not well characterized. Here, by characterizing the sequences and functions of LTRs from endemic Chinese HIV-1 subtypes, we showed that nucleotide variances of Sp1 core promoter and NF- κ B element are associated with varied LTR capacity for driving viral gene transcription. The greater responsiveness of Chinese HIV-1 B'-LTR for driving viral gene transcription upon stimulation is associated with an increased level of viral reactivation. Moreover, we demonstrated that the introduction of CRISPR/dead Cas9 targeting Sp1 or NF- κ B element suppressed viral gene expression. Taken together, our study characterized LTRs from endemic HIV-1 subtypes in China and suggests a potential target for the suppression of viral gene expression and a novel strategy that facilitates the accomplishment of a functional cure.

HIV-1 long terminal repeat (LTR) promoter plays an essential role in driving viral transcription and productive infection¹⁻³. The HIV-1 LTR is approximately 640 base pairs in length. It is segmented into the U3, R, and U5 regions. The U3 can be further subdivided according to transcription factor sites and the impact on LTR activity and viral gene expression. The enhancer regions usually consist of two highly conserved NF- κ B binding sites, and the core region of LTR contains three tandem GC-rich binding sites for the Sp transcription factor family members, and the immediate TATA box for the target of TATA binding protein (TBP)⁴⁻⁶.

The sequence and function of LTRs from many HIV-1 subtypes have been well characterized. Subtype-specific genetic diversity in the LTR has been observed, and the subtle variance in LTR can have a significant impact on promoter/enhancer activity, replication kinetics and the virulence of different HIV-1 subtypes^{2,7-12}. LTRs from HIV-1 B subtypes contain two NF- κ B binding sites, whereas HIV-1 LTR from subtype C, which is the predominant clade in Africa, contains an additional NF- κ B-like binding site and shows higher activity in activating viral gene transcription¹³⁻¹⁵. In contrast, CRF01_AE-LTR contains a single NF- κ B site and the upstream NF- κ B site usually presents in other subtypes is replaced by a GABP (GA binding protein) binding site. The NF- κ B to GABP conversion in CRF01_AE-LTR might account for the increased level of basal transcription and the reduced response upon TNF- α stimulation, and thus may likely be a contributor to latent infection^{9,10,16}. Additionally, CRF01_AE contains a different TBP binding site with the sequence TAAAA, whereas all other subtypes have a TATAA sequence; this single nucleotide variance reduces the efficiency of assembly of the TBP-TFIIB-TATA complex for the recruitment of RNA polymerase II^{2,16,17}.

Notably, the subtle alterations in LTRs may affect disease progression^{8,11}. An analysis in HIV-infected patients revealed that the C-to-T change at the Sp1 site III is positively correlated with disease progression¹¹.

¹CAS Key Laboratory of Molecular Virology and Immunology, Institut Pasteur of Shanghai, Chinese Academy of Sciences, Shanghai, China. ²Key laboratory of Prevention and Treatment with Traditional Chinese Medicine on Viral Infection Disease, the First Affiliated Hospital of Henan University of Traditional Chinese Medicine, Zhengzhou, China. Correspondence and requests for materials should be addressed to J.-H.W. (email: jh_wang@sibs.ac.cn)

The additional creation of a new NF- κ B binding site and the insertion of a duplicating transcription factor site for RBEIII in LTRs of subtype C have been observed in therapeutic failure patients from Brazil and Mozambique; this may partially explain the greater fitness of subtype C and the substitution of subtype B in those regions¹⁸. This LTR sequence alteration may affect the modulation of viral latency. A systematic genetic screen for HIV LTR promoter elements revealed that single point mutations of the Sp1 site III at the 4th guanine nucleotide and TATA box at the 2nd adenine increased the frequency of switching from latency to productive HIV-1 infection¹².

The indispensable roles of HIV-1 LTR in driving viral gene expression and maintenance of viral latency make it an attractive target for antiviral therapy. Several gene editing technologies, such as zinc finger nuclease (ZFN), transcription activator-like effector nuclease (TALEN) and clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) 9 (CRISPR/Cas9), have been used to remove the proviral HIV-1 DNA from the host cell genome by targeting its highly conserved LTRs^{19–28}.

Previous studies have mainly characterized LTRs from HIV-1 B, C and AE subtypes that are more prevalent in Europe and America^{2,29}, whereas the LTR and viral profiles from endemic Chinese HIV-1 subtypes have not been carefully characterized³⁰. Several HIV-1 subtypes are co-circulating among various high-risk groups in China. Subtypes B' (Thailand's variant of subtype B) and CRF01_AE were introduced into China from Thailand, and subtype C was primarily transmitted from India^{29,31}. CRF07_BC and CRF08_BC are two Chinese subtypes that most likely have originated from Yunnan province of China³². In the early stages of HIV epidemic, subtype B' was mainly circulating among former plasma donors in Central China (mostly in Henan province), CRF01_AE was predominant among heterosexuals and injection drug users (IDUs), and CRF07_BC and CRF08_BC were prevalent among IDUs. In recent years, various HIV-1 subtypes have rapidly disseminated among different high-risk groups, an increased prevalence of CRF01_AE and CRF07_BC has been observed among men who have sex with men (MSM).

In this study, we cloned LTRs from Chinese specific subtypes B', CRF07/08_BC, and CRF01_AE, and compared their activities with those of global HIV-1 subtypes. We found that the HIV-1 B'-LTR sequence has the strongest capacity for driving viral gene expression. Sequence analysis revealed that, in addition to the variance in NF- κ B binding sites, the variance in Sp1 core promoter determined its activity. Finally, we demonstrated that the Sp1 core promoter nucleotides and the NF- κ B binding sites can serve as targets for CRISPR/Cas9 gene editing, and the depletion of these sites led to suppressed viral gene expression.

Results

The 5'-LTR from HIV-1 B' subtype has the strongest capacity for driving gene expression. The 5'-LTR of HIV-1 has been shown to maintain subtype-specific genetic diversity^{7,10,12,18}. To confirm this in prevalent HIV-1 subtypes in China, we downloaded 300 LTR sequences of circulating HIV-1 strains including B' (54 sequences), CRF07_BC (42 sequences), CRF08_BC (36 sequences) and CRF01_AE (168 sequences) from the Los Alamos HIV database (<http://www.hiv.lanl.gov/content/index>, on April 21, 2014) (Table S1), and aligned these sequences to discern genetic variability. For comparison, LTR sequences from worldwide distributed subtype B (300 sequences) and subtype C (300 sequences) were from the same database and aligned by Mega software using Clustal W parameters. As expected, subtype-specific genetic diversity in the LTR was observed as shown in the phylogenetic tree analysis (Fig. S1). The 5'-LTR consensus sequences were highlighted (Fig. S2). CRF07_BC and CRF08_BC show mosaic patterns in a subtype C genomic backbone inserted by several subtype B' fragments in *gag*, *pol* and *env* genes^{33,34}. Because they have the same LTR consensus and are thus combined for the latter analysis. The LTR of B' and B have two NF- κ B binding sites; the LTR of CRF07/08_BC and CRF07/08_C have two NF- κ B binding sites and one NF- κ B binding-like site, respectively; the LTR of CRF01_AE has a single NF- κ B enhancer. Besides the number of NF- κ B binding sites, the other transcription factor elements such as TCF-1 (T-cell-specific transcription factor 1), Sp1 and TAR (Trans-activating region) also showed marked variations.

Previous studies showed that the minor variance in LTR can have a significant impact on its activity^{2,8,9,11,12}. To investigate the potential difference in LTR activity from HIV-1 subtypes predominant in China, we synthesized these consensus LTR sequences and inserted them into a luciferase-based reporter vector pGL3 (Fig. 1a). These plasmids containing various LTR sequences were transfected into HEK293T cells and then the luciferase activities were measured. B'-LTR showed the strongest capacity for driving basal gene expression and elevated gene expression upon stimulation with either TNF- α or PMA/Ionomycin. In comparison, CRF07/08_BC-, C- and B-LTR displayed a moderate response to the same stimuli, whereas CRF01_AE-LTR exhibited the least activity under these experimental conditions (Fig. 1b,c).

To further investigate the activity of subtype-specific LTRs variants, we inserted synthesized LTR consensus nucleotides into pHIV-eGFP and replaced the 5'-LTR of this plasmid (Fig. 1d), and then co-transfected HEK293T cells with the expression plasmid for vesicular stomatitis virus G (VSV-G) protein. Culture supernatant containing the pseudotyped single-cycle infectious HIV-eGFP/VSV-G was used to infect Jurkat CD4⁺ T cells. The pseudotyped virus containing HIV-1 B'-LTR from Chinese isolates showed the most activity for driving viral gene transcription upon stimulation with TNF- α (Fig. 1e). Taken together, these data confirmed the variation in the function of subtype-specific LTRs, and that the 5'-LTR from HIV-1 B' has the strongest capacity for driving gene expression.

The (-138)-0 fragment containing the enhancer/core promoter sequences determines the difference in LTR activity. Multiple nucleotide alterations distributed throughout different transcription factor elements have been observed, and the most noted alterations are mainly found in the (-138)-0 fragment containing the enhancer/core promoter sequences (Fig. S2). To investigate whether variances in the (-138)-0 region determine the differences in LTR activity, we synthesized the (-138)-0 region of B'-LTR and used it to replace the homologous fragment in CRF01_AE-LTR. The (-452)-(-295) fragment, (-294)-(-139) fragment containing

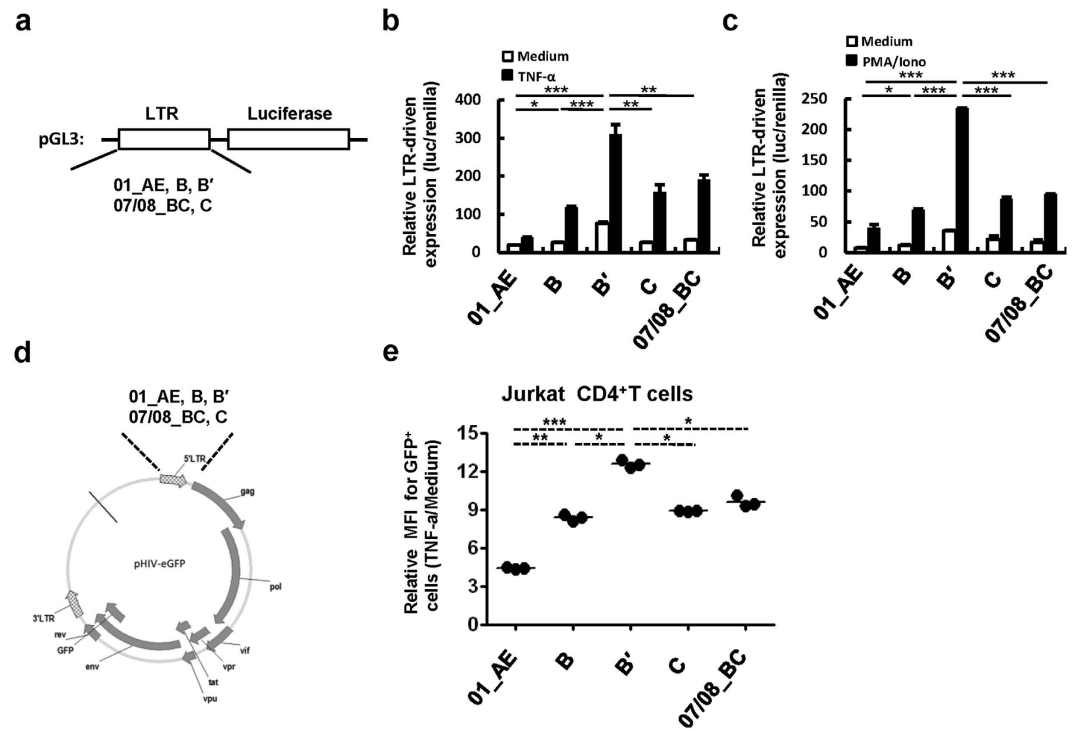


Figure 1. The 5'-LTR from HIV-1 B' subtype demonstrates the strongest capacity for driving gene expression. (a) Schematic construction of pGL3 containing LTR-consensus-driven luciferase reporter. (b,c) Assay for LTR-driven gene expression upon stimulation. HIV-1 LTR-luciferase-reporter plasmids and pRenilla-luc-TK were co-transfected into HEK293T cells for 24 h and stimulated with TNF- α (b) or PMA/Ionomycin (c) for additional 24 h. The cells were harvested, and luciferase activity was measured and calculated by using dual luciferase reporter assay. (d) Schematic construction of pHIV-eGFP containing LTR consensus. (e) Infection of Jurkat CD4⁺ T cells. Jurkat CD4⁺ T cells were infected with pHIV-eGFP/VSV-G (1 ng p24^{gag} amount viruses for 10⁶ cells) for 24 h, with or without TNF- α (2 ng/ml) stimulation, and viral infection was determined by measuring GFP expression with flow cytometry, and the mean fluorescence intensity (MFI) was calculated. *P < 0.05, **P < 0.01 and ***P < 0.001 are considered significant difference in an unpaired t-test.

TCF-1 element variance and 1–182 fragment containing TAR variance were also synthesized for homologous replacement. The replacement with the (-138)-0 but not the (-452)-(-295), (-294)-(-139) or 1–182 region of B'-LTR significantly increased CRF01_AE-LTR activity upon TNF- α or PMA/Ionomycin stimulation (Fig. 2a). In contrast, when the (-138)-0 region of B'-LTR was replaced with homologous segment of CRF01_AE, B'-LTR significantly reduced its activity (Fig. 2b). Similarly, B-LTR acquired significantly enhanced activity after homologous introduction of B'-LTR (-138)-0 region (Fig. 2c). Conversely, B'-LTR significantly reduced its activity after the (-138)-0 region was homologously replaced with that of the B-LTR.

The 1–182 fragment contains TAR element (nt +1 to +59), which binds to HIV-1 Tat protein to allow transcription elongation. To further confirm that the 1–182 region of HIV-1 LTR is dispensable for the observed differences in LTR activity, we co-transfected HEK293T cells with HIV-1-tat expressing plasmid (pCMV-tat, cloned from HIV-1 NL4-3, B subtype) and plasmids containing various LTRs, to examine the variation in Tat-driven *trans*-activation. As expected, different LTRs showed similar response upon Tat activation (Fig. 2d), indicating that variation of LTR activity was not attributable to Tat-driven *trans*-activation and transcription elongation, and that the (-138)-0 region containing enhancer elements for transcription initiation was the key element responsible for the functional variability associated with LTR variations. Taken together, these data demonstrate that the (-138)-0 nucleotides in B'-LTR confer the variation in LTR' ability for driving gene expression.

The variances in Sp1 core promoter are associated with HIV-1 B'-LTR activity for driving gene expression.

We then concentrated on the (-138)-0 nucleotides of LTRs to investigate how its diversity affected gene expression using a reporter system. This (-138)-0 region contains multiple elements for transcription factor binding, including the NF- κ B enhancer and Sp1 core promoter. The AE subtype contains only one NF- κ B binding site, whereas B' and B contain two (Fig. S2). Another apparent difference is that B' and B each contains a TATA box in the core promoter, whereas CRF01_AE contains a TAAA sequence in this area (Fig. 3a). Indeed, the number of NF- κ B binding sites influence subtype specific HIV-1 LTR activity, as the depletion of one NF- κ B element from B'-LTR, or the insertion of an additional NF- κ B element to CRF01_AE-LTR significantly altered LTR activity upon TNF- α or PMA/Ionomycin stimulation (Fig. 3b,c). To further prove the role of NF- κ B binding sites on HIV-1 LTR capacity for driving viral gene expression, dCas9/gRNA system was used to silence NF- κ B

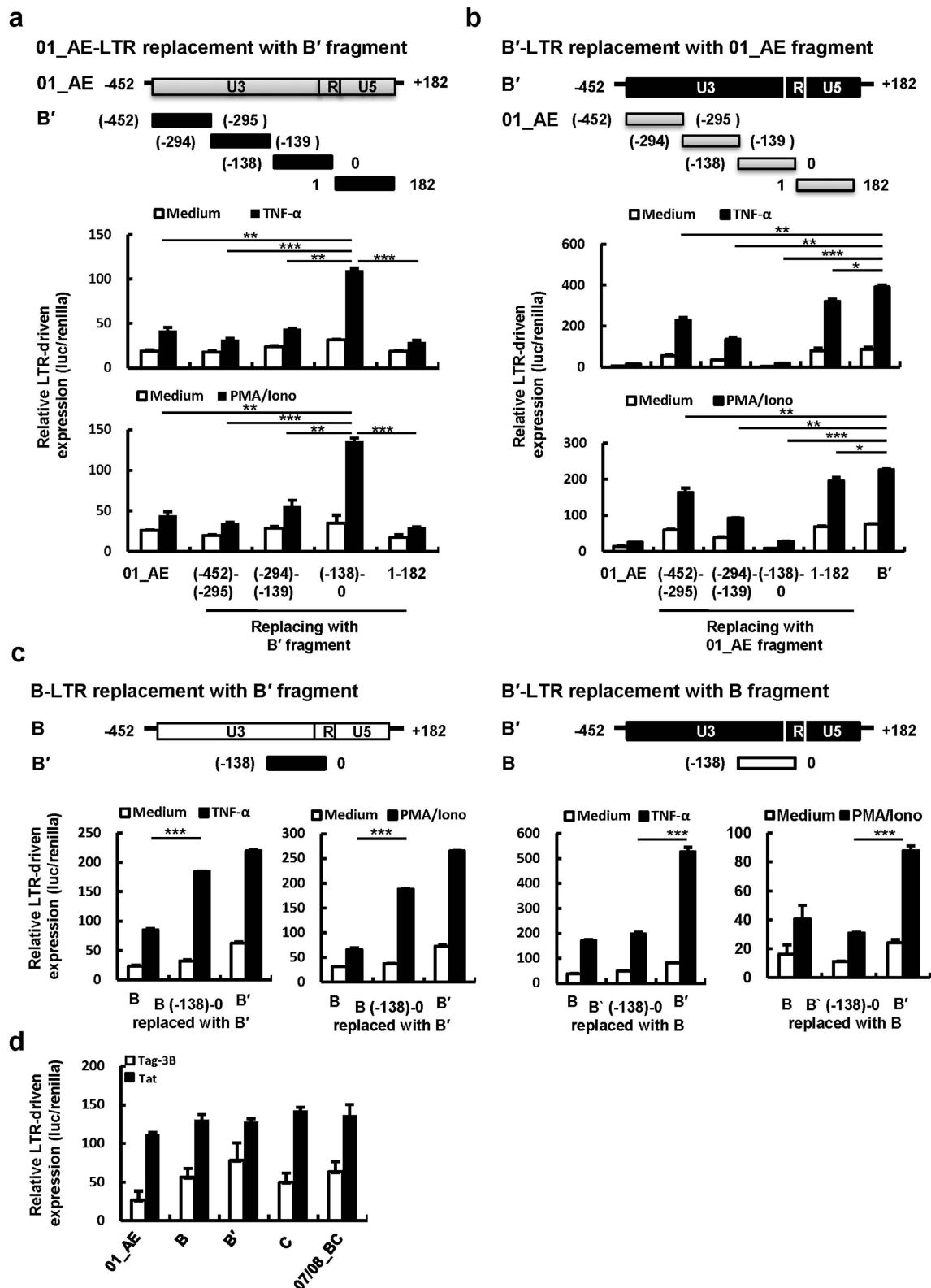


Figure 2. The (-138)-0 fragment determines the efficiency of LTR activity. (a–c) Assay for LTR-driven gene expression in transfected HEK293T cells upon TNF- α or PMA/Ionomycin stimulation. The homologous fragments were replaced as depicted in schematic graphs. Dual luciferase reporter system was used. (e) Investigation on Tat-driven *trans*-activation. HIV-1-*tat* expressing plasmid and plasmids containing various LTRs were co-transfected into HEK293T cells for 48 h and luciferase activities were measured. Data are mean \pm SD. Results are representative of four independent experiments (a–c) and of three independent experiments (d). * $P < 0.05$, ** $P < 0.01$ and *** $P < 0.001$ are considered significant difference in an unpaired t-test.

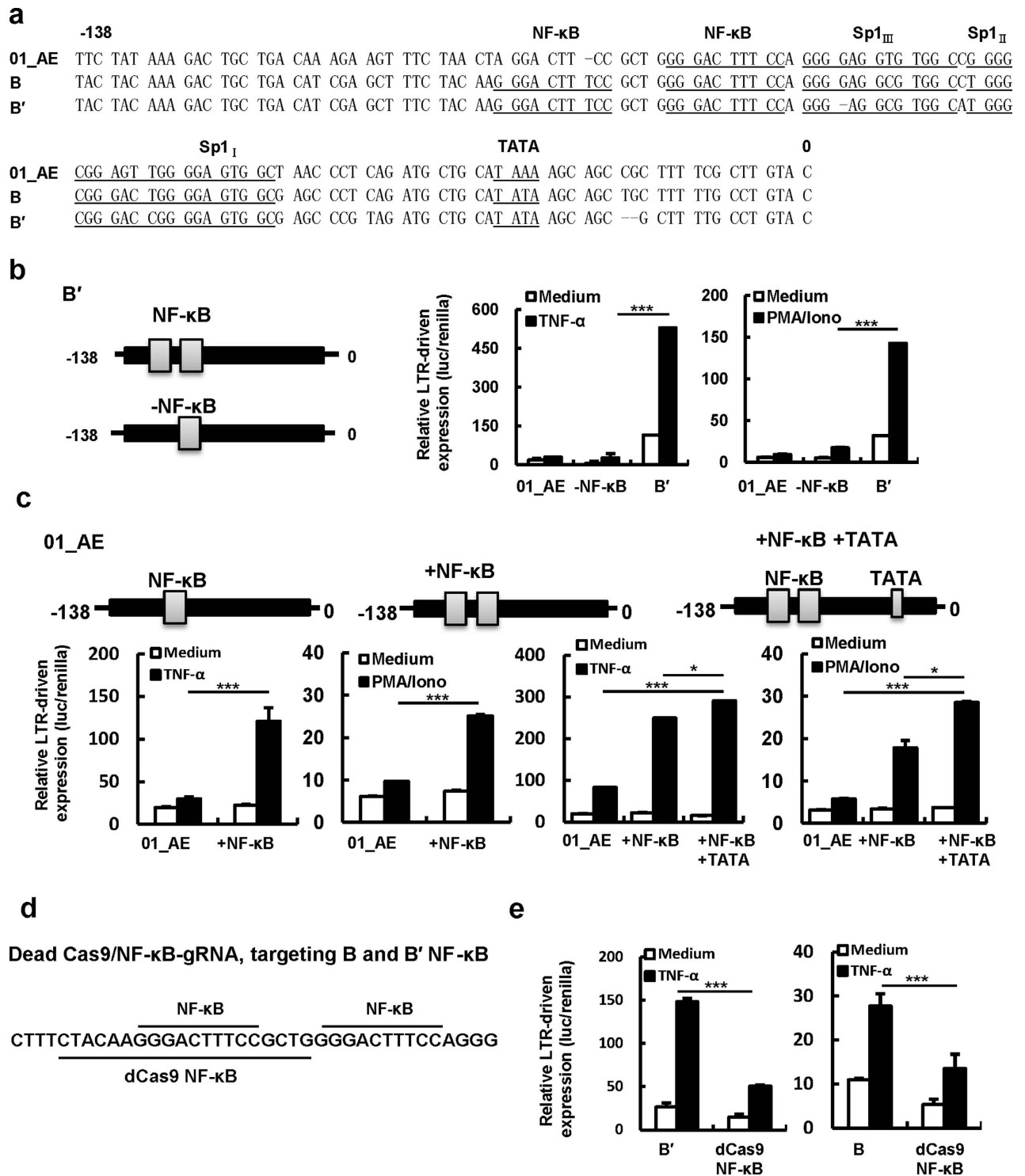


Figure 3. The frequency of NF-κB binding sites and the TATA regulatory element are essential for HIV-1-LTR activity. (a) Alignment of the core elements in (-138-0) fragment of CRF_01AE-, B- and B'-LTR. (b,c) Assay for LTR-driven gene expression. The first NF-κB binding site was depleted from B'-LTR (-NF-κB) (b). A NF-κB binding site was added (+NF-κB), or a NF-κB binding site and the TATA box were added (+NF-κB +TATA) to 01AE-LTR (c). (d) Design of a gRNA targeting the first NF-κB binding site of B'. (e) Assay for LTR-driven gene expression. HIV-LTR-specific dCas9/gRNAs and LTR-reporter plasmids were co-transfected into HEK293T cells, and cells were harvested for quantifying LTR activity after being stimulated with TNF-α for 24 h. Data are mean ± SD. Results are representative of four independent experiments. * $P < 0.05$, ** $P < 0.01$ and *** $P < 0.001$ are considered significantly different.

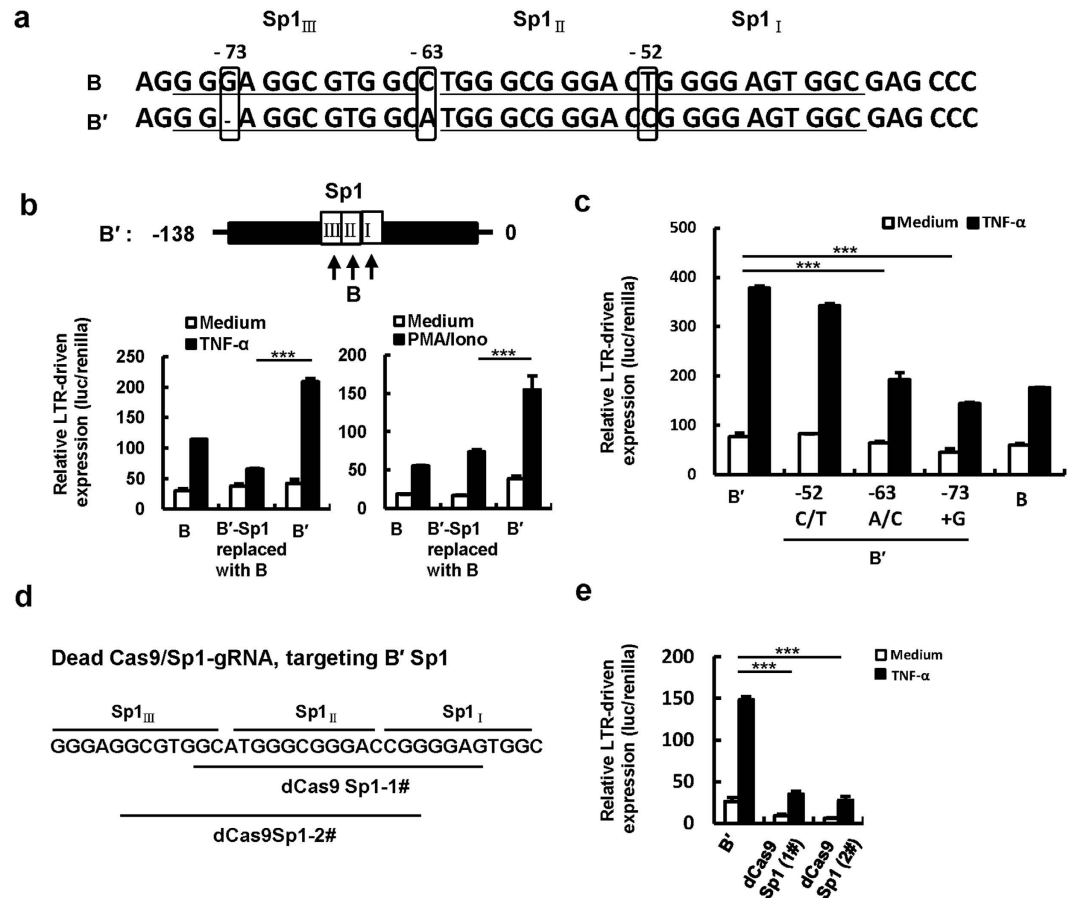


Figure 4. The variances in Sp1 element confer HIV-1 B'-LTR greater ability for driving gene expression and the introduction of CRISPR/dCas9 targeting Sp1 core elements suppressed viral gene expression. (a) Alignment of Sp1 element sequences. (b) The Sp1 elements of B'-LTR were replaced with B homologous fragments, and LTR activity was quantified after TNF- α or PMA/Ionomycin stimulation. (c) The varied bases in HIV- B'-LTR Sp1 element were mutated and the capacity for driving gene expression was measured. (d) Design of gRNAs targeting Sp1 transcription element of LTR. (e) Assay for LTR-driven gene expression. HIV-LTR-specific dCas9/gRNAs and LTR-reporter plasmids were co-transfected into HEK293T cells, and cells were harvested for quantifying LTR activity after stimulated with TNF- α for 24 h. Data are mean \pm SD. Results are representative of three independent experiments. ***P < 0.001 are considered significant difference.

activation (Fig. 3d). We constructed NF- κ B-gRNA first, and then co-transfected it with either B'- or B-LTR into HEK293T cells. As expected, the LTR activity was significantly reduced (Fig. 3e). Moreover, the addition of TATA box into CRF01_AE-LTR further enhanced LTR activity (Fig. 3c).

B'-LTR showed greater capacity for driving gene expression than B-LTR (Fig. 1b,c,e). Although B- and B'-LTR show high conservation in (-138)-0 regions (Fig. 3a), there are subtle variances between them in the Sp1 core promoter sequences. Compared with B'-LTR, cytosine is changed to adenine at -63 position, thymine is changed to cytosine at -52 position, and one guanine is depleted at -73 position in the Sp1_{III} and Sp1_{II} regions of B-LTR (Fig. 4a).

The replacements of B'-LTR (-138)-0 region or Sp1_{III} and Sp1_{II} elements with B homologous fragments significantly lowered its activity (Fig. 4b). To determine which base in the Sp1 element contributes to the stronger ability of B'-LTR, we mutated specific bases in B'-LTR, and found that the replacement of adenine at -63 position with cytosine and the adding of guanine at -73 position, but not the replacement of cytosine at -52 position with thymine, significantly impaired B'-LTR responses to stimulation (Fig. 4c), demonstrating the uniqueness of B'-LTR Sp1_{III} region is the key for enhanced activity. Taken together, these data prove that the variances in Sp1 elements contribute to the stronger capacity of HIV-1 B'-LTR to induce activation.

The introduction of CRISPR/dCas9 targeting Sp1 core elements suppressed viral gene expression.

The indispensable roles of HIV-1 LTR in driving viral gene expression make it an attractive target for designing new antiviral strategy. The CRISPR/Cas9 system is known to play a major role in adaptive immune responses against foreign pathogens in prokaryotes, and recently, it has been adopted to disrupt HIV-1 entry coreceptors C-C chemokine receptor 5 (CCR5), or excise HIV-1 provirus in order to suppress viral gene expression and eradicate viral latency^{19–26,28,35,36}. To further prove the role of Sp1 core promoter on HIV-1 LTR driving viral gene expression in Chinese HIV-1 B' subtype, dCas9/gRNA system was used to inactivate the activation of Sp1

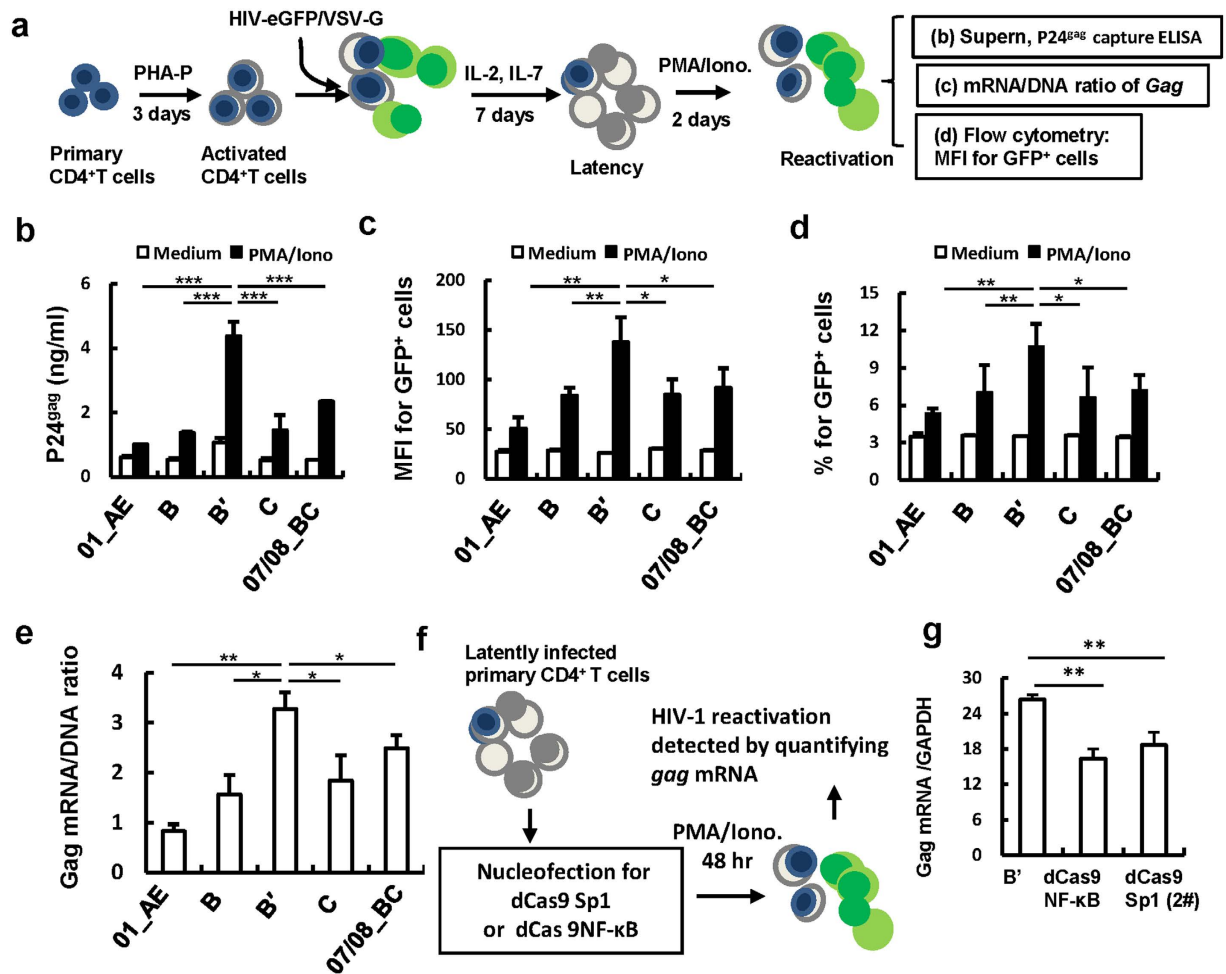


Figure 5. HIV-1 B' virus shows greater capacity to induce reactivation from latency. (a) Establishment of HIV-1 latently infected primary CD4⁺ T cells. PHA-P-activated CD4⁺ T cells were infected with pseudotyped HIV-eGFP/VSV-G, and cells were cultured in the presence of IL-2 and IL-7 for 7 days. Viral latency was reactivated by PMA/Ionomycin stimulation, and newly synthesized viruses released in culture supernatant were detected with p24^{gag} capture ELISA (b). Viral reactivation from latency was also monitored by measuring GFP expression with flow cytometry (c,d), or quantifying the relative Gag mRNA/DNA ratio with quantitative (RT-) PCR (e). (f,g) The HIV-1 latently infected CD4⁺ T cells can be targeted by the Cas9/gRNAs system to suppress reactivation. dCas9-Sp1 or dCas9- NF-κB were nucleofected into HIV-1 latently infected primary CD4⁺ T cells (f), and viral reactivation from latency upon PMA/Ionomycin stimulation was quantified with quantitative (RT-) PCR to measure the Gag mRNA/DNA ratio (g).

transcription elements²⁸. Two gRNAs were designed to target Sp1 transcription elements in LTRs (Fig. 4d). Transfection of these HIV-LTR-specific dCas9/gRNAs into HEK293T cells silenced HIV-1 B'-LTR-driven gene expression (Fig. 4e). Together, these data demonstrate that Sp1 core elements can be targeted by the CRISPR/Cas9 system for blocking HIV-1 gene expression.

Virus containing HIV-1 B'-LTR shows the greatest responsiveness to experimental reactivation from latency. Silencing HIV-1 LTR-driven transcription of proviral DNA is the key for maintaining HIV-1 post-integration latency, but the process is reversible³⁷⁻⁴⁰. To investigate whether LTR variation affects viral reactivation from latency, we used pseudotyped HIV-eGFP/VSV-G viruses containing variable synthesized LTR consensus nucleotides to infect activated primary CD4⁺ T cells to establish viral latency first, and then used PMA/Ionomycin to stimulate cells for viral reactivation (Fig. 5a). HIV-1 B'-LTR possessed the strongest capacity for driving viral reactivation from latency, as measured by either quantifying newly synthesized virions in culture supernatant or detecting Gag-GFP expression (Fig. 5b-d). And HIV-eGFP/VSV-G containing HIV-1 B'-LTR displayed the highest mRNA/DNA ratio in reactivated primary CD4⁺ T cells (Fig. 5e). Moreover, these LTRs in latent infected cells could be targeted by the dCas9/gRNAs system and resulted in a blockade of viral reactivation from latency (Fig. 5f,g). Together, these data demonstrate that HIV-1 B'-LTR possesses the strongest ability in reactivating viral transcription from latency, and the predominant LTR sequences that promote this activity may provide key targets for gene editing to achieve modulation of viral latency.

Discussion

In this study, we characterized the sequences and functions of LTRs from major endemic HIV-1 subtypes in China, and discovered that Chinese HIV-1 B'-LTR consensus sequences possessed the strongest capacity for driving viral gene expression. Specifically, the nucleotide variances of Sp1 core promoter were associated with LTR capacity for driving viral gene transcription. Additionally, the higher responsiveness of B'-LTR to external stimulation determined the increased reactivation from latency. This study reveals a small number of nucleotides that have profound influence on LTR activity, and thus permitting the identification of specific target sites for manipulation to suppress viral gene expression.

LTR variances have been known to significantly impact on promoter/enhancer activity, viral replication kinetics, virulence and dissemination^{2,8,9}. The mutation of Sp1 element in LTR, particularly the point mutation of Sp1 site III at the 4th guanine, has been proven in multiple studies to significantly impair Tat-mediated transactivation^{6,12,41–43}. Whether the same is true for HIV-1 subtypes prevalent in China has not been carefully studied. HIV-1 subtype B' is mainly circulating among former plasma donors in Central China. Its Sp1 element includes two nucleotide mutations, which appeared to have conferred B'-LTR the greatest capacity for driving gene expression upon stimulation with TNF- α or PMA/Ionomycin. The increased basal transcription and enhanced sensitivity to TNF- α stimulation suggest that the B' subtype is less likely to become latent. Also, B'-LTR's strong activity and responsiveness to stimulation may confer this virus higher replication fitness in patients. Therefore, more attention should be paid to this subtype during HIV surveillance.

An increased prevalence of CRF01_AE was recently observed among MSM with more rapid disease progression. Multiple cellular and viral factors have been found to be correlated with rapid disease progression in CRF01_AE infection⁴⁴. The NF- κ B to GABP conversion in LTR in CRF01_AE confers this virus more fitness to out-compete other HIV subtypes in new infections¹⁶, although this conversion reduces viral responses to TNF- α -stimulation¹⁰. Whether the enhanced fitness is a direct result of the NF- κ B to GABP conversion and a correlate of rapid disease progression in CRF01_AE infection are yet to be confirmed.

The evolution of LTR may be compartmentalized in tissues, and varied among individual patients. One study investigated the phylogenetic relationship among HIV-1 proviral LTR quasispecies existed in an AIDS patient's postmortem tissues, including lymph node, spleen, lung, dorsal root ganglion and spinal cord, as well as PBMCs obtained 2 months prior death. The study found that the spinal cord and dorsal root ganglion harbored HIV-LTRs that were genetically distinct from that existed in other organs. Thus, LTR variance may reflect the adaptation to transcription factors in specific tissue environments⁴⁵.

The indispensable roles of HIV-1 LTR in driving viral gene expression make it an attractive target for combating infection. The conserved regions in LTR have been targeted by gene editing techniques including ZFN, TALENs and CRISPR/Cas9 for editing or removing integrated proviral DNA to archive a functional cure^{19–28}. We demonstrated that the Sp1 core element could be targeted by CRISPR/Cas9 targeting system for silencing gene expression. Our results are in agreement with a recent study demonstrating that the Sp1 binding sites could be targeted by the host restrictive factor tripartite motif-containing 22 for silencing HIV-1 transcription⁴⁶.

Taken together, our study characterized the property of LTRs from major endemic HIV-1 subtypes in China, and confirmed the critical role of Sp1 and NF- κ B elements for LTR activity. These findings permit the identification of potential target sites for manipulation to suppress viral gene expression.

Methods

Ethics statement. Peripheral blood mononuclear cells from health donors were purchased from Shanghai Blood Center, Shanghai, China. The usage of human cells, and the related methods and experimental protocols have been approved by the Medical Ethics Review Committee of Institut Pasteur of Shanghai, Chinese Academy of Sciences. All experiments were performed in accordance with relevant national guidelines and regulations.

Cells. Peripheral blood mononuclear cells (PBMCs) were separated from buffy coats obtained from healthy donors using Ficoll-Paque density gradient centrifugation, and CD4⁺ T cells were further purified from PBMCs using anti-CD4 antibody-coated magnetic beads (Miltenyi Biotec). Resting CD4⁺ T cells were treated with 5 μ g/ml phytohaemagglutinin-P (PHA-P) (Sigma) for 3 days in the presence of IL-2 (20 IU/ml) for activation. The HEK293T cells were kindly provided by Dr. Li Wu (The Ohio State University, USA). The cells were maintained in Dulbecco's modified Eagle's medium (Hyclone) containing 10% fetal bovine serum (Hyclone), 100 U/ml penicillin and 100 μ g/ml streptomycin.

Plasmid DNA. HIV-1 backbone-based vector pHIV-eGFP and vesicular stomatitis virus G (VSV-G) expression plasmids were gifts from Dr. Li Wu (The Ohio State University, USA). Plasmids containing B'-, B-, CRF07/08_BC-, C-LTR and CRF01_AE-LTR, were constructed by inserting synthesized consensus LTR sequences into the multiple-cloning site of luciferase-based reporter vector pGL3 (Promega). Recombinants and mutations were constructed based on LTR-luciferase-reporter plasmids using ClonExpress II One Step Cloning Kit and Mut Express II Fast Mutagenesis Kit (Vazyme). To insert synthesized LTR consensus nucleotides into pHIV-eGFP vector, two restriction enzyme sites (XbaI and ClaI) were constructed into pHIV-eGFP vector using Mut Express II Fast Mutagenesis Kit (Vazyme). Consensus LTR sequences were inserted into pHIV-eGFP vector through these two restriction enzyme sites.

The construction of CRISPR/dCas9. The pX330 plasmid containing two expression cassettes, hSpCas9 and the chimeric guide RNA and Dead Cas9 plasmid were gifts from Dr. Huan-Zhang Zhu (Fudan University, China). The CRISPR/dCas9 system targeting transcription modulators has been previously used to regulate HIV-1 gene expression^{47,48}. One target sequence for NF- κ B binding site (5'-CTACAAGGGACTTTCCGCTG-3') and two target sequences for the Sp1 transcription element (5'-GCATGGGCGGGACCGGGGAG-3');

5'-GCGTGGCATGGGCGGGACCG-3') were cloned into pX330 plasmids according to the genome engineering toolbox⁴⁹. These target sequences were tested using the <http://www.crispr.mit.edu> tool to check for off-target effects. The scores showed that all of three designed targets had high quality guide. BLAST search against human whole genome confirmed the non-specific binding of these designed gRNAs to other host genes.

Virus Stock. Lipofectamine 2000 (Invitrogen) transfection of HEK293T cells was used to generate virus stock according to the manufacturer's instructions. Pseudotyped single-cycle pHIV-eGFP/VSV-G viruses were generated by co-transfection with different LTR-containing pHIV-eGFP vectors and VSV-G expression plasmids. Supernatants harvested from transfected cells containing viral particles were titrated with a p24^{gag} capture ELISA.

Dual luciferase reporter assay. HEK293T cells were grown in a 24-well cell culture plate. Transfection was performed using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. HIV-1 LTR-luciferase-reporter plasmids and pRenilla-luc-TK were used. At 6 hr post-transfection, medium containing a mixture of plasmids and transfection reagent was replaced with fresh DMEM supplemented with 10% FBS. At 24 h post-transfection, cells were treated with TNF- α (50 ng/ml) or PMA (20 nM)/Ionomycin (1.5 μ M). The cells were harvested for luciferase activity after 24 hr of treatment. Briefly, cells were washed with phosphate buffered saline (PBS) and homogenized with 100 μ l lysis buffer (Promega); lysate (10 μ l) was mixed with luciferase assay buffer (10 μ l) and the firefly luciferase activity was measured first on a Veritas luminometer (Turner BioSystems). The plate was then removed from the luminometer and mixed with Stop & Glo Reagent (10 μ l) and Renilla luciferase activity was measured. Data recorded on the luminometer were analyzed.

The establishment of viral latency in primary CD4⁺ T cells. PHA-P-activated primary CD4⁺ T cells (0.5×10^6) were incubated with pHIV-eGFP/VSV-G (25 ng p24^{gag}) for 5 h at 37°C. Cells were further cultured for 7 days in the presence of IL-7 cytokine for the establishment of viral latency^{50,51}. Viral reactivation was induced by PMA (20 nM)/Ionomycin (1.5 μ M) stimulation and detected by analyzing Gag-GFP expression with a Fortessa flow cytometer (BD Pharmingen); or alternatively, HIV-1 Gag mRNA/DNA production quantified with Real Time-(RT-) PCR, or HIV-1 p24 production with a p24^{gag} capture ELISA.

Real time-(RT-) PCR analysis. Total cellular RNA from infected cells was extracted with TRIzol reagent (Life Technologies) and then reverse transcribed to cDNA with ReverTra Ace qPCR RT Master Mix with gDNA Remover Kit (TOYOBO). Genome DNA (gDNA) from infected cells was extracted using QIAamp MiniElute DNA Kit (Qiagen) according to the manual. RT-PCR was performed using the Thunderbird SYBR qPCR Mix (TOYOBO) as described previously with an initial denaturation step for 10 min at 95°C, amplification with 40 cycles of denaturation (95°C, 30 s), annealing (55°C, 30 s), and extension (72°C, 30 s), followed by a final extension at 72°C for 6 min. The primers used were as follows: GAPDH: forward, 5'-GGGAAATCGTGCCTGACAT-3' and reverse, 5'-GTCAGGCAGCTCGTAGCTCTT-3'; Gag: Forward, 5'-GTGTGGAAAATCTCTAGCAGTGG-3' and Reverse, 5'-CGCTCTCGACCCATCTC-3'.

Phylogenetic analysis. For phylogenetic analysis, the complete 5'-LTR nucleotide sequences of HIV-1 subtypes endemic in China and elsewhere in the world were obtained from the Los Alamos HIV database (<http://www.hiv.lanl.gov/content/index>, on April 21, 2014) (Table S1). Sequence alignment was performed with the ClustalW program in the Molecular Evolutionary Genetics Analysis (MEGA) 5 software. A phylogenetic tree was constructed with the Neighbor-Joining method, using MEGA. The bootstrap test was performed with 500 replications.

Statistics. Statistical analysis was performed using an unpaired t test with SigmaStat 2.0 (Systat Software, San Jose, CA, USA).

References

- Schiralli Lester, G. M. & Henderson, A. J. Mechanisms of HIV Transcriptional Regulation and Their Contribution to Latency. *Mol Biol Int.* **2012**, 614120, doi: 10.1155/2012/614120 (2012).
- Jeeninga, R. E. *et al.* Functional differences between the long terminal repeat transcriptional promoters of human immunodeficiency virus type 1 subtypes A through G. *J Virol.* **74**, 3740, doi: 10.1128/JVI.74.8.3740-3751 (2000).
- Tripathy, M. K., Abbas, W. & Herbein, G. Epigenetic regulation of HIV-1 transcription. *Epigenomics.* **3**, 487, doi: 10.2217/epi.11.61 (2011).
- Hiscott, J., Kwon, H. & Genin, P. Hostile takeovers: viral appropriation of the NF-kappaB pathway. *J Clin Invest.* **107**, 143, doi: 10.1172/JCI11918 (2001).
- Kilareski, E. M., Shah, S., Nonnemacher, M. R. & Wigdahl, B. Regulation of HIV-1 transcription in cells of the monocyte-macrophage lineage. *Retrovirology.* **6**, 118, doi: 10.1186/1742-4690-6-118 (2009).
- Berkhout, B. & Jeang, K. T. Functional roles for the TATA promoter and enhancers in basal and Tat-induced expression of the human immunodeficiency virus type 1 long terminal repeat. *J Virol.* **66**, 139 (1992).
- van Opijnen, T. *et al.* Human Immunodeficiency Virus Type 1 Subtypes Have a Distinct Long Terminal Repeat That Determines the Replication Rate in a Host-Cell-Specific Manner. *J Virol.* **78**, 3675, doi: 10.1128/jvi.78.7.3675-3683.2004 (2004).
- Duverger, A. *et al.* An AP-1 binding site in the enhancer/core element of the HIV-1 promoter controls the ability of HIV-1 to establish latent infection. *J Virol.* **87**, 2264, doi: JVI.01594-12 (2013).
- Montano, M. A., Nixon, C. P. & Essex, M. Dysregulation through the NF-kappaB enhancer and TATA box of the human immunodeficiency virus type 1 subtype E promoter. *J Virol.* **72**, 8446 (1998).
- van der Sluis, R. M., Pollakis, G., van Gerven, M. L., Berkhout, B. & Jeeninga, R. E. Latency profiles of full length HIV-1 molecular clone variants with a subtype specific promoter. *Retrovirology.* **8**, 73, doi: 10.1186/1742-4690-8-73 (2011).
- Nonnemacher, M. R., Irish, B. P., Liu, Y., Mauger, D. & Wigdahl, B. Specific sequence configurations of HIV-1 LTR G/C box array result in altered recruitment of Sp isoforms and correlate with disease progression. *J Neuroimmunol.* **157**, 39, doi: 10.1016/j.jneuroim.2004.08.021 (2004).

12. Miller-Jensen, K., Skupsky, R., Shah, P. S., Arkin, A. P. & Schaffer, D. V. Genetic selection for context-dependent stochastic phenotypes: Sp1 and TATA mutations increase phenotypic noise in HIV-1 gene expression. *PLoS Comput Biol.* **9**, e1003135, doi: 10.1371/journal.pcbi.1003135 (2013).
13. Montano, M. A. *et al.* Divergent transcriptional regulation among expanding human immunodeficiency virus type 1 subtypes. *J Virol.* **71**, 8657 (1997).
14. Montano, M. A. *et al.* Elevated tumor necrosis factor- α activation of human immunodeficiency virus type 1 subtype C in Southern Africa is associated with an NF- κ B enhancer gain-of-function. *J Infect Dis.* **181**, 76, doi: 10.1086/315185 (2000).
15. Roof, P. *et al.* Differential regulation of HIV-1 clade-specific B, C, and E long terminal repeats by NF- κ B and the Tat transactivator. *Virology.* **296**, 77, doi: 10.1006/viro.2001.1397 (2002).
16. Verhoef, K., Sanders, R. W., Fontaine, V., Kitajima, S. & Berkhout, B. Evolution of the human immunodeficiency virus type 1 long terminal repeat promoter by conversion of an NF- κ B enhancer element into a GABP binding site. *J Virol.* **73**, 1331 (1999).
17. Gao, F. *et al.* The heterosexual human immunodeficiency virus type 1 epidemic in Thailand is caused by an intersubtype (A/E) recombinant of African origin. *J Virol.* **70**, 7013 (1996).
18. Boulosa, J. *et al.* Genetic diversity in HIV-1 subtype C LTR from Brazil and Mozambique generates new transcription factor-binding sites. *Viruses.* **6**, 2495, doi: 10.3390/v6062495 (2014).
19. Qu, X. T. Zinc-finger-nucleases mediate specific and efficient excision of HIV-1 proviral DNA from infected and latently infected human T cells. *Nucleic Acids Res.* **41**, 7771, doi: 10.1093/nar/gkt571 (2013).
20. Hu, W. *et al.* RNA-directed gene editing specifically eradicates latent and prevents new HIV-1 infection. *Proc Natl Acad Sci USA* **111**, 11461, doi: 10.1073/pnas.1405186111 (2014).
21. Ebina, H., Misawa, N., Kanemura, Y. & Koyanagi, Y. Harnessing the CRISPR/Cas9 system to disrupt latent HIV-1 provirus. *Sci Rep.* **3**, 2510, doi: 10.1038/srep02510 (2013).
22. Dampier, W., Nonnemacher, M. R., Sullivan, N. T., Jacobson, J. M. & Wigdahl, B. HIV Excision Utilizing CRISPR/Cas9 Technology: Attacking the Proviral Quasispecies in Reservoirs to Achieve a Cure. *MOJ Immunol.* **1**, doi: 10.15406/moji.2014.01.00022 (2014).
23. Zhu, W. *et al.* The CRISPR/Cas9 system inactivates latent HIV-1 proviral DNA. *Retrovirology.* **12**, 22, doi: 10.1186/s12977-015-0150-z (2015).
24. Liao, H. K. *et al.* 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 (2015).
25. Saayman, S. M. *et al.* Potent and targeted activation of latent HIV-1 using the CRISPR/dCas9 activator complex. *Mol Ther.* doi: 10.1038/mt.2015.202 (2015).
26. Zhang, Y. *et al.* CRISPR/gRNA-directed synergistic activation mediator (SAM) induces specific, persistent and robust reactivation of the HIV-1 latent reservoirs. *Sci Rep.* **5**, 16277, doi: 10.1038/srep16277 (2015).
27. Reynolds, L. *et al.* Repression of the HIV-1 5' LTR promoter and inhibition of HIV-1 replication by using engineered zinc-finger transcription factors. *Proc Natl Acad Sci USA* **100**, 1615, doi: 10.1073/pnas.252770699 (2003).
28. Ji, H. *et al.* Specific reactivation of latent HIV-1 by dCas9-SunTag-VP64-mediated guide RNA targeting the HIV-1 promoter. *Mol Ther.* doi: 10.1038/mt.2016.7 (2016).
29. Buonaguro, L., Tornesello, M. L. & Buonaguro, F. M. Human immunodeficiency virus type 1 subtype distribution in the worldwide epidemic: pathogenetic and therapeutic implications. *J Virol.* **81**, 10209, doi: 10.1128/JVI.00872-07 (2007).
30. Wang, Y. *et al.* A small set of succinct signature patterns distinguishes Chinese and non-Chinese HIV-1 genomes. *PLoS One.* **8**, e58804, doi: 10.1371/journal.pone.0058804 (2013).
31. Lau, K. A., Wang, B. & Saksena, N. K. Emerging trends of HIV epidemiology in Asia. *AIDS Rev.* **9**, 218 (2007).
32. Tee, K. K. *et al.* Temporal and spatial dynamics of human immunodeficiency virus type 1 circulating recombinant forms 08_BC and 07_BC in Asia. *J Virol.* **82**, 9206, doi: 10.1128/JVI.00399-08 (2008).
33. Su, L. *et al.* Characterization of a virtually full-length human immunodeficiency virus type 1 genome of a prevalent intersubtype (C/B') recombinant strain in China. *J Virol.* **74**, 11367, doi: 10.1128/JVI.74.23.11367-11376 (2000).
34. Piyasirisilp, S. *et al.* A recent outbreak of human immunodeficiency virus type 1 infection in southern China was initiated by two highly homogeneous, geographically separated strains, circulating recombinant form AE and a novel BC recombinant. *J Virol.* **74**, 11286, doi: 10.1128/JVI.74.23.11286-11295 (2000).
35. Wang, W. *et al.* CCR5 gene disruption via lentiviral vectors expressing Cas9 and single guided RNA renders cells resistant to HIV-1 infection. *PLoS One.* **9**, e115987, doi: 10.1371/journal.pone.0115987 (2014).
36. Li, C. *et al.* Inhibition of HIV-1 infection of primary CD4+ T-cells by gene editing of CCR5 using adenovirus-delivered CRISPR/Cas9. *J Gen Virol.* **96**, 2381, doi: 10.1099/vir.0.000139 (2015).
37. Strain, M. C. *et al.* Heterogeneous clearance rates of long-lived lymphocytes infected with HIV: intrinsic stability predicts lifelong persistence. *Proc Natl Acad Sci USA* **100**, 4819, doi: 10.1073/pnas.0736332100 (2003).
38. Kumar, A., Darcis, G., Van Lint, C. & Herbein, G. Epigenetic control of HIV-1 post integration latency: implications for therapy. *Clin Epigenetics* **7**, 103, doi: 10.1186/s13148-015-0137-6 (2015).
39. Bruner, K. M., Hosmane, N. N. & Siliciano, R. F. Towards an HIV-1 cure: measuring the latent reservoir. *Trends Microbiol.* **23**, 192, doi: S0966-842X(15)00027-X (2015).
40. Archin, N. M., Sung, J. M., Garrido, C., Soriano-Sarabia, N. & Margolis, D. M. Eradicating HIV-1 infection: seeking to clear a persistent pathogen. *Nat Rev Microbiol.* **12**, 750, doi: 10.1038/nrmicro3352 (2014).
41. Kamine, J., Subramanian, T. & Chinnadurai, G. Sp1-dependent activation of a synthetic promoter by human immunodeficiency virus type 1 Tat protein. *Proc Natl Acad Sci USA* **88**, 8510, doi: 10.1073/pnas.88.19.8510 (1991).
42. Harrich, D. *et al.* Role of SP1-binding domains in *in vivo* transcriptional regulation of the human immunodeficiency virus type 1 long terminal repeat. *J Virol.* **63**, 2585 (1989).
43. Das, A. T., Harwig, A. & Berkhout, B. The HIV-1 Tat protein has a versatile role in activating viral transcription. *J Virol.* **85**, 9506, doi: 10.1128/JVI.00650-11 (2011).
44. Li, Y. *et al.* CRF01_AE subtype is associated with X4 tropism and fast HIV progression in Chinese patients infected through sexual transmission. *AIDS.* **28**, 521, doi: 10.1097/QAD.0000000000000125 (2014).
45. Ait-Khaled, M., McLaughlin, J. E., Johnson, M. A. & Emery, V. C. Distinct HIV-1 long terminal repeat quasispecies present in nervous tissues compared to that in lung, blood and lymphoid tissues of an AIDS patient. *AIDS.* **9**, 675, doi: 10.1097/00002030-199507000-00002 (1995).
46. Turrini, F. *et al.* HIV-1 transcriptional silencing caused by TRIM22 inhibition of Sp1 binding to the viral promoter. *Retrovirology.* **12**, 104, doi: 10.1186/s12977-015-0230-0 (2015).
47. Ji, H. *et al.* Specific Reactivation of Latent HIV-1 by dCas9-SunTag-VP64-mediated Guide RNA Targeting the HIV-1 Promoter. *Mol Ther.* **24**, 508, doi: 10.1038/mt.2016.7 (2016).
48. Perez-Pinera, P. *et al.* RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat Methods.* **10**, 973, doi: 10.1038/nmeth.2600 (2013).
49. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science.* **339**, 819, doi: 10.1126/science.1231143 (2013).
50. Marini, A., Harper, J. M. & Romero, F. An *in vitro* system to model the establishment and reactivation of HIV-1 latency. *J Immunol.* **181**, 7713, doi: 10.1111/117713 (2008).
51. Spina, C. A. *et al.* An in-depth comparison of latent HIV-1 reactivation in multiple cell model systems and resting CD4+ T cells from aviremic patients. *PLoS Pathog.* **9**, e1003834, doi: 10.1371/journal.ppat.1003834 (2013).

Acknowledgements

We thank Dr. Li Wu and Huan-Zhang Zhu for plasmids and cells, and Prof. Ronald Veazey (Tulane National Primate Research Center, USA) and Prof. Xia Jin (Institut Pasteur of Shanghai, Chinese Academy of Sciences, China) for editorial assistances and suggestions. This work was supported by grants for Interdisciplinary and Collaboration Team of the Chinese Academy of Sciences (J.-H.W), the Natural Science Foundation of China (No. 81572001), and the National Mega Program Project Grant on Key Infectious Diseases (2014ZX10001003).

Author Contributions

J.-H.W. and D.Q. conceived the project, designed the study and wrote the manuscript. D.Q., C.L. and F.S. conducted the experiments. J.-H.W. and D.Q. analyzed the results. Q.L., Z.-Q.J., L.-R.X., H.-J.G. and C.Z. advised the study. All authors reviewed the results and approved the final version of the manuscript.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Qu, D. *et al.* The variances of Sp1 and NF- κ B elements correlate with the greater capacity of Chinese HIV-1 B'-LTR for driving gene expression. *Sci. Rep.* **6**, 34532; doi: 10.1038/srep34532 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016