CellPress

Trends in Genetics

Review

Viral *cis*-regulatory elements as sensors of cellular states and environmental cues

Jaice T. Rottenberg[1,6], Tommy H. Taslim[1,2,6], Luis F. Soto-Ugaldi[3], Lucia Martinez-Cuesta[1,4], Camila Martinez-Calejman[5], and Juan I. Fuxman Bass[1,2,*]

To withstand a hostile cellular environment and replicate, viruses must sense, interpret, and respond to many internal and external cues. Retroviruses and DNA viruses can intercept these cues impinging on host transcription factors via *cis*-regulatory elements (CREs) in viral genomes, allowing them to sense and coordinate context-specific responses to varied signals. Here, we explore the characteristics of viral CREs, the classes of signals and host transcription factors that regulate them, and how this informs outcomes of viral replication, immune evasion, and latency. We propose that viral CREs constitute central hubs for signal integration from multiple pathways and that sequence variation between viral isolates can rapidly rewire sensing mechanisms, contributing to the variability observed in patient outcomes.

Highlights

Viruses sense cellular states and environmental signals by recruiting host transcription factors to *cis*-regulatory elements (CREs) in viral genomes.

Viral CREs act as signal integration hubs that impart specificity to viral gene regulation.

Viral CREs are highly evolvable due to their high mutation rate and ability to rapidly change sensing mechanisms, affecting infection spread and patient outcomes.

Latent viral infections can be treated by targeting viral CREs for reactivation or permanent silencing.

Viral sensing strategies

Viruses are obligate pathogens that need viable cells to complete their replication cycle (Figure 1). To leverage the cell machinery, evade antiviral mechanisms, and initiate viral replication, viruses must recognize a variety of cell states and environmental cues [1–3]. Viruses have evolved a variety of sensing strategies according to their specific replication cycles and cellular niches, including those mediated by viral proteins that sense activation states or that interact with host proteins involved in cell cycle, cell death, and stress. Given their small genomes, viruses cannot encode the diversity of protein sensors needed to detect the myriad cell states they could encounter. One of the main strategies to address this issue involves the sensing of host transcription factors (TFs) by viral CREs. Eukaryotic cells have a complex system of signal transduction cascades to detect changes in the extracellular and intracellular environment that ultimately impinge on different sets of TFs depending on the trigger signal. In the host, these TFs regulate the activity of promoters and enhancers leading to the upregulation or downregulation of gene expression programs [4]. Viruses can intercept and exploit these host sensing mechanisms by binding different TFs to their own viral CREs (Figure 2). This type of sensing mechanism benefits from the compactness of TF binding sites, the rapid evolutionary gain and loss of binding sites in viral genomes, the complex repertoire of TFs that respond to multiple signals and states, and the direct link between sensing and transcriptional regulation of viral effector proteins [5,6].

In a collective effort spanning more than four decades of research, the field has identified hundreds of CREs from retroviruses and double-stranded DNA viruses with many studies describing the signal inputs and regulatory mechanisms by which these CREs function. Key contributions in our overall understanding of viral gene regulation include the impact of chromatin states on viral CRE usage [7–11] as well as the downstream mechanisms of viral gene expression, including RNA polymerase II (Pol II) pause release and elongation [12,13]. Given the breadth of this field, this review cannot discuss the outstanding work produced by hundreds of laboratories. Instead, we aim to provide a broad and synthesized overview of the current knowledge of viral CREs

[1]Department of Biology, Boston University, Boston, MA, USA
[2]Molecular and Cellular Biology and Biochemistry Program, Boston University, Boston, MA, USA
[3]Tri-Institutional Program in Computational Biology and Medicine, New York, NY, USA
[4]Department of Biochemistry, University of Nebraska-Lincoln, Lincoln, NE, USA
[5]Centro de Estudios Farmacologicos y Botanicos (CEFYBO) CONICET-UBA, Buenos Aires, Argentina
[6]Equal contributions

*Correspondence: fuxman@bu.edu (J.I. Fuxman Bass).

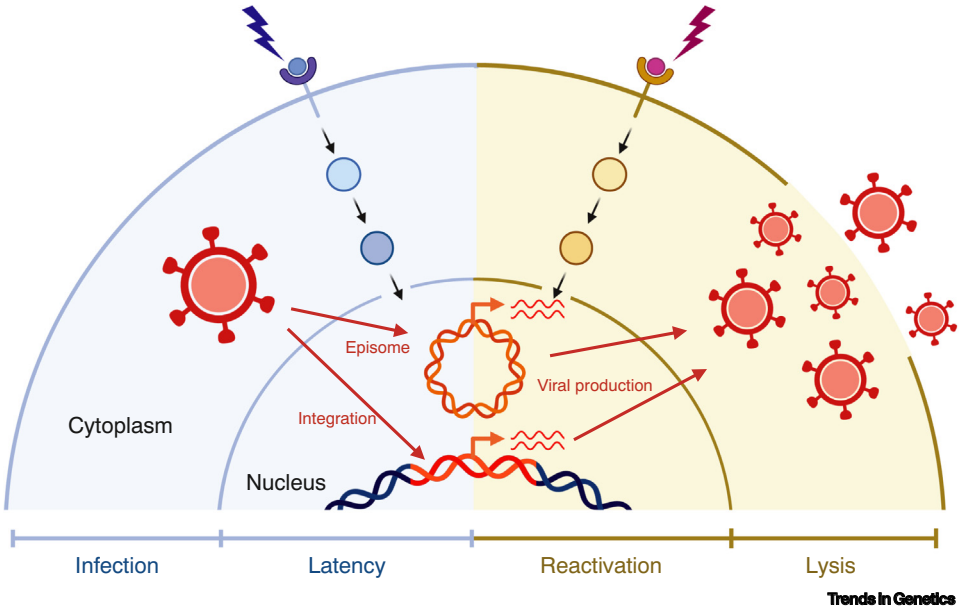

Figure 1. Viral replication cycle and transcriptional control. Schematic of a representative replication cycle for retroviruses and dsDNA viruses. Following initial infection, retroviruses and most dsDNA viruses localize to the nucleus. Here, these viruses may transition to latency through episome formation (e.g., herpesviruses) or via integration into the host genome (e.g., retroviruses, nonenveloped dsDNA viruses). Internal or external cues can promote reactivation from latency or direct entry to the lytic cycle, marked by activation of viral transcription followed by the assembly and release of new virions.

focusing on their function as sensors of the cellular environment. We describe the main signals sensed and the TFs involved, how viral CREs act as signal integration hubs, and how they can rapidly evolve to acquire or rewire sensing strategies. Finally, we discuss the potential targeting of viral CREs for therapeutic purposes.

Overview of viral CREs

Retroviruses and most dsDNA viruses express their genes using the host transcriptional machinery in the nucleus of infected cells [14,15]. This process is controlled by CREs in the viral genome that activate or repress the transcription of viral genes depending on the stages in the replication cycle as well as the states of the infected cells. Given the compactness of viral genomes, most of the identified CREs act as promoter elements; however, some CREs have been shown to act distally as enhancers upstream or downstream of transcription start sites (TSSs) [16–18]. For example, the origin of latent replication (oriP) region in Epstein–Barr virus (EBV) is known to act as an enhancer to the LMP1 gene that is 10 kbp upstream of the oriP [16].

Viral CREs have been identified and characterized using an array of approaches that determine their location in the viral genome, measure their transcriptional activity under different infection and cellular conditions, and identify the molecular mechanisms involved in their regulation (Box 1). Given the mechanistic parallels between viral and host gene expression, viral CREs often share sequence and functional similarities with host CREs. Many of the seminal studies in the 1980s on mammalian transcription used viral CREs, such as the SV40 and BKV enhancer elements and gene promoters from different viruses, as models for CRE structure and regulation processes [19–25]. Similar to host cell CREs, viral CREs also utilize TFs to remodel the chromatin avoiding host repression, recruit RNA Pol II, and promote transcriptional

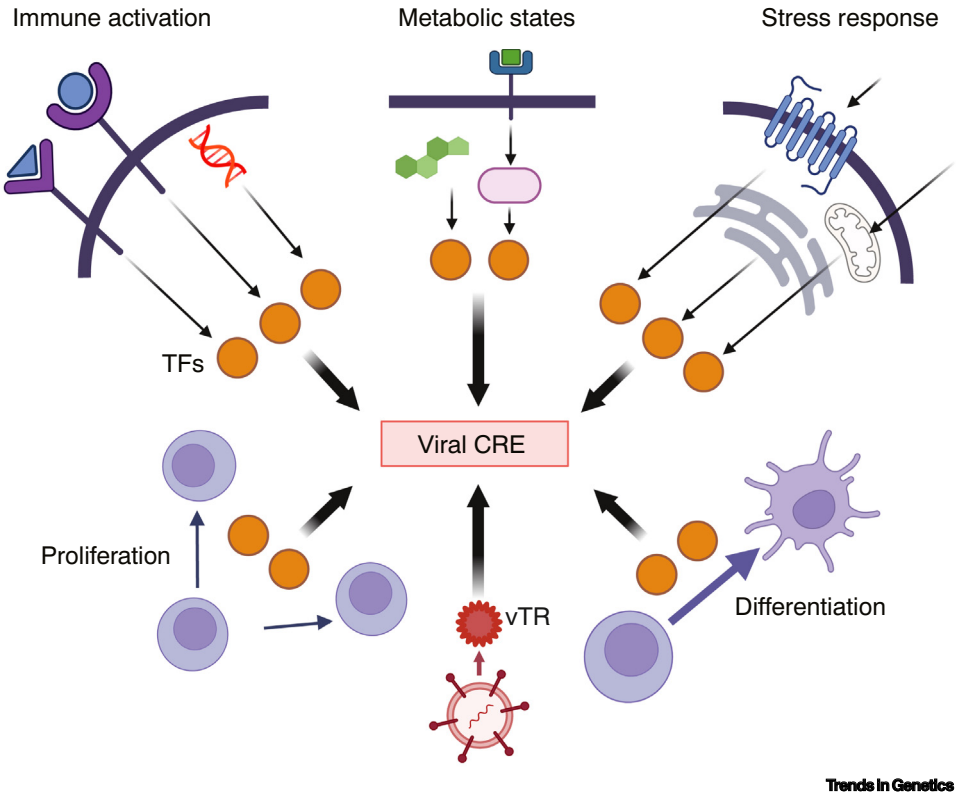

Figure 2. Pathways sensed by viral *cis-regulatory elements* (CREs). Viral CREs sense immune activation, stress, and metabolic states, as well as cell proliferation and differentiation by recruiting host transcription factors (TFs) downstream of these pathways. Abbreviation: vTR, viral transcriptional regulator.

elongation to regulate the expression of the viral genes [10,11]. The hijacking of host TFs enables the virus to transcribe viral genes immediately after infection, which is necessary to produce viral proteins such as viral transcriptional regulators (vTRs) that further directly or indirectly impact the expression of additional viral genes [26]. Ultimately, the recruitment of specific sets of TFs and vTRs onto the viral CREs determines the appropriate timing and cellular states in which each viral gene is expressed, leading to productive replication cycles and evasion of antiviral responses.

The number of CREs across viruses often scales with the number of viral genes and the regulatory complexity; however, gene-to-CRE relationships are often not one to one [27]. Small viruses, such as polyomavirus and retroviruses, have one major CRE region, which is generally multifunctional and controls the transcription of the entire viral genome. For example, the long terminal repeats (LTRs) of HIV and human T lymphotropic virus (HTLV) are involved in the activation and silencing of viral transcription controlling both replication and latency [11,28,29]. In these retroviruses, different genes are produced through alternative splicing of this primary transcript rather than through different CREs [30]. The noncoding control region (NCCR) of the circular polyomavirus also has a dual role, in this case acting as a bidirectional promoter driving the expression of early genes when transcribing in one direction and late genes in the other, depending on the TFs recruited to this CRE [31]. For most other dsDNA viruses, the expression of viral genes is often achieved through independent CREs. For example, human papillomaviruses (HPVs) control the expression of early and late genes through at least two separate promoters [32–34]. Larger

Box 1. Identification and characterization of viral CREs

Identification of viral CREs

Prediction of viral CREs has been performed based on the location of viral coding sequences and by mapping TSSs using 5′RACE for individual genes or genome-wide using cap analysis of gene expression (CAGE), a high-throughput approach to map and quantify TSS activity at single-base resolution (Figure I). This unbiased mapping of TSSs is particularly important for large dsDNA viruses that encode multiple coding and noncoding genes and that have complex patterns of regulation depending on the stage in the viral replication cycle [36]. To determine whether the identified genomic regions are sufficient to drive transcription, reporter assays are commonly used. Bashing experiments, evaluating progressive deletions, internal deletions, mutations, or nucleotide replacements, have been instrumental determining the bounds of CREs and identifying elements within CREs that activate or repress expression.

Identification of host TFs regulating viral CREs:

To determine the TFs that bind and activate or repress a viral CRE, an array of DNA binding and functional assays is generally used. Direct binding is often identified using electrophoretic mobility shift assays using purified TF or nuclear extracts followed by supershift assays, DNase footprinting assays to identify regions in a CRE that are protected from DNase cleavage using nuclear extracts or purified TFs, or yeast one-hybrid assays, a method that can evaluate in parallel the binding of hundreds of TFs to a DNA element of interest. The functional role of these TFs is generally determined using reporter assays where TF binding sites are mutated or by knocking down TF expression. These studies in non-native contexts are often complemented by ChIP and knockdown of specific TFs in infected cells to identify the cellular and infection stage contexts in which the TF binds/regulates the viral CRE.

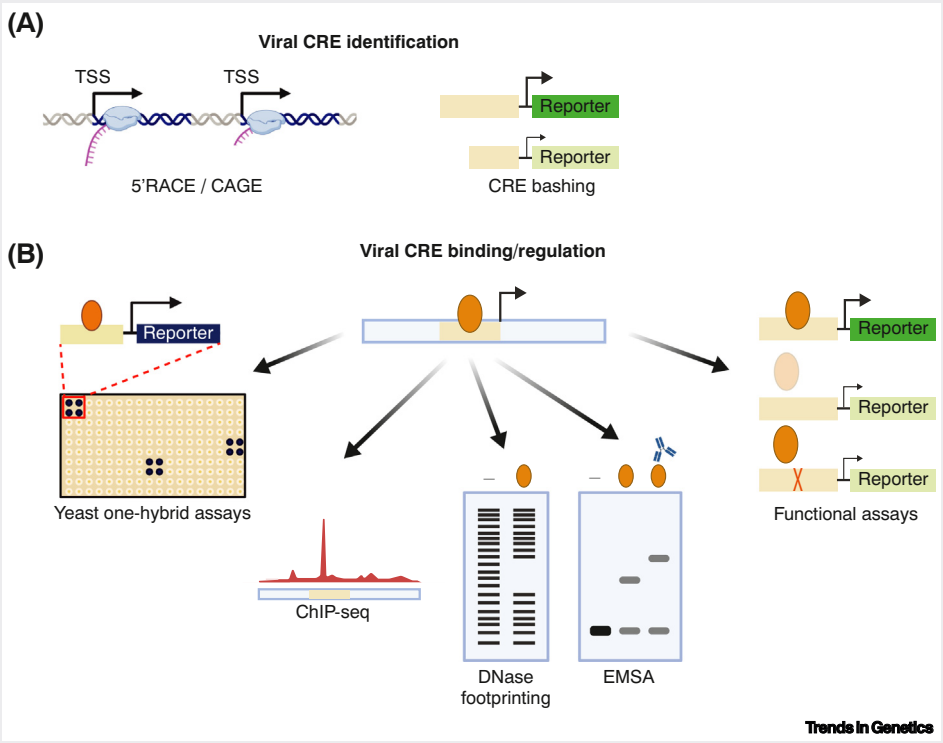

Figure I. Methods to identify and characterize viral *cis-regulatory elements* (CREs). (A) Viral CREs are often predicted based on their location upstream of protein coding sequences, using 5′RACE or cap analysis of gene expression (CAGE) to identify transcription start sites and using CRE bashing reporter assay experiments. (B) Binding of transcription factors (TFs) to viral CREs can be determined using yeast one-hybrid assays, ChIP or ChIP-seq, DNase footprinting in the presence of a TF or nuclear extract, or an electrophoretic mobility shift assay (EMSA). The regulatory effect of these TFs is often measured in reporter or functional assays overexpressing or knocking down the TF or mutating TF binding sites.

dsDNA viruses, such as adenovirus and herpesviruses, have a much more complex CRE landscape, containing tens to hundreds of CREs to control the convoluted cascade of viral gene expression required for different stages in viral replication and latency. For instance, adenovirus 5 with ~35 kb is known to contain 13 CREs, HSV-1 with ~150 kb contains at least 190 CREs [35], and EBV with ~170 kb contains about 322 CREs [36]. This leads to an estimate of one CRE per 0.5–5 kb depending on the virus, a similar or higher CRE density than the human genome [37] and more biased towards promoter rather than enhancer elements.

Viral CREs as sensors of cellular states

Viruses infect different cellular niches and need to sense and respond to diverse cellular and environmental cues. This is often achieved by recruiting to their CREs an array of TFs that are constitutive and lineage specific and those that respond to immune activation, stress, metabolic states, or cell proliferation.

Constitutive and lineage-specific TFs

Viral replication often leads to high cellular stress and elicits immune responses, which may cause apoptosis and early viral clearance. Certain cell differentiation states are less likely to trigger these events and may constitute permissive environments for viral replication. Viruses, therefore, often sense cell differentiation states through the binding of cell- and stage-specific TFs to their CREs to regulate the expression of latent and lytic genes (Figure 2). For example, this is achieved by recruiting XBP1, upregulated during B cell differentiation, to the EBV major lytic gene BZLF1 and the KHSV ORF50 gene, both of which lead to active viral replication [38,39]. In addition to lineage-specific TFs, constitutive TFs such as SP1, SP3, and YY1 that contribute to maintaining cellular identity also regulate gene expression in many retroviruses and dsDNA viruses [40,41]. The activity of these TFs is often controlled by interacting cofactors and post-translational modifications, which can be leveraged by many viruses that depend on these TFs for replication. For instance, SP1-dependent viruses such as HSV-1 and HCMV trigger SP1 phosphorylation soon after initial infection, which leads to lower SP1 activity and reduced expression of target HSV-1 early and immediately early genes [42]. This primes the virus for latency and avoids immune detection, while expressing sufficient viral factors necessary for genome integration. YY1, instead, has a dual function in viral gene inhibition and activation controlled by post-translational modifications and interactions with cofactors, which may be present/active in certain cell stages or conditions [43,44]. This illustrates how constitutive TFs can also lead to context-specific viral gene regulation.

Immune activation

Sensing of host immune activation states is essential for the regulation of entry and exit from latency in many viral species. Depending on the virus, immune activation can either stimulate or repress reactivation, which is likely to reflect different strategies for immune evasion – exit from an immune-stimulated cell versus avoiding detection altogether. Many viral CREs are regulated by the same TFs activated during infection, including AP-1, NF-κB, IRFs, NFAT, and STATs (Figure 2), allowing the virus to differentiate between immune signals and potentially coordinate specific evasion responses [45]. For example, AP-1 family members such as FOS, JUN, and ATFs activated by different immune and stress signals regulate the expression of many CREs from adenovirus, herpesvirus, polyomavirus, and retroviruses and can be involved in both latency and reactivation [45]. Similarly, NF-κB, a TF complex that responds to signals transduced via TNF- and Toll-like receptors as well as nucleic acid sensing pathways, is also recruited to many viral CREs. Binding of NF-κB to CREs can have different regulatory outcomes, ranging from promotion to prevention of reactivation in polyomavirus and KSHV, respectively [46,47]. This sensing of the upstream and downstream antiviral response pathways may afford viruses the ability to more precisely time their own gene expression to counteract host immunity during exit from latency.

Cellular stress

Cellular stress often triggers viral reactivation as a way to avoid hostile cell conditions and enable viral dispersal (Figure 2). Reactivation is induced by hypoxic stress in KSHV and EBV [48,49], fever and UV light in HSV-1 [50], and DNA damage in HCMV [51]. TFs that act downstream of oxidative stress pathways, such as HIF1A and EPAS1, activate CREs of KSHV genes [48,52], including the major regulator of KSHV latency LANA [53], as well RTA and ORF34, both of which are associated with latency switching and viral production [54]. Stress response factors such as MYC and TP53 have overall repressive effects on viral CREs, including the LCR of HPV 16 and 18, the NCRR of BK polyomavirus, and the X gene promoter of hepatitis B virus (HBV) [55–57]. In many of these cases, early viral genes involved in reactivation (i.e., HPV E6, polyomavirus large T-antigen, HBV X protein) inhibit TP53 function through a variety of mechanisms including degradation and transcriptional repression [55–57]. This suggests that TP53 binding to viral CREs may act as a regulatory switch, which, on repression/degradation by viral factors, leads to full-scale viral reactivation. Some viruses, such as HSV-1 and HPV, sense hormone-mediated stress – for example, by binding the glucocorticoid receptor NR3C1 to CREs of immediate early or early genes – promoting viral reactivation [58–60]. Taking these findings together, stress-related TFs can have both positive and negative effects on viral transcription, highlighting the specificity of stress signals sensed by viral CREs.

Metabolic states

Cellular metabolism is also an essential state that viruses leverage to control their own latency and replication (Figure 2). A hallmark example of this metabolic sensing involves HBV, whose main niche is metabolically active liver hepatocytes. HBV can re-enter the lytic phase in response to metabolic stress signals sensed through CREs that recognize metabolically related nuclear hormone receptors such as PPARA, RXRA, and NR1H4 [61]. In addition to nuclear hormone receptors, viral CREs also respond to multiple TFs that are modulated by metabolic signaling pathways such as Akt, TORC, and ERK rather than by direct ligand binding, including CREB1 and FOXO1/3 [62,63]. These pathways sense nutrient availability, growth signals, and metabolic stress and impinge on CREs from HBV and multiple herpesviruses.

Cell cycle and proliferation

Viruses require the host machinery and resources to replicate. These resources, such as nucleotides, amino acids, lipids, and biosynthetic enzymes, often fluctuate during the cell cycle and are more readily available in proliferating cells [64]. In addition, replication during cell division may favor viral production and allow viruses to evade immune sensing mechanisms such as cytosolic DNA sensing [65]. Therefore, among other mechanisms, viruses can sense and initiate viral reactivation by binding cell cycle-regulated TFs such as E2F, MYC, TP53, EGR1, and TEADs to their CREs (Figure 2).

Viral CREs integrate cellular and viral signals

Viral gene expression is a mutistep process that, depending on the virus and gene, can involve chromatin remodeling, DNA demethylation, and RNA Pol II recruitment, pause-release, and productive elongation (Figure 3A) [8–14]. TFs and vTRs are at the center of all of these processes and mediate the integration of the different signaling pathways into a finely tuned transcriptional output. Therefore, viral CREs, similar to many host CREs regulating immune, stress, and cell differentiation genes, function as signal integration hubs by binding TFs that sense different signaling cascades and cellular states [45]. This is achieved through multiple mechanisms that include DNA binding cooperativity or antagonism between different TFs (and vTRs), cooperative recruitment of cofactors, and recruitment of different cofactors with synergistic or antagonistic transcriptional effects (Figure 3B). Collectively, these processes ensure that the regulation of viral

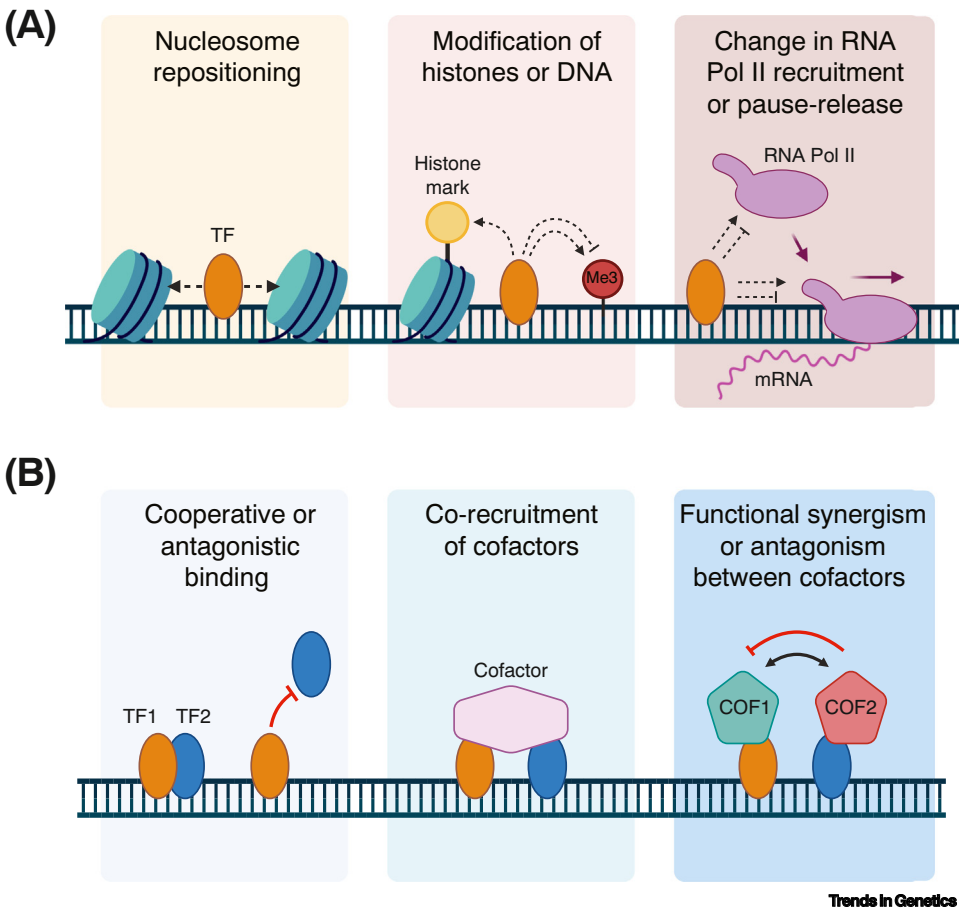

Figure 3. Viral *cis-regulatory elements* (CREs) as signal integration hubs. (A) Consequences of transcription factor (TF) and viral transcriptional regulator (vTR) recruitment to viral CREs, including remodeling of local chromatin structure, modification of histone marks or DNA methylation, and alteration of RNA polymerase II (Pol II) recruitment, pause-release, and elongation. (B) Mechanisms for signal integration including cooperative or antagonistic recruitment of TFs to viral CREs, co-recruitment of cofactors (COFs) by different TFs, and functional synergism or antagonism between cofactors.

gene expression depends on simultaneous activation or inhibition signals. This not only provides specificity but also leads to failsafe mechanisms that prevent unintended viral gene activation, safeguarding against premature immune detection, unproductive replication, or triggering of cell death.

Multiple studies have reported cooperative and antagonistic TF binding to viral CREs. Cooperativity is often associated with heterodimeric TFs but also with TFs that do not necessarily interact but facilitate each other's recruitment [66,67]. This cooperativity can lead to robust viral gene expression when independent signals are present, diminishing the impact of noise in signaling pathway spurious activation. Cooperativity has also been extensively reported between host TFs and vTRs, which enables coordination of host cell states with specific stages in viral latency or reactivation cascades [26]. In addition to cooperativity, a few examples of antagonism due to competition for binding sites or TF sequestration have been reported [68–70]. Synergism between TFs and signaling pathways can also occur by the cooperative recruitment of cofactors that remodel the chromatin, recruit RNA Pol II, or promote transcriptional elongation (Figure 3B). In addition to cooperative

recruitment, TFs that bind a viral CRE may recruit different cofactors that may have additive or synergistic effects on target gene expression by acting on different steps in transcriptional initiation and elongation. Conversely, TFs binding to a CRE may have antagonistic effects if they recruit cofactors with opposing functions, such as a histone acetyltransferase and a histone deacetylase. Both synergistic and antagonistic recruitment of cofactors are observed in the HIV LTR. During reactivation, TFs including NFAT, NF-κB, and AP-1 family members recruit histone acetyl-transferases, increasing local chromatin accessibility and facilitating RNA Pol II recruitment and elongation mediated by P-TEFb [71–73]. By contrast, SP1 and YY1 can recruit HDACs and methyltransferases such as SUV39H1 to maintain latency [74].

Evolution of viral CREs

The gain of new viral proteins is rare within a particular virus species and evolution of viral protein sequences to rewire sensing mechanisms often requires multiple nucleotide changes. By contrast, sensing through viral CREs is highly evolvable as single nucleotide variants (SNVs) and indels frequently lead to gain or loss of TF binding sites given their short nucleotide lengths (Figure 4A). This is similar to what is observed with host CREs, where SNVs affecting a single TF binding site can cause severe developmental malformations or drive cancer [75–77], but supercharged by the high mutation rate and large population sizes of viruses. In the case of viral CREs, sequence variants can lead to sensing variability across isolates both within and across individuals, having a large impact on viral dispersal and the course of infection. This includes a variation in the number of binding sites for a TF within the population of viruses, which may be responsible for differences in patient outcomes (Figure 4B). For example, variation in

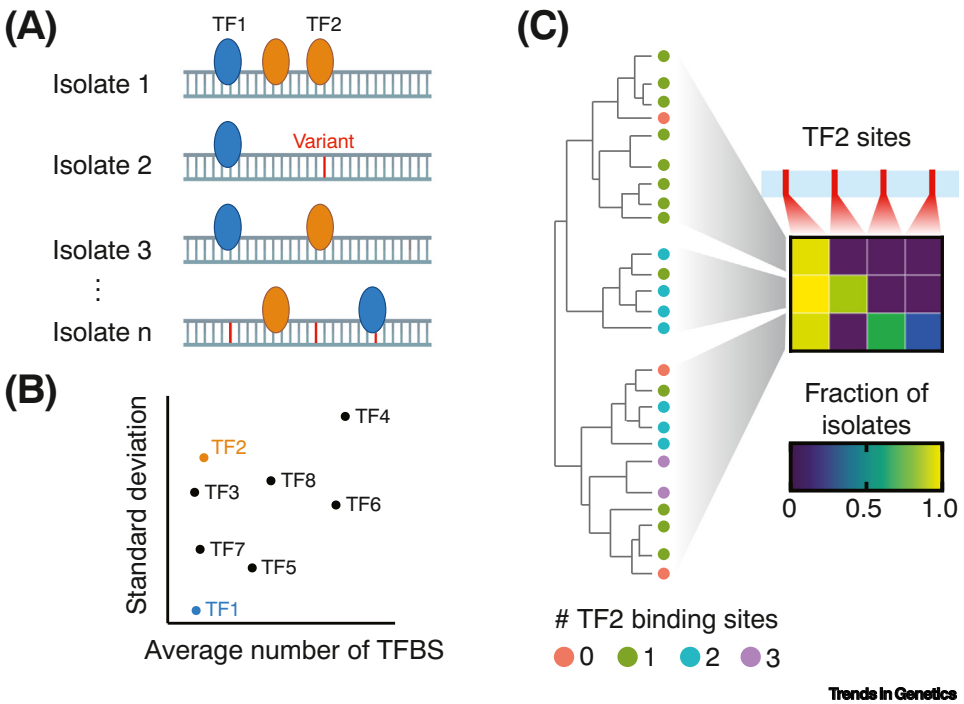

Figure 4. Evolution of viral *cis-regulatory elements* (CREs). (A) Schematic of transcription factor (TF) binding variation between CREs from different isolates due to single nucleotide or indel variants. (B) Scatter plot showing the average number of TF binding sites (TFBSs) versus the standard deviation across isolates for TFs that bind to a viral CRE. (C) Viral dendrogram where isolates are colored based on the number of TF2 binding sites in a CRE. The heatmap shows the fraction of isolates in each subclade that contain a binding site for TF2 at each position in the CRE.

the number of NF-κB and NFAT binding sites across HIV-1 clades has been associated with differences in LTR transcriptional activities, in responses to the upstream signaling pathways, and in pathogenesis and transmissibility across isolates [78–81]. Similarly, variation in NFATC3 binding sites in the BK polyomavirus NCCR has been associated with differences in the recruitment of RELA and AP-1, in transcriptional activity, in viral production, and in viral reactivation in renal transplant patients [82,83]. Further, viral CREs from closely related strains can not only differ in the number of binding sites of a TF but also in their location in the CRE (Figure 4C). This provides an opportunity for novel signal integration outcomes as gain of binding sites at other locations in a CRE can create new cooperative or antagonistic relationships with other TFs and cofactors. This leads to high plasticity and variation in viral CREs, which together with the large viral numbers within individuals and across the population and the strong selective pressures viruses face, leads to the selection of variants that can rapidly rewire sensing mechanisms.

Viral CREs as therapeutic targets

Latent infections have been a longstanding medical challenge, which often require lifetime treatment to avoid or manage recurrent infections. In HIV-1, which exhibits latent infection, the standard course of antiretroviral therapy is unable to eliminate replication-competent provirus residing in $CD4^+$ T cells [84–86]. Exploratory treatment strategies dubbed 'shock and kill' have been developed to eliminate latently infected cells through reactivation and subsequent immune clearance of the virus using latency reversing agents (LRAs) [87]. LRAs include compounds that promote open chromatin such as HDAC, BRD4, or HMT inhibitors, thereby derepressing latent HIV-1 [88]. Other LRAs involve PKC and TLR agonists, which reactivate HIV-1 via stimulation of immune signaling pathways, given that the HIV LTR binds immune response TFs such as NF-κB and AP-1 [88,89]. While shock and kill can provide a means to target latent viruses, current methods prove challenging in practice as viral reactivation must be balanced with limited host tissue damage. An alternative approach, termed 'block and lock' seeks to silence HIV-1 using latency promoting agents (LPAs) that insulate the LTR from activating cellular signals for prolonged drug-free remission [87]. However, the efficiency of both shock-and-kill and block-and-lock strategies is limited due to the intrinsic noise in LTR activation, the variation in HIV LTR sequences within and across patients, and the influence of integration sites in reactivation or silencing dynamics. Current research into the implementation of LRAs and LPAs previously aimed towards retroviral infections have since been adapted to treat latent herpesvirus infections such as EBV and HCMV [90,91]. A recently proposed method for the treatment of latent HCMV involving HDAC inhibition leads to transient expression of lytic immediate early genes without reactivating the whole virus [92]. While overall these avenues of treatment offer promise, broad challenges such as toxicity and applicability across latent viral types remain unsolved. Further, the high evolvability of viral CREs may render shock-and-kill and block-and-lock strategies largely ineffective, as pressures imposed by treatment may rapidly select for escape variants.

Targeting individual or select groups of viral CREs that control genes that do not lead to reactivation (e.g., late genes) may prove an effective strategy to signal the immune system for clearance without reactivating the entire virus and initiating lytic replication. Many TFs, such as nuclear hormone receptors, have specific ligands that activate or repress transcriptional effector functions on DNA binding. Other TFs can also be directly drugged, their function can be inhibited through signaling pathway perturbations, or they can be selectively targeted for degradation [93]. Unlike existing methods targeting cofactors, targeting single TFs or small groups of TFs may reduce the overall likelihood of off-target effects. By leveraging the druggability of such TFs and signaling pathways, CREs regulating the expression of highly immunogenic viral proteins that are not involved in lytic replication (e.g., capsid, envelope) can be selectively targeted for a precision sensitizing rather than shock-and-kill approach. Alternatively, the disruption of negative regulatory

loops can be effective to increase viral gene expression to cytotoxic levels, inhibiting viral replication and resulting in a high genetic barrier to resistance [94]. Another approach to prevent viral replication is to interfere with the typical order of events in the lytic signaling cascade, by targeting viral CREs to induce the expression of several lytic viral proteins out of signaling order or by activating single lytic components known to be dependent on other viral proteins for their activity. Together, these precision sensitizing strategies may be able to circumvent some of the issues and limitations of shock-and-kill and block-and-lock approaches.

Concluding remarks

Ongoing research is shedding light on the complex interactions between viruses and hosts and how viruses sense, respond, and adapt to host cellular and organismal states. As discussed in this review, viral CREs are at the center of these sensing and effector mechanisms and constitute promising targets for antiviral development. However, many challenges remain in the path of identifying, characterizing, and targeting viral CREs for therapeutics (see Outstanding questions). High-throughput approaches such as massively parallel reporter assays or STARR-seq can be used for systematic identification of viral CREs and the cellular contexts in which they are active [95,96], whereas high-throughput protein–DNA binding assays may lead to the identification of the TFs and mechanisms involved in viral CRE regulation [97,98]. Leveraging this information for therapeutic targeting is likely to prove more challenging given the shared regulatory strategies used by host and viral CREs and the rapid evolution of viral CREs within and across patients, which could limit broad applicability and may lead to resistance. Viral CRE studies are also critical for the development of gene therapy, as cryptic CREs present in viral vectors may result in off-target payload expression [99]. Future research must therefore prioritize not only the discovery and characterization of viral CREs but also the development of innovative strategies to circumvent the issues of sequence diversity and specificity, ensuring that therapeutic interventions can be both effective and broadly applicable.

Outstanding questions

What is the full repertoire of CREs for each virus and in which viral stages are they active?

Which cellular and environmental signals impinge on the different viral CREs and how are these signals integrated?

What is the impact of different environmental exposures (e.g., drugs, xenobiotics, metabolites, other infections) on viral CRE activity and reactivation?

How variable are viral CREs across isolates and how does this affect viral transmission and the course of infection?

Which CRE-targeting strategies would be most effective to treat latent viral infections?

Acknowledgments

This work was funded by National Institutes of Health grant R35 GM128625 awarded to J.I.F.B. L.M-C. was supported by the Program RAICES from the Ministry of Science and Technology. C.M-C. was supported by a grant from the Agencia Nacional de Promoción Científica y Tecnológica (PICT-2021-I-INVI-00846).

Declaration of interests

The authors declare no competing interests.

References

1. Barber, G.N. (2001) Host defense, viruses and apoptosis. *Cell Death Differ.* 8, 113–126
2. Dimitrov, D.S. (2004) Virus entry: molecular mechanisms and biomedical applications. *Nat. Rev. Microbiol.* 2, 109–122
3. Palmer, C.S. (2022) Innate metabolic responses against viral infections. *Nat. Metab.* 4, 1245–1259
4. Weidemüller, P. *et al.* (2021) Transcription factors: bridge between cell signaling and gene regulation. *Proteomics* 21, 2000034
5. Lambert, S.A. *et al.* (2018) The human transcription factors. *Cell* 172, 650–665
6. Pai, A. and Weinberger, L.S. (2017) Fate-regulating circuits in viruses: from discovery to new therapy targets. *Annu. Rev. Virol.* 4, 469–490
7. Lieberman, P.M. (2008) Chromatin organization and virus gene expression. *J. Cell. Physiol.* 216, 295–302
8. Abbate, E.A. *et al.* (2006) Structure of the papillomavirus DNA-tethering complex E2:Brd4 and a peptide that ablates HPV chromosomal association. *Mol. Cell* 24, 877–889
9. Barbera, A.J. *et al.* (2006) The nucleosomal surface as a docking station for Kaposi's sarcoma herpesvirus LANA. *Science* 311, 856–861
10. Tsai, K. and Cullen, B.R. (2020) Epigenetic and epitranscriptomic regulation of viral replication. *Nat. Rev. Microbiol.* 18, 559–570
11. Kato, H. *et al.* (1991) Repression of HIV-1 transcription by a cellular protein. *Science* 251, 1476–1479
12. Core, L. and Adelman, K. (2019) Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation. *Genes Dev.* 33, 960–982
13. Devlin, A.M. *et al.* (2022) The PNUTS-PP1 complex acts as an intrinsic barrier to herpesvirus KSHV gene expression and replication. *Nat. Commun.* 13, 7447
14. Varmus, H. (1988) Retroviruses. *Science* 240, 1427–1435
15. Goff, S.P. (2007) Host factors exploited by retroviruses. *Nat. Rev. Microbiol.* 5, 253–263
16. Gahn, T.A. and Sugden, B. (1995) An EBNA-1-dependent enhancer acts from a distance of 10 kilobase pairs to increase expression of the Epstein–Barr virus LMP gene. *J. Virol.* 69, 2633–2636
17. Gu, W. *et al.* (1995) Multiple tandemly repeated binding sites for the YY1 repressor and transcription factors AP-1 and SP-1 are clustered within intron-1 of the gene encoding the IE110 transactivator of herpes simplex virus type 1. *J. Biomed. Sci.* 2, 203–226

18. Arvey, A. *et al.* (2012) An atlas of the Epstein–Barr virus transcriptome and epigenome reveals host-virus regulatory interactions. *Cell Host Microbe* 12, 233–245
19. Rosenthal, N. *et al.* (1983) BK viral enhancer element and a human cellular homolog. *Science* 222, 749–755
20. Banerji, J. *et al.* (1981) Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. *Cell* 27, 299–308
21. Seto, E. *et al.* (1993) Interaction between transcription factors Sp1 and YY1. *Nature* 365, 462–464
22. Ondek, B. *et al.* (1988) The SV40 enhancer contains two distinct levels of organization. *Nature* 333, 40–45
23. Imbra, R.J. and Karin, M. (1986) Phorbol ester induces the transcriptional stimulatory activity of the SV40 enhancer. *Nature* 323, 555–558
24. Horikoshi, M. *et al.* (1988) Transcription factor ATF interacts with the TATA factor to facilitate establishment of a preinitiation complex. *Cell* 54, 1033–1042
25. Dynan, W.S. and Tjian, R. (1983) The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. *Cell* 35, 79–87
26. Liu, X. *et al.* (2020) Human virus transcriptional regulators. *Cell* 182, 24–37
27. Ho, J.S.Y. *et al.* (2021) Unconventional viral gene expression mechanisms as therapeutic targets. *Nature* 593, 362–371
28. Nabel, G. and Baltimore, D. (1987) An inducible transcription factor activates expression of human immunodeficiency virus in T cells. *Nature* 326, 711–713
29. Siekevitz, M. *et al.* (1987) Activation of the HIV-1 LTR by T cell mitogens and the trans-activator protein of HTLV-I. *Science* 238, 1575–1578
30. Sertznig, H. *et al.* (2018) Behind the scenes of HIV-1 replication: alternative splicing as the dependency factor on the quiet. *Virology* 516, 176–188
31. Moens, U. *et al.* (1995) Noncoding control region of naturally occurring BK virus variants: sequence comparison and functional analysis. *Virus Genes* 10, 261–275
32. Smotkin, D. and Wettstein, F.O. (1986) Transcription of human papillomavirus type 16 early genes in a cervical cancer and a cancer-derived cell line and identification of the E7 protein. *Proc. Natl. Acad. Sci. U. S. A.* 83, 4680–4684
33. Ozbun, M.A. and Meyers, C. (1998) Temporal usage of multiple promoters during the life cycle of human papillomavirus type 31b. *J. Virol.* 72, 2715–2722
34. Thierry, F. *et al.* (1987) Characterization of a transcriptional promoter of human papillomavirus 18 and modulation of its expression by simian virus 40 and adenovirus early antigens. *J. Virol.* 61, 134–142
35. Whisnant, A.W. *et al.* (2020) Integrative functional genomics decodes herpes simplex virus 1. *Nat. Commun.* 11, 2038
36. Fülöp, Á. *et al.* (2022) Integrative profiling of Epstein–Barr virus transcriptome using a multiplatform approach. *Virol. J.* 19, 7
37. Gasperini, M. *et al.* (2020) Towards a comprehensive catalogue of validated and target-linked human enhancers. *Nat. Rev. Genet.* 21, 292–310
38. Bhende, P.M. *et al.* (2007) X-box-binding protein 1 activates lytic Epstein–Barr virus gene expression in combination with protein kinase D. *J. Virol.* 81, 7363–7370
39. Dalton-Griffin, L. *et al.* (2009) X-box binding protein 1 contributes to induction of the Kaposi's sarcoma-associated herpesvirus lytic cycle under hypoxic conditions. *J. Virol.* 83, 7202–7209
40. Jones, K. and Tjian, R. (1985) Sp1 binds to promoter sequences and activates herpes simplex virus 'immediate-early' gene transcription *in vitro*. *Nature* 317, 179–182
41. Liu, R. *et al.* (1994) The transcription factor YY1 binds to negative regulatory elements in the human cytomegalovirus major immediate early enhancer/promoter and mediates repression in non-permissive cells. *Nucleic Acids Res.* 22, 2453–2459
42. Iwahori, S. *et al.* (2007) Enhanced phosphorylation of transcription factor Sp1 in response to herpes simplex virus type 1 infection is dependent on the ataxia telangiectasia-mutated protein. *J. Virol.* 81, 9653–9664
43. Zalani, S. *et al.* (1997) The cellular YY1 transcription factor binds a *cis*-acting, negatively regulating element in the Epstein–Barr virus BRLF1 promoter. *J. Virol.* 71, 3268–3274
44. Khalil, M.I. *et al.* (2014) Cellular transcription factor YY1 mediates the varicella-zoster virus (VZV) IE62 transcriptional activation. *Virology* 449, 244–253
45. Kropp, K.A. *et al.* (2014) Viral enhancer mimicry of host innate-immune promoters. *PLoS Pathog.* 10, e1003804
46. Ranganathan, P.N. and Khalili, K. (1993) The transcriptional enhancer element, xB, regulates promoter activity of the human neurotropic virus, JOV, in cells derived from the CNS. *Nucleic Acids Res.* 21, 1959–1964
47. Gorrill, T.S. and Khalili, K. (2005) Cooperative interaction of p65 and C/EBPβ modulates transcription of BKV early promoter. *Virology* 335, 1–9
48. Davis, D.A. *et al.* (2001) Hypoxia induces lytic replication of Kaposi sarcoma-associated herpesvirus. *Blood* 97, 3244–3250
49. Jiang, J.-H. *et al.* (2006) Hypoxia can contribute to the induction of the Epstein–Barr virus (EBV) lytic cycle. *J. Clin. Virol.* 37, 98–103
50. Roizman, B. and Whitley, R.J. (2013) An inquiry into the molecular basis of HSV latency and reactivation. *Ann. Rev. Microbiol.* 67, 355–374
51. Merchut-Maya, J.M. *et al.* (2022) Human cytomegalovirus hijacks host stress response fueling replication stress and genome instability. *Cell Death Differ.* 29, 1639–1653
52. Haque, M. *et al.* (2003) Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8) contains hypoxia response elements: relevance to lytic induction by hypoxia. *J. Virol.* 77, 6761–6768
53. Cai, Q. *et al.* (2006) Kaposi's sarcoma-associated herpesvirus latent protein LANA interacts with HIF-1α to upregulate RTA expression during hypoxia: latency control under low oxygen conditions. *J. Virol.* 80, 7965–7975
54. Nishimura, M. *et al.* (2017) Kaposi's sarcoma-associated herpesvirus ORF34 is essential for late gene expression and virus production. *Sci. Rep.* 7, 329
55. Sichero, L. *et al.* (2012) Identification of novel cellular transcription factors that regulate early promoters of human papillomavirus types 18 and 16. *J. Infect. Dis.* 206, 867–874
56. Takada, S. *et al.* (1996) Hepatitis B virus X gene expression is activated by X protein but repressed by p53 tumor suppressor gene product in the transient expression system. *Virology* 216, 80–89
57. Shivakumar, C.V. and Das, G.C. (1996) Interaction of human polyomavirus BK with the tumor-suppressor protein p53. *Oncogene* 13, 323–332
58. Ostler, J.B. *et al.* (2019) The glucocorticoid receptor (GR) stimulates herpes simplex virus 1 productive infection, in part because the infected cell protein 0 (ICP0) promoter is cooperatively transactivated by the GR and Krüppel-like transcription factor 15. *J. Virol.* 93, e02063-18
59. Ostler, J.B. and Jones, C. (2021) Stress induced transcription factors transactivate the herpes simplex virus 1 infected cell protein 27 (ICP27) transcriptional enhancer. *Viruses* 13, 2296
60. Gloss, B. *et al.* (1987) The upstream regulatory region of the human papilloma virus-16 contains an E2 protein-independent enhancer which is specific for cervical carcinoma cells and regulated by glucocorticoid hormones. *EMBO J.* 6, 3735–3743
61. Li, L. *et al.* (2009) Limited effects of fasting on hepatitis B virus (HBV) biosynthesis in HBV transgenic mice. *J. Virol.* 83, 1682–1688
62. Hale, A.E. *et al.* (2020) FOXO transcription factors activate alternative major immediate early promoters to induce human cytomegalovirus reactivation. *Proc. Natl. Acad. Sci. U. S. A.* 117, 18764–18770
63. Lang, D. *et al.* (1995) Functional interaction between the human cytomegalovirus 86-kilodalton IE2 protein and the cellular transcription factor CREB. *J. Virol.* 69, 6030–6037
64. Diehl, F.F. *et al.* (2024) The bidirectional relationship between metabolism and cell cycle control. *Trends Cell Biol.* 34, 136–149
65. Zhong, L. *et al.* (2020) Phosphorylation of cGAS by CDK1 impairs self-DNA sensing in mitosis. *Cell Discov.* 6, 26
66. Spitz, F. and Furlong, E.E.M. (2012) Transcription factors: from enhancer binding to developmental control. *Nat. Rev. Genet.* 13, 613–626
67. Morgunova, E. and Taipale, J. (2017) Structural perspective of cooperative transcription factor binding. *Curr. Opin. Struct. Biol.* 47, 1–8
68. Berenson, A. *et al.* (2023) Paired yeast one-hybrid assays to detect DNA-binding cooperativity and antagonism across transcription factors. *Nat. Commun.* 14, 6570

69. Jolma, A. *et al.* (2015) DNA-dependent formation of transcription factor pairs alters their binding specificity. *Nature* 527, 384–388
70. Zhou, X. and O'Shea, E.K. (2011) Integrated approaches reveal determinants of genome-wide binding and function of the transcription factor Pho4. *Mol. Cell* 42, 826–836
71. Lusic, M. (2003) Regulation of HIV-1 gene expression by histone acetylation and factor recruitment at the LTR promoter. *EMBO J.* 22, 6550–6561
72. Zhu, Y. *et al.* (1997) Transcription elongation factor P-TEFb is required for HIV-1 Tat transactivation *in vitro*. *Genes Dev.* 11, 2622–2632
73. Mancebo, H.S.Y. *et al.* (1997) P-TEFb kinase is required for HIV Tat transcriptional activation *in vivo* and *in vitro*. *Genes Dev.* 11, 2633–2644
74. Chéné, I.D. *et al.* (2007) Suv39H1 and HP1γ are responsible for chromatin-mediated HIV-1 transcriptional silencing and post-integration latency. *EMBO J.* 26, 424–435
75. Maurano, M.T. *et al.* (2012) Systematic localization of common disease-associated variation in regulatory DNA. *Science* 337, 1190–1195
76. Maurano, M.T. *et al.* (2015) Large-scale identification of sequence variants influencing human transcription factor occupancy *in vivo*. *Nat. Genet.* 47, 1393–1401
77. Hindorff, L.A. *et al.* (2009) Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl. Acad. Sci. U. S. A.* 106, 9362–9367
78. Montano, M.A. *et al.* (1997) Divergent transcriptional regulation among expanding human immunodeficiency virus type 1 subtypes. *J. Virol.* 71, 8657–8665
79. Jeeninga, R.E. *et al.* (2000) Functional differences between the long terminal repeat transcriptional promoters of human immunodeficiency virus type 1 subtypes A through G. *J. Virol.* 74, 3740–3751
80. Montano, M.A. *et al.* (1998) Dysregulation through the NF-kB enhancer and TATA box of the human immunodeficiency virus type 1 subtype E promoter. *J. Virol.* 72, 8446–8452
81. Bachu, M. *et al.* (2012) Multiple NF-κB sites in HIV-1 subtype C long terminal repeat confer superior magnitude of transcription and thereby the enhanced viral predominance. *J. Biol. Chem.* 287, 44714–44735
82. Jordan, J.A. *et al.* (2010) Transcriptional regulation of BK virus by nuclear factor of activated T cells. *J. Virol.* 84, 1722–1730
83. Li, Y.J. *et al.* (2012) Cyclophilin A and nuclear factor of activated T cells are essential in cyclosporine-mediated suppression of polyomavirus BK replication. *Am. J. Transplant.* 12, 2348–2362
84. Ho, Y.-C. *et al.* (2013) Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell* 155, 540–551
85. Finzi, D. *et al.* (1997) Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. *Science* 278, 1295–1300
86. Siliciano, J.D. *et al.* (2003) Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting $CD4^+$ T cells. *Nat. Med.* 9, 727–728
87. Yeh, Y.-H.J. and Ho, Y.-C. (2021) Shock-and-kill versus block-and-lock: targeting the fluctuating and heterogeneous HIV-1 gene expression. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2103692118
88. Spivak, A.M. and Planelles, V. (2018) Novel latency reversal agents for HIV-1 cure. *Annu. Rev. Med.* 69, 421–436
89. Nixon, C.C. *et al.* (2020) Systemic HIV and SIV latency reversal via non-canonical NF-κB signalling *in vivo*. *Nature* 578, 160–165
90. Keck, K.M. *et al.* (2017) Bromodomain and extraterminal inhibitors block the Epstein–Barr virus lytic cycle at two distinct steps. *J. Biol. Chem.* 292, 13284–13295
91. Groves, I.J. *et al.* (2021) Bromodomain proteins regulate human cytomegalovirus latency and reactivation allowing epigenetic therapeutic intervention. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2023025118
92. Krishna, B.A. *et al.* (2016) Transient activation of human cytomegalovirus lytic gene expression during latency allows cytotoxic T cell killing of latently infected cells. *Sci. Rep.* 6, 24674
93. Patalano, S.D. *et al.* (2023) Transcription factors in the development and treatment of immune disorders. *Transcription* 1–23
94. Chaturvedi, S. *et al.* (2022) Disrupting autorepression circuitry generates "open-loop lethality" to yield escape-resistant antiviral agents. *Cell* 185, 2086–2102.e22
95. Melnikov, A. *et al.* (2012) Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. *Nat. Biotechnol.* 30, 271–277
96. Arnold, C.D. *et al.* (2013) Genome-wide quantitative enhancer activity maps identified by STARR-seq. *Science* 339, 1074–1077
97. Fuxman Bass, J.I. *et al.* (2015) Human gene-centered transcription factor networks for enhancers and disease variants. *Cell* 161, 661–673
98. Robertson, G. *et al.* (2007) Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat. Methods* 4, 651–657
99. Logan, G.J. *et al.* (2017) Identification of liver-specific enhancer–promoter activity in the 3′ untranslated region of the wild-type AAV2 genome. *Nat. Genet.* 49, 1267–1273