

Dynamics of Viral and Host 3D Genome Structure upon Infection

Meyer J. Friedman¹, Haram Lee², Young-Chan Kwon³, and Soohwan Oh^{2*}

¹Department and School of Medicine, University of California, San Diego, La Jolla, CA 92093, USA

²College of Pharmacy, Korea University, Sejong 30019, Republic of Korea

³Center for Convergent Research of Emerging Virus Infections, Korean Research Institute of Chemical Technology, Daejeon 34114, Republic of Korea

Eukaryotic chromatin is highly organized in the 3D nuclear space and dynamically regulated in response to environmental stimuli. This genomic organization is arranged in a hierarchical fashion to support various cellular functions, including transcriptional regulation of gene expression. Like other host cellular mechanisms, viral pathogens utilize and modulate host chromatin architecture and its regulatory machinery to control features of their life cycle, such as lytic versus latent status. Combined with previous research focusing on individual loci, recent global genomic studies employing conformational assays coupled with high-throughput sequencing technology have informed models for host and, in some cases, viral 3D chromosomal structure re-organization during infection and the contribution of these alterations to virus-mediated diseases. Here, we review recent discoveries and progress in host and viral chromatin structural dynamics during infection, focusing on a subset of DNA (human herpesviruses and HPV) as well as RNA (HIV, influenza virus and SARS-CoV-2) viruses. An understanding of how host and viral genomic structure affect gene expression in both contexts and ultimately viral pathogenesis can facilitate the development of novel therapeutic strategies.

Keywords: Chromatin structure, EBV, KSHV, HPV, HIV, influenza virus, SARS-CoV-2

Introduction

Distal cis-regulatory enhancers control appropriate transcription of genes in eukaryotes through their interaction with proximal cis-regulatory promoters. These DNA regulatory elements, which exhibit distinctive epigenetic features, can be separated by tens to hundreds of kilobases (kb) and contain binding sequences for specific transcription factors (TFs) that drive gene expression in part by recruiting co-factors endowed with various enzymatic activities responsible for altering the epigenetic landscape. Advances in the gene regulation field have revealed that 3D chromatin (re-)organization plays an important role in bringing enhancers in close proximity to their target promoters [1]. The structure of the 3D genome largely relies on architectural proteins such as CTCF, cohesin, and YY1 [2]. These architectural proteins contribute not only to topologically associated domains (TADs), which are demarcated by boundaries, but also to smaller chromatin loops within TADs, such as enhancer-promoter (E-P) interactions. In addition to the characteristic epigenetic properties of enhancers and promoters, CTCF/cohesin effects on 3D chromatin organization are a key determinant of E-P interaction specificity and thus spatiotemporal regulation of gene expression [3, 4]. Intriguingly, depletion of host architectural proteins, including CTCF or cohesin, also affects reactivation and replication of many viruses (Table 1). These results imply that the viral life cycle can be regulated by chromosome architecture and further suggest viral mechanisms involving genome organization during infection.

Multiple lines of evidence indicate that the infection state (*i.e.*, lytic vs. latent) significantly impacts chromosome structure as well as E-P interactions of the host and can involve distinct viral-host genomic interplay. Recent advances in genome-wide chromatin biology have revealed a crucial role for chromatin architecture and epigenetic signatures in the regulation of viral infection. It is now apparent that the genomic conformations of both the host and the virus collectively influence the respective gene expression profiles. DNA viruses, some of which can integrate whereas others are maintained as multicopy minichromosomes (or episomes), have several mechanisms that control genome organization of the host and themselves. RNA viruses, a subset of which can also integrate, utilize additional mechanisms for modulating host genome organization involving viral gene products or hijacked host regulatory machinery (Table 2). This review seeks to provide an informative summary of our current understanding of host and viral genomic structural dynamics and its relationship with gene expression in infected cells.

Received: August 12, 2022
Accepted: September 23, 2022

First published online:
September 30, 2022

*Corresponding author
Phone: +82-44-860-1621
Fax: +82-44-860-1606
E-mail: soohwanoh@korea.ac.kr

pISSN 1017-7825
eISSN 1738-8872

Copyright © 2022 by the authors.
Licensee KMB. This article is an
open access article distributed
under the terms and conditions
of the Creative Commons
Attribution (CC BY) license.

Table 1. Involvement of CTCF/cohesin in viral replication, reactivation, and gene expression.

Virus		Reported viral effect(s) of:	
		CTCF knockdown	Cohesin knockdown
Herpesviruses	HSV1 (HHV1)	In lytic infection [40]: ↓ viral gene expression, genome copy number, and viral yield ↓ RNA pol II viral genome occupancy ↑ H3K9me ³ , H3K27me ³ viral genome deposition In latent infection [101]: ↑ ICP0 expression, ↑ viral reactivation	In lytic infection [41]: ↓ viral gene expression, genome copy number, and viral yield ↓ RNA pol II viral genome occupancy ↑ RNA pol II P-ser5 on viral genes ↑ H3K27me ³ viral genome deposition
	EBV (HHV4)	No effect on viral reactivation [102]	Loss of long-range interactions within the viral genome in latently infected cells ↑ latent viral gene expression [16]
	KSHV (HHV8)	Modified viral chromatin conformation Altered lytic and latent viral gene expression [103] ↑ viral yield altered viral gene expression [34] lytic gene expression unaffected viral DNA replication unaffected [21]	Modified viral chromatin conformation Altered lytic and latent viral gene expression [103] ↑ viral yield altered viral gene expression [34] ↑ lytic viral gene expression ↑ viral DNA replication [21]
	HCMV (HHV5)	↑ immediate early and early lytic gene expression ↑ viral yield [38]	
	HPV	↓ viral replication/amplification [56]	↓ viral replication/amplification [56]
	HIV-1	↓ establishment of latency [104]	
	HBV	↑ viral gene expression [105]	None with wild-type virus ↑ viral gene expression in the absence of viral transcriptional regulator HBx [106]
Adenovirus		↓ viral replication ↓ late but not early viral gene expression [107]	

Table 2. Viral mechanisms involving chromatin architecture.

	Genome	Integration into host	Regulatory mechanisms involving chromatin architecture
HHVs	DNA	No Persists as an episome	<ul style="list-style-type: none"> • Regulation of viral gene expression via CTCF/cohesin-dependent changes in episomal structure • Alterations of epigenetic/chromatin organization by viral gene products to modulate host gene expression • Viral-host chromosomal interactions
HPVs	DNA	Yes Maintained either as an episome or by integration into the host genome	<ul style="list-style-type: none"> • Regulation of viral gene expression via CTCF/cohesin/YY1-dependent changes in episomal structure • Activation of viral oncogenes due to disruption of CTCF binding sites during integration • Hijacking of host enhancers to drive viral oncogene expression • Alteration of host gene expression by generation of host chromosome SVs upon viral integration
HIVs	RNA	Yes Stable integration into the host genome	<ul style="list-style-type: none"> • Preferential integration into active open chromatin • Proviral gene expression is correlated with the epigenomic landscape of the host chromosomal loci that contact the viral integration site
Influenza virus	RNA	No	<ul style="list-style-type: none"> • CTCF/cohesin-dependent host chromosomal structure affects viral pathogenesis • Disruption of host genomic organization by viral protein-mediated transcriptional read-through • Activation of host genes by viral protein-induced cohesin recruitment to the host genomic locus
SARS-CoV-2	RNA	No	<ul style="list-style-type: none"> • CTCF/cohesin-dependent host chromosomal structure affects viral pathogenesis • Infection-induced attenuation of chromatin compartmentalization

Technologies for Investigating Chromosome Architecture

The advent of chromosome conformation capture (3C)-based techniques coupled with next-generation sequencing (NGS) technology has dramatically elevated our appreciation of chromosomal organization (Fig. 1). A series of 3C-derivative sequencing methods involving crosslinking, enzymatic digestion of chromosomes, and

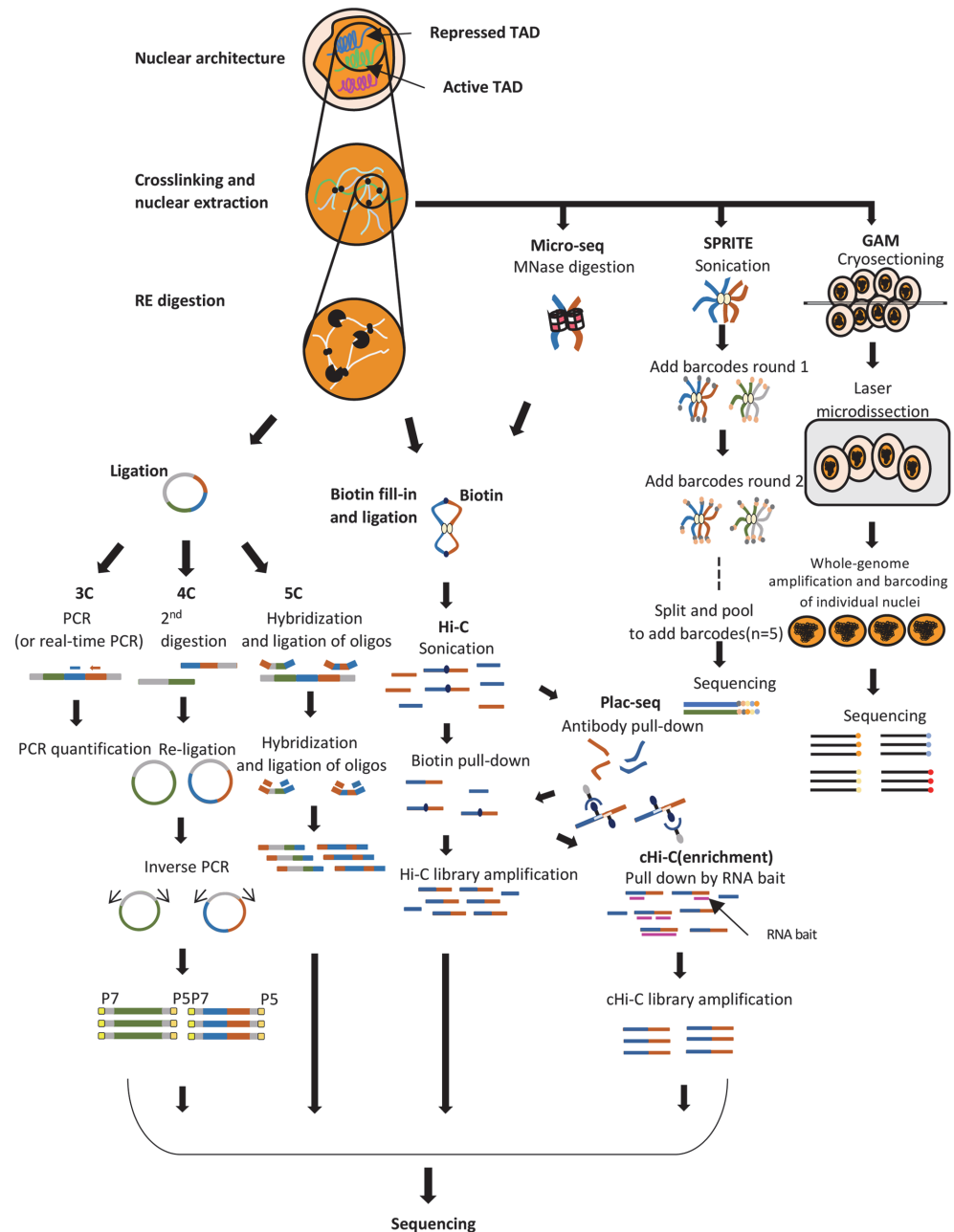


Fig. 1. Schematic representation of assays for studying chromosome architecture. For 3C-based methods, nuclei are first treated with appropriate fixatives (*i.e.*, formaldehyde, DSG, etc.). In GAM, cryosections are cut from paraformaldehyde-fixed and sucrose-embedded samples. In 3C, 4C and 5C, fixed nuclei are treated with restriction enzymes, ligated, and the ligation frequency is measured by PCR or NGS. For Hi-C, Plac-seq, and cHi-C, chromosomal DNA is digested by restriction enzymes while micro-C uses MNase for finer resolution. The digested DNA ends are repaired with biotin-labeled nucleotides followed by blunt-end ligation. The ligated biotin-labeled contacts are sheared and purified with streptavidin beads prior to NGS sequencing. In Plac-seq and cHi-C, antibody pull-down or RNA oligo-mediated DNA pull-down is performed for target enrichment, respectively. In the SPRITE method, crosslinked chromatin is fragmented by sonication, each interacting complex is uniquely tagged by multiple rounds of split-pool barcoding, and the final material is sequenced. In the GAM method, the DNA contents from cryosections are extracted, fragmented, and sequenced. Appropriate computational analysis of the sequencing data from each approach is necessary to detect physical interactions between genomic loci.

proximity ligation have enabled high-throughput and genome-wide detection of contact frequency between genomic loci, including Circular Chromosome Conformation Capture (4C) [5] (one to all), Chromosome Conformation Capture Carbon Copy (5C) [6] (many to many), capture-HiC [7], Proximity Ligation-Assisted ChIP-Seq (PLAC-seq) [8], Chromatin Interaction Analysis by Paired-End Tag Sequencing (ChIA-PET) [9] (many to all), Hi-C [10] and Micro-C [11] (all to all). All of these 3C-based techniques, except for Micro-C, use restriction enzymes to achieve the desired resolution. In principle, a more frequent cutter will yield higher resolution. However, even with a 4-cutter enzyme, which recognizes a tetranucleotide sequence, information on finer contacts between genomic loci, such as E-P, is very rare. In Micro-C, improved resolution to the level of individual nucleosomes (147 bp) is achieved by employing micrococcal nuclease (MNase) in the digestion step. Accordingly, Micro-C structural information shows overall consistency with that of Hi-C but with increased power for detecting short-range associations, including E-P interactions. Several studies have interrogated host and, when applicable, viral chromosomal architecture upon infection with 3C-based techniques, but none have reported data with nucleosomal resolution.

Ligation-independent methods that retain a crosslinking step also have been introduced to assess 3D genomic organization, including split-pool recognition of interactions by tag extension (SPRITE) [12] and genome architecture mapping (GAM) [13]. SPRITE relies on the sequencing of barcoded DNA following multiple rounds of splitting and pooling, such that each interacting chromatin complex is expected to have a unique barcode. The interaction map can be determined from the extracted DNA segments with the same barcode. In the GAM technique, interaction information is derived from micro-dissected slices of nuclei in fixed cells. Because DNA loci in close proximity have a higher probability of being in the same slice, the frequency of the genomic regions in a given slice is calculated as the interaction map. Currently, there are no published studies using SPRITE or GAM to interrogate viral and host genomic organization.

In some cases, the viral components of NGS datasets have not been analyzed yet. For example, lymphoblastoid cell lines (LCLs), which are widely-used, immortalized B cells, are established by infection with the human herpesvirus (HHV) EBV. Because of the convenience of immortalization, the HapMap project [14] used LCLs to amass genotype and gene expression data. In addition, extensively characterized tier 1 ENCODE cell lines include an LCL designated as GM12878 [15]. In LCLs, ~1% of mappable sequencing reads from RNA-seq, ChIP-seq, and other NGS datasets can be aligned to the EBV genome [16]. Thus, a substantial amount of NGS data pertaining to epigenetic profiles and chromosome organization have been collected for the EBV genome through HapMap, ENCODE, and subsequent studies [17, 18]. By virtue of these publicly available datasets, EBV chromatin biology has been relatively well studied compared to other viruses.

Herpesviruses

Herpesviruses are double-stranded DNA viruses that have the potential to establish lifelong infection and cause various diseases. The eight HHVs (HHV1-8) feature relatively large DNA genomes of more than 100kb that encode 100-200 genes.

After infection, the HHV genomes circularize and undergo chromatinizing events, such as histone binding and DNA methylation. The chromatinized viral genome acts as a minichromosome, which is referred to as an episome, and recruits host chromosome regulatory machinery [19, 20]. Because HHV episomes behave like host chromosomes, the role of 3D chromatin structure on viral events and pathogenesis has been extensively studied. In EBV (HHV4) and KSHV (HHV8), the chromosome architectural proteins CTCF and cohesin co-occupy several genomic loci. CTCF/cohesin co-bound sites are most enriched at terminal repeats and in the vicinity of key latent transcripts [16, 21]. Since CTCF/cohesin binding to chromosomes facilitates the formation of chromatin loops, including TADs and E-P interactions, a number of studies have attempted to uncover the role of chromatin conformation in transcriptional control of HHV infection-associated gene expression (Tables 1, 2), as highlighted below.

EBV Episomal Structure

EBV-infected B cells are transformed to LCLs via EBV nuclear antigens (EBNAs). Of the six EBNAs, EBNA2 is specifically important for the expression of essential host and viral genes required for the establishment and proliferation of LCLs and hence EBV persistence [22]. Chromatin conformation has a significant role in the regulation of EBNAs. A loop between the shared EBNA gene promoter Cp and the viral origin of replication (OriP) has an enhancer-blocking effect that silences EBNA2 expression upon transition to restricted latency, which is characterized by increasingly limited latent viral transcription [23, 24]. Formation of this loop depends on CTCF occupancy of a binding site in the EBV genome [24], deletion of which results in upregulation of EBNA2 expression [23]. Conversely, CTCF-dependent transcriptional activation also has been reported in the EBV genome. Disruption of the CTCF binding site in the latent Q promoter (Qp) attenuates associated transcription [25]. In addition, mutagenesis of the CTCF binding site positioned within the overlapping EBV *LMP1* and *LMP2A* genes reduces their expression as part of the observed deregulation of the perturbed latent viral locus [26].

Modulation of viral gene expression through episome structural reconfiguration can affect host chromosome conformation and transcriptional regulation. Latent EBNA proteins serve as transcription factors not only for the viral genome but also for host genes, as exemplified by EBNA2 induction of the *MYC* proto-oncogene to promote B cell proliferation [27, 28]. EBNA2 binds to distal enhancers located hundreds of kb from the transcriptional start site (TSS) of *MYC* [29]. Notably, inactivation of EBNA2 reduces *MYC* expression by weakening chromatin looping between EBNA2-binding enhancers and the *MYC* TSS [30]. Furthermore, *MYC* upregulation increases its recruitment to cognate E-box motifs located proximal to the origin of lytic replication (*oriLyt*) in the EBV

genome. MYC occupancy prevents *oriLyt* interaction with the promoter of the immediate early lytic factor BZLF1 to preserve viral latency [31].

KSHV Episomal Structure

Studies of KSHV show strong CTCF/cohesin binding at an intergenic region between the latent gene *LANA* and the lytic gene *K14* [32]. Disruption of this CTCF/cohesin binding induces the expression of a set of lytic genes, including *K14*, *ORF74*, *ORF57*, and *ORF6* [33, 34]. CTCF/cohesin binding is also detected at the promoter of the reactivation gene *ORF50/RTA*, which is upregulated upon depletion of cohesin [21]. These results raise the possibility that CTCF/cohesin-mediated chromosome architecture can restrict KSHV lytic gene expression and replication. Like EBV, KSHV also increases proto-oncogene *MYC* expression in an enhancer-dependent manner; however, the viral activator of *MYC* enhancers in the host genome remains elusive [35].

Episomal Structure of Other Herpesviruses

In HSV-1 (HHV1) and CMV (HHV5), analogous to other HHVs, CTCF interacts with the viral genome, and it can serve as an insulator to regulate gene expression in the latent and reactivation states [36–39]. In addition, CTCF is recruited to the genome of lytic HSV-1 at multiple binding sites whereupon it promotes lytic transcription by facilitating RNA Pol II elongation and by preventing the deposition of silencing chromatin marks [40]. While cohesin has been reported to support HSV1 lytic transcription in a similar fashion to CTCF [41], any contribution of CTCF/cohesin-mediated episome structural effects has not been rigorously evaluated in the lytic context.

NGS-Based Studies of Herpesvirus and Host Chromatin Structure

Recently, a number of studies have sought to interrogate HHV episomal structure by application of unbiased, genome-wide chromosomal architecture methods, such as Hi-C.

Insights into herpesvirus episomal structure. Re-analysis of EBV-aligned reads in Hi-C data from the LCL cell line GM12878 [17], which features the most transcriptionally active latency program, latency III, confirmed clustering of the major EBNA gene promoter Cp, the latent promoter Qp, and OriP [42]. Further comparison of GM12878 with the largely transcriptionally silent LCL Mutu, representing latency I, revealed significantly different viral episome structures, including substantially more intragenomic contacts in type III than type I latency. Interestingly, EBV episomal structure relies on poly(ADP-ribose) polymerase (PARP) enzymatic activity, which stabilizes CTCF occupancy but may antagonize cohesin binding. Inhibition of PARP1, a transcriptional co-factor and the most active PARP family member, leads to fewer loops within the EBV episome and, curiously, decreased CTCF but increased cohesin at their established co-binding sites in the viral genome [42].

Capture-HiC of the KSHV episome revealed a 3D viral genomic map with a number of dynamic, long-range interactions within the viral chromosome [43]. This study revealed a tendency for gene clustering within similar functional genomic domains, such as the distinct latency and early-lytic gene clusters. Upon lytic protein K-Rta-induced reactivation, genomic interaction of regions encoding early- and immediate-early lytic genes is augmented. A subsequent capture-HiC study with finer resolution (500 bp) reported the KSHV TAD structure, which is restricted by CTCF/cohesin binding [44]. During reactivation, K-Rta is recruited to TAD boundaries, leading to the formation of a larger regulatory unit with a shift from repressive B compartments to active A compartments. However, in these studies, terminal repeats (TRs, >30 repeats of 0.8kb genomic sequence for a total length of ~30 kb), which are likely key regulatory and architectural genomic features, were excluded during analysis due to the technical challenge of mapping repetitive sequences. Therefore, future studies will need to address the impact of TRs on episomal conformation.

Assessment of host and herpesvirus chromosome interaction. Extensive chromosomal interaction between the host and HHV genomes also has been documented recently. In situ Hi-C of EBV-infected Burkitt lymphoma (BL) cell lines revealed inter-chromosomal interactions between host chromosomes and the viral episome during latent infection [45], including preferential association of the EBV genome with inactive and low-gene density regions of the host genome. In addition, 4C-seq using EBV sequence as the viewpoint identified genome-wide association of EBV episomes and host chromosomes in BL [46]. In latently infected BL cells, the EBV episome favors interaction with host genomic sites bound by EBNA1, which tend to be transcriptionally silenced. In infected cells of different origins, the latent EBV episome also makes frequent contacts with lamina-associated domains (LADs) in the host genome, which generally constitute heterochromatic environments with limited transcriptional activity [45, 47]. Thus, the nuclear lamina may be a critical topological and epigenetic regulator of the EBV episome [48]. Interestingly, EBV-host chromosomal interactions are altered during viral reactivation from latency, as the viral episomes reposition from transcriptionally repressive heterochromatin toward active euchromatin [45].

Genome-wide analysis of 3D chromatin topologies in EBV-infected gastric cancer (GC) cells indicated a transition from heterochromatin to euchromatin at EBV-interacting host genomic loci [47]. EBV episomes associate with repressed H3K9me3-marked host enhancers, which are reprogrammed to engage and activate nearby GC-related proto-oncogenes. In other studies, re-analyzed Hi-C [17] and 4C-seq datasets from LCLs showed evidence of EBV episomes interacting with typical and super enhancers decorated with active epigenetic marks in the host genome [49]. Inconsistencies regarding the reported localization as well as functional significance of host-viral chromosomal interactions in BL, GC, and LCLs might be attributable, at least in part, to cell type- and/or latency type-specific effects and regulatory strategies [46].

Similar to EBV, in situ Hi-C analysis suggested that latent KSHV episomes preferentially associate with

heterochromatic, low-gene density regions of the host genome and tend to move toward active euchromatin upon reactivation [45]. In addition, capture-HiC performed in primary effusion lymphoma (PEL) cells indicated that host interaction sites for KSHV episomes are enriched near centromeres [50]. While the significance of this localization remains unclear, overlay of ChIP-seq results for identified cellular interactors of the KSHV latency protein LANA, which has a well-established role in episome maintenance in infected, replicating cells that it achieves by binding KSHV TRs in addition to host chromatin factors [51], suggested an intriguing regulatory strategy for viral latency. The data support a model in which latent KSHV episomes colocalize with LANA and cellular CHD4/ADNP-containing ChAHP complexes on host chromosomes through interactions involving KSHV TRs [50]. This interplay could promote viral latency by allowing the strong repressor protein CHD4 to suppress KSHV lytic genes. Furthermore, a viral lncRNA that is robustly upregulated during KSHV reactivation, known as PAN RNA, may titrate CHD4 from KSHV episomes to de-repress viral lytic genes [50].

Herpesvirus Summary

Studies of HHVs have revealed multiple mechanisms by which chromatin organization allows for maintenance and regulatory control of viral genomes as extrachromosomal DNA, or episomes, that extend beyond viral replication during interphase. Furthermore, host chromosome structural alterations, and their functional effects, are also an important feature of HHV-infected cells (Table 2). Possible molecular mechanisms involving chromosome conformation consequent to HHV infection include: 1) use of host chromosome architectural proteins by viruses to achieve changes in viral gene expression by modulating 3D episomal structure (Fig. 2A); 2) association of viral gene products with regulatory elements of host target genes to induce epigenetic/chromatin organization alterations that impact gene expression (Fig. 2B); and 3) interaction of virus and host chromosomes, the pattern of which can vary during the viral life cycle and selectively affects viral and host gene expression (Fig. 2C).

Human Papillomavirus

Human papillomaviruses (HPVs) are a group of small, non-enveloped dsDNA viruses that are responsible for 99% of cervical cancer (CC) cases [52]. HPV genomes can be maintained in the nucleus either as a chromatinized episome or by integration into the host genome. HPV infection modulates host gene expression by various mechanisms, including direct effects on host chromosomal architecture and through the actions of viral oncoproteins E6 and E7 (E6/E7) [53, 54]. The HPV protein E2 acts as a negative regulator of E6/E7, and its disruption leads to overexpression of E6/E7 that promotes oncogenesis [55].

HPV Episomal Structure

HPV viral gene expression is also regulated by host chromosome architectural proteins. A cohesin subunit, SMC1, is constitutively activated by HPV infection and binds to the HPV genome in a CTCF-dependent manner to facilitate genome amplification [56, 57]. Notably, the HPV genome harbors a conserved CTCF binding site in the E2 ORF that recruits CTCF to regulate viral gene expression [58]. Disruption of this CTCF binding site results in decreased E2 expression as well as epigenetic activation of an HPV enhancer, which collectively upregulates E6/E7 to enhance cell proliferation. During keratinocyte differentiation, loss of CTCF/YY1-dependent suppressive looping in the HPV18 episome results in activation of viral enhancer-mediated E6/E7 expression [58, 59] and modulation of viral transcript splicing [60].

HPV Integration

HPV integration into the host genome dramatically impacts chromosomal architecture with pathological consequences. Viral integration, which is a feature of most HPV-induced CCs, results in induction of oncogenic E6/E7 due to downregulation of E1 and E2 [61]. Uniquely, HPV integration can cause structural variants (SV) of host chromosomes that result in gene expression changes [62], which are tightly linked to tumorigenesis [63]. Moreover, as demonstrated in HeLa cells, an extensively studied CC model system, HPV integration is directly associated with a long-range 3D genome organization that favors cancer development [64].

Although HPV integration sites are dispersed throughout the host genome, they tend to be enriched in transcriptionally active regions [65, 66]. In addition, several integration hotspots, which are associated with cancer-related gene expression programs, have been reported [62, 67]. HPV integration breakpoints occur more frequently near super enhancers, which can be amplified at sites of recurrent integration. Some of these super enhancers control the expression of cell-identity genes and may be co-opted to upregulate integrated viral oncogenes [68]. Moreover, integration events can yield super enhancer-like elements composed of tandem, interspersed copies of a viral and a cellular enhancer that drive viral oncogene expression [69]. Comprehensive evaluation of viral integration sites using Nanopore long-read sequencing uncovered various configurations of tandemly integrated HPV sequences at high resolution [70]. Analysis of 16 HPV-positive CCs revealed four HPV integration types, some of which can induce genomic disruptions through chromosomal translocation in addition to known effects on genes located proximal to integration breakpoints. Interestingly, the long-read sequencing data demonstrated that HPV can excise host genomic DNA along with integrated viral DNA to generate a viral-host hybrid extrachromosomal circular DNA (eccDNA) species. Because eccDNA may play an important role in CC by regulating oncogene expression [71], follow-up studies are needed for further characterization of these genomic hybrids.

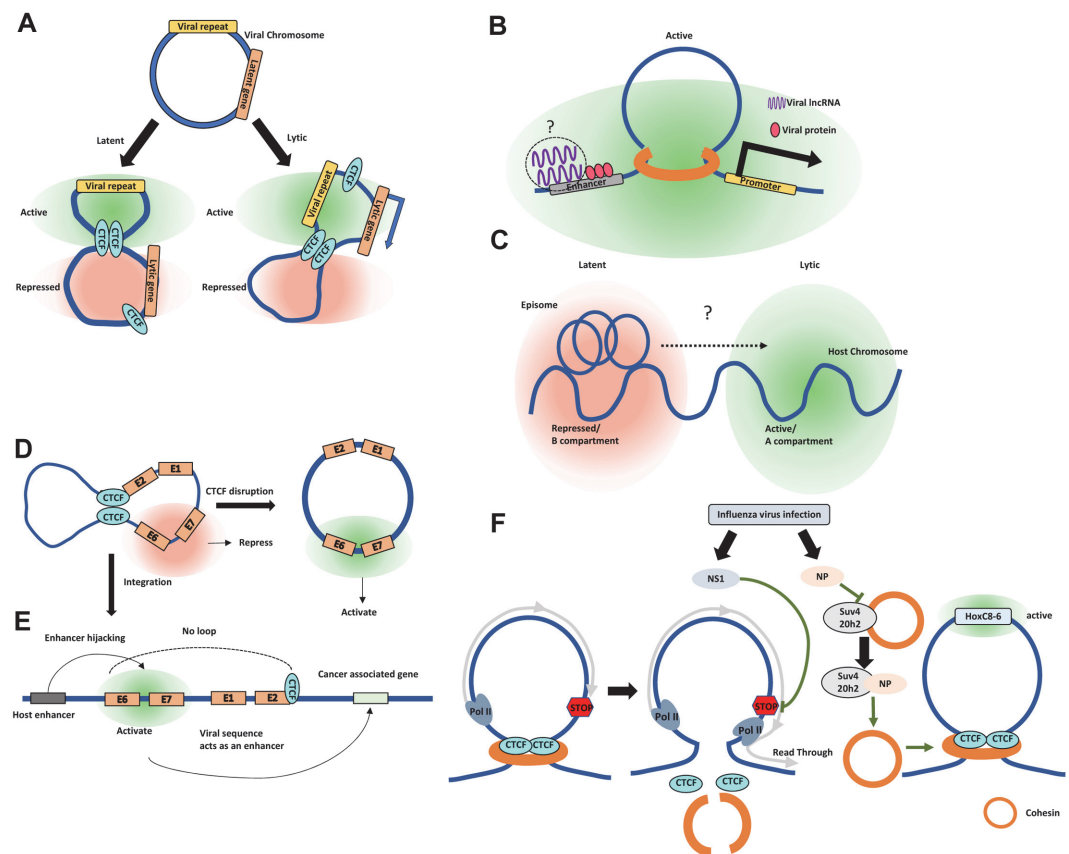


Fig. 2. Models of viral and host regulation related to chromatin conformation. (A) In herpesviruses, host architectural proteins (*i.e.*, CTCF and cohesin) bind the viral episome and influence viral gene expression depending on the state of the virus (*i.e.*, latent vs. lytic) via conformational changes to the episome. (B) Viral gene products can modulate host gene expression by targeting host enhancers or promoters (*i.e.*, EBNA and LANA). (C) The herpesvirus episome is associated with a repressed compartment in the host genome during latency. Upon reactivation, the interaction between the episome and euchromatin may increase. This viral-host chromosomal interplay could modulate both viral and host gene expression. (D) CTCF binding at the E2 promoter acts as a suppressor of viral oncogene E6/E7 by modulating HPV chromatin architecture. Disruption of CTCF binding causes transcriptional activation of E6/E7 in the HPV episome. (E) Viral integration into the host genome weakens looping between the E2 and E6/E7 loci, which also results in activation of the latter. The integrated viral sequence can hijack host enhancers for transcription of viral gene products or can act as an enhancer itself to stimulate neighboring cancer-associated genes. (F) Upon IAV infection, the viral protein NS1 inhibits the termination of transcription in a subset of highly transcribed genes. This inhibition leads to RNA Pol II-dependent dissociation of CTCF from its binding sites and disruption of chromatin loops (left). IAV protein NP competitively binds the host protein Suv4-20h2 to release Suv4-20h2-bound cohesin. Liberated cohesin is then recruited to the *HoxC8-HoxC6* locus where it induces gene expression by forming an active chromatin loop (right).

HPV-Induced Host Genome Structural Variants

Host genome SVs induced by HPV integration have been explored in an unbiased, genome-wide manner using 3C-based approaches. Hi-C data suggested that TAD alterations can be associated with enhancer hijacking, leading to changes in gene expression [54]. HPV integration at a hotspot in the *CCDC106* locus splits the affected TAD into two smaller TADs and, due to the newly imposed TAD boundaries, redirects an enhancer from the tumor suppressor gene *PEG3* to *CCDC106*, causing downregulation of the former and upregulation of the latter, which encodes a pro-proliferative factor. Another Hi-C study interrogating CCs also demonstrated significant alteration in 3D genome architecture, with A/B compartments switching in 24% of the host genome as compared to normal tissues [72]. Consistent with previous findings, genes neighboring HPV integration breakpoints were differentially expressed and strongly associated with various cancer-related pathways [54, 72]. Finally, modulation of host chromosome 3D structure by viral integration may be partially attributable to CTCF recruitment to conserved binding sites in the viral genome, resulting in reorganization of chromatin interactions [73].

In summary, HPV episomal structure is largely dependent on CTCF occupancy, disruption of which can be sufficient to upregulate viral E6/E7 oncogenes (Fig. 2D). Notably, CTCF binding sites are often mutated upon viral integration, causing overexpression of E6/E7. The preference of viral integration sites for actively transcribing host genomic regions enables HPV to hijack host regulatory elements and machinery to drive viral gene

expression. Alternatively, the integrated viral sequence can act as an enhancer or boundary element, leading to dysregulation of host gene expression via alteration of short- and long-range chromosomal interactions. The host genes that are impacted by the above viral mechanisms are often cancer-associated, and the expression changes of host and viral (onco)genes collectively contribute to the transformation of infected cells (Fig. 2E).

HIV-1

Human immunodeficiency virus type 1 (HIV-1) is a retrovirus with a positive-sense, single-stranded RNA genome that can result in Acquired Immunodeficiency Syndrome (AIDS). Long-term maintenance of HIV-1 requires stable integration of the viral genome into a host chromosome via a viral integrase-dependent mechanism during its life cycle.

HIV-1 Integration

HIV-1 preferentially integrates in active genes in euchromatin that selectively localizes in spatial proximity to the nuclear pore compartment [74, 75]. The integrated HIV-1 provirus can enter a latent state that is facilitated by the deposition of repressive chromatin marks (*i.e.*, H3K27 and H3K9 trimethylation) and concomitant transcriptional silencing. Alternatively, integrated HIV-1 proviruses can remain transcriptionally and translationally active, irrespective of replication capacity, which most have lost due to acquired deletions and other mutations. Differential transcription of individual proviruses has been assessed systematically with a sequencing-based approach, dubbed barcoded HIV ensembles (B-HIVE), which revealed position effects [76]. In these studies, proviral expression levels correlated best with the presence of a nearby host genomic enhancer (< 5kb). Furthermore, inhibiting the interaction between HIV-1 integrase and the chromatin-tethering factor LEDGF/p75, which associates with the H3K36me3 chromatin mark characteristic of actively transcribed genes, attenuated proviral expression and increased the amount of silent provirus, which tended to integrate further away from H3K36me3-marked sites [77]. These results indicate HIV-1 proviral integration and gene expression are regulated by the host epigenomic landscape and raise the possibility of reciprocal effects involving host genomic architecture.

HIV-1 Integration Sites and Host Chromatin Organization

3C-based approaches were used to elucidate viral and host chromatin organization and their relationship to viral gene expression in HIV-1-infected cells. An early 3C study of HIV-1 provirus structure showed loop formation between the 5' and 3' long terminal repeats (LTRs), which did not require proviral insertion but was strongly linked to HIV-1 transcription [78]. In addition, 4C-seq data using HIV-1 provirus as the viewpoint demonstrated a frequent chromosomal interaction between integrated HIV-1 and the pericentromeric region of chromosome 12, which correlated with transcriptional silencing in the latent state [79]. Recent Hi-C analysis of HIV-1-infected cells revealed that genes recurrently targeted by the virus are proximal to super-enhancer genomic elements that tend to cluster spatially in the nuclei of CD4⁺ T cells [80]. Because HIV-1 can invade the central nervous system, productively infecting macrophages and microglia therein, the HIV-1 integration pattern in relation to the 3D genomic architecture of microglia cells was also investigated by Hi-C [81]. The results further strengthened the notion that HIV-1 integration preferentially occurs at active, open chromatin, including CTCF-bound regions with active histone marks positioned proximal to TAD boundaries. Finally, a comprehensive analysis of epigenomic, transcriptomic, and 3D genome architecture (Hi-C) datasets in combination with machine learning was undertaken to discern patterns characterizing HIV-1 proviral transcription. Specific chromatin states, active nuclear sub-compartments, and unique positions as well as orientations with respect to human genes and regulatory elements were all found to correlate with proviral transcriptional activity in CD4⁺ T cells [82].

Influenza Virus

Influenza viruses feature a negative-sense, single-stranded RNA genome and are a common cause of respiratory infections. Transcription and replication of influenza virus occur in the nucleus of host cells and can affect cell function. Interestingly, although the influenza viral genome does not integrate and is not known to impact host chromatin organization via direct interactions, an *in situ* Hi-C time course of influenza A (IAV)-infected human monocyte-derived macrophages demonstrated that infection significantly alters host 3D genomic architecture [83]. In a subset of highly transcribed host genes, the viral NS1 protein induces read-through transcription by inhibition of transcription termination, which enables RNA polymerase II to disrupt downstream chromatin loops by discharging cohesin from CTCF-bound sites that serve as chromatin anchors, leading to locus decompaction that is more permissive to TF binding (Fig. 2F, left). Another study, using 4C-seq and DNA FISH data, reported an alternative viral strategy targeting cohesin, which is mediated by the IAV protein NP and impacts host chromosomal organization to enhance viral replication [84]. In this scenario, the viral protein NP binds to the host histone H4 methyltransferase Suv4-20h2, disrupting the association of the latter with cohesin. Suv4-20h2 normally suppresses transcription of the *HoxC8-HoxC6* loci by preventing cohesin-dependent looping that has a stimulatory effect on gene expression. Because the encoded HoxC8 and HoxC6 proteins promote IAV replication, active loop formation at the *HoxC8-HoxC6* loci, as a consequence of NP-dependent disruption of the Suv4-20h2-cohesin interaction, contributes to influenza-induced pathology [85] (Fig. 2F, right).

SARS-CoV-2

Since its emergence in late 2019, SARS-CoV-2, a positive-sense, single-stranded RNA virus, has resulted in >600 million cases of its associated disease COVID-19 and more than ~6.5 million deaths worldwide as of

September 2022. Given that the pandemic virus can elicit long-term sequelae in a significant subset of infected individuals, even in the absence of severe disease initially, it is crucial to elucidate the relevant viral-host regulatory strategies in infected cells.

Host Chromatin Organization and COVID-19 Severity

Although a direct association of the SARS-CoV-2 genome or viral gene products with host chromatin has not been described yet, a few reports suggest a functional interplay between the virus and host chromatin structure. An example of host genomic organization affecting SARS-CoV-2 pathogenesis has been characterized for the *IL-6* locus. The cytokine *IL-6* contributes to hyper-inflammation and its levels strongly correlate with the severity of COVID-19. CTCF binding at the *IL-6* enhancer promotes *IL-6* expression. A SNP that disrupts the conserved CTCF-binding site has been shown to attenuate *IL-6* induction in COVID-19, thereby providing protection from severe outcomes [86]. Similarly, genome-wide association studies (GWAS) have identified 3p21.31 as the most robust risk locus for severe COVID-19 [87]. At this locus, a collection of chemokine receptors (*CCR1*, *CCR2*, and *CCR5*) and multiple enhancer elements were found to cluster as a CTCF-dependent active chromatin hub [88]. Rapid degradation of CTCF by an auxin-inducible degron (mAID) decommissioned the chromatin hub and selectively reduced expression of the constituent chemokine receptors in macrophages. Because high expression of chemokine receptor *CCR2* in monocytes and macrophage is an established pathologic hallmark of severe COVID-19 [89, 90], CTCF-mediated 3D chromatin interactions play a critical role in the acute respiratory symptoms of COVID-19 and may also directly or indirectly impact other affected tissues.

Effect of SARS-CoV-2 Infection on Host Chromosome Architecture

Recent reports also suggest that SARS-CoV-2 can affect the host epigenomic landscape and genomic organization. Comprehensive transcriptomic analysis of multiple infected cell lines and COVID-19 patient-derived samples combined with pathway enrichment analysis implicated over 60 epigenetic response proteins that are relevant to SARS-CoV-2 infection [91]. Additionally, differential DNA methylation patterns in immune cells that distinguish COVID-19 convalescents from uninfected controls have been reported [92]. Finally, changes in host 3D chromatin architecture upon SARS-CoV-2 infection have been assessed by in situ Hi-C. Quantification of the pairwise interactions between chromosome regions showed that SARS-CoV-2 infection alters host chromosome organization to induce certain pro-inflammatory genes and suppress antiviral interferon-responsive genes [93]. These changes are accompanied by weakening/switching of A/B compartments and intra-TAD chromatin contacts. In this study, alteration of chromatin compartmentalization was observed in as much as 30% of genomic regions. The molecular basis of this widespread impact of SARS-CoV-2 on host 3D genome architecture may be attributable, at least in part, to selective loss of cohesin at intra-TAD regions, but further investigation will be required to uncover the mechanistic details.

Conclusion and Outlook

The application of evolving 3C-based technologies combined with high-throughput sequencing in recent years has yielded substantial insights into 3D genomic architectural changes consequent to viral infection, but much remains to be learned. While early findings clearly show the potential for DNA viruses, whether as chromatinized episomes or upon integration, to impact host genomic architecture, RNA viruses, even in the absence of a capacity to integrate, also can elicit alterations in host chromatin structure (Table 2). The molecular mechanisms employed by RNA viruses, which are typically limited to acute infection, are likely varied but may still converge on common host architectural factors, such as CTCF and cohesin. Indeed, recent evidence suggests that IAV and SARS-CoV-2 infections can perturb CTCF and/or cohesin distribution to cause alterations in host chromosomal topology [83, 84, 93], while CTCF/cohesin-dependent host genome organization may contribute to the pathology associated with these RNA viruses [86, 88]. Nevertheless, further investigation is needed to elucidate the mechanistic details by which non-integrating RNA viruses affect host chromatin architecture to modulate gene expression and promote disease. Although RNA viruses generally feature smaller genomes than their DNA counterparts and therefore tend to encode fewer gene products, both viral proteins and non-coding RNAs could contribute to alterations in host chromatin profiles and chromosomal organization. Notably, CTCF localization can be influenced by its RNA-binding capacity, which could have concomitant effects on chromosome conformation [94]. While molecular mechanisms underlying conformational changes in host as well as viral genomes have frequently implicated CTCF and the associated cohesin complex, alternative strategies, involving, for example, condensin, which unlike cohesin does not rely on CTCF to determine its genomic distribution in mammalian genomes [95], YY1 [59] and other factors with possibly unappreciated architectural roles likely await future discovery.

The functional significance of DNA virus-elicited changes in host genomic structure has been linked to oncogenesis and best characterized in cancer cells but almost certainly applies to other pathological states, such as neurological disease. Notably, there is accumulating evidence of a viral etiology for at least a subset of multiple sclerosis (MS) [96] and Alzheimer's disease (AD) [97, 98] cases, possibly involving specific herpesviruses in each pathological context and likely manifesting in combination with certain genetic susceptibility loci. The potential effects of viral genomes or gene products on host genomic organization, which may be cell-type specific within the brain, and the concomitant altered expression of susceptibility genes have not been addressed in these central nervous system (CNS) pathologies.

The 3C-derived strategies that have informed our current understanding of the impact of viral infection on host 3D genome structure and possible reciprocal effects on viral genomic architecture have clear limitations, primarily related to the necessary inclusion of crosslinking and ligation steps that can introduce artifacts as well as

the logistics and cost considerations that allow for sampling of only one or, at best, a few time points. Alternative methods, such as SPRITE and GAM, still require crosslinking and have not been widely adopted yet, possibly due to technical challenges and accessibility. Future insights will likely involve additional approaches, including advanced microscopy methods and analytical pipelines that, when combined with clustered, regularly interspaced, short palindromic repeats (CRISPR)-based genomic modification and visualization strategies [99, 100], can yield simultaneous, real-time data on specific transcriptional events as well as nuclear spatial organization and the structural dynamics of both the host and viral genomes in infected cells.

Acknowledgments

Soohwan Oh is supported by a Korea University Grant (K2223251) and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2022R1C1C101069911).

Conflict of Interest

The authors have no financial conflicts of interest to declare.

References

- Benabdallah NS, Williamson I, Illingworth RS, Kane L, Boyle S, Sengupta D, *et al.* 2019. Decreased enhancer-promoter proximity accompanying enhancer activation. *Mol. Cell* **76**: 473–484.e7.
- Zheng H, W Xie. 2019. The role of 3D genome organization in development and cell differentiation. *Nat. Rev. Mol. Cell Biol.* **20**: 535–550.
- Kubo N, Ishii H, Xiong X, Bianco S, Meitinger F, Hu R, *et al.* 2021. Promoter-proximal CTCF binding promotes distal enhancer-dependent gene activation. *Nat. Struct. Mol. Biol.* **28**: 152–161.
- Oh S, Shao J, Mitra J, Xiong F, D'antonio M, Wang R, *et al.* 2021. Enhancer release and retargeting activates disease-susceptibility genes. *Nature* **595**: 735–740.
- Simonis M, Klous P, Splinter E, Moshikin Y, Willemsd R, Wit Ed, *et al.* 2006. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat. Genet.* **38**: 1348–1354.
- Dostie J, Richmond TA, Arnaout RA, Selzer RR, Lee WL, Honan TA, *et al.* 2006. Chromosome conformation capture carbon copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res.* **16**: 1299–1309.
- Jäger R, Migliorini G, Henrion M, Kandaswamy R, Speedy HE, Heindl A, *et al.* 2015. Capture Hi-C identifies the chromatin interactome of colorectal cancer risk loci. *Nat. Commun.* **6**: 6178.
- Fang R, Yu M, Li G, Chee S, Liu T, Schmitt AD, *et al.* 2016. Mapping of long-range chromatin interactions by proximity ligation-assisted ChIP-seq. *Cell Res.* **26**: 1345–1348.
- Fullwood MJ, Wei C-L, Ruan LY. 2009. Next-generation DNA sequencing of paired-end tags (PET) for transcriptome and genome analyses. *Genome Res.* **19**: 521–532.
- Sati S, G Cavalli. 2017. Chromosome conformation capture technologies and their impact in understanding genome function. *Chromosoma* **126**: 33–44.
- Hsieh TH, Weiner A, Lajoie B, Dekker J, Fridman N, Rando OJ. 2015. Mapping nucleosome resolution chromosome folding in yeast by micro-C. *Cell* **162**: 108–119.
- Quinodoz SA, Bhat P, Chovanec P, Jachowicz JW, Ollikainen N, Detmar E, *et al.* 2022. SPRITE: a genome-wide method for mapping higher-order 3D interactions in the nucleus using combinatorial split-and-pool barcoding. *Nat. Protoc.* **17**: 36–75.
- Beagrie RA, Scialdone A, Schueler M, Kraemer DCA, Chotalia M, Xie SQ, *et al.* 2017. Complex multi-enhancer contacts captured by genome architecture mapping. *Nature* **543**: 519–524.
- Consortium IH. 2003. The international HapMap project. *Nature* **426**: 789–796.
- Consortium EP. 2021. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**: 57–74.
- Arvey A, Tempera I, Tsai K, Chen H-S, Tikhmyanova n, Klichinsky M, *et al.* 2012. An atlas of the Epstein-Barr virus transcriptome and epigenome reveals host-virus regulatory interactions. *Cell Host Microbe*. **12**: 233–245.
- Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, *et al.* 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**: 1665–1680.
- Quinodoz SA, Ollikainen N, Tabak B, Palla A, Schmidt JM, Detmar E, *et al.* 2018. Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. *Cell* **174**: 744–757.e24.
- Buschle A, W Hammerschmidt. 2020. Epigenetic lifestyle of Epstein-Barr virus. *Semin Immunopathol.* **42**: 131–142.
- Toth Z, K Brulois, JU Jung. 2013. The chromatin landscape of Kaposi's sarcoma-associated herpesvirus. *Viruses* **5**: 1346–1373.
- Chen HS, Wikramasinghe P, Showe L, Lieberman PM. 2012. Cohesins repress Kaposi's sarcoma-associated herpesvirus immediate early gene transcription during latency. *J. Virol.* **8**: 9454–9464.
- Zhao B, Zou J, Wang H, Jphannsen E, Peng C-W, Quackenbush Q, *et al.* 2011. Epstein-Barr virus exploits intrinsic B-lymphocyte transcription programs to achieve immortal cell growth. *Proc. Natl. Acad. Sci. USA* **108**: 14902–14907.
- Chau CM, Zhang X-Y, McMahon SB, Lieberman PM. 2006. Regulation of Epstein-Barr virus latency type by the chromatin boundary factor CTCF. *J. Virol.* **80**: 5723–5732.
- Hughes DJ, Marendy Em, Dickerson CA, Yetming KD, Sample CE, Sample JT. 2012. Contributions of CTCF and DNA methyltransferases DNMT1 and DNMT3B to Epstein-Barr virus restricted latency. *J. Virol.* **86**: 1034–1045.
- Tempera I, Wiedmer A, Dheekollu D, Lieberman PM. 2010. CTCF prevents the epigenetic drift of EBV latency promoter Qp. *PLoS Pathog.* **6**: e1001048.
- Chen HS, Martin KA, Lu F, Lupey LN, Mueller JM, Lieberman PM, *et al.* 2014. Epigenetic deregulation of the LMP1/LMP2 locus of Epstein-Barr virus by mutation of a single CTCF-cohesin binding site. *J. Virol.* **88**: 1703–1713.
- Molyneux EM, Rochford R, Griffin B, Newton R, Jackson G, Menon G, *et al.* 2012. Burkitt's lymphoma. *Lancet* **379**: 1234–1244.
- Wang LW, Shen H, Nobre I, Ersing I, Paulo JA, Trudeau S, *et al.* 2019. Epstein-Barr-v-induced one-carbon metabolism drives B cell transformation. *Cell Metab.* **30**: 539–555.e11.
- Wood CD, Veenstra H, Khasnis S, Gunnell A, Webb HM, Shannon-Lowe C, *et al.* 2016. MYC activation and BCL2L1 silencing by a tumour virus through the large-scale reconfiguration of enhancer-promoter hubs. *Elife* **5**: e18270.
- Jiang S, Zhou H, liang J, Gerdt C, Wang C, Ke L, *et al.* 2017. The Epstein-Barr virus regulome in lymphoblastoid cells. *Cell Host Microbe* **22**: 561–573.e4.
- Guo R, Jiang C, Zhang Y, Govande A, Trudeau SJ, Chen F, *et al.* 2020. MYC controls the epstein-barr virus lytic switch. *Mol. Cell* **78**: 653–669.e8.
- Knipe DM, Lieberman PM, Jung JU, McBride AA, Morris KV, Ott M, *et al.* 2013. Snapshots: chromatin control of viral infection. *Virology* **435**: 141–156.

33. Stedman W, Kang H, Lin S, Kissil JL, Bartolomei MS, Lieberman PM. 2008. Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators. *EMBO J.* **27**: 654-666.
34. Li DJ, Verma D, Mosbrugger T, Swaminathan. 2014. CTCF and Rad21 act as host cell restriction factors for Kaposi's sarcoma-associated herpesvirus (KSHV) lytic replication by modulating viral gene transcription. *PLoS Pathog.* **10**: e1003880.
35. Park A, Oh S, Jung KL, Choi UY, Lee H-R, Rosenfeld MG, et al. 2020. Global epigenomic analysis of KSHV-infected primary effusion lymphoma identifies functional MYC superenhancers and enhancer RNAs. *Proc. Natl. Acad. Sci. USA* **117**: 21618-21627.
36. Bloom DC, NV Giordani, DL Kwiatkowski. 2010. Epigenetic regulation of latent HSV-1 gene expression. *Biochim. Biophys. Acta* **1799**: 246-256.
37. Ertel MK, Cammarata AL, Hron RJ, Neumann DM. 2012. CTCF occupation of the herpes simplex virus 1 genome is disrupted at early times postreactivation in a transcription-dependent manner. *J. Virol.* **86**: 12741-12759.
38. Martínez FP, Cruz R, Lu F, Plasschaert R, Deng Z, Rivera-Molina YA, et al. 2014. CTCF binding to the first intron of the major immediate early (MIE) gene of human cytomegalovirus (HCMV) negatively regulates MIE gene expression and HCMV replication. *J. Virol.* **88**: 7389-7401.
39. Lee JS, Raja P, Pan D, Pesola JM, Coen DM, Knioe DM. 2018. CCCTC-binding factor acts as a heterochromatin barrier on herpes simplex viral latent chromatin and contributes to poised latent infection. *mBio* **9**: e02372-17.
40. Lang F, Li X, Vladimirova O, Hu B, Chen G, Xiao Y, Singh V, et al. 2017. CTCF interacts with the lytic HSV-1 genome to promote viral transcription. *Sci. Rep.* **7**: 39861.
41. Li X, Yu Y, Lang F, Chen G, Wang E, Li L, et al. 2021. Cohesin promotes HSV-1 lytic transcription by facilitating the binding of RNA Pol II on viral genes. *Virol. J.* **18**: 26.
42. Morgan SM, Tanizawa H, Caruso LB, Hulse M, Kossenkova A, Madzo J, et al. 2022. The three-dimensional structure of Epstein-Barr virus genome varies by latency type and is regulated by PARP1 enzymatic activity. *Nat. Commun.* **13**: 187.
43. Campbell M, Watanabe T, Nakano K, Davis R, Lyu Y, Tepper CG, et al. 2018. KSHV episomes reveal dynamic chromatin loop formation with domain-specific gene regulation. *Nat. Commun.* **9**: 49.
44. Campbell M, Chantarasrivong C, Yanagihashi Y, Inagaki T, Davis RR, Nakano K, et al. 2022. KSHV topologically associating domains in latent and reactivated viral chromatin. *J. Virol.* **96**: e0056522.
45. Moquin SA, Thomas S, Whalen S, Warburton, Fernandez SG, McBride AA, et al. 2018. The Epstein-Barr virus episome maneuvers between nuclear chromatin compartments during reactivation. *J. Virol.* **92**: e01413-17.
46. Kim, KD, Tanizawa H, De Leo A, Vladimirova O, Kossenkova A, Lu F, et al. 2020. Epigenetic specifications of host chromosome docking sites for latent Epstein-Barr virus. *Nat. Commun.* **11**: 877.
47. Okabe A, Huang KK, Matsusaka K, Fukuyo M, Xing M, Ong X, et al. 2020. Cross-species chromatin interactions drive transcriptional rewiring in Epstein-Barr virus-positive gastric adenocarcinoma. *Nat. Genet.* **52**: 919-930.
48. Caruso LB, Gou R, Keith K, Madzo J, Maestri D, Boyle S, et al. 2022. The nuclear lamina binds the EBV genome during latency and regulates viral gene expression. *PLoS Pathog.* **18**: e1010400.
49. Wang L, Laing J, Yan B, Zhou H, Ke L, Wang C, et al. 2020. Epstein-Barr virus episome physically interacts with active regions of the host genome in lymphoblastoid cells. *J. Virol.* **94**: e01390-20.dk.
50. Kumar A, Lyu Y, Yanagihashi Y, Chantarasrivong C, Majerciak V, Salem M, et al. 2022. KSHV episome tethering sites on host chromosomes and regulation of latency-lytic switch by CHD4. *Cell Rep.* **39**: 110788.
51. Ballestas ME, KM Kaye. 2011. The latency-associated nuclear antigen, a multifunctional protein central to Kaposi's sarcoma-associated herpesvirus latency. *Future Microbiol.* **6**: 1399-1413.
52. Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, Snijder PJ, et al. 1999. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J. Pathol.* **189**: 12-19.
53. McLaughlin-Drubin ME, K Münger. 2009. Oncogenic activities of human papillomaviruses. *Virus Res.* **143**: 195-208.
54. Cao C, Hong P, Huang X, Lin D, Cao G, Wang L, et al. 2020. HPV-CCDC106 integration alters local chromosome architecture and hijacks an enhancer by three-dimensional genome structure remodeling in cervical cancer. *J. Genet. Genomics* **47**: 437-450.
55. Zhou, C., Z.K. Tuong, and I.H. Frazer, 2019. Papillomavirus immune evasion strategies target the infected cell and the local immune system. *Front. Oncol.* **2019**: 9: 682.
56. Mehta K, Gunasekharan V, Satsuka A, Laimins LA. 2015. Human papillomaviruses activate and recruit SMC1 cohesin proteins for the differentiation-dependent life cycle through association with CTCF insulators. *PLoS Pathog.* **11**: e1004763.
57. Mehta K, Laimins. 2018. Human papillomaviruses preferentially recruit DNA repair factors to viral genomes for rapid repair and amplification. *mBio* **9**: e00064-18.
58. Paris C, P Pe ntland I, Groves I, Roberts DC, Powis SJ, Coleman N, et al. 2015. CCCTC-binding factor recruitment to the early region of the human papillomavirus 18 genome regulates viral oncogene expression. *J. Virol.* **89**: 4770-4785.
59. Pentland I, Compo-Leon K, Cotic M, Davies K-J, Wood CD, Groves IJ, et al. 2018. Disruption of CTCF-YY1-dependent looping of the human papillomavirus genome activates differentiation-induced viral oncogene transcription. *PLoS Biol.* **16**: e2005752.
60. Ferguson J, Campos-Leon K, Pe ntland I, Stockton JD, Cunther T, Beggs AD, et al. 2021. The chromatin insulator CTCF regulates HPV18 transcript splicing and differentiation-dependent late gene expression. *PLoS Pathog.* **17**: e1010032.
61. McBride AA, A Warburton. 2017. The role of integration in oncogenic progression of HPV-associated cancers. *PLoS Pathog.* **13**: e1006211.
62. Ojesina AI, Lichtenstein L, Freeman SS, Pedamallu CS, Imaz-Rosshandler I, Pugh TJ, et al. 2014. Landscape of genomic alterations in cervical carcinomas. *Nature* **506**: 371-375.
63. Dixon JR, Xu J, Dileep V, Zhan Y, Song F, Le VT, et al. 2018. Integrative detection and analysis of structural variation in cancer genomes. *Nat. Genet.* **50**: 1388-1398.
64. Shen C, Liu Y, Shi S, Zhang R, Zhang T, Xu Q, et al. 2017. Long-distance interaction of the integrated HPV fragment with MYC gene and 8q24.22 region upregulating the allele-specific MYC expression in HeLa cells. *Int. J. Cancer* **141**: 540-548.
65. Bodelon C, Untereiner ME, Machiela MJ, Vinokurova, Wentzensen N. 2016. Genomic characterization of viral integration sites in HPV-related cancers. *Int. J. Cancer* **139**: 2001-2011.
66. Kelley DZ, Flam EL, Izumchenko E, Danilova LV, Wulf HA, Guo T, et al. 2017. Integrated analysis of whole-genome ChIP-Seq and RNA-Seq data of primary head and neck tumor samples associates HPV integration sites with open chromatin marks. *Cancer Res.* **77**: 6538-6550.
67. Hu Z, Zhu D, Wang W, Li W, Jia W, Zeng X, et al. 2015. Genome-wide profiling of HPV integration in cervical cancer identifies clustered genomic hot spots and a potential microhomology-mediated integration mechanism. *Nat. Genet.* **47**: 158-163.
68. Warburton A, Markowitz TE, Katz JP, Pipas JM, McBride A. 2021. Recurrent integration of human papillomavirus genomes at transcriptional regulatory hubs. *NPJ Genom. Med.* **6**: 101.
69. Warburton A, Redmond CJ, Dooley KE, Fu H, Gillson ML, Akagi K, et al. 2018. HPV integration hijacks and multimerizes a cellular enhancer to generate a viral-cellular super-enhancer that drives high viral oncogene expression. *PLoS Genet.* **14**: e1007179.
70. Zhou L, Qiu Q, Zhou Q, Li J, Yu M, Li K, et al. 2022. Long-read sequencing unveils high-resolution HPV integration and its oncogenic progression in cervical cancer. *Nat. Commun.* **13**: 2563.

71. Turner KM, Deshpande V, Beyter D, Koga T, Risert J, Lee C, *et al.* 2017. Extrachromosomal oncogene amplification drives tumour evolution and genetic heterogeneity. *Nature* **543**: 122-125.
72. Adeel MM, Jiang H, Arega Y, Cao K, Lin D, Cao C, *et al.* 2021. Structural variations of the 3D genome architecture in cervical cancer development. *Front. Cell Dev. Biol.* **9**: 706375.
73. Mehran Karimzadeh, Christopher Arlidge, Ariana Rostami, Mathieu Lupien, Scott V. Bratman, Michael M. Hoffman. 2022. Human papillomavirus integration transforms chromatin to drive oncogenesis. <https://doi.org/10.1101/2020.02.12.942755>.
74. Lusic M, RF Siliciano. 2017. Nuclear landscape of HIV-1 infection and integration. *Nat. Rev. Microbiol.* **15**: 69-82.
75. Marini B, Kertesz-Farkas A, Ali H, Lucic B, Lisek K, Manganaro L, *et al.* 2015. Nuclear architecture dictates HIV-1 integration site selection. *Nature* **521**: 227-231.
76. Chen H-C, Martinez JB, Zorita E, Meyerhans A, Filion GJ. 2017. Position effects influence HIV latency reversal. *Nat. Struct. Mol. Biol.* **24**: 47-54.
77. Vansant G, Chen H-C, Zorita E, Trejbalova K, Millik D, Filion G, *et al.* 2020. The chromatin landscape at the HIV-1 provirus integration site determines viral expression. *Nucleic Acids Res.* **48**: 7801-7817.
78. Perkins KJ, Lusic M, Mitar I, Giacca M, Proudfoot NJ. 2008. Transcription-dependent gene looping of the HIV-1 provirus is dictated by recognition of pre-mRNA processing signals. *Mol. Cell* **29**: 56-68.
79. Dieudonné M, Maiuri P, Biancotto C, Knezevich A, Kula A, Lusic M, *et al.* 2009. Transcriptional competence of the integrated HIV-1 provirus at the nuclear periphery. *EMBO J.* **28**: 2231-2243.
80. Lucic B, Chen H-C, Kuzman M, Zorita E, Wegner J, Minneker V, *et al.* 2019. Spatially clustered loci with multiple enhancers are frequent targets of HIV-1 integration. *Nat. Commun.* **10**: 4059.
81. Rheinberger M, Costa AL, Kampmann M, Glavas D, Shytaj LL, Penzo C, *et al.* 2022. Genomic profiling of HIV-1 integration in microglia links viral insertions to TAD organization. *bioRxiv* doi: <https://doi.org/10.1101/2022.02.14.480322>.
82. Ruess H, Lee J, Guzman C, Malladi V, D'Orso I. 2020. An integrated genomics approach towards deciphering human genome codes shaping HIV-1 proviral transcription and fate. *bioRxiv* doi: <https://doi.org/10.1101/2020.05.01.072231>.
83. Heinz S, Texari L, Hayes MCB, Urbanowski M, Chang MW, Givarkes N, *et al.* 2018. Transcription elongation can affect genome 3D structure. *Cell* **174**: 1522-1536.e22.
84. Shiimori M, Ichida Y, Nukiwa R, Sakuma T, Abe H, Kajitani R, *et al.* 2021. Suv4-20h2 protects against influenza virus infection by suppression of chromatin loop formation. *iScience* **24**: 102660.
85. Shiimori M, Nukiwa R, Imai Y. 2021. Dynamics of the host chromatin three-dimensional response to influenza virus infection. *Int. Immunol.* **33**: 541-545.
86. Chen T, Lin Y-X, Zha Y, Sun Y, Tian J, Yang Z, *et al.* 2021. A low-producing haplotype of interleukin-6 disrupting CTCF binding is protective against severe COVID-19. *mBio* **12**: e0137221.
87. Nakanishi T, Pigaxxini S, Degehardt F, Cordi, Butler-Laporte G, Maya-Miles D, *et al.* 2021. Age-dependent impact of the major common genetic risk factor for COVID-19 on severity and mortality. *J. Clin. Invest.* **131**: e152386.
88. Stikker BS, Stik G, Ouwerkerk AFV, Trap L, Spicuglia S, Hendriks RW, *et al.* 2022. Severe COVID-19-associated variants linked to chemokine receptor gene control in monocytes and macrophages. *Genome Biol.* **23**: 96.
89. Merad M, Martin JC. 2020. Pathological inflammation in patients with COVID-19: a key role for monocytes and macrophages. *Nat. Rev. Immunol.* **20**: 355-362.
90. Szabo PA, Dogra P, Gray JJ, Wells SB, Connors TJ, Weisberg SP, *et al.* 2021. Longitudinal profiling of respiratory and systemic immune responses reveals myeloid cell-driven lung inflammation in severe COVID-19. *Immunity* **54**: 797-814.e6.
91. Salgado-Albarrán M, Navarro-Delgado EI, Moral-Morales AD, Alcaraz N, Baumbach J, Gonzalez-Barrios R, *et al.* 2021. Comparative transcriptome analysis reveals key epigenetic targets in SARS-CoV-2 infection. *NPJ Syst. Biol. Appl.* **7**: 21.
92. Huoman J, sayyab S, Apostolou E, Karlsson L, Procile L, Rizwan M, *et al.* 2021. Epigenome-wide DNA methylation profiling of healthy COVID-19 recoverees reveals a unique signature in circulating immune cells. *MedRxiv* doi: <https://doi.org/10.1101/2021.07.05.21260014>.
93. Wang R, Lee J-H, Xiong F, Kim J, Hasani LA, Yuan X, *et al.* 2021. SARS-CoV-2 Restructures the host chromatin architecture. *bioRxiv* doi: [10.1101/2021.07.20.453146](https://doi.org/10.1101/2021.07.20.453146).
94. Merkenschlager M, Nora EP. 2016. CTCF and cohesin in genome folding and transcriptional gene regulation. *Annu. Rev. Genomics Hum. Genet.* **17**: 17-43.
95. Yuen, K.C. and J.L. Gerton, 2018. Taking cohesin and condensin in context. *PLoS Genet.* **14**: e1007118.
96. Bjornevik K, Cortese M, Healy BC, Kuhle J, Mina MJ, Leng Y, *et al.* 2022. Longitudinal analysis reveals high prevalence of Epstein-Barr virus associated with multiple sclerosis. *Science* **375**: 296-301.
97. Itzhaki RF. 2022. Does antiherpetic antiviral therapy reduce the risk of dementia? *Nat. Rev. Neurol.* **18**: 63-64.
98. Young-Xu Y, Powell EI, Zwain GM, Yazdi MT, Gui J, Shiner B. 2021. Symptomatic herpes simplex virus infection and risk of dementia in US veterans: a cohort study. *Neurotherapeutics* **18**: 2458-2467.
99. Ma H, Naseri A, Reyes-Gutierrez P, Wolfe S, Zhang S, Pederson T. 2015. Multicolor CRISPR labeling of chromosomal loci in human cells. *Proc. Natl. Acad. Sci. USA* **112**: 3002-3007.
100. Chen B, Gillbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li G-W, *et al.* 2013. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* **155**: 1479-1491.
101. Washington SD, Edenfield SI, Lieux C, Watson ZL, Taasan SM, Dhumakupt A, *et al.* 2018. Depletion of the insulator protein CTCF results in herpes simplex virus 1 reactivation. *J. Virol.* **92**: e00173-18.
102. Lupey-Green LN, Moquin SA, Martin KA, McDevitt SM, Hulse M, Caruso LB, *et al.* 2017. PARP1 restricts Epstein Barr Virus lytic reactivation by binding the BZLF1 promoter. *Virology* **507**: 220-230.
103. Kang H, Wiedmer A, Yuan Y, Robertson E, Lieberman PM. 2011. Coordination of KSHV latent and lytic gene control by CTCF-cohesin mediated chromosome conformation. *PLoS Pathog.* **7**: e1002140.
104. Jefferys SR, Burgos SD, Peterson JJ, Selitsky SR, Turner AMW, James LI, *et al.* 2021. Epigenomic characterization of latent HIV infection identifies latency regulating transcription factors. *PLoS Pathog.* **17**: e1009346.
105. D'Arienzo V, Ferguson J, Giraud G, ChapusF, Harris JM, Wing PAC, *et al.* 2021. The CCCTC-binding factor CTCF represses hepatitis B virus enhancer I and regulates viral transcription. *Cell. Microbiol.* **23**: e13274.
106. Decorsière A, Mueller H, van Breugel PC, Abdul F, Gerossier L, Beran RK, *et al.* 2016. Hepatitis B virus X protein identifies the Smc5/6 complex as a host restriction factor. *Nature* **531**: 386-389.
107. Komatsu T, Sekiya T, Nagata K. 2013. DNA replication-dependent binding of CTCF plays a critical role in adenovirus genome functions. *Sci. Rep.* **3**: 2187.