BMB Reports - Manuscript Submission

Manuscript Draft

Manuscript Number: BMB-22-154

Title: Topological implications of DNA tumor viral episomes

Article Type: Mini Review

Keywords: DNA tumor viruses; viral episome; 3D genomics; epigenetics; viral latency

Corresponding Author: Kyoung-Dong Kim

Authors: Kyoung-Dong Kim[3,*], Eui Tae Kim[1,2]

Institution: [1]Department of Microbiology and Immunology, Jeju National University College of Medicine, Jeju, 63241, Republic of Korea,
[2]Department of Biomedicine & Drug Development, Jeju National University, Jeju, 63241, Republic of Korea,
[3]Department of Systems Biotechnology, Chung-Ang University, Anseong, 17546, Republic of Korea,

Manuscript Type: Mini Review

Title: Topological implications of DNA tumor viral episomes

Author's name: Eui Tae Kim[1,2] and Kyoung-Dong Kim[3*]

Affiliation:

[1]Department of Microbiology and Immunology, Jeju National University College of Medicine, Jeju, 63241, Republic of Korea

[2]Department of Biomedicine & Drug Development, Jeju National University, Jeju, 63241, Republic of Korea

[3]Department of Systems Biotechnology, Chung-Ang University, Anseong, 17546, Republic of Korea

Running Title: Interplay of viral episomes and host chromosomes

Keywords: DNA tumor viruses, viral episome, 3D genomics, epigenetics, viral latency

Corresponding Author's Information:

+82-31-670-3359; kdkim0122@cau.ac.kr

ABSTRACT

A persistent DNA tumor virus infection transforms normal cells into cancer cells by either integrating its genome into host chromosomes or retaining it as an extrachromosomal entity called episome. Viruses have evolved mechanisms for attaching episomes to infected host cell chromatin to efficiently segregate the viral genome during mitosis. It has been reported that viral episome can affect the gene expression of the host chromosomes through interactions between viral episomes and epigenetic regulatory host factors. This mini review summarizes our current knowledge of the tethering sites of viral episomes, such as EBV, KSHV, and HBV, on host chromosomes analyzed by three-dimensional genomic tools.

INTRODUCTION

DNA viruses mainly maintain their genome as episomal DNA, which is important for viral replication and gene expression (1, 2). The tethered episomes affect the gene expression patterns on host chromosomes, leading to pathological consequences such as cancer development (3-6). Therefore, it is important to elucidate the tethering sites of viral episomes in host chromosomes. Recently developed chromosome conformation capture (3C) derived Next Generation Sequencing (NGS) methods allowed us to examine the association between viral episomes and human chromosomes. In this mini review, we summarized the basic features of DNA tumor viruses, viral episomes, and their positions on human chromosomes identified by 3C-derived methods to get an insight into the tethering mechanisms and impacts on host gene expression.

EPISOME OF TUMOR VIRUSES

DNA tumor viruses

Certain viruses can transform infected cells into cancerous ones. In order to gain an opportunity for tumorigenesis, viral genetic materials must persist within the host cells, which they typically do by forming an episomal structure. Tumor viruses include DNA viruses like Epstein–Barr virus (EBV), Kaposi's sarcoma-associated herpesvirus (KSHV), human papillomavirus (HPV), hepatitis B virus (HBV), and Merkel cell polyomavirus (MCPyV), as well as a few RNA viruses. This review will mainly discuss the episomal structure of DNA tumor viruses and their typical positions on host chromosomes related to gene expression regulation.

Specification of viral episomes

An episome is a segment of genetic material that can exist independently or integrate into the

host chromosome. Viral genomes exhibit remarkable diversity in terms of nucleic acid types, sizes, and complexity. DNA tumor viruses are double-stranded or partially single-stranded; they can be either linear or circular. The viral genomes are maintained in episomal form after infection, or some viral genomes are maintained by integration into the host chromosome.

EBV, which causes Burkitt lymphoma, has a large linear double-stranded DNA genome. The genome is around 172 kbp and encodes 80 proteins and 46 noncoding RNAs. EBV maintains its latency by keeping its chromatinized episomes in sync with the replication of the host chromosomes. The circularized viral chromosome is not integrated into the host genome and retains genomic stability while allowing the expression of a few viral genes essential for replication. Epigenetic status profoundly influences the expression of genes on episomes. Epigenetic modifications of the EBV genome occur during initial infection, latency, lytic replication, and virion production (7). Prior to the first round of EBV genome replication, the incoming EBV DNA rapidly circularizes and acquires nucleosomes in the infected cells. Episome assembly occurs during the G1 phase of host cells, long before the start of EBV-induced viral DNA replication (8).

KSHV was discovered as the causative agent of AIDS-associated Kaposi sarcoma (9). The KSHV genome is a linear double-stranded DNA. Upon infection, the linear viral DNA rapidly circularizes after entering the nucleus and is maintained as an episome (10). During latency, KSHV maintains 50–100 genome copies per infected cell (11). KSHV genomes replicate once every cell cycle in latent cells and are segregated into daughter cells. Episomal modification and nucleosome positioning play a role in both activation and inactivation of latent genes (12). On silenced episomes, transcription activation of the ORF50 immediate early gene (Rta) can initiate reactivation of the KSHV lytic cycle (13). ORF50 expression is repressed by the KSHV latency-associated nuclear antigen (LANA) during latency (12).

Activated ORF50 triggers the expression of early genes required for viral DNA replication, followed by the expression of late genes (14). On the other hand, for latent infection, KSHV episomes undergo methylation at CpG nucleotides in conjunction with particular histone modification marks, resulting in the rapid establishment of latency and suppression of lytic gene expression (15).

Due to the fact that HBV infection can result in liver cirrhosis, liver failure, hepatocellular carcinoma, and even death, it is considered one of the top 20 causes of human mortality (12). HBV comprises a partially double-stranded, 3.2 kbp circular DNA genome covalently linked to a multifunctional polymerase, with both RNA- and DNA-dependent polymerase functions as well as an RNase H function. HBV virions infect hepatocytes, and then the relaxed circular DNA (rcDNA) is transported to the nucleus. This form is converted into covalently closed circular DNA (cccDNA) that exists in an episomal state, some of which are not necessary for the viral replication cycle but are integrated into the host genome (16) The host RNA polymerase II then uses cccDNA as a template to make all viral RNAs. rcDNA is transported to the nucleus to convert and amplify cccDNA via an intracellular pathway (17). cccDNA does not appear to be attached to the host chromosome during mitosis; consequently, cccDNAs are randomly distributed between daughter cells, and some are lost during cell division (18).

Long-lasting infections with high-risk HPVs can develop cancer in areas where HPV infects cells, such as the cervix and oropharynx. HPV has a circular, chromatinized double-stranded DNA genome in a non-enveloped capsid. Unlike the two herpesviruses previously introduced, HPV is a virus that completes its entire production life cycle with a circular episome in infected basal epithelial cells. HPV genomes are maintained as a low copy as circular episome replicated alongside cellular DNA (19). Integration of HPV DNA is commonly

reported in related cancer genomes. However, both the integrated and episomal HPV genomes appear to be implicated in invasive cervical cancer (5). The mechanisms by which the HPV genome integrates into the host chromosome are still unknown.

MCPyV causes aggressive Merkel cell carcinoma (MCC), a rare skin cancer. MCPyV has a typical circular double-stranded DNA genome. The MCPyV genome is maintained as a replication-competent episome in persistently infected cells. During persistent infection, the virus resides and replicates as an episome in infected non-malignant cells. However, it has frequently been observed that viral DNA found in MCCs is integrated into the cellular genome.

The diversity of episomal maintenance is closely related to the viral life cycle, including DNA replication, transcriptional modulation, and genome segregation. In addition, studies of viral episome structure will provide a direction for potential therapeutic strategies because they are involved in cancer development or various immune responses by causing the regulation of host physiology.

Maintenance of viral episome

The goal of viruses is to replicate themselves, and some viruses pass their genomes on to the next generation along with the division of host cells. In order to accomplish this, viruses have developed various strategies to replicate their genomes and attach them to host chromosomes. Viral genome tethering is required for transporting the incoming viral genome into the nucleus or maintaining the genome as an episome in persistently infected cells.

For tethering of the viral episomes, proteins that bind the episome and the host chromosome are required. In addition to their functions for tethering, these viral episome maintenance proteins (EMPs) may also be involved in viral replication and transcription. Although well-known EMPs are usually encoded by viruses, cellular proteins involved in

organizing the chromosome architecture of host cells also play a role in viral episome maintenance. EBNA1 of EBV, LANA1 of KSHV, and E2 of HPV share common structural features and have an integrated function for stable segregation of episomes.

Episomal maintenance of EBV and KSHV has been well studied. EBNA1 is a viral protein expressed in all EBV-related tumors. It is necessary for viral DNA replication and episome maintenance while latently infected cells grow and divide (20). EBNA1 has two major domains in the amino (N)-terminal region with chromosome-tethering domains (CTDs) that bind to the minor groove of the AT-rich scaffold-associated region of the host chromosome (Fig. 1). A DNA-binding domain (DBD) existed in the carboxy (C)-terminal region of EBNA1 is responsible for sequence-specific DNA binding. It recognizes an 18 bp palindromic sequence found in several copies at the viral origin of plasmid replication (*oriP*) (21-23). EBV genome tethering can be achieved not only through direct recognition of the specific DNA sequences of these two domains but also through association with chromosome-binding proteins such as chromosome-associated EBP2, BRD4, RCC1, HMGB2, and PARP1(24-28).

LANA is the KSHV EMP. The C-terminal DBD of LANA1 binds to the terminal repeat region of the viral episome (29) (Fig. 1). In addition, LANA binds to the core histones H2A and H2B on the nucleosomal surface (30) and their interaction is essential for KSHV genome replication and persistence (31, 32). Cellular BUB1, DEK, NUMA, PARP1, and CHD4 appear to be involved in the tethering of the KSHV episome (33-37), but further studies are needed to clarify whether this is a direct role.

The HBV X (HBx) protein is essential to initiate and maintain viral replication after infection. HBx is mostly cytoplasmic, but a minor variable fraction is in the nucleus and recruited to the cccDNA episome and participates in the initiation of cccDNA-driven transcription. However, HBx does not bind DNA directly; rather, it seems to interact with the

host transcriptional machinery proteins that do (38).

The HPV E2 protein binds viral episomes to mitotic host chromosomes during cell division for partitioning and maintenance. E2 is composed of three regions: the N-terminal trans-activating domain (TAD), a hinge region, and the C-terminal DBD, which binds to several E2 binding sites on the viral episomes (Fig. 1) (39). The TAD and hinge region interact with host proteins on cellular chromosomes, and viral episomes are tethered to and stably maintained on mitotic chromosomes (40, 41).

Large and small T Antigens (LT- and ST-Ag, respectively) of MCPyV are expressed immediately upon nuclear delivery of viral episomes. These drive the cell cycle into S-phase, favorable for viral episome propagation (42). In addition, LT-Ag also possesses helicase activity and recruits host replication factors to the viral episome, functions that are essential for viral DNA replication (42). However, due to the greater focus on viral genome integration in MCC, episome tethering in MCPyV remains unexplored.

TETHERING SITES OF VIRAL EPISOMES ON HOST CHROMOSOMES

3C-derived methods to detect viral episomes tethering sites

The attachment site for episomal DNA cannot be detected by linear whole genome sequencing (WGS) since the episome is separated from the host chromosome, unlike viral integration sites. Advances in microscopic methods and the NGS technology made it possible to identify the position of viral episomes in the nucleus over the past decade. Fluorescence in situ hybridization (FISH) was used to identify the attachment sites of viral episomes on host chromosomes (43, 44). However, microscopic images can detect only partial sites among whole episomal attachment sites, which can be a piece of the puzzle.

3C-derived methods detect the topological structure of chromosomes (45). Briefly,

cells are fixed with formaldehyde and digested with a 4 bp cutter enzyme, and then fragmented DNAs are ligated with excessive ligase. The proximity of DNA fragments can be detected by PCR with a set of primers in 3C or by the NGS technology in Hi-C (Fig. 2) (46, 47). Hi-C provides ligation frequencies between whole genomic loci that can be computationally reconstructed into three-dimensional (3D) genomic organization. In addition to cellular genomic association, information about interactions between viral episomes and host chromosomes can be extracted from Hi-C data for cells infected with episomal viruses. Circular chromosome conformation capture (4C) is the method to detect the genomic association of one locus with whole genomic regions (48, 49). Because 4C only amplify specific associations between viral episomes and host chromosomes, approximately 100-fold fewer sequencing reads compared to the Hi-C method is required. Capture Hi-C (CHi-C) also enriches specific genomic positions linked to bait (50) similarly to 4C method. To detect the tethering sites of viral episome, CHi-C uses the biotinylated RNA bait library derived from the viral genome, allowing deep sequencing information for specific target loci linked to the viral genome (Fig. 2) (33, 50). Therefore, 4C and CHi-C are useful methods for the detection of tethering sites of viral episomes on the host chromosome. Moreover, these methods require less intensive computational works than the Hi-C analysis. Nevertheless, Hi-C would be beneficial if the tethering sites of viral episomes should be understood in the context of three-dimensional structure of the host genome (Fig. 2).

The position of viral episomes on host chromosomes

Viral gene expression is regulated by epigenetic changes in viral episomes through EMPs and associated proteins. The position of viral episomes on host chromosomes is being identified in several viral cases, including EBV, KSHV, and HBV, through the 3C-derived NGS methods (6,

33, 44, 51-56) and are summarized in Table 1.

The Burkitt lymphoma cell line showed the enrichment of EBV episomes on the transcriptionally repressed genomic region that coexisted with heterochromatic marker H3K9me3 (44). Moreover, transcriptional expression of EBV tethering genes was de-repressed when EBV episomes were dissociated from the linked genes. Mechanistically, the enrichment of H3K9me3 was significantly decreased in shEBNA1, which induces the dissociation of EBV episomes from host chromosomes (44). Therefore, EBV episome represses host gene expression mediated to heterochromatin complexes in the Burkitt lymphoma cell line. The tethering sites of EBV episomes in lymphoblastoid cell lines (LCLs) GM12878 differed from that of Burkitt lymphoma. The position in LCLs GM12878 was analyzed by 4C-seq and Hi-C methods and reproducibly confirmed to be located in active promoters and active histone markers such as H3K27ac, H3K4me1, and H3K4me3 (44, 53). The Burkitt lymphoma only expresses viral protein EBNA1, and represses other viral proteins, which belongs to the latency type I. On the contrary, LCL expresses all of the EBNAs as well as LMPs of EBV, belonging to the latency type III and similar conditions for viral reactivation of Akata-Zta cell, an EBV-positive BL (2, 57). Thus, the position of EBV episomes on host chromosomes may depend on the viral latency type.

The tethering sites of EBV in gastric carcinoma have been intensively examined through 4C-seq and Hi-C methods (6). The comparison analyses between EBV-associated gastric cancer cell lines (EBVaGC) and normal gastric epithelial cell lines revealed that EBV episome attachment induces heterochromatin to euchromatin transition. Furthermore, EBV-infected MKN7 and GES1 cells reproducibly demonstrated epigenetic redistribution from heterochromatin to euchromatin by association with EBV episomes. These results suggest that EBV episomes of gastric cancer cells are associated with the active enhancer region, which can

induce epigenetic reprogramming through an unknown mechanism.

The tethering sites of KSHV episomes in primary effusion lymphoma (PEL) cell lines, such as BC-1, BC-3, and BCBL-1, were identified by the CHi-C method (33). KSHV episomes are preferentially associated with near centromeric regions in all three KSHV infected PEL cell lines, which is consistent with the results that KSHV episomal maintenance protein LANA interacts and co-localizes with centromeric protein CENP-F and kinetochore protein BUB1 (33, 37). Kumar *et al.* identified CHD4 as an interaction partner with LANA through a proximity biotin labeling assay and observed that both KSHV episomal signals and LANA signals were enriched at CHD4 binding sites. Since LANA knockout KSHV results in aberrant lytic gene expression and dysregulation of host genes involved in cell cycle and proliferation pathways (58), considering the oncogenic function of CHD4, it is thought that there is a link between KSHV episome tethering and cancer development.

The tethering sites of HBV episomes are also not randomly distributed but localized at the specific genomic region The HBV episomes tend to be localized at active chromatin, such as CpG islands (CGIs), transcription start sites, and enhancers in HBV-infected hepatocytes (51, 55, 56). HBV protein HBx has central roles in the viral life cycle, including viral transcription, replication, and pathogenesis (54, 55, 59, 60). Recent studies analyzed the role of HBx in the tethering of viral episomes on the host chromosome (54-56). Moreu *et al.* observed that HBV tethering at CGIs was not affected in HBV-ΔX, suggesting that HBx does not drive the tethering of HBV cccDNA in the host chromosomes (56). However, Hensel *et al.* showed that HBx protein apparently co-localized with HBV cccDNA in the host chromosomes and that HBV-ΔX dropped the stability of HBV episomes (55). More recently, Tang *et al.* showed alterations in the episomal tethering sites of the host chromosomes caused by the HBV-ΔX mutation. Interestingly, HBV-ΔX mutant episomes, in which genes were transcriptionally

repressed, were preferentially linked to the five heterochromatic regions of chromosome 19. Differently from HBV-ΔX mutant episomes, the HBV wild-type episomes have reduced enrichment in those five heterochromatic regions of chromosome 19 and are preferentially associated with active chromatin regions known as compartment A (54). Thus, transcriptionally inactive cccDNA preferentially co-localizes in chromosome 19, whereas activated cccDNA associates localized to transcriptionally active regions.

Viral episomal tethering sites appear to be favorable for viral replication or transcription. Interestingly, the tethering sites of EBV episomes depend on viral latency type (44). EBV episomes in Burkitt lymphoma cell lines belonging to latency type I, indicating that most viral genes were repressed, tended to be associated with repressive chromatin regions. On the contrary, EBV episomes in LCLs belonging to latency type III, indicating that most viral genes were actively expressed, tended to be associated with active chromatin regions (44). Therefore, EBV viral episomes are positioned in the favorable region for their replication and transcription. Thus, the expression of EBV genes might be effectively controlled by host chromatin environments according to latency types. As another example, HBV cccDNA could have been affected by the cellular chromatin environment mediated to cellular protein CFP1. CFP1 binds to CGIs and recruits the methyltransferase SET1 responsible for H3K4me3 deposition. Interestingly, CFP1 also binds to HBV cccDNA and is required to enrich H3K4me3 in HBV cccDNA. The enrichment of H3K4me3 was significantly decreased in both cellular chromatin and HBV cccDNA by the depletion of CFP1. Therefore, the cellular active chromatin environment linked to HBV cccDNA can influence viral replication or transcription through the host factor, such as CFP1(56).

PERSPECTIVES OF 3D GENOMICS WITH VIRAL EPISOMES

NGS-based genomic technologies and microscopic analyses have revealed that genomes are not randomly distributed but organized in hierarchical order in the nucleus (61). The genome structures range from gene loops to topologically associating domains (TADs) and compartment A/B (61, 62). The tethering sites of viral episomes should be understood in the context of the three-dimensional organization of the host genome since the host genome is not a one-dimensional linear structure. A recent study has shown that transcriptionally inactive HBV cccDNA is associated with the inactive compartment B and transcriptionally active HBV cccDNA preferentially interacts with the active compartment A (54).

Our knowledge of the tethering sites of viral episomes is limited to the mean score for the cell population. In the case of EBV-positive Burkitt lymphoma, approximately 1,000 significant 4C peaks were identified despite the presence of 50–100 episomes in a single nucleus (44). Thus, the combination of tethering sites in a single nucleus is unknown. Single-cell Hi-C technology has been developed to determine the genomic contacts in an individual nucleus (63, 64). However, the limited number of associations between viral episomes and host chromosomes in a single nucleus might be impossible to capture through the single-cell Hi-C analysis. Instead, future 3D genomic methods combined with single-cell technology and 4C-seq or CHi-C may allow us to amplify the specific association between viral episomes and host chromosomes in an individual nucleus.

Cellular factors for the maintenance of viral episomes have not been well addressed. It has been reported that cellular genome organizer CTCF associates with viral episomes and regulates the viral latency through forming a 3D organization of the viral genome in EBV, KSHV, herpes simplex virus (HSV), cytomegalovirus (CMV), and HPV (2, 65-67). However, the role of CTCF in the tethering of viral episomes has not been reported, and the binding sites of CTCF do not appear to correlate with the tethering sites of EBV episomes (44). The cellular

protein complex SMC5/6 is known to be a host restriction factor for HBV infection by repressing the transcription of viral genes, and HBx also antagonizes the role of SMC5/6 by the SMC5/6 protein degradation (68). The expression of viral genes and thereby tethering sites of HBV episomes were regulated by cellular protein complex SMC5/6, implying that the chromatin environment of the viral episome, rather than the presence of HBx itself, is important for the position of viral episomes on host chromosomes (54). Therefore, the role of host factors, including genome organizers such as CTCF and SMC protein complexes, in the tethering of viral episomes could be an interesting research topic.

We summarized the tethering sites of viral episomes according to viruses, host cells, and viral latency types. In order to comprehensively understand the mechanism of episomal attachment on host chromosomes under these various conditions, collective information acquired from different host cells and viral latency types is necessary. In addition, most research has been conducted with virus-infected cell lines, not clinical samples. Therefore, for clinical application, it is necessary to further study the changes in epigenetic features and tethering sites of viral episomes with clinical samples.

ACKNOWLEDGEMENTS

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Ministry of Science and ICT, Republic of Korea (No. 2019R1F1A1061826, 2019R1A4A1024764, 2022R1A2C100442311 and 2021R1A2C1010313) and the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI) funded by the Ministry of Health & Welfare, Republic of Korea (No. HI22C1510).

322 **CONFLICTS OF INTEREST**

323 The authors have no conflicting interests.

324

325

FIGURES

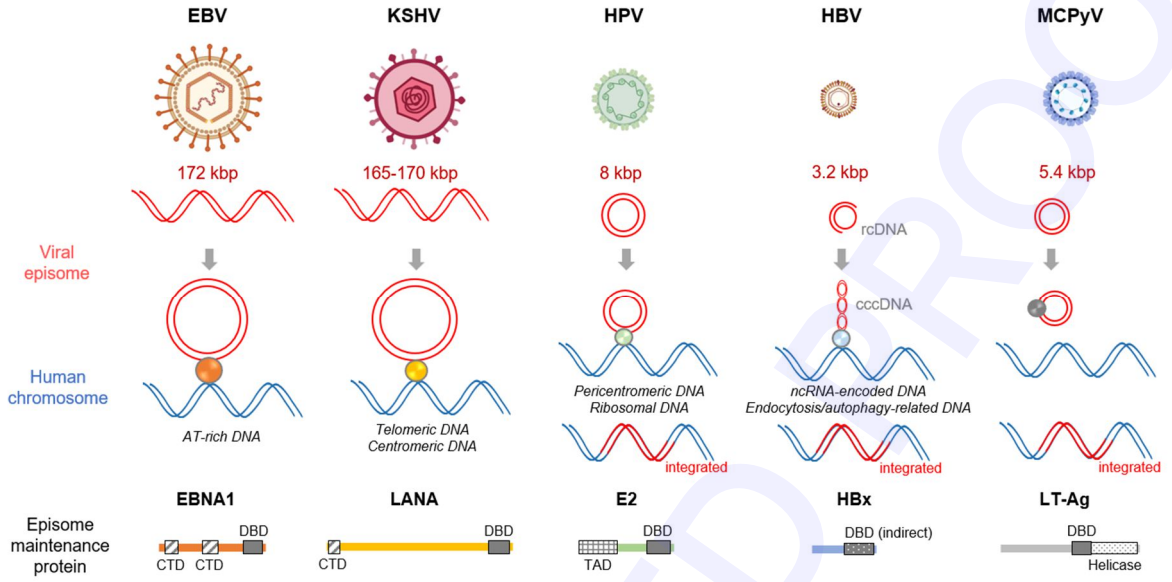

Figure 1. Episomes and episome maintenance proteins of DNA tumor viruses. The characteristic episomal form of each virus and the DNA binding domains of EMPs are depicted. CTD, chromosome-tethering domain; DBD, DNA-binding domain; TAD, Trans-activating domain. For more specific information on episome and host chromosome tethering sites, see Table 1.

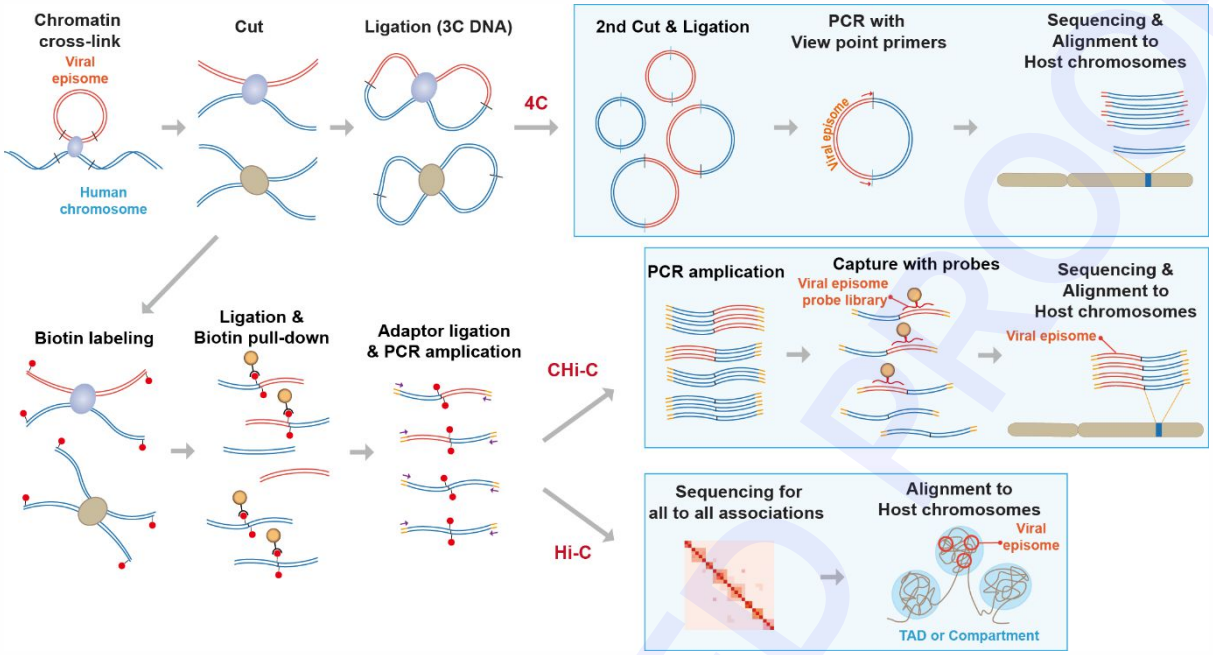

Figure 2. Schematic of 3C derived methods to detect the associations between viral episomes and host chromosomes. Blue-colored boxes indicates specific procedures for 4C, CHi-C, and Hi-C. Each sequencing results were aligned to the chromosomes and presented as a linear context for 4C and CHi-C and a three-dimensional context for Hi-C.

TABLE

Table 1. Tethering sites of viral episome analyzed by 3C derived methods

Virus	Host	Method (resolution)	Tethering sites on host chromosome	Co-localized factors	Ref.
EBV	Burkitt's lymphoma cell line (Daudi, KemIII, RaeI, Raji)	Hi-C (chromosome level)	gene-poor chromosomes (latent); gene-rich chromosomes (reactivation)	ND*	(52)
	Lymphoblastoid cell line (GM12878)	Hi-C (chromosome level)	gene-poor chromosomes	ND	(52)
	Lymphoblastoid cell line (GM12878)	Hi-C (10 kb), 4C for validation	typical or super enhancers and active markers	EBNA2/3 (EBV), IKZF1/ RUNX3, HDGF, NBS1/ NFIC	(53)
	Burkitt's lymphoma cell line (MutuI, Raji)	4C (10 kb)	heterochromatin, silent neuronal genes	EBNA1 (EBV), EBF1, RBP-jK, H3K9me3, AT-rich flanking sequence	(44)
	Lymphoblastoid cell line (Mutu-LCL, GM12878)	4C, Hi-C (10 kb)	active chromatin	EBNA2 (EBV), H3K27ac, H3K4me1/3	(44)
	Gastric cell lines (14 EBV associated Gastric cancer cell lines, 2 normal gastric epithelial cell lines)	Hi-C (25 kb), 4C for validation	heterochromatin to euchromatin transition	H3K9me3 to H3K4me1/H3K27ac	(6)
HBV	primary human hepatocyes (0, 7 days after infection)	Hi-C, capture Hi-C (400 kb)	active chromatin, CpG islands (higly expressed genes)	Cfp1	(56)
	HepaRG hepatocytes	4C (2/ 10/ 50/ 250 kb)	nuclear subdomain associated with open chromatin	HBx (HBV)	(55)
	HepG2-NTCP	3C-HTGTS	transcription start sites, enhancers, CpG islands	H3K4me2/3, H3K9ac, H3K27ac, H3K36me3	(51)
	HepG2-NTCP	4C, Hi-C	HBV-ΔX: Chr9 heterochromatin hub; HBV-wt: compartment A	HBV-ΔX: H3K9me3; HBV-wt: active chromatin; controlled by HBx and SMC5/6	(54)
KSHV	PEL cell line (BC-1)	Hi-C (chromosome level)	gene-poor chromosomes (latent)	ND	(52)
	PEL cell lines (BC-1, BC-3, BCBL-1)	capture Hi-C (10 kb)	near Centromere (1% of total)	LANA (KSHV), ADNP, CHD4	(33)

*ND: Not Determined

REFERENCES

1. Lieberman PM (2013) Keeping it quiet: chromatin control of gammaherpesvirus latency. Nat Rev Microbiol 11, 863-875
2. Lieberman PM (2016) Epigenetics and Genetics of Viral Latency. Cell Host Microbe 19, 619-628
3. Crosbie EJ, Einstein MH, Franceschi S and Kitchener HC (2013) Human papillomavirus and cervical cancer. Lancet 382, 889-899
4. Goncalves PH, Ziegelbauer J, Uldrick TS and Yarchoan R (2017) Kaposi sarcoma herpesvirus-associated cancers and related diseases. Curr Opin HIV AIDS 12, 47-56
5. Matsukura T, Koi S and Sugase M (1989) Both episomal and integrated forms of human papillomavirus type 16 are involved in invasive cervical cancers. Virology 172, 63-72
6. Okabe A, Huang KK, Matsusaka K et al (2020) Cross-species chromatin interactions drive transcriptional rewiring in Epstein-Barr virus-positive gastric adenocarcinoma. Nat Genet 52, 919-930
7. Chakravorty A, Sugden B and Johannsen EC (2019) An Epigenetic Journey: Epstein-Barr Virus Transcribes Chromatinized and Subsequently Unchromatinized Templates during Its Lytic Cycle. J Virol 93
8. Mrozek-Gorska P, Buschle A, Pich D et al (2019) Epstein-Barr virus reprograms human B lymphocytes immediately in the prelatent phase of infection. Proc Natl Acad Sci U S A 116, 16046-16055
9. Chang Y, Cesarman E, Pessin MS et al (1994) Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. Science 266, 1865-1869
10. Purushothaman P, Dabral P, Gupta N, Sarkar R and Verma SC (2016) KSHV Genome Replication and Maintenance. Front Microbiol 7, 54
11. Ueda K, Sakakibara S, Ohsaki E and Yada K (2006) Lack of a mechanism for faithful partition and maintenance of the KSHV genome. Virus Res 122, 85-94
12. Lu F, Day L, Gao SJ and Lieberman PM (2006) Acetylation of the latency-associated nuclear antigen regulates repression of Kaposi's sarcoma-associated herpesvirus lytic transcription. J Virol 80, 5273-5282
13. Zhu FX, Cusano T and Yuan Y (1999) Identification of the immediate-early transcripts of Kaposi's sarcoma-associated herpesvirus. J Virol 73, 5556-5567
14. Jenner RG, Alba MM, Boshoff C and Kellam P (2001) Kaposi's sarcoma-associated herpesvirus latent and lytic gene expression as revealed by DNA arrays. J Virol 75, 891-902
15. Toth Z, Brulois K, Lee HR et al (2013) Biphasic euchromatin-to-heterochromatin transition on the KSHV genome following de novo infection. PLoS Pathog 9, e1003813
16. Tu T, Budzinska MA, Vondran FWR, Shackel NA and Urban S (2018) Hepatitis B Virus DNA Integration Occurs Early in the Viral Life Cycle in an In Vitro Infection Model via Sodium Taurocholate Cotransporting Polypeptide-Dependent Uptake of Enveloped Virus Particles. J

Virol 92

17. Tuttleman JS, Pourcel C and Summers J (1986) Formation of the pool of covalently closed circular viral DNA in hepadnavirus-infected cells. Cell 47, 451-460
18. Allweiss L, Volz T, Giersch K et al (2018) Proliferation of primary human hepatocytes and prevention of hepatitis B virus reinfection efficiently deplete nuclear cccDNA in vivo. Gut 67, 542-552
19. Pyeon D, Pearce SM, Lank SM, Ahlquist P and Lambert PF (2009) Establishment of human papillomavirus infection requires cell cycle progression. PLoS Pathog 5, e1000318
20. Lieberman PM (2014) Virology. Epstein-Barr virus turns 50. Science 343, 1323-1325
21. Jones CH, Hayward SD and Rawlins DR (1989) Interaction of the lymphocyte-derived Epstein-Barr virus nuclear antigen EBNA-1 with its DNA-binding sites. J Virol 63, 101-110
22. Rawlins DR, Milman G, Hayward SD and Hayward GS (1985) Sequence-specific DNA binding of the Epstein-Barr virus nuclear antigen (EBNA-1) to clustered sites in the plasmid maintenance region. Cell 42, 859-868
23. Ambinder RF, Shah WA, Rawlins DR, Hayward GS and Hayward SD (1990) Definition of the sequence requirements for binding of the EBNA-1 protein to its palindromic target sites in Epstein-Barr virus DNA. J Virol 64, 2369-2379
24. Morgan SM, Tanizawa H, Caruso LB 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
25. Jourdan N, Jobart-Malfait A, Dos Reis G et al (2012) Live-cell imaging reveals multiple interactions between Epstein-Barr virus nuclear antigen 1 and cellular chromatin during interphase and mitosis. J Virol 86, 5314-5329
26. Deschamps T, Bazot Q, Leske DM et al (2017) Epstein-Barr virus nuclear antigen 1 interacts with regulator of chromosome condensation 1 dynamically throughout the cell cycle. J Gen Virol 98, 251-265
27. Lin A, Wang S, Nguyen T, Shire K and Frappier L (2008) The EBNA1 protein of Epstein-Barr virus functionally interacts with Brd4. J Virol 82, 12009-12019
28. Shire K, Ceccarelli DF, Avolio-Hunter TM and Frappier L (1999) EBP2, a human protein that interacts with sequences of the Epstein-Barr virus nuclear antigen 1 important for plasmid maintenance. J Virol 73, 2587-2595
29. Cotter MA, 2nd and Robertson ES (1999) The latency-associated nuclear antigen tethers the Kaposi's sarcoma-associated herpesvirus genome to host chromosomes in body cavity-based lymphoma cells. Virology 264, 254-264
30. Barbera AJ, Chodaparambil JV, Kelley-Clarke B et al (2006) The nucleosomal surface as a docking station for Kaposi's sarcoma herpesvirus LANA. Science 311, 856-861
31. Lim C, Choi C and Choe J (2004) Mitotic chromosome-binding activity of latency-associated nuclear antigen 1 is required for DNA replication from terminal repeat sequence of Kaposi's

sarcoma-associated herpesvirus. J Virol 78, 7248-7256

32. Barbera AJ, Ballestas ME and Kaye KM (2004) The Kaposi's sarcoma-associated herpesvirus latency-associated nuclear antigen 1 N terminus is essential for chromosome association, DNA replication, and episome persistence. J Virol 78, 294-301

33. Kumar A, Lyu Y, Yanagihashi Y et al (2022) KSHV episome tethering sites on host chromosomes and regulation of latency-lytic switch by CHD4. Cell Rep 39, 110788

34. Ohsaki E, Ueda K, Sakakibara S, Do E, Yada K and Yamanishi K (2004) Poly(ADP-ribose) polymerase 1 binds to Kaposi's sarcoma-associated herpesvirus (KSHV) terminal repeat sequence and modulates KSHV replication in latency. J Virol 78, 9936-9946

35. Si H, Verma SC, Lampson MA, Cai Q and Robertson ES (2008) Kaposi's sarcoma-associated herpesvirus-encoded LANA can interact with the nuclear mitotic apparatus protein to regulate genome maintenance and segregation. J Virol 82, 6734-6746

36. Krithivas A, Fujimuro M, Weidner M, Young DB and Hayward SD (2002) Protein interactions targeting the latency-associated nuclear antigen of Kaposi's sarcoma-associated herpesvirus to cell chromosomes. J Virol 76, 11596-11604

37. Xiao B, Verma SC, Cai Q et al (2010) Bub1 and CENP-F can contribute to Kaposi's sarcoma-associated herpesvirus genome persistence by targeting LANA to kinetochores. J Virol 84, 9718-9732

38. Belloni L, Pollicino T, De Nicola F et al (2009) Nuclear HBx binds the HBV minichromosome and modifies the epigenetic regulation of cccDNA function. Proc Natl Acad Sci U S A 106, 19975-19979

39. Piirsoo M, Ustav E, Mandel T, Stenlund A and Ustav M (1996) Cis and trans requirements for stable episomal maintenance of the BPV-1 replicator. EMBO J 15, 1-11

40. Bastien N and McBride AA (2000) Interaction of the papillomavirus E2 protein with mitotic chromosomes. Virology 270, 124-134

41. Ilves I, Kivi S and Ustav M (1999) Long-term episomal maintenance of bovine papillomavirus type 1 plasmids is determined by attachment to host chromosomes, which Is mediated by the viral E2 protein and its binding sites. J Virol 73, 4404-4412

42. Sullivan CS and Pipas JM (2002) T antigens of simian virus 40: molecular chaperones for viral replication and tumorigenesis. Microbiol Mol Biol Rev 66, 179-202

43. Calattini S, Sereti I, Scheinberg P, Kimura H, Childs RW and Cohen JI (2010) Detection of EBV genomes in plasmablasts/plasma cells and non-B cells in the blood of most patients with EBV lymphoproliferative disorders by using Immuno-FISH. Blood 116, 4546-4559

44. Kim KD, Tanizawa H, De Leo A et al (2020) Epigenetic specifications of host chromosome docking sites for latent Epstein-Barr virus. Nat Commun 11, 877

45. Kempfer R and Pombo A (2020) Methods for mapping 3D chromosome architecture. Nat Rev Genet 21, 207-226

46. Zhao Z, Tavoosidana G, Sjolinder M et al (2006) Circular chromosome conformation capture

(4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nat Genet 38, 1341-1347

47. Lieberman-Aiden E, van Berkum NL, Williams L et al (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326, 289-293

48. van de Werken HJ, Landan G, Holwerda SJ et al (2012) Robust 4C-seq data analysis to screen for regulatory DNA interactions. Nat Methods 9, 969-972

49. van de Werken HJ, de Vree PJ, Splinter E et al (2012) 4C technology: protocols and data analysis. Methods Enzymol 513, 89-112

50. Mifsud B, Tavares-Cadete F, Young AN et al (2015) Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. Nat Genet 47, 598-606

51. Yang B, Li B, Jia L et al (2020) 3D landscape of Hepatitis B virus interactions with human chromatins. Cell Discov 6, 95

52. Moquin SA, Thomas S, Whalen S et al (2018) The Epstein-Barr Virus Episome Maneuvers between Nuclear Chromatin Compartments during Reactivation. J Virol 92

53. Wang L, Laing J, Yan B et al (2020) Epstein-Barr Virus Episome Physically Interacts with Active Regions of the Host Genome in Lymphoblastoid Cells. J Virol 94

54. Tang D, Zhao H, Wu Y et al (2021) Transcriptionally inactive hepatitis B virus episome DNA preferentially resides in the vicinity of chromosome 19 in 3D host genome upon infection. Cell Rep 35, 109288

55. Hensel KO, Cantner F, Bangert F, Wirth S and Postberg J (2018) Episomal HBV persistence within transcribed host nuclear chromatin compartments involves HBx. Epigenetics Chromatin 11, 34

56. Moreau P, Cournac A, Palumbo GA et al (2018) Tridimensional infiltration of DNA viruses into the host genome shows preferential contact with active chromatin. Nat Commun 9, 4268

57. Heslop HE (2020) Sensitizing Burkitt lymphoma to EBV-CTLs. Blood 135, 1822-1823

58. Toth Z, Papp B, Brulois K, Choi YJ, Gao SJ and Jung JU (2016) LANA-Mediated Recruitment of Host Polycomb Repressive Complexes onto the KSHV Genome during De Novo Infection. PLoS Pathog 12, e1005878

59. Benhenda S, Cougot D, Buendia MA and Neuveut C (2009) Hepatitis B virus X protein molecular functions and its role in virus life cycle and pathogenesis. Adv Cancer Res 103, 75-109

60. Leupin O, Bontron S, Schaeffer C and Strubin M (2005) Hepatitis B virus X protein stimulates viral genome replication via a DDB1-dependent pathway distinct from that leading to cell death. J Virol 79, 4238-4245

61. Dowen JM, Fan ZP, Hnisz D et al (2014) Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. Cell 159, 374-387

62. Risca VI and Greenleaf WJ (2015) Unraveling the 3D genome: genomics tools for multiscale exploration. Trends Genet 31, 357-372
63. Nagano T, Lubling Y, Stevens TJ et al (2013) Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. Nature 502, 59-64
64. Nagano T, Lubling Y, Varnai C et al (2017) Cell-cycle dynamics of chromosomal organization at single-cell resolution. Nature 547, 61-67
65. Tempera I, Wiedmer A, Dheekollu J and Lieberman PM (2010) CTCF prevents the epigenetic drift of EBV latency promoter Qp. PLoS Pathog 6, e1001048
66. Pentland I and Parish JL (2015) Targeting CTCF to Control Virus Gene Expression: A Common Theme amongst Diverse DNA Viruses. Viruses 7, 3574-3585
67. Bloom DC, Giordani NV and Kwiatkowski DL (2010) Epigenetic regulation of latent HSV-1 gene expression. Biochim Biophys Acta 1799, 246-256
68. Decorsiere A, Mueller H, van Breugel PC et al (2016) Hepatitis B virus X protein identifies the Smc5/6 complex as a host restriction factor. Nature 531, 386-389