University of Vermont

UVM ScholarWorks

Dissertations and Theses

2021

Characterization Of Higher Order Chromatin Structures And Chromatin States In Cell Models Of Human Herpesvirus Infection

Michael Mariani
University of Vermont

CHARACTERIZATION OF HIGHER ORDER CHROMATIN
STRUCTURES AND CHROMATIN STATES IN CELL MODELS OF
HUMAN HERPESVIRUS INFECTION

A Dissertation Presented

by

Michael Mariani

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfilment of the Requirements
For the Degree of Doctor of Philosophy
Specializing in Cellular, Molecular, and Biomedical Sciences

May, 2021

Defense Date: March 24, 2021
Dissertation Examination Committee:

Seth Frietze, Ph.D., Advisor
Matthew Poynter, Ph.D., Chairperson
David Pederson Ph.D.
Stephen Keller Ph.D.
Cynthia J. Forehand, Ph.D., Dean of the Graduate College

COPYRIGHT

ABSTRACT

Human herpesviruses are ubiquitous pathogens worldwide with 90% of the global population infected with one or more Human herpesviruses (HHV's) by adulthood. All herpesviruses have three unique life cycle stages. Upon resolution of a primary acute stage infection, they can establish a latent stage infection within the host cell nucleus. This stage is characterized primarily by transcriptional quiescence of the viral genome. Specific physiological conditions (e.g., cell stress) can cause the latent virus to enter the reactivation stage, often many years after resolution of the acute infection, in which the virus becomes replicationally active again. HHV's are known to cause disease in humans in all three stages of their lifecycle and chronic infection is becoming increasingly associated with a wide range of human morbidities. The field has been met with challenges in establishing tractable cell models to study HHV infection *in vitro* and the exact mechanisms regulating the maintenance of, and transition between, the life cycle stages are largely undefined; however, chromatin structure and function are known to play a role. In this dissertation I employ both molecular and computational approaches to study the higher order chromatin structures and chromatin states in cell models of infection for two Human herpesviruses at specific stages in their life cycles.

In the first part of this dissertation, I successfully demonstrate the establishment of a latent Human betaherpesvirus 6A (HHV-6A) infection, which requires integration of its genome into the host cell genome, in an *in vitro*-derived HEK-293 cell model. I then employ unbiased epigenomic methods along with bioinformatics techniques to identify three-dimensional virus-host contact regions and characterize the chromatin states of the latent virus. I then demonstrate the above in both a patient-derived cell model and in infected primary cells. Finally, I developed a novel and innovative computational approach to identify the sites of HHV-6A integration in host cell chromosomes using next generation sequencing data: to my knowledge, a first in the field.

In the second part of this work, I employ a human fetal lung fibroblast (HFL) cell model of acute infection to study Varicella Zoster Virus (VZV) chromatin biology - specifically with regards to the transcription factor CTCF. CTCF is a key organizer of chromatin three-dimensional structure and plays a role in chromatin organization and transcriptional regulation in certain HHV's. Herein, I present evidence that CTCF is likely involved in virus-host structural interactions – an important finding that is countervailing to the current suppositions of the Human alphaherpesvirus field.

Overall, the work presented herein offers novel insight into the complex and dynamic relationship between HHV's and host cell chromatin. With a more comprehensive characterization of the higher order chromatin structures and chromatin states that define this relationship, we can better define the HHV life cycle. Such knowledge will allow for the development of improved treatments against these insidious infections.

CITATIONS

Material from this thesis (or dissertation) has been published in the following form:

Mariani, M., Zimmerman, C., Rodriguez P., Hasenohr, E., Aimola, G., Gerrard, D.L., Richman, A., Dest, A., Flamand, L., Kaufer, B., Frietze, S.. (2021). Higher-Order Chromatin Structures of Chromosomally Integrated HHV-6A Predict Integration Sites. Frontiers in Cellular and Infection Microbiology, 11, 67.

DEDICATION

I dedicate this dissertation to my family: Mom, Dad and Rob. If it is said that I have accomplished anything in this life, it is only by dint of your unwavering love and support.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

CHAPTER 1

Human herpesviruses

Human herpesviruses are ubiquitous pathogens with greater than 90% of the world population having one or more herpesviruses by adulthood (Lan and Luo, 2017). All Human herpesviruses (HHVs) were originally referred to as *herpesviruses hominis* (Lusso et al., 2007) (Bastian et al., 1972) and as simplex by 1973 (Baringer and Swoveland, 1973). At that time Varicella Zoster Virus, or VZV, was not known to be the in the same family as herpes simplex and was not known to be the same pathogen that causes shingles or herpes zoster (HZ): 'Although varicella and HZ were known to be related, it was not until 1965 that the British general practitioner, Dr Robert E. Hope-Simpson, suggested that "*herpes zoster is a spontaneous manifestation of varicella infection*".' (Hope-Simpson, 1964) (Depledge et al., 2018b). Thus, it was only in the 1960s that the capacity for herpesviruses to remain dormant in their hosts subsequent to initial infection, and become transcriptionally competent again at a future timepoint, was discovered.

These different "states" of HHV infection are now referred to as stages of the HHV life cycle which are characterized by periods of differential transcriptional and replicational activity. There are three stages to the life cycle: acute, latent, and reactivation, and this ability to transition to different life cycle stages is the most defining characteristic of all herpesviruses, even those not endemic to humans (Wang et al., 2020). The ability to establish a latent infection is related to the very etymology of the word "herpes", which is

1

derived from the classical Greek verb "*herpein*", meaning "to creep". The "creeping" nature of this virus is evidenced by its ability to evade the host cell immune response during the latent stage of its life cycle (De Bolle et al., 2005). To achieve this ability to hide during latency, HHVs have evolved to take up residence in the nucleus of a variety of human tissue types and have evolved impressive mechanisms to avoid the host immune system, leading to their pervasive presence across the globe.

Health effects acute and chronic

HHVs cause a wide range of health issues in humans including genital lesions, chicken pox, shingles, encephalitis, cerebrovascular disease and even cancer. Most of the readily visible manifestations of HHV infection are found during the acute and reactivation stages of the virus life cycle. The Human alphaherpesviruses HSV-1 and HSV-2 are the primary causes of cold sores and genital herpes, while VZV is the cause of chickenpox. Approximately 90% of all adults harbor the latent HHV-6 virus - primarily spread via saliva, and the initial infection usually occurs between 6 months and 2 years of age. The human betaherpesvirus HHV-6B is the primary cause of febrile seizures in otherwise healthy children and is the causative agent of roseola infantum (exanthema subitem or Sixth's disease). Both HHV-6A and HHV-6B primary infection can lead to viral encephalitis in both children and adults – an oftentimes serious medical condition (Mullins and Krishnamurthy, 2020) (Gewurz et al., 2008) as well as thyroiditis (Seyyedi et al., 2019). Additionally, HHV-6A and B are now known to greatly increase the likelihood and severity of developing acute graft-versus-host-disease (aGVHD) in transplant recipient

patients. aGVHD is a serious medical condition that can lead to multi-systemic organ failure (Gotoh et al., 2014; Agut et al., 2017). Furthermore, approximately 1% of the world's population harbors a germline integrated form of the virus - iciHHV-6, resulting in these individuals harboring a copy of the HHV-6 genome in every cell in their body (Collin and Flamand, 2017). Thus, individuals with this condition are at a much greater risk for developing morbidities such as post-transplant complications. HHV-7 infection generally presents with few if any symptoms but is also associated with roseola and encephalitis - particularly in immune compromised individuals (Wolz et al., 2012). The Human gammaherpesviruses Epstein-Barr Virus (EBV) is the causative agent of mononucleosis or "mono" (Dunmire et al., 2015), while human cytomegalovirus (HCMV) can lead to health complications in immunocompromised individuals as well (Britt, 2008). Kaposi Sarcoma-associated Herpesvirus (KSHV or HHV-8) generally does not cause disease in healthy (non-immunocompromised) individuals during the acute and replication stages of infection, but as its name implies, it does cause Kaposi sarcoma in some individuals, particularly in those that are immunocompromised (Dupin). Gammaherpesviruses such as EBV also been known to contribute to periodontal disease as well (Binshabaib et al., 2018).

While the health complications arising from the acute and replication stages of HHV infections are fairly well characterized at this point, the long term, chronic health effects of latent and reactivating infection are less well characterized overall. In some cases, such as infection with the human gammaherpesviruses EBV and KSHV, the low grade chronic expression of the latent stage viral transcripts are known now to be oncogenic in humans, particularly in immunocompromised individuals - the only two HHVs known to

be oncogenic (De Bolle et al., 2005; Luo and Ou, 2015; Goncalves et al., 2017; Song et al., 2019). VZV, the causative agent of chickenpox, or varicella, can reactivate years or decades later to cause shingles (Arvin, 1996; Sauerbrei). In fact, about one third of all patients with VZV will get shingles in their lifetime, sometimes with neurological sequelae (Depledge et al., 2018b). Additionally, the long terms effects of all the Human alphaherpesviruses: HSV-1, HSV-2, and VZV, the primary neurotrophic HHVs, are increasingly being associated with neurological complications such as Alzheimer's disease (Mancuso et al., 2019) and cerebrovascular disease (Nagel and Bubak, 2018). Overall, numerous studies in recent years have begun to associate chronic HHV infection with a wide range of human diseases ranging from cardiovascular disease, and stroke, to neuro-degenerative disorders and autoimmune disease.

HHV virology

As discussed above, HHVs are associated with disease at all stages of their life cycle and can lead to both acute and long-term health complications. While much of the biology surrounding the external causes of transition between life cycle stages, as well as the exact mechanisms governing these dynamic transitions are still under investigation, many aspects of HHV virology, fortunately, are well understood at this point. In fact, the viral transcriptional landscapes of HHVs are among the best characterized of the double-stranded DNA viruses (Alba et al., 2001). Many HHV genomes have been fully sequenced to date, and this genetic and molecular information has allowed taxonomists to accurately classify HHVs together in the family Herpesviridae and subsequently into three subfamilies

based on their gene content, cellular tropism and genome organization (Alba et al., 2001). In addition to their more common clinical names, which vary for historical reasons, the international committee on the taxonomy and nomenclature of viruses (ICTV) employs a genus-species-number naming system, e.g., Human herpesvirus 1-8 (HHV-1 – HHV-8) for purposes of the standardization of nomenclature. The exception in naming with regards to HHVs is HHV-6A/B, which, as of 2012, was officially reclassified from a single HHV-6 virus into two separate Human betaherpesviruses: HHV-6A and HHV-6B (Adams and Carstens, 2012). Furthermore, HHVs are placed into the Double-stranded DNA viruses of the Baltimore classification of viruses.

Based on the above information, Human herpesviruses are now known to represent a distinct group of the herpesvirus family, comprising nine species across three sub families (alpha, beta, and gamma). The three sub-families are thought to have diverged over 200 million years ago, before mammalian speciation (Aswad et al., 2020) and infected a wide variety of animal genera, but restricted groups of species within each taxon (Aswad et al., 2020) (McGeoch et al., 1995). Thus, these viruses have evolved along with pre-human primates, then humans, for millions of years.

There are important similarities and differences in terms of structure and function between individual HHV species as well. Structurally, all HHVs share a great deal of similarity. They are all large, double-stranded DNA viruses. The average length of a herpesvirus genome is approximately 130kbp and the average diameter of a herpesvirus virion is around 200nm. The double-stranded genetic material is contained within a viral capsid displaying pleiomorphic icosahedral (T-20) symmetry. All herpesviruses also contain a structural tegument layer of viral-encoded proteins and RNAs that sit between

the nucleo-capsid and the outer viral membrane. This tegument has shown to be ordered and important to herpesvirus virion assembly and host-cell egress. The outer viral envelope contains glycoproteins that maculate the surface of the virion and are responsible for host cell identification and binding (Guo et al., 2010).

All HHVs typically have between 70 and 120 open reading frames (ORFs), which have become synonymous with viral genes, that encode hundreds of potential gene products (Alba et al., 2001). All HHVs contain a set of core genes, a subset of sub-family specific genes, and a variable number of ORFs unique to the individual virus (Alba et al., 2001). While new virus genes are still being discovered, many of the core proteins regulating the initial stages of acute infection as well as many of the genes that are expressed solely during latency have been identified, and homology does exist across these identified genes (Arvey et al., 2012). In terms of key differences between the HHV species, it is the unique ORF's mentioned above that allow the individual HHV species to operate within their own specific ecological niches.

As discussed above, both the many commonalities and distinctly unique elements of the individual HHVs have enabled the successful taxonomic classification of these viruses, and much of the epidemiology and pathophysiology of HHV infection has been uncovered. Above all, the presence of the distinct life cycle stages is perhaps the single most defining and uniting characteristic of all HHVs. However, the specific mechanisms that regulate the maintenance of, and transition between, the different life cycle stages, both within and across individual HHVs, still remains largely a mystery. For example, while it is currently understood that HHVs are known to reactivate from their latent life cycle stages (regaining transcriptional and replicational competency) under certain physiological

stressors such as hormonal imbalances (De Bolle et al., 2005), the precise cellular mechanisms governing the reactivation of latent HHVs, as well as those governing the acute to latent transition, have yet to be uncovered. However, because herpesviruses are nuclear-only replicating DNA viruses, host cell chromatin biology was initially suspected, and has since been confirmed, to play a role in regulating the HHV life cycle.

Early studies with the Human gammaherpesviruses EBV as well as with the human alphaherpesvirus HSV-1 were among the first to demonstrate the role that host cell chromatin plays in the HHV life cycle. Human gammaherpesviruses were chosen because their preferred cell tropism for replication are actively dividing CD8+ and CD4+ T cells (Long et al., 2019). Furthermore, EBV was well known to the public as the causative agent of mononucleosis and has been found to be oncogenic (Yin et al., 2019). Thus, the facts that they infect regularly dividing cells, along with their public presence and carcinogenic potential, put them at the forefront of research. The alphaherpesviruses HSV-1 (and HSV-2 and VZV) were also early contenders for the development of models to study the HHV life cycle, as cold sores, genital herpes, chickenpox, and shingles are well known to the public. In addition, there were early successes in the establishment of animal models of infection with HSV-1 and successful reproduction of components of human disease in these models (Webre et al., 2012). In contrast, the study of betaherpesviruses lagged that of the other HHVs, largely due to their more recent discovery and formerly unknown role in disease etiology. Furthermore, it is puzzling on the surface why the cellular study of VZV infection, the HHV in which the presence of life cycle stages was first confirmed, has not been at the forefront of HHV chromatin biology research. This is largely due to two factors. Firstly, researchers were not able to establish and reactivate infection in small

animal models like they were with HSV-1, and secondly, the biology of VZV has lent itself to the successful creation of vaccines (Eshleman et al., 2011). For a discussion of why the development of a VZV vaccine has been successful compared to other herpesviruses (particularly vs. HSV-1) see Rouse et al. (Rouse and Kaistha, 2006).

To summarize, EBV is arguably the HHV about which the most chromatin biology is currently understood, followed in second place by HSV-1. In fact, these two viruses have been so well established relative to other HHVs, it is often within the contextual framework of their respective chromatin biology that hypotheses for other HHVs, such as HHV-6A and VZV – which are focus of this dissertation, are developed, tested, and compared against. Furthermore, the betaherpesviruses are somewhat unique in terms of their chromatin biology in that HHV-6A/B and possibly HHV-7 (Prusty et al.) require integration of their genome into the host cell genome to establish latency. The site of integration is generally the subtelomeric regions near the ends of host chromosomes, regions that have been difficult to assay due to their repetitive nature (Aimola et al., 2020).

Regardless of the popularity of specific HHVs within the chromatin research world, all herpesviruses are nucleus-only replicating DNA viruses, and their biology is closely intertwined with that of the host cell nucleus. As host chromatin structure and function, as well as cell and viral factors, have been demonstrated to work in concert to control transcription and replication in HHVs (such as EBV and HSV-1), it therefore follows that a basic understanding of the chromatin biology of both the host and virus, as well as of the cell models and experimental methodology employed to uncover said biology, is essential to gaining clear understanding of the complex dynamics of the HHV life cycle.

Chromatin biology of the host

As mentioned above, HHVs are nuclear-only replicating viruses, with transcription replication, and encapsidation, all occurring predominantly within the nucleus of the infected host cell (Heming et al., 2017). Organization of host DNA within the human cell nucleus is organized into chromatin. Chromatin is comprised of concatenations of nucleosomes, each nucleosome itself being comprised of DNA wrapped 1.46 times around an octamer of core histone proteins. This core octamer is in turn composed of two copies each of histone proteins H2A, H2B, H3, and H4. The core nucleosome is also bound by the H1 histone protein via a 46 bp length of DNA. In total, the core-plus-linker histone protein complex is wrapped approximately two times by a total of 200bp of DNA. Each nucleosome is in turn connected via the DNA strand to other nucleosomes in a chain-like fashion. This linear chain of nucleosomes is then wrapped and folded into a higher order structure called the 30nm fiber (Hubner et al., 2013). These fibers and the three-dimensional conformations that they can assume within the nucleus are referred to in this work as higher order chromatin structures.

This higher order structuring allows the human genome, which is approximately one meter in length, to fit within a human cell nucleus (approximately seven µm in diameter). In addition, alteration of post-translational modification of the tails of histone proteins allows for structural loosening and tightening of the chromatin into open/loose or closed/compact regions. This regulation of compaction in turn allows for the regulation of access of the DNA to transcriptional machinery (Conn and Schang, 2013). These post-translational modifications of the amino-terminal tail of the histone proteins also provide

docking sites for proteins and contribute to the structural stability of the three-dimensional chromatin conformation (Conn et al., 2013).

In addition to the compactness of the host chromatin the three-dimensional structure of chromatin is highly organized within the larger nuclear architectural framework. The three-dimensional structure of the nucleus is partitioned into topologically associating domains or TADs. Genomic regions within a TADs are more likely to interact with each other than with regions in other TADs. Thus, this partitioning leads to three-dimensional conformational organization within the host cell nucleus that comprises the higher order chromatin structures of the genome. The adoption of such conformational states adds another level at which the cell can regulate transcription of its genome (Beagan and Phillips-Cremins, 2020).

While chromatin structure itself an important aspect of genomic regulation. A variety of chromatin accessory factors such as histone modifying enzymes, transcription factors, and members of transcriptional regulating protein complexes play a role in regulating the level of compactness and the three-dimensional orientation of the host cell chromatin. The regions of the genome that encode for or regulate the above elements (promoters, enhancers, repetitive regions etc.) are referred to, within a spatial context, as chromatin states (Ernst and Kellis, 2015). Taken together, the higher order chromatin structures, chromatin states, and chromatin accessory factors all work together to regulate the host cellular genome.

The transcription factor CTCF

With a brief review of host (human) chromatin now complete, special attention will be paid to the transcription factor CTCF. CTCF is a ubiquitous transcription factor encoded by the *CTCF* gene, also referred to as CCCTC-binding factor. It contains 11-zinc finger motifs for binding to DNA and performs many roles in the cell. While transcriptionally, it generally acts as a repressor (Filippova et al., 1996), it also plays a key role as an organizer of chromatin architecture/conformation (Phillips and Corces, 2009). It achieves this ability via serving as a structural linking, or anchoring, protein between the cohesin protein complex and chromatin (Rubio et al., 2008). To exert its transcriptional repressive effects, CTCF acts as a transcriptional insulator. It sterically prevents specific regions of DNA from spatially interacting with other each other. For example, CTCF can sterically prevent an enhancer region of the genome from reaching its regulatory target. Besides its chromatin organizing and transcriptional regulatory roles, CTCF is even involved in VDJ recombination in the immune system, where it organizes the formation of chromatin loops to effect the chromatin conformational changes necessary for VDJ recombination in B and T-cells (Chaumeil and Skok, 2012).

CTCF clearly plays many important roles in organizing and regulating the host cell genome, but it has also been found to be heavily involved in the chromatin biology of HHVs as well. As will be discussed in more detail further in this chapter, Human gammaherpesviruses, such as EBV, are known to utilize CTCF to tether their own genomes to the host cell genome – an essential feature for the establishment and maintenance of the latent life cycle stage in these viruses, as well as for the coordination of viral replication

during the acute and reactivation stages of infection. In addition, both EBV and HSV-1 have been found to harbor CTCF binding regions across their respective genomes, and utilize CTCF to regulate genomic transcription, much in the same way that the uninfected host cell does. Thus, CTCF is an important factor in both host and virus chromatin biology.

Genomics methodology

As chromatin has been demonstrated to play a role in the dynamic life cycle of HHV, methods used to assay host cell chromatin biology can be employed within the context of HHV infection. Because the host nucleus is a dynamic and complexly structured organelle, a lot of information must be simultaneously evaluated when seeking to comprehensively interrogate nuclear structure and function. The development of genomic methodologies in recent years have become powerful tools in addressing this challenge. These methods can capture a great deal of structural and functional information regarding the biological system under investigation. Some methods lend themselves to elucidating genome or transcriptome sequence information without consideration of chromatin structures or accessory factors with which the genome or transcriptome may be interacting at the time of experiment. Other methods, conversely, are less concerned with sequence information and seek to capture information concerning chromatin structure, including, histone modifications and three-dimensional structure, or seek to capture the effects of particular chromatin remodeling enzymes or identify the presence and/or distribution of particular transcription factors as they interact with the host cell genome. Currently, it is

becoming more common to employ multiple techniques simultaneously to capture information about various aspects of the host cell chromatin biology across different treatment conditions and timepoints. Such combinational approaches are allowing researchers to investigate how chromatin structural organization and interactions with cellular factors work together to dynamically regulate the genome. Each method possesses its own strengths and weaknesses and must be considered for dispatch within the context of the goals of the experiment and the hypotheses under consideration (Kempfer and Pombo, 2020).

To capture transcriptomic information, RNA-sequencing allow the researcher to access far greater quantities of genetic and transcriptomic information more efficiently and more comprehensively than the earlier molecular biology techniques that preceded them. The entire mRNA of a sample of cells can be extracted, converted, to cDNA libraries and undergo the same sequencing processes used for DNA sequencing. This provides an unbiased method for quantification of transcripts from which sequence alteration information and differential splicing abundance can be identified and quantified. More recently, single-cell RNA sequencing allows researchers to assay the transcriptomes of heterogenous cell populations to look at transcriptional differences arising between individual cells (Goodwin et al., 2016).

While sequencing data lends itself to addressing a variety of hypotheses, regarding questions pertaining to chromatin biology specifically, other powerful genomics methods have been developed to interrogate the state of chromatin conformation and higher order chromatin structure as well as the association of transcription factors and chromatin remodeling enzymes with the host cell chromatin.

Chromatin immuno-precipitation sequencing, or ChIP-Seq, allows researchers to pull down chromatin associating factors using antibodies specific for them. Antibodies are created against the DNA-binding factors and unbound regions of the DNA are not extracted. These antibodies are then isolated from the digested cellular material and unlinked form the DNA-to which they are bound. This DNA can then undergo high-throughput sequencing, as mentioned earlier, to gain insight as to the quantity and distribution of specific DNA and chromatin associated factors, such as transcription factors, under specific experimental conditions (Beagan et al., 2020). In fact, ChIP-seq can be performed sequentially with antibodies to proteins that are hypothesized to be complexed together. This so-called "ChIP-ReChIP" provides the researcher insight into host factors that may be associating together, such as two members of a protein complex that may be working in concert.

With ChIP-Seq, the DNA region under investigation must be contained within chromatin that is sterically open enough (in a euchromatic state) to allow the DNA-binding factor of interest to "land" on the DNA sequence to which it normally binds. However, depending on the hypothesis, it can also be informative to identify regions of the genome that are contained within compact heterochromatin where such transcription factors would have difficulty binding due to steric restrictions. These regions are said to be in a heterochromatic state or "chromatinized". In such a case, we can turn to digestion methods such as MNase-Seq or ATAC-Seq to assess which regions of the genome under investigation are contained within open or closed chromatin (Beagan et al., 2020).

To identify closed or heterochromatic regions, MNase sequencing uses micrococcal nuclease to digest all DNA within the host genome that is not bound to histone proteins.

After digestion, the chromatin histone proteins are unlinked from the DNA to which they are attached, and the DNA can be isolated and undergo sequencing. Conversely, ATAC sequencing can be applied to uncover regions that are not bound to histone proteins and can give insight into which regions of the genome under investigation are contained within euchromatin and are thus more likely to be transcriptionally active (Beagan et al., 2020).

While MNase-Seq and ATAC-Seq are useful for gaining insight into regions of the genome that are contained in open or closed chromatin, they do not provide spatial or three-dimensional conformational information about the chromatin. To assay the three-dimensional chromatin conformation or higher order chromatin structures within a cell nucleus, we turn to the "C" family of genomics methodologies. The "C's" usually refer to either chromatin or conformation or both. These techniques aim to capture information about the three-dimensional chromatin structure of the nucleus under investigation. Higher order chromatin structural information is important because phenomena like DNA-looping are known to be involved with transcriptional regulation and the organization of the genome into the three-dimensional segmented compartments that define TADs, which also contribute to the overall regulation of the cellular genome (Beagan et al., 2020).

The first "C" methodology to be developed was "3C-Seq", which stands for chromosome conformation capture sequencing. In 3C-Seq, a primer is designed for a specific region of interest in the host cell genome. This region is usually referred to as a bait. A primer is defined for two regions of interest that are suspected of interacting, or contacting, one another, spatially. The primers anneal to the region(s) of interest and are crosslinked with a protein that protects this region of interest from enzymatic digestion. The protected region can then be de-protected (generally a de-crosslinking procedure) and

the researcher is left with a region of the genome with which this region was interacting (Beagan et al., 2020).

The second "C' method in the lineup is chromosome conformation capture circular sequencing or 4C-Seq. In this case, the cells are fixed (generally with paraformaldehyde) and then subject to multiple rounds of nuclease digestion, enzymatic ligation and decrosslinking, that ultimately result in circular DNA amplicons of genetic regions that were in spatial proximity to each other at the time of fixation. Primers are made for a specific region of interest and will anneal to the specific amplicons that contain their hybrid sequence. The hybrid amplicons then undergo PCR elongation and amplification before being sent for sequencing (Krijger et al., 2020). Ultimately, the researcher obtains DNA sequences that can undergo computational alignment to a reference genome in order to identify the regions of the host genome under investigation to which the bait, or viewpoint, primer region was proximal in three-dimensional space. Thus, 4CSeq provides a snapshot of the higher order chromatin conformation of the sample genome in the vicinity of the viewpoint region at the time of fixation. As a final note, 4C-Seq lends itself particularly well to investigating the higher order structural interactions between small exogenous chromosomes, such as an infecting HHV genome, with the host cell genome. Because the HHV genome is so small relative to the human genome, 4C-Seq results for a small number of viral viewpoint regions can identify virus-host interactions for much a viral genome.

Building on the 4C-Seq framework, 5C can use an array of bait or viewpoint primers to get spatial information for multiple regions simultaneously and HiC will digest the entire genome into fragments to find interacting regions across an entire genome. In the case of HiC over 5C, one will gain additional breadth of genomic information (i.e., the

entire genome) at the cost of some specificity of the identified interactions (Beagan et al., 2020). Taken together, the "C" family of technologies are powerful tools for identifying higher-order spatial genomic interactions and providing information about the higher order chromatin conformation of the nuclei being investigated.

There are many variations of the above genomics methodologies and new permutations and iterations seem to appear almost daily at times. The various alterations to the main methods described above are generally implemented in order to evaluate a biological hypothesis under very specific experimental conditions. Furthermore, these techniques can be combined with any number of other molecular biology techniques - providing a level of experimental customization tailored to the needs of the researchers' goals. But at their core, these methods all aim to allow researchers to answer specific questions about the chromatin states, higher-order interactions, and association between chromatin and DNA-binding factors and chromatin remodeling enzymes, that dynamically work together to regulate chromatin organization and the transcription of the cell genome. With a brief overview of both the host cell chromatin biology and current state of genomics methodologies complete, I will now review the literature surrounding the HHV chromatin biology that is germane to the main body of work presented in chapters 2 and 3.

Human gammaherpesviruses and host chromatin (EBV)

Most of what initially became known about the role of chromatin in Human herpesviruses began with studies of Human gammaherpesviruses, particularly the well-characterized oncogenic virus EBV using LCLs, or human lymphoblastoid cell lines

(Arvey et al., 2012). These cell lines are commonly employed as cell models in gammaherpesvirus studies, and they are created via infection of primary human B-lymphocytes with EBV (Neitzel, 1986; Hussain and Mulherkar, 2012). The genetically stable, proliferating cell lines contain multicopy episomal EBV genomes that then persist across cell divisions. Specific factors, that will be discussed more in the next section, can be added to reactivate the latent virus after it has transitioned from an acute to a latent infection (Chau et al., 2006; Wang et al., 2006).

With the establishment of these competent cell models, the early chromatin studies of EBV focused on assaying the virus and host genomes for specific transcriptionally activating (H3K27ac, H3K4me3) or repressive (H3K9me3, H3K27me3) histone marks or for chromatin factors such as CTCF that were known to play key roles on chromatin organization in healthy cells. Active EBV genes during acute infection were found to contain nucleosome-free promoter regions, demonstrating not only that host chromatin plays a role in regulating the HHV life cycle but that acute infection appeared to be associated with open, euchromatin regions of the genome (Zhou et al., 2005). It was not long before it was discovered that specific histone modifications are associated with different cell cycle stages and that chromatin boundary factors such as CTCF were identified as key contributors to the latent stage of the EBV life cycle (Arvey et al., 2012).

In addition, besides its chromatin organizing functionality, CTCF was found to act as a chromatin insulator in EBV - where it serves as a barrier to the transcription of particular EBV genes (Amelio et al., 2006; Arvey et al., 2012; Pentland and Parish, 2015). It has since been identified as the key insulator in Human gammaherpesviruses that

separates or "insulates" distinct temporal transcription clusters from one another (Tempera et al., 2010; Arvey et al., 2012).

These spatially separated viral gene clusters that are expressed at temporally distinct periods are referred to as comprising a particular "kinetic class" and are found in all the HHVs. Although, they typically are defined solely for the acute stage infection. But with EBV (and potentially HSV-1), disparate kinetic classes for the latent life cycle stage have also been identified and have been deemed "latency programs" (Arvey et al., 2012). These different latency programs have been identified within different cell types, giving rise to the delineation of "sub-stages" within the latent stage of EBV. To add to the complexity, in addition to viral proteins, miRNAs, ncRNAs and snoRNAs have been found to be expressed as part of these acute and latent expression profiles (Hutzinger et al., 2009; Arvey et al., 2012).

Thus, kinetic clusters comprise a variety of gene products, and it is not just CTCF insulator activity that is associated with their transcriptional regulation. Histone modifications have been found to be enriched at clusters of microRNAs (miRNAs) and non-coding RNAs in EBV as well, indicating that other chromatin factors besides CTCF likely play a role in regulating these RNAs. Potentially, these non-coding RNAs could themselves play a role in regulating the HHV life cycle (Arvey et al., 2012).

While CTCF has been shown to be a chromatin organizer and transcriptional insulator, perhaps its most important role in EBV infection was the discovery that CTCF binds with cohesin protein to form a CTCF-cohesin protein complex that allows the EBV genome to tether itself to the host genome during both an active acute (lytic) infection, and also during latency (Arvey et al., 2012). This tethering is a defining characteristic of latency

in gammaherpesviruses and underlines the critical importance of CTCF to the regulation of the gammaherpesvirus life cycle (it was also identified in KSHV). To date, Gammaherpesvirus researchers have been able to fully characterize how the cohesin complex, containing CTCF, regulates the three-dimensional conformation of the EBV virus episome (Arvey et al., 2012).

In addition to the CTCF-mediated higher order chromatin structure, the histone modifications and transcription factor binding sites that comprise the chromatin states of the EBV infection have been well-characterized as well. One study found that 26 of 60 total surveyed transcription factors had 109 bindings sites across the entire EBV genome. These binding sites were found to be arranged in clusters along the virus genome. Such clusters are likely regulatory hotspots and are found in other model organisms as well (Gerstein et al., 2010), and have been identified in HSV-1 (Amelio et al., 2006). Thus, EBV likely utilizes a variety of host cell gene products to coordinate its own life cycle.

Overall, HHV genomes are dense and contain numerous overlapping ORFs. Additionally, the identification of alternative donor and acceptor splice sites led to the discover of numerous alternatively spliced viral genes in EBV. Alternative spliceoforms could potentially contribute to modulation of the virus life cycle via chromatin involvement as well.

Overall, our understanding of the expression profiles that largely define the HHV life cycle have only become more complex in recent years, as have the number of ways that chromatin could be interacting with the viral genome to regulate its life cycle. The first full systems biology approach to the study of HHV chromatin and life cycle was by Arvey et al., in 2012, which involved multiple simultaneous genomic experiments, along with the

incorporation of publicly available DNA, and this study paved the way for the incorporation of systems genomics methods in the herpes virology field (Arvey et al., 2012). Such studies are rapidly improving the resolution at which the boundaries between (and within) the EBV life cycle are being defined. EBV studies have led the way in the uncovering the role of chromatin biology in the HHV life cycle and it remains the HHV in which we have the clearest understanding of the mechanisms governing the dynamic HHV life cycle. But another HHV has also provided researchers with an enormous wealth of information pertaining to HHV chromatin biology.

Human alphaherpesviruses and host chromatin (HSV-1)

The human gammaherpesviruses, particularly EVB, remains the HHV about which we know the most chromatin biology, but a great deal is known about the role of chromatin during HSV-1 (herpes simplex 1 or human alphaherpesvirus 1) infection as well. As was seen above with EBV, the HSV-1 life cycle appears to be regulated by a complex interplay of chromatin structures, chromatin states, and other host and virus factors.

While the chromatin biology is not as well understood as that of EBV, quite a bit is known relative to other HHVs. HSV-1 researchers have benefitted from the fact that many *in vivo* cell types and animal models have been successfully employed in the study of this virus, whereas the same cannot be said of another alphaherpesvirus, VZV, about which my work is covered in chapter 3 of this dissertation.

MNase studies have found that host chromatin associates with viral HSV-1 genes at all stages of infection and this finding was found to be independent of DNA replication status (Cliffe and Knipe, 2008; Conn et al., 2008; 2011; 2013; Cabral et al., 2018). It has been shown that chromatin regulates nuclear DNA accessibility through mechanisms such as remodeling enzymes that can control heterochromatin state, nucleosome stability, and/or recruitment of transcription factors (Hu et al., 2019). Such chromatin regulation in fact plays a major role in the normal host DNA repair, DNA replication, and DNA transcription processes (Turner, 2000; Berger, 2002; Cosgrove et al., 2004; Henikoff and Grosveld, 2008; Hu et al., 2019). Further evidence in recent years has increasingly emphasized the dynamic role that chromatin biology plays in the acute and latent (as well as reactivated) stages of the HSV-1 life cycle.

As with EBV acute, infection, viral gene expression proceeds in a temporal fashion defined by the sequential expression of different kinetic classes of genes (as all HHVs are believed to follow). For DNA replication to occur, "IE" or "immediate early" proteins must initiate expression of "E" or "early" proteins that must be expressed. After competent replication has begun, L or "late" genes are only then transcribed (Tombacz et al., 2017a).

The initial virion protein, which is found in the tegument, called virion protein 16 or VP16, is the first and foremost activating viral protein involved in transcription of the immediate-early (hereafter IE) HSV-1 genes. The IE proteins such as ICP0 and ICP4 then go on to recruit chromatin remodelers, along with transcription activators, to their respective promoter regions. The subsequent recruitment of said remodelers by ICP0 and ICP4 then leads to transcription of the early and late viral genes. Intracellular viral DNA, isolated with EdU-labeling and click chemistry, has been shown to associate with VP16,

ICP4, and ICP22 (Dembowski and DeLuca, 2015). Emphasizing the role of potential feedback mechanisms in the early stages of acute HSV-1 infections. It is known that VP16 can initiate HSV-1 viral transcription by recruiting both methyltransferases and demethylases, as well as histone acetylases (Herrera and Triezenberg, 2004; Kristie et al., 2010; Vogel and Kristie, 2013; Lang et al., 2017). Furthermore, VP16 recruits histone acetyl transferases (HATs) to HSV-1 associated chromatin. These HATs include CBP/p300, GCN5, and PCF (Memedula and Belmont, 2003; Wysocka and Herr, 2003; Herrera and Triezenberg, 2004), but interestingly, they have been shown via knockdown to not have a strong overall effect on transcription (Kutluay et al., 2009; Kutluay and Triezenberg, 2009).

Following activation by VP16, ICP4 will then recruit chromatin remodelers to the HSV-1 genome (Kalamvoki and Roizman, 2011; Wagner and DeLuca, 2013) which then induces histone dynamics (Gibeault et al., 2016) to activate the early (E) and late (L) genes (Gu and DeLuca, 1994). In turn, ICPO, enhances immediate early (IE), early (E) and late (L) gene expression. ICP0 also induces degradation of specific cellular factors which then leads to the silencing of histone variants. This silencing in turn alters histone occupancy (Lomonte et al., 2001; Lomonte and Morency, 2007; Cliffe and Knipe, 2008; Ferenczy and DeLuca, 2009; 2011). In terms of chromatin and transcription activators, ICP0 blocks histone deacetylation. It performs this feat through interactions with the REST/Co-REST complexes (Gu et al., 2005; Poon et al., 2006; Cliffe and Knipe, 2008; Coleman et al., 2008; Ferenczy and DeLuca, 2009) as well as interacting with chromatin remodelers. Furthermore, ICP0 (an E3-ubiquitin ligase) can cause histone silencing proteins to be degraded. Examples of such variants include Cenp-A and Cenp-B (Lomonte et al., 2001;

23

Lomonte and Morency, 2007). Additionally, previous studies have shown that ICP4 will increase chromatin dynamics of all the core histones as well as interact with chromatin remodelers (such as the CLOCK complex) (Conn et al., 2008; 2011; Kalamvoki and Roizman, 2011; Conn et al., 2013; Gibeault et al., 2016). Overall, the role of chromatin in the initial stages of HSV-1 acute infection is a complex and dynamic process involving an interplay of host chromatin structures, and chromatin states.

As mentioned earlier, HSV-1 is known to associate with host chromatin at all stages of its life cycle and it has been found to be associated with silenced chromatin during latency (Deshmane and Fraser, 1989; Wang et al., 2005; Cliffe et al., 2009; Bloom et al., 2010). Mutations in ICP4, ICP0 and VP16 have led to transcriptional silencing in fibroblasts (the hallmark of latency) (Conn and Schang, 2013). Both the acetylation, as well as methylation, of HSV-1-bound chromatin have been found to occur in distinct regions of the host cell nucleus to contribute to latency (Kennedy et al., 2015). In the part of the nucleus where the virus is associated with permissive H3K9ac, H3K14ac, and H3K4me2 actively transcribed HSV-1 DNA is seen (Kubat et al., 2004a; Kubat et al., 2004b; Neumann et al., 2007; Knipe and Cliffe, 2008). H3K9me2, H3K9me3, and H3K27me3 are the modifications found on nucleosomes of the promoters of virus genes that are transcriptionally silent (Cliffe and Knipe, 2008; Knipe and Cliffe, 2008; Cliffe et al., 2009). Such activating and permissive histone marks cause changes in the conformational state of the host chromatin, further emphasizing the importance of chromatin spatial conformation during HSV-1 infection.

As with EBV, various factors leading to cell stress can induce reactivation of latent HSV-1. Hypoxia (Bloom et al., 2010), inhibition or alteration of the mTOR pathway

(Kobayashi et al., 2012; Wilson and Mohr, 2012), odium butyrate (NaB), infected tissue explanation (Ertel et al., 2012), and dexamethasone treatment (Du et al., 2012), can all cause reactivation of cells infected with HSV-1. Thus, cellular pathways, epigenetic modifications, chemical exposure, and changes in physiological environment have all been found to contribute to the reactivation of latent HSV-1 virus (Kennedy et al., 2015).

As with EBV, CTCF is an important transcription factor that been found to play a key role in regulating the HSV-1 life cycle. CTCF will bind many regions of the HSV-1 genome during the acute, latent, and reactivation stages (Lang et al., 2017). Concerning the acute stage (lytic infection), previous ChIP-Seq experiments have shown that the CTCF interacts extensively with the virus (Lang et al., 2017). When CTCF was knocked down using CTCF siRNA in HeLa cells, a reduction in overall virus yield, virus genome copy number, and a reduction of viral transcription was observed (Lang et al., 2017). In addition, H3K9me3 and H3K27me3 deposition was increased on viral genes, and RNA Pol-II deposition, was decreased following knockdown. RNA POL-II elongation was also found to be facilitated by CTCF. This fact, along with the reduction in repressive chromatin deposition associated with CTCF, demonstrates its role in promoting the transcription of acute stage lytic viral genes. Furthermore, higher levels of RNA POL-II modified with CTDSer2P was found relative to levels of the CTDSer5P modification (Lang et al., 2017), indicating that differential POL-II phosphorylation, possibly regulated by CTCF, could also impact the virus life cycle. Again, CTCF likely plays a crucial role in regulating the HHV life cycle.

It has been shown that CTCF occupies specific DNA binding sites across the HSV-1 genome, and earlier work had characterized the functionality of these sites as so-called

"functional enhancer-blocking insulators", essentially meaning that they acted as a transcriptional roadblock - blocking enhancer DNA elements from reaching certain viral gene promoter regions (Amelio et al., 2006; Washington et al., 2018). Furthermore, HSV-1 evidenced CTCF eviction post-reactivation (Washington et al., 2018).

The above mentioned CTCF binding sites, or motifs, are conserved into clusters in HSV-1. As with EBV, such clustering partitions the transcription profiles into different domains (Amelio et al., 2006). Seven clusters of DNA binding motifs have been identified from whole genome analysis of HSV-1 (Amelio et al., 2006) and were found to be organized around each of the immediate early genes and found flanking the portion of the latency-associated transcript (LAT) gene that is critical to latency and reactivation (Amelio et al., 2006). Whether CTCF insulator activity first regulates transcriptional activity, or whether it is transcriptional activity that first establishes CTCF insulator activity, remains unclear (Hu et al., 2019). In mouse dorsal root ganglia (DRG) tissue, all seven of the CTCF binding clusters were found to be occupied by CTCF during latency. Recombinant viruses lacking IE transcription were used to demonstrate that IE expression evicts CTCF - upon administration of a reactivation signal in these recombinant viruses there is no CTCF eviction (Ertel et al., 2012). Thus, CTCF is shown to have a latency-maintaining role and loss of CTCF is essential to the reactivation of HSV-1 *in-vivo* (Ertel et al., 2012). Finally, CTCF is also able to interact with nuclear structures like lamin (Silva et al., 2008), and as with EBV, CTCF-cohesin complexes were found to be responsible for the formation of chromatin looping that regulates acute and latent transcription profiles (Kang et al., 2011; Arope et al., 2013; Hansen et al., 2017; Schwartz et al., 2017).

CTCF is again seen to play a critical role in regulating the higher order chromatin structure in HSV-1 infection. Overall, the critical role of CTCF in the HSV-1 life cycle must be emphasized. To recapitulate, CTCF is associated with nucleosome shifts, delineates the euchromatic from heterochromatic viral genome regions, blocks enhancer-dependent gene transcription, and complexes with cohesin protein to maintain DNA loops (Amelio et al., 2006; Gaszner and Felsenfeld, 2006; Chen et al., 2007; Rubio et al., 2008). Thus, as with EBV, we have mounting evidence that CTCF is critical to HSV-1 life cycle dynamics and likely works in concert with a variety of host and viral structures and factors to serve as the main chromatin switch in the regulation of the HHV life cycle.

Human betaherpesviruses and host chromatin (HHV-6A)

While the chromatin biology of the Human betaherpesviruses is not as developed as that of the other HHVs, much has been discovered in recent years about the role of chromatin in these viruses. Because T-cells are highly infectable by HHV-6A and because CTCF is involved in VJD recombination (Chaumeil and Skok, 2012), there were initial suspicions of a putative connection between CTCF and HHV-6 chromatin biology.

HHV-6A, along with HHV-6B, and possibly HHV-7, are unique among herpesviruses in that they require the integration of their genome into the host genome to establish latency and transcriptional quiescence. Thus, the chromatin biology surrounding HHV-6A is expected to diverge somewhat from other HHVs. These, differences notwithstanding, HHV-6A still shares many similarities with other viruses of its family.

27

Like other HHVs, it is nucleus-only replicating virus and uses cytoplasmic organelles for viral protein production and protein maturation (Naqvi et al., 2018). As with EBV and HSV-1, the transcriptome of HHV-6A is turning out to be more complex than initially thought. (Gompels et al., 1995; Dominguez et al., 1999; Gatherer et al., 2011; Gravel et al., 2013; O'Grady et al., 2016; Balazs et al., 2017; Tombacz et al., 2017b; Balazs et al., 2018; Depledge et al., 2019; Kara et al., 2019; O'Grady et al., 2019).

HHV-6A can infect a wide variety of cell types due to its ingress via the cell surface receptor CD46 (De Bolle et al., 2005). But CD4+ T lymphocytes and particularly CD8+ T cells are preferentially infected *in vitro* (Grivel et al., 2003). Other than T lymphocytes, cells that have been successfully infected *in vitro* include natural killer cells, oligodendrocytes, epithelial cells, astrocytes, fibroblasts, continuous liver cells, endothelial cells, and embryonic kidney cells (Aimola et al.).

In vivo, it has been demonstrated that HHV-6A can infect brain tissue (Luppi et al., 1994; Chan et al., 2001; Donati et al., 2003), liver tissue (Ozaki et al., 2001; Ishikawa et al., 2002; Harma et al., 2003) tonsillar tissue (Roush et al., 2001) salivary glands (Fox et al., 1990), and endothelium (Caruso et al., 2002). Bone marrow progenitors have been found to harbor latent HHV-6, which may be transmitted longitudinally to differentiated blood cells of different lineages (Luppi et al., 1999). This is confirmed by the demonstration of HHV-6 in monocytes/macrophages (Kondo et al., 1991; Burd and Carrigan, 1993; Kondo et al., 2002a) and dendritic cells (Kakimoto et al., 2002).

Patient-derived cells have become increasingly the subject of study as they provide a physiologic environment like that of the *in vivo* infection. Lymphoblastoid cell lines or LCLs are one such model that, as mentioned earlier, have been heavily used in the study

of EBV, and they can be readily derived from patients' cells and subsequently used to model latent (and other) stages of the HHV-6A virus life cycle.

For the study of iciHHV-6A, which occurs in about 1 percent of the population, a patient-derived iciHHV-6A cell line has been developed from the umbilical arteries of patients with iciHHV-6A (Wallaschek et al., 2016). This patient-derived cell model thus contains a cellular physiology that more closely emulates that of the *in vivo* physiology of infection.

In addition to these patient-derived cell models, *in vitro*-derived cell models have been developed. The best current *in vitro* model is a HEK293T cell model which contains a recombinant HHV-6A genome (Uganda isolate 107 RefSeq id: NC_00164.4) with a green fluorescent reporter protein (GFP) under transcriptional control of a human cytomegalovirus (hCMV) promoter. The hCMV promoter is known to respond to specific phorbol esters, such as TPA, that can be administered by the experimenter to reactivate the virus after it has entered the transcriptionally quiescent latent stage of infection (Saviola et al., 2019a). Thus, with this cell line – deemed ciHHV-6A, the experimenter has some control over the HHV-6A life cycle stage.

When employing the above ciHHV-6A cell model, transcriptional activity is evidenced during the acute lytic stage of infection via immunofluorescence microscopy, which captures the GFP signal. As the virus proceeds to integrate its genome and enter latency, over the course of its incubation period (~1-2 weeks), the attenuation of the GFP signal can be seen to decrease over time, culminating in its complete abrogation with the establishment of latent integration and the associated transcriptional quiescence that is the hallmark of viral latency. Upon stimulation with the phorbol ester 12-O-

Tetradecanoylphorbol-13-acetate (TPA) the recrudescence of the GFP signal indicates a return to transcriptional competency that defines the reactivated stage of the HHV life cycle. RNA sequencing can also be performed over a time course to confirm the establishment of latent infection in this cell model (Saviola et al., 2019a). Fluorescent *in situ* hybridization, or FISH, has been used to identify integrated viral genomes within cell models (Saviola et al., 2019a; Aimola et al., 2020; Aswad et al., 2020) and ChIP-Seq experiments have been used to assay the chromatin deposition along the viral genome during latent infection - demonstrating the enrichment of repressive histone marks (Saviola et al., 2019a).

With cell models in place to study HHV-6A chromatin biology, it is important to remember that integration is a defining characteristic of the HHV-6A life cycle. Structurally, the genome is flanked by direct terminal repeat sequences or "DR" sequences. Each direct repeat is approximately 8 kb in length and consists of two cleavage-packaging motifs called pac-1 and pac-2. There are ORFs related to the direct repeats, ORF1 to ORF7 (in the traditional ORF numbering system). The most important feature of the terminal direct repeats are their imperfect and perfect repeat arrays on the 5' and 3' ends of the viral genome, respectively. The perfect repeat sequence on the 3' end contains hexanucleotide repeats that are identical to the canonical mammalian repeats (TTAGGG)n. This homology is the feature that allows for the integration of the viral genome into the host genome via homologous recombination. The recombination proceeds via either the base excision repair pathway or the single strand annealing pathway. For a thorough review of HHV-6 integration see Aimola et al., 2020 (Aimola et al., 2020). Because HHV-6A has evolved to exploit the chromatin region of the telomere, and because it is believed that the virus needs

to excise itself to undergo competent transcription and replication, telomere chromatin biology is hypothesized to play an important role in coordinating integration and maintaining the stages of the HHV-6A life cycle. The virus has likely evolved to use the higher order chromatin structure of the host to regulate its life cycle, while simultaneously evading the host immune response. It has been further suggested (Aimola et al., 2020) that the site of integration could be directly within the telomere loop that forms, along with chaperone proteins, at the end cap of the human telomere. The nature of this t-loop, present at the telomere ends, allows it to unravel during periods of chromatin restructuring, and the virus could take advantage of this period to integrate itself (Aimola et al., 2020).

Looking to the future, researchers are still unclear as to the exact mechanisms, such as the recruitment of the accessory enzymes necessary to procure integration, that lead to homologous recombination of the infecting HHV-6A virus and the establishment of latency. Fortunately, in the case of HHV-6A, cell models are available to study its chromatin biology. I employ such cell models in the work that I outline in chapter 2 of this dissertation. With that in mind, I will now move on to review the chromatin biology of VZV.

VZV and host chromatin biology

Much more is known about HSV-1 chromatin biology within the alphaherpesvirus subfamily than with its cousin VZV. HSV-1 has subsequently served as a hypothesis-generating template of sorts for the investigation of VZV chromatin biology. Upon the

resolution of an initial VZV infection, which typically presents as chickenpox, the virus takes up residency exclusively in the dorsal root and cranial nerve (trigeminal nerve) ganglia of the peripheral sensory nervous system. VZV is believed to enter ganglia via retrograde axonal transport and hematogenous transport (Kennedy et al., 2015).

Like the other HHVs, the acute stage of the viral infection proceeds with the sequential expression of viral genes by kinetic class: alpha, gamma, and beta (not to be confused with the subfamily designations of the HHV taxonomic nomenclature). Like other HHVs, almost all of the viral expression takes place in the host cell nucleus and so again, host cell chromatin is heavily implicated in regulating the transcriptional profiles that define the VZV life cycle stages. But as mentioned earlier, much less is understood about the role of chromatin in the regulation of these life cycle stages, due to difficulty in establishing successful animal and cell models that could capture the full spectrum of infection, but also because VZV is the only HHV for which a vaccine exists, somewhat mitigating the exigency for a more thorough understanding of the cellular biology of the virus.

Unlike HSV-1, VZV reactivation is more prevalent in the elderly and tends not to be "recurrent" - meaning that reactivation does not present periodically (Kennedy and Steiner, 1994). Such differences could result from fundamental differences in alphaherpesvirus chromatin biology (Kennedy et al.). With such hypotheses in mind, researchers needed suitable animal or cell models with which to evaluate them.

While VZV can infect a variety of cell types and is known take up latent infection in cranial, dorsal root, and autonomic ganglia (Mahalingam et al., 2019), researchers are currently lacking a suitable animal model to study the life cycle stages of VZV. Many

attempts have been made over the years to establish a competent animal model of infection and disease. Mice, cotton rats, and guinea pigs in particular (which were successfully employed in the study of HSV-1), have been employed in attempts to establish a viable animal model of infection (Takahashi et al., 1975; Myers et al., 1980; Matsunaga et al., 1982; Wroblewska et al., 1982; Myers et al., 1985; Walz-Cicconi et al., 1986; Myers et al., 1991; Wroblewska et al., 1993).

Interestingly, it has been found that simian varicella virus (SVV) infection in non-human primates produces varicella, but the main drawback to this particular non-human primate animal model, although it has provided many insights, is that it is a very expensive model to study (Mahalingam et al., 2019). Unfortunately, while a cost-effective animal model capturing the full spectrum of infection and disease has not been successfully created, attempts undertaken to use post-mortem autopsied human ganglia have also been left wanting. Relying on tissue from recently deceased donors presents its own challenges as the decomposition process has already begun and obtaining adequate amounts of viral DNA from samples can prove challenging (Kennedy et al., 2015).

Concerning *in vitro* models, quite a few attempts have been made with varied success. Gershon et al., created a model with enteric neurons isolated from adult guinea pigs and from fetal mice (Gershon et al., 2008). ORF61 protein is capable of inducing reactivation in this model (Kennedy et al., 2015). For as of yet not understood reasons, the viral transcriptome of non-productive neuron infection is similar to that of productive infection of human lung fibroblasts (Kennedy et al., 2015) – a model that will be discussed in chapter 3. Yet for some reason in the neuronal cells, there is no evidence for the accumulation of VZV DNA as there is with the HFL cells (Baird et al., 2014).

Unfortunately, a latent state has not yet been achieved for HFL. In the last few years, human embryonic stem cells have become successful cell models to study VZV infection (Depledge et al., 2018b). Perhaps most excitingly, a cell model cultured from human dorsal root ganglia tissue has been recently employed, demonstrating successful control over the virus life cycle stage by the experimenter (Thellman et al., 2017). These HD10.6 cells may be the most promising candidate model yet for an *in vitro* model of VZV infection.

With these improved models we can be sure that the biology mimics that of the host as much of possible. This is especially important when trying to define the transcriptional profiles of VZV that are subsequently used to delineate stages of the life cycle. It is from the vista of these delineated stages from which we can survey the relevant chromatin biology. The current paradigm holds that VZV, like HSV-1, is likely epigenetically regulated (Kennedy et al., 2015). With the availability of cell models for particular stages of infection, e.g. HFL cells for acute infection, the VZV transcriptional profiles, or kinetic classes, of acute infection have been able to be characterized into immediate early (IE), early (E), and late (L) kinetic classes primarily based on what is known about the more thoroughly understood biology of the alphaherpesvirus HSV-1 as well as from some experiments directly with VZV (Reichelt et al., 2009; Lenac Rovis et al., 2013; Depledge et al., 2018b). It is hypothesized, as with HSV-1, that each set of genes in the kinetic class is reliant on expression of the preceding class in a temporal fashion. Additionally, as with HSV-1, IE genes serve a transcriptional regulatory role, E genes regulate DNA replication, and L genes encode the structural proteins that are required for the formation of virions and their subsequent departure out of the host cell (Depledge et al., 2018b). While studies have been able to categorize VZV genes, difficulties have been encountered when trying to

prepare VZV preparations with adequately high-titer to verify the existence of temporal groups in VZV to the same degree as with HSV-1 (Cohrs et al., 2003; Kennedy et al., 2005; Kennedy et al., 2015).

In VZV, ORF4, ORF10, ORF62, and ORF63 are transcriptional transactivators (Kennedy et al., 2015) and are hypothesized to interact with host cell chromatin. Kitchington et al. first identified them as being present in the VZV virion tegument (Kinchington et al., 1992; Kinchington et al., 1995b). The formation of heterochromatin structures on VZV gene promoters leads to viral gene transcriptional silencing (Everett et al., 2006; Everett, 2010; Newhart et al., 2012). Additionally, ORF66, a VZV viral serine/threonine kinase, can become phosphorylated to ORF66p, which, when activated, will deactivate histone deacetylases. Thus, it has already been demonstrated that key viral proteins do interact with host chromatin factors.

Interestingly, and importantly, ORF62 which encodes the VZV IE62 protein, is described by Depledge et al. as homologous to "The dominant transcriptional regulator and possibly only true immediate-early protein by Varicellovirus" (Kinchington et al., 1992; Perera et al., 1992; Perera et al., 1993; Moriuchi et al., 1994; Tyler et al., 1994; Kinchington and Turse, 1998; Ruyechan et al., 2003; Yang et al., 2006; Khalil et al., 2016). This is an important claim indicating the prevailing sentiment that there is in fact a dominant transcriptional regulator for each HHV. In the case of VZV, IE62 is indeed a major transactivator protein, and can activate genes in all three kinetic classes of the acute infection stage. It can perform this feat even with no other viral protein present (Depledge et al., 2018b). The viral protein can do this during productive infection (acute or reactivated stages) in a variety of cell types, underlining the role of common transcription factors and

or viral proteins in coordinating infections (Depledge et al., 2018a; Depledge et al., 2018b) (Baird et al., 2014; Jones et al., 2014; Markus et al., 2014; Sadaoka et al., 2016). In fact, there are currently four suggested immediate early genes, all of which could potentially interact with host chromatin factors to modulate transcription and thus viral life cycle. They are ORF4, ORF61, ORF62, and ORF63 (Depledge et al., 2018b). The respective protein products IE4, IE61, and IE63, have been shown to have a small to negligible effect on the ORF62 promoter, providing further evidence that viral protein IE62 (encoded by ORF62) is indeed the true major transactivator protein for VZV. (Kost et al., 1995; Michael et al., 1998; Wang et al., 2009; Ruyechan, 2010; Depledge et al., 2018b).

Like the latency-associated transcript (LAT) of HSV-1, IE63 is thought to play an important role in the latent stage of the VZV life cycle where, as with all HHVs, viral transcription is mostly silent. IE63 is nucleus-localized and highly phosphorylated, and likely represses transcription through histone modification, indicating an important role for chromatin interactions with this viral gene product. IE63 has only been observed in tissue culture cells but its transcripts have been found in ganglia, suggesting the protein itself may be present in ganglia where it could inhibit VZV-induced apoptosis (Kinchington et al., 1995a; Sadzot-Delvaux et al., 1998; Di Valentin et al., 2005; Hood et al., 2006; Ambagala et al., 2009; Mueller et al., 2009; Mueller et al., 2010; Pugazhenthi et al., 2011; Ouwendijk et al., 2016). VZV IE63 protein, most likely represses transcription through histone modification (Kinchington et al., 1995a; Sadzot-Delvaux et al., 1998) (Di Valentin et al., 2005; Ambagala et al., 2009; Mueller et al., 2009; Mueller et al., 2010). However, mostly transcripts, and not the protein itself (with few exceptions) have been found in autopsied trigeminal ganglia with a PMI (post-mortem interval) under nine hours (Mahalingam et al.,

1996; Kennedy et al., 2000; Zerboni et al., 2010a; Zerboni et al., 2010b; Ouwendijk et al., 2012a; Ouwendijk et al., 2012b; Kennedy et al., 2015). Because ORF63 is also a key IE transactivator it would have to be altered in some way to change its function from activating virus transcription to repressing virus transcription – and chromatin structural alteration is one means by which this could be achieved.

Even more recently, a Biorxiv preprint by Depledge et al., from 2017, explains that a newly identified transcript, that they call VLT (for VZV latency transcript), could play a larger role in VZV latency than IE63. VLT may be able to limit IE61's histone deacetylase inhibitor activity (Wang et al., 2009; Boutell and Everett, 2013) which would contribute to maintaining chromatin repression during VZV infections. In turn, this chromatinization would limit viral gene transcription in VZV infection (Depledge et al., 2017, Biorxiv preprint 2017). This newly identified VLT, in one or more of its isoforms could be the key gene in maintaining latency. If this is the case, looking for chromatin interactions with this newly identified transcript, particularly with regards to CTCF, could greatly improve our understanding of the role that host cell chromatin biology plays in the switch between acute and latent VZV infections. But it should also be noted that VLT differs markedly from the key latency transcript of its cousin HSV-1. The HSV-1 latency associated transcript (LAT), unlike VLT (at least demonstrated *in vitro* at this point), encodes miRNAs, is involved in blocking apoptosis, and can downregulate the expression of the IE transactivator ICP0. Whereas, we do not see such similarities with VLT as of yet (Depledge et al., 2017, Biorxiv preprint)(Perng et al., 2000a; Perng et al., 2000b; Umbach et al., 2008).

A variety of reactivating stimuli have been identified in different HHVs across different time points and models. Some HHV reactivating stimuli were discussed in earlier

sections of this chapter and in the case of the newly employed HD10.6 model of VZV infection, depletion of nerve growth factor or inhibition of PI3Knase excitingly leads to reactivation of the virus. PI3Kinase has been shown to modulate levels of a variety of histone marks (Spangle et al., 2017) which presumably led to decreases in repressive marks and repressing viral transcription during latency. For example, in the Human gammaherpesvirus Kaposi sarcoma-associated herpesvirus (KSHV/HHV-8), stress has been shown to lead to the dissociation of the CTCF-cohesin ring that binds the episomal virus genome to the host cell genome. This action then leads to the release of paused RNA polymerase II during latency (Kang and Lieberman, 2011; Chen et al., 2017). As mentioned earlier CTCF-cohesin ring complexes can bring together portions of different chromosomes, transforming the three-dimensional higher order chromatin structure in a manner that allows for the regulation of viral transcription (Kagey et al., 2010). The recruitment of RNA polymerase II to CTCF/cohesin complexes, and/or the unpausing thereof, may contribute to reactivation in alphaherpesviruses such as VZV (Chernukhin et al., 2007; Wada et al., 2009; Kennedy et al., 2015). Stress has long been indicated in the reactivation of HHVs (Rooney et al., 2019) and reactivation of VZV could potentially proceed via a similar route to KSHV. Again, as better cell models are developed, and interrogative techniques refined, these more macroscopic questions can begin to be addressed within the context of HHV chromatin biology. The aforementioned development of the HD10.6 human dorsal root ganglion-derived *in vitro* cell model will allow alphaherpesvirus researchers to finally have a viable cell model that closely mimics the *in vivo* infection physiology with which to study the acute, latent, and reactivated forms of VZV infection.

With such a model, our understanding of the complex interplay of chromatin biology and VZV in the host cell nucleus will likely accelerate. We will be able to better understand the higher order chromatin structures, epigenetic factors (e.g., histone marks and the regulation of histone modifying enzymes), and overall chromatin states, that likely work in tandem with the complex cellular and viral pathways to regulate the VZV life cycle.

Concluding remarks

In this chapter, I have introduced the basic virology of Human herpesviruses and their role in human disease. I then proceeded with a brief overview of human/host cell chromatin biology and specifically highlighted the importance of the transcription factor CTCF. Following a brief overview of current genomics methodologies employed in the study of chromatin biology, I moved to a review of the literature encompassing our current understanding of the role of chromatin biology in the HHV life cycle. I began largely chronologically with a discussion of the literature relevant to EBV and HSV-1, about which most of our knowledge of chromatin and HHVs is derived, and ended with an examination of the state of our understanding of the chromatin biology of HHV-6A and VZV infections. Within the context of HHV chromatin biology, comprehension largely recapitulates chronology. I have also tried to arrange the discussion within the boundaries of the individual life cycle stages and kinetic transcriptional classes that are characteristic of all HHVs.

While there is a relatively less known about HHV-6A and VZV, this is precisely why these viruses require the attention of virologists and chromatin biologists. The fact that HHV-6A integrates into subtelomeric regions of the host genome or that tractable animal models of VZV infection have not been achieved, are to emphasize that these viruses as all the more deserving of our scientific endeavor in my opinion.

In the ensuing chapters of this dissertation, I will present my contributions in uncovering the higher order chromatin interactions and chromatin states in HHV-6A and VZV in their latent and acute life cycle stages, respectively. In chapter 2, I demonstrate the successful establishment of latent infection in the best current *in vivo* cell model we have for HHV-6A infection – the HEK293/ciHHV-6A model. I characterize the chromatin state as well as the higher order chromatin interactions that the virus makes with the host cell genome. In addition, I characterize the chromatin landscape of the HHV-6A latent infection in both the ciHHV-6A cell model and in the patient-derived cell model as well as in primary lymphocytes from patients with iciHHV-6A. Finally, I present a novel computational genomics methodology that I developed to identify sites of HHV-6A integration using high-throughput sequencing data, the first time such a task has been achieved. In chapter 3 I successfully characterize the higher order chromatin structures and chromatin states of the HFL cell model of acute VZV infection. Building off genomics and computational techniques that I developed for my work in chapter 2, I proceed to employ the above methodology, along with other genomics technologies, to identify higher order chromatin associations and patterns of CTCF occupancy in acute VZV infection that is countervailing to the current understanding of the role that CTCF plays during an acute VZV infection

and hints at its potential role in regulating the chromatin structure that is hypothesized to modulate the VZV life cycle.

CHAPTER 2

Title

Higher-order chromatin structures of chromosomally-integrated HHV-6A predict Integration Sites

Authors

Michael Mariani[1], Cosima Zimmerman[2], Princess Rodriguez[1], Ellie Hasenohr[1], Diana Gerrard[1], Giulia Aimola[2], Alyssa Richman[1], Andrea Dest[1], Louis Flamand[3], Benedikt Kaufer[2]* and Seth Frietze[1,4]*

Affiliations

[1]Department of Biomedical and Health Sciences, College of Nursing and Health Sciences, University of Vermont, Burlington, VT, USA

[2]Institute of Virology, Freie Universität Berlin, Berlin, Germany

[3]Division of Infectious Disease and Immunity, CHU de Québec Research Center-Université Laval Quebec City, Quebec Canada G1V 4G2

[4]University of Vermont Cancer Center, Burlington, VT, USA

Correspondence

seth.frietze@med.uvm.edu

benedikt.kaufer@fu-berlin.de

Abstract

Human herpesvirus -6A and 6B (HHV-6A/B) can integrate their genomes into the telomeres of human chromosomes. Viral integration can occur in several cell types, including germinal cells, resulting in individuals that harbor the viral genome in every cell of their body. The integrated genome is efficiently silenced but can sporadically reactivate resulting in various clinical symptoms. To date, the integration mechanism and the subsequent silencing of HHV-6A/B genes remains poorly understood. Here we investigate the genome-wide chromatin contacts of the integrated HHV-6A in latently-infected cells. We show that HHV-6A becomes transcriptionally silent upon infection of these cells over the course of seven days. In addition, we established an HHV-6–specific 4C-seq approach,

43

revealing that the HHV-6A 3D interactome is associated with quiescent chromatin states in cells harboring integrated virus. Furthermore, we observed that the majority of virus chromatin interactions occur towards the distal ends of specific human chromosomes. Exploiting this finding, we established a 4C-seq method that accurately detects the chromosomal integration sites. We further implement long-read minION sequencing in the 4C-seq assay and developed a method to identify HHV-6A/B integration sites in clinical samples.

Introduction

Human herpesvirus 6A (HHV-6A) and 6B (HHV-6B) are two closely related and ubiquitous betaherpesviruses (Zerr et al., 2005). Most people are infected with HHV-6B as infants, while the etiology of HHV-6A remain poorly defined (Okuno et al., 1989). Following primary infection, HHV-6A/B establishes latency in the host for life (Flamand, 2018). During this latent state, viral gene expression and viral load are not detectable (Saviola et al., 2019b). However, the virus can sporadically reactivate in the host, resulting in the production of infectious virions and viral transmission (Endo et al., 2014). HHV-6A/B reactivation has been linked to a variety of pathologies including encephalitis, graft rejection and a spectrum of other diseases (Aimola et al., 2020).

HHV-6A/B are unique among human herpesviruses as they are able to integrate their genomes into the telomeres of human chromosomes (Arbuckle et al., 2010; Osterrieder et al., 2014). Chromosomal integration of HHV-6A/B occurs very efficiently

in vitro and is dependent on telomere sequence arrays present at the ends of the virus genome (Wallaschek et al., 2016; Gravel et al., 2017). HHV-6A/B can also integrate into the germline, resulting in individuals harboring the virus in every nucleated cell of the body and inherit it to their offspring. This inherited chromosomally-integrated HHV-6 (iciHHV-6) is present in about 1% of the human population (Kaufer and Flamand, 2014) with several independent integration events occurring thousands of years ago (Aswad et al., 2020). Integration into the telomeres allows maintenance of the virus genome in iciHHV-6 individuals and latency (Arbuckle et al., 2010); however, it remains unknown if integration into the telomeres contributes to the silencing of the virus genome.

Here, we explored if, how, and when the HHV-6A genome is silenced upon integration using an *in vitro* integration model and RNA sequencing. We demonstrate that HHV-6A genes are silenced upon integration over the course of 7 days. We employ circular chromosome conformation capture assays (4C-seq) to assess the higher order chromatin structures formed between the virus and host genome in these cells. Because HHV-6A/B integrate into highly repetitive telomeric regions of individual human chromosomes, we used 4C-seq-based analysis as a tool to identify integration sites. Our 4C-seq approach employs novel scoring method as well as Oxford Nanopore minION long-read sequencing to effectively identify integration sites in human telomeres. Overall, this study provides a better understanding of the chromatin programs that may regulate HHV-6A latency and provide novel diagnostic methods to determine chromosomal integration sites.

Materials and Methods

Ethics

Specimens were obtained from the Fred Hutch Research Cell Bank, which prospectively collects and cryopreserves peripheral blood mononuclear cells (PBMCs) from donors and recipients. The University of Washington Institutional Review Board approved use of the iciHHV-6 specimens from the Fred Hutchinson Cancer Research Center and use of anonymized excess HHV-6-positive samples submitted for testing at the University of Washington Virology lab.

Cell lines and virus

SMC cells harboring integrated HHV-6A virus (iciHHV-6A cells) were immortalized following transduction with a lentiviral vector (pLenti CMV/TO SV40 small + large T (W612-1), a gift from Eric Campeau and obtained from Addgene (plasmid #22298)) and expressing the SV40 T antigens. SMCs were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 20% fetal bovine serum (FBS), 1% penicillin streptomycin (Pen/Strep), and 1% L-Glutamine. Human epithelial kidney 293 T (293 T, ATCC CRL-11268) cells were cultured in the same medium but supplemented with 10% FBS. All cells were maintained in 10 cm² flasks as a monolayer culture in a humidified 5% CO2 air incubator at 37°C. Bacterial artificial chromosome (BAC)-derived

HHV6A (strain U1102) expressing green fluorescent protein (GFP) under the control of the HCMV major immediate early (IE) promoter (293-HHV-6A virus) was propagated in J-Jhan cells as described previously (Tang et al., 2010b). 293 T cells were infected with 293-HHV-6A virus and GFP positive cells were isolated using a FACS AriaIII cell sorter (BD Biosciences). Clones harboring the integrated 293-HHV-6A genome (ciHHV-6A cells) were identified by quantitative PCR (qPCR) and HHV-6A integration was confirmed by fluorescent *in situ* hybridization (FISH). iciHHV-6B+ human lymphocytes (NCBI GenBank: KY315552) were isolated from an infected patient as described by Aswad et al. (Aimola et al., 2020).

Fluorescence *in situ* hybridization

FISH was performed with digoxigenin-labeled bacterial artificial chromosome (BAC) probes against both the HHV-6A genome and individual chromosomal reads essentially as described previously (Kaufer et al., 2011; Kaufer, 2013). Slides were mounted using DAPI Vectashield (Vector Laboratories) and images were taken with an Axio Imager M1 (Zeiss).

RNA-Seq analysis

RNA was extracted from an HHV-6A infected 293T cell line using Trizol Reagent (Life Technologies) and purified using Direct-zol RNA MicroPrep Kit (Zymo Research

#R2060) following the manufacturer's instructions. One microgram of total RNA was depleted of ribosomal RNA (rRNA) using the KAPA RiboErase Kit (#KR1142) and libraries were prepared using the KAPA Stranded RNA-seq Library Preparation Kit (#KR0934). Libraries were quantified using Qubit (Life Technologies), and quality was assessed using the Agilent Bioanalyzer High-Sensitivity DNA kit (Agilent Technologies). Barcoded libraries were pooled and sequenced on an Illumina HiSeq 2000 to obtain 100-bp paired-end reads. RNA-seq data were processed using STAR as described previously (Dobin et al., 2013) using human (hg38) and a custom HHV-6A reference genomes (NC_001664.4) modified to contain GFP transgenes. BAM files were sorted and converted to SAM files using SAMtools (Li et al., 2009) and reads were counted with featureCounts (Liao et al., 2014) against the corresponding viral GTF file and normalized in DESeq2 (Love et al., 2014). Alignment of viral reads across the HHV-6A genome was visualized using the "seqsetvis" R library (J, 2020) and in-house scripts.

4C-Seq library preparation and sequencing

In order to identify the higher order chromatin structure and virus-host DNA contacts, 4C-seq libraries were prepared using both ciHHV-6A and iciHHV-6A samples as described previously (Krijger et al., 2020), using HindIII/DpnII restriction enzymes. Cells were cultured as previously described (Saviola et al., 2019b) and fixed with 1% formaldehyde prior to being snap frozen. For each 4C library, ten million cells were used unless otherwise noted. Inverse PCR was performed in 50 uL reactions using 200ng 3C

library template, 25 uL Q5 Hot Start High-Fidelity DNA Polymerase 2X Mastermix (New England Biolabs), 1.5 ul of 10uM forward and reverse inverse PCR primers and water under the following conditions: 30s 98°C, 10s 98°C, 30s 55°C, 2m 72°C, 5m 72°C for a total of 25 cycles. PCR products were cleaned up using the PCR purification kit (Thermofisher Scientific) using the B3 reagent to exclude fragments less than 300 base pairs or with Ampure XP beads (Beckman Coulter) and eluted in water. Library indexing was performed using the DNA HT Dual Index Kit (Illumina) or with custom dual indexing primers with an additional 8 cycles of PCR. Indexing of nanopore libraries was performed with custom 12-nucleotide dual index primers with 8 additional PCR cycles. All primers sequences are listed in **Supplemental Table 2.1**. Libraries were pooled and sequenced with either HiSeq2000 or MinION platforms. The experiments were performed with samples from separate cell culture dates. For the Illumina viewpoint comparison libraries, two replicates were performed for two separate viewpoints. For the ONP titration and viewpoint comparison studies, one and two replicates were performed, respectively.

4C-Seq primer design

We developed a 4C-seq primer design tool[1] to facilitate the generation of the inverse PCR primers necessary for 4C-Seq assays in a genome-agnostic manner for use with experiments involving custom genomes (i.e., chromosomally integrated HHV-6A). A fasta file containing the HHV-6A genome (NC_001664) was obtained from NCBI and

---

[1] https://github.com/FrietzeLabUVM/4c_primer

uploaded into the design tool, selecting HindIII and DpnII restriction endonucleases, and minimum fragment sizes 500 bp for the first fragment and 300 bp for the second fragment were used (van de Werken et al.).

4C-Seq analysis

4C-seq *trans* interactions were determined using a "window" method as described by Kim et al. (Kim et al., 2020). Briefly, the virus viewpoint regions at the beginning of the reads were trimmed off using Cutadapt (Martin) software and the remaining human portions were aligned to the human UCSC hg38 reference genome using bowtie2 (version 2.3.5) (Langmead et al., 2009; Langmead and Salzberg, 2012). P-values for each mapped base were then calculated for each 10kb window across all chromosomes using a Poisson formula. Final, significant peak regions are then called using the MACS2 (version 2.2.7) bdgpeakcall function (Zhang et al., 2008). Significant interacting *trans* peak regions were identified using the above window method for each replicate, for each viewpoint, in both the ciHHV-6A and iciHHV-6A samples. These regions were plotted as circos plots using the Circlize R library (Gu et al., 2014). Bedtools (Quinlan and Hall, 2010) was used to find the intersection of the *trans* regions and the 127 epigenomes 25-state imputation based chromatin state model created with ChromImpute (Ernst and Kellis, 2012; Ernst and Kellis, 2015) and made available from the Wang Lab[2] (Jin Wook Lee). For ciHHV-6A cells, chromatin state annotation was also performed using fetal kidney and fetal adrenal gland

---

[2] https://egg2.wustl.edu/roadmap/web_portal/imputed.html#chr_imp

tissues – also made available from the Wang lab[2]. 4C *trans* peaks and signals were also compared to signals for the repressive histone marks H3K9me3, NCBI GEO: GSE66530 (Hattori et al., 2016), and H3K27me3, NCBI GEO: (Lamb et al., 2019), in HEK293 cells and compared to the same marks, NCBI GEO: GSE121984, in iciHHV-6A cells (Saviola et al., 2019b).

4C-seq *cis* interactions were determined using the program peakC (Geeven et al., 2018). Telomere content comparison was assessed by TelomereHunter (Feuerbach et al., 2019). Illumina data was simulated using ART (Huang et al., 2012) set to HiSeq2500 data. Reads were produced to cover the entire UCSC hg38 reference genome at 1X coverage. Then, a number of simulated reads equivalent to the number of reads sequenced between 4C data were randomly sampled from the simulated data. TelomereHunter was run on all samples and a non-parametric Wilcoxon test was performed to compare 4C-seq reads to simulated reads.

To identify the chromosomes harboring HHV-6A integration, a scoring system was developed and coded using the R programming language. Accordingly, trimmed reads were aligned to the hg38 genome using bowtie2, or BWA-MEM (Li, 2013) for MinION reads and each read that maps within 500 kb of any autosome terminus was scored. The total scores for reads within these 500 kb regions are tallied and then compared to each other in one of two ways. If replicates are available, a two-way ANOVA is first performed with chromosome end sum scores as the outcome variable and individual chromosome ends as the categorical predictor variable. TukeyHSD pairwise-post hoc tests are then performed and the chromosome end with the highest mean -$\log_{10}$ p-value is determined to be the most likely candidate. A significance of 0.05 was used for this approach. If replicates

were not available for a specific viewpoint/cell type combination, the chromosome end scores were all compared to each other using Wilcoxon nonparametric tests. The chromosomal end with the largest $-\log_{10}$ p-value adjusted for multiple test comparison was chosen as the most likely candidate and an alpha of 0.05 was used as a cutoff of significance. Furthermore, statistical ranking via the aforementioned pairwise post-hoc tests from the algorithm score sums (for each chromosome end) to identify the ends of each chromosome as most significant candidate for HHV-6A integration.

Results

Dynamics of HHV-6A gene expression during the establishment of latency

We recently reported that the chromosomally integrated HHV-6A genome exists in a compacted transcriptionally silent state (Saviola et al., 2019b). To study the kinetics of HHV-6A gene silencing upon HHV-6A infection, we performed a 7-day time course in human 293T cells. 293T cells are susceptible to HHV-6A infection and were previously established as a model for studying virus integration (Arbuckle et al., 2010; Gravel et al., 2017; Saviola et al., 2019b). Cells were infected with recombinant HHV-6A virus expressing GFP under the control of the major immediate-early (IE) HCMV promoter (Tang et al., 2010a; Wallaschek et al., 2016). Infected GFP-positive cells were isolated by FACS sorting 16 h post infection and subsequently cultured. The presence of the HHV-6A genome was confirmed by qPCR and integration was validated by FISH at day seven post-

infection. Infected cells were collected daily until 7 days post-infection (dpi) and processed for gene expression profiling via RNA-sequencing (**Figure 2.1.A**). The expression of all HHV-6A genes progressively decreases over the 7-day period as visualized via a clustered heatmap (**Figure 2.1.B**). The observed gene silencing over this period course can be grouped into 3 clusters based on expression levels of all viral genes (early, mid and late). The collective expression of genes is reduced over a 3-day period, followed by near complete silencing from 5-7 dpi (**Figure 2.1.C**). Thus, HHV-6A gene expression upon infection is silenced within 7 days in 293T cells harboring the integrated virus genome.

Higher-order chromatin interactions of chromosomally integrated HHV-6A

The epigenetic mechanisms that regulate gene silencing in chromosomally integrated HHV-6A/B remain poorly characterized. We previously demonstrated that the integrated viral genome forms repressed heterochromatin domains (Saviola et al., 2019b). We further hypothesized that higher-order chromatin interactions play a role in virus gene silencing. To investigate higher-order chromatin interactions between chromosomally integrated HHV-6A, we designed 4C-seq assays using distinct viewpoint regions designed against distinct HHV-6A regions (**Figure 2.2.A**). To obtain the inverse PCR primers required for 4C analysis, we developed a genome agnostic 4C primer design tool (see methods). Two distinct viewpoints (vp1 vs. vp2) were designed based on optimal fragment length and restriction enzyme sequence along the length of the ~160 kb HHV-6A genome. The HHV-6A genomic regions targeted by viewpoint vp1 and vp2 primers are located adjacent to the

U39 gene-encoding glycoprotein B (gB) (position: 64 kb) and the U95 gene (position: 148 kb), respectively (Nicholas and Martin, 1994). With this approach, we assessed the HHV-6A chromatin conformation in a previously described 293T cells line harboring the chromosomally integrated HHV-6A (ciHHV-6A) (Saviola et al., 2019b) and in iciHHV-6A+ patient-derived cells (iciHHV-6A). We obtained approximately 2 hundred thousand to 2 million reads per viewpoint replicate across 2 independent 4C-seq assays, of which 85.93% of reads mapped to the combined reduced human+HHV-6A genome (hg38 + NC_001664.4 or recombinant GFP genome, depending on the sample) (**Supplemental Table 2.2.1**). Significant '*trans*' interacting regions between the HHV-6A genome and the human genome were identified for both *in vitro*-derived (ciHHV-6A) and patient-derived iciHHV-6A cells (Figure 2B). For ciHHV-6A, we found a total of 6 high-confidence interactions that intersected between viewpoints, and a total of 3 interactions were between viewpoints for iciHHV-6A cells. In addition, a number of *cis*-interacting regions were identified for both viewpoints (**Figure 2.2.C**; **Supplemental Table 2.2.2 and 2.2.3**). In line with other 4C-seq results (Krijger et al., 2020) there were relatively high densities of interactions that occur near and between viewpoints. Overall, there were a number of *cis* interactions identified in each cell types suggesting that virus genome may reside in a highly folded compartment similar to a topologically associated domain.

   To assess the chromatin features of the identified interaction regions, we annotated the identified interacting human regions using available chromatin state segmentations compiled from multiple human tissue types (Ernst and Kellis, 2015). This annotation dataset is derived from 12 epigenetic features and from 127 original reference epigenomes. We found that the human chromatin that interacted with HHV-6A is largely composed of

quiescent and heterochromatin states, whereas enhancer and transcriptional activation annotations occur to a lesser degree (**Figure 2.2.D**). Similarly, chromatin state annotations derived from single representative cell types exhibit similar chromatin state patterns at the HHV-6A interacting genomic regions (Supplemental Figure 3). We further inspected the enrichment of the repressive histone modifications H3K9me3 and H3K27me3 at both ciHHV-6A and iciHHV-6A 4C-seq peaks using ChIP-seq data from each respective cell-line (**Supplemental Figure 2.2**). In both ciHHV6-A and iciHHV6-A, the majority of the significant 4C-seq peaks overlap directly with these repressive histone marks. These results indicate that the integrated HHV-6A genome forms higher order chromatin structures within the viral genome as well as with regions of the human genome that are enriched with repressive chromatin. Overall, these results provide insight into HHV-6 chromatin structures within the host cell nucleus.

Proximity chromatin ligation reveals integration sites

HHV-6A/B integration sites were first detected by FISH (Nacheva et al., 2008), and the virus-host junction of three different integrations was sequenced using a PCR-based approach. Although PCR amplification and Sanger sequencing provided sequence information, this approach requires previous knowledge of the chromosomal location of the virus genome (Huang et al., 2014; Tweedy et al., 2016) and is prone to amplification biases and sequencing errors due to the repetitive nature of the region (Arbuckle et al., 2010; Arbuckle et al., 2013; Huang et al., 2014; Ohye et al., 2014; Gulve et al., 2017).

Genome-wide mapping of short read sequence data to human subtelomeres is challenging due the repetitive nature of subtelomeric regions. It is well-documented that higher order chromatin interactions largely occur in *cis* in large megabase pair (Mbp)-sized topologically associated domains (TADs) (Szabo et al., 2019). Because HHV-6A integrates into a host chromosome, the extra-chromosomal *trans* 4C-seq interactions that are identified should behave like endogenous intrachromosomal *cis* interactions, i.e., 4C-seq should show distinct interacting regions between the host and HHV-6A genomes. With this in mind, we hypothesized that 4C-seq data could identify HHV-6A integration sites, and visualization of 4C alignments to the human genome reveals particular telomeric ends that are enriched with 4C signal. **Figure 2.3.A** shows enrichment at the distal end of chromosome 15q relative to the rest of the genome in ciHHV-6A cells. In addition, clustering of reads at the distal ends of chromosome 15q and chromosome 19q in ciHHV-6A and iciHHV-6A, respectively, can be seen in **Figure 2.3.B**. Further, we identified significantly more telomeric sequences in our 4C-seq data relative to simulated reads of equal size and GC content (**Supplementary Figure 2.1**). To enable systematic analysis of HHV-6A integration sites using 4C-seq, we established a scoring procedure to assess the chance of integration across all chromosomes based on clustered read mapping (see methods). The ends of chromosome 15q and chromosome 19q rank as the highest scoring chromosome ends for ciHHV-6A and iciHHV-6A, respectively, in terms of read mapping density and integration probability (**Figures 2.3.C and 2.3.D**). This indicates that these two chromosome ends were the most likely candidates for harboring integrated HHV6A in the ciHHV-6A and iciHHV-6A cell models, respectively. We confirmed that these chromosomal regions harbor the integrated virus by FISH using probes specific for the

virus genome and the respective human chromosomes (**Figure 2.3.E**). The integrated HHV-6A genome indeed was present in the respective chromosomal loci, highlighting that the 4C-seq analysis is an unbiased method to identify HHV-6 chromosomal integration sites.

4C-seq integration site analysis using Nanopore sequencing

Next generation sequencing technologies have made it possible to generate large quantities of sequence data required for a variety of genomic assays including 4C-seq. However, 4C-seq library sizes derived from inverse PCR are relatively large and often difficult to cluster on Illumina flowcells. Long read sequencing platforms, including the Oxford Nanopore Technologies (ONT) minION, have become common an option for long read sequencing and have been applied for structural genome mapping approaches (De Coster et al., 2019). We therefore evaluated minION sequencing on the 4C-seq libraries as a method to identify integration sites from HHV-6 samples. We sequenced the same 4C-seq libraries with minION flowcells and obtained 150,646,076 bp of demuxable reads with a mean read length of 604.0 bp for the ciHHV-6A and iciHHV-6A sequencing run. For the cell counts titration sequencing run we obtained 419,121,238 reads with a mean read length of 730.4 bp. Alignment of reads that pass quality filters to the human genome reveals a similar grouping of telomere proximal reads at 15q and 19q for ciHHV-6A and iciHHV-6A samples, respectively (**Figure 2.4.A**). Current 4C-seq peak calling software are not well-suited for long-read with low quality scores. We thus adapted our own method based

off clustering of adjacent 4C-seq reads (methods). Using this approach, we found an overall comparable number of *trans* peaks called between replicates (**Figure 2.4.B**). Thus, minION sequencing data represents a sequencing method to generate 4C-seq read data.

We applied the minION 4C-seq reads to investigate the required input material required for identifying HHV-6A integration sites by titrating the number of cells for 4C library construction. We compared a range of input cell quantities for 4C library construction (10^4 to 10^7 cells) using the iciHHV-6A cells combined with minION sequencing (**Figure 2.4.C**). These results confirm that 19q harbors the integrated HHV-6A genome for iciHHV-6A SMC cells, and this can be reliably detected in as few as 100,000 cells for the 4C-seq assay. However, using just 10,000 cells 19q also scores highest, however there appears to be some ambiguity as 14q also scores highly. Finally, we also applied minION 4C-seq to detect HHV-6A integration sites using frozen patient-derived B cells from iciHHV-6A+ individuals for this assay. Like the iciHHV-6A SMC samples, 19q also scores as the likely integration site for HHV-6 integration in additional patient-derived lymphocyte samples (**Figure 2.4.D**). FISH validation using chromosome 19q probes of the patient samples confirms HHV-6A integration into chromosome 19q (Aswad et al., 2020; Wight et al., 2020). Thus, 4C-seq libraries can be generated as little as 10,000 cells and using cryopreserved lymphocyte cell samples. In summary, this workflow allowed us to efficiently identify the chromosomal ends that most likely harbor the integrated virus genome.

Discussion

HHV-6A and HHV-6B, potentially along with HHV-7 (Prusty et al., 2017) are the only known pathogens to integrate into human telomeres (Gulve et al., 2017; Prusty et al., 2017). While the chromatin mechanisms that govern HHV-6A latency and reactivation remain overall poorly characterized, it has been suggested that integration of the HHV-6A/B genome facilitates the maintenance of the virus genome latency (Arbuckle et al., 2010; Huang et al., 2014). The establishment quiescent/latent states has been investigated in a number of cell lines including 293T cells (Gravel et al., 2017). In these *in vitro*-generated cells and patient-derived iciHHV-6A cells, the virus genome transcriptionally silent (Saviola et al., 2019b). These global RNA-sequencing data are in contrast to Kondo et al., who detected a few latency-associated transcripts by qRT-PCR (Kondo et al., 2002b).

To investigate the establishment of a latent-like expression profile, we infected cells and performed RNA-sequencing during the first days of integration. Indeed, in early time points (days 1-3, **Figure 2.1**), we detect high expression levels of all HHV-6A RNA transcripts. However, over the period of 5 days post infection the viral transcripts were significantly reduced and became virtually undetectable by 7 days post infection. It remains unclear whether the RNA-expression is derived from an integrated or extrachromosomal virus genome. However, after 7 days post infection only integrated virus genome were detected when gene expression is absent.

Our prior analyses of the heterochromatin enrichment profiles of integrated HHV-6A genomes revealed that the integrated viral genome was resistant to MNase digestion, suggesting that the viral genome forms a compact chromatin domain (Saviola et al.,

2019b). Consistent with these findings, they found that the viral genome was enriched with the repressive posttranslational histone modifications H3K27me3 and H3K9me3. These results suggested that repressive chromatin structures likely play a role in viral gene silencing.

To further explore the chromatin-mediated mechanisms associated with silent chromatin, we applied the 4C-seq to determine the genome-wide chromatin contacts of HHV-6A based on nuclear proximity ligation. 4C-seq is designed to examine the genome for sequences contacting a selected genomic site of interest. Accordingly, we designed 4C-seq assays with viral viewpoint primers to identify the HHV-6 chromatin interactome. We generated high-resolution contact profiles for distinct viral viewpoints and detail the chromosomally integrated HHV-6 genomic organization. We find that clonally expanded *in vitro* HHV-6A (ciHHV-6A) cells display unique virus-host interactions relative to patient-derived iciHHV-6A cells. Overall, there are not many significantly detected inter-chromosomal or *trans* contacts across the human genome. However, the interactions that we detected were largely enriched with chromatin annotated as being quiescent or heterochromatin. Overall, this analysis indicates that the HHV-6A genome folds in a manner that it is capable of physically interacting with repressive chromatin and may explain some of the mechanisms used to silence virus gene expression in chromosomally integrated HHV-6A cells. To our knowledge, these results were the first reported chromatin interaction for a betaherpesvirus, and it is possible that these interactions between the virus and repressive chromatin elements play distinctive functions in HHV-6A latency.

We also found that the integrated HHV-6A genome forms several shared intra-chromosomal or *cis* contacts between the two ciHHV-6A cell types and that the virus

genome exists in a highly folded state. It is possible that the HHV-6 genome forms a topologically associated domain within the virus and neighboring human sub-telomeric regions. Topologically associated domains or TADs are large regions of local intrachromosomal interactions. In support of viral TAD formation, we find a strong enrichment of telomeric sequence content in HHV-6A 4C-seq data. Interestingly, these domains appear to extend past the non-unique subtelomeric regions into the unique portion of the human genome and aligned reads cluster to the ends of 15q and 19q stand out in this data. Importantly, these regions score as the strongest interaction sites and confirm as being the integration chromosomes in independent FISH assays using probes to both virus and the identified chromosomal regions. Further experimental and computational methods including HiC and TAD calling algorithms are required to comprehensively determine HHV-6A TAD structures.

The identified *trans* interacting human sites as well as the strong signals along the distal ends of distinctive human chromosomes indicated that the 4C assay may prove useful to identify integration sites. Current 4C-seq peak calling methods are designed to identify interacting peaks based on monotonic shape as well as proximity to viewpoint regions. In addition, most 4C-seq peak calling methods are not designed to detect *trans* interactions between two unrelated genomes, although recent methodological improvements have been reported in the case of Epstein Barr Virus (Kim et al., 2020) and we were able to successfully apply such a method to our Illumina sequencing data.

To further establish 4C assays for calling integration sites, we developed a simple scoring method that scores and statistically evaluates candidate chromosomal integration sites. This method bins read alignments that occur at the distal 500 kb ends of each

chromosome and performs statistical analysis to provide a useful score to evaluate chromosomal integration. For the ciHHV-6A and iciHHV-6A samples, this method accurately predicts the integration site. We further used this method to evaluate reads derived from a minION flow cell from a titration experiment with single replicates for each titration level. In each case, this method accurately identified validated integration regions as top scoring candidate regions and helped to demonstrate that as few as $10,000 - 100,000$ cells can be used as input for 4C-seq assays.

Chromatin conformation capture methods have been previously applied to the study of physical chromatin interactions between host and viral genomes. For herpesvirus latency, HiC was used to demonstrate that the latent EBV will disassociate from repressive heterochromatin compartments and form new associations with transcriptionally permissive euchromatin upon reactivation (Moquin et al., 2018). 4C-seq was recently used with Burkitt's lymphoma cells, showing that the latent EBV episomes make contact with transcriptionally repressive H3K9me3 sites as well as attachment sites associated with transcriptionally silent genes (Kim et al., 2020). Finally, 4C-seq has been applied to study the chromatin structures of the latent HIV proviral genome (Dieudonne et al., 2009).

Chromatin conformation capture libraries generally require a large amount of starting material, i.e. as many as 10 million cells (Nilsen, 2014). To enable the 4C assay towards a routine and throughput method to identify integration sites in clinically relevant samples, we performed the assay using a titration of cell numbers ranging from 10^4 to 10^7 cells. Indeed, our titration results demonstrate that we can reliably detect sites as low as 10,000 cells (Figure 4). However, with this low number of cells we identified a few potential spurious hits including chromosome 14. Another adaptation to the assay, is the

use of the Nanopore sequencing platform. This newer sequencing platform has several advantages over Illumina, including cost and speed of data generation. In particular relevance to this study, 4C libraries generated via inverse PCR can be greater than 1 kb in size. The resulting Illumina libraries can be difficult to cluster, resulting in a poor output from the sequencing run at a higher cost. We therefore used Nanopore to generate data for integration analysis and find that using our scoring method that this is adequate for integration analysis. Due to the reduced read quality and relatively lower output, it remains to be determined if Nanopore-derived 4C-seq analysis compares to Illumina analysis for interaction peak analysis.

Finally, it is important to note that the resolution of 4C-seq can be as high as 1-2 Mbp for 4C-seq interactions (Krijger et al., 2020) and the annotation of the corresponding physical interactions can be challenging. Higher resolution 3C methods, including capture HiC can be applied in future studies to better study the nature of these identified physical interactions. These issues of scale are particularly important in the case of HHV-6A integration because the interactions between the virus and the host genomes are akin to *cis-trans* interactions. The HHV-6A genome is simultaneously acting like an independent chromosome (*trans*) that integrates into a particular human chromosome (*cis*) and we need to consider these interactions as a special case of 4C interactions, i.e., something like a "*cis-trans*" interaction. Because of this simultaneous "*cis*" nature, we chose to focus on the *trans* interactions in the 1-2 Mb range that is typical of the size of *cis* interactions reported in literature. Future studies aimed at more precisely defining the cellular factors and human contact points made by integrated HHV-6 will facilitate a more mechanistic understanding of viral gene silencing.

In summary, we have utilized a 4C-seq framework to identify both *trans* (virus-host) and *cis* (virus-virus) interactions that are formed within the human host cell nucleus. We further utilized this assay towards the unbiased identification of HHV-6A integration sites in human chromosomes. This optimally complements our recently developed optical mapping approach for the integrated HHV-6 genome (Wight et al., 2020). While research into the mechanisms and factors governing HHV-6A integration is an active area of ongoing research (Aimola et al., 2020), implementation of the methodology presented in this study provides important information on the silencing of the integrated virus genome and lays the foundation for high throughput detection of HHV-6A/B integration sites in clinical samples.

Funding

Acknowledgements

Author contributions

BK and SF designed and supervised the study and acquired funding. CZ performed cell culture, qPCR, FISH, and flow cytometry experiments. GA performed FISH experiments. MM, EH, PR, AR, and DG performed 4C-seq and RNA-seq experiments. MM and SF performed the analysis of the sequencing data. LF generated and provided cells. MM, AD and SF wrote the manuscript. MM, BK, SF, LF, DW, and GA revised the manuscript.

References

Aimola, G., Beythien, G., Aswad, A., and Kaufer, B.B. (2020). Current understanding of human herpesvirus 6 (HHV-6) chromosomal integration. *Antiviral Res* 176**,** 104720. doi: 10.1016/j.antiviral.2020.104720.

Arbuckle, J.H., Medveczky, M.M., Luka, J., Hadley, S.H., Luegmayr, A., Ablashi, D., et al. (2010). The latent human herpesvirus-6A genome specifically integrates in telomeres of human chromosomes in vivo and in vitro. *Proc Natl Acad Sci U S A* 107(12)**,** 5563-5568. doi: 10.1073/pnas.0913586107.

Arbuckle, J.H., Pantry, S.N., Medveczky, M.M., Prichett, J., Loomis, K.S., Ablashi, D., et al. (2013). Mapping the telomere integrated genome of human herpesvirus 6A and 6B. *Virology* 442(1)**,** 3-11. doi: 10.1016/j.virol.2013.03.030.

Aswad, A., Aimola, G., Wight, D., Roychoudhury, P., Zimmermann, C., Hill, J., et al. (2020). Evolutionary history of endogenous Human Herpesvirus 6 reflects human migration out of Africa. *Mol Biol Evol*. doi: 10.1093/molbev/msaa190.

Ben-Fredj, N., Ben-Selma, W., Rotola, A., Nefzi, F., Benedetti, S., Frih-Ayed, M., et al. (2013). Prevalence of human herpesvirus U94/REP antibodies and DNA in Tunisian multiple sclerosis patients. *J Neurovirol* 19(1)**,** 42-47. doi: 10.1007/s13365-012-0138-6.

De Bolle, L., Naesens, L., and De Clercq, E. (2005). Update on human herpesvirus 6 biology, clinical features, and therapy. *Clin Microbiol Rev* 18(1)**,** 217-245. doi: 10.1128/CMR.18.1.217-245.2005.

De Coster, W., De Rijk, P., De Roeck, A., De Pooter, T., D'Hert, S., Strazisar, M., et al. (2019). Structural variants identified by Oxford Nanopore PromethION sequencing of the human genome. *Genome Res* 29(7)**,** 1178-1187. doi: 10.1101/gr.244939.118.

Dieudonne, 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(15)**,** 2231-2243. doi: 10.1038/emboj.2009.141.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29(1)**,** 15-21. doi: 10.1093/bioinformatics/bts635.

Endo, A., Watanabe, K., Ohye, T., Suzuki, K., Matsubara, T., Shimizu, N., et al. (2014). Molecular and virological evidence of viral activation from chromosomally integrated human herpesvirus 6A in a patient with X-linked severe combined immunodeficiency. *Clin Infect Dis* 59(4)**,** 545-548. doi: 10.1093/cid/ciu323.

Ernst, J., and Kellis, M. (2012). ChromHMM: automating chromatin-state discovery and characterization. *Nat Methods* 9(3)**,** 215-216. doi: 10.1038/nmeth.1906.

Ernst, J., and Kellis, M. (2015). Large-scale imputation of epigenomic datasets for systematic annotation of diverse human tissues. *Nature Biotechnology* 33(4)**,** 364-376. doi: 10.1038/nbt.3157.

Feuerbach, L., Sieverling, L., Deeg, K.I., Ginsbach, P., Hutter, B., Buchhalter, I., et al. (2019). TelomereHunter - in silico estimation of telomere content and composition from cancer genomes. *BMC Bioinformatics* 20(1)**,** 272. doi: 10.1186/s12859-019-2851-0.

Flamand, L. (2018). Chromosomal Integration by Human Herpesviruses 6A and 6B. *Adv Exp Med Biol* 1045**,** 209-226. doi: 10.1007/978-981-10-7230-7_10.

Geeven, G., Teunissen, H., de Laat, W., and de Wit, E. (2018). peakC: a flexible, non-parametric peak calling package for 4C and Capture-C data. *Nucleic Acids Res* 46(15)**,** e91. doi: 10.1093/nar/gky443.

Gerstein, M.B., Lu, Z.J., Van Nostrand, E.L., Cheng, C., Arshinoff, B.I., Liu, T., et al. (2010). Integrative analysis of the Caenorhabditis elegans genome by the modENCODE project. *Science* 330(6012)**,** 1775-1787. doi: 10.1126/science.1196914.

Gravel, A., Dubuc, I., Wallaschek, N., Gilbert-Girard, S., Collin, V., Hall-Sedlak, R., et al. (2017). Cell Culture Systems To Study Human Herpesvirus 6A/B Chromosomal Integration. *J Virol* 91(14). doi: 10.1128/JVI.00437-17.

Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). circlize implements and enhances circular visualization in R. *Bioinformatics* 30(19)**,** 2811-2812. doi: 10.1093/bioinformatics/btu393.

Gulve, N., Frank, C., Klepsch, M., and Prusty, B.K. (2017). Chromosomal integration of HHV-6A during non-productive viral infection. *Sci Rep* 7(1)**,** 512. doi: 10.1038/s41598-017-00658-y.

Hattori, T., Lai, D., Dementieva, I.S., Montano, S.P., Kurosawa, K., Zheng, Y., et al. (2016). Antigen clasping by two antigen-binding sites of an exceptionally specific antibody for histone methylation. *Proc Natl Acad Sci U S A* 113(8)**,** 2092-2097. doi: 10.1073/pnas.1522691113.

Huang, W., Li, L., Myers, J.R., and Marth, G.T. (2012). ART: a next-generation sequencing read simulator. *Bioinformatics* 28(4)**,** 593-594. doi: 10.1093/bioinformatics/btr708.

Huang, Y., Hidalgo-Bravo, A., Zhang, E., Cotton, V.E., Mendez-Bermudez, A., Wig, G., et al. (2014). Human telomeres that carry an integrated copy of human herpesvirus 6 are often short and unstable, facilitating release of the viral genome from the chromosome. *Nucleic Acids Res* 42(1)**,** 315-327. doi: 10.1093/nar/gkt840.

J, B. (2020). "Boyd J (2020). seqsetvis: Set Based Visualizations for Next-Gen Sequencing Data.". R package version 1.8.0. ed.).

Jin Wook Lee, W.M.a.A.K. *Imputed signal tracks* [Online]. Washington University in St. Louis: wang Lab. Available: https://egg2.wustl.edu/roadmap/web_portal/imputed.html#chr_imp [Accessed 27 September 2020].

Kaufer, B.B. (2013). Detection of Integrated Herpesvirus Genomes by Fluorescence In Situ Hybridization (FISH). *Methods in molecular biology* 1064**,** 141-152. doi: 10.1007/978-1-62703-601-6_10.

Kaufer, B.B., and Flamand, L. (2014). Chromosomally integrated HHV-6: impact on virus, cell and organismal biology. *Curr Opin Virol* 9C**,** 111-118. doi: 10.1016/j.coviro.2014.09.010.

Kaufer, B.B., Jarosinski, K.W., and Osterrieder, N. (2011). Herpesvirus telomeric repeats facilitate genomic integration into host telomeres and mobilization of viral DNA during reactivation. *J Exp Med* 208(3)**,** 605-615. doi: 10.1084/jem.20101402.

Kim, K.D., Tanizawa, H., De Leo, A., Vladimirova, O., Kossenkov, A., Lu, F., et al. (2020). Epigenetic specifications of host chromosome docking sites for latent Epstein-Barr virus. *Nat Commun* 11(1)**,** 877. doi: 10.1038/s41467-019-14152-8.

Kondo, K., Shimada, K., Sashihara, J., Tanaka-Taya, K., and Yamanishi, K. (2002). Identification of human herpesvirus 6 latency-associated transcripts. *J Virol* 76(8)**,** 4145-4151. doi: 10.1128/jvi.76.8.4145-4151.2002.

Krijger, P.H.L., Geeven, G., Bianchi, V., Hilvering, C.R.E., and de Laat, W. (2020). 4C-seq from beginning to end: A detailed protocol for sample preparation and data analysis. *Methods* 170**,** 17-32. doi: 10.1016/j.ymeth.2019.07.014.

Lamb, K.N., Bsteh, D., Dishman, S.N., Moussa, H.F., Fan, H., Stuckey, J.I., et al. (2019). Discovery and Characterization of a Cellular Potent Positive Allosteric Modulator of the Polycomb Repressive Complex 1 Chromodomain, CBX7. *Cell Chem Biol* 26(10)**,** 1365-1379 e1322. doi: 10.1016/j.chembiol.2019.07.013.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9(4)**,** 357-359. doi: 10.1038/nmeth.1923.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10(3)**,** R25. doi: 10.1186/gb-2009-10-3-r25.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997*.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25(16)**,** 2078-2079. doi: 10.1093/bioinformatics/btp352.

Liao, R., Zhang, R., Guan, J., and Zhou, S. (2014). A New Unsupervised Binning Approach for Metagenomic Sequences Based on N-grams and Automatic Feature Weighting. *IEEE/ACM Trans Comput Biol Bioinform* 11(1)**,** 42-54. doi: 10.1109/TCBB.2013.137.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15(12)**,** 550. doi: 10.1186/s13059-014-0550-8.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011* 17(1)**,** 3. doi: 10.14806/ej.17.1.200.

Moquin, S.A., Thomas, S., Whalen, S., Warburton, A., Fernandez, S.G., McBride, A.A., et al. (2018). The Epstein-Barr Virus Episome Maneuvers between Nuclear Chromatin Compartments during Reactivation. *J Virol* 92(3). doi: 10.1128/JVI.01413-17.

Nacheva, E.P., Ward, K.N., Brazma, D., Virgili, A., Howard, J., Leong, H.N., et al. (2008). Human herpesvirus 6 integrates within telomeric regions as evidenced by five different chromosomal sites. *J Med Virol* 80(11)**,** 1952-1958. doi: 10.1002/jmv.21299.

Nicholas, J., and Martin, M.E. (1994). Nucleotide sequence analysis of a 38.5-kilobase-pair region of the genome of human herpesvirus 6 encoding human cytomegalovirus immediate-early gene homologs and transactivating functions. *J Virol* 68(2)**,** 597-610. doi: 10.1128/jvi.68.2.597-610.1994.

Nilsen, T.W. (2014). Preparation of cross-linked cellular extracts with formaldehyde. *Cold Spring Harb Protoc* 2014(9)**,** 1001-1003. doi: 10.1101/pdb.prot080879.

Ohye, T., Inagaki, H., Ihira, M., Higashimoto, Y., Kato, K., Oikawa, J., et al. (2014). Dual roles for the telomeric repeats in chromosomally integrated human herpesvirus-6. *Sci Rep* 4**,** 4559. doi: 10.1038/srep04559.

Okuno, T., Takahashi, K., Balachandra, K., Shiraki, K., Yamanishi, K., Takahashi, M., et al. (1989). Seroepidemiology of human herpesvirus 6 infection in normal children and adults. *J Clin Microbiol* 27(4)**,** 651-653. doi: 10.1128/jcm.27.4.651-653.1989.

Osterrieder, N., Wallaschek, N., and Kaufer, B.B. (2014). Herpesvirus Genome Integration into Telomeric Repeats of Host Cell Chromosomes. *Annual Review of Virology* 1(1)**,** 215-235. doi: 10.1146/annurev-virology-031413-085422.

Prusty, B.K., Gulve, N., Rasa, S., Murovska, M., Hernandez, P.C., and Ablashi, D.V. (2017). Possible chromosomal and germline integration of human herpesvirus 7. *J Gen Virol* 98(2)**,** 266-274. doi: 10.1099/jgv.0.000692.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26(6)**,** 841-842. doi: 10.1093/bioinformatics/btq033.

Saviola, A.J., Zimmermann, C., Mariani, M.P., Signorelli, S.A., Gerrard, D.L., Boyd, J.R., et al. (2019). Chromatin Profiles of Chromosomally Integrated Human Herpesvirus-6A. *Front Microbiol* 10**,** 1408. doi: 10.3389/fmicb.2019.01408.

Szabo, Q., Bantignies, F., and Cavalli, G. (2019). Principles of genome folding into topologically associating domains. *Sci Adv* 5(4)**,** eaaw1668. doi: 10.1126/sciadv.aaw1668.

Tang, H., Kawabata, A., Yoshida, M., Oyaizu, H., Maeki, T., Yamanishi, K., et al. (2010a). Human herpesvirus 6 encoded glycoprotein Q1 gene is essential for virus growth. *Virology* 407(2)**,** 360-367. doi: http://dx.doi.org/10.1016/j.virol.2010.08.018.

Tang, H., Sadaoka, T., and Mori, Y. (2010b). [Human herpesvirus-6 and human herpesvirus-7 (HHV-6, HHV-7)]. *Uirusu* 60(2)**,** 221-235. doi: 10.2222/jsv.60.221.

Tweedy, J., Spyrou, M.A., Pearson, M., Lassner, D., Kuhl, U., and Gompels, U.A. (2016). Complete Genome Sequence of Germline Chromosomally Integrated Human Herpesvirus 6A and Analyses Integration Sites Define a New Human Endogenous Virus with Potential to Reactivate as an Emerging Infection. *Viruses* 8(1). doi: 10.3390/v8010019.

van de Werken, H.J., de Vree, P.J., Splinter, E., Holwerda, S.J., Klous, P., de Wit, E., et al. (2012). 4C technology: protocols and data analysis. *Methods Enzymol* 513**,** 89-112. doi: 10.1016/B978-0-12-391938-0.00004-5.

Wallaschek, N., Sanyal, A., Pirzer, F., Gravel, A., Mori, Y., Flamand, L., et al. (2016). The Telomeric Repeats of Human Herpesvirus 6A (HHV-6A) Are Required for Efficient Virus Integration. *PLoS Pathog* 12(5)**,** e1005666. doi: 10.1371/journal.ppat.1005666.

Wight, D.J., Aimola, G., Aswad, A., Jill Lai, C.Y., Bahamon, C., Hong, K., et al. (2020). Unbiased optical mapping of telomere-integrated endogenous human herpesvirus 6. *Proc Natl Acad Sci U S A* 117(49)**,** 31410-31416. doi: 10.1073/pnas.2011872117.

Zerr, D.M., Meier, A.S., Selke, S.S., Frenkel, L.M., Huang, M.L., Wald, A., et al. (2005). A population-based study of primary human herpesvirus 6 infection. *N Engl J Med* 352(8)**,** 768-776. doi: 10.1056/NEJMoa042207.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., et al. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 9(9)**,** R137. doi: 10.1186/gb-2008-9-9-r137.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or

financial relationships that could be construed as a potential conflict of interest.

Figures and figure legends
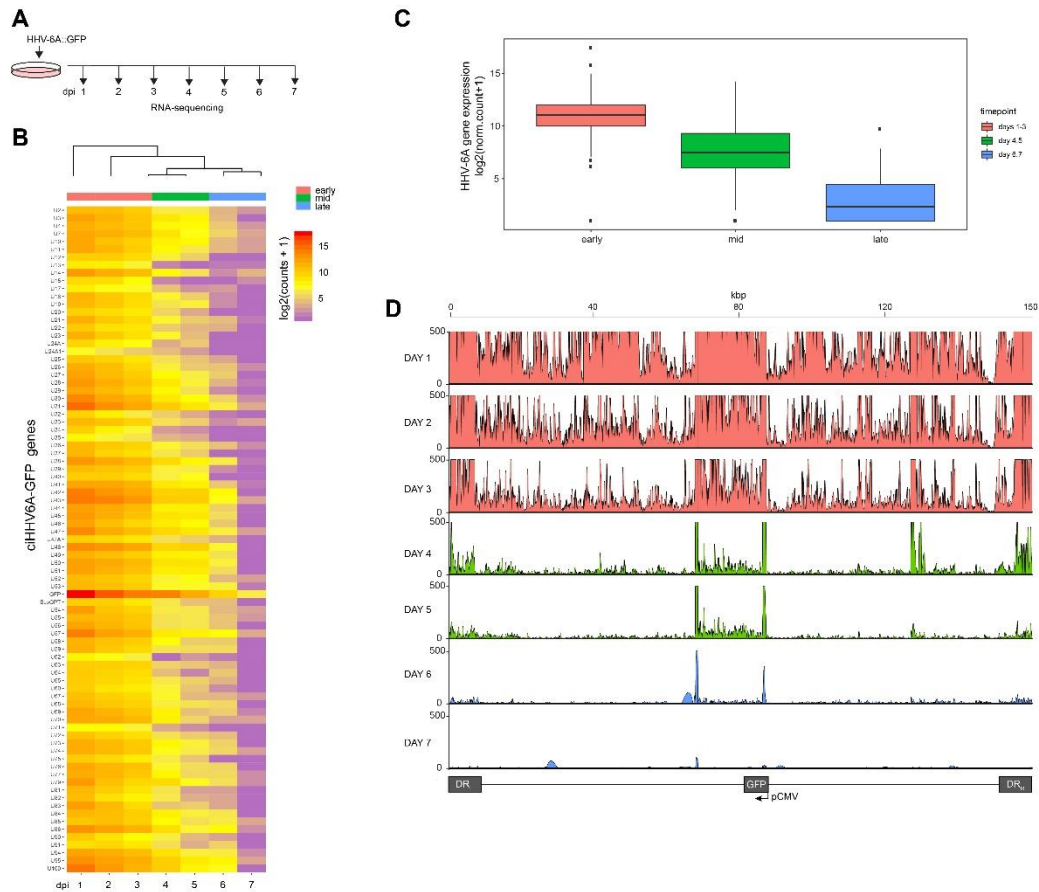
Figure 2.1 RNA-Seq time course reveals the establishment of transcriptionally silent state upon HHV6A infection. A. 293T cells were infected with ciHHV-6A and GFP expressing cells were isolated by FACS. RNA-seq analysis was performed on the cells during the integration phase at the indicated days post infection (dpi). **B.** Cluster analysis of RNA-seq expression of HHV-6A GFP genes, showing expression values of normalized counts with a pseudo-count of one **C.** Boxplots of mean viral gene expression shown for each time point group. **D.** RNA-seq signal plotted across the HHV-6A genome for each day. Viral expression is almost completely absent by day 7 indicating the establishment of

a transcriptionally silent latent infection. At the later timepoints (e.g., days 4-6), two prominent peaks can be seen flanking the U41 and U42 genes, the origin of lytic DNA replication (oriLyt) and the CMV promoter and GFP regions.
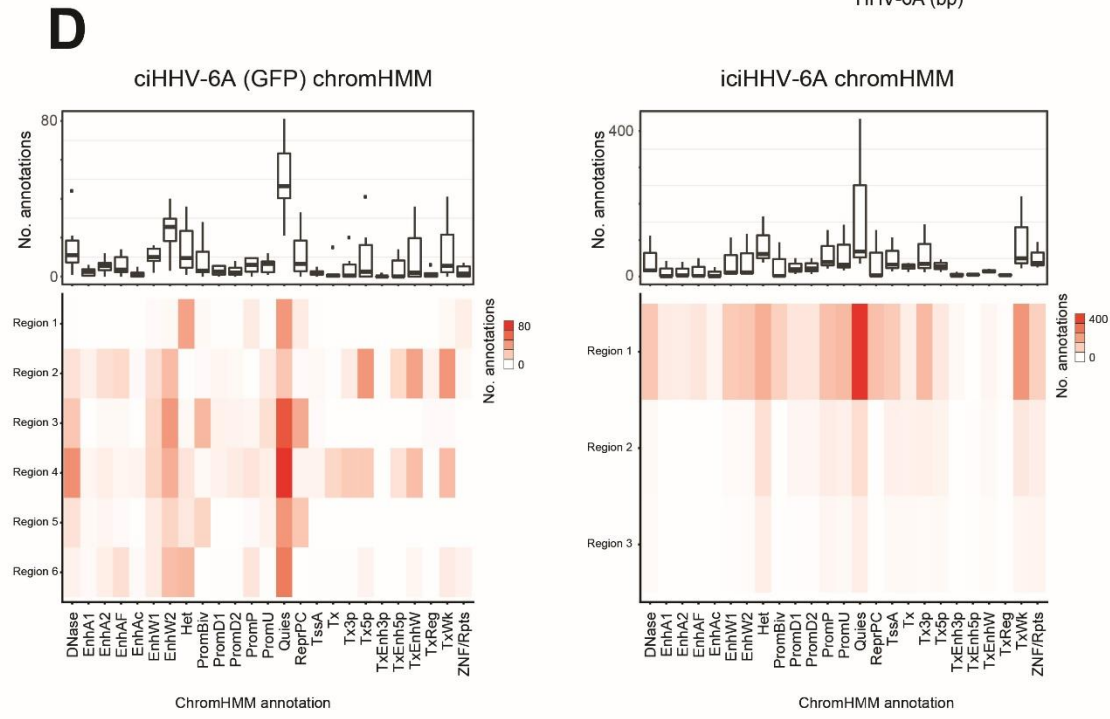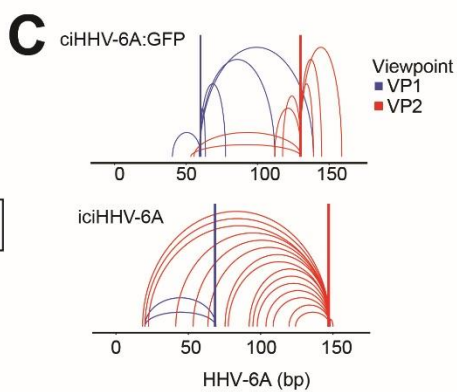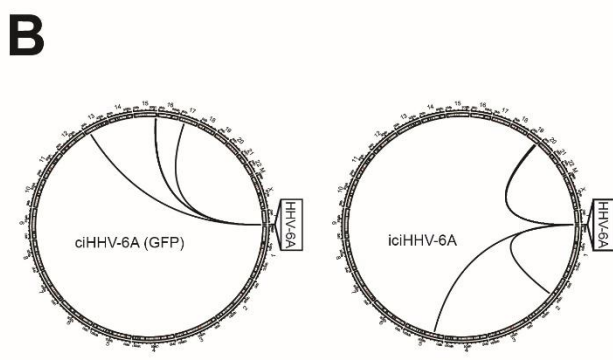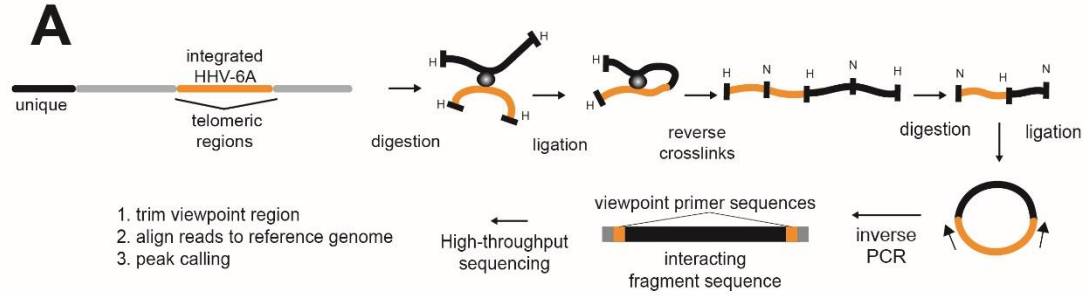
Figure 2.2. Chromatin interactions of chromosomally integrated HHV-6A. Circular Chromosome Conformation Capture (4C-seq) assay was employed using chromosomally integrated HHV-6A cell models to investigate virus-host chromatin interactions. **A.** 4C-seq assay with ciHHV-6A cells. The integrated HHV-6A genome (yellow) is shown with its flanking telomeric regions (gray) and telomere-proximal unique portions of a human chromosome (black). Indicated are the restriction endonucleases sites used in the 4C-seq assay (HindIII and DpnII are indicated as 'H' and 'D', respectively). Inverse PCR was performed with distinct viewpoint primers and analysis to identify interaction regions was performed as shown. **B.** HHV-6A *trans* interactions for the *in vitro* integration model (ciHHV-6A) and the chromosomally inherited iciHHV-6A cell model (iciHHV-6A). Significant interactions are visualized using circos plots - both viewpoints are shown for ciHHV-6A cells (left) and for iciHHV-6A cells(right). **C.** HHV-6A *cis* interactions identified for two separate HHV-6A viewpoints. Inverse primers were designed for two viewpoint regions chosen at ~ 64 kb and 148 kb along the HHV-6A genome. Note that the HHV6A-GFP genome is slightly larger than the canonical HHV-6A genome due to the additional presence of recombinant elements. **D**. Host interaction regions (*trans)* are annotated using the 127 epigenomes 25-state chromatin state model database for ciHHV-6A (left) and iciHHV-6A (right). The host interaction regions contain the indicated chromatin database annotations. Displayed in boxplots (top) and as a heatmap (below).
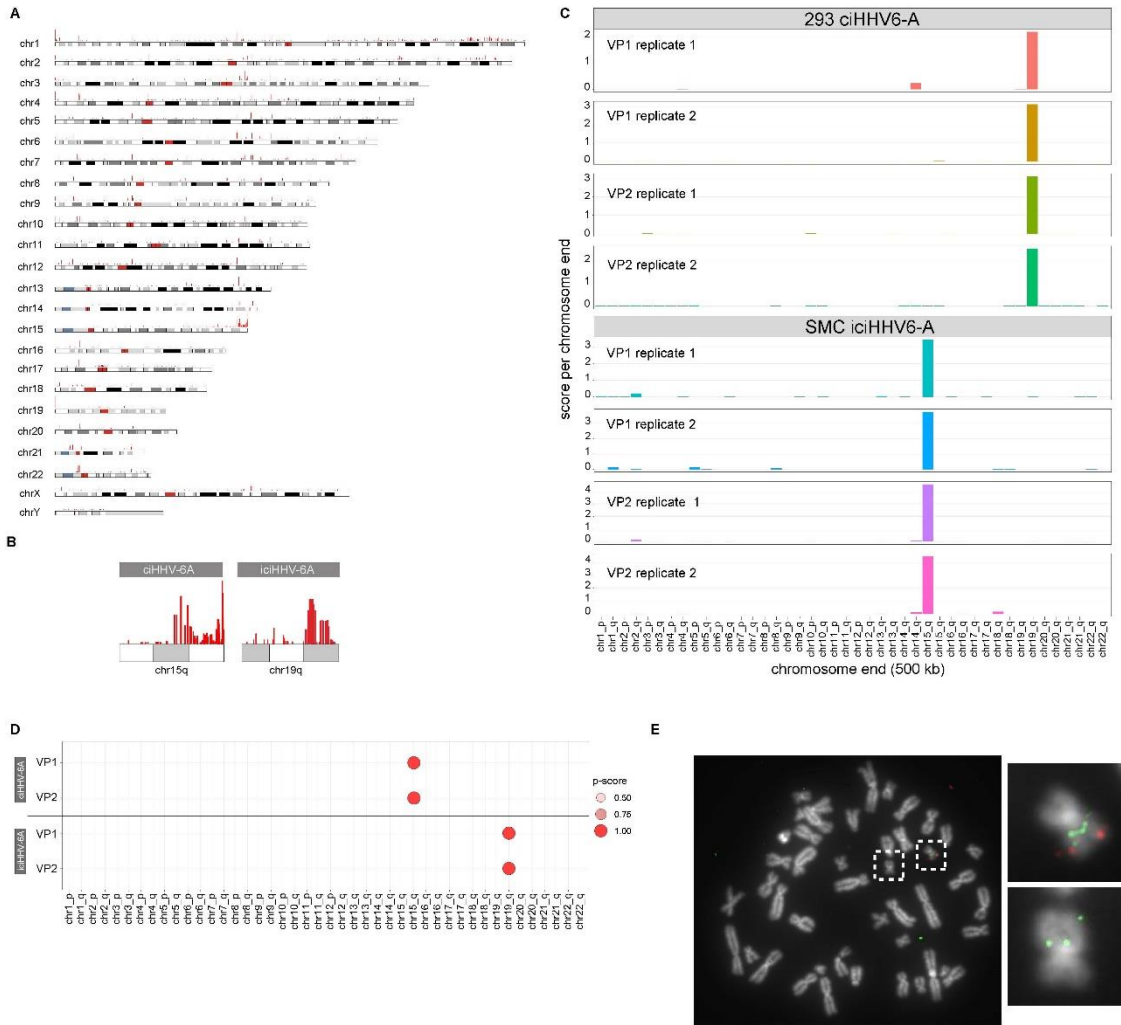
Figure 2.3. Identification of chromosome ends harboring integrated HHV-6A. A.
Genome-wide karyotype plot showing 4C-seq read mapping density for iciHHV-6A,
viewpoint 6a10 replicate 1, across each human chromosome. **B**. Karyotype plots showing
4C-seq read mapping density for chromosomes 15 and 19 for each cell-line. **C.** 4C-seq
signal clustering method to score all the reads within 500 kb from each chromosomal end.
The sum of these scores is represented in the bar plots for each replicate for each viewpoint
in both ciHHV-6A and iciHHV-6A cell types. **D.** statistical measure of the likelihood that
a specific chromosome harbors integrated HHV-6A. For datasets that contain replicates,
an ANOVA was performed using the score from our algorithm across all chromosome ends
followed by pairwise Tukey's HSD post-hoc tests. The "p-score" is the mean -log10(p-
value) value for all pairwise comparisons divided by (normalized by) the maximum mean

73

-log10(p-value). **E.** Integration sites were confirmed via fluorescent in situ hybridization using probes complimentary to chr19q (Gerstein et al.) and the HHV-6A genome (Ben-Fredj et al.) in iciHHV-6A (SMC) cells.
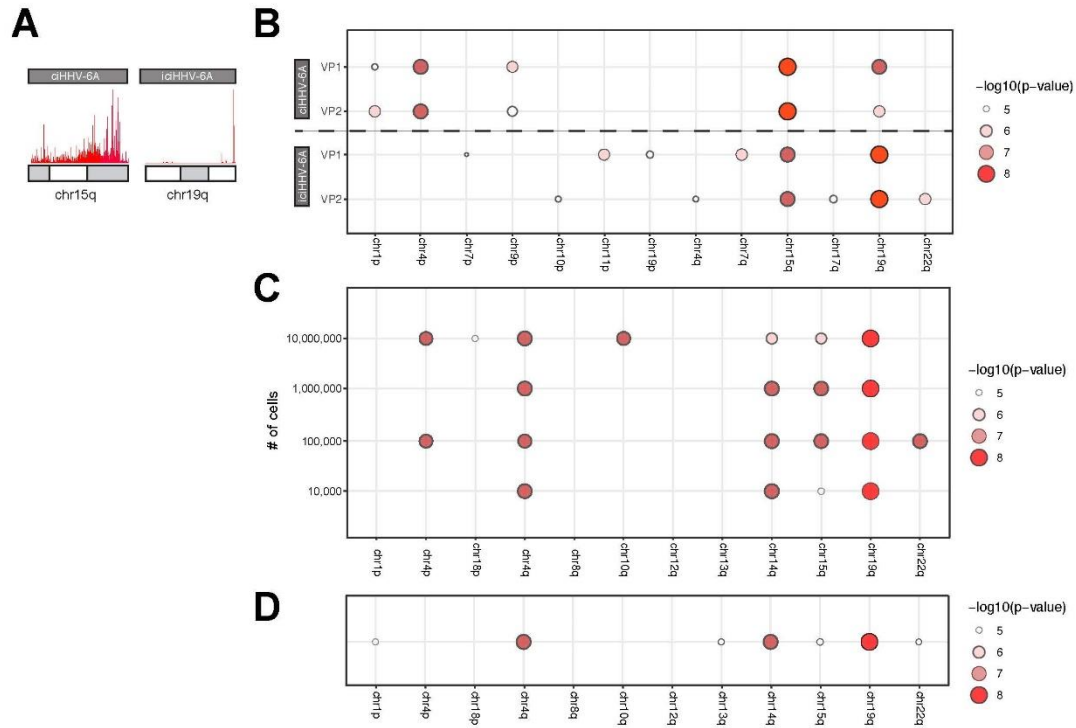
Figure 2.4. MinION-based 4C-seq HHV-6A integration analysis. A. Karyotype plots showing 4C-seq MinION read mapping density at chr15q and chr19q for ciHHV-6A and iciHHV-6A cell models. **B.** Chromosomal end integration site analysis using 4C-seq with MinION reads. As in Figure 3D, a statistical measure of the likelihood that a specific chromosome harbors integrated HHV-6A. Plotted are only the top scoring chromosome ends. **C.** A titration series of cell dilutions was used for 4C-seq MinION analysis and was scored with statistical testing to identify potential HHV-6A integration chromosome ends. **D.** MinION 4C-seq analysis results of a frozen iciHHV-6B+ human lymphocyte sample.

CHAPTER 3

Introduction

Varicella-zoster virus, or VZV, is a ubiquitous Human alphaherpesvirus that primarily causes childhood varicella (chickenpox) and zoster (shingles) upon reactivation in adults. Like all herpesviruses, VZV can establishes a latent infection after resolution of primary infection: VZV does this within host cell neurons of the peripheral nervous system. To further uncover chromatin states and structures underlying the acute VZV infection, as well as the chromatin dynamics hypothesized to mediate the establishment of latency, a thorough characterization of the chromatin landscape within the nucleus of acutely infected cells is required.

Alphaherpesviruses, like VZV, have a relatively short replication cycle that is part of the acute and reactivation stages of the virus life cycle. During this acute replication cycle, virus gene expression is exquisitely choreographed resulting in immune evasion, the reprogramming of host cell functions to express various classes of virus genes, genome replication, packaging of said genome into virions, and the release of these virions form the host cell to proliferate the infection (Kennedy et al., 2015). For herpes simplex virus type 1 (HSV-1) the prototype Human alphaherpesvirus, these events typically are completed within 8 hours and involve coating of the incoming protein free virus DNA with mobile histones that soon acquire posttranslational markings indicative of heterochromatin. While forays into the understanding of the chromatin control of alphaherpesviruses like VZV and HSV-1 have begun to be unraveled over the last decade (Cliffe et al., 2013), most

of the findings have been identified in HSV-1, while the picture remains far less clear for VZV in terms of the epigenetic programs, and chromatin structures and states comprising the chromatin biology that mediates this complex, dynamic process.

To begin to glean insight into the dynamic chromatin biology mentioned above, I primarily focus on employing 4C sequencing and ChIP sequencing to advance our knowledge of this complex and dynamic relationship – particularly with regards to the chromatin organizer and transcriptional insulator CTCF. Chromosome confirmation Capture Circular sequencing or "4C-Seq" was first developed almost a decade ago (van de Werken et al., 2012), yet its potential to unravel the viral interactions within host cells has yet to be realized in the realm of Human alphaherpesviruses. In fact, to my knowledge, 4C-Seq has only been applied to the gammaherpesvirus EBV and only as recently as 2020 (Kim et al., 2020).

4C-seq occupies a specific niche between the worlds of 3C-seq and 5C-seq. In the former, a single primer pair is developed for a bait and target region whereas in the later, a large set of primers are developed and analyzed relative to numerous interacting loci. 4C-Seq however, probes the DNA-DNA interactions between a single bait or viewpoint primer sequence and all the other regions of the genome with which this select sequence interacts; thus, it is a considered a "one-vs-all" analysis, whereas 3C is considered "one-vs-one" and 5C is considers "many-vs-many" (Krijger et al., 2020).

By designing probes along the VZV genome, we can catch a glimpse of the interactions that the virus genome is making with the host cell genome in three-dimensional space. To identify these interactions during acute VZV infections, a cell model derived from human fetal-lung fibroblasts was employed. These cells resemble the *in-vivo*

fibroblasts which are the primary sites of entry for VZV infection; furthermore, these cells have proven to be a tractable model for the study of VZV acute infection *in-vitro*. The infecting VZV virus will establish successful replication in a subset of the cells and will subsequently establish viral replication centers that are essentially factories for producing and packaging new virions. Eventually, the cell will lyse and virions will be released into the extracellular environment, allowing the infection to proliferate. These replication centers take over the host nuclear space, pushing the host cell nuclear material to the periphery of the nucleus. It is thus obvious that the infecting virus is reorganizing chromatin to make room for the establishment of these replication centers within the nucleus.

The virus is likely taking advantage of the initial chromatin reorganization to regulate host and viral gene transcription in a way that allows it to evade or downregulate the host immune system as was discussed in chapter 1. The chromatin states and structures involved with an acute VZV infection, as well as those of the latent and reactivated stages, are not well understood. To begin to understand the complex chromatin biology regulating the latent and reactivated stages of VZV, we must first gain a solid foothold in in our attempts to traverse the biology of the acute infection. With a better understanding of the acute stage VZV chromatin biology, researchers can then move to better understand latency and reactivation – a condition that can lead to serious health conditions such as shingles, severe neuralgia, and stroke (Depledge et al., 2018b).

Given that there is such a pressing need to understand how the latent virus so effectively takes up residence in its host cells and reactivates later in life, causing significant health complications, our study begins to answer vital questions pertaining to

the role that chromatin biology plays in the establishment and maintenance of the acute stages of the VZV life cycle.

To begin my investigations, the establishment of an acute, lytic infection in HFL cells was first achieved and these infected cells were then assayed with 4C sequencing to identify the host-virus interactions that comprise the higher order chromatin structures of the infection, as well as to identify chromatin states associated with the acute infection. I hypothesized that the host-virus interactions would be associated with open chromatin during infection, contributing to a transcriptionally permissive environment for the acutely infecting virus. Furthermore, chromatin immunoprecipitation sequencing (ChIP-Seq) was performed for the ubiquitous transcription factor CTCF. Because CTCF has been identified as a key player in the chromatin biology of other herpesviruses, as discussed in chapter 1, it was hypothesized to be associated with the host-virus interactions, thus contributing to the permissive chromatin state of the acute VZV infection.

Together, in this study I have employed genomics methodologies and cutting-edge bioinformatics analyses to probe the hidden world of chromatin host-virus interactions. The results obtained and discussion entertained herein shed new light on components of the complex, elegant world of molecular virology. In turn, it is hoped that the biology that I have begun to uncover will allow researchers to begin to better characterize the transcriptional landscape regulated by these higher order chromatin structures. Such insights could ultimately lead to the development of novel therapeutics, such as small molecule inhibitors, that could either prevent VZV (or other HHV) infections from successfully establishing latency, or at least prevent the virus from reactivating, reducing the virus' potential to contribute significantly to human disease morbidity.

Varicella zoster virus (VZV) is a ubiquitous human pathogen to which >90% of the world's population has been exposed (Virgin et al., 2009). As an alphaherpesvirus, the VZV life-cycle is characterized by lytic replication where all of the approximately 70 genes are expressed (Cohrs et al., 2003; Nagel et al., 2011) and a latent stage where transcription is limited to at a set of latency associated genes (Nagel et al., 2011; Depledge et al., 2018b). Reactivation from latency, resulting in clinically significant disease, occurs with decline of virus specific CD8+ T cell decline that occurs with aging and immunosuppressive disease or therapy (Gershon et al., 2008). VZV is a highly cell associated virus and time course analysis of virus transcription during one-step growth cannot be performed, but all indications suggest VZV transcription mirrors herpes simplex virus type 1 (HSV-1) in that virus gene expression follows a highly orchestrated process including immediate-early, early, and late gene transcription (Wagner and DeLuca, 2013). However, like HSV-1, this process is altered during virus reactivation where initial virus transcription is deregulated and includes genes of all kinetic classes (Roizman et al., 2011; Kobayashi et al., 2012; Wilson and Mohr, 2012).

Knowledge concerning the initiation of virus gene transcription during the transition between latency and reactivation is critical to understanding the entire picture of VZV reactivation. Promoters for VZV genes that are expressed during early stages of virus reactivation contain histones with euchromatic post-translational marks (H3K9ac) that are absent on histones that coat promoters for untranscribed genes (Gary et al., 2006). Such findings support the overall hypothesis that host chromatin plays a key role in regulating the VZV life cycle.

To begin properly investigating the chromatin biology of the first, or acute stage, of VZV infection, tractable cell models must be employed – i.e., models in which the researcher has control over the stage of infection, and in which the cellular physiology must mimic the *in vivo* physiology as much as possible. HFL cells are employed herein as their physiology is similar to the dermal fibroblasts that are a common site of VZV entry, and acute infection in these cells also mirrors some aspects of a latent *in vivo* infection as well (as mentioned in chapter 1).

Within the nucleus of acutely infected cells, it is known that loops formed by chromatin fibers are further looped into larger regulatory loops creating insulated chromatin "neighborhoods". Furthermore, groups of these regulatory loops can band together to form topologically associated domains or TADs, introduced in chapter 1 (Mishra and Hawkins, 2017). TADs are typically 500KB to 1MB in size and it is currently held that TADs are formed by the looping of long strands of DNA via anchor proteins such as the CTCF–Cohesin complex, which also forms boundaries between TADs (Insulator proteins like CTCF can also form barriers at the sub-TAD level as well). Gene expression is primarily modulated within TADs by sub-TAD looping, but some gene regulation can occur across TAD boundaries – called inter-TAD regulatory loops (at the meta-TAD level) (Mishra and Hawkins, 2017). Groupings of similar TADs at the meta-TAD level form chromosomal compartments commonly referred to as "A" or "B" compartments. A compartments are associated with transcriptionally active euchromatin states, found internally within the nucleus, whereas B compartments are associated with transcriptionally silent or closed chromatin (typically lacking enrichment of significant

histone modifications). B compartments are usually found at the periphery of the nucleus whereas A tend to be found internally (Mishra and Hawkins, 2017).

The related Human alphaherpesvirus, HSV-1, contains CTCF binding sequences that have been shown to play a role in regulating viral gene transcription and are thought to play a role in transitions between HHV life cycle stages, e.g., a transition between an acute and latent stage or between a latent and reactivated stage (Washington et al.). Thus, I believe that it is reasonable to suspect that VZV also interacts with host CTCF. Such interactions could result in phenomena like chromatin looping and/or anchoring that is associated with TAD organization and thence gene expression. To test the hypothesis that VZV higher order chromatin structure associates with CTCF during the acute stage of infection, I first interrogated VZV-infected human fetal lung fibroblast (HFL) cells for the presence of host-virus DNA interactions using 4C-Seq. I then mapped these interaction sites to CTCF binding sites using ChIP-Seq. Taken together, the results from the above experiments indicate that during acute VZV infection in the HFL cell model, VZV DNA forms topologically linked loops with host DNA that are located near host CTCF binding sites. Such looping likely plays a role in modulating host gene expression in a manner that facilitates transitions between, or maintenance of, stages of the VZV life cycle.

Methods

Cell Lines and Virus

81

Clonally expanded human fetal lung fibroblast cells (HFLs) (ATCC, Manassas, VA infected with varicella zoster virus (VZV, Human alphaherpesvirus 3, strain Gilden; GenBank: MH379685.1; https://www.ncbi.nlm.nih.gov/nuccore/MH379685.1) were procured from the laboratory of Randall Cohrs at the University of Colorado Denver - Anschutz Medical Campus and were propagated at the desired multiplicity of infection (MOI) as described previously (Grose et al., 2013; Baird et al., 2014). Briefly, infected HFL cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 1% penicillin streptomycin (Pen/Strep), and 1% L-Glutamine. All cells were maintained in 75 cm2 flasks (Sigma-Aldrich Corp.) as monolayer cultures in a humidified 5% CO_2 air incubator at 37°C. VZV infection proceeded to an MOI equal to 1×10^{-3} cell-free VZV virions.

Confocal Microscopy

VZV-infected HFL cells were prepared in 24-well u-plates (ibidi, Martinsried, Germany) described above. At 3 dpi, cells were fixed with 4% paraformaldehyde for 20 mins and immunostained for the major VZV DNA binding protein ORF29. Briefly, cells were permeabilized with 0.15% Triton-X, blocked in 10% normal donkey serum, incubated in mouse anti-VZV ORF29 overnight at 4°C. The following day, secondary antibodies, Alexa Fluor 488 donkey anti-mouse (1:500, Invitrogen), were applied for 2 hours followed by 2 ug/mL 4', 6-diamidino-2-phenylindole (DAPI; Vector Labs) for 5 minutes. Cells were

visualized by confocal microscopy (using a 3I Marianas inverted spinning disk on Zeiss Axio observer Z1; Oberkochen, Germany) and analyzed using 3I Slidebook 6 software.

Chromatin Immunoprecipitation

The antibody used for ChIP-Seq experiments was polyclonal rabbit anti-Human CTCF (D31H2) antibody (cat: 3418S from Cell Signaling). ChIP sequencing was performed as described previously (O'Geen et al., 2011) on the VZV infected HFL cells. For all immunoprecipitations, 20 ug of chromatin was incubated with 10 ul of antibody overnight at 4°C. The complexes were precipitated with 20 ul of Protein G Magnetic Beads, (Cat: 70024S Cell Signaling), followed by extensive washing, and final elution of the immunoprecipitated chromatin complexes in 50 uL of ChIP elution buffer for subsequent DNA purification and library construction. Libraries were prepared using the ThruPLEX DNA-seq High Performance Library Preparation kit for Illumina NGS Platform following the manufacturer's instructions. Library DNA concentration and quality were assessed on a Qubit 2.0 Fluorometer and an Agilent 2100 Bioanalyzer Instrument, respectively. Pooled libraries were sequenced on an Illumina NovaSEQ 6000 instrument using paired end sequencing at the University of Colorado Denver – Anschutz Medical Campus to obtain 5x10 million 101 bp reads per sample. Sequence reads were aligned to the human (hg38) and VZV (NC_001348.1) genomes using Bowtie2 (Langmead and Salzberg, 2012). The resulting SAM alignment files were used for peak calling using MACS2 against the control (no antibody) inputs (Zhang et al., 2008) and using default parameters.

4C-Seq library preparation and sequencing

VZV infected HFL cells were prepared as described above and 4C-seq libraries were prepared according to methodology as described previously (van de Werken et al., 2012; Krijger et al., 2020), using HindIII/DpnII restriction enzymes. VZV infected HFL cells were cultured as described above and fixed with 1% formaldehyde prior to being snap frozen. For each 4C library, ten million cells were used unless otherwise noted. Inverse PCR was performed in 50 uL reactions using 200ng 3C library template, 25 uL Q5 Hot Start High-Fidelity DNA Polymerase 2X Mastermix (New England Biolabs), 1.5 ul of 10uM forward and reverse inverse PCR primers and water under the following conditions: 30s 98°C, 10s 98°C, 30s 55°C, 2m 72°C, 5m 72°C for a total of 25 cycles. PCR products were cleaned up using the PCR purification kit (Thermofisher Scientific) using the B3 reagent to exclude fragments less than 300 base pairs or with Ampure XP beads (Beckman Coulter) and eluted in water. Library indexing was performed using the DNA HT Dual Index Kit (Illumina) or with custom dual indexing primers with an additional 8 cycles of PCR. All primers sequences are listed in Supplemental Table 4. Libraries were pooled and sequenced on an Illumina HiSeq2500 at the Vermont Integrative Genomics (VIGR) Core Facility at the University of Vermont.

4C-Seq primer design

Our previously designed 4C-Seq primer design tool (https://github.com/FrietzeLabUVM/4c_primer), which was developed in-house to facilitate the creation and generation of the inverse PCR primers necessary for 4C-Seq library generation, was employed here along with Primer-Blast (primer3) (Rozen and Skaletsky, 2000; Ye et al., 2012) to develop the inverse PCR primers. A fasta file containing the reference genome of interest (RefSeq: NC_001348.1), from which the viewpoint primers were chosen, was uploaded into the interactive tool. Then the restriction endonucleases of choice (HindIII and DpnII) were selected. Finally, the minimum fragment sizes following the first and second rounds of digestion were selected using the sliding bars. A minimum size of 500 bp for the first fragment and 300 bp for the second fragment were selected as parameters for our experiments (van de Werken et al., 2012).

4C-Seq analysis

4C-seq *trans* interactions were determined using a "window" method as described by Kim et al. (Kim et al., 2020). The virus viewpoint regions at the beginning of the reads were trimmed off using Cutadapt (Martin) software and the remaining human portions were aligned to the human UCSC hg38 reference genome using bowtie2 (version 2.3.5) (Langmead et al., 2009; Langmead and Salzberg, 2012). P-values for each mapped base were then calculated for each 10kb window across all chromosomes using a Poisson formula. At the end, significant peak regions are filtered using MACS2 (version 2.2.7) bdgpeakcall function (Zhang et al., 2008). Significant interacting *trans* peak regions were

identified using the above window method for each replicate and for all four viewpoint regions. These regions were plotted as circos plots using the Circlize R library (Gu et al., 2014). Bedtools (Quinlan and Hall, 2010) was used to find the intersection of the *trans* regions and the IMR90 25-state imputation-based chromatin state model created with ChromImpute (Ernst and Kellis, 2012; Ernst and Kellis, 2015) and made available from the Wang Lab (Jin Wook Lee). 4C *trans* peaks and signals were also compared to CTCF ChIP-Seq peak results identified using in-house scripts along with the programs Bowtie2 and MACS2. 4C-seq *cis* interactions were determined using the program peakC (Geeven et al., 2018). Finally, publicly available IMR90 HiC compartment call data (Fortin and Hansen, 2015) and IMR90 TAD boundary data (Akdemir and Chin, 2015) were used to identify which regions of the overlapping 4C *trans* peaks and CTCF ChIP-Seq peaks overlapped with which regions of open and closed chromatin, along with TAD boundary regions. HiC plotter (Akdemir and Chin, 2015) was used to visualize the HiC expression matrices and contact maps relative to the 4C-Seq and ChIP-Seq peaks.

Results

Establishment of acute VZV infection and 4C sequencing results

Latent VZV infection was successfully established *in vitro* using HFL cells. Successful acute infection was confirmed via confocal scanning laser microscopy. The acute, lytic VZV infection evidences discrete replication centers within the host cell nucleus which are readily visibly with the confocal scanning laser microscopy upon treatment of samples with

green florescent probes for the ORF-29 VZV major DNA binding protein region (**Fig. 3.1.A-B**). Viewpoint primers were successful generated against the VZV genome (**Fig. 3.1.C.**) and 4C-Seq libraries were constructed. The data returned after sequencing was aligned to the UCSC hg38 reference genome and produced signal across regions of the host cell chromosomes indicating the successful capture of higher order chromatin contact points between the virus and the host cell nucleus (**Fig. 3.1.D).**

Higher order chromatin interactions reveal transcriptionally permissive states

The significant 4C *trans* peaks that were concordant across replicates showed some overlap between viewpoints (**Fig. 3.2.A**), with a total of 334 peaks concordant across all four viewpoints (**Fig. 3.2.B).** When annotated with chromatin states that were determined using databases of publicly available IMR90 data along with the software chromHMM and ChromImpute (Ernst and Kellis, 2015), heterochromatin states were less enriched than others. Hierarchical clustering followed by K-means clustering confirmed the heterochromatin states were the least enriched in the high confidence 4C peaks as can be seen at the far right of cluster 3 - large purple box (**Fig. 3.2.C and 3.2.D**).

Virus-host contact regions overlap with CTCF sites which are in turn localized around TAD boundaries.

Chromosome 12 harbored the highest number of significant 4C peaks (**Fig. 3.3.A. and Fig. 3.3.B).** The HiC contact maps that are shown in the top two panels demonstrate that regions of the host chromosome that are closer to one another are more likely to interact. The 4C and CTCF peaks – center panels – demonstrate that these peaks are found across the genome but tend to cluster with the CTCF peaks, which themselves are more concentrated near the TAD boundaries (the peripheries of the TADs). The right panel shows a zoom-in on chr12:48,000,000-58,00,000 where it is more easily seen that the CTCF peaks concentrate on the periphery of the TAD boundaries (bottom panel). The distribution of 4C peaks is not as limited to these peripheries but there is overlap between the 4C peaks and the CTCF regions as hypothesized.

CTCF and virus-host interactions are significantly associated with one another and are enriched in open, infected chromatin.

Fig. 3.4.A and **Fig. 3.4.B** show the overlap between 4C peaks, CTCF ChIP peaks and open and closed regions of IMR90 chromatin. In both cases there are more CTCF and 4C peaks overlapping with open regions of the genome than with closed regions. But most importantly, there is a significant difference between the number of 4C/CTCF colocalized peaks in the infected data relative to the mock data, and this difference is not present when comparing CTCF and closed region overlaps between mock and infected data or when comparing CTCF open and closed regions in the infected data. **Fig. 3.2.C**. Taken together, these results indicate that CTCF and virus-host 4C peaks are colocalizing to a significantly

greater degree in the infected data vs the mock, and this is not seen when looking at CTCF alone with regard to either infection status or open/closed chromatin status. Thus, we can deduce that the significant CTCF and 4C-Seq host-virus interactions are significantly associated in the open chromatin regions of the infected samples.
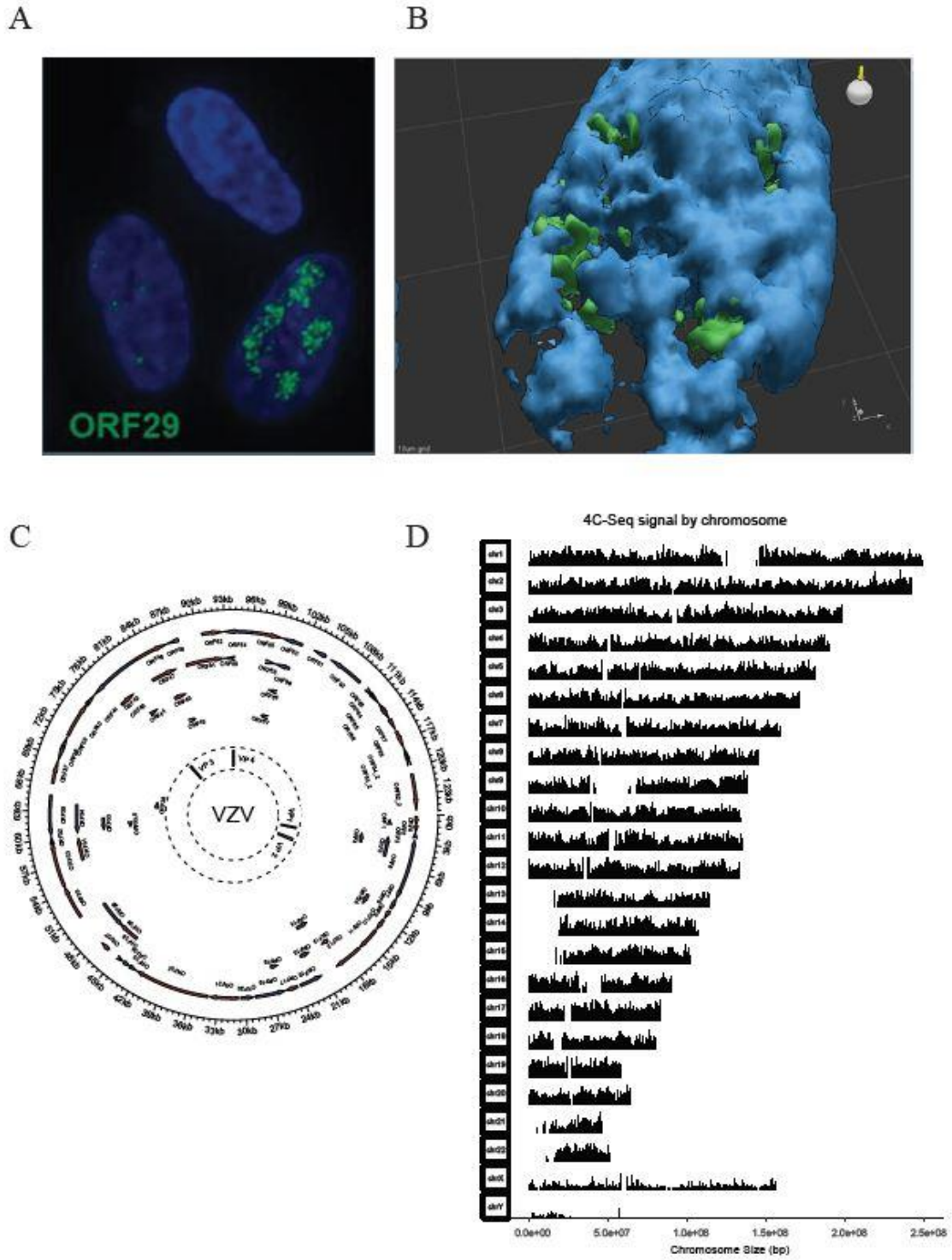
Figure 3.1. HFL cells provide a tractable model to study VZV acute infection *in vitro*.

Fig. 3.1.A Surface image of confocal scanning microscopy imaging of HFL nuclei acutely

infected with VZV. Stained with DAPI (blue) and the major VZV DNA-binding protein ORF29 (Gerstein et al.). **Fig. 3.2.B**. Three-dimensional reconstruction of scanning confocal laser microscopy images of acutely infected HFL cells (nucleus: DAPI, blue; VZV replications centers, green). **Fig 3.2.C**. The VZV genome presented visually as a circos plot, with the locations of the four viral viewpoints relative to the VZV genome. **Fig 3.2.D.** 4C-Seq signal for all replicates and viewpoints across the viral genome faceted by chromosome, demonstrating successful 4C-Seq library preparation and sequencing results.
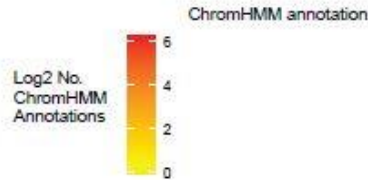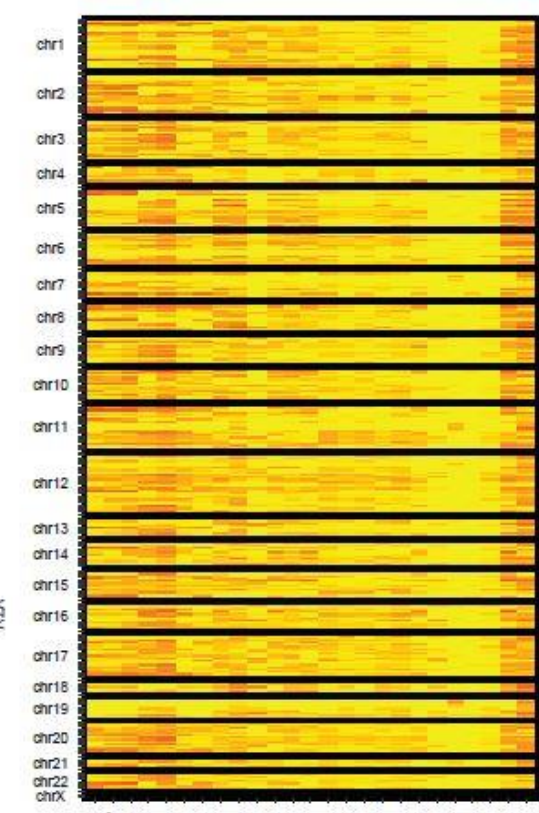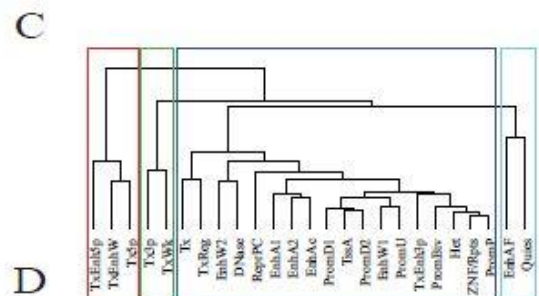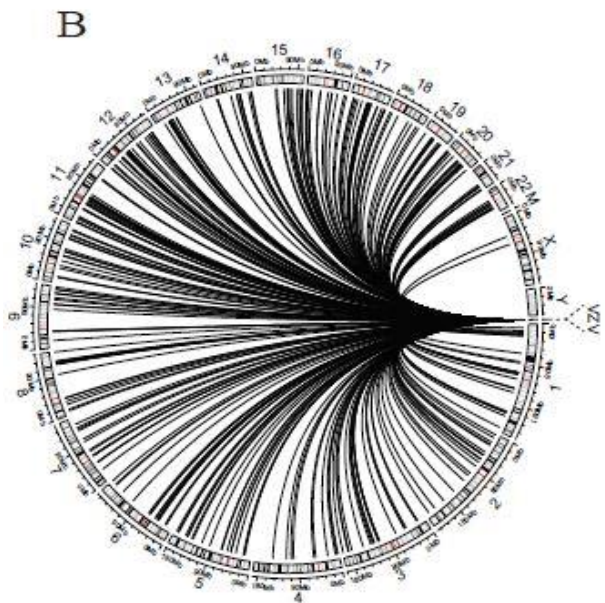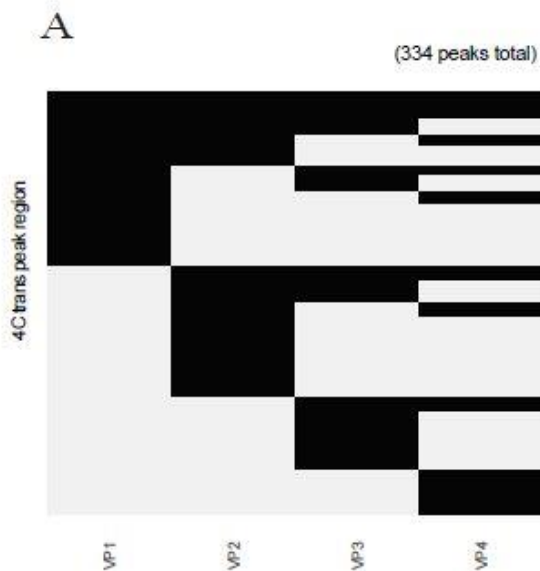
Figure 3.2. High confidence 4C-Seq *trans* interactions and chromatin state annotation. Fig 3.2.A. binary heatmap of the overlapping peaks by viewpoint. The y-axis represents significant peak regions overlapping across replicates. Overlaps of the overlaps can be seen by identifying the regions of the heatmap "black" that overlap between viewpoints. The 334 peaks that overlap across the viewpoints (VP1-VP4) are shown as a circos plot in **Fig. 3.2.B**. The acutely infecting VZV genomes make contact points with multiple chromosomes, with denser regions indicating "hotspots" of interaction activity. **Fig. 3.2.C-D**. The 334 identified most significant 4C peaks annotated with chromatin state information from publicly available IMR90 data. Hierarchical clustering dendrogram (top tree) was applied to the number of annotations for the different chromatin states (x-axis) followed by k-means clustering – colored boxes, identifying heterochromatin states – far right of the center purple box, as the least enriched states associated with the 4C-Seq regions.
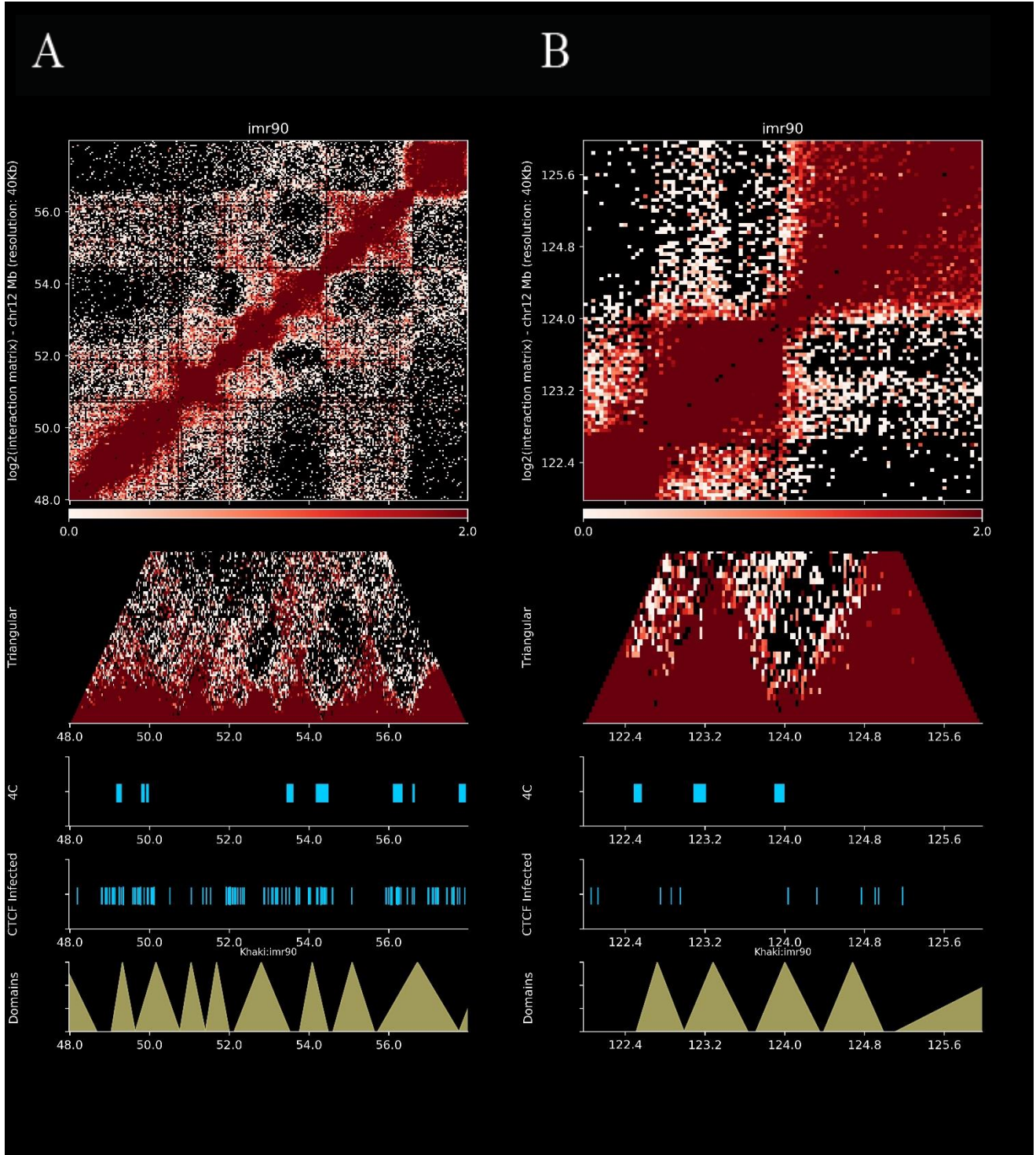
Figure 3.3. Visualization of the higher order chromatin structures of VZV with the host cell genome. Fig. 3.3.A. Interaction matrix of IMR90 data across chromosome 12. Both square and triangular representations demonstrate the organization of the host genome into TADs. CTCF peaks, as expected, are concentrated more to the periphery of TADs and 4C-Seq peaks are overlapping with numerous CTCF ChIP-Seq peaks. **Fig. 3.3.B.** Zoom in on regions chr12:48,000,000-58,000,000. Again, we see CTCF peaks more to the periphery of TAD boundaries with CTCF peaks within the vicinity.

A

Mock CTCF high confidence peaks — IMR90 open chromatin regions

1270
17604
10990
0
25
228
7
4
18298
19
50
7217
1
15
22

4C high confidence peaks — IMR90 closed chromatin regions

B

Infected CTCF high confidence peaks — IMR90. open chromatin regions

608
13512
10150
0
33
187
3
4
19997
11
91
4019
4
13
24

4C high confidence peaks — IMR90 closed chromatin regions

C

Peak number comparisons between mock and infected by venn category

Two Proportion Z-Tests

ns   Padj > 0.05
*   Padj ≤ 0.05
**   Padj ≤ 0.01
***   Padj ≤ 0.001
**** Padj ≤ 0.0001

Condition
Infected
Mock

Figure 3.4. CTCF is associated with host-virus contact points in open chromatin regions. Fig. 3.4.A. There are more overlaps with CTCF and 4C-Seq regions in open chromatin regions in the mock data. **Fig. 3.4.B**. A similar pattern is observed for the overlaps in the infected data. **Fig. 3.4.C**. There is a significant difference between mock and infected peaks (left bar). The infected overlap counts from the Venn diagrams in **Fig. 3.4.B** were then scaled up so that the relative proportion of mock and infected overlaps were equal for the three subsequent comparisons. Most importantly, the CTCF-4C-open chromatin overlapping regions are significantly increased in the infected data relative to CTCF-4C-open chromatin regions in the mock data (middle-left bar); however, there is not a significant difference between CTCF-4C-closed chromatin overlapping regions in mock vs. infected data (middle-right bar). Finally, there is no significant difference between CTCF-open chromatin overlaps in the mock vs. infected data (right bar).

Discussion

Here I employ human fetal lung fibroblast (HFL) cells infected with VZV, along with unbiased epigenomic profiling methods to identify higher order chromatin structures and chromatin states present in the acutely infected samples. I identified both virus-host contacts (*trans)* interactions as well as virus-virus (*cis*) (**Appendix B-1, supplemental figure 3.1**) interactions using a variety of bioinformatic and statistical methods applied to our 4C-sequencing results. *Trans* interactions were found to correspond preponderantly to open chromatin regions within topologically associated domains or TADs. I also performed

97

ChIP-Seq for the transcription factor CTCF – known to play a role in maintenance of alphaherpesvirus life-cycle stage. The acute chromatin states of the host-virus interactions were found to be "permissive" rather than "silencing" (**Fig. 3.2**), and perhaps most importantly, I identified a significant association between virus-host contact regions (4C-Seq regions) and CTCF peak regions in open chromatin in infected data relative to uninfected (mock) data (**Fig. 3.4**). This association suggests that CTCF plays a role in mediating the higher order chromatin structures of the host-virus interactions during an acute infection. This finding hints at the potential presence of a CTCF-mediated host-virus structural mechanism at play in the acute infection. It remains to be shown if such a mechanism, akin to the tethering mechanism in gammaherpesviruses such as EBV, is indeed present at time points during any or all of the alphaherpesvirus (VZV and HSV-1, etc.) life cycle stages.

The results from these experiments, provide an in-depth characterization of the higher order chromatin interactions of VZV with the human host genome that occur during acute VZV infection. While studies with giardia gels, which separate out large molecules, found that HSV-1 was not found to be associated covalently with any chromatin in its latent state, it is known that the ends of the latent virus are fused together covalently as an episome, and the same is the case with VZV. Thus, a CTCF-cohesin tether is possible, but there is no essential requirement that this be the case. Unlike with EBV, where the virus maintains a tether to the host genome in order to successfully replicate and equally partition its genome into daughter cells, neurons don't replicate (generally speaking) so the neurotropic alphaherpesviruses like HSV-1 and VZV do not need to be concerned with the partitioning of their latent genome during host cell replication. So, there appears to be no

essential need for a tether at the acute (or any other) stage of VZV infection in the manner of the gammaherpesvirus infection. But it has also been hard, until recently, to gather enough latent material from cell, animal, and human samples to adequately address this issue during the latent state. The recent arrival of the LUHMES cell model of infection for HSV-1 and the HD10.6 cell model of infection for VZV will allow us to collect enough genomic material to adequately settle the question of CTCF-cohesin tethering in alphaherpesviruses once and for all. In addition, it is possible that some degree of tethering happens in acute alphaherpesvirus infection, simply due to the nature of the host-virus interactions, but there is no end achieved through this tethering as the cell does not live very long before succumbing to infection.

Regardless of whether viral tethering during acute (or latent/reactivation) VZV infection exists, and has a distinct biological purpose or not, I have demonstrated that we can identify the sites of host-virus interactions during acute infection and that there does appear to be evidence indicating that CTCF is associated with these higher order chromatin structures. This important finding goes against the current status quo in the alphaherpesvirus field. The methodology developed here should be applied to the newer, and more comprehensive, HD10.6 cell model of VZV infection in future studies. With the proper techniques and cell models in place, and with the results herein hinting at the presence of a CTCF-related structural mechanism involved in acute VZV infection, we can soon establish whether tethering exists at the various stages of the alphaherpesvirus life cycle. Such findings will provide much deeper insight into the chromatin structure and states, particularly with regards to CTCF, that contribute to regulation of the VZV life cycle.

CHAPTER 4

Discussion

Chromatin is an essential part of the herpesvirus life cycle. Host-virus chromatin interactions, as well as the recruitment of chromatin modifying enzymes and the epigenetic and genetic pathways involved, work in a concerted manner to establish, maintain, and regulate the transition of the herpesvirus life cycle (Kennedy et al., 2015). Furthermore, a variety of cell and animal models are being investigated in which to study the host and cellular factors regulating the herpesvirus life cycle across different tissues and over time. As was discussed in chapter 1 and reemphasized via the results in the main body of this dissertation, there is growing evidence of to support the role of chromatin factors playing a role in regulating the herpesvirus life cycle across various tissue types, physiological conditions and across specific viruses (Yalcin, 2012).

While there are sub-families of Human herpesviruses, they all share similarities, and harbor their own unique characteristics with which they have coevolved with their hosts over hundreds of millions of years (Aswad et al., 2020). For example, the unique portions of each species' genome contains both conserved genes and gene cluster unique to each viral species (Wang et al., 2020). In addition, the different HHVs contain varying types and numbers of repetitive regions of their respective genomes. In the case of HHV-6 (A and B) the repetitive regions at the end of the genomes are known to be necessary for integration and the establishment of latency; whereas, with other regions, ORFs are

contained therein such as with VZV ORF4 (Wang et al., 2020), and for some repetitive regions, the function remains unknown.

In this dissertation I focused on the role that chromatin biology plays in two separate HHVs at two separate life cycle stages in relevant cell models. In chapter 2 I presented my work on latent HHV-6A infection before proceeding to discuss my experimental findings with regards to acute VZV infection in chapter 3. Both viruses are relatively less well characterized, in terms of their chromatin biology, than the gammaherpesvirus EBV and the alphaherpesvirus HSV-1, about which we have the most comprehensive understanding of the role of chromatin structure and function in establishing and maintaining the various HHV life cycle stages.

I successfully investigated the chromatin biology of two HHV-6 cell models of infection in chapter 2: the *in vitro*-derived HEK29T cell lines infected with recombinant HHV-6A (ciHHV-6A), as well as a cell line derived from the umbilical arterial smooth muscle cells of patients with inherited chromosomally integrated HHV-6A infection (iciHHV-6A). These two cell models, although somewhat limited, remain the best cell models that we have for studying the HHV-6A life cycle. Cells that are more physiologically relevant to the HHV-6A latent cell tropism such as J-Jhan cells, have not been capable of sustaining a latent infection *in-vitro*. Primary cells - such as lymphocytes isolated from iciHHV-6A positive patients, are of course useful in that they contain the physiologically relevant cellular environment found in an actual infection, but they do not present the researcher with the ability to regulate the transition between the different HHV-6A life cycles stages.

The ciHHV-6A cells establish latency over a time course as shown using RNA-Seq in chapter 2. Our lab had previously demonstrated that the genome would also reactivate with the administration of the phorbol ester TPA, although only GFP was transcribed (Saviola et al.). This is the main disadvantage of the ciHHV-6A cell model. The model was initially demonstrated as capable of full reactivation of the virus genome by Arbuckle et al. (Arbuckle et al., 2010), but attempts to replicate these findings have not been successful by other researchers. The reactivation potential is somewhat unstable and may itself be subject to degradation over time with increasing cell passage number or the accumulation of mutations that render the chromatin states that we seek to investigate unstable and further removed from the *in vivo* scenario. This should highlight to the reader the difficulty in developing tractable *in vitro* models for studying the HHV life cycle. With that caveat in mind, the ciHHV-6A cell model nonetheless represents a notable achievement towards the development of tractable cell models to study HHV-6 infections.

To assay the chromatin landscape surrounding the latent, integrated HHV-6A virus I chose to employ 4C-Seq because the small size of the viral genome allows a comprehensive picture of three-dimensional host-virus interactions to be painted with a relatively small number of viewpoint regions. However, even though software was available for generic 4C analysis, i.e., analyses utilizing human viewpoint regions, I learned after trial and error that the current 4C software, for the most part, is not suitable to assaying virus host interactions. This is because most software expects the viewpoint region to be on a human chromosome. Thus current methods had to be modified to take into account the presence of an exogenous chromosome, such as an infecting HHV genome.

Eventually, I found that a recently published method developed in Paul Lieberman's lab by Kim et al. (Kim et al., 2020) allowed for the most flexibility for assaying 4C interactions within the constraints of our experiments and offered the best resolution of identified significant 4C peaks. However, the statistical and computational methodology was only briefly discussed in their publication, and was not published as a stand-alone software, and neither was, to my knowledge, the computer code. Thus, I had to elucidate how the p-value equation that they outline in their paper was detailing how to calculate a probability for each nucleotide in every chromosome in each 4C sample. I believe this scenario emphasizes the importance of technological proficiency as virology enters the genomics era. Methods have always driven science in my opinion and computational genomics is rapidly becoming a powerful tool for biological discovery and such methodology is in most cases not simply "push button". The genomics researcher will need to have more than a superficial knowledge of biology, computer science, and statistics simultaneously, in order to make the most out of these powerful new techniques.

I showed in chapter 2 that both the Illumina and Nanopore platforms produced similar results, and with the increasing accuracy of long read Nanopore technology, coupled with its inexpensive costs, I believe this will become an increasingly popular sequencing alternative in the future relative to the more traditional sequencing systems. But the more established systems such as Illumina still provide greater base pair resolution (Wang et al., 2020) and read depth at the time of writing. In the case of VZV acute infection, where we have many copies of the viral genome in a single infected nucleus, the added resolution of the traditional Illumina system is still preferable.

I will again emphasize that is was because of the increasingly powerful techniques available to the genomics researcher that I was able to develop and apply my novel 4C-Seq framework to interrogate the higher order chromatin structure of HHV-6A. This work builds off the previous findings to which I contributed where we showed via ChIP-Seq experiments, and bioinformatics techniques, that the repressive histone marks H3K9me3 and H3K27me3 were broadly enriched across the latent HHV-6A genome (Saviola et al., 2019a). I was able to build on this foundational work and create an innovative and reproducible method for utilizing 4C sequencing data to statistically rank the chromosome ends in terms of their probability for harboring an integrated HHV-6A genome.

As discussed in chapter 2, the integrated chromosome technically represents a "*cis-trans*" event because we are looking for virus-host interactions (*trans*) on the chromosome of which the virus is now a part via integration (*cis*). Because *cis* interaction signal tends to drop off monotonically across approximately 1 MBp, and because the sub-telomeric regions are approximately 500kb in size (2014 genome biology reference), 500 KBp was chosen as the size with which to compare the chromosome ends. I developed an algorithm to score the reads that were closest in proximity to each other, and these scores could then be evaluated with the aforementioned statistical ranking system.

In fact, the parameters chosen for the above computational components themselves could be refined in computational experiments potentially leading to further improvements in the detection capabilities of my methodology. I mention this fact because I believe that it is important to remember that computational experiments, such as parameter sweeps, are in themselves scientific experiments in which a hypothesis is generated, tested, and evaluated within a framework of rational empiricism and falsifiability.

With the parameters that I currently use, it was seen in chapter 2 that across viewpoints and replicates, chromosome 15q was identified in ciHHV-6A cells as the prime candidate for integration by far. While chrom15q has not yet been widely reported as a site of integration for the HHV-6A virus, studies on a population level have only just begun (Aswad et al., 2020). Another strong possibility is that the end of chromosome 15q underwent a translocation event, either pre- or post-viral integration, leading to a recombinant chromosome harboring HHV-6A integrated at the end of one arm of its chromosome. Interestingly, there is recent evidence in the literature for such translocation events in 293 cells identified via molecular cytogenetics techniques (Binz et al., 2019). A future study utilizing a three-probe fish: one for virus, one for original chromosome, and one for secondary chromosome, could be used to confirm whether a translocation event has occurred in the chromosome harboring the integrated HHV-6A virus.

Turning our attention to the iciHHV-6A cell samples, it was clear both from my 4C-Seq workflow (from library building through computational genomics analyses and FISH confirmation) that chromosome 19q harbored the integrated virus. Interestingly, chromosome 19q has been reported as a site of integration across numerous collated studies confirmed via FISH (Aswad et al., 2020). While there does not yet appear to be a clear pattern as to a specific chromosome preference in which the virus will integrate, the application of novel computational methods to high-throughput data, coupled with improving bioinformatics methods for data collation and mining, will allow researchers to more easily answer questions such as "is there a chromosomal preference for integration across HHV-6A strains and across human and cell populations?"

In chapter 2 we saw that the integrated virus made few interactions with other chromosomes. The few "*trans*" interactions that it did make were found to be associated with "quiescent" and "heterochromatin" annotations along with some weaker evidence for some enhancer and transcription factor activity. Because quiescent regions were most associated and because the latent virus is itself transcriptionally quiescent, it makes sense that HHV-6A has evolved its integration site preference based not only on its ability to undergo homologous recombination with the canonical mammalian repeat regions (TTAGGG)n of the subtelomeric region, but also because this is an overall transcriptionally quiet region of the host genome. Thus, the virus is disrupting the host transcriptional machinery weakly if at all, disruptions which could signal intracellular stress responses that might jeopardize its survival.

Additionally, the heterochromatin annotation accords well with the previous studies showing chromatinization of the HHV-6A genome with repressive histone marks (Saviola et al., 2019a). Here, further evidence is provided that HHV-6A is heavily chromatinized during its latent state, contributing to the evidence that chromatinization, or the deposition of heterochromatin along the virus genome, is rendering the virus transcriptionally quiescent during latency. A finding that corroborates the role of chromatinization patterns in HHV life cycle regulation that has been discovered in other HHVs (as discussed in chapter 1). One major outstanding question is how this chromatinization is established in the first place.

Future studies should focus on reproducing the *trans* interactions that were identified and striving to improve the resolution of said interactions. While it appears likely that the virus does not want to disrupt the host cell transcriptional machinery too much,

which again may give away its presence to the cell, confirmation or identification of additional *trans* regions can be compared and contrasted with studies from acute infections to identify regulatory regions, specific genes and/or pathways that are common between the two life cycle stages, or that at least point researchers towards possible mechanistic bridges between the two stages. Thus, an excellent next step would be to consider chromatin conformation assays such as 4C-Seq across a time course spanning acute to latent stages of infection: ~1 to 7 days. Such a study could be combined with RNA-seq data in a "systems biology" approach to identify associations between higher order chromatin structures and gene regulation. Gene products or regulatory regions implicated in the transition could then be probed via knockdown or targeted deletion to confirm whether these factors play an essentially role to the life cycle transition.

Overall, the process of uncovering the mechanisms governing the establishment and maintenance of both acute and latent HHV-6A infections, as well as the mechanisms that control reactivation of the latent virus is an ongoing and challenging endeavor. Employing both cell models of primary infections and patient samples, I have developed methodology that utilizes unbiased epigenomic profiling methods to identify the site of HHV-6A integration within the host genome and to assess the higher order chromatin structures of which it is part. I have then used this information along with data from other epigenomic profiling methods, such as MNase and ChIP-seq to more fully characterize the chromatin state of the latent HHV-6A virus in multiple cell models of infection as well as in primary patient cells. Such methodology provides a high-throughput means to uncovering the higher-order host-virus chromatin biology hypothesized to regulate the virus life cycle and contributes to the groundwork for the application of these technologies

to the realm of clinical diagnostics, e.g., for screening transplant recipients for the presence of HHV-6 integration.

To summarize, in chapter two I: 1.) demonstrated the validity of the best current "controlled" cell model for studying the chromatin biology of HHV-6, 2.) established a novel high-throughput methodology for assaying higher order chromatin structure in cells infected with HHVs, and 3.) Contributed to the current understanding of the overall chromatin landscape (structure and state) of latent, integrated HHV-6A in these cells, as well as in patient-derived iciHHV-6A cells, and in lymphocyte cells harboring integrated HHV-6A harvested directly from patients.

With the methodology worked out for HHV-6A, I was then able to apply the same techniques to the study of the higher order chromatin structures and chromatin states of acute VZV infection in HFL cells. This information could then be applied along with CTCF ChIP sequencing results and computational methods in a systems biology approach to gain insight into the role that CTCF is playing in acute VZV infection.

In terms of cell models, as discussed in chapter 1, some past initial success was obtained with the use of human fetal fibroblast lung (HFL) cells, which are similar physiologically to the dermal fibroblasts that are a primary infection site for VZV. These cells are also able to recapitulate some elements of *in vivo* neural latency transcription patterns, but the establishment of true latency, and reactivation therefrom, is not possible in this cell model. Again, many cells are capable of infection with VZV *in-vivo*, *ex-vivo*, and *in-vitro*, but the creation of a viable cell model in which the researcher can control the life cycle stage of the infection has only been realized recently in the HD10.6 cell model of infection (Thellman et al., 2017). This cell model derived from dorsal root ganglion cells

is to date the best model for studying the HHV life cycle and the pertinent chromatin biology *in vivo*. Thus, future studies utilizing the epigenomics and systems biology approaches employed in this dissertation will likely lead to even greater breakthroughs in our understanding of the role the host chromatin plays at all stages of the VZV life cycle, in a more physiologically relevant environment.

It is known for example that during acute infection HSV-1 phosphatidylinositol 3-kinase (PI3K) signaling proceeds through protein kinase B or AKT. One such pathway involved mTOR complex 1 (mTORC01 - mammalian target of rapamycin), which mediates the hyper phosphorylation of 4E-binding protein (4E-BP) which can mediate cap-dependent translation. NGF in turn can maintain latency through activation of mTORC-1 and its removal can lead to reactivation (Camarena et al., 2010; Kobayashi et al., 2012; Kennedy et al., 2015). In the new HD10.6 cells, NGF depletion leads to reactivation as is the case above for HSV-1 (Thellman et al., 2017).

For my work with HFL cells, efforts were focused on using 4C sequencing to begin uncovering the host-virus chromatin interactions during acute infection. 4C sequencing has been applied to numerous viruses including EBV and HIV, and the "higher C" studies are becoming more and more common in general. Whereas 3C-Seq allows the researcher to identify interactions that appear between a viewpoint or bait region and a single target region (one vs. one), and 5C allows numerous bait regions to be assay relative to the entire genome (all vs. all), 4C provides a flexible middle ground in terms of both comprehensiveness and specificity. We can choose a particular viral region of interest – the viewpoint(s) – and design primers against this region. Upon inverse PCR we can create

amplicons that can then be sequenced via high throughput sequencing platforms and from the sequencing data we can identify the host-virus *trans* interactions.

When the *trans* peaks that were identified as concordant across replicates and viewpoints, were then annotated with chromatin state models developed using chromHMM (Ernst and Kellis, 2012) along with the 127 genomes from the roadmap epigenomics project (Wang et al., 2020), various chromatin states were found to be associated with these contact regions. The annotations were of permissive chromatin states associated with the open chromatin that was hypothesized and confirmed to harbor the virus-host interactions. The heterochromatin state was among the least enriched – the opposite of what was seen for the transcriptionally repressed, latent HHV-6A infection. Not surprisingly, there was also some enhancer and transcription factor chromatin state annotation with the VZV acute infection, providing further evidence that the virus is coordinating with open regions of the host genome involved in transcriptional activation.

Furthermore, I also identified particular regions of CTCF ChIP-Seq peaks that were concordant across replicates, which were in turn concordant with the VZV 4C peaks shared across replicates. Thus, I was able to identify a number of high-confidence infected 4C and CTCF ChIP peaks that colocalized along the host genome. This is an important finding for a few reasons. First, as discussed in the literature review in the introduction, CTCF has been identified as a key component of the CTCF-cohesin complex coordinating three-dimensional host virus chromatin architecture in Human gammaherpesviruses. Second, CTCF has been identified as a key insulator protein in both Human gamma herpesviruses and alphaherpesvirus such as HSV-1, where it has been shown to spatially divide the virus genome into transcriptional competent and quiescent regions. Finally,

CTCF is known to coordinate chromatin boundaries and the establishment of topologically associated domains via looping (Wang et al., 2020). Thus, the CTCF and 4C interactions indicate that CTCF is also playing a role in the chromatin biology of acute VZV infection.

It is thus becoming apparent that not only is CTCF a critical transcription factor and chromatin organizer in uninfected cells and plays a key role in the more distantly related gammaherpesviruses and the more closely related HSV-1, but that it is also an important chromatin factor in the VZV life cycle, possibly working to organize chromatin during acute infection in order to regulate VZV transcription and subsequently the viral life cycle.

Future studies could also look at differential CTCF occupancy between infected and uninfected samples as well as Differential RNA sequencing between mock and infected samples. The subsequent genes and pathways identified could reveal several genes and signaling pathways that are differentially regulated between conditions and such signaling pathways are thought to work in concert with epigenetic alterations to facilitate the establishment of infection and ultimately latency (Kennedy et al., 2015).

Another genomics approach to unravelling the role of host chromatin in the acute infection is to look at the differential occupancy between mock and infected cells and compare this to both the 4C-Seq results from the infected data as well as the differential RNA-Seq results. As mentioned earlier, CTCF has many roles in the uninfected and infected host cell biology, and alterations in the position or frequency of CTCF occupation between mock and infected samples can indicate how CTCF is dynamically changing across conditions, potentially contributing to virus life cycle regulation.

The identification of specific genes and promoters that are in the vicinity of differentially occupied CTCF sites could provide us with likely candidates for genes involved in the infection response and these genes are potentially coopted by the virus to aid in the regulation of its life cycle. The top differentially occupied CTCF regions could be investigated for overlap with the promoters of differentially expressed genes from an RNA-Seq study. Significant genes identified in more than one experiment (including genes with promoters proximal to identified 4C and/or CTCF sites) could be run through pathway analysis algorithms to confirm inflammation response pathways due to infection as well as to identify possible signaling pathways relevant to chromatin organization and function.

The above concordance between various chromatin-based assays and differential transcription assays would shed light on the role that chromatin dynamics play in fundamentally altering the host nuclear transcriptional and regulatory landscape. Genes and pathways identified could be corroborated in literature and the identification of novel markers serve as excellent seed points for the generation of subsequent hypotheses pertaining to the interplay of chromatin biology and the host transcriptional regulatory pathways involved in the HHV life cycle. As such, future studies will need to proceed further towards molecular confirmation of the key gene markers identified by such studies, to provide assurance that the pathways identified and corroborated are indeed significant to the biology under consideration. Such assurance will allow us to confidently decipher which markers involved in the pathways are most relevant to the chromatin component of the underlying acute stage HHV biology.

To summarize the VZV component of my work, I have obtained novel findings pertaining to the higher order chromatin structures and chromatin states involved in the

acute VZV life cycle stage. Such information contributes to building a secure foundation for our understanding of VZV chromatin biology. Researchers can build on this foundation by continuing to refine the cell models and methodology discussed herein and then carrying out novel experiments aimed at building a complete picture of the role that chromatin plays in the VZV life cycle.

The immediate and exciting next steps would be to transfer our experimental and virological knowledge from the less physiologically relevant, but hitherto more tractable, HFL cell model, to the new DRG-derived HD10.6 cell model of VZV infection. We also have autopsied samples of trigeminal ganglia from patient donors to interrogate as well. We can apply the same high-throughput methodology that I have developed to answer the same and similar hypothesis that I discussed in chapter 3 relating to chromatin and the VZV life cycle. With the proper methods and cell models in place, and insights from the results that we have obtained thus far, we are certain to make exciting and far-reaching discoveries into the mechanisms regulating the VZV life cycle in the human host.

Through the course of this work, I have successfully developed and employed high throughput methodology and computational genomics techniques to gain further understanding of the higher order chromatin structure and chromatin states in two viable cell models for studying HHV-6A and VZV: two herpesviruses about which much of the chromatin biology remains a mystery.

Thus, both VZV and HHV-6A are two HHVs about which much headway needed to be made in terms of uncovering the role that chromatin biology plays in regulating their respective life cycles. Due to the rapidly accelerating advances in high-throughput genomics techniques, hardware, and chemistry in recent decades, coupled with

concomitant advances in both software availability and computational resources, I was in an excellent position to adapt these methodologies to the study of chromatin biology in these two herpesviruses about which little was known.

It should again be emphasized that new methods developed herein will contribute to the foundation for the application of ever improving and evolving high-throughput sequencing technologies to uncover the complex interplay of nuclear replicating viruses within their host cells. Such methods will almost assuredly lead to faster, more precise methods for uncovering chromatin biology during HHV infection, as well as lend themselves to improved clinical methodology.

During the Course of my work, I was able to implement and further characterize existing cell models, both *in vitro*-derived and patient-derived, as model systems in which to develop and apply my innovative computational genomics methods in order to assay the higher order chromatin structures between of HHV infection, as well as add to the growing body of literature pertaining to the role that host chromatin structure and states play in coordinating and maintaining the HHV life cycle stages.

Furthermore, methodological developments such as the ones that I present herein, can lead to the development of new fast, accurate, and cost-effective methodology for identifying individuals that have iciHHV-6A infection. This would be an important milestone in allowing clinicians to more effectively pre-screen and triage transplant recipients that may harbor HHV-6A infection effectively prior to undergoing surgery.

It is my hope that other herpes virologists will be able to incorporate my findings into their own work as well as draw upon and improve the methodology that I have developed. We all seek to build upon the understanding of chromatin biology to which I

have contributed that will ultimately lead to a complete picture of the precise role(s) that host chromatin plays in regulating the HHV life cycle. It is my hope that soon the elusive cellular mechanisms that govern the acute, latent, and reactivation stages of HHV infection will finally be uncovered in cell models that adequately capture the *in vivo* human physiology.

Only through a thorough understanding of the chromatin biology of HHV infection will we gain a complete picture of the mechanisms regulating the life cycle of these tenacious and troublesome viruses. Thus, with a thorough understanding of host cell chromatin biology we will be able to piece together a complete model to explain the mechanisms and structures involved in the establishment and maintenance of, as well as the transition between, the various stages of the HHV life cycle. In future studies, we will seek to further elucidate the higher-order chromatin structures and chromatin states of HHV infection and how these chromatin factors work with other host and viral factors to precisely regulate said life cycle stages.

With this knowledge, scientists and physicians can work to identify novel therapeutic targets for which treatments can be developed to effectively treat HHV infection at various the various life cycle stages. Ideally, efficient and precise ways of eradicating the latent virus from its *in vivo* latent reservoirs could be identified as well. As mentioned at the beginning of this work, the word "herpes" relates to this pathogen's ability to stealthily creep along throughout the human body well after resolution of the initial infection. Through persistent hard work, professional interdisciplinary cooperation, intellectual rigor, and the thrill of scientific endeavor, I believe that humanity will be able to eradicate these insidious viral infections for good.

COMPREHENSIVE BIBLIOGRAPHY

Adams, M.J., and Carstens, E.B. (2012). Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses (2012). *Arch Virol* 157(7)**,** 1411-1422. doi: 10.1007/s00705-012-1299-6.

Agut, H., Bonnafous, P., and Gautheret-Dejean, A. (2017). Update on infections with human herpesviruses 6A, 6B, and 7. *Med Mal Infect* 47(2)**,** 83-91. doi: 10.1016/j.medmal.2016.09.004.

Aimola, G., Beythien, G., Aswad, A., and Kaufer, B.B. (2020). Current understanding of human herpesvirus 6 (HHV-6) chromosomal integration. *Antiviral Res* 176**,** 104720. doi: 10.1016/j.antiviral.2020.104720.

Akdemir, K.C., and Chin, L. (2015). HiCPlotter integrates genomic data with interaction matrices. *Genome Biol* 16**,** 198. doi: 10.1186/s13059-015-0767-1.

Alba, M.M., Das, R., Orengo, C.A., and Kellam, P. (2001). Genomewide function conservation and phylogeny in the Herpesviridae. *Genome Res* 11(1)**,** 43-54. doi: 10.1101/gr.149801.

Ambagala, A.P., Bosma, T., Ali, M.A., Poustovoitov, M., Chen, J.J., Gershon, M.D., et al. (2009). Varicella-zoster virus immediate-early 63 protein interacts with human antisilencing function 1 protein and alters its ability to bind histones h3.1 and h3.3. *J Virol* 83(1)**,** 200-209. doi: 10.1128/JVI.00645-08.

Amelio, A.L., McAnany, P.K., and Bloom, D.C. (2006). A chromatin insulator-like element in the herpes simplex virus type 1 latency-associated transcript region binds CCCTC-binding factor and displays enhancer-blocking and silencing activities. *J Virol* 80(5)**,** 2358-2368. doi: 10.1128/JVI.80.5.2358-2368.2006.

Arbuckle, J.H., Medveczky, M.M., Luka, J., Hadley, S.H., Luegmayr, A., Ablashi, D., et al. (2010). The latent human herpesvirus-6A genome specifically integrates in telomeres of human chromosomes in vivo and in vitro. *Proc Natl Acad Sci U S A* 107(12)**,** 5563-5568. doi: 10.1073/pnas.0913586107.

Arbuckle, J.H., Pantry, S.N., Medveczky, M.M., Prichett, J., Loomis, K.S., Ablashi, D., et al. (2013). Mapping the telomere integrated genome of human herpesvirus 6A and 6B. *Virology* 442(1)**,** 3-11. doi: 10.1016/j.virol.2013.03.030.

Arope, S., Harraghy, N., Pjanic, M., and Mermod, N. (2013). Molecular characterization of a human matrix attachment region epigenetic regulator. *PLoS One* 8(11)**,** e79262. doi: 10.1371/journal.pone.0079262.

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(2)**,** 233-245. doi: 10.1016/j.chom.2012.06.008.

Arvin, A.M. (1996). Varicella-zoster virus. *Clin Microbiol Rev* 9(3)**,** 361-381. doi: 10.1128/CMR.9.3.361-381.1996.

Aswad, A., Aimola, G., Wight, D., Roychoudhury, P., Zimmermann, C., Hill, J., et al. (2020). Evolutionary history of endogenous Human Herpesvirus 6 reflects human migration out of Africa. *Mol Biol Evol*. doi: 10.1093/molbev/msaa190.

Baird, N.L., Bowlin, J.L., Cohrs, R.J., Gilden, D., and Jones, K.L. (2014). Comparison of varicella-zoster virus RNA sequences in human neurons and fibroblasts. *J Virol* 88(10)**,** 5877-5880. doi: 10.1128/JVI.00476-14.

Balazs, Z., Tombacz, D., Szucs, A., Csabai, Z., Megyeri, K., Petrov, A.N., et al. (2017). Long-Read Sequencing of Human Cytomegalovirus Transcriptome Reveals RNA Isoforms Carrying Distinct Coding Potentials. *Sci Rep* 7(1)**,** 15989. doi: 10.1038/s41598-017-16262-z.

Balazs, Z., Tombacz, D., Szucs, A., Snyder, M., and Boldogkoi, Z. (2018). Dual Platform Long-Read RNA-Sequencing Dataset of the Human Cytomegalovirus Lytic Transcriptome. *Front Genet* 9**,** 432. doi: 10.3389/fgene.2018.00432.

Baringer, J.R., and Swoveland, P. (1973). Recovery of herpes-simplex virus from human trigeminal ganglions. *N Engl J Med* 288(13)**,** 648-650. doi: 10.1056/NEJM197303292881303.

Bastian, F.O., Rabson, A.S., Yee, C.L., and Tralka, T.S. (1972). Herpesvirus hominis: isolation from human trigeminal ganglion. *Science* 178(4058)**,** 306-307. doi: 10.1126/science.178.4058.306.

Beagan, J.A., Pastuzyn, E.D., Fernandez, L.R., Guo, M.H., Feng, K., Titus, K.R., et al. (2020). Three-dimensional genome restructuring across timescales of activity-induced neuronal gene expression. *Nat Neurosci* 23(6)**,** 707-717. doi: 10.1038/s41593-020-0634-6.

Beagan, J.A., and Phillips-Cremins, J.E. (2020). On the existence and functionality of topologically associating domains. *Nat Genet* 52(1)**,** 8-16. doi: 10.1038/s41588-019-0561-1.

Ben-Fredj, N., Ben-Selma, W., Rotola, A., Nefzi, F., Benedetti, S., Frih-Ayed, M., et al. (2013). Prevalence of human herpesvirus U94/REP antibodies and DNA in Tunisian multiple sclerosis patients. *J Neurovirol* 19(1)**,** 42-47. doi: 10.1007/s13365-012-0138-6.

Berger, S.L. (2002). Histone modifications in transcriptional regulation. *Curr Opin Genet Dev* 12(2)**,** 142-148. doi: 10.1016/s0959-437x(02)00279-4.

Binshabaib, M., SS, A.L., Salehpoor, D., Michelogiannakis, D., and Javed, F. (2018). Contribution of herpesviruses in the progression of periodontal and peri-implant diseases in systemically healthy individuals. *Rev Med Virol* 28(5)**,** e1996. doi: 10.1002/rmv.1996.

Binz, R.L., Tian, E., Sadhukhan, R., Zhou, D., Hauer-Jensen, M., and Pathak, R. (2019). Identification of novel breakpoints for locus- and region-specific translocations in 293 cells by molecular cytogenetics before and after irradiation. *Sci Rep* 9(1)**,** 10554. doi: 10.1038/s41598-019-47002-0.

Bloom, D.C., Giordani, N.V., and Kwiatkowski, D.L. (2010). Epigenetic regulation of latent HSV-1 gene expression. *Biochim Biophys Acta* 1799(3-4)**,** 246-256. doi: 10.1016/j.bbagrm.2009.12.001.

Boutell, C., and Everett, R.D. (2013). Regulation of alphaherpesvirus infections by the ICP0 family of proteins. *J Gen Virol* 94(Pt 3)**,** 465-481. doi: 10.1099/vir.0.048900-0.

Britt, W. (2008). Manifestations of human cytomegalovirus infection: proposed mechanisms of acute and chronic disease. *Curr Top Microbiol Immunol* 325**,** 417-470. doi: 10.1007/978-3-540-77349-8_23.

Burd, E.M., and Carrigan, D.R. (1993). Human herpesvirus 6 (HHV-6)-associated dysfunction of blood monocytes. *Virus Res* 29(1)**,** 79-90. doi: 10.1016/0168-1702(93)90127-9.

Cabral, J.M., Oh, H.S., and Knipe, D.M. (2018). ATRX promotes maintenance of herpes simplex virus heterochromatin during chromatin stress. *Elife* 7. doi: 10.7554/eLife.40228.

Camarena, V., Kobayashi, M., Kim, J.Y., Roehm, P., Perez, R., Gardner, J., et al. (2010). Nature and duration of growth factor signaling through receptor tyrosine kinases regulates HSV-1 latency in neurons. *Cell Host Microbe* 8(4)**,** 320-330. doi: 10.1016/j.chom.2010.09.007.

Caruso, A., Rotola, A., Comar, M., Favilli, F., Galvan, M., Tosetti, M., et al. (2002). HHV-6 infects human aortic and heart microvascular endothelial cells, increasing their ability to secrete proinflammatory chemokines. *J Med Virol* 67(4)**,** 528-533. doi: 10.1002/jmv.10133.

Chan, P.K., Ng, H.K., Hui, M., and Cheng, A.F. (2001). Prevalence and distribution of human herpesvirus 6 variants A and B in adult human brain. *J Med Virol* 64(1)**,** 42-46. doi: 10.1002/jmv.1015.

Chau, C.M., Zhang, X.Y., McMahon, S.B., and Lieberman, P.M. (2006). Regulation of Epstein-Barr virus latency type by the chromatin boundary factor CTCF. *J Virol* 80(12)**,** 5723-5732. doi: 10.1128/JVI.00025-06.

Chaumeil, J., and Skok, J.A. (2012). The role of CTCF in regulating V(D)J recombination. *Curr Opin Immunol* 24(2)**,** 153-159. doi: 10.1016/j.coi.2012.01.003.

Chen, C.P., Lyu, Y., Chuang, F., Nakano, K., Izumiya, C., Jin, D., et al. (2017). Kaposi's Sarcoma-Associated Herpesvirus Hijacks RNA Polymerase II To Create a Viral Transcriptional Factory. *J Virol* 91(11). doi: 10.1128/JVI.02491-16.

Chen, Q., Lin, L., Smith, S., Huang, J., Berger, S.L., and Zhou, J. (2007). CTCF-dependent chromatin boundary element between the latency-associated transcript and ICP0 promoters in the herpes simplex virus type 1 genome. *J Virol* 81(10)**,** 5192-5201. doi: 10.1128/JVI.02447-06.

Chernukhin, I., Shamsuddin, S., Kang, S.Y., Bergstrom, R., Kwon, Y.W., Yu, W., et al. (2007). CTCF interacts with and recruits the largest subunit of RNA polymerase II to CTCF target sites genome-wide. *Mol Cell Biol* 27(5)**,** 1631-1648. doi: 10.1128/MCB.01993-06.

Cliffe, A.R., Coen, D.M., and Knipe, D.M. (2013). Kinetics of facultative heterochromatin and polycomb group protein association with the herpes simplex viral genome during establishment of latent infection. *mBio* 4(1). doi: 10.1128/mBio.00590-12.

Cliffe, A.R., Garber, D.A., and Knipe, D.M. (2009). Transcription of the herpes simplex virus latency-associated transcript promotes the formation of facultative

heterochromatin on lytic promoters. *J Virol* 83(16)**,** 8182-8190. doi: 10.1128/JVI.00712-09.

Cliffe, A.R., and Knipe, D.M. (2008). Herpes simplex virus ICP0 promotes both histone removal and acetylation on viral DNA during lytic infection. *J Virol* 82(24)**,** 12030-12038. doi: 10.1128/JVI.01575-08.

Cohrs, R.J., Hurley, M.P., and Gilden, D.H. (2003). Array analysis of viral gene transcription during lytic infection of cells in tissue culture with Varicella-Zoster virus. *J Virol* 77(21)**,** 11718-11732. doi: 10.1128/jvi.77.21.11718-11732.2003.

Coleman, H.M., Connor, V., Cheng, Z.S.C., Grey, F., Preston, C.M., and Efstathiou, S. (2008). Histone modifications associated with herpes simplex virus type 1 genomes during quiescence and following ICP0-mediated de-repression. *J Gen Virol* 89(Pt 1)**,** 68-77. doi: 10.1099/vir.0.83272-0.

Collin, V., and Flamand, L. (2017). HHV-6A/B Integration and the Pathogenesis Associated with the Reactivation of Chromosomally Integrated HHV-6A/B. *Viruses* 9(7). doi: 10.3390/v9070160.

Conn, K.L., Hendzel, M.J., and Schang, L.M. (2008). Linker histones are mobilized during infection with herpes simplex virus type 1. *J Virol* 82(17)**,** 8629-8646. doi: 10.1128/JVI.00616-08.

Conn, K.L., Hendzel, M.J., and Schang, L.M. (2011). Core histones H2B and H4 are mobilized during infection with herpes simplex virus 1. *J Virol* 85(24)**,** 13234-13252. doi: 10.1128/JVI.06038-11.

Conn, K.L., Hendzel, M.J., and Schang, L.M. (2013). The differential mobilization of histones H3.1 and H3.3 by herpes simplex virus 1 relates histone dynamics to the assembly of viral chromatin. *PLoS Pathog* 9(10)**,** e1003695. doi: 10.1371/journal.ppat.1003695.

Conn, K.L., and Schang, L.M. (2013). Chromatin dynamics during lytic infection with herpes simplex virus 1. *Viruses* 5(7)**,** 1758-1786. doi: 10.3390/v5071758.

Cosgrove, M.S., Boeke, J.D., and Wolberger, C. (2004). Regulated nucleosome mobility and the histone code. *Nat Struct Mol Biol* 11(11)**,** 1037-1043. doi: 10.1038/nsmb851.

De Bolle, L., Naesens, L., and De Clercq, E. (2005). Update on human herpesvirus 6 biology, clinical features, and therapy. *Clin Microbiol Rev* 18(1)**,** 217-245. doi: 10.1128/CMR.18.1.217-245.2005.

De Coster, W., De Rijk, P., De Roeck, A., De Pooter, T., D'Hert, S., Strazisar, M., et al. (2019). Structural variants identified by Oxford Nanopore PromethION sequencing of the human genome. *Genome Res* 29(7)**,** 1178-1187. doi: 10.1101/gr.244939.118.

Dembowski, J.A., and DeLuca, N.A. (2015). Selective recruitment of nuclear factors to productively replicating herpes simplex virus genomes. *PLoS Pathog* 11(5)**,** e1004939. doi: 10.1371/journal.ppat.1004939.

Depledge, D.P., Ouwendijk, W.J.D., Sadaoka, T., Braspenning, S.E., Mori, Y., Cohrs, R.J., et al. (2018a). A spliced latency-associated VZV transcript maps antisense to the viral transactivator gene 61. *Nat Commun* 9(1)**,** 1167. doi: 10.1038/s41467-018-03569-2.

Depledge, D.P., Sadaoka, T., and Ouwendijk, W.J.D. (2018b). Molecular Aspects of Varicella-Zoster Virus Latency. *Viruses* 10(7). doi: 10.3390/v10070349.

Depledge, D.P., Srinivas, K.P., Sadaoka, T., Bready, D., Mori, Y., Placantonakis, D.G., et al. (2019). Direct RNA sequencing on nanopore arrays redefines the transcriptional complexity of a viral pathogen. *Nat Commun* 10(1), 754. doi: 10.1038/s41467-019-08734-9.

Deshmane, S.L., and Fraser, N.W. (1989). During latency, herpes simplex virus type 1 DNA is associated with nucleosomes in a chromatin structure. *J Virol* 63(2), 943-947. doi: 10.1128/JVI.63.2.943-947.1989.

Di Valentin, E., Bontems, S., Habran, L., Jolois, O., Markine-Goriaynoff, N., Vanderplasschen, A., et al. (2005). Varicella-zoster virus IE63 protein represses the basal transcription machinery by disorganizing the pre-initiation complex. *Biol Chem* 386(3), 255-267. doi: 10.1515/BC.2005.031.

Dieudonne, 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(15), 2231-2243. doi: 10.1038/emboj.2009.141.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29(1), 15-21. doi: 10.1093/bioinformatics/bts635.

Dominguez, G., Dambaugh, T.R., Stamey, F.R., Dewhurst, S., Inoue, N., and Pellett, P.E. (1999). Human herpesvirus 6B genome sequence: coding content and comparison with human herpesvirus 6A. *J Virol* 73(10), 8040-8052. doi: 10.1128/JVI.73.10.8040-8052.1999.

Donati, D., Akhyani, N., Fogdell-Hahn, A., Cermelli, C., Cassiani-Ingoni, R., Vortmeyer, A., et al. (2003). Detection of human herpesvirus-6 in mesial temporal lobe epilepsy surgical brain resections. *Neurology* 61(10), 1405-1411. doi: 10.1212/01.wnl.0000094357.10782.f9.

Du, T., Zhou, G., and Roizman, B. (2012). Induction of apoptosis accelerates reactivation of latent HSV-1 in ganglionic organ cultures and replication in cell cultures. *Proc Natl Acad Sci U S A* 109(36), 14616-14621. doi: 10.1073/pnas.1212661109.

Dunmire, S.K., Hogquist, K.A., and Balfour, H.H. (2015). Infectious Mononucleosis. *Curr Top Microbiol Immunol* 390(Pt 1), 211-240. doi: 10.1007/978-3-319-22822-8_9.

Dupin, N. (2020). Update on oncogenesis and therapy for Kaposi sarcoma. *Curr Opin Oncol* 32(2), 122-128. doi: 10.1097/CCO.0000000000000601.

Endo, A., Watanabe, K., Ohye, T., Suzuki, K., Matsubara, T., Shimizu, N., et al. (2014). Molecular and virological evidence of viral activation from chromosomally integrated human herpesvirus 6A in a patient with X-linked severe combined immunodeficiency. *Clin Infect Dis* 59(4), 545-548. doi: 10.1093/cid/ciu323.

Ernst, J., and Kellis, M. (2012). ChromHMM: automating chromatin-state discovery and characterization. *Nat Methods* 9(3), 215-216. doi: 10.1038/nmeth.1906.

Ernst, J., and Kellis, M. (2015). Large-scale imputation of epigenomic datasets for systematic annotation of diverse human tissues. *Nature Biotechnology* 33(4), 364-376. doi: 10.1038/nbt.3157.

Ertel, M.K., Cammarata, A.L., Hron, R.J., and Neumann, D.M. (2012). CTCF occupation of the herpes simplex virus 1 genome is disrupted at early times postreactivation in a transcription-dependent manner. *J Virol* 86(23), 12741-12759. doi: 10.1128/JVI.01655-12.

Eshleman, E., Shahzad, A., and Cohrs, R.J. (2011). Varicella zoster virus latency. *Future Virol* 6(3)**,** 341-355. doi: 10.2217/fvl.10.90.

Everett, R.D. (2010). Depletion of CoREST does not improve the replication of ICP0 null mutant herpes simplex virus type 1. *J Virol* 84(7)**,** 3695-3698. doi: 10.1128/JVI.00021-10.

Everett, R.D., Rechter, S., Papior, P., Tavalai, N., Stamminger, T., and Orr, A. (2006). PML contributes to a cellular mechanism of repression of herpes simplex virus type 1 infection that is inactivated by ICP0. *J Virol* 80(16)**,** 7995-8005. doi: 10.1128/JVI.00734-06.

Ferenczy, M.W., and DeLuca, N.A. (2009). Epigenetic modulation of gene expression from quiescent herpes simplex virus genomes. *J Virol* 83(17)**,** 8514-8524. doi: 10.1128/JVI.00785-09.

Ferenczy, M.W., and DeLuca, N.A. (2011). Reversal of heterochromatic silencing of quiescent herpes simplex virus type 1 by ICP0. *J Virol* 85(7)**,** 3424-3435. doi: 10.1128/JVI.02263-10.

Feuerbach, L., Sieverling, L., Deeg, K.I., Ginsbach, P., Hutter, B., Buchhalter, I., et al. (2019). TelomereHunter - in silico estimation of telomere content and composition from cancer genomes. *BMC Bioinformatics* 20(1)**,** 272. doi: 10.1186/s12859-019-2851-0.

Filippova, G.N., Fagerlie, S., Klenova, E.M., Myers, C., Dehner, Y., Goodwin, G., et al. (1996). An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. *Mol Cell Biol* 16(6)**,** 2802-2813. doi: 10.1128/mcb.16.6.2802.

Flamand, L. (2018). Chromosomal Integration by Human Herpesviruses 6A and 6B. *Adv Exp Med Biol* 1045**,** 209-226. doi: 10.1007/978-981-10-7230-7_10.

Fortin, J.P., and Hansen, K.D. (2015). Reconstructing A/B compartments as revealed by Hi-C using long-range correlations in epigenetic data. *Genome Biol* 16**,** 180. doi: 10.1186/s13059-015-0741-y.

Fox, J.D., Briggs, M., Ward, P.A., and Tedder, R.S. (1990). Human herpesvirus 6 in salivary glands. *Lancet* 336(8715)**,** 590-593. doi: 10.1016/0140-6736(90)93392-3.

Gary, L., Gilden, D.H., and Cohrs, R.J. (2006). Epigenetic regulation of varicella-zoster virus open reading frames 62 and 63 in latently infected human trigeminal ganglia. *J Virol* 80(10)**,** 4921-4926. doi: 10.1128/JVI.80.10.4921-4926.2006.

Gaszner, M., and Felsenfeld, G. (2006). Insulators: exploiting transcriptional and epigenetic mechanisms. *Nat Rev Genet* 7(9)**,** 703-713. doi: 10.1038/nrg1925.

Gatherer, D., Seirafian, S., Cunningham, C., Holton, M., Dargan, D.J., Baluchova, K., et al. (2011). High-resolution human cytomegalovirus transcriptome. *Proc Natl Acad Sci U S A* 108(49)**,** 19755-19760. doi: 10.1073/pnas.1115861108.

Geeven, G., Teunissen, H., de Laat, W., and de Wit, E. (2018). peakC: a flexible, non-parametric peak calling package for 4C and Capture-C data. *Nucleic Acids Res* 46(15)**,** e91. doi: 10.1093/nar/gky443.

Gershon, A.A., Chen, J., and Gershon, M.D. (2008). A model of lytic, latent, and reactivating varicella-zoster virus infections in isolated enteric neurons. *J Infect Dis* 197 Suppl 2**,** S61-65. doi: 10.1086/522149.

Gerstein, M.B., Lu, Z.J., Van Nostrand, E.L., Cheng, C., Arshinoff, B.I., Liu, T., et al. (2010). Integrative analysis of the Caenorhabditis elegans genome by the modENCODE project. *Science* 330(6012)**,** 1775-1787. doi: 10.1126/science.1196914.

Gewurz, B.E., Marty, F.M., Baden, L.R., and Katz, J.T. (2008). Human herpesvirus 6 encephalitis. *Curr Infect Dis Rep* 10(4)**,** 292-299. doi: 10.1007/s11908-008-0048-1.

Gibeault, R.L., Conn, K.L., Bildersheim, M.D., and Schang, L.M. (2016). An Essential Viral Transcription Activator Modulates Chromatin Dynamics. *PLoS Pathog* 12(8)**,** e1005842. doi: 10.1371/journal.ppat.1005842.

Gompels, U.A., Nicholas, J., Lawrence, G., Jones, M., Thomson, B.J., Martin, M.E., et al. (1995). The DNA sequence of human herpesvirus-6: structure, coding content, and genome evolution. *Virology* 209(1)**,** 29-51. doi: 10.1006/viro.1995.1228.

Goncalves, P.H., Uldrick, T.S., and Yarchoan, R. (2017). HIV-associated Kaposi sarcoma and related diseases. *AIDS* 31(14)**,** 1903-1916. doi: 10.1097/QAD.0000000000001567.

Goodwin, S., McPherson, J.D., and McCombie, W.R. (2016). Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet* 17(6)**,** 333-351. doi: 10.1038/nrg.2016.49.

Gotoh, M., Yoshizawa, S., Katagiri, S., Suguro, T., Asano, M., Kitahara, T., et al. (2014). Human herpesvirus 6 reactivation on the 30th day after allogeneic hematopoietic stem cell transplantation can predict grade 2-4 acute graft-versus-host disease. *Transpl Infect Dis* 16(3)**,** 440-449. doi: 10.1111/tid.12229.

Gravel, A., Ablashi, D., and Flamand, L. (2013). Complete Genome Sequence of Early Passaged Human Herpesvirus 6A (GS Strain) Isolated from North America. *Genome Announc* 1(3). doi: 10.1128/genomeA.00012-13.

Gravel, A., Dubuc, I., Wallaschek, N., Gilbert-Girard, S., Collin, V., Hall-Sedlak, R., et al. (2017). Cell Culture Systems To Study Human Herpesvirus 6A/B Chromosomal Integration. *J Virol* 91(14). doi: 10.1128/JVI.00437-17.

Grivel, J.C., Santoro, F., Chen, S., Faga, G., Malnati, M.S., Ito, Y., et al. (2003). Pathogenic effects of human herpesvirus 6 in human lymphoid tissue ex vivo. *J Virol* 77(15)**,** 8280-8289. doi: 10.1128/jvi.77.15.8280-8289.2003.

Grose, C., Yu, X., Cohrs, R.J., Carpenter, J.E., Bowlin, J.L., and Gilden, D. (2013). Aberrant virion assembly and limited glycoprotein C production in varicella-zoster virus-infected neurons. *J Virol* 87(17)**,** 9643-9648. doi: 10.1128/JVI.01506-13.

Gu, B., and DeLuca, N. (1994). Requirements for activation of the herpes simplex virus glycoprotein C promoter in vitro by the viral regulatory protein ICP4. *J Virol* 68(12)**,** 7953-7965. doi: 10.1128/JVI.68.12.7953-7965.1994.

Gu, H., Liang, Y., Mandel, G., and Roizman, B. (2005). Components of the REST/CoREST/histone deacetylase repressor complex are disrupted, modified, and translocated in HSV-1-infected cells. *Proc Natl Acad Sci U S A* 102(21)**,** 7571-7576. doi: 10.1073/pnas.0502658102.

Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). circlize implements and enhances circular visualization in R. *Bioinformatics* 30(19)**,** 2811-2812. doi: 10.1093/bioinformatics/btu393.

Gulve, N., Frank, C., Klepsch, M., and Prusty, B.K. (2017). Chromosomal integration of HHV-6A during non-productive viral infection. *Sci Rep* 7(1)**,** 512. doi: 10.1038/s41598-017-00658-y.

Guo, H., Shen, S., Wang, L., and Deng, H. (2010). Role of tegument proteins in herpesvirus assembly and egress. *Protein Cell* 1(11)**,** 987-998. doi: 10.1007/s13238-010-0120-0.

Hansen, A.S., Pustova, I., Cattoglio, C., Tjian, R., and Darzacq, X. (2017). CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *Elife* 6. doi: 10.7554/eLife.25776.

Harma, M., Hockerstedt, K., and Lautenschlager, I. (2003). Human herpesvirus-6 and acute liver failure. *Transplantation* 76(3)**,** 536-539. doi: 10.1097/01.TP.0000069233.13409.DF.

Hattori, T., Lai, D., Dementieva, I.S., Montano, S.P., Kurosawa, K., Zheng, Y., et al. (2016). Antigen clasping by two antigen-binding sites of an exceptionally specific antibody for histone methylation. *Proc Natl Acad Sci U S A* 113(8)**,** 2092-2097. doi: 10.1073/pnas.1522691113.

Heming, J.D., Conway, J.F., and Homa, F.L. (2017). Herpesvirus Capsid Assembly and DNA Packaging. *Adv Anat Embryol Cell Biol* 223**,** 119-142. doi: 10.1007/978-3-319-53168-7_6.

Henikoff, S., and Grosveld, F. (2008). Welcome to epigenetics & chromatin. *Epigenetics Chromatin* 1(1)**,** 1. doi: 10.1186/1756-8935-1-1.

Herrera, F.J., and Triezenberg, S.J. (2004). VP16-dependent association of chromatin-modifying coactivators and underrepresentation of histones at immediate-early gene promoters during herpes simplex virus infection. *J Virol* 78(18)**,** 9689-9696. doi: 10.1128/JVI.78.18.9689-9696.2004.

Hood, C., Cunningham, A.L., Slobedman, B., Arvin, A.M., Sommer, M.H., Kinchington, P.R., et al. (2006). Varicella-zoster virus ORF63 inhibits apoptosis of primary human neurons. *J Virol* 80(2)**,** 1025-1031. doi: 10.1128/JVI.80.2.1025-1031.2006.

Hope-Simpson, R.E. (1964). The Nature of Herpes Zoster. *Practitioner* 193**,** 217-219.

Hu, M., Depledge, D.P., Flores Cortes, E., Breuer, J., and Schang, L.M. (2019). Chromatin dynamics and the transcriptional competence of HSV-1 genomes during lytic infections. *PLoS Pathog* 15(11)**,** e1008076. doi: 10.1371/journal.ppat.1008076.

Huang, W., Li, L., Myers, J.R., and Marth, G.T. (2012). ART: a next-generation sequencing read simulator. *Bioinformatics* 28(4)**,** 593-594. doi: 10.1093/bioinformatics/btr708.

Huang, Y., Hidalgo-Bravo, A., Zhang, E., Cotton, V.E., Mendez-Bermudez, A., Wig, G., et al. (2014). Human telomeres that carry an integrated copy of human herpesvirus 6 are often short and unstable, facilitating release of the viral genome from the chromosome. *Nucleic Acids Res* 42(1)**,** 315-327. doi: 10.1093/nar/gkt840.

Hubner, M.R., Eckersley-Maslin, M.A., and Spector, D.L. (2013). Chromatin organization and transcriptional regulation. *Curr Opin Genet Dev* 23(2)**,** 89-95. doi: 10.1016/j.gde.2012.11.006.

Hussain, T., and Mulherkar, R. (2012). Lymphoblastoid Cell lines: a Continuous in Vitro Source of Cells to Study Carcinogen Sensitivity and DNA Repair. *Int J Mol Cell Med* 1(2)**,** 75-87.

Hutzinger, R., Feederle, R., Mrazek, J., Schiefermeier, N., Balwierz, P.J., Zavolan, M., et al. (2009). Expression and processing of a small nucleolar RNA from the Epstein-Barr virus genome. *PLoS Pathog* 5(8)**,** e1000547. doi: 10.1371/journal.ppat.1000547.

Ishikawa, K., Hasegawa, K., Naritomi, T., Kanai, N., Ogawa, M., Kato, Y., et al. (2002). Prevalence of herpesviridae and hepatitis virus sequences in the livers of patients with fulminant hepatitis of unknown etiology in Japan. *J Gastroenterol* 37(7)**,** 523-530. doi: 10.1007/s005350200081.

J, B. (2020). "Boyd J (2020). seqsetvis: Set Based Visualizations for Next-Gen Sequencing Data.". R package version 1.8.0. ed.).

Jin Wook Lee, W.M.a.A.K. *Imputed signal tracks* [Online]. Washington University in St. Louis: Wang Lab. Available: https://egg2.wustl.edu/roadmap/web_portal/imputed.html#chr_imp [Accessed 27 September 2020].

Jones, M., Dry, I.R., Frampton, D., Singh, M., Kanda, R.K., Yee, M.B., et al. (2014). RNA-seq analysis of host and viral gene expression highlights interaction between varicella zoster virus and keratinocyte differentiation. *PLoS Pathog* 10(1)**,** e1003896. doi: 10.1371/journal.ppat.1003896.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., et al. (2010). Mediator and cohesin connect gene expression and chromatin architecture. *Nature* 467(7314)**,** 430-435. doi: 10.1038/nature09380.

Kakimoto, M., Hasegawa, A., Fujita, S., and Yasukawa, M. (2002). Phenotypic and functional alterations of dendritic cells induced by human herpesvirus 6 infection. *J Virol* 76(20)**,** 10338-10345. doi: 10.1128/jvi.76.20.10338-10345.2002.

Kalamvoki, M., and Roizman, B. (2011). The histone acetyltransferase CLOCK is an essential component of the herpes simplex virus 1 transcriptome that includes TFIID, ICP4, ICP27, and ICP22. *J Virol* 85(18)**,** 9472-9477. doi: 10.1128/JVI.00876-11.

Kang, H., and Lieberman, P.M. (2011). Mechanism of glycyrrhizic acid inhibition of Kaposi's sarcoma-associated herpesvirus: disruption of CTCF-cohesin-mediated RNA polymerase II pausing and sister chromatid cohesion. *J Virol* 85(21)**,** 11159-11169. doi: 10.1128/JVI.00720-11.

Kang, H., Wiedmer, A., Yuan, Y., Robertson, E., and Lieberman, P.M. (2011). Coordination of KSHV latent and lytic gene control by CTCF-cohesin mediated chromosome conformation. *PLoS Pathog* 7(8)**,** e1002140. doi: 10.1371/journal.ppat.1002140.

Kara, M., O'Grady, T., Feldman, E.R., Feswick, A., Wang, Y., Flemington, E.K., et al. (2019). Gammaherpesvirus Readthrough Transcription Generates a Long Non-Coding RNA That Is Regulated by Antisense miRNAs and Correlates with Enhanced Lytic Replication In Vivo. *Noncoding RNA* 5(1). doi: 10.3390/ncrna5010006.

Kaufer, B.B. (2013). Detection of Integrated Herpesvirus Genomes by Fluorescence In Situ Hybridization (FISH). *Methods in molecular biology* 1064**,** 141-152. doi: 10.1007/978-1-62703-601-6_10.

Kaufer, B.B., and Flamand, L. (2014). Chromosomally integrated HHV-6: impact on virus, cell and organismal biology. *Curr Opin Virol* 9C**,** 111-118. doi: 10.1016/j.coviro.2014.09.010.

Kaufer, B.B., Jarosinski, K.W., and Osterrieder, N. (2011). Herpesvirus telomeric repeats facilitate genomic integration into host telomeres and mobilization of viral DNA during reactivation. *J Exp Med* 208(3)**,** 605-615. doi: 10.1084/jem.20101402.

Kempfer, R., and Pombo, A. (2020). Methods for mapping 3D chromosome architecture. *Nat Rev Genet* 21(4)**,** 207-226. doi: 10.1038/s41576-019-0195-2.

Kennedy, P.G., Grinfeld, E., and Bell, J.E. (2000). Varicella-zoster virus gene expression in latently infected and explanted human ganglia. *J Virol* 74(24)**,** 11893-11898. doi: 10.1128/jvi.74.24.11893-11898.2000.

Kennedy, P.G., Rovnak, J., Badani, H., and Cohrs, R.J. (2015). A comparison of herpes simplex virus type 1 and varicella-zoster virus latency and reactivation. *J Gen Virol* 96(Pt 7)**,** 1581-1602. doi: 10.1099/vir.0.000128.

Kennedy, P.G., and Steiner, I. (1994). A molecular and cellular model to explain the differences in reactivation from latency by herpes simplex and varicella-zoster viruses. *Neuropathol Appl Neurobiol* 20(4)**,** 368-374. doi: 10.1111/j.1365-2990.1994.tb00982.x.

Kennedy, P.G.E., Grinfeld, E., Craigon, M., Vierlinger, K., Roy, D., Forster, T., et al. (2005). Transcriptomal analysis of varicella-zoster virus infection using long oligonucleotide-based microarrays. *J Gen Virol* 86(Pt 10)**,** 2673-2684. doi: 10.1099/vir.0.80946-0.

Khalil, M.I., Che, X., Sung, P., Sommer, M.H., Hay, J., and Arvin, A.M. (2016). Mutational analysis of varicella-zoster virus (VZV) immediate early protein (IE62) subdomains and their importance in viral replication. *Virology* 492**,** 82-91. doi: 10.1016/j.virol.2016.02.012.

Kim, K.D., Tanizawa, H., De Leo, A., Vladimirova, O., Kossenkov, A., Lu, F., et al. (2020). Epigenetic specifications of host chromosome docking sites for latent Epstein-Barr virus. *Nat Commun* 11(1)**,** 877. doi: 10.1038/s41467-019-14152-8.

Kinchington, P.R., Bookey, D., and Turse, S.E. (1995a). The transcriptional regulatory proteins encoded by varicella-zoster virus open reading frames (ORFs) 4 and 63, but not ORF 61, are associated with purified virus particles. *J Virol* 69(7)**,** 4274-4282. doi: 10.1128/JVI.69.7.4274-4282.1995.

Kinchington, P.R., Hougland, J.K., Arvin, A.M., Ruyechan, W.T., and Hay, J. (1992). The varicella-zoster virus immediate-early protein IE62 is a major component of virus particles. *J Virol* 66(1)**,** 359-366. doi: 10.1128/JVI.66.1.359-366.1992.

Kinchington, P.R., and Turse, S.E. (1998). Regulated nuclear localization of the varicella-zoster virus major regulatory protein, IE62. *J Infect Dis* 178 Suppl 1**,** S16-21. doi: 10.1086/514263.

Kinchington, P.R., Vergnes, J.P., and Turse, S.E. (1995b). Transcriptional mapping of varicella-zoster virus regulatory proteins. *Neurology* 45(12 Suppl 8)**,** S33-35. doi: 10.1212/wnl.45.12_suppl_8.s33.

Knipe, D.M., and Cliffe, A. (2008). Chromatin control of herpes simplex virus lytic and latent infection. *Nat Rev Microbiol* 6(3)**,** 211-221. doi: 10.1038/nrmicro1794.

Kobayashi, M., Wilson, A.C., Chao, M.V., and Mohr, I. (2012). Control of viral latency in neurons by axonal mTOR signaling and the 4E-BP translation repressor. *Genes Dev* 26(14)**,** 1527-1532. doi: 10.1101/gad.190157.112.

Kondo, K., Kondo, T., Okuno, T., Takahashi, M., and Yamanishi, K. (1991). Latent human herpesvirus 6 infection of human monocytes/macrophages. *J Gen Virol* 72 (Pt 6)**,** 1401-1408. doi: 10.1099/0022-1317-72-6-1401.

Kondo, K., Kondo, T., Shimada, K., Amo, K., Miyagawa, H., and Yamanishi, K. (2002a). Strong interaction between human herpesvirus 6 and peripheral blood monocytes/macrophages during acute infection. *J Med Virol* 67(3)**,** 364-369. doi: 10.1002/jmv.10082.

Kondo, K., Shimada, K., Sashihara, J., Tanaka-Taya, K., and Yamanishi, K. (2002b). Identification of human herpesvirus 6 latency-associated transcripts. *J Virol* 76(8)**,** 4145-4151. doi: 10.1128/jvi.76.8.4145-4151.2002.

Kost, R.G., Kupinsky, H., and Straus, S.E. (1995). Varicella-zoster virus gene 63: transcript mapping and regulatory activity. *Virology* 209(1)**,** 218-224. doi: 10.1006/viro.1995.1246.

Krijger, P.H.L., Geeven, G., Bianchi, V., Hilvering, C.R.E., and de Laat, W. (2020). 4C-seq from beginning to end: A detailed protocol for sample preparation and data analysis. *Methods* 170**,** 17-32. doi: 10.1016/j.ymeth.2019.07.014.

Kristie, T.M., Liang, Y., and Vogel, J.L. (2010). Control of alpha-herpesvirus IE gene expression by HCF-1 coupled chromatin modification activities. *Biochim Biophys Acta* 1799(3-4)**,** 257-265. doi: 10.1016/j.bbagrm.2009.08.003.

Kubat, N.J., Amelio, A.L., Giordani, N.V., and Bloom, D.C. (2004a). The herpes simplex virus type 1 latency-associated transcript (LAT) enhancer/rcr is hyperacetylated during latency independently of LAT transcription. *J Virol* 78(22)**,** 12508-12518. doi: 10.1128/JVI.78.22.12508-12518.2004.

Kubat, N.J., Tran, R.K., McAnany, P., and Bloom, D.C. (2004b). Specific histone tail modification and not DNA methylation is a determinant of herpes simplex virus type 1 latent gene expression. *J Virol* 78(3)**,** 1139-1149. doi: 10.1128/jvi.78.3.1139-1149.2004.

Kutluay, S.B., DeVos, S.L., Klomp, J.E., and Triezenberg, S.J. (2009). Transcriptional coactivators are not required for herpes simplex virus type 1 immediate-early gene expression in vitro. *J Virol* 83(8)**,** 3436-3449. doi: 10.1128/JVI.02349-08.

Kutluay, S.B., and Triezenberg, S.J. (2009). Regulation of histone deposition on the herpes simplex virus type 1 genome during lytic infection. *J Virol* 83(11)**,** 5835-5845. doi: 10.1128/JVI.00219-09.

Lamb, K.N., Bsteh, D., Dishman, S.N., Moussa, H.F., Fan, H., Stuckey, J.I., et al. (2019). Discovery and Characterization of a Cellular Potent Positive Allosteric Modulator of the Polycomb Repressive Complex 1 Chromodomain, CBX7. *Cell Chem Biol* 26(10)**,** 1365-1379 e1322. doi: 10.1016/j.chembiol.2019.07.013.

Lan, K., and Luo, M.H. (2017). Herpesviruses: epidemiology, pathogenesis, and interventions. *Virol Sin* 32(5)**,** 347-348. doi: 10.1007/s12250-017-4108-2.

Lang, F., Li, X., Vladimirova, O., Hu, B., Chen, G., Xiao, Y., et al. (2017). CTCF interacts with the lytic HSV-1 genome to promote viral transcription. *Sci Rep* 7**,** 39861. doi: 10.1038/srep39861.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9(4)**,** 357-359. doi: 10.1038/nmeth.1923.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10(3)**,** R25. doi: 10.1186/gb-2009-10-3-r25.

Lenac Rovis, T., Bailer, S.M., Pothineni, V.R., Ouwendijk, W.J., Simic, H., Babic, M., et al. (2013). Comprehensive analysis of varicella-zoster virus proteins using a new monoclonal antibody collection. *J Virol* 87(12)**,** 6943-6954. doi: 10.1128/JVI.00407-13.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997*.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25(16)**,** 2078-2079. doi: 10.1093/bioinformatics/btp352.

Liao, R., Zhang, R., Guan, J., and Zhou, S. (2014). A New Unsupervised Binning Approach for Metagenomic Sequences Based on N-grams and Automatic Feature Weighting. *IEEE/ACM Trans Comput Biol Bioinform* 11(1)**,** 42-54. doi: 10.1109/TCBB.2013.137.

Lomonte, P., and Morency, E. (2007). Centromeric protein CENP-B proteasomal degradation induced by the viral protein ICP0. *FEBS Lett* 581(4)**,** 658-662. doi: 10.1016/j.febslet.2007.01.027.

Lomonte, P., Sullivan, K.F., and Everett, R.D. (2001). Degradation of nucleosome-associated centromeric histone H3-like protein CENP-A induced by herpes simplex virus type 1 protein ICP0. *J Biol Chem* 276(8)**,** 5829-5835. doi: 10.1074/jbc.M008547200.

Long, H.M., Meckiff, B.J., and Taylor, G.S. (2019). The T-cell Response to Epstein-Barr Virus-New Tricks From an Old Dog. *Front Immunol* 10**,** 2193. doi: 10.3389/fimmu.2019.02193.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15(12)**,** 550. doi: 10.1186/s13059-014-0550-8.

Luo, G.G., and Ou, J.H. (2015). Oncogenic viruses and cancer. *Virol Sin* 30(2)**,** 83-84. doi: 10.1007/s12250-015-3599-y.

Luppi, M., Barozzi, P., Maiorana, A., Marasca, R., and Torelli, G. (1994). Human herpesvirus 6 infection in normal human brain tissue. *J Infect Dis* 169(4)**,** 943-944. doi: 10.1093/infdis/169.4.943.

Luppi, M., Barozzi, P., Morris, C., Maiorana, A., Garber, R., Bonacorsi, G., et al. (1999). Human herpesvirus 6 latently infects early bone marrow progenitors in vivo. *J Virol* 73(1)**,** 754-759. doi: 10.1128/JVI.73.1.754-759.1999.

Lusso, P., Crowley, R.W., Malnati, M.S., Di Serio, C., Ponzoni, M., Biancotto, A., et al. (2007). Human herpesvirus 6A accelerates AIDS progression in macaques. *Proc Natl Acad Sci U S A* 104(12)**,** 5067-5072. doi: 10.1073/pnas.0700929104.

Mahalingam, R., Gershon, A., Gershon, M., Cohen, J.I., Arvin, A., Zerboni, L., et al. (2019). Current In Vivo Models of Varicella-Zoster Virus Neurotropism. *Viruses* 11(6). doi: 10.3390/v11060502.

Mahalingam, R., Wellish, M., Cohrs, R., Debrus, S., Piette, J., Rentier, B., et al. (1996). Expression of protein encoded by varicella-zoster virus open reading frame 63 in latently infected human ganglionic neurons. *Proc Natl Acad Sci U S A* 93(5)**,** 2122-2124. doi: 10.1073/pnas.93.5.2122.

Mancuso, R., Sicurella, M., Agostini, S., Marconi, P., and Clerici, M. (2019). Herpes simplex virus type 1 and Alzheimer's disease: link and potential impact on treatment. *Expert Rev Anti Infect Ther* 17(9)**,** 715-731. doi: 10.1080/14787210.2019.1656064.

Markus, A., Waldman Ben-Asher, H., Kinchington, P.R., and Goldstein, R.S. (2014). Cellular transcriptome analysis reveals differential expression of pro- and antiapoptosis genes by varicella-zoster virus-infected neurons and fibroblasts. *J Virol* 88(13)**,** 7674-7677. doi: 10.1128/JVI.00500-14.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011* 17(1)**,** 3. doi: 10.14806/ej.17.1.200.

Matsunaga, Y., Yamanishi, K., and Takahashi, M. (1982). Experimental infection and immune response of guinea pigs with varicella-zoster virus. *Infect Immun* 37(2)**,** 407-412. doi: 10.1128/IAI.37.2.407-412.1982.

McGeoch, D.J., Cook, S., Dolan, A., Jamieson, F.E., and Telford, E.A. (1995). Molecular phylogeny and evolutionary timescale for the family of mammalian herpesviruses. *J Mol Biol* 247(3)**,** 443-458. doi: 10.1006/jmbi.1995.0152.

Memedula, S., and Belmont, A.S. (2003). Sequential recruitment of HAT and SWI/SNF components to condensed chromatin by VP16. *Curr Biol* 13(3)**,** 241-246. doi: 10.1016/s0960-9822(03)00048-4.

Michael, E.J., Kuck, K.M., and Kinchington, P.R. (1998). Anatomy of the varicella-zoster virus open-reading frame 4 promoter. *J Infect Dis* 178 Suppl 1**,** S27-33. doi: 10.1086/514273.

Mishra, A., and Hawkins, R.D. (2017). Three-dimensional genome architecture and emerging technologies: looping in disease. *Genome Med* 9(1)**,** 87. doi: 10.1186/s13073-017-0477-2.

Moquin, S.A., Thomas, S., Whalen, S., Warburton, A., Fernandez, S.G., McBride, A.A., et al. (2018). The Epstein-Barr Virus Episome Maneuvers between Nuclear Chromatin Compartments during Reactivation. *J Virol* 92(3). doi: 10.1128/JVI.01413-17.

Moriuchi, M., Moriuchi, H., Straus, S.E., and Cohen, J.I. (1994). Varicella-zoster virus (VZV) virion-associated transactivator open reading frame 62 protein enhances the infectivity of VZV DNA. *Virology* 200(1)**,** 297-300. doi: 10.1006/viro.1994.1190.

Mueller, N.H., Graf, L.L., Orlicky, D., Gilden, D., and Cohrs, R.J. (2009). Phosphorylation of the nuclear form of varicella-zoster virus immediate-early protein 63 by casein kinase II at serine 186. *J Virol* 83(23)**,** 12094-12100. doi: 10.1128/JVI.01526-09.

Mueller, N.H., Walters, M.S., Marcus, R.A., Graf, L.L., Prenni, J., Gilden, D., et al. (2010). Identification of phosphorylated residues on varicella-zoster virus immediate-early protein ORF63. *J Gen Virol* 91(Pt 5)**,** 1133-1137. doi: 10.1099/vir.0.019067-0.

Mullins, T.B., and Krishnamurthy, K. (2020). "Roseola Infantum," in *StatPearls*. (Treasure Island (FL)).

Myers, M.G., Connelly, B.L., and Stanberry, L.R. (1991). Varicella in hairless guinea pigs. *J Infect Dis* 163(4)**,** 746-751. doi: 10.1093/infdis/163.4.746.

Myers, M.G., Duer, H.L., and Hausler, C.K. (1980). Experimental infection of guinea pigs with varicella-zoster virus. *J Infect Dis* 142(3)**,** 414-420. doi: 10.1093/infdis/142.3.414.

Myers, M.G., Stanberry, L.R., and Edmond, B.J. (1985). Varicella-zoster virus infection of strain 2 guinea pigs. *J Infect Dis* 151(1)**,** 106-113. doi: 10.1093/infdis/151.1.106.

Nacheva, E.P., Ward, K.N., Brazma, D., Virgili, A., Howard, J., Leong, H.N., et al. (2008). Human herpesvirus 6 integrates within telomeric regions as evidenced by five different chromosomal sites. *J Med Virol* 80(11)**,** 1952-1958. doi: 10.1002/jmv.21299.

Nagel, M.A., and Bubak, A.N. (2018). Varicella Zoster Virus Vasculopathy. *J Infect Dis* 218(suppl_2)**,** S107-S112. doi: 10.1093/infdis/jiy425.

Nagel, M.A., Choe, A., Traktinskiy, I., Cordery-Cotter, R., Gilden, D., and Cohrs, R.J. (2011). Varicella-zoster virus transcriptome in latently infected human ganglia. *J Virol* 85(5)**,** 2276-2287. doi: 10.1128/JVI.01862-10.

Naqvi, A.R., Shango, J., Seal, A., Shukla, D., and Nares, S. (2018). Herpesviruses and MicroRNAs: New Pathogenesis Factors in Oral Infection and Disease? *Front Immunol* 9**,** 2099. doi: 10.3389/fimmu.2018.02099.

Neitzel, H. (1986). A routine method for the establishment of permanent growing lymphoblastoid cell lines. *Hum Genet* 73(4)**,** 320-326. doi: 10.1007/BF00279094.

Neumann, D.M., Bhattacharjee, P.S., Giordani, N.V., Bloom, D.C., and Hill, J.M. (2007). In vivo changes in the patterns of chromatin structure associated with the latent herpes simplex virus type 1 genome in mouse trigeminal ganglia can be detected at early times after butyrate treatment. *J Virol* 81(23)**,** 13248-13253. doi: 10.1128/JVI.01569-07.

Newhart, A., Rafalska-Metcalf, I.U., Yang, T., Negorev, D.G., and Janicki, S.M. (2012). Single-cell analysis of Daxx and ATRX-dependent transcriptional repression. *J Cell Sci* 125(Pt 22)**,** 5489-5501. doi: 10.1242/jcs.110148.

Nicholas, J., and Martin, M.E. (1994). Nucleotide sequence analysis of a 38.5-kilobase-pair region of the genome of human herpesvirus 6 encoding human cytomegalovirus immediate-early gene homologs and transactivating functions. *J Virol* 68(2)**,** 597-610. doi: 10.1128/jvi.68.2.597-610.1994.

Nilsen, T.W. (2014). Preparation of cross-linked cellular extracts with formaldehyde. *Cold Spring Harb Protoc* 2014(9)**,** 1001-1003. doi: 10.1101/pdb.prot080879.

O'Geen, H., Echipare, L., and Farnham, P.J. (2011). Using ChIP-seq technology to generate high-resolution profiles of histone modifications. *Methods Mol Biol* 791**,** 265-286. doi: 10.1007/978-1-61779-316-5_20.

O'Grady, T., Feswick, A., Hoffman, B.A., Wang, Y., Medina, E.M., Kara, M., et al. (2019). Genome-wide Transcript Structure Resolution Reveals Abundant Alternate Isoform Usage from Murine Gammaherpesvirus 68. *Cell Rep* 27(13)**,** 3988-4002 e3985. doi: 10.1016/j.celrep.2019.05.086.

O'Grady, T., Wang, X., Honer Zu Bentrup, K., Baddoo, M., Concha, M., and Flemington, E.K. (2016). Global transcript structure resolution of high gene density genomes

through multi-platform data integration. *Nucleic Acids Res* 44(18)**,** e145. doi: 10.1093/nar/gkw629.

Ohye, T., Inagaki, H., Ihira, M., Higashimoto, Y., Kato, K., Oikawa, J., et al. (2014). Dual roles for the telomeric repeats in chromosomally integrated human herpesvirus-6. *Sci Rep* 4**,** 4559. doi: 10.1038/srep04559.

Okuno, T., Takahashi, K., Balachandra, K., Shiraki, K., Yamanishi, K., Takahashi, M., et al. (1989). Seroepidemiology of human herpesvirus 6 infection in normal children and adults. *J Clin Microbiol* 27(4)**,** 651-653. doi: 10.1128/jcm.27.4.651-653.1989.

Osterrieder, N., Wallaschek, N., and Kaufer, B.B. (2014). Herpesvirus Genome Integration into Telomeric Repeats of Host Cell Chromosomes. *Annual Review of Virology* 1(1)**,** 215-235. doi: 10.1146/annurev-virology-031413-085422.

Ouwendijk, W.J., Choe, A., Nagel, M.A., Gilden, D., Osterhaus, A.D., Cohrs, R.J., et al. (2012a). Restricted varicella-zoster virus transcription in human trigeminal ganglia obtained soon after death. *J Virol* 86(18)**,** 10203-10206. doi: 10.1128/JVI.01331-12.

Ouwendijk, W.J., Flowerdew, S.E., Wick, D., Horn, A.K., Sinicina, I., Strupp, M., et al. (2012b). Immunohistochemical detection of intra-neuronal VZV proteins in snap-frozen human ganglia is confounded by antibodies directed against blood group A1-associated antigens. *J Neurovirol* 18(3)**,** 172-180. doi: 10.1007/s13365-012-0095-0.

Ouwendijk, W.J., Getu, S., Mahalingam, R., Gilden, D., Osterhaus, A.D., and Verjans, G.M. (2016). Characterization of the immune response in ganglia after primary simian varicella virus infection. *J Neurovirol* 22(3)**,** 376-388. doi: 10.1007/s13365-015-0408-1.

Ozaki, Y., Tajiri, H., Tanaka-Taya, K., Mushiake, S., Kimoto, A., Yamanishi, K., et al. (2001). Frequent detection of the human herpesvirus 6-specific genomes in the livers of children with various liver diseases. *J Clin Microbiol* 39(6)**,** 2173-2177. doi: 10.1128/JCM.39.6.2173-2177.2001.

Pentland, I., and Parish, J.L. (2015). Targeting CTCF to Control Virus Gene Expression: A Common Theme amongst Diverse DNA Viruses. *Viruses* 7(7)**,** 3574-3585. doi: 10.3390/v7072791.

Perera, L.P., Mosca, J.D., Ruyechan, W.T., Hayward, G.S., Straus, S.E., and Hay, J. (1993). A major transactivator of varicella-zoster virus, the immediate-early protein IE62, contains a potent N-terminal activation domain. *J Virol* 67(8)**,** 4474-4483. doi: 10.1128/JVI.67.8.4474-4483.1993.

Perera, L.P., Mosca, J.D., Sadeghi-Zadeh, M., Ruyechan, W.T., and Hay, J. (1992). The varicella-zoster virus immediate early protein, IE62, can positively regulate its cognate promoter. *Virology* 191(1)**,** 346-354. doi: 10.1016/0042-6822(92)90197-w.

Perng, G.C., Jones, C., Ciacci-Zanella, J., Stone, M., Henderson, G., Yukht, A., et al. (2000a). Virus-induced neuronal apoptosis blocked by the herpes simplex virus latency-associated transcript. *Science* 287(5457)**,** 1500-1503. doi: 10.1126/science.287.5457.1500.

Perng, G.C., Slanina, S.M., Yukht, A., Ghiasi, H., Nesburn, A.B., and Wechsler, S.L. (2000b). The latency-associated transcript gene enhances establishment of herpes

simplex virus type 1 latency in rabbits. *J Virol* 74(4)**,** 1885-1891. doi: 10.1128/jvi.74.4.1885-1891.2000.

Phillips, J.E., and Corces, V.G. (2009). CTCF: master weaver of the genome. *Cell* 137(7)**,** 1194-1211. doi: 10.1016/j.cell.2009.06.001.

Poon, A.P., Gu, H., and Roizman, B. (2006). ICP0 and the US3 protein kinase of herpes simplex virus 1 independently block histone deacetylation to enable gene expression. *Proc Natl Acad Sci U S A* 103(26)**,** 9993-9998. doi: 10.1073/pnas.0604142103.

Prusty, B.K., Gulve, N., Rasa, S., Murovska, M., Hernandez, P.C., and Ablashi, D.V. (2017). Possible chromosomal and germline integration of human herpesvirus 7. *J Gen Virol* 98(2)**,** 266-274. doi: 10.1099/jgv.0.000692.

Pugazhenthi, S., Nair, S., Velmurugan, K., Liang, Q., Mahalingam, R., Cohrs, R.J., et al. (2011). Varicella-zoster virus infection of differentiated human neural stem cells. *J Virol* 85(13)**,** 6678-6686. doi: 10.1128/JVI.00445-11.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26(6)**,** 841-842. doi: 10.1093/bioinformatics/btq033.

Reichelt, M., Brady, J., and Arvin, A.M. (2009). The replication cycle of varicella-zoster virus: analysis of the kinetics of viral protein expression, genome synthesis, and virion assembly at the single-cell level. *J Virol* 83(8)**,** 3904-3918. doi: 10.1128/JVI.02137-08.

Roizman, B., Zhou, G., and Du, T. (2011). Checkpoints in productive and latent infections with herpes simplex virus 1: conceptualization of the issues. *J Neurovirol* 17(6)**,** 512-517. doi: 10.1007/s13365-011-0058-x.

Rooney, B.V., Crucian, B.E., Pierson, D.L., Laudenslager, M.L., and Mehta, S.K. (2019). Herpes Virus Reactivation in Astronauts During Spaceflight and Its Application on Earth. *Front Microbiol* 10**,** 16. doi: 10.3389/fmicb.2019.00016.

Rouse, B.T., and Kaistha, S.D. (2006). A tale of 2 alpha-herpesviruses: lessons for vaccinologists. *Clin Infect Dis* 42(6)**,** 810-817. doi: 10.1086/500141.

Roush, K.S., Domiati-Saad, R.K., Margraf, L.R., Krisher, K., Scheuermann, R.H., Rogers, B.B., et al. (2001). Prevalence and cellular reservoir of latent human herpesvirus 6 in tonsillar lymphoid tissue. *Am J Clin Pathol* 116(5)**,** 648-654. doi: 10.1309/Y2HH-B1CK-0F5L-U7B8.

Rozen, S., and Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* 132**,** 365-386. doi: 10.1385/1-59259-192-2:365.

Rubio, E.D., Reiss, D.J., Welcsh, P.L., Disteche, C.M., Filippova, G.N., Baliga, N.S., et al. (2008). CTCF physically links cohesin to chromatin. *Proc Natl Acad Sci U S A* 105(24)**,** 8309-8314. doi: 10.1073/pnas.0801273105.

Ruyechan, W.T. (2010). Roles of cellular transcription factors in VZV replication. *Curr Top Microbiol Immunol* 342**,** 43-65. doi: 10.1007/82_2010_42.

Ruyechan, W.T., Peng, H., Yang, M., and Hay, J. (2003). Cellular factors and IE62 activation of VZV promoters. *J Med Virol* 70 Suppl 1**,** S90-94. doi: 10.1002/jmv.10328.

Sadaoka, T., Depledge, D.P., Rajbhandari, L., Venkatesan, A., Breuer, J., and Cohen, J.I. (2016). In vitro system using human neurons demonstrates that varicella-zoster vaccine virus is impaired for reactivation, but not latency. *Proc Natl Acad Sci U S A* 113(17)**,** E2403-2412. doi: 10.1073/pnas.1522575113.

Sadzot-Delvaux, C., Arvin, A.M., and Rentier, B. (1998). Varicella-zoster virus IE63, a virion component expressed during latency and acute infection, elicits humoral and cellular immunity. *J Infect Dis* 178 Suppl 1**,** S43-47. doi: 10.1086/514259.

Sauerbrei, A. (2016). Varicella-zoster virus infections - antiviral therapy and diagnosis. *GMS Infect Dis* 4**,** Doc01. doi: 10.3205/id000019.

Saviola, A.J., Zimmermann, C., Mariani, M.P., Signorelli, S.A., Gerrard, D.L., Boyd, J.R., et al. (2019a). Chromatin profiles of chromosomally integrated human herpesvirus-6A. *Frontiers in microbiology* 10**,** 1408.

Saviola, A.J., Zimmermann, C., Mariani, M.P., Signorelli, S.A., Gerrard, D.L., Boyd, J.R., et al. (2019b). Chromatin Profiles of Chromosomally Integrated Human Herpesvirus-6A. *Front Microbiol* 10**,** 1408. doi: 10.3389/fmicb.2019.01408.

Schwartz, C., Fischer, M., Mamchaoui, K., Bigot, A., Lok, T., Verdier, C., et al. (2017). Lamins and nesprin-1 mediate inside-out mechanical coupling in muscle cell precursors through FHOD1. *Sci Rep* 7(1)**,** 1253. doi: 10.1038/s41598-017-01324-z.

Seyyedi, N., Dehbidi, G.R., Karimi, M., Asgari, A., Esmaeili, B., Zare, F., et al. (2019). Human herpesvirus 6A active infection in patients with autoimmune Hashimoto's thyroiditis. *Braz J Infect Dis* 23(6)**,** 435-440. doi: 10.1016/j.bjid.2019.10.004.

Silva, L., Cliffe, A., Chang, L., and Knipe, D.M. (2008). Role for A-type lamins in herpesviral DNA targeting and heterochromatin modulation. *PLoS Pathog* 4(5)**,** e1000071. doi: 10.1371/journal.ppat.1000071.

Song, H., Lim, Y., Im, H., Bae, J.M., Kang, G.H., Ahn, J., et al. (2019). Interpretation of EBV infection in pan-cancer genome considering viral life cycle: LiEB (Life cycle of Epstein-Barr virus). *Sci Rep* 9(1)**,** 3465. doi: 10.1038/s41598-019-39706-0.

Spangle, J.M., Roberts, T.M., and Zhao, J.J. (2017). The emerging role of PI3K/AKT-mediated epigenetic regulation in cancer. *Biochim Biophys Acta Rev Cancer* 1868(1)**,** 123-131. doi: 10.1016/j.bbcan.2017.03.002.

Szabo, Q., Bantignies, F., and Cavalli, G. (2019). Principles of genome folding into topologically associating domains. *Sci Adv* 5(4)**,** eaaw1668. doi: 10.1126/sciadv.aaw1668.

Takahashi, M., Okuno, Y., Otsuka, T., Osame, J., and Takamizawa, A. (1975). Development of a live attenuated varicella vaccine. *Biken J* 18(1)**,** 25-33.

Tang, H., Kawabata, A., Yoshida, M., Oyaizu, H., Maeki, T., Yamanishi, K., et al. (2010a). Human herpesvirus 6 encoded glycoprotein Q1 gene is essential for virus growth. *Virology* 407(2)**,** 360-367. doi: http://dx.doi.org/10.1016/j.virol.2010.08.018.

Tang, H., Sadaoka, T., and Mori, Y. (2010b). [Human herpesvirus-6 and human herpesvirus-7 (HHV-6, HHV-7)]. *Uirusu* 60(2)**,** 221-235. doi: 10.2222/jsv.60.221.

Tempera, I., Wiedmer, A., Dheekollu, J., and Lieberman, P.M. (2010). CTCF prevents the epigenetic drift of EBV latency promoter Qp. *PLoS Pathog* 6(8)**,** e1001048. doi: 10.1371/journal.ppat.1001048.

Thellman, N.M., Botting, C., Madaj, Z., and Triezenberg, S.J. (2017). An Immortalized Human Dorsal Root Ganglion Cell Line Provides a Novel Context To Study Herpes Simplex Virus 1 Latency and Reactivation. *J Virol* 91(12). doi: 10.1128/JVI.00080-17.

Tombacz, D., Balazs, Z., Csabai, Z., Moldovan, N., Szucs, A., Sharon, D., et al. (2017a). Characterization of the Dynamic Transcriptome of a Herpesvirus with Long-read Single Molecule Real-Time Sequencing. *Sci Rep* 7**,** 43751. doi: 10.1038/srep43751.

Tombacz, D., Csabai, Z., Szucs, A., Balazs, Z., Moldovan, N., Sharon, D., et al. (2017b). Long-Read Isoform Sequencing Reveals a Hidden Complexity of the Transcriptional Landscape of Herpes Simplex Virus Type 1. *Front Microbiol* 8**,** 1079. doi: 10.3389/fmicb.2017.01079.

Turner, B.M. (2000). Histone acetylation and an epigenetic code. *Bioessays* 22(9)**,** 836-845. doi: 10.1002/1521-1878(200009)22:9<836::AID-BIES9>3.0.CO;2-X.

Tweedy, J., Spyrou, M.A., Pearson, M., Lassner, D., Kuhl, U., and Gompels, U.A. (2016). Complete Genome Sequence of Germline Chromosomally Integrated Human Herpesvirus 6A and Analyses Integration Sites Define a New Human Endogenous Virus with Potential to Reactivate as an Emerging Infection. *Viruses* 8(1). doi: 10.3390/v8010019.

Tyler, J.K., Allen, K.E., and Everett, R.D. (1994). Mutation of a single lysine residue severely impairs the DNA recognition and regulatory functions of the VZV gene 62 transactivator protein. *Nucleic Acids Res* 22(3)**,** 270-278. doi: 10.1093/nar/22.3.270.

Umbach, J.L., Kramer, M.F., Jurak, I., Karnowski, H.W., Coen, D.M., and Cullen, B.R. (2008). MicroRNAs expressed by herpes simplex virus 1 during latent infection regulate viral mRNAs. *Nature* 454(7205)**,** 780-783. doi: 10.1038/nature07103.

van de Werken, H.J., de Vree, P.J., Splinter, E., Holwerda, S.J., Klous, P., de Wit, E., et al. (2012). 4C technology: protocols and data analysis. *Methods Enzymol* 513**,** 89-112. doi: 10.1016/B978-0-12-391938-0.00004-5.

Virgin, H.W., Wherry, E.J., and Ahmed, R. (2009). Redefining chronic viral infection. *Cell* 138(1)**,** 30-50. doi: 10.1016/j.cell.2009.06.036.

Vogel, J.L., and Kristie, T.M. (2013). The dynamics of HCF-1 modulation of herpes simplex virus chromatin during initiation of infection. *Viruses* 5(5)**,** 1272-1291. doi: 10.3390/v5051272.

Wada, Y., Ohta, Y., Xu, M., Tsutsumi, S., Minami, T., Inoue, K., et al. (2009). A wave of nascent transcription on activated human genes. *Proc Natl Acad Sci U S A* 106(43)**,** 18357-18361. doi: 10.1073/pnas.0902573106.

Wagner, L.M., and DeLuca, N.A. (2013). Temporal association of herpes simplex virus ICP4 with cellular complexes functioning at multiple steps in PolII transcription. *PLoS One* 8(10)**,** e78242. doi: 10.1371/journal.pone.0078242.

Wallaschek, N., Sanyal, A., Pirzer, F., Gravel, A., Mori, Y., Flamand, L., et al. (2016). The Telomeric Repeats of Human Herpesvirus 6A (HHV-6A) Are Required for Efficient Virus Integration. *PLoS Pathog* 12(5)**,** e1005666. doi: 10.1371/journal.ppat.1005666.

Walz-Cicconi, M.A., Rose, R.M., Dammin, G.J., and Weller, T.H. (1986). Inoculation of guinea pigs with varicella-zoster virus via the respiratory route. *Arch Virol* 88(3-4), 265-277. doi: 10.1007/BF01310880.

Wang, B., Saito, Y., Nishimura, M., Ren, Z., Tjan, L.H., Refaat, A., et al. (2020). An Animal Model That Mimics Human Herpesvirus 6B Pathogenesis. *J Virol* 94(6). doi: 10.1128/JVI.01851-19.

Wang, J., Lindner, S.E., Leight, E.R., and Sugden, B. (2006). Essential elements of a licensed, mammalian plasmid origin of DNA synthesis. *Mol Cell Biol* 26(3), 1124-1134. doi: 10.1128/MCB.26.3.1124-1134.2006.

Wang, L., Sommer, M., Rajamani, J., and Arvin, A.M. (2009). Regulation of the ORF61 promoter and ORF61 functions in varicella-zoster virus replication and pathogenesis. *J Virol* 83(15), 7560-7572. doi: 10.1128/JVI.00118-09.

Wang, Q.Y., Zhou, C., Johnson, K.E., Colgrove, R.C., Coen, D.M., and Knipe, D.M. (2005). Herpesviral latency-associated transcript gene promotes assembly of heterochromatin on viral lytic-gene promoters in latent infection. *Proc Natl Acad Sci U S A* 102(44), 16055-16059. doi: 10.1073/pnas.0505850102.

Washington, S.D., Edenfield, S.I., Lieux, C., Watson, Z.L., Taasan, S.M., Dhummakupt, A., et al. (2018). Depletion of the Insulator Protein CTCF Results in Herpes Simplex Virus 1 Reactivation In Vivo. *J Virol* 92(11). doi: 10.1128/JVI.00173-18.

Washington, S.D., Singh, P., Johns, R.N., Edwards, T.G., Mariani, M., Frietze, S., et al. (2019). The CCCTC Binding Factor, CTRL2, Modulates Heterochromatin Deposition and the Establishment of Herpes Simplex Virus 1 Latency In Vivo. *J Virol* 93(13). doi: 10.1128/JVI.00415-19.

Webre, J.M., Hill, J.M., Nolan, N.M., Clement, C., McFerrin, H.E., Bhattacharjee, P.S., et al. (2012). Rabbit and mouse models of HSV-1 latency, reactivation, and recurrent eye diseases. *J Biomed Biotechnol* 2012, 612316. doi: 10.1155/2012/612316.

Wight, D.J., Aimola, G., Aswad, A., Jill Lai, C.Y., Bahamon, C., Hong, K., et al. (2020). Unbiased optical mapping of telomere-integrated endogenous human herpesvirus 6. *Proc Natl Acad Sci U S A* 117(49), 31410-31416. doi: 10.1073/pnas.2011872117.

Wilson, A.C., and Mohr, I. (2012). A cultured affair: HSV latency and reactivation in neurons. *Trends Microbiol* 20(12), 604-611. doi: 10.1016/j.tim.2012.08.005.

Wolz, M.M., Sciallis, G.F., and Pittelkow, M.R. (2012). Human herpesviruses 6, 7, and 8 from a dermatologic perspective. *Mayo Clin Proc* 87(10), 1004-1014. doi: 10.1016/j.mayocp.2012.04.010.

Wroblewska, Z., Devlin, M., Reilly, K., van Trieste, H., Wellish, M., and Gilden, D.H. (1982). The production of varicella Zoster virus antiserum in laboratory animals. Brief report. *Arch Virol* 74(2-3), 233-238. doi: 10.1007/BF01314717.

Wroblewska, Z., Valyi-Nagy, T., Otte, J., Dillner, A., Jackson, A., Sole, D.P., et al. (1993). A mouse model for varicella-zoster virus latency. *Microb Pathog* 15(2), 141-151. doi: 10.1006/mpat.1993.1064.

Wysocka, J., and Herr, W. (2003). The herpes simplex virus VP16-induced complex: the makings of a regulatory switch. *Trends Biochem Sci* 28(6), 294-304. doi: 10.1016/S0968-0004(03)00088-4.

Yang, M., Peng, H., Hay, J., and Ruyechan, W.T. (2006). Promoter activation by the varicella-zoster virus major transactivator IE62 and the cellular transcription factor USF. *J Virol* 80(15)**,** 7339-7353. doi: 10.1128/JVI.00309-06.

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., and Madden, T.L. (2012). Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* 13**,** 134. doi: 10.1186/1471-2105-13-134.

Yin, H., Qu, J., Peng, Q., and Gan, R. (2019). Molecular mechanisms of EBV-driven cell cycle progression and oncogenesis. *Med Microbiol Immunol* 208(5)**,** 573-583. doi: 10.1007/s00430-018-0570-1.

Zerboni, L., Reichelt, M., and Arvin, A. (2010a). Varicella-zoster virus neurotropism in SCID mouse-human dorsal root ganglia xenografts. *Curr Top Microbiol Immunol* 342**,** 255-276. doi: 10.1007/82_2009_8.

Zerboni, L., Sobel, R.A., Ramachandran, V., Rajamani, J., Ruyechan, W., Abendroth, A., et al. (2010b). Expression of varicella-zoster virus immediate-early regulatory protein IE63 in neurons of latently infected human sensory ganglia. *J Virol* 84(7)**,** 3421-3430. doi: 10.1128/JVI.02416-09.

Zerr, D.M., Meier, A.S., Selke, S.S., Frenkel, L.M., Huang, M.L., Wald, A., et al. (2005). A population-based study of primary human herpesvirus 6 infection. *N Engl J Med* 352(8)**,** 768-776. doi: 10.1056/NEJMoa042207.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., et al. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 9(9)**,** R137. doi: 10.1186/gb-2008-9-9-r137.

Zhou, J., Chau, C.M., Deng, Z., Shiekhattar, R., Spindler, M.P., Schepers, A., et al. (2005). Cell cycle regulation of chromatin at an origin of DNA replication. *EMBO J* 24(7)**,** 1406-1417. doi: 10.1038/sj.emboj.7600609.

APPENDICES
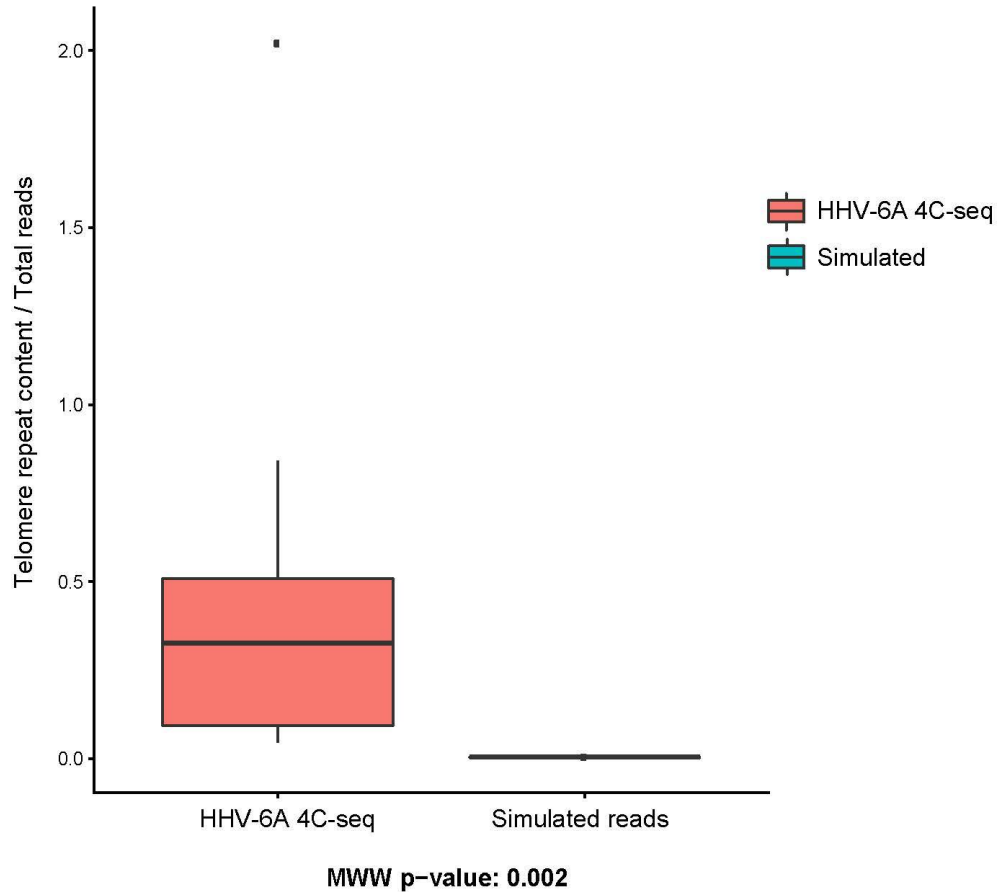
Appendix A-1: Supplemental Figure 2.1. HHV-6A 4C-seq reads are enriched with telomeric repeats. Telomere content comparison was assessed by TelomereHunter (Feuerbach et al., 2019). Illumina data was simulated using ART (Huang et al., 2012). Reads were produced to cover the entire UCSC hg38 reference genome at 1X coverage. Then, several simulated reads equivalent to the number of reads sequenced between ciHHV-6A 4C-seq data were randomly sampled from the simulated data. TelomereHunter was run on all samples and a non-parametric Wilcoxon test was performed to compare 4C-seq reads to simulated reads (p-value = 0.002).
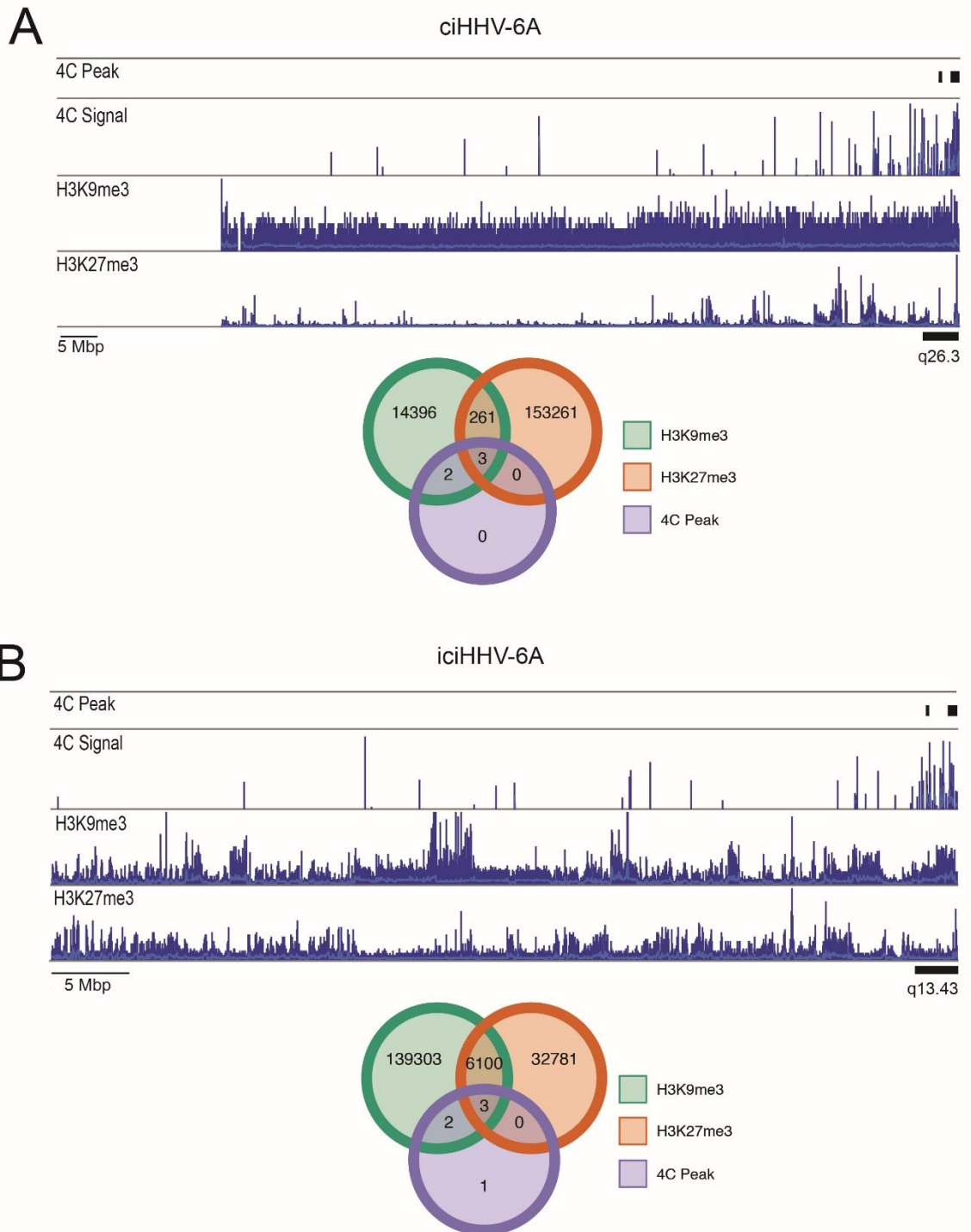
Appendix A-2: Supplemental Figure 2.2. HHV-6A *trans* interaction regions are enriched with repressive histone modifications H3K9me3 and H3K27me3. A. Genome tracks across chromosome 15 for ciHHV-6A showing significant 4C *trans* peak regions (top track), 4C signal coverage, HEK293T H3K9me3 ChIP signal and HEK293T

137

H3K27me3 ChIP signal (bottom track). **B**. The same as **A,** but across chromosome 19 for iciHHV-6A samples.

ciHHV-6A chromHMM Annotation
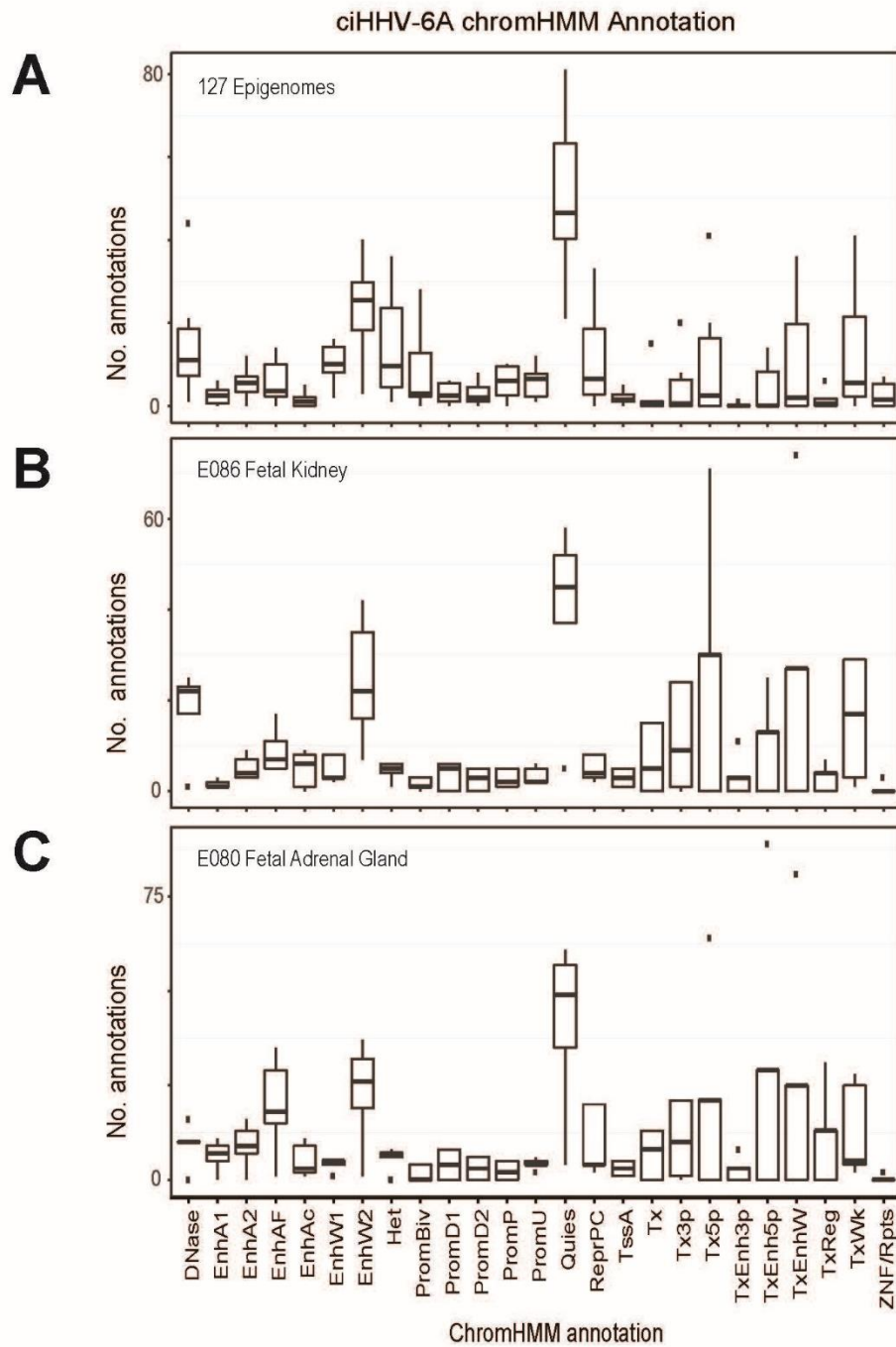
Appendix A-3: Supplemental Figure 2.3. **Similarities in chromatin state annotation in tissues related to HEK293 cells**. **A**. The significant *trans* regions returned by the 4C window method were annotated using the 127 epigenomes chromHMM 25-state model as in **Figure 2D**. **B and C**. Annotations for fetal kidney and fetal adrenal gland tissues were then added for comparison with **A**.

Appendix A-4: Supplemental table 2.1. viewpoint primers designed for HHV-6A 4C-Seq experiments presented in chapter 2.

Experiment	Type	Cell	Viewpoint	Viewpoint forward primer	Viewpoint reverse primer
4C-Seq	illumina	ciHHV-6A (HHV-6A-GFP HEK293T)	VP1	ACTTAGGGT ACATAAAGC TT	CCACAAAACC CCATTATCTC
4C-Seq	illumina	ciHHV-6A (HHV-6A-GFP HEK293T)	VP2	CCCACCCTG ACATAAAGC TT	AAGAAACAGT AAATCTCTCG GT
4C-Seq	illumina	ciHHV-6A (HHV-6A-GFP HEK293T)	VP1	ACTTAGGGT ACATAAAGC TT	CCACAAAACC CCATTATCTC
4C-Seq	illumina	ciHHV-6A (HHV-6A-GFP HEK293T)	VP2	CCCACCCTG ACATAAAGC TT	AAGAAACAGT AAATCTCTCG GT
4C-Seq	illumina	iciHHV-6A (iciHHV-6A+ SMC)	VP1	ACTTAGGGT ACATAAAGC TT	CCACAAAACC CCATTATCTC
4C-Seq	illumina	iciHHV-6A (iciHHV-6A+ SMC)	VP2	CCCACCCTG ACATAAAGC TT	AAGAAACAGT AAATCTCTCG GT
4C-Seq	illumina	iciHHV-6A (iciHHV-6A+ SMC)	VP1	ACTTAGGGT ACATAAAGC TT	CCACAAAACC CCATTATCTC
4C-Seq	illumina	iciHHV-6A (iciHHV-6A+ SMC)	VP2	CCCACCCTG ACATAAAGC TT	AAGAAACAGT AAATCTCTCG GT
4C-Seq	ONP	ciHHV-6A (HHV-6A-GFP HEK293T)	VP1	ACTTAGGGT ACATAAAGC TT	CCACAAAACC CCATTATCTC
4C-Seq	ONP	ciHHV-6A (HHV-6A-GFP HEK293T)	VP2	CCCACCCTG ACATAAAGC TT	AAGAAACAGT AAATCTCTCG GT

4C-Seq	ONP	iciHHV-6A (iciHHV-6A+ SMC)	VP1	ACTTAGGGT ACATAAAGC TT	CCACAAAACC CCATTATCTC
4C-Seq	ONP	iciHHV-6A (iciHHV-6A+ SMC)	VP2	CCCACCCTG ACATAAAGC TT	AAGAAACAGT AAATCTCTCG GT
4C-Seq	ONP	iciHHV-6A (iciHHV-6A+ SMC)	VP1	ACTTAGGGT ACATAAAGC TT	CCACAAAACC CCATTATCTC
4C-Seq	ONP	iciHHV-6A (iciHHV-6A+ SMC)	VP1	ACTTAGGGT ACATAAAGC TT	CCACAAAACC CCATTATCTC
4C-Seq	ONP	iciHHV-6A (iciHHV-6A+ SMC)	VP1	ACTTAGGGT ACATAAAGC TT	CCACAAAACC CCATTATCTC
4C-Seq	ONP	iciHHV-6A (iciHHV-6A+ SMC)	VP1	ACTTAGGGT ACATAAAGC TT	CCACAAAACC CCATTATCTC
4C-Seq	ONP	iciHHV-6A (iciHHV-6A+ SMC)	VP1	ACTTAGGGT ACATAAAGC TT	CCACAAAACC CCATTATCTC
4C-Seq	ONP	Lymphocytes (iciHHV-6B+ Lymphocytes)	VP1	ACTTAGGGT ACATAAAGC TT	CCACAAAACC CCATTATCTC
4C-Seq	ONP	iciHHV-6A (iciHHV-6A+ SMC)	VP2	CCCACCCTG ACATAAAGC TT	AAGAAACAGT AAATCTCTCG GT
4C-Seq	ONP	iciHHV-6A (iciHHV-6A+ SMC)	VP2	CCCACCCTG ACATAAAGC TT	AAGAAACAGT AAATCTCTCG GT
4C-Seq	ONP	iciHHV-6A (iciHHV-6A+ SMC)	VP2	CCCACCCTG ACATAAAGC TT	AAGAAACAGT AAATCTCTCG GT
4C-Seq	ONP	iciHHV-6A (iciHHV-6A+ SMC)	VP2	CCCACCCTG ACATAAAGC TT	AAGAAACAGT AAATCTCTCG GT
4C-Seq	ONP	iciHHV-6A (iciHHV-6A+ SMC)	VP2	CCCACCCTG ACATAAAGC TT	AAGAAACAGT AAATCTCTCG GT
4C-Seq	ONP	Lymphocytes (iciHHV-6B+ Lymphocytes)	VP2	CCCACCCTG ACATAAAGC TT	AAGAAACAGT AAATCTCTCG GT

Appendix A-5: Supplemental table 2.2.1. 4C-Seq alignment mapping results to human reference genome (UCSC Hg38).

Experiment	Virus aligned	Cell	Viewpoint	Sample	Mapped	Unmapped	Percent Mapped
Illumina 4C ciHHV-6A for window method (trans peaks)	GFP	ciHHV-6A (HHV-6A-GFP HEK293T)	vp1	33_S33_L002_R1_001	521861	614726	45.914743
Illumina 4C ciHHV-6A for window method (trans peaks)	GFP	ciHHV-6A (HHV-6A-GFP HEK293T)	vp2	34_S34_L002_R1_001	871036	541435	61.66753158
Illumina 4C ciHHV-6A for window method (trans peaks)	GFP	ciHHV-6A (HHV-6A-GFP HEK293T)	vp1	35_S35_L002_R1_001	687221	767298	8.220028492
Illumina 4C ciHHV-6A for window method (trans peaks)	GFP	ciHHV-6A (HHV-6A-GFP HEK293T)	vp2	36_S36_L002_R1_001	82113	681939	10.74704339
Illumina 4C iciHHV-6A for window method	NC_001664.4	iciHHV-6A (iciHHV-6A+SMC)	vp1	29_S29_L002_R1_001	92493	83829	52.45686868

(trans peaks)							
Illumina 4C iciHHV-6A for window method (trans peaks)	NC_001664.4	iciHHV-6A (iciHHV-6A+SMC)	vp2	30_S30_L002_R1_001	243915	101992	70.51461809
Illumina 4C iciHHV-6A for window method (trans peaks)	NC_001664.4	iciHHV-6A (iciHHV-6A+SMC)	vp1	31_S31_L002_R1_001	941709	677226	58.16842554
Illumina 4C iciHHV-6A for window method (trans peaks)	NC_001664.4	iciHHV-6A (iciHHV-6A+SMC)	vp2	32_S32_L002_R1_001	557669	342514	61.9506256
Illumina 4C ciHHV-6A for peakC (cis peaks)	GFP	ciHHV-6A (HHV-6A-GFP HEK293T)	vp1	33_S33_L002_R1_001	1097794	38793	96.58688688
Illumina 4C ciHHV-6A for peakC (cis peaks)	GFP	ciHHV-6A (HHV-6A-GFP HEK293T)	vp2	34_S34_L002_R1_001	1297841	114630	91.88443515
Illumina 4C ciHHV-6A for	GFP	ciHHV-6A (HHV-6A-	vp1	35_S35_L002_R1_001	804374	31645	96.21479895

peakC (cis peaks)		GFP HEK293T)					
Illumina 4C ciHHV-6A for peakC (cis peaks)	GFP	ciHHV-6A (HHV-6A-GFP HEK293T)	vp2	36_S36_L002_R1_001	738157	25895	96.61083277
Illumina 4C iciHHV-6A for peakC (cis peaks)	NC_001664.4	iciHHV-6A (iciHHV-6A+SMC)	vp1	29_S29_L002_R1_001	132506	43816	75.15000964
Illumina 4C iciHHV-6A for peakC (cis peaks)	NC_001664.4	iciHHV-6A (iciHHV-6A+SMC)	vp2	30_S30_L002_R1_001	308590	37317	89.21184018
Illumina 4C iciHHV-6A for peakC (cis peaks)	NC_001664.4	iciHHV-6A (iciHHV-6A+SMC)	vp1	31_S31_L002_R1_001	1098379	520556	67.84577515
Illumina 4C iciHHV-6A for peakC (cis peaks)	NC_001664.4	iciHHV-6A (iciHHV-6A+SMC)	vp2	32_S32_L002_R1_001	701451	198732	77.92315563
ONP titrations	NC_001664.4	iciHHV-6A (iciHHV-6A+SMC)	vp1	sample_6a_6_10e7_35X	97825	46172	67.9354431

ONP titrations	NC_001664.4	iciHHV-6A (iciHHV-6A+SMC)	vp1	sample_6a_6_10e7_25X	118350	64174	64.84078806
ONP titrations	NC_001664.4	iciHHV-6A (iciHHV-6A+SMC)	vp1	sample_6a_6_10e6	208776	18522	52.98746129
ONP titrations	NC_001664.4	iciHHV-6A (iciHHV-6A+SMC)	vp1	sample_6a_6_10e5	244779	18294	57.23002829
ONP titrations	NC_001664.4	iciHHV-6A (iciHHV-6A+SMC)	vp1	sample_6a_6_10e4	139770	10692	56.64585192
ONP titrations	NC_001664.4	iciHHV-6A (iciHHV-6A+SMC)	vp1	sample_6a_6_pbmc	7102	3516	66.88641929
ONP titrations	NC_001664.4	iciHHV-6A (iciHHV-6A+SMC)	vp2	sample_6a_10_10e7_35X	14338	8905	61.68738975
ONP titrations	NC_001664.4	iciHHV-6A (iciHHV-6A+SMC)	vp2	sample_6a_10_10e7_25X	43438	27536	61.20269394
ONP titrations	NC_001664.4	iciHHV-6A (iciHHV-6A+SMC)	vp2	sample_6a_10_10e6	9044	7192	55.70337522
ONP titrations	NC_001664.4	iciHHV-6A (iciHH	vp2	sample_6a_10_10e5	5723	2868	66.61622628

		V-6A+ SMC)					
ONP titrations	NC_001664.4	iciHHV-6A (iciHHV-6A+ SMC)	vp2	sample_6a_10_10e4	2769	1295	68.13484252
ONP titrations	NC_001664.4	iciHHV-6A (iciHHV-6A+ SMC)	vp2	sample_6a_10_pbmc	4371	2360	64.93834497

Appendix A-6: Supplemental table 2.2.2. Identified significant *cis* interactions.

| sample	cell	viewpoint	chrom	chrom Start	chrom End
bed.peakc.vp1.293	ciHHV-6A (HHV-6A-GFP HEK293T)	vp1	hhv6a-gfp	40483	40544
bed.peakc.vp1.293	ciHHV-6A (HHV-6A-GFP HEK293T)	vp1	hhv6a-gfp	63477	63554
bed.peakc.vp1.293	ciHHV-6A (HHV-6A-GFP HEK293T)	vp1	hhv6a-gfp	77596	77683
bed.peakc.vp1.293	ciHHV-6A (HHV-6A-GFP HEK293T)	vp1	hhv6a-gfp	112004	112081
bed.peakc.vp1.293	ciHHV-6A (HHV-6A-GFP HEK293T)	vp1	hhv6a-gfp	138929	139012
bed.peakc.vp2.293	ciHHV-6A (HHV-6A-GFP HEK293T)	vp2	hhv6a-gfp	53292	53353
bed.peakc.vp2.293	ciHHV-6A (HHV-6A-GFP HEK293T)	vp2	hhv6a-gfp	55502	55579
bed.peakc.vp2.293	ciHHV-6A (HHV-6A-GFP HEK293T)	vp2	hhv6a-gfp	111993	112084
bed.peakc.vp2.293	ciHHV-6A (HHV-6A-GFP HEK293T)	vp2	hhv6a-gfp	117493	117578
bed.peakc.vp2.293	ciHHV-6A (HHV-6A-GFP HEK293T)	vp2	hhv6a-gfp	138993	139013
bed.peakc.vp2.293	ciHHV-6A (HHV-6A-GFP HEK293T)	vp2	hhv6a-gfp	144665	144746
bed.peakc.vp2.293	ciHHV-6A (HHV-6A-GFP HEK293T)	vp2	hhv6a-gfp	158849	158876
bed.peakc.vp1.smc	iciHHV-6A (iciHHV-6A+ SMC)	vp1	NC_001664.4	19887	19946

bed.peakc.vp1.smc	iciHHV-6A (iciHHV-6A+ SMC)	vp1	NC_001664.4	19887	19946
bed.peakc.vp2.smc	iciHHV-6A (iciHHV-6A+ SMC)	vp2	NC_001664.4	18299	18380
bed.peakc.vp2.smc	iciHHV-6A (iciHHV-6A+ SMC)	vp2	NC_001664.4	19951	20014
bed.peakc.vp2.smc	iciHHV-6A (iciHHV-6A+ SMC)	vp2	NC_001664.4	22304	22373
bed.peakc.vp2.smc	iciHHV-6A (iciHHV-6A+ SMC)	vp2	NC_001664.4	41338	41478
bed.peakc.vp2.smc	iciHHV-6A (iciHHV-6A+ SMC)	vp2	NC_001664.4	53323	53419
bed.peakc.vp2.smc	iciHHV-6A (iciHHV-6A+ SMC)	vp2	NC_001664.4	63528	63671
bed.peakc.vp2.smc	iciHHV-6A (iciHHV-6A+ SMC)	vp2	NC_001664.4	75356	75450
bed.peakc.vp2.smc	iciHHV-6A (iciHHV-6A+ SMC)	vp2	NC_001664.4	77651	77732
bed.peakc.vp2.smc	iciHHV-6A (iciHHV-6A+ SMC)	vp2	NC_001664.4	92097	92180
bed.peakc.vp2.smc	iciHHV-6A (iciHHV-6A+ SMC)	vp2	NC_001664.4	94690	94831
bed.peakc.vp2.smc	iciHHV-6A (iciHHV-6A+ SMC)	vp2	NC_001664.4	97662	97750
bed.peakc.vp2.smc	iciHHV-6A (iciHHV-6A+ SMC)	vp2	NC_001664.4	103725	103881
bed.peakc.vp2.smc	iciHHV-6A (iciHHV-6A+ SMC)	vp2	NC_001664.4	108670	108757
bed.peakc.vp2.smc	iciHHV-6A (iciHHV-6A+ SMC)	vp2	NC_001664.4	119614	119637
bed.peakc.vp2.smc	iciHHV-6A (iciHHV-6A+ SMC)	vp2	NC_001664.4	124122	124273
bed.peakc.vp2.smc	iciHHV-6A (iciHHV-6A+ SMC)	vp2	NC_001664.4	149965	150054

Appendix A-7: Supplemental table 2.2.3. Identified significant *trans* interactions.

sample	cell	viewpoint	chrom	chromStart	chromEnd
illumina ciHHV-6A vp1 rep1	ciHHV-6A (HHV-6A-GFP HEK293T)	vp1	chr15	98596798	98961988

illumina ciHHV-6A vp1 rep1	ciHHV-6A (HHV-6A-GFP HEK293T)	vp1	chr15	99920132	100253774
illumina ciHHV-6A vp1 rep1	ciHHV-6A (HHV-6A-GFP HEK293T)	vp1	chr15	101094961	101947367
illumina ciHHV-6A vp1 rep1	ciHHV-6A (HHV-6A-GFP HEK293T)	vp1	chr17	25870638	26110959
illumina ciHHV-6A vp2 rep1	ciHHV-6A (HHV-6A-GFP HEK293T)	vp2	chr13	16986159	17402007
illumina ciHHV-6A vp2 rep1	ciHHV-6A (HHV-6A-GFP HEK293T)	vp2	chr15	99812578	100152996
illumina ciHHV-6A vp2 rep1	ciHHV-6A (HHV-6A-GFP HEK293T)	vp2	chr15	101127425	101947364
illumina iciHHV-6A vp1 rep1	iciHHV-6A (iciHHV-6A+ SMC)	vp1	chr19	57954477	58297244
illumina iciHHV-6A vp1 rep1	iciHHV-6A (iciHHV-6A+ SMC)	vp1	chr5	46491480	46786604
illumina iciHHV-6A vp2 rep1	iciHHV-6A (iciHHV-6A+ SMC)	vp2	chr19	56650087	56887714
illumina iciHHV-6A vp2 rep1	iciHHV-6A (iciHHV-6A+ SMC)	vp2	chr19	58073924	58583680
illumina iciHHV-6A vp1 rep2	iciHHV-6A (iciHHV-6A+ SMC)	vp1	chr19	51751008	52140370
illumina iciHHV-6A vp1 rep2	iciHHV-6A (iciHHV-6A+ SMC)	vp1	chr19	55751644	56222163
illumina iciHHV-6A vp1 rep2	iciHHV-6A (iciHHV-6A+ SMC)	vp1	chr19	56336437	56889554
illumina iciHHV-6A vp1 rep2	iciHHV-6A (iciHHV-6A+ SMC)	vp1	chr19	56999717	57734180

illumina iciHHV-6A vp1 rep2	iciHHV-6A (iciHHV-6A+ SMC)	vp1	chr19	57840298	58583760
illumina iciHHV-6A vp1 rep2	iciHHV-6A (iciHHV-6A+ SMC)	vp1	chr2	93383343	93662225
illumina iciHHV-6A vp2 rep2	iciHHV-6A (iciHHV-6A+ SMC)	vp2	chr19	55875090	56282506
illumina iciHHV-6A vp2 rep2	iciHHV-6A (iciHHV-6A+ SMC)	vp2	chr19	56399232	57787075
illumina iciHHV-6A vp2 rep2	iciHHV-6A (iciHHV-6A+ SMC)	vp2	chr19	57920333	58591574
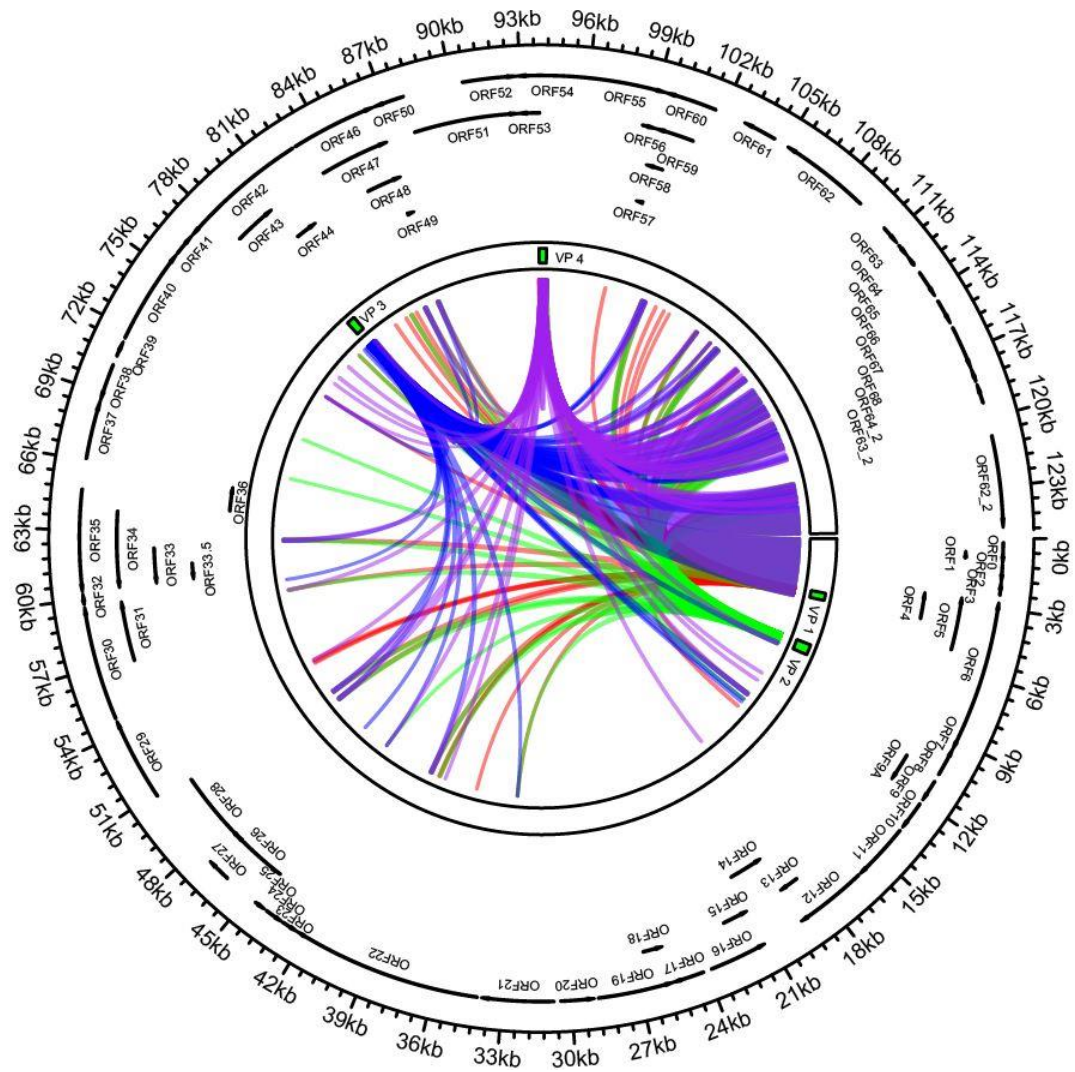
Appendix B-1: Supplemental figure 3.1. Identified significant *cis* interactions across the VZV genome. The virus-virus (*cis*) three-dimensional interactions are visualized above for each of the four viewpoints selected. There is overlap in the contact regions across viewpoints. Regions that overlap between two (or more viewpoints) indicate regions of the virus genome that are in physical proximity to one another in three-dimensional space.

149

Appendix B-2: Supplemental table 3.1. Number of high confidence 4C-Seq *trans* **peaks by chromosome (334 total).**

chrom	peak count
chr12	28
chr1	24
chr11	21
chr17	20
chr2	20
chr3	19
chr5	19
chr10	16
chr6	16
chr20	15
chr15	14
chr7	14
chr8	14
chr9	14
chr16	13
chr14	12
chr13	11
chr19	10

chr4	10
chr22	9
chr18	7
chr21	6
chrX	2

GLOSSARY

HHV(s)	Human Herpesvirus(es)
HSV-1/2	Herpes simplex 1/2 (HHV-1/2, Human alphaherpesvirus 1/2)
VZV	Varicella Zoster Virus (HHV-3, Human alphaherpesvirus 3)
EBV	Epstein Barr Virus (HHV-4, Human gammaherpesvirus 4)
HCMV	Human cytomegalovirus (HHV-5, human betaherpesvirus 5)
HHV-6A/B	Human herpesvirus 6A/B (Human betaherpesvirus 6A/B)
iciHHV-6A	The inheritable germline integrated form of HHV-6A infection. Patients will have a copy of the HHV-6A genome in every cell in their body. Affects ~1% of the world population. Also refers to the cell model presented in this work derived from umbilical artery SMC cells of patients infected with iciHHV-6A.
ciHHV-6A	Chromosomally integrated HHV-6A. HHV-6A infection requires integration of its genome within the host cell genome to establish the latent stage of its life cycle. Also refers to the *in vitro*-derived cell model of HHV-6A infection presented in this work. This cell model harbors a recombinant GFP-expressing HHV-6A virus under control of an hCMV promoter. Reactivation of the latent virus can be achieved via administration of the phorbol ester TPA to these cells.
HHV-7	Human herpesvirus 7 (Human betaherpesvirus 7)
KSHV	Kaposi sarcoma-associated herpesvirus (HHV-8, Human gammaherpesvirus 8)
RNA-Seq	RNA sequencing
ChIP-Seq	Chromatin immunoprecipitation sequencing
MNase-Seq	Micrococcal nuclease digestion with deep sequencing
ATAC-Seq	Assay for transposase-accessible chromatin using sequencing
3C-Seq	Chromosome conformation capture sequencing
4C-Seq	Chromosome conformation capture circular sequencing
5C-Seq	Chromosome conformation capture carbon copy sequencing
Hi-C-Seq	Genome-wide chromatin conformation capture using deep sequencing
High-throughput sequencing	Synonymous herein with "sequencing", "next-generation sequencing", and "deep-sequencing" (e.g., Illumina or Nanopore brand sequencing platforms)
Deep-sequencing	Synonymous herein with "sequencing", "next-generation sequencing", and "high-throughput sequencing" (e.g., Illumina or Nanopore brand sequencing platforms)
Sequencing	Generally, refers to sequencing by synthesis or long-read sequencing technologies in this work. (e.g., Illumina or Nanopore brand sequencing technologies)