CELL TYPE DEPENDENT EFFECTS ON HSV TRANSCRIPTION

by

Justine Harkness

B.S. Carnegie Mellon University, 2010

Submitted to the Graduate Faculty of

the University of Pittsburgh School of Medicine in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2015

UNIVERSITY OF PITTSBURGH

SCHOOL OF MEDICINE

This dissertation was presented

by

Justine M. Harkness

It was defended on

July 16th 2015

and approved by

Fred Homa, Associate Professor, Department of Microbiology and Molecular Genetics

Nicolas Sluis-Cremer, Associate Professor, Division of Infectious Diseases

Karen Arndt, Professor, Department of Biological Sciences

Saleem Khan, Professor, Department of Microbiology and Molecular Genetics

Dissertation Advisor: Neal DeLuca, Professor, Department of Microbiology and Molecular Genetics

CELL TYPE DEPENDENT EFFECTS ON HSV TRANSCRIPTION

Justine M. Harkness

University of Pittsburgh, 2015

Herpes Simplex Virus type 1 (HSV-1) infects both epithelial cells and neurons during the course of its lifecycle. Transcription of viral genes relies heavily on the cellular transcription machinery and chromatin remodeling complexes. Differences in the expression of transcription factors and in the epigenetic landscape in these two cell types could affect the ability of the virus to transcribe its genes and to reactivate from latency. To determine how cell type affects viral gene expression, RNA sequencing was used to quantify differences in viral transcription from wild type and mutant viruses in neurons and MRC5 cells. To ascertain if differences in the cellular epigenetic landscape affect the ability of the virus to reactivate, an *in vitro* model of latency was used in which viral genomes persist in a quiescent state and acquire chromatin modifications similar to *in vivo* latency. Adenovirus vectors were used to deliver the viral activators ICP0 and ICP4 to quiescently infected cells to determine if they could induce transcription from quiescent genomes. RNA sequencing and ChIP were used to quantify viral transcription and the binding of ICP4 to the genome. From these studies, it was concluded that i) viral gene expression from a mutant virus devoid of immediate early gene expression was highly repressed and dysregulated, ii) expression from this virus was greater in neurons than in MRC5 cells, iii) ICP4 was unable to bind to or induce expression from quiescent genomes in MRC5 cells in the absence of ICP0, iv) ICP0 removes repressive chromatin, enabling ICP4 to function on quiescent genomes in MRC5 cells, and v) ICP4 can bind to and induce transcription from quiescent genomes in neurons in the absence of ICP0.

The results of these studies suggest that in the absence of immediate early gene expression, transcription from the genome is highly dysregulated and repressed. However, the degree of repression differs between cell types. Neurons were more permissive to transcription, suggesting that they are less able to repress the viral genome. Furthermore, the viral transcription factor ICP4 was unable to access repressed genomes in MRC5 cells due to the presence of chromatin on the genome. However, ICP4 was able to access some regions of repressed genomes in neurons, suggesting that the form of chromatin on these genomes is less repressive and is not uniform throughout the genome. Together, these data imply that neurons are less efficient in repressing the viral genomes, and this may allow for a finer balance of latency vs. reactivation based on the expression of viral activators and the latency associated transcript.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

HSV-1: Herpes Simplex Virus Type 1

ICP: Infected Cell Polypeptide

tk: thymidine kinase

g(B-M): glycoprotein (B-M)

VZV: Varicella zoster virus

EBV: Epstein–Barr virus

KSHV: Kaposi's sarcoma-associated herpesvirus

HHV: Human herpes virus

HCMV: Human Cytomegalovirus

VP: Viral Protein

LAT: Latency Associated Transcript

OriS: Origin S

OriL: Origin L

HIV: Human Immunodeficiency Virus

HCF: Host Cell Factor

Oct1: Octamer binding protein 1

IE: Immediate Early

E: Early

L: Late

ER: Endoplasmic Reticulum

TAF: TBP Associated Factor

RNA PolII: RNA Polymerase II

SUMO: Small Ubiquitin-like Modifier

UL: Unique Long

US: Unique Short

HVEM: HerpesVirus Entry Mediator

TNF: Tumor Necrosis Factor

NGF: Nerve Growth Factor

NF1: Nuclear Factor 1

Sp1: Specificity Protein 1

TFII (A-H): Transcription Factor II (A-H)

ATM: Ataxia Telangectasia Mutated

INR: Initiator Element

DAS: Downstream Activation Sequence

MOI: Multiplicity of Infection

NuRD: Nucleosome Remodeling and Deacetylase

USP7: Ubiquitin-Specific-Processing Protease 7

hDaxx: Death-domain Associated Protein

ATRX: Alpha Thalassemia Syndrome X-linked

PML: ProMyelocytic Leukemia

ND10: Nuclear Domain 10

Asf1: Anti-silencing function 1 histone chaperone

IFI16: InterFeron gamma-Inducible protein 16

NFkb: Nuclear Factor kappa B

CHD: Chromodomain/helicase/DNA-binding domain

REST: RE1 Silencing Transcription Factor

CoREST: REST Co-repressor protein 1

SWI/SNF: Switch/Sucrose Non Fermantable

CBP: CREB-Binding Protein

TLR: Toll Like Receptor

IRF3: Interferon Regulatory Factor 3

ATR: ATM and Rad3 related

DNA-PK: DNA dependent protein kinase

miRNA: micro RNA

PRC: polycomb repressive complex

CTCF: CCCTC-Binding Factor

DRG: dorsal root ganglia

TG: trigeminal ganglia

EZH: Enhancer of Zeste Homolog

PDK1: Pyruvate Dehydrogenase Kinase 1

mTOR: Mechanistic Target of Rapamycin

4E-BP: eukaryotic translation initiation factor 4E binding protein 1

PAA: PhosphonoAcetic Acid

HAT: Histone Acetyl Transferase

PCAF: P300/CBP-Associated Factor

CLOCK: Circadian Locomotor Output Cycles Kaput

MLL: Mixed-Lineage Leukemia

PHD: Plant HomeoDomain

MBD: Methyl-CpG Binding Domain protein

WD: tryptophan-aspartic acid

RBBP: RetinoBlastoma Binding Protein

MTA: Metastasis Associated

BHC80: BRAF-HDAC complex

CtBP: C-terminal-Binding Protein

PCGF: Polycomb Group Ring Finger

CBX: ChromoBox Homolog

RNAseq: RNA sequencing

ChIP: Chromatin ImmunoPrecipitation

ChIPseq: Chromatin ImmunoPrecipitation Sequencing

GDNF: glial cell-derived neurotrophic factor

RT-PCR: reverse transcription PCR

GFP: Green Fluorescent Protein

Min: minutes

Sec: seconds

mRNA: messenger RNA

cDNA: DNA copy of mRNA

WT: Wild Type

1.0 INTRODUCTION

Herpes Simplex Virus (HSV) infection causes a wide variety of pathologies, ranging from the common cold sore to a potentially fatal encephalitis. A significant percentage of the global population is infected, largely due to the ability of the virus to persist throughout the life time of the host and reactivate from latency to initiate another round of productive infection and spread. Both the lytic and latent portions of the virus lifecycle are regulated by the complex interplay of cellular and viral factors to promote or repress gene expression. During all phases of its lifecycle, cellular chromatin is present on the viral genome and modulates gene expression. During lytic infection, nucleosomes are acetylated and randomly dispersed on the viral genome, characteristic of the active form of chromatin known as euchromatin. However, during latent infection, the viral genome is repressed by densely packed nucleosomes rich in histone methylation, characteristic of the repressive from of chromatin known as heterochromatin. During reactivation of the virus, the epigenetic state of the virus shifts from heterochromatic to euchromatic to enable viral lytic gene expression to resume and likely is mediated by a combination of viral and cellular proteins. An understanding of the ability of viral activators to function on partially heterochromatic genomes will provide greater insight into the process of reactivation. Thus, this introduction will discuss the two distinct stages of the viral life cycle and reactivation, with a particular emphasis on the chromatin control of the viral lifecycle. Additionally, an overview of the different states of chromatin and the complexes responsible for their establishment will be discussed.

1.1 PATHOLOGY AND VIRUS STRUCTURE

1.1.1 The Herpesviridae Family

Herpes Simplex Virus 1 is a member of the family Herpesviridae, which encompasses the herpes viruses that infect mammals, birds, and reptiles. To date, nine herpes viruses have been identified that infect humans, but greatly differ in their host cell range and disease presentation. These nine herpes viruses can be further subdivided into three groups based on their biological properties: alphaherpesvirinae, betaherpesvirinae, and gammaherpesvirinae. Alphaherpesviruses are characterized by a wide host range, short reproductive cycle, killing of host cells, and ability to establish latent infection, mainly in sensory ganglia. HSV-1 and HSV-2 as well as the chicken pox virus Varicella Zoster Virus (VZV) are in this subfamily. Betaherpesviruses tend to have a more restrictive host range and a longer replication cycle, and establish latency in multiple tissues including lymphoid tissue, secretory glands, and kidneys. Infected cells often become enlarged and carrier cells can be easily established. Members of this family include cytomegalovirus (HCMV), the two variants of Roseolovirus (HHV-6A&B) and HHV-7. Gammaherpesviruses primarily infect lymphoid cells, but some members can also replicate in specific types of epithelial cells and fibroblasts *in vitro*. These viruses will usually infect and establish latency in either T or B cells. Epstein-Barr virus (EBV), the causative agent of mononucleosis, and Kaposi's sarcoma-associated herpesvirus (KSHV or HHV-8), an oncogenic virus which commonly induces sarcomas in immunocompromised patients, are members of this family.

All herpesviruses share 4 significant biological properties: i. they specify their own set of enzymes involved in nucleic acid metabolism, DNA replication, and protein processing; ii. they transcribe genes, replicate DNA, and assemble nucleocapsids in the host cell nucleus; iii. lytic replication and production of progeny virions kills the infected cell; and iv. they establish a latent infection and persist throughout the life of the host. The remainder of this introduction will focus specifically on HSV-1.

1.1.2 HSV-1 Viral and Genome Structure

The general structure of a Herpes Simplex Virus is shown in cartoon form in Figure 1. All herpesviruses share a similar structure, but vary in overall size due to differences in tegument thickness. Cyro-Electron Microscopy studies have estimated the size of the HSV-1 virion to be 168nm in diameter, 225nm with the glycoprotein spikes included [1]. The viral particle consists of 4 main parts. At the very core of the virus is the genome, a strand of double stranded DNA that is densely packed in a toroid structure within the viral capsid with polyamines to reduce electrostatic repulsion from the DNA backbone [2, 3]. The dense packing of the genome is theorized to aid the injection of the viral DNA into the nucleus during uncoating [4]. The viral capsid has a T=16 icosahedral symmetry and consists of 162 capsomeres, including 150 hexons, 11 pentons, and the portal protein [5]. It is composed of four main structural proteins, VP5, VP26, VP23, and VP19C [6]. Between the capsid and the envelope is the viral tegument, an unstructured layer consisting of at least 18 proteins. The most notable of these proteins are VP16, the inducer of viral immediate early transcription, and virion host shutoff protein (VHS), which ceases translation of host cell proteins. The exterior layer of the virus is the viral envelope which

is studded with at least 13 glycoproteins, many of which are involved in cell entry. The lipids composing the viral envelope are believed to originate from the cytoplasmic membrane [7].

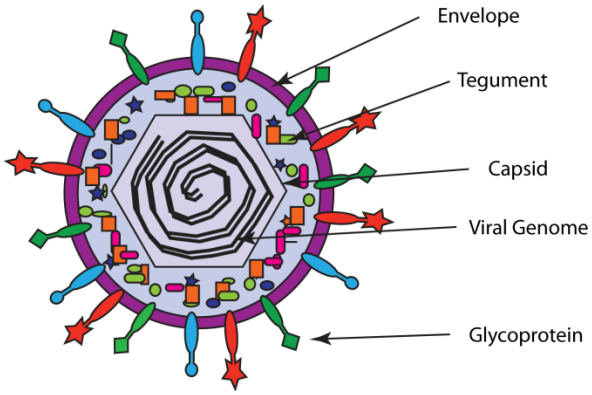

Figure 1: HSV Virion Structure
Cartoon representation of HSV-1 viral particle, adapted from [8]

The herpes virus genome is a linear double strand of DNA of approximately 150 kilobase pairs [9]. The exact length of the genome varies between isolates due to differences in the number of a sequences and length of the inverted repeats. A schematic of the structural composition of the genome is shown in Figure 2. The genome can be thought of as two segments, the unique long and unique short regions, flanked and joined by repeated regions [10]. The unique long and short regions contain the majority of viral genes, and each gene on these segments is present only once. The repeat regions contain sequences for DNA packaging and also contain some viral genes, mainly those that encode immediate early proteins. Since these regions are present twice in the genome, genes within the repeat regions are present in duplicate. Furthermore, the presence of the repeated regions enables recombination during replication, leading to 4 possible orientations of the L and S segments in equimolar amounts [11]. The internal repeat sequences can be deleted to prevent recombination, but this does prohibit replication in cell culture [12].

The viral genome contains 85 protein coding genes that are very densely packed. Despite the dense packing, these genes have their own unique promoters responsive to both viral and cellular transcription factors and are transcribed by the host cell RNA polymerase II [13]. To fit all of these genes in such a short sequence of DNA, many genes are oriented antiparallel and transcribed from opposite strands of DNA. Some genes will also produce a transcript that encodes multiple proteins that are proteolytically cleaved upon translation [14]. Other genes may also share a common 3' terminal end, but initiate transcription at different start sites [15-18]. Unlike cellular genes, most viral genes do not require splicing; only the ICP0, ICP27, ICP22, UL15, and LAT transcripts are spliced.

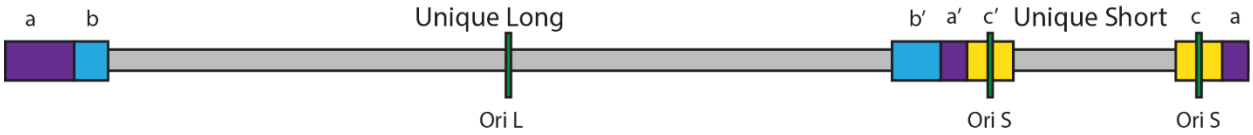

Figure 2: Structure of HSV-1 Genome.
The genome contains unique long and unique short regions, which are flanked by repeated sequences. During DNA replication, recombination occurs between these repeated sequences, yielding four different configurations of the genome in equimolar amounts, adapted from [19]

1.1.3 HSV-1 Prevalence and Pathology

HSV-1 has been documented in the human population as early as the ancient Greek times. Hippocrates first described the cutaneous spreading HSV lesions, leading Greek scholars to define the term herpes to mean "to creep or crawl." However, the transmission of the virus from one individual to another was not recognized until 1893. HSV research blossomed in the twentieth century and in 1919, the infectious nature of HSV was experimentally confirmed by Lowenstein [20].

Herpes simplex virus remains highly prevalent in the human population to this day. It has been estimated that approximately 90% of the population is seropositive and nearly 85-90% of trigeminal ganglia contain herpes DNA at autopsy [21, 22]. This high prevalence among the population can be attributed to the ability of the virus to establish a persistent infection in the host, where the virus remains latent in the ganglia and periodically reactivates to cause recrudescent disease and spread. HSV can cause a primary infection in epithelial cells, including the oral mucosa, skin, and genital tract. Infection of the oral mucosa leads to the development of herpes labialis, which manifests as small blisters on the lips and mouth. Recently, HSV-1 has become the more prevalent cause of herpes genitalis in the younger population, likely due to changes in sexual behavior, reviewed in [23]. The virus can also infect other cutaneous areas of the body. Before the wearing of gloves was common practice, herpes whitlow, an infection of the fingers, was common in dentists and dental hygienists [24]. Additionally, many wrestlers are on prophylactic valacyclovir treatment to prevent the development of herpes gladiatorum, an infection of the skin. It has been estimated that approximately 40% of wrestlers are infected [25].

The virus can cause more severe conditions, depending on both the site of infection and immune competency of the host. When inoculated through an ocular route, the virus can induce inflammation of the cornea through repeated rounds of reactivation, causing stromal keratitis [26]. Herpes stromal keratitis is the leading cause of infectious blindness, affecting approximately 1.5 per 1000 people [27]. While affected individuals require corneal transplants, there is no difference in transplantation efficacy between herpetic and non-herpetic stromal keratitis, indicating that the presence of infection does not negatively impact the transplant in the short term [28].

In both ocular and oral routes of infection, the virus can spread from the primary site of infection through the ganglia to infect the central nervous system and cause a potentially fatal encephalitis. It has been estimated that herpes virus encephalitis constitutes approximately 20% of the total cases of encephalitis and affects 1 in 250,000-500,000 individuals annually [29, 30]. Left untreated, the infection has a 70% mortality rate. The standard treatment is intravenous acyclovir, but mortality remains high at 30% and many survivors are left with neurological impairments [31].

In addition to causing pathology on its own, HSV can synergize with HIV in co-infected individuals, exasperating the replication and spread of both viruses. Studies of co-infected patients have shown that HIV shedding is greatly increased during clinical HSV reactivation in the genital tract [32]. The open sores caused by HSV also enhance the transmission of HIV. Furthermore, the enhanced replication of HIV during active HSV disease has been shown to be independent of the immune state of the host, implicating HSV as the driving factor [33]. HIV infection also enhances the replication and spread of HSV through disruption of tight and adherence junctions at the oral mucosa [34]. Immunocompromised individuals, such as those with HIV, are at an increased risk of developing drug resistant HSV strains, making treatment of these individuals more challenging [35].

1.1.4 HSV as a Gene Therapy Vector

Despite all of the sometimes severe diseases the virus can induce, several decades of research into the HSV virus and progress in recombinant DNA technology has led to the development of HSV as a vector for multiple applications to benefit human health. Uses for HSV vectors

include: i delivery and expression of genes in neuronal tissues; ii targeting and destruction of cancer cells; iii immunotherapy against tumors; and iv potential vaccine development [36-38]. Because the genome is very large and encodes many non-essential genes, regions of the genome can be deleted and replaced with exogenous DNA sequences, up to 40-50kb, for multiple purposes [39]. However, the deletion of viral genes to improve safety can result in a loss of viral activities necessary for efficient gene delivery and prolonged expression in the host. There are currently three different types of vectors used that have been designed to overcome some of these challenges: amplicon vectors, replication-deficient vectors, and replication competent vectors with a restricted host cell range [40].

Amplicon vectors are structurally identical to wild type virus, containing all of the same structural proteins, but have an entirely different genome devoid of any viral genes. Amplicons are made by transfecting cells will the amplicon plasmid, then superinfecting with a HSV-1 helper virus which will encode all of the proteins necessary for viral packaging. To prevent contamination of amplicon stock with helper virus, the genome of the helper virus lacks packaging sites to prevent encapsidation of the genome [41]. Additional improvements on this system include deletion of y34.5 and ICP4, a virulence factor and essential transcription factor, respectively [42]. Since these vectors are effective at delivering genes, but are diluted out as cells divide, they are best suited for acute applications, and have been used for pain management and anti-cancer therapies [43, 44].

Replication deficient vectors are created by knocking out viral genes that are essential to the replication of the virus. Deletions of the viral immediate early proteins ICP4 and ICP27 is sufficient to render the virus incompetent for replication, and these knockout viruses can be grown on complementing cell lines. Replication deficient viruses will infect cells and establish a

lifelong infection, enabling long term expression of genes. Furthermore, insertion of genes into the LAT locus can promote lasting specific expression in neuronal cells [45]. Recombinant viruses have been shown to prevent lethal challenges of infection when used as a vaccine and can promote a lasting immunity [46, 47]. Preclinical studies have also shown that these vectors can be used to deliver growth factors to nerves to treat peripheral neuropathy caused by several factors [48].

Replication competent viruses are the most effective at delivering a transgene in that a small population of infected cells is sufficient to allow spread to more cells. In these viruses, the essential genes are left intact while some virulence genes – including thymidine kinase, ribonucleotide reductase, y34.5, and virion host shutoff protein – are deleted. The first attenuated HSV virus to be characterized and analyzed as a vaccine was NV1020 [49, 50]. Since then, attenuated viral strains have become an attractive means of creating oncolytic viruses. The surface glycoproteins can be modified to target the virus to specific cell types. Recent studies have shown that an oncolytic virus specific to hepatocarcinoma cells can result in tumor regression in a mouse model [51]. Furthermore, several promising oncolytic viruses are currently in clinical trials [52-54].

1.2 LYTIC REPLICATION

1.2.1 General Overview

A schematic of the lytic replication cycle is shown in Figure 3. The virus first enters the cell through attachment to various receptors followed by fusion of the viral envelope to the cell plasma membrane or endocytosis. The viral capsid is released into the cytoplasm and travels along microtubules in an energy dependent manner to the nuclear pore where it docks, releasing viral DNA and some viral proteins into the cell nucleus. Viral gene expression begins when the viral transactivator VP16 complexes with two cellular proteins, Oct-1 and HCF, and induces expression of immediate early proteins, including ICP0 and ICP4 [55]. ICP0 functions to counteract the cellular innate immune response and maintain the viral genome in a euchromatic state. ICP4, the main viral transcription factor, induces the expression of early genes and represses the expression of immediate early genes [56, 57]. Early gene products mainly include proteins required for replication of the viral genome. DNA replication begins approximately 4 hours post infection and proceeds by a rolling circle or recombination based mechanism, producing long concatamers of viral DNA.

Following DNA replication, ICP4 induces the expression of late genes, which encode the structural proteins of the virus. The immediate early protein ICP27 also functions to promote the expression of late genes and some early genes. Late gene products are transported back into the nucleus and viral capsids are assembled. The viral DNA is cleaved into genome length segments and packaged into preformed capsids [58]. The capsids bud into the ER and acquire their envelope studded with viral glycoproteins. The viral particles then exit the ER through budding, acquiring a second membrane, travel through the golgi complex, and then exit the cell through fusion with the plasma membrane. The mature virus will consist of a viral capsid, tegument, and viral envelope. Following replication in epithelial cells, the virus can spread to surrounding epithelial cells or gain access to the ganglia through neurons innervating the primary site of

infection. A more detailed description of the different stages of the viral lifecycle including entry, viral gene transcription, DNA replication, capsid assembly, and chromatin dynamics is presented in the following sections.

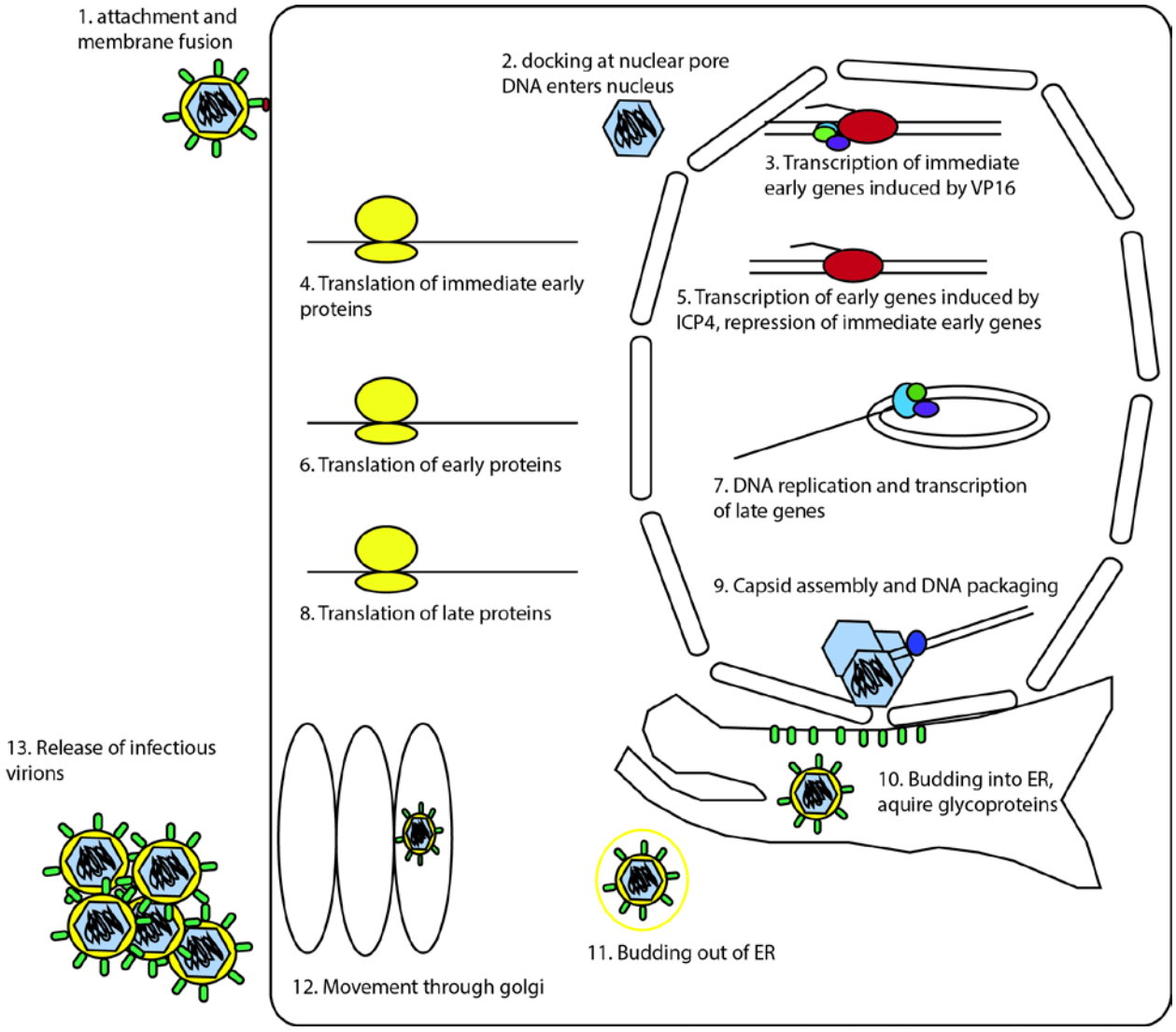

Figure 3: Lytic Replication Cycle of HSV-1.
1. The viral particle binds to surface proteins and enters the cell through fusion of the viral envelope to the plasma membrane. 2. The viral capsid and associated tegument proteins are transported to the nuclear pore. 3&4. Immediate early genes are expressed and translated. 5&6. Early genes are expressed and translated. 7. Viral DNA is replicated, inducing the expression of late genes. 8. Late gene transcripts are translated. 9. Capsid assembles, followed by DNA packaging. 10. Viral capsids bud into ER and acquire glycoproteins. 11. Viral particles bud out

of ER. 12. Viral particles move through the golgi complex. 13. Infectious virions are released from the cell.

1.2.2 Viral Entry

HSV-1 has a very broad range of cells that it can infect due to multiple receptors that the viral glycoproteins can bind: heparin sulfate, HVEM, and nectin-1. The HSV-1 envelope is studded with 12 different glycoproteins. However, only 7 of these 12 are necessary and sufficient for virus entry; these glycoproteins include gB, gC, gD, gH, gK, gL, and gM. The first step in entry is the attachment of the virion to the cell surface through interactions of gB and gC with heparin sulfate on the cell surface. Heparin sulfate is a proteoglycan that is found on the cell surface of most cell types [59]. It is composed of a polyanionic carbohydrate chain of 10-200 disaccharides covalently linked to a protein core and can be further modified by multiple glycosyltransferases in the cell [60]. Both gB and gC can bind heparin sulfate, but gB binds with a higher affinity [61]. While gC can enhance the binding, it is not necessary for attachment. This interaction with heparin sulfate mainly functions to tether the virion to the cell surface and alone is not sufficient for cell entry. However, the 3-O-sulfated isoform of heparin sulfate can directly interact with gD to function as a true receptor for cell entry [62].

In addition to the 3-O-sulfated isoform of heparin sulfate, gD can also interact with herpes virus entry mediator (HVEM) and nectin-1 to induce membrane fusion and cell entry. HVEM is a member of the TNF/NGF receptor family and is primarily expressed on the surface of immune cells [37, 63]. While this receptor is sufficient for viral entry, T cells are not a target of the virus. Rather, the interaction of HSV-1 (and HSV-2) with this receptor is believed to modulate the host immune response to the virus [40]. Nectin-1, a cell adhesion molecule that is

abundantly found on epithelial cells and neuronal synapses, is the more commonly used receptor for viral entry [64, 65]. It is sufficient for viral spread from the cornea to the ganglia [40], but the presence of only nectin-1 leads to attenuation of viral spread [66].

The binding of gD to its receptor triggers conformational changes in the protein, exposing an interaction site to bind to the dimeric gL/gH complex. This interaction induces a conformational change in gL/gH that triggers the conversion of gB into its fusogenic state. The fusion loops of gB are inserted into the cellular plasma membrane, leading to the fusion of the cell membrane with the viral envelope, reviewed in [41]. Following fusion, the viral capsid is released into the cytoplasm. The viral capsid is transported to the nucleus through interactions with microtubules via the motor protein dynein [46]. The capsid docks to the nuclear pore, likely through interactions between tegument proteins and cellular proteins. The mechanism of uncoating at the nuclear pore is still unclear [44]. The viral DNA is then translocated into the nucleus where transcription of viral genes occurs.

1.2.3 Viral Gene Expression

The expression of viral genes occurs through a very tightly controlled cascade of RNApol II mediated transcription [13], depicted below in Figure 4. The tight regulation of transcription is partly due to differences in the architecture of promoters of genes from different kinetic classes. A schematic of these promoters is shown in Figure 5.

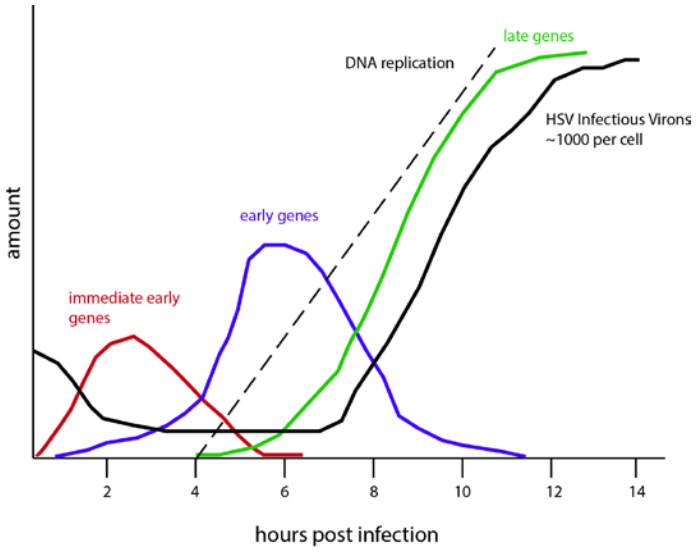

Figure 4: Lytic Gene Expression Cascade
Viral genes are expressed in a highly regulated transcriptional cascade in which immediate early genes are first expressed followed by early and late genes, adapted from [67]

The first genes to be expressed are immediate early genes, which peak in synthesis between 2 and 4 hours post infection and are expressed in the absence of de novo viral protein synthesis [68]. Immediate early promoters have binding sites for VP16 as well as several cellular transcription factors including NF1, Sp1 and TFIID [69, 70]. They are induced when the viral tegument protein VP16 enters the nucleus and forms a complex with cellular proteins HCF and Oct-1 and binds to TAATGARAT sequence elements on the promoter [71]. Approximately 500-1000 copies of VP16 are found in the tegument [72]. VP16 forms contacts with several cellular activators to promote transcription. These specific interactions are discussed in a later section.

VP16 functions to activate transcription of the five immediate early proteins ICP0, ICP4, ICP27, ICP47 and ICP22 which play a role in promoting viral transcription and disarming the host immune response. ICP4 is the main transcription factor of the virus and will activate both early and late genes and repress immediate early genes.

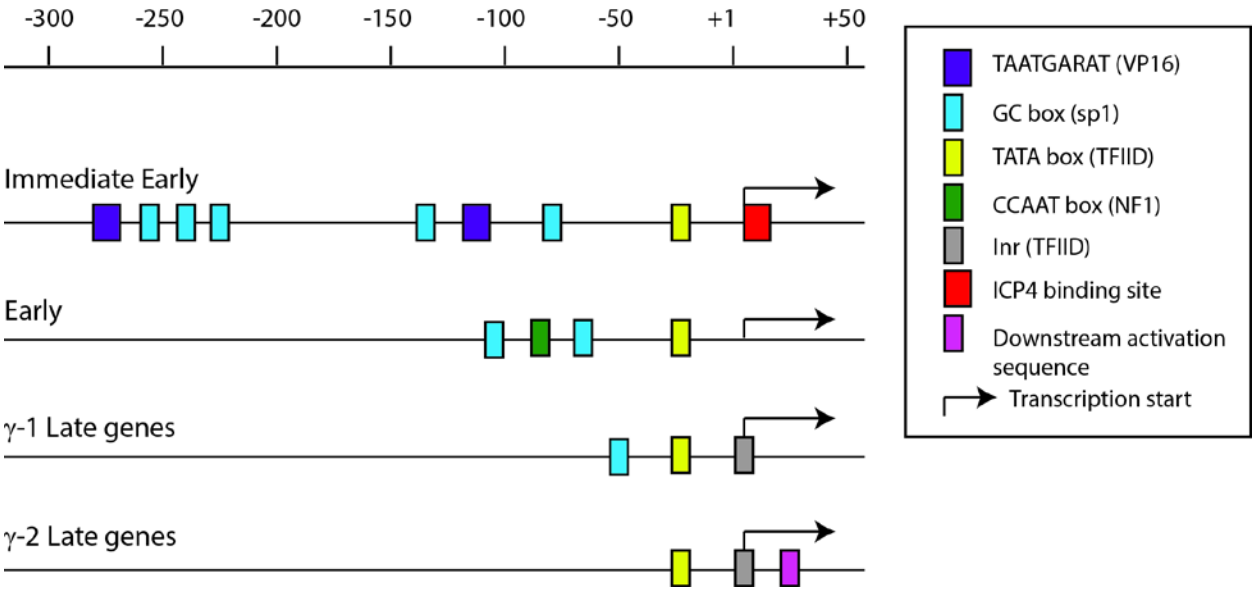

Figure 5: HSV Promoter Architecture
Representation of promoter elements of each kinetic class of viral genes, adapted from *[67]*

Following the expression of immediate early genes, early and late genes will be induced by ICP4. Concurrent with the activation of early genes is the ICP4 mediated repression of immediate early genes. ICP4 represses its own expression through binding to a strong binding site in the promoter downstream of the TATA box and forming a tripartite complex with TBP and TFIIB that blocks formation of the transcription initiation complex [73]. ICP4 binding sites on the ICP0 promoter upstream of the TATA box have also been identified and shown to be important for the repression of ICP0; however, the exact mechanism of repression of this promoter has not yet been elucidated [74]. Furthermore, viruses in which ICP4 has been deleted highly overexpress the other immediate early genes, indicating that other mechanisms exist for repression [56].

Early genes are maximally expressed between 4 and 6 hours post infection. The promoters of early genes lack TAATGARAT elements and are thus not responsive to VP16 activation. ICP4 is the strongest inducer of expression of early genes, but some cellular

transcription factors are also involved. The promoters of early genes contain binding sites for the cellular transcription factors Sp1 and NF1, as well as a TATA box. Removal of the Sp1 sites leads to a 2-3 fold reduction in thymidine kinase expression, a representative early gene. Sp1 enhances transcription by stabilizing the binding of TFIID through interactions with an unidentified protein [75]. However, the absence of NF1 does not alter expression from thymidine kinase, indicating that NF1 may not be directly required for promoting early gene expression [76]. The TATA box is the most important element for ICP4 mediated transcription. ICP4 binds to the DNA in a less specific manner through oligimerization at low affinity sites [77] and functions to stabilize the interaction of TFIID with the TATA box [78]. ICP4 also recruits mediator and other chromatin remodeling complexes to promote early gene expression [79].

Similar to immediate early genes, early genes are repressed following their expression. However, a mechanism for this repression has not yet been identified, though it is likely to depend on changes in the modification or abundance of cellular and viral transcription factors that occur during DNA replication. One possible mechanism could be changes in the phosphorylation of Sp1 that occur during viral DNA replication. During DNA replication, the cell recognizes viral DNA ends and activates the DNA damage response, consequently, the DNA damage response protein ATM phosphorylates Sp1 [80]. Phosphorylated Sp1 does not activate transcription as efficiently *in vitro* [81] and would likely be less efficient in activating early gene expression. Furthermore, the cellular general transcription factor TFIIA is degraded between 4 and 8 hours post infection [82]. During transcription initiation, TFIIA functions to stabilize the binding of TBP with the promoter [83]. TFIIA has been shown to be required for the expression of early genes, but dispensable for the expression of late genes as the INR element can functionally substitute for the absence of TFIIA [84]. Because the INR element is only present in

the promoters of late genes, late genes but not early genes can be expressed when TFIIA levels are low in the cell. ICP27 also enhances the expression of late genes, possibly through linking PolII to progeny DNA through interactions with both the DNA binding protein ICP8 and the RNA polII holoenzyme [85].

In addition to changes in cellular transcription factors, ICP4 differently associates with cellular transcription factors throughout the course of lytic infection. ICP4 primarily associates with components of TFIID early in infection to promote gene expression. The mediator complex does not associate with ICP4 until later in infection, around 6 hours post infection when viral DNA is actively replicating. Furthermore, this form of mediator contains the kinase domain and lacks the component Med26 [79]. This form of mediator is generally associated with gene repression [86]. The transition from early gene expression to late gene expression could in part be due to changes in ICP4 interaction partners. ICP4 bound to TFIID can be a potent activator of early gene expression, but when bound to mediator, could be unable to activate early genes and instead transition to activate late genes.

Late genes are expressed to their peak levels only after the viral DNA has been replicated after 6 hours post infection. Late genes can further be subdivided into γ1 and γ2 classes based on their promoter structure. Leaky late genes, γ1, can be expressed to low levels prior to DNA replication. Their promoters contain Sp1 sites, TATA boxes, and INR elements. True late genes, γ2, require DNA replication for their expression. Their promoters contain INR elements, TATA boxes, and downstream activation sequences (DAS). The level of expression of late genes is most highly correlated to the similarity of the promoter to the consensus sequence of the TATA box and the INR element [87]. The INR elements specify TAF1/2 binding and are required for the expression of the true late protein gC [88]. The DAS is not found in the promoter of all late

genes and is not required for the expression of genes with strong TATA boxes and INR elements. Rather, it functions to promote expression of genes with weak TATA boxes and INR elements [89], such as UL38 which has a non-consensus TATA box and requires this element for expression [90].

1.2.4 DNA Replication

The HSV genome has three origins of replication. Two are located within the short repeats between ICP4 and ICP22 or ICP27 and the third one is located between UL29 and UL30 in the unique long region (Figure 2 [91]). Although all three origins are nested between coding regions, they appear to play no role in transcriptional regulation of the surrounding genes [92]. Both the OriS origins and OriL origin contain A/T rich sequences flanked by binding sites for the origin binding protein UL9, but differ in their exact structure. OriS contains a 45 base pair imperfect palindrome flanked by two UL9 binding sites and a third weak binding site [93]. The OriS sites have been shown to be dispensable for DNA replication, as a mutant virus in which both OriS sites were deleted replicated DNA with only slightly delayed kinetics compared to wild type virus [94]. OriL has a 144 base pair perfect palindrome containing four binding sites for UL9 [95, 96]. It also appears to play a more major role in DNA replication during *in vivo* infection, as mutations in OriL but not in OriS reduce pathogenesis [97].

DNA replication begins around 4 hours post infection after the transcription and translation of early genes, many of which are involved in the replication of the genome. Seven virally encoded genes have been identified to be required for viral DNA replication from both the OriL and OriS [98]. Of these seven proteins, six of them are conserved across all herpes

viruses and include the single strand binding protein ICP8, the two subunit DNA polymerase UL30 and UL42, and the helicase primase complex consisting of UL5, UL8, and UL52. The origin binding protein UL9 is specific to HSV-1 and HSV-2 and no clear homologs have been found in β and γ herpesviruses [99].

DNA replication begins with the recognition of one or several origins of replication by UL9. Since UL9 is required early in replication, but not at later times, a two stage model of DNA replication has been suggested [100]. In the first step of replication, UL9 in complex with ICP8 induces the destabilization of the origin of replication through slightly separating the DNA strands and the activation of the helicase activity of UL9, enabling recruitment of viral replication factors [101]. Once the origin has been activated, the helicase primase complex is recruited, likely through protein-protein interactions; UL9 has been shown to interact with UL8 subunit [102]. The helicase primase complex will then unwind the double stranded DNA and synthesize a short RNA primer. This complex has also been suggested to function as a scaffold for the recruitment of additional replication factors, specifically the DNA polymerase complex [103]. Once recruited, polymerase will catalyze DNA synthesis on the leading and lagging strands. Once DNA synthesis has begun at one or more origins of replication, the process becomes UL9 independent and proceeds by a rolling circle and/or DNA recombination dependent mechanism [19, 104].

1.2.5 Capsid Assembly and DNA Packaging

Assembly of the capsid and DNA packing both occur in the nucleus. The virus can form three different forms of capsids: A, B, and C. The C capsid is the only type that has undergone DNA

packaging and is found in infectious virions. A and B capsids are both dead ends in the packaging process, having either failed to package DNA or failed to release the scaffolding protein [58]. The viral capsid is composed of multiple proteins, the majority of which are true late proteins. The four main structural proteins that compose the capsid are the major capsid protein VP5 and three less abundant proteins: VP19C, VP23, and VP26. The capsid also contains minor proteins including UL6, UL15, UL17, UL28, and UL33 which are required for the cleavage and packaging of viral DNA into preassembled capsids. The capsid has a T=16 icosahedral lattice and is composed of 162 capsomers – 150 hexons, 11 pentons, and the portal where DNA enters the capsid.

The capsid is assembled through the use of a scaffolding protein. The scaffold protein, UL26.5 (VP22a), forms a core scaffold inside the capsid and interacts directly with VP5 though hydrophobic interactions with the N terminus of VP5. The scaffold protein has the ability to self interact and multimerize which drives the assembly of the capsid. As the capsid assembles, one portal complex, composed of 12 UL6 protein subunits, will be added through interactions with the scaffold [105]. The triplex, composed mainly of V19C and VP23, is added as the hexons and pentons assemble. VP26 will localize to the distal tips of VP5 in hexons, but not pentons. The formed capsid containing scaffolding protein is referred to as the procapsid, an unstable conformation. The procapsid can undergo a structural change prior to DNA packaging and form a B capsid or DNA packaging proteins can associate with the portal to begin packaging.

Five proteins have been shown to be necessary for DNA packaging, two of which are of unknown function. UL15, UL28, and UL33 comprise the terminase complex, an ATPase, which associates with the portal and is responsible for inserting the DNA into the capsid and cleaving the DNA into genome lengths [106]. The complex recognizes and cleaves specific packaging

signals in the viral DNA, denoted pac1 and pac2. Following packaging of the DNA, UL25 and UL17, the two components of the C-capsid specific component (CCSC), associate with the viral capsid and function to stabilize the DNA within the capsid to form a mature C capsid [107, 108]. In the absence of this CCSC, the DNA leaves the capsid, forming an A capsid, another dead end of virus assembly.

The mature capsid will exit the nucleus through exocytosis and acquire its envelope and glycoproteins while budding through the inner nuclear membrane. The viral envelope fuses with the outer nuclear membrane, releasing the virus into the cytosol, where it will pass through the trans-golgi network. It will be packaged within two membranes as it leaves the golgi and will exit the cell through fusion of the second membrane to the plasma membrane [109]. The mature virion will go on to infect other epithelial cells or can enter the nerve tissue through neurons innervating the primary site of infection.

1.2.6 Chromatin Dynamics during Lytic Infection

The viral DNA in fully assembled virions is completely devoid of cellular histones. HSV particles were found to contain the polyamines spermine and spermidine. Spermidine was found to be primarily associated with the viral envelope, where as spermine was mainly found in the nucleocapsids, possibly due to a higher affinity for viral DNA. Furthermore, there is a constant molar ratio between DNA phosphate and polyamine nitrogen, suggesting that the DNA is packaged with the polyamines, primarily spermine, to stabilize the viral DNA through neutralization of the negatively charged phosphate backbone, similar to T4 phages [3].

However, when the viral DNA enters the nucleus, cellular histones are deposited onto the genome. Histone deposition and removal is regulated by several factors, including localization of the incoming viral genome, cellular chromatin modifiers, and viral proteins. The viral genome is localized to the nuclear periphery through interactions with type A lamins. This localization is correlated to a reduction in heterochromatin deposition, measured by H3K9 and H3K27 trimethylation, on viral genomes early in infection [110].

Two different cellular complexes can localize to viral genomes and deposit histones: hDaxx/ATRX and HIRA/Asf1. Early during infection, the major ND10 complex proteins PML, Sp100, and hDaxx rapidly re-localize to viral genomes [111]. HDaxx functionally interacts with the SWI/SNF chromatin remodeler ATRX to deposit H3.3 containing histones on viral DNA. [112]. hDaxx functionally interacts with HDAC2 and Dek, a protein that has been shown to change the topology of chromatin and block DNA replication and could possibly impair transcription [113, 114]. Chromatin deposition by the hDaxx ATRX complex ultimately leads to the formation of heterochromatin, likely aided by other proteins localized to ND10 such as Sp100.

On the other hand, Asf1 and HIRA mediated histone deposition promotes HSV lytic infection. Depletion of Asf1 from cells led to an increase in the expression of immediate early genes, but not early and late genes. Despite the increase in immediate early gene expression, there was a decrease in viral DNA replication and viral yield, indicating that Asf1 is important in promoting HSV DNA replication and virus production [115]. Asf1 forms a complex with HIRA to promote the deposition of H3.3 in a DNA replication independent manner [116]. HIRA has been shown to deposit H3.3 histones onto viral DNA early in infection and the incorporation of H3.3 correlates with higher gene expression. Knockdown of HIRA leads to a decrease in H3.3

occupancy on the genome and lower levels of gene expression and viral replication. HIRA is not required for the incorporation of H3.1 containing nucleosomes, which are added to the genome in a DNA replication dependent manner [117]. The localization of the viral genomes to either of these two complexes may determine the initial epigenetic state of incoming genomes.

There are also several cellular complexes that modify viral chromatin to either activate or repress viral transcription. These complexes are shown in Figure 6. Expression from the viral genome is silenced through the deposition of repressive heterochromatin marks, notably H3K9 and H3K27 trimethylation. The polycomb repressive complex 2 methylates H3K27 residues during latent infection [118] and is likely, but has not shown to be, responsible for this methylation early in lytic infection. The methylase responsible for methylating H3K9 during infection has not yet been identified, though Suv39h1 is a potential candidate. The chromodomain helicase DNA-binding (CHD) nucleosome remodeler family member proteins CHD1 and CHD3 can then bind to trimethylated H3K9 and H3K27 and promote the compaction of viral chromatin [119]. In addition to depositing silencing marks onto the genome, other cellular complexes remove activating marks for viral chromatin. The CoREST complex consists of the REST and CoREST proteins in complex with HDACs1 and 2 as well as LSD1. In this complex, HDAC1 stimulates the demethylase activity of LSD1 to remove the activating H3K4 methylation [120, 121].

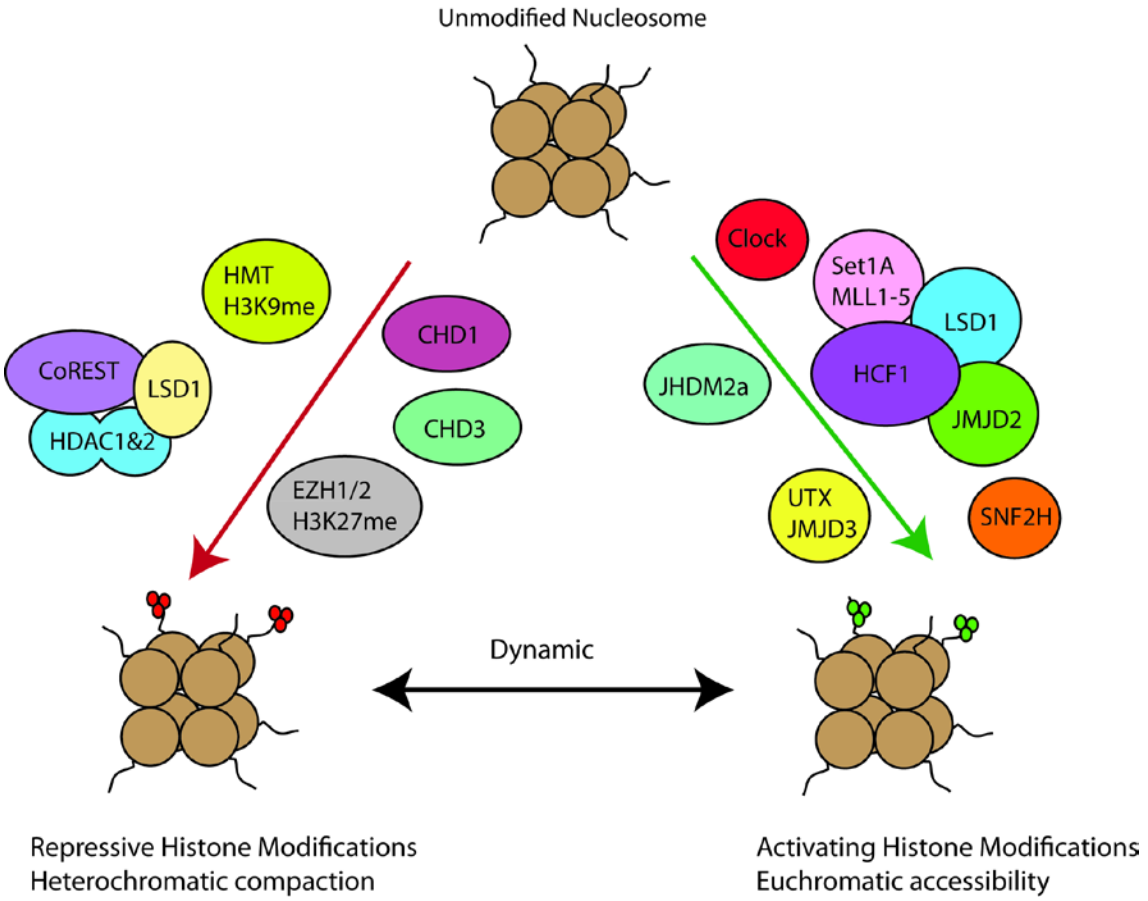

Figure 6: Chromatin Modifiers
Complexes involved in depositing or removing silencing marks on the HSV genome during lytic replication, adapted from [122]

To counteract the deposition of silencing chromatin early in infection, viral transcription factors actively recruit cellular histone remodeling complexes to viral genomes and disable silencing factors to remove silencing modifications and add activating modifications. VP16 interacts with Oct-1 and HCF to induce expression of immediate early genes, while removing repressive chromatin from the genome. HCF-1 serves as a scaffold for the recruitment of various viral and cellular proteins and is required for effective transcription of immediate early genes [123]. HCF-1 interacts with and recruits the Set1 and MLL1 complexes as well as the demethylases JMJD2 and LSD1 to catalyze the removal of methylation from H3K9 and promote the methylation of H3K4. The proteins in the Set1/MLL/JMJD2/LSD1 complex cooperate in a

step wise manner to first demethylate H3K9 and then methylate H3K4. The demethylation of H3K9 is driven by JMJD2 and LSD1. The JMJD2 family of proteins can demethylate trimethylated H3K9, where as LSD1 can only demethylate mono and dimethylated H3K9, thus JMJD2 function is a prerequisite for the action of LDS1 [124]. Knockdown of either of these proteins leads to an increase in repressive chromatin on viral promoters and a decrease in viral gene expression [125, 126]. Once H3K9 methylation has been removed, the Set1 or MLL1 complex will methylate H3K4. Both complexes are recruited by HCF-1 to immediate early promoters [127]. Knockdown of Set1 led to a decrease in H3K4 methylation, demonstrating Set1 was required for this methylation [128].

Other chromatin remodelers are recruited to viral genomes independent of interactions with the VP16 activator complex. SNF2H, a member of the ISWI family of chromatin remodeling enzymes, was found to be localized to viral replication compartments [129]. SNF2H was found to be associated with the viral genome when immediate early genes are expressed and declined as infection progressed to late gene expression. Furthermore, in the absence of SNF2H, immediate early gene expression was significantly decreased while histone association increased, suggesting that SNF2H functions early in infection to enhance the expression of immediate early genes to drive lytic replication [130].

The CLOCK acetyltransferase has also been shown to be recruited to the viral genome at ND10 bodies and is stabilized during infection. Furthermore, knock down of CLOCK during infection led to a decrease in gene expression indicating that it promotes viral gene expression. Overexpression of CLOCK, but not a catalytically dead mutant was also shown to partially compensate for ICP0 during infection with an ICP0 null virus, indicating that the histone acetylase activity of CLOCK is required to promote gene expression [131]. ICP0 also functions

independent of VP16 to counteract cellular repressors through catalyzing the addition of uniquitin to substrates and targeting them for degradation by the cellular proteosome. The many proteins that ICP0 targets are discussed in the next section under the function of ICP0 as a viral activator.

The synergistic functioning of these viral activators and cellular chromatin remodeling complexes leads to a euchromatic viral genome. Multiple studies using micrococcal nuclease, a nuclease that cleaves regions of DNA between nucleosomes, have shown that during lytic infection HSV DNA was not organized into nucleosomes [132, 133]. More recent studies have shown that HSV DNA is associated with nucleosomes-like complexes throughout lytic infection, but they are unstable [134]. Furthermore, histones associated with viral genomes are modified in ways that promote viral transcription. Histones are acetylated at positions H3K9 and H3K14 during lytic infection, and H3K4 is methylated [135].

1.3 VIRAL ACTIVATORS

1.3.1 VP16

VP16 is a 490 amino acid protein containing 2 functional domains, the core domain and the C terminal activation domain, Figure 7A. The core domain resides in the center of the protein, is required for DNA binding, and is conserved among many alpha herpes viruses [55]. It has both structured and unstructured regions; the structured region recognizes and binds the TAATGARAT consensus sequence, while the unstructured region interacts with Oct-1 and HCF to stabilize the binding of VP16 to the DNA [136, 137]. The activation domain is enriched in

acidic residues and resides in the 81 C terminal amino acids. It can further be divided into two distinct regions, H1 and H2, based on interaction partners [138].

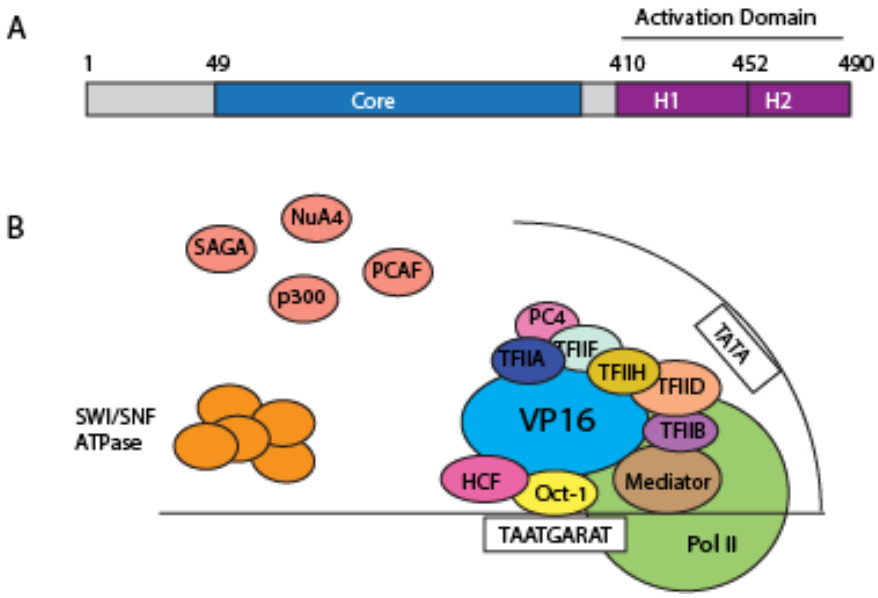

Figure 7: VP16 Structure and Function
A. Functional domains of the VP16 protein. B. Schematic of interaction of VP16 with transcriptional activators and DNA, adapted from *[139]*

The VP16 activation domain is one of the most potent activation domains known due to its ability to interact with multiple transcriptional activators and chromatin modifiers (Figure 7B). It is required for the recruitment TBP and RNA polymerase II to viral immediate early promoters and for maintaining immediate early promoters and reading frames in a relatively histone free state [140, 141]. In vitro studies using fusion proteins showed that the VP16 activation domain recruits the RNA polymerase through direct interactions with the basal transcription factors TFIIA, TFIIB, TFIIF, TFIIH and several components of TFIID [142-146].To further promote transcription, it recruits the mediator complex through direct interactions with two subunits, MED25 and MED17 [147, 148]. Recent ChIP studies have shown

that VP16 can interact with and recruit the histone deacetylases CBP and p300 and the SWI/SNF complex to immediate early promoters [140].

1.3.2 ICP4

ICP4 is the master regulator of viral gene expression. The gene encoding ICP4 is present in the genome in duplicate in the inverted repeat regions of the virus. ICP4 is expressed early in infection and controls the expression of all other viral genes. The protein is 1294 amino acids long with a predicted molecular weight of 133kDa, however, due to extensive phosphorylation and ADP-ribosylation, its actual size is closer to 175kDa [149, 150]. There are several domains of ICP4 that are critical for its functioning. These domains are depicted in Figure 8A. ICP4 has a DNA binding domain that is required for its function as ICP4 must bind to DNA to recruit cellular transcription factors. It also has a multimerization domain which enables it to self interact and multimerize on DNA, increasing its binding affinity. ICP4 has both N-terminal and C-terminal activation domains.

ICP4 binds to the viral DNA through the recognition of the loose consensus sequence RTCGTCNNYNYSG where R is a purine, Y is a pyramidine, S is C or G, and N is any base [151]. There are strong ICP4 binding sites in the promoters of OrfP/L/ST, ICP4, and LAT which are required for their repression. However, the majority of the ICP4 binding sites on the genome are low affinity, and several have been shown to be dispensable for ICP4 transactivation [76]. ICP4 is able to bind to these weak sites through oligmerization on DNA which increases its overall affinity to the viral genome [77].

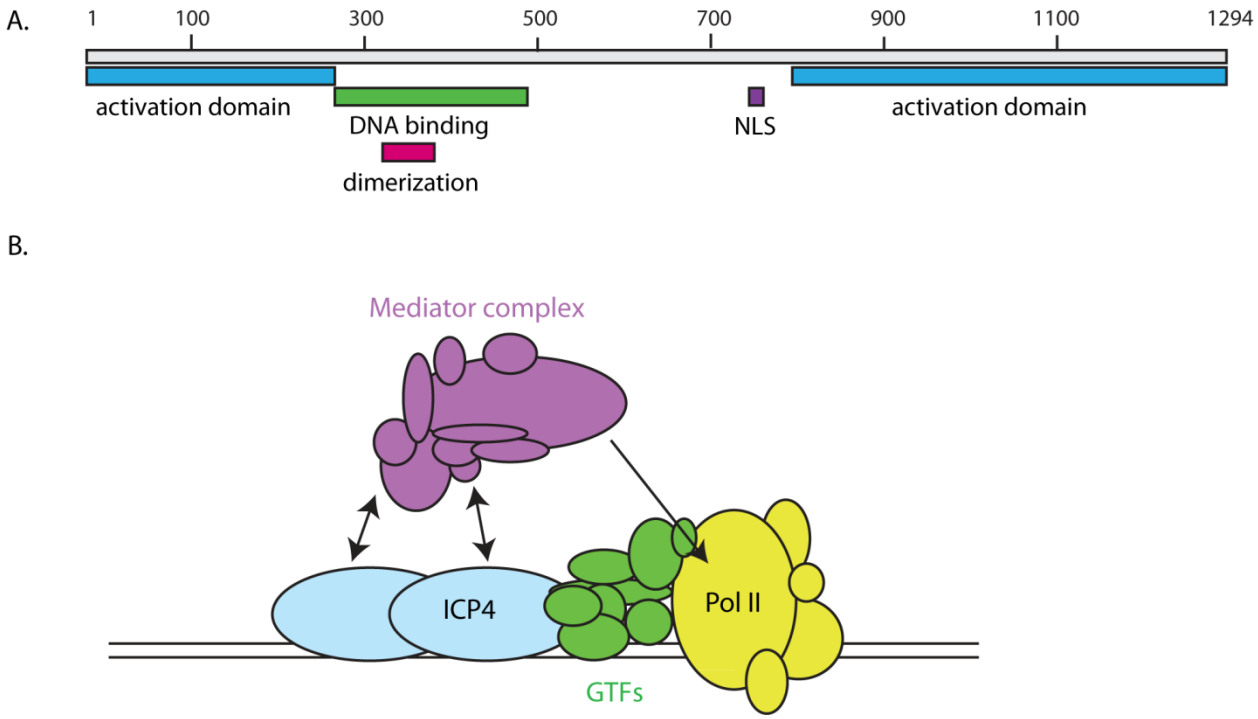

Figure 8: **ICP4 Structure and Function**
A. Domain schematic of ICP4 protein. B. Interactions of ICP4 with cellular transcription machinery, adapted from [152].

Once bound to DNA, ICP4 promotes transcription through interactions of its N and C terminal activation domains with multiple cellular transcription factors, most notably TFIID and mediator. Different mutant viruses with deletions within the ICP4 coding regions have been used to tease apart the functions of the two activation domains. When the 520 C-terminal amino acids are deleted, the virus is able to activate early, but not late gene expression [153]. The C terminal domain of ICP4 has been shown to interact with TFIID through the TAF1 subunit [154]. When the N-terminal 30-210 amino acids are further deleted, the resultant ICP4 is unable to activate early or late genes or repress immediate early genes [155]. Deletion mutants where just the N terminal region is deleted have a similar phenotype, indicating that the N terminal domain is a true and required activation domain. The N-terminal 210 amino acids are required for interaction with TBP, TAF1, and Med1, a component of the mediator complex [79]. The structure of the N-terminal regions is unstructured and elongated [156], which may enable it to form long range

contact with multiple transcription factors. While the C terminus is not a true activation domain, it may stabilize interactions of the N-terminal domain with transcription factors or enhance the binding of ICP4 to viral DNA to promote transcription.

In addition to its interactions with TFIID and mediator, ICP4 has also been shown to interact with several chromatin remodeling complexes including SWI/SNF, NuRD, and Ino80 throughout the lytic lifecycle [79]. While NuRD has been shown to be a repressor of gene expression, Ino80 and SWI/SNF can slide and eject nucleosomes to increase the accessibility of DNA, suggesting that ICP4 may function in maintaining an active chromatin state during active transcription.

1.3.3 ICP0

ICP0 has multiple functions in both lytic replication and reactivation that range from maintaining the viral chromatin in a euchromatic state permissive to transcription to disarming the cellular innate immune response. The gene encoding ICP0 is present in the genome twice and resides within the coding region on the latency associated transcript. ICP0 is a 775 amino acid long protein with several functional domains - most important of which is the ring finger domain which is required for its function. The domains of ICP0 are depicted below in Figure 9A. The ring finger domain has a characteristic fold seen in E3 ubiquitin ligases and is stabilized by coordination with 2 zinc atoms [157]. Mutations in the ring finger domain that abolish its activity create an identical phenotype to an ICP0 null mutant, indicating that this domain is necessary for its function. ICP0 is localized to the nucleus early in infection through a NLS around residue 500 [158]. The NLS is essential for ICP0 function in both lytic replication and reactivation from

latency [159]. ICP0 also interacts with the cellular protein USP7, which removes ubiquitin modifications from ICP0, enhancing its stability [160].

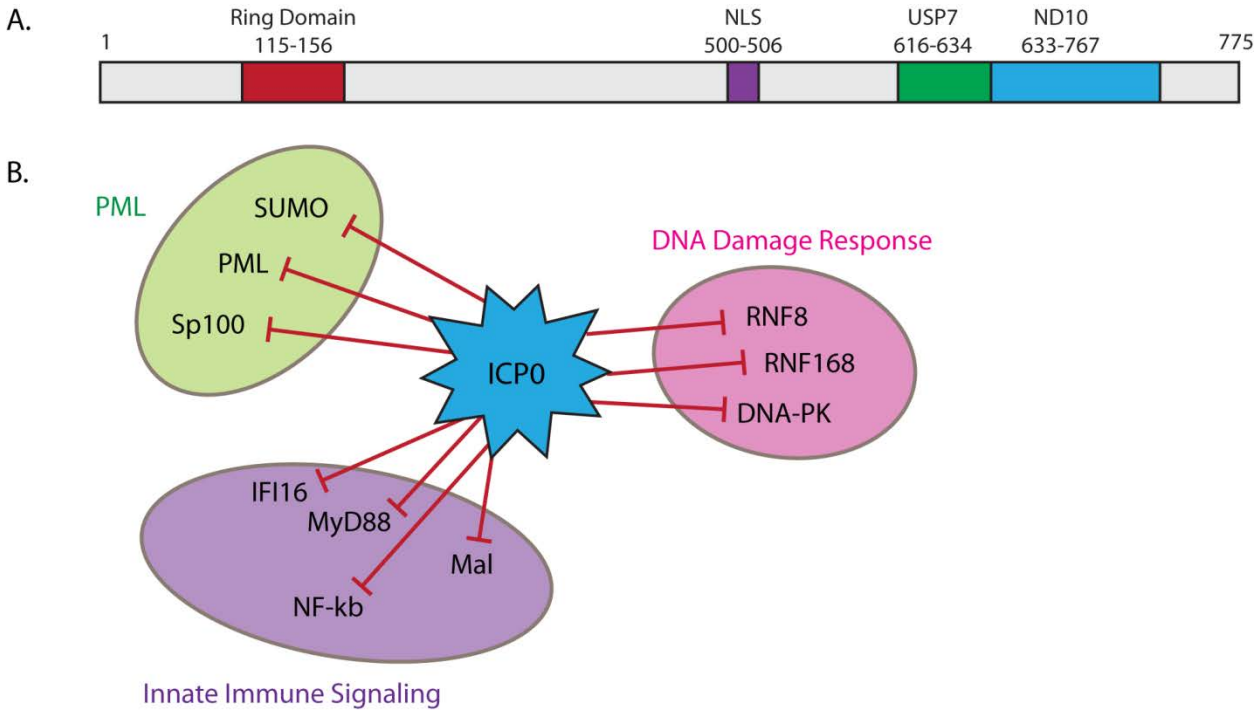

Figure 9: **ICP0 Structure and Function**
A. Domains of ICP0. B. Targets of ICP0 to counteract the host cell innate immune response on several fronts, adapted from [161]

Interestingly, ICP0 null viruses can replicate to different efficiencies in different cell types. U2OS cells are permissive to infection to an ICP0 null virus, enabling the virus to replicate to the same efficiency as wild type virus [162], but replication in human diploid fibroblasts leads to a reduction of plaque formation of approximately 1000 fold [163]. However, this difference in replication efficiency only exists at low MOIs; as a larger MOI is used, the disparity in replication deficiency becomes smaller, indicating that ICP0 promotes viral replication for low numbers of incoming viral particles and that restrictive host cell factors can be titrated out with increasing concentrations of virus. Several innate suppressors of HSV

replication have been identified including ND10 bodies, chromatin modifiers, and innate immune signaling molecules – all of which are targeted for degradation or inactivated by ICP0.

ND10 bodies, also referred to as promyelocytic leukaemia protein (PML) bodies, are small nuclear substructures that are defined by the presence of the PML. PML is the primary component of ND10 bodies and serves as the scaffold for their assembly [164]. PML is SUMO modified, which aids in the recruitment of other proteins to the complex [165]. Over 50 proteins have been shown to associate with ND10, but not all are localized constitutively [166]. Some of these proteins, most notably PML, Sp100, hDaxx and ATRX have been shown to mediate repression of HSV genomes. Furthermore, siRNA knockdown of these proteins has been shown to partially restore replication of an ICP0 null mutant, but did not affect replication of wild type virus. Sp100 is an interferon inducible protein and major component of ND10. All isoforms of Sp100 have been shown to interact with the heterochromatin binding protein HP1, a protein that has also been shown to recruit histone methyltransferases that target H3K9 residues [167]. Some isoforms of Sp100 have also been directly linked to repression of immediate early gene expression [168]. ICP0 targets Sp100 through interactions with the SUMOlyated residues with its SIM domain for proteosomal degradation through its ring finger domain [169].

In addition to Sp100, hDaxx and ATRX also are involved in the silencing of incoming HSV genomes. ATRX and hDaxx are constitutive components of ND10 and are localized through interactions of hDaxx with SUMO-modified PML [164]. These two proteins strongly interact and can be found in complex with several other chromatin modifying proteins. hDaxx has been shown the be a repressor of transcription and an interaction partner of histone deacetylase enzymes (HDACs) [113, 170]. hDaxx and ATRX have also been shown to restrict the replication of an ICP0 null HSV mutant, but not wild type virus indicating a role in

restricting viral replication. The interaction between these two proteins is required for this restrictive ability [112]. ICP0 has not been shown to destabilize these two proteins. However, it does alter their localization by inducing the degradation of the PML scaffold.

In addition to ND10 components, ICP0 targets several cellular proteins that are responsible for activating the host cell innate immune response, including IFI16, NF-κβ, and toll like receptor (TLR) signaling adapter proteins. IFI16 is a nuclear DNA sensor that has been shown to induce an IFN-β response to HSV-1 infection. IFI16 has also been shown to induce heterochromatin deposition, promoting the methylation of H3K9 and demethylation of H3K4 [171]. ICP0 promotes the relocalization of IFI16 to the cytoplasm and its subsequent proteosome mediated degradation. ICP0 also functions to sequester the downstream signaling protein IRF-3 early in nuclear compartments during infection to block IFN-β expression [172].

ICP0 also targets signaling adapters for TLRs. HSV glycoproteins are recognized by TLR2 and TLR4 which both reside in the plasma membrane. Signaling through these receptors requires the adapter proteins MyD88 and Mal and results in the activation of transcription factors NF-κβ and AP1. ICP0 targets both of these adapter proteins for degradation by the proteosome in a ring finger dependent manner, resulting in a decreased inflammatory response [173]. ICP0 also directly targets NF-κβ to block immune activation. The NF-κβ transcription factor is composed of two fragments, p50 and p65. ICP0 targets p50 for proteosomal degradation, preventing the nuclear translocation of p65 and thus blocking NF-κβ signaling [174].

ICP0 also blocks the deposition of repressive chromatin on the viral genome through interactions with HDACs. HDACs remove activating acetylation marks from chromatin, reducing transcription. ICP0 has been shown to directly interact with the N terminal region of class 2 HDACs. ICP0 does not appear to induce degradation of these proteins; rather it changes

their localization in the nucleus of infected cells and counteracts HDAC mediated transcriptional repression [175]. ICP0 also targets the CoREST/REST/HDAC1 repressor complex early in infection. ICP0 has a domain of approximately 76 amino acids that is homologous to the 80 C-terminal amino acids of CoREST. ICP0 induces the dissociation of HDAC1 from the complex [176].

Furthermore, ICP0 deactivates the DNA damage response, which has been recently shown to activate immune signaling in response to DNA viruses [177]. The DNA damage signaling cascade is mediated by three kinases, DNA dependent protein kinase (DNA-PK), ataxia telangectasia mutated (ATM) kinase, and ATM and Rad3 related (ATR) kinase. Infection with HSV leads to the activation of the ATM pathway [178]. ICP0 directs the proteosomal degradation of DNA-PK and both of the downstream signaling proteins of the ATM pathway, RNF8 and RNF 168 to block activation of this response [179, 180].

Through disarming multiple facets of the host cell innate immune response to viral infection, ICP0 promotes viral replication, especially when there is a low number of incoming virions. A summary of ICP0 targets is depicted in Figure 9B. ICP0 also functions to remove chromatin from the viral genome during reactivation, but that will be discussed in a later section.

1.4. LATENCY

1.4.1. Viral Entry into the Ganglia

Unlike infection of epithelial cells where the viral capsid and associated tegument proteins travel a short distance from the cell membrane to the nucleus, when the virus enters neurons through the axon termini, it must travel a much greater distance up the length of the axon to reach the soma and nucleus. This distance differential may affect the ability of the virus to enter productive infection. In one study, infection of neurons at the soma yielded a productive infection, where as infection at the axon termini largely yielded a quiescent latency-like state. However, the genome remained responsive to VP16 transactivation through delivery by helper virus, indicating that the viral genome transited to the nucleus, but failed to initiate lytic replication and virus production [181]. The inability of virus to enter lytic replication can be attributed to a loss of VP16 from the viral tegument as the viral particle travels up the axon, through the soma and into the nucleus. A study using gold labeled VP16 showed that as viral particles moved through the axon, VP16 was lost from the tegument and was dispersed in the axon and soma cytoplasm [182]. Furthermore, unlike in epithelial cells, the HCF protein which is required for VP16 function is localized to the cytoplasm [183]. The loss of VP16 from viral capsids during entry coupled with cytoplasmic localization of HCF may account for the less permissive state of neurons and the propensity for the virus to form a latent reservoir.

Furthermore, recent evidence has shown that a neuron specific micro RNA (miRNA), miR-138, targets the ICP0 transcript at two distinct sites, resulting in a decrease in ICP0 expression and lytic gene transcription. Deletion of these target sites in ICP0 resulted in an increase in viral gene expression and replication and viral spread in the ganglia leading to encephalitic infections in mice [184]. The neuron specific expression of this miRNA likely reduces lytic gene expression through decreasing ICP0 expression, limiting the initial round of lytic replication in the ganglia.

Latency is established in approximately 5-20% of the sensory neurons in the ganglia and the large majority of these neurons express the surface antigen recognized by the A5 antibody [185]. Specific elements in the LAT region may play a role in determining cell tropism for the establishment of latency [186]. During latency, the viral genome persists in an endless form, probably as a circular episome in the nuclei of infected cells and is repressed by cellular heterochromatin. The large majority of the genome is transcriptionally silent, except for the latency associated transcript and is correlated to this epigenetic state of local chromatin, discussed in the next section.

1.4.2. Chromatin Dynamics during Latency

Both euchromatic and heterochromatic marks have been identified on the viral genome during latency. Euchromatic modifications, including acetylated H3 at positions 9 and 14 as well as trimethylated H3K4, have been detected specifically within the LAT locus, partially explaining its high level of transcription during latency [187]. However, the LAT locus also has been found to have markers of heterochromatin, including trimethylated H3K9 and H3K27 and incorporation of the repressive histone variant macroH2A into nucleosomes [188]. This suggests that not all latent viral genomes are equal in their chromatin modification and may explain why only 1/3 of latently infected neurons express LAT.

While the LAT locus can be enriched in both euchromatin and heterochromatin, lytic genes have been shown to be exclusively associated with repressive heterochromatin during latent infection. These genes bear marks of both constitutive and facultative heterochromatin. Constitutive heterochromatin exists in tightly condensed 30nm fibers and is often considered an

irreversible form of heterochromatin as it cannot be reverted back to euchromatin. The other form of heterochromatin, facultative heterochromatin, is a more reversible form of chromatin and exists in multiple states of compaction. These two forms can be distinguished from each other based on the histone modifications made by different sets of chromatin modifying enzymes.

Constitutive heterochromatin is characterized by the di and trimethylation of H3K9. Studies have shown that H3K9 methylation increases in abundance on the viral genome in the trigeminal ganglia of mice as latency is established in vivo *[189]*. Furthermore, constitutive heterochromatin deposition on the genome has been demonstrated in quiescent models of latency. HSV viruses in which immediate early genes have been deleted enter a non-productive persistent infection in primary human fibroblast cells *[190]*. The genomes of these viruses become enriched in repressive chromatin marks including trimethylated H3K9 and heterochromatin protein 1 (HP1) *[191]*. HP1has been shown to drive condensation of chromatin through dimer formation bridging two nucleosomes *[192]*. HP1 also interacts with and recruits the H3K9 methyltransferase Suv39h1 to promote the spread of heterochromatin *[193]*. The location of the genome during latency may promote the deposition of constitutive heterochromatin. The viral genome is localized to ND10 bodies and centromeres of chromosomes during latent infection *[194]*. HP1 also interacts with the ND10 component Sp100 and is also localized to centromeric heterochromatin, possibly linking constitutive heterochromatin deposition to localization at these two areas *[167, 195]*.

Facultative heterochromatin is characterized by the presence of trimethylated H3K27 and the incorporation of the histone variant macroH2A. The trimethylation of H3K27 is mediated by polycomb repressive complex 2 (PRC2) and the maintenance of this modification as well as further compaction of chromatin is maintained by polycomb repressive complex 1 (PRC1). Both

trimethylated H3K27 and macroH2a have been observed on the genome during latency [196]. Additionally Suz12, a component of PRC2, has been shown to be bound to the viral genome during the establishment of latency, implicating PRC2 in the deposition of facultative heterochromatin [118]. Furthermore, Bmi1, a component of PRC1, occupies the viral genome during latency, suggesting PRC1 maintains facultative heterochromatin during latency [196].

Since there are separate regions of euchromatin and heterochromatin during latency, boundaries are necessary to prevent unwanted spread of heterochromatin into areas of euchromatin and vice versa. Insulator elements are DNA sequences that protect genes from signals emanating from their surrounding environment. Two main classes of insulator elements have been identified: enhancer blockers and barriers [197]. Enhancer blockers are located between enhancers and promoters and prohibit the action of an enhancer on a nearby promoter, enabling the enhancer to activate specific promoters without off target effects. The second class of insulator, a barrier, separates regions of different chromatin structure, reviewed in [198]. The CTCF protein was the first protein to be demonstrated to bind insulator elements in vertebrate genomes and function as an enhancer blocker/barrier [199]. Computational analysis of the human genome has identified nearly 15,000 potential CTCF binding sites, of which 13,804 have been confirmed through chromatin immunoprecipitation-microarray experiments [200, 201]. Nucleosomes near CTCF binding sites are well positioned and are modified with trimethylated H3K27 or acetylated H2AK5, depending on which side of the element the nucleosomes falls on [202]. Thus, CTCF plays an important role in maintaining chromatin by serving as a boundary between euchromatin and facultative heterochromatin.

Seven clusters of CTCF elements have been identified in the HSV-1 genome and have been shown to be occupied by CTCF protein during latent infection. One cluster lies upstream of

the LAT promoter near the unique long-repeat short junction. A second cluster, CTRL2, is located within the LAT intron downstream from the LAT enhancer and upstream of ICP0. Other clusters were found to be located within the "a" sequences, the repeat short region, and in the unique short region. The CTRL2 insulator element has been shown to function as an enhancer blocker and blocks the effects of the LAT enhancer on the ICP0 promoter [203]. Other labs have confirmed this finding and have also shown that CTRL2 can also function as a barrier when inserted into the *Drosophila* genome [204]. Together, these results indicate that insulator elements function to both separate areas of euchromatin and heterochromatin during latency and also block activation of ICP0 by the LAT promoter.

The latency associated transcript has been implicated in the regulation of chromatin modification during latent infection, but the effect differs between viral strains. Deletion of the LAT promoter or sequences within the transcript led to decreased H3K27 methylation in the KOS strain [188]. The LAT transcript may function to target PRC2 to lytic genes during latency through interaction with EZH2. Non-coding RNAs such as Xist have been shown to mediate targeting of PRC2 to genes through interaction of the EZH2 subunit with RNA [205].

However, similar deletions in the strain 17 led to an increase in H3K27 methylation [196]. This discrepancy warrants further investigation to determine if the difference is due to the site of latency, as Cliffe studied the TG [188] and Kwiatkowski [196] studied the DRG as latent models, or the specific strain, since these viruses do differ in their pathogenicity.

1.4.3. Gene Expression and the Latency Associated Transcript

During latency, the main transcript expressed from the virus is the latency associated transcript (LAT) while the expression of other genes is heavily attenuated. The LAT locus is located within the inverted repeat long region and this is present in the genome twice. It runs antiparallel to ICP0, γ34.5 and ICP4. The LAT is expressed as a 8.5kb polyadenylated transcript, consisting of two exons and one intron, and can be spliced to form a stable 2kb intron. The 2kb intron can be further spliced to 1.5kb in some neurons [206]. The predominant form of the LAT transcript is the 2kb intron since it is incredibly stable, with a half life of 24 hours due to its lariat structure [207]. The LAT transcript has been shown to have multiple functions in latency including repressing lytic gene expression through its encoded miRNAs, preventing cell apoptosis, and establishing and reactivating from latency.

To date, 17 mature microRNAs originating from the virus have been identified in infected cells. Of these 17 microRNAs, only 9 of them have been shown to be loaded into the RISC complex, the complex responsible for targeting miRNAs to their respective targets. These microRNAs include miR-H1-H8 and miR-H11 [208]. Many of these functional microRNAs are located within the LAT region, suggesting that LAT functions as a microRNA precursor, reviewed in [209]. MicroRNAs originating from LAT include H1-5, H7, H8, H14 and H15. Many of these LAT encoded microRNAs have been shown to repress lytic gene expression to maintain latency. MiR-H2 has been shown to downregulate ICP0 expression in cell culture, but does not target other immediately early genes [210, 211]. Furthermore, in vivo studies in mice showed that when this miRNA was knocked out, neurovirulence and the frequency of reactivation was enhanced [212]. These results suggest that miR-H2 functions to decrease neurovirulence to promote survival of the host and maintain latency through targeting ICP0.

1.5. REACTIVATION

HSV can reactivate from latency following stress to the host. Common stressors include UV exposure, psychological stress, and immune suppression [213]. It has also been known for over 100 years that severing the trigeminal nerve leads to reactivation of HSV [214]. Until recently however, the molecular mechanism underlying this observation was unknown. Reactivation induced by cutting the trigeminal nerve or explanting the ganglia proceeds via a similar mechanism – disruption of the NGF signaling pathway which maintains latency. NGF signals through the TrkA receptor and can activate multiple pathways. However, only the PI3-K signaling pathway was shown to be necessary to maintain latency, as inhibitors targeting this pathway but not others resulted in reactivation of virus. Even transient disruption of this signaling pathway was sufficient to induce reactivation. Active PI3-K recruits PDK1 to the plasma membrane, leading to its phosphorylation and the phosphorylation of Akt. The duration of signaling through Akt was found to be directly correlated to maintenance of latency, with NGF being the most potent activator of long lasting activation of the pathway. However, the downstream targets of Akt were not identified in this study [215]. The cellular mTOR kinase is a target of the PI3-K Akt pathway and has been implicated in the maintenance of latency. Inhibition of this protein or raptor, its regulatory subunit responsible for complex assembly and substrate specificity, led to reactivation of latent virus. mTOR functions to inhibit 4E-BP, a translational repressor, suggesting that proteins whose translation is impeded by 4E-BP may be involved in controlling latency [216]. Additional studies are needed to identify the proteins targeted by this protein involved in the reactivation pathway.

Disabling the NGF signaling pathway through inhibition of PI3-K leads to a biphasic transcription pattern during reactivation. The first phase occurs between 0 and 20 hours following the reactivation stimulus and is marked by a transient induction of global gene expression. During this phase, genes from all kinetic classes are expressed independent of protein synthesis and DNA replication, as gene expression was similar with and without cyclohexamide or PAA treatment. Coincident with this gene expression is the cytoplasmic accumulation of VP16 protein. Between 20 and 25 hours, gene expression declines and is followed by a second phase of expression. The second phase begins with a relocalization of VP16 and HCF to the nucleus, leading to the formation of the VP16 Oct-1 HCF transcription complex which stimulates the expression of immediate early genes. Interestingly, inhibition of the NGF signaling pathway induces the nuclear translocation of HCF which resides in the cytoplasm under normal conditions. During the second phase of transcription, gene expression is dependent on DNA replication, indicating this pattern is more similar to the lytic cascade. It is also during this phase when infectious virus is produced [217].

Additional studies have also confirmed this observation. Early following reactivation stimulus, lytic gene expression is induced independent of protein expression. Concurrently, expression of the LAT and its associated miRNAs also declined. The accumulation of viral lytic transcripts and continued expression of LAT was mutually exclusive. Also, unlike the lytic transcripts, the decrease in LAT expression is dependent on protein expression, suggesting that a viral protein may be involved in the degradation of the LAT transcript during reactivation [218]. Since the burst of transcription during reactivation was independent of protein production, it is likely mediated my existing cellular proteins, possibly including chromatin remodeling complexes.

During reactivation from latency, the viral genome must transition from a heterochromatic to euchromatic state to enable lytic viral transcription to resume. Several cellular chromatin remodeling complexes have been implicated in removal of repressive chromatin during reactivation. In an *in vitro* model of latency using anti NGF antibodies to induce reactivation, chemical inhibition of the two H3K27 demethylases JMJD3 and UTX resulted in decreased expression from various kinetic classes of viral transcripts along with a decrease in the production of infectious virus. This decrease in transcription and viral titer was correlated to high levels of trimethylated H3K27 on the viral genome, indicating that removal of facultative heterochromatin from the viral genome by JMJD3 and UTX was required for maximal reactivation [219]. Furthermore, in *in vivo* models using explant induced reactivation, specific inhibition of LSD1 led to a significant reduction in viral yield compared to mock treated controls, indicating that LSD1 is also important for reactivation of the virus [220]. These two studies taken together suggest a role of cellular chromatin remodeling complexes in removing repressive chromatin from the viral genome to enable reactivation of the virus.

In addition to changes in histone modification, there are also changes in CTCF occupancy on the genome during reactivation. As discussed in the previous section, CTCF elements are located around the LAT, ICP0, and ICP4 genes. During latency, these sites are occupied by CTCF and serve as enhancer blockers, preventing the activation of lytic genes by the LAT enhancer. Following a reactivation stimulus, CTCF occupancy at these sites rapidly decreased, with a 50% reduction within the first hour. Furthermore, the loss of CTCF from some sites during reactivation was dependent on IE gene expression, indicating that IE genes may contribute to the loss of CTCF [221]. The disruption of these boundaries may enable activation

of ICP0 by the LAT enhancer or the spread of euchromatin from the LAT region to other areas of the genome.

Additional studies have focused on determining which viral activators are sufficient to induce reactivation in quiescent models of latency. In quiescent models of latency, a virus lacking immediate early genes is repressed in cell culture and acquires chromatin modifications similar to *in vivo* latency [191]. Studies using a B-galactosidase reporter virus showed that expression of ICP0, but not an ICP0 mutant lacking the ring finger domain, was sufficient to induce expression of the reporter gene [159]. Further studies showed that provision of ICP0 to quiescently infected cells leads to the derepression of silencing chromatin, including removal of H3 and H4, reduction in H3K9 methylation, and increases in histone acetylation [222]. Later studies showed that the ring finger domain of ICP0 was required for the ability to induce changes in chromatin [223].

However, cellular stress has been shown to be sufficient to reactivate quiescent genomes in the absence of ICP0. Treatment of cells with either sodium arsenite - a general enzyme inhibitor - or gramicidin D - a pore forming toxin that renders the membrane permeable to monovalent cations - led to reactivation, measured by the expression of a reporter gene. Furthermore, toxin treatment induced the expression of cell stress proteins and led to a reduction in PML foci, suggesting a link between cellular stress and PML mediated repression of viral genomes, as viral genomes are localized to PML bodies during quiescence [111, 224]. These results are consistent with another study that showed cellular stress enhanced the replication of ICP0 null viruses in cell culture [225].

Overall, there are multiple factors that contribute to reactivation from latency including NGF signaling, cellular stress, chromatin remodelers, and viral factors. As there are many

mechanisms to induce latency, the exact proteins required for reactivation may change based on the reactivation stimulus.

1.6. FORMS OF CHROMATIN

1.6.1 Nucleosomes

The basic building blocks of chromatin are nucleosomes, which consist of the two copies of each of the core histone proteins – H3, H4, H2A and H2B - DNA, and sometimes the linker histone H1. Histone proteins are relatively conserved from yeast to higher mammals [226]. During formation of a nucleosome, 146-147 base pairs of DNA are wrapped around the histone octamer [227]. The nucleosome unit is a disc-like structure 5nm tall and 11nm wide [228]. Histone proteins are small, only 11-15kDa, and are basic, enabling association with the negatively charged DNA. Each core histone protein, although varied in amino acid sequence, is composed of two domains, a globular domain rich in alpha helices which serves as the site of histone-histone and histone-DNA interactions, and an unstructured tail domain which can be modified by various chromatin modifying enzymes to augment the structure of chromatin [229].

The histone proteins interact to from dimers, tetramers and finally octamers. The histone folds in the globular domain mediate the formation of heterodimers of H3-H4 and H2A-H2B. H3-H4 dimers then interact to form stable tetramers through interactions along the face of H3. Once the tetramer has formed, H2A-H2B dimers associate through interactions between the faces of H2B and H4. However, the interaction between H2B and H4 is much weaker than the

interaction between H3-H3 in tetramers, so the full histone octamer is not formed unless in contact with DNA or in high salt concentration. [230]. Nucleosomes evenly spaced on a DNA template make up the 11nm fiber [231]. The linker histone H1 binds to the histone octamer to protect DNA between neighboring nucleosomes and aids in the formation of higher order chromatin structures by compacting the chromatin into 30nm fibers [232].

There are several forms of chromatin including euchromatin, constitutive heterochromatin and facultative heterochromatin. These forms differ in their accessibility to transcription factors and in the modifications present on their histone tails, deposited by different chromatin modifying enzymes. The histone tail modifications also serve as docking sites for different proteins that may further alter the chromatin structure. A general overview of the structure of these three forms of chromatin is presented below in Figure 10. Each form of chromatin will be discussed in greater detail in the following sections.

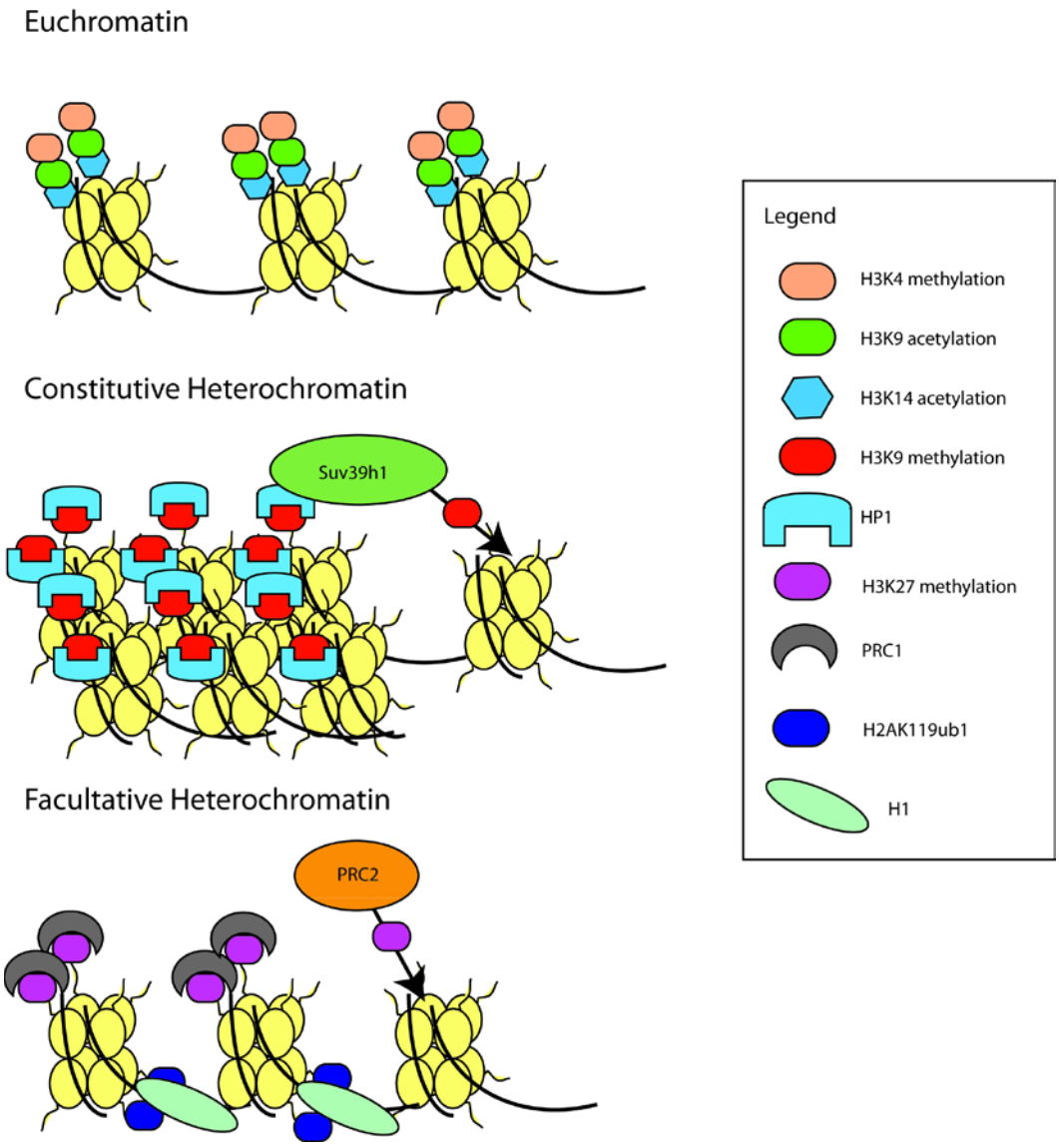

Figure 10: Forms of Chromatin
A. Euchromatin is rich in histone acetylation and H3K4 methylation. The acetylation neutralizes the positive charge of the lysine residues, decreasing the affinity of DNA for the histone proteins. For this reason, euchromatin is more open and permissive to transcription. B. Constitutive heterochromatin is characterized by densely packed nucleosomes rich in H3K9 methylation. Due to the compact structure, regions of the genome occupied by constitutive heterochromatin are not expressed. HP1 recognizes and is recruited to H3K9 trimethylation and aids in the further compaction of chromatin and spread of H3K9 methylation through recruitment of the methyltransferase Suv39h1. C. Facultative heterochromatin is a silencing form of chromatin, but can transition back to euchromatin. It is enriched in H3K27 trimethylation, deposited by PRC2 and H2AK119 ubiquitination, deposited by PRC1.

1.6.2. Euchromatin

Euchromatin has a looser, more unfolded and less structured conformation than heterochromatin to allow access to transcription factors access to DNA, making it the transcriptionally active form of chromatin. For this reason, actively transcribed genes tend to be enriched in euchromatin. Euchromatin is characterized by wide spread histone acetylation and specific activating methylation marks. Acetylation of H3 at positions K9 and K14 along with trimethylated H3K4 are enriched at the transcription initiation site of eukaryotic genes, but these modification are greatly decreased downstream of the transcription start site [233]. Additional studies have also shown that the acetylation of H2AK9, H2BK5, H3K18, H3K36 and H4K91 are also localized to transcription start sites, while acetylation of H2BK12, H2BK20, H2BK120, H3K4, H4K5, H4K8, H4K12, and H4K16 are found both in the promoter and transcribed regions of active genes [234]. Acetylation of histones has been shown to enhance the binding of transcription factors to nucleosomal DNA in vitro [235, 236]. In vivo studies have been less fruitful in establishing a level of acetylation required for efficient transcription factor binding since there are redundancies in acetyltransferase function and knockdown of several factors is detrimental to cells.

The acetylation of histone tails is the main determinant of compaction, as the addition of an acetyl group to lysine neutralizes the positive charge and disrupts electrostatic interactions between the histones and negatively charged backbone of DNA, but does not directly lead to unfolding of individual nucleosomes [237, 238]. There are small increases in temporary unwrapping of DNA near the edges of hyperacetylated nucleosomes [239]. The main change in chromatin upon histone acetylation is the formation of higher order structures. Acetylation of histones leads to the adoption of an extended chromatin structure [240]. One possible mechanism

for this change is the weakening of internucleosomal contacts. H4 makes extensive contacts with the H2A-H2B dimer of neighboring nucleosomes, and the amino acids of H4 that mediate this interaction are subject to acetylation [241]. Specifically, acetylation of H4K16 inhibits the formation of the compacted 30nm fiber [242]. However, histone acetylation alone is not sufficient to overcome nucleosome blockage during transcription elongation, suggesting additional factors are involved [243].

In addition to preventing the formation of higher order chromatin structure, histone acetylation can also enhance the function chromatin remodelers, including the Swi/Snf complex. The Swi/Snf complex is a chromatin modifying ATPase that uses energy from ATP hydrolysis to disrupt and slide nucleosomes on DNA [244]. This complex contains a bromodomain which specifies binding to acetylated histones [245]. Studies have shown that enhanced action of chromatin remodelers on acetylated templates is mainly due to an increase in affinity, not enhanced catalytic function [246].

The concerted action of acetyl transferases along with the recruitment and action of histone remodeling complexes maintains euchromatin in a loose structure that is permissive to gene expression.

1.6.3. Constitutive Heterochromatin

Constitutive heterochromatin is localized mainly to telomeres and pericentromeric regions of chromosomes, which are usually gene poor areas consisting of tandem repeats or satellites that vary in size from 5 to a few hundred base pairs [247]. This form of chromatin is very densely packed and transcriptionally inert. The most prominent histone modification is the trimethylation

of H3K9, which is deposited in a sequential manner by several histone methyl transferases. The first step is the monomethylation of H3K9 which is mediated by three enzymes Prdm3, Prdm16, and SETBD1 and may occur prior to the deposition of histones onto DNA [248, 249]. The two isoforms of Suv39h1/2 mediate the di- and trimethylation of H3K9 specifically at pericentromeric sites, as knock down of these two proteins leads to a decrease in H3K9 methylation at pericentromeric regions, but not other locations on chromosomes, suggesting a specific targeting mechanism [250]. Studies in *Schizosaccharomyces pombe* have implicated RNAi in directing H3K9 methylation [251], but how Suv39h proteins are targeted initially to chromatin in mammals is not clear, but the mechanism for their maintenance on constitutive heterochromatin is mediated by interactions between their chromodomain and H3K9me3. Trimethylated H3K9 also serves as a docking site for heterochromatin protein 1 (HP1), an interaction partner of Suv39h [252]. HP1 can bind to two separate H3K9me3 residues and has the ability to dimerize, linking to nucleosomes occupying different loci to promote compaction of chromatin [253].

Another hallmark of constitutive heterochromatin is the di- or trimethylation of H4K20, however, the function of this specific modification in silencing is unclear. The trimethylation of H3K9 is a prerequisite for this modification, as HP1 has been shown to interact with and recruit Suv420h, the enzyme responsible for this modification. Suv420h has been shown to be involved in the binding of cohesins to heterochromatin, suggesting a role in sister-chromatid cohesion and chromosome segregation during mitosis [254, 255]. Additional modifications including H3K27 monomethylation and H3K64 trimethylation have been observed on constitutive heterochromatin, but their function is not yet clear. The methylases PRC2 and G9a have been suggested to methylate H3K27, but have not been proven.

Differential incorporation of histone variants has also been observed at sites of constitutive heterochromatin. The swi/snf chromatin remodeler ATRX has been found to be enriched at pericentromeric heterochromatin and to recruit Daxx in a cell cycle dependent manner [256, 257]. These two proteins are involved in the deposition of histone variant H3.3 in repetitive regions. This incorporation may serve to help transcription, which is correlated to the recruitment of H1 and is critical for the formation of heterochromatin in developing mice [258]. The other histone variant is H2A.Z [259], which contains a unique acidic patch located on its surface. This region binds to the tail of H4 with greater affinity and likely contributes to enhanced chromatin folding. Furthermore, HP1 is recruited to H2A.Z containing nucleosomes and stabilizes the interaction of H2A.Z with H4 of neighboring nucleosomes to further promote the packing of chromatin into dense fibers. HP1 has a 2.5 fold higher affinity for H2A.Z than H2A suggesting that different heterochromatic regions containing exclusively one variant over the other may differ in their structure [260]. H1 and its variant H5 are enriched and function to promote higher order structures by clamping DNA fibers to nucleosomes. H5 tends to be more enriched with highly compacted and repressed chromatin [261].

Through the trimethylation of H3K9 and recruitment of HP1, constitutive heterochromatin adopts a highly condensed state which is refractory to transcription.

1.6.4. Facultative Heterochromatin

Like constitutive heterochromatin, facultative heterochromatin is also transcriptionally silent, but retains the ability to transition back to euchromatin to allow transcription in different contexts. This transition can occur during: i) temporally regulated processes such as development or cell

division, ii) changes in localization in the nucleus, and iii) expression of specific genes. The mechanisms regulating this transition are still under investigation, but include changes in histone components, chromatin modification, action of cellular proteins, and nuclear localization [262]. For this reason, facultative heterochromatin is differentially localized in different cell types, depending on which genes must be silenced at a particular time. This form of chromatin is also present in the inactive X chromosome.

Facultative heterochromatin can adopt a wide variety of structures, ranging from the confined compaction of the 11nm fiber to the condensed 30nm fiber and different forms of higher order chromatin structure. The most common model to study facultative heterochromatin structure and function is the inactive X chromosome. Studies of this chromosome have shown the heterochromatin has a unique ultra structure - a porous structure containing chromatin substructures with intervening spaces ranging in length from 30-400nm. The level of compaction is between that of constitutive heterochromatin and euchromatin. Regions of facultative heterochromatin are localized to the nuclear periphery in structures known as Barr bodies, which is correlated to repression of transcription [263, 264].

X inactivation has also shown sequential changes in histone modifications and histone variant incorporation into nucleosomes. These changes are diagramed in Figure 11. The Xist RNA has been shown to direct the inactivation of the X chromosome. The earliest changes in the formation of facultative heterochromatin include the loss of euchromatin markers including H3K9 acetylation and H3K4 di- and trimethylation followed quickly by H4 hypoacetylation [265, 266]. The exact mechanism underlying the loss of these modifications has not yet been clearly elucidated. Following the loss of euchromatic marks, repressive marks are deposited including H3K27 trimethylation, H3K9 dimethylation, H4K20 methylation, and H2A

ubiquitination [267, 268]. However, these modifications are not consistent throughout the inactivated x chromosomes as there are differences in their distribution [269].

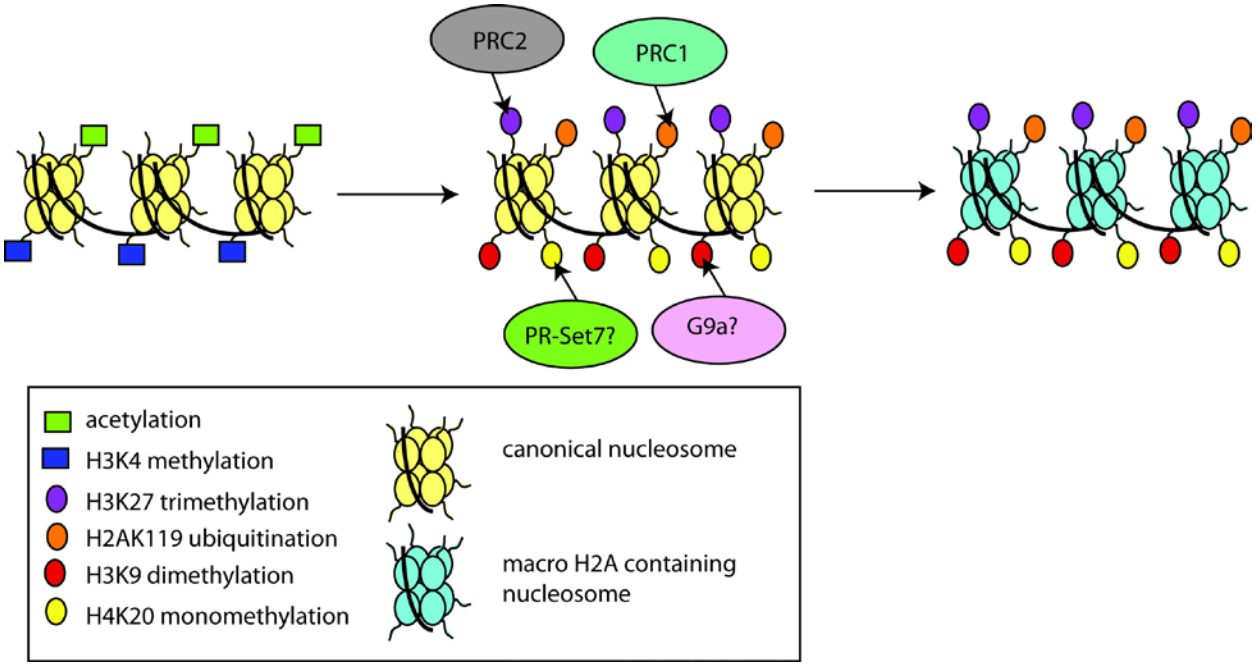

Figure 11: Formation of Facultative Heterochromatin
Summary of the changes in histone modifications during the transition of euchromatin to facultative heterochromatin, adapted from [270]

Many of the enzymes responsible for these modifications have been identified, or likely candidates are suspected. The methyltransferase responsible for H4K20 mono-methylation is unclear, but PR-Set7 is a strong candidate [271]. G9a is also suspected to be the H3K9 methyltransferase, but this has not been experimentally proven. However, the proteins responsible for H3K27 methylation and H2A ubiquitination have been identified as the two polycomb complexes. PRC2 is recruited to the DNA through Xist RNA mediated targeting and methylates H3K27. These methylation marks serve as a docking site for PRC1 which ubiquitinates H2A at lysine 119. These two complexes are reviewed in greater detail in later sections.

The histone variant macro H2A is recruited to facultative heterochromatin through interactions with the Xist RNA [272]. The properties of nucleosomes formed in the presence of macro H2A differ from those of the canonical nucleosome. Histone octamers containing macro H2A are more stable than canonical nucleosomes in the absence of DNA due to more internucleosomal contacts and are resistant to chaperone assisted H2A-H2B dimer exchange [273, 274], making them more repressive.

1.7. CHROMATIN REMODELING COMPLEXES

As there are many chromatin remodeling and modifying complexes, this section will focus mainly on the complexes that have been shown to be or hypothesized to be responsible for modulating chromatin dynamics on the HSV genome during lytic and latent infection.

1.7.1. Histone Acetyltransferase Complexes

CBP/p300

CBP and its paralog p300 have overlapping function as they often play similar roles in the cell, but sometimes have distinct functions. Since their functionality is so similar, they are often collectively referred to as CBP/p300 [275]. These proteins are important transcription coactivators that function to both recruit additional coactivators and remodel chromatin. They share similar domains, referred to as transactivation domains, which mediate their interactions with transcription factors, components of the general transcription machinery, and other

coactivators. The presence of 4 TADs enables CPB/p300 to simultaneously interact with multiple proteins. CBP/p300 also has a HAT domain that enables it to acetylate both histones and other proteins, as well as a bromodomain which mediates its interactions with acetylated histones [276]. Figure 12 shows some of the interaction partners of CBP/p300 and how it can bridge transcription factors to the general transcription machinery.

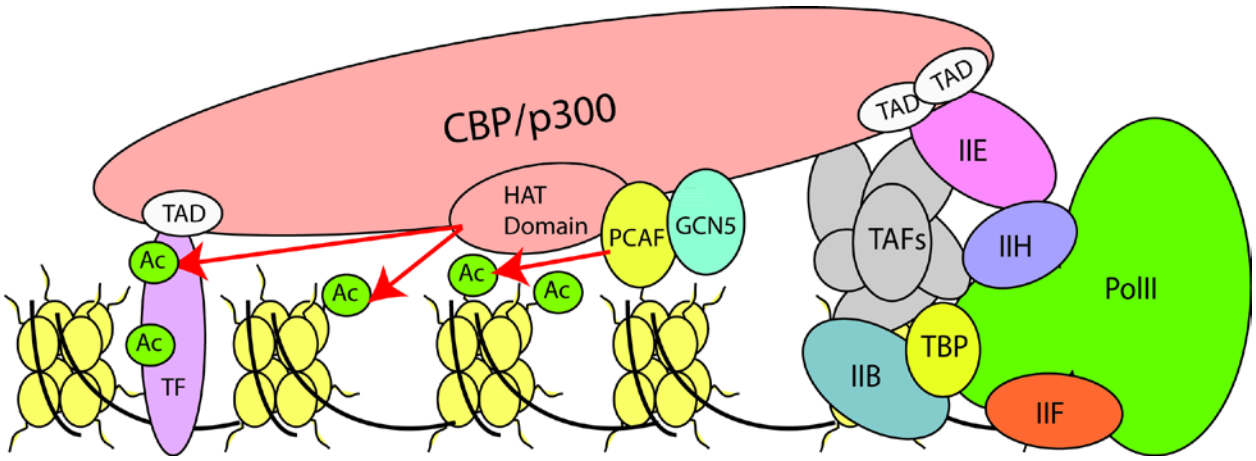

Figure 12: CBP/p300 Interactions
Functions of the CBP/p300 complex include bridging transcription factors (TFs) to the general transcription machinery and recruiting the coactivators PCAF and GCN5. CBP/p300 also has its own acetyltransferase activity and can acetylate both histones and transcription factors to augment their function, adapted from [277].

CBP/p300 has been shown to be part of the RNA PolII holoenzyme, and the association is enhanced by the presence of RNA helicase A [278, 279]. Depletion of CBP/300 from the holoenzyme decreased its association with TFIIB and TBP, indicating that it may provide binding sites for these proteins [280]. CBP/p300 may also form transient interactions with TFIIE and TFIIF to mediate their acetylation [281]. It also interacts with and recruits the HATs PCAF and GCN5, which are more diverse in their structure and substrate specificity. These two proteins often synergistically act with CBP/p300.

In addition to being a protein interaction hub, CBP/p300 possesses its own intrinsic HAT enzymatic activity and will acetylate histones in its general vicinity [282]. It will acetylate all possible residues on histones 2A and 2B and will specifically acetylate K14, K18 and K27 on H3 as well as K5 and K8 of histone 4 [283, 284]. It also targets H3K56, but acetylation of this residue is associated with DNA compaction rather than relaxation of chromatin structure [285]. CBP/p300 also acetylates non-histone substrates such as transcription factors to either enhance or decrease their activity. The effect of acetylation varies between transcription factors and is reviewed in [277].

CLOCK

The CLOCK protein is most well known as the master regulator of circadian rhythms, but it has also been shown to be involved in the HSV lifecycle. During lytic infection, CLOCK is recruited to ND10 bodies where incoming genomes localize and facilitates viral transcription through modification of viral chromatin [131]. ICP0 has been shown to directly interact with BMAL1, an interaction partner of CLOCK, and form a transcriptionally active complex, capable of enhancing expression from a reporter gene [286]. HSV infection has been shown to stabilize the CLOCK protein, as it usually undergoes rapid turnover in a proteosome dependent manner [131]. The CLOCK protein has intrinsic acetyl transferase activity which is further enhanced by the presence of its interaction partner BMAL1. This protein heterodimer acetylates histone 3 [287]. MLL1, a H3K4 histone methyltransferase, also associates with CLOCK and is recruited to target genes in a circadian dependent manner [288].

1.7.2. Histone deacetylase complexes

Histone deacetylases 1 and 2 (HDAC1 and HDAC2) are components of multiple deacetylase complexes including the CoREST and NuRD. These two proteins cannot bind to DNA independently and likely require interacting partners for catalytic activity [289]. CoREST has been implicated in silencing incoming HSV genomes by removing activating acetyl and methyl marks [120]. NuRD has not been directly implicated in the HSV lifecycle, but has been shown to co-purify with ICP4 during immunoprecipitation experiments in lytic infection [79]. A schematic showing the protein components of each of these complexes is depicted below in Figure 13. Both of these complexes will be discussed in greater detail.

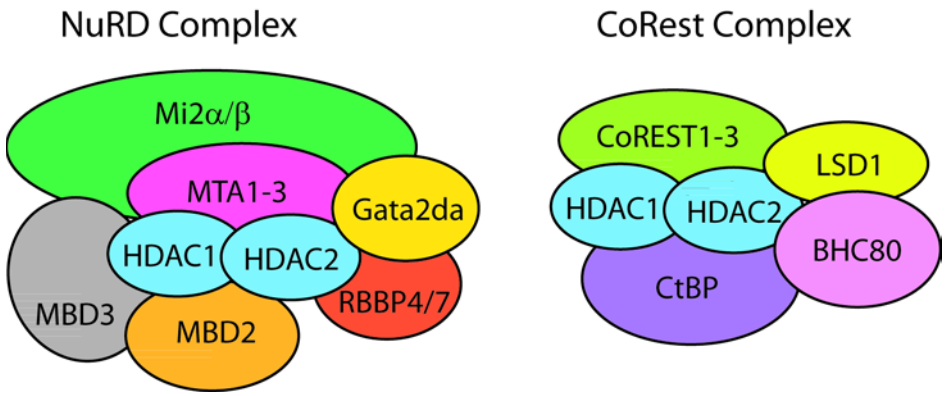

Figure 13: **HDAC Complexes**
Protein components of the two HDAC1/2 containing complexes NuRD and CoREST, adapted from [290]

NuRD Complex

The NuRD complex possesses both deacetylase activity and ATP dependent nucleosome remodeling activity. The Mi2α/β proteins, also known as CHD3/CHD4, are the ATPase subunits and can function as ATP dependent chromatin remodelers independent of their association with the NuRD complex, but less efficiently [291]. These proteins also have PHD chromodomains

that mediate their interaction with histone H3 tails. Furthermore, the modification of H3 tails affects the affinity of the PHD domain; trimethylated H3K9 increases the affinity, while trimethylated H3K4 abolishes it [292, 293]. The chromodomains have also been shown to modulate the ATPase activity by steric hindrance, either by blocking the DNA binding domain or ATP binding domain, providing a mechanism for control. Structural rearrangements in CHD4 enable catalytic activity [294]. CHD3/4 is also in direct contact with HDAC1 and likely serves as a scaffold for complex assembly.

The MBD2 and 3 proteins are members of the methyl cytosine-guanine-binding domain family and bind to methylated DNA. MBD2 strongly associates with methylated DNA, whereas the affinity of MDB3 is much lower [295]. MBD2/3 can also mediate the targeting of the NuRD complex to different subsets of genes [289]. RBBP4 and 7 are WD repeat histone chaperone proteins that bind to histone H4. When bound to RBBP4/7, histone H4 undergoes unfolding to be accessible, which may be essential for ATP dependent remodeling [296]. The MTA proteins have been shown to interact directly with transcriptional repressors and the estrogen receptor to mediate gene silencing [297]. The GATA2 subunit is involved in binding to unmodified histone tails [298].

REST CoREST Complex

The REST-CoREST complex has both deacetylase activity mediated by HDACs1 and 2, as well as demethylase activity driven by the histone demethylase LSD1. This complex is distinctly different from the NuRD complex in that is lacks the RBBP4/7 subunits. The CoREST protein directly interacts with HDACs 1 and 2 through its N terminal domain, which is required for transcriptional repression [299]. The CoREST protein also appears to be the scaffold protein for

complex assembly as it also interacts with the REST and LSD1 subunits [300]. The interaction of CoREST with LSD1 also stimulates its H3K4 demethylase activity [301]. Together, LSD1 and the HDACs make up the catalytic component of the complex. The other proteins associated with the complex are involved in targeting and activation of the complex.

The BHC80 subunit binds to unmethylated H3K4 through its PHD domain, specifically interacting with the first 21 residues of histone 3, but not any other the other histone proteins. BHC80 directly interacts with LSD1 [302], and is required for the targeting of LSD1 to unmethylated histones, as H3K4 methylation greatly decreases the affinity of BHC80 to histone 3 [303]. Less is currently known about the function of CtBP subunit and its requirement for activity of the complex. CtBP is a binding partner of multiple transcription factors and has been shown to direct the repression of target genes through the recruitment of HDACs, possibly implicating it in the targeting of the complex, reviewed in [304].

1.7.3. Histone Methyltransferases

Set 1 Methyltransferase

Set1 is the H3K4 methyltransferase that is responsible for methylating histones on the HSV genome during lytic infection [128]. A schematic of the core Set1/MLL1 complex is shown below in Figure 14. Both the catalytic Set 1 and MLL1 components can interact with HCF-1, an interaction partner of VP16, which recruits them to immediate early promoters where they function to promote immediate early gene expression.

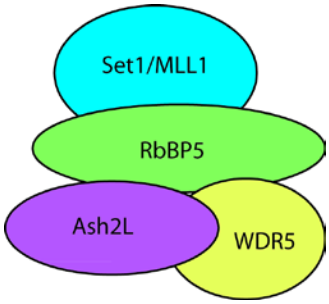

Figure 14: Set1 Methyltransferase Complex
Protein components of the Set1/MLL1 histone H3K4 methyl transferase, adapted from [305]

The non-enzymatic core components include AshL2, WDR5 and RbBP5, which form a structural scaffold independent of the catalytic components and but can interact with and recruit different catalytic subunits. Both RbBP5 and WDR5 are essential for the formation of stable complexes with MLL proteins. Additionally, RbBP5 plays a critical role in the organization of the complex as it interacts with both Ash2L and WDR5, whereas Ash2L is not necessary for the formation of the complex. All three core components are necessary for maximal enzymatic action of the histone methyltransferase complex [306]. WDR5 has been shown by crystal structure to interact with H3 and is likely involved in the targeting of the complex to nucleosomes [307]. Point mutations of this protein have shown it to have two functions: maintaining structural integrity of the complex through interactions with RbBP5 and MLL and directly contacting H3 for complex recruitment [306].

Suv39h1

Suv39h1/2 is one of the main H3K9 methyltransferases and is responsible for the formation of pericentromeric and telomeric heterochromatin. The human isoforms of this protein were first discovered in 2000 and were shown to have two separate domains, the chromodomain and SET domain. In general, chromodomains mediate the targeting of chromatin modifiers to euchromatin

or heterochromatin [308]. The SET domain of Suv39h1 is the catalytic domain, is essential for H3K9 trimethylation, and is specific only to lysine 9 of H3. Additionally, existing histone modifications, especially H3K9 acetylation and H3S10 phosphorylation, can block H3K9 methylation [309].

Immunoprecipitation studies of Suv39h1 have shown that it is part of a larger complex containing the additional H3K9 methyltransferases G9a, GLP, and SETDB1, as well as two isoforms of HP1, HP1β and HP1γ. The formation of this complex could allow for sequential methylation for the establishment of constitutive heterochromatin where SETDB1 or G9a/GLP dimethylate H3K9, creating a substrate for further methylation by suv39h1 [310]. Additionally, the binding of Suv39h to HP1 increases its affinity for trimethylated H3K9, which would increase the affinity of the complex for heterochromatin to enhance targeting [311].

Polycomb complex 1 and 2

The polycomb repressive complexes are responsible for creating and maintaining facultative heterochromatin through the trimethylation of H3K27 and compaction of chromatin. PRC2 deposits the methyl groups on the histone, while PRC1 binds to trimethylated H3K27 and ubiquitinates histone 2A to promote chromatin compaction. The core components of these complexes are shown in Figure 15. PRC2 is discussed in detail first since its enzymatic function on histones precedes PRC1 recruitment.

PRC2 is composed of 4 core subunits that are conserved from Drosophila to mammals, including EZH1/2, SUZ12, EED and RbAp46/48 (also known as RBBP7/4). The catalytic subunit of this complex is the EZH protein, which has two isoforms that differ in their expression and function: EZH2 but not EZH1 is expressed in actively dividing cells. EZH2 is the active

methyltransferase of PRC2 responsible for methylating H3K27. Unlike other methylases, EZH2 can successively add three methyl groups. Polycomb complexes containing EZH1 lack this ability and instead use a deferent mechanism to repress and compact chromatin [312]. The non-catalytic subunits EED and SUZ12 are required for complex activity. EED is a WD repeat protein that recognizes repressive trimethylated lysines and aids in the targeting of PRC2 to trimethylated H3K27 [313]. EED also contributes to the catalytic activity of EZH2 [314]. SUZ12 contains a C-terminal VEFS domain which promotes assembly of PRC2 and is required for the proper methylating function of PRC2 [315]. RbAp46/48 is not required for catalytic activity of the complex, but plays a role in complex assembly through interactions with SUZ12 and histone binding [316].

In addition to these four core components, PRC2 also interacts with other proteins for proper functioning. Jarid2 is a member of the Jumonji C family of proteins, but is catalytically inactive [317]. It promotes the targeting of PRC2, but functionally represses its activity, possibly to allow finer tuning of histone methylation. However, Jarid2 is required for PRC2 function since knockdown of the protein greatly reduces overall H3K27 methylation [318]. PCL has been suggested to stimulate the methylase activity of the complex and has been shown to recruit the complex to target genes through its PHD finger domain [319, 320]. AEBP2 has also been identified as a binding partner of PRC2; however, its function has not yet been determined.

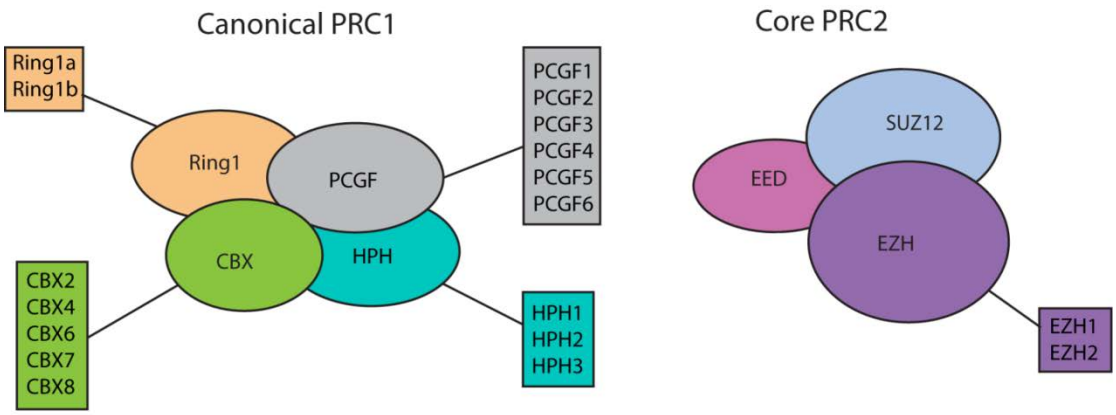

Figure 15: Polycomb Complexes
Core components of the mammalian polycomb complexes 1 and 2, adapted from *[321]*

Canonical PRC1 is composed of one isoform of each of the four main proteins: Ring1, PCGF, CBX, and HPH. The number of isoforms varies between species, but mammalian complexes can consist of one of 5 CBX, 2 Ring1, 6 PCGF, and 3 HPH different isoforms [322]. Depending on the composition, the function and targeting of the complex is different. Only polycomb complexes containing CBX2 and PCGF2 (Mel18) and/or PCGF4 (Bmi1) as well as Ring1a and Ring1b were found to bind to trimethylated H3K27, consistent with other data showing that CBX2 was the only CBX protein to show affinity for trimethylated H3K27 [323, 324]. Ring1 is the catalytic component of PRC1 that ubiquinates H2AK119, and its activity is enhanced by PCGF4 [325]. Ubiquitination of H2AK119 has been shown to be essential for chromatin compaction and repression of gene expression [326].

1.7.4. Histone Demethylases

The demethylases responsible for removing repressive methylation of H3K9 and H3K27 are localized within the same complex. Figure 16 shows the protein components of this complex.

The components of this complex include Set1A/MLL1-5, HCF-1, LSD1, JMJD2, and UTX/JMJD3. JMJD2 and LSD1 cooperatively function to demethylate trimethylated H3K9, while UTX/JMJD3 demethylates trimethylated H3K79 [124, 327]. The Set1A/MLL1-5 components are H3K4 methylases and are discussed in the above section. The H3K4 methylation mark is an activating mark, so formation of this complex may enable a coupling of repressive marker removal to the deposition of activating chromatin modifications.

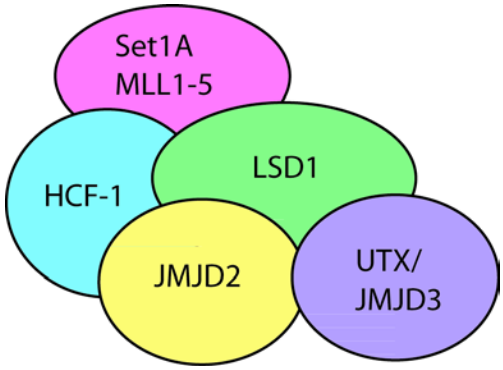

Figure 16: **Demethylase Complex**
Constituents of the LSD1 demethylase complex, adapted from [122]

LSD1 and JMJD2

LSD1 can demethylate both H3K9 and H3K4 to either activate or repress transcription, depending on its interaction partners, functioning as a H3K4 demethylase in the REST-CoREST complex, and a H3K9 demethylase when in the Set1A/MLL1 complex. Its function in the REST-CoREST complex was discussed in an earlier section, so this section will focus on its role as a H3K9 demethylase.

While LSD1 is a H3K9 demethylase, it can only demethylate mono or dimethylated H3K9, not trimetylated H3K9 [328], indicating that it cannot function alone in H3K9 demethylation. Rather, it functionally interacts with JMJD2, a second demethylase specific for H3K9me3. The JMJD2 protein can only remove one methyl group, leaving a dimethylated H3K9

product [327], a substrate for LSD1. Thus, trimethylated H3K9 is processively demethylated by the action of first JMJD2 then LSD1.

UTX/JMJD3

UTX and JMJD3 are specific to H3K27 methylation marks and can remove the first two methyl groups from trimethylated H3K27, but cannot efficiently remove the final methyl group. Furthermore, the function of these proteins, when purified, is much less efficient on nucleosomal substrates, indicating that there may be additional cofactors or modifications that are necessary to stimulate their activity [329]. These two proteins are likely involved in the activation of gene expression as monomethylation of H3K27, the product of their enzymatic function, is commonly associated with actively expressed genes [330]. UTX and JMJD3 can associate with MLL proteins to become incorporated into the H3K4 methylase complex [331].

1.8 RATIONALE

Herpes Simplex Virus infects multiple cell types during the stages of its lifecycle. Productive infection begins when the virus contacts the oral mucosa and enters epithelial cells. The virus replicates to high titers and gains access to nervous tissue through sensory neurons innervating the primary site of infection. The virus then travels up the length of the axon and enters the cell body where it replicates to low levels and establishes a lifelong latent infection in the ganglia. The virus periodically reactivates to reestablish productive infection, enabling the virus to spread. For this reason, HSV affects a large subset of the general population, and poses an

especial risk to immunocompromised patients who are more likely to develop resistance to acyclovir, the standard treatment. Second line drugs must be administered intravenously and have more severe side effects. Currently, there is no treatment to target the latent reservoir or permanently block viral reactivation. The potential to develop novel therapies underscores the need for a better understanding of the mechanism of reactivation and cellular factors that may affect viral transcription.

During the HSV lifecycle, the virus must transcribe and replicate its genome in two cell types – epithelial cells and neurons. These two cell types are vastly different, as epithelial cells are constantly dividing, whereas neurons are terminally differentiated and have exited the cell cycle. HSV transcription is mediated by host RNA polymerase II and various cellular transcription factors and activators. As neurons and epithelial cells have differences in their metabolic activity and transcriptional needs, they likely have different abundances of transcription factors that would likely affect HSV gene expression.

Furthermore, it has been shown that neurons and epithelial cells express different histone variants, and that the incorporation of these different variants augments chromatin structure in different ways. Chromatin plays an important role in the HSV lifecycle and is modified by various cellular complexes. During lytic infection in epithelial cells, the genome is in a euchromatic state so that viral transcription proceeds robustly. However, during latency, the viral genome is repressed in densely packed heterochromatin which greatly restricts transcription. For reactivation to occur, the genome must transition back to a euchromatic state, but during the early events of reactivation, viral activators must likely function on genomes that are at least partially heterochromatic.

Therefore, the goals of this study were to determine if HSV transcription varied between cell types and if HSV transactivators could function on heterochromatinized templates. More specifically, the two goals of this project were to:

1. **Characterize the transcriptome of wild type and different HSV mutants in neurons and MRC5 cells.**
2. **Determine if the main transcription factor ICP4 could access and activate transcription from quiescent genomes in neurons and MRC5 cells.**

These two goals will be discussed in the following chapters.

2.0 TRANSCRIPTION OF THE HERPES SIMPLEX VIRUS, TYPE 1 GENOME DURING PRODUCTIVE AND QUIESCENT INFECTION OF NEURONAL AND NON-NEURONAL CELLS.

2.1 SUMMARY

Herpes simplex virus, type 1 (HSV-1) can undergo a productive infection in non-neuronal and neuronal cells such that the genes of the virus are transcribed in an ordered cascade. HSV-1 can also establish a more quiescent or latent infection in peripheral neurons, where gene expression is substantially reduced relative to productive infection. HSV mutants defective in multiple immediate early (IE) gene functions are highly defective for later gene expression and model some aspects of latency in vivo. We compared the expression of wild-type (wt) virus and IE gene mutants in non-neuronal cells (MRC5) and adult murine trigeminal ganglion (TG) neurons using the Illumina platform for RNA sequencing (RNAseq). RNAseq analysis of wild type virus revealed that expression of the genome mostly followed the previously established kinetics, validating the method, while highlighting variations in gene expression within individual kinetic classes. The accumulation of immediate early transcripts differed between MRC5 cells and neurons, with a greater abundance in neurons. Analysis of a mutant defective in all five IE genes

(d109) showed dysregulated genome wide low-level transcription that was more highly attenuated in MRC5 cells than in TG neurons. Furthermore, a subset of genes in d109 was more abundantly expressed over time in neurons. While the majority of the viral genome became relatively quiescent, the latency-associated transcript was specifically upregulated. Unexpectedly, other genes within repeat regions of the genome, as well as the unique genes just adjacent the repeat regions also remained relatively active in neurons. The relative permissiveness of TG neurons to viral gene expression near the joint region is likely significant during the establishment and reactivation of latency.

2.2 IMPORTANCE

During productive infection the genes of HSV-1 are transcribed in an ordered cascade. HSV can also establish a more quiescent or latent infection in peripheral neurons. HSV mutants defective in multiple immediate early (IE) genes establish a quiescent infection that models aspects of latency in vivo. We simultaneously quantified the expression of all the HSV genes in non-neuronal and neuronal cells by RNAseq analysis. The results for productive infection shed further light on the nature of genes and promoters of different kinetic classes. In quiescent infection, there was greater transcription across the genome in neurons relative to non-neuronal cells. In particular, the transcription of the latency-associated transcript (LAT), IE genes, and genes in the unique regions adjacent to the repeats, persisted in neurons. The relative activity of this region of the genome in the absence of viral activators suggests a more dynamic state for quiescent genomes persisting in neurons.

2.3 INTRODUCTION

Herpes Simplex Virus 1 (HSV-1) productively replicates in non-neuronal cells at peripheral sites of individuals where it gains access to nerve endings of the peripheral nervous system. In sensory neurons, the virus can either replicate or it can establish life-long latency, where the viral genome is relatively quiescent compared to productive infection. Therefore, the viral genome can be expressed to produce progeny virions in both non-neuronal and neuronal cells, but can also establish a reversible silent state in some neuron populations. As some studies suggest [332, 333], there may be differences in how the genome is expressed in neuronal and non-neuronal cells, however some aspects of the former study may be explained by viral spread in vivo [334].

The transcription of HSV genes in non-neuronal cells by RNA pol II [13] is sequentially and coordinately regulated [68, 335], generally being described as a cascade where three classes of genes are expressed, the Immediate Early (IE), Early (E), and Late (L) genes. IE gene transcription is activated in the absence of prior viral protein synthesis by VP16 present in the infecting virion [336-338]. The products of IE genes are required for the transcription of the E genes, which encode the DNA synthetic machinery. IE and E proteins along with viral DNA replication are required for efficient late (L) transcription, the protein products of which mostly comprise the virion structure or are required for its assembly.

Promoters for the genes of the 3 kinetic classes are generally thought to share common features. IE gene promoters contain a TATA box [339, 340], elements that are responsive to VP16 in the incoming virion, and upstream binding sites for cellular transcription factors [336, 337, 341]. Early gene promoters that have been characterized contain a TATA box, and upstream

binding sites for cellular transcription factors, such as found in the thymidine kinase promoter [342, 343]. Late genes that are tightly dependent on viral DNA synthesis contain a TATA box and an initiator element [344] at the start sites of transcription [345]. Despite these similarities, genes of a particular class may be expressed quite differently.

While viral cis- and trans-acting elements are a determinant of how individual viral genes are expressed, the transcriptional environment in different cell types likely influences the expression of the genome as well. When HSV enters sensory neurons, gene expression may be repressed and the virus can establish a latent infection. The genome of the latent virus persists in a state repressed by heterochromatin [189, 346], where the predominant gene that is expressed is the latency associated transcript or LAT [347]. The genome can periodically reactivate to produce new virus by a poorly understood process that may involve the initial low-level dysregulated expression of the genome [217, 218].

We and others have utilized a system that employs viral mutants, which are defective for the expression of IE genes to model some aspects of latency in vitro [190, 348]. d109 is a mutant that does not express the IE proteins, is nontoxic to most cells in culture and its genome is relatively quiescent in infected Vero cells or diploid human fibroblasts [190]. d109 genomes persist in an endless form in cultured cells [349], as do latent virus genomes in vivo [350]. Moreover, the expression of select viral genes from d109 genomes persisting in human fibroblasts is several orders of magnitude less than wild-type virus and the genomes are bound by heterochromatin [191, 222] as they are in latency. While it is reasonable to pursue this model of latency given the similarities with latency in vivo, the expression of mutants such as d109 in cells where latency is naturally established has not been studied in detail.

The intent of this study was to simultaneously quantify the expression of all the viral genes during productive infection with wt virus and during quiescent infection of non-neuronal (diploid fibroblasts) and neuronal (adult trigeminal neurons) cells to provide insight into: i) what determines the kinetics of expression of individual viral genes, ii) neuronal specific differences in viral gene expression, and iii) the expression of the viral genome as it persists in TG neurons. Previous approaches to simultaneously quantify the expression of the viral genes have used hybridization to microarrays of viral sequences representing the viral genes [351]. In the present study we used RNA-sequencing (RNAseq) of cDNA derived from infected MRC5 cells and neurons. This method proved to be sufficiently sensitive to measure the expression of all the viral genes in the absence of viral activators in MRC5 cells and TG neurons.

2.4 MATERIALS AND METHODS

Ethics statement: This study was carried out in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. All animal procedures were performed according to a protocol approved by the Institutional Animal Care and Use Committee of the University of Pittsburgh (protocol number: 1201001B, most recently approved January 8, 2014). Appropriate sedatives, anesthetics and analgesics were used during handling and surgical manipulations to ensure minimal pain, suffering and distress to animals.

Virus and Cells. Experiments were performed using human embryonic lung cells (MRC5) and primary mouse trigeminal neurons. MRC5 cells were obtained from and propagated as recommended by the American Type Culture Collection. Neurons were isolated from the trigeminal ganglia of 6 week old CD1 mice, as described below. The viruses used in this study were d109 (ICP4- ICP0- ICP22- ICP27- ICP47-), d106 (ICP4- ICP22- ICP27- ICP47-), n12 (ICP4-), and the wild type virus KOS. d109, d106, n12, and KOS viruses were propagated on FO6F1, E11, E5, and Vero cells respectively [153, 190].

Generation of neuronal cultures. The trigeminal ganglia (TG) were dissected from six week old CD1 mice and neurons were isolated following a protocol similar to a previously described procedure [352]. Neurons from the TG of 10 mice were used to seed a 24 well plate. Following establishment of cultures, the cells were maintained in Neurobasal-A media supplemented with 2% B27, 1% Penicillin-Streptavidin, .5mM L-glutamine, 50ng/mL nerve growth factor (NGF), glial cell-derived neurotrophic factor (GDNF), and neurturin. Media was changed once per week.

RNA Isolation and Reverse Transcription. 2×10^6 MRC-5 cells in 60-mm plates were infected at a MOI of 10 PFU/cell of d109, d106, n12, or KOS. For the infection of neurons, 10^6 PFUs of d109 or KOS were added to each well. The infection was performed at room temperature for 1 hour with rocking every 10 minutes. Following infection, the inoculum was removed, the cells were washed, and fresh pre-warmed media was added. RNA was isolated with the Ambion RNaqueous-4PCR kit by following the included protocol. RNA was harvested at the indicated time points by removing the medium and adding 500µl lysis/binding buffer to MRC5 cell cultures and by adding 100µl lysis/binding buffer per well to neuronal cell culture. For the

neuron cultures, 8 wells were pooled per sample. The following steps are the same for both neuron and MRC5 samples. Cells were collected and vortexed. An equal volume of 67% ethanol was added. The solution was applied to a filter spin column and centrifuged at 12,500 rpm for 1 min. The bound RNA was washed with wash buffers 1 and 2/3. The column was centrifuged dry to remove residual wash buffer from the column. RNA was eluted in two steps with 60µl and 20µl 75°C elution solution. The RNA was treated with DNase I at 37°C for 30 min to degrade remaining DNA. The DNAse was inactivated with the supplied reagent.

To generate cDNA, two micrograms of total RNA was reverse transcribed in a reaction volume of 20µl containing 0.5µlRiboguard RNase inhibitor (40 U/ul), 1µl 10pM OligodT Primer, .5µl MMLV High Performance Reverse Transcriptase (200U/µL), 2µl 100 mM dithiothreitol (DTT), 4µl nucleotide mix (2.5mM concentration of each dNTP) and 2µl 10x reaction buffer. The reaction tube was incubated at 65°C for 2 min to remove RNA secondary structure, and the RT reaction was carried out for 1 h at 37°C. Following completion of the RT reaction, the reaction tube was incubated at 85°C for 5 min to inactivate the reverse transcriptase.

Quantitative PCR. RNA was diluted to 1µg in 60µl and cDNA was diluted 1:6 in ultrapure water. A master mix containing 0.3µl of each primer (stock concentration, 100 µM), 5µl Applied Biosystems SYBR green mix with 1.0µM 6-carboxy-X-rhodamine and .4µl of water for a total of 6µl for each reaction was made. The primers used in this study were: for gC, gtgacgtttgcctggttcctgg, and gcacgactcctgggccgtaacg; for ICP27, gggccctttgacgccgagaccaga, and atggccttggcggtcgatgcg. The TK primers were the same primers used in previous studies [191]. A 96 well plate was prepared with 6µl master mix and 4µl of sample or standard. All samples were run in triplicate. Purified viral DNA was used to create a standard curve of 1:10 dilutions from

4000000 to 40 copies, which covers the threshold cycle value range for the samples tested. qPCR was run on a StepOne Plus real-time PCR machine. The conditions for the run were 95°C for 10 min and 40 cycles of 95°C for 15 s and 60°C for 1 min. At the end of the run, a dissociation curve was completed to determine the purity of the amplified products. Results were analyzed using the StepOne v2.1 software from Applied Biosystems and compiled in Microsoft Excel.

RNA-Sequencing. RNA was harvested using the Ambion kit as described above and prepared for sequencing following the IlluminaTruSeq RNA Sample Preparation v2 Guide and accompanying kit (Illumina, San Diego, CA). The libraries were analyzed for length and concentration using the Agilent Bioanalyzer. Samples for an experiment were mixed in equimolar concentration and sent to the Tufts University Core Facility, Boston, MA, for sequencing.

Following sequencing, the Illumina reads were uploaded onto the Galaxy server for analysis (1,2,3). The reads were first groomed using the Fastq groomer function and trimmed from the 3' end to remove low quality score base reads. The reads were then mapped to the appropriate reference genome using the TopHat read mapper. For wt HSV, the strain KOS sequence was used (Genbank accession number: JQ780693). For alignments, the terminal repeats (TR_L and TR_S) were removed since they are redundant with the internal repeats. For d109, the plasmids used to generate the virus were sequenced [190], and the changes were incorporated into the KOS sequence. The terminal repeats were also removed for the purpose of alignment. Cufflinks was used to quantify the mapped reads, outputting data in FPKM format. Since the number of reads and mapped reads varied between samples, the data was renormalized to the total number of reads.

2.5 RESULTS

RNAseq analysis of productive infection in human fibroblasts and trigeminal neuronal cultures.

In order to determine the kinetics of expression of all the HSV genes in non-neuronal and neuronal cells, monolayers of MRC5 cells and cultures of adult TG neurons were infected with HSV-1, strain KOS. Total RNA was collected at different times post infection, and processed and subjected to Illumina sequencing as described in the Materials and Methods. Illumina reads were processed and aligned to a modified KOS sequence. Figure 17A shows the percentage of total reads that map to the viral genome. The percent of viral reads in the cell rapidly increases from less than 5% at 1hpi to over 60% by 6 hpi. This accumulation begins to plateau after 8 hpi, where approximately 80% of the total mRNA in the cell is virally derived. Similar results were obtained in trigeminal neuron cultures, although the percent viral reads were somewhat reduced compared to MRC5 cells. The results in MRC5 cells are consistent with previous studies using metabolically labeled RNA from infected Hela cells and hybridization techniques [353], where it was found that 25-30% of newly synthesized nuclear polyA+ RNA in the nucleus was HSV specific at 2-3 hpi, and this increased to 33% between 5-6 hpi. The same study found that most of the polysome associated polyA+ RNA was virus specific at both times. While the previous methods differ considerably from ours, the similarity with our results helps validate the approach.

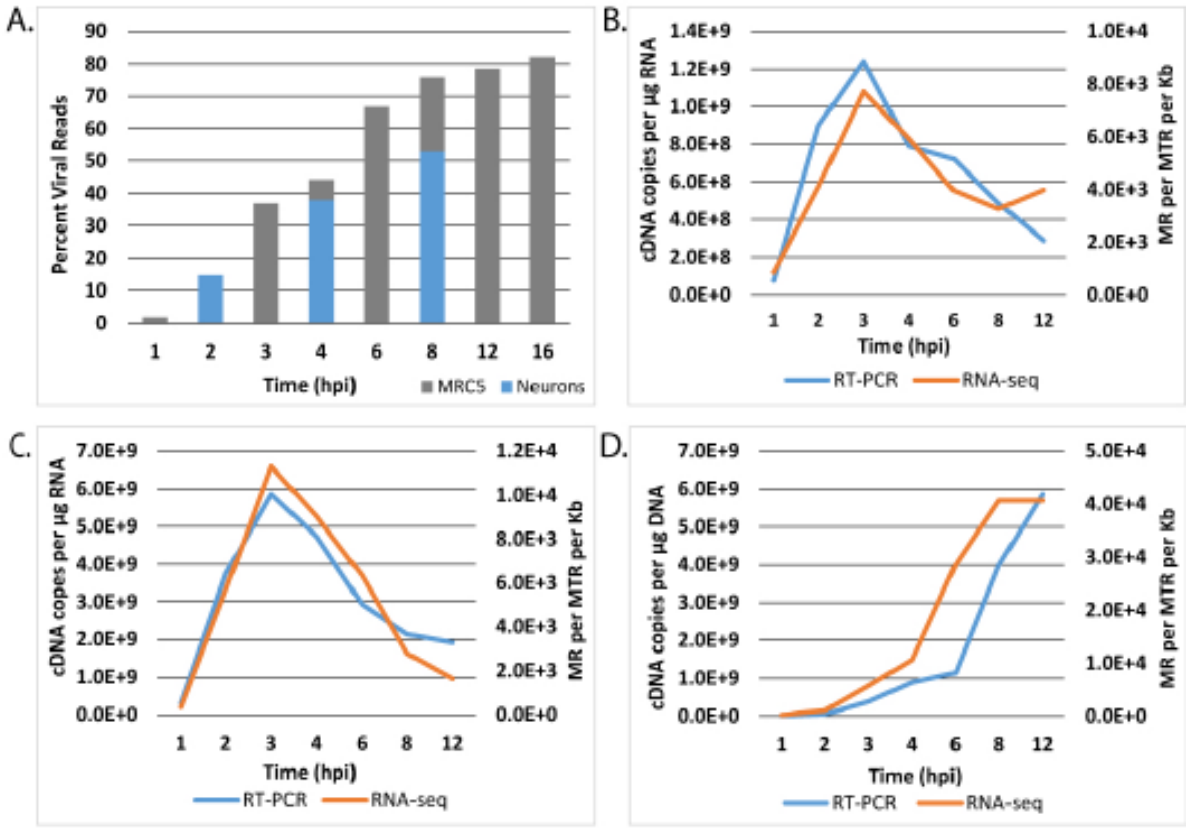

Figure 17: Evaluation of RNAseq reads from infected MRC5 cells.
Monolayers of MRC5 cells and cultures of adult trigeminal neurons were infected with HSV-1, strain KOS at an MOI of 10 PFU/cell. Total RNA was collected from MRC5 cells at 1, 2, 3, 4, 6, 8, 12, and 16 hours post infection, and from TG neurons at 2, 4, and 8 hours post infection. cDNA was prepared and subject to Illumina sequencing as described in the Materials and Methods. Illumina reads were processed and aligned to a modified KOS sequence using the TopHat mapper in the Galaxy Cloud software package. **A.** The percent of sequencing reads mapping to the HSV genome relative to the total number of reads. **B.** Comparison of kinetics of ICP27 (UL54) accumulation by RNAseq and RT-PCR. **C.** Comparison of kinetics of tk (UL23) accumulation by RNAseq and RT-PCR. **D.** Comparison of kinetics of gC (UL44) accumulation by RNAseq and RT-PCR. The units for RT-PCR were cDNA copies per ug RNA as determined by comparison to standards using CsCl-purified viral genomic DNA. The units for RNAseq are reads mapped to the viral genome per million total reads (viral+cell) per kilobasepair (MR per MTR per Kb).

To further validate our approach, we compared the amount of signal from representative IE, E and L genes as determined by RNAseq and quantitative reverse transcription PCR (RT-PCR). Figures 17 B, C and D represent the results from ICP27 (IE), tk or UL23 (E), and gC or UL44 (L), respectively. The accumulation of these three genes was similar when measured by

RNAseq and RT-PCR, demonstrating that RNA-sequencing is sufficiently sensitive and accurate to detect changes in transcript abundance over time, with the obvious additional ability to quantitatively compare differences in the expression of many genes.

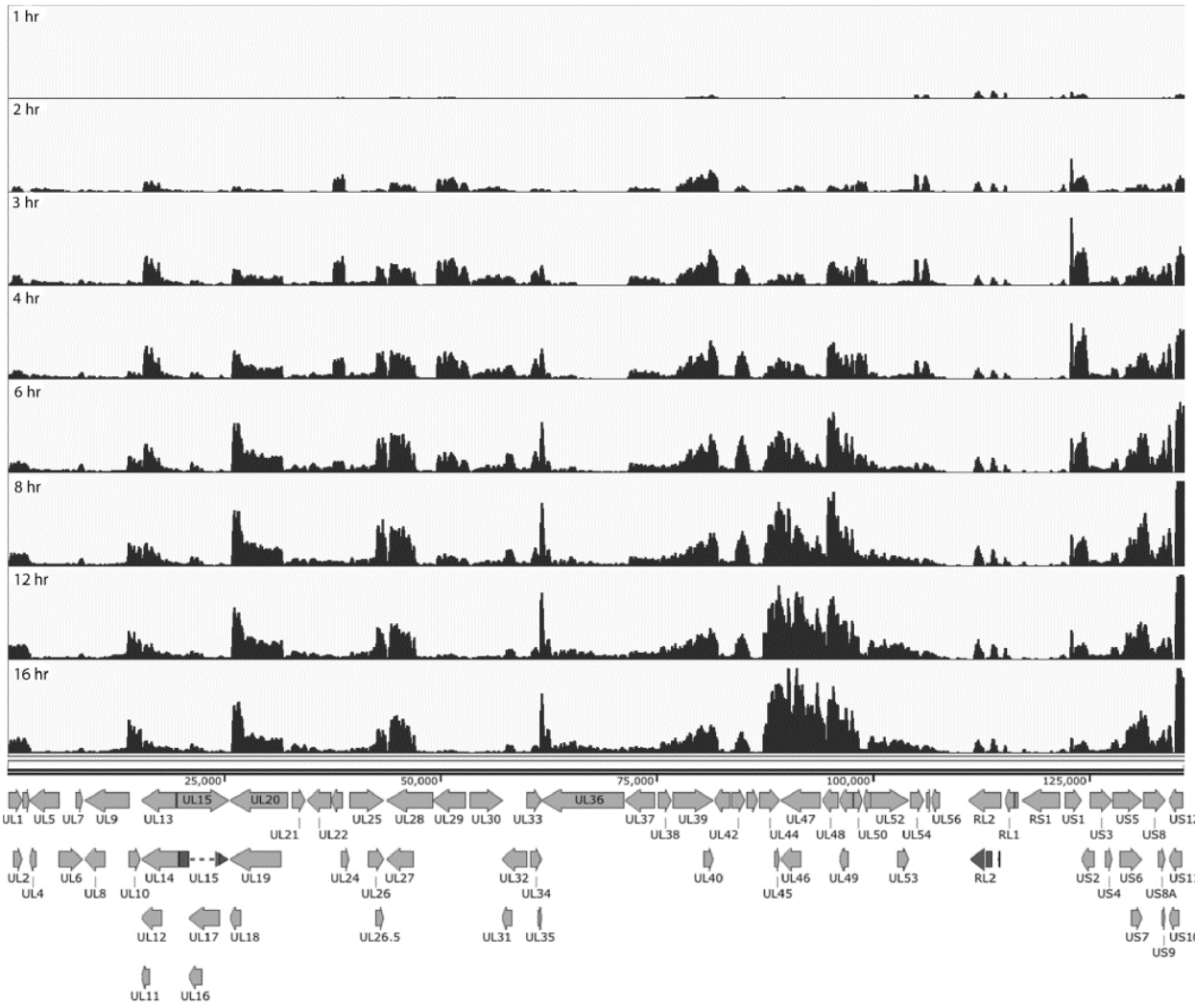

Figure 18: Locations of RNAseq reads across the HSV genome as a function of time post infection in MRC5 cells.
The locations of reads for each of the time points are shown relative to a map of the mRNA for each of the HSV genes. The scales of the graphs were set to normalize to the number of total reads to enable qualitative comparison of transcript abundance between different time points. The reads and the genome are represented without the long and short terminal repeats since their sequences are represented within the internal repeats. The exons of RL2 and UL15 are also shown by darker shading.

The reads that mapped to the viral genome from the RNA-seq of wt-virus (KOS) infection were visualized using the Integrated Genomics Viewer [354]. Figure 18 shows the reads as a function of time post infection mapped to the KOS genome. Changes in the expression of individual genes over time can readily be seen and in many cases appear to reflect the established transcriptional cascade of viral genes. At the 1 and 2 hour time points, signals for the immediate early genes, ICP27 (UL54), ICP0 (RL2), ICP22 (US1) and ICP47 (US12) are evident. The accumulation of ICP4 is not readily seen at this level of sensitivity. The three peaks between 111-115kb map to the exons of the ICP0 locus, further validating the precision of RNA sequencing for HSV mRNA profiling.

At 4 hours post infection, the abundance of ICP27 and ICP22 mRNA begins to decline. ICP0 abundance continues to increase throughout the infection, but constitutes a lower proportion of expressed genes. This is consistent with previous observations [355]. Concurrently, early genes begin to be expressed. Thymidine kinase (UL23), located between 37 and 39kb, peaks between 3-4 hours post infection. Expression of DNA replication machinery genes ICP8 (UL29) and polymerase (UL30), between 49 and 52kb and 53.5 and 57kb, respectively, also peak at a similar time to UL23. Multiple other early genes show similar expression patterns.

Replication of viral DNA begins around 4 hours post infection, which enables or enhances the expression of late genes. Many late genes, particularly in the region between the genes for gC (UL44) and UL49, predominate at this time. The peak for the mRNA encoding the tegument protein VP16 (UL48) appears between 94.2 and 95.7kb. It is expressed as early as 2 hours post infection, but its abundance continues to increase following DNA replication. UL46 also shows a similar expression pattern. By contrast, UL44, a true late gene, is not expressed until 4 hpi, after which its expression increases significantly. UL10, another true late gene

located between 14 and 15.4 kb, does not have a distinguishable peak until 4 hours post infection and becomes much more abundantly expressed later in infection.

We next aligned the reads from the infections of trigeminal neurons to the virus genome. The aligned reads of cDNA from mRNA derived from TG neurons infected for 2, 4 and 8 hpi are shown in Figure 19. In addition to the reduced abundance of total viral transcripts in TG neurons, there were some notable differences in the accumulation of some individual transcripts over the 8h. VP16 (UL48) accumulation is more abundant early in infection of TG neurons relative to MRC5 cells. It has been suggested that VP16 possess promoter elements conferring neuron specific transcription [356]. Additionally, there was a larger proportion of the immediate early transcript region expressed, including ICP4 (RS1), ICP22 (US1), and ICP47 (US12).

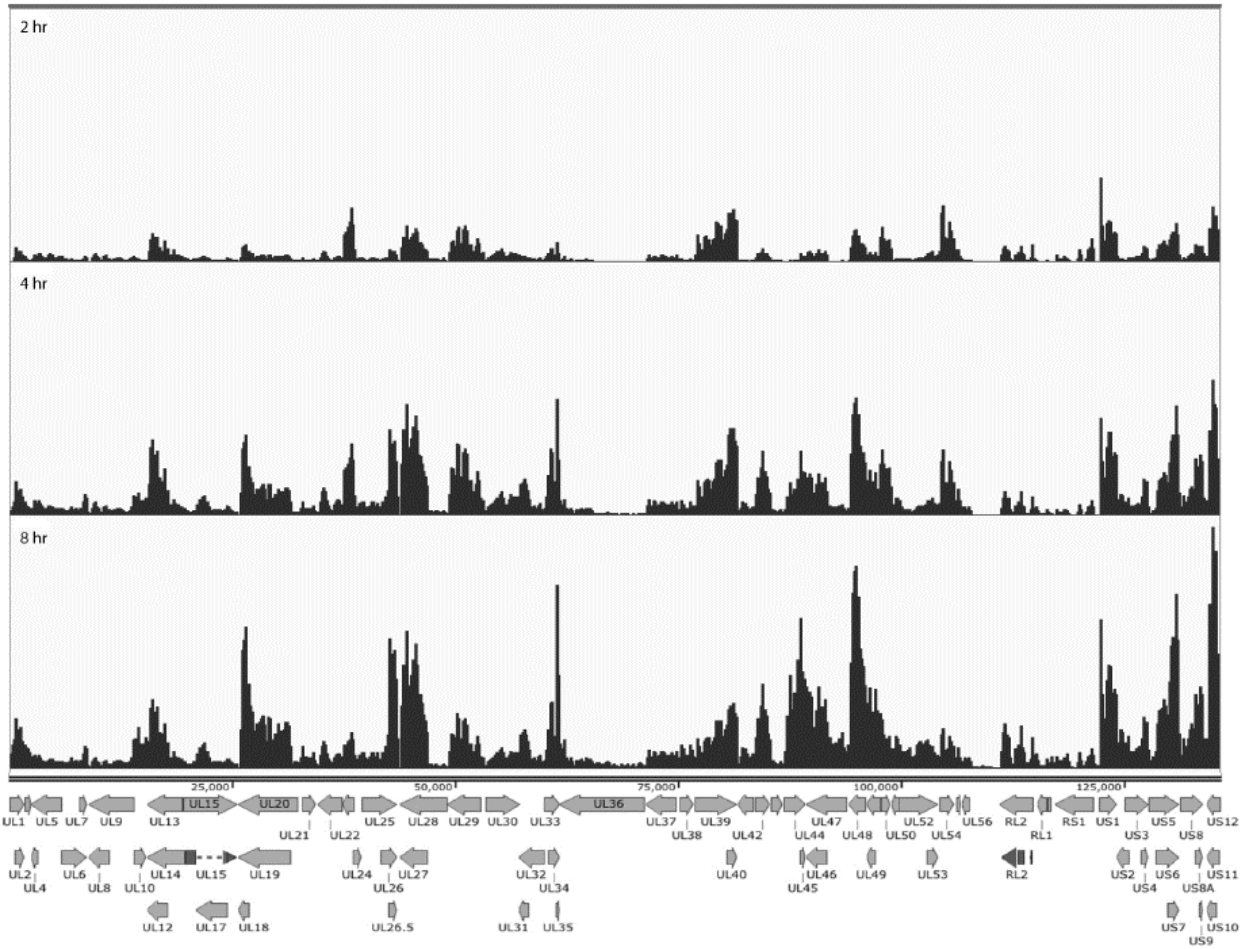

Figure 19: Locations of RNAseq reads across the HSV genome as a function of time post infection in cultured TG neurons.
As described in the legend for Fig. 18.

The visual representations of the kinetics of mRNA accumulation (Figures 18 and 19) provide a picture of the simultaneous transcription of all the HSV genes on the genome, but they lack the sensitivity to accurately reflect the transcription of many of the HSV genes. For example, an abundant transcript, such as UL44, may give rise 2 $x10^5$ reads, and its peak is clearly evident, while poorly transcribed genes, such as UL28 or UL36, are barely visually evident (Fig. 18). However, there are approximately 2 x 10^7 total reads in each of the multiplexed samples in Fig. 18, providing sufficient sensitivity to quantify the poorly expressed genes such as UL28 and UL36. Therefore, the expression of a subset of immediate early, early and late genes was quantified in MRC5 cells and neurons using the Cufflinks software package on the galaxy server. The accumulation of transcripts from the IE genes in MRC5 cells and neurons are shown in Figure 20A and 20B, respectively. The expression kinetics for ICP27 and ICP0 was quantitatively similar in both cells types. The kinetics of ICP4 expression was similar between the two cell types, however ICP4 was 2-5 fold more abundant in neurons than in MRC5 cells. There was a considerable difference in the expression profile of ICP22 (US1). In MRC5 cells, ICP22 peaked at 4hpi, however transcript abundance continues to increase in neurons.

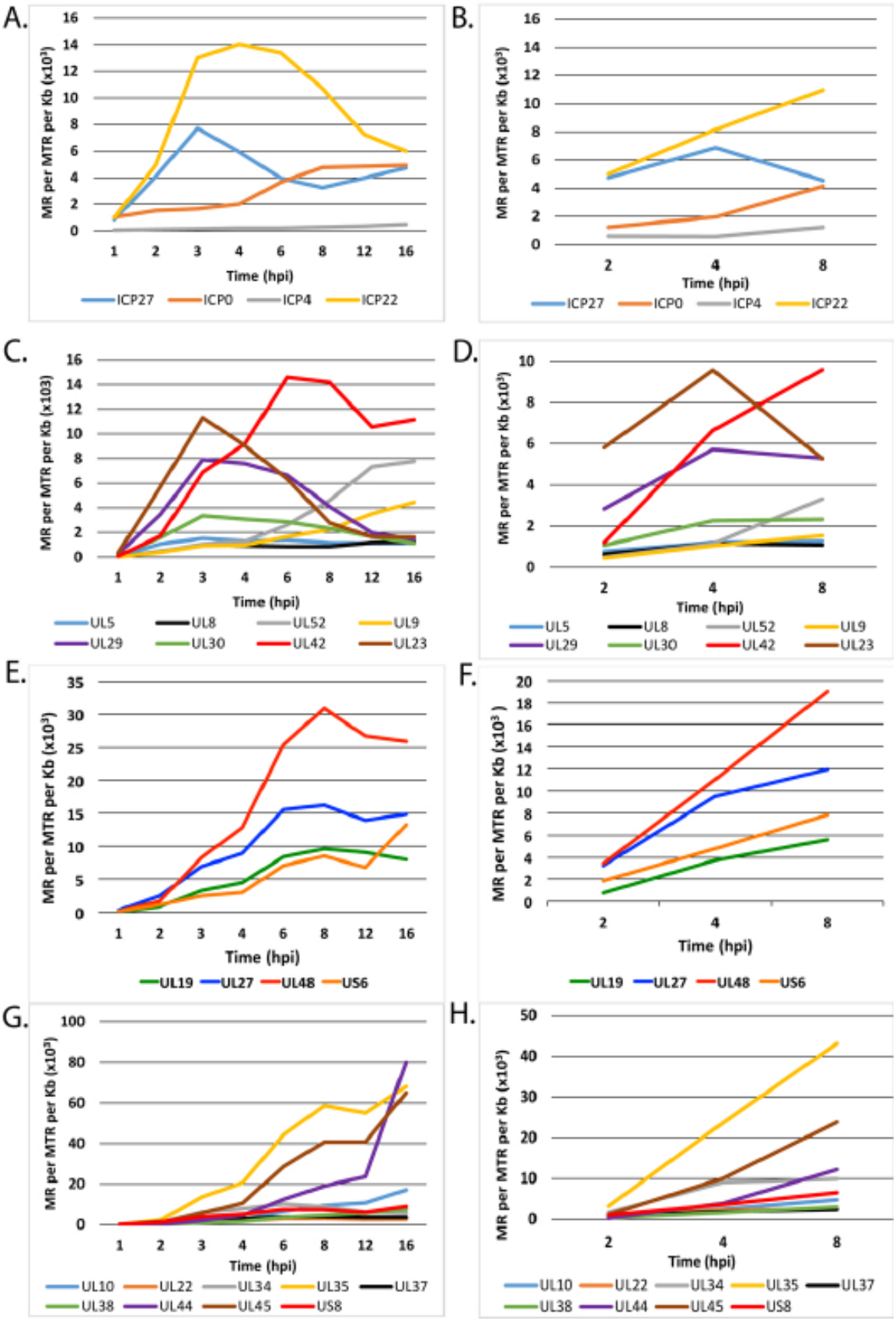
A.
MR per MTR per Kb (x10³)
Time (hpi)
ICP27
ICP0
ICP4
ICP22
B.
MR per MTR per Kb (x10³)
Time (hpi)
ICP27
ICP0
ICP4
ICP22
C.
MR per MTR per Kb (x103)
Time (hpi)
UL5
UL8
UL52
UL9
UL29
UL30
UL42
UL23
D.
MR per MTR per Kb (x10³)
Time (hpi)
UL5
UL8
UL52
UL9
UL29
UL30
UL42
UL23
E.
MR per MTR per Kb (x10³)
Time (hpi)
UL19
UL27
UL48
US6
F.
MR per MTR per Kb (x10³)
Time (hpi)
UL19
UL27
UL48
US6
G.
MR per MTR per Kb (x10³)
Time (hpi)
UL10
UL22
UL34
UL35
UL37
UL38
UL44
UL45
US8
H.
MR per MTR per Kb (x10³)
Time (hpi)
UL10
UL22
UL34
UL35
UL37
UL38
UL44
UL45
US8

Figure 20: The accumulation of HSV transcripts in MRC5 cells and TG neurons.
Quantification of IE (α) (**A and B**), E (β) (**C and D**), L (γ1) (**E and F**), and L (γ2) (**G and H**) transcripts in MRC5 cells (**A, C, E, and G**) and neurons (**B, D, F, and H**). The units for RNAseq are reads mapped to the viral genome per million total reads (viral+cell) per kilobasepair (MR per MTR per Kb).

Genes specifying the proteins involved in viral DNA synthesis are considered early genes, as is the viral thymidine kinase. Figures 20C and 20D show the kinetics of expression of these genes in MRC5 cells and neurons, respectively. UL23, UL29, and UL30 followed the established expression profile for early genes, with an increase in mRNA abundance early in infection followed by a decrease after 3-4 hpi. Expression of UL42 was the highest expressed DNA replication gene reaching maximal expression by 6 hpi and subsequently remaining high through the duration of the time course. The number of UL42 specific reads indicates that the abundance of this mRNA comprised over 1% of the total polyadenylated RNA in the cell. The somewhat delayed peak in the accumulation of UL42 mRNA is consistent with earlier studies [357]. Unlike immediate early genes, the expression patterns of these early genes in MRC5 cells and neurons were similar, although transcript abundance in neurons was less overall. The read abundance of UL5, 8, 52 and UL9 was considerably less than the others and their accumulation continued to increase following replication of DNA, reaching their maximal levels late after infection. Their accumulation appears to be more consistent with late gene expression kinetics.

Late genes (γ) require DNA synthesis for maximum expression and continue to accumulate throughout infection. Some L genes, the γ1, also have characteristics of E genes and are some of the most abundantly expressed genes of the virus. Figure 20E and 20F shows the accumulation of 4 γ1 genes, UL19 (major capsid protein), UL27 (gB), UL48 (VP16), and US6 (gD) in MRC5 cells and trigeminal neurons, respectively. The transcripts for these genes began to accumulate at 2 hpi and continued to accumulate to relatively high levels. VP16 and gB in

particular began to accumulate very early in neurons each eventually comprising 0.4% ($4x10^3$ MR/MTR/Kb) of the total mRNA in the cell. γ2 genes require DNA synthesis for their transcription. The accumulation of transcripts for 9 γ2 genes in MRC5 cells and TG neurons is shown in Figures 20 G and H, respectively. Like the transcripts for early genes (Figures 20C and 20D), the expression patterns of these late genes were similar in neurons and MRC5 cells. However, there was a large difference in the level of accumulation of the different transcripts in this group of genes. The three most highly expressed late genes were UL44, UL45, and UL35. While displaying similar expression kinetics, the accumulation of UL10, 22, 34, 37, 38, and US8 were considerably less.

Immediate Early Gene Mutants

Immediate early proteins of HSV are required for the expression of the remainder of the HSV genome [335]. Specifically, ICP4 is required for the expression of viral early and late genes [358-360]. Mutants deleted in the ICP4 gene are defective for transcription beyond the IE phase [56, 153], such that the expression of a prototypic early gene, UL23 (tk) for example, is approximately two orders of magnitude less than that seen in wt virus infection [76]. ICP0 promotes virus growth and reactivation from latency [361-363]. The deletion of ICP0 from HSV already deleted for ICP4 results in the further restriction of viral gene expression [190, 191]. Abrogating the expression of all IE genes results in a virus that is not toxic to cell, and persists for long periods of time [190] in a repressed heterochromatic state [191, 222] resembling that seen with wt virus in vivo [188, 346]. To examine the expression across the entire genome of key IE gene mutants, MRC5 cells were infected with KOS, n12, d106, and d109 at an MOI of 10 PFU/cell for 4h and processed for RNAseq. n12 contains a nonsense mutation in the coding

region of ICP4, rendering it non-functional [153]. d106 is mutated such that ICP0 is the only intact IE gene and d109 is defective for the expression of all the IE genes [190]. The mapped reads from each virus relative to the KOS genome are shown in the same scale in Figure 21A to illustrate the disparity in gene expression between the 4 different viruses.

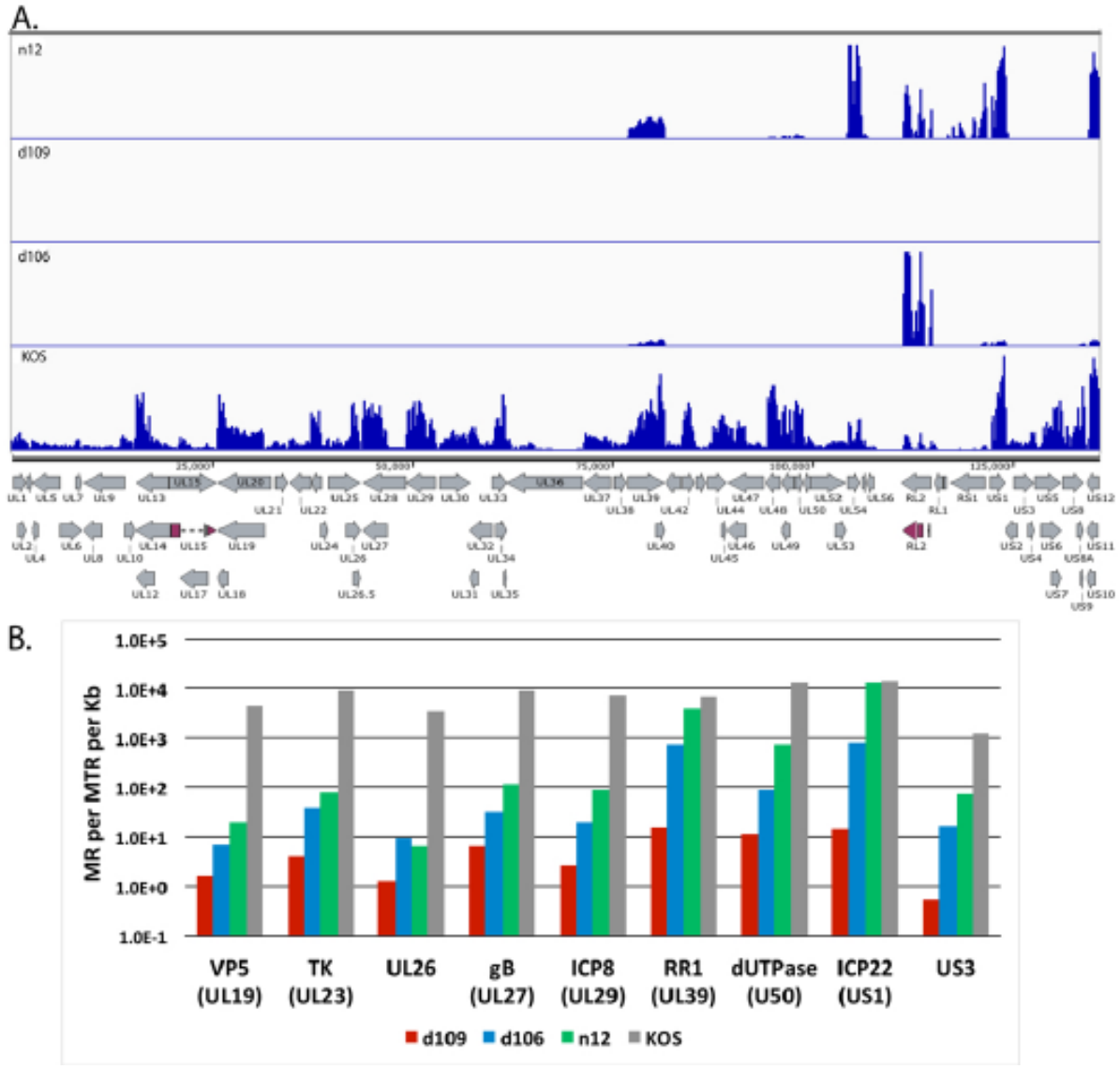

Figure 21: Transcript accumulation in IE mutant-infected MRC5 cells.
MRC5 cells were infected with KOS, n12, d106, and d109 at a moi of 10 PFU/cell for 4 hr, and processed for RNA seq as previously stated. **A.** Locations of mapped reads relative to the KOS genome. The Illumina reads were aligned to the modified KOS sequence as before. The maximum on the y-axis is 30,000 reads. **B.** Quantification of the numbers of reads for select HSV genes.

Transcription of numerous genes can be seen across the entire KOS genome, while transcription of the n12 genome appears to be restricted to the UL39, 40 loci, ICP27 (UL54), ICP0 (RL2), ICP4 (RS1), ICP22 (US1), and ICP47 (US12). The levels of ICP0, ICP27 and ICP4 transcripts are considerably higher in n12-infected cells compared to KOS, consistent with the previously defined phenotype of n12 [153]. The expression profiles in Fig 21A of d106 and d109 are also consistent with their published phenotypes [190]. However, this representation does not quantitatively describe the expression of the genome in the absence of the IE proteins. Considering there were ~20 million reads in each sample and the percent viral reads were 44, 13, 4.9 and .07, for KOS, n12, d106, and d109, respectively, then there are a considerable number of reads to accurately quantify the relatively low abundance transcription across the genome in these mutant backgrounds. For example, while no signal is visible for d109 in Figure 21A, there are still 1.4×10^4 viral reads in the d109 sample. Fig. 21B shows the number of reads mapped to a particular viral gene per million total reads per kilobase pair (MR per MTR per Kb) for selected viral genes in the different mutant backgrounds. The genes are for ICP22 (IE), tk (E), ICP8 (E), RR1 (E), dUTPase (E), US3 (L), VP5 (L), UL26 (L) and gB (L). Several observations can be made from the results; i. IE genes (when present) are highly expressed in the absence of ICP4, ii. In the absence of ICP4 (n12), early and late gene transcripts accumulate to 0.1% to 5.0% the level seen in wt virus infection. An exception is RR1 (UL39), the transcription of which is relatively independent of ICP4, iii. Early and late gene expression is further reduced 3 - 5 fold in the d106 background at this time post infection, and iv. In the absence of IE proteins early and late transcription is 3 to 4 orders of magnitude less than wt virus at 4 hpi in MRC5 cells. For example the total number of reads for the tk mRNA is 3×10^5 and 1.5×10^2, for KOS and d109, respectively.

Expression from the d109 genome in MRC5 cells is further reduced with time due to the ongoing formation of heterochromatin [191]. To examine how this repression affects expression across the viral genome, and to further compare expression in MRC5 cells to that in trigeminal neurons in the absence of IE proteins, MRC5 cells and TG neuronal cultures were infected with d109 at a MOI of approximately 10 PFU/cell. mRNA was isolated at 4, 8 and 24 hours and 7 days post infection, and analyzed by RNAseq as before. The percent of viral reads from the MRC5 samples decreased from 0.07% at 4h to just under 0.01% at 7 days, while the percent viral reads in trigeminal neurons was 0.5 % at 4h and a little more than 0.1% at 7days, Figure 22. Therefore, the number of viral reads was about 10-fold higher in d109-infected neurons than in d109-infected MRC5 cells, suggesting a more transcriptionally active genome in trigeminal neurons. This is in contrast to what was seen in KOS infected cells, where the proportion of viral mRNA in neurons was less than in MRC5 cells. The locations of the reads are shown in Figure 22 aligned to a modified d109 genome. Unlike Fig. 21A, the scale of the graphs was chosen to view the relatively low level expression across the entire genome. As a consequence the signals for UL39, 40, the internal repeats and GFP are off the scale. They are represented in Figure 23. Regions where there were no reads are represented by gray areas.

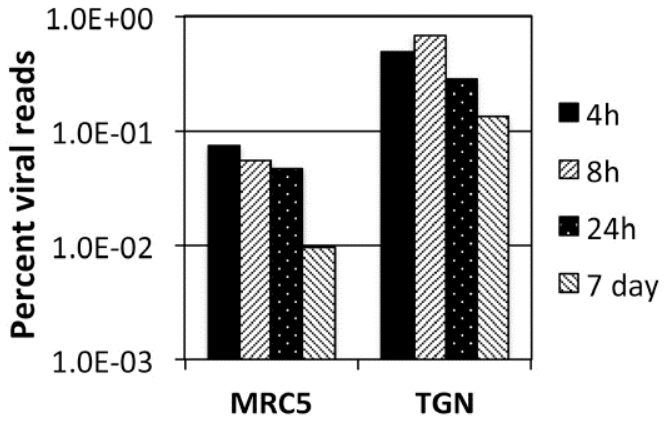

Figure 22: Percent viral reads in d109 infected MRC5 cells and TG neurons.

The numbers of reads that mapped to the d109 genome were divided by the total number of reads and multiplied by 100.

The accumulation of gene specific reads can readily be seen. Consistent with the higher level of transcription overall, there was a general trend toward higher levels of individual transcripts in neurons than in MRC5 cells. For example, the levels of glycoprotein B (UL27), ICP8 (UL29), UL30, UL41, UL42 were higher in neurons than in MRC5 cells, even at 7 days post infection. While transcription across the d109 genome was detected in neurons at 7 days post infection, no transcription was detected from many loci in MRC5 cells at this time as indicated by the gray shaded regions. This is not due to a lack of persisting d109 genomes, since it can be demonstrated that the provision of ICP0 can quantitatively “reactivate” the quiescent d109 genomes in MRC5 cells in this system [222].

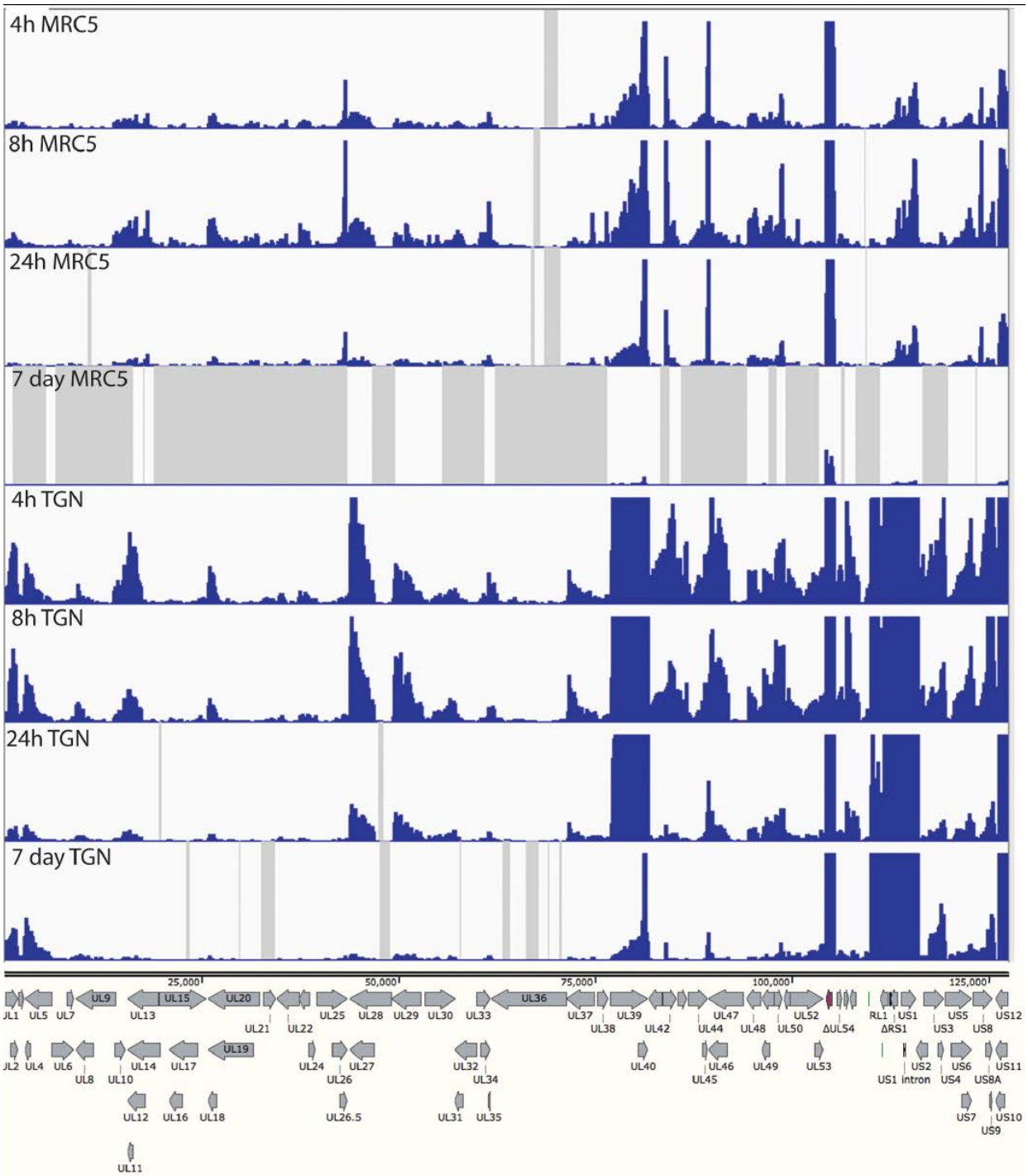

Figure 23: HSV transcripts synthesized in d109-infected MRC5 cells and TG neurons. MRC cells and cultures of TG neurons were infected with d109 (MOI = 10 PFU/cell) for 4h, 8h, 24h and 7d. The RNA synthesized in the cells was analyzed by RNAseq as before, and the reads were mapped to a modified d109 genome. The reads were aligned to the d109 genome. The GFP gene is shown in red in deleted ICP27 locus. The maximum on the y-axes of the graphs was set to 100 reads. The gray shaded areas indicate regions were there were no reads.

One interesting observation is the pattern of expression in neurons of genes just outside the long and short repeat regions. In the short unique region of the genome, expression of the US3/4, US5/6/7, and US8/9 clusters is evident, and decreases over the course of 24h. However, the abundance of the US3/4 and US8/9 increase from 24h to 7days, as that of US5/6/7 decreases, as most of the gene do. Likewise UL1/2 and UL4/5 mRNAs decrease over the course of 24 h, but increase at 7 days. This pattern is also in contrast to that of all the other HSV genes in U_L.

In neurons, the pattern of expression of the genes adjacent to the long and short repeat regions is also seen within the repeats. Figure 24A shows the region of the d109 genome from the LAT promoter proximal region through to US2. Figure 24A also shows the RNAseq reads in this region from d109 in neurons. These signals were off the scale in Figure 23. Note the only place in this region where there were no reads was in the US1 intron. The expression of this region overall was considerably greater than that of genes in the unique long and short regions (with the exception of UL39 and GFP). While the numbers of reads in this region generally decreased over 24h, they were sustained or increased from 24 h to 7 days. The signal just immediately downstream from the previously defined LAT mRNA start site that contains an ICP4 binding site [364, 365] increased considerably from 1 to 7 days post infection. The LAT start site is indicated by the arrow to the left in Figure 24A and by the 2 small arrows (previously defined by primer extension analysis [366]) in Figure 24B. Figure 24B also shows the ICP4 binding site and the TATA box for the LAT transcript along with the RNAseq reads for d109 in neurons at 1 and 7 days post infection. The elevated RNAseq signal begins precisely at the previously described LAT mRNA start site, strongly indicating that it is transcribed from the LAT promoter.

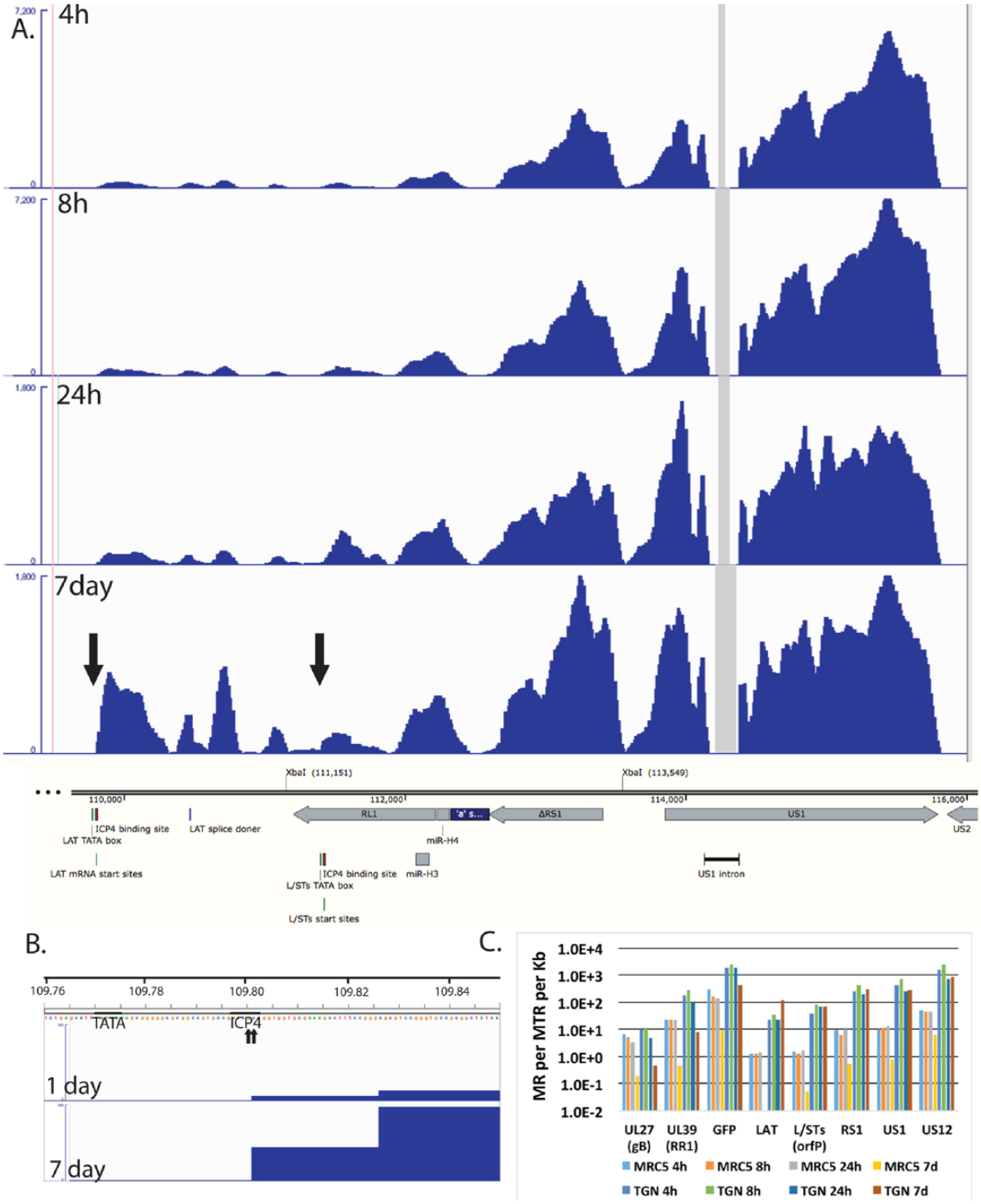

Figure 24: Transcription of the internal repeat region from persisting genomes.
A. The RNAseq reads from d109-infected TG neurons mapping to the region encoding the LAT promoter to US2 are shown relative to features in this region of the genome. The Xba1 site at 111,151 marks the locus of the ICP0 deletion and the site 113,549 marks the locus of the deletion that removed the VP16 responsive elements from the region.The numbers shown are in bp from the 5' end of the modified d109 genome in the parental orientation. The maximum of the Y-axes

for the 4 and 8h graphs was set at 7,200 reads, while those for the 24h and 7 day graphs were set at 1,800 reads. **B.** The RNAseq reads for the 1 and 7 day d109-infected TG neuron samples are shown in the region of the LAT TATA box, mRNA start sites, and ICP4 binding site. The maximum on the Y-axis is 100 reads. The beginning of the signal corresponds to arrow to the left in the 7 d graph of panel A. **C.** The quantification of the reads from the samples in figure 5 for select viral genes are shown.

The abundance of transcripts from this region of d109 in MRC5 cells and neurons were quantified and plotted in Figure 23C. The XbaI site at 111,151, Figure 24A marks the locus of the ICP0 deletion in d109 [190], therefore the sequence from the LAT mRNA start site to the XbaI site (111,151) is collinear with the 5'end of the LAT primary transcript [367]. The region quantified and designated LAT in Figure 23C corresponds to the region from the LAT mRNA start site to 1 kB downstream. The transcription of this region is quite complicated with the RL1 gene and the L/STs (orf P) transcripts antiparallel to each other. The L/STs TATA box, mRNA start site and ICP4 binding sites are indicated (Figure 24A) [368, 369]. For the sake of clarity and because there is a distinct increase in reads on the L/STs start site (right arrow), the quantification in Figure 24C represented as L/STs are the reads from the L/STs start site to 1 kb downstream. Also quantified in Figure 24C are ΔRS1, the second exons of US1 and US12, GFP, UL39, and UL27. Figure 234 shows several results; i. The abundance of all the transcripts in this region is higher in TG neurons than in MRC5 cells. ii. The abundance of all the HSV transcripts in MRC5 cells decrease with time such that the maximum abundance is less than 1 in 10^6 at seven days (US1) and LAT is undetectable (<1 in 2 x 10^7). iii. In neurons, the abundance of UL27, UL39 and GFP decrease from 1 to 7 days by ~10-fold, whereas the abundance of ΔRS1, US1 and US12, and LAT all increase from 1 to 7 days post infection, with LAT increasing the most. iv. The numbers of reads for LAT, L/STs, RS1, US1, and US12 reflect a transcript abundance for each that is between 1 in $3x10^3$ and 1 in $10^4$of all transcripts in the cell. This suggest that this region of the genome is considerable more transcriptionally active in TG

neurons than in MRC5 cells, despite repression of the remainder of the viral genome, even in the absence of viral activators.

2.6 DISCUSSION

Next generation sequencing has emerged as a high throughput methodology to quantify changes in transcriptome wide gene expression. RNA-sequencing has been shown to be more sensitive for detection of low abundance transcripts and to have a wider range of detection than traditional microarrays [370]. The goal of this study was to determine how the presence and absence of immediate early proteins impacted viral gene expression in neuronal (TG neurons) and non-neuronal cells (MRC5). In particular, we were interested in the expression from a virus that establishes a persistent quiescent infection in trigeminal neurons as a model for what might be occurring during latency. From the RNAseq experiments performed, we have made the following findings; i) during productive infection, there are deviations from the established cascade paradigm for early genes, ii) the level of expression of late genes largely a function of the TATA box and INR, iii) in the absence of immediate early proteins, transcription from the viral genome occurs at low level independent of kinetic class, iv) neurons are less restrictive to viral transcription than MRC5 cells, and v) genes near the joint/repeat region of the genome, and the LAT locus in particular, are more transcriptionally active during persistence in neurons.

Productive Infection Gene Regulation and Promoter Architecture

The promoters of HSV early genes are characterized by the presence of upstream elements for cellular transcription factors. Several early genes were quantified in Figure 20C and D. ICP8 and TK were abundantly expressed. The promoter for tk gene contains three distinct upstream elements, two Sp1 binding sites and one CCAAT element [342]. ICP8 also contains two Sp1 binding sites. The DNA polymerase promoter contains a degenerate TATA box, and one Sp1 binding site [371]. Expression of this transcript is reduced approximately 4 fold relative to tk and 2 fold relative to ICP8. UL42 was the most abundantly expressed gene, comprising approximately 1% of the total cellular RNA and peaking at 6 hours post infection, consistent with previous results [357]. The abundance of these 4 transcripts decreased late after infection.

The mRNA for the components of the helicase-primase complex UL5, UL8, and UL52 [372] and the origin binding protein UL9 [373] were expressed with similar kinetics and intensity during early time points prior to DNA replication. Expression of these four transcripts slowly accumulated throughout the time course to low levels. The UL8 and UL9 mRNA start sites have been mapped [374]. UL8 and UL9 both have degenerate TATA boxes and UL8 has a Sp1 site 22 bp upstream of the TATA box. UL 9 has been reported to have an upstream Sp1 site, however none are evident in the KOS sequence [374]. The UL5 and UL52 mRNA start sites have not been determined making it difficult to localize transcription signals, however the only potential TATA boxes within 400 bp of these genes consist of 4 A and T bp.

Therefore, genes with promoters possessing good TATA boxes, TATA boxes that matched the consensus sequence, and multiple sites for upstream factors are abundantly expressed early after infection and their decline later in infection is pronounced. This represents the early gene paradigm. However, it is difficult to fit some genes whose products are clearly required for DNA synthesis into this paradigm. Those with poor TATA boxes, lacking upstream

elements, are expressed at a low-level, their decline later in infection is not as pronounced, and in some cases accumulation continues at a low level (UL52, UL9). This is reminiscent of some previous studies with mutant and chimeric promoters [366, 375, 376].

Late genes (γ2) require viral DNA synthesis for their expression and their promoters consist of at TATA box and a initiator element [377]. The promoter elements of the late genes represented in Figures 20G and H are shown in Table 1. The expression of these genes varies considerably. There are 4 transcripts with abundance above 10,000 mapped fragments per million reads, representing the most highly expressed late genes. The three most highly expressed genes, UL35, UL45 and UL44, have a TATA box that completely matches the consensus sequence and an initiator element with a perfect match to the first 4 bases of the consensus. UL10 is expressed to a lesser extent, and has a mismatch in the TATA box. Genes expressed to levels below 10,000 have multiple mismatches the TATA box, the first 4 nucleotides of the initiator element, or a combination of both. There was no apparent correlation to the downstream activation sequence, though it likely plays an important role on a gene by gene basis. Therefore, the expression of late genes varies considerably, and their levels of transcription correlate with the degree of match to both the TATA box and INR element consensus, suggesting that the ability of TFIID to bind to the promoter is the main determinant in late gene transcription.

	TATA Box	*INR*	*DAS*
**Consensus*	TATAAA	YYANWYY	GAGNNYGAG
UL10	TAAAA	**G**CACAC	None
UL22	**A**ATAAAA	TCATA**AA**	GAGTGC**C**AG
UL34	TATAAA	**G**CAGAC	**C**A**C**GGCGAG
UL35	TATAAA	CTAATC**G**	None
UL37	TATAA**C**A	TC**G**CA**A**C	GAGGGC**CGC**
UL38	T**T**TAAA	CCAGTC**G**	GAGGCG**GT**AG
UL44	TATAAA	CTAC**CGA**	GAGGCCG**CA**
UL45	TATAAA	TCAT**C**CT	G**GC**GT**G**GAG
US8	**A**T**T**TAA	CT**T**CT**GG**	GAGCCC**A**A**C**

Table 1: Promoter elements of γ genes.
The bold letters designate mismatches between the specific gene element and the consensus sequence.

γ1 genes are transcribed early after infection and their synthesis increases after DNA replication. Four γ1 genes are presented in Figures 20 E and F. This group of genes is more uniform in its expression relative to the other groups of genes and is represented by a minimum of 1% of the total reads (virus + cell). The early and sustained high level of expression is most likely due to the presence of a TATA box INR and binding sites for upstream factors in their promoters [378]. The binding sites for upstream factors enables expression early after infection, and the INR allows for the continued expression late after infection.

We hypothesize that late after infection, cellular activator function and the activity of ICP4 on early promoters becomes less active. We have shown that Sp1 gets phosphorylated after DNA replication and is less active in vitro [81], and that TFIIA, which is required for the activation of tk by ICP4 in vitro, decreases late after infection [82, 84]. In addition, upstream activators require a form of the mediator complex for activation, and we have shown that as infection proceeds, ICP4 interacts with a form of the mediator complex that is often associated with repression of transcription [79]. Therefore, weaker promoters, such as UL8 or UL9, which are not activated to the same extent by upstream factors such as Sp1, might not show as dramatic

a decrease, or any at all, in activity late after infection. The sustained high-level activity of late genes is most likely a function of the ability of ICP4 to activate transcription through the INR sequence present at the start site of late genes [81, 88].

Gene Expression during quiescent infection.

HSV mutants engineered to not express viral activators can establish a quiescent infection in cells in culture such that their genomes are transcriptionally repressed and persist in cells [190, 348, 379]. Heterochromatin forms on d109 genomes persisting in MRC5 cells [191, 222], reminiscent of the structure of latent viral genomes [346, 380]. We have also shown that d109 genomes may not be as tightly repressed in TG neurons as in non-neuronal cells in culture based on GFP transgene expression and the ability to reactivate quiescent virus [381]. Therefore, we sought to compare the expression of the entire d109 genome in MRC5 cells, as a model for non-neuronal cells, and TG neurons.

Trigeminal neurons were significantly more permissive for the expression of genes across the persisting d109 genome than MRC5 cells, Figures 23 and 24. The MRC5 cells are presumably more homogeneous than these TG cultures, which are known to contain populations of neurons that differ in their permissiveness for LAT expression [352]. Therefore it is possible that some infected neurons are more permissive for HSV gene expression than others. Moreover, expression bore little relation to kinetic class. This “dysregulated” pattern of HSV gene expression has been seen in two different systems when reactivation stimuli were applied to latent HSV [217, 218]. In our experiments, the persisting genomes are void of genes that are thought to activate viral early and late genes, and since the infected neuronal cultures are maintained in NGF, known reactivation stimuli were not applied.

The elevated level of transcription of persisting vial genomes in neurons relative to non-neuronal cells most likely is due to the different transcriptional environment in neurons. In mammalian cells, there are three different histone H3 variants, H3.1 H3.2 and H3.3. During differentiation of neurons, the abundance of variants H3.1 and H3.2 decrease while H3.3 accumulates [382]. H3.1 contains a combination of both activating and repressive modifications, H3.2 is enriched in inactive marks, and H3.3 is enriched in activating markers and is localized to transcriptionally active regions [383-385]. The association of H3.3 with the viral genome has been shown to facilitate lytic gene expression [117]. Perhaps H3.3 facilitates transcription from d109 genomes in TG neurons, while H3.1 and 3.2 restrict gene expression in MRC5 cells. This remains to be studied.

While the d109 genome in general is more transcriptionally active in TG neurons than in MRC5 cells, a large majority of the transcripts in neurons at 7 days post infection map to loci around the joint regions and in the unique short region of the virus, Figures 23 and 24. There are several possible explanations for this. i) Neuronal specific promoters could be directing the expression of all these genes. With the exception of LAT, which is generally thought to possess a neuronal specific promoter, it is unlikely that all of these genes (UL1, 2, 4, 5, US1, 3, 4, 8, 9, 12 RS1, and LAT) have neuron specific promoters. The underlying basis for their enhanced expression more likely has to do with a general structural property of the region. ii) Perhaps the high G+C content of the internal repeats relative to rest of the genome results in reduced occupancy of repressive chromatin in TG neurons. iii) It is also possible that chromatin insulators, DNA elements that function to separate euchromatic and heterochromatic areas of the genome [386, 387] are involved. Seven distinct regions containing CTCF binding sites have been identified in the herpes genome [388]. These are all located in the repeat region of the genome

and 5 remain in d109. Two specific CTCF binding sites surrounding the LAT enhancer have been shown to function as enhancer-blocking elements and maintain the localization of euchromatin to the LAT locus [204]. The presence of these elements could explain the specific upregulation of LAT expression during persistence of d109 in neurons. Other insulator elements are located surrounding other immediate early genes near in the repeat regions [203]. The presence of these insulators could slow the accumulation of heterochromatin in the regions, enabling low-level transcription of IE genes in the absence of VP16.

The physiological consequences of the elevated low-level expression of genes in the joint region are not known. However, an intriguing possibility involves the interplay between HSV gene products and microRNAs encoded in this region [210, 389]. During latent infection, the primary latency associated transcript serves as a precursor to several miRNAs [210], which may target ICP4 and ICP0 [208]. The low-level expression of the IE genes and miRNAs in this region could provide the basis for the ability of the virus to reactivate in a regulated way, independent of viral activators.

3.0 FUNCTION OF HERPES SIMPLEX VIRUS TYPE 1 ICP4 ON QUIESCENT GENOMES

3.1 SUMMARY

During the process of reactivation from latency, the herpes simplex virus genome undergoes epigenetic changes to transition from a heterochromatic to a euchromatic state. Therefore, during process of reactivation, it is likely that viral activators function on genomes that are partially heterochromatic. ICP4 is promiscuous activator of polII transcription that functions in part by binding to the genome. To study the interactions of ICP4 with a heterochromatic genome, we used a quiescent model of latency in which a virus that does not express immediate early genes persists in cells and acquires epigenetic markers similar to *in vivo* latency, along with adenovirus vectors to deliver the activators ICP4 and ICP0 to quiescently infected MRC5 cells and TG neurons. Changes in transcription, chromatin and ICP4 binding, were quantified using RNAseq, ChIP and ChIPseq. We found that in MRC5 cells, ICP4 alone was not able to bind to the viral genome or induce transcription, suggesting that viral chromatin blocked access of ICP4 to the genome and prevented it from inducing the expression of viral genes. When ICP0 was present, ICP4 binding was significantly enhanced along with a decrease in histone H3 occupancy and an

increase in viral transcription. However, in neurons ICP4 was more efficient in inducing transcription from the quiescent genome in the absence of ICP0. These results suggest that: i. ICP4 cannot access and therefore cannot activate genomes repressed by heterochromatin, such as in quiescently infected MRC5 cells. ii. ICP0 can relieve the repression enabling ICP4 to bind to the genome and activate transcription. iii. Quiescent genomes in trigeminal neurons exist in a state that allows ICP4 to access the genome and activate transcription independent of ICP0.

3.2 IMPORTANCE

During the process of reactivation, the viral genome transitions from a heterochromatic to a euchromatic state. During this process, viral activators likely function on genomes that are partially heterochromatic. An HSV mutant defective in immediate early gene expression establishes a quiescent infection and acquires chromatin modifications similar to *in vivo* latency. We quantified gene expression induced from the quiescent state in both neuronal and non-neuronal cells following adenovirus mediated delivery of ICP4 and ICP0 by RNAseq and quantified ICP4 and histone occupancy on the viral genome by ChIP and ChIPseq. We found that ICP4 alone was able to bind to and activate transcription from quiescent genomes in neurons but not in non-neuronal cells, suggesting that neurons more loosely repress the genomes, possibly due a different type of chromatin deposited on the genome. The looser repression in neurons may allow for a more delicate regulation of reactivation where small increases in ICP0 or ICP4 expression are balanced by miRNA silencing to maintain the latent state.

3.3 INTRODUCTION

Herpes simplex virus 1 (HSV-1) establishes a latent infection, which is characterized by a relatively low abundance of viral transcription, in the ganglia innervating the primary site of productive infection [390]. The mechanism of viral genome silencing and repression has been an area of continuing study. Several early key observations identified epigenetics as a major regulatory mechanism, citing histone deposition and modification of histone tails as mechanisms for affecting the expression of latent genomes [187, 189, 346]. These modifications and their biological significance have been studied in relation to the regulation of viral transcription, but are still not fully characterized.

The epigenetic state of the viral genome differs between the productive and latent phases of infection. Lytic replication is characterized by relatively nucleosome free promoter regions, and genes with loosely spaced nucleosomes rich in histone acetylation [134, 135]. VP16 and ICP0 are thought to establish this euchromatic state early in infection through interactions with multiple cellular chromatin modifiers [140, 141, 380, 391-393]. However, when these viral proteins cannot counteract cell mediated silencing, the virus establishes a latent infection in which the histones on the viral DNA become deacetylated and methylated, and the viral genome is repressed in heterochromatin [388]. Markers of both facultative [118, 188] and constitutive [189] heterochromatin have been reported, as well as components of the polycomb repressive complex and HP1, which maintain the heterochromatic architecture [196]. Latency is not permanent, as cellular stress and other unknown environmental factors can induce reactivation [394, 395].

During reactivation of the virus, the viral chromatin becomes acetylated, dispersed, and lytic transcription resumes [203, 396], yet the mechanism underlying this process is not well

understood. Numerous studies have focused on the reactivation phenotype of different mutant viruses and have yielded varying results, likely due to the complex interplay between viral activators during this process. We and others have shown that ICP0 can derepress the viral genome through reversing constitutive heterochromatin marks, dissociating HP1, and increasing histone acetylation [222, 223, 397], which may or may not be required for other activators, including the main viral transactivator ICP4, to function. ICP4 is the main transcriptional activator of HSV-1 as it regulates the transcription of most viral genes [56, 77, 153]. It binds to the genome and interacts with and recruits multiple cellular transcriptional complexes including Mediator, TFIID, and chromatin remodeling proteins during lytic replication [79, 155].

In the early stages of reactivation, it is likely that viral activators, such as ICP0 and ICP4, function on genomes with heterochromatic character to both cause derepression of the viral chromatin and activate transcription. Previously, we have created and characterized a mutant virus, d109, which is repressed in heterochromatin similar to latency *in vivo* [190, 191]. To study these activators, we adopted an adenoviral expression system to deliver these proteins into d109 quiescently infected MRC5 cells and murine trigeminal neurons, since previous studies in our lab have shown a difference in d109 transcription between the two cell types. The aim of this study is to investigate the ability of ICP4 to bind to and activate transcription from a heterochromatic genome in neuronal and non-neuronal cells.

3.4 MATERIAL AND METHODS

Virus and Cells. Experiments were performed using MRC5 cells (human embryonic lung cells), which were obtained from and propagated as recommended by the American Type Culture Collection. Neurons were isolated from the trigeminal ganglia of CD1 mice as described below. The viruses used in this study were the IE mutant virus, d109, and several adenovirus vectors. The d109 virus was propagated on complementing F06F1 cells as previously described [190]. Adenovirus vectors expressing ICP0, ICP4, and beta-galactosidase were created using the Invitrogen pAd/CMV/V5-DEST system (V493-20), referred to as Ad0, Ad4, and AdZ respectively. The adenoviruses were E1 and E3 deficient and were propagated on the complementing 293A cell line.

Neuron Isolation. Six week old CD1 mice were euthanized and perfused with saline solution. The trigeminal ganglia were dissected and neurons were isolated following a previously established protocol [87]. Neurons from 10 mice ganglia were pooled and seeded into 24 well plates. Following establishment of cultures, the cells were maintained in Neurobasal-A media supplemented with 2% B27, 1% Penicillin-Streptavidin, .5mM L-glutamine, 50ng/mL nerve growth factor (NGF), glial cell-derived neurotrophic factor (GDNF), and neurturin. Media was changed once per week.

Infections. MRC5 cells were infected with d109 at a multiplicity of infection (MOI) of 10PFU/cell in TBS (recipe) at room temperature for 1 hour with rocking every 10 minutes. After adsorption, the inoculum was removed and DME (Dulbecco's modified Eagle's medium) containing 5% fetal bovine serum (FBS) at 37º was added. Infected cells were maintained at 37º for 24 hours. After 24 hours, the media was changed and the cells were moved to 34º to slow the

division of cells while quiescence was established. The media was changed at 3 and 6 days post infection. At day 7, the media was removed and saved and the cells were then superinfected with adenovirus expressing ICP0, ICP4, and B-galactosidase, termed Ad0, Ad4, and AdZ respectively as diagramed in Figure 25A. All adenovirus infections were done at a MOI of 10 PFU/cell in TBS. Following infection, the inoculum was removed and the saved media (kept at 37º) was added back to the cells. This point was considered time 0. Cells were harvested at the indicated times for transcriptional or epigenetic analyses. Infection of neurons was similar, except that they were done in neuron media instead of TBS to maintain NGF in the media. Additionally, the cells were washed with warm media following infection, and the first media change occurred at 6 days post infection. All infections were done with 10^7 PFU/mL using 100uL of inoculum.

ChIP. Chromatin immunoprecipitation (ChIP) was carried out as previously described, with some modifications [191, 222]. A total of 5e6 MRC5 cells were seeded into 100mm dishes and infected as described above. Cells were prepared for ChIP using the following protocol. Media was aspirated from cells and replaced with prewarmed crosslinking buffer: 5% media supplemented with 1% formaldehyde and .15% sodium bicarbonate. Crosslinking was done for a total of 10 minutes at 37°. The cells were washed 2 times with TBS + Protease inhibitors (Protease inhibitors (PI) are .4uM aprotinin, 2uM pepstatin, 0.15mM tosyllysinecholormethyl ketone hydrochloride (TLCK), and 1mM phenyl methyl sulfonyl fluoride (PMSF). Cells were then scraped into TBS+PI and centrifuged for 15 minutes at 4 degrees at 3000rpm. The supernatant was aspirated and the cells were resuspended in 500uL sodium dodecyl sulfate (SDS) lysis buffer (1% SDS, 50mM Tris-HCl pH8.1, 10mM + PI). The cells were incubated for 15 minutes on ice, transferred to Eppendorf tubes, and incubated on ice for 15 minutes. The cells

were sonicated for 6 10 second pulses at 30% output. The cells were then centrifuged for 15 minutes at 14000rpm at 4 degrees. The supernatant was diluted 1:11 in dilution buffer (.001% SDS, 1.1% Triton X-100, 167mM NaCl, 1.2mM EDTA, 16.7mM Tris-HCl pH 8.1 + PI). The samples were precleared by adding 120uL protein A agarose beads (Millipore) and rotating the samples for 2 hours, centrifuging 15 minutes 3000rpm, and collecting supernatant. This process was repeated for a total of 2 times. The samples were then added to 50uL of protein A agarose beads containing either H3 (Abcam ab1791) or N15 antibody. The samples were incubated with the beads overnight with rotation at 4º. The beads were washed twice for 5 min each in low-salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mMTris-HCl pH 8.1, 150 mMNaCl), twice for 5 min each in high-salt buffer (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20 mMTris-HCl pH 8.1, 500 mMNaCl), once for 5 min in LiCl buffer (0.25 mM LiCl, 1% NP-40, 1% deoxycholate, 1 mM EDTA, 10 mMTris-HClpH 8.1), and three times for 5 min each in Tris-EDTA (TE) buffer (10 mMTris-HClpH 8.0,1 mM EDTA). The samples were eluted from the beads in two steps with 250 μl of elution buffer (1% SDS, 0.1 M sodium bicarbonate) and rotation for 20 minutes. Twenty microliters of 5 M NaCl was added to each sample, and the cross links were reversed overnight through incubation at 65 degrees. Thirty-one microliters of a proteinase K mixture (1:2:1 0.5 M EDTA–Tris-HCl pH 6.5–20 mg/ml proteinase K) was added to each tube, and the samples were incubated at 55°C for 2 h. The DNA was purified using phenol-chloroform-isoamyl alcohol (25:24:1) extractions. The purified DNA was precipitated with the addition of 1mL 100% ethanol, washed with 70% ethanol and resuspended in 60uL water. The immunoprecipitated DNA was quantified through qPCR.

ChIPseq: immunoprecipitated DNA was processed for the generation of Illumina sequencing libraries according to the Illumina TruSeq ChIP preparation protocol. ChIPseq libraries were sent to the Tufts University Core Facilty, Boston, MA, for sequencing. The analysis of the sequencing data was similar to that described previously [87] except that Bowtie was used instead of TopHat.

RNA isolation. MRC-5 cell were seeded into 60mm dishes at a density of 2e6 cells per dish and infected with d109 at an MOI of 10 PFU/cell. The adenovirus infections were the same as described above. RNA was isolated with the Ambion RNaqueous-4PCR kit by following the included protocol. RNA was harvested by adding 500µl lysis/binding buffer to MRC5 cell cultures and by adding 100µl lysis/binding buffer per well to neuronal cell culture. For the neuron cultures, 6 wells were pooled per sample. The following steps are the same for both neuron and MRC5 samples. The cells were scraped into the lysis buffer and collected in a 1.5mL Eppendorf tube. An equal volume of 67% ethanol was added. The solution was pipetted into a filter spin column and centrifuged at 13000 rpm for 1 min. The column was washed with wash buffers 1 and 2/3 and centrifuged dry to remove residual wash buffer. RNA was eluted in two steps with 60µl and 20µl 75°C elution solution. The RNA was treated with DNase I at 37°C for 30 min to degrade remaining DNA. The DNAse was inactivated with the supplied reagent in the kit.

Reverse Transcription and qPCR. Two micrograms of RNA were mixed with 1µl 10pM OligodT Primer in a total volume of 11µl. The tube was incubated at 65°C for 2 min to remove RNA secondary structure. Following the incubation, the tube was cooled and the following

reagents were added: .5µlRiboguard RNAse inhibitor, .5µl MMLV High Performance Reverse Transcriptase (200U/µL), 2µl 100 mM dithiothreitol (DTT), 4µl nucleotide mix (2.5mM concentration of each dNTP) and 2µl 10x reaction buffer. The RT reaction was carried out for 1 h at 37°C. Following completion of the RT reaction, the reaction tube was incubated at 85°C for 5 min to inactivate the reverse transcriptase.

RNA was diluted to 1µg in 60µl and cDNA was diluted 1:6 in ultrapure water. A master mix containing 0.3µl of each primer (stock concentration, 100 µM), 5µl Applied Biosystems SYBR green mix with 1.0µM 6-carboxy-X-rhodamine and .4µl of water for a total of 6µl for each reaction was made. The primers used in this study were: VP16ds, TKds, DNApolds, ICP8ds, gCds, and RR1ds. The sequences of the primers used are:

RR1ds (F: CGATCCAGATTCCAAAGTGCCCR: CGGTTTCCAGCCAGCGGTTA),

VP16ds (F: CAGGCCCTCGATGGTAGACCCGTAAR:

GGTGTTGATTCGGGCGAAGTTGGAC),

ICP8ds (F: CGAACTTGCGGGTGCGGTCAAAR: CATGGTCGTGTTGGGGTTGAGCATC)

DNA polds (F: GCTTCGATATCGAATGCAAGGCGGR:

CGAGCGAAAACAGGAGGACGTGCTC)

TKds and gCds primers were the same primers used in previous studies [191]. A 96 well plate was prepared with 6µl master mix and 4µl of sample or standard. All samples were run in triplicate. D106 viral DNA was used to create a standard curve of 1:10 dilutions from 4000000 to 40 copies, which covers the threshold cycle value range for the samples tested. qPCR was run on a StepOne Plus real-time PCR machine. The conditions for the run were 95°C for 10 min and 40 cycles of 95°C for 15 s and 60°C for 1 min. At the end of the run, a dissociation curve was

completed to determine the purity of the amplified products. Results were analyzed using the StepOne v2.1 software from Applied Biosystems and compiled in Microsoft Excel.

RNAseq: RNA was isolated using the procedure described above and prepared for deep sequencing according to the manufacturer's instructions (Illumina) and as described previously [87].

3.5 RESULTS

Optimization of Infection

In this study, adenoviruses expressing ICP0, ICP4, and Beta-galactosidase driven by the HCMV immediate early promoter were used to infect MRC5 cells and TG neurons that were quiescently infected with d109 to determine how ICP4 in particular functions on quiescent genomes repressed by chromatin. Time courses of protein expression from the adenoviruses were performed to determine the kinetics of protein expression (data not shown). Using this scheme, we next showed that the simultaneous presence of ICP4 and ICP0 did not greatly affect the abundance of either protein relative to the single infections (Figure 25B). While ICP0 slightly enhanced the transcription of ICP4, (Figure 25C) it did not lead to an increase in ICP4 protein abundance (Figure 25B). ICP4 did cause a small reduction in the amount of ICP0 protein expressed (Figure 25B).

As a first step to determine the effects of adenovirus-expressed ICP4 and ICP0 on the expression of quiescent d109 genomes, MRC5 cells were infected with d109 and the indicated

adenoviruses as outlined in Figure 25A. The HCMV immediate early promoter drives the expression of GFP in the d109 genome. The cells were visualized by fluorescence microscopy to monitor the expression of GFP, Figure 25D. Very few GFP-expressing cells were found in the mock, AdZ-, and Ad4- infected cultures. Consistent with previous results [222], infection by Ad0 resulted in the induction of GFP expression. The simultaneous presence of ICP4 did not significantly affect the abundance of GFP.

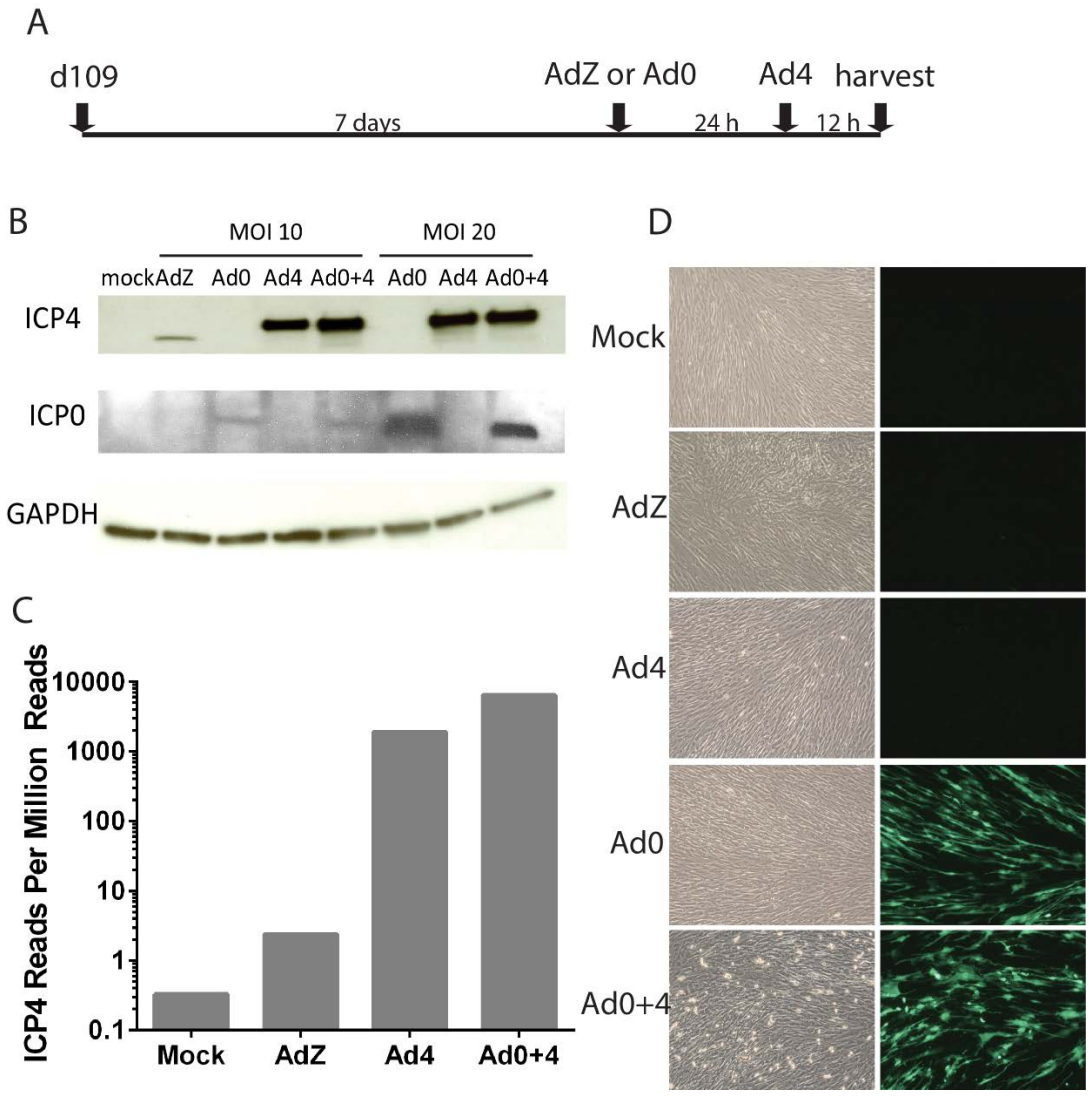

Figure 25: Adenovirus Infection Protocol
A. The schematic for the infection protocol is shown. MRC5 cells or neurons were infected with d109 for 7 days to establish quiescence. Cells were then superinfected with Adenoviruses expressing ICP0, ICP4, or LacZ. Time points were taken at 36 hours for Ad0, AdZ and mock,

and at 12 hours for Ad4. **B.** Western blot analysis of protein expression from adenovirus vectors, with GAPDH as a loading control. **C.** Quantification of ICP4 RNA expression from adenovirus determined by RNA-seq analysis. **D.** Phase contrast and fluorescence microscopy images of quiescent cells superinfected with adenovirus. GFP is constitutively expressed in the absence of epigenetic repression.

Activation of Quiescent Genomes in MRC5 cells by ICP4

MRC5 cells were infected with d109 for 1 week and then superinfected with adenovirus as described above. RNA was extracted and prepared for Illumina sequencing. Individual reads were mapped to a modified d109 genome in which duplicated sequences and sequences that were also present in the adenovirus vector were removed. The aligned reads are displayed on equivalent scales and are shown in Figure 26A. The gray regions represent areas that are devoid of reads. GFP and areas in the joint region, located around 108kb and 115kb respectively, were robustly transcribed in the Ad0 and Ad0+4 samples and therefore their signals fall outside of the range of the scale presented. The scale was chosen so that lower abundance reads could be easily visualized.

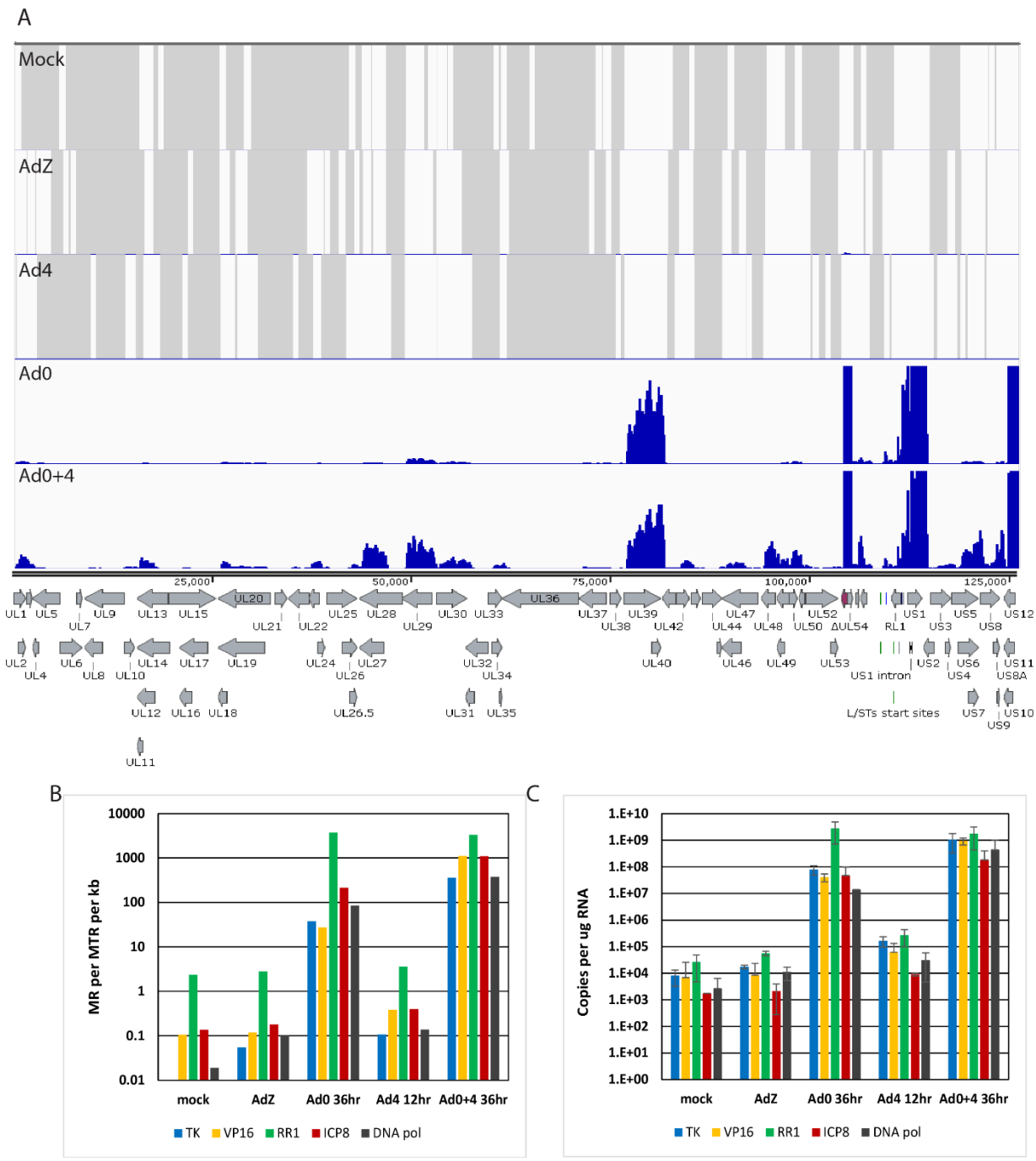

Figure 26: RNAseq of Quiescently Infected MRC5 Cells Superinfected with Adenovirus
Monolayers of MRC5 cells were infected with d109 at a MOI of 10 PFU/cell. Quiescence was established for 1 week and the cells were then superinfected with adenovirus at a MOI of 10 PFU/cell. Cells were harvested at 12 hours post infection for Ad4, and at 36 hours for Ad0, AdZ, mock, and Ad0+4 staggered. cDNA was prepared and subjected to Illumina sequencing as described in the Materials and Methods. Illumina reads were processed and aligned to a modified d109 sequence using the TopHat mapper in the GalaxyCloud software package. **A.** Reads from the RNAseq are shown mapped to the d109 genome. The reads were normalized to the total

number of reads to reduce variation from read density during sequencing. Grey regions represent areas of the genome completely devoid of transcription **B**. Quantification of a few representative genes from the RNAseq data set using the cufflinks software package. The Y axis is graphed as Mapped Fragments per Kilobase pair per Million Total Reads (MR per MTR per Kb). **C.**qRT-PCR quantification of viral transcripts.

The viral genome was essentially silenced during quiescence in MRC5 cells, as evidenced by the near lack of viral transcription, Figure 26A mock, consistent with previously published data [87]. When AdZ was added, there was no significant change in the expression of viral genes, evidenced by a lack of change in the mapped reads profile and an insignificant increase in the percent viral reads mapped, demonstrating that the viral vector alone did not induce changes in transcription. When Ad4 was added alone, it also did not induce any noticeable changes in gene expression. Therefore, there was no difference between the read profile of mock, AdZ and Ad4 infected cells, indicating that ICP4 alone could not activate the quiescent genome in MRC5 cells.

However, gene expression increased following the expression of ICP0, consistent with previous studies. The percent mapped increased to 8.2%, with ribonucleotide reductase, GFP and genes around the joint being the most abundantly expressed, Figure 26A and 26B. When Ad4 was added to quiescent cells that had been infected with Ad0, viral transcription increased above levels seen with ICP0 alone, Figure 26A. Components of the DNA replication machinery were highly induced by ICP4 including UL29, UL30 and UL42. Some leaky late genes, most notably UL19 (ICP5), UL27 (gB), UL48 (VP16), US6 (gD), US7 (gI) and US8.5/US9 were also induced. However, true late gene expression was not significantly induced, as seen in the traces for UL36 (VP1/2), UL41 (vhs) and UL44 (gC).

The cufflinks software was used to quantify expression of a few viral genes, TK, VP16, ICP8, UL39 (RR1), and DNA polymerase, Figure 26B. The most highly expressed gene in Ad0

and Ad0+4 treated cells was ribonucleotide reductase (UL39), however ICP4 did not induce its expression above ICP0 alone. Some increases in viral transcription between AdZ and Ad4 treated cells could be detected, however, since the number of viral reads was so few, these changes were not statistically significant. RT-PCR was used to validate the cufflinks data and is shown in Figure 26C. Increases in gene expression between mock and Ad0 samples were consistent between the two methods. For TK, VP16, ICP8 and polymerase the fold increase in mRNA resulting from the addition of ICP4 on top of ICP0 was approximately 10-20 as measured by both mapped reads and RT-PCR.

The data in Figures 26B and 26C also demonstrate a small increase in the abundance of TK, ICP8, VP16 and pol mRNA by the addition of ICP4 alone relative to mock superifected cells. However, the absolute level of transcripts is many orders of magnitude less than with ICP0 and ICP0+ICP4. We believe this represents transcription that is occurring in small subpopulation of cells and will be discussed in the next section.

Chromatin restricts the binding of ICP4 to the quiescent genome in MRC5 cells

The inability of ICP4 to activate quiescent genomes on a global scale is most likely related to its ability to access and/or function on the quiescent genomes. To investigate this, MRC5 cells were infected with d109 and the adenovirus vectors as done in the gene expression experiments above. Cells were harvested and ChIP was performed for ICP4 and H3 abundance at the TK and gC promoters, representative early and late gene promoters, respectively. ICP4 binding to the genome was undetectable above background at both promoters when ICP4 alone was present, Figure 27A&B. However, when ICP0 was present, ICP4 association with the viral genome significantly increased. H3 occupancy remained high when the Ad4 virus was added,

but in the presence of ICP0, H3 occupancy decreased at both promoters, Figure 27C&D, consistent with previous results [222]. Therefore, the binding of ICP4 was inversely correlated to histone 3 occupancy, suggesting that the presence of chromatin on the viral genome blocked the binding of ICP4 and that ICP0 enabled ICP4 to activate quiescent genomes by removing repressive chromatin.

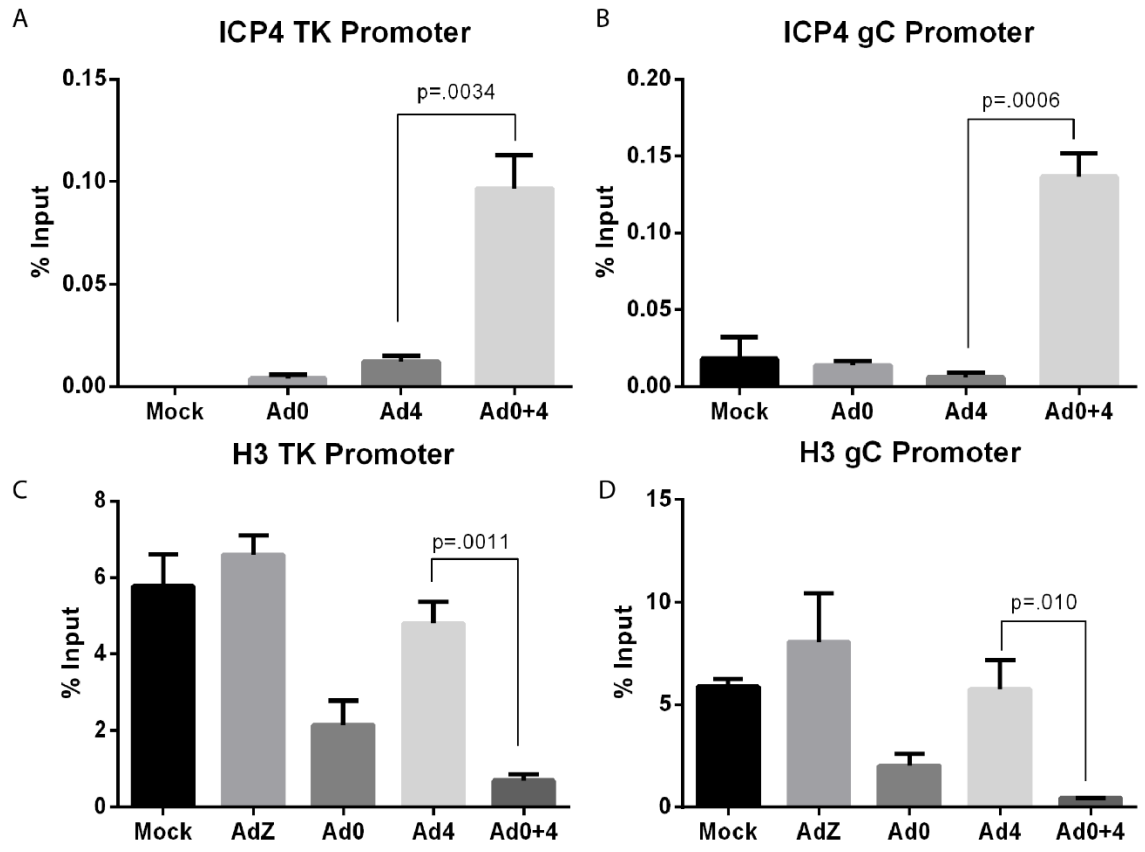

Figure 27: ICP4 binding and Histone Occupancy on Quiescent Genomes in MRC5 Cells
The four panels are from three independent chromatin immunoprecipitation experiments on adenovirus superinfected quiescently infected cells. An ANOVA analysis was run to determine significance between the multiple conditions and then a t-test was done to determine significance between Ad4 and Ad0+4 treated samples. The data is shown as a percent immunoprecipitated relative to the input. The promoters analyzed were thymidine kinase (TK), a representative early gene and glycoprotein C (gC), a representative late gene. **A.** H3 IP on TK promoter **B.** H3 IP on gC promoter **C.** ICP4 IP on TK promoter **D.** ICP4 IP on gC promoter.

The binding of ICP4 was also measured by ChIPseq in which immunoprecipitated chromatin was prepared for Illumina sequencing. Reads were mapped to the viral genome and

are shown in Figure 28. The grey regions represent areas devoid of reads and hence ICP4 binding. When ICP4 was expressed alone, it was unable to bind to the genome, as there was no difference in peak density between the mock and Ad4 treated samples. However, when ICP0 was present, ICP4 binding increased genome wide, confirming the ChIP results.

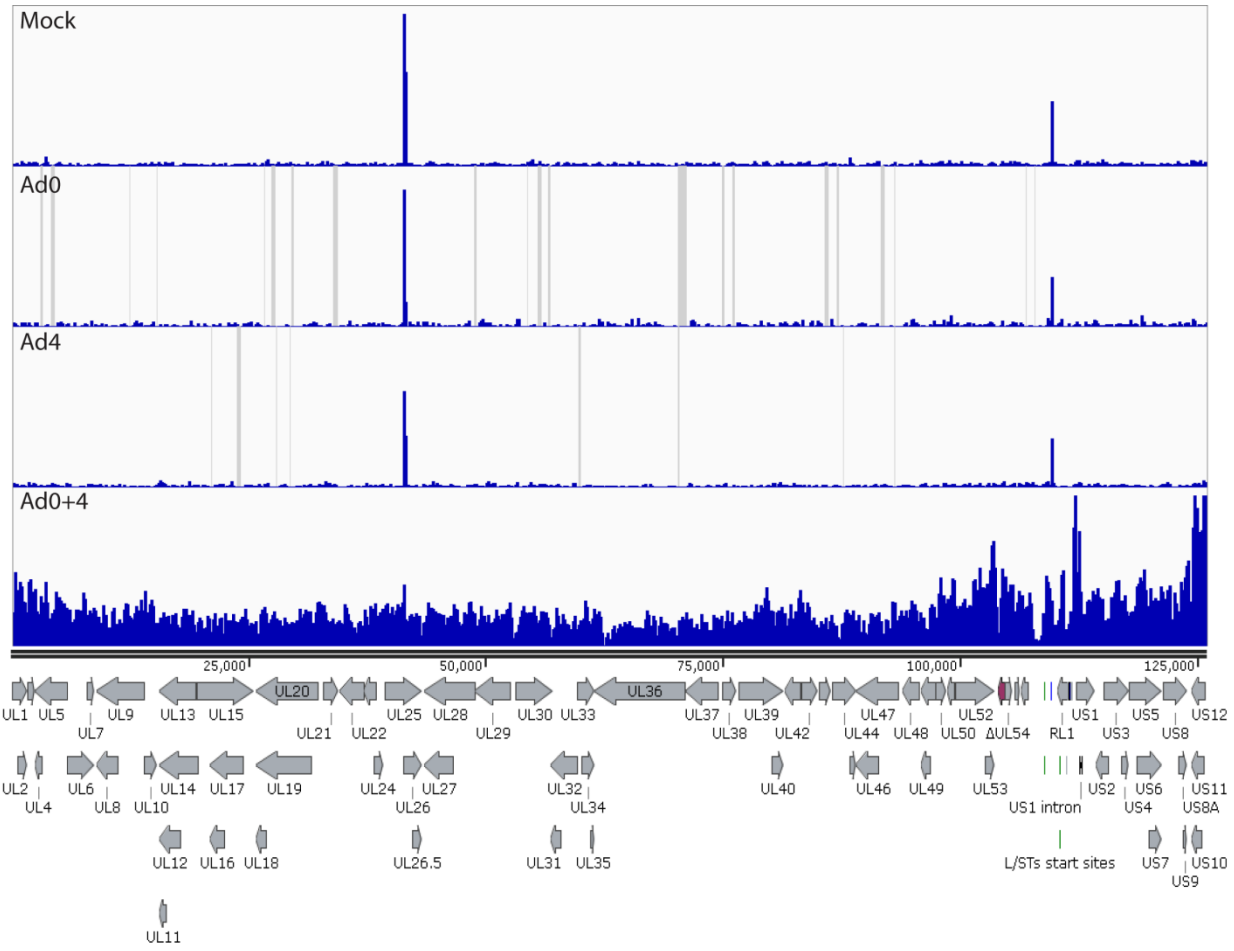

Figure 28: **Genome Wide ICP4 Binding to Quiescent Genomes in MRC5 Cells**
Chromatin immunoprecipiation was used as before to isolate DNA bound by ICP4 following adenovirus superinfection of quiescently infected cells. The chromatin was prepared for alumina sequencing and aligned to the viral genome. Grey areas represent areas devoid of ICP4 occupancy. The prominent peaks in the Mock, Ad0, and Ad4 samples correspond to poly T regions that likely map cellular reads.

Activation of quiescent genomes in neurons by ICP4.

Since previous studies showed that neurons were relatively more permissive for viral transcription during quiescence, we wanted to determine if ICP4 alone could activate quiescent

d109 genomes in neurons. Therefore, the experiment depicted in Figure 25A was repeated in primary murine adult trigeminal neurons. The mapped reads aligned to the d109 genome are shown in Figure 29. When Ad0 was added to quiescently infected neurons, viral transcription was enhanced, similar to that seen in MRC5 cells, Figure 29. When ICP4 was added to ICP0 expressing cells, viral transcription was greatly enhanced. The induction of viral genes by ICP4 in the presence of ICP0 was more robust in neurons than MRC5 cells; while in MRC5 cells, viral transcription only increased by 1%, in neurons the percent viral transcripts increased from 2% to over 7% (data not shown).

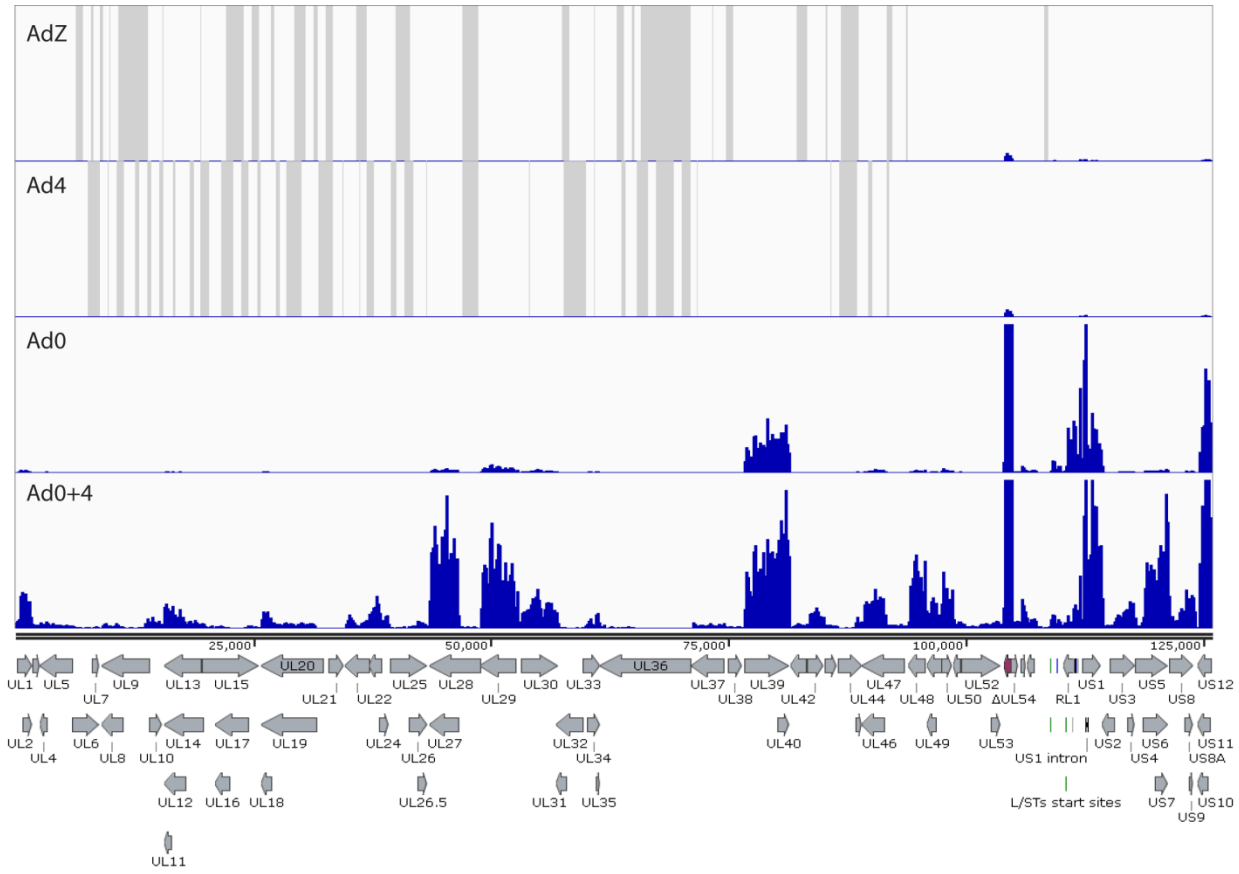

Figure 29: RNAseq of Quiescently Infected TG Neurons Superinfected with Adenovirus
Trigeminal murine neurons were infected with d109, superinfected with adenovirus, and harvested as done with MRC5 cells. RNA was prepared for Illumina sequencing. Reads were mapped to the modified d109 genome as before.

At 24 hours, ICP4 functioned to more efficiently induce transcription from the quiescent genome in neurons, but not in MRC5 cells, despite a higher ICP4 transcript abundance in MRC5 cells, Figure 30B. The mapped reads between mock and Ad4 infected cells are shown in Figure 30A. The reads for mock and Ad4 in each cell type have been normalized to account for differences in the number of reads and the scales been modified to show similar profiles in the mock conditions so that differences in the Ad4 treatment can be visualized. The grey regions are areas of the genome devoid of transcripts. In neurons, provision of ICP4 led to an increase in viral transcription across the genome as the density of grey regions decreased, indicating more of the genome was being expressed. Additionally, several loci on the genome had notable increases in the density of viral reads. There were more mapped reads at UL1-4, UL27, UL29, UL30, UL39, UL40, and UL46, Figure 30A. However, in MRC5 cells, Ad4 did not lead to an increase in transcription, since the mapped read profiles of the mock and Ad4 treated cells were nearly identical. Additionally the fold increase between mock and Ad4 treatment was calculated for neurons and MRC5 cells and is shown in Figure 30C. The fold increase in viral reads is much larger in neurons than MRC5 cells, consistent with the read profiles.

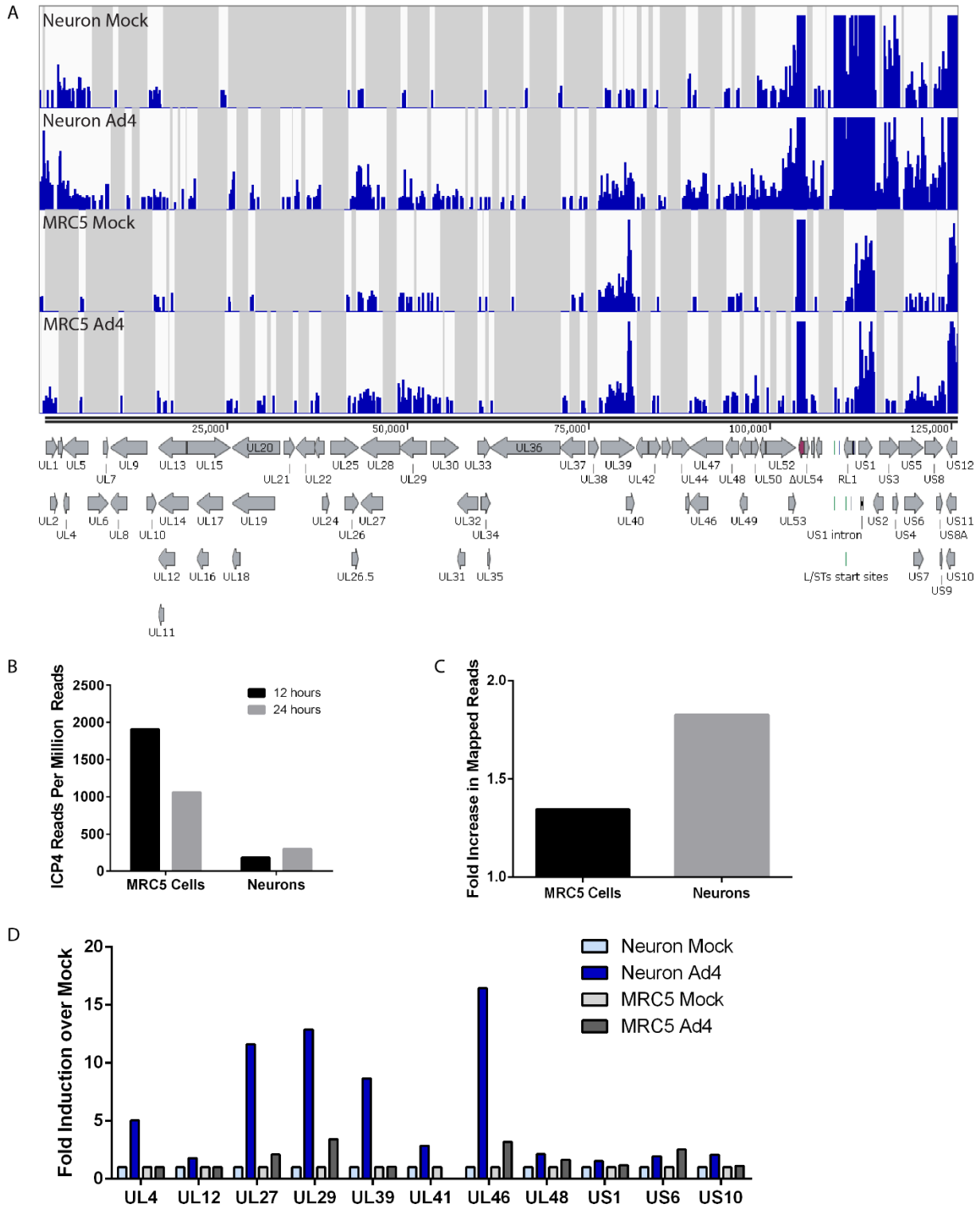

Figure 30: Activation of Gene Expression by ICP4 in TG Neurons and MRC5 Cells
Viral gene expression 24 hours post infection from quiescently infected MRC5 cells and neurons mock and Ad4 infected. **A**. Reads mapping to the viral genome. The reads for mock and Ad4 in each cell type have been normalized to account for differences in the number of reads. The scales

for neurons and MRC5 cells have been modified to show similar profiles in the mock conditions so that differences in the Ad4 treatment can be visualized. The grey regions are areas of the genome devoid of transcripts. **B**. Number of reads mapping to the ICP4 locus normalized to the total number of reads. **C.** Fold increase in the percentage of viral reads, calculated by dividing the percent reads of Ad4 by mock. **D.** Cufflinks analysis of viral transcripts. The mock samples were set to 1, and the values for the Ad4 conditions were obtained by dividing the number of reads in Ad4 by mock to determine the fold increase in expression.

The cufflinks software package was used to quantify the number of viral transcripts that mapped to each viral locus. The fold increase in gene expression between mock and Ad4 treatments is shown from a few viral loci in Figure 30D. In neurons, ICP4 induced a statistically significant increase in the expression of several genes including UL4, UL12, UL29, UL39, UL41, UL46, UL48, US1, US6, and US10 along with several other genes (data not shown). The genes activated were mainly from the early and leaky late gene classes. Activation of late genes was not observed and was likely due to insufficient DNA replication. In MRC5 cells, the only transcript that was statistically significantly upregulated by ICP4 was UL29. ICP4 induced the expression of many viral genes in neurons but not in MRC5 despite the higher transcript abundance in MRC5 cells relative to neurons, demonstrating that ICP4 can function on quiescent genomes in neurons, but not in MRC5 cells.

Binding of ICP4 to Quiescent Genomes in Neurons

ICP4 did not activate or detectably bind to quiescent genomes in MRC5 cells. However, ICP4 alone was able to activate some viral genes to a relatively small extent in cultured neurons. Therefore, we hypothesized that ICP4 could bind to the quiescent genome. To test this, ChIP was performed on quiescent neurons to assess the binding of ICP4 in the presence and absence of ICP0, Figure 31. ICP4 bound to the genome in the absence of ICP0 at the TK promoter, but not the gC promoter. When ICP0 was present, the binding of ICP4 was seen at both the tk and gC

promoters. Therefore, ICP4 bound to the quiescent genome to some extent in the absence of ICP0 to activate transcription from the quiescent genome. As in MRC5 cells, the presence of ICP0 further promoted the binding of ICP4 to increase transcription.

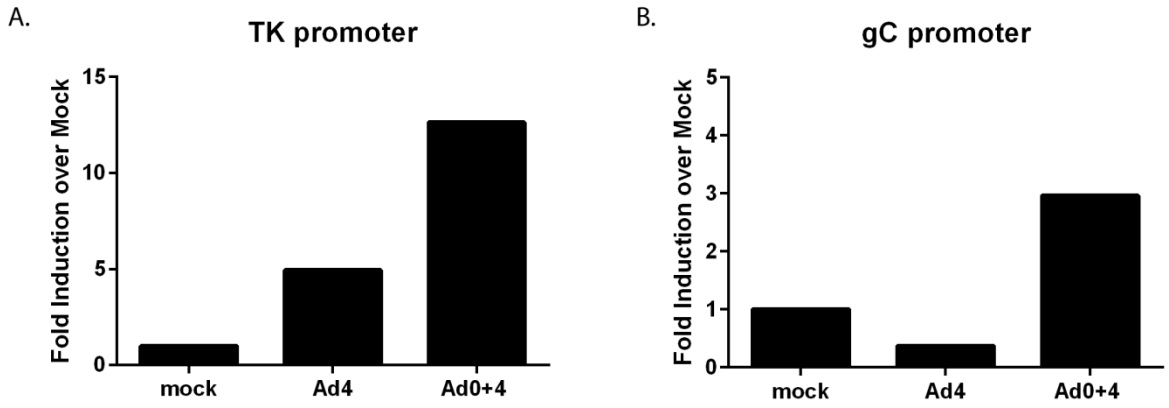

Figure 31: ICP4 Binding to Quiescent Genomes in TG Neurons
Association of ICP4 with the A. TK and B. gC promoters, representative early and late gene promoters respectively, in quiescently infected TG neurons.

3.6 DISCUSSION

We have previously established RNAseq as a viable method to quantify transcription from the herpes genome during both productive and quiescent infection and have shown that cell type dependent differences in transcription from the quiescent state exist [87]. In these studies, we investigated the ability of the viral transcription factor ICP4 to bind to and initiate transcription from the quiescent state in MRC5 cells and TG neurons in the presence and absence of ICP0.

Since neurons were more permissive to d109 transcription, we hypothesized that ICP4 may by more efficient in activating transcription from quiescent genomes in neurons than in MRC5 cells. From these experiments, we have made the following conclusions: 1) ICP4 alone cannot induce transcription from the repressed genomes in MRC5 cells since it cannot access to quiescent genomes, 2) ICP0 enables ICP4 to bind to and transactivate viral genomes through inducing the removal of repressive chromatin, and 3) ICP4 can induce transcription from the quiescent genome in neurons in the absence of ICP0 since it can access these genomes in the absence of ICP0.

ICP4 requires the presence of ICP0 to activate transcription from quiescent genomes in MRC5 cells

During the establishment of quiescence, the viral genome acquires cellular histones that are modified through trimethylation of H3 at lysines 9 and 27, markers of heterochromatin [191]. Simultaneously, histone acetylation, a marker of active chromatin, decreases[346]. This epigenetic repression of the viral genome during latency may prohibit the binding of ICP4. ICP4 binds to the viral genome through multimerization on several low affinity binding sites [77] and recruits cellular transcription factors and complexes to promote viral gene expression. Previous studies have shown that ICP4 can form complexes with multiple chromatin remodeling proteins, including the SWI/SNF and Ino80 complexes [79]. The Ino80 complex has the ability to slide nucleosomes and eject histones [398]. It has also been shown to be involved in transcription through chromatin and maintenance of an open chromatin state [399, 400]. The human SWI/SNF complex has been associated with both transcriptional activation and repression [401]. Despite its ability to interact with chromatin remodeling complexes, ICP4 did not activate transcription

from the quiescent state in MRC5 cells. RNAseq of Ad4 infected quiescently infected cells indicated no activation of quiescent genomes by ICP4, compared to AdZ control and mock infections, Figure 26A. The inability to activate transcription was due to the inability of ICP4 to bind to quiescent genomes. When ICP4 was present alone, ChIP at 2 viral promoters did not show association of ICP4 above background levels, Figure 27. Furthermore, when the association of ICP4 with the genome was assayed genome wide using ChIPseq, there was no ICP4 occupancy detected on the genome above background levels, Figure 28. The ChIP and ChIPseq data indicate that ICP4 cannot gain access to quiescent genomes.

While ICP4 cannot activate transcription from genomes that have been silenced by the cells, it will function on genomes that have not been repressed. Previous studies using d109 have shown that GFP expressing cells can persist as long as 28 days [190]. GFP expressing MRC5 cells can be seen in culture after 7 days, Figure 25A, but constitute a very small percentage of the total culture of less than .1%. As GFP is a marker of the epigenetic state of the viral genome, this population of GFP-expressing cells likely originates from viral genomes that have not been repressed by cellular heterochromatin or have been repressed to a much less extent. Detection of transcription from these cells using RNAseq is difficult since the vast majority of transcripts are cellular; in both AdZ and Ad4 treated cells, only a few hundred of the twenty million reads were virally derived. However, RT-PCR is transcript specific, and can detect very small changes in transcript abundance. RT-PCR analysis of a few viral genes showed that ICP4 alone induced the expression TK and VP16, (data not shown). Since this transcription is very low, and since ICP4 cannot access genomes bound in chromatin, these transcripts are likely derived from the population of GFP expressing cells whose genomes have not been repressed in chromatin.

ICP0 enhances ICP4 transcription in MRC5 cells through epigenetic derepression of the genome

Although ICP4 was unable to transactivate repressed genomes, it induced transcription from the quiescent genomes when ICP0 was present, Figure 26A. Studies have shown that when ICP0 is provided to quiescently infected cells, epigenetic silencing of the virus was removed: levels of trimethylated H3K9 decreased, histone occupancy was reduced, and histones remaining on the genome became more acetylated [222]. We found that ICP4 alone could not access the viral genome, but in the presence of ICP0, the binding of ICP4 to the viral genome was greatly enhanced. Additionally, the binding of ICP4 to the viral genome was inversely correlated to histone occupancy. When ICP4 was present alone, histone occupancy on the genome remained high, however, in the presence of ICP0, histone occupancy was significantly reduced. These results demonstrate that chromatin restricts the binding of ICP4 to the genome during quiescence and that ICP0 enables ICP4 to function through epigenetic derepression of the genome.

ICP4 Activates Quiescent Genomes in Neurons

Previous studies have shown that TG neurons are more permissive than MRC5 cells to transcription from the d109 genome during quiescence. We had hypothesized that this difference could be either due to differences in the epigenetic landscape of the cells or from the presence of insulator elements in the genome [87]. Insulator elements consist of core sequence of CCCTC that is bound by the CTCF protein [199]. Global studies of the human genome have shown CTCF elements to be a site of demarcation between active and repressed regions of chromatin [202]. Recently, seven CTCF elements have been identified in the HSV genome and are concentrated near the joint regions of the genome [203]. These insulator regions have been

shown to maintain chromatin boundaries between LAT and ICP0 during latency [204]. The presence of insulators would create a euchromatic region near the LAT locus, leaving the remainder of the genome largely heterochromatic. We have shown the regions near the insulator elements were more highly expressed in quiescently infected neurons [87], suggesting that these elements are responsible for the higher transcription near the joint regions in the genome. In these studies, however, there appears to be no bias as to what regions of the genome ICP4 can transactivate. ICP4 activated transcription globally in quiescent neurons, Figure 30A. There were increases in transcription near the joint regions and also within the unique long region, far away from the CTCF sites. While the CTCF elements may maintain higher transcription during quiescence, chromatin outside those regions did not prohibit activation by ICP4. This suggests that differences in repression efficiency exist between the two cell types.

The ability of ICP4 to activate transcription from quiescent genomes in neurons but not MRC5 cells is likely due to a difference in the epigenetic landscape of the two cell types. We propose a model, Figure 32, in which MRC5 cells have restricted the binding of ICP4 through constitutive heterochromatin occupancy on the genome, whereas in TG neurons, the genome is repressed in facultative heterochromatin, a less restrictive form of chromatin. This difference in chromatin structure may be due to differential expression of histone variants in the two cell types. The main H3 variants in non-neuronal cells (MRC5) are H3.1 and H3.2, which have been shown to be enriched in repressive markers [402]. However, during the differentiation of neurons, the expression of histone variants H3.1 and H3.2 become down regulated while H3.3 expression is up regulated, making H3.3 the main histone variant present in neuronal chromatin [403, 404]. H3.3 has been shown to be associated with transcribed regions of the genome and is

absent from the inactivated X chromosome [402]. The different histone variants may also influence the type of heterochromatin deposited onto the quiescent genome.

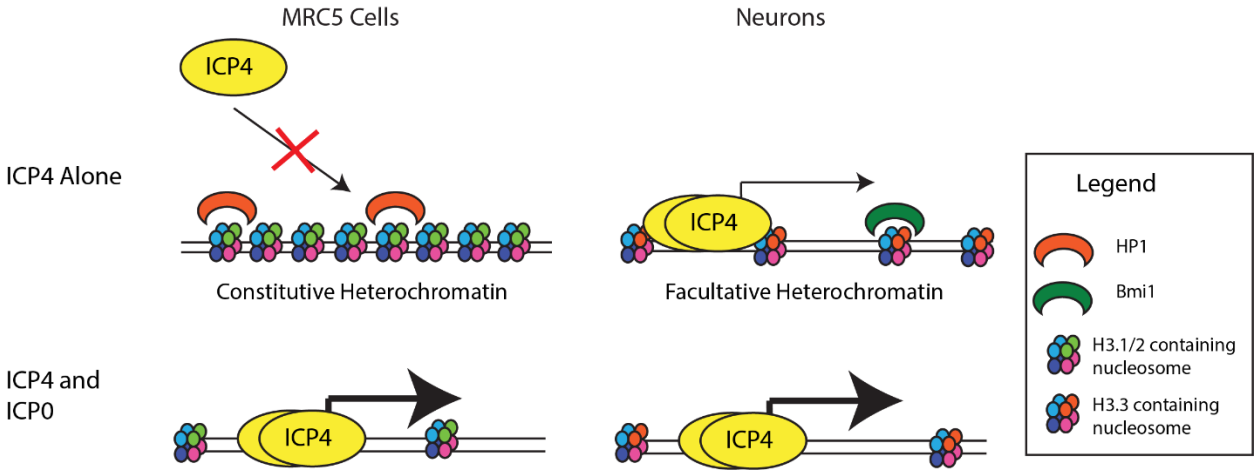

Figure 32: **Model of ICP4 Action on Quiescent Genomes in MRC5 cells and TG Neurons.** In MRC5 cells, constitutive heterochromatin blocks access of ICP4, preventing transcription of viral genes. When ICP0 is present, it derepresses and disperses chromatin, enabling binding and function of ICP4. In TG Neurons, facultative chromatin reduces, but not completely blocks, the binding of ICP4 to the genome. Thus, ICP4 can induce low levels of gene expression. When ICP0 is present, it removes chromatin, enhancing the function of ICP4.

Different types of heterochromatin have been observed on quiescent/latent genomes in MRC5 cells and neurons. During quiescence in MRC5 cells, the genome is bound by histones with trimentylated H3K9. HP1 association has also been observed [191]. HP1 recruits the methyltransferase Suv39h1 to promote the spread of heterochromatin [193, 405]. These modifications are indicative of constitutive heterochromatin, which is believed to be a more permanent form of repressive chromatin [406]. However, in neurons, in vivo studies have shown that as latency is established, trimethylation of H3K27 but not H3K9 is observed. Furthermore, the polycomb repressive complex 1, a complex responsible for maintaining facultative heterochromatin, subunit Bmi1 has been observed on genomes in latently infected neurons [118]. Facultative heterochromatin is a less restrictive and more regulatable form of chromatin [262].

The more relaxed repression in neurons may enable more regulated control of the latent state. We have previously shown that there was a low level of transcription and specific upregulation of LAT in quiescently infected neurons. This study has shown that when ICP4 is present in quiescently infected neurons, it can bind to the quiescent genome and activate transcription across the quiescent genome. ICP0 can further induce transcription from the latent state by making the genome more accessible to ICP4. Uncontrolled expression of these two proteins could lead to frequent reactivation if left uncontrolled. The LAT transcript encodes several micro RNAs that target the ICP0 and ICP4 transcripts so that latency can be maintained [208, 210]. Fluctuations in the expression of LAT may lead to changes in ICP0 and ICP4 expression to promote reactivation from latency.

4.0 SUMMARY OF THESIS AND GENERAL DISCUSSION

4.1 SUMMARY OF THESIS

HSV transcription heavily relies on cellular transcription machinery and chromatin remodeling and modifying complexes. During its lifecycle, HSV infects both epithelial and neuronal cells, productively replicating in epithelial cells and establishing and reactivating from latency in neurons. As these two cell types are in different stages in the cell cycle and transcribe different genes, they differ in their abundance of transcription factors and chromatin remodeling complexes. These differences may impact both viral transcription, and the ability of the virus to re-initiate lytic gene expression during reactivation from latency. A comprehensive study of the expression of viral genes in these two cell types had not been previously investigated. Nor has the ability of individual viral activators to induce reactivation from latency been studied. Thus, the aims of this study were to use RNA sequencing and ChIP to characterize wild type and mutant virus gene expression and to investigate the ability of ICP4 and ICP0 to induce transcription from quiescent genomes in these two cell types.

To determine how cell type affected transcription of viral genes, RNAseq was used to quantify the viral transcription of both wild type and mutant viruses in MRC5 cells and TG neurons. Studies of productive infection with wild type virus showed that gene expression was similar in kinetics between neurons and MRC5 cells, but overall transcription was attenuated in neurons evidenced by less early and late gene transcription, possibly due to a decreased abundance of some cellular transcriptional activators used by the virus. The expression from d109, a viral mutant lacking immediate early genes, was highly attenuated and dysregulated in

both cell types. However, in the absence of immediate early genes, transcription was markedly different. Viral gene expression was approximately 10 times greater in neurons than MRC5 cells at all time points assayed.

To investigate how the level of repression of quiescent genomes affected the ability of the virus to reactivate in these two cell types, adenoviruses were used to deliver ICP0 and ICP4 to cells quiescently infected with d109 and RNAseq and ChIP/ChIPseq were used to quantify changes in gene expression and ICP4 binding, respectively. RNAseq studies showed that ICP4 was unable to activate transcription from quiescent genomes in MRC5 cells. ICP0 enabled ICP4 to induce gene expression through inducing the removal of chromatin from the genome, suggesting that in MRC5 cells, chromatin restricts access of ICP4 to the genome. However, in neurons, ICP4 alone induced a low level of transcription from the genome and was able to access some promoters and not others, suggesting that the heterochromatin on quiescent genomes in neurons is less repressive and not uniform. The results of these studies suggest that latency is a more dynamic than previously thought, which allows a more delicate balance between the maintenance of latency and reactivation of the virus. These two areas will be discussed in greater detail in the following sections.

4.2 GENERAL DICSUSSION

Latency – A Dynamic State

As d109 establishes quiescence in neurons and MRC5 cells, gene expression is highly dysregulated and attenuated. However, gene expression is approximately 10 times greater in

neurons than MRC5 cells. The majority of gene expression in neurons originated from loci within the unique short region and areas surrounding the joint region of the viral genome. The increased expression of these regions may be due to the presence of CTCF insulator elements in the genome, which serve as boundaries between areas of euchromatin and heterochromatin [202]. To date, seven CTCF elements have been identified in the HSV genome [203] and five of these elements are present in the d109 genome. It would be of interest to mutate these insulator elements and determine if gene expression changes during quiescence to establish a direct role of these elements in regulating expression from the quiescent genome.

The greater level of transcription in neurons could be due to differences in the epigenome which would affect the ability of the cells to repress the viral gene expression. Studies have shown that as neurons differentiate, they express different histone variants, shifting from H3.1 and H3.2 to H3.3 [403]. Nucleosomes containing the H3.3 variant have been shown to be enriched in euchromatic areas, associating them with active chromatin, while variants H3.1 and H3.2 are associated with repressed regions of chromatin [402]. The incorporation of this histone variant may promote the formation of facultative heterochromatin, rather than constitutive heterochromatin, on quiescent viral genomes. This form of chromatin is less repressive and can be reversed back to euchromatin following certain stimuli. The increased permissiveness of neurons to quiescent viral transcription may provide some insight as to why HSV evolved to establish a latent infection in neurons instead of epithelial cells. The expression of LAT and cellular repressors could serve to maintain the latent state, but not repress the genome too tightly to completely block the ability to reactivate.

The latency associated transcript is specifically up-regulated in neurons, likely due to the presence of a neuron specific enhancer element in its promoter [364]. Early studies showed that

LAT expression reduced the abundance of immediate early transcripts in neuron cell lines [407]. Later studies showed that LAT codes for several microRNAs, several of which have been shown to target the ICP0 and ICP4 transcripts [210, 211]. More recent studies have shown that the deletion of miR-H2, the microRNA that targets ICP0, leads to increased neurovirulence and an increased rate of reactivation [212, 408]. The increased reactivation in the absence of miR-H2 suggests that it may function to maintain the latent state.

The specific upregulation of LAT in neurons may provide a mechanism to control lytic gene expression through targeting immediate early gene transcripts. However, fluctuations in LAT transcription could lead to a transient decline in LAT and thus microRNA abundance and could allow low levels of immediate early gene transcripts to accumulate. Since ICP4 can bind to the LAT promoter and inhibit LAT expression, this could begin to tip the balance to favor reactivation.

Reactivation – Tipping the Balance

Reactivation from latency has been an area of intense study, but a mechanism for this process has yet to be clearly elucidated. Recent data has suggested a two phase model of reactivation in which a burst of dysregulated global transcription is followed by the normal lytic cascade of viral gene expression culminating in the production of viral progeny. Additionally, the first phase of gene expression does not always lead to initiation of lytic expression and may serve as an additional point of control [217, 218]. However, this model and other models of reactivation use thermal stress and NGF withdrawal, and while they do efficiently induce reactivation of latent

virus in animal models, they may apply a much stronger reactivation stimulus than is what is experienced by individuals with recurring HSV breakouts.

Our studies have shown that neurons are more permissive to HSV gene expression during quiescence and that LAT is specifically upregulated. In vivo studies of latency have shown that at least one lytic gene is expressed in two thirds of infected cells and that half of the infected cells have expression from more than one locus. Furthermore, lytic gene expression was directly correlated with increased expression of cellular antiviral genes, suggesting that the cell can detect increases in viral transcription and respond through activation of the cellular innate immune response to repress transcription. This opens the possibility that there is a low level of abortive reactivation during latency [409]. This suggests that stimuli that either decrease the cellular immune response or increase viral transcription could tip this balance to favor reactivation of the virus.

Viral gene expression could be favored by transient decreases in LAT expression during latency. LAT expresses multiple microRNAs that have been shown to target the ICP0 and ICP4 transcripts, preventing their accumulation [210]. Decreased expression of LAT would lead to a decrease in the abundance of these microRNAs, enabling ICP4 and ICP0 transcripts to accumulate and possibly be translated. Our studies have shown that ICP4 alone can bind to and activate transcription from the quiescent genome in neurons, which would lead to an increase in lytic gene transcription. ICP0 further enhances the activity of ICP4 by removing repressive chromatin from the genome, increasing ICP4 binding.

In response to increases in viral transcription, the cell upregulates the expression of immune genes, resulting in an abortive reactivation event. However, ICP0 also targets multiple arms of the cellular innate immune response to prevent silencing of the genome. Studies have

shown that ICP0 is not required for the burst of gene expression preceding the lytic cascade, but is necessary for the production of infectious virus. It is possible that ICP0 functions during the second stage to block the innate immune response to prevent silencing of transcription and reestablishment of latency. It would be of interest to determine if the ability of ICP4 to induce transcription from the latent genome is dependent on cessation or reduction in innate immune signaling.

Taken together these studies provide valuable insight into viral gene expression in the latent state and a possible mechanism for the reactivation of virus in the absence of external stressors. The increased permissiveness to viral transcription in neurons enables finer tuning of reactivation, through balancing lytic gene expression with the expression of the latency associated transcript and innate immune response. Increased expression of ICP4 and ICP0 could be caused from a decrease in LAT expression. ICP4 could function to promote lytic gene expression, while ICP0 could function to enhance ICP4 activity and block the cellular innate immune response to promote reactivation and the production of infectious virions.

5.0 FUTURE DIRECTIONS

The next step in this project is to determine a reason for why neurons are more permissive to transcription than MRC5 cells. We have hypothesized that this difference may be due to differences in the chromatin deposited on the genome, in which neurons deposit facultative heterochromatin and MRC5 cells deposit constitutive heterochromatin. Facultative heterochromatin, but not constitutive heterochromatin, can be converted back to euchromatin under certain conditions. Recent studies have begun to link stress, both psychological and physiological, to changes in gene expression due to changes in chromatin modification [410]. If the genome is repressed by facultative heterochromatin in neurons, it is possible that cellular stress could induce a change in the epigenetic state of the viral genome from heterochromatic to euchromatin, facilitating reactivation of the virus.

Another possibility could be differences in the activation of the cellular innate immune response. In the absence of viral gene expression, HSV-1 induces the interferon response, leading to the upregulation of PML and Sp100, two components of the cellular ND10 complex. ICP0 degrades these complexes during lytic replication, but in latency, the viral genome is localized to these expanded structures and becomes silenced through the deposition of cellular heterochromatin [411]. If neurons are less able to induce the interferon response, they would express less of these restrictive proteins, leading to a decreased ability to repress the virus.

We propose to use the d109 virus to determine both how the genomes are differentially repressed in neurons and MRC5, focusing on differences in the types of chromatin deposited onto the genome and in the induction of the host cell immune response.

Epigenetic Differences between Cell Types

In chapter 3, we suggested that the differential expression of histone variants may affect the type of chromatin deposited onto genomes, but due to the unavailability of antibodies specific to H3.1, H3.2, and H3.3, it is not possible to test this hypothesis without the use of transgenic cell lines expressing tagged histone variants. However, we can investigate the deposition of different forms of heterochromatin – facultative heterochromatin and constitutive heterochromatin in neurons and MRC5 cells, respectively. Chromatin can be prepared from quiescently infected cells and immunoprecipitated using antibodies to trimethylated H3K27 and trimethylated H3K9, markers of facultative and constitutive heterochromatin, respectively. An enrichment of trimethylated H3K27 in neurons and trimethylated H3K9 in MRC5 would support this hypothesis.

Additionally, it would also be of interest to determine why ICP4 could bind to some loci on the d109 genome in neurons but not others. This is most likely due to differences in local chromatin structure. To explore this possibility, ChIP using antibodies against histone H3, trimethylated H3K27 and trimethylated H3K9, as well as ICP4 will be used to correlate ICP4 binding to chromatin medication. It would be especially interesting to do ChIP seq on neurons to compare genome wide histone occupancy, modification, and ICP4 binding. However, due to the low number of cells obtained from trigeminal ganglia culture, and the large number of cells necessary for this technique, this is not yet feasible.

Furthermore, if we find that neurons deposit mainly facultative heterochromatin and MRC5 cells deposit mainly constitutive heterochromatin, it would be of great interest to study how cellular stress can contribute to reactivation of the virus in neurons. NGF withdrawal has been shown to be sufficient to induce reactivation of latent virus. Quiescently infected neurons

will be treated with either LY294002, an inhibitor of the NGF signaling pathway, or vector control. RNA will be collected at 0, 4, 8, and 24 hours post infection and prepared for Illumina sequencing. If an upregulation of viral gene expression is detected, ChIP using antibodies against H3 and trimethylated H3K27 will be performed to determine if the increase in gene expression is correlated to changes in the chromatin state.

Differences in Interferon Response

To determine how the cellular response differs between cell types, both neurons and MRC5 cells will be infected with d109, and RNA will be harvested at 4, 8, and 24 hours post infection. A mock infected sample will also be collected for a baseline comparison. RNA will be prepared for Illumina sequencing. The reads will be aligned to the human genome for MRC5 cell samples and to the mouse genome for neuron samples using the cufflinks software package on the Galaxy server. The cuffcompare software package will be used to detect changes in gene expression between the mock infected and d109 infected samples and a list of genes that are significantly upregulated between mock and infected samples will be compiled for each cell type.

Since cuffcompare cannot compare differences in expression between samples from different species, the gene lists from neurons and MRC5 cells will need to be manually compared. To streamline this process and to determine if one particular set of genes is upregulated more in one cell type, genes will be grouped according to function, specifically focusing on interferon inducible genes and genes involved in chromatin modification. The focus will be looking for genes highly upregulated in MRC5 cells, but not in neurons. To determine if these genes contribute to the increased silencing of d109 genomes in MRC5 cells, highly induced

genes can be knocked down with siRNA prior to infection to determine the effect of their absence.

6.0 BIBLIOGRAPHY

1. Grunewald, K., et al., *Three-dimensional structure of herpes simplex virus from cryo-electron tomography.* Science, 2003. **302**(5649): p. 1396-8.

2. Furlong, D., H. Swift, and B. Roizman, *Arrangement of herpesvirus deoxyribonucleic acid in the core.* J Virol, 1972. **10**(5): p. 1071-4.

3. Gibson, W. and B. Roizman, *Compartmentalization of spermine and spermidine in the herpes simplex virion.* Proc Natl Acad Sci U S A, 1971. **68**(11): p. 2818-21.

4. Bauer, D.W., et al., *Herpes virus genome, the pressure is on.* J Am Chem Soc, 2013. **135**(30): p. 11216-21.

5. Schrag, J.D., et al., *Three-dimensional structure of the HSV1 nucleocapsid.* Cell, 1989. **56**(4): p. 651-60.

6. Zhou, Z.H., et al., *Seeing the herpesvirus capsid at 8.5 A.* Science, 2000. **288**(5467): p. 877-80.

7. van Genderen, I.L., et al., *The phospholipid composition of extracellular herpes simplex virions differs from that of host cell nuclei.* Virology, 1994. **200**(2): p. 831-6.

8. Flint, S.J., *Principles of Virology*. 2 ed2004, Washington DC: ASM press.

9. McGeoch, D.J., et al., *The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1.* J Gen Virol, 1988. **69 (Pt 7)**: p. 1531-74.

10. Wadsworth, S., R.J. Jacob, and B. Roizman, *Anatomy of herpes simplex virus DNA. II. Size, composition, and arrangement of inverted terminal repetitions.* J Virol, 1975. **15**(6): p. 1487-97.

11. Hayward, G.S., et al., *Anatomy of herpes simplex virus DNA: evidence for four populations of molecules that differ in the relative orientations of their long and short components.* Proc Natl Acad Sci U S A, 1975. **72**(11): p. 4243-7.

12. Jenkins, F.J. and B. Roizman, *Herpes simplex virus 1 recombinants with noninverting genomes frozen in different isomeric arrangements are capable of independent replication.* J Virol, 1986. **59**(2): p. 494-9.

13. Alwine, J.C., W.L. Steinhart, and C.W. Hill, *Transcription of herpes simplex type 1 DNA in nuclei isolated from infected HEp-2 and KB cells.* Virology, 1974. **60**(1): p. 302-7.

14. Liu, F. and B. Roizman, *Characterization of the protease and other products of amino-terminus-proximal cleavage of the herpes simplex virus 1 UL26 protein.* J Virol, 1993. **67**(3): p. 1300-9.

15. Carter, K.L. and B. Roizman, *Alternatively spliced mRNAs predicted to yield frame-shift proteins and stable intron 1 RNAs of the herpes simplex virus 1 regulatory gene alpha 0 accumulate in the cytoplasm of infected cells.* Proc Natl Acad Sci U S A, 1996. **93**(22): p. 12535-40.

16. Liu, F.Y. and B. Roizman, *The herpes simplex virus 1 gene encoding a protease also contains within its coding domain the gene encoding the more abundant substrate.* J Virol, 1991. **65**(10): p. 5149-56.

17. Liu, F.Y. and B. Roizman, *The promoter, transcriptional unit, and coding sequence of herpes simplex virus 1 family 35 proteins are contained within and in frame with the UL26 open reading frame.* J Virol, 1991. **65**(1): p. 206-12.

18. Poon, A.P. and B. Roizman, *Herpes simplex virus 1 ICP22 regulates the accumulation of a shorter mRNA and of a truncated US3 protein kinase that exhibits altered functions.* J Virol, 2005. **79**(13): p. 8470-9.

19. Boehmer, P.E. and I.R. Lehman, *Herpes simplex virus DNA replication.* Annu Rev Biochem, 1997. **66**: p. 347-84.

20. Philip E. Pallett, B.R., in *Fields Virology*, P.M.H. David M. Knipe, Editor 2013.

21. Kennedy, P.G. and A. Chaudhuri, *Herpes simplex encephalitis.* J Neurol Neurosurg Psychiatry, 2002. **73**(3): p. 237-8.

22. Smith, J.S. and N.J. Robinson, *Age-specific prevalence of infection with herpes simplex virus types 2 and 1: a global review.* J Infect Dis, 2002. **186 Suppl 1**: p. S3-28.

23. Lippelt, L., R.W. Braun, and J.E. Kuhn, *Genital herpes simplex virus type 1 infection: new fields for an old acquaintance?* Intervirology, 2002. **45**(1): p. 2-5.

24. Rubright, J.H. and A.B. Shafritz, *The herpetic whitlow.* J Hand Surg Am, 2011. **36**(2): p. 340-2.

25. Anderson, B.J., *The effectiveness of valacyclovir in preventing reactivation of herpes gladiatorum in wrestlers.* Clin J Sport Med, 1999. **9**(2): p. 86-90.

26. Pepose, J.S., *Herpes simplex keratitis: role of viral infection versus immune response.* Surv Ophthalmol, 1991. **35**(5): p. 345-52.

27. Liesegang, T.J., et al., *Epidemiology of ocular herpes simplex. Incidence in Rochester, Minn, 1950 through 1982.* Arch Ophthalmol, 1989. **107**(8): p. 1155-9.

28. Halberstadt, M., et al., *The outcome of corneal grafting in patients with stromal keratitis of herpetic and non-herpetic origin.* Br J Ophthalmol, 2002. **86**(6): p. 646-52.

29. Solomon, T., et al., *Management of suspected viral encephalitis in adults--Association of British Neurologists and British Infection Association National Guidelines.* J Infect, 2012. **64**(4): p. 347-73.

30. Granerod, J., et al., *Causes of encephalitis and differences in their clinical presentations in England: a multicentre, population-based prospective study.* Lancet Infect Dis, 2010. **10**(12): p. 835-44.

31. Whitley, R.J., et al., *Vidarabine versus acyclovir therapy in herpes simplex encephalitis.* N Engl J Med, 1986. **314**(3): p. 144-9.

32. Mole, L., et al., *The impact of active herpes simplex virus infection on human immunodeficiency virus load.* J Infect Dis, 1997. **176**(3): p. 766-70.

33. LeGoff, J., et al., *Cervicovaginal HIV-1 and herpes simplex virus type 2 shedding during genital ulcer disease episodes.* AIDS, 2007. **21**(12): p. 1569-78.

34. Sufiawati, I. and S.M. Tugizov, *HIV-associated disruption of tight and adherens junctions of oral epithelial cells facilitates HSV-1 infection and spread.* PLoS One, 2014. **9**(2): p. e88803.

35. Lolis, M.S., et al., *Drug-resistant herpes simplex virus in HIV infected patients.* Acta Dermatovenerol Croat, 2008. **16**(4): p. 204-8.

36. Glorioso, J.C. and D.J. Fink, *Herpes vector-mediated gene transfer in treatment of diseases of the nervous system.* Annu Rev Microbiol, 2004. **58**: p. 253-71.

37. Karpoff, H.M., et al., *Efficient cotransduction of tumors by multiple herpes simplex vectors: implications for tumor vaccine production.* Cancer Gene Ther, 2000. **7**(4): p. 581-8.

38. Maguire-Zeiss, K.A., W.J. Bowers, and H.J. Federoff, *HSV vector-mediated gene delivery to the central nervous system.* Curr Opin Mol Ther, 2001. **3**(5): p. 482-90.

39. Roizman, B., *The function of herpes simplex virus genes: a primer for genetic engineering of novel vectors.* Proc Natl Acad Sci U S A, 1996. **93**(21): p. 11307-12.

40. Epstein, A.L., *HSV-1-derived recombinant and amplicon vectors for preventive or therapeutic gene transfer: an overview.* Gene Ther, 2005. **12 Suppl 1**: p. S153.

41. Logvinoff, C. and A.L. Epstein, *Intracellular Cre-mediated deletion of the unique packaging signal carried by a herpes simplex virus type 1 recombinant and its relationship to the cleavage-packaging process.* J Virol, 2000. **74**(18): p. 8402-12.

42. Zaupa, C., V. Revol-Guyot, and A.L. Epstein, *Improved packaging system for generation of high-level noncytotoxic HSV-1 amplicon vectors using Cre-loxP site-specific recombination to delete the packaging signals of defective helper genomes.* Hum Gene Ther, 2003. **14**(11): p. 1049-63.

43. Zou, W., et al., *Microinjection of HSV-1 amplicon vector-mediated human proenkephalin into the periaqueductal grey attenuates neuropathic pain in rats.* Int J Neurosci, 2012. **122**(4): p. 189-94.

44. Ho, I.A., et al., *Targeting human glioma cells using HSV-1 amplicon peptide display vector.* Gene Ther, 2010. **17**(2): p. 250-60.

45. Burton, E.A., et al., *Replication-defective genomic herpes simplex vectors: design and production.* Curr Opin Biotechnol, 2002. **13**(5): p. 424-8.

46. Kaur, A., et al., *Ability of herpes simplex virus vectors to boost immune responses to DNA vectors and to protect against challenge by simian immunodeficiency virus.* Virology, 2007. **357**(2): p. 199-214.

47. Lauterbach, H., et al., *Protection from bacterial infection by a single vaccination with replication-deficient mutant herpes simplex virus type 1.* J Virol, 2004. **78**(8): p. 4020-8.

48. Chattopadhyay, M., *Targeted delivery of growth factors by HSV-mediated gene transfer for peripheral neuropathy.* Curr Gene Ther, 2013. **13**(5): p. 315-21.

49. Meignier, B., R. Longnecker, and B. Roizman, *In vivo behavior of genetically engineered herpes simplex viruses R7017 and R7020: construction and evaluation in rodents.* J Infect Dis, 1988. **158**(3): p. 602-14.

50. Meignier, B., et al., *In vivo behavior of genetically engineered herpes simplex viruses R7017 and R7020. II. Studies in immunocompetent and immunosuppressed owl monkeys (Aotus trivirgatus).* J Infect Dis, 1990. **162**(2): p. 313-21.

51. Wang, J., et al., *Treatment of human hepatocellular carcinoma by the oncolytic herpes simplex virus G47delta.* Cancer Cell Int, 2014. **14**(1): p. 83.

52. Markert, J.M., et al., *A phase 1 trial of oncolytic HSV-1, G207, given in combination with radiation for recurrent GBM demonstrates safety and radiographic responses.* Mol Ther, 2014. **22**(5): p. 1048-55.

53. Harrington, K.J., et al., *Phase I/II study of oncolytic HSV GM-CSF in combination with radiotherapy and cisplatin in untreated stage III/IV squamous cell cancer of the head and neck.* Clin Cancer Res, 2010. **16**(15): p. 4005-15.

54. Fong, Y., et al., *A herpes oncolytic virus can be delivered via the vasculature to produce biologic changes in human colorectal cancer.* Mol Ther, 2009. **17**(2): p. 389-94.

55. Wysocka, J. and W. Herr, *The herpes simplex virus VP16-induced complex: the makings of a regulatory switch.* Trends Biochem Sci, 2003. **28**(6): p. 294-304.

56. DeLuca, N.A., A.M. McCarthy, and P.A. Schaffer, *Isolation and characterization of deletion mutants of herpes simplex virus type 1 in the gene encoding immediate-early regulatory protein ICP4.* J Virol, 1985. **56**(2): p. 558-70.

57. DeLuca, N.A. and P.A. Schaffer, *Activation of immediate-early, early, and late promoters by temperature-sensitive and wild-type forms of herpes simplex virus type 1 protein ICP4.* Mol Cell Biol, 1985. **5**(8): p. 1997-2008.

58. Homa, F.L. and J.C. Brown, *Capsid assembly and DNA packaging in herpes simplex virus.* Rev Med Virol, 1997. **7**(2): p. 107-122.

59. Willis, R.A., et al., *Dendritic cells transduced with HSV-1 amplicons expressing prostate-specific antigen generate antitumor immunity in mice.* Hum Gene Ther, 2001. **12**(15): p. 1867-79.

60. Chen, X., et al., *HSV amplicon-mediated neurotrophin-3 expression protects murine spiral ganglion neurons from cisplatin-induced damage.* Mol Ther, 2001. **3**(6): p. 958-63.

61. Calakos, N., et al., *Cortical MRI findings associated with rapid correction of hyponatremia.* Neurology, 2000. **55**(7): p. 1048-51.

62. Jiang, C., et al., *Immobilized cobalt affinity chromatography provides a novel, efficient method for herpes simplex virus type 1 gene vector purification.* J Virol, 2004. **78**(17): p. 8994-9006.

63. Palese, P. and B. Roizman, *Genetic engineering of viruses and of virus vectors: a preface.* Proc Natl Acad Sci U S A, 1996. **93**(21): p. 11287.

64. Cuchet, D., et al., *Characterization of antiproliferative and cytotoxic properties of the HSV-1 immediate-early ICPo protein.* J Gene Med, 2005. **7**(9): p. 1187-99.

65. Ferrera, R., et al., *Efficient and non-toxic gene transfer to cardiomyocytes using novel generation amplicon vectors derived from HSV-1.* J Mol Cell Cardiol, 2005. **38**(1): p. 219-23.

66. Handal, G., et al., *Prophylaxis against respiratory syncytial virus in high-risk infants: administration of immune globulin and epidemiological surveillance of infection.* Tex Med, 2000. **96**(5): p. 58-61.

67. Wagner, L., *HSV-1 ICP4, A MULTIFACETED RNA POLII TRANSCRIPTION FACTOR*, 2012, University of Pittsburgh.

68. Heine, J.W., et al., *Proteins specified by herpes simplex virus. XII. The virion polypeptides of type 1 strains.* J Virol, 1974. **14**(3): p. 640-51.

69. Jones, K.A. and R. Tjian, *Sp1 binds to promoter sequences and activates herpes simplex virus 'immediate-early' gene transcription in vitro.* Nature, 1985. **317**(6033): p. 179-82.

70. O'Rourke, D. and P. O'Hare, *Mutually exclusive binding of two cellular factors within a critical promoter region of the gene for the IE110k protein of herpes simplex virus.* J Virol, 1993. **67**(12): p. 7201-14.

71. Goding, C.R. and P. O'Hare, *Herpes simplex virus Vmw65-octamer binding protein interaction: a paradigm for combinatorial control of transcription.* Virology, 1989. **173**(2): p. 363-7.

72. Roizman, B.a.K., D.M., *Herpes Simplex Viruses and their Replication*, in *Fields Virology*2001, Lippincott, Williams, and Wilkens.

73. Gu, B., R. Kuddus, and N.A. DeLuca, *Repression of activator-mediated transcription by herpes simplex virus ICP4 via a mechanism involving interactions with the basal transcription factors TATA-binding protein and TFIIB.* Mol Cell Biol, 1995. **15**(7): p. 3618-26.

74. Lium, E.K., et al., *Repression of the alpha0 gene by ICP4 during a productive herpes simplex virus infection.* J Virol, 1996. **70**(6): p. 3488-96.

75. Pugh, B.F. and R. Tjian, *Mechanism of transcriptional activation by Sp1: evidence for coactivators.* Cell, 1990. **61**(7): p. 1187-97.

76. Imbalzano, A.N., D.M. Coen, and N.A. DeLuca, *Herpes simplex virus transactivator ICP4 operationally substitutes for the cellular transcription factor Sp1 for efficient expression of the viral thymidine kinase gene.* J Virol, 1991. **65**(2): p. 565-74.

77. Kuddus, R.H. and N.A. DeLuca, *DNA-dependent oligomerization of herpes simplex virus type 1 regulatory protein ICP4.* J Virol, 2007. **81**(17): p. 9230-7.

78. Grondin, B. and N. DeLuca, *Herpes simplex virus type 1 ICP4 promotes transcription preinitiation complex formation by enhancing the binding of TFIID to DNA.* J Virol, 2000. **74**(24): p. 11504-10.

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

80. Iwahori, S., et al., *Enhanced phosphorylation of transcription factor sp1 in response to herpes simplex virus type 1 infection is dependent on the ataxia telangiectasia-mutated protein.* J Virol, 2007. **81**(18): p. 9653-64.

81. Kim, D.B. and N.A. DeLuca, *Phosphorylation of transcription factor Sp1 during herpes simplex virus type 1 infection.* J Virol, 2002. **76**(13): p. 6473-9.

82. Zabierowski, S. and N.A. DeLuca, *Differential cellular requirements for activation of herpes simplex virus type 1 early (tk) and late (gC) promoters by ICP4.* J Virol, 2004. **78**(12): p. 6162-70.

83. DeJong, J., R. Bernstein, and R.G. Roeder, *Human general transcription factor TFIIA: characterization of a cDNA encoding the small subunit and requirement for basal and activated transcription.* Proc Natl Acad Sci U S A, 1995. **92**(8): p. 3313-7.

84. Zabierowski, S.E. and N.A. Deluca, *Stabilized binding of TBP to the TATA box of herpes simplex virus type 1 early (tk) and late (gC) promoters by TFIIA and ICP4.* J Virol, 2008. **82**(7): p. 3546-54.

85. Olesky, M., et al., *Evidence for a direct interaction between HSV-1 ICP27 and ICP8 proteins.* Virology, 2005. **331**(1): p. 94-105.

86. Paoletti, A.C., et al., *Quantitative proteomic analysis of distinct mammalian Mediator complexes using normalized spectral abundance factors.* Proc Natl Acad Sci U S A, 2006. **103**(50): p. 18928-33.

87. Harkness, J.M., M. Kader, and N.A. DeLuca, *Transcription of the Herpes Simplex Virus 1 Genome during Productive and Quiescent Infection of Neuronal and Nonneuronal Cells.* J Virol, 2014. **88**(12): p. 6847-6861.

88. Gu, B. and N. DeLuca, *Requirements for activation of the herpes simplex virus glycoprotein C promoter in vitro by the viral regulatory protein ICP4.* J Virol, 1994. **68**(12): p. 7953-65.

89. Woerner, A.M. and J.P. Weir, *Characterization of the initiator and downstream promoter elements of herpes simplex virus 1 late genes.* Virology, 1998. **249**(2): p. 219-30.

90. Guzowski, J.F., J. Singh, and E.K. Wagner, *Transcriptional activation of the herpes simplex virus type 1 UL38 promoter conferred by the cis-acting downstream activation sequence is mediated by a cellular transcription factor.* J Virol, 1994. **68**(12): p. 7774-89.

91. Knopf, C.W., B. Spies, and H.C. Kaerner, *The DNA replication origins of herpes simplex virus type 1 strain Angelotti.* Nucleic Acids Res, 1986. **14**(21): p. 8655-67.

92. Summers, B.C. and D.A. Leib, *Herpes simplex virus type 1 origins of DNA replication play no role in the regulation of flanking promoters.* J Virol, 2002. **76**(14): p. 7020-9.

93. Weir, H.M. and N.D. Stow, *Two binding sites for the herpes simplex virus type 1 UL9 protein are required for efficient activity of the oriS replication origin.* J Gen Virol, 1990. **71 (Pt 6)**: p. 1379-85.

94. Igarashi, K., et al., *Construction and properties of a recombinant herpes simplex virus 1 lacking both S-component origins of DNA synthesis.* J Virol, 1993. **67**(4): p. 2123-32.

95. Gray, C.P. and H.C. Kaerner, *Sequence of the putative origin of replication in the UL region of herpes simplex virus type 1 ANG DNA.* J Gen Virol, 1984. **65 (Pt 12)**: p. 2109-19.

96. Hardwicke, M.A. and P.A. Schaffer, *Cloning and characterization of herpes simplex virus type 1 oriL: comparison of replication and protein-DNA complex formation by oriL and oriS.* J Virol, 1995. **69**(3): p. 1377-88.

97. Balliet, J.W. and P.A. Schaffer, *Point mutations in herpes simplex virus type 1 oriL, but not in oriS, reduce pathogenesis during acute infection of mice and impair reactivation from latency.* J Virol, 2006. **80**(1): p. 440-50.

98. Wu, C.A., et al., *Identification of herpes simplex virus type 1 genes required for origin-dependent DNA synthesis.* J Virol, 1988. **62**(2): p. 435-43.

99. Weller, S.K. and D.M. Coen, *Herpes simplex viruses: mechanisms of DNA replication.* Cold Spring Harb Perspect Biol, 2012. **4**(9): p. a013011.

100. Blumel, J. and B. Matz, *Thermosensitive UL9 gene function is required for early stages of herpes simplex virus type 1 DNA synthesis.* J Gen Virol, 1995. **76 (Pt 12)**: p. 3119-24.

101. Aslani, A., M. Olsson, and P. Elias, *ATP-dependent unwinding of a minimal origin of DNA replication by the origin-binding protein and the single-strand DNA-binding protein ICP8 from herpes simplex virus type I.* J Biol Chem, 2002. **277**(43): p. 41204-12.

102. McLean, G.W., et al., *The herpes simplex virus type 1 origin-binding protein interacts specifically with the viral UL8 protein.* J Gen Virol, 1994. **75 (Pt 10)**: p. 2699-706.

103. Carrington-Lawrence, S.D. and S.K. Weller, *Recruitment of polymerase to herpes simplex virus type 1 replication foci in cells expressing mutant primase (UL52) proteins.* J Virol, 2003. **77**(7): p. 4237-47.

104. Wilkinson, D.E. and S.K. Weller, *The role of DNA recombination in herpes simplex virus DNA replication.* IUBMB Life, 2003. **55**(8): p. 451-8.

105. Newcomb, W.W., et al., *Assembly of the herpes simplex virus capsid: identification of soluble scaffold-portal complexes and their role in formation of portal-containing capsids.* J Virol, 2003. **77**(18): p. 9862-71.

106. Beard, P.M., N.S. Taus, and J.D. Baines, *DNA cleavage and packaging proteins encoded by genes U(L)28, U(L)15, and U(L)33 of herpes simplex virus type 1 form a complex in infected cells.* J Virol, 2002. **76**(10): p. 4785-91.

107. Toropova, K., et al., *The herpes simplex virus 1 UL17 protein is the second constituent of the capsid vertex-specific component required for DNA packaging and retention.* J Virol, 2011. **85**(15): p. 7513-22.

108. McNab, A.R., et al., *The product of the herpes simplex virus type 1 UL25 gene is required for encapsidation but not for cleavage of replicated viral DNA.* J Virol, 1998. **72**(2): p. 1060-70.

109. Baines, J.D., *Herpes simplex virus capsid assembly and DNA packaging: a present and future antiviral drug target.* Trends Microbiol, 2011. **19**(12): p. 606-13.

110. Silva, L., et al., *Role for A-type lamins in herpesviral DNA targeting and heterochromatin modulation.* PLoS Pathog, 2008. **4**(5): p. e1000071.

111. Everett, R.D. and J. Murray, *ND10 components relocate to sites associated with herpes simplex virus type 1 nucleoprotein complexes during virus infection.* J Virol, 2005. **79**(8): p. 5078-89.

112. Lukashchuk, V. and R.D. Everett, *Regulation of ICP0-null mutant herpes simplex virus type 1 infection by ND10 components ATRX and hDaxx.* J Virol, 2010. **84**(8): p. 4026-40.

113. Hollenbach, A.D., et al., *Daxx and histone deacetylase II associate with chromatin through an interaction with core histones and the chromatin-associated protein Dek.* J Cell Sci, 2002. **115**(Pt 16): p. 3319-30.

114. Alexiadis, V., et al., *The protein encoded by the proto-oncogene DEK changes the topology of chromatin and reduces the efficiency of DNA replication in a chromatin-specific manner.* Genes Dev, 2000. **14**(11): p. 1308-12.

115. Oh, J., N. Ruskoski, and N.W. Fraser, *Chromatin assembly on herpes simplex virus 1 DNA early during a lytic infection is Asf1a dependent.* J Virol, 2012. **86**(22): p. 12313-21.

116. Tang, Y., et al., *Structure of a human ASF1a-HIRA complex and insights into specificity of histone chaperone complex assembly.* Nat Struct Mol Biol, 2006. **13**(10): p. 921-9.

117. Placek, B.J., et al., *The histone variant H3.3 regulates gene expression during lytic infection with herpes simplex virus type 1.* J Virol, 2009. **83**(3): p. 1416-21.

118. Cliffe, A.R., D.M. Coen, and D.M. Knipe, *Kinetics of facultative heterochromatin and polycomb group protein association with the herpes simplex viral genome during establishment of latent infection.* MBio, 2013. **4**(1).

119. Arbuckle, J.H. and T.M. Kristie, *Epigenetic repression of herpes simplex virus infection by the nucleosome remodeler CHD3.* MBio, 2014. **5**(1): p. e01027-13.

120. Lee, M.G., et al., *An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation.* Nature, 2005. **437**(7057): p. 432-5.

121. Lee, M.G., et al., *Functional interplay between histone demethylase and deacetylase enzymes.* Mol Cell Biol, 2006. **26**(17): p. 6395-402.

122. Kristie, T.M., *Dynamic modulation of HSV chromatin drives initiation of infection and provides targets for epigenetic therapies.* Virology, 2015. **479-480C**: p. 555-561.

123. Narayanan, A., et al., *Combinatorial transcription of herpes simplex virus and varicella zoster virus immediate early genes is strictly determined by the cellular coactivator HCF-1.* J Biol Chem, 2005. **280**(2): p. 1369-75.

124. Wissmann, M., et al., *Cooperative demethylation by JMJD2C and LSD1 promotes androgen receptor-dependent gene expression.* Nat Cell Biol, 2007. **9**(3): p. 347-53.

125. Liang, Y., et al., *Targeting the JMJD2 histone demethylases to epigenetically control herpesvirus infection and reactivation from latency.* Sci Transl Med, 2013. **5**(167): p. 167ra5.

126. Liang, Y., et al., *Inhibition of the histone demethylase LSD1 blocks alpha-herpesvirus lytic replication and reactivation from latency.* Nat Med, 2009. **15**(11): p. 1312-7.

127. Narayanan, A., W.T. Ruyechan, and T.M. Kristie, *The coactivator host cell factor-1 mediates Set1 and MLL1 H3K4 trimethylation at herpesvirus immediate early promoters for initiation of infection.* Proc Natl Acad Sci U S A, 2007. **104**(26): p. 10835-40.

128. Huang, J., et al., *Trimethylation of histone H3 lysine 4 by Set1 in the lytic infection of human herpes simplex virus 1.* J Virol, 2006. **80**(12): p. 5740-6.

129. Taylor, T.J. and D.M. Knipe, *Proteomics of herpes simplex virus replication compartments: association of cellular DNA replication, repair, recombination, and chromatin remodeling proteins with ICP8.* J Virol, 2004. **78**(11): p. 5856-66.

130. Bryant, K.F., R.C. Colgrove, and D.M. Knipe, *Cellular SNF2H chromatin-remodeling factor promotes herpes simplex virus 1 immediate-early gene expression and replication.* MBio, 2011. **2**(1): p. e00330-10.

131. Kalamvoki, M. and B. Roizman, *Circadian CLOCK histone acetyl transferase localizes at ND10 nuclear bodies and enables herpes simplex virus gene expression.* Proc Natl Acad Sci U S A, 2010. **107**(41): p. 17721-6.

132. Leinbach, S.S. and W.C. Summers, *The structure of herpes simplex virus type 1 DNA as probed by micrococcal nuclease digestion.* J Gen Virol, 1980. **51**(Pt 1): p. 45-59.

133. Muggeridge, M.I. and N.W. Fraser, *Chromosomal organization of the herpes simplex virus genome during acute infection of the mouse central nervous system.* J Virol, 1986. **59**(3): p. 764-7.

134. Lacasse, J.J. and L.M. Schang, *Herpes simplex virus 1 DNA is in unstable nucleosomes throughout the lytic infection cycle, and the instability of the nucleosomes is independent of DNA replication.* J Virol, 2012. **86**(20): p. 11287-300.

135. Kent, J.R., et al., *During lytic infection herpes simplex virus type 1 is associated with histones bearing modifications that correlate with active transcription.* J Virol, 2004. **78**(18): p. 10178-86.

136. Liu, Y., et al., *Crystal structure of the conserved core of the herpes simplex virus transcriptional regulatory protein VP16.* Genes Dev, 1999. **13**(13): p. 1692-703.

137. Babb, R., et al., *DNA recognition by the herpes simplex virus transactivator VP16: a novel DNA-binding structure.* Mol Cell Biol, 2001. **21**(14): p. 4700-12.

138. Walker, S., R. Greaves, and P. O'Hare, *Transcriptional activation by the acidic domain of Vmw65 requires the integrity of the domain and involves additional determinants distinct from those necessary for TFIIB binding.* Mol Cell Biol, 1993. **13**(9): p. 5233-44.

139. Hirai, H., T. Tani, and N. Kikyo, *Structure and functions of powerful transactivators: VP16, MyoD and FoxA.* Int J Dev Biol, 2010. **54**(11-12): p. 1589-96.

140. Herrera, F.J. and S.J. Triezenberg, *VP16-dependent association of chromatin-modifying coactivators and underrepresentation of histones at immediate-early gene promoters during herpes simplex virus infection.* J Virol, 2004. **78**(18): p. 9689-96.

141. Kutluay, S.B. and S.J. Triezenberg, *Regulation of histone deposition on the herpes simplex virus type 1 genome during lytic infection.* J Virol, 2009. **83**(11): p. 5835-45.

142. Kobayashi, N., T.G. Boyer, and A.J. Berk, *A class of activation domains interacts directly with TFIIA and stimulates TFIIA-TFIID-promoter complex assembly.* Mol Cell Biol, 1995. **15**(11): p. 6465-73.

143. Lin, Y.S., et al., *Binding of general transcription factor TFIIB to an acidic activating region.* Nature, 1991. **353**(6344): p. 569-71.

144. Xiao, H., et al., *Binding of basal transcription factor TFIIH to the acidic activation domains of VP16 and p53.* Mol Cell Biol, 1994. **14**(10): p. 7013-24.

145. Stringer, K.F., C.J. Ingles, and J. Greenblatt, *Direct and selective binding of an acidic transcriptional activation domain to the TATA-box factor TFIID.* Nature, 1990. **345**(6278): p. 783-6.

146. Zhu, H., V. Joliot, and R. Prywes, *Role of transcription factor TFIIF in serum response factor-activated transcription.* J Biol Chem, 1994. **269**(5): p. 3489-97.

147. Ito, M., et al., *Identity between TRAP and SMCC complexes indicates novel pathways for the function of nuclear receptors and diverse mammalian activators.* Mol Cell, 1999. **3**(3): p. 361-70.

148. Mittler, G., et al., *A novel docking site on Mediator is critical for activation by VP16 in mammalian cells.* EMBO J, 2003. **22**(24): p. 6494-504.

149. Wilcox, K.W., et al., *Herpes simplex virus phosphoproteins. I. Phosphate cycles on and off some viral polypeptides and can alter their affinity for DNA.* J Virol, 1980. **33**(1): p. 167-82.

150. Preston, C.M. and E.L. Notarianni, *Poly(ADP-ribosyl)ation of a herpes simplex virus immediate early polypeptide.* Virology, 1983. **131**(2): p. 492-501.

151. DiDonato, J.A., J.R. Spitzner, and M.T. Muller, *A predictive model for DNA recognition by the herpes simplex virus protein ICP4.* J Mol Biol, 1991. **219**(3): p. 451-70.

152. Deluca, N.A., *Functions and Mechanism of Action of the Herpes Simplex Virus Type Regulatory Protein, ICP4*, in *Alphaherpesviruses*, S.K. Weller, Editor 2011, Caister Academic Press.

153. DeLuca, N.A. and P.A. Schaffer, *Physical and functional domains of the herpes simplex virus transcriptional regulatory protein ICP4.* J Virol, 1988. **62**(3): p. 732-43.

154. Carrozza, M.J. and N.A. DeLuca, *Interaction of the viral activator protein ICP4 with TFIID through TAF250.* Mol Cell Biol, 1996. **16**(6): p. 3085-93.

155. Lester, J.T. and N.A. DeLuca, *Herpes simplex virus 1 ICP4 forms complexes with TFIID and mediator in virus-infected cells.* J Virol, 2011. **85**(12): p. 5733-44.

156. Wagner, L.M., et al., *The N terminus and C terminus of herpes simplex virus 1 ICP4 cooperate to activate viral gene expression.* J Virol, 2012. **86**(12): p. 6862-74.

157. Joazeiro, C.A. and A.M. Weissman, *RING finger proteins: mediators of ubiquitin ligase activity.* Cell, 2000. **102**(5): p. 549-52.

158. Ciufo, D.M., M.A. Mullen, and G.S. Hayward, *Identification of a dimerization domain in the C-terminal segment of the IE110 transactivator protein from herpes simplex virus.* J Virol, 1994. **68**(5): p. 3267-82.

159. Everett, R.D., M.L. Parsy, and A. Orr, *Analysis of the functions of herpes simplex virus type 1 regulatory protein ICP0 that are critical for lytic infection and derepression of quiescent viral genomes.* J Virol, 2009. **83**(10): p. 4963-77.

160. Canning, M., et al., *A RING finger ubiquitin ligase is protected from autocatalyzed ubiquitination and degradation by binding to ubiquitin-specific protease USP7.* J Biol Chem, 2004. **279**(37): p. 38160-8.

161. Lanfranca, M.P., H.H. Mostafa, and D.J. Davido, *HSV-1 ICP0: An E3 Ubiquitin Ligase That Counteracts Host Intrinsic and Innate Immunity.* Cells, 2014. **3**(2): p. 438-54.

162. Yao, F. and P.A. Schaffer, *An activity specified by the osteosarcoma line U2OS can substitute functionally for ICP0, a major regulatory protein of herpes simplex virus type 1.* J Virol, 1995. **69**(10): p. 6249-58.

163. Everett, R.D., C. Boutell, and A. Orr, *Phenotype of a herpes simplex virus type 1 mutant that fails to express immediate-early regulatory protein ICP0.* J Virol, 2004. **78**(4): p. 1763-74.

164. Ishov, A.M., et al., *PML is critical for ND10 formation and recruits the PML-interacting protein daxx to this nuclear structure when modified by SUMO-1.* J Cell Biol, 1999. **147**(2): p. 221-34.

165. Jensen, K., C. Shiels, and P.S. Freemont, *PML protein isoforms and the RBCC/TRIM motif.* Oncogene, 2001. **20**(49): p. 7223-33.

166. Negorev, D. and G.G. Maul, *Cellular proteins localized at and interacting within ND10/PML nuclear bodies/PODs suggest functions of a nuclear depot.* Oncogene, 2001. **20**(49): p. 7234-42.

167. Seeler, J.S., et al., *Interaction of SP100 with HP1 proteins: a link between the promyelocytic leukemia-associated nuclear bodies and the chromatin compartment.* Proc Natl Acad Sci U S A, 1998. **95**(13): p. 7316-21.

168. Negorev, D.G., O.V. Vladimirova, and G.G. Maul, *Differential functions of interferon-upregulated Sp100 isoforms: herpes simplex virus type 1 promoter-based immediate-early gene suppression and PML protection from ICP0-mediated degradation.* J Virol, 2009. **83**(10): p. 5168-80.

169. Muller, S. and A. Dejean, *Viral immediate-early proteins abrogate the modification by SUMO-1 of PML and Sp100 proteins, correlating with nuclear body disruption.* J Virol, 1999. **73**(6): p. 5137-43.

170. Michaelson, J.S. and P. Leder, *RNAi reveals anti-apoptotic and transcriptionally repressive activities of DAXX.* J Cell Sci, 2003. **116**(Pt 2): p. 345-52.

171. Orzalli, M.H., et al., *Nuclear interferon-inducible protein 16 promotes silencing of herpesviral and transfected DNA.* Proc Natl Acad Sci U S A, 2013. **110**(47): p. E4492-501.

172. Orzalli, M.H., N.A. DeLuca, and D.M. Knipe, *Nuclear IFI16 induction of IRF-3 signaling during herpesviral infection and degradation of IFI16 by the viral ICP0 protein.* Proc Natl Acad Sci U S A, 2012. **109**(44): p. E3008-17.

173. van Lint, A.L., et al., *Herpes simplex virus immediate-early ICP0 protein inhibits Toll-like receptor 2-dependent inflammatory responses and NF-kappaB signaling.* J Virol, 2010. **84**(20): p. 10802-11.

174. Zhang, J., et al., *Herpes simplex virus 1 E3 ubiquitin ligase ICP0 protein inhibits tumor necrosis factor alpha-induced NF-kappaB activation by interacting with p65/RelA and p50/NF-kappaB1.* J Virol, 2013. **87**(23): p. 12935-48.

175. Lomonte, P., et al., *Functional interaction between class II histone deacetylases and ICP0 of herpes simplex virus type 1.* J Virol, 2004. **78**(13): p. 6744-57.

176. Gu, H., et al., *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, 2005. **102**(21): p. 7571-6.

177. Ferguson, B.J., et al., *DNA-PK is a DNA sensor for IRF-3-dependent innate immunity.* Elife, 2012. **1**: p. e00047.

178. Shirata, N., et al., *Activation of ataxia telangiectasia-mutated DNA damage checkpoint signal transduction elicited by herpes simplex virus infection.* J Biol Chem, 2005. **280**(34): p. 30336-41.

179. Parkinson, J., S.P. Lees-Miller, and R.D. Everett, *Herpes simplex virus type 1 immediate-early protein vmw110 induces the proteasome-dependent degradation of the catalytic subunit of DNA-dependent protein kinase.* J Virol, 1999. **73**(1): p. 650-7.

180. Lilley, C.E., et al., *A viral E3 ligase targets RNF8 and RNF168 to control histone ubiquitination and DNA damage responses.* EMBO J, 2010. **29**(5): p. 943-55.

181. Hafezi, W., et al., *Entry of herpes simplex virus type 1 (HSV-1) into the distal axons of trigeminal neurons favors the onset of nonproductive, silent infection.* PLoS Pathog, 2012. **8**(5): p. e1002679.

182. Aggarwal, A., et al., *Ultrastructural visualization of individual tegument protein dissociation during entry of herpes simplex virus 1 into human and rat dorsal root ganglion neurons.* J Virol, 2012. **86**(11): p. 6123-37.

183. Kristie, T.M., J.L. Vogel, and A.E. Sears, *Nuclear localization of the C1 factor (host cell factor) in sensory neurons correlates with reactivation of herpes simplex virus from latency.* Proc Natl Acad Sci U S A, 1999. **96**(4): p. 1229-33.

184. Pan, D., et al., *A neuron-specific host microRNA targets herpes simplex virus-1 ICP0 expression and promotes latency.* Cell Host Microbe, 2014. **15**(4): p. 446-56.

185. Yang, L., C.C. Voytek, and T.P. Margolis, *Immunohistochemical analysis of primary sensory neurons latently infected with herpes simplex virus type 1.* J Virol, 2000. **74**(1): p. 209-17.

186. Margolis, T.P., et al., *Herpes simplex virus type 2 (HSV-2) establishes latent infection in a different population of ganglionic neurons than HSV-1: role of latency-associated transcripts.* J Virol, 2007. **81**(4): p. 1872-8.

187. Kubat, N.J., et al., *The herpes simplex virus type 1 latency-associated transcript (LAT) enhancer/rcr is hyperacetylated during latency independently of LAT transcription.* J Virol, 2004. **78**(22): p. 12508-18.

188. Cliffe, A.R., D.A. Garber, and D.M. Knipe, *Transcription of the herpes simplex virus latency-associated transcript promotes the formation of facultative heterochromatin on lytic promoters.* J Virol, 2009. **83**(16): p. 8182-90.

189. Wang, Q.Y., et al., *Herpesviral latency-associated transcript gene promotes assembly of heterochromatin on viral lytic-gene promoters in latent infection.* Proc Natl Acad Sci U S A, 2005. **102**(44): p. 16055-9.

190. Samaniego, L.A., L. Neiderhiser, and N.A. DeLuca, *Persistence and expression of the herpes simplex virus genome in the absence of immediate-early proteins.* J Virol, 1998. **72**(4): p. 3307-20.

191. Ferenczy, M.W. and N.A. DeLuca, *Epigenetic modulation of gene expression from quiescent herpes simplex virus genomes.* J Virol, 2009. **83**(17): p. 8514-24.

192. Azzaz, A.M., et al., *Human heterochromatin protein 1alpha promotes nucleosome associations that drive chromatin condensation.* J Biol Chem, 2014. **289**(10): p. 6850-61.

193. Yamamoto, K. and M. Sonoda, *Self-interaction of heterochromatin protein 1 is required for direct binding to histone methyltransferase, SUV39H1*. Biochem Biophys Res Commun, 2003. **301**(2): p. 287-92.

194. Catez, F., et al., *HSV-1 genome subnuclear positioning and associations with host-cell PML-NBs and centromeres regulate LAT locus transcription during latency in neurons.* PLoS Pathog, 2012. **8**(8): p. e1002852.

195. Hayakawa, T., et al., *Cell cycle behavior of human HP1 subtypes: distinct molecular domains of HP1 are required for their centromeric localization during interphase and metaphase.* J Cell Sci, 2003. **116**(Pt 16): p. 3327-38.

196. Kwiatkowski, D.L., H.W. Thompson, and D.C. Bloom, *The polycomb group protein Bmi1 binds to the herpes simplex virus 1 latent genome and maintains repressive histone marks during latency.* J Virol, 2009. **83**(16): p. 8173-81.

197. Donze, D. and R.T. Kamakaka, *Braking the silence: how heterochromatic gene repression is stopped in its tracks.* Bioessays, 2002. **24**(4): p. 344-9.

198. West, A.G., M. Gaszner, and G. Felsenfeld, *Insulators: many functions, many mechanisms.* Genes Dev, 2002. **16**(3): p. 271-88.

199. Bell, A.C., A.G. West, and G. Felsenfeld, *The protein CTCF is required for the enhancer blocking activity of vertebrate insulators.* Cell, 1999. **98**(3): p. 387-96.

200. Xie, X., et al., *Systematic discovery of regulatory motifs in conserved regions of the human genome, including thousands of CTCF insulator sites.* Proc Natl Acad Sci U S A, 2007. **104**(17): p. 7145-50.

201. Kim, T.H., et al., *Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome.* Cell, 2007. **128**(6): p. 1231-45.

202. Cuddapah, S., et al., *Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains.* Genome Res, 2009. **19**(1): p. 24-32.

203. Amelio, A.L., P.K. McAnany, and D.C. Bloom, *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, 2006. **80**(5): p. 2358-68.

204. Chen, Q., et al., *CTCF-dependent chromatin boundary element between the latency-associated transcript and ICP0 promoters in the herpes simplex virus type 1 genome.* J Virol, 2007. **81**(10): p. 5192-201.

205. Zhao, J., et al., *Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome.* Science, 2008. **322**(5902): p. 750-6.

206. Farrell, M.J., A.T. Dobson, and L.T. Feldman, *Herpes simplex virus latency-associated transcript is a stable intron.* Proc Natl Acad Sci U S A, 1991. **88**(3): p. 790-4.

207. Thomas, D.L., et al., *The 2-kilobase intron of the herpes simplex virus type 1 latency-associated transcript has a half-life of approximately 24 hours in SY5Y and COS-1 cells.* J Virol, 2002. **76**(2): p. 532-40.

208. Flores, O., et al., *Mutational inactivation of herpes simplex virus 1 microRNAs identifies viral mRNA targets and reveals phenotypic effects in culture.* J Virol, 2013. **87**(12): p. 6589-603.

209. Sun, L. and Q. Li, *The miRNAs of herpes simplex virus (HSV).* Virol Sin, 2012. **27**(6): p. 333-8.

210. Umbach, J.L., et al., *MicroRNAs expressed by herpes simplex virus 1 during latent infection regulate viral mRNAs.* Nature, 2008. **454**(7205): p. 780-3.

211. Umbach, J.L., et al., *Analysis of human alphaherpesvirus microRNA expression in latently infected human trigeminal ganglia.* J Virol, 2009. **83**(20): p. 10677-83.

212. Jiang, X., et al., *A herpes simplex virus type 1 mutant disrupted for microRNA H2 with increased neurovirulence and rate of reactivation.* J Neurovirol, 2015. **21**(2): p. 199-209.

213. Glaser, R., et al., *Stress, loneliness, and changes in herpesvirus latency.* J Behav Med, 1985. **8**(3): p. 249-60.

214. Cushing, H., *The surgical aspects of major neuralgia of the trigeminal nerve.* The Journal of the American Medical Association, 1905.

215. Camarena, V., et al., *Nature and duration of growth factor signaling through receptor tyrosine kinases regulates HSV-1 latency in neurons.* Cell Host Microbe, 2010. **8**(4): p. 320-30.

216. Kobayashi, M., et al., *Control of viral latency in neurons by axonal mTOR signaling and the 4E-BP translation repressor.* Genes Dev, 2012. **26**(14): p. 1527-32.

217. Kim, J.Y., et al., *Transient reversal of episome silencing precedes VP16-dependent transcription during reactivation of latent HSV-1 in neurons.* PLoS Pathog, 2012. **8**(2): p. e1002540.

218. Du, T., G. Zhou, and B. Roizman, *HSV-1 gene expression from reactivated ganglia is disordered and concurrent with suppression of latency-associated transcript and miRNAs.* Proc Natl Acad Sci U S A, 2011. **108**(46): p. 18820-4.

219. Messer, H.G., et al., *Inhibition of H3K27me3-specific histone demethylases JMJD3 and UTX blocks reactivation of herpes simplex virus 1 in trigeminal ganglion neurons.* J Virol, 2015. **89**(6): p. 3417-20.

220. Liang, Y., et al., *A novel selective LSD1/KDM1A inhibitor epigenetically blocks herpes simplex virus lytic replication and reactivation from latency.* MBio, 2013. **4**(1): p. e00558-12.

221. Ertel, M.K., et al., *CTCF occupation of the herpes simplex virus 1 genome is disrupted at early times postreactivation in a transcription-dependent manner.* J Virol, 2012. **86**(23): p. 12741-59.

222. Ferenczy, M.W. and N.A. DeLuca, *Reversal of heterochromatic silencing of quiescent herpes simplex virus type 1 by ICP0.* J Virol, 2011. **85**(7): p. 3424-35.

223. Ferenczy, M.W., D.J. Ranayhossaini, and N.A. Deluca, *Activities of ICP0 involved in the reversal of silencing of quiescent herpes simplex virus 1.* J Virol, 2011. **85**(10): p. 4993-5002.

224. Preston, C.M. and M.J. Nicholl, *Induction of cellular stress overcomes the requirement of herpes simplex virus type 1 for immediate-early protein ICP0 and reactivates expression from quiescent viral genomes.* J Virol, 2008. **82**(23): p. 11775-83.

225. Bringhurst, R.M. and P.A. Schaffer, *Cellular stress rather than stage of the cell cycle enhances the replication and plating efficiencies of herpes simplex virus type 1 ICP0-viruses.* J Virol, 2006. **80**(9): p. 4528-37.

226. White, C.L., R.K. Suto, and K. Luger, *Structure of the yeast nucleosome core particle reveals fundamental changes in internucleosome interactions.* EMBO J, 2001. **20**(18): p. 5207-18.

227. Luger, K., et al., *Characterization of nucleosome core particles containing histone proteins made in bacteria.* J Mol Biol, 1997. **272**(3): p. 301-11.

228. Richmond, T.J., et al., *Structure of the nucleosome core particle at 7 A resolution.* Nature, 1984. **311**(5986): p. 532-7.

229. Arents, G., et al., *The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-handed superhelix.* Proc Natl Acad Sci U S A, 1991. **88**(22): p. 10148-52.

230. Cutter, A.R. and J.J. Hayes, *A brief review of nucleosome structure.* FEBS Lett, 2015.

231. Wolffe, A., *Chromatin: structure and function*1998: Academic Press.

232. Robinson, P.J. and D. Rhodes, *Structure of the '30 nm' chromatin fibre: a key role for the linker histone.* Curr Opin Struct Biol, 2006. **16**(3): p. 336-43.

233. Liang, G., et al., *Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome.* Proc Natl Acad Sci U S A, 2004. **101**(19): p. 7357-62.

234. Wang, Z., et al., *Combinatorial patterns of histone acetylations and methylations in the human genome.* Nat Genet, 2008. **40**(7): p. 897-903.

235. Lee, D.Y., et al., *A positive role for histone acetylation in transcription factor access to nucleosomal DNA.* Cell, 1993. **72**(1): p. 73-84.

236. Vettese-Dadey, M., et al., *Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro.* EMBO J, 1996. **15**(10): p. 2508-18.

237. Simpson, R.T., *Structure of chromatin containing extensively acetylated H3 and H4.* Cell, 1978. **13**(4): p. 691-9.

238. Ausio, J. and K.E. van Holde, *Histone hyperacetylation: its effects on nucleosome conformation and stability.* Biochemistry, 1986. **25**(6): p. 1421-8.

239. Anderson, J.D., P.T. Lowary, and J. Widom, *Effects of histone acetylation on the equilibrium accessibility of nucleosomal DNA target sites.* J Mol Biol, 2001. **307**(4): p. 977-85.

240. Garcia-Ramirez, M., C. Rocchini, and J. Ausio, *Modulation of chromatin folding by histone acetylation.* J Biol Chem, 1995. **270**(30): p. 17923-8.

241. Luger, K., et al., *Crystal structure of the nucleosome core particle at 2.8 A resolution.* Nature, 1997. **389**(6648): p. 251-60.

242. Shogren-Knaak, M., et al., *Histone H4-K16 acetylation controls chromatin structure and protein interactions.* Science, 2006. **311**(5762): p. 844-7.

243. Carey, M., B. Li, and J.L. Workman, *RSC exploits histone acetylation to abrogate the nucleosomal block to RNA polymerase II elongation.* Mol Cell, 2006. **24**(3): p. 481-7.

244. Peterson, C.L. and J.W. Tamkun, *The SWI-SNF complex: a chromatin remodeling machine?* Trends Biochem Sci, 1995. **20**(4): p. 143-6.

245. Hassan, A.H., S. Awad, and P. Prochasson, *The Swi2/Snf2 bromodomain is required for the displacement of SAGA and the octamer transfer of SAGA-acetylated nucleosomes.* J Biol Chem, 2006. **281**(26): p. 18126-34.

246. Ferreira, H., A. Flaus, and T. Owen-Hughes, *Histone modifications influence the action of Snf2 family remodelling enzymes by different mechanisms.* J Mol Biol, 2007. **374**(3): p. 563-79.

247. Schueler, M.G. and B.A. Sullivan, *Structural and functional dynamics of human centromeric chromatin.* Annu Rev Genomics Hum Genet, 2006. **7**: p. 301-13.

248. Pinheiro, I., et al., *Prdm3 and Prdm16 are H3K9me1 methyltransferases required for mammalian heterochromatin integrity.* Cell, 2012. **150**(5): p. 948-60.

249. Loyola, A., et al., *The HP1alpha-CAF1-SetDB1-containing complex provides H3K9me1 for Suv39-mediated K9me3 in pericentric heterochromatin.* EMBO Rep, 2009. **10**(7): p. 769-75.

250. Dillon, N. and R. Festenstein, *Unravelling heterochromatin: competition between positive and negative factors regulates accessibility.* Trends Genet, 2002. **18**(5): p. 252-8.

251. Volpe, T.A., et al., *Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi.* Science, 2002. **297**(5588): p. 1833-7.

252. Peters, A.H., et al., *Partitioning and plasticity of repressive histone methylation states in mammalian chromatin.* Mol Cell, 2003. **12**(6): p. 1577-89.

253. Seum, C., et al., *Ectopic HP1 promotes chromosome loops and variegated silencing in Drosophila.* EMBO J, 2001. **20**(4): p. 812-8.

254. Hahn, M., et al., *Suv4-20h2 mediates chromatin compaction and is important for cohesin recruitment to heterochromatin.* Genes Dev, 2013. **27**(8): p. 859-72.

255. Greeson, N.T., et al., *Di-methyl H4 lysine 20 targets the checkpoint protein Crb2 to sites of DNA damage.* J Biol Chem, 2008. **283**(48): p. 33168-74.

256. McDowell, T.L., et al., *Localization of a putative transcriptional regulator (ATRX) at pericentromeric heterochromatin and the short arms of acrocentric chromosomes.* Proc Natl Acad Sci U S A, 1999. **96**(24): p. 13983-8.

257. Ishov, A.M., O.V. Vladimirova, and G.G. Maul, *Heterochromatin and ND10 are cell-cycle regulated and phosphorylation-dependent alternate nuclear sites of the transcription repressor Daxx and SWI/SNF protein ATRX.* J Cell Sci, 2004. **117**(Pt 17): p. 3807-20.

258. Santenard, A., et al., *Heterochromatin formation in the mouse embryo requires critical residues of the histone variant H3.3.* Nat Cell Biol, 2010. **12**(9): p. 853-62.

259. Rangasamy, D., et al., *Pericentric heterochromatin becomes enriched with H2A.Z during early mammalian development.* EMBO J, 2003. **22**(7): p. 1599-607.

260. Fan, J.Y., et al., *H2A.Z alters the nucleosome surface to promote HP1alpha-mediated chromatin fiber folding.* Mol Cell, 2004. **16**(4): p. 655-61.

261. Grigoryev, S.A., *Higher-order folding of heterochromatin: protein bridges span the nucleosome arrays.* Biochem Cell Biol, 2001. **79**(3): p. 227-41.

262. Trojer, P. and D. Reinberg, *Facultative heterochromatin: is there a distinctive molecular signature?* Mol Cell, 2007. **28**(1): p. 1-13.

263. Rego, A., et al., *The facultative heterochromatin of the inactive X chromosome has a distinctive condensed ultrastructure.* J Cell Sci, 2008. **121**(Pt 7): p. 1119-27.

264. Taddei, A., et al., *The function of nuclear architecture: a genetic approach.* Annu Rev Genet, 2004. **38**: p. 305-45.

265. Okamoto, I., et al., *Epigenetic dynamics of imprinted X inactivation during early mouse development.* Science, 2004. **303**(5658): p. 644-9.

266. Chaumeil, J., et al., *Integrated kinetics of X chromosome inactivation in differentiating embryonic stem cells.* Cytogenet Genome Res, 2002. **99**(1-4): p. 75-84.

267. de Napoles, M., et al., *Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation.* Dev Cell, 2004. **7**(5): p. 663-76.

268. Kohlmaier, A., et al., *A chromosomal memory triggered by Xist regulates histone methylation in X inactivation.* PLoS Biol, 2004. **2**(7): p. E171.

269. Rougeulle, C., et al., *Differential histone H3 Lys-9 and Lys-27 methylation profiles on the X chromosome.* Mol Cell Biol, 2004. **24**(12): p. 5475-84.

270. Heard, E., *Delving into the diversity of facultative heterochromatin: the epigenetics of the inactive X chromosome.* Curr Opin Genet Dev, 2005. **15**(5): p. 482-9.

271. Nishioka, K., et al., *PR-Set7 is a nucleosome-specific methyltransferase that modifies lysine 20 of histone H4 and is associated with silent chromatin.* Mol Cell, 2002. **9**(6): p. 1201-13.

272. Gilbert, S.L., J.R. Pehrson, and P.A. Sharp, *XIST RNA associates with specific regions of the inactive X chromatin.* J Biol Chem, 2000. **275**(47): p. 36491-4.

273. Chakravarthy, S. and K. Luger, *The histone variant macro-H2A preferentially forms "hybrid nucleosomes".* J Biol Chem, 2006. **281**(35): p. 25522-31.

274. Changolkar, L.N. and J.R. Pehrson, *Reconstitution of nucleosomes with histone macroH2A1.2.* Biochemistry, 2002. **41**(1): p. 179-84.

275. Kalkhoven, E., *CBP and p300: HATs for different occasions.* Biochem Pharmacol, 2004. **68**(6): p. 1145-55.

276. Ragvin, A., et al., *Nucleosome binding by the bromodomain and PHD finger of the transcriptional cofactor p300.* J Mol Biol, 2004. **337**(4): p. 773-88.

277. Wang, F., C.B. Marshall, and M. Ikura, *Transcriptional/epigenetic regulator CBP/p300 in tumorigenesis: structural and functional versatility in target recognition.* Cell Mol Life Sci, 2013. **70**(21): p. 3989-4008.

278. Nakajima, T., et al., *RNA helicase A mediates association of CBP with RNA polymerase II.* Cell, 1997. **90**(6): p. 1107-12.

279. Cho, H., et al., *A human RNA polymerase II complex containing factors that modify chromatin structure.* Mol Cell Biol, 1998. **18**(9): p. 5355-63.

280. Kim, T.K., T.H. Kim, and T. Maniatis, *Efficient recruitment of TFIIB and CBP-RNA polymerase II holoenzyme by an interferon-beta enhanceosome in vitro.* Proc Natl Acad Sci U S A, 1998. **95**(21): p. 12191-6.

281. Imhof, A., et al., *Acetylation of general transcription factors by histone acetyltransferases.* Curr Biol, 1997. **7**(9): p. 689-92.

282. Ogryzko, V.V., et al., *The transcriptional coactivators p300 and CBP are histone acetyltransferases.* Cell, 1996. **87**(5): p. 953-9.

283. Jin, Q., et al., *Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation.* EMBO J, 2011. **30**(2): p. 249-62.

284. Schiltz, R.L., et al., *Overlapping but distinct patterns of histone acetylation by the human coactivators p300 and PCAF within nucleosomal substrates.* J Biol Chem, 1999. **274**(3): p. 1189-92.

285. Das, C., et al., *CBP/p300-mediated acetylation of histone H3 on lysine 56.* Nature, 2009. **459**(7243): p. 113-7.

286. Kawaguchi, Y., et al., *Herpes simplex virus 1 alpha regulatory protein ICP0 functionally interacts with cellular transcription factor BMAL1.* Proc Natl Acad Sci U S A, 2001. **98**(4): p. 1877-82.

287. Doi, M., J. Hirayama, and P. Sassone-Corsi, *Circadian regulator CLOCK is a histone acetyltransferase.* Cell, 2006. **125**(3): p. 497-508.

288. Katada, S. and P. Sassone-Corsi, *The histone methyltransferase MLL1 permits the oscillation of circadian gene expression.* Nat Struct Mol Biol, 2010. **17**(12): p. 1414-21.

289. Zhang, Y., et al., *Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation.* Genes Dev, 1999. **13**(15): p. 1924-35.

290. Kelly, R.D. and S.M. Cowley, *The physiological roles of histone deacetylase (HDAC) 1 and 2: complex co-stars with multiple leading parts.* Biochem Soc Trans, 2013. **41**(3): p. 741-9.

291. Wang, H.B. and Y. Zhang, *Mi2, an auto-antigen for dermatomyositis, is an ATP-dependent nucleosome remodeling factor.* Nucleic Acids Res, 2001. **29**(12): p. 2517-21.

292. Musselman, C.A., et al., *Bivalent recognition of nucleosomes by the tandem PHD fingers of the CHD4 ATPase is required for CHD4-mediated repression.* Proc Natl Acad Sci U S A, 2012. **109**(3): p. 787-92.

293. Musselman, C.A., et al., *Binding of the CHD4 PHD2 finger to histone H3 is modulated by covalent modifications.* Biochem J, 2009. **423**(2): p. 179-87.

294. Watson, A.A., et al., *The PHD and chromo domains regulate the ATPase activity of the human chromatin remodeler CHD4.* J Mol Biol, 2012. **422**(1): p. 3-17.

295. Hendrich, B. and A. Bird, *Identification and characterization of a family of mammalian methyl-CpG binding proteins.* Mol Cell Biol, 1998. **18**(11): p. 6538-47.

296. Murzina, N.V., et al., *Structural basis for the recognition of histone H4 by the histone-chaperone RbAp46.* Structure, 2008. **16**(7): p. 1077-85.

297. Mazumdar, A., et al., *Transcriptional repression of oestrogen receptor by metastasis-associated protein 1 corepressor.* Nat Cell Biol, 2001. **3**(1): p. 30-7.

298. Brackertz, M., et al., *p66alpha and p66beta of the Mi-2/NuRD complex mediate MBD2 and histone interaction.* Nucleic Acids Res, 2006. **34**(2): p. 397-406.

299. You, A., et al., *CoREST is an integral component of the CoREST- human histone deacetylase complex.* Proc Natl Acad Sci U S A, 2001. **98**(4): p. 1454-8.

300. Andres, M.E., et al., *CoREST: a functional corepressor required for regulation of neural-specific gene expression.* Proc Natl Acad Sci U S A, 1999. **96**(17): p. 9873-8.

301. Shi, Y.J., et al., *Regulation of LSD1 histone demethylase activity by its associated factors.* Mol Cell, 2005. **19**(6): p. 857-64.

302. Iwase, S., et al., *Characterization of BHC80 in BRAF-HDAC complex, involved in neuron-specific gene repression.* Biochem Biophys Res Commun, 2004. **322**(2): p. 601-8.

303. Lan, F., et al., *Recognition of unmethylated histone H3 lysine 4 links BHC80 to LSD1-mediated gene repression.* Nature, 2007. **448**(7154): p. 718-22.

304. Chinnadurai, G., *CtBP, an unconventional transcriptional corepressor in development and oncogenesis.* Mol Cell, 2002. **9**(2): p. 213-24.

305. Kristie, T.M., Y. Liang, and J.L. Vogel, *Control of alpha-herpesvirus IE gene expression by HCF-1 coupled chromatin modification activities.* Biochim Biophys Acta, 2010. **1799**(3-4): p. 257-65.

306. Dou, Y., et al., *Regulation of MLL1 H3K4 methyltransferase activity by its core components.* Nat Struct Mol Biol, 2006. **13**(8): p. 713-9.

307. Ruthenburg, A.J., et al., *Histone H3 recognition and presentation by the WDR5 module of the MLL1 complex.* Nat Struct Mol Biol, 2006. **13**(8): p. 704-12.

308. Platero, J.S., T. Hartnett, and J.C. Eissenberg, *Functional analysis of the chromo domain of HP1.* EMBO J, 1995. **14**(16): p. 3977-86.

309. Rea, S., et al., *Regulation of chromatin structure by site-specific histone H3 methyltransferases.* Nature, 2000. **406**(6796): p. 593-9.

310. Fritsch, L., et al., *A subset of the histone H3 lysine 9 methyltransferases Suv39h1, G9a, GLP, and SETDB1 participate in a multimeric complex.* Mol Cell, 2010. **37**(1): p. 46-56.

311. Eskeland, R., A. Eberharter, and A. Imhof, *HP1 binding to chromatin methylated at H3K9 is enhanced by auxiliary factors.* Mol Cell Biol, 2007. **27**(2): p. 453-65.

312. Margueron, R., et al., *Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms.* Mol Cell, 2008. **32**(4): p. 503-18.

313. Margueron, R., et al., *Role of the polycomb protein EED in the propagation of repressive histone marks.* Nature, 2009. **461**(7265): p. 762-7.

314. Ketel, C.S., et al., *Subunit contributions to histone methyltransferase activities of fly and worm polycomb group complexes.* Mol Cell Biol, 2005. **25**(16): p. 6857-68.

315. Pasini, D., et al., *Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity.* EMBO J, 2004. **23**(20): p. 4061-71.

316. Nowak, A.J., et al., *Chromatin-modifying complex component Nurf55/p55 associates with histones H3 and H4 and polycomb repressive complex 2 subunit Su(z)12 through partially overlapping binding sites.* J Biol Chem, 2011. **286**(26): p. 23388-96.

317. Cloos, P.A., et al., *Erasing the methyl mark: histone demethylases at the center of cellular differentiation and disease.* Genes Dev, 2008. **22**(9): p. 1115-40.

318. Peng, J.C., et al., *Jarid2/Jumonji coordinates control of PRC2 enzymatic activity and target gene occupancy in pluripotent cells.* Cell, 2009. **139**(7): p. 1290-302.

319. Casanova, M., et al., *Polycomblike 2 facilitates the recruitment of PRC2 Polycomb group complexes to the inactive X chromosome and to target loci in embryonic stem cells.* Development, 2011. **138**(8): p. 1471-82.

320. Li, X., et al., *Mammalian polycomb-like Pcl2/Mtf2 is a novel regulatory component of PRC2 that can differentially modulate polycomb activity both at the Hox gene cluster and at Cdkn2a genes.* Mol Cell Biol, 2011. **31**(2): p. 351-64.

321. Di Croce, L. and K. Helin, *Transcriptional regulation by Polycomb group proteins.* Nat Struct Mol Biol, 2013. **20**(10): p. 1147-55.

322. Levine, S.S., et al., *The core of the polycomb repressive complex is compositionally and functionally conserved in flies and humans.* Mol Cell Biol, 2002. **22**(17): p. 6070-8.

323. Gao, Z., et al., *PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes.* Mol Cell, 2012. **45**(3): p. 344-56.

324. Kaustov, L., et al., *Recognition and specificity determinants of the human cbx chromodomains.* J Biol Chem, 2011. **286**(1): p. 521-9.

325. Buchwald, G., et al., *Structure and E3-ligase activity of the Ring-Ring complex of polycomb proteins Bmi1 and Ring1b.* EMBO J, 2006. **25**(11): p. 2465-74.

326. Endoh, M., et al., *Histone H2A mono-ubiquitination is a crucial step to mediate PRC1-dependent repression of developmental genes to maintain ES cell identity.* PLoS Genet, 2012. **8**(7): p. e1002774.

327. Whetstine, J.R., et al., *Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases.* Cell, 2006. **125**(3): p. 467-81.

328. Metzger, E., et al., *LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription.* Nature, 2005. **437**(7057): p. 436-9.

329. Lan, F., et al., *A histone H3 lysine 27 demethylase regulates animal posterior development.* Nature, 2007. **449**(7163): p. 689-94.

330. Barski, A., et al., *High-resolution profiling of histone methylations in the human genome.* Cell, 2007. **129**(4): p. 823-37.

331. Cho, Y.W., et al., *PTIP associates with MLL3- and MLL4-containing histone H3 lysine 4 methyltransferase complex.* J Biol Chem, 2007. **282**(28): p. 20395-406.

332. Kosz-Vnenchak, M., et al., *Evidence for a novel regulatory pathway for herpes simplex virus gene expression in trigeminal ganglion neurons.* J Virol, 1993. **67**(9): p. 5383-93.

333. Nichol, P.F., et al., *Herpes simplex virus gene expression in neurons: viral DNA synthesis is a critical regulatory event in the branch point between the lytic and latent pathways.* J Virol, 1996. **70**(8): p. 5476-86.

334. Pesola, J.M., et al., *Herpes simplex virus 1 immediate-early and early gene expression during reactivation from latency under conditions that prevent infectious virus production.* J Virol, 2005. **79**(23): p. 14516-25.

335. Honess, R.W. and B. Roizman, *Regulation of herpesvirus macromolecular synthesis: sequential transition of polypeptide synthesis requires functional viral polypeptides.* Proc Natl Acad Sci U S A, 1975. **72**(4): p. 1276-80.

336. Batterson, W. and B. Roizman, *Characterization of the herpes simplex virion-associated factor responsible for the induction of alpha genes.* J Virol, 1983. **46**(2): p. 371-7.

337. Campbell, M.E., J.W. Palfreyman, and C.M. Preston, *Identification of herpes simplex virus DNA sequences which encode a trans-acting polypeptide responsible for stimulation of immediate early transcription.* J Mol Biol, 1984. **180**(1): p. 1-19.

338. Post, L.E., S. Mackem, and B. Roizman, *Regulation of alpha genes of herpes simplex virus: expression of chimeric genes produced by fusion of thymidine kinase with alpha gene promoters.* Cell, 1981. **24**(2): p. 555-65.

339. Gannon, F., et al., *Organisation and sequences at the 5' end of a cloned complete ovalbumin gene.* Nature, 1979. **278**(5703): p. 428-34.

340. Wasylyk, B., et al., *Specific in vitro transcription of conalbumin gene is drastically decreased by single-point mutation in T-A-T-A box homology sequence.* Proc Natl Acad Sci U S A, 1980. **77**(12): p. 7024-8.

341. Mackem, S. and B. Roizman, *Structural features of the herpes simplex virus alpha gene 4, 0, and 27 promoter-regulatory sequences which confer alpha regulation on chimeric thymidine kinase genes.* J Virol, 1982. **44**(3): p. 939-49.

342. Eisenberg, S.P., D.M. Coen, and S.L. McKnight, *Promoter domains required for expression of plasmid-borne copies of the herpes simplex virus thymidine kinase gene in virus-infected mouse fibroblasts and microinjected frog oocytes.* Mol Cell Biol, 1985. **5**(8): p. 1940-7.

343. Coen, D.M., S.P. Weinheimer, and S.L. McKnight, *A genetic approach to promoter recognition during trans induction of viral gene expression.* Science, 1986. **234**(4772): p. 53-9.

344. Smale, S.T. and D. Baltimore, *The "initiator" as a transcription control element.* Cell, 1989. **57**(1): p. 103-13.

345. Guzowski, J.F. and E.K. Wagner, *Mutational analysis of the herpes simplex virus type 1 strict late UL38 promoter/leader reveals two regions critical in transcriptional regulation.* J Virol, 1993. **67**(9): p. 5098-108.

346. Kubat, N.J., et al., *Specific histone tail modification and not DNA methylation is a determinant of herpes simplex virus type 1 latent gene expression.* J Virol, 2004. **78**(3): p. 1139-49.

347. Stevens, J.G., et al., *RNA complementary to a herpesvirus alpha gene mRNA is prominent in latently infected neurons.* Science, 1987. **235**(4792): p. 1056-9.

348. Preston, C.M. and M.J. Nicholl, *Repression of gene expression upon infection of cells with herpes simplex virus type 1 mutants impaired for immediate-early protein synthesis.* J Virol, 1997. **71**(10): p. 7807-13.

349. Jackson, S.A. and N.A. DeLuca, *Relationship of herpes simplex virus genome configuration to productive and persistent infections.* Proc Natl Acad Sci U S A, 2003. **100**(13): p. 7871-6.

350. Rock, D.L. and N.W. Fraser, *Detection of HSV-1 genome in central nervous system of latently infected mice.* Nature, 1983. **302**(5908): p. 523-5.

351. Stingley, S.W., et al., *Global analysis of herpes simplex virus type 1 transcription using an oligonucleotide-based DNA microarray.* J Virol, 2000. **74**(21): p. 9916-27.

352. Bertke, A.S., et al., *A5-positive primary sensory neurons are nonpermissive for productive infection with herpes simplex virus 1 in vitro.* J Virol, 2011. **85**(13): p. 6669-77.

353. Stringer, J.R., et al., *Quantitation of herpes simplex virus type 1 RNA in infected HeLa cells.* J Virol, 1977. **21**(3): p. 889-901.

354. Thorvaldsdottir, H., J.T. Robinson, and J.P. Mesirov, *Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration.* Brief Bioinform, 2013. **14**(2): p. 178-92.

355. Weinheimer, S.P. and S.L. McKnight, *Transcriptional and post-transcriptional controls establish the cascade of herpes simplex virus protein synthesis.* J Mol Biol, 1987. **195**(4): p. 819-33.

356. Thompson, R.L., C.M. Preston, and N.M. Sawtell, *De novo synthesis of VP16 coordinates the exit from HSV latency in vivo.* Plos Pathogens, 2009. **5**(3): p. e1000352.

357. Goodrich, L.D., F.J. Rixon, and D.S. Parris, *Kinetics of expression of the gene encoding the 65-kilodalton DNA-binding protein of herpes simplex virus type 1.* J Virol, 1989. **63**(1): p. 137-47.

358. Watson, R.J. and J.B. Clements, *A herpes simplex virus type 1 function continuously required for early and late virus RNA synthesis.* Nature, 1980. **285**: p. 329-330.

359. Preston, C.M., *Control of herpes simplex virus type 1 mRNA synthesis in cells infected with wild-type virus or the temperature-sensitive mutant tsK.* J Virol, 1979. **29**(1): p. 275-84.

360. Dixon, R.A. and P.A. Schaffer, *Fine-structure mapping and functional analysis of temperature-sensitive mutants in the gene encoding the herpes simplex virus type 1 immediate early protein VP175.* J Virol, 1980. **36**(1): p. 189-203.

361. Leib, D.A., et al., *Immediate-early regulatory gene mutants define different stages in the establishment and reactivation of herpes simplex virus latency.* J Virol, 1989. **63**(2): p. 759-68.

362. Sacks, W.R. and P.A. Schaffer, *Deletion mutants in the gene encoding the herpes simplex virus type 1 immediate-early protein ICP0 exhibit impaired growth in cell culture.* J Virol, 1987. **61**(3): p. 829-39.

363. Stow, N.D. and E.C. Stow, *Isolation and characterization of a herpes simplex virus type 1 mutant containing a deletion within the gene encoding the immediate early polypeptide Vmw110.* J Gen Virol, 1986. **67 (Pt 12)**: p. 2571-85.

364. Batchelor, A.H. and P. O'Hare, *Regulation and cell-type-specific activity of a promoter located upstream of the latency-associated transcript of herpes simplex virus type 1.* J Virol, 1990. **64**(7): p. 3269-79.

365. Farrell, M.J., et al., *Effect of the transcription start region of the herpes simplex virus type 1 latency-associated transcript promoter on expression of productively infected neurons in vivo.* J Virol, 1994. **68**(9): p. 5337-43.

366. Rivera-Gonzalez, R., et al., *The role of ICP4 repressor activity in temporal expression of the IE-3 and latency-associated transcript promoters during HSV-1 infection.* Virology, 1994. **202**(2): p. 550-64.

367. Devi-Rao, G.B., et al., *Relationship between polyadenylated and nonpolyadenylated herpes simplex virus type 1 latency-associated transcripts.* J Virol, 1991. **65**(5): p. 2179-90.

368. Bohenzky, R.A., et al., *Identification of a promoter mapping within the reiterated sequences that flank the herpes simplex virus type 1 UL region.* J Virol, 1993. **67**(2): p. 632-42.

369. Yeh, L. and P.A. Schaffer, *A novel class of transcripts expressed with late kinetics in the absence of ICP4 spans the junction between the long and short segments of the herpes simplex virus type 1 genome.* J Virol, 1993. **67**(12): p. 7373-82.

370. Zhao, S., et al., *Comparison of RNA-Seq and Microarray in Transcriptome Profiling of Activated T Cells.* PLoS One, 2014. **9**(1): p. e78644.

371. Yager, D.R. and D.M. Coen, *Analysis of the transcript of the herpes simplex virus DNA polymerase gene provides evidence that polymerase expression is inefficient at the level of translation.* J Virol, 1988. **62**(6): p. 2007-15.

372. Crute, J.J., et al., *Herpes simplex virus 1 helicase-primase: a complex of three herpes-encoded gene products.* Proc Natl Acad Sci U S A, 1989. **86**(7): p. 2186-9.

373. Olivo, P.D., N.J. Nelson, and M.D. Challberg, *Herpes simplex virus DNA replication: the UL9 gene encodes an origin-binding protein.* Proc Natl Acad Sci U S A, 1988. **85**(15): p. 5414-8.

374. Baradaran, K., C.E. Dabrowski, and P.A. Schaffer, *Transcriptional analysis of the region of the herpes simplex virus type 1 genome containing the UL8, UL9, and UL10 genes and identification of a novel delayed-early gene product, OBPC.* J Virol, 1994. **68**(7): p. 4251-61.

375. Cook, W.J., et al., *Induction of transcription by a viral regulatory protein depends on the relative strengths of functional TATA boxes.* Mol Cell Biol, 1995. **15**(9): p. 4998-5006.

376. Imbalzano, A.N. and N.A. DeLuca, *Substitution of a TATA box from a herpes simplex virus late gene in the viral thymidine kinase promoter alters ICP4 inducibility but not temporal expression.* J Virol, 1992. **66**(9): p. 5453-63.

377. Smale, S.T., et al., *The initiator element: a paradigm for core promoter heterogeneity within metazoan protein-coding genes.* Cold Spring Harb Symp Quant Biol, 1998. **63**: p. 21-31.

378. Lieu, P.T. and E.K. Wagner, *Two leaky-late HSV-1 promoters differ significantly in structural architecture.* Virology, 2000. **272**(1): p. 191-203.

379. Jamieson, D.R., et al., *Quiescent viral genomes in human fibroblasts after infection with herpes simplex virus type 1 Vmw65 mutants.* J Gen Virol, 1995. **76 (Pt 6)**: p. 1417-31.

380. Cliffe, A.R. and D.M. Knipe, *Herpes simplex virus ICP0 promotes both histone removal and acetylation on viral DNA during lytic infection.* J Virol, 2008. **82**(24): p. 12030-8.

381. Terry-Allison, T., C.A. Smith, and N.A. DeLuca, *Relaxed repression of herpes simplex virus type 1 genomes in Murine trigeminal neurons.* J Virol, 2007. **81**(22): p. 12394-405.

382. Meshorer, E., *Chromatin in embryonic stem cell neuronal differentiation.* Histol Histopathol, 2007. **22**(3): p. 311-9.

383. Ahmad, K. and S. Henikoff, *The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly.* Mol Cell, 2002. **9**(6): p. 1191-200.

384. Chow, C.M., et al., *Variant histone H3.3 marks promoters of transcriptionally active genes during mammalian cell division.* EMBO Rep, 2005. **6**(4): p. 354-60.

385. Hake, S.B., et al., *Expression patterns and post-translational modifications associated with mammalian histone H3 variants.* J Biol Chem, 2006. **281**(1): p. 559-68.

386. Felsenfeld, G., et al., *Chromatin boundaries and chromatin domains.* Cold Spring Harb Symp Quant Biol, 2004. **69**: p. 245-50.

387. Zhao, H. and A. Dean, *An insulator blocks spreading of histone acetylation and interferes with RNA polymerase II transfer between an enhancer and gene.* Nucleic Acids Res, 2004. **32**(16): p. 4903-19.

388. Bloom, D.C., N.V. Giordani, and D.L. Kwiatkowski, *Epigenetic regulation of latent HSV-1 gene expression.* Biochim Biophys Acta, 2010. **1799**(3-4): p. 246-56.

389. Jurak, I., et al., *Numerous conserved and divergent microRNAs expressed by herpes simplex viruses 1 and 2.* J Virol, 2010. **84**(9): p. 4659-72.

390. Rock, D.L., et al., *Detection of latency-related viral RNAs in trigeminal ganglia of rabbits latently infected with herpes simplex virus type 1.* J Virol, 1987. **61**(12): p. 3820-6.

391. Nedialkov, Y.A. and S.J. Triezenberg, *Quantitative assessment of in vitro interactions implicates TATA-binding protein as a target of the VP16C transcriptional activation region.* Arch Biochem Biophys, 2004. **425**(1): p. 77-86.

392. Hall, D.B. and K. Struhl, *The VP16 activation domain interacts with multiple transcriptional components as determined by protein-protein cross-linking in vivo.* Journal of Biological Chemistry, 2002. **277**(48): p. 46043-50.

393. Gu, H. and B. Roizman, *Herpes simplex virus-infected cell protein 0 blocks the silencing of viral DNA by dissociating histone deacetylases from the CoREST-REST complex.* Proc Natl Acad Sci U S A, 2007. **104**(43): p. 17134-9.

394. Blondeau, J.M., F.Y. Aoki, and G.B. Glavin, *Stress-induced reactivation of latent herpes simplex virus infection in rat lumbar dorsal root ganglia.* J Psychosom Res, 1993. **37**(8): p. 843-9.

395. Jenkins, F.J. and A. Baum, *Stress and reactivation of latent herpes simplex virus: a fusion of behavioral medicine and molecular biology.* Ann Behav Med, 1995. **17**(2): p. 116-23.

396. Neumann, D.M., et al., *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, 2007. **81**(23): p. 13248-53.

397. Coleman, H.M., et al., *Histone modifications associated with herpes simplex virus type 1 genomes during quiescence and following ICP0-mediated de-repression.* J Gen Virol, 2008. **89**(Pt 1): p. 68-77.

398. Shen, X., et al., *Involvement of actin-related proteins in ATP-dependent chromatin remodeling.* Mol Cell, 2003. **12**(1): p. 147-55.

399. Shen, X., et al., *A chromatin remodelling complex involved in transcription and DNA processing.* Nature, 2000. **406**(6795): p. 541-4.

400. Wang, L., et al., *INO80 Facilitates Pluripotency Gene Activation in Embryonic Stem Cell Self-Renewal, Reprogramming, and Blastocyst Development.* Cell Stem Cell, 2014. **14**(5): p. 575-91.

401. Sudarsanam, P. and F. Winston, *The Swi/Snf family nucleosome-remodeling complexes and transcriptional control.* Trends Genet, 2000. **16**(8): p. 345-51.

402. Snyers, L., et al., *Distinct chromatin signature of histone H3 variant H3.3 in human cells.* Nucleus, 2014. **5**(5).

403. Pina, B. and P. Suau, *Changes in histones H2A and H3 variant composition in differentiating and mature rat brain cortical neurons.* Dev Biol, 1987. **123**(1): p. 51-8.

404. Bosch, A. and P. Suau, *Changes in core histone variant composition in differentiating neurons: the roles of differential turnover and synthesis rates.* Eur J Cell Biol, 1995. **68**(3): p. 220-5.

405. Schotta, G., et al., *A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin.* Genes Dev, 2004. **18**(11): p. 1251-62.

406. Bartova, E., et al., *Histone modifications and nuclear architecture: a review.* J Histochem Cytochem, 2008. **56**(8): p. 711-21.

407. Mador, N., et al., *Herpes simplex virus type 1 latency-associated transcripts suppress viral replication and reduce immediate-early gene mRNA levels in a neuronal cell line.* J Virol, 1998. **72**(6): p. 5067-75.

408. Jiang, X., et al., *Increased neurovirulence and reactivation of the herpes simplex virus type 1 latency-associated transcript (LAT)-negative mutant dLAT2903 with a disrupted LAT miR-H2.* J Neurovirol, 2015.

409. Ma, J.Z., et al., *Lytic gene expression is frequent in HSV-1 latent infection and correlates with the engagement of a cell-intrinsic transcriptional response.* PLoS Pathog, 2014. **10**(7): p. e1004237.

410. Sasaki, A., W.C. de Vega, and P.O. McGowan, *Biological embedding in mental health: an epigenomic perspective.* Biochem Cell Biol, 2013. **91**(1): p. 14-21.

411. Glass, M. and R.D. Everett, *Components of promyelocytic leukemia nuclear bodies (ND10) act cooperatively to repress herpesvirus infection.* J Virol, 2013. **87**(4): p. 2174-85.