Regulation of Host and Viral RNA Expression during Gammaherpesvirus Infection

Susan Priscilla Canny
Washington University in St. Louis

Follow this and additional works at: http://openscholarship.wustl.edu/art_sci_etds
Part of the Biology Commons

Recommended Citation
Canny, Susan Priscilla, "Regulation of Host and Viral RNA Expression during Gammaherpesvirus Infection" (2015). Arts & Sciences Electronic Theses and Dissertations. 462.
http://openscholarship.wustl.edu/art_sci_etds/462

This Dissertation is brought to you for free and open access by the Arts & Sciences at Washington University Open Scholarship. It has been accepted for inclusion in Arts & Sciences Electronic Theses and Dissertations by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.
Regulation of Host and Viral RNA Expression during Gammaherpesvirus Infection

by

Susan Priscilla Canny

A dissertation presented to the Graduate School of Arts and Sciences of Washington University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

May 2015

St. Louis, Missouri
Table of Contents

List of Figures ........................................................................................................................... vii
List of Tables ............................................................................................................................ viii
List of Abbreviations ............................................................................................................... viii
Acknowledgments ................................................................................................................... x
Abstract ................................................................................................................................... xiii

Chapter 1: Introduction ............................................................................................................. 1
  1.1 Herpesviridae Family and Viral Life Cycle ................................................................. 2
    1.1.1 Murine Gammaherpesvirus 68 Model System ...................................................... 3
    1.1.2 In vivo infection and pathogenesis ......................................................................... 4
    1.1.3 Pervasive Transcription of MHV68 ...................................................................... 4
    1.1.4 ORF 50: The Molecular Latent-to-Lytic Switch ..................................................... 6
    1.1.5 Role of Interferons in MHV68 Infection ................................................................. 7
    1.1.6 Gammaherpesvirus Latency in Pathology and Protection from Disease ............. 8
  1.2 Functions of the Products of Pervasive Transcription ............................................... 9
    1.2.1 Functions of Noncoding RNAs in Herpesviruses ................................................. 10
    1.2.2 Herpesvirus ncRNAs expressed during lytic infection ......................................... 11
    1.2.3 Herpesvirus ncRNAs expressed during latent infection ........................................ 12
  1.3 References ....................................................................................................................... 14

Chapter 2: Pervasive transcription of a herpesvirus genome generates functionally important RNAs .............................................................................................................................. 28
  2.1 Abstract ......................................................................................................................... 29
  2.2 Introduction .................................................................................................................... 30
  2.3 Results ........................................................................................................................... 32
    2.3.1 Establishment of a system to target MHV68 transcripts: antisense targeting of a known essential gene alters late viral protein expression ........................................... 32
    2.3.2 Targeting EGR transcripts alters surface expression of the late viral protein encoded by ORF 4 .................................................................................................................. 33
    2.3.3 Mapping 5' and 3' ends of EGR 26 and EGR 27 transcripts .................................. 34
    2.3.4 EGR 26 and EGR 27 transcripts are enriched in the nucleus of infected cells ........ 36
2.3.5 Targeting EGR 27 transcripts alters surface expression of ORF 4 protein ........................................ 37
2.3.6 Targeting of EGR 27 transcripts alters expression of multiple viral genes ........................................... 39
2.3.7 EGR 26c ASO decreases specific genes of all kinetic classes ................................................................. 40
2.4 Discussion ...................................................................................................................................................... 41
  2.4.1 Implications of transcriptional complexity of EGR-encoded RNAs ......................................................... 42
  2.4.2 Potential mechanism for EGRs 26 and 27 transcript function ............................................................... 43
2.5 Materials and Methods ............................................................................................................................... 45
  2.5.1 Cells, virus, and virus assays .................................................................................................................. 45
  2.5.2 Antisense oligonucleotides .................................................................................................................... 45
  2.5.3 Western blot analysis, antisera and antibodies ....................................................................................... 46
  2.5.4 Flow cytometric analysis for detection of surface expression of ORF 4 protein .................................. 46
  2.5.5 RNA sequencing (RNA-Seq) library construction and expression analysis ........................................ 47
  2.5.6 Identification of polyadenylated reads from RNA-Seq data sets ........................................................... 48
  2.5.7 Rapid Amplification of cDNA Ends (RACE) .......................................................................................... 48
  2.5.8 Northern blot analysis ............................................................................................................................ 49
  2.5.9 Subcellular Fractionation ........................................................................................................................ 50
  2.5.10 Quantitative reverse transcriptase-PCR (qRT-PCR) ........................................................................... 50
  2.5.11 Statistical analysis ...................................................................................................................................... 51
2.6 Acknowledgements ......................................................................................................................................... 52
2.7 References ...................................................................................................................................................... 53

Chapter 3: Stat1 regulates transcription of the murine gammaherpesvirus 68 latent-to-lytic switch gene, ORF 50 ...................................................................................................................................................... 78
3.1 Abstract ......................................................................................................................................................... 79
3.2 Introduction .................................................................................................................................................... 80
3.3 Results ......................................................................................................................................................... 83
  3.3.1 IFNγ suppresses viral replication in bone marrow-derived macrophages .................................................. 83
  3.3.2 Phosphorylation of Stat1 is required for IFNγ-mediated suppression of viral replication in bone marrow-derived macrophages ........................................................................................................... 83
  3.3.3 IFNγ-mediated regulation of ORF 50 transcript levels is dependent on Stat1 .......................................... 83
  3.3.4 IFNγ suppresses ORF 50 promoter activity ............................................................................................ 84
  3.3.5 Stat1 is required for IFNγ-mediated regulation of ORF 50 promoter activity .......................................... 84
3.3.6 IFNγ-mediated suppression of ORF 50 promoter activity is independent of predicted GAS elements ................................................................. 85
3.3.7 IFNγ-independent role for Stat1 in ORF 50 promoter activity .................................................. 85
3.4 Discussion ............................................................................................................................. 87
3.4.1 Regulation of ORF 50 Expression ................................................................................... 87
3.4.2 Mechanisms of Stat1 regulation of MHV68 ................................................................. 88
3.4.3 Implications of Stat1 regulation of ORF 50 for viral pathogenesis ............................. 89
3.5 Materials and Methods ....................................................................................................... 91
3.5.1 Cells, virus and viral assays .......................................................................................... 91
3.5.2 Quantitative reverse transcriptase-PCR (qRT-PCR) .................................................. 91
3.5.3 Reporter plasmids and transfections .............................................................................. 92
3.5.4 Luciferase promoter assay ............................................................................................ 92
3.5.5 Statistical analyses ........................................................................................................ 92
3.6 Acknowledgements ............................................................................................................ 93
3.7 References .......................................................................................................................... 94

Chapter 4: Latent gammaherpesvirus 68 infection induces distinct transcriptional changes in different organs ......................................................................................................................... 111
4.1 Abstract .................................................................................................................................. 112
4.2 Text ......................................................................................................................................... 113
4.2.1 Chronic MHV68 alters the host transcriptional profile ................................................. 114
4.2.2 Distinct pathways are regulated by latent viral infection in livers and spleens......... 117
4.2.3 Differentially expressed genes are significantly associated with genes regulated by the cytokine, interferon gamma .......................................................... 119
4.3 Acknowledgements ............................................................................................................... 122
4.4 References ............................................................................................................................ 123

Chapter 5: Summary and Future Directions .............................................................................. 138
5.1 Summary .................................................................................................................................. 139
5.2 Function of products of pervasive transcription in MHV68 .............................................. 141
5.2.1 Further functional analysis of EGR 26 and EGR 27 transcripts ....................................... 141
5.2.2 Transcriptional architecture and functional analysis of other EGR-encoded transcripts . 145
5.2.3 Role of EGR-encoded transcripts during infection in vivo ........................................... 147
5.3 Stat1 regulation of gammaherpesvirus 68 ORF 50 transcription ........................................ 148
5.4 Latent infection alters host transcription ................................................................. 151
  5.4.1 Relative contribution of latency establishment versus reactivation to latency-induced phenotypes .......................................................................................................................................................................................... 151
  5.4.2 Effect of MHV68 latency on specific cell types and tissues ........................................ 152
  5.4.3 Effect of latent infection on T cells ............................................................................. 153
5.5 References .................................................................................................................. 155
List of Figures

Figure 2.1  An antisense oligonucleotide to ORF 6 decreases ORF 6 transcript and protein expression and late gene expression ................................................................. 60
Figure 2.2  Targeting of EGRs 9, 16, 18, 23, 26, and 27 reduces ORF4 surface expression on infected cells .................................................................................... 62
Figure 2.3  EGR 26 - 27 transcriptional architecture ....................................................................................................................... 63
Figure 2.4  EGR 26 and 27 transcripts are enriched in nucleus of infected cells ............... 65
Figure 2.5  Effect of targeting EGR 26 and EGR 27 transcripts on ORF 4 surface expression ........................................................................................................ 66
Figure 2.6  Effect on immediate-early, early, and late genes by EGR27 ........................................ 68
Figure 2.7  EGR 26c ASO decreases specific genes of all kinetic classes ......................... 70
Figure 2.8  Genome-wide RNA expression in MHV-68 as detected by RNA-Seq ............ 72

Figure 3.1  IFNγ-mediated suppression of lytic replication in bone marrow-derived macrophages is mediated by STAT1 ................................................................. 101
Figure 3.2  Stat1 Y701 residue is necessary for IFNγ-mediated suppression of lytic replication in bone marrow-derived macrophages .......................................... 103
Figure 3.3  IFNγ suppresses MHV68 lytic switch gene 50 in a Stat1-dependent manner .... 104
Figure 3.4  IFNγ suppresses ORF 50 promoter activity .................................................... 105
Figure 3.5  Stat1 is necessary for IFNγ-mediated suppression of ORF 50 promoter activity .................................................................................................................. 107
Figure 3.6  IFNγ-mediated suppression of ORF 50 promoter activity is contained within the minimal 410bp promoter and does not require predicted Stat binding elements .................................................................................................. 108
Figure 3.7  Stat1 regulates ORF 50 promoter activity in the absence of IFNγ ................... 110

Figure 4.1  Latent MHV68 alters host gene expression ..................................................... 126
Figure 4.2  Validation of microarray analyses by qRT-PCR ............................................ 128
Figure 4.3  Pathway analysis of differentially expressed genes in latently infected mice .... 130
Figure 4.4  Overlap of genes regulated by IFNγ and latent viral infection ....................... 132

Figure 5.1  Multiple polyadenylated transcripts emanate from EGR 23 ......................... 161
List of Tables

Table 2.1: Antisense oligonucleotide probes.................................................................74
Table 2.2: PCR primers for northern blot probes ............................................................75
Table 2.3: 3’ ends of MHV68 transcripts identified by RNA sequencing.........................76

Table 4.1: Primers used in this study for qRT-PCR ..........................................................134
Table 4.2: Genes most differentially expressed in spleens of latently-infected mice.........135
Table 4.3: Genes most differentially expressed in livers of latently-infected mice ..........136
Table 4.4: Genes most differentially expressed in brains of latently-infected mice...........137
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASO</td>
<td>Antisense oligonucleotide</td>
</tr>
<tr>
<td>ChIP</td>
<td>Chromatin immunoprecipitation</td>
</tr>
<tr>
<td>EAE</td>
<td>Experimental autoimmune encephalomyelitis</td>
</tr>
<tr>
<td>EBV</td>
<td>Epstein-Barr Virus</td>
</tr>
<tr>
<td>EGR</td>
<td>Expressed Genomic Region</td>
</tr>
<tr>
<td>GAS</td>
<td>Gamma-Activated Sequence</td>
</tr>
<tr>
<td>HCMV</td>
<td>Human cytomegalovirus</td>
</tr>
<tr>
<td>HSV</td>
<td>Herpes simplex virus</td>
</tr>
<tr>
<td>Hpi</td>
<td>Hours postinfection</td>
</tr>
<tr>
<td>IFN</td>
<td>Interferon</td>
</tr>
<tr>
<td>IRF</td>
<td>Interferon Regulatory Factor</td>
</tr>
<tr>
<td>KSHV</td>
<td>Kaposi's sarcoma-associated herpesvirus</td>
</tr>
<tr>
<td>LANA</td>
<td>Latency-associated nuclear antigen</td>
</tr>
<tr>
<td>LAT</td>
<td>Latency-associated transcript</td>
</tr>
<tr>
<td>MHV68</td>
<td>Murine gammaherpesvirus 68</td>
</tr>
<tr>
<td>MMP-9</td>
<td>Matrix Metalloproteinase 9</td>
</tr>
<tr>
<td>MOI</td>
<td>Multiplicity of infection</td>
</tr>
<tr>
<td>ncRNA</td>
<td>Noncoding RNA</td>
</tr>
<tr>
<td>ORF</td>
<td>Open Reading Frame</td>
</tr>
<tr>
<td>PAN</td>
<td>Polyadenylated nuclear</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>PRC2</td>
<td>Polycomb Repressive Complex 2</td>
</tr>
<tr>
<td>PFU</td>
<td>Plaque forming unit</td>
</tr>
<tr>
<td>qRT-PCR</td>
<td>Quantitative reverse transcription polymerase chain reaction</td>
</tr>
<tr>
<td>RACE</td>
<td>Rapid Amplification of cDNA Ends</td>
</tr>
<tr>
<td>RIP</td>
<td>RNA Immunoprecipitation</td>
</tr>
<tr>
<td>RLU</td>
<td>Relative light unit</td>
</tr>
<tr>
<td>RTA</td>
<td>Replication and Transcription Activator</td>
</tr>
<tr>
<td>Seq</td>
<td>Sequencing (via to next-generation sequencing technologies)</td>
</tr>
<tr>
<td>STAT</td>
<td>Signal Transducer and Activator of Transcription</td>
</tr>
</tbody>
</table>
Acknowledgments

First, I would like to thank my dissertation committee: Drs. Wayne Yokoyama, Michael Diamond, Deborah Lenschow, Eugene Oltz, and Jean Schaffer. Their guidance and advice has been an integral part of my training. Both as a committee and individually, they have provided valuable insights and thoughtful suggestions. I truly appreciate the time and dedication that they have invested in my training and the opportunity to interact with a group of such outstanding scientists.

Next, I would like to thank my dissertation mentor, Dr. Herbert (Skip) Virgin. I am especially grateful for his constant support of every aspect of my training, the time that he has devoted to advise and guide my studies, and the investment that he has made to help me develop as a scientist. I appreciate his boundless enthusiasm for science (especially new data and interesting questions), good humor, and patience throughout the course of my training. His advice and insights on matters both large and small have been invaluable. I admire his ability to articulate the critical questions and his gift for perceiving novel connections between broad topics. He has created an excellent lab environment in which to train.

I am grateful to all past and present members of the Virgin lab with whom I have had the good fortune to train. Both as colleagues and as friends, they have had a profound impact on my training and helped me through the rough spots. I would especially like to thank Drs. Steve Johnson and Tiffany Reese for all of their guidance, mentorship and advice. I appreciate their patience, encouragement and honest feedback throughout my training. I have learned a lot from working with both of them and I thank them for their continued dedication to my training.
I am especially thankful to my MSTP colleagues, Nicole Maloney and Christine Yokoyama, who have become close friends. Nicole has always been an excellent source of advice and a great sounding board and has always made time to discuss data and lab issues even with her now busy clinical schedule. Christine has been a wonderful addition to the lab and has taught me everything I know about Illustrator. She has been patient, thoughtful and encouraging and gotten me out of the lab when needed. She always asks insightful, perceptive questions, and I appreciate her willingness to talk through experiments and perplexing results with me.

I am grateful to the members of the gammaherpesvirus latency and gene expression/virology small group in all its various iterations over the years, especially Drs. Erin Willert, Joy Loh, Megan Goodwin, Donna MacDuff, and Tim Nice among many others. These scientists have provided helpful feedback on experimental design and interpretation of results and encouragement when needed. I am very thankful for the help and support of Amal Kambal, Xin Zhang, Erning Duan, and Cathy Liu, who have kindly and patiently helped with experiments especially in these final months as I sought to finish several projects simultaneously.

One of many wonderful aspects of Washington University is the collaborative atmosphere that is fostered. I am thankful for the opportunity to collaborate both inside and outside the university and have learned much from each experience. Dr. Jean Schaffer and Ben Scruggs have provided helpful experimental advice, protocols and reagents as I ventured in to studies of RNA. Drs. Ramnik Xavier and Gautam Goel have been wonderful collaborators on microarray analyses and kindly welcomed me in to their lab. Gautam has worked with me patiently and persistently to figure out the best way to analyze and present these data. I have learned much about both computational biology and working effectively across disciplines with him. Although our project with Drs. David Holtzman and Jungsu Kim, Mary Beth Finn and Floy
Stewart did not yield results included in this dissertation, working with them showed me early on what it was like to work with truly fantastic collaborators. Thanks to the Washington University Genome Technology Access Center (GTAC), especially Mike Heinz and Twyla Juehne, for performing microarrays and Illumina sequencing and offering technical advice.

Finally, I would like to thank my friends from the MSTP, WUSM, Stanford and beyond, and my family, especially my parents, whose support, encouragement, advice and understanding have been invaluable.

The work in this dissertation was supported by grants from the National Institutes of Health, including R01 CA096511 and U54 AI057160 ARRA Administrative Supplement (to H.W.V.). S.P.C. is supported by an individual training grant from the National Heart, Lung and Blood Institute (F30 HL099019). GTAC is partially supported by NCI Cancer Center Support Grant P30 CA91842 to the Siteman Cancer Center and by ICTS/CTSA Grant UL1 TR000448 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH), and NIH Roadmap for Medical Research.

Susan Priscilla Canny

Washington University in St. Louis

May 2015
ABSTRACT OF THE DISSERTATION
Regulation of Host and Viral RNA Expression during Gammaherpesvirus Infection
by
Susan Priscilla Canny
Doctor of Philosophy in Biology & Biomedical Sciences
(Immunology)
Washington University in St. Louis, 2015
Professor Herbert Virgin, Chairperson

Gammaherpesviruses are oncogenic viruses that establish latent infection and can periodically reactivate. Replication, reactivation from latency and gammaherpesvirus-mediated pathology are regulated by viral and host factors. Long-standing interests in the field include understanding viral gene expression and regulation during lytic and latent infection and the effects of chronic viral infection on the host. Murine gammaherpesvirus 68 (MHV68) is used to model features of human gammaherpesvirus biology.

In this dissertation, I investigated: (1) the function of antisense viral transcripts; (2) Stat1 regulation of MHV68 gene 50 promoters; and (3) the effect of chronic infection on host gene expression. Our laboratory identified extensive transcription outside of annotated open reading frames (ORFs) in the MHV68 genome, which we termed expressed genomic regions (EGRs). Similar pervasive transcription is also observed in the related human gammaherpesvirus, Kaposi's sarcoma-associated herpesvirus (KSHV). Despite their prevalence, the functions of these novel transcripts are largely unknown. To study their function in MHV68 gene expression, we targeted thirteen EGRs using strand-specific antisense oligonucleotides and found targeting of six EGRs reduced late viral gene expression. We identified three transcripts enriched in the nucleus of infected cells, the first putative nuclear noncoding RNA to be identified in MHV68.

xiii
ORF50, the latent-to-lytic switch gene in MHV68 and KSHV, is sufficient to induce lytic reactivation from latently infected cells. We determined that interferon-gamma (IFNγ), which controls viral reactivation of MHV68, suppressed ORF50 transcription by regulating its promoters in a Stat1-dependent manner, suggesting that the virus maintains Stat1-responsive elements to control the viral life cycle.

To assess the effect of chronic viral infection on host gene expression, we analyzed the transcriptional profile of tissues from latently infected mice and identified over 600 differentially expressed genes. We identified organ-specific transcriptional changes in gene expression. Most differentially expressed genes were involved in immune signaling, immune response, or cell cycle pathways and many were associated with IFNγ signaling.

Taken together, these studies provide the first functional evidence for the importance of regions of pervasive transcription in MHV68 and add to our understanding of how the cytokine, IFNγ, controls viral infection and may modulate gene transcription in the host.
Chapter 1:

Introduction
1.1 *Herpesviridae* Family and Viral Life Cycle

The *Herpesviridae* family consists of three subfamilies, *Alpha*-, *Beta*-, and *Gammaherpesvirinae*, which share common virion architecture and large double-stranded DNA genomes and are able to establish lifelong latency in their hosts (96). During productive infection, herpesviruses replicate in the nucleus; release of progeny virus occurs with destruction of the infected cell. There is an initial production of virus during acute infection of a naïve host. The immune system is unable to completely clear herpesvirus infection, resulting in chronic infection. There are three components of chronic herpesvirus infections: (1) latent infection, (2) reactivation from latency and (3) persistent infection (134). Latency is characterized experimentally as the presence of viral genome without preformed infectious virus. Latent virus retains the capacity to resume production of infectious virus and return to the lytic viral life cycle. The emergence of virus from latently infected cells is termed reactivation. Persistent infection is the production of infectious virus during chronic infection and can result from either continuous replication or reactivation from latency (134).

Herpesviruses have co-evolved with their hosts and exhibit restricted host range (80, 96). Following entry into host cells, viral particles enter the nucleus where viral replication and virion assembly occur. Viral genes are expressed in a regulated fashion; immediate-early genes require no viral protein synthesis for expression; early genes are independent of viral DNA synthesis, but require new viral protein synthesis; and late genes require viral DNA synthesis and are often structural. Release of infectious progeny virus causes destruction of the infected cell (96, 97). Most genes are expressed during lytic replication whereas only a restricted set of transcripts is produced during latent infection (44).
The gammaherpesviruses are characterized by their limited host range and ability to establish latency in immune cells (96). Epstein-Barr Virus (EBV, HHV-4) and Kaposi's sarcoma-associated herpesviruses (KSHV, HHV-8) are the human gammaherpesviruses. EBV and KSHV are oncogenic and are associated with lymphomas, carcinomas and Kaposi's sarcoma in infected individuals. Although many cases of EBV infection are clinically asymptomatic, acute EBV infection is one of the causes of infectious mononucleosis (23, 114). EBV is also associated with Burkitt's lymphoma, post-transplant lymphoproliferative disease, nasopharyngeal carcinoma, Hodgkin's lymphoma, and gastric adenocarcinoma (50, 52, 55). KSHV was first identified using representational difference analysis of Kaposi's sarcoma lesions from patients with AIDS in comparison to normal tissue (16) and is strongly implicated as required for the development of Kaposi's sarcoma (44, 83). Interestingly, a role for lytic replication in Kaposi's sarcoma tumorigenesis has been suggested (49). KSHV is also linked to HIV-associated multicentric Castleman's disease and primary effusion lymphoma (8, 87). The association of KSHV and EBV with a variety of cancers in immunosuppressed patients suggests an important role for immune regulation in control of pathologies caused by these viral infections.

1.1.1 Murine Gammaherpesvirus 68 Model System

Murine gammaherpesvirus 68 (MHV68, γHV-68, MuHV4) is genetically related to the human gammaherpesviruses, KSHV and EBV (36, 37, 116, 135). Infection with KSHV and EBV is species specific and de novo infection in vitro has a proclivity for latent infection, limiting the ability of investigators to study the dynamics of productive infection (44, 66). MHV68 provides a tractable small animal model system to investigate features of gammaherpesvirus pathogenesis in vivo and productive primary infection in vitro (36, 37, 116, 135). During productive infection
of fibroblasts and replication during acute infection of mice, all known ORFs are detectably expressed (62, 78, 103). During latent infection, a restricted set of transcripts is expressed, including mK3, ORFs 73 and 74 (latency-associated nuclear antigen, LANA, and viral G-protein coupled receptor homolog, GPCR, respectively) and many of the MHV68-specific genes, such as M2, M3, M8, M9 and M11 (viral Bcl-2) (78, 103, 136).

1.1.2 *In vivo* infection and pathogenesis

MHV68 infects multiple cell types *in vivo* during acute infection and establishes latent infection in macrophages, dendritic cells and B cells (38-41, 67, 124, 125, 140). It causes lymphoproliferative disease, splenic and pulmonary fibrosis and large-vessel vasculitis in immunocompromised mice (33-35, 74, 126, 127, 129, 139). The development of these pathologies in mice requires the presence of persistent virus (25, 86).

Numerous viral genes have been identified as playing a role in establishment of, or reactivation from, latency, including viral-cyclin (ORF 72), viral-GPCR (ORF 74), viral-bcl-2 (M11), and the LANA homolog (ORF 73) (28, 43, 45, 84, 85, 92, 133). While most viral genes that have been examined to date are important for promoting latency or the development of disease, M1 is reported to suppress reactivation and the development of atypical lymphoid hyperplasia (19, 127).

1.1.3 Pervasive Transcription of MHV68

*In vitro*, MHV68 can infect multiple cell types including fibroblasts, macrophages, and endothelial cells. Lytic replication is typically studied *in vitro* in fibroblast cell lines which are highly permissive for infection. The complete sequence of the 119,450 base pair genome of
MHV68 was originally reported to contain at least 80 open reading frames (ORFs) defined as methionine-initiated ORFs of at least 100 amino acids [135]. Such an annotation, however, fails to identify noncoding RNAs, spliced exons and small protein-coding genes [62]. Indeed, subsequent studies have identified splice forms of several ORFs [2, 29, 48]. Moreover, a tiled array designed in our laboratory reveals that many regions of the MHV68 genome that do not correspond to annotated ORFs are actively transcribed during productive infection in fibroblasts at 18 hours post-infection [62]. We identified over 73,000 nucleotides of unexpected transcription, representing a 31% increase in the RNA coding potential of this virus. The presence of pervasive transcription of MHV68 has also been confirmed by others [17]. We termed regions of unexpected transcription as Expressed Genomic Regions (EGRs) when RNA expression is: (1) at least 700 nucleotides 5′or 3′of an annotated ORF or, if not associated with an annotated ORF, at least 700 nucleotides in length, (2) greater than one standard deviation above the mean signal from mock-infected cells, and (3) detectable at each nucleotide in at least two of three independent experiments. Using these criteria, 30 EGRs are expressed in fibroblast cells. MHV68 EGRs range in size from about 700 nucleotides to nearly 10,000 nucleotides and can encode multiple transcripts [62].

This same phenomenon of extensive antisense transcription has recently been described in the human gammaherpesvirus, KSHV, and has also been recognized in the betaherpesvirus, human cytomegalovirus (HCMV) [14, 147]. The functional significance of this pervasive transcription is one area of focus in this dissertation. By comparing ORFs greater than 20 amino acids in all possible reading frames of MHV68, KSHV, EBV and *Herpesvirus saimiri*, Johnson et al. identified gammaherpesvirus conserved putative peptide sequences in eight of 30 (27%) EGRs. These conserved regions are small (less than 60 amino acids on average) and are limited
to regions of EGRs where the protein-coding ORF on the opposite strand is also highly conserved, suggesting that the conservation noted may be due to sequence constraints of the opposite strand ORFs (62). In addition, unannotated ORFs greater than 20 amino acids were compared to the GenBank database. However, none of these sequences are conserved with sequences for annotated peptides or proteins currently available in GenBank (Dr. L. Steven Johnson, personal communication). Therefore, we hypothesized that few of the EGRs encode proteins and propose that a substantial portion of these newly identified transcripts may be noncoding RNAs (ncRNAs). Further characterization and functional analyses of EGR-encoded transcripts is presented in Chapter 2. Examples of known ncRNA function from other herpesviruses will be presented at the end of this Chapter.

1.1.4 ORF 50: The Molecular Latent-to-Lytic Switch

MHV68 open reading frame 50 (ORF 50), an immediate-early gene that encodes the replication and transcription activator (RTA), is the latent-to-lytic switch gene, a feature that is shared by all gammaherpesvirus RTAs (118). ORF 50 is both necessary and sufficient to induce lytic reactivation in latently infected cells (76, 93, 94, 118, 143). ORF 50 induces lytic reactivation through the transactivation of other viral promoters (26, 93, 118, 143). Since gammaherpesvirus transmission and disease often occur following reactivation from latency, understanding ORF 50 regulation is important to understanding disease pathogenesis (66). When ORF 50 is constitutively expressed (as driven by the MCMV IE1 promoter), the virus is unable to establish latency (79), suggesting that elements within the ORF 50 promoter may be important in regulating the transition between latent and lytic infection. Several negative regulators of KSHV ORF 50 have been previously identified, including the transcription factors, NF-κB and
RBP-Jκ and the viral gene, LANA (ORF73) (9, 61, 69, 70). NF-κB is also known to inhibit MHV68 lytic replication (9). We determined that interferon-gamma (IFNγ) suppresses ORF 50 transcript expression in bone marrow derived macrophages (47), suggesting that IFNγ may regulate the transition between lytic replication and latency by suppressing RTA expression (Chapter 3).

1.1.5 Role of Interferons in MHV68 Infection

Interferons (IFNs) were first described for their ability to interfere with viral infection (58). In MHV68, both type I and type II IFNs play important, but distinct, roles in controlling viral infection. Type I IFNs are critically important to controlling replication during acute infection; mice unresponsive to IFNα/β succumb to dose-dependent mortality after infection (32). Although type II IFN is typically unnecessary for control of replication during the early stages of MHV68 infection, a recent study identified a role for IFNγ responsiveness in the BALB/c background to protect mice from acute bronchopneumonia (73, 112, 139).

Both type I and type II IFNs regulate reactivation of MHV68 from latency; ex vivo reactivation frequency is increased in IFNαR-, IFNγ-, and IFNγR-deficient mice compared to wild-type immunosufficient control mice (3, 32, 130). Furthermore, depletion of IFNγ in vivo increases the frequency of reactivation while ex vivo treatment of latently infected cells with IFNγ potently suppresses the frequency of reactivation (3, 119). Of note, IFNα, but not IFNγ, controls the presence of persistent virus and the development of a large-vessel arteritis (3, 130, 139), suggesting that type I and type II IFNs may act to control MHV68 via distinct mechanisms.

There are several studies that have investigated the role of IFNs and their effectors on specific MHV68 genes. Recent studies, including our own, have demonstrated a role for both
type I and type II IFNs in controlling ORF 50 transcription ((47, 142) and Chapter 3). ORF 36 is a viral kinase that initiates the DNA damage response in infected cells by promoting phosphorylation of H2AX and inhibits the type I interferon response by preventing IRF-3-mediated recruitment of RNA polymerase II to the IFNβ promoter (56, 128). ORF 54, a viral dUTPase, inhibits IFNβ signaling by promoting degradation of its receptor (72). M2 also disrupts IFN signaling by downregulating STAT2 expression (75). Interestingly, IFNβ-induced IRF2 plays a role in regulating M2 gene expression in B cells specifically; mice infected with a virus in which the IFN-stimulated response element (ISRE) is mutated in the M2 promoter have elevated titers in the lung and spleen during acute infection, increased reactivation from latency and greater mortality in a model of acute viral pneumonia (77). It is clear that IFNs play a critical role in control of MHV68 infection and that the virus has evolved mechanisms to manipulate this system.

1.1.6 Gammaherpesvirus Latency in Pathology and Protection from Disease

Latent viral infection can profoundly alter the host response to a range of infectious and chronic inflammatory diseases. Previous studies identified a role for latent MHV68 or the betaherpesvirus, murine cytomegalovirus, in protecting mice from lethal bacterial challenges with *Listeria monocytogenes* or *Yersinia pestis* (4, 145). Similarly, latent MHV68 reduces the severity of some parasitic and viral infections, such as *Plasmodium yoelii*, influenza A virus, and mouse adenovirus type 1, but not West Nile virus (4, 53, 90, 108). Latent MHV68 also protects mice from a lethal challenge with lymphoma cells (141). When mice are lytically infected with MHV68 and *P. yoelii*, they develop severe lung inflammation (53), suggesting that the timing and stage of viral life cycle are critically important in the host response to coinfection.
Resistance to many of these challenges is associated with activation of macrophages and natural killer cells (4, 108, 141), suggesting that activation of the innate immune system is critically important in mediating many of these protective responses. Moreover, both type I and type II IFNs as well as tumor necrosis factor alpha (TNFα) are elevated in mice latently infected with MHV68 (4, 77, 113), suggesting a central role for cytokines in the orchestrating cellular responses in the setting of latent MHV68 infection.

While latent MHV68 has been found to play an important role in protection from a range of infectious challenges, it has also been shown to exacerbate certain chronic inflammatory diseases, including experimental autoimmune encephalitis (EAE), a model for multiple sclerosis, and a model of inflammatory bowel disease (12, 89, 95). EBV has long been suggested as a possible trigger for multiple sclerosis (reviewed in (91)). In contrast, chronic MHV68 infection is associated with decreased autoantibody production and prevents kidney disease in lupus-prone mice (71). It is evident that latent MHV68 infection can have diverse effects in shaping the immune response to a range of infections and autoimmune diseases. Together these studies formed the basis for our investigation of the effect of latent gammaherpesvirus infection on host gene expression in several different organs (Chapter 4).

1.2 Functions of the Products of Pervasive Transcription

Transcription outside of known ORFs, or pervasive transcription, is increasingly recognized in a wide range of species (1, 5, 11, 18, 27, 51, 54, 63, 64, 115, 144, 147). Current estimates are that between 83.7-85.2% of the human genome is transcribed, although many transcripts are only produced in specific cell types, suggesting that the cellular context is important (10, 31, 51).
These regions of pervasive transcription are thought to predominately represent noncoding RNAs (ncRNAs). Proposed functions of ncRNAs include epigenetic modification, regulation of transcription (for example, by acting as co-factors to regulate transcription factors) as well as regulation of enzymatic pathways and modulation of post-transcriptional processing by regulation of mRNA splicing, transport, or translation (57, 65, 82, 131, 146).

Some of the best studied long ncRNAs have been found to play roles in regulating gene transcription (reviewed in (101)). For example, XIST, which regulates X chromosome inactivation, and HOTAIR both repress gene transcription via interaction with Polycomb Repressive Complex 2 (PRC2) (102, 148). In contrast, other ncRNAs, like HOTTIP, have been found to promote transcription by binding to proteins, such as WDR5, to catalyze trimethylation of histone H3 lysine 4 (H3K4me3) (137). Interestingly, another long ncRNA, NeST, also binds to WDR5, promoting H3K4me3, and thereby inducing transcription of IFNγ (46).

1.2.1 Functions of Noncoding RNAs in Herpesviruses

Noncoding RNAs have been identified in several herpesviruses and it is clear that herpesvirus ncRNAs can play important and diverse roles in viral pathogenesis. Viral ncRNAs are currently divided into microRNAs (miRNAs) and "everything else" (121). Viral miRNAs can regulate gene expression and promote viral evasion of host defenses (121). There is a known cluster of viral miRNAs in MHV68 at the left end of the viral genome (6, 30, 98, 99). However, since the focus in this dissertation will be on long ncRNA, I will present examples of herpesvirus long ncRNAs which have been found to be expressed during either lytic or latent infection. HCMV and KSHV are both known to express ncRNAs during lytic infection while EBV,
Herpesvirus saimiri, and Herpes Simplex virus (HSV) are all known to produce ncRNAs during latent infection.

1.2.2 Herpesvirus ncRNAs expressed during lytic infection

During lytic infection, HCMV encodes a 2.7-kilobase noncoding RNA, β2.7, that can protect cells from apoptotic cell death (100, 121). β2.7 interacts with mitochondrial enzyme complex I, inhibiting its relocalization following apoptotic stimulation and promoting continued generation of adenosine triphosphate which is thought to be important for supporting viral replication in the face of signals for apoptotic cell death (100, 121).

Extensive antisense transcription in KSHV has recently been described and several regions have been confirmed to encode long transcripts by northern blot, one of which is not associated with polysomes, suggesting that it does not encode a protein or peptides (14). The best-studied KSHV ncRNA is the polyadenylated nuclear (PAN) RNA which is a 1-kilobase nuclear ncRNA produced during lytic infection (21, 122, 149). The function of PAN RNA is currently unknown, but it is the focus of several recent reports. PAN RNA is activated by ORF 50 and is the most abundant lytic transcript produced by the virus, so it is likely to play an important role in the viral life cycle or viral pathogenesis (15, 21, 117, 123). Recent studies to investigate the role of PAN RNA in viral replication have used antisense oligonucleotide (ASO) knockdown or a viral deletion mutant. Using ASO-mediated knockdown of PAN RNA, late, but not immediate-early or early, genes were decreased, suggesting a role for PAN RNA in viral DNA replication (7). In contrast, genes of all kinetic classes, including ORF 50, were reduced in a PAN RNA deletion virus (105). The discrepancy between these two studies may be due to incomplete knockdown of PAN RNA by ASOs or disruption of an antisense transcript in the
deletion virus. Recent studies have further identified an association of PAN RNA with the ORF 50 promoter as well as the demethylases, JMJD3 and UTX, and methyltransferase, MLL2, and SUZ12 and EZH2, protein components of PRC2 (105, 106). It is proposed that PAN RNA can activate as well as repress KSHV gene expression by directing epigenetic modification of the ORF 50 promoter (105, 106). Interestingly, using next-generation sequencing in combination with chromatin isolation by RNA purification (ChIRP-Seq), investigators found that PAN RNA interacts with a number of viral and host genes, including host genes involved in the immune response (106), suggesting that this long ncRNA plays a crucial role in orchestrating viral gene expression and modulating the immune response to infection.

1.2.3 Herpesvirus ncRNAs expressed during latent infection

Noncoding RNAs known to be produced during latent infection include EBV-encoded RNAs (EBERs), *Herpesvirus saimiri* U RNAs (HSURs), and HSV latency associated transcript (LAT). EBV encodes two short nonpolyadenylated RNAs (EBER1 and EBER2) that are expressed at high levels during latency and reactivation of lytic replication (66, 111). These RNAs are known to prevent IFN-α-induced apoptosis (88, 107). One proposed, but controversial, mechanism is that EBERs prevent apoptosis by binding to PKR and preventing its phosphorylation (88, 111). However, these studies used transient transfection of EBERs or *in vitro* assays to show binding of EBERs to PKR (81, 88). In infected cells, EBERs are predominately detected in the nucleus and are reported not to inhibit PKR phosphorylation (42, 66, 107, 121). Other activities attributed to EBERs include induction of interleukin-10 and type 1 IFNs via RIG-I-mediated signaling and induction of insulin-like growth factor-1 (59, 60, 68, 109-111).
*Herpesvirus saimiri*, an oncogenic gammaherpesvirus of primates that causes malignant transformation of T cells, encodes seven small noncoding RNAs (HSURs 1-7) (21). Expression of HSURs 1 and 2 is correlated with expression of host mRNAs associated with T-cell receptor (TCR) signaling (22). HSURs 1 and 2 induce upregulation of TCR signaling chains as well as signaling molecules, such as DAP10, suggesting that these viral ncRNAs contribute to T cell activation (22). HSURs 1 and 2 can bind to host miRNAs and HSUR1 expression is sufficient to downregulate one of these, miR-27 (13). Downregulation of miR-27 is associated with increased levels of its target protein, FOXO1, suggesting that a viral noncoding RNA can modulate host protein expression by regulating expression of a host miRNA (13).

HSV, an alphaherpesvirus, encodes a latency-associated transcript (LAT) of about 9 kb which has been the subject of extensive study (reviewed in (104)). LAT is thought to prevent neurons from cell death, although its mechanism of action is unclear. Two proposed mechanisms are silencing of lytic gene expression and inhibition of apoptosis (104). A recent study has suggested that two small ncRNAs encoded by LAT inhibit apoptosis (24). The contribution of LAT to establishment of and reactivation from latency remains controversial (104). However, it has been shown that small ncRNAs, including miRNAs, can arise from LAT and regulate the transcriptional activator, *ICP0*, which supports viral replication and may promote reactivation (120, 132). Recent studies demonstrated that small ncRNAs encoded by LAT regulate IFNβ activity and have implicated LAT in gene silencing (20, 24, 138).
1.3 References


42. Fok, V., K. Friend, and J. A. Steitz. 2006. Epstein-Barr virus noncoding RNAs are confined to the nucleus, whereas their partner, the human La protein, undergoes nucleocytoplasmic shuttling. J Cell Biol 173:319-325.


Xu, Y., and D. Ganem. 2010. Making sense of antisense: seemingly noncoding RNAs antisense to the master regulator of Kaposi's sarcoma-associated herpesvirus lytic


Chapter 2:

Pervasive transcription of a herpesvirus genome generates functionally important RNAs

This chapter is from:


*denotes equal contribution.
2.1 Abstract

Pervasive transcription is observed in a wide range of organisms, including humans, mice and viruses, but the functional significance of these transcripts remains uncertain. Current genetic approaches are often limited by their emphasis on protein coding open reading frames (ORFs). We previously identified extensive pervasive transcription from the murine gammaherpesvirus (MHV68) genome outside of known ORFs and antisense to known genes (termed expressed genomic regions, EGRs). Similar antisense transcripts have been identified in many other herpesviruses, including Kaposi's sarcoma-associated herpesvirus and human and murine cytomegalovirus. Despite their prevalence, whether these RNAs have any functional importance in the viral life cycle is unknown, and one interpretation is that these are merely "noise" generated by functionally unimportant transcriptional events. To determine whether pervasive transcription of a herpesvirus genome generates RNA molecules that are functionally important, we used a strand-specific functional approach to target transcripts from thirteen EGRs in MHV68. We found that targeting transcripts from six EGRs reduced viral protein expression, proving that pervasive transcription can generate functionally important RNAs. We characterized transcripts emanating from EGRs 26 and 27 in detail using several methods, including RNA sequencing, to finely map several novel polyadenylated transcripts that were enriched in the nucleus of infected cells. These data provide the first evidence of the functional importance of regions of pervasive transcription emanating from MHV68 EGRs and demonstrate that transcripts from genome regions outside of known ORFs in herpesviruses can play critical roles in the viral life cycle.
2.2 Introduction

Gammaherpesviruses are oncogenic herpesviruses that undergo productive replication and can establish latent infection in their hosts. Human gammaherpesviruses, Epstein-Barr Virus (EBV, HHV-4) and Kaposi’s sarcoma-associated herpesvirus (KSHV, HHV-8), are associated with malignancies, including Burkitt’s and primary effusion lymphomas, nasopharyngeal carcinoma, and Kaposi’s sarcoma. Murine gammaherpesvirus 68 (MHV68, γHV68, MuHV4) is genetically related to EBV and KSHV and causes lymphomas and lymphoproliferative disease in immunocompromised mice (13, 34, 51, 52, 58), providing a tractable model system in which to study productive infection in vitro and viral pathogenesis in vivo.

MHV68 and other herpesviruses can serve not only as models for viral infection and pathogenesis, but also as systems in which to unravel the complexity of pervasive transcription and to probe the function of its products. Using high density tiled arrays and RNA sequencing, we and others have shown that widespread transcription occurs outside of annotated open reading frames (ORFs) during lytic MHV68 infection, generating regions of transcription termed expressed genomic regions (EGRs) (8, 25). We termed these EGRs rather than genes because the signal in a tiled array analysis may represent multiple different transcripts, and because they were not initially investigated for functional importance. These EGRs contain no ORFs with significant homology to known proteins and, as such, may encode noncoding RNAs (ncRNAs), spliced transcripts, and/or novel small polypeptides (25). While there are several well-studied examples of herpesvirus ncRNAs expressed during lytic or latent infection (reviewed in (50)), new sequencing and array technologies reveal a substantially more complex transcriptional landscape for both betaherpesviruses and gammaherpesviruses than was previously appreciated (7, 8, 10, 14, 25, 26, 36, 62, 64). Intergenic and antisense transcription has also been widely
described in a range of other organisms including humans and mice (2, 11, 20, 28). However, the function of these viral transcripts remains an open question. The implications of these findings for genomic mutagenesis studies targeting ORFs within herpesvirus genomes are incompletely understood. Importantly, while initial findings of pervasive transcription of the MHV68 genome (25) have been independently confirmed (8), whether these RNAs have any function is unknown.

One reasonable interpretation is that the products of pervasive transcription are irrelevant to the viral life cycle and might merely be the result of failed termination of functionally important, often protein-coding, transcripts. Studies relying on disruption of the viral genome in ways that alter transcripts from both genomic strands cannot resolve this question.

In this study, we used a flow cytometry-based screen combined with strand-specific knockdown of candidate RNAs to assess the importance of RNAs encoded within EGRs. We found that targeting several EGR-encoded transcripts altered the expression of other viral genes and selected two adjacent regions (EGRs 26 and 27) for detailed analysis. We characterized the transcript architecture, localization, and effects on other viral genes for RNAs emanating from EGRs 26 and 27. Herein, we report that EGR 26 and 27 transcripts were enriched in the nucleus of infected cells and that targeting EGR 27 transcripts altered expression of multiple viral genes. To our knowledge, this is the first proof of functional significance of novel viral transcripts identified using a transcriptome-based approach.
2.3 Results

2.3.1 Establishment of a system to target MHV68 transcripts: antisense targeting of a known essential gene alters late viral protein expression.

Since many EGRs are antisense to other EGRs or known ORFs, it was necessary to design a strategy to determine the functionality of transcripts from a single strand of the viral genome. To disrupt viral transcripts in a strand-specific manner, we designed single-stranded antisense oligonucleotide (ASO) gapmer probes to known MHV68 viral genes. Gapmer ASOs, which are reported to disrupt gene expression by RNase-H mediated degradation of their target transcript (60), have been previously used to target herpesvirus ncRNAs (3, 6).

An ASO to the single-stranded DNA binding protein, ORF 6 (53), significantly reduced transcript (Fig. 2.1A) and protein levels (Fig. 2.1B) of its target. An ASO to a non-cellular, non-viral transcript (GFP) did not alter ORF 6 transcript or protein abundance (Fig. 2.1A-B) and served as a negative control. Additionally, the ORF 6 ASO reduced the abundance of the ORF 4-ORF 6 bicistronic transcript (Fig. 2.1A, (27)). We observed a smaller (~1kb) band in cells transfected with the ASO to ORF 6, which may represent a stable product of the degraded ORF 6 transcript (Fig. 2.1A). These data demonstrated that targeting a viral gene with an ASO decreased both transcript and protein levels of its target, confirming this as an effective strategy for knocking down specific transcripts in MHV68 infected cells.

We next validated a flow cytometry-based assay to rapidly assess the effect of knockdown of transcripts on viral protein expression. We selected the late viral gene ORF 4, the viral complement regulatory protein (v-RCA), expressed on the surface of infected cells (27), to measure the effect of viral transcript knockdown on a late stage of the viral life cycle. An ASO to the essential gene, ORF 6 (53), significantly inhibited the surface expression of ORF 4 (Fig. 32).
2.1C). An ASO designed to be complementary to the nonessential gene, M3 (56), did not reduce ORF 4 surface expression (Fig. 2.1C). Since ORF 4 is transcribed as a bicistronic transcript with ORF 6, it is possible that the reduction in ORF 4 surface expression observed was due to direct knockdown of the transcript that encodes ORF 4 by the ASO designed to target ORF 6. To confirm that knocking down ORF 6 altered expression of other late proteins, we evaluated protein expression of ORF 26 and M9 (ORF 65) (4, 12) and confirmed that knocking down ORF 6 decreased expression of multiple late proteins (Fig. 2.1D). These data demonstrated that disrupting the expression of an essential gene transcript was detectable by a change in ORF 4 surface expression and indicated that this method could be used as a screening tool for the function of EGR transcripts.

2.3.2 Targeting EGR transcripts alters surface expression of the late viral protein encoded by ORF 4.

To determine whether transcripts encoded within EGRs played a role in productive infection, we designed ASOs to regions within 13 EGRs (EGR nomenclature as in (25)). We selected 12 EGRs that might be important for viral replication (EGRs 9, 10, 11, 15, 16, 18, 20, 21, 23, 25, 26, and 27) and one EGR that might be dispensable for viral replication (EGR 30) based on results from a genome-wide transposon-based mutagenesis screen of the MHV68 genome (48). Since transposon mutagenesis disrupts genetic elements on both DNA strands, some of the phenotypes previously attributed to transposon insertions in known ORFs could also be explained by a function of EGR-encoded transcripts from the opposite strand. Cells were transfected with ASOs to regions of the 13 selected EGRs or known viral genes, ORF 6 or M3, infected with MHV68, and then analyzed for ORF 4 surface expression by flow cytometry (see
Table 2.1 for ASO sequences and targeted genomic coordinates). ASO targeting of EGRs 9, 16, 18, 23, 26 and 27 showed statistically significant reductions in ORF 4 surface expression (Fig. 2.2). For EGRs 9, 18, and 23, which were targeted with multiple ASOs, not all ASOs altered ORF 4 expression, suggesting that these ASOs may differ in knockdown efficiency or may target distinct transcripts encoded by the EGR. In fact, we have found that there is extensive splicing within the MHV68 transcriptome, including within EGR 23, which may account for the differences observed (Johnson, L.S., Canny, S.P., Virgin, H.W., unpublished data). Cells transfected with ASOs targeting EGRs 11 and 20 had increased ORF 4 surface expression (Fig. 2.2), which might suggest a role for transcripts emanating from these EGRs in suppressing viral replication.

Taken together, these results show that several EGR-encoded transcripts were important for ORF 4 protein expression. ASOs to the adjacent EGRs 26 and 27 (EGR 26c and EGR 27b in Table 2.1) on the negative strand of the viral genome had the largest effect on cell surface expression of ORF 4, comparable to the effect of targeting the essential gene, ORF 6 (Fig. 2.2). For this reason, we chose to further characterize EGR 26-27 encoded transcript(s) and to assess the effect of targeting RNAs from this region on viral gene expression.

2.3.3 Mapping 5' and 3' ends of EGR 26 and EGR 27 transcripts.

To define the transcript(s) targeted by ASOs to EGRs 26 and 27, we used a combination of RNA sequencing, 5' and 3' RACE, and northern blot analysis. First, we confirmed active RNA expression within EGRs 26 and 27 in lytically infected fibroblasts by RNA sequencing and identified a transcriptional signal similar to the signal that we had previously detected using tiled array technology (Fig. 2.8 and (25)). Spearman correlations calculated between our previous
tiled array data and RNA sequencing datasets from 18 hours post infection (hpi) yielded correlation coefficients ranging from 0.78 to 0.85.

To determine the size and relative position of transcripts emanating from EGRs 26 and 27, we used northern blot analysis. We identified a ~12 kb transcript (hereafter EGR 26 transcript A) that overlapped EGRs 26 and 27 and ORFs 65-67; an ~8 kb transcript (hereafter EGR 27 transcript A) that overlapped EGR 27 and ORFs 65-67; and a ~3 kb transcript (hereafter EGR 27 transcript B) that overlapped EGR 27 (Fig. 2.3). We also identified a transcript > 12 kb that was detected by probe 1 which we termed EGR 26 transcript B, but did not map further. Using a probe to genomic coordinates 101170-101208 (probe 12, Fig. 2.3A), we detected a ~ 8 kb transcript and a ~3 kb transcript on the opposite side of the 100-bp internal repeat (Fig. 2.3B), suggesting that both EGR 27 transcripts A and B overlap the 100-bp internal repeat. We tested two gene-specific primers in the PCR step of 5' RACE for EGR 27 and obtained comparable results. We identified 9/9 colonies within one nucleotide of 101225. We were unable to detect transcripts using probe 13, suggesting that ~101225 was the 5' end for both EGR 27 transcripts A and B (Fig. 2.3B). Probe 8, but not probes 9 or 10, detected a ~ 12 kb transcript, comparable in size to EGR 26 transcript A (Fig. 2.3A-B). 5' RACE analysis identified 6/6 colonies within four nucleotides of 97790 (1 colony at 97786, 1 colony at 97789, 3 colonies at 97790, and 1 colony at 97791), suggesting that the 5' end of EGR26 transcript A was ~97790.

EGRs 26 and 27 encode polyadenylated transcripts as shown by northern blot analysis of polyadenylated RNA (data not shown and Fig. 2.3) and confirmed by RNA sequencing of polyadenylated RNA (Fig. 2.8). To identify 3' ends of EGR 26 and 27 transcripts, we mined RNA sequencing data from four experiments, two from 18 hpi and two from 6 hpi, using a custom analysis pipeline. By mining RNA sequencing data for polyadenylated reads, we
identified 3' ends within the EGRs 26 and 27 at 85207, 93842, and 98510 and within the intervening ORFs at 95322, 95758, and 95895 (Fig. 2.8 and Table 2.3). It is worth noting that these 3' ends at 95322, 95758, and 95895 were not attributable to transcripts detected by northern blot, suggesting that there may be additional less abundant transcripts within this region. This is important because it indicates that northern blot analysis may not be sensitive enough to detect all transcripts encoded within EGRs that might be targeted by specific ASOs. We performed 3' RACE analysis to confirm 3' ends at 98510 (6/6 colonies) and within 3 nucleotides of 93842 (5/5 colonies). In addition, our RNA sequencing approach confirmed six previously known polyadenylation sites as well as 53 novel sites downstream of known protein-coding and EGR-encoded transcripts; the majority of these were detected at both 6 and 18 hpi (Table 2.3). Together, these data identified three transcripts that overlap the EGR 26-27 region and highlight the profound transcriptional complexity across the MHV68 genome (Fig. 2.3 and (8, 25, 32)).

2.3.4 **EGR 26 and EGR 27 transcripts are enriched in the nucleus of infected cells.**

To determine if transcripts that overlap EGRs 26 and 27 were retained in the nucleus or exported to the cytoplasm of infected cells, we separated nuclear and cytoplasmic fractions, extracted RNA, and probed for EGR 26 transcript A using probe 2 and EGR 27 transcripts A and B using probe 11. We confirmed adequate separation of nuclei from cytoplasm by western and northern blot analyses of known cytoplasmic and nuclear proteins and RNAs with only modest contamination of the cytoplasmic fraction with nuclear protein and RNA (Fig. 2.4). We found that the EGR 26 transcript A, EGR 27 transcript A, and EGR 27 transcript B were enriched in the nuclear fraction (Fig. 2.4B). Nuclear enrichment of EGR 26 and EGR 27 transcripts was comparable to the well-characterized mouse nuclear ncRNA, nuclear enriched abundant
transcript 2 (Neat2, also known as Malat1, (23), Fig. 2.4B). Analysis of representative viral (ORF 6) and host (actin) protein-coding transcripts demonstrated that these translated transcripts accumulated in the cytoplasm (Fig. 2.4B). The nuclear distribution of EGR 26 transcript A, EGR 27 transcript A, and EGR 27 transcript B was consistent with a known nuclear ncRNA and not with protein-coding transcripts, suggesting that these EGR transcripts may be nuclear ncRNA.

2.3.5 Targeting EGR 27 transcripts alters surface expression of ORF 4 protein.

To determine the effect of knockdown of RNAs encoded within EGR 27 on ORF 4 surface expression, we targeted this region with four ASOs, two of which we predicted would target both EGR 27 transcripts detectable by northern blot and two of which we predicted would specifically target EGR 27 transcript A. Because EGR 27 transcript A overlaps the entire length of EGR 27 transcript B, it was not possible to specifically target the smaller transcript. EGR 27a and EGR 27b ASOs reduced the abundance of both EGR 27 transcripts as predicted (Fig. 2.5B-C). Surprisingly EGR 27d ASO also significantly reduced EGR 27 transcript B although it was not predicted to target this transcript (Fig. 2.5C), raising the possibility that an additional transcript might be derived from this region. In support of this hypothesis, we identified additional 3’ ends ORFs 66-67 at 95322, 95758, and 95895 that have not been ascribed to specific transcripts (Table 2.3) that were not attributed to RNAs detected by northern blot (see above), suggesting the presence of additional as-yet-unmapped transcripts within this region. As predicted, EGR 27c ASO reduced EGR 27 transcript A, but not EGR 27 transcript B (Fig. 2.5B-C). All EGR 27 ASOs reduced ORF 4 surface expression (Fig. 2.5E). While we cannot attribute the effect on ORF 4 surface expression to a specific transcript because the transcripts overlap,
these data confirm data in our initial screen (Fig. 2.2) by showing that RNAs encoded in this region of the genome are functionally important for viral gene expression.

We also targeted EGR 26 with six ASOs, five predicted to target EGR 26 transcript A (EGR 26a-e) and one directed 3' of EGR 26 transcript A (EGR 26f). EGR 26b-d ASOs significantly reduced EGR 26 transcript A levels (Fig. 2.5D). Although EGR 26d ASO reduced EGR 26 transcript A, it did not significantly reduce ORF 4 surface expression (Fig. 2.5E), suggesting that the effect on ORF 4 surface expression by other EGR 26 ASOs may be due to other smaller transcripts in this region (see Fig. 2.3B). The effect of EGR 26f ASO, which did not target EGR 26 transcript A, but likely targeted other RNAs derived from EGR 26, such as EGR 26 transcript B, further supports that RNAs derived from EGR 26 are important for viral protein expression. These data suggest that similar to those transcripts emanating from EGR 27, RNAs from EGR 26 are also functionally relevant and that the functional significance of EGR transcription is not restricted to a single region of the genome.

To determine whether the observed changes in ORF 4 protein expression were associated with a detectable change in viral replication, we analyzed viral titers at 24 hpi in ASO-transfected cells. We detected a modest reduction in viral titer in cells transfected with EGR 27a, EGR 27b, EGR 27d, and EGR 26c ASOs, comparable to the reduction in viral titer in cells transfected with an ASO to the essential gene, ORF 6 (Fig. 2.5F). Although the effects were modest, likely due to partial knockdown of the target message, it is notable that the EGR ASOs significantly reduced viral titers.
2.3.6 Targeting of EGR 27 transcripts alters expression of multiple viral genes.

Having shown that RNAs derived from EGRs can be functionally important as measured by the effects on a viral late protein and viral replication, we next assessed the extent of ASO effects on different aspects of viral transcription and protein expression. To confirm that EGR 27 ASOs reduced late gene expression, we assessed the effect of EGR 27 ASOs on ORF 26 and M9 (ORF 65) protein expression by western blot analysis and spliced ORF 29 transcripts by qRT-PCR (35). EGR 27a-d ASOs reduced ORF 26 and M9 protein expression and ORF 29 transcript expression (Fig. 2.6A-B). These data show that targeting EGR 27 transcripts broadly altered the expression of multiple late genes and did not exert effects restricted to surface expression of the late protein encoded by ORF 4.

Next, we tested the effect of EGR 27 ASOs on the expression of the early gene, ORF 6 (59), or the immediate-early gene, ORF 50, also known as the replication and transcription activator (RTA) (37). We found that EGR 27a, 27b and 27d ASOs, but not EGR 27c ASO, significantly reduced ORF 6 transcript levels (Fig. 2.6C). However, all EGR 27 ASOs reduced ORF 6 protein levels (Fig. 2.6D). Furthermore, we found that EGR 27a and 27d ASOs, but not EGR 27b or 27c ASOs, significantly reduced the abundance of spliced ORF 50 transcripts (Fig. 2.6E). In summary, these data suggest that a transcript or transcripts targeted by EGR 27a and/or 27d ASOs acts early in the viral lifecycle, altering gene expression of specific viral genes from each kinetic class. Furthermore, a transcript(s) targeted by EGR 27c ASO may act later in the viral lifecycle.
2.3.7 **EGR 26c ASO decreases specific genes of all kinetic classes.**

We evaluated EGR 26c ASO, the ASO used in our initial screen (Fig. 2.2), for its effect on the viral lifecycle. To confirm that EGR 26c ASO altered the expression of multiple late genes, we analyzed its effect on the expression of early-late (M3) and late (M9) genes (46, 57, 59) using northern and western blot analyses. We found that EGR 26c ASO reduced M9 and M3 protein levels (Fig. 2.7A and 2.7C) as well as M9 and M3 transcript levels (Fig. 2.7B and data not shown). Taken together with the effect of EGR 26c ASO on ORF 4 cell surface expression, these data demonstrate that a transcript targeted by EGR 26c ASO was important for expression of several early-late and late genes and suggested that it acted upstream of their expression. Interestingly, EGR 26c ASO did not alter the level of EGR 27 transcript A, detected by the northern blot probe 7 (data not shown), indicating that EGR 26c ASO altered the levels of specific viral RNAs rather than overall transcription of the viral genome.

Next, we tested the effect of EGR 26 ASOs on the expression of ORF 6 and found that EGR 26a-c ASOs reduced ORF 6 transcript levels, while EGR 26d-f ASOs did not (Fig. 2.7D), suggesting that a transcript targeted by EGR 26a-c ASOs acts to reduce viral gene expression. EGR 26c ASO also reduced ORF 6 protein levels (Fig. 2.7E). Finally, we tested whether EGR 26c ASO altered the expression of ORF 50 by qRT-PCR. We found that EGR 26c ASO significantly reduced the abundance of spliced ORF 50 transcripts (Fig. 2.7F). Northern blot analysis using probe 2 on RNA selected for polyadenylated transcripts revealed the presence of smaller transcripts that might be targeted by EGR 26c ASO (see Fig. 2.3B). The fact that EGR 26a-c ASOs all reduce ORF 6 transcript abundance (Fig. 2.7D) suggests that a transcript that overlaps genomic coordinates 85576 to 85891 acts early in the viral life cycle, affecting gene expression of multiple specific viral genes.
2.4 Discussion

In this paper, we provide the first data demonstrating that pervasive transcription of a herpesvirus genome, identified by pan-genomic analysis of RNA expression using both tiled arrays and RNA sequencing can generate functionally important RNAs for viral gene and protein expression. This is significant because it was unknown whether extensive antisense transcription observed by many groups from various herpesvirus genomes is functionally important or is due to "read-through" transcription that results in large amounts of functionally irrelevant RNA. We identified six EGRs that generate transcripts that altered ORF 4 surface expression. Upon examining transcripts encoded by two EGRs in more detail, we confirmed that targeting transcripts from the strand of the genome antisense to known protein coding genes alters multiple parts of the viral transcriptional and translational program.

These data have fundamental implications for the approach to and interpretation of genetic studies and highlight the importance of mapping transcripts derived from the strand opposite to ORFs. A limitation of traditional genetic approaches is the emphasis on ORFs and the utilization of genetic approaches that alter the sequence of both strands of the viral genome. This approach has been very fruitful, but misses an important layer of complexity to viral gene regulation revealed here. One of our most notable findings is the extreme complexity of transcripts emanating from EGRs and the fact that many of these transcripts are polyadenylated but remain concentrated in the nucleus. Furthermore, our data suggest that in some cases there may be important contributions of less abundant transcripts whose presence may be difficult to detect using classical methods (see EGR 26 as an example). Taken together, our results support the use of sensitive technologies, such as next-generation RNA sequencing, in transcript analysis and the necessity of considering antisense transcripts in genetic analyses of herpesviruses and
designing functional experiments to evaluate the function of each strand of the viral genome independently.

2.4.1 Implications of transcriptional complexity of EGR-encoded RNAs

It is interesting to reconsider previous findings in view of the transcriptional complexity underlying the MHV68 genome. Here, we have reported a region of the MHV68 genome in which detailed analysis reveals at least five transcripts overlapping the late capsid protein, M9 (ORF 65), each of which may play distinct roles for the virus (5, 41). Previous studies using RNase protection assays and qRT-PCR have identified the M9 region as: (i) a candidate region of latent gene expression; (ii) a component of the virion; and (iii) producing an RNA with immediate-early kinetics (22, 46, 59). We note that previous analyses are complicated by the presence of multiple overlapping transcripts because probes designed to M9 detect five independent transcripts (including EGR 26 transcript A and EGR 27 transcript A).

Additionally, our results suggest that caution should be used for interpretation of transposon screens (39, 48). Because ~90% of EGR nucleotides overlap with transcription on the opposite strand ((25) and Fig. 2.8), any mutagenesis strategy that disrupts both strands may alter functionally relevant transcripts derived from either strand. For example, transcripts encoded by EGRs 26 and 27 are antisense to known protein-coding ORFs 63, 64, 68 and 69. ORFs 63 and 64 are predicted to be essential tegument proteins, ORF 68 an essential glycoprotein, and ORF 69 an essential gene of unknown function (48, 58). Thus, while these transposon screens provide invaluable functional data, defining the genetic elements responsible for observed phenotypes will need to incorporate the transcriptional complexity of each region and subsequently utilize strand-specific approaches for functional analysis.
2.4.2 Potential mechanism for EGRs 26 and 27 transcript function.

Our studies show that transcripts encoded by EGRs are functionally important and effect multiple parts of the viral life cycle, but do not identify specific viral transcripts with individual functions relevant to viral transcription and protein expression. The reduction in ORF 50 transcript expression that we observed following targeting with EGR 27a and 27d ASOs and the nuclear localization of EGR 27 transcripts is consistent with transcripts targeted by these ASOs functioning to epigenetically regulate the ORF 50 promoter(s). There is a growing consensus in the mammalian ncRNA literature that epigenetic modification of chromatin is a key function of long ncRNAs (1, 29, 44, 45, 54). Interestingly, the KSHV ncRNA, polyadenylated nuclear (PAN) RNA, associates with the ORF 50 promoter as well as the demethylases, JMJD3 and UTX, and the methyltransferase, MLL2 (47). Many studies highlight the importance of epigenetic modifications in controlling herpesvirus gene expression, and herpesvirus proteins can interact with histone deacetylases (HDACs) to prevent gene silencing (9, 19, 30, 42, 49). We and others have identified several important modifications that silence ORF 50 in macrophages, B cells, and cells from latently infected mice, including methylation of the distal ORF 50 promoter (16, 17) and recruitment of HDACs and the nuclear receptor corepressor (NCoR) to the core ORF 50 promoter (15, 63). It is intriguing to speculate whether transcripts targeted by EGR 27a, 27d, and/or 26c ASO may act as a molecular scaffold for chromatin modifying proteins or prevent the association of known repressive complexes with ORF 50 promoters in lytically infected fibroblasts.

In contrast, EGR 27c ASO, which specifically targeted EGR 27 transcript A, did not reduce ORF 50 transcript expression, suggesting that this RNA may act later in the viral life cycle, playing a role in viral DNA replication or virion assembly. Interestingly, a role for an
EBV RNA, BHLF1 transcript, has been recently described in DNA replication; BHLF1 transcript forms an RNA-DNA hybrid molecule at the origin of lytic replication (OriLyt) and is important for recruitment of the viral single stranded binding protein, BALF2, to OriLyt (43). These results suggest that RNAs derived from EGRs may play a range of different roles in viral replication and merit further investigation and consideration in functional studies.

Many new transcripts have been identified not only in MHV68 but also in important human pathogens, including KSHV and HCMV (7, 14, 36, 62, 64). Our results suggest that transcripts in each of these viruses may play crucial roles in the viral life cycle and/or in viral pathogenesis, and that strand-specific approaches combined with detailed transcriptional analysis similar to ours will allow for identification of novel transcripts critical for viral gene expression and infection.
2.5 Materials and Methods

2.5.1 Cells, virus, and virus assays.

NIH 3T12 fibroblasts (ATCC CCL-164) were grown in Dulbecco's modified Eagle medium (DMEM) containing 5% fetal calf serum (FCS), 2 mM L-glutamine, and 10 mM HEPES. Cells were infected with MHV68 clone WUMS (ATCC VR-1465) for 1 hour at 37°C at a multiplicity of infection (MOI) of 5 or 10 as indicated. Viral passage and titer were performed on NIH 3T12 fibroblasts as described (61) except that cells were overlaid with 2% methylcellulose in MEM supplemented with 5% FCS and 2 mM L-glutamine.

2.5.2 Antisense oligonucleotides.

Custom antisense gapmer oligonucleotides (ASOs) containing a phosphorothioate backbone and locked nucleic acid (LNA) residues to increase probe stability and knockdown efficiency (18, 24, 31) were designed and synthesized by Exiqon (Woburn, MA) to target viral transcripts (Table 2.1). Gapmers contain LNA bases at their ends surrounding a central stretch of DNA enabling RNase H-mediated cleavage of their targets (60). Transfections were performed with 40 pmol of ASO per 10^5 cells using Lipofectamine 2000 (Life Technologies, Grand Island, NY) according to the manufacturer's instructions. An ASO designed to target GFP was used as a negative control. Cells were infected with MHV68 six hours after transfection and harvested at the indicated times post infection. Toxicity was assessed by alamarBlue (Life Technologies) at 30 hours post transfection according to the manufacturer's instructions. There was no association of toxicity relative to untransfected cells as assessed by alamarBlue and phenotype as assessed by ORF 4 surface staining (Spearman's Correlation, p=0.8250).
2.5.3 Western blot analysis, antisera and antibodies.

For MHV68 western blot analysis, samples were lysed at 18 hpi in 2x Laemmeli buffer, subjected to protein electrophoresis on 4-15% or 4-20% Tris gradient gels (Bio-Rad), and then transferred to polyvinylidene difluoride membrane. For western blot analysis of subcellular fractions, 2x Laemmeli buffer was added to samples lysed in the relevant buffer as described below (Subcellular Fractionation). 4 μg of protein per sample was subjected to protein electrophoresis on a 10% Tris gel, then transferred to polyvinylidene difluoride membrane. Antibodies used were anti-Lamin B1 (AbCam), anti-GAPDH (clone GAPDH-71.1) (Sigma, St. Louis, MO), anti-calreticulin (clone 16/Calreticulin) (BD Biosciences), polyclonal rabbit antisera generated to ORF 6 (40), M3 (55), ORF 26 (4) or M9 (4) and a goat anti-rabbit or goat anti-mouse HRP-conjugated secondary antibody as appropriate (Jackson Immunoresearch, West Grove, PA). MHV68 western blots were stripped and reprobed with anti-beta-actin (clone AC-74) (Sigma) and then with a goat anti-mouse HRP-conjugated secondary antibody (Jackson Immunoresearch) to control for loading. Blots were developed with ECL Plus chemiluminescent reagent (GE Healthcare Life Sciences) or Pierce ECL chemiluminescent reagent (Thermo Scientific) and imaged using film or a Storm 840 phosphorimager. Bands were quantitated using ImageJ (NIH). For each sample, the indicated protein was normalized to the actin loading control.

2.5.4 Flow cytometric analysis for detection of surface expression of ORF 4 protein.

At 24 hpi, cells were removed from tissue culture dishes by gentle scraping after incubation in a 0.02% EDTA solution for 10 minutes at 4°C. Cells were fixed in 2% formaldehyde and stained using anti-ORF 4 antiserum (27) at 1:10,000 or preimmune rabbit
serum (Cocalico Biologicals, Reamstown, PA), followed by donkey anti-rabbit DyLight 649 secondary antibody (Biolegend, San Diego, CA). Flow cytometry was performed with a FACSCalibur (BD Biosciences, San Jose, CA). Data were analyzed by FlowJo (Tree Star; Ashland, OR).

2.5.5 RNA sequencing (RNA-Seq) library construction and expression analysis.

Total RNA was isolated from TRIzol (Life Technologies) as described (25) at 6 or 18 hpi as indicated. Poly(A) selected RNA was chemically fragmented using RNA Fragmentation Reagents (Life Technologies) and purified using RNeasy MinElute Cleanup columns (Qiagen). Directional RNA-Seq libraries were generated using 600 nanograms of fragmented 18 hpi RNA according to Illuminaâ€™s directional mRNA-Seq protocol or 200 nanograms of 6 hpi RNA according to Illumina's directional Tru-Seq protocol. Nucleotides with a Phred quality score less than 20 were trimmed from the 3' end of raw Illumina reads. Trimmed reads with a mean quality score less than 10 or a length less than 20 were discarded. Reads were then mapped to mouse rRNA, MHV68 genome (GenBank accession number: U97553), and mouse genome sequences (Build 37) with the short read aligning program Bowtie-0.12.5 (33). Bowtie alignments were performed using the parameter settings â€“bestâ€œ and " -e 420" or "-e 600" settings for 76 nucleotide or 100 nucleotide reads, respectively. Bowtie output from the MHV68 genome mapping was converted to wig files of read depth coverage, on a log\(_2\) scale, and visualized using Gbrowse (www.gbrowse.org). Correlation coefficients between read depth coverage and tiled array signals were calculated using Spearmanâ€™ ranked correlation coefficient in the R statistical environment. Sequencing data are available at the National Center for Biotechnology Information.
Information (NCBI) Sequence Read Archive (SRA) accession numbers SRX403400 to SRX403403 for samples at 6 hpi and SRX403404 to SRX403407 for samples at 18 hpi.

**2.5.6 Identification of polyadenylated reads from RNA-Seq data sets.**

3'-terminal adenosine (A) residues were trimmed from filtered reads with greater or equal to five 3'-terminal A's. Both trimmed and untrimmed reads were mapped to MHV68 with Bowtie with default "-e" settings. Mapped reads that contained at least five non-genomic A's were considered putative polyadenylated reads. If a potential polyadenylation cleavage site was located within or downstream of a stretch of genomic A's, the coordinates of this potential site was moved upstream of the genomic poly(A) stretch to maintain consistency. Sites that were located adjacent to a stretch of seven or more genomic A's were excluded due to the possibility that during RACE validation such sites could provide false internal priming of the oligo-dT primer. Those sites with an average of at least five supporting reads between biologic replicates were clustered. The site reported in Table 2.3 is that with the most supporting reads in 30 nucleotide sliding windows. Raw read quality filtering and the identification of putative polyadenylated reads were done using a combination of Linux utilities and custom Perl scripts, available upon request.

**2.5.7 Rapid Amplification of cDNA Ends (RACE).**

5' and 3' transcript ends were identified by RACE using Invitrogen's 5' and 3' RACE systems according to the manufacturer's instructions. cDNA was generated from total RNA extracted 18 hpi using a gene-specific primer (for 5' RACE) or an oligo(dT)-containing adapter primer (for 3' RACE). PCR amplification was performed using gene-specific primers with
Invitrogen's amplification primers. The following gene-specific primers were used for reverse transcription in 5' RACE: EGR 26 transcript A: 5'-CGATCAGGTGGCTCAACTGG-3'; EGR 27: 5'-GCGAGGAGCAGCAGCACAGCAGA-3'. For 5' RACE reactions, the PCR primers used were as follows: EGR 26 transcript A: 5'-CTGCTCACATAAAGGTATCTGG-3'; EGR 27: 5'-GCAGAGGTCCGGTCCAGTACAGGA-3' and 5'- GGTCCGTCCAGTAGCGA-3'. For 3' RACE reactions, the PCR primers used were as follows: EGR 27 transcript A: 5'-GCCAGACATTGCACACAACAC-3'; EGR 27 transcript B (B1): 5'-CGAGATACATGTTGAAGCATTCA-3'. The resulting PCR products were gel purified and ligated into a pCR4-TOPO TA sequencing vector (Life Technologies). Universal M13 forward and reverse primers were used for sequencing.

2.5.8 Northern blot analysis.

Total RNA was isolated as described (25) at 14 or 18 hours post infection (hpi) as indicated. Templates for northern probes were amplified by PCR from viral or mouse genomic DNA (using PCR primers listed in Table 2.2), prepared as described in (25) for probes 2 and 7 (formerly EGR 26 probes 1 and 4 (25), respectively), or obtained from a commercial vendor for actin (Life Technologies). Northern blotting using Ambion's NorthernMax kit (Life Technologies, Grand Island, NY) and generation of single-stranded P$^{32}$-labeled RNA probes using the Maxiscript Sp6/T7 kit (Life Technologies) were performed as described (25). Probe 12 was generated using the mirVANA miRNA Probe Construction kit (Life Technologies) according to the manufacturer's instructions. Five micrograms of total RNA were used for all northern blot analysis unless otherwise stated. Membranes were scanned using a Storm 840 Phosphorimager and quantitated using ImageQuant TL (GE Healthcare Biosciences, Pittsburgh,
PA). For quantitation of each sample, the indicated transcript signal was normalized to the signal from the actin loading control.

2.5.9 Subcellular Fractionation.

Nuclei were separated from cytoplasm using a protocol adapted from published methods (21). Briefly, 3T12 cells were removed from plates by trypsinization at 18 hpi and centrifuged to pellet cells. After washing with PBS, the pellet was resuspended in 150 mM NaCl, 50 mM HEPES, 1% NP40 and 1 U/μl SUPERase-In RNase Inhibitor (Ambion) to disrupt non-nuclear membranes and incubated on ice for 30 minutes. The lysate was centrifuged at 7000 RCF to pellet the nuclei, and the supernatant removed for RNA and protein analysis (cytoplasmic fraction). After washing with PBS, the nuclear pellet was resuspended in cold 150 mM NaCl, 50 mM HEPES, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate, and 1 U/μl SUPERase-In RNase Inhibitor (Life Technologies) to disrupt the nuclear membrane and incubated for two hours at 4°C with rotation. RNA was extracted from lysates using TRIzol LS (Life Technologies) and analyzed by northern blot as described above. Approximate cellular equivalents were calculated based on the relative amounts of RNA recovered from nuclear and cytoplasmic fractions. For comparison, unfractionated cells were collected in TRIzol (Life Technologies) and then processed to isolate RNA according to the manufacturer's instructions. Samples were also analyzed by western blot to ensure adequate separation of fractions as described above.

2.5.10 Quantitative reverse transcriptase-PCR (qRT-PCR).

For analysis of ORF 50 and ORF 29 transcripts, cDNA was synthesized from 1 μg of RNA using SuperScript III (Life Technologies) and random hexamers (Life Technologies) as
described (35). qPCR was performed using Power SYBR Green Master Mix (Applied Biosystems) and the following primer sequences: 5'-TGCCCCCATGTTTGTGATG-3' and 5'-TGTGGTCA T GAGCCCTTCC-3' for glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (15), 5'-GATTC C C CTTCCAGCGATAAG-3' and 5'- CAGACATTGTAGAAGTCAGGTC-3' for spliced ORF 50 transcript, and 5'- TTCTCA T TGGCA TCTTTGAGG-3' and 5'-GGAAATGGGGTGATCCTGT-3' for spliced ORF 29 transcript (35) on the StepOnePlus System (Life Technologies). Transcript levels were normalized to GAPDH within each sample and compared to untransfected cells using the \( \Delta \Delta^Ct \) method, where Ct is the threshold cycle (38).

2.5.11 Statistical analysis.

Data were analyzed statistically with Prism 6 software (GraphPad Software, La Jolla, CA). All experimental conditions were compared to the corresponding untransfected or GFP ASO transfected controls as noted. Data were analyzed by two-tailed paired t-test or one-way ANOVA as indicated. All significant differences are noted.
2.6 Acknowledgements.

This work was supported by R01 CA096511 and U54 AI057160 ARRA Administrative Supplement (to H.W.V.). S.P.C. is supported by an NIH F30 fellowship grant HL099019 and T.A.R. was supported by the Damon Runyon Cancer Research Foundation.

Thanks to L. Steven Johnson for analysis of RNA sequencing data; Tiffany Reese, Xin Zhang, and Amal Kambal for generating RACE data, Erning Duan for assistance with western blots and Xin Zhang for assistance with qPCR data.

We thank Washington University Genome Technology Access Center (GTAC) for Illumina sequencing and technical advice, Jean E. Schaffer and Ben S. Scruggs for helpful experimental advice and discussions, Catherine Liu and Lori Xu for assistance with experiments, the Sun lab for kind gifts of M9 and ORF 26 antiserum, Christine Yokoyama for assistance with figure design, and members of the Virgin lab for helpful review of the manuscript.
2.7 References


47. **Rossetto, C. C., and G. Pari.** 2012. KSHV PAN RNA associates with demethylases UTX and JMJD3 to activate lytic replication through a physical interaction with the virus genome. PLoS Pathog 8:e1002680.


Figure 2.1: An antisense oligonucleotide to ORF 6 decreases ORF 6 transcript and protein expression and late gene expression. 3T12 cells transfected with ASOs targeting ORF 6, M3, or GFP (negative control) or left untransfected (No ASO) were infected with MHV68. (A) Representative northern blot for ORF 6 or actin transcripts at 14 hpi and corresponding quantification of ORF 6 monocistronic transcript levels normalized to those of actin (MOI=10,
mean +/- S.E.M., 3 experiments, ** p<0.01 by paired t-test). (B) Representative western blot for ORF 6 and actin protein at 18 hpi and corresponding quantification of ORF 6 normalized to actin (MOI=10, mean +/- S.E.M., 8 experiments, *** p<0.001 by paired t-test). Relevant lanes of representative blots are shown. (C) Flow cytometry analysis of ORF 4 surface expression at 24 hpi (MOI= 5 or 10). Flow cytometry data are graphed as the percentage of ORF 4 positive cells for each condition normalized to the value from untransfected cells (mean of 34-35 replicates +/- S.E.M, statistically significant results relative to GFP are indicated, **** p<0.0001 by one-way ANOVA with Dunnett's post-test). (D) Western blot analysis of M9 or ORF 26 at 18 hpi. Relevant lanes of representative blots are shown. Data are representative of four independent experiments.
Figure 2.2: Targeting of EGRs 9, 16, 18, 23, 26, and 27 reduces ORF4 surface expression on infected cells. 3T12 cells transfected with the indicated ASOs or left untransfected were analyzed for cell surface expression of ORF 4 at 24 hpi by flow cytometry. Cells were infected at MOI = 5 or 10. Controls were included in parallel in each experiment in which EGRs were analyzed. The percentage of ORF 4 positive cells for each condition was normalized to the value from untransfected cells. Data are presented as mean of the pooled data +/- S.E.M; n=number of independent experiments. Statistically significant results relative to GFP are shown. * p<0.05; ** p<0.01; *** p<0.001; ****p<0.0001 by one-way ANOVA with Dunnett’s post-test. Black bars: controls (data reproduced from Fig. 2.1C); gray bars: EGRs.
Figure 2.3: EGR 26 - 27 transcriptional architecture. (A) Schematic representation of EGR 26-27/M9 region transcripts. Approximate size and transcript name are listed below the relevant transcript. (B) Northern blot detection of EGR 26-27/M9 region transcripts. RNA harvested at 18 hpi (MOI=10) was analyzed by northern blot analysis using the indicated probes. 28S and 18S rRNA bands visualized by ethidium bromide staining to demonstrate equal loading are shown below the corresponding northern blots. Northern blots analyzed with probe 12 and probe 2 (at far right of panel) had 500 ng of polyA-selected RNA per lane. * indicates virus-specific bands that are of comparable size to EGR 27 transcripts A and B. ** indicates smaller EGR 26 transcripts referenced in text.
Figure 2.4: EGR 26 and 27 transcripts are enriched in nucleus of infected cells. Protein or RNA extracted from nuclear or cytoplasmic fractions or unfractionated cells (Total) was analyzed by western blot (A) or northern blot (B) at 18 hpi (MOI=10). (A) 4 μg of protein was loaded per sample. (B) 5 μg of RNA for each total sample and approximate cellular equivalents for nuclear (0.5 μg) or cytoplasmic (4.5 μg) fractions were analyzed. Data are representative of three independent experiments. EGR 26-A indicates EGR 26 transcript A (as shown in Fig. 2.3) detected using northern blot probe 2. EGR 27-A indicates EGR 27 transcript A and EGR 27-B indicates EGR 27 transcript B (as shown in Fig. 2.3) detected using northern blot probe 11.
Figure 2.5: Effect of targeting EGR 26 and EGR 27 transcripts on ORF 4 surface expression. (A) Schematic of EGR 26-27 region. See Table 2.1 for specific coordinates targeted by ASOs. (B-D) 3T12 cells transfected with the indicated ASOs were analyzed for EGR 27 transcript A (B), EGR 27 transcript B (C), or EGR 26 transcript A (D) at 14 hpi by northern blot. EGR 27 transcripts A and B were detected using northern blot probe 11 and EGR 26 transcript A was detected using northern blot probe 2. Depending on the experiment, 1.5–5 μg RNA was used per sample. The graphs are quantitation of signal intensity of the indicated transcript normalized to actin and presented as a fraction of the signal from untransfected cells. Representative northern blots are shown. The actin for EGR 27 transcripts is reproduced for panels B and C as the same blot is displayed. # indicates the location of a tear in the agarose gel. Data are presented as mean of the pooled data (3–7 experiments) +/- S.E.M. (E) 3T12 cells transfected with the indicated ASOs were analyzed for cell surface expression of ORF4 at 24 hpi by flow cytometry. As in Figure 2.2, the percentage of ORF4 positive cells for each condition was normalized to the value from untransfected cells. Data are presented as mean of the pooled data (5–35 experiments) +/- S.E.M. (F) 3T12 cells transfected with the indicated ASOs were analyzed for viral titer at 24 hpi by plaque assay. Data are presented as mean (3 experiments) +/- S.E.M. Statistically significant results relative to GFP are shown (* p<0.05; ** p<0.01; *** p<0.001; **** p<0.0001 by one-way ANOVA with Dunnett’s post-test). Black bars: controls; gray bars: EGR 26-27 ASOs.
Figure 2.6: Effect on immediate-early, early, and late genes by EGR27. 3T12 cells transfected with GFP or ASOs targeting EGR 27 or left untransfected (No ASO) were infected with MHV68 (MOI=10) and analyzed for protein (A and D) or transcript levels (B-C and E). See figure 2.5 and Table 2.1 for ASO locations. (A) Representative western blot for M9 and ORF 26.
proteins at 18 hpi (2-3 experiments). (B) ORF 29 transcript levels at 14 hpi. 1 μg RNA was reverse transcribed (RT) and cDNA was analyzed by qPCR using primers designed to detect spliced ORF 29 transcripts or GAPDH. Data are presented as relative ORF 29 abundance normalized to GAPDH and compared to untransfected cells by the ∆∆Ct method (mean +/- S.E.M.; 3-8 experiments). (C) Representative northern blot for ORF 6 and actin at 14 hpi and corresponding quantification of ORF 6 transcript levels normalized to actin and compared to the value in untransfected cells (mean +/- S.E.M., 5-7 experiments). (D) Representative western blot for ORF 6 and actin at 18 hpi (3 experiments). The representative experiment is the same as shown in (A). (E) ORF 50 transcript levels at 14 hpi measured by qRT-PCR as in (B). Statistical analyses were performed by one-way ANOVA with Dunnett’s post-test. * p<0.05; ** p<0.01; *** p<0.001; **** p<0.0001; ns = not significant.
Figure 2.7: EGR 26c ASO decreases specific genes of all kinetic classes. 3T12 cells transfected with GFP or EGR 26c ASO or untransfected (No ASO) were infected with MHV68 (MOI=10) and analyzed for protein (A, C, and E) or transcript levels (B, D and F). (A) Representative western blots for M9 and actin at 18 hpi (2 experiments). (B) Representative northern blot using a probe to M3 or actin at 14 hpi and corresponding quantification of M3 transcript levels normalized to those of actin (mean +/- S.E.M., 3 experiments). 0.5 μg RNA was used per lane for M3 northern blot. (C) Representative western blot for M3 protein at 18 hpi and corresponding quantification of M3 protein levels normalized to those of actin (mean +/- S.E.M.,
4 experiments). (D) Representative northern blot for ORF 6 and actin at 14 hpi and corresponding quantification of ORF 6 transcript levels normalized to actin for cells transfected with GFP or EGR 26a-f ASOs (mean +/- S.E.M., 3-5 experiments). (E) Representative western blot for ORF 6 and actin at 18 hpi and corresponding quantification of ORF 6 protein normalized to actin (mean +/- S.E.M., 5 experiments). Representative western blots for ORF 6 are the same as presented in figure 2.1. (F) ORF 50 transcript levels at 14 hpi. 1 μg RNA was reverse transcribed and cDNA was analyzed by qPCR using primers designed to detect spliced ORF 50 transcripts or GAPDH. Data are presented as relative ORF 50 abundance normalized to GAPDH and compared to untransfected cells by the \( \Delta \Delta \text{Ct} \) method (mean +/- S.E.M.; 4 experiments). Statistical analyses were performed by paired t-test (A-C and E) or one-way ANOVA with Dunnett’s post-test (D and F). * p<0.05; ** p<0.01; *** p<0.001; ns = not significant.
Figure 2.8: Genome-wide RNA expression in MHV-68 as detected by RNA-Seq. RNA-Seq RNA expression measurements are highly reproducible, strongly correlate with our earlier tiled array results, and confirmed the expression of all 30 expressed genomic regions (EGRs). Transcriptional signal from infected 3T12 cells at 18 hpi was plotted on a log2 scale underneath the annotated gene positions. Read depth from two independent RNA-Seq experiments was plotted, on a log2 scale, as blue/light blue and red/light red lines for the positive and negative genomic strands, respectively. Blue or red arrows represent annotated genes on the positive or negative strand of the genome, respectively. Gray boxes represent the two internal repeat regions of the MHV-68 genome that were excluded from our analysis. EGRs 1-30 are shown as hatched arrows. (A) Genomic positions 1-40 kb. (B) Genomic positions 40-80 kb. (C) Genomic positions 80-119.45 kb.
Table 2.1: Antisense oligonucleotide probes

<table>
<thead>
<tr>
<th>Probe Name</th>
<th>Sequence (5'-3')</th>
<th>Start Position</th>
<th>End Position</th>
<th>Target Strand$^1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>as-GFP</td>
<td>GGAGGTGTGGAGGGTT</td>
<td>74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>as-ORF 6</td>
<td>CCCCCAACAGCCCTTTTC</td>
<td>12242</td>
<td>12257</td>
<td></td>
</tr>
<tr>
<td>as-M3</td>
<td>CTCCACACCAGGTAAA</td>
<td>6617</td>
<td>6632</td>
<td></td>
</tr>
<tr>
<td>as-EGR 9a</td>
<td>CTGCCCTGCACCCCT</td>
<td>1224</td>
<td>28576</td>
<td></td>
</tr>
<tr>
<td>as-EGR 9b</td>
<td>CTCAATGCCCACCATC</td>
<td>29191</td>
<td>29206</td>
<td></td>
</tr>
<tr>
<td>as-EGR 10a</td>
<td>ATACGCTATACCGAC</td>
<td>30044</td>
<td>30059</td>
<td></td>
</tr>
<tr>
<td>as-EGR 10b</td>
<td>TGATCTGCTGCCTGTT</td>
<td>30529</td>
<td>30544</td>
<td></td>
</tr>
<tr>
<td>as-EGR 11a</td>
<td>ACCATGCGAGACATCA</td>
<td>30812</td>
<td>30827</td>
<td></td>
</tr>
<tr>
<td>as-EGR 11b</td>
<td>GAAGACTGCTAGGGAAT</td>
<td>31578</td>
<td>31593</td>
<td></td>
</tr>
<tr>
<td>as-EGR 15</td>
<td>AGAGACAAGAGACTGGAA</td>
<td>46703</td>
<td>46718</td>
<td></td>
</tr>
<tr>
<td>as-EGR 16</td>
<td>ACCAGCCTTTTTGGATT</td>
<td>50819</td>
<td>50834</td>
<td></td>
</tr>
<tr>
<td>as-EGR 18a</td>
<td>AACTGTTGTGTTGCTCA</td>
<td>56313</td>
<td>56328</td>
<td></td>
</tr>
<tr>
<td>as-EGR 18b</td>
<td>ATTTCTAGCAGGCTTT</td>
<td>56450</td>
<td>56465</td>
<td></td>
</tr>
<tr>
<td>as-EGR 20</td>
<td>GAATAGCACGCAACGA</td>
<td>60121</td>
<td>60136</td>
<td></td>
</tr>
<tr>
<td>as-EGR 21</td>
<td>TAGACCAGGATATGAC</td>
<td>62546</td>
<td>62561</td>
<td></td>
</tr>
<tr>
<td>as-EGR 23a</td>
<td>ACACATACACTGGA</td>
<td>68533</td>
<td>68548</td>
<td></td>
</tr>
<tr>
<td>as-EGR 23b</td>
<td>CCCTCTCTCAGGACCA</td>
<td>68704</td>
<td>68719</td>
<td></td>
</tr>
<tr>
<td>as-EGR 23c</td>
<td>CCAATCCACCGTAGCC</td>
<td>69268</td>
<td>69283</td>
<td></td>
</tr>
<tr>
<td>as-EGR 23d</td>
<td>AACAGAGACGCAGCAT</td>
<td>70737</td>
<td>70752</td>
<td></td>
</tr>
<tr>
<td>as-EGR 25</td>
<td>TCTACTGCTGCTCTAAC</td>
<td>74083</td>
<td>74098</td>
<td></td>
</tr>
<tr>
<td>as-EGR 26a</td>
<td>CCTCCAGTGGAGACAGAT</td>
<td>85876</td>
<td>85891</td>
<td></td>
</tr>
<tr>
<td>as-EGR 26b</td>
<td>AAACCTTTCCTGATAG</td>
<td>85596</td>
<td>85611</td>
<td></td>
</tr>
<tr>
<td>as-EGR 26c$^2$</td>
<td>ACATCATCAACGGGAC</td>
<td>85576</td>
<td>85591</td>
<td></td>
</tr>
<tr>
<td>as-EGR 26d</td>
<td>ACTTATCTAGGAGCC</td>
<td>85357</td>
<td>85372</td>
<td></td>
</tr>
<tr>
<td>as-EGR 26e</td>
<td>CATTGTCTCATTACC</td>
<td>85256</td>
<td>85271</td>
<td></td>
</tr>
<tr>
<td>as-EGR 26f</td>
<td>ACACACTGATGCTCGT</td>
<td>85006</td>
<td>85021</td>
<td></td>
</tr>
<tr>
<td>as-EGR 27a</td>
<td>TGCTCGATGCCTGTT</td>
<td>98843</td>
<td>98858</td>
<td></td>
</tr>
<tr>
<td>as-EGR 27b$^2$</td>
<td>GCACGACTCTACCTTT</td>
<td>98705</td>
<td>98720</td>
<td></td>
</tr>
<tr>
<td>as-EGR 27c</td>
<td>TACCATGCCGCTCAAAC</td>
<td>98057</td>
<td>98072</td>
<td></td>
</tr>
<tr>
<td>as-EGR 27d</td>
<td>AACTACGCTACCCACT</td>
<td>98004</td>
<td>98019</td>
<td></td>
</tr>
<tr>
<td>as-EGR 30</td>
<td>CGACCACCTCCCTTATT</td>
<td>116835</td>
<td>116850</td>
<td></td>
</tr>
</tbody>
</table>

$^1$assuming genome orientation as in (25, 58).

$^2$EGR 26 and 27 ASOs used in screen (Fig. 2.2).
Table 2.2: PCR primers for northern blot probes

<table>
<thead>
<tr>
<th>Target Gene</th>
<th>Probe Name</th>
<th>Orientation</th>
<th>Primer Sequence (5'-3')&lt;sup&gt;1&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGR 26</td>
<td>Probe 1</td>
<td>Forward</td>
<td>ATTTAGGTGACACTATAGAAGCTTGCCTGATCTGACG</td>
</tr>
<tr>
<td>EGR 26</td>
<td>Probe 3</td>
<td>Forward</td>
<td>ATTTAGGTGACACTATAGAAGCTTGCCTGATCTGACG</td>
</tr>
<tr>
<td>EGR 26</td>
<td>Probe 4</td>
<td>Forward</td>
<td>ATTTAGGTGACACTATAGAAGCTTGCCTGATCTGACG</td>
</tr>
<tr>
<td>EGR 26</td>
<td>Probe 5</td>
<td>Forward</td>
<td>ATTTAGGTGACACTATAGAAGCTTGCCTGATCTGACG</td>
</tr>
<tr>
<td>EGR 26</td>
<td>Probe 6</td>
<td>Forward</td>
<td>ATTTAGGTGACACTATAGAAGCTTGCCTGATCTGACG</td>
</tr>
<tr>
<td>EGR 26</td>
<td>Probe 7</td>
<td>Forward</td>
<td>ATTTAGGTGACACTATAGAAGCTTGCCTGATCTGACG</td>
</tr>
<tr>
<td>EGR 27</td>
<td>Probe 8</td>
<td>Forward</td>
<td>ATTTAGGTGACACTATAGAAGCTTGCCTGATCTGACG</td>
</tr>
<tr>
<td>EGR 27</td>
<td>Probe 9</td>
<td>Forward</td>
<td>ATTTAGGTGACACTATAGAAGCTTGCCTGATCTGACG</td>
</tr>
<tr>
<td>EGR 27</td>
<td>Probe 10</td>
<td>Forward</td>
<td>ATTTAGGTGACACTATAGAAGCTTGCCTGATCTGACG</td>
</tr>
<tr>
<td>M3</td>
<td>M3</td>
<td>Forward</td>
<td>ATTTAGGTGACACTATAGAAGCTTGCCTGATCTGACG</td>
</tr>
<tr>
<td>Neat2</td>
<td>Neat2&lt;sup&gt;3&lt;/sup&gt;</td>
<td>Forward</td>
<td>ATTTAGGTGACACTATAGAAGCTTGCCTGATCTGACG</td>
</tr>
</tbody>
</table>

<sup>1</sup>Sp6 and T7 promoter sequences are underlined.

<sup>2</sup>adapted from (27).

<sup>3</sup>adapted from (23).
Table 2.3: 3' ends of MHV68 transcripts identified by RNA sequencing. The average number of reads per position at 6 hpi and 18 hpi is indicated in bold in the 2nd and 3rd columns, respectively. The numbers in parentheses indicate the number of reads from specific experiments at the indicated time point. The numbers in parentheses indicate the number of reads detected in individual experiments at the indicated time point. Gene region refers to a possible gene associated with the indicated polyadenylation site selected based on the gene with the closest 3' end to a given site.

<table>
<thead>
<tr>
<th>Position</th>
<th>Number of Reads at 6 hpi</th>
<th>Number of Reads at 18 hpi</th>
<th>Strand</th>
<th>Gene Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>3378</td>
<td>0</td>
<td>84 (124, 45)</td>
<td>-</td>
<td>M2</td>
</tr>
<tr>
<td>6007</td>
<td>678 (1030, 326)</td>
<td>328 (297, 359)</td>
<td>-</td>
<td>M3</td>
</tr>
<tr>
<td>9866</td>
<td>246 (315, 178)</td>
<td>104 (44, 163)</td>
<td>+</td>
<td>M4</td>
</tr>
<tr>
<td>14909</td>
<td>424 (534, 315)</td>
<td>300 (314, 287)</td>
<td>+</td>
<td>ORF 6</td>
</tr>
<tr>
<td>19123</td>
<td>49 (64, 34)</td>
<td>848 (836, 861)</td>
<td>+</td>
<td>ORF 8</td>
</tr>
<tr>
<td>23293</td>
<td>14 (18, 9)</td>
<td>14 (2, 25)</td>
<td>+</td>
<td>ORF 10</td>
</tr>
<tr>
<td>24698</td>
<td>192 (272, 113)</td>
<td>109 (80, 138)</td>
<td>+</td>
<td>ORF 11</td>
</tr>
<tr>
<td>24698</td>
<td>28 (41, 15)</td>
<td>71 (25, 117)</td>
<td>-</td>
<td>K3</td>
</tr>
<tr>
<td>28264</td>
<td>348 (527, 169)</td>
<td>777 (588, 966)</td>
<td>-</td>
<td>ORF 17</td>
</tr>
<tr>
<td>30458</td>
<td>18 (25, 12)</td>
<td>0</td>
<td>+</td>
<td>ORF 18b</td>
</tr>
<tr>
<td>30700</td>
<td>136 (213, 59)</td>
<td>86 (40, 133)</td>
<td>-</td>
<td>ORF 19</td>
</tr>
<tr>
<td>30785</td>
<td>322 (529, 114)</td>
<td>102 (80, 125)</td>
<td>+</td>
<td>ORF 18b</td>
</tr>
<tr>
<td>34847</td>
<td>142 (203, 80)</td>
<td>78 (18, 138)</td>
<td>+</td>
<td>ORF 21</td>
</tr>
<tr>
<td>34959</td>
<td>36 (51, 21)</td>
<td>31 (9, 53)</td>
<td>+</td>
<td>ORF 21</td>
</tr>
<tr>
<td>35524</td>
<td>10 (17, 4)</td>
<td>34 (18, 51)</td>
<td>+</td>
<td>ORF 21</td>
</tr>
<tr>
<td>36936</td>
<td>0</td>
<td>28 (18, 39)</td>
<td>+</td>
<td>ORF 22</td>
</tr>
<tr>
<td>36936</td>
<td>39 (60, 18)</td>
<td>228 (195, 260)</td>
<td>-</td>
<td>ORF 23</td>
</tr>
<tr>
<td>37023</td>
<td>68 (93, 44)</td>
<td>119 (99, 139)</td>
<td>+</td>
<td>ORF 22</td>
</tr>
<tr>
<td>37029</td>
<td>0</td>
<td>6 (6, 6)</td>
<td>-</td>
<td>ORF 23</td>
</tr>
<tr>
<td>39039</td>
<td>6 (6, 5)</td>
<td>6 (6, 6)</td>
<td>+</td>
<td>EGR 13</td>
</tr>
<tr>
<td>44726</td>
<td>0</td>
<td>14 (3, 24)</td>
<td>+</td>
<td>ORF 25</td>
</tr>
<tr>
<td>45541</td>
<td>0</td>
<td>15 (11, 19)</td>
<td>+</td>
<td>ORF 26</td>
</tr>
<tr>
<td>46402</td>
<td>30 (36, 24)</td>
<td>254 (191, 317)</td>
<td>+</td>
<td>ORF 28</td>
</tr>
<tr>
<td>49667</td>
<td>0</td>
<td>18 (11, 26)</td>
<td>+</td>
<td>ORF 32</td>
</tr>
<tr>
<td>50905</td>
<td>22 (22, 21)</td>
<td>46 (47, 44)</td>
<td>+</td>
<td>EGR 16</td>
</tr>
<tr>
<td>55759</td>
<td>0</td>
<td>11 (15, 7)</td>
<td>+</td>
<td>ORF 37</td>
</tr>
<tr>
<td>55775</td>
<td>82 (110, 55)</td>
<td>7769 (9263, 6275)</td>
<td>-</td>
<td>ORF 39</td>
</tr>
<tr>
<td>55815</td>
<td>0</td>
<td>102 (63, 142)</td>
<td>+</td>
<td>ORF 38</td>
</tr>
<tr>
<td>58881</td>
<td>5 (6, 4)</td>
<td>0</td>
<td>+</td>
<td>ORF 40</td>
</tr>
</tbody>
</table>
Table 2.3 (cont): 3' ends of MHV68 transcripts identified by RNA sequencing

<table>
<thead>
<tr>
<th>Position</th>
<th>Number of Reads at 6 hpi</th>
<th>Number of Reads at 18 hpi</th>
<th>Strand</th>
<th>Gene Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>63621</td>
<td>104 (137, 71)</td>
<td>612 (679, 544)</td>
<td>-</td>
<td>ORF 45</td>
</tr>
<tr>
<td>65405</td>
<td>45 (57, 33)</td>
<td>164 (161, 166)</td>
<td>-</td>
<td>ORF 48</td>
</tr>
<tr>
<td>66116</td>
<td>5 (5, 5)</td>
<td>0</td>
<td>-</td>
<td>ORF 49</td>
</tr>
<tr>
<td>69462</td>
<td>8 (10, 7)</td>
<td>40 (11, 70)</td>
<td>+</td>
<td>ORF 50</td>
</tr>
<tr>
<td>70937</td>
<td>206 (303, 109)</td>
<td>2544 (1775, 3313)</td>
<td>-</td>
<td>ORF 52</td>
</tr>
<tr>
<td>70938</td>
<td>38 (56, 19)</td>
<td>230 (217, 243)</td>
<td>+</td>
<td>M7</td>
</tr>
<tr>
<td>71293</td>
<td>0</td>
<td>10 (10, 10)</td>
<td>-</td>
<td>ORF 53-2</td>
</tr>
<tr>
<td>72561</td>
<td>16 (25, 6)</td>
<td>40 (17, 62)</td>
<td>-</td>
<td>ORF 55</td>
</tr>
<tr>
<td>72764</td>
<td>52 (69, 35)</td>
<td>18 (2, 34)</td>
<td>+</td>
<td>ORF 54</td>
</tr>
<tr>
<td>76559</td>
<td>6 (7, 4)</td>
<td>0</td>
<td>+</td>
<td>M8</td>
</tr>
<tr>
<td>76613</td>
<td>4 (7, 2)</td>
<td>12 (10, 15)</td>
<td>+</td>
<td>M8</td>
</tr>
<tr>
<td>76801</td>
<td>18 (26, 11)</td>
<td>0</td>
<td>+</td>
<td>M8</td>
</tr>
<tr>
<td>77186</td>
<td>0</td>
<td>100 (96, 104)</td>
<td>-</td>
<td>ORF 58</td>
</tr>
<tr>
<td>77187</td>
<td>86 (107, 66)</td>
<td>62 (19, 105)</td>
<td>+</td>
<td>ORF 57</td>
</tr>
<tr>
<td>77192</td>
<td>114 (152, 77)</td>
<td>0</td>
<td>-</td>
<td>ORF 58</td>
</tr>
<tr>
<td>77922</td>
<td>784 (1099, 468)</td>
<td>374 (240, 509)</td>
<td>-</td>
<td>ORF 59</td>
</tr>
<tr>
<td>80483</td>
<td>801 (998, 604)</td>
<td>214 (146, 283)</td>
<td>-</td>
<td>ORF 61</td>
</tr>
<tr>
<td>85207</td>
<td>25 (26, 24)</td>
<td>956 (1437, 475)</td>
<td>-</td>
<td>EGR 26</td>
</tr>
<tr>
<td>93842</td>
<td>528 (664, 391)</td>
<td>6305 (5254, 7356)</td>
<td>-</td>
<td>M9</td>
</tr>
<tr>
<td>93938</td>
<td>16 (24, 9)</td>
<td>14 (10, 17)</td>
<td>+</td>
<td>ORF 64</td>
</tr>
<tr>
<td>95322</td>
<td>30 (38, 22)</td>
<td>54 (39, 69)</td>
<td>-</td>
<td>ORF 67</td>
</tr>
<tr>
<td>95758</td>
<td>0</td>
<td>5 (6, 4)</td>
<td>-</td>
<td>ORF 67</td>
</tr>
<tr>
<td>95895</td>
<td>5 (6, 4)</td>
<td>0</td>
<td>-</td>
<td>ORF 67</td>
</tr>
<tr>
<td>98510</td>
<td>148 (233, 63)</td>
<td>260 (109, 410)</td>
<td>-</td>
<td>EGR 27</td>
</tr>
<tr>
<td>98953</td>
<td>43 (59, 27)</td>
<td>190 (115, 266)</td>
<td>+</td>
<td>ORF 69</td>
</tr>
<tr>
<td>102412</td>
<td>0</td>
<td>17 (8, 26)</td>
<td>-</td>
<td>ORF 72</td>
</tr>
<tr>
<td>103787</td>
<td>10 (11, 8)</td>
<td>4 (3, 6)</td>
<td>-</td>
<td>ORF 73</td>
</tr>
<tr>
<td>105812</td>
<td>216 (296, 136)</td>
<td>849 (672, 1026)</td>
<td>-</td>
<td>ORF 75c</td>
</tr>
<tr>
<td>106644</td>
<td>5 (8, 2)</td>
<td>56 (63, 49)</td>
<td>+</td>
<td>ORF 74</td>
</tr>
<tr>
<td>117906</td>
<td>0</td>
<td>16 (9, 23)</td>
<td>+</td>
<td>EGR 30</td>
</tr>
</tbody>
</table>
Chapter 3:

Stat1 regulates transcription of the murine gammaherpesvirus 68 latent-to-lytic switch gene, ORF 50

In addition to unpublished data, portions of this chapter are adapted from:

3.1 Abstract

Interferons (IFN) are critical regulators of chronic murine gammaherpesvirus 68 infection, suppress reactivation from latency, and are elevated in the serum of latently infected mice. In this chapter, we defined a role for IFNγ and Stat1 in suppressing viral replication in bone marrow-derived macrophages and transcription of the latent-to-lytic switch gene, ORF50. We found that IFNγ decreased activity of the two known ORF 50 promoters and identified a role of the key IFN signaling molecule, Stat1, in regulating the promoters of the latent-to-lytic switch gene, ORF 50. The suppressive activity of IFNγ was independent of canonical Stat binding sites. Promoter studies in Stat1-, IFNβ- and IFNγR-deficient macrophages in the absence of IFNγ suggested that type I IFNs also repressed ORF 50 promoter activity. Together, these data suggest that gammaherpesvirus 68 ORF 50 promoters have evolved elements responsive to Stat1 to control the transition between lytic and latent life cycles.
3.2 Introduction

Herpesviruses undergo acute infection followed by the establishment of a life-long latent infection during which the virus can periodically reactivate. Lytic gene expression following reactivation is thought to be an important component of pathology caused by herpesviruses (16, 21, 37, 46, 47). Both host and viral factors are critical for control of viral latency and reactivation. Murine gammaherpesvirus 68 (MHV68, \( \beta \)HV-68, MuHV4) is a rodent virus that provides a tractable model system for the oncogenic human gammaherpesviruses, Epstein-Barr Virus (EBV) and Kaposi's sarcoma associated herpesvirus (KSHV). Type I interferons (IFNs) are important to control acute MHV68 infection; infection of IFN\( \alpha/\beta \)R\(-/-\) mice leads to elevated viral titers in the lungs and dose-dependent mortality (14). While IFN\( \gamma \) is dispensable for control of acute MHV68 infection in C57BL/6 mice, it is required in BALB/c mice; IFN\( \gamma \)R\(-/-\) mice on the BALB/c background succumb to an acute bronchopneumonia (30, 49). For MHV68, both type I and type II interferons (IFNs) are essential for regulating reactivation from latency, whereas type II (IFN\( \gamma \)), but not type I, IFNs are important for controlling persistent viral replication and MHV68-induced vasculitis (1, 10, 15, 51, 54, 59).

While both type I and type II IFNs can signal through STAT1, type I IFNs can also activate STAT2, STAT3, and STAT5 to form homodimers or heterodimers (reviewed in (43)). Classically, type I IFNs induce STAT1-STAT2 heterodimers which form a complex with interferon regulatory factor (IRF) 9 and binds to IFN-stimulated response elements (ISRE) (reviewed in (43)). IFN\( \gamma \) signaling induces STAT1 phosphorylation, dimerization of STAT1 and translocation to the nucleus to transactivate promoter elements (44). The tyrosine at position 701 is required for STAT1 phosphorylation and dimerization (50). IFN\( \gamma \) activation of gene
transcription is primarily mediated by STAT1 homodimers binding to IFN-activated site (GAS) elements (11). Type I IFN-stimulated complexes can also bind to GAS elements (43).

The main effectors of type I IFNs are IRF3 and IRF7 (reviewed in (23)). Interestingly, IRF3 and IRF7 have been found to interact with gammaherpesvirus proteins. KSHV vIRF3, EBV BGLF4, and MHV68 ORF 36 all bind cellular IRF3 (24, 35, 56). IRF7 interacts with and regulates EBV proteins, LMP-1 and EBNA-1, as well as KSHV ORF 57 transactivation by RTA (40, 57, 66, 67). Interestingly, IRF7 is also targeted for degradation by RTA, suggesting that tight control of its levels during gammaherpesvirus infection is important (64). The key effectors of IFNγ are IRF1 and CIITA (reviewed in (3)). KSHV encodes a vIRF1 which inhibits type I IFN and acetyltransferase activity (6, 31, 32). Less is known about how IFNs and STAT1 suppress gene transcription.

The molecular mechanism by which IFNγ controls viral replication and reactivation is unknown. We found that IFNγ could suppress viral replication of MHV68 in macrophages in vitro and reduced the abundance of transcripts of the key latent-to-lytic switch gene, ORF 50, which is essential for viral replication (17, 42). IFNγ-mediated suppression of viral replication in macrophages was dependent on its receptor and Stat1, but was independent of Irf1, CIITA, Irf3 and Irf7 (17). To determine whether IFNγ reduced ORF 50 transcript levels by acting directly on the known promoters, we evaluated ORF 50 promoter activity by luciferase assay in the presence of varying doses of IFNγ. We found that IFNγ regulated promoter activity of the known ORF 50 promoters in a dose- and Stat1-dependent manner, but that predicted GAS elements in the ORF 50 promoters were dispensable. We also identified a role for Stat1 in the regulation of ORF 50 promoter activity in the absence of IFNγ. This activity was dependent on the IFNγ receptor and IFNγ cytokine, suggesting a role for type I IFNs in controlling ORF 50
promoter activity and viral replication in macrophages. Taken together, these results demonstrate a molecular mechanism by which IFNγ and type I IFNs control ORF 50 gene expression and suggest that the virus has maintained Stat1-repressible elements perhaps to aid in the delicate balance necessary to preserve chronic infection in a long-lived host.
3.3 Results

3.3.1 IFNγ suppresses viral replication in bone marrow-derived macrophages.

To determine the role of IFNγ in regulating viral replication in macrophages, we infected bone marrow-derived macrophages from wild-type C57BL/6 mice as well as mice deficient in various IFN signaling molecules in the presence or absence of IFNγ and assessed viral growth at various times post infection. IFNγ significantly inhibited viral replication at 24 hours post infection (hpi), 48 hpi and 72 hpi (Fig. 3.1). As expected, IFNγ-mediated suppression of viral replication was dependent on its receptor and the key signaling molecule, Stat1. IFNγ-mediated suppression of viral replication was independent of the type II IFN effectors, Irf1 or CIITA, as well as the IFNα/βR and the type I IFN effectors, Irf3 or Irf7 (Fig. 3.1).

3.3.2 Phosphorylation of Stat1 is required for IFNγ-mediated suppression of viral replication in bone marrow-derived macrophages.

To determine whether phosphorylation of Stat1 was necessary for IFNγ to regulate viral replication, we evaluated the effect of IFNγ on viral replication in bone marrow macrophages derived from mice in which the tyrosine at position 701 in Stat1 was mutated to a phenylalanine (Y701F). IFNγ was unable to suppress viral replication in these cells (Fig. 3.2), suggesting that canonical IFNγ signaling, involving Stat1 phosphorylation and dimerization, is required to regulate viral replication.

3.3.3 IFNγ-mediated regulation of ORF 50 transcript levels is dependent on Stat1.

To determine whether IFNγ regulated the lytic switch gene RTA encoded by ORF 50, we evaluated ORF 50 transcript levels by quantitative reverse transcription-PCR (qRT-PCR) at 12
hpi, a time point at which no difference in viral titer was observed in the presence or absence of IFNγ (Fig. 3.1). ORF 50 transcript levels were significantly reduced by IFNγ at 12 hpi (Fig. 3.3A). Importantly, ORF 73 transcript levels were unchanged in the presence of IFNγ, suggesting that early events in virus binding and entry were unaffected by IFNγ treatment (Fig. 3.3B). As expected, IFNγ-mediated suppression of ORF 50 RNA was dependent on Stat1 (Fig. 3.3A). Moreover, in the absence of IFNγ, ORF 50 transcript levels were elevated in Stat1-deficient macrophages, suggesting a role for Stat1 in basal or IFNα/β-mediated regulation of ORF 50 transcript expression (Fig. 3.3A).

### 3.3.4 IFNγ suppresses ORF 50 promoter activity.

To determine whether IFNγ regulates ORF 50 transcript levels by acting directly on ORF 50 promoters, we assessed the activity of the known ORF 50 promoters in the presence of IFNγ using a luciferase assay. We found that IFNγ reduced promoter activity of the two known promoters, a 410 bp promoter for the primary ORF 50 transcript (33) and a 250 bp promoter for an alternative ORF 50 transcript (20) in a dose dependent fashion (Fig. 3.4). At 10 units/ml, IFNγ decreased the 410 bp promoter activity ~32-fold and the 250 bp promoter ~15-fold. These data suggest that IFNγ regulated ORF 50 expression by directly suppressing its promoters.

### 3.3.5 Stat1 is required for IFNγ-mediated regulation of ORF 50 promoter activity.

Much of the activity of IFNγ is mediated by Stat1 signaling (44). To determine whether Stat1 was required for IFNγ-mediated suppression of ORF 50 promoter activity, we evaluated ORF 50 promoter activity in Stat1-deficient macrophages in the presence or absence of IFNγ. We found that IFNγ did not suppress the activity of either ORF 50 promoters in Stat1-deficient...
macrophages, demonstrating that Stat1 was necessary for IFNγ to suppress ORF 50 promoter activity (Fig. 3.5A-B).

3.3.6 IFNγ-mediated suppression of ORF 50 promoter activity is independent of predicted GAS elements.

We used MatInspector to predict potential Stat1 binding elements within the ORF 50 promoters (Fig. 3.6A). To determine whether these predicted elements in the ORF 50 promoters were necessary for IFNγ's action, we assessed ORF 50 promoter activity in luciferase vectors in which we had mutated these elements to sequences known to disrupt Stat1 binding (12, 19, 27). We found that IFNγ suppressed ORF 50 promoter activity even when GAS elements were mutated (Fig. 3.6B-C). To define the location of IFNγ-responsive elements in the core ORF50 promoter, we used luciferase reporter constructs containing fragments of the 410 bp ORF50 promoter (33). We found that IFNγ responsive elements in the 410 bp ORF50 promoter resided within the minimal promoter region contained within nucleotides 66442–66552 (Fig. 3.6D). These results suggest that IFNγ mediated its suppressive effect on ORF 50 promoter activity through potentially novel Stat1 binding elements or via an additional transcription factor.

3.3.7 IFNγ-independent role for Stat1 in ORF 50 promoter activity.

In addition to the role of Stat1 in IFNγ-mediated suppression of ORF 50 promoter activity, we observed that ORF 50 promoter activity was elevated in Stat1-deficient macrophages (Fig. 3.7). To determine whether the role of Stat1 in modulating ORF 50 promoter activity in the absence of IFNγ was due to type I IFN signaling or basal regulation by Stat1, we assessed ORF 50 promoter activity in macrophages deficient in the IFNγ receptor (IFNγR) or the cytokine,
IFNβ. ORF 50 promoter activity was enhanced in macrophages deficient in Stat1, IFNα/βR, or IFNβ compared to wild type macrophages (Fig 7), suggesting that Stat1 regulation of ORF 50 promoter activity in the absence of IFNγ was due to type I IFN signaling.
3.4 Discussion

In this chapter, we describe the role of Stat1 in controlling ORF 50 expression. We found that Stat1 was required to control ORF 50 promoter activity in both the presence and absence of IFNγ. The suppressive activity of IFNγ in macrophages was dependent on Stat1 phosphorylation, but not predicted GAS elements within ORF 50 promoters. These results indicate that Stat1 may regulate ORF 50 promoters by binding to a noncanonical motif or by inducing or repressing another host factor which in turn acts to suppress ORF 50 promoter activity. Stat1-mediated regulation of ORF 50 promoter activity in the absence of IFNγ was likely mediated through type I IFN signaling through Stat1 rather than a basal effect of Stat1. This finding has recently been confirmed by others (61). Taken together, these results provide a molecular mechanism for IFN-mediated regulation of viral replication in vitro and suggest a mechanism for IFN-mediated control of viral reactivation in vivo.

3.4.1 Regulation of ORF 50 Expression.

As the latent-to-lytic switch gene, it is important that ORF 50 expression be tightly controlled. MHV68 is unable to establish latency when ORF 50 is constitutively expressed by driving gene expression from a constitutively active promoter, suggesting that elements within the ORF 50 promoter are important in regulating the transition between latent and lytic infection (38). ORF 50 regulation has been best studied to date in KSHV. Known activators of ORF 50 transcription include Oct-1, viral G protein-coupled receptor (ORF 74), XBP-1, HMGB1, Ras/MEF/ERK/Ets-1 pathway (4, 22, 48, 60, 63). Several known negative regulators of KSHV ORF 50 have also been identified: the transcription factors, NF-κB, RBP-Jκ, and YY1; cellular cohesins and the viral gene, latency-associated nuclear antigen (LANA), encoded by ORF 73 (5,
In MHV68, NF-κB inhibits lytic replication (5). We and others have identified a role for a complex of histone deacetylases (HDACs 3 and 4) with the nuclear receptor corepressor (NCoR) in regulating MHV68 ORF 50 expression (18, 62). Here, we have identified Stat1 as an additional negative regulator of MHV68 ORF 50 promoter activity. Whether there is any interaction between Stat1 and the HDAC-NCoR complex in ORF 50 regulation is currently unknown.

Similar to MHV68, IFN-γ is known to be important for controlling EBV infection. In case reports and small clinical studies, IFN-γ deficiency or IFN-γ polymorphisms that result in low production of IFN-γ are associated with a range of EBV-associated malignancies, including nasopharyngeal carcinoma, post-transplant lymphoproliferative disease, B-cell non-Hodgkin lymphoma (2, 13, 55, 65). Interestingly, while IFN-γ negatively regulates MHV68 ORF 50 expression and lytic replication, this cytokine has been found to induce infectious virus production from latently infected KSHV BCBL-1 cells (7, 39). How IFN-γ has opposing roles in regulating ORF 50 in KSHV and MHV68 and whether IFN-γ induces replication of KSHV by acting on the ORF 50 promoters has not been studied to date. Our studies in MHV68 suggest that herpesviruses have evolved promoter elements responsive to inflammatory cytokines in order to control the delicate interchange between latent and lytic life cycles.

### 3.4.2 Mechanisms of Stat1 regulation of MHV68.

We have demonstrated that Stat1 is a key molecule for IFN-mediated suppression of ORF 50 expression. However, mutating putative Stat1 binding elements in the known ORF 50 promoters did not alter IFN-γ-dependent promoter activity. Similarly, IFN-γ-mediated regulation of the bullous pemphigoid antigen gene was also found to be independent of GAS elements, but...
dependent on ISRE and IRF consensus sequences in the gene's promoter (26, 41, 52). These results suggest that Stat1 suppression of ORF 50 expression may be mediated by noncanonical binding elements or via a second effector, although IFNβ-dependent suppression of MHV68 replication is independent of the classic effector genes, Irf1 and CIITA (3, 17). In contrast, for type I IFN regulation of ORF 50, a recent report has implicated the type I IFN transcription factor, Irf3, in mediating type I IFN-induced Stat1 regulation of ORF 50 transcription (61).

While IFNγ primarily induces gene transcription, there are several examples of genes that are repressed by IFNγ, including genes involved in regulation of extracellular matrix, cell cycle or chemokine signaling (44, 45). In one well-studied example, Stat1 represses promoter activity of matrix metalloproteinase 9 (MMP-9) by binding and sequestering the coactivator cyclic AMP response element-binding protein-binding protein (CREB-binding protein or CBP) and p300 proteins (36). This results in reduced H3/H4 acetylation and RNA polymerase II recruitment to the promoter, thereby suppressing MMP-9 transcription (36). Whether Stat1 binds directly to the ORF 50 promoters through a novel inhibitory element or acts to sequester coactivators or recruit corepressors is unknown and will be the basis of future studies.

### 3.4.3 Implications of Stat1 regulation of ORF 50 for viral pathogenesis.

The role of Stat1 in mediating suppression of ORF 50 promoter activity by both type I and type II IFNs suggests that ORF 50 promoters have evolved IFN responsive elements to control viral replication and establish the delicate balance necessary to maintain chronic infection in the face of an active immune response (17, 61). It is intriguing to speculate that such elements would be crucial to this function and if mutated, would result in a virus that maintains high levels of persistent viral replication and would be limited in its ability to establish and sustain a latent
infection. It may well be that the virus is benefited by maintaining elements responsive to immune regulation such that it can establish a latent infection that the immune system is unable to clear. Whether the virus maintains elements to respond to other immune signals to promote ORF 50 expression and viral reactivation is the basis of ongoing research.
3.5 Materials and Methods

3.5.1 Cells, virus and viral assays

Primary bone marrow macrophages were derived as described (53) from C57BL/6, Stat1-/-, Irf1-/-, Irf3-/-, Irf7-/-, CIITA-/-, IFNδR-/-, IFNβ-/-, 129, or Stat1 Y701F mice from internal breeding colonies at Washington University in St. Louis or purchased from Jackson Laboratory (Bar Harbor, ME). Mice were handled in accordance with federal and Washington University regulations. Cells were infected with MHV68 clone WUMS (ATCC VR-1465) for 1 hour at 37°C at a multiplicity of infection (MOI) of 10. Viral passage and titer were performed on NIH 3T12 fibroblasts (ATCC CCL-164) as described (58) except that cells were overlaid with 2% methylcellulose in MEM supplemented with 5% FCS and 2 mM L-glutamine.

3.5.2 Quantitative reverse transcriptase-PCR (qRT-PCR)

Total RNA was isolated 12 hpi using TRIzol reagent (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. RNA was treated with DNase I (Ambion, Austin, TX) before reverse transcription using oligo(dT)12-18 and Superscript II (Invitrogen, Carlsbad, CA). qRT-PCR was performed with SYBR green (Invitrogen, Carlsbad, CA) and the following primer sequences: 5′-TGCCCCATGTTTGTGATG and 5′-TGTGGTCATGAGCCCTTCC for glyceraldehyde-3-phosphate dehydrogenase (GAPDH), 5′-AGAAACCCACAGCTCGCACTT and 5′-CAATATGCTGGACAGGCCTATC for gene 50, and 5′-CCAGAAGCTTTGTACTTTGTGATG and 5′-AATACCACAGCAGCGTAGAAGGT for
gene 73. Transcript levels were normalized to GAPDH within each sample and compared to untransfected cells using the ΔΔCt method, where Ct is the threshold cycle (34).

3.5.3 Reporter plasmids and transfections

DNA fragments of the 410 bp ORF50 promoter were cloned into the luciferase reporter vector, pGL2 (Promega, Madison, WI), as described previously (33). The 250 bp ORF50 promoter described in (20) was cloned into the pGL2 luciferase reporter vector. Point mutations were introduced to mutate predicted GAS sites using QuikChange site-directed mutagenesis (Stratagene, La Jolla, CA). Cesium chloride preparations of plasmids were transfected into bone marrow derived macrophages using an Amaxa Nucleofector II device (Lonza, Basel, CH).

3.5.4 Luciferase promoter assay

1 x 10^6 bone marrow derived macrophages were incubated with IFNγ for 12 hours prior to co-transfection with 2 µg of the relevant pGL2 vector and 2 µg of pUB/V5-His/lacZ (Life Technologies, Grand Island, NY). Luciferase activity was measured at 12 hours post-transfection with the luciferase assay system (Promega) and normalized to β-galactosidase activity. The value from cells transfected with empty pGL2 vector was subtracted for each condition.

3.5.5 Statistical analyses.

Data were analyzed statistically with Prism 5 software (GraphPad Software, La Jolla, CA). Data were analyzed by two-tailed t-test or one-way ANOVA with Dunnett's post-test as indicated. All significant differences are noted.
3.6 Acknowledgements

This work was supported by NIH grant R01 CA 96511 (to H.W.V.). S.P.C. was supported by an NIH F30 fellowship grant HL099019, and M.M.G. was supported by a National Science Foundation graduate research fellowship. IRF1, IRF3, and IRF7-deficient mice were kind gifts from Dr. Michael Diamond. STAT1 Y701F and IFNβ-deficient mice were kind gifts from Dr. Robert Schreiber. We thank Darren Kreamalmeyer for expert animal care and members of the Virgin lab for helpful discussions.

Thanks to Megan Goodwin for viral titer and qRT-PCR data (as noted) and collaboration on ORF 50 promoter analyses.
3.7 References


94


promotes viral persistence by inhibiting the IRF-3-mediated type I interferon response. Cell Host. Microbe. 5:166-178.


Figure 3.1: IFNγ-mediated suppression of lytic replication in bone marrow-derived macrophages is mediated by STAT1. Primary bone marrow-derived macrophages from the...
indicated genotypes were incubated in the presence or absence of 10 units/ml of IFNγ for 12 hours before infection with MHV68 (MOI = 10). After 1 hour of absorption, cells were washed and IFNγ was added back to treated cultures. At the indicated times postinfection, the viral titer was measured by plaque assay on 3T12 fibroblasts. The limit of detection was 50 plaque forming unit (PFU)/ml. Data were collected from 2-5 independent experiments and are presented as mean +/- standard error of the mean (SEM). Statistical analyses were performed by a Student’s t-test. *, p < 0.05; **, p < 0.01; ***, p < 0.001. Data generated by Dr. Megan Goodwin.
Figure 3.2: Stat1 Y701 residue is necessary for IFNγ-mediated suppression of lytic replication in bone marrow-derived macrophages. Bone marrow-derived macrophages from 129 wild-type mice (A) or Stat1 Y701F mutant mice (B) were incubated in the presence or absence of 10 units/ml of IFNγ for 12 hours before infection with MHV68 (MOI = 10). After 1 hour of absorption, cells were washed and IFNγ was added back to treated cultures. At the indicated times postinfection, the viral titer was measured by plaque assay on 3T12 fibroblasts. The limit of detection was 50 plaque forming unit (PFU)/ml. Data were collected from 2 independent experiments and are presented as mean +/- SEM. Statistical analyses were performed by a Student’s t-test. *, p < 0.05.
**Figure 3.3: IFNγ suppresses MHV68 lytic switch gene 50 in a Stat1-dependent manner.** (A) IFN-γ treatment decreased gene 50 transcript levels compared to medium alone in wild-type but not Stat1⁻/⁻ bone marrow-derived macrophages. (B) In contrast, IFN-γ did not alter gene 73 transcript levels. Day 10 bone marrow-derived macrophages were pretreated for 12 hours with 10 units/ml of IFN-γ and then infected with MHV68 (MOI=10). Total RNA isolated 12 hours postinfection was reverse transcribed using oligo(dT)12-18 and Superscript II. ORF 50, ORF 73, and GAPDH transcript levels were determined by qRT-PCR. Transcript levels were normalized to GAPDH within each sample. Data were collected from 4 to 7 independent experiments and calculated using the \( \Delta\Delta CT \) method. Data are presented as mean ± SEM, and statistical analyses were performed by a Student's \( t \) test. Data generated by Dr. Megan Goodwin.
Figure 3.4: IFNγ suppresses ORF 50 promoter activity. (A) Schematic of known ORF 50 transcripts and their promoters. Transcription of the predominant ORF 50 transcript is driven by a 410bp promoter located at 66242-66652 in the MHV68 genome. Transcription of the ORF 50 transcript including exon 0 is driven by a 250bp promoter located at 65667-65920. (B-C) Bone marrow derived macrophages from C56BL/6 mice were transiently transfected with 2 μg of a pGL2-luciferase vector containing the 410bp (B) or 250bp (C) ORF 50 promoter and 2 μg of a β-galactosidase reporter driven by the ubiquitin promoter following pretreatment for 12 hours with varying doses of IFNγ. Luciferase assay was measured at 12 hours post-transfection and normalized to β-galactosidase expression, and the value from cells transfected with empty pGL2 vector was subtracted for each condition. Data were pooled from 5 to 11 independent
experiments and are presented as mean +/- standard error of mean (SEM). Statistical analyses were performed by a Student's t-test. * p < 0.05; ** p < 0.01. RLU, relative light units; OD, optical density.
Figure 3.5: Stat1 is necessary for IFNγ-mediated suppression of ORF 50 promoter activity.

(A-B) Bone marrow derived macrophages from Stat1-deficient mice were transiently transfected with 2 μg of a pGL2-luciferase vector containing the 410bp (A) or 250bp (B) ORF 50 promoter and 2 μg of a β-galactosidase reporter driven by the ubiquitin promoter following pretreatment for 12 hours with 10 units/ml of IFNγ. Luciferase assay was measured at 12 hours post-transfection and normalized to β-galactosidase expression, and the value from cells transfected with empty pGL2 vector was subtracted for each condition. Data were pooled from 4 to 5 independent experiments and are presented as mean +/- SEM. Statistical analyses were performed by a Student’s t-test. RLU, relative light units; OD, optical density.
Predicted GAS Sites

410bp Promoter
nt 66356-66363  TTCTTGAA → TCCTTGGA
nt 66394-66403  TTCCCTGGAA → TCCAATGGGA

250bp Promoter
nt 65686-65694  TTCCTAAA → TCCATTAGA

B

C

D

<table>
<thead>
<tr>
<th>Reporter construct</th>
<th>Genome coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>66242-66652</td>
</tr>
<tr>
<td>2</td>
<td>66442-66552</td>
</tr>
<tr>
<td>3</td>
<td>66402-66552</td>
</tr>
<tr>
<td>4</td>
<td>66242-66552</td>
</tr>
</tbody>
</table>
Figure 3.6: IFNγ-mediated suppression of ORF 50 promoter activity is contained within the minimal 410bp promoter and does not require predicted Stat binding elements. (A) The nucleotide (nt) position and sequence of MatInspector-predicted GAS sites in the ORF 50 promoters are listed. Bold, underlined nucleotides indicate nucleotides mutated to disrupt the predicted sites. (B-C) Bone marrow derived macrophages from C56BL/6 mice were transiently transfected with 2 μg of a pGL2-luciferase vector containing the 410bp (B and D) or 250bp (C) ORF 50 promoter or truncations of the 410bp promoter (D) and 2 μg of a β-galactosidase reporter driven by the ubiquitin promoter following pretreatment for 12 hours with varying doses (B-C) or 10 units/ml (D) of IFNγ. Luciferase assay was measured at 12 hours post-transfection and normalized to β-galactosidase expression, and the value from cells transfected with empty pGL2 vector was subtracted for each condition. Data were pooled from 3 to 8 independent experiments and are presented as means +/- SEM. Statistical analyses were performed by a Student’s t-test. *, p < 0.05; **, p < 0.01. RLU, relative light units; OD, optical density.
Figure 3.7: Stat1 regulates ORF 50 promoter activity in the absence of IFNγ. Bone marrow derived macrophages from C56BL/6, Stat1-/-, IFNβ-/- or IFNαβR-/- mice were transiently transfected with 2 μg of a pGL2-luciferase vector containing the 410bp (A) or 250bp (B) ORF 50 promoter and 2 μg of a β-galactosidase reporter driven by the ubiquitin promoter. Luciferase assay was measured at 12 hours post-transfection and normalized to β-galactosidase expression, and the value from cells transfected with empty pGL2 vector was subtracted for each condition. Data were pooled from 3 to 4 independent experiments and are presented as means +/- SEM. Statistical analyses were performed by one-way analysis of variance (ANOVA) with Dunnett’s post-test compared to C57BL/6. *, p < 0.05; **, p < 0.01. RLU, relative light units; OD, optical density.
Chapter 4:

Latent gammaherpesvirus 68 infection induces distinct transcriptional changes in different organs

This chapter is adapted from:

*denotes equal contribution.
4.1 Abstract

Previous studies identified a role for latent herpesvirus infection in cross-protection to infection and exacerbation of chronic inflammatory diseases. To assess how latent viral infection alters host gene expression, we compared the transcriptional signature in spleens, livers, and brains from mice latently infected with gammaherpesvirus 68 to uninfected mice or mice infected with a virus that is defective in latency establishment. We identified over 600 genes that were differentially expressed in the tissues of latently infected mice and found distinct sets of genes linked to different pathways were altered in spleen compared to liver. Several of the most differentially expressed latency-specific genes (e.g. IFNγ, Cxcl9, Ccl5) are associated with known latency-specific phenotypes.
4.2 Text

Herpesviruses transit between lytic and latent life cycles and cause lifelong infection of their hosts. Following acute lytic infection of various cell types, herpesviruses establish latency, defined experimentally as the presence of viral genome without production of infectious virus (23). Humans are chronically infected with numerous viruses, including several herpesviruses (23). However, the effect of chronic viral infection on the host is not well understood. We and others showed that latent viral infection can protect hosts from lethal challenges with bacteria and lymphoma cells (1, 25, 27). This latency-induced cross protection is associated with activation of macrophages, production of the inflammatory cytokine, interferon-gamma (IFNγ), and arming of natural killer cells (1, 25). Latent murine gammaherpesvirus 68 (MHV68, γHV-68, MuHV-4) also modulates adenovirus and plasmodium infections (7, 13). MHV68 is implicated in exacerbating experimental autoimmune encephalomyelitis (EAE), a rodent model for multiple sclerosis (2, 14), as well as a mouse model of inflammatory bowel disease (12). Together, these studies are consistent with chronic or latent herpesvirus infection having substantial effects on host immune responses to unrelated antigens. This led us to determine whether latent herpesvirus infection induces distinct transcriptional changes in specific organs by analyzing microarray data from RNA extracted from spleens, livers, and brains of mice infected with MHV68 or a mutant MHV68 virus with a stop codon in open reading frame (ORF) 73 (ORF73.stop) that undergoes acute replication, but is defective in the establishment of latency (5, 11).
4.2.1 Chronic MHV68 alters the host transcriptional profile.

To determine whether chronic MHV68 infection alters the host transcriptional profile in various organs, we infected 7-9 week old male C57BL/6J mice intranasally with wild-type (WT) MHV68, ORF73.stop mutant virus or mockulum (mock) as previously described (1). After cardiac perfusion with PBS to diminish tissue numbers of peripheral blood cells, brains, livers and spleens were harvested at 28 days post infection (dpi), a time point at which we know that latency is established and at which cross-protection to bacteria is apparent (1, 21, 25, 27).

To identify the gene signature of latent infection, we analyzed the microarray data using two methods: (1) factorial analysis to identify genes with significant fold-change differences between WT MHV68 infected mice and ORF73.stop infected mice (after comparison of each to mock-infected mice) and (2) a two-way ANOVA to identify genes that were significantly differentially expressed due to type of tissue, virus, or the combination of these two factors (p-value < 0.05 for reported genes in each analysis). Differential expression analysis was performed in MATLAB. For both analyses, differentially expressed genes were identified by first comparing mock-infected mice to mice infected with ORF73.stop or WT MHV68. Next, we compared the gene expression profiles of differentially expressed genes from mice infected with ORF73.stop virus to mice infected with WT MHV68 to identify a set of genes specifically differentially expressed in mice that were infected with the latency-sufficient WT MHV68. We considered this set of genes to be specific to latent infection. 604 genes were identified by both methods as differentially expressed in virus-infected hosts and we considered this our high confidence gene set (Fig. 4.1). Of these 604 genes, 601 genes were specifically differentially expressed in latently infected animals and 3 genes were specifically differentially expressed in ORF73.stop infected mice.
Some genes are regulated with oscillations in the circadian rhythm. In fact, a recent study delineated a role for the circadian clock in expression and function of an innate immune response gene, Toll-like receptor 9, highlighting the importance of controlling for diurnal gene regulation in studies of infection and immunity (17). Since two of the liver and spleen data sets were consistently collected in either the morning or afternoon, we included a third data set for which all samples were collected within two hours. We also computationally tested whether circadian rhythm genes were present in our list of genes regulated by latent infection. Only 3 of 67 genes related to circadian rhythm in MSigDB were present in our list of latency-specific host genes, indicating that time of harvest did not significantly confound our results.

To confirm the reproducibility of the latency-specific differential gene expression that we observed, we compared our gene expression signature in the spleens of latently infected mice to a complementary data set of gene expression in spleens of latently infected mice, independently generated by another group (26). There were 309 genes in common between these two data sets (p-value = 1.8876e-24 by a hypergeometric test). Furthermore, we used a direct quantitative correlation computation to compare the data sets. Since the data sets were profiled on distinct platforms (Illumina vs Affymetrix), we identified the most similar probes on the two platforms using Smith-Waterman local alignment algorithm to compute a sequence nucleotide alignment score. After selecting matched pairs of probes, we evaluated the correlation between fold-change in expression as observed in the two data sets. The regression coefficient between the comparable sets of data was 0.67. Together, these analyses demonstrate a significant overlap between spleen data sets generated using distinct platforms at different institutions and indicate that the gene expression changes that we identified are robust. Our data extend these findings by analyzing gene expression in additional tissues, specifically liver and brain.
To further validate our microarray findings, we selected 10 genes that we identified as specifically differentially expressed in latently infected animals to confirm by quantitative real-time PCR (qRT-PCR, Fig. 4.2). Genes were selected to encompass a range of fold changes and differential expression patterns (some specific to spleen, some specific to liver and some differentially expressed in both tissues), and relationship to known pathways (e.g. IFNγ signaling) or diseases (e.g. genome wide association studies of multiple sclerosis or celiac disease) of interest. Our analyses predicted that basic leucine zipper transcription factor, ATF-like (Batf), IFNγ, and interleukin 18 receptor accessory protein (Il18rap) were upregulated in the spleens of latently infected mice; vascular cell adhesion molecular 1 (Vcam1), spleen tyrosine kinase (Syk) and pleckstrin (Plek) were upregulated in the livers of latently infected mice; interferon gamma induced GTPase (Igtp), interferon inducible GTPase 1 (Iigp1), and guanylate binding protein 1 (Gbp1) were upregulated in both livers and spleens while macrophage receptor with collagenous structure (Marco) was upregulated in livers, but downregulated in spleens. We confirmed latency-specific differential expression for 7 of 10 genes and 14 of 18 predictions. We found that Gbp1, Igtp, Iigp1, Marco, Syk and Vcam1, but not Batf or Plek, were differentially expressed in the livers of latently infected mice and that Gbp1, IFNγ, Igtp, Iigp1, Marco, and Syk, but not Batf, Il18rap, Plek or Vcam1, were differentially expressed in the spleens of latently infected mice (Fig. 4.2). These results demonstrate that our computational analyses specifically identified many genes differentially expressed in latently infected animals.

To identify the genes with the largest differential gene expression in latently-infected animals, we calculated the signal-to-noise ratio for each gene in each tissue by determining the difference between the mean fold change of MHV68 vs. mock and ORF73.stop vs. mock scaled by the variance between these (i.e. the sum of the standard deviations). For genes detected by
multiple probes, we averaged the fold change across the probes and then computed the signal-to-
noise ratio using the average fold change values. Genes were ranked in order of the absolute
value of the signal-to-noise ratio. We selected the top 50 genes identified in this signal-to-noise
analysis and identified those that overlapped with the genes detected as statistically significant in
our previous analyses. This conservative approach generated a set of genes (Tables 4.2-4.4) that
exhibit infection-related changes in expression which we then used as a substrate for considering
the biological pathways that might be altered by chronic viral infection in different organs.

4.2.2 Distinct pathways are regulated by latent viral infection in livers and spleens.

Using hierarchical clustering, we found that the differentially expressed genes covaried
by organ (Fig. 4.1). We identified a subset of genes that were primarily differentially expressed
in the liver and a separate subset of genes that were primarily differentially expressed in the
spleen with modest overlap between the two tissues. To determine the specific cellular pathways
altered in the organs of latently infected mice, we used MSigDB to evaluate pathway enrichment
by Gene Set Enrichment Analysis (19). To determine the pathways enriched by latent viral
infection, we used factorial analysis to determine the differentially expressed genes in each organ
in WT MHV68 infected mice compared to mock mice and then removed any genes that were
also differentially expressed in ORF73.stop-infected mice. The number of genes that were
differentially expressed in both WT MHV68 and ORF73.stop-infected mice compared to mock
was modest (65 for liver, 15 for spleen and 1 for brain). Of the 736 genes differentially expressed
in the spleens of latently infected mice, we found that most genes were associated with cell cycle
pathways (Fig. 4.3A). This is consistent with a previous report of differential gene expression in
the spleens of mice infected with MHV68, the betaherpesvirus, murine cytomegalovirus
(MCMV), or the human alphaherpesvirus, herpes simplex virus (HSV) (26). In contrast, of the 538 genes differentially expressed in the livers of latently infected mice, most differentially expressed genes were associated with immune signaling and immune response pathways, including interferon signaling (Fig. 4.3B). The immune-related gene expression changes detected in the liver were consistent with histological sections in which foci of inflammatory cells were visible from mice infected with WT MHV68, but not ORF73.stop virus or mock-infected mice, and likely reflected the immigration of inflammatory cells rather than changes in resident hepatocytes (data not shown). Because of the potential of diurnal gene regulation as a confounding factor and the fact that the circadian rhythm was identified as a significant pathway for differentially expressed genes in WT MHV68 livers when compared to livers from mock mice, we identified the circadian genes (Clock, Cpt1a, Elovl3, Per1 and Per2) that were differentially expressed in the livers of latently infected mice and repeated pathway analysis with these removed to test their effect on latency-specific pathways. The pathway of genes involved in platelet activation, signaling and aggregation was no longer statistically significant. The number of genes in the overlap with the pathway of genes involved in metabolism of lipids and lipoproteins decreased by three and the log of the q-value decreased to 4.0915 from 5.9586. All of the other pathways were insensitive to the removal of the circadian clock genes. Although there were few genes whose expression changed in the brains of latently infected mice (32), many of the genes altered in the brain were also differentially expressed in the liver or spleen of latently infected mice (see Fig. 4.1). The pathways significantly enriched in brains of latently infected mice were predominantly related to the immune system and interferon signaling (Fig. 4.3C). Finally, we tested what pathways were enriched across multiple tissues. We found that there were 109 genes that were differentially expressed in both livers and spleens of latently
infected mice. The immune system generally and interferon signaling specifically were the most significant pathways for these overlapping genes (data not shown). Taken together, these results demonstrate that latent viral infection differentially altered host gene expression specifically by tissue type, suggesting that there are distinct tissue responses to the host inflammatory response induced by chronic infection and that the modest number of common gene expression changes across tissues may be largely attributable to chronic interferon signaling (1).

4.2.3 Differentially expressed genes are significantly associated with genes regulated by the cytokine, interferon gamma.

IFNγ is required for controlling MHV68 reactivation from latency and persistent viral replication and is elevated in the serum of latently infected mice (1, 6, 18, 22). Previous work suggested that IFNγ is critical to latency-induced cross-protection to bacteria challenges and that CNS-infiltrating T cells in MHV68-exacerbated EAE produce elevated levels of IFNγ (1, 2). Our microarray pathway analysis also confirmed that the IFNγ pathway was one of the most significantly enriched pathways in livers and brains of latently infected mice. To confirm that host genes induced in the setting of chronic MHV68 infection are also regulated by IFNγ, we compared our list of differentially expressed genes in mice infected with WT MHV68 (relative to mock-infected mice) to a list of IFNγ-regulated genes in murine bone marrow-derived macrophages (10). We found that there was significant overlap between IFNγ-regulated genes in bone marrow-derived macrophages and tissues from MHV68-infected mice. Of the 359 genes regulated by IFNγ, 59 were regulated in the liver (p-value = 2.4390e-27), 72 were regulated in the spleens (p-value = 7.4127e-33) and 12 were regulated in the brain (p-value = 1.8151e-13) of latently infected mice (Fig. 4.4). Many of the most differentially expressed genes in all three
tissues from latently infected mice (Tables 4.2-4.4) are known to be induced by IFNγ, including Igtp, Igip1, Gbps, Ccl5, Cxcl9, and MHC molecules. IFNγ itself was significantly upregulated in the spleens of latently infected mice (Fig. 4.2 and Table 4.2), an organ in which significantly reduced *Listeria* and *Yersinia* titers are observed in latently-infected animals, suggesting a mechanism by which latently-infected mice resist lethal bacterial challenges (1). Interestingly, the key IFNγ signaling molecule, Stat1, was one of the most upregulated genes in both livers and spleens of latently-infected mice (Tables 4.2-4.3). This suggests that active interferon signaling was also present in the livers of latently-infected mice where *Listeria* titers are reduced (1). These results suggest that many of the differentially expressed genes in latently infected mice were a result of elevated IFNγ levels in the host which have been detected during chronic MHV68 infection and are mechanistically important to latency-induced cross-protection (1).

EAE exacerbation by latent MHV68 is not only associated with IFNγ, but also with increased inflammatory infiltration into the CNS and elevated serum levels of chemokines, Ccl5 and Cxcl9, which are known to attract T cells (2). Although the authors were unable to detect these chemokines in the CNS of latently infected mice, Ccl5 and Cxcl9 are significantly elevated in mouse sera from latently infected mice at the onset of EAE symptoms (2). Importantly, we detected Ccl5 and Cxcl9 as two of the most differentially expressed transcripts in the brains of latently infected mice (Table 4.4). These data suggest that elevated Ccl5 and Cxcl9 expression in the brain during latent infection may induce T cell migration into the CNS leading to augmented EAE symptoms.

Activated macrophages have been implicated in several models of latency-induced cross-protection, including challenges with bacteria and influenza virus (1, 16). Specifically, these models have identified activated macrophages as those expressing increased MHC class II
molecules. MHC class II molecules, specifically H2-Aa and H2-Eb1, were among the most highly differentially expressed genes in the livers and brains of latently-infected mice (Tables 4.3-4.4). These data confirm that latent MHV68 infection can induce upregulation of MHC class II molecules in multiple organs with implications for cross-protection to a range of infectious challenges.

Here, we show that infection with a chronic herpesvirus infection can profoundly alter gene expression in the host in a tissue-specific manner. The observed host transcriptional changes induced by latent infection were predominantly immune-related genes. However, the specific genes and pathways induced in latently infected animals differed by tissue type. Whether this was due to local differences in the inflammatory milieu or differential tissue responses to the same stimulus is unknown. However, given the differential effects of interferons and interferon-regulated molecules in different cell types (3, 4, 8, 15, 20, 24), it is reasonable to hypothesize that the differences observed here were the result of differential cell responses and cell composition in specific organs.

It is interesting to speculate what effects the gene expression changes induced by latent viral infection might have on the function of host tissues. Previous studies delineating a role for MHV68 in exacerbating EAE pathology (2, 14) in addition to the importance of IFNγ in EAE pathogenesis (9) and in the induction of many of the changes observed here, suggest that the gene expression changes induced in the setting of latent viral infection may alter the course of chronic disease pathogenesis. In a previous study of splenic gene expression changes in response to HSV, MCMV, or MHV68 infection, the authors identified virus-specific changes (26), suggesting that there may be important, dynamic and complex gene expression changes in a human host infected with multiple viruses. Taken together, these results demonstrate that chronic
viral infection can have profound and distinct effects on the host with potential implications for tissue function and pathogenesis of chronic disease in susceptible hosts.

Microarray data have been deposited in GeoArchive, series number GSE51365.

4.3 Acknowledgements

This work was supported by NIH grants CA096511 (to H.V.W.) and DK043351 (to R.X.). S.P.C. is supported by NIH F30 HL099019, and T.A.R. was supported by the Damon Runyon Cancer Research Foundation. We thank the Genome Technology Access Center in the Department of Genetics at Washington University School of Medicine for performing microarrays. The Center is partially supported by NCI Cancer Center Support Grant P30 CA91842 to the Siteman Cancer Center and by ICTS/CTSA Grant UL1 TR000448 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH), and NIH Roadmap for Medical Research. This publication is solely the responsibility of the authors and does not necessarily represent the official view of NCRR or NIH.

We thank Darren Kreamalmeyer for expert animal care, members of the Virgin lab for helpful discussions and review of the manuscript, Weixiong Zhang and Michael Stevens for initial discussion of the data, and Erik Barton for sharing data. Final computational analyses presented in this chapter were completed by Gautam Goel, PhD of the Broad Institute. Thanks to Xin Zhang for assistance with some qRT-PCR analysis.
4.4 References


Figure 4.1: Latent MHV68 alters host gene expression. Male C57Bl/6 mice were infected intranasally with MHV68, ORF73.stop mutant virus (ORF73) or mockulum (mock) and spleens, livers and brains (3 experiments) were harvested at 28 d.p.i. RNA was extracted using the ToTALLY RNA kit (Ambion) from tissues incubated in RNAlater-ICE (Ambion). Equal amounts of RNA were pooled from three to four mice per condition for each microarray chip. RNA was labeled and hybridized to Affymetrix M430 2.0 microarrays at the Multiplexed Gene Analysis Core Facility at Washington University in St. Louis. Data were normalized using Robust Multi-array Average normalization routine in Affymetrix® Expression Console software. Microarray data were analyzed by both factorial analysis and two-way ANOVA. Differentially expressed genes identified by both analyses are shown in the heat map after hierarchical clustering of the genes. The heat map represents the fold change relative to mock-infected mice with red representing the row maximum and blue the row minimum value. Statistical significance was performed by standard student's t-test in the case of factorial analysis. P-values were corrected for multiple hypothesis testing. Gene expression differences were considered significant if the adjusted p-value < 0.05 and > 1.5-fold change was observed in expression levels. Mice were handled in accordance with federal and Washington University regulations. Figure generated by Dr. Gautam Goel
Figure 4.2: Validation of microarray analyses by qRT-PCR. RNA pooled from three to four mice per condition was treated with TURBO DNase (Life Technologies) before cDNA synthesis using 1 ug RNA, oligo(dT)12-18 and Superscript II reverse transcriptase (Life Technologies). qPCR was performed on cDNA from livers (A) or spleens (B) using Power SYBR Green Master Mix (Applied Biosystems) with the indicated gene specific primers (Table 4.1). Data were collected from 4 independent experiments. Transcript levels were normalized to the level of GAPDH within each sample and compared to the level in mock-infected mice using the \( \Delta\Delta C_T \) method (where \( C_T \) is the threshold cycle). Data are presented as mean ± S.E.M. Statistical
analyses were performed by Mann-Whitney U test. * p<0.05; ns = not significant; nd = could not reliably detect target.
Figure 4.3: Pathway analysis of differentially expressed genes in latently infected mice.

Gene overlap analyses of differentially expressed genes in spleens (A), livers (B), or brains (C)
of MHV68-infected mice. A list of latent infection-specific differentially expressed genes was identified by comparing MHV68-infected mice to mock-infected mice and then removing genes that overlapped with those differentially expressed from ORF73.stop-infected mice compared to mock-infected mice. Pathways were identified using MSigDB’s Canonical Pathway gene set. The 50 most significant pathways for spleens (A) and livers (B) are displayed. The x-axis represents the q-value, indicating the significance of enrichment for a given gene set. The values are plotted on a negative log_{10} scale. Gene sets with values of > 1.3 (q-value < 0.05) were significantly enriched. Figure generated by Dr. Gautam Goel.
Figure 4.4: Overlap of genes regulated by IFNγ and latent viral infection. Heat map of differentially expressed genes in livers, spleens, or brains from MHV68-infected mice (compared to mock-infected mice) that significantly overlapped with genes altered by IFNγ in bone marrow-derived macrophages. The heat map displays the fold change relative to mock-infected mice with red representing the row maximum and blue the row minimum value. Statistical significance was assessed by a hypergeometric test. Nominal p-values were used. Figure generated by Dr. Gautam Goel.
Table 4.1: Primers used in this study for qRT-PCR

<table>
<thead>
<tr>
<th>Target</th>
<th>F primer</th>
<th>R primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Batf</td>
<td>TCAAGAAGGGAGCCAGCTA</td>
<td>GCAGCGATGCGATTCTTCTC</td>
</tr>
<tr>
<td>Gbp1</td>
<td>AAGAAGTGAAAGCAGGGACA</td>
<td>GCGGCTTCTGCTTTTATACG</td>
</tr>
<tr>
<td>IFNγ</td>
<td>CCTGCGGCCTAGCTCTGA</td>
<td>TGGCAGTAACAGCCAGAAACAG</td>
</tr>
<tr>
<td>Igtp</td>
<td>CCTGTTTTTAGCCCCGTCTTTTCACG</td>
<td>TCCTTCAAGAGAGACGCCCTATGCT</td>
</tr>
<tr>
<td>Igp1</td>
<td>CATTGAAAAGAAGCGGAATT</td>
<td>CAGTGCGGTAGAATTTCATGCT</td>
</tr>
<tr>
<td>Il18rap</td>
<td>CAGATCATCTCTACCGGCCTC</td>
<td>AAACACCCAGCCCAAACAGA</td>
</tr>
<tr>
<td>Marco</td>
<td>TCATGGCCATCCACCTGATC</td>
<td>GCCCCACTGCAGCGAGAAAGAA</td>
</tr>
<tr>
<td>Plek</td>
<td>TCCTGGATAGCCCTGATGCT</td>
<td>TCACCCCCAGCAGGATCATA</td>
</tr>
<tr>
<td>Syk</td>
<td>GCCTTGGTTAAAGTGCAGACTTG</td>
<td>GGCTATGCAGCGCTTTAGGCTTT</td>
</tr>
<tr>
<td>Vcam1</td>
<td>CTGGGAAGGCTGGAACGAAGT</td>
<td>GCCAAACACTTGACCGTGAC</td>
</tr>
<tr>
<td>GAPDH</td>
<td>TGCCCCCATGTGTTAGATG</td>
<td>TGTGGTCATGAGCCCTTCC</td>
</tr>
<tr>
<td>Gene ID</td>
<td>Gene Name</td>
<td>Signal to Noise Ratio</td>
</tr>
<tr>
<td>---------</td>
<td>---------------------------------------------------------------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>CXCL9</td>
<td>chemokine (C-X-C motif) ligand 9</td>
<td>22.9176</td>
</tr>
<tr>
<td>SAA3</td>
<td>serum amyloid A 3</td>
<td>14.5842</td>
</tr>
<tr>
<td>TTC39C</td>
<td>tetratricopeptide repeat domain 39C</td>
<td>10.3055</td>
</tr>
<tr>
<td>GBP1</td>
<td>guanylhylate binding protein 1, interferon-inducible</td>
<td>10.1534</td>
</tr>
<tr>
<td>IGTP</td>
<td>interferon gamma induced GTPase</td>
<td>10.0122</td>
</tr>
<tr>
<td>TIMP1</td>
<td>tissue inhibitor of metalloproteinase 1</td>
<td>9.8577</td>
</tr>
<tr>
<td>GZMB</td>
<td>granzyme B</td>
<td>8.3171</td>
</tr>
<tr>
<td>MELK</td>
<td>maternal embryonic leucine zipper kinase</td>
<td>7.8645</td>
</tr>
<tr>
<td>INCENP</td>
<td>inner centromere protein</td>
<td>7.499</td>
</tr>
<tr>
<td>GPNMB</td>
<td>glycoprotein (transmembrane) nmb</td>
<td>7.1763</td>
</tr>
<tr>
<td>AA467197</td>
<td>expressed sequence AA467197</td>
<td>6.9617</td>
</tr>
<tr>
<td>OSBPL3</td>
<td>oxysterol binding protein-like 3</td>
<td>6.7681</td>
</tr>
<tr>
<td>MKI67</td>
<td>antigen identified by monoclonal antibody Ki 67</td>
<td>6.5786</td>
</tr>
<tr>
<td>FABP5</td>
<td>fatty acid binding protein 5, epidermal</td>
<td>6.5125</td>
</tr>
<tr>
<td>IFNG</td>
<td>interferon gamma</td>
<td>6.4947</td>
</tr>
<tr>
<td>RRM2</td>
<td>ribonucleotide reductase M2</td>
<td>5.999</td>
</tr>
<tr>
<td>MARCO</td>
<td>macrophage receptor with collagenous structure</td>
<td>-5.7155</td>
</tr>
<tr>
<td>GBP2</td>
<td>guanylhylate binding protein 2</td>
<td>5.6835</td>
</tr>
<tr>
<td>2410003J06RIK</td>
<td>Mageb16 melanoma antigen family B, 16</td>
<td>5.6572</td>
</tr>
<tr>
<td>XDH</td>
<td>xanthine dehydrogenase</td>
<td>5.5757</td>
</tr>
<tr>
<td>STX3</td>
<td>syntaxin 3</td>
<td>5.2772</td>
</tr>
<tr>
<td>BHLHE40</td>
<td>basic helix-loop-helix family, member e40</td>
<td>5.2416</td>
</tr>
<tr>
<td>IL1RL1</td>
<td>interleukin 1 receptor-like 1</td>
<td>5.1771</td>
</tr>
<tr>
<td>CDC45L</td>
<td>cell division cycle 45</td>
<td>5.1685</td>
</tr>
<tr>
<td>TRAFD1</td>
<td>TRAF type zinc finger domain containing 1</td>
<td>5.126</td>
</tr>
<tr>
<td>TAP1</td>
<td>transporter 1, ATP-binding cassette, sub-family B (MDR/TAP)</td>
<td>5.1003</td>
</tr>
<tr>
<td>PARP9</td>
<td>poly (ADP-ribose) polymerase family, member 9</td>
<td>4.9727</td>
</tr>
<tr>
<td>CDC6</td>
<td>cell division cycle 6</td>
<td>4.9208</td>
</tr>
<tr>
<td>IIGP1</td>
<td>interferon inducible GTPase 1</td>
<td>4.9031</td>
</tr>
<tr>
<td>IRG1</td>
<td>immunoresponsive gene 1</td>
<td>4.8747</td>
</tr>
<tr>
<td>PDCD1</td>
<td>programmed cell death 1</td>
<td>4.8335</td>
</tr>
<tr>
<td>RACGAP1</td>
<td>Rac GTPase-activating protein 1</td>
<td>4.7826</td>
</tr>
<tr>
<td>HK2</td>
<td>hexokinase 2</td>
<td>4.766</td>
</tr>
<tr>
<td>CDKN1A</td>
<td>cyclin-dependent kinase inhibitor 1A (P21)</td>
<td>4.6531</td>
</tr>
<tr>
<td>KIF4</td>
<td>kinesin family member 4</td>
<td>4.6377</td>
</tr>
<tr>
<td>UBAC1</td>
<td>ubiquitin associated domain containing 1</td>
<td>4.6097</td>
</tr>
<tr>
<td>UPP1</td>
<td>uridine phosphorylase 1</td>
<td>4.5889</td>
</tr>
<tr>
<td>DTX1</td>
<td>deltx 1 homolog</td>
<td>-4.5719</td>
</tr>
<tr>
<td>CLEC4D</td>
<td>C-type lectin domain family 4, member d</td>
<td>4.5384</td>
</tr>
</tbody>
</table>
Table 4.3: Genes most differentially expressed in livers of latently-infected mice

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Gene Name</th>
<th>Signal to Noise Ratio</th>
<th>Up or Down relative to ORF73.stop</th>
</tr>
</thead>
<tbody>
<tr>
<td>IGTP</td>
<td>interferon gamma induced GTPase</td>
<td>7.3767</td>
<td>Up</td>
</tr>
<tr>
<td>MS4A4B</td>
<td>membrane-spanning 4-domains, subfamily A, member 4B</td>
<td>6.1604</td>
<td>Up</td>
</tr>
<tr>
<td>IIIGP1</td>
<td>interferon inducible GTPase 1</td>
<td>5.726</td>
<td>Up</td>
</tr>
<tr>
<td>STAT1</td>
<td>signal transducer and activator of transcription 1</td>
<td>5.4494</td>
<td>Up</td>
</tr>
<tr>
<td>PTPN18</td>
<td>protein tyrosine phosphatase, non-receptor type 18</td>
<td>5.4011</td>
<td>Up</td>
</tr>
<tr>
<td>FYB</td>
<td>FYN binding protein</td>
<td>5.0647</td>
<td>Up</td>
</tr>
<tr>
<td>H2-Q7</td>
<td>histocompatibility 2, Q region locus 7</td>
<td>5.013</td>
<td>Up</td>
</tr>
<tr>
<td>LY6A</td>
<td>lymphocyte antigen 6 complex, locus A</td>
<td>4.7926</td>
<td>Up</td>
</tr>
<tr>
<td>LCK</td>
<td>lymphocyte protein tyrosine kinase</td>
<td>4.7378</td>
<td>Up</td>
</tr>
<tr>
<td>SAA3</td>
<td>serum amyloid A 3</td>
<td>4.6449</td>
<td>Up</td>
</tr>
<tr>
<td>GZMK</td>
<td>granzyme K</td>
<td>4.6062</td>
<td>Up</td>
</tr>
<tr>
<td>GZMA</td>
<td>granzyme A</td>
<td>4.2374</td>
<td>Up</td>
</tr>
<tr>
<td>GPR65</td>
<td>G-protein coupled receptor 65</td>
<td>4.142</td>
<td>Up</td>
</tr>
<tr>
<td>BC023105</td>
<td>cDNA sequence BC023105</td>
<td>4.0748</td>
<td>Up</td>
</tr>
<tr>
<td>RGS1</td>
<td>regulator of G-protein signaling 1</td>
<td>4.0444</td>
<td>Up</td>
</tr>
<tr>
<td>MPA2L</td>
<td>guanylate binding protein 6</td>
<td>4.022</td>
<td>Up</td>
</tr>
<tr>
<td>LST1</td>
<td>leukocyte specific transcript 1</td>
<td>3.9534</td>
<td>Up</td>
</tr>
<tr>
<td>SAMHD1</td>
<td>SAM domain and HD domain, 1</td>
<td>3.9515</td>
<td>Up</td>
</tr>
<tr>
<td>S100A4</td>
<td>S100 calcium binding protein A4</td>
<td>3.9262</td>
<td>Up</td>
</tr>
<tr>
<td>GBP8</td>
<td>guanylate-binding protein 8</td>
<td>3.8212</td>
<td>Up</td>
</tr>
<tr>
<td>ZBP1</td>
<td>Z-DNA binding protein 1</td>
<td>3.8156</td>
<td>Up</td>
</tr>
<tr>
<td>PITPNM1</td>
<td>phosphatidylinositol transfer protein, membrane-associated 1</td>
<td>3.752</td>
<td>Up</td>
</tr>
<tr>
<td>FCGR4</td>
<td>Fc receptor, IgG, low affinity IV</td>
<td>3.678</td>
<td>Up</td>
</tr>
<tr>
<td>TAP1</td>
<td>transporter 1, ATP-binding cassette, sub-family B (MDR/TAP)</td>
<td>3.5956</td>
<td>Up</td>
</tr>
<tr>
<td>PTGER4</td>
<td>prostaglandin E receptor 4 (subtype EP4)</td>
<td>3.5583</td>
<td>Up</td>
</tr>
<tr>
<td>CAR14</td>
<td>carbonic anhydrase 14</td>
<td>-3.4574</td>
<td>Down</td>
</tr>
<tr>
<td>ADCY7</td>
<td>adenylate cyclase 7</td>
<td>3.4362</td>
<td>Up</td>
</tr>
<tr>
<td>GBP2</td>
<td>guanylate binding protein 2</td>
<td>3.373</td>
<td>Up</td>
</tr>
<tr>
<td>CCL5</td>
<td>chemokine (C-C motif) ligand 5</td>
<td>3.3726</td>
<td>Up</td>
</tr>
<tr>
<td>H2-AA</td>
<td>histocompatibility 2, class II antigen A, alpha</td>
<td>3.3344</td>
<td>Up</td>
</tr>
<tr>
<td>PSMB8</td>
<td>proteasome (prosome, macropain) subunit, beta type 8 (large multifunctional peptidase 7)</td>
<td>3.2619</td>
<td>Up</td>
</tr>
</tbody>
</table>
### Table 4.4: Genes most differentially expressed in brains of latently-infected mice

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Gene Name</th>
<th>Signal to Noise Ratio</th>
<th>Up or Down relative to ORF73.stop</th>
</tr>
</thead>
<tbody>
<tr>
<td>IIGP1</td>
<td>interferon inducible GTPase 1</td>
<td>8.358</td>
<td>Up</td>
</tr>
<tr>
<td>GBP2</td>
<td>guanylate binding protein 2</td>
<td>6.2147</td>
<td>Up</td>
</tr>
<tr>
<td>H2-Q7</td>
<td>histocompatibility 2, Q region locus 7</td>
<td>5.3892</td>
<td>Up</td>
</tr>
<tr>
<td>H2-K1</td>
<td>histocompatibility 2, K1, K region</td>
<td>4.5909</td>
<td>Up</td>
</tr>
<tr>
<td>GBP6</td>
<td>guanylate binding protein 6</td>
<td>4.0888</td>
<td>Up</td>
</tr>
<tr>
<td>CCL5</td>
<td>chemokine (C-C motif) ligand 5</td>
<td>3.6399</td>
<td>Up</td>
</tr>
<tr>
<td>MS4A4B</td>
<td>membrane-spanning 4-domains, subfamily A, member 4B</td>
<td>3.2078</td>
<td>Up</td>
</tr>
<tr>
<td>IGTP</td>
<td>interferon gamma induced GTPase</td>
<td>3.1533</td>
<td>Up</td>
</tr>
<tr>
<td>H2-EB1</td>
<td>histocompatibility 2, class II antigen E beta</td>
<td>3.1481</td>
<td>Up</td>
</tr>
<tr>
<td>MPA2L</td>
<td>guanylate binding protein 6</td>
<td>2.5005</td>
<td>Up</td>
</tr>
<tr>
<td>CXCL9</td>
<td>chemokine (C-X-C motif) ligand 9</td>
<td>2.4388</td>
<td>Up</td>
</tr>
<tr>
<td>H2-AA</td>
<td>histocompatibility 2, class II antigen A, alpha</td>
<td>2.414</td>
<td>Up</td>
</tr>
<tr>
<td>PSMB9</td>
<td>proteasome (prosome, macropain) subunit, beta type 9 (large multifunctional peptidase 2)</td>
<td>1.9199</td>
<td>Up</td>
</tr>
<tr>
<td>PLAC8</td>
<td>placenta-specific 8</td>
<td>0.9556</td>
<td>Up</td>
</tr>
</tbody>
</table>
Chapter 5:

Summary and Future Directions
5.1 Summary

Herpesviruses can undergo both productive replication and the establishment of a latent infection. Understanding the role of viral genes in both stages of the viral life cycle and how the immune system works to control viral infection are areas of interest in the field. The unifying theme of this dissertation is investigation of gene expression and regulation during the course of gammaherpesvirus 68 infection. Many viral genes are expressed during productive infection; in fact, recent work by our laboratory and others has identified many new regions of active transcription outside of the known genes, suggesting that viral RNA expression during productive replication is more extensive than previously appreciated (5, 22). The function of such RNAs in MHV68 has not been investigated to date. During latent infection, few transcripts are produced. The regulation of viral gene expression to transition between productive replication and latency and the involvement of host factors, such as IFNs, in this process are areas of current investigation. Moreover, chronic MHV68 infection fundamentally alters the host response to a range of infectious and autoimmune challenges; an open question is what effect chronic immune stimulation has on host gene expression. In this dissertation, I sought to address three questions:

(1) Are transcripts from EGRs functionally important for viral gene expression during productive infection?

(2) How does the cytokine IFNγ regulate productive infection of macrophages?

(3) What is the effect of chronic viral infection on host gene expression?

In Chapter 2, we used new mass sequencing approaches to confirm pervasive transcription by MHV68 and characterized several regions of complex, antisense transcription. We found that several novel EGR-encoded transcripts altered late gene expression and that
transcripts within EGRs 26 and 27 altered the expression of the lytic-switch gene encoded by ORF 50. These data suggest that these newly identified RNAs may play an important and previously unrecognized role in regulating viral gene expression.

In Chapter 3, we sought to understand the mechanism of cell extrinsic control of the virus by the host immune response in macrophages, specifically by IFNs. We have defined a role for IFNγ and its key signaling molecule, Stat1, in regulating transcript expression and promoter activity of the latent-to-lytic switch gene, RTA, encoded by ORF 50. Stat1 is also a key signaling molecule for type I IFNs, and Stat1-deficiency in the absence of IFNγ also increased viral replication, ORF 50 expression, and ORF 50 promoter activity, suggesting a role for type I IFNs in regulating ORF 50 expression in macrophages. Together, these results suggest that MHV68 has evolved to regulate viral gene expression at least in part based on cell extrinsic signals, specifically IFNs.

In Chapter 4, we investigated the effect of latent viral infection on host gene expression in latently infected mice. We found that many genes were differentially expressed in the livers, spleens and brains of latently infected mice compared to uninfected mice or mice infected with a latency-defective virus. Many of the pathways of the genes differentially expressed genes in latent mice are immune-related pathways and specifically IFN signaling pathways. Interestingly, we observed distinct transcriptional signatures in livers and spleens of latently infected mice; in fact, many of the differentially expressed genes in spleens were in cell cycle pathways. Together, these results suggest that the immune system in a latently infected host is profoundly different from that of an uninfected host or one that has cleared an acute infection. These findings further support the proposed immune-mediated mechanisms of latency-induced phenotypes (such as the important role of activated macrophages in protection against bacterial or influenza virus...
infection (1, 38)) and have implications for understanding the complexity of the immune responses in latently infected hosts.

5.2 Function of products of pervasive transcription in MHV68

We and others identified extensive pervasive transcription in MHV68 (Chapter 2 and (5, 22)). In this dissertation, I investigated the function of EGR-encoded transcripts outside of known ORFs and found evidence that transcripts encoded by EGRs alter viral gene expression, suggesting a role for these transcripts during productive infection in vitro.

There are many remaining avenues of future investigation for this topic, including: (1) identification and functional characterization of other transcripts encoded by EGR 27 and EGR 26; (2) examination of the transcriptional architecture of other EGRs and characterization of encoded transcripts; and (3) testing the role of EGR-encoded transcripts during acute and chronic infection in vivo. Moreover, it will be interesting to investigate the function of similar transcripts in other herpesviruses. Our results suggest that pervasive transcription described in KSHV and HCMV (4, 11, 53, 54) and likely present in other human herpesviruses may well have important functions in viral replication and viral pathogenesis.

5.2.1 Further functional analysis of EGR 26 and EGR 27 transcripts

To date, we have found that ASOs targeting EGRs 26 and 27 alter viral gene expression. However, there are some questions still to be resolved such as why EGR 26a-c ASOs, but not EGR 26d-e ASOs, reduce ORF 6 transcript levels (Chapter 2). These data suggest that the effects that we have described for EGR 26a-c ASOs were the result of targeting a smaller transcript or spliced isoform within EGR 26. Left unexplored by our work to date is the significance of the
additional 3’ ends detected by RNA sequencing within EGR 27 and ORFs 66-67 which may correspond to additional, less abundant EGR 27 transcripts. In addition, I have performed 3’ RACE in the EGR 27 region that further suggests the presence of multiple RNAs within EGR 27 that are less abundant than EGR 27 transcripts A and B (data not shown). Analysis of our RNA sequencing data by Dr. Steve Johnson suggests splicing is common throughout the virus, especially in EGRs 23, 26 and 27 (unpublished observations). The fact that EGR 27d ASO reduced levels of EGR 27 transcript B and exerted the same effect on ORF 6 and ORF 50 transcript levels as EGR 27a ASO suggests that transcription in this region may be even more complex than depicted in Chapter 2 and suggests that there may be an additional transcript encoded by EGR 27 or a spliced isoform that is functionally important. Additional detailed mapping of the transcripts in this region would help to further elucidate the complex transcriptional architecture in this region. Together, these data suggest that the northern blot analysis and transcriptional mapping to date may not completely capture the complexity of the MHV68 transcriptome.

It is worthwhile considering the implications for analysis of mammalian transcriptomes with our data in mind. We were able to achieve a great depth of sequencing coverage given the small size of the MHV68 genome relative to a mammalian genome. Analysis of the human transcriptome has revealed similar complex and pervasive transcription across the genome (6, 18). Some of these transcripts have only been detected with a single read (18), suggesting that the current sequencing depth may only scratch the surface of the complexity of the mammalian transcriptome.

As described in Chapter 2, we have identified transcripts emanating from EGRs 26 and 27 that overlap the capsid protein, M9 (ORF 65) (3, 31). Previous studies have shown that the
M9 region plays a number of divergent roles in the MHV68 life cycle, including a candidate region of latent gene expression and a component of the virion (19, 34, 48). It is intriguing to consider whether transcripts encoded by EGRs 26-27 may be produced during latency. Determining whether EGR 26-27 transcripts or other EGR-derived transcripts are produced during latency would be an interesting area for future research. However, it is also possible that the detection of the M9 region in latently infected cells is the result of low level reactivation events from a highly expressed region of the virus.

Our data suggest that transcripts encoded by EGRs 26 and 27 regulate ORF 50 expression since ORF 50 transcript levels are decreased in cells transfected with EGR 27a, 27d or 26c ASOs (Chapter 2). This has important implications for regulation of viral replication given the central role of ORF 50 in viral gene expression. Defining the precise transcripts encoded by EGRs 26 and 27 that modulate ORF 50 expression is an important next step. Additionally, understanding how such transcripts regulate ORF 50 expression is a critical topic for future investigation. It is intriguing to speculate whether EGR 26 or 27 transcripts regulate ORF 50 expression by acting as a molecular scaffold for chromatin-modifying complexes like KSHV PAN RNA or might sequester proteins that act to repress ORF 50 promoter activity (35). One avenue to investigate the role of chromatin proteins and EGR-derived transcripts would be to use RNA immunoprecipitation, in which RNA binding proteins are immunoprecipitated, in combination with RNA sequencing (RIP-Seq) to identify RNA-protein interactions. For example, it would be interesting to determine whether the HDAC3/4/NCoR complex known to repress ORF 50 promoter activity interacts with EGR-encoded transcripts (14). Although activating complexes or modifications of ORF 50 promoters are less well-characterized at present, several recent studies have identified activating histone modifications in KSHV (15, 45). Such studies in MHV68
could lead to candidate proteins that interact with ORF 50 promoters and EGR-encoded transcripts. Additional discussion on ORF 50 regulation by EGR-encoded transcripts is included in the section on ORF 50 regulation below.

Our data suggest that EGR 27 transcript A may act later in the viral life cycle than EGR 27 transcript B and therefore may have a distinct function. Further work will be needed to clarify the role of EGR 27 transcript A including targeting this transcript with additional ASOs. The second ASO that we designed to this region reduced the abundance of EGR 27 transcript B in addition to the levels of EGR 27 transcript A even though it was not predicted to target the former. The one ASO that specifically targeted EGR 27 transcript A (EGR 27c ASO) reduced late gene expression as well as protein levels of the early gene, ORF 6, but not ORF 6 transcript levels or ORF 50 transcript levels. The lack of an effect on the immediate-early gene, ORF 50, suggests that this RNA acts later in the viral life cycle than EGR 27 transcript B and may play a role in viral protein expression or viral DNA replication, rather than a role in regulating viral transcription. Interestingly, a role in viral DNA replication has been described for the EBV BHLF1 transcript, which forms an RNA-DNA hybrid molecule at the origin of lytic replication (OriLyt) and is important for recruitment of the viral single stranded binding protein, BALF2, to OriLyt (33).

Importantly, while our data demonstrate that RNAs derived from EGRs are functionally important for the virus, our strategy has been limited to transient knockdown with single-stranded ASOs. As is to be expected with such a strategy, we observed partial transcript knockdown and as a result, subtle phenotypes. Several options are available for further study of EGR-encoded transcripts, including transfecting cells with multiple ASOs targeting the same transcript or the generation of a knockout virus. Again, given the transcriptional complexity of
MHV68, such a strategy necessitates thorough transcript mapping. Given the complex nature of MHV68 transcription with multiple overlapping and antisense transcripts, the feasibility of making a knockout virus will depend on the precise transcriptional architecture of the region of interest. Future work could also utilize overexpression of transcripts of interest to confirm the phenotypes reported here and to investigate the function(s) of other EGR-encoded transcripts. It will be important, however, to consider the contribution of antisense transcripts of known ORFs in any such experiments. For example, EGR 26 transcript A encompasses the entire transcript and putative promoter for ORF 64 as well as most of ORF 63. One option would be to mutate the sequence for the antisense ORF so that overexpression findings are not confounded by it.

5.2.2 Transcriptional architecture and functional analysis of other EGR-encoded transcripts

The work in this dissertation has primarily focused on EGRs 26 and 27 as targeting of transcripts encoded by these EGRs had the greatest effect on ORF 4 surface expression, making them the best candidates for mechanistic studies with the available reagents. However, it is clear from our ORF 4 flow cytometry screen that other EGR-encoded transcripts may function in viral gene expression (Chapter 2). Moreover, there are sixteen EGRs that we did not target in our screen and have not been investigated by any means; these EGRs could form the basis of additional screens.

One region of particular interest is EGR 23, which is antisense to ORF 50. A transcript in KSHV has also been described from a sytenically-conserved region of the KSHV genome, though this RNA was found not to alter KSHV ORF 50 transcript levels and instead was proposed to encode small peptides (53). In MHV68, we have performed northern blot analysis of
EGR 23. A northern blot probe detected at least five overlapping polyadenylated transcripts between nucleotides 68313 and 68765 (Fig. 5.1). Additionally, Dr. Steve Johnson identified many splicing events by computational analysis of RNA sequencing data, confirming the presence of multiple overlapping transcripts in this region. The northern probe encompassed the targets of EGR 23a and 23b ASOs which significantly reduced ORF 4 surface expression (Chapter 2). Mapping transcripts encoded by EGR 23 and determining the role of each in viral gene expression is an area of interest for future research and could be achieved using similar methods to those used for analyzing transcripts derived from EGRs 26 and 27. Preliminary western blot analysis suggests that, unlike transcripts encoded by EGRs 26 and 27, transcripts targeted by EGR 23b ASO did not reduce protein levels of the early gene, ORF 6 (data not shown). An effect specifically on late genes suggests a role for EGR 23 transcripts in viral DNA replication and suggests that distinct EGR-encoded transcripts function in different stages of the viral life cycle.

Two of the ASOs used in our ORF 4 flow cytometry screen increased ORF 4 surface expression suggesting that the targeted RNAs may have a role in suppressing viral gene expression (Chapter 2). It will be interesting to overexpress these RNAs and confirm that these RNAs act to restrict viral gene expression and to further assess the role of these RNAs *in vivo*. It is intriguing to hypothesize that RNAs that suppress viral gene expression *in vitro* may play a role in establishing or maintaining latency.

Finally, while EGR 26-27 transcripts are enriched in the nucleus of infected cells, suggesting that they are nuclear noncoding RNAs, the localization of other EGR-encoded transcripts have not been evaluated thus far. If other EGR-encoded transcripts were found to be expressed in the cytoplasm, these RNAs may be cytoplasmic ncRNAs or may encode peptides or
small proteins. In addition to determining nuclear versus cytoplasmic localization for other EGR-encoded transcripts, the best option for evaluating which RNAs might encode proteins or peptides would be ribosomal profiling. In ribosomal profiling, RNAs associated with ribosomes are isolated using a sucrose gradient and then sequenced to determine their identity. A recent report has defined a pattern of ribosome release that is characteristic of ncRNAs (17). Using a combination of subcellular fractionation and ribosome profiling, it will be interesting to determine which EGR-encoded transcripts are likely to be noncoding and which might be translated.

5.2.3 Role of EGR-encoded transcripts during infection in vivo

An advantage of studying MHV68 is the opportunity to investigate the role of viral and host genes in the context of in vivo infection. Studies to date of herpesvirus ncRNA have necessarily focused on the effect of such RNAs in vitro (Chapter 2 and (2, 8, 20, 21, 23, 24, 28, 30, 35-37, 39-41)). Depending on the transcriptional architecture of the specific EGR-encoded transcript in question, it may be possible to generate a mutant virus and then to use this virus to infect mice. Infection of mice with a mutant virus would allow us to determine the role of a specific RNA in acute viral infection, establishment of and reactivation from latency, and the production of persistent virus as well as virus-induced pathologies, such as large-vessel vasculitis and the development of lymphoma. Because of the evident transcriptional complexity in many regions, an alternative is to generating a mutant virus is to target a specific EGR-encoded transcript in vivo using ASOs shown to knockdown the transcript of interest in vitro (Chapter 2 and (16, 49, 50)).
Finally, for RNAs hypothesized to play a role in modulating the host response to infection, it may be possible to deliver such RNAs individually and investigate their role in the host immune response outside of the context of viral infection. As an example, an RNA could be overexpressed in vitro and the effect on host gene expression evaluated by microarray or RNA sequencing. If such an RNA substantially altered a host pathway, such as IFN signaling, this would indicate that it may have a role in modulating the host response to infection and would be a candidate for evaluation in vivo. This could be accomplished by generating a transgenic mouse as has been done for the mammalian ncRNA, NeST (13).

5.3 Stat1 regulation of gammaherpesvirus 68 ORF 50 transcription

In this dissertation, I have presented evidence that ORF 50 promoters are regulated by the host cytokine, IFNγ, and its key signaling molecule, Stat1, and that type I IFNs may also regulate ORF 50 promoters via an IFNγ-independent role of Stat1. Preliminary chromatin immunoprecipitation (ChIP) experiments by Dr. Megan Goodwin suggest that Stat1 does not bind directly to the ORF 50 promoters in the presence or absence of IFNγ (data not shown), suggesting that Stat1-mediated repression is indirect. Whether Stat1 sequesters activating complexes as it does to regulate MMP-9 (26) or acts indirectly via another transcription factor is unknown and could form the basis for future studies. There are several possible approaches to investigate the role of Stat1 in regulating ORF 50 promoters in the presence or absence of IFNγ. The best approach would be to screen for proteins that regulate ORF 50 promoter activity using a cDNA overexpression or siRNA knockdown library. Once additional transcriptional factors are
identified by such a screening approach, it would be possible to assess whether these factors bind directly to the ORF 50 promoters by ChIP. Another possibility is to analyze the ORF 50 promoters for putative transcription factor binding sites and then assessing the role of candidate transcription factors in regulating ORF 50 promoters. Using MatInspector for analysis of the core and distal ORF 50 promoters (data not shown), we have identified many candidate binding elements, but as was shown for Stat1, many of these elements may not bind their predicted binding elements. Candidate proteins will need to be each tested empirically. Moreover, given that the mechanisms of IFNγ-mediated gene repression are less well-studied, we favor a screen over a candidate-based approach to better capture novel means of regulation.

Type I IFN regulation of viral replication in macrophages is mediated by IRF3 (52). Type I IFN treatment of bone macrophages has been shown to directly reduce ORF 50 transcript levels, but the effect of type I IFN on ORF 50 promoter activity has not been formally tested to date and is an important next step (52). Further, it will be important to specifically test whether IRF3 mediates type I IFN control of ORF 50 promoter activity (52). Whether IRF3 binds directly to the ORF 50 promoters is unknown and could be tested by ChIP. If IRF3 does not bind directly, the screening described above can be used to identify additional candidate molecules for both type I and type II IFN regulation of ORF 50.

It would be interesting to evaluate the role of IFN-repressible elements within the ORF 50 promoters in the context of viral infection. We hypothesize that mutation of IFNγ-induced regulatory elements would generate a virus that persists and is insensitive to IFNγ regulation, reactivating even in the presence of an intact immune response, perhaps generating pathologies such as large-vessel vasculitis and lymphoma typically only observed in immunodeficient hosts (7, 25, 43, 44, 47). However, a complication to generating such a virus is that the Speck
laboratory has recently identified two additional ORF 50 promoters in addition to the two ORF 50 promoters described here and has confirmed that these promoters are also responsive to IFN-γ (personal communication). Moreover, the role of type I IFNs in regulating these novel promoters has not been evaluated to date. It is clear that the regulation of transcription of this key viral molecule is complex.

Another area for future investigation is to evaluate whether type I or type II IFNs interact with other regulators of ORF 50. One area of interest is determining whether IFNs regulate transcripts that overlap EGRs 26-27. Drs. Erin Willert and Steve Johnson investigated the role of IFN-γ on viral gene expression by tiled array and found that nearly all viral genes, aside from ORF 73, were silenced by treatment with this cytokine (personal communication). Given the importance of ORF 50 in regulating viral gene expression, it is likely that many genes were suppressed due to decreased ORF 50 levels. However, in light of our data that EGR-encoded transcripts may regulate ORF 50 transcript expression (Chapter 2), it will be interesting to evaluate the effect of type I and type II cytokines on the promoters of EGR 26 and 27 transcripts. Furthermore, it is intriguing to speculate whether IFN-γ, EGR 26 and 27-encoded transcripts, and the HDAC3/4/NCoR complex interact to regulate ORF 50 expression. This could be tested by RIP. Finally, in view of the role of KSHV PAN RNA in regulating ORF 50 gene expression, it will be interesting to test whether cytokines regulate the promoter of PAN RNA (35, 36).

Finally, other cytokines are known to be important in MHV68 infection. Like IFN-γ, TNF-α is upregulated during latent infection (1). However, the role of this cytokine in regulating ORF 50 promoter activity or MHV68 viral gene expression has not been evaluated. Moreover, recent research in our laboratory has uncovered a role for the TH2 cytokines, IL-4 and IL-13, in promoting viral replication in macrophages and ORF 50 promoter activity (Tiffany Reese,
personal communication). It will be interesting to define the mechanism of such cytokines on the ORF 50 promoters and viral gene expression.

5.4 Latent infection alters host transcription

Previous work has defined a role for cells of the innate immune system, specifically macrophages and natural killer cells, in latency-induced phenotypes, including resistance to lethal challenges with bacterial pathogens, influenza virus, and lymphoma cells (1, 38, 51). We have now uncovered immune signatures in the brains, livers and spleens of latently infected MHV68. Interesting questions for future research include: (1) Are latency-specific host gene expression changes and latency-induced phenotypes, such as cross-protection to Listeria monocytogenes (1), attributable to establishment of latency or depend on the capacity to reactivate? (2) What are the latency-specific gene expression changes in specific cell types (e.g. macrophages) or other tissues (e.g. lung)? and (3) What is the effect of such transcriptional changes on cells of the adaptive immune system, specifically T cells?

5.4.1 Relative contribution of latency establishment versus reactivation to latency-induced phenotypes

Our work and the work of others have identified a role for latency in several latency-induced phenotypes by comparing wild-type infection to infection with latency-defective MHV68 (Chapter 4 and (1, 9)). As a next step, it will be interesting to determine whether the capacity of the virus to reactivate influences latency-induced phenotypes. For example, we could compare infection of mice with ORF73.stop mutant virus to a virus such as ORF72.stop mutant
virus and then evaluate gene expression changes and resistance to lethal challenges with *Listeria monocytogenes* and influenza virus. The ORF72.stop mutant virus is able to establish a latent infection, but is profoundly defective in its ability to reactivate from latency whereas, when administered intranasally, the ORF73.stop mutant virus fails to establish a latent infection (10, 29, 46). A crucial first experiment will be to test establishment of latency by limiting dilution PCR and *ex vivo* reactivation from latency by limiting dilution assay for mice infected with ORF72.stop mutant virus delivered intranasally to confirm that ORF72.stop mutant virus establishes latency, but is defective in reactivation from latency. Previous studies have demonstrated that the route of infection (intranasal vs. intraperitoneal) alters whether ORF73.stop mutant virus can establish a latent infection (29, 32).

### 5.4.2 Effect of MHV68 latency on specific cell types and tissues

We have reported here distinct gene expression profiles for three specific tissues (brain, liver and spleen) from latently infected mice. It may be worthwhile to determine the gene expression profile of other tissues as well. Tissues of interest include the lung given the observed latency-induced protection to influenza virus infection (38). One limitation of our study was that we determined the gene expression profile for an entire organ, composed of multiple distinct cell types. Since macrophages are important in several models of latency-induced cross protection (1, 38), it will be interesting to determine macrophage gene expression signatures from various tissues during latent infection. A recent comprehensive study has characterized macrophage-specific markers and these could be used to isolate macrophages from tissues of interest (12).
5.4.3 Effect of latent infection on T cells

During latent MHV68 infection, several cytokines, including type I and type II IFNs and TNFα, are produced (1, 27). A number of studies have suggested that the presence of cytokines, including IL-12 and type I IFNs, can alter effector T cell differentiation and responses (reviewed in (42)). Interestingly, there is a prominent type I IFN gene signature in bystander CD8+ memory T cells from mice chronically infected with lymphocytic choriomeningitis virus (LCMV), and this gene expression signature significantly overlaps with our signature from latently-infected mice (Erietta Stelekati and E. John Wherry, personal communication), suggesting that type I IFNs may play an important role in shaping bystander T cell responses during MHV68 latency as well. While type II IFN is unlikely to play a role in T cell conditioning, it may play some role in mediating B cell responses. In fact, antibody has been identified as required in a mouse model of MHV68-induced thrombocytopenia (9).

To directly test whether latent MHV68 infection altered T cells, we transferred naïve OT1 T cells into congenic mice prior to infection with wild type MHV68 or ORF73.stop mutant virus and assessed the transcriptional profile of the transferred cells by microarray analysis at 28 days post infection. In a pilot experiment, we found that there were a number of differentially expressed genes between latently infected mice and mice infected with the ORF73.stop mutant virus (data not shown) suggesting that chronic viral infection altered CD8 T cells. One limitation of this experiment was that only a small percentage of the transferred T cells remained after five weeks. An alternative approach would be to transfer T cells after latency is established so that the cells would not need to survive as long in vivo before analysis. In addition to investigating the role of latent MHV68 on naïve bystander CD8+ T cells, it would be interesting to evaluate its role on naïve CD4+ T cells as well as effector and memory cells. If a distinctive gene signature is
confirmed in bystander T cells, it would be interesting to evaluate whether other herpesvirus infections, such as MCMV, have similar or distinct effects on T cell gene signatures. Whether type I IFN plays a critical role in the observed phenotypes could be evaluated by using IFNAR-/-transgenic T cells as bystander cells. Experiments such as these have implications for understanding how chronic viral infection may shape the mature immune response in humans.
5.5 References


replication do not regulate that transcript but serve as mRNAs encoding small peptides. J Virol \textbf{84}:5465-5475.

Figure 5.1: Multiple polyadenylated transcripts emanate from EGR 23. 200 ng of poly(A)-selected or 5 µg of total RNA isolated from MHV68-infected 3T12 fibroblasts (MOI=10) at 18 hpi was analyzed by northern blot analysis using a probe to EGR 23 (nucleotides 68313 and 68765).