Analysis of the viral gene expression of *Cercopithecine herpesvirus 2* (SA8) by DNA microarray

by

Claire Sevenhuysen,

A thesis submitted to the Faculty of Graduate Studies of the University of Manitoba
in partial fulfillment of the requirements of the degree of

MASTER OF SCIENCE

Department of Medical Microbiology
University of Manitoba
Winnipeg, Manitoba, Canada

Copyright © 2006 by Claire Sevenhuysen

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION

Analysis of the viral gene expression of *Cercopithecine herpesvirus* 2 (SA8) by DNA microarray

by

Claire Sevenhuysen

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree

of

Master of Science

Claire Sevenhuysen © 2006

Permission has been granted to the Library of the University of Manitoba to lend or sell copies of this thesis/practicum, to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film, and to University Microfilms Inc. to publish an abstract of this thesis/practicum.

This reproduction or copy of this thesis has been made available by authority of the copyright owner solely for the purpose of private study and research, and may only be reproduced and copied as permitted by copyright laws or with express written authorization from the copyright owner.

TABLE OF CONTENTS

TABLE OF CONTENTS ii
ACKNOWLEDGEMENTS v
LIST OF FIGURES vi
LIST OF TABLES vi
LIST OF COPYRIGHTED MATERIAL vii
LIST OF ABBREVIATIONS vii
ABSTRACT 1
1.1. Herpes virology 2
1.1.1. Clinical syndrome & virology 2
1.1.2. Virology and taxonomy 4
1.1.3. Viral genes involved in pathogenesis 6
1.1.3.1. ICP34.5 6
1.1.3.2. US11 8
1.1.4. Viral genome structure and homology 9
1.1.4.1. Transcripts of the L/S junction 12
1.1.5. Viral replication 18
1.1.6. Temporal HSV gene expression 19
1.1.7. Immediate-early gene products 23
1.1.7.1. RL02 – ICP0 23
1.1.7.2. RS01 – ICP4 25
1.1.7.3. UL54 – ICP27 26
1.1.7.4. US01 – ICP22 27
1.1.7.5. US12 – ICP47 28
1.2. DNA Microarrays 28
1.2.1. Principle 28
1.2.2. Applications 29
1.2.3. Platform selection 30
1.2.4. Experimental design 32
1.2.4.1. Data 'cleaning' 32
1.2.4.2. Replicate measurements 33
1.2.4.3. Randomization 34
1.2.4.4. Reference selection 34
1.2.5. Determining differential expression 37
1.2.5.1. Low-level analysis 38
1.2.5.2. Statistical tests 39
1.2.5.3. High-level analysis / pattern recognition 40
1.2.6. Validation of microarray results 42
1.2.7. SA8 microarray data analysis 43
1.3. Herpesvirus microarrays 44
2. OBJECTIVES 46
3. HYPOTHESIS 46
4. MATERIALS AND METHODS 47
4.1. Cells and virus 47

4.2. Construction of SA8 microarray 48
4.3. Viral genomic DNA preparation & hybridization 50
4.4. PFGE of replicating viral DNA & Southern blot 52
4.5. RNA preparation 55
4.7. Array hybridization, washing & scanning 56
4.8. Microarray quality control 57
4.9. Data processing and analysis 58
4.9.1. Scanning & quantification 58
4.9.2. Image verification 59
4.9.3. Treatment of replicates and variants 60
4.9.4. Inter-slide normalization using a gDNA standard 61
4.9.5. Inter-slide normalization using the Lucidea Scorecard 61
4.9.5.1. Lucidea Calibration Curve normalization 61
4.9.5.2. Average Lucidea intensity normalization 62
4.9.6. Clustering (k-means) 62
4.9.7. Classification analysis 63
4.10. Quantitative RT-PCR 64
5. RESULTS 67
5.1. Development of microarray normalization protocol 67
5.1.1. Simian simplex virus microarray design 67
5.1.2. Test Arrays 67
5.1.3. Microarray Quality Control 70
5.1.4. Simian simplex virus infection kinetics 71
5.1.5. Development of a reliable microarray normalization protocol 71
5.1.5.1. Inter-slide normalization using viral genomic DNA as reference 73
5.1.5.2. Inter-slide normalization using Lucidea Scorecard Calibration curve 74
5.1.5.3. Normalization protocol for small, variable microarray data 76
5.2. Analysis of simplexvirus immediate-early gene expression 78
5.2.1. Microarray analysis of HSV-1 α-gene expression 78
5.2.2. Microarray analysis of SA8 α-gene expression 80
5.2.3. Validation by RT-PCR of SA8 α-gene expression 80
5.2.4. Analysis of the sequences of SA8 α-gene promoters 87
5.3. Analysis of SA8 early & late gene expression 88
5.3.1. Microarray analysis of SA8 β- and γ-gene expression 88
5.3.2. Temporal kinetics of SA8 gene expression 90
5.4. Alternate analysis protocol for viral gene expression microarrays 90
5.4.1. Analysis protocol for the expression of viral genes 90
5.4.1.1. Selection of non-coterminal, non-overlapping transcripts 93
5.4.1.2. Relative difference of signal between experimental conditions 94
5.4.1.3. Clustering (k-means) 95
5.4.1.4. Classification analysis 95
5.4.1.5. Analysis protocol for the expression of coterminal viral genes 96
5.4.1.6. Interpretation 96
5.4.2. Microarray analysis of SA8 α-gene expression 96
5.4.3. Microarray analysis of SA8 β-gene expression 106
5.4.4. Similarities in expression kinetics of SA8 genes 110

5.4.4.1. Microarray analysis of gene expression in the presence of CHX 110
5.4.4.2. Comparison of microarray results from 5h and 8h 113
5.4.4.3. Microarray analysis of gene expression in the presence of ACV 114
5.4.5. Temporal kinetics of SA8 gene expression 115
5.5. Summary of SA8 gene expression 121
6. DISCUSSION 124
6.1. Comparison of SA8 IE gene kinetics with HSV-1 homologues 124
6.2. Analysis of SA8 IE gene expression by qRT-PCR 126
6.3. Comparison of SA8 E and L gene kinetics with HSV-1 homologues 127
6.4. Presence of transcripts in the RL1 region of SA8 130
6.5. Challenges presented by microarray analysis of herpesvirus gene expression 130
7. REVIEW OF OBJECTIVES 132
8. FURTHER WORK 132
9. APPENDICES 134
9.1. Appendix A – DNA Oligonucleotide microarray probes 134
9.2. Appendix B – qRT-PCR results of SA8 gene expression analysis 141
9.3. Appendix C – Additional observations of similar kinetic characteristics 144
9.4. Appendix D –Microarray cluster membership in the presence of CHX 147
9.5. Appendix E – Mean microarray cluster membership in the presence of ACV 154
REFERENCES 157

ACKNOWLEDGEMENTS

I thank the members of my advisory committee, Dr. Alberto Severini, Dr. Graham Tipples, Dr. Kevin Coombs, and Dr. Sabine Mai, for their valuable input. I thank my supervisor Dr. Alberto Severini for his "open-door" availability and willingness to answer questions, as well as for allowing me the liberty to pursue my own ideas and learn from my mistakes.

I express thanks to all members of the Viral STD Laboratory. I would particularly like to thank Dr. Tim Dumonceaux for his indispensable assistance in establishing the qRT-PCR validation of immediate-early gene expression in SA8. I would also like to thank Ravinder Singh for his analysis of SA8 gene expression by northern blot, and Suzanne Gibbons for her guidance and support.

I would like to thank all present and past members of the Viral Exanthema lab for their support and advice as well as the use of "their" lab and office space. I am particularly grateful to Joanne Hiebert, for her wealth of practical knowledge of herpesvirus microarrays.

I express my appreciation for the hard work of the members of the NML DNACore Facility, in particular Claude Ouelette, Candice Fisher, Shari Tyson, and Travis Murphy. I would also like to thank Dr. Morag Graham for her extensive advice on the subjects of microarray experimental design and data analysis.

I thank the Special Pathogens section for the preliminary work with herpes B virus which does not appear in this thesis. I extend my gratitude particularly to Dr. Lisa Fernando for her long hours in the BSL4 lab.

Thanks to the many wonderful people on the 5th floor, in particular Sharon Simon, Debra Parchaliuk, and Hannah Richards for their advice on microarrays, graduate studies, career goals, and just for lightening many days of 'data crunching' with laughter.

I would like to thank my husband, Philippe Lagacé-Wiens for his unwavering support, for providing advice and an alternate perspective for my scientific musings, and for bolstering my confidence in my abilities any time it was needed. I thank my parents and in-laws for sending words of encouragement from near and far, as well as not asking too often in the final months "how's your thesis going?".

I would especially like to thank my father, Dr. Gustaaf Sevenhuysen for his unabated enthusiasm for this project and for his academic advice, but particularly for his invaluable assistance in the refinement of the statistical analysis procedures without which this project may have never come to fruition.

This work was supported by the Public Health Agency of Canada, and a Health Canada Genomics Research and Development Grant.

LIST OF FIGURES

Figure 1 – Map of the open reading frames and features of the SA8 genome. 10
Figure 2 – SA8 oligonucleotide microarray images. 68
Figure 3 – Box plot of signal intensity of different probe types. 69
Figure 4 – Southern blot of SA8 DNA at various times post-infection. 72
Figure 5 – Comparison of mean signal intensity of Lucidea and viral probes 77
Figure 6 – HSV-1 microarray signal intensity 79
Figure 7 – SA8 microarray signal intensity 81
Figure 8 – Northern blot analysis of SA8 US12 85
Figure 9 – Northern blot analysis of SA8 UL54. 86
Figure 10 – SA8 microarray intensity signals in the presence and absence of ACV 89
Figure 11 – SA8 microarray intensity signals at 5hpi and 8hpi 91
Figure 12 – SA8 microarray intensity signals at 5hpi and 13hpi 92
Figure 13 – SA8 microarray clusters in the presence of CHX (Exp #1). 98
Figure 14 – Classification tree (Exp #1) 99
Figure 15 – SA8 microarray clusters in the presence of CHX (Exp #2) 100
Figure 16 – SA8 microarray clusters in the presence of CHX (Exp #3A) 102
Figure 17 – SA8 microarray clusters in the presence of CHX (Exp #3B) 103
Figure 18 – Summary of SA8 IE gene expression in the presence of CHX. 105
Figure 19 – SA8 microarray gene clusters in the presence of ACV. 107
Figure 20 – SA8 microarray gene clusters at 5h and 8hpi. 116
Figure 21 – SA8 microarray gene clusters at 5h and 13hpi. 117

LIST OF TABLES

Table 1 – Features of select human and simian simplexviruses 5
Table 2 – Open reading frames and other features of the SA8 genome. 13
Table 3 – Primers used for qRT-PCR validation. 66
Table 4 – Expression of SA8 genes analysed by qRT-PCR. 83
Table 5 – Cluster membership of SA8 genes in the presence of ACV 108
Table 6 – Groups of SA8 transcripts with similar expression 111
Table 7 – Clusters of SA8 transcripts at 5h and 8hpi. 118
Table 8 – Clusters of SA8 transcripts at 5h and 13hpi. 119
Table 9 – Summary of SA8 gene kinetics based on microarray 123

LIST OF COPYRIGHTED MATERIAL FOR WHICH PERMISSION WAS OBTAINED

Figure 1	Map of the open reading frames and features of the SA8 genome Tyler, S. D., G.A. Peters, and A. Severini. 2005. Complete sequence of cercopithecine herpesvirus 2 (SA8) and comparison with other herpesviruses. Virology **331**:429-440, figure 1.	Page 10
Table 2	Open reading frames and other features of the SA8 genome. Tyler, S. D., G.A. Peters, and A. Severini. 2005. Complete sequence of cercopithecine herpesvirus 2 (SA8) and comparison with other herpesviruses. Virology **331**:429-440, table 1.	Page 13

LIST OF ABBREVIATIONS

A	Adenine
ACV	Acyclovir
ANOVA	Analysis of Variance
bp	Base pair(s)
°C	Degrees Celsius
C	Cytosine
cDNA	Single-strand DNA transcribed from RNA
CeHV-1	*Cercopithecine herpesvirus 1* ; B virus
CeHV-2	*Cercopithecine herpesvirus 2* ; SA8
CHX	Cycloheximide
CPE	Cytopathic effect
CT	Crossing threshold
DEPC	Diethylpyrocarbonate
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTPs	Mixture of the four deoxynucleotides
dsRNA	Double-strand ribonucleic acid
dUTP	Deoxyuridine triphosphate
E	Early (β)
EDTA	Ethylenediaminetetraacetic acid
eIF2α	Eukaryotic translation initiation factor 2α
ER	Endoplasmic reticulum
f.c.	Final concentration
FBS	Fetal Bovine Serum
G	Guanine
G+C	Guanine-cytosine content (%)
gDNA	Genomic DNA
x g	Times gravity

HCMV	Human cytomegalovirus, (*Human herpesvirus 5*)
hpi	Hours post-infection
HSV-1	Herpes Simplex Virus, type 1 (*Human herpesvirus 1*)
HSV-2	Herpes Simplex Virus, type 2 (*Human herpesvirus 2*)
HVP-2	Herpes virus papio, type 2 (*Cercopithecine herpesvirus 16*)
IE	Immediate-early (α)
IFN	Interferon
L	Late (γ)
L-Glu	L-glutamine
M	Molar
MEM	Minimal Essential Medium
MOI	Multiplicity of infection
MOPS	3-(N-morpholino) propanesulfonic acid
mRNA	Messenger RNA
NML	National Microbiology Laboratory
Oligo	Oligonucleotide
ORF	Open reading frame
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PenStrep	Penicillin-streptomycin
PFGE	Pulsed Field Gel Electrophoresis
PKR	Protein kinase R
PP1α	Protein phosphatase 1α
qRT-PCR	Quantitative reverse transcriptase polymerase chain reaction
RNA	Ribonucleic acid
RT	Reverse transcription
SA8	Simian agent 8 (*Cercopithecine herpesvirus 2*)
SD	Standard deviation
SDS	Sodium dodecyl sulfate
SOP	Standard Operating Procedure
T	Thymidine
TIGR	The Institute for Genomic Research
Tm	Melting temperature

ABSTRACT

Simian agent 8 (SA8) has considerable genetic similarity to other primate herpes simplexviruses, but differences in virulence are observed during infection of xenogeneic hosts. Simplexvirus genes are expressed in a tightly controlled cascade, depending on the requirement for *de novo* protein synthesis or viral DNA replication for their expression. Although SA8 is genetically similar to its more virulent counterparts, critical differences may be found at the level of viral gene expression.

An oligonucleotide microarray was developed in order to study SA8 viral gene expression. A novel protocol for the analysis of viral gene expression microarrays is presented, as well as preliminary analysis of patterns in viral gene expression.

Results from the SA8 microarray, as well as analysis of transcription by quantitative RT-PCR and northern blot, revealed that the kinetics of expression of the SA8 immediate-early genes differs greatly from that of the prototype simplexvirus, HSV-1.

1. INTRODUCTION

1.1. Herpes virology

Human infection with herpes simplex virus type 1 (HSV-1), and simian herpesviruses *Cercopithecine herpesviruses 1, 2,* and *16* (CeHV-1 (B virus), CeHV-2 (SA8), and CeHV-16 (HVP-2)) result in disease with a wide range of severity. These four viruses have significant gene homology despite differences in pathogenicity in humans which ranges from asymptomatic (SA8) to fatal encephalitis (B virus). A complete description and comparison of the gene expression patterns of these viruses in the human host will add to our understanding of the mechanism of pathogenesis of lethal infection with B virus. The primary objective of this project is to study the gene expression of SA8 by DNA oligonucleotide microarrays and compare expression in SA8 to the well-studied prototype simplexvirus, HSV-1.

Because of their close genetic similarities, a comparative analysis of human and simian simplexviruses may identify the subtle differences that are responsible for their dramatically different virulence in xenogeneic hosts. Candidate virulence genes may be identified and this would contribute to our understanding of the pathogenic mechanisms of simplexvirus infection in humans and the development of attenuated HSV strains for vaccine purposes.

1.1.1. Clinical syndrome & virology

Monkeys of the genus *Macaca* are the natural host of B virus (named "B" after the family name of the first person infected) (175, 180). Baboons (genus *Papio*) are the natural hosts of both SA8 (Simian Agent 8) (104), and HVP-2 (Herpesvirus Papio 2)

(43). Humans are the natural host of both type 1 and type 2 herpes simplex viruses (HSV-1 and HSV-2). More than 90% of baboons are seropositive for HVP-2 (41, 42). The prevalence of B virus in captive macaques is 45-50% (6, 176, 180), which presents an occupational risk for humans in contact with infected animals. Current seroprevalence rates in humans of HSV-1 and HSV-2 in developed nations are 57.7% and 17.0%, respectively (186).

In their natural hosts, all four of these simplexviruses cause mucosal and neuronal infections with a range of symptoms similar to HSV in humans. They are transmitted through sexual or oral contact (including bites). Adult primates tend to acquire genital lesions, while oral lesions are more common among the young (96). In humans, HSV-1 is predominantly found in oral lesions, while HSV-2 is more often the cause of genital infection, although the reverse does occur (186). Disseminated disease caused by SA8, B virus & HVP-2 is sometimes observed in young animals, similar to neonatal infection with HSV-2 (42). In humans, disseminated neonatal herpesvirus infection affects many organ systems including the liver, lungs, skin and/or mouth, and CNS (encephalitis), resulting in mortality of 61% (179).

All herpesviruses are capable of establishing latent infection in their natural host and infected individuals suffer periodic reactivation of disease. Simplexviruses establish latency in the sensory ganglia which innervate the site of primary infection. Reactivation can result in asymptomatic shedding or recurrence of lesions (145, 179).

The severity of disease caused by these four primate herpesviruses depends on the infected host. In xenogeneic primate hosts, both HSV-1 and B virus cause severe infection. Human infection with B virus causes encephalitis which, untreated, leads to

70-80% mortality (175, 180), and is therefore characterized as a biosafety level 4 agent (138). SA8 and HVP-2 are not known to cause disease in primates other than their natural host.

Different levels of pathogenicity can also be observed during infection of mice with these primate alphaherpesviruses (140). All but SA8 cause severe encephalitis in mice, although for B virus and HVP-2 disease severity varies by strain (141). Interestingly, different strains of HVP-2 do not display a range of virulence in mice, but fall into two distinct phenotypes, either highly neurovirulent or entirely non-pathogenic (141). Table 1A illustrates the genome and virulence traits of these simplexviruses.

The variable disease outcomes of primate alphaherpesvirus infection in different hosts provides a number of options for analysis of virus-host interactions in order to better understand the mechanisms of pathogenesis (140). For example, lethal HVP-2 infection of mice has been proposed as a model system for study of B virus infection in humans (142).

1.1.2. Virology and taxonomy

Despite the variable clinical presentations, members of the family Herpesviridae have remarkable structural homology. They are icosahedral, enveloped, DNA viruses with a 100-110nm diameter capsid surrounded by an amorphous protein tegument layer. The double-strand DNA genome of the five primate viruses previously mentioned is 152-157kb in size, with a G+C content of 68% to 76% (see Table 1) (144). Herpesviruses are characteristically widespread in their natural hosts and have been identified in a great variety of animal species. They also share the ability to initiate a latent infection capable of periodic reactivation (144).

Table 1 – Features of select human and simian simplexviruses

A – Comparison of genome characteristics and pathogenicity of two human and three simian herpesviruses.

Designation	Natural primate host	Genome (kbp)	G+C (%)	Pathogenicity	
				in other primates	in mice
HSV-1 (Human herpesvirus 1)	Human	152	68.3	Severe infection	Encephalitis
HSV-2 (Human herpesvirus 2)	Human	152	69	Severe infection	Encephalitis
B virus (Cercopithecine herpesvirus 1)	Macaque	157	74.5	Encephalitis	Encephalitis
SA8 (Cercopithecine herpesvirus 2)	Baboon	151	76	None	None
HVP-2 (Cercopithecine herpesvirus 16)	Baboon	156	76.5	None	Encephalitis

B – Comparison of sequence homology among five primate simplexviruses.

Sequence homology (%)	HSV-1	HSV-2	B virus	SA8	HVP-2
HSV-1	100				
HSV-2	75.7	100			
B virus	59.4	60.5	100		
SA8	64.1	68.8	83.3	100	
HVP-2	60.1	61.4	79.1	85.5	100

HSV = Herpes Simplex Virus; SA8 = Simian Agent 8; HVP-2 = Herpes Virus Papio 2

Three subfamilies of Herpesviridae, *Alpha-*, *Beta-*, and *Gammaherpesvirinae* have been identified based on their speed of infection in culture, host-specificity, duration of the replicative cycle, and site of latency within the host. Herpes simplex viruses 1 and 2 (HSV-1 and HSV-2), cercopithecine herpesviruses 1, 2, and 16 (B virus, SA8, and HVP-2, respectively) are all members of the genus *Simplexvirus* within the *Alphaherpesvirinae* subfamily (144).

1.1.3. Viral genes involved in pathogenesis

Herpesviruses manage a sophisticated balancing act with the antiviral host response. Numerous viral functions contribute to evasion of the immune response, as well as repression of cellular pathways unfavourable to viral replication, such as apoptosis and activation of the interferon response (144). Three genes are of particular interest in the analysis of differential gene expression between the simplexviruses, ICP34.5, US11, and ICP0 (see section g. for discussion of the immediate-early proteins).

1.1.3.1. ICP34.5

One of the cellular host responses to viral infection is activation of protein kinase R (PKR) (52). PKR phosphorylates cellular translation initiation factor 2α (eIF2α), and phosphorylated eIF2α inhibits protein synthesis (36). PKR is activated in the presence of dsRNA as well as by the IFNα/β response (11). Although interferon has limited effect on HSV replication in culture, it is a key factor in development of infection *in vivo* (94, 119).

Although many viruses evade the host response by inhibiting the activation of PKR, HSV-1 infection stimulates PKR while producing proteins to prevent the PKR-induced

phosphorylation of translation factor eIF2α (28, 59). In HSV-1, the RL1 gene encodes infected-cell protein 34.5 (ICP34.5, or $\gamma_1$34.5), which is expressed with late gene kinetics during lytic infection (143). ICP34.5 recruits cellular protein phosphatase 1α (PP1α) to dephosphorylate eIF2α. This negates the antiviral PKR response and allows the production of viral proteins (28, 61, 62).

Interestingly, the C-terminal domain of ICP34.5 contains a domain which is homologous to both a cellular growth arrest and DNA-damage protein, GADD34, and a gene which induces differentiation of myeloid leukemia cells, MyD116 (16, 30, 60). The stress-response protein, GADD34, acts in a feedback loop with eIF2α to dephosphorylate eIF2α and permit continued protein translation (120). In addition to the production of ICP34.5, HSV-1 infection also produces an increase in levels of GADD34 protein, either by inducing its expression (independent of phosphorylation of eIF2α), or by influencing its degradation (25).

In addition to PKR activation, HSV-1 infection also induces a stress response in the endoplasmic reticulum which induces the ER resident kinase, PERK to phosphorylate eIF2α in order to shutoff protein synthesis (25).

The importance of maintaining protein synthesis in an infected cell is illustrated by the variety of mechanisms used by HSV-1 to counteract this antiviral response (115).

ICP34.5 has been designated a neurovirulence gene in HSV-1, as mutants display a phenotype of attenuated virulence *in vivo* (13, 29). ICP34.5 mutants are replication deficient in cells with a functional PKR system, but achieve a near-wild-type level of infection in $PKR^{-/-}$ cells, demonstrating the specific target of this viral protein (95). HSV-1 viruses carrying large deletions or the introduction of a stop codon within the

ICP34.5 gene (29) were used to illustrate the sensitivity of different cell lines to the loss of ICP34.5. Chou *et al* noted that neuroblastoma cells suffered arrest of protein synthesis and cell death when infected with ICP34.5 mutant HSV-1 viruses, while in cells of non-neuronal origin, the infection with mutant viruses was indistinguishable from that of wild-type HSV-1. In the absence of ICP34.5, and coincident with viral DNA synthesis, a complete cessation of both viral and cellular protein synthesis occurs which is reminiscent of apoptotic cell death (31).

1.1.3.2. US11

US11 is also an inhibitor of PKR (110, 143), but uses a different mechanism than that of ICP34.5. US11 encodes an RNA-binding protein (86) present in the tegument at approximately 2,000 copies per virion (143).

US11 is normally expressed with late kinetics, following the replication of viral DNA (73). In HSV-1, the deletion of ICP34.5 can be compensated by a change in the expression kinetics of US11 from late to immediate-early, thus restoring protein synthesis (21, 22, 59, 114). ICP34.5 deletion mutants without a compensatory mutation are unable to synthesize γ2 proteins, including US11 (31).

A surprising finding following the sequencing of B virus is that despite its lethal phenotype, B virus lacks the RL1 open reading frame (125). Subsequent sequencing of SA8 and HVP-2 in our laboratory reveals that this is the case for all three of these simian simplexviruses (166, 167). These viruses may have alternate mechanisms for disabling the PKR response which may involve alterations in the expression kinetics of US11 or other viral proteins. Despite the lack of RL1 ORF in the simian simplexviruses, this study employs several microarray probes targeted to this region of the genome, in order

to detect transcripts from unrecognized open reading frames. The possible detection of transcripts from this region, as well as information about the kinetics of US11 will add to the understanding of the pathogenesis mechanisms of SA8 and B virus.

1.1.4. Viral genome structure and homology

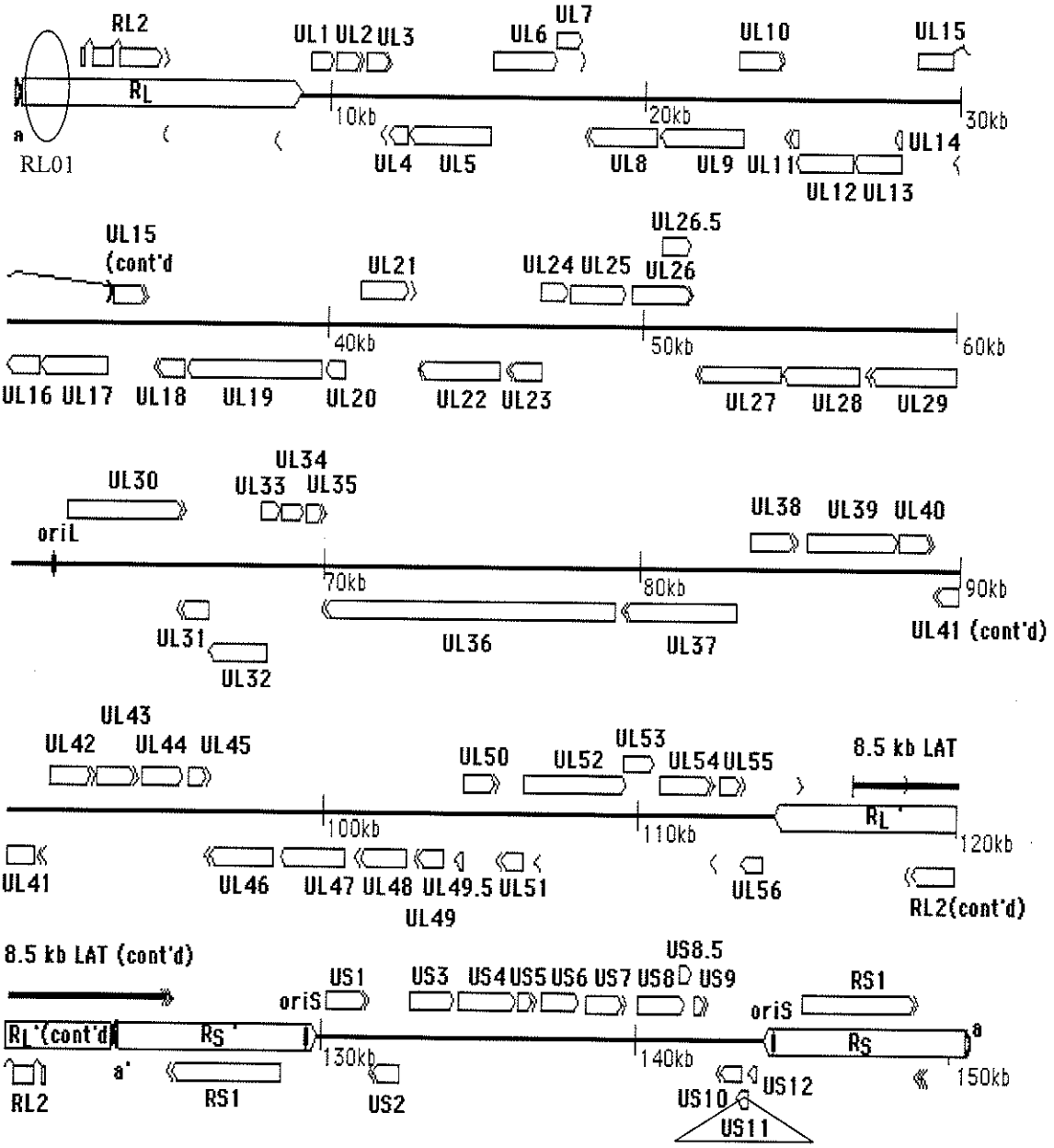
Complete genome sequences are now available for the simian herpesviruses B virus (125), SA8 (166) and HVP2 (167). HSV-1, SA8, B virus, and HVP-2 have linear dsDNA genomes of approximately 155kbp with a GC content of 68-76%. These primate herpesviruses all have a "type E" genome arrangement, consisting of two unique segments (unique long (UL) and unique short (US)) each flanked by a pair of inverted repeats (144). B virus is unique in that it contains duplicate origins of replication, resulting in six origins compared to three in the other alphaherpesviruses (125), although this may be a strain-specific feature (R. Eberle, personal communication). Figure 1 illustrates the features and ORFs of the SA8 genome.

Only two HSV-1 genes are spliced within the coding region, RL02 and UL15. The intron of RL02 is transported to the cytoplasm, although the reason for this is unknown (47). The intron of UL15 contains ORFs UL16 and UL17 in antisense orientation (106). In HSV-2 a third gene contains a splice site within its ORF, RL1, which encodes the virulence protein ICP34.5 (105, 133).

A few HSV-1 genes overlap, or are collinear with other genes, including UL26 and UL26.5 in which the mRNA of UL26.5 is initiated within the UL26 ORF (143). Some genes in HSV-1 are produced from an overlapping antisense transcript, for example HSV-1 RL1 and ORFs O and P (131, 143).

Figure 1 – Map of the open reading frames and features of the SA8 genome. Polyadenylation sites are indicated by single chevrons pointing in the predicted direction of transcription. The region upstream of RL02 where RL1 is located in HSV-1 is noted by an open oval. The US11 gene name is framed by an open triangle.

Reproduced from Tyler *et al.*, 2005 (166) with permission from Elsevier.



Nearly half of all simplexvirus genes are coterminal with at least one other ORF. Transcription is initiated at the appropriate promoter for each gene, but it terminates at a shared poly(A) signal, sometimes resulting in very long mRNA molecules.

Although the majority of HSV-1 ORFs encode a single protein, UL03 is a notable exception. UL01, UL02 and UL03 are coterminal transcripts, but while both UL01 and UL02 have specific initiation sites, the mechanism of transcription of UL03 has yet to be described (153).

Prior to the publication of the complete genome sequences, similarities were recognized between the simian and human simplexviruses. The first gene homologies observed between the HSV-1 and the simian viruses were those of the DNA polymerase, major DNA binding protein, p40 capsid antigen, and thymidine kinase (66). The antigenic determinants (66) as well as DNA and predicted amino acid sequence (44) of several segments of these alphaherpesviruses were compared. An antigenic subgroup was identified based on the similarity of response to HSV-1, HSV-2, SA8, and B virus, but the human herpesviruses continue to show the closest relationship (66). Recently, a newly described simplexvirus isolated from a chimpanzee (ChHV) was shown to be the closest relative of HSV-2 (100).

Individual genes of the unique short (US) region were compared to those of HSV by Eberle *et al.* in 1993 (44), and these results were confirmed by Tyler *et al.* in 2005 with sequencing of the entire SA8 genome (166). SA8 genes are collinear and homologous to those of HSV with only a few differences. An additional polyadenylation signal is found between US5 and US6, and none is present between US3 and US4, contrary to their locations in HSV. This alters the composition of the co-terminal transcript sets of SA8

compared to HSV. In SA8, US3, US4, and US5 form one co-terminal transcript set, and US5 and US6 another. In both viruses, the US8 and US9 ORFs are also co-terminal (44). The sequencing of SA8 also revealed a lack of polyadenylation signal between UL24 and UL25, thus extending the co-terminal transcript set to include genes UL24, UL25, UL26 and UL26.5. In addition, no poly-(A) signal was identified following UL53, suggesting a co-terminal transcript of three genes: UL52 – UL54. The genome location and predicted function of the SA8 genes, and their homology to the simian viruses and HSV-1 are listed in Table 2.

GeneMark (gene prediction software) analysis of B virus suggested a putative ORF antisense to UL53, designated UL53a (125). No such putative ORF was identified in SA8. However, a polyadenylation signal was identified on the untranscribed strand complementary to UL52 (166), suggesting the possible presence of a yet unidentified transcript.

1.1.4.1. Transcripts of the L/S junction

The repeat sequences flanking the UL and US regions of the HSV-1 genome contain a number of ORFs transcribed in both directions. The two immediate-early transcripts mentioned previously, ICP0 (RL02) and ICP4 (RS01), as well as ICP34.5 (RL1) are transcribed in the same direction on either side of the 'a' sequence which separates the two repeats. Three latency-associated transcripts (LATs) of different lengths are transcribed antisense to RL02. The two smaller LATs, 1.45kb and 2kb overlap with RL02. Transcription of the 8kb LAT continues through the junction into the RS region (92).

Table 2 – Open reading frames and other features of the SA8 genome.

Feature name	Position Start	End	G/C	Strand	Length (codons)	Protein homology (%) B virus	HSV-1	HSV-2	Gene product predicted function[a]
a seq	1	236	88.5						Terminal direct repeat
TRL	237	9207	79.7						Terminal long repeat region
RL2			79.0	+	631	63.0	36.5	39.8	Immediate-early protein ICP0; multifunctional regulatory protein
exon 1	2178	2225							
exon 2	2448	3020							
exon 3	3398	4669							
UL	9208	114322	74.5						Unique long region
UL1	9469	10167	66.2	+	233	65.2	48.7	46.9	Virion membrane glycoprotein L; in complex with gH; membrane fusion
UL2	10235	10990	72.4	+	252	86.1	72.9	71.7	Uracil-DNA glycosylase; DNA repair
UL3	11130	11813	74.5	+	228	74.9	52.9	52.4	Colocolization with ICP22 and UL4 in small, dense nuclear bodies
UL4	11897	12511	74.5	-	205	85.7	55.8	54.2	Colocolization with ICP22 and UL3 in small, dense nuclear bodies
UL5	12536	15163	69.3	-	876	93.3	80.9	81.6	Component of helicase-primase complex
UL6	15162	17198	75.6	+	679	90.3	65.4	67.0	Capsid protein; DNA cleavage/packaging
UL7	17149	18039	73.1	+	297	85.1	62.8	63.5	Capsid protein; DNA cleavage/packaging
UL8	18213	20489	80.7	-	759	77.6	55.7	58.5	Component of helicase-primase complex
UL9	23050	24402	72.2	+	451	90.7	75.7	78.0	*ori* binding protein; helicase activity
UL10	23050	24402	74.6	+	451	86.7	64.2	63.6	Virion membrane glycoprotein M; proposed role in capsid envelopment
UL11	24649	24912	74.3	-	88	78.2	44.8	48.3	Myristylated tegument protein; capsid envelopment
UL12	24855	26702	76.8	-	616	79.8	62.1	60.8	DNase; endoexonuclease; processing of DNA replication intermediates

UL13	26702	28246	72.6	-	515	90.9	65.0	66.9	Virion protein kinase
UL14	28018	28662	73.4	-	215	90.7	64.0	68.2	Minor tegument protein
UL15			69.5	+	736	96.6	87.5	87.2	DNA cleavage/packaging; transiently associated with maturing capsids
Exon 1	28761	29789							
Exon 2	33312	34490							
UL16	29948	31036	77.0	-	363	87.8	62.7	64.6	Capsid associated; DNA cleavage/packaging; located in UL15 intron
UL17	31057	33168	79.3	-	704	86.7	65.9	67.7	Tegument protein; DNA cleavage/packaging
UL18	34642	35595	74.8	-	318	94.3	80.4	81.7	Capsid protein VP23; forms triplexes with VP19C that connect pentons and hexons in capsids
UL19	35743	39876	72.7	-	1378	95.5	86.8	86.5	Major capsid protein VP5; forms pentons and hexons of capsid shell
UL20	40057	40734	71.7	-	226	89.2	65.3	65.8	Virion membrane protein; virion egress; *syn5* locus
UL21	41301	42881	74.9	+	527	85.9	61.2	66.0	Nucleotidylylated phosphoprotein; interacts with microtubules and facilitates intracellular transport of the virus
UL22	43029	45620	77.0	-	864	81.6	57.4	59.7	Virion membrane glycoprotein H; in complex with gL; membrane fusion, entry, cell-to-cell spread
UL23	45797	46933	72.0	-	379	80.5	55.3	57.4	Thymidine kinase
UL24	46868	47650	71.2	+	261	83.8	63.1	61.9	Nonglycosylated membrane-associated protein; *syn5* locus
UL25	47847	49580	75.0	+	578	91.3	79.7	80.1	Minor capsid protein; DNA packaging; possible role in DNA anchoring
UL26	49793	51577	78.7	+	595	76.4	60.4	58.2	Capsid maturation protease
UL26.5	50699	51577	80.2	+	293	66.1	43.8	49.7	Scaffolding protein
UL27	51943	54600	70.7	-	886	88.4	77.4	77.9	Virion membrane glycoprotein B; cell entry; contains *syn3* locus
UL28	54631	56985	74.2	-	785	92.3	81.4	82.4	DNA cleavage/packaging; transiently associated with maturing capsids
UL29	57464	61054	71.5	-	1197	94.4	82.2	81.9	Single-strand DNA-binding protein; key role in

									assembly of DNA replication proteins
oriL	61445	61542	47.8						Inverted repeat region of origin *oriL*
UL30	61764	65444	71.3	+	1227	90.1	77.2	77.7	DNA polymerase catalytic subunit; complexes with UL42
UL31	65389	66303	72.0	-	305	90.0	75.3	74.3	Nuclear phosphoprotein; interacts with UL34; capsid egress from nucleus
UL32	66296	68068	76.0	-	591	87.8	73.9	73.4	DNA packaging; not associated with capsids
UL33	68067	68474	70.6	+	136	88.6	71.5	71.5	DNA packaging; not associated with capsids
UL34	68524	69324	75.0	+	267	86.8	65.4	68.4	Type II nuclear membrane-associated phosphoprotein; interacts with UL31; capsid egress from nucleus
UL35	69442	69786	76.2	+	115	79.8	58.0	54.5	Basic phosphorylated capsid protein VP26
UL36	69965	79177	78.4	-	3071	79.9	58.2	58.7	Very large tegument protein; interacts with UL19 and UL37
UL37	79388	83041	76.9	-	1218	85.3	68.2	67.1	Minor tegument protein
UL38	83508	84887	76.8	+	460	89.7	68.8	66.9	Capsid protein VP19C, forms triplexes with VP23 that connect pentons and hexons in capsids
UL39	85252	88179	70.3	+	976	83.9	66.7	65.4	Large subunit of ribonucleotide reductase
UL40	88237	89214	66.4	+	326	89.5	83.4	83.4	Small subunit of ribonucleotide reductase
UL41	89390	90844	70.7	-	485	90.7	70.7	71.5	Tegument phosphoprotein; virion-associated host shutoff (vhs) protein
UL42	91272	92588	74.8	-	439	65.1	42.9	45.2	Double-stranded DNA-binding protein, DNA polymerase subunit
UL43	92733	93878	80.0	+	382	72.3	42.8	42.5	Predicted membrane-associated protein
UL44	94112	95506	78.9	+	465	76.1	47.8	48.9	Virion membrane glycoprotein C; cell attachment; blocking host immune response
UL45	95717	96241	79.8	+	175	83.3	61.0	56.4	Type II membrane protein; possible role in cell fusion
UL46	96429	98471	76.8	-	681	83.7	53.2	54.7	Tegument phosphoprotein VP11/12; modulates alpha *trans*-inducing factor activity
UL47	98655	100679	79.3	-	675	81.8	58.3	54.2	Tegument phosphoprotein VP13/14; O-glycosylated; modulate -TIF activity; RNA binding
UL48	101176	102642	72.0	-	489	90.8	67.6	67.0	Major tegument protein VP16; trans-activator of genes

UL49	102951	103778	75.9	-	276	73.1	40.4	35.6	Major tegument protein VP22; binds RNA; carrier of mRNA from infected to uninfected cells
UL49.5	104165	104398	70.6	-	78	58.4	37.7	35.1	Envelope protein
UL50	104413	105516	75.9	+	368	83.1	55.9	53.1	Deoxyuridine triphosphatase
UL51	105673	106359	74.8	-	229	73.6	64.9	63.6	Capsid/tegument-associated phosphoprotein
UL52	106397	109558	75.2	+	1054	90.2	70.7	69.3	Component of helicase-primase complex
UL53	109510	110514	71.1	+	335	90.4	66.8	68.6	Membrane glycoprotein K; virion egress; contains *syn1* locus
UL54	110790	112322	75.9	+	511	75.9	50.2	51.2	Immediate-early protein ICP27; regulates some early and all late gene expression
UL55	112612	113184	71.2	+	191	84.2	61.8	62.4	Nuclear matrix-associated protein
UL56	113344	114024	73.8	-	227	56.6	29.2	30.1	Type II membrane protein; involved in virus pathogenicity
IRL	114323	123293	79.7						Internal copy of large inverted repeat region
RL2			79.0		631	63.0	36.5	39.8	Immediate-early protein ICP0; multifunctional regulatory protein
Exon 3	118861	120132							
Exon 2	120510	121082							
Exon 1	121305	121352							
a' seq	123294	123529	88.5						Inverted copy of *a* sequence
IRS	123530	129918	82.2						Internal copy of small inverted repeat region
RS1	125207	128707	84.1	-	1167	76.5	60.7	62.4	Immediate-early protein ICP4; regulator of gene expression
OriS	129452	129623	47.7						Inverted repeat region of replication origin *oriS*
US	129919	144385	76.3						Unique small region
US1	130088	131383	73.3	+	432	60.1	34.3	38.5	Immediate-early protein ICP22; required for optimal ICP0 expression
US2	131610	132512	78.8	-	301	78.0	54.3	50.2	Tegument protein
US3	132837	134201	73.8	+	455	82.6	59.3	60.1	Protein kinase; antiapoptotic activity
US4	134314	136128	80.7	+	605	58.6	17.6	33.1	Virion membrane glycoprotein G; entry into polarized cells

US5	136332	136652	78.2	+	107	54.7	27.2	17.4	Glycoprotein J; block apoptosis
US6	136982	138169	75.7	+	396	81.7	53.6	54.5	Virion membrane glycoprotein D; cell entry; interacts with cellular receptors
US7	138375	139574	77.6	+	400	68.2	43.8	46.5	Virion membrane glycoprotein I; in complex with gE; basolateral viral spread
US8	139886	141508	76.2	+	541	73.5	43.9	44.3	Virion membrane glycoprotein E; in complex with gI; basolateral viral spread
US8.5	141453	141761	76.7	+	103	79.4	41.2	45.1	Nucleolar phosphoprotein
US9	141852	142127	75.0	+	92	66.7	48.9	52.8	Tegument protein
US10	142577	143407	75.0	-	277	69.6	37.3	42.0	Tegument protein
US11	143205	143552	79.4	-	116	53.0	36.5	33.9	RNA-binding tegument protein; interacts with protein kinase R
US12	143585	143821	70.5	-	79	57.7	29.5	25.6	Immediate-early protein ICP47; inhibits antigen presentation
TRS	144083	150479	82.2						Terminal copy of small repeat region
oriS	144386	144557	47.7						Inverted repeat region of replication origin *oriS*
RS1	145302	148805	84.1	+	1168	76.5	60.7	62.4	Immediate-early protein ICP4; regulator of gene expression
a seq	150480	150715	88.5						Terminal direct repeat

[a] The predicted function of the SA8 gene products are the same as reported in Perelygina et al., 2003 (125) for the homologous genes of B virus. Reproduced from Tyler *et al.*, 2005 (166) with permission from Elsevier.

Other transcripts, known as L/STs, also span the overlap between long and short segments of the genome. ORF O and ORF P are overlapping transcripts, the former being a frameshift of the latter (132). Both are transcribed antisense to RL1, and ORF P has the unique distinction of being almost entirely antisense to ICP34.5 (92).

One of the benefits of oligonucleotide microarrays is that probes can be specifically designed for each orientation of the genome. We have constructed probes for exploration of the expression of LAT transcripts in SA8, as well as the presence of any transcript from the region between the 'a' sequence and the start of RL02 (where RL1 cannot be identified).

1.1.5. Viral replication

Attachment and penetration of the virus is mediated by the interaction of several viral glycoproteins with cellular receptors and coreceptors. Following fusion of the viral and cellular membranes, the nucleocapsid is transported to the nuclear membrane along the microtubular network. Tegument proteins thus released with the capsid may remain in the cytoplasm, continue to be associated with the capsid, or be transported to the nucleus. These proteins play important roles in establishing a cellular environment favourable to viral replication (143).

Viral DNA circularizes upon introduction into the nucleus. The cellular RNA polymerase II (RNApolII) is responsible for transcription of viral genes, but various viral factors also play a role (143). The following section describes the cascade of viral gene expression in more detail.

Seven viral proteins are required for viral DNA synthesis, which takes place in replication compartments within the nucleus. HSV has three viral origins of replication,

OriL and two copies of OriS. DNA synthesis proceeds by a combination of rolling-circle replication and extensive homologous recombination, resulting in progeny head-to-tail concatemers of the four viral genome isomers by inversion of the two unique sequences between the inverted repeat sequences (143).

Capsid assembly, cleavage of DNA concatemers, and packaging occur in the nucleus. The nucleocapsid acquires its tegument and envelope by budding from the inner nuclear membrane and is transported through the Golgi network where final maturation of the viral glycoproteins takes place. It is unclear whether the initial viral envelope is maintained throughout this transport, or whether the virus de-envelops at the outer nuclear membrane and is subsequently re-enveloped at entry into the Golgi network. Viral particles may be released into the extracellular environment or guided to intercellular junctions by the viral glycoproteins (143).

During SA8 infection, envelopment of capsids is observed at the inner nuclear membrane as well as during virus budding into the endoplasmic reticulum or cytoplasmic vesicles, the latter occurring more frequently at late times during infection when the nuclear membrane started to degrade. SA8 capsids containing viral DNA are first detected at 5-6 hours post-infection (hpi). Enveloped, mature virions are observed in the extracellular environment no earlier than 8hpi (14). An increase in B virus DNA replication is noted by 4hpi. Mature virions are detected at 6 to 10hpi, and peak levels occur at 24 to 28hpi (67, 180).

1.1.6. Temporal HSV gene expression

Regulation of herpesvirus gene expression is very complex and reflects the complex interactions with the host cell and the sophisticated viral pathogenic mechanism. HSV

gene expression is a tightly regulated cascade and genes can be generally divided into three categories (77, 143). These categories are identified by the requirement for protein synthesis or viral DNA replication, as determined by inhibitors of these processes. Indeed, their expression in an unmodified system does not follow clear temporal boundaries, but rather illustrates a continuum of expression.

Genes that do not require *de novo* synthesis of viral proteins for transcriptional activation are designated α (alpha), or immediate-early (IE) genes. Alpha gene products are all transcriptional regulators and they are required for the expression of β (early) genes. Beta genes include the proteins required for viral DNA replication and nucleotide metabolism. Within this kinetic class some genes are expressed almost simultaneously with the immediate-early genes while other transcripts appear later, these kinetics are described as β_1 and β_2, respectively. The third category of genes can be divided generally into γ_1 and γ_2 (or late (L) genes). The transcription of late genes is either enhanced by (γ_1), or entirely dependent on (γ_2), synthesis of viral DNA. These genes encode many of the structural viral proteins. And finally, an additional category of expression is composed of genes expressed during viral latency, known as latency-associated transcripts (LATs).

A number of inhibitors are commonly used to distinguish the kinetic classes of gene expression based on the requirement for *de novo* protein synthesis or initiation of viral DNA replication. Cycloheximide is commonly used as an inhibitor of protein synthesis in order to distinguish immediate-early gene expression from that of early genes. Acyclovir is often used to interrupt viral DNA synthesis, thus inhibiting the expression of many late genes.

The virion tegument protein VP16 stimulates the transcription of the five immediate-early genes in HSV-1: ICP0 (RL02), ICP4 (RS01), ICP27 (UL54), ICP22 (US01), and ICP47 (US12) (18, 131, 143). Promoters of the immediate-early viral transcripts contain a number of binding sites for cellular transcription factors, while the early, and late (both γ_1 and γ_2) transcript promoters have progressively fewer such sites (143).

A key feature of α promoter sequences in HSV-1 is the TAATGARAT sequence upstream of the cap site (51, 102, 131). This sequence serves as a binding site for a transcription complex composed of viral protein VP16, and cellular factors Oct-1, host cell factor 1 (HCF-1), among others (89, 127, 143). The initial interaction of VP16 and HCF promotes recruitment of Oct-1 to the complex, and binding to the viral promoter sequence (90). HCF does not interact with viral DNA but stabilizes the enhancer complex (79, 89).

VP16 binds the GARAT region of the promoter (157), while Oct-1 binds the preceding TAAT sequence. Oct-1 binding is further enhanced by an ATGC sequence immediately upstream of TAAT (7, 38). This complete octamer site (ATGCTAAT) is found in the promoter sequence of HSV-1 ICP0 and is bound more tightly by the VP16 enhancer complex than that of ICP4 which lacks the additional ATGC (38). Additional cellular factors CREBP (cAMP response element binding factor) and transcription factor F2 are able to bind regions of the ICP0 promoter (178).

GA-binding protein (GABP) and SP1 are also involved in transcription enhancement (75, 93, 165). GA-rich regions consisting of tandem repeats of the CGGAAR sequence (165), and SP1 binding sites are observed in IE promoters adjacent to the TAATGARAT binding site. The short distance between the GA-rich element and the TAATGARAT

sequence is critical to its enhancer role (9). In addition, GABP has been shown to interact with HCF, thus supporting the idea that HCF is the coordinator of the IE transcription enhancing complex (171).

VP16 is also necessary for the recruitment of other cellular factors such as TFIIA, TFIIB, TFIID, TFIIH, and TATA-binding protein (TBP) to viral IE promoters (88, 185). The 3bp sequence CTT, known as the determinant (or D) element, is found downstream of the GARAT sequence and influences the selectivity of the VP16 interaction in HSV-1 (70). It has been proposed that although the DNA-binding capacity of VP16 is dependent on Oct-1, VP16 does contact viral DNA at the D element which stimulates formation of the transcriptional enhancer complex (8).

VP16 also interacts with histone acetyltransferase (HAT) proteins p300, CBP, and the hGCN5 complex (108). Although viral DNA is not associated with nucleosomes during lytic infection, histone H3 was observed at γ promoters (64). It is proposed that VP16 is responsible for removing or preventing the accumulation of histones on viral IE promoters (64).

The promoters of β and some γ_1 genes contain CAAT (also known as CCATT) sequences (131). Strict late (γ_2) gene promoters do not have activator elements upstream of the TATA box. However, there are at least two downstream elements, including an initiator element at the start of transcription. Leaky-late (γ_1) promoters have features of both early and late gene promoters, including regulatory elements upstream of the TATA box as well as the aforementioned initiator element of late promoters (178).

ICP4 interacts with the TATA box of viral gene promoters in order to activate transcription (72). However, there are no apparent structural differences between the

TATA boxes of promoters from different kinetic classes and this interaction seems to influence transcript abundance rather than expression kinetics (33).

Viral genes may have numerous functions, possibly determined by their posttranslational processing. Although the majority of herpes simplex virus genes may be dispensable for replication in culture, few are so for efficient infection or reactivation from latency in animal models. Many herpesvirus proteins play a role in changing the intracellular environment to favour viral replication.

1.1.7. Immediate-early gene products

The immediate-early genes of HSV-1 are clustered around the termini of the unique genome segments, with RL02 and RS01 ORFs within the inverted repeat, thus these two genes are present in two copies per viral genome (76, 103, 143). An additional IE gene, not observed in the simian herpesviruses is US01.5, a collinear gene which produces a protein identical to the C-terminal domain of the larger protein, ICP22 (US01) (143). IE gene transcripts are observed between 2 and 4 hours post-infection in cell culture (143).

1.1.7.1. RL02 – ICP0

ICP0 is present in two copies per HSV-1 genome, owing to its location in the repeat sequence flanking the UL region. As one of the few spliced simplexvirus RNAs, ICP0 contains three exon sequences (143). The introns are important however, as regulatory sequences have been observed in intron 1 (55).

ICP0 is not required for HSV-1 infection in cell culture. ICP0 mutant viruses are able to express α genes, but fail to progress through a complete cycle of lytic infection. Virus yield from low-MOI (multiplicity of infection) infections is reduced 10- to 100-fold,

however this deficiency can be overcome at higher multiplicities of infection (48, 147). A challenge presented by analysis of ICP0 mutants is that these viruses are by definition double-mutants for ICP0 and LAT transcripts since the LAT sequence is antisense to ICP0 (143). It is therefore difficult to attribute the specific loss of ICP0 to reduced efficiency of reactivation from latency.

ICP0 stimulates transcription of α, β, and γ genes, and ICP0 interacts with the transcriptional regulator, ICP4 (190). ICP0 also interacts with cyclin D proteins, and elongation factor 1δ (eF-1δ) (80) among other cellular proteins, and may produce histone deacetylation (68), possibly during reactivation from latency.

In the absence of ICP0, the lytic viral cycle is repressed. In the presence of ICP0, viral genomes, either recently uncoated or existing in the cell in latent form, can be replicated (109). It is proposed that among other functions during lytic infection, ICP0 interferes with the rearrangement of viral DNA into a nucleosomal structure which would otherwise prevent access of transcription factors to the viral genome (56).

During lytic infection, viral DNA is localized to nuclear domain 10 (ND10) structures for replication (143). These structures are composed mainly of the PML (promyelocytic leukemia) protein, which recruits other proteins such as Daxx, CBP, Sp100, including a number of transactivators and transrepressors (135). The interaction of viral DNA with ND10 is mediated by association of ICP27 and ICP4 with ND10 component Daxx (161). ICP0 mutants are sensitive to interferon (112), and the success of this antiviral response is dependent on PML (24).

The structure of ND10 is not required for viral replication, and indeed it is actively disrupted by ICP0. Newly synthesized ICP0 is transported to the nucleus early in

infection where it participates in the disruption of ND10 structures, and the degradation of several other proteins (56). One of the roles of ICP0 is as an E3 ubiquitin ligase and it contains two sites capable of this activity. In keeping with this function, it possesses the expected ring-finger domain necessary for ubiquination of target proteins for proteasomal degradation (56). Disruption of the ND10 structures requires a functional ring finger domain in ICP0, and can be arrested by proteasome inhibitors (46, 123).

Interactions have been noted between ICP0 and several cellular proteins, leading to the hypothesis that ICP0 plays a key role in the degradation of specific cellular targets, thus accelerating the recycling of proteins needed for viral replication (178). ICP0 exists in at least six major isoforms during the course of viral infection. It undergoes a variety of posttranslational modifications, including phosphorylation by the viral kinase UL13 and cellular kinase cdc2 (1, 122). These modifications take place in different cellular compartments, and it has been proposed that these modifications may allow ICP0 to perform multiple roles (1). The diversity of effects caused by ICP0 suggests that this immediate-early protein plays many roles in establishing a cellular environment conducive to viral gene expression and productive HSV-1 infection (143).

1.1.7.2. RS01 – ICP4

ICP4 represses its own transcription (113, 143), and may down-regulate the transcription of other IE genes, such as ICP0 (53). ICP4 is a potent transactivator of early and late genes (54, 143, 174). ICP4 binding sites have been found upstream of certain β genes, but there is contradictory evidence that these sites are solely responsible for ICP4 regulation of transcription (143, 151, 152).

ICP4 contains functional domains involved in transactivation, repression, dimerization, nuclear localization, and DNA binding (178). ICP4 is extensively modified, including phosphorylation at multiple sites (184), as well as nucleotidylylation, the latter possibly performed by β or γ proteins (12, 143).

ICP4 also has binding affinity for cellular transcription factors such as TFIIB (154), and TFIID (20, 143), the latter occurring by interaction with the TBP-associated factor (TAF) 250 (20).

Immediately after entry of the viral genomes into the nucleus, both ICP4 and ICP27 as well as ICP0 (as previously mentioned), colocalize with viral DNA at the ND10 domains. Indeed, there is evidence that ND10 substructures may be formed *de novo* or recruited from elsewhere in the nucleus to the viral DNA-protein complexes (49).

In the porcine alphaherpesvirus pseudorabies virus (PRV), only the ICP4 homolog (IE180) is expressed with IE kinetics (71). Genes homologous to HSV-1 ICP0 and ICP27 are expressed with early kinetics.

1.1.7.3. UL54 – ICP27

ICP27 is an essential regulatory protein in HSV-1 (146) which stimulates the expression of β genes, particularly those involved in DNA replication (168), as well as enhancing γ gene expression (137). This viral regulator also plays a role in inhibiting host gene expression, and RNA splicing (57, 156). Interestingly, although four HSV-1 mRNAs are spliced, three of these are IE genes and, as such are proposed to be expressed before ICP27 exerts its full effect, and the fourth, UL15, is a late gene, and its splicing may occur after removal of the ICP27 block or by another mechanism (178). The

presence of two RNA-binding motifs, preferential binding of intronless viral mRNA and a nuclear export signal implicates ICP27 in translocation of viral mRNA (178).

1.1.7.4. US01 – ICP22

ICP22 is responsible for the activation of cellular cyclin-dependent kinase 1 (cdk1, also known as cdc2) (3). Cdk1 subsequently binds the viral dsDNA binding protein, UL42, and this complex binds topoisomerase IIα in an interaction that continues to depend on ICP22 (2).

ICP22 is phosphorylated by the viral kinases US03 and UL13, but other interactions also occur. ICP22 and UL13 are both required for altered phosphorylation of RNA Pol II, as well as for the optimal expression of a subset of viral genes with late kinetics including US11, UL41, and UL38 (3, 128). In addition, ICP22 interacts with another cyclin-dependent kinase, cdk9. The complex formed of these (and possibly other) proteins binds RNA Pol II in a US3-dependent manner (40).

The HSV-1 US1.5 gene is located within the carboxy-terminal region of the US01 coding region, which encodes the immediate-early protein ICP22. Mutation of the N-terminal (ICP22-specific) region does not affect viral gene expression but attenuates virulence in mice. Deletion or mutation of the shared C-terminal region of these proteins reduces expression of some γ_2 proteins, as well as alterations in mRNA stability such as that of ICP0 (reduced) and cellular cyclin proteins (increased) (143).

1.1.7.5. US12 – ICP47

The only reported function of US12 (ICP47) is to inhibit transport of antigens to the endoplasmic reticulum and prevent their presentation on the surface of infected cells. This is achieved by ICP47 binding to the TAP1/TAP2 complex (65).

1.2. DNA Microarrays

1.2.1. Principle

Although the microarray platform is being implemented for numerous other analyses of the structure & function of DNA, RNA and proteins (69), microarrays are most commonly used for sequence and transcriptional profile analysis. DNA microarrays have been likened to a miniature version of the Southern blot technique. However, microarray technology provides a much higher density of fixed probe sequences than would a membrane array, up to the tens of thousands per experiment (17). In addition, two samples labelled with different fluorescent dyes may be hybridized to the same slide. This effectively doubles the amount of data collected and allows direct comparison of samples. Other advantages include the use of less hybridization solution due to the smaller, non-porous surface of the array; a good tolerance of glass for high temperatures during hybridization and washing; and limited background due to low inherent fluorescence of the surface (116). Microarrays are however the reverse of a Southern blot since DNA probes representing the sequences (or open reading frames) of interest are fixed to the solid surface and challenged with the RNA sample under study. Fluorescently labelled cDNA, reverse transcribed from an RNA sample, is applied to the

array and individual transcripts hybridize to their complementary 'spot' of DNA at specific locations on the array surface. Quantification of the fluorescent intensity provides information on the relative amount of the RNA species of interest at that site (17).

1.2.2. Applications

DNA microarrays have been used to study host gene expression, host-pathogen interactions, as well as genomic and gene expression patterns of bacteria, viruses, and fungi. The advantage of the microarray platform for these studies is that whole and even multiple genomes can be studied simultaneously and the exact sequence of a particular strain does not have to be known. Diagnostic microarrays for clinical or epidemiological investigation can allow detection of a wide variety of different organisms or strains. Pathways related to virulence, novel classification of genes based on common patterns of regulation, as well as prognostic indicators of disease may be examined (for further review, see (17)). Microarray analysis allows simultaneous study of multiple pathways (& host interactions), including genes of unknown function and those expressed at low levels of expression (17).

In the host, both adaptive and innate immune responses can be studied, as well as stress or cell death pathways, and tumorigenesis after infection. In microbial arrays, genetic polymorphisms, indicators of pathogen evolution, drug interactions and synergies, as well as identification of novel gene targets can be studied (17).

1.2.3. Platform selection

Microarrays exist in two general formats: *in situ* synthesized oligonucleotide, and printed (or 'spotted') microarrays. The best known *in situ* synthesized oligonucleotide microarrays are those created by Affymetrix (Affymetrix, CA, USA). Oligonucleotides of the exact sequence in question are synthesized on a silicon chip, as well as 'mismatch' oligonucleotides containing single base-pair changes which allow discrimination between 'true' signal and non-specific background hybridization (17). Affymetrix arrays are single-channel (i.e. one colour) and are distinct from two-colour arrays in their production, sample preparation, image acquisition, and data analysis. Therefore, for the remainder of this document I will limit the discussion to the two-colour spotted microarray technology which was employed in this study.

Spotted microarrays can be created with single-strand oligonucleotides (50-70bp), or PCR products (300-800bp) in solution (17). Robotic deposition by microspotting or ink jetting allows 'printing' of the DNA solution at specific locations on a glass slide with great precision (up to 20,000 spots per cm^2 (116)). Spotted DNA arrays are convenient in that they are readily customizable with probes from any organism. However, the equipment, facility, and trained personnel necessary to print microarrays of good quality may make commercially available printed slides more feasible for smaller-scale investigations.

An advantage of using oligonucleotides is that they are single-strand and therefore bind only one possible transcript, unlike PCR products which are necessarily double-stranded and may produce misleading results in cases where the organism of study has antisense overlapping ORFs (173).

Care should be taken in planning the array layout in order to minimize spatial bias that may affect the accurate quantification of signal intensity. For example, printing an array from one side of the slide to the other, and placing replicate arrays above and below in the same orientation may introduce spatial bias between probe measurements. If the signal intensity varies from one side of the slide to the other, the same probes will be affected and this may create false differences in intensity between two probes. However, if replicate arrays are printed in a different (e.g. mirror-image) orientation, probe measurements from replicate arrays will incorporate spatial differences in intensity across the slide.

Consideration should also be given to experimental design, replicate measurements and statistical analyses. These topics will be discussed in the next section.

Once the slides are printed, sample preparation must be addressed. Harvest of RNA from cells or tissues must be done as quickly as possible in order to preserve a 'snap shot' of RNA abundance as close to the desired biological condition as possible. In addition, precautions should be taken to preserve the integrity of the RNA for optimal labelling and hybridization efficiencies. Amplification of small (limiting) amounts of RNA must be linear, so as not to distort the representation of RNA species in the original pool (116).

Either oligo(dT) or random primers can be used to reverse transcribe RNA into cDNA, depending on the presence of poly(A) tails on all transcripts of interest and whether total RNA or mRNA is being used as a template. Unlike radioactive detection, a relatively large amount of RNA is necessary to overcome the background noise present in most microarray systems ($\geq$ 10 μg of total RNA is needed for each microarray sample, compared to only 0.1ug for radioactive detection) (116).

Fluorophores can be incorporated into cDNA either directly, by addition of a dye-tagged nucleotide to the reverse transcriptase (RT) reaction, or indirectly, via incorporation of an amino-allyl-modified nucleotide followed by binding of the dye molecule in a subsequent reaction. Indirect labelling reduces the difference in incorporation efficiencies between the two dyes caused by different degrees of steric hindrance during cDNA synthesis (10, 139, 173).

1.2.4. Experimental design

Many factors contribute to variation among microarray samples. Heterogeneity of the slide surface, excessive background hybridization, spot size and shape, DNA concentration per spot, as well as scanning and quantification parameters can all affect the measurement of a hybridized signal at a given location on the slide. The intensity of this signal is also influenced during the production of the sample, including efficiency of the reverse transcription, labelling and hybridization reactions (50, 83, 116).

In order to save an experimenter's time, expensive reagents, and extensive repetition of sample preparation, it is expeditious to determine an experimental design which provides statistically significant answers to the questions under investigation before conducting any lengthy experiments.

1.2.4.1. Data 'cleaning'

Some degree of data filtering is often needed in preliminary analysis of microarray data. The majority of intensities in any system are low, and these data can cause considerable error during further manipulations, such as log ratio comparisons (see section 1.2.5.1). At the same time, flaws on the array image such as smeared spots, fleck

of dust, scratches, or high background from insufficient washing can alter the calculation of signal intensity. Most image analysis programs are equipped with 'outlier' detection protocols which allow the user to flag and discard spots whose signal is uneven across the spot (implying a flaw) or for which the immediate background is unusually high compared to other spots. Once these obvious errors have been excluded, questions still remain about the statistical significance of signal intensity. Some form of threshold is applied, such as spot signal greater than two standard deviations above the local background of the spot, in order to flag (or exclude) data which is unlikely to be true signal (129). The algorithms used for adjusting data based on background signal or negative controls can affect the conclusions drawn from an array experiment (50).

1.2.4.2. Replicate measurements

Scientists are familiar with the notion of technical replicates, and this principle is applied to microarray analysis in the form of repeated measurements of the same RNA sample, labelled with different dyes, and hybridized to different slides in order to assess the amount of variability introduced by the microarray protocol. Multiple replicate arrays are generally included on each slide, which allows measurement of the same probe under identical conditions of sample preparation, hybridization efficiency, and slide effects. Probe variants of a different sequence but targeting the same transcript are often included when space permits. Technical replicates allow measurement of the 'noise' in the experimental design in order to better distinguish a true change in expression values (50, 83, 116).

The average of replicate values may be used for subsequent analysis in order to reduce the complexity of calculation (129). However, there are calls for better use of the

information provided by replicate data points, for example analysis of their correlation (155). Recently, powerful microarray analysis programs are available that maintain replicate data information due to their advanced computational capacity (e.g. Partek Genomics Suite, Partek Inc.).

Replicate measurements from different batches of cultured cells, different tissue samples, or different experimental animals constitute biological replicates. Given the high variability of the microarray technique, both technical and biological replicates are necessary to draw reliable conclusions about gene expression, and there are many possible calculations to determine the appropriate number of replicates to achieve statistical validity (116).

1.2.4.3. Randomization

Sources of technical variation in the microarray procedure are numerous and can significantly affect the resulting data. Variation can be random, such as heterogeneity of the slide surface or chemistry, or variation among printer pins, and array spots may be different shapes or sizes due to printing and DNA concentration inaccuracies (17). Any factors which may influence signal intensities and cannot be controlled through experimental design must be reduced to another unbiased source of experimental 'noise' through careful randomization by the experimenter.

1.2.4.4. Reference selection

Analyses by microarray often present a logical reference to which samples can be compared. For example, infected cells compared to uninfected cells, treated compared to untreated individuals in an animal model system, comparative genomic hybridization of a

variant to the prototype strain. In a two-colour system these are generally hybridized to the same slide in order to minimize technical effects and facilitate comparison of the two RNA samples.

There are two general models of comparison of sample to the selected 'control'. A reference design compares all samples to the same 'standard' condition. A loop design may aim to compare all samples to one reference condition, but it does so through comparison of samples to each other and some of these to the reference. This is particularly useful if there are many different sample types and/or material for the reference is limited (83).

In the analysis of microbial gene expression, however, it may not be easy to identify an appropriate reference sample. For example, the comparison of viral gene transcripts between infected and uninfected samples will yield only information on the increase in viral expression over 'background' levels of technical noise present on the viral gene spots.

Pooled RNA samples prepared from samples known to express all the genes of interest can replace a single biological standard. However, the same preparation must be used for all arrays related to this comparison as valid microarray results hinge on its unchanging composition, as well as adequate representation of all transcripts (188). Commercially available RNA preparations exist for some organisms, such as the human and mouse Universal Reference RNAs from Stratagene (La Jolla, USA), but for the majority of analyses, in-house preparations are created which limits comparison of results from one laboratory to another.

Genomic DNA (gDNA) has been used as a microarray reference condition, particularly for the study of microbial gene expression including *Mycobacterium tuberculosis* (160), *Borrelia burgdorferi* (136), and *Vibrio cholerae* (187). There are conflicting results on whether using gDNA or pooled RNA as a reference is preferable (87, 177, 181), but the choice falls mainly to the availability of reliable pooled RNA standards. The representation of genes of interest in these pools is also of concern if analysis is directed towards genes which are not always expressed such as developmental factors, or genes specific to minor tissue components (181).

However, the use of gDNA can present technical complications during normalization and analysis of the data, since the gDNA and cDNA to be compared are by definition not treated with precisely the same techniques prior to hybridization. Precision in preparation and quantitation of the gDNA is critical as it is necessary to make the assumption that gDNA signals on all slides (and in both Cy3 and Cy5 channels) are identical.

In addition to a reference sample, exogenous control RNA may be added in order to provide another benchmark for microarray accuracy (169). The use of exogenous spikes has advantages over other reference systems, in that different sequences can be used to generate known ratios of control RNA in different samples, which can help to normalize data during fold-change analysis. The variation in the amount of different controls can also be designed to span the full range of signal intensities, creating a dose-response relationship that may facilitate assessment of the precision of the array system, as well as illustrate the sensitivity of each individual slide (169).

For microarray investigations of systems in which broad changes in gene expression are known to occur, exogenous controls can be invaluable for normalization (170). During herpesvirus infection, cellular mRNA synthesis is inhibited following rapid degradation of existing cellular mRNA. In addition, viral genes are vigorously transcribed immediately after infection by the action of viral and cellular transactivators. These two characteristics of RNA abundance during herpesvirus infection make any use of 'invariant' cellular genes or global array signals for normalization highly inaccurate. As described in Results, we employed the Lucidea Universal Scorecard (GE Healthcare, Canada) system of exogenous RNA spikes for normalization of SA8 microarray samples.

There are a variety of experimental design choices as well as statistical analyses commonly used for microarray data, and misuse of these can drastically alter conclusions (130).

The interpretation of microarray data requires extensive use of statistical and pattern recognition techniques. The commonly used proprietary software associated with DNA microarray equipment can perform many of these techniques, but additional statistical analyses are needed to find higher order patterns in the data. Appropriate choices include non-linear procedures such as cluster and classification tree analyses. One objective of this project is to determine an appropriate protocol for analysis of small, highly variable microarray data such as that generated by the array for SA8.

1.2.5. Determining differential expression

The guiding principle for analysis of gene expression by 2-colour DNA microarray is generally a comparison of the expression between a control (or untreated) sample labelled with one dye, and that of a test (or treated) sample labelled with the other dye. The

difference in the fluorescent intensities of labelled cDNA to any given probe on the array provides information on the up- or down-regulation of that gene in the test sample relative to the control sample (17). All repeated measurements, dye-swap experiments, and data normalization strategies are performed in an effort to separate this comparison in RNA abundance from confounding effects of the experimental procedure.

1.2.5.1. Low-level analysis

As previously mentioned, the majority of signals in any microarray are in the low range of intensities, producing a heavily skewed distribution. Since many statistical tests require a normal (Gaussian) distribution of data, applying a logarithmic transformation can facilitate analysis by producing a 'more normal' distribution (116).

A simple comparison of data from the two samples on each slide is to measure the ratio of the intensity of one over the other. Combined with the expression of signals in a logarithmic scale, this leads to the expression of microarray data in terms of log-ratios. This is convenient because log-ratios are more intuitive than simple ratio calculations from intensity values because up- or down-regulated ratios are presented in the same way. The most commonly used log transformation in microarray analysis is log base 2. The expression of a gene whose intensity is unchanged between two samples has a ratio of 1 and a log ratio of 0. A gene whose expression is upregulated by a factor of two has a log ratio of 1. Similarly a gene downregulated by a factor of 2 has a log ratio of -1, whereas its signal ratio would be 0.5 (129).

The necessity for data normalization, and the type of bias observed in the data is often illustrated by a scatter plot of log ratio values on the Y axis ($Log_2(R/G)$) compared to log intensities on the X axis ($Log_{10}(R*G)$) where R = red signal intensity and G = green

signal intensity (otherwise known as an RI plot). This plot illustrates intensity-dependent bias in the data based on the distribution of scatter plot which bends one way or another from the theoretical centre around a log ratio of 0 (34, 129).

A common method for normalization of intensity-dependent bias in 2-colour microarray data is LOWESS (LOcally WEighted Scatterplot Smoothing) regression. LOWESS estimates the dependence of the log ratio on log intensity, and transforms each value to minimize this effect. LOWESS normalization is often conducted on repeated subsets of points, such as those printed with the same pin during slide printing in order to accommodate elements of spatial bias in the correction (129).

However both 'correction' of the data based on the RI plot, and LOWESS normalization depend on the assumption that the two samples in each channel are equal in their intensity values (i.e. globally the samples have the same level of expression for all genes) (130). These tools are valid for microarrays designed to analyse vast numbers of genes, such as those in a mammalian genome, within which only a handful are expected to have altered expression following the selected treatment.

1.2.5.2. Statistical tests

Many microarrays, particularly in the early years of the development of this technology aim to compare two conditions. As described above, data is 'cleaned' of flawed spots, and multiple data sets from different dyes, slides, and/or biological samples are normalized by an appropriate method. Statistical tests such as t-test (or their non-parametric equivalents, such as the Mann-Whitney test) are subsequently employed to determine which genes are statistically different between the two conditions under study (35, 50, 116).

Comparison of different amounts of RNA present in two sample types has also been commonly expressed as fold-change (35). However, investigators sometimes disagreed on what constitutes a statistically significant fold-change of expression. There is often no biological significance for the selection of a fold-change cut-off, and "significant" differences in expression are an arbitrary determination. For example, a change in log-ratio in excess of two standard deviations from the mean has been commonly used to identify differentially regulated genes. This method guarantees that some genes will be identified as statistically significant even if the differences observed are only due to random variation. For this reason, many researchers are steering away from this analysis (116, 173).

Further refinement of microarray analysis protocols has established measurements of absolute intensity values as well as improvements to ratiometric methods (50).

For the comparison of multiple conditions by microarray, analysis of variance (ANOVA) procedures have been used (32, 84).

The difficulty with the application of traditional statistical techniques to microarray analysis is that these tests effectively deal with many data points and few variables. However due to sample and cost limitations, microarray data often has numerous variables (genes) to consider, but relatively few measurements (data sets) (116).

1.2.5.3. High-level analysis / pattern recognition

Explorative methods such as k-means clustering, self-organizing maps (SOM), and principle component analysis (PCA) are being applied more frequently in microarray analysis (50, 116).

Principal Components Analysis (PCA) identifies the key variables that 'explain' the greatest variation among observations within the data set (116). Both clustering and classification analysis are used widely in microarray data analysis, often in the context of identifying samples based on their characteristics (e.g. distinguishing patterns of gene expression associated with malignancy) (17, 116).

Clustering algorithms seek to group data into categories based on similar behaviour in some variable. The dissimilarity matrix is the measure of difference between groups used to divide them into clusters. Generally speaking, clustering algorithms are of two types. Partitioning methods (such as k-means) seek to minimize differences between a given number of clusters, and the optimal number of clusters is explored during the clustering process. Hierarchical methods produce 'nested' clusters where each division of the data is made once, and subsequent distinctions are identified within the previous clusters, thus producing the commonly recognized image of a dendrogram (17, 27, 116).

Intuitively, classification is similar to clustering but the user provides 'guidance' to the algorithm by identifying known features about the data to be grouped. Classification is described as "supervised learning" because predefined classes are used in an algorithm designed to identify the differences between them. Given this *a priori* knowledge about the data, the aim of classification algorithms is to identify the variables which best distinguish one class of data from another (39). Novel assignment of genes can provide information on possible functions according to the similarity of these genes to other members of their class.

1.2.6. Validation of microarray results

Even in a perfectly planned and executed microarray experiment, there are some factors over which the experimenter has little control. Contamination of the printing plates (or multiple PCR products in a reaction used to generate probes), uneven hybridization across the slide, or very low signal intensity values are some of a number of factors that can impair analysis of microarray data despite appropriate normalization protocols. For this reason, microarray studies are often complemented by other techniques such as northern blot, RT-PCR, ribonuclease protection assays or in situ hybridization (particularly for analysis of mixed tissue samples) (17, 116).

Cellular house-keeping genes are often used for normalization during RT-PCR quantification. HSV produces marked changes in cellular gene expression due to the virion host shut-off (vhs) protein encoded by UL41 which induces the degradation of both viral and cellular mRNA as well as the inhibition of cellular mRNA synthesis (143). It has been demonstrated that 18S rRNA transcript abundance is less variable than other house-keeping genes such as GAPDH and β-actin (121).

Dr. Michael Pfaffl has described a new mathematical model for relative quantification of real-time RT-PCR results based on comparison of the PCR reaction efficiency and deviation of crossing threshold (CT) for both target and reference gene detection (126). The advantages of this method are that a) it does not rely on constant expression of a house-keeping gene, and b) no development of an exogenous calibration curve nor precise quantification of standard cDNA is required.

The inverse relationship of increasing CT values to amount of input cDNA in a series of control reactions is linear. The PCR efficiency (E) of the reaction is defined by formula 1, which employs the slope of this linear relationship.

$$E = 10^{(-1/\text{slope})} \qquad (1)$$

A reference gene is selected based on the expectation that its level of expression will vary little under different experimental conditions. Reference and target PCR reactions may be performed under different conditions, so long as the PCR efficiency calculation is made using the same reaction conditions. A control condition is identified among the samples to be tested (e.g. mock-infected, or a standardized pooled sample of RNA). The results are presented as a ratio of differential expression between the control and sample conditions using formula 2 (126).

$$\text{Ratio} = (E_{\text{target}})^{\Delta CT_{\text{target}}^{(\text{control-sample})}} / (E_{\text{ref}})^{\Delta CT_{\text{ref}}^{(\text{control-sample})}} \qquad (2)$$

By this method, differences in expression of the house-keeping reference gene can be incorporated into the calculation of differences in expression of target gene expression between the control and sample conditions. An ideal reference gene would have no ΔCT difference among experimental conditions. In this case, the denominator of formula 2 would become 1, and could thus be ignored (126).

Although this method was developed using the LightCycler (Roche Applied Science, Canada), it is equally applicable to TaqMan (Applied Biosystems, USA) technology, using either gene-specific probes or SYBR Green (Applied Biosystems, USA).

1.2.7. SA8 microarray data analysis

Normalization of viral gene probe intensity data between slides was achieved by scaling data based on the signal intensity of the Lucidea Scorecard calibration spots (see

Results section 5.1.5.3). Mean intensity data provided an easily interpreted picture of viral gene expression in the presence and absence of cycloheximide and acyclovir. In addition, a combination of cluster and classification analysis was developed in order to identify similarities among viral genes, of both signal intensity as well as reduction of expression in the presence of an inhibitor (see Results section 5.4.1). This protocol was conceived by applying statistical techniques previously employed in microarray analysis to the challenge of identifying different kinetic patterns of viral gene expression. Briefly, the microarray data from non-overlapping, non-coterminal genes were subjected to cluster analysis in order to identify patterns of signal intensity which could be characteristic of a particular type of expression kinetics (IE, E, or L). Classification analysis was then used to assign genes, the kinetic pattern of which may be complicated by overlapping or coterminal transcripts, into the previously defined clusters. This protocol provides an objective analysis of similarities in viral gene expression.

1.3. Herpesvirus microarrays

DNA microarray technology has been used to study the gene expression of several herpesviruses, including HSV-1, cytomegalovirus (CMV), and Kaposi's sarcoma-associated herpesvirus (KSHV). DNA microarrays are widely used to study pathogen gene expression patterns as well as host responses to infection. Oligonucleotide microarrays have the advantage of providing DNA strand-specificity in detection of RNA transcripts, which can be of critical importance when studying complex viral genomes.

The first herpesvirus DNA microarray was developed by Chambers *et al.* in 1999 for the study of human cytomegalovirus (HCMV), demonstrating the utility of microarrays

for simultaneous analysis of all herpesvirus transcripts (23). In 2000, Stingley *et al.* constructed an oligo-nucleotide microarray for HSV-1, containing probes for both viral and select cellular genes (158). This HSV-1 array was used for analyses of gene expression in HSV-1 mutants, as well as to expand on previous knowledge of HSV-1 temporal expression kinetics (158, 189). Analysis of HCMV microarray data was conducted by comparison of signal intensity of cellular probe spots between infected and uninfected samples, which is possible since HCMV does not repress host transcription. In the presence of either cycloheximide or gancyclovir, differentially expression viral genes demonstrated signal intensities statistically significantly greater ($\geq$ 3-fold) than signals from mock-infected cells. HSV-1 oligonucleotide microarray analysis employed two techniques: 1) Ranking viral signals between conditions (e.g. treatment with inhibitors) and establishing significant changes by application of a modified Mann-Whitney non-parametric test (158); or 2) Scaling of viral signal intensities between slides, calculation of the relative abundance of each transcript compared to the total viral signal, and subsequent measure of statistically significant changes in expression using Student's t-test (189).

Microarray technology has been applied to several members of the alpha-, beta-, and gammaherpesvirinae subfamilies. A variety of studies of both cellular and viral gene expression have been conducted, including analysis of promoter and cofactor requirements (117, 159, 192); cellular antiviral transcriptional response to infection or reactivation (82, 111, 134, 193); alteration of the cellular environment by means of inhibitors (23, 78, 158); analysis and comparison of temporal gene expression patterns of

the whole viral genome (4, 5, 23, 81, 124, 158); and even detection of RNA packaged within virions (15, 149).

Oligonucleotide microarrays have also been used to analyse the expression of other large dsDNA viruses with distinct temporal classes of gene transcription, such as bacteriophage T4 (101), Red Sea Bream Iridiovirus (99), and White Spot Syndrome Virus (85).

2. OBJECTIVES

- To develop a DNA microarray and analysis algorithm for the study of the viral gene expression kinetics of the simian simplexvirus SA8.
- To compare the expression pattern of SA8 with that of the well-studied prototype simplexvirus HSV-1.

3. HYPOTHESIS

Since HSV-1 and SA8 are closely related and all their genes are homologous and collinear, we expect that SA8 virus will have gene expression kinetics similar to those of HSV-1. However, given the difference in virulence of these two viruses in humans and in mice, it is possible that subtle differences in gene expression may identify candidate pathogenesis genes.

4. MATERIALS AND METHODS

4.1. Cells and virus

Simian simplexvirus SA8 (strain B264, ATCC VR-936) was generously donated by Dr. Heinz Feldman (National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, Canada) for use in this study. SA8 stock was prepared in a biosafety level 2 laboratory by culture in Vero cells at a multiplicity of infection (MOI) of 0.1 PFU/cell, and harvested at 50 hpi. Vero cells were passaged in MEM supplemented with 5% calf serum and 2 mM L-glutamine (L-Glu). For virus propagation the serum concentration was reduced to 2% FBS, and 1% penicillin-streptomycin (PenStrep, final concentration (f.c.) 100 ug/mL penicillin, and 100 IU/mL streptomycin) were added (all reagents from Gibco/Invitrogen, Canada). Infected-cell supernatant was collected after 3 freeze-thaw cycles in a dry-ice/acetone bath. Plaque assays were performed in Vero cells, overlaid with 0.75% agar and stained with 1.5% neutral red. The working titre of SA8 viral stock was 1.6×10^9 PFU/mL.

For microarray studies, human lung fibroblasts, Mrc-5 cells (#CCL-171, ATCC, USA), of a passage number between 10 and 20 were cultured in 185 cm^2 flasks (Nunc, VWR, Canada). Cells were mock-infected or infected at an MOI of 5 PFU/cell. Mrc-5 cells were passaged in minimum essential medium (MEM) supplemented with 10% FBS, 2 mM L-Glu, 1 mM sodium pyruvate, 0.1 mM non-essential amino acids. In the final passage before infection, 1% PenStrep was added to the medium (all reagents from Gibco/Invitrogen, Canada). During infection, the serum concentration was reduced to 2% FBS.

Inhibition of *de novo* protein synthesis, or of viral DNA synthesis, was achieved by treatment of cells with 50 or 100 μg/mL cycloheximide or 10 μg/mL acyclovir, respectively. Stock solutions of 10 mg/mL cycloheximide (Fisher Scientific Canada, Ottawa, Canada) and 1 mg/mL acyclovir (Sigma-Aldrich, Canada) were prepared in 100% DMSO (Fisher Scientific, Canada). 1% DMSO was added to medium without inhibitors as a control.

Cells were treated with medium containing DMSO and/or inhibitors (as necessary), and chilled to 4°C in a refrigerated room for at least 30 mins prior to infection. Cells were maintained at 4°C during virus adsorption (1hr). Time 0 represents the addition of warm medium to the monolayer and start of incubation at 37°C, 5% CO_2. Cells were harvested using 0.05% Trypsin/EDTA (Gibco/Invitrogen, Canada), and all wash & harvest solutions were collected in the same tube in order to maximize yield and pelleted by centrifugation at 500 xg for 10 mins at 4°C. The cell pellet was washed once in PBS pH7.2 prior to addition of RNAlater (Qiagen Inc., Canada, f.c. 90%) in a total volume of approximately 500 μL. The concentration of DMSO and/or inhibitors was maintained in all solutions (PBS, trypsin/EDTA, medium) during harvest, until the addition of RNAlater.

4.2. Construction of SA8 microarray

Two oligonucleotides between 50 and 70 bp in length were designed for each of the 73 viral ORFs as well as for 106 cellular genes of interest. In addition, 5 oligonucleotides (3 sense, 2 antisense) were designed spanning the region of the missing "RL1" gene in the region of 237 to 2177 bp in the SA8 genome.

Array Designer (Premier Biosoft International, California, USA) provides an efficient platform for the selection of microarray probes by simultaneously minimizing secondary structures such as hairpins and dimers, avoiding sequences with cross-homology, and designing multiple oligonucleotides for the same transcript. Oligonucleotides of 50-70 bp in length were chosen within 400 bp of the 3' end of the desired ORF, with a Tm of 75°C +/- 14°C, and a minimum of 70 bp between probes. Viral gene sequences were submitted to a BLAST search (National Library of Medicine, National Institutes of Health, Bethesda, MD, USA) against human mRNA prior to probe design in order to avoid sequences with homology to human gene transcripts. Oligonucleotides were synthesized by Sigma Genosys Canada (Oakville, ON, Canada). Probe sequences and characteristics are found in appendix A.

Oligonucleotides were resuspended in 50% DMSO at a concentration of 0.1 mM. 384-well master plates of oligonucleotides were prepared using the Biomek FX, Laboratory Automation Workstation (Beckman Coulter, Canada). These master plates were used for subsequent array printing. Arrays were printed on Corning UltraGAPS (Gamma Amino Propyl Silane) coated microarray slides (Fisher, Canada), and slides were stored desiccated at room temperature both prior to, and after printing. Slides were printed using a VersArray Chipwriter Pro (BioRad Laboratories, USA), producing spots of approximately 100 μm in diameter, by the NML DNACore Facility (National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, Canada).

Lucidea Scorecard oligonucleotides were prepared and spotted according to the instructions provided by GE Healthcare for use in normalization and calibration of the arrays.

Viral and cellular gene oligonucleotides were grouped within each array. Within the viral and cellular gene areas of the array, oligonucleotide placement was randomly assigned. In the initial array layout, each array consisted of eight sub-grids and printed in quadruplicate on each slide. Housekeeping cellular gene oligonucleotides were located at the upper left-hand corner of each sub-grid in the array in order to facilitate grid alignment. Lucidea control oligonucleotides were spotted in the last three rows of each sub-grid, resulting in 32 data sets of Lucidea spot intensities per slide, evenly distributed throughout the array area. Therefore, each array contained 151 viral probes, 212 cellular probes, 8 replicates of the Lucidea Scorecard containing calibration, ratio, negative, and blank spots, as well as spots containing only print buffer. This resulted in 604 viral and 848 cellular gene probe spots, 640 Lucidea standardization spots per slide. In the final array layout, after the design of additional viral probes, the layout was altered to reduce spatial bias, resulting in six subgrids per array.

Array data was obtained from three independent experiments using cycloheximide, and one experiment using acyclovir. In every experiment, each sample was labelled with both Cy3 and Cy5 in different hybridizations to account for differences in spot intensity due to differences in the chemistry or incorporation efficiency of the fluorophores (also know as a "dye-swap").

4.3. Viral genomic DNA preparation & hybridization

Viral DNA was obtained by extraction from infected cells and purification on a density gradient. Four 150 cm^2 flasks of confluent Vero cells were infected at MOI=1 PFU/cell with SA8 in MEM supplemented with 2% FBS, 1% PS, and 2 mM L-Glu. At 100% CPE (47 hpi), cells were harvested by banging the flask sharply with the heel of

one's hand. After centrifugation for 5mins at 500 xg the cell pellet from each 150 cm^2 flask of infected cells was resuspended in approximately 5 ml TE buffer (10 mM Tris pH 7.5, 10 mM EDTA). SDS (10%) was added to 0.4% (f.c.), proteinase k (20 mg/mL) to 200 ug/mL (f.c.), and cell lysates were incubated overnight at 37°C with rotation.

Viral DNA was subsequently purified based on its density by separation on a NaI gradient. A saturated solution of NaI containing 240 μg/ml ethidium bromide was mixed at a ratio of 1.5:1 with infected cell lysate. Following ultracentrifugation at 173,502 xg for approximately 48 hours (Optima L70K, Beckman Coulter, Canada), a single band of viral DNA was visualized under UV illumination that could easily be distinguished from a less dense band containing the cellular DNA. Approximately 8 mL of the NaI gradient containing the viral DNA was removed by carefully puncturing the side of the centrifuge tube with a needle. Ethidium bromide was removed by repeated extractions with water-saturated butanol until the butanol phase was the same colour as the aqueous phase. Remaining salts were removed by dialysis overnight at 4°C in approximately 2 L of circulating dialysis buffer (10 mM Tris pH 7.5, 10 mM EDTA, 100 mM NaCl), using a 3-12 mL Slide-a-Lyzer cassette (Fisher Scientific Canada). Purified viral DNA was precipitated by adding 2 volumes of 95% ethanol at 4°C, chilling the tubes briefly and subjecting them to ultracentrifugation at 72,128 xg at 4°C for 5 mins (Optima L70K, Beckman Coulter, Canada). DNA pellets were washed in cold 70% ethanol, dried overnight at room temperature, and resuspended in 750 μL TE buffer at room temperature over the course of several days.

Genomic viral DNA (gDNA) was labelled directly by incorporation of Cy3- or Cy5-dCTP in a reaction with Klenow, using a protocol from The Institute for Genomic

Research (TIGR) modified to allow direct rather than indirect labelling ("Microbial Genomic DNA Aminoallyl Labelling for Microarrays" SOP #M009, Rev#0.3, TIGR, USA, www.tigr.org). Briefly, 8 μg of purified viral DNA was passed through a Qiashredder column (Qiagen Inc., Canada) in a volume of 30 μL in order to reduce the size of RNA fragments in order to achieve more efficient labelling. In order to achieve a 2:3 ratio of dye-labelled to unlabeled nucleotide, a labelling mix was prepared containing 2.78 mM dATP, dGTP, and dTTP (each 100mM, Invitrogen, Canada), and 1.67 mM dCTP (100 mM, Invitrogen, Canada). 6.6 μL of random hexamer primers (3 mg/mL, Invitrogen, Canada) were added to the gDNA (f.c. 0.27 mg/mL), and the reaction volume was brought up to 72.4 μL with MilliQ water. Labelling reaction was incubated at 100°C for 5 mins, snap-frozen on a dry ice / ethanol bath, and centrifuged briefly at 10,600 xg. Subsequent steps were performed in the dark, in order to protect the integrity of the fluorophore. To the 72.4 μL DNA/hexamer solution, we added 10 μL of 10x Klenow buffer, 4.8 μL of labelling mix, and 4.8 μL of Cy3- or Cy5-labeled dCTP, and 8 μL Klenow fragment (3 to 5' exonuclease negative, 5,000 U/mL, NEB, USA) for a final volume of 100 μL. After overnight incubation at 37°C, the labelled viral DNA was purified using Microcon YM-30 columns following the same protocol as described for labelled-cDNA purification. Quantification and efficiency of labelling of gDNA was determined by spectrophotometry using the Nanodrop (Nanodrop Technologies, USA).

4.4. PFGE of replicating viral DNA & Southern blot

Infected cells were suspended and lysed in agarose plugs and subsequently subjected to pulsed field gel electrophoresis (PFGE) and Southern blot to visualize the kinetics of

increase in replicating viral DNA. Mrc-5 cells were grown in 75 cm^2 flasks (Nunc, VWR, Canada) as described above and mock-infected or infected at a MOI of 5 PFU/cell. Cells were harvested at multiple time points post-infection using 0.05% Trypsin/EDTA (Gibco/Invitrogen, Canada), and plugs for PFGE were prepared essentially as described earlier (150). Briefly, cells were pelleted by centrifugation at 500 xg, at 4°C for 5 mins. Low-melting point agarose (1.3%) in PBS was melted and maintained at 40°C. 500 µL of molten agarose was added to the cell pellet, quickly mixed, and plugs were cast containing approximately 100 µL of agarose-cell suspension using the PFGE disposable plug moulds (BioRad, USA). Moulds were chilled for 6-10 mins at 4°C to set the plugs, and the 5-6 plugs per sample were transferred to 3 mL lysis buffer (10mM Tris, 10mM EDTA, 1% SDS). Proteinase k (20 mg/mL, Invitrogen, Canada) was added to the lysis buffer to a final concentration of 1mg/ml (150 µL), and samples were incubated for 48 hours at 37°C before plugs were transferred to fresh lysis buffer.

Agarose plugs containing lysed SA8-infected cells were washed in three changes of 10ml TE buffer, the first two at room temperature over 6 hours, the third at 4°C for 14 hours. A 1% pulsed field certified agarose (BioRad, USA) gel was prepared in 0.5x tris-borate buffer (TBE, 89 mM tris-base, 89 mM boric acid, 2 mM EDTA), and plugs were cut in half in order to be inserted into the wells. The PFGE was run at 6 V/cm, with switch times of the 120° field changes increasing from 2 sec to 70 sec during the 18 hour run. Circulating 0.5x TBE was cooled to 14°C throughout.

DNA fragments corresponding to monomers of the viral genome were observed by ethidium bromide staining of the PFGE gel. The PFGE gel was treated with each of the following solutions (in order): depurination buffer (0.5 M HCl) for exactly 30 mins,

denaturing buffer (1.5 M NaCl, 0.25 M NaOH), and neutralizing buffer (0.7M HCl, 1.5M NaCl, 0.5 M Tris base) for at least 30 mins each. The gel was subsequently placed "face down" on a solid support covered with a sheet of 3MM Whatman paper (Whatman, UK) whose ends were bathed throughout the transfer in 7x SSC buffer. A piece of positively-charged nylon membrane (Roche, Canada), pre-wetted with water, followed by several sheets of 3MM Whatman paper (Whatman, UK), pre-wetted with 2x SSC buffer were placed on top of the membrane. The capillary transfer system was completed by adding a stack of paper towel of the same dimensions as the gel, membrane and previous wicking papers to several inches in height and placing a weight on top of the stack. Capillary transfer was allowed to continue for 20-24 hours, and the membrane was cross-linked to covalently bind transferred DNA using a UVC Crosslinker (120 mJ/cm^2; Hoefer, USA).

Purified genomic SA8 (& CeHV-1) DNA was labelled with digoxigenin (DIG) as described in the DIG High Prime DNA Labelling & Detection Starter Kit II (Roche, Canada). The membrane was incubated with hybridization buffer (6x SSC, 7% SDS) for approximately 2 hours at 55°C. Heat-denatured Vero cell DNA was added to the hybridization buffer at a final concentration of 20 μg/mL and the membrane was incubated a further 5mins at 55°C. Heat denatured DIG-labelled SA8 DNA was added at 50 ng/mL and the membrane was incubated at 55°C for approximately 19 hours. Washing was performed in 0.5x SSC, 0.5% SDS. Two short washes at room temperature were followed by two 30 min washes at 68°C. Detection of the DIG-labelled DNA was completed as per the manufacturer's instructions for the DIG Luminescent Detection Kit (Roche, Canada), and hybridized DNA was visualized by fluorography.

4.5. RNA preparation

Immediately following harvest, the cell pellet was stored in 90% RNAlater for up to 7 days. Removal of the cells from RNAlater, and total RNA extraction was performed using the RNeasy Protect Mini kit, including the optional on-column digestion with the RNase-free DnaseI kit (both kits from Qiagen Inc., Canada). Purified RNA was stored at -80°C.

RNA integrity was analyzed by denaturing agarose gel as well as by capillary electrophoresis. Briefly, all materials were treated with RNaseZAP (Sigma Canada), and all water was treated with DEPC. A denaturing 1.5% agarose gel (2% formaldehyde, 1x MOPS, 8 mM sodium acetate, 1 mM EDTA) was prepared, and run in 1x MOPS buffer at 75 V for 1-2 hours. Loading buffer was prepared containing 1x MOPS, 50% formamide, 7% formaldehyde, and 0.04% bromophenol blue. Capillary electrophoresis was performed using the RNA 6000 Nano LabChip on the 2100 Bioanalyzer system (Agilent Technologies, Canada). Briefly, the capillary passages of the chip are filled with polymer to produce a molecular sieve. Up to 5 ng of total RNA and a fluorescent dye are added to each channel. The Bioanalyzer measures fluorescence in each channel during the run, and produces an electropherogram. RNA quality was good in several early experiments, confirming the extraction technique. RNA quality was not verified for subsequent experiments.

RNA for microarray analysis was harvested from three independent experiments of SA8 infection in Mrc-5 cells in the presence and absence of cycloheximide. Preliminary microarray data was gathered on the effect of acyclovir during SA8 infection from one

independent experiment. Time course data was gathered based on two time points within the same experiment, both infections in which cycloheximide treatment was also studied.

4.6. cDNA synthesis and labelling

Aminoallyl-labelled cDNA was generated from 10 μg of total RNA using random hexamer primers (Invitrogen, Canada) in a reaction with SuperScript III reverse transcriptase (Invitrogen, Canada), and amino-allyl-modified dUTP (Sigma, Canada). cDNA synthesis, purification and labelling reactions were performed according to the TIGR protocol "Aminoallyl Labelling of RNA for Microarrays" (SOP #M004, Rev#2, TIGR, USA). Some modifications to the protocol include the addition of Lucidea Universal Scorecard RNA (GE Healthcare, Canada), which was added to the reverse transcription reaction at the recommended amount of 2 μL per 10 μg of total RNA sample. After cDNA purification by Microcon (Millipore, Canada), Cy3 or Cy5 dye esters (GE Healthcare, Canada) were incorporated into the cDNA product during the labelling reaction. The yield of cDNA and the efficiency of fluorescent labelling was determined by spectrophotometry using the Nanodrop (Nanodrop Technologies, USA) and calculated according to the TIGR protocol (SOP #M004, Rev#2, TIGR, USA).

4.7. Array hybridization, washing & scanning

The entire product of the cDNA labelling reaction was used for hybridization. Samples from both 'test' and 'control' conditions were mixed following purification of fluorescently-labelled cDNA. Hybridization was performed as described in the TIGR protocol "Microarray Labelled Probe Hybridization" (SOP #M004, Rev#2, TIGR, USA). However, for hybridization, the buffer described in the protocol was replaced with DIG Easy Hyb buffer (Roche, Canada), containing 0.33 μg/μL salmon sperm DNA (10

mg/mL, Invitrogen, Canada), 0.33 μg/μL yeast tRNA (10 mg/mL, Roche, Canada), and 0.66 μg/μL human Cot-1 DNA (20 mg/mL, Invitrogen, Canada). Lifter slips (Erie Scientific, Fisher Scientific Canada) were placed so as to cover the array area, and 30 μL of hybridization solution was added by capillary action. Slides were placed in a humidified hybridization chamber (Genetix, UK) and hybridized at 42°C overnight.

Slides were washed according to the same TIGR protocol (SOP #M004, Rev#2, TIGR, USA). Briefly, two washes of 4 mins each were performed in low stringency buffer (1x SSC, 0.2% SDS), followed by high stringency buffer (0.1x SSC, 0.2% SDS) at 64°C. Subsequent washes in 0.1x SSC (two washes of 2.5 mins each) and in water (briefly) were performed at room temperature. Slides were scanned within 4 hours of completion on an Agilent DNA Microarray scanner (G2565BA, Agilent, Canada). Hybridized slides were stored desiccated at room temperature, protected from light.

4.8. Microarray quality control

In order to identify probes with excessive cross-hybridization (either cellular cDNA hybridizing to viral gene probes, or vice versa), the net intensity values from relevant data sets were ranked in descending order, and probes with a mean rank value below the threshold were redesigned.

Cross-hybridization data for viral probe spots was calculated based on four hybridizations (two Cy3 and two Cy5) of RNA from uninfected Mrc-5 cells. Cellular cDNA signal intensity on array spots representing viral genes was ranked in descending order for each data set. Probes were sorted based on their mean rank across several data sets and those with a mean rank value of <100 were redesigned using Array Designer according to the criteria previously described.

Cross-hybridization of viral cDNA on cellular probes was determined based on similar calculations of signal rank. Signal intensity of cellular probe spots was measured following hybridization with genomic viral DNA. Data from 12 hybridizations (six Cy3 and six Cy5) was ranked and mean values sorted as previously described. Probes with a mean rank value of <100 were redesigned using Array Designer under the conditions previously described. Cross-hybridization thus identified could also be confirmed by visual inspection of the array image.

In addition to cross-hybridization, viral probes that responded with weak signal to hybridization with genomic viral DNA were also redesigned. Probes that did not consistently achieve signal intensity significantly above the mean background were identified. Using 12 hybridizations of viral genomic DNA, all net intensity values (local background subtracted signal) on each slide were filtered to exclude measurements smaller than the mean background intensity plus two standard deviations. Probes were redesigned if greater than 10% of spot measurements across all slides were excluded. All probes, both initial and redesigned continued to be included in the final array.

4.9. Data processing and analysis

4.9.1. Scanning & quantification

Images were stored in .TIF format for quantification using the Feature Extraction software (version 8.1; Agilent Technologies, Canada).

Grids containing the appropriate layout and orientation of spots were placed on slide images. Sub-grid location was adjusted manually to best reflect the orientation of each image. Default criteria were used for many Feature Extraction measurements. Modifications included restricting the deviation limit to 30 μm from the spot centroid,

which limited the area on the array grid within which the program would select the feature for quantification.

Spots were analyzed using the "cookie cutter" method (default setting) that measures the intensity for the centre of the spot rather than judging the exact location of the edge of the spot. Background subtraction was performed using local background for each spot, measured in a concentric ring immediately outside the spot.

Data for each slide was exported as a .JPEG image file, and as a .TXT data file. Subsequent data analysis was completed using Microsoft Excel (2000, Microsoft, USA).

4.9.2. Image verification

The Feature Extraction software (Agilent Technologies, Canada) calculates many parameters of array intensity values and reports these for both red (Cy5) and green (Cy3) data.

Non-uniformity outlier spots were defined using default settings for both feature and background measurements for the three terms: CV^2, Poisson Noise, and Background, which describe the pixel 'noise' thresholds for a spot whose intensity is uniform. Visually, it was confirmed that features identified as outliers often had smearing of the spot, streaks, flecks and other imperfections within the spot or the immediate background. Population outliers were not identified since the SA8 array contains only four replicates per data set, rather than the minimum of five required.

Features on the array are described as "positive and significant" if the intensity is significantly greater than that of the background (local background of the feature, in the case of the SA8 array), based on a two-sided Student's t-test. In addition, features are identified as "well above background" if they have been described as "positive and

significant" and if the background-subtracted intensity is larger than 2.6 times the standard deviation of background pixel intensity. In the SA8 array data, features were excluded from analysis if they were identified as non-uniformity outliers, or if they did not meet the "well above background" criteria.

4.9.3. Treatment of replicates and variants

Probe replicates are repeats of the identical probe sequence at different locations in the array. For both the SA8 and CeHV-1 arrays, there are 4 probes replicates on each slide since the array is printed in quadruplicate.

Probes that have different sequences but hybridize to the same cDNA target in the sample are identified as probe variants. Two to five probe variants (noted "a" through "e" as necessary) were included for each gene. Probe variants were analyzed separately throughout the data processing.

Four replicate spots of each viral probe are present on each slide, and many slides (in both Cy3 and Cy5 channels) are hybridized using RNA from the same sample type. The mean value for each probe location on the slide (maintaining the 4 replicates) was calculated prior to cluster analysis. Four mean values were used for each probe, each of which was calculated from all available data sets of the same sample type. Final intensity values reported are calculated for the probe with the highest quality score among the probes for a given gene or transcript group. This mean intensity is calculated from all values in one experiment, both replicate spots and multiple slides.

4.9.4. Inter-slide normalization using a gDNA standard

Genomic DNA (gDNA), labelled with either Cy3 or Cy5 was used as a reference signal for cDNA hybridization to viral gene probes. Seventy-five nanograms of gDNA was selected as the optimal amount of gDNA for hybridization. Genomic DNA was labelled in large batches, sufficient for several array experiments in order to reduce variation between different arrays. Since the Lucidea RNA spikes could not be added to the gDNA preparation, Lucidea RNA was also reverse-transcribed and labelled in large batches for pairing with a gDNA sample. During the hybridization of the arrays, 75 ng of labelled gDNA and 60 ng Lucidea cDNA were mixed with a cDNA sample labelled with the other Cy dye when mixing of two samples is indicated, prior to probe denaturing and hybridization to the array.

4.9.5. Inter-slide normalization using the Lucidea Scorecard

Because we cannot make the assumption for these arrays that the majority of genes will have equal expression in every sample, normalization between slides must be calculated based on a set of invariant 'control' genes, in this case, the Lucidea Scorecard probes. Lucidea Universal Scorecard RNA was added to each RNA sample prior to reverse transcription and labelling.

4.9.5.1. Lucidea Calibration Curve normalization

The 32 calibration points quantified on each array provided ample data for creation of a Lucidea Calibration curve of the logarithm of picograms of RNA to mean signal intensity (arbitrary units) following the anticipated 4^{th}- or 5^{th}-parameter logistic curve. This curve was applied to sample data intensity values in order to normalize between

slides, since the Lucidea intensity would be representative of the sample intensity on each slide. In addition, Lucidea calibration was intended to normalize for the influence of different dye on signal intensity since the dose-response curve for cDNA labelled with Cy3 or Cy5 has a slightly different shape.

4.9.5.2. Average Lucidea intensity normalization

The mean intensities of the Lucidea Scorecard calibration and ratio spots were used to calculate a normalization factor (NF) as described by formulae 3 and 4. Data sets (Cy3 or Cy5) were analyzed individually.

$$NF(A_j) = \frac{X_{Luc}(A_j)}{X_{Luc}(A)} \qquad (3)$$

$$NS_{ij} = S_{ij} / NF(A_j) \qquad (4)$$

Where NF is the normalization factor for an individual array A_j; $X_{Luc}(A_j)$ is the mean of the Lucidea calibration and ratio values for a given array A_j; $X_{Luc}(A)$ is the mean of Lucidea values across all data sets from a series of arrays (A); NS_{ij} is the normalized signal for spot i on array j; and S_{ij} is the (unnormalized) signal for the same spot.

Gene intensity values are 'scaled' between slides by the application of NF in order to be able to compare viral gene signal intensities between slides.

4.9.6. Clustering (k-means)

K-means clustering was performed using two data variables: intensity of signals per probe replicate from infected cells treated with the inhibitor, as well as the percentage difference between signals from cells treated and untreated with inhibitor (F, see equation 5).

$$F = (T-Ti)/T \qquad (5)$$

Where F is the percentage of mean signal change observed between a sample treated with an inhibitor compared to a non-inhibited sample at the same time post-infection; T is the mean net intensity from an infected-cell sample; and Ti is the mean net intensity from an infected-cell sample treated with an inhibitor.

Only probes whose signal represented a single transcript were used for clustering, excluding probes that hybridized to coterminal or overlapping transcripts. Clustering was performed using NCSS (NCSS and PASS, Hintze, J., 2001, Number Cruncher Statistical Systems, Kaysville, Utah, www.NCSS.com). The number of clusters was selected based on the percentage of variation between the sum of squares of the values of the two variables within each cluster and the sum of squares of the variables with no clustering. In short, when repeated iterations of the clustering algorithm are performed using the same number of clusters, and the percentage of variation remains stable, no further optimization of the cluster membership is needed (i.e. the movement of data values between clusters increases the variation). As more clusters are added, the "tightness" of the clusters increases, and so the variation of data within each cluster is an increasingly small percentage of the variation in the data when no clusters are applied. When the addition of further clusters does not substantially decrease the percentage of variation, the data are grouped into an "optimal" number of clusters.

4.9.7. Classification analysis

In order to identify threshold values that distinguish clusters of different probes, we used classification analysis. Classification And Regression Trees (CART, version 5.0, Salford Systems, San Diego, CA, USA) software was used to predict cluster membership, given the same two variables used for cluster analysis. The threshold criteria thus

described were subsequently applied to all viral gene probe values in order to assign cluster membership to those probes not included in the clustering process because of coterminal or overlapping transcripts in that region of the genome.

4.10. Quantitative RT-PCR

Primers of 18-22 bp in length were designed using VectorNTI Suite (version 10, Invitrogen) and Primer Express (v2.0.0, Applied Biosystems, Foster City, CA, USA) in order to amplify a 150-500 bp region of the selected gene sequence. Table 3 lists the primers used for qRT-PCR.

18S rRNA qRT-PCR was performed using TaqMan Universal PCR Master Mix, Ribosomal RNA Forward and Reverse primers (50 nM), and Ribosomal RNA Probe, labeled with the VIC^{TM} fluorophore (200 nM), following the manufacturer's protocol (Applied Biosystems). Cycling parameters were 2 mins at 50°C, followed by 10 mins at 95°C, and 40 cycles of 15 sec at 95°C and 1 min at 60°C (combined annealing and extension temperature). 18S rRNA was used as an endogenous control for qRT-PCR reactions because it was determined by Nystrom *et al* to be the house-keeping gene with the most stable level of expression during herpesvirus infection (121). Relative quantification of mRNA between samples was calculated using the method described by Pfaffl (126).

Pooled RNA samples from five experiments, representing four cycloheximide treatments at an early time post-infection, as well as three acyclovir treatments at a late time post-infection, were analysed by qRT-PCR. Total RNA (5 μg) was used as a template for reverse transcription (SuperScript III, Invitrogen) with random hexamer

primers to generate cDNA for qRT-PCR. The detection of RS01 required the use of oligo-dT-primed cDNA due to limited abundance of this transcript.

qRT-PCR reactions were performed in 96-well plate format using SYBR Green dye on the 7500 Fast Real-Time PCR System (Applied Biosystems). The reaction mix contained: 1 x SYBR Green buffer (Applied Biosystems); 3 mM $MgCl_2$ (except for US12 qRT-PCR for which 1.5 mM $MgCl_2$ was used); 0.4 μM dUTP and 0.2 μM dNTPs; 5% DMSO; 0.5 μM each, forward & reverse primers; and 0.625 U AmpliTaq Gold enzyme (Applied Biosystems). Table 3 describes the product size, application, %GC content, melting temperature (Tm), and sequence of all primers used. Cycling parameters for RL02, US01, US06, and US12 (the latter for both HSV-1 and SA8) were as follows: 10 min at 95°C, followed by 40 cycles of 1 min at 95°C, 30 sec at 60°C, and 1 min at 72°C. Amplification of RS01 was performed as above, but used a combined annealing and extension step of 1 min at 68°.

In order to determine the number of genome copies of IE transcripts in the presence and absence of cycloheximide, a standard curve was generated using SA8 genomic DNA. Using the PCR method described above, ten-fold dilutions of purified SA8 gDNA from 0.4 ng/μL to 4 x 10^{-6} ng/μL were amplified using each of the primer pairs for RL02, US01, RS01, and US12. The linear relationship of the logarithm of gDNA per reaction to mean CT value (mean of duplicate measurements) was used calculate the number of nanograms of cDNA in each RNA sample based on the CT of the PCR products. There are approximately 6.14 x 10^6 copies of the SA8 genome per nanogram of gDNA, so it was subsequently possible to determine the number of copies of each IE gene in RNA samples, expressed as copies per microgram of total RNA. In order to standardize the

data between experiments, the number of copies of each of the SA8 IE genes under different conditions was compared to the number of copies in the sample at 5hpi, using the calculation of percentage change expressed in formula 6.

$$F = (T-Ti)/T \qquad (6)$$

Where F is the percentage change in copies of an IE gene per microgram of total RNA in samples under different conditions; T is the gene copies / μg of total RNA at 5hpi; and Ti is the gene copies / μg of total RNA from cells treated with an inhibitor or at a different time post-infection. Student's t-test was used to determine statistically significant changes in copy number under different conditions.

Table 3 – Primers used for qRT-PCR validation.

SA8 ORF	**Product size (bp)**	**Effc** [a]	**Use**	**Tm**	**% G+C**	**Sequence (5'-3')**
RL02	160	2.12	qRT-PCR	87.0	66.0	TCGACGAGACCCAGCTCT ATGGTGCTGTACGAGCCG
RS01	267	1.80	qRT-PCR	89.0	76.0	CGCCAGTACGCCCTCATCAC GTCGTAGCGGCGGCTCAT
US01	405	1.60	qRT-PCR	89.5	76.3	CGTCTGGTCCGCGACTGCTA AAACTCCCGCCGCAGCCT
UL54	375	N/A [b]	Northern blot	88.6	74.4	CCCGAAAACATCGACCAG GCACATCTTGCACCACGA
US12	166	1.69	qRT-PCR	85.7	72.9	CGAGGTGCGAGTGGTTGG GCGTTCCTACGGAATCCG
US06	307	2.08	qRT-PCR	87.8	73.3	GTGCCCGTCGCCGTGTACT GCATGAGGAACCCCAGGTCG
HSV-1 ORF	**Product size (bp)**	**Effc**	**Use**	**Tm**	**% G+C**	**Sequence (5'-3')**
US12	101	2.03	qRT-PCR	78.4	61.4	ATGTCGTGGGCCCTGGA CGCCCCCTTTTATTGATCTCA

[a] Effc = efficiency; [b] N/A = not applicable

5. RESULTS

5.1. Development of microarray normalization protocol

5.1.1. Simian simplex virus microarray design

In order to study the expression of all viral genes, as well as features of the host response to infection, DNA microarrays were designed with both viral and cellular gene probes.

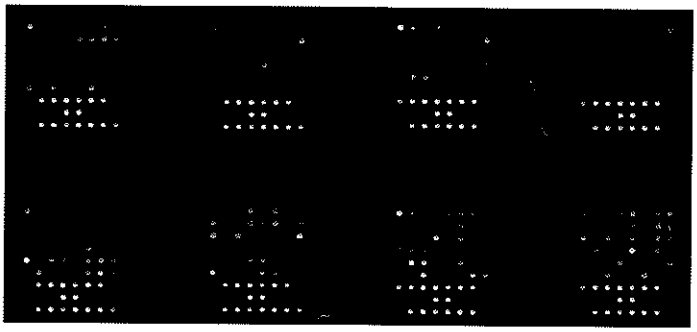
Oligonucleotides were designed for each of the 73 viral ORFs, as well as 106 cellular genes of interest. The Lucidea Universal Scorecard (GE Healthcare) provides a set of 18 control DNA solutions for spotting within the array, as well as RNA spikes which are added to sample RNA for use in assessing factors influencing signal intensity such as reaction yield and hybridization efficiency. Figure 2 shows a representative image (A) of the initial array layout.

5.1.2. Test Arrays

Seven independent experiments were conducted to test the efficiency of the array probes as well as gain preliminary understanding of ideal time points for study and development of data analysis methods. Hybridizations were conducted using RNA samples from mock-infected cells, as well as viral genomic DNA, in order to assess the specificity and sensitivity of the oligonucleotide probes. Figure 3 shows the distribution of microarray signal intensity for each type of data point: viral gene, cellular gene, Lucidea control, or negative (no DNA) spots. The range of Lucidea control intensity values covers all possible values from 0 to saturation at 65,000. Viral probe signals

Figure 2 – SA8 oligonucleotide microarray images.

A – The initial array layout contains 2 probes for each of 73 viral ORFs and 106 cellular ORFs. Probes from the Lucidea Universal Scorecard (GE Healthcare) were spotted in the bottom three rows of each array subgrid. Each array consists of 8 subgrids and 4 copies of the array were printed on each slide. The slide depicted here was hybridized with cDNA generated from Mrc-5 cells infected with SA8 at 8hpi (Cy5, red) and 8hpi+CHX (Cy3, green).



B – Thirty-nine additional probes for viral ORFs were added to the array following assessment of cross-hybridization and probe performance. All initial probes were maintained, and the layout of the array was changed (see text for details). Probes from the Lucidea Universal Scorecard (GE Healthcare) were spotted at the top of each array subgrid. Each array consists of 6 larger subgrids and 4 copies of the array were printed on each slide. This figure shows two arrays with the new array layout hybridized with cDNA generated from uninfected Mrc-5 cells (Cy5, red) and cells infected with SA8 at 13hpi (Cy3, green).

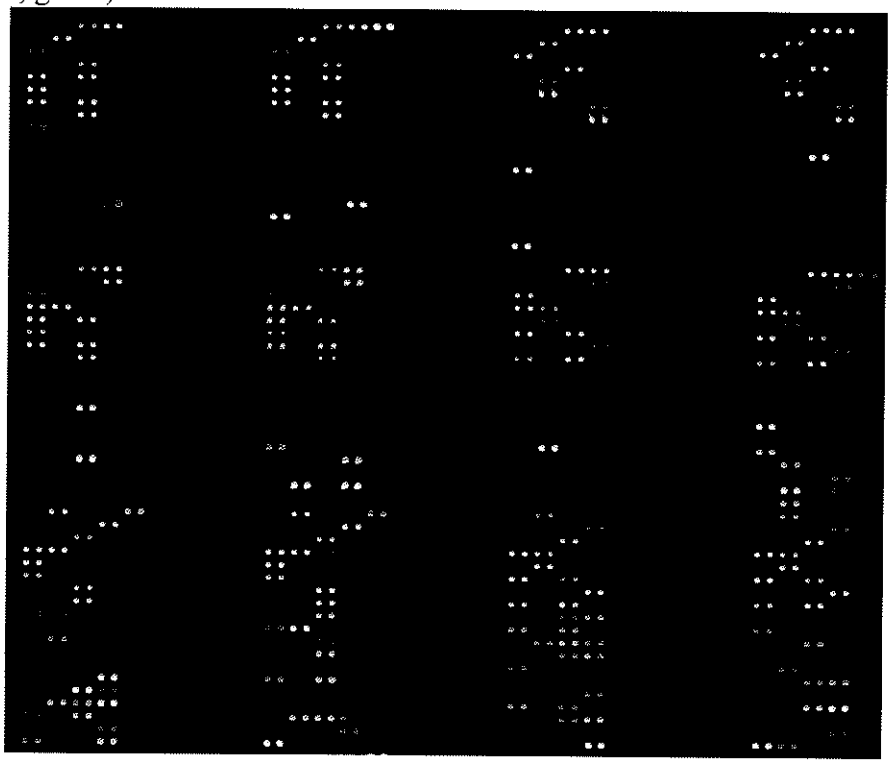
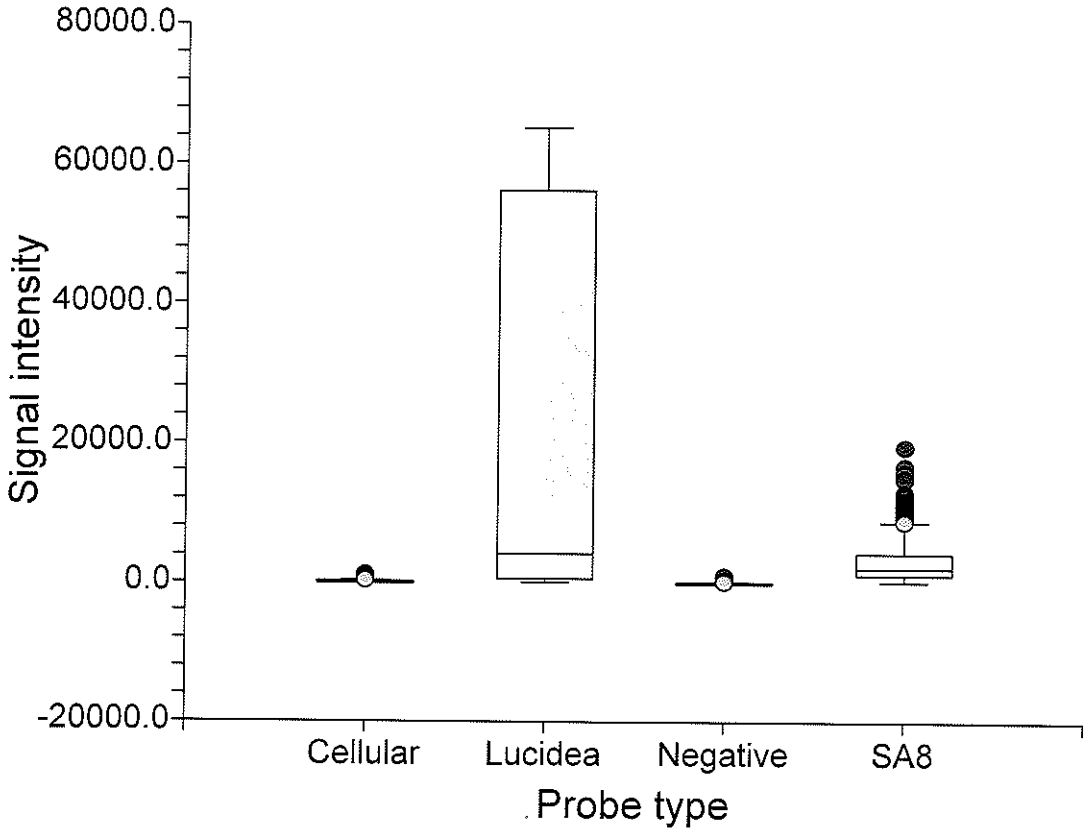


Figure 3 – Box plot of signal intensity of different probe types. Representative data shown is from a single microarray slide hybridized with a Cy3-labelled cDNA sample from SA8-infected cells at 5hpi. The central line inside the box indicates the median signal intensity. The bottom and top of the box indicate the 25th and 75th percentiles, respectively. The adjacent values (T-shaped lines) extend to the furthest observation that is not more than 1.5 times the inter-quartile range (IQR) from the edge of the box. Outlier values are outside the adjacent values and are identified as mild (light grey) for values within 3 IQR of the 25th or 75th percentiles, or severe (dark grey) for more extreme values.



are considerably brighter than those of either cellular gene probes or negative control spots.

Unfortunately, analysis of cellular gene expression was not possible due to the low signal intensity of these probes. This is because the cellular transcripts selected for analysis are not genes that are commonly expressed at a high level. Alterations in hybridization temperature did not substantially improve the cellular gene signals and analysis of these genes was not performed.

5.1.3. Microarray Quality Control

During initial experiments, some probes showed considerable cross-hybridization and/or poor signal. Additional probes were designed and these were included, along with all of the initial probes in the new SA8 array. Figure 2 (B) shows a representative image of the new slide layout.

The complete set of SA8 probes was subsequently analyzed to quantify probe specificity and provide a relative score that measures non-specific hybridization. First, the percentage of signal change observed in a mock-infected sample compared to infected cells was calculated using the formula:

$$Q = (T-M)/T \qquad (7)$$

Where Q is the percentage change between the infected-cell signal and the signal from a mock-infected sample; T is the net intensity from an infected-cell sample; and M is the net intensity from a mock-infected cell sample.

Probe cross-hybridization data were compiled from three independent experiments. The score for each probe was expressed as a percentage - the higher the score, the more reliable the probe. Probe quality scores are listed in appendix A.

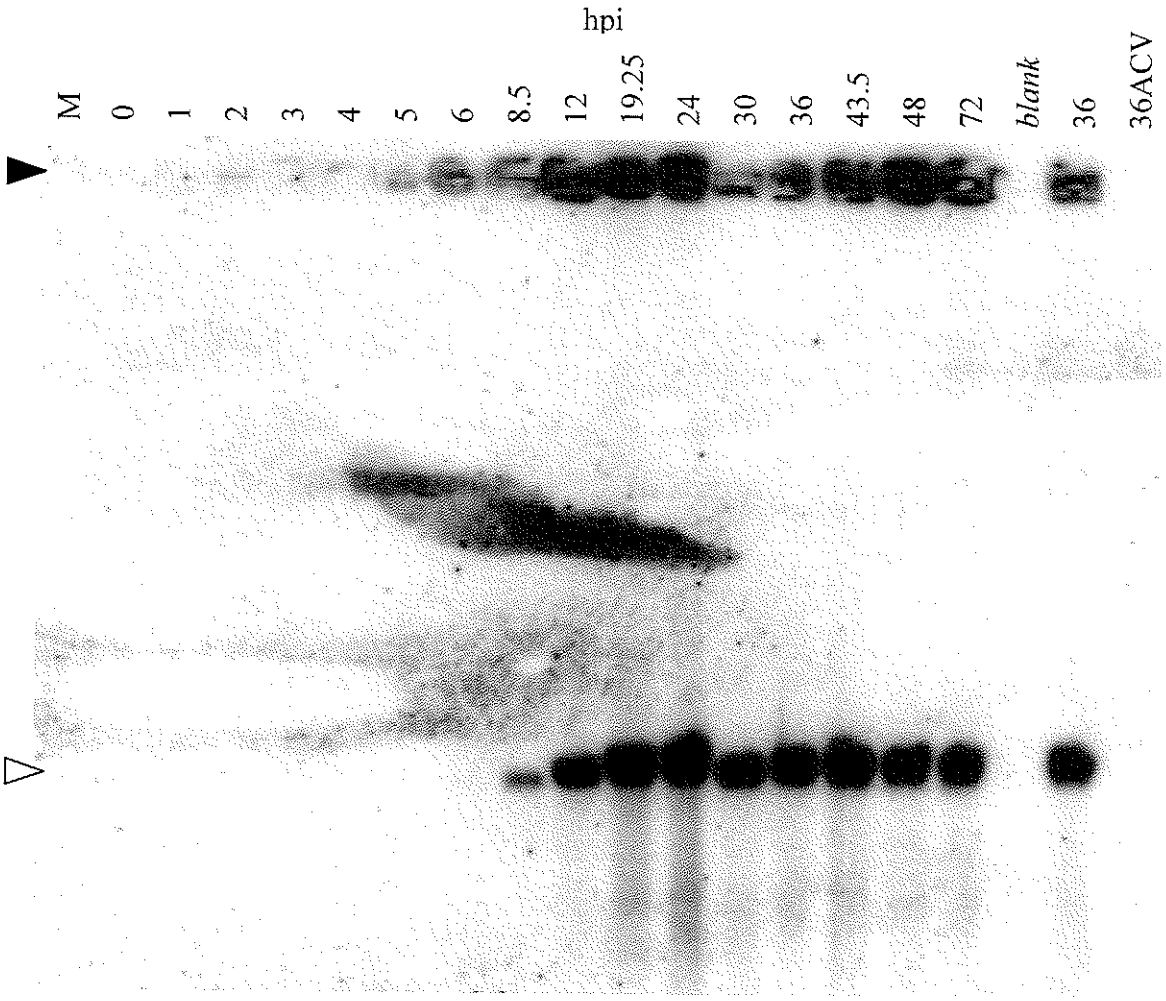
5.1.4. Simian simplex virus infection kinetics

In order to determine the optimal time points post-infection for the study of expression kinetics, replicating viral DNA was detected by pulsed-field gel electrophoresis (PFGE) and Southern blot (figure 4). At early times post-infection, amounts of viral DNA are undetectable in the infected cells by this method. At this time, preceding viral DNA replication, immediate-early genes are expressed. At approximately 5hpi, the start of viral DNA replication is observed by the appearance of very high molecular weight DNA which is retained in the wells during PFGE (solid arrow, figure 4). At this time, early genes are expressed. Shortly after, viral DNA fragments corresponding to the size of a single genome length are observed (open arrow, figure 4). This indicates the beginning of virus maturation and expression of late genes. In order to assess viral transcription of IE and E genes (and CHX treatment) by microarray, SA8-infected cells were harvested at 5hpi. For comparison of late gene transcriptional profiles (and the effect of ACV), 9hpi was chosen as a late time point for microarray analysis.

5.1.5. Development of a reliable microarray normalization protocol

Analysis of SA8 viral gene expression by microarray presents two difficulties, a small number of data points for analysis (only 172 probes), and highly variable signals due to the dramatic change in viral gene expression during lytic infection. The common ratiometric method of comparing signal to a control condition (such as mock-infected cells) cannot be applied to viral gene expression as mock-infected cells would provide an intensity of zero. It would be difficult to prepare a representative pool of viral mRNA for use as a standard, but purified viral DNA provided a possible reference which would

Figure 4 – Southern blot of SA8 DNA at various times post-infection. SA8-infected cells were cast in agarose plugs and lysed prior to pulsed-field gel electrophoresis. Following capillary transfer to a nylon membrane, hybridization was performed using DIG-labelled SA8 viral DNA. Times indicated are hours post-infection; M = mock-infected sample; *blank* = unused lane; ACV = sample treated with acyclovir. The solid arrow indicates the gel wells, and the open arrow indicates the size of monomer SA8 viral DNA.



contain all viral genes. We also included the Lucidea Universal Scorecard probes as an exogenous control for use in normalization.

The Lucidea Universal Scorecard (GE Healthcare) is a commercially available set of artificial genes which are employed as an exogenous control for microarrays in order to facilitate data normalization. The Lucidea Scorecard includes a set of DNA microarray probes which are spotted within the microarray, and corresponding RNA sequences which are added to the sample RNA prior to reverse transcription. Ten sequences are provided in the format of a calibration curve. Each RNA spike mix contains a range of corresponding RNA fragments, from 1 to 30,000 pg RNA per reaction.

For each spot on the array, intensity values were measured as described in the methods. A number of algorithms were explored in order to effectively normalize data between different slides, different dyes, and different sample types (mock-infected, time points post-infection, and/or treatment of infected cells with an inhibitor). Three of these algorithms are described in the following sections:

5.1.5.1. Inter-slide normalization using viral genomic DNA as reference

In order to find a common reference signal for viral gene expression in the SA8 microarray, viral genomic DNA was purified, labelled with either Cy3 or Cy5 and hybridized with a cDNA sample labelled with the other dye. Different quantities of gDNA were assayed, and 75 ng was selected as the optimal amount based on equivalent signal intensity of gDNA and cDNA. gDNA was labelled using direct labelling (see Methods). The indirect labelling technique used to label cDNA proved inefficient for gDNA.

However, analysis of this data proved too challenging for this gDNA comparison to be of practical use in microarray analysis. The genomic DNA signals could not be compared to the cDNA signals due to their different composition, nor could the Lucidea Scorecard RNA spike controls be added to the gDNA reference for normalization purposes. Due to the lack of exogenous hybridization control for gDNA, analysis required the assumption of equal hybridization of gDNA to each slide. Comparison of cDNA / gDNA signal ratios did not lower the inter-slide variability. The gDNA normalization protocol was deemed unfeasible due to difficulties in the comparative analysis between cDNA samples and the gDNA standard.

The SA8 array contains two or more probes for each gene, and the signal intensities of these probes show wide variation due to differences in hybridization efficiency. We developed a technique for applying a normalization factor to cDNA intensity values of different probes based on the ratio of gDNA signal intensity of each probe to the total signal gDNA intensity. However, the variation between probe signals for the same gene was not significantly reduced by employing this manipulation, and it was not applied to the SA8 microarray analysis.

5.1.5.2. Inter-slide normalization using Lucidea Scorecard Calibration curve

Because the Lucidea Scorecard RNA spikes are added directly to the sample RNA, experimental factors that influence cDNA yield or labelling efficiency can be observed by changes in the signal intensity of the Lucidea controls. The calibration curve provided in the Lucidea Scorecard follows a sigmoidal dose-response curve of increasing signal to amount of Lucidea RNA as a proportion of total RNA in the

reaction. In addition, the dose-response curves are slightly different for cDNA labelled with Cy3 compared to Cy5.

Our goal was to 'calibrate' sample data using the Lucidea dose-response curve in each dye channel in order to normalize data taking into account dye, sample, slide, and treatment effects. The Lucidea Scorecard sequences were printed 32 times on each slide in the initial layout, and 24 times in the final layout. Following image quantitation, data cleaning, and background subtraction, the mean value for each calibration point was calculated and plotted against the corresponding relative picograms of Lucidea RNA. The relationship of spot intensity to Log(pg spike RNA / µg sample RNA) follows a 4^{th}- or 5^{th}-parameter logistic calibration curve (also known as a sigmoidal dose-response curve). The formula for this curve is as follows:

$$Y = B + \frac{(T-B)}{1 + 10^{(LogEC50-X)*HS}} \qquad (8)$$

Where Y is the signal intensity; X is the logarithm of the amount of RNA; T is the maximum response (or signal saturation), where a plateau of signal is reached despite increasing amounts of RNA; B is the minimum (baseline) response, when no RNA is present; LogEC50 is the amount of RNA which corresponds to signal at 50% signal saturation; and HS is the hill slope, a measurement of the steepness of the dose-response curve.

A critical feature for selection of exogenous controls is that they have the same length and sequence characteristics as the mRNA (169). The Lucidea control products used are approximately 1000bp in length, compared to the 50-70mer oligonucleotide probes on our array. The number of probe molecules per spot is therefore considerably reduced, and Lucidea controls reach saturation at lower

intensity value than those of oligonucleotide probes. As a result, application of the Lucidea calibration curve to viral gene signals increased the variation in the data by exaggerating signal values plotted on the upper asymptote of the Lucidea curve.

5.1.5.3. Normalization protocol for small, variable microarray data

Although the Lucidea calibration curve cannot be used for inter-slide normalization, the variation of average intensity of Lucidea spots is still proportional to the inter-slide sample variation. The Lucidea RNA spikes are affected by factors which influence sample yield as well as labelling and hybridization efficiency. For this reason, mean Lucidea intensities for each sample on each slide were used to normalize data for comparison.

We applied a technique similar to the one described by Forster in 2003 (50), which uses a normalization factor to 'scale' data between multiple slides based on the mean of their Lucidea control values. The equation used for mean Lucidea normalization is as follows:

$$NF(A_j) = \frac{X_{Luc}(A_j)}{X_{Luc}(A)} \qquad (3)$$

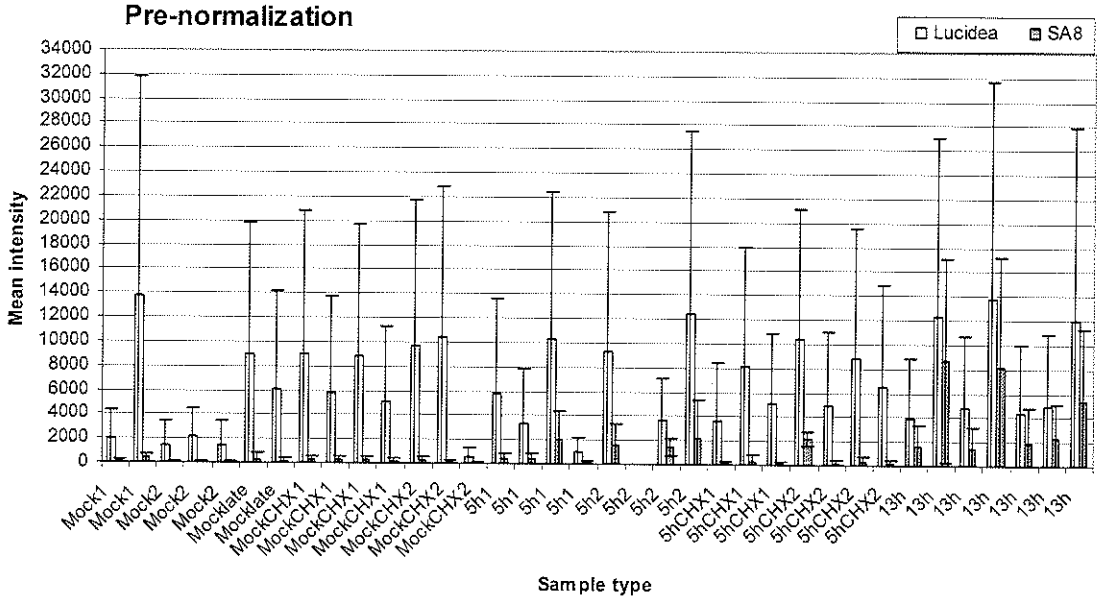
$$NS_{ij} = S_{ij} / NF(A_j) \qquad (4)$$

Where NF is the normalization factor for an individual array A_j; $X_{Luc}(A_j)$ is the mean of the Lucidea calibration and ratio values for a given array A_j; $X_{Luc}(A)$ is the mean of Lucidea values across all data sets from a series of arrays (A); NS_{ij} is the normalized signal for spot i on array j; and S_{ij} is the (unnormalized) signal for the same spot.

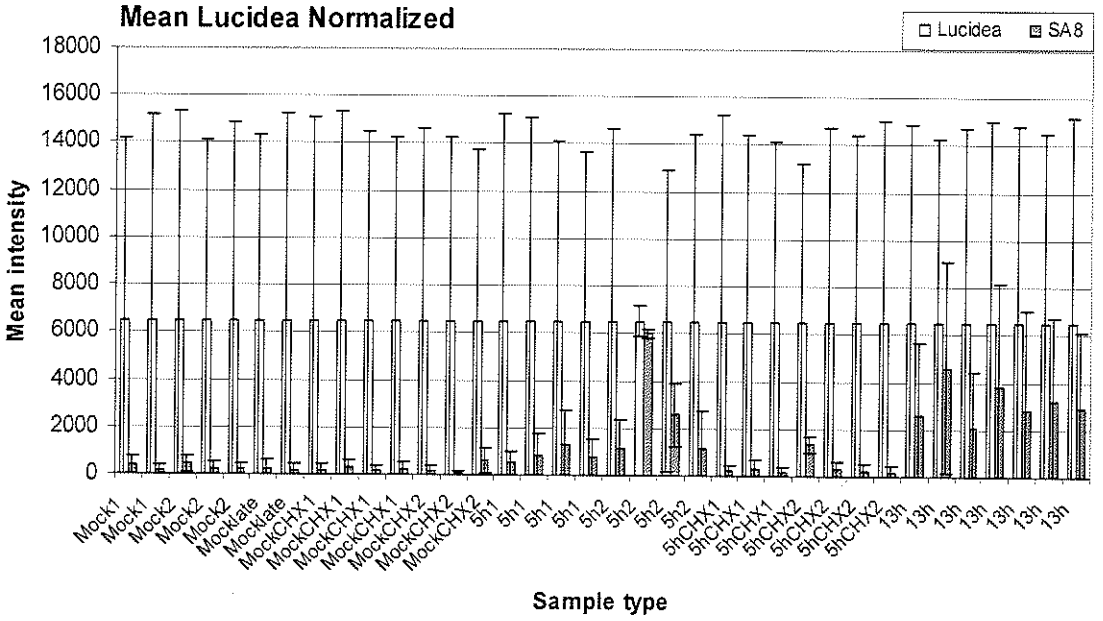
Figure 5 shows the data from one experiment before (fig. 5A) and after (fig. 5B) normalization. In figure 5A, the Lucidea intensity bars (white) show considerable

Figure 5 – Comparison of mean signal intensity of Lucidea and viral probes before and after normalization for each slide in experiment 1. Mean Lucidea intensity (± 1 SD) on each slide is indicated by the white bars, and viral SA8 signal intensity by the dark bars. Mean signal intensities before normalization (A), and after normalization using the mean Lucidea value for each slide (B). Sample type is noted on the X axis: Mock indicates uninfected cells; 5h and 13h are samples from 5 and 13 hours post-infection, respectively; Mock or 5h samples noted as "CHX" were treated with 100 µg/mL cycloheximide.

A.



B.



variation, reflecting the usual inter-slide differences of microarray experiments. In figure 5B, the Lucidea bars are all normalized to the same intensity as a result of the application of formula 4.

Comparison of figure 5A and B also shows that the intensities of the viral signals (dark bars) become considerably less variable. This is particularly obvious for the set of mock-infected samples (first 14 bars), and for the 13h slides (last 7 bars). Aberrant data sets can also be identified, for example viral signals in third "5h2" slide have been skewed by application of mean Lucidea normalization because the signals from both Lucidea and viral probes in this data set were unusually small, suggesting a problem with that array.

5.2. Analysis of simplexvirus immediate-early gene expression

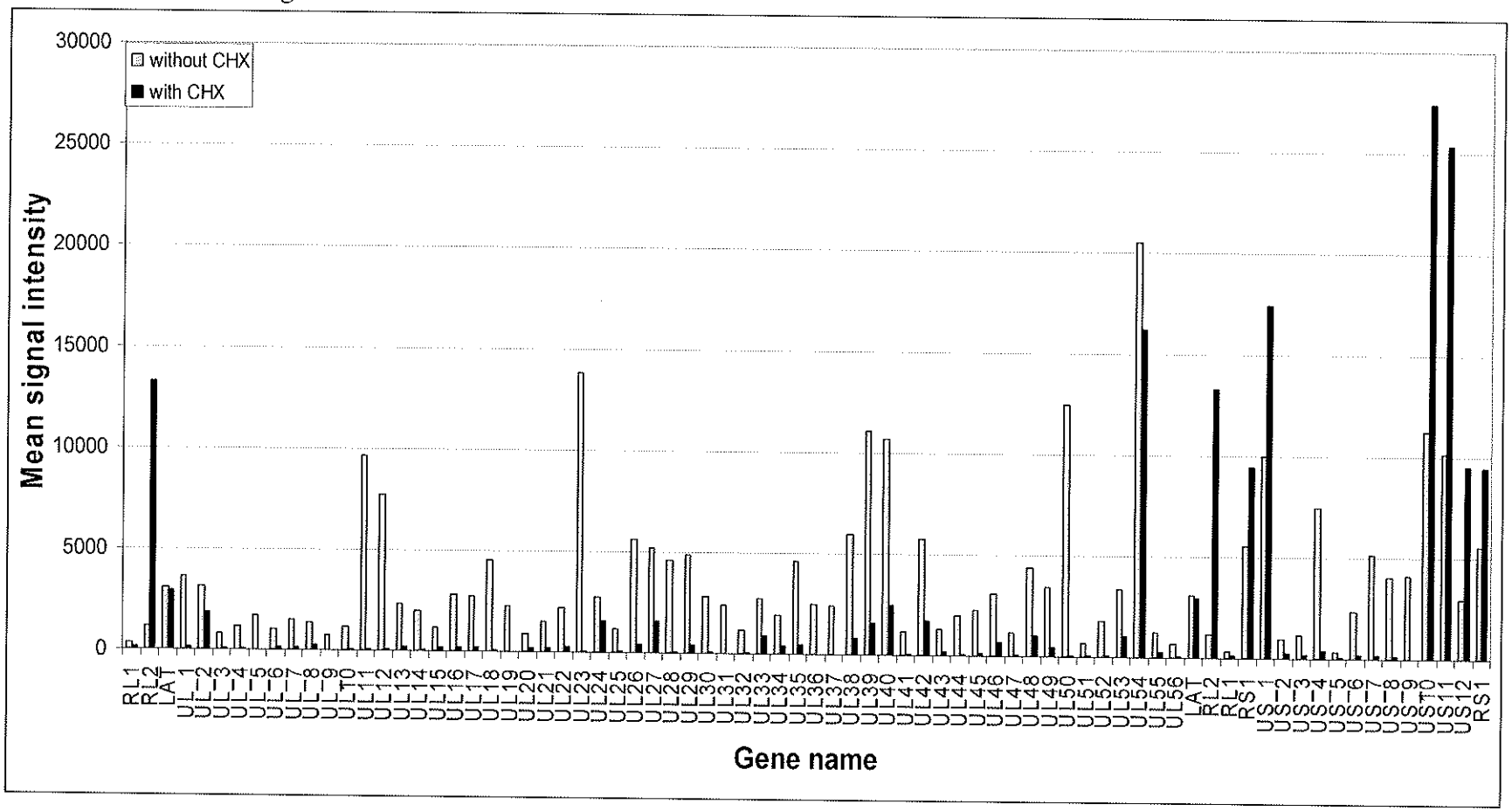
Viral gene expression kinetics of SA8 infection *in vitro* were analyzed by microarray analysis as described in section 5.1.5.3. Cycloheximide (CHX) is an inhibitor of protein synthesis, which prevents progression of productive infection by preventing the synthesis of the early gene products responsible for DNA replication and other non-structural viral functions. Treatment with CHX was used to distinguish the immediate-early from early classes of viral gene expression. SA8 gene expression was compared to that of HSV-1.

5.2.1. Microarray analysis of HSV-1 α-gene expression

Previous microarray analysis from our laboratory in the study of HSV-1 infection in Mrc-5 cells in the presence and absence of cycloheximide illustrates the predicted pattern of HSV-1 IE gene expression (figure 6). The five immediate-early genes, RL02, RS01, US01, US12, and UL54 are not significantly inhibited by cycloheximide and, except for

Figure 6 – HSV-1 microarray signal intensity in the presence and absence of cycloheximide. DNA oligonucleotide microarray analysis of viral gene expression with and without cycloheximide treatment following infection of Mrc-5 cells with HSV-1 at an MOI of 5. Removal of outliers was performed manually, and probe intensities from a mock-infected sample were subtracted from infected-cell signals. Values presented are the average of all replicate probes for each gene. Genes are listed in their genome order.

Thanks to Steven Harrington and Joanne Hiebert.



UL54, they show elevated signal in the presence of cycloheximide, probably due to a lack of feedback inhibition of transcription by the IE gene products. Note that US10, US11, and US12 display "immediate-early" kinetics due to their existence as a coterminal transcript, in which US12 is the largest member. The US12 mRNA would therefore hybridize to the probes of US10 and US11.

5.2.2. Microarray analysis of SA8 α-gene expression

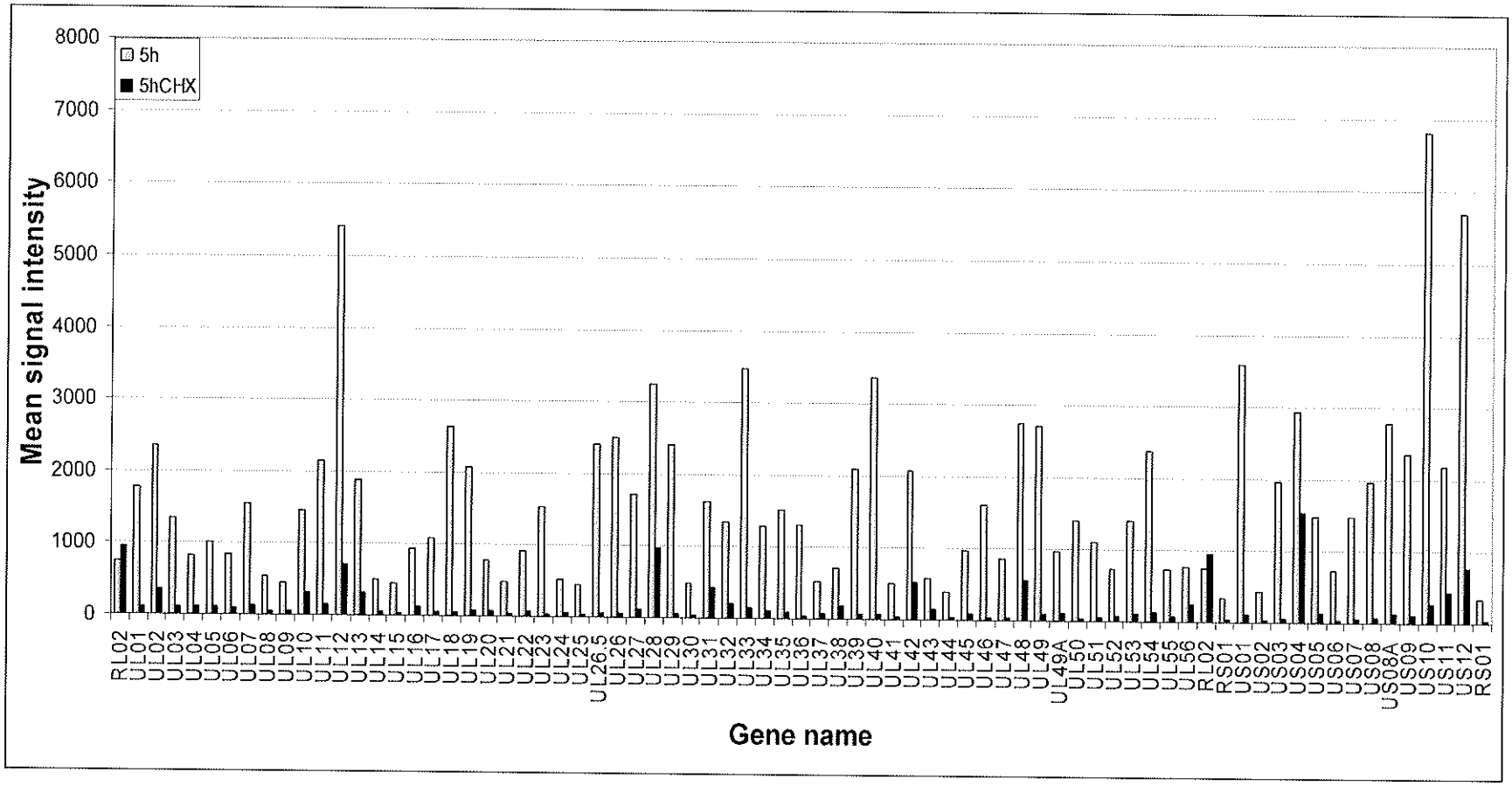
Experiments #1-3: Three independent experiments were performed to study the effect of cycloheximide (CHX) at 50 μg/mL (experiment #2) and 100 μg/mL (experiments #1 and 3) on SA8 infection in Mrc-5 cells. Cells were infected in the presence and absence of CHX and harvested at 5 or 8 hours post-infection (hpi).

In comparison to the HSV-1 α-gene expression pattern shown in figure 6, our SA8 microarray data shows that only the expression of RL02 is maintained in the presence of cycloheximide (figure 7). Although the US04 probes show relatively strong signal, it is reduced considerably in the presence of CHX. Other than RL02, all SA8 genes are inhibited by treatment with CHX. This surprising finding prompted further investigation of SA8 IE gene expression.

5.2.3. Validation by RT-PCR of SA8 α-gene expression

Experiments #1, 2, 4, 5 and 6: Four independent experiments were performed to study the effect of cycloheximide (CHX) at 50 μg/mL (experiment #2) and 100 μg/mL (experiments #1, 5, 6) on SA8 infection in Mrc-5 cells. Three independent experiments were performed to study the effect of acyclovir (ACV) at 10 μg/mL (experiments #3, 5, 6). Cells were infected in the presence and absence of CHX and harvested at 5 hours

Figure 7 – SA8 microarray signal intensity in the presence and absence of cycloheximide. DNA oligonucleotide microarray analysis of SA8 viral gene expression in the presence and absence of 100 µg/ml cycloheximide treatment at 5hpi following infection of Mrc-5 cells with SA8 at MOI of 5. Microarrays were hybridized as described in the methods, and the data analysis protocol outlined in the results section was applied. Representative data from one experiment is shown and genes are listed in their genome order. Values are mean normalized intensity of all data points in both Cy3 and Cy5 channels across several slides. For each gene, the value of the probe with the best quality score is shown.



post-infection (hpi). Cells treated with ACV were harvested at later times post-infection, either 9 or 27hpi.

In order to confirm the microarray results, we performed quantitative reverse transcriptase PCR (qRT-PCR) of four of the five SA8 genes with immediate-early homologues in HSV-1. Relative as well as specific quantification of mRNA abundance for RL02, RS01, US01 and US12 was calculated as described in the methods.

Only three of the five genes expressed with immediate-early kinetics in HSV-1 have a specific poly(A) signal (i.e. are not coterminal): RL02, US01, and RS01. US12 (ICP47) can be detected by amplifying a specific region of the gene, as it is the first gene in a coterminal series of three (with US10 and US11, both tegument proteins). The mean relative expression of RL02 from a sample 5hpi in the presence of cycloheximide (CHX) showed nearly a 4-fold increase in the amount of RNA detected. Table 4 presents the mean values from all data sets. In contrast, RS01, US01 and US12 were considerably inhibited in the presence of CHX, showing fold-reductions in expression of 28, 17, and 34, respectively. This supports the conclusion drawn from the microarray data, that RL02 is the only "immediate-early" gene in SA8 which truly displays this kinetic pattern in the presence of CHX.

RS01 presented the same challenges for qRT-PCR analysis as it did for microarray analysis, given that detection of this transcript is difficult due to its limited abundance. RNA samples for RS01 analysis had to be 'enriched' by performing reverse transcription using oligo(dT) primers rather than random hexamers, in order to achieve reliable RT-PCR results. This is consistent with the observation by microarray analysis that RS01 has very weak signals compared to other viral genes.

Table 4 – Expression of SA8 genes analysed by qRT-PCR. Relative expression, fold-change in relative expression, and percentage change in the number of gene copies per sample for SA8 (and HSV-1) genes are presented.

Gene	Relative expression, 5h Mean	Relative expression, 5h+CHX Mean (SD)	Fold-change, 5h±CHX	Change in copy number, 5h±CHX Mean (SD) % *
SA8 RL02	1	3.83 (1.54)	3.83	126.22 (67.24)
SA8 RS01	1	0.06 (0.04)	-28.25	-87.05 (6.92)
SA8 US01	1	0.05 (0.01)	-16.63	-92.49 (1.06)
SA8 US06	1	0.01 (0.02)	-525.71	-89.43 (7.96)
SA8 US12	1	0.03 (0.01)	-34.12	-96.17 (0.88)
Gene	**Relative expression, 4h**	**Relative expression, 4h+CHX**	**Fold-change, 4h±CHX**	
HSV-1 US12	1	0.21	-4.69	

NB: Relative expression and percentage change in quantified expression were calculated as described in the text. Values presented are mean and standard deviation (in parentheses) for SA8 genes. Values from a single are presented for HSV-1. Fold change illustrates the magnitude of change in signal in treated compared to untreated samples at the same time post-infection, therefore negative fold-change values indicate a reduction in signal. RS01 qRT-PCR was conducted using oligo(dT)-primers in order to achieve reliable signal, whereas random primers were used for other genes.

* All changes in the number of gene copies per microgram of total RNA are significantly different from no change at $p<0.05$.

The conclusion that US12 follows early or leaky-late gene kinetics is also consistent with data from our lab on the kinetics of the US10-12 coterminal transcripts, as analyzed by northern blot (figure 8). The ORFs US12, US11, and US10 share a termination site from which US12 is the most distal ORF and therefore generates the longest mRNA. Figure 8 shows the relative size of the three transcripts. In addition, northern blot data has been gathered in our lab for the coterminal transcript group of UL52-54 (figure 9), and this gene also appears to have early, and not immediate-early, kinetics in SA8. UL54 shares a termination site with both UL52 and UL53. UL54 is closest to the poly(A) signal, and is therefore represented by the smallest of the RNA fragments seen in figure 9. RL02 expression was also analysed by northern blot, and consistent with the microarray and qRT-PCR results, it displays sustained expression in the presence of cycloheximide (data not shown). These tests confirm that RL02 is the only gene expressed with immediate-early kinetics in SA8.

The relative expression of HSV-1 US12 was also analysed by qRT-PCR in order to compare the gene expression of SA8 with that of the prototype simplexvirus. Although some reduction in transcript abundance is observed in the CHX-treated sample, the change in expression (4-fold) is not as large as that seen with SA8 US12 (34-fold). These results do not refute the belief that HSV-1 US12 is expressed with immediate-early kinetics, and they support the proposal that there is a significant difference in relative transcript abundance of US12 between HSV-1 and SA8. Further experiments to investigate the abundance of other IE transcripts of HSV-1 by qRT-PCR are planned.

Amplification of a dilution series of SA8 gDNA in order to produce a standard curve allowed conversion of the CT values into gene copies per microgram of total RNA.

Figure 8 – Northern blot analysis of SA8 US12. Briefly: Vero cells were infected with SA8 at MOI of 5 PFU/cell; 5 μg of total RNA from each sample was separated in a 1% formaldehyde gel; and capillary transfer on to nylon membrane was performed overnight. A PCR product was generated as a probe for the coterminal transcript of US10-11. Briefly, SA8 genomic DNA was amplified using the primer pairs described for qRT-PCR (Table 3); DIG-labelled dUTP was incorporated during the PCR reaction; and the blot was developed using the DIG chemiluminescent detection system.

Thanks to Ravinder Singh and Tim Dumonceaux.

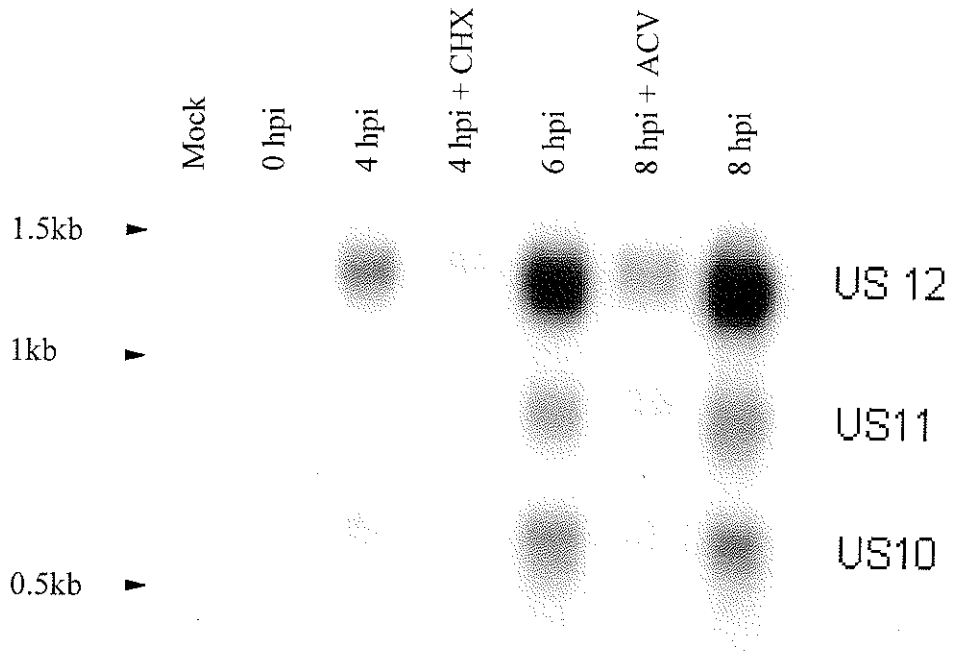
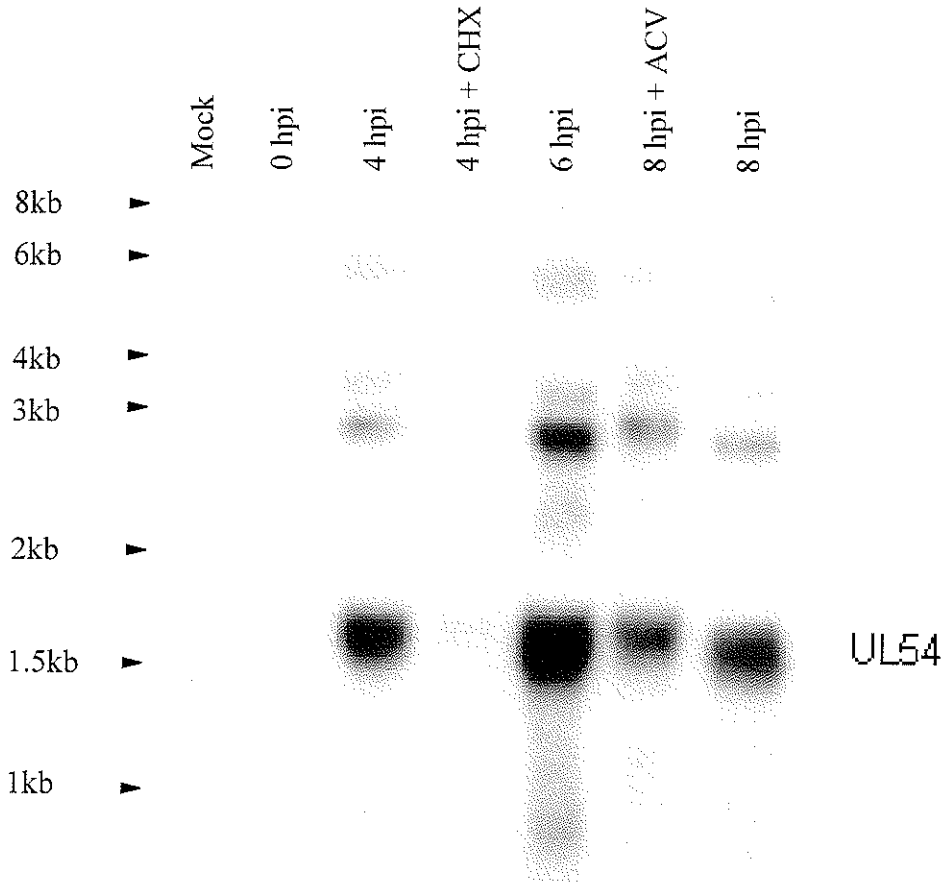


Figure 9 – Northern blot analysis of SA8 UL54. Briefly: Vero cells were infected with SA8 at MOI of 5 PFU/cell; 5 μg of total RNA from each sample was separate in a 1% formaldehyde gel; and capillary transfer on to nylon membrane was performed overnight. A PCR product was generated as a probe for the coterminal transcript UL52-54. Briefly, SA8 genomic DNA was amplified using the primer pairs described for qRT-PCR (Table 3); DIG-labelled dUTP was incorporated during the PCR reaction; and the blot was developed using the DIG chemiluminescent detection system. The bands of higher molecular weight than UL54 are the other coterminal transcripts, UL52 and UL53.

Thanks to Ravinder Singh and Tim Dumonceaux.



Differences in the level of absolute expression between experiments were normalized by calculating the mean percentage change in the number of cDNA copies under different conditions. The relative expression, quantitative expression (number of gene copies per microgram of total RNA), and percentage change in gene copy number is summarized in table 4. Results for each of the five data sets analysed can be found in appendix B.

5.2.4. Analysis of the sequences of SA8 α-gene promoters

Promoter sequences of SA8 genes RL02, RS01, US01, US12, and UL54 were compared to those of HSV-1, in order to assess the presence of the critical IE promoter consensus site, TAATGARAT and any other features of HSV-1 IE promoters. The SA8 RL02 promoter has at least one sequence similar to the TAATGARAT at approximately the same distance from the TATA-box and the start of the coding region as is observed in HSV-1. The sequence is TGATTGAGAT, located from -484 to -475, and is also preceded by the ATGC Oct-1-binding enhancer sequence found in HSV-1, lending credence to its role in transcriptional activation.

In the promoter of the RS01 gene of SA8, two sites similar to TAATGARAT are observed, slightly more proximal to the start codon than in HSV-1. The pair of TAATGCGAT sequences are found at -351 to -342 and -312 to -303, each preceded by the ATGC sequence, which remains of potential interest even though this enhancer region is not found in the ICP4 gene of HSV-1.

The analysis of the US01 and US12 promoter regions is complicated somewhat by sequence of the origin of replication OriS within the RS repeats. Nevertheless, the modified TAATGARAT sequence seen in the RS01 promoter (TAATGCGAT) is found

in tandem repeats upstream of US12. No similar sequence is observed in the US01 promoter.

5.3. Analysis of SA8 early & late gene expression

Viral gene expression kinetics of SA8 infection *in vitro* were analyzed using the microarray analysis protocol described in section 5.1.5.3. Acyclovir (ACV) is an inhibitor of viral DNA replication. Apart from the interruption of DNA synthesis, ACV prevents progression of lytic infection by preventing the expression of the late gene products encoding many structural viral proteins. Treatment with ACV was used to distinguish early from late classes of viral gene expression. In addition, the strength of inhibition by ACV identifies genes whose expression may be 'leaky-late' (γ_1). SA8 gene expression was compared to the prototype for simplexvirus expression kinetics, HSV-1.

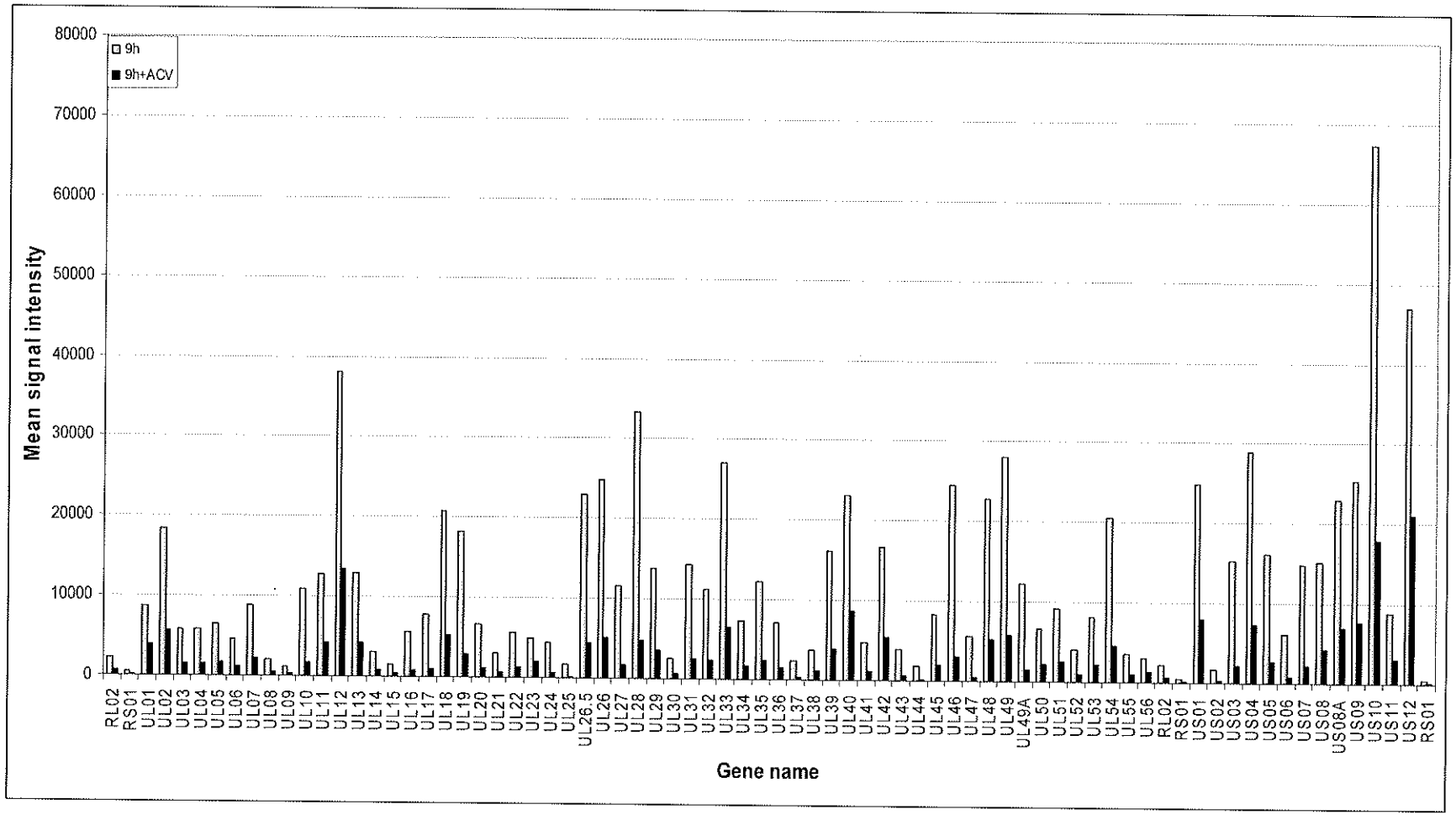
5.3.1. Microarray analysis of SA8 β- and γ-gene expression

Experiment #4: One experiment was performed to study the effect of 10 μg/ml acyclovir (ACV) on SA8 infection in Mrc-5 cells. This concentration of ACV exceeds the K_i value for HSV-1 polymerase (45). Cells were infected in the presence and absence of ACV and harvested at 9hpi.

Signal intensity in the presence and absence of ACV shows that all SA8 genes are inhibited by ACV to some extent (figure 10). It is difficult to determine whether genes with predicted early kinetics based on their HSV-1 homologues are more tolerant of ACV treatment than other genes. Further analysis is required for interpretation of these results.

Despite ambiguity of the microarray results, at least one SA8 transcript is confirmed to have late gene kinetics. The US06 gene encoding SA8 glycoprotein D was analysed

Figure 10 –SA8 microarray intensity signals in the presence (dark bars) and absence (light bars) of acyclovir (ACV) at 9hpi following infection of Mrc-5 cells with SA8 at MOI of 5 PFU/cell. Values are the normalized mean intensity of values in both Cy3 and Cy5 channels across several slides. For each gene, the value of the probe with the best quality score is shown. Genes are listed in their genome order.



by qRT-PCR in order to establish the reliability of this method to detect different kinetic classes of expression. Analysis of US06 showed a 341-fold reduction in expression in the presence of ACV. As expected for a structural gene, the transcript abundance increases, 57-fold, between early and late times during SA8 infection (appendix B).

5.3.2. Temporal kinetics of SA8 gene expression

Experiment #5 – Additional comparisons were conducted using data collected during SA8 infection in Mrc-5 cells in experiments 1 and 3. Pairs of time points: 5hpi and 8hpi (figure 11), and 5hpi and 13hpi (figure 12), were compared in order to determine relative changes in viral gene expression.

The gene kinetics of simplexviruses are defined primarily by their response to CHX and ACV. The temporal cascade of transcription is really a continuum of expression throughout infection. Therefore the complex patterns of gene expression observed in figures 11 and 12 are not surprising. Methods of analysis that identify patterns of similar gene expression would facilitate the interpretation of these data.

5.4. Alternate analysis protocol for viral gene expression microarrays

5.4.1. Analysis protocol for the expression of viral genes

As illustrated by the variability of microarray signal intensity in the presence and absence of CHX (figure 7) and especially ACV (figure 10), it is difficult to determine which differences in gene expression truly represent inhibition. The method of ranking signal intensity data, previously described for the analysis HSV-1 oligonucleotide microarray data (158), was not used in this study due to a global inhibition of viral signal intensities in treated (cycloheximide or acyclovir) compared

Figure 11 –SA8 microarray intensity signals at 5hpi (light bars) and 8hpi (dark bars) following infection of Mrc-5 cells with SA8 at MOI of 5. Values are the normalized mean intensity of values in both Cy3 and Cy5 channels across several slides. For each gene, the value of the probe with the best quality score is shown. Genes are listed in their genome order.

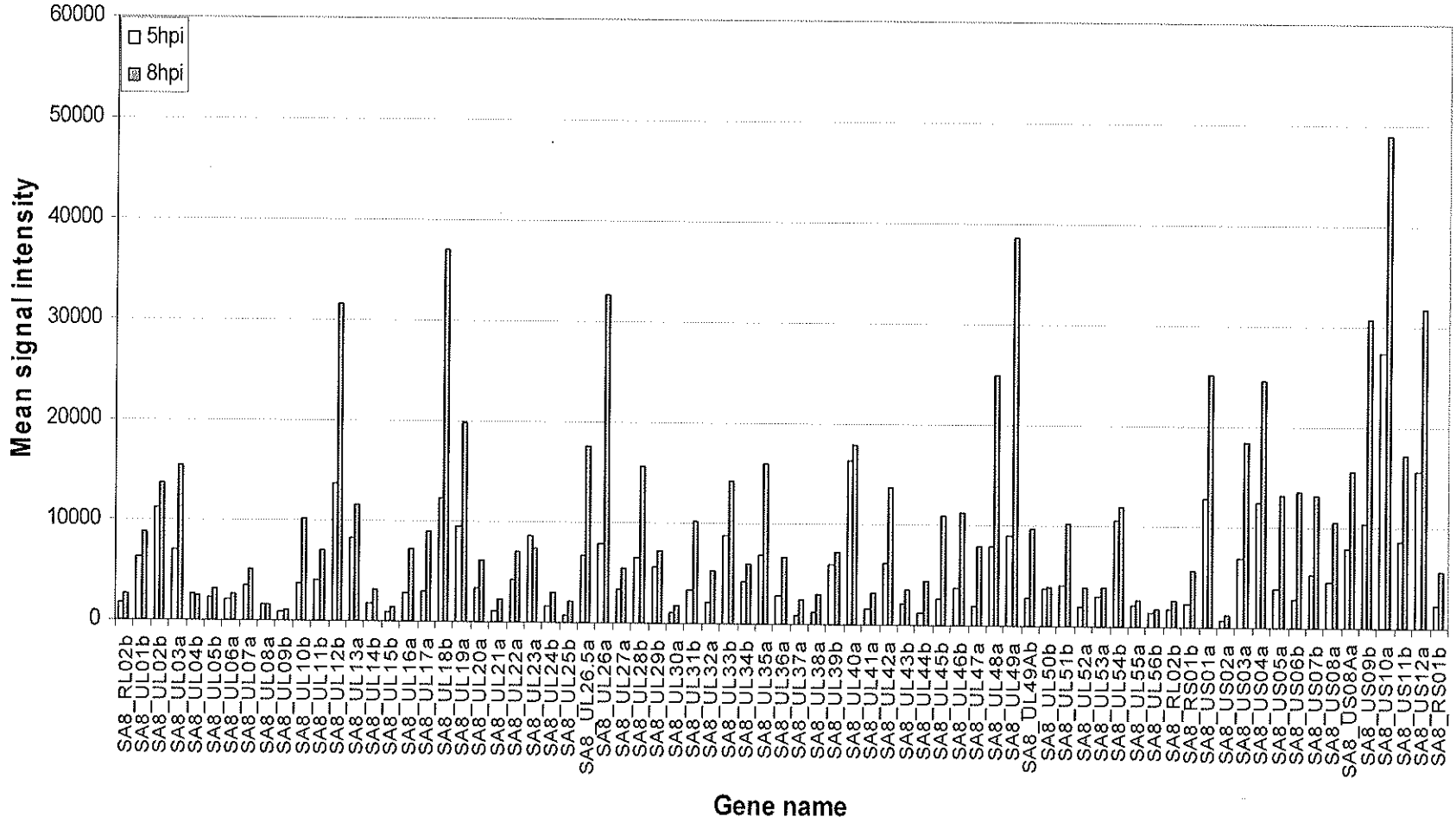
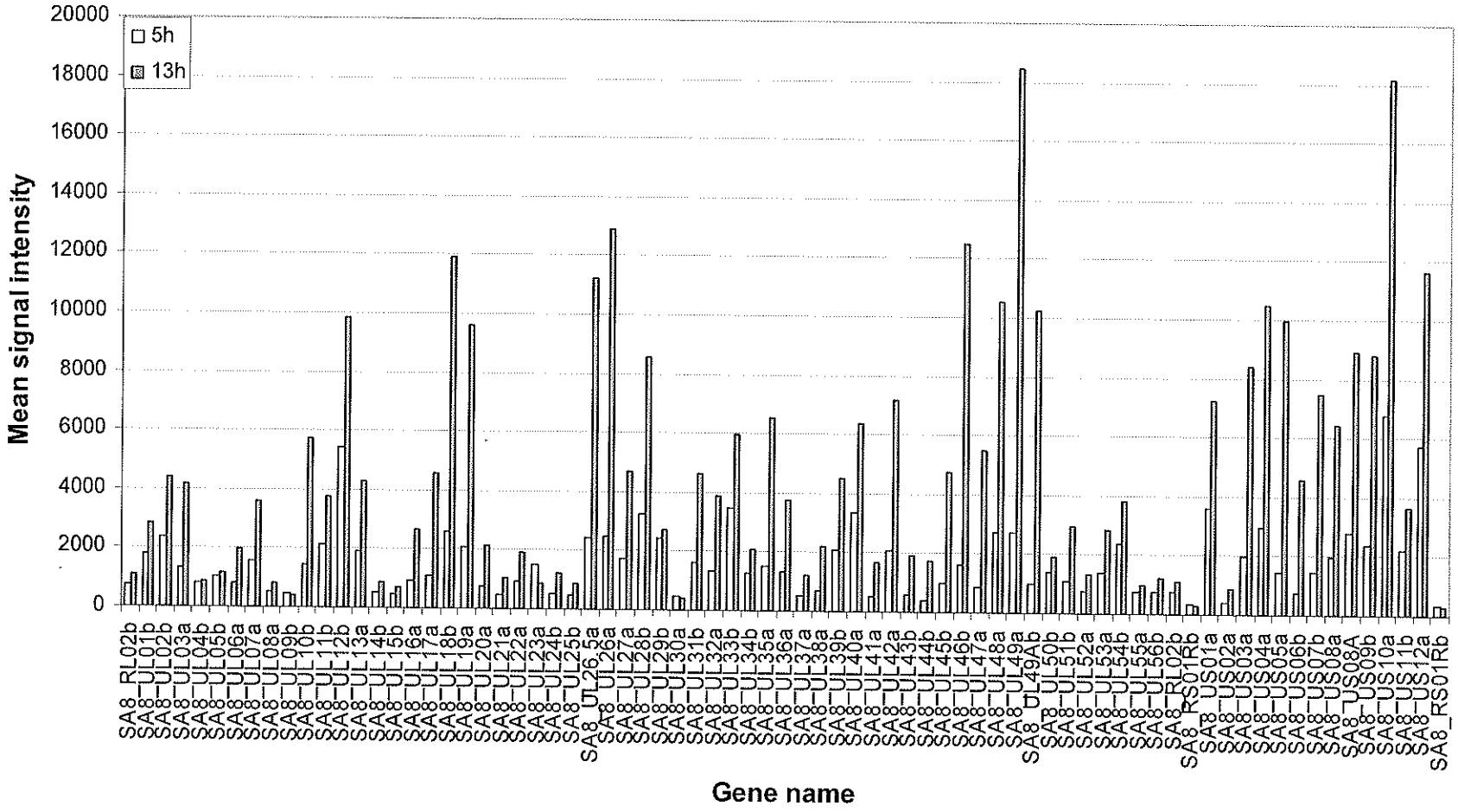


Figure 12 –SA8 microarray intensity signals at 5hpi (light bars) and 13hpi (dark bars) following infection of Mrc-5 cells with SA8 at MOI of 5. Values are the normalized mean intensity of values in both Cy3 and Cy5 channels across several slides. For each gene, the value of the probe with the best quality score is shown. Genes are listed in their genome order.



to untreated samples. In order to identify patterns in the magnitude of inhibition observed in different gene signals, we developed a novel protocol to identify genes with similar kinetic profiles. A combination of cluster and classification analysis was used in order to provide an objective method for analysis of viral gene expression. Briefly, transcripts having only a single kinetic pattern (i.e. non-coterminal, non-overlapping genes) were selected for cluster analysis. Cluster analysis divided these transcripts into groups with similar characteristics of microarray signal intensity. The intent was to identify characteristic kinetic patterns, based on the expression of transcripts encoding a single ORF. In order to establish criteria by which all remaining transcripts (those having more than one kinetic pattern) could be assigned to one of the identified clusters, classification analysis was performed. Finally, SA8 transcript kinetics were compared to those of their HSV-1 homologues.

5.4.1.1. Selection of non-coterminal, non-overlapping transcripts

The presence of coterminal and overlapping transcripts in SA8 makes it difficult to identify patterns that may be representative of different kinetic classes. In particular, coterminal groups such as UL52-54 and US10-12 contain genes with different predicted kinetics. For this reason, SA8 genes were divided into two categories: "A" genes are non-coterminal, non-overlapping transcripts, while other genes whose transcripts are overlapping or coterminal are designated "B" genes. "A" gene signals were used for cluster analysis, based on the principle that their kinetic behaviour would be specific to one kinetic class.

5.4.1.2. Relative difference of signal between experimental conditions

The variation in signal intensity among genes makes it necessary to consider the relative strength of the untreated signal intensity of a gene when interpreting its inhibitor-treated signal. For example, very small signals from CHX-treated cells do not automatically mean that a gene is inhibited by CHX if its untreated signal is equally small. For example, if gene X has a signal intensity of 1000 at 5hpi, and a signal intensity of 750 at 5hpi in the presence of cycloheximide, the percentage of signal change illustrates a reduction in signal of (1000-750)/1000, which equals 25%. Similarly, gene Y with a much smaller signal in both samples, but influenced to the same degree by CHX, may have a percentage change calculation of F = (100-75)/100 = 25%.

The percentage change in signal intensity (F) between inhibitor-treated and untreated samples, was calculated for each "A" gene using the formula:

$$F = (T-Ti)/T \qquad (9)$$

Where F is the percentage of mean signal change for a gene between a sample treated with an inhibitor, and a non-inhibited sample at the same time post-infection; T is the mean net intensity from an infected-cell sample; and Ti is the mean net intensity from an infected-cell sample treated with an inhibitor.

Simply put, F is the percentage change in signal in the treated sample compared to the untreated. It is equal to 0% if the signal intensity of the gene is the same in both treated and untreated samples, and it approaches 100% as the signal intensity in the inhibitor-treated sample becomes negligible compared to that of the untreated sample.

Formula 9 can also be applied to time course data in order to measure the change in signal at early and late times during infection.

5.4.1.3. Clustering (k-means)

Cluster analysis is often used in microarray analysis to identify differences in the patterns of gene expression between different samples (e.g. cells undergoing malignant transformation versus normal cells). However, in this protocol, we already know that the samples are different (treated with inhibitors or harvested at different times during infection). Our aim is to identify differences in viral gene expression between the genes within these samples (e.g. genes that maintain their expression in the presence of cycloheximide while others are inhibited).

Similar characteristics among "A" genes were identified by k-means clustering based on their signal intensity in the presence of an inhibitor, and their percentage change value (F).

5.4.1.4. Classification analysis

We used classification analysis as a tool to identify threshold values of signal intensity and percentage change (F) which characterize the clusters previously identified so that "B" genes (coterminal or overlapping transcripts) could be assigned to the same clusters as "A" genes based on their kinetic behaviour in the presence of inhibitors.

Classification is a 'supervised' method of analysis, which requires identification of "classes" within the data. The algorithms separate these classes according to the variables provided and aim to identify features common to each class. Each "A" gene

was given a cluster number as it's "class", and classification was performed using the same variables as were used for cluster analysis. This results in identical grouping of the "A" genes, but additional information about the cluster characteristics is provided by the classification algorithm.

5.4.1.5. Analysis protocol for the expression of coterminal viral genes

Despite their potentially complex kinetic patterns "B" genes were grouped into the same clusters as "A" genes using the threshold values provided by classification analysis. Genes in the same cluster have similar expression kinetics. The cluster characteristics may shed light on whether the expression of these genes is similar to their HSV-1 homologues. Clusters containing genes with different predicted kinetics provides a starting point for further analysis of SA8 gene expression.

5.4.1.6. Interpretation

Four values for each probe were assigned to clusters according to their intensity and percentage change (F) values. When multiple probes for the same gene are located in different clusters, the probe with the highest quality score (see Results section 5.1.3) is the one that defined the cluster of the gene. Coterminal transcripts were assigned to a cluster of the 'best' probe of all the probes for all the ORFs in that coterminal transcription unit.

5.4.2. Microarray analysis of SA8 α-gene expression

Experiments #1-3: Three independent experiments were performed to study the effect of cycloheximide (CHX, 50 and 100 μg/mL) on SA8 infection in Mrc-5 cells. Cells were infected in the presence and absence of CHX and harvested at 5 or 8 hours post-

infection (hpi). A summary of the cluster membership of each gene and each transcript, as well as mean signal intensity values in the presence and absence of CHX, are presented in appendix DA (experiments 1 and 2), and DB (experiment 3).

The clusters identified by analysis of "A" gene kinetics in experiment 1 are illustrated in the scatterplot in figure 13. Mean signal intensity in the presence of cycloheximide (5h+CHX) is plotted on the Y axis, and the percentage change between inhibited and uninhibited signals is plotted on the X axis. In this way, clusters of genes that are strongly inhibited by CHX are near the bottom-right corner of the graph. These genes have weak signal in the presence of CHX, and the signal intensity from CHX-treated cells is negligible compared to the untreated signal, resulting in a percentage change value very close to 100%. In contrast, genes which have greater intensity signal in the presence of CHX have negative percentage change values, and this often also corresponds to significant CHX signal intensity, resulting in clusters which are located in the upper left section of the graph. The clusters containing the five predicted IE genes are labelled accordingly on each graph.

Classification was used to determine values of signal intensity and percentage change which are characteristic of the clusters. Figure 14 shows a representative classification tree from experiment 1. Each box highlights a division of the data. For example, in experiment 1, the "A" gene signals are best distinguished at box 1 based on whether their signal intensity at 5hpi+CHX is greater or smaller than 808.45. Probes that have greater signal are further distinguished based on whether their value of percentage change is greater or smaller than -25.37%. These two criteria separate clusters 3 (figure 14,

Figure 13 – SA8 microarray clusters in the presence of cycloheximide treatment. Clusters in experiment 1 were identified using signal intensity of "A" genes (Y axis) in the presence of 100 μg/mL cycloheximide treatment at 5hpi compared to the difference between treated and untreated samples (X axis). This difference is expressed as a percentage of untreated intensity, therefore 0% indicates equality of treated and untreated signal intensities and 100% indicates minimal signal after cycloheximide treatment compared to the signal in an untreated sample.

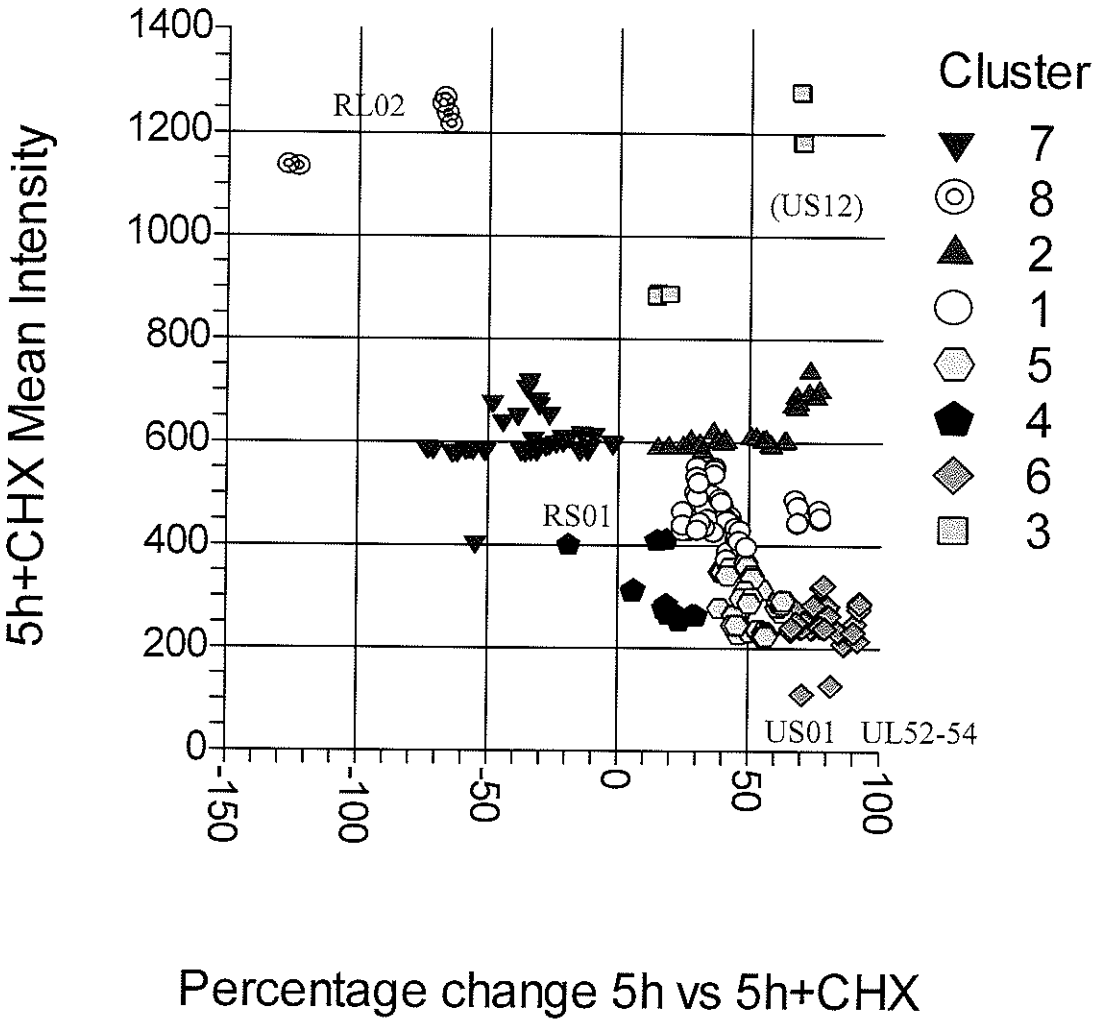
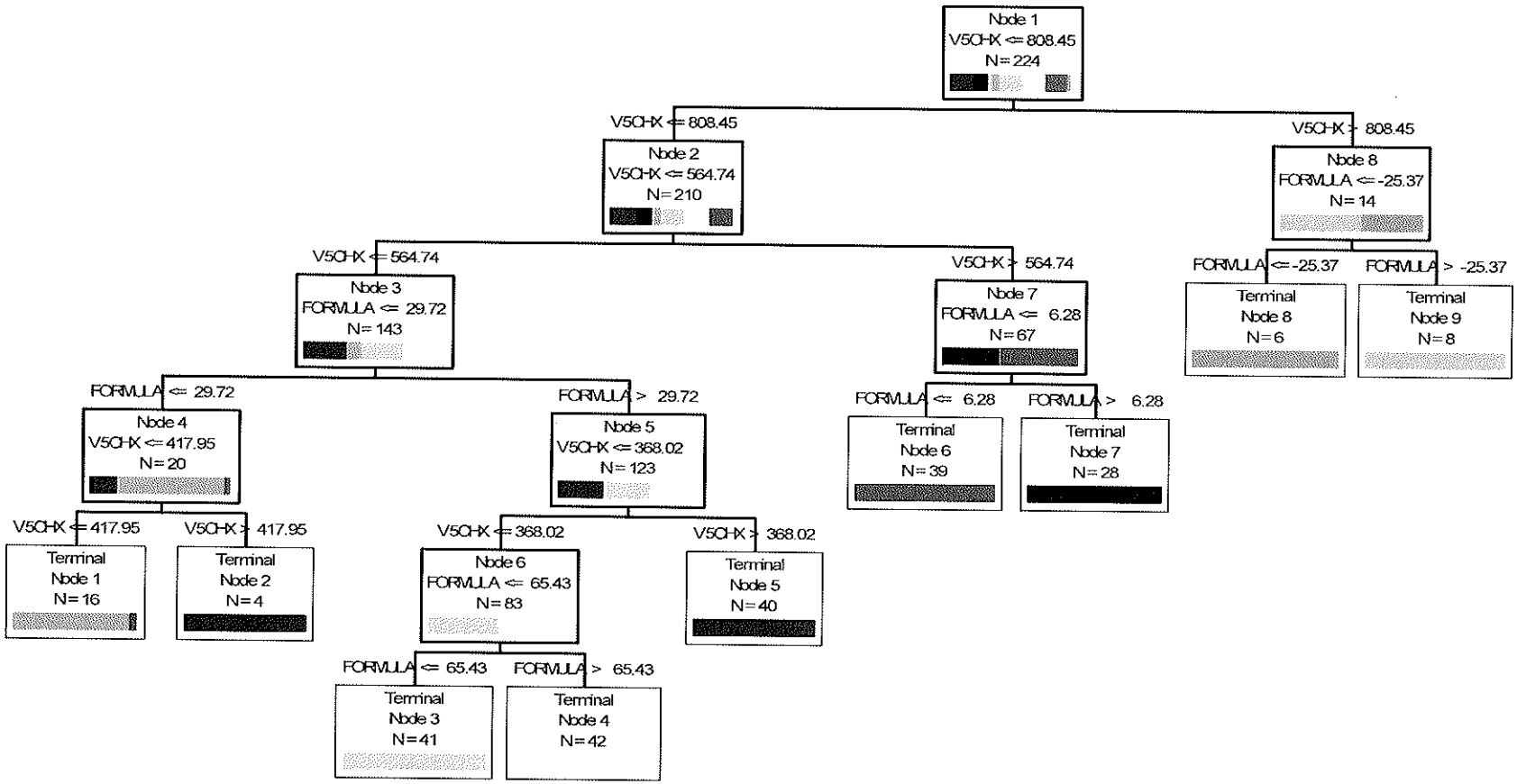


Figure 14 – Classification tree based on SA8 gene expression in the presence of cycloheximide. Signal intensity at 5hpi +CHX and the percentage of signal change between CHX-treated and untreated samples (F) for "A" genes were used for classification (experiment #1 data shown). The coloured bars represent the distribution of probe values from different clusters in different nodes. Cluster colours, 1 (red) through 8 (grey) can be identified sequentially in the uppermost node and the sequence of colours is maintained throughout the tree. Note that terminal node (TN) 1 contains values from clusters 4 (orange) and 7 (pink), while TN 2 contains values from cluster 1 (red).



"terminal node 9") and 8 (figure 14, "terminal node 8") from the other clusters as well as from each other.

Figure 15 shows the clusters identified in experiment 2 (50 μg/mL CHX), and experiment 3 shows analysis of CHX treatment at both 5hpi (figure 16) and 8hpi (figure 17).

The results of experiment 1 show that the predicted "immediate-early" (IE) genes of SA8 do not function in the same way as their homologues in HSV-1. The hypothesis that SA8 genes have the same pattern of expression as their HSV-1 homologues is contradicted by these results. All four data sets show different behaviour in the presence of cycloheximide of the three non-coterminal IE genes: RL02, RS01, and US01. Only RL02 (ICP0) clearly shows the anticipated pattern of IE gene kinetics characterized by strong signal intensity in both treated and untreated samples, demonstrating the resistance of RL02 expression to inhibition by CHX. In fact, RL02 consistently gives a higher signal in samples treated with CHX. This confirms the data presented in figure 10, where RL02 is the only gene not inhibited by CHX.

US01 (ICP22) on the other hand, is strongly inhibited by CHX treatment, suggesting early gene kinetics. RS01 (ICP4) provides a mixed kinetic picture during CHX treatment as it is located in a cluster between RL02 and RS01. In all samples, RS01 expression is reduced following CHX treatment, and the signal obtained from treated samples is small, suggesting limited expression. However the percentage of signal change in CHX-treated samples is not substantially less compared to untreated samples, suggesting limited effect of the inhibitor. The low signal intensity observed from RS01 makes it difficult to

Figure 15 – SA8 microarray clusters in the presence of cycloheximide treatment. Clusters in experiment 2 were identified using signal intensity of "A" genes in the presence of 50 μg/mL cycloheximide treatment at 5hpi (Y axis) compared to the difference between treated and untreated samples (X axis). This difference is expressed as a percentage of untreated intensity, therefore 0% indicates equality of treated and untreated signal intensities and 100% indicates minimal signal after cycloheximide treatment compared to the signal in an untreated sample.

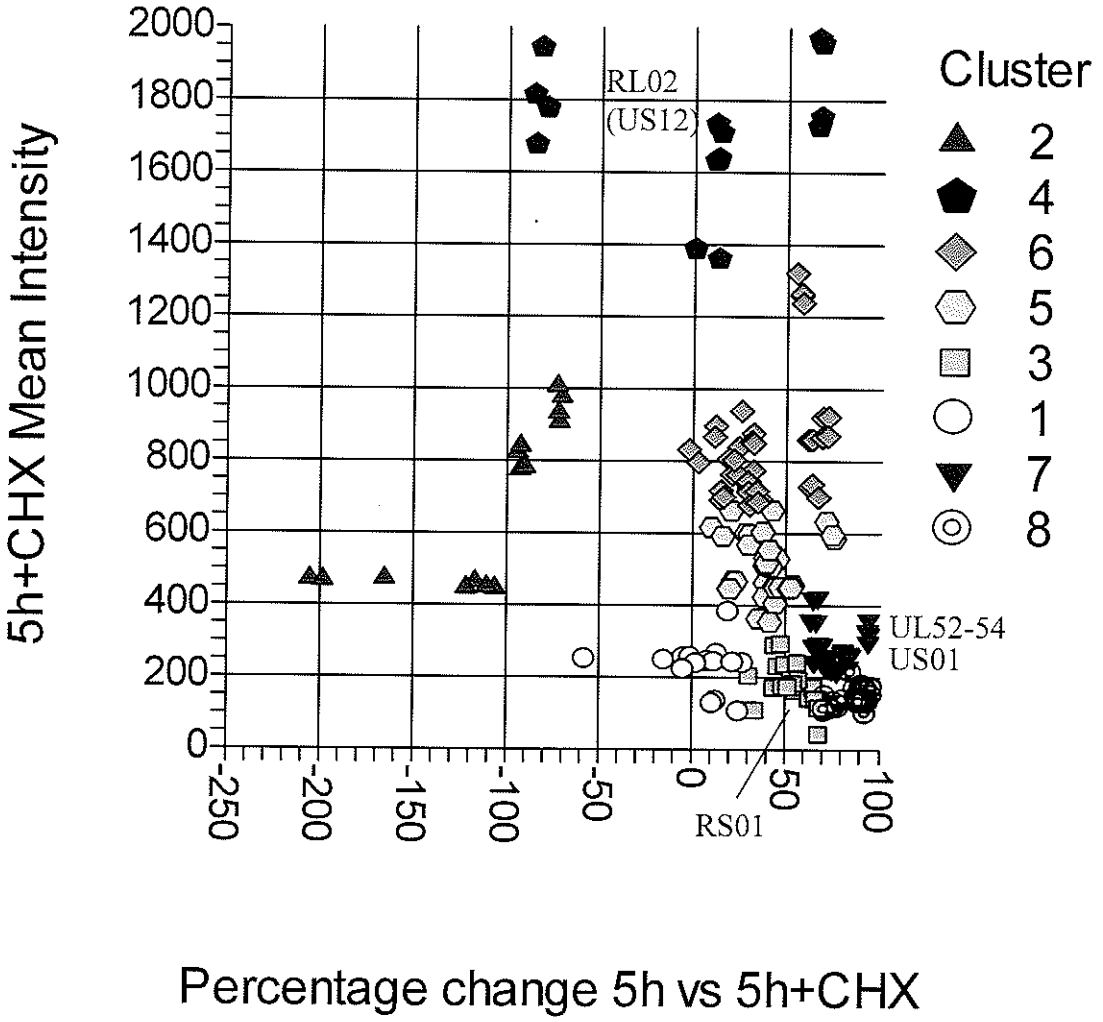


Figure 16 – SA8 microarray clusters in the presence of cycloheximide treatment. Clusters in experiment 3A identified using signal intensity of "A" genes in the presence of 100 μg/mL cycloheximide treatment at 5hpi (Y axis) compared to the difference between treated and untreated samples (X axis). This difference is expressed as a percentage of untreated intensity, therefore 0% indicates equality of treated and untreated signal intensities and 100% indicates minimal signal after cycloheximide treatment compared to the signal in an untreated sample.

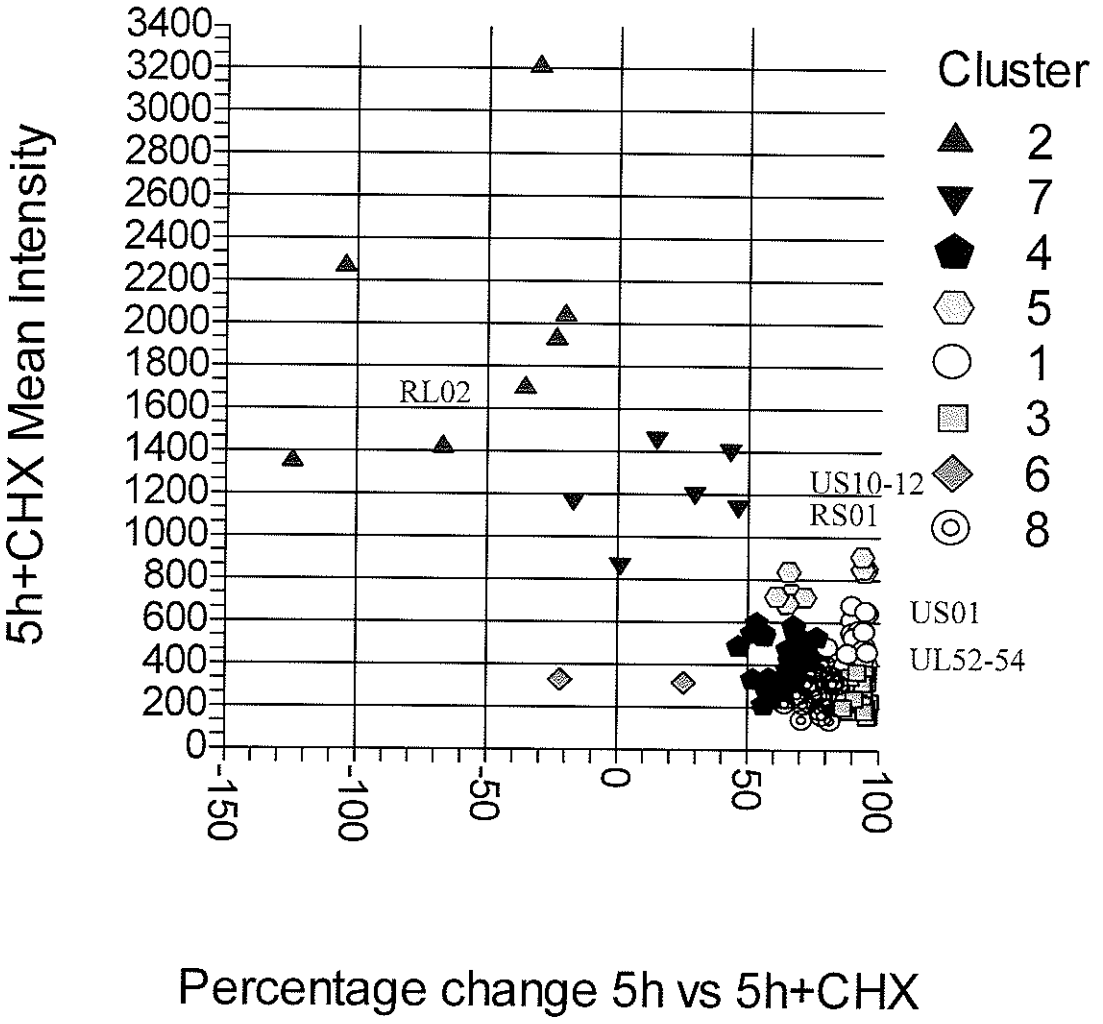
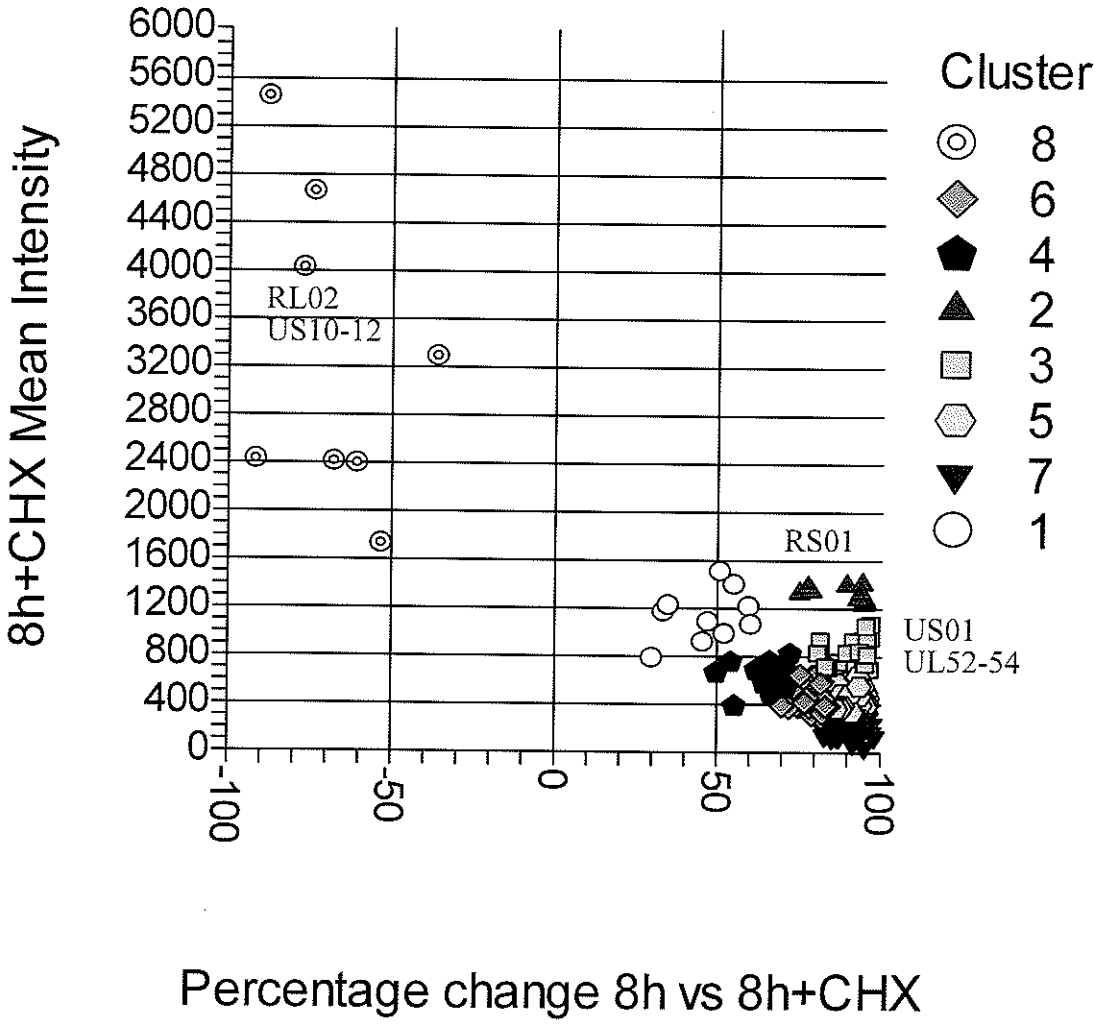


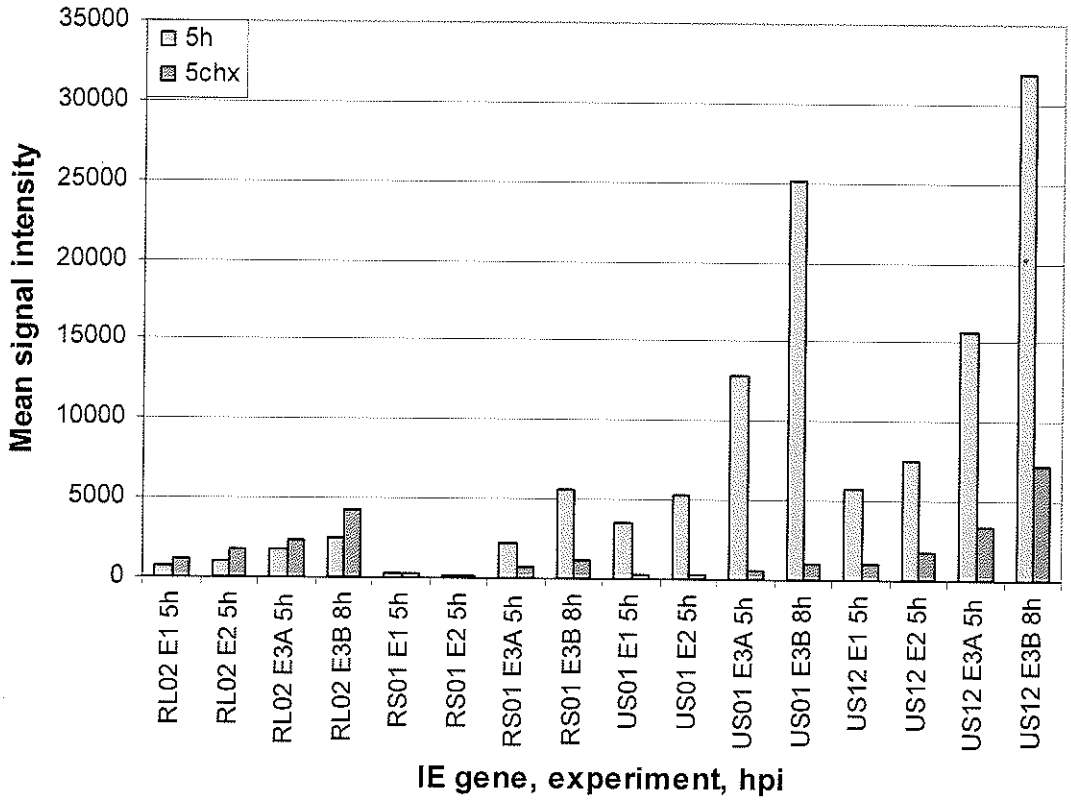
Figure 17 – SA8 microarray clusters in the presence of cycloheximide treatment. Clusters from experiment 3B identified using signal intensity of "A" genes in the presence of 100 μg/mL cycloheximide treatment at 8hpi (Y axis) compared to the difference between treated and untreated samples (X axis). This difference is expressed as a percentage of untreated intensity, therefore 0% indicates equality of treated and untreated signal intensities and 100% indicates minimal signal after cycloheximide treatment compared to the signal in an untreated sample.



determine if RS01 displays the kinetic pattern of an immediate-early gene during SA8 infection.

Although US12 belongs to a coterminal group, it is the ORF most distal from the poly(A) signal, and the genome location of the probes for US12 suggests that the results should be specific for this transcript as they do not overlap the anticipated start site of US11. In all comparisons at 5hpi, the US10-12 transcript displays a signal in the presence of CHX almost as strong as that of RL02. However, its expression is reduced much more than that of RL02 by CHX treatment. Data set 3B, comparing expression at 8hpi ± CHX, seems to contradict this conclusion as US10-11 and US12 are all present in the same cluster as RL02 and the cluster characteristics suggest that this transcript is upregulated in the presence of CHX. Looking at the mean intensity values of the non-coterminal "IE" genes in the presence and absence of CHX (figure 18), we see the true picture of US12 expression. Indeed it presents a strong signal but is it substantially reduced by CHX. The location of US10-12 in the same cluster as RL02 in data set 3B is due to the similarly 'extreme' values of these genes compared to others in the presence of CHX rather than any relationship in their degree of inhibition by CHX. If the data had been separated into more clusters, the kinetics of these genes may have been more clearly defined. A minimum number of clusters were chosen based on the percentage variation between the sum of squares (see methods section 4.9.6) in order to maintain broad clusters which would contain multiple transcripts.

Figure 18 – Summary of SA8 IE gene expression in the presence of CHX. The mean signal intensity of the four non-coterminal immediate-early genes in each of four microarray experiments following treatment with (dark grey bars) and without (light grey bars) cycloheximide is shown. X axis labels include the gene name, experiment number (E), and time of sample harvest (5h or 8h post-infection).



5.4.3. Microarray analysis of SA8 β-gene expression

Experiment #4: One experiment was performed to study the effect of acyclovir (ACV, 10 μg/ml) on SA8 infection in Mrc-5 cells. Cells were infected in the presence and absence of ACV and harvested at 9hpi.

The mean signal intensity in the presence of inhibitors and the percentage change (F) between inhibited and uninhibited signals for each probe replicate are presented in figure 19. A summary of the genes in each cluster illustrated in figure 19 by their predicted kinetic class is outlined in table 5. Cluster membership of each gene and/or transcript as well as mean signal intensity values in the presence and absence of CHX are presented in appendix E.

The results of experiment 4 (figure 19) show that the predicted "early" genes of SA8 do not function entirely in the same way as their homologues in HSV-1. Of the five non-coterminal "early" (E) genes, UL23 (thymidine kinase) is least affected by treatment with acyclovir, as illustrated by a less dramatic reduction in its signal intensity in the ACV-treated sample compared to the other E genes.

UL29 and UL42 belong to a cluster illustrating a slightly greater inhibition by ACV, but with stronger signal intensity in the ACV-treated sample compared to UL23. This cluster (cluster 4) also contains two coterminal transcript groups of which one member is proposed to be expressed with early kinetics, UL01-02 (glycoprotein L, and uracil DNA glycosylase), and UL52-54 (helicase/primase protein, glycoprotein K, ICP27). In fact, based on the transcription pattern of UL54 observed in experiments #1-3, the latter coterminal transcript contains two genes with early kinetics.

Figure 19 – SA8 microarray gene clusters in the presence of acyclovir treatment. Clusters from experiment 4 were identified using signal intensity of "A" genes in the presence of 10 μg/mL acyclovir treatment at 9hpi (Y axis) compared to the difference between treated and untreated samples (X axis). This difference is expressed as a percentage of untreated intensity, therefore 0% indicates equality of treated and untreated signal intensities and 100% indicates minimal signal after acyclovir treatment compared to the signal in an untreated sample.

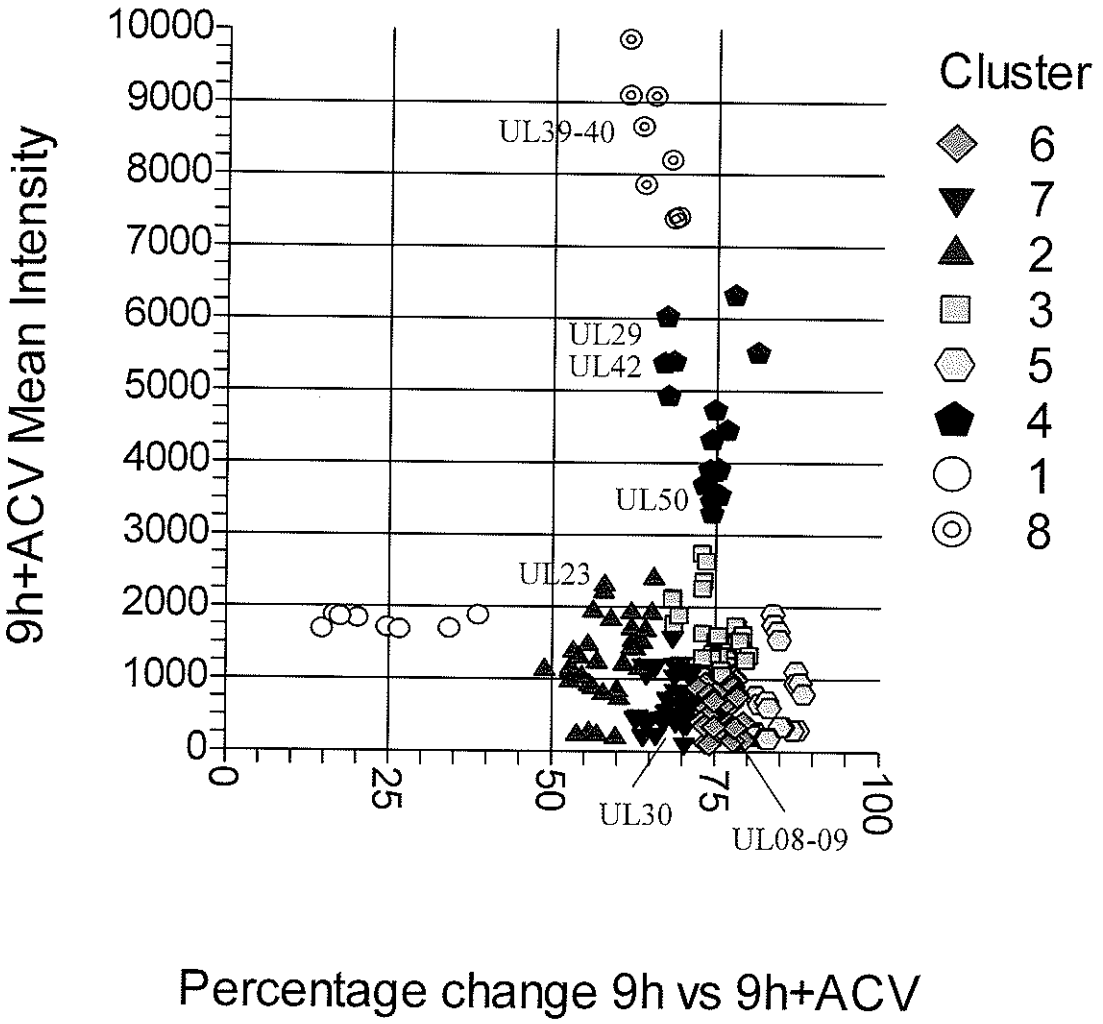


Table 5 – Cluster membership of SA8 genes in the presence of acyclovir (experiment 4), grouped according to the kinetic class predicted by their HSV-1 homologues.

Cluster	**1**	**2**	**8**	**4**	**3**	**7**	**6**	**5**
IE / E		UL23	UL39 -40	UL29		UL30	UL08-09	
			US01	UL42		RL02	RS01	
			US10 -12					
Partial E			UL11 -14	UL01-02				
			US03 -05	UL52-54				
L		UL56	US08 -09	UL18-20	UL03	UL15	UL21	UL10
				UL24-26.5	UL04-05	UL38	UL55	UL16-17
				UL27-28	UL06-07		US02	UL31-32
				UL33-35	UL22			UL37
				UL48	UL36			UL43
					UL41			UL46-47
					UL44-45			US06-07
					UL51			
Unknown	RL1a	RL1 anti-b					RL1b	RL1 anti-e
	RL1e	RL1 anti-c					RL1c	
		RL1 anti-d						
		RL1d						

NB: Transcripts are divided based on their predicted kinetics: IE / E = immediate-early or early; Partial E = Coterminal transcript group in which one or more ORFs is predicted to have early kinetics; L = late; Unknown = probes for which no transcript has been predicted for SA8.

Neighbouring cluster 8 is characterized by a similar reduction in transcript abundance in the presence of ACV but even greater signal intensity in the ACV-treated sample compared to cluster 4. Cluster 8 contains a coterminal pair of "early" genes, UL39-40 (ribonucleotide reductase, large and small subunits). This transcript shares cluster membership with two immediate-early transcripts, US01 and US12, as well as two coterminal transcript groups of which one member gene is anticipated to display early kinetics, UL11-14 and US03-05 (if one adopts the notion that the US03 kinase is expressed as a β gene as suggested by Wagner (172)). These transcript groups may be in this cluster, distinguished by considerable signal intensity in the presence of ACV, primarily due to the influence of their 'early' ORF. Their substantial inhibition with ACV may be due to the contribution of their 'late' gene ORFs which create a vast difference in signal intensity between samples with and without ACV.

UL30 (DNA polymerase) has the smallest mean signal intensity of the non-coterminal E genes, in both the ACV-treated and untreated samples. The weak signal intensity, coupled with a substantial reduction in signal in the presence of ACV makes uncertain the inclusion of UL30 in the list of SA8 "early" genes.

There is a coterminal pair of β transcripts, UL08 and UL09, encoding a protein of the helicase-primase complex, and the origin-binding protein, respectively. This transcript shows slightly more inhibition with acyclovir treatment than is observed for UL30, and relatively small signal intensity.

Both UL30 and UL08-09 share their respective clusters with RL02 and RS01 as well as a number of genes anticipated to be expressed with late kinetics. Clusters 3, 5, 6, and

7 have similar characteristics, so it may be logical to conclude that genes in these clusters are not expressed with early kinetics in SA8.

Conclusive data was not available for UL50 (dUTPase) in this experiment, as its cluster membership is mixed equally between cluster 2 (containing several 'early' genes), and cluster 3 (mainly 'late' genes) in figure 19.

5.4.4. Similarities in expression kinetics of SA8 genes

Microarray analysis always generates more data than just the answer to the initial question, and this is one of the advantages of using such a broad analysis tool. Additional observations can be made of similarities in kinetic behaviour in the presence of inhibitors beyond the differentiation of immediate-early and early transcripts. Of particular interest is any detection of transcripts from probes in the region of SA8 corresponding to the location of the RL1 (ICP34.5) in HSV-1 since no ORF has been predicted in this region of SA8.

5.4.4.1. Microarray analysis of gene expression in the presence of cycloheximide

Grouping transcripts located in clusters with similar kinetic behaviour may provide information on similarities of regulation or abundance of these transcripts. Multiple clusters are identified when the transcripts are located in different (often adjacent) clusters. Cluster characteristics are illustrated in figures 13, 15, 16, and 17. Table 6 summarizes the composition of the 'similarity' groups A through J, including the cluster membership of each group in the experiments #1 through 3.

Three groups of SA8 transcripts with similar kinetic characteristics in the presence of cycloheximide (see figures 13, 15, 16, and 17) are noted below.

Table 6 – Groups of SA8 transcripts with similar expression in the presence of cycloheximide.

Group	Transcripts with similar kinetics	Cluster Exp 1	Cluster Exp 2	Cluster Exp 3A	Cluster Exp 3B
A	RL1-anti-a, *-anti-e, -c,* -b	7	2	2	8
B	*LATCDS3,* RL1-d, -e	4	1	6	4
C	UL11-14, UL27-28, US03-05, US12	3	4	5	8, 2 & 5
D	US02, US08-09, UL30	7 & 4	3	8	4, 6 & 7
E	UL15, UL37, UL38, UL43	5	3 & 5	4 & 8	4, 6 & 5
F	UL41, UL23, UL29, UL36	7, 2, & 6	2 & 8	1 & 3	5
G	UL10, UL18-20, UL46-47	1	6 & 8	3	8 & 5
H	UL03, 04-05, 06-07, 22, 33-35, 39-40, 44-45, 49-49A, 52-54, 55, US01, US08-09	6	7	1 & 3	3 & 5
I	UL50, UL51	6	8	3	3 & 5

NB: Probes indicated in italics were present in experiments #1 and 2, but not #3.

Additional groups are described in appendix C. The latter consist mainly of genes predicted to have late kinetics, and are generally located in clusters inhibited by CHX, which supports this kinetic prediction. These groupings may provide preliminary information for further analysis of regulation of SA8 transcription or levels of viral protein synthesis.

Group A – RL1-anti-a, -anti-e, -c, -b

Both sense and antisense probes were designed in the region of the SA8 genome upstream of RL02, within the region ascribed in HSV-1 to RL1 (ICP34.5) in order to determine if any transcripts are generated from this region. Several of these probes have cluster membership in an adjacent cluster to RL02: RL1-anti-a, -anti-e, -c, and -b. RL1-b and -c probes are located 179 bp and 550 bp upstream of the start codon of RL02 and may detect the leader sequence of this transcript, thus explaining the similarity of their signal. RL1-anti-a and –anti-e probes are located in a region upstream of RL02 which better approximates the location of the HSV-1 RL1 gene. However, being on the antisense strand, a transcript in this region would correspond to either ORF O or ORF P in HSV-1 (92), which have not been deleted in the SA8 genome (166).

Group B – LAT CDS3, RL1-d and -e

In experiments #1 and 2, in which these probes were included, they are found in a cluster with low signal intensity in a CHX-treated sample, and little difference between CHX-treated and untreated samples. Given that these probes were not designed for a known transcript, and given the error in strand-selectivity of LAT CDS3, it is likely that no transcripts are specifically detected by these probes. These

transcripts are located 901bp (RL1-e) and 1843bp (RL1-d) upstream of the RL02 coding region, approaching the genomic 'a' sequence. Given that RL1-anti-e may represent expression of an ORF P-like transcript, RL1-e (its complementary sequence) does not appear to detect an RL1-like transcript in this region of SA8.

Group C – UL11-14, UL27-28, US03-05, US12 (IE protein ICP47)

Certain "late" genes in SA8 demonstrate kinetics expected from immediate-early genes, a strong signal in the presence of CHX, and a small difference in signal between CHX-treated and untreated cells. The coterminal series UL11-14 (membrane-associated protein, alkaline exonuclease, protein kinase, tegument protein), UL27-28 (glycoprotein B, capsid-associated protein), and US03-05 (protein kinase, glycoproteins G and J) consistently give stronger signals than other genes with predominantly late kinetics in the presence of CHX at 5hpi. However, the percentage of signal reduction compared to that of untreated samples varies between experiments (mean reduction of 40%). In addition, transcripts US10-11 & UL01-02 share a cluster distinct from that of Group C in experiments #1 and 2, although these genes are all present in one cluster in experiment #3A, suggesting some similarity of expression.

5.4.4.2. Comparison of microarray results from 5h and 8h following cycloheximide treatment

It is interesting to note that the cluster results from experiment #3B (8hpi ± CHX) often deviate from those of the experiments performed at 5hpi. This may provide insight into the changing dynamics of transcript synthesis and accumulation when the inhibition of protein synthesis continues for a longer period of time. For example, in

the 8hpi ± CHX comparison, the similarity of behaviour noted at 5hpi has dissipated among transcripts UL11-14, UL27-28, US03-05, and US10-12. At 8hpi, US03-05 and US10-12 cluster with RL02. These two genes have much larger signals in the presence of CHX compared to other genes at 8hpi, and these signals are larger than the signal intensity at 8hpi without CHX (resulting in a negative mean difference for this cluster). This may be the result of altered expression of one of the coterminal ORFs, such as increased expression of US12 after prolonged exposure to CHX and increased expression of the US03 kinase, for example. Unfortunately, microarray analysis does not allow us to determine which transcript may have altered the kinetics of the group.

5.4.4.3. Microarray analysis of gene expression in the presence of acyclovir

The late transcripts UL18-20 (capsid proteins VP23 and VP5, and membrane protein), UL24-26.5 (membrane protein, capsid protein, polyprotein (protease), and capsid-associated protein ICP35), UL27-28 (glycoprotein B, capsid-associated protein), UL33-35 (capsid/packaging, membrane-associated, and capsid VP26 proteins), UL48 (tegument protein VP16), UL49-49A (tegument protein VP22, and a membrane-associated protein), and US08-09 (glycoprotein E, and a tegument protein) share cluster membership in clusters 8 and 4 in experiment 4 (figure 19). These two clusters also contain a number of the early transcripts, and partially early coterminal groups mentioned in section 5.4.3. Given no apparent explanation for the altered kinetics of so many late transcripts, these ORFs may be interesting targets for further investigation in order to clarify these preliminary results.

The behaviour of two transcripts in the RL1 region of SA8 is also of interest following treatment with acyclovir (ACV). Probes RL1-a and -e share membership in cluster 1 (figure 19), which demonstrates the lowest level of inhibition by ACV of all clusters (24% reduction), although their signal in the presence of acyclovir is not very strong. These probes are antisense to those noted to hybridize in the presence and absence of cycloheximide treatment. This combination of results may therefore suggest a RL1-like transcript in SA8 which is expressed later in infection, following expression of an ORF P-like transcript from the opposite strand of the same region.

5.4.5. Temporal kinetics of SA8 gene expression

Experiment #3 – Additional comparisons were conducted using data collected during SA8 infection in Mrc-5 cells in experiments #1 and 3. Pairs of time points: 5hpi and 8hpi (experiment #3), and 5hpi and 13hpi (experiment #1), were compared in order to determine relative changes in viral gene expression. Gene expression at 5h represents early gene transcription, and both 8h and 13h illustrate late gene expression.

The mean signal intensity at the later time point (8h or 13h) and the percentage change between signals (F) are presented in figure 20 for the 5hpi vs 8hpi comparison, and in figure 21 for comparison of 5hpi and 13hpi. Tables 7 (5-8hpi) and 8 (5-13hpi) list all transcripts in each cluster, categorized by their predicted function (structural or non-structural, mixed, or unknown) based on their HSV-1 homologues. The differences in signal intensity at 8h and 13h cannot be directly compared since these data were derived from different experiments. The larger variation of values within each cluster observed in the comparison of 5h and 8h signals may stem from these experimental differences.

Figure 20 – SA8 microarray gene clusters at 5h and 8hpi. Clusters in experiment #3 were identified using signal intensity of "A" genes at 8hpi (Y axis) compared to the difference between 8hpi and 5hpi signal intensity (X axis). This difference is expressed as a percentage of untreated intensity, therefore 0% indicates equality of signal intensities at the two time points and 80% indicates minimal signal at 5hpi compared to 8hpi.

A

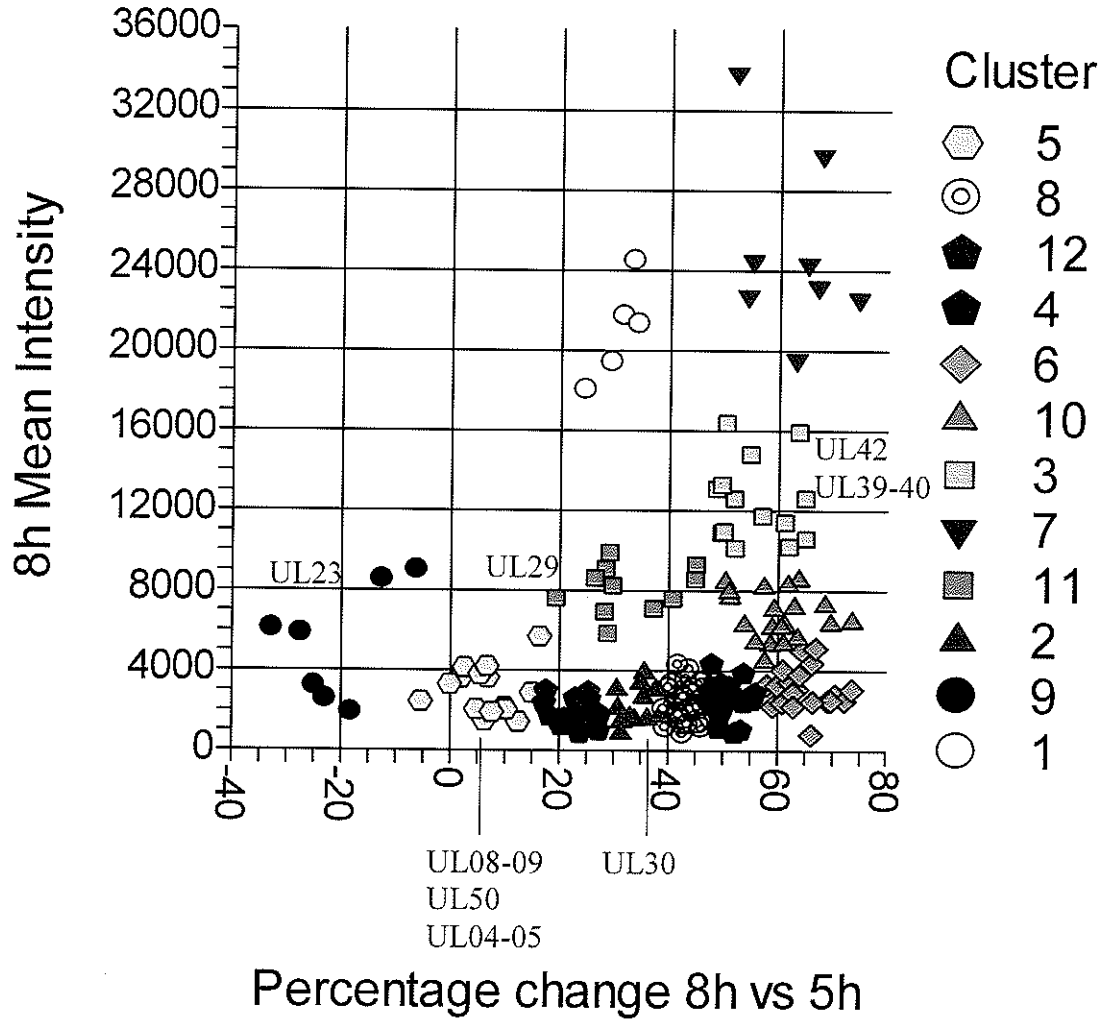


Figure 21 – SA8 microarray gene clusters at 5h and 13hpi. Clusters in experiment 1 were identified using signal intensity of "A" genes at 13hpi (Y axis) compared to the difference between 13hpi and 5hpi signal intensity (X axis). This difference is expressed as a percentage of untreated intensity, therefore 0% indicates equality of signal intensities at the two time points and 100% indicates minimal signal at 5hpi compared to 13hpi.

A

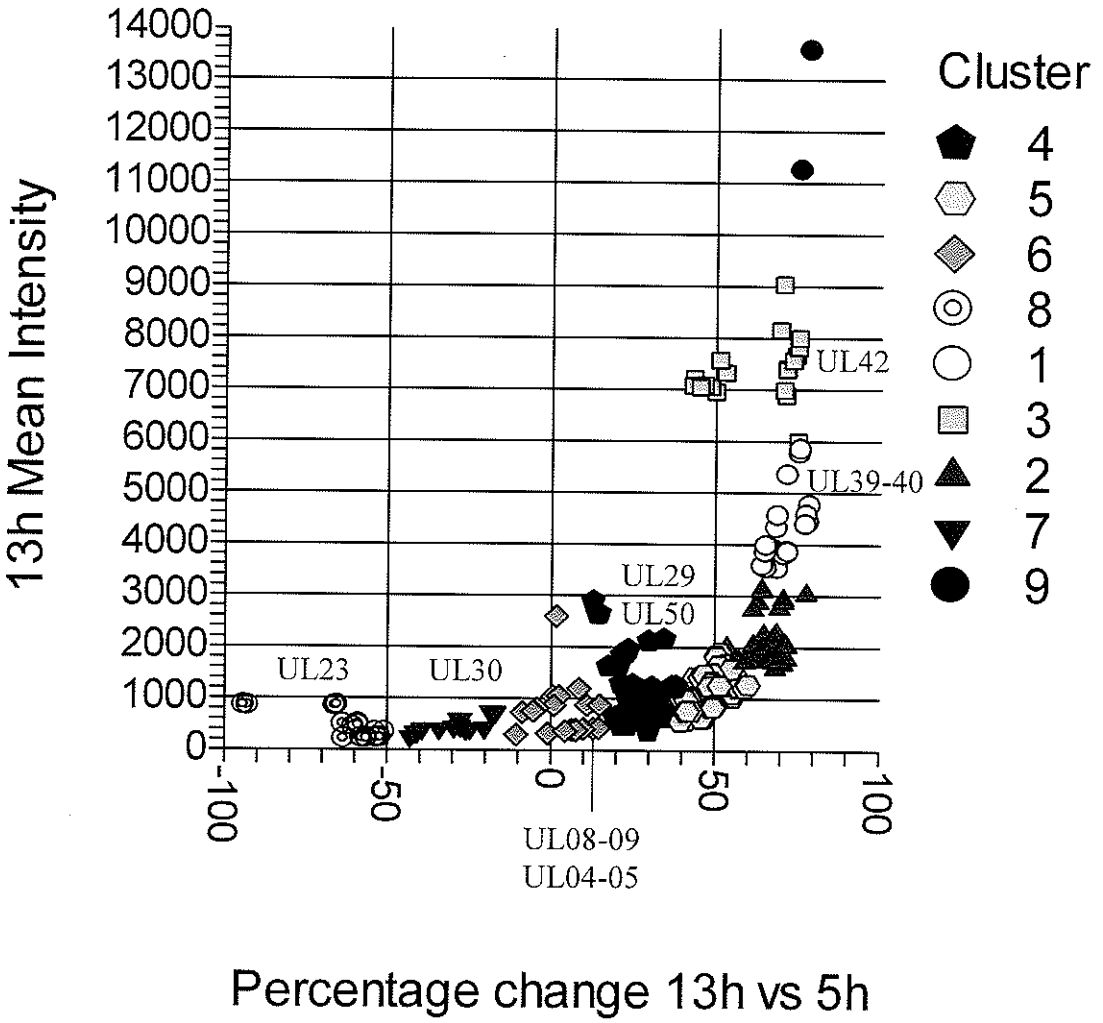


Table 7 – Clusters of SA8 transcripts at 5h and 8hpi.

Cluster	**9**	**5**	**12**	**2**	**8**	**11**	**4**	**6 / 10(5%)**	**10**	**3**	**7**
Non-Struct	UL23	UL08-09 UL50		UL30		UL29			RS01	UL42	US01
Struct			UL55 UL56	UL15	UL43 US02	UL06-07 UL22	UL41	UL37 UL38 UL44-45	UL16-17 UL36 UL31-32	UL03 UL10 UL33-35 UL39-40 UL46-47 US06-07	UL18-20 UL27-28 UL48 UL49-49A US08-09
Mixed		UL04-05								UL01-02 UL52-54	UL11-14 UL24-26.5 US03-05 US10-12
Unknown		RL1d RL1e		RL1b			RL1anti-a	RL1anti-c UL21			

Cluster membership of all transcripts is categorized by gene function and clusters are presented in the order seen in figure 20.
NB: Struct = structural

Table 8 – Clusters of SA8 transcripts at 5h and 13hpi.

Cluster	8	7	6	4	5	2 / 5(3%)	1	3	9
Non-Struct	UL23	UL30	UL08-09	RL02 UL29 UL50 RS01			UL39-40	UL42 US01	
Struct				UL15 UL55 UL56	UL22 UL37 US02	UL38 UL41 UL43 UL48	UL03 UL06-07 UL10 UL16-17 UL31-32 UL36 UL44-45	UL27-28 UL33-35 US06-07 US08-09	UL18-20 UL46-47 UL49-49A
Mixed			UL04-05		UL52-54		UL01-02 UL11-14	US03-05	UL24-26.5 US10-12
Unknown	RL1anti-d RL1e	RL1anti-e RL1c RL1d	RL1a RL1anti-b		RL1anti-a RL1anti-c RL1b UL21	UL51			

Cluster membership of all transcripts is categorized by gene function and clusters are presented in the order seen in figure 21.
NB: Struct = structural

Since the difference of intensity values between 5h and 13h samples is greater than between 5h and 8h signals, the data is spread across a wider range of X axis values in figure 21.

Comparison of both 5hpi and 8hpi and 5hpi and 13hpi reveals that UL23 shows greater abundance at 5hpi than at either 8h or 13h post-infection. The predicted early transcript UL50 (dUTPase), and early or partially early coterminal series UL08-09 (helicase/primase, origin-binding protein), and UL04-05 (structural protein, helicase/primase protein) are present in an adjacent cluster to UL23 in experiment #1. These transcripts have substantial expression at 8hpi, which changes little in comparison to their expression at 5hpi. Transcripts UL04-05 and UL08-09 also maintain this relative equality of expression when comparing transcript abundance at 5hpi and 13hpi. UL50 is expressed at a slightly lower level at 5hpi compared to 13h, and it shares this characteristic with genes UL55 and UL56 (both nuclear proteins).

UL30 (DNA polymerase) belongs to the same cluster (cluster 2) as UL15 (capsid-associated protein) which also share or abuts the cluster containing UL55 and UL56. Cluster 2 is defined by moderate signal intensity at 8hpi and greater intensity at this later time point. Having approximately the same difference in expression between 5h and 8hpi, UL29 (ssDNA-binding protein) is found in a nearby cluster (cluster 11) which is differentiated from cluster 2 by stronger signal intensity at 8hpi.

Early genes that are expressed to a greater degree at 8hpi compared to 5hpi include UL42, and UL39-40 as well as the four coterminal transcript groups containing at least one 'early' gene member: UL01-02, UL11-14, UL52-54, US03-05. Interestingly, UL11-14 and US03-05 share a cluster with US01 and US10-12, as they did in experiment #2.

Consistent with their strong signal in the presence of acyclovir at 9hpi, the two clusters to which all these genes belong display some of the brightest signals at 8hpi.

In the comparison of gene expression both at 5 vs. 8hpi, and 5 vs. 13hpi, these early and 'partially early' coterminal groups are joined by many other 'late' transcripts in the three clusters of each data set which display the largest signal intensity, and the most dominance of signal from the later time point over that of 5hpi. Since many of these genes encode structural viral proteins, their expression at high levels at later times post-infection is consistent with the kinetics of their HSV-1 homologues.

5.5. Summary of SA8 gene expression

Microarray analysis of SA8 viral gene expression revealed a surprising difference between SA8 and HSV-1. SA8 expresses only one of the five HSV-1 "immediate-early" (IE) regulatory proteins during infection in a human cell line. It was confirmed by qRT-PCR and northern blot in several independent experiments that the SA8 RL02 gene is expressed with IE kinetics while RS01, US01, US12, and UL54 are all expressed with early or leaky-late kinetics. A combination of techniques were used to study the gene expression kinetics of SA8 and all assays support the unexpected conclusion that the transcription of IE genes during SA8 differs from that in HSV-1.

Development of a novel analysis technique for viral gene expression microarray data allowed us to draw preliminary conclusions on the expression of SA8 early and late gene expression. A summary of SA8 gene kinetics in comparison to those of HSV-1 is presented in table 9. The most striking difference is between the HSV-1 IE genes and their SA8 homologues. Transcripts from the region of the SA8 genome corresponding to HSV-1 RL1 seem to demonstrate different kinetics than those in HSV-1, but further

analysis is required. Microarray analysis of SA8 early and late gene expression shows that the majority of SA8 genes have similar kinetics to their HSV-1 homologues. Two notable exceptions are UL08-09, UL30, and these transcripts may be useful targets for further study.

Table 9 – Summary of SA8 gene kinetics based on microarray, qRT-PCR and northern blot analysis, compared to the kinetics of their HSV-1 homologues (143, 172).

Transcript[a]	SA8 Expression	HSV Expression	Transcript	SA8 Expression	HSV Expression
RL02	IE	IE	UL30	L	E
RL1a	E	L	UL31-32	L2	L2, L2
RL1b	?[b]	unknown	UL33-35	E or L1	L, L1, L2
RL1c	?	unknown	UL36	L	L2, L2
RL1d	?	unknown	UL37	L2	L1
RL1e	E	L	UL38	L	L2, L2
RL1anti-a	IE	pre-IE[c]	UL39-40	E or L1	E, E
RL1anti-b	?	none	UL41	L	L2
RL1anti-c	?	none	UL42	E or L1	E
RL1anti-d	?	none	UL43	L2	L
RL1anti-e	IE	pre-IE	UL44-45	L	L2, L
RS01	E	IE	UL46-47	L2	L, L1
UL01-02	E or L1	L, E	UL48	E or L1	L
UL03	L	L	UL49-49A	E or L1	L2, L
UL04-05	L	L, E	UL50	?	E
UL06-07	L	L, L	UL51	L	L
UL08-09	L	E, E	UL52-54	E or L1	E, L, IE
UL10	L2	L	(UL54)	E	IE
UL11-14	E or L1	L, E, L, L2	UL55	L	L1
UL15	L	L	UL56	E	L2
UL16-17	L2	L, L	US01	E	IE
UL18-20	E or L1	L1, L1, L	US02	L	L2
UL21	L	L1	US03-05	E or L1	L, L, L
UL22	L	L	US06-07	L	L, L
UL23	E	E	US08-09	E or L1	L, L, L
UL24-26.5	E or L1	L, L2, L, L	US10-12	E or L1	L1, L2, IE
UL27-28	E or L1	L, L	(US12)	E	IE
UL29	E or L1	E			

[a] Transcript names in parentheses belong to coterminal units, but their expression has been analyzed by specific methods, so their kinetics are listed separately. [b] SA8 genes with yet undetermined expression are noted "?". [c] Pre-IE kinetics are described for ORF O and P in HSV-1, although these transcripts are inhibited by the production of ICP4 (RS01).

6. DISCUSSION

6.1. Comparison of SA8 IE gene kinetics with HSV-1 homologues

DNA microarray, as well as qRT-PCR and northern blot analysis of the SA8 homologues of HSV-1 immediate-early genes revealed a unique transcription pattern in this simian herpesvirus. RL02 (ICP0) was the only gene expressed with immediate-early kinetics during SA8 infection of a human cell line, compared to the five IE genes expressed during HSV-1 infection. The expression of SA8 US01 (ICP22), RS01 (ICP4), UL54 (ICP27), and US12 (ICP47) is consistent with early gene kinetics in that their expression is greatly reduced in the presence of the protein synthesis inhibitor, cycloheximide.

RL02 (ICP0) is conserved to varying degrees among a number of members of the alphaherpesviridae subfamily (26, 162, 182). HSV-2 ICP0 differs somewhat from that of HSV-1 in that intron 1 is shorter (and lacks regulatory sequences), and there is sequence divergence in the N-terminal portion of the protein (105).

The choice of cell line can significantly influence HSV transcription (158) which might lead to speculation on whether this has influenced the differences observed in SA8 viral IE gene expression. Mrc-5 cells were previously used in our lab to study viral and cellular expression during HSV-1 and HSV-2 infection and the expected pattern of IE gene expression was observed, so these cells clearly support typical HSV transcription. However, it is possible that human cells lack transcription factors required for the efficient expression of RS01, US01, US12, and UL54 in the absence of *de novo* protein synthesis of other viral factors, but SA8 infection may proceed in cell culture by other yet unknown mechanisms. Preliminary data from our laboratory show that RL02 is still the

only immediate-early gene expressed by SA8 in culture in a green monkey cell lines (Vero cells), which suggests that the unique expression of RL02 prior to viral protein synthesis is not the result of restrictive conditions in human cells.

Further investigation of gene expression and phenotypes of SA8 deletion mutants *in vivo* will provide more insight into the mechanisms of simplexvirus pathogenesis. It has been observed that some viral genes are critical for replication in the presence of an intact interferon response which are dispensable for replication in cell culture (143). For example, the localization of different isoforms of Sp100 within the cell have been implicated in the antiviral interferon response and repression of ICP0 (118). This may explain the non-virulent phenotype of SA8 in humans although lytic infection can occur in a human-derived cell line.

Both ICP0 and ICP4 increase transcription of early genes during HSV-1 infection (148). However, some genes depend more heavily on one transactivator over the other. For example, ICP6 is almost entirely transactivated by ICP0 (37). It is possible that SA8 infection in human cells is inefficient, crippled by the delayed transcription of the "immediate-early" proteins, or alternatively, SA8 genes may rely more heavily on the general transactivation by ICP0 (and presumably VP16).

Analysis of the transcriptional profile of cells infected with an HSV-1 mutant virus which expresses only ICP0 reveals p53-independent cell cycle arrest as well as the induction of a subset of cellular genes involved in cell-cycle regulation, including p21, gadd45 and mdm-2 (68). These mutant viruses display high levels of gene expression, but suffer the arrest of cellular DNA synthesis and eventually die (183). Similarly, ICP4-mutant HSV-1 viruses are cytotoxic (74). SA8 may survive the negative effects observed

during infection with HSV-1 IE deletion mutants because it is capable of eventual expression of ICP4, and the other essential IE genes. Further exploration of promoter usage and transcription factors required during SA8 infection may provide new avenues for the creation of HSV-1 mutant viruses for use in gene therapy.

Interestingly, the Gadd45 protein is able to alter nucleosome structure (19), suggesting that its induction by ICP0 may facilitate the protection of viral DNA from cellular silencing (68). Combined with the activity of ICP0 in preventing the accumulation of histones on viral DNA, as well as its other activities, it may be possible that SA8 completes all the necessary functions for initiation of the lytic cycle in the absence of immediate expression of ICP4 and the other "IE" genes.

Beta-gene expression down-regulates HSV-1 α-gene transcription (54). In particular, ICP8 (UL29) represses expression of all viral IE transcripts. Given the importance of the "immediate-early" transcripts, especially ICP4 (RS01), SA8 may have developed mechanisms to allow the continued transcription of these key regulatory proteins in the presence of other β-gene expression.

Preliminary analysis of SA8 promoters of RL02, RS01, US01, UL54 and US12 did not reveal an obvious association with the expression kinetics observed. A clear TAATGARAT site was not found in the RL02 promoter. Modified TAATGARAT-like sequences are found in the RL02, and RS01 promoters of SA8, however it is not clear if these motifs are functional transcription enhancer binding sites.

6.2. Analysis of SA8 IE gene expression by qRT-PCR

Microarray gene expression results for the 3 non-coterminal immediate-early SA8 genes as well as US12 have been confirmed by qRT-PCR. Expression of UL54 (ICP27)

was not analyzed by qRT-PCR because it would not yield meaningful results on the kinetics of this gene, since UL54 is in a coterminal series, downstream of UL52 (helicase/primase protein) and UL53 (glycoprotein K). However, northern blot data from our lab shows that UL54 is also strongly inhibited by CHX. Relative quantification of gene expression during SA8 infection with and without cycloheximide treatment reveals that only RL02 has sustained transcription in the presence of cycloheximide and is thus the only SA8 gene with "immediate-early" kinetics. Consistent with the observation that IE transcript abundance is increased following inhibition of *de novo* protein synthesis (58), greater expression of RL02 is observed in CHX-treated samples by microarray as well as qRT-PCR analysis.

Although their kinetics differ from that of HSV-1, increased expression of these four SA8 genes between 5hpi and 9hpi is consistent with previous literature on the kinetics of their HSV-1 homologues. RL02 (ICP0) sustains relatively high levels of transcription at late times, while RS01 (ICP4) is expressed at low levels (58). QRT-PCR analysis of SA8 RL02 indicates 5-fold more expression at 9h than at 5h. By way of comparison, SA8 US06 (glycoprotein D) is expressed with late kinetics (strongly inhibited by ACV), and increases in abundance 57-fold between early and late times post-infection.

6.3. Comparison of SA8 E and L gene kinetics with HSV-1 homologues

Analysis of early and late gene expression was performed using cluster analysis. The majority of predicted 'early' genes in SA8 show kinetic characteristics compatible with this description, but the degree of inhibition by ACV and the strength of signal observed in the presence of ACV vary considerably. UL23 (thymidine kinase), UL29 (ssDNA-binding protein), UL42 (dsDNA –binding protein), and UL39-40 (small and large

subunits of the ribonucleotide reductase) are clearly expressed with early gene kinetics. The presence of coterminal transcript groups UL01-02, UL11-14, UL52-54 and US03-05 in these clusters lends evidence to one or more of their member genes having early kinetics, presumably ORFs UL02 (uracil DNA glycosylase), UL12 (alkaline nuclease), UL52 (helicase/primase) and US03 (protein kinase), based on literature regarding HSV-1 transcription.

The early gene kinetics of UL29 (ssDNA-binding protein, ICP8) is consistent with the kinetics of its HSV-1 homologue which is implicated in, and required for efficient late gene transcription (107). The involvement of proteins used in DNA replication for activation of late gene transcription is also seen in bacteriophage T4 (63, 163).

Analysis of transcript abundance between two time points during SA8 infection supports the hypothesis that the expression of SA8 genes is similar to that of their HSV-1 homologues. Many early (and 'partially early' coterminal transcription units) are found in clusters with moderate signal intensity, and less dramatic decline in expression at later times. The majority of structural genes have considerable signal intensity at later times post-infection, and this corresponds to huge increases in signal intensity compared to the earlier time point. Further analysis of SA8 gene expression following the time course of infection would clarify these findings and highlight specific transcripts whose kinetics vary from those of HSV-1. Time limitations did not permit this additional analysis for this thesis.

UL30 (DNA polymerase) and UL08-09 (helicase/primase, and origin binding-protein) share their kinetic properties with several late genes. HSV-1 UL30 has been

described as having early or late kinetics (143), so this result may be consistent with HSV kinetics.

Reduced expression of UL23 at later time points is consistent with the literature on HSV-1 which suggests that thymidine kinase gene expression is turned off a few hours after infection (97).

Further analysis of "late" genes that share cluster membership with "early" genes observed in experiment #2 may be due to misclassification of the kinetics of these genes in HSV-1. Although they were designated as late genes, some discrepancy exists in the literature on their HSV-1 homologues as to the precise kinetics. In contrast, a number of other genes also described as "late" in HSV-1 have kinetic behaviour more characteristic of late gene expression by microarray analysis, and these may in fact be the true complement of late genes expressed by SA8. The analysis of early and late gene expression in SA8 by DNA microarray and cluster analysis is based on a single experiment, and the statistical significance of these results could not be tested. Therefore, these results should be confirmed by further experiments, including qRT-PCR.

Changes in transcription kinetics can play a significant role in viral fitness. For example, Tran *et al.* observed that changing the 'leaky-late' (γ_1) kinetics of virion protein 5 (VP5, UL18) in HSV-1 resulted in a reduction in virulence as well as viral yield in vivo (164), while studies in cell culture revealed no reduction in replication (97).

These data suggest that SA8 has different gene kinetics than HSV-1, and suggests a number of genes which may prove to be interesting targets for further analysis.

6.4. Presence of transcripts in the RL1 region of SA8

The possible identification of ORF O and/or P upstream (antisense) of RL02 in SA8 is very interesting as these transcripts are expressed in HSV-1 only in the absence of ICP4 due to repression of transcription at the ORF P promoter by ICP4 binding (191). These transcripts are designed pre-α as their transcription is induced by VP16 but inhibited following expression of the IE proteins (91). Lagunoff *et al.* have shown that overexpression of ORF P leads to productive infection, and that the expression of ORF P does not require *de novo* protein synthesis (91). Similarly the subsequent detection of a transcript from the sense strand later in infection (and showing limited reduction in expression in the presence of cycloheximide) might suggest an RL1-like transcript which would be expressed following the "early" kinetic control over ICP4. However, the signal intensities measured by microarray analysis are relatively small. As well as the difficulties in differentiating low signals from experimental "noise", the measurement of "percentage of signal change" used for interpretation of the level of expression under different conditions is highly susceptible to fluctuations of low intensity values. Although this technique is useful for the analysis of low-level genes which are known to be transcribed, it may be misleading for the interpretation of "unknown" transcripts. Further analysis of this region is warranted before conclusions can be drawn on the existence of ORFs in this region, but the locations identified by the RL1-a and -e probes may provide useful targets for analysis.

6.5. Challenges presented by microarray analysis of herpesvirus gene expression

There are a number of challenges for microarray analysis of herpesvirus expression which affect experimental design to different degrees. First, host-transcription is shut-off

by simplexvirus infection, which limits the use of uninfected cells as a baseline for cellular gene transcription. In addition, for analysis of viral gene expression, an appropriate control condition is necessary in order to accurately compare array results among replicate data sets. The small number of data points available for statistical analysis of viral gene expression (due to a limited number of ORFs) presents yet another hurdle for analysis. Many common protocols for microarray analysis must be applied to data sets containing tens of thousands of largely indistinguishable signals. We believe that normalization of data sets using the exogenous Lucidea control is the most reliable way to compare viral gene data.

Recent developments in the field of microarray analysis of limited time-course experiments provide alternatives to the protocol presented here, such as the quadratic regression analysis by Liu *et al.* (98). In order to apply this technique to SA8 gene expression, additional time points would have to be selected as the quadratic regression protocol is designed for 4 to 10 time points. ANOVA-based methods are still suggested for fewer time points. Other methods that treat the variable of time as continuous, such as B spline-based methods are best applied to a large series of time points (98). During further analysis of SA8 (or other simian simplexvirus) gene expression, it may be valuable to compare these methods to the protocol presented here.

Coterminal and overlapping viral genes makes interpretation of microarray results difficult as probe signals may be due to the abundance of multiple different transcripts. This obstacle was overcome by applying cluster analysis only to gene signals from non-coterminal, non-overlapping genes, classifying the remaining genes into the same

clusters, and subsequently interpreting the kinetic behaviour based on the best quality probe of the multiple ORFs within each transcript.

7. REVIEW OF OBJECTIVES

A gene expression DNA microarray was developed for SA8, containing probes for both cellular and viral genes. A number of challenges were presented by the microarray format and features of herpesvirus transcription. As such, analysis was only completed for viral gene expression in SA8. However, this revealed a fascinating difference between this simian herpesvirus and the prototype simplexvirus HSV-1. Although HSV-1 expresses 5 genes with immediate-early kinetics, SA8 expresses only one. The remaining "immediate-early" genes are expressed with early or leaky-late kinetics. In addition to this contribution to the knowledge of herpesvirus transcription, a novel protocol was developed for the analysis of viral gene expression microarray data.

8. FURTHER WORK

Analysis of the "immediate-early" gene expression of other simian simplexviruses such as B virus and HVP-2 will reveal whether the differences observed between SA8 and HSV-1 are characteristic of simian viruses or whether the gene expression of SA8 relates to its inability to produce productive infection in humans. qRT-PCR analysis of "IE" gene expression during simplexvirus infections of both human cell lines and monkey-derived cells lines will contribute to answering this question.

Investigation of a selection of SA8 β and γ genes by qRT-PCR as well as northern blot analysis of other coterminal transcript groups would confirm the kinetics observed by microarray analysis. This remains the simplest technique used to distinguish the

kinetics of many viral genes, particularly genes which share a transcription termination site but which are expressed under different conditions during infection. In addition, a more comprehensive analysis of SA8 promoter structure as well as cellular transcription factors required for expression in both human and simian cells will complete our understanding of transcriptional regulation in SA8.

Following these targeted approaches to simian simplexvirus gene expression, microarray analysis may be used to complete the global picture of viral gene regulation. As suggested by Stingley *et al.* in 2000, microarray analysis of cellular or tissue transcriptional responses to herpesvirus infection are of great value in the study of anti-viral and cell cycle regulatory pathways in both human and animal model systems (158). Further analysis of both viral and cellular gene expression during simian simplexvirus infection will clarify the relationship of transcriptional regulation to pathogenicity in different hosts.

9. APPENDICES

9.1. Appendix A – DNA Oligonucleotide microarray probes for SA8 viral genes

Probe name	Pos.	Sequence	Tm	GC %	Score
RL1a	821	GCCTGCGATCCTGGGGTGAGTCGGTGGTTTTGGGGTCGGTGGTCGTTCCC	82.5	66	**5.8**
RL1b	1999	CATATAACGCCGGGGACACCATCGGCACCCCAGCCAGTCGCCGTCGGACC	83.3	68	**36.7**
RL1c	1628	TGATTAAAAAATCGATGCGCCCGCCCAGCGGGGAGAGCCCGGAAGATCCC	80.8	60	**10.0**
RL1d	335	CAGCGTCCGCGCGCTCGCCACCCTCACTCTCGGCCGCCGTCGGTCTCTCT	86.4	74	**17.5**
RL1e	1277	GCCGAGGGGGAGGCGCCCGGCGGAAGCCCGGATATGCTGATGCGCCCCTC	87.3	76	**15.0**
RL1anti-a	821	GGGAACGACCACCGACCCCAAAACCACCGACTCACCCCAGGATCGCAGGC	82.5	66	**29.2**
RL1anti-b	1999	GGTCCGACGGCGACTGGCTGGGGTGCCGATGGTGTCCCCGGCGTTATATG	83.3	68	**20.0**
RL1anti-c	1628	GGGATCTTCCGGGCTCTCCCCGCTGGGCGGGCGCATCGATTTTTTAATCA	80.8	60	**18.3**
RL1anti-d	335	AGAGAGACCGACGGCGGCCGAGAGTGAGGGTGGCGAGCGCGCGGACGCTG	86.4	74	**21.7**
RL1anti-e	1277	GAGGGGCGCATCAGCATATCCGGGCTTCCGCCGGGCGCCTCCCCCTCGGC	87.3	76	**15.0**
RL02a	4343	CTCAACAAGACCGTCACGGGCGACTGCCTGCCGGTCCTCGACATGGAGAC	81.3	64	36.7
RL02b	4611	GCAGCTTCCACAGCCTGGACTCGCGCCACCCCTGGACCCCCGCCGAGGGC	87.3	76	**39.2**
RL02c	4490	CTCCCCGAGGCCGCCCCCGGGTGCGTGTCCCCGCCCGAGTACCCGGGAGA	89	80	30.8
LAT CDS1a	5135	CCAGAGCGACGGGGGACCGAGGCCCGGCCAGAGACACAGGGGACAGAGGC	85.5	74	**15.8**
LAT CDS1b	5882	CGAGGGGAAGGGGGTCGCGGGGAAGGCGGGGAAGGCGGCCGAGGGGAAGG	87.7	78	6.7
LAT CDS2a	2801	CTATCTCATCGTGGGCGTGAAGCCCGACGGCTCGTACAGCACCATCCCGG	81.1	64	**47.5**
LAT CDS2b	4339	CTACCTCAACAAGACCGTCACGGGCGACTGCCTGCCGGTCCTCGACATGG	81.2	64	46.7
LAT CDS3a	6740	TTCGAGCTGAGGTAGACGTGTTGTCGTGTGCTGTGTGTTGTGTGTCGTTC	76.9	52	**16.7**

LAT					
CDS3b	7530	GAGGAACCTACCTGTCCCGCCGGGCGTCTCGGTGGCTCTCTCGGGCCGGG	85.3	74	13.3
RS01a	148667	CGCCGACACGCTGCTCGCCGACGCCGAGGACGCCGCCGCCCTGGAGCTGG	89	80	20.8
RS01b	148074	CACGTCTACCTGCCCTGCGCCGTGCTGCCCGCCGCGCAGTGCGACGTCCG	87.5	76	25.0
RS01Ra	148647	GTGCGCTGGGCCGCCGGCCGCGCCGACACGCTGCTCGCCGACGCCGAGGAC	90.8	82.4	20.0
RS01Rb	148741	CCCTGGGGGGCGCGACCCCGGACTGGGACCCGGACGGCGAGCCGGCGGCC	91.6	86	**26.7**
UL01a	10105	GATCGCAAGCCCCGACCGAAGCCCAGAAGAAACTCCACCGAGCCCTCCCG	82.3	66	41.7
UL01b	9940	GTCAGCTCTGAGTACTCTCGGACGCGGCGGTGTCTTGGCAGAGACACGCC	81	64	58.3
UL02a	10908	CCTTCGGCTCCTGTCGCCATTTCTGCCTCGCGAACCAGTACCTCCGGGAG	81.2	64	46.7
UL02b	10769	GTCGTCCGTCGGCTCGCCGCATCGCGCCCCGGTTTGGTCTTTATGCTCTG	82.8	66	**63.3**
UL02c	10618	CAACTGCTGTCCGGAGGCGACGATGAGCGGCCACGGGTGTCTGGAAAAGTG	81.5	62.7	46.7
UL03a	11764	GCACGTCCACCCGCCTCGCCGATCCCGCCAGTGAAGACCCATGGCAATAA	82.6	64	**51.7**
UL03b	11618	CTCGCGCAAGCCGCGCAAATACTACACGCGCTCCTCCGACGGACGCATCTC	82.5	64.7	46.7
UL04a	11929	GGCGGCGGCGGACCTGCTGGTGGAGGTGATGAAGGAGATCCAACTGAGCC	82.4	66	20.0
UL04b	12087	GTGGTGTGGGGCGAGGACTCCATCTACATCACGGTCACCGTCTACGGCAG	79.9	62	**46.7**
UL04c	12245	GTCGGACGGAGTTCTGCGCCTACCGGTTCGTGGGCCGGTCGGAGCGCCAG	85	72	43.3
UL05a	12561	CACGAGCGCGATGATGTCATCAGCGGGCACATACTGTCGGCCCTCCGTGA	81.3	62	34.2
UL05b	12823	CTTCATGTCGGTCGTCAACACCAACGTGAGCGAGTTCGTGGAGTCCGTGG	78.7	58	45.8
UL06a	16965	GTGCAGACTTACCTGACAGACCTGGCCGCCCTCTTCGTCGCCGACGTCCG	82	66	44.2
UL06b	16803	CGCGAGACGGCCATCTTCTACTCCAGCGGGGCCATCGCCCTGTTCGTGGC	83.4	68	36.7
UL07a	17778	CCTGCCCCACTGGTGTATGGCCCGCTACCTCCTGGATCACAACCCGCCCC	83.2	68	**53.3**
UL07b	17987	CGTCGGACGCCCCCAAGACCCAGACGGCCGCGCTCTTCTACCGCTTCGTC	84	70	37.5
UL08a	18263	CTTCGCGTATCCCTTCGAGGAGCGATTCCTCTTCCTGTTCGATCCCAGCG	78.2	58	**36.7**
UL08b	18397	GTGCCGATCGACCTCTACCCCCGCCCCCTCGTCCTGCCCCCCATCGACTG	84.5	72	10.0
UL08c	18550	GACGCCGAGTACTGGGCGGTCGCGGAGGGCCCGTCCGACCCCGGAGACAC	86.4	76	19.2
UL09a	20715	CACGCACCGGCAGCTCTACGCGCTGCTCATGGCGCACAAGAGGGACTTCG	82.9	66	32.5
UL09b	20544	CGGAGGCCTGGCCCATGACGCACGGCGCGCTCGACTTCAGCAGGCTATAA	83.4	66	**36.7**
UL10a	24342	CATCCAACACCACCCCCGGAGGGTCGAGGAGCCCATCTACGAGACCGTGG	81.6	66	47.5
UL10b	24205	CCCGTCGAGCCCGTGTACGCCGAGATCAGGTACGCCGGGGAGTCTGACGT	83.1	68	**56.7**
UL11a	24849	CTCCCGCCTCGGACCCTGCTGTTGCAGACGGAACGTCCTCGTGACCGACC	82.9	68	46.7
UL11b	24728	TTCTACATCTCCCCGGAGCTGCGGGTGGTGACGCAGCCCCCGAGCCAGCC	84.9	70	66.7

UL12a	24921	GTCTACGACCTGATCCGCCGCAACAGCCGCCTGGCCTTCGACGAGACGCTC	82.9	66.7	41.7
UL12b	25078	GAGCCCATCTTCGCCAACCCCCGCCACCCCAACTTCAAGCAGATCCTGGT	81.7	62	**87.5**
UL12c	25201	CACCGCACGCCCGACGAGGAGGGGCTGAGCCTGCGCCTGGGGGACGCCCC	89	80	36.7
UL13a	26730	CTGAGGCCGCTGCTGATCCTGGTGTCCCGCCTCTGTCACGCCAACCCGTC	83.3	68	60.0
UL13b	26902	CCTCTTCGTGAACTCGGTCGAGACCAACGCCCACGGCCCGCCCTGCGACG	84.2	70	55.8
UL13cRa	26813	CTCCAGCGCCACTTCGCCGGGCAGTATAAGAACTACCAGCGCACCCACGAG	81.2	62.7	37.5
UL13cRb	26702	GCCTCTGTCACGCCAACCCGTCCGCGCGCCACGCCCAGTTCCAAGTGTGA	84.7	68	34.2
UL14a	28204	GAGGACGGATGGATGAGCGCCGAGGACAGCGACCTGCTGATCATGTGGCAG	81.1	62.7	35.0
UL14b	28019	GACCCAGCTTCCGAGAGCGACTGCGGGCCGGGTTCTCGCGGTGGCGACTG	85.1	72	36.7
UL15a	34388	CGCCTCGGACGACCTCATGGTCGCCGTCATCATGGCCGTCTACCTGGCGG	83	68	30.0
UL15b	34266	GTCTCGGTCACCATGCGCCTCCAGACGGACCCGGTCGAGTATCTGACGGC	81.7	66	**36.7**
UL16a	29972	CGCTCTCGTCCTACGCGAGCCGTCTCTTCGCCACCAGCTGCCCGACGATC	82.8	68	41.7
UL16b	30273	CTGCGAGACGTGCGTCCCGTCGGCCCACGCGGACTCGGTCAACGCCCTGTG	85.8	72.5	37.5
UL17a	31203	GTACCTGCCGGGGGAGAGCTACGCCTACCTCTGCCTGGGGTTCAACCGCC	82.8	68	**47.5**
UL17b	31361	CACTACGCTTTCGACTACTACAGCACCGGCGGGGCGACCGTGCGGACGAC	82.2	66	32.5
UL18a	34855	CGCAGGACGGCTACGTGAACACGCTGCTGCAGATGCAGAGCGCGACGCGG	83.9	68	48.3
UL18b	34642	CCGCGGTCGTGCGGTCGGGGGACCTGGCCCCGGTCATCCGATACCACTAA	84.5	70	**70.8**
UL19a	35758	CCGAGACGCACTTCATGCAGTACCTCATCTACGACGCATCTCCGCTGCGG	79.4	60	65.8
UL19b	35950	CTCATCGTGGAGACGGGCTCTGCGGTGTCAACGGCGACGGCGACCAGCGAC	83.9	68.6	46.7
UL20a	40057	CTGGCTCGCTTCTGGGCCAGGACCATCCTAAACGCGCCCGTGGCCTTTTAG	81.6	62.7	46.7
UL20b	40230	GGCCGCGCAGGAGACGTTTGGGGCGGACGGCGTCTTTCGCCTCGTGGCCG	86.7	74	31.7
UL21a	42746	CGTGGAAGACGTGAACCGCCTGTCGGCCTTCGACCGCGGGCTGGTGGGCG	85.5	72	**36.7**
UL21b	42624	CCTCCCCGGCCCACGCCACCTTCGTCGGGGCGTTCGCGGAACTGGTGTAC	85.3	72	30.0
UL22a	43200	GAGGGGCCGGACGTCTTCTCCAGCGACGTGCCCTCCACCGCGCTCCTGCTG	85.5	72.5	**46.7**
UL22b	43332	CAGAGACGCGCCGCGACCTGGGGCTCGTGGGGGCGGTGTTTATGCGGTAC	84.4	70	40.0
UL23a	45988	GTGGACGCTCGACCTGCTGGCCGAGCGCCTGCGGTCCATGCACCTCTTCG	84.5	70	**53.3**
UL23b	45867	ACCCTGGTCACGACCCCCGGGTCGTTGACAGTGCTGCGCACGCTCGCGCG	86.2	72	42.5
UL24a	47589	GAGGCGGGGTACTCCAGAAGATCGCCGCGCTCTTCTGCGTCCCCGTCTCG	82.7	68	27.5
UL24b	47285	CACCGCACCCTCTACCTGTGTCCGGTCCTGCTGTTCGTCTCGCAGAAGAC	80.1	62	38.3
UL24c	47417	CTGTCGACGTACGCCGTGCCCCTCGCGGCGACGAAGACGCGCAGAGGGCG	86	74	5.0
UL25a	49500	GCGTTTAACGTCGGCAGCGACTACGACCTGTTGTACTTCCTGTGTCTGGG	77.7	56	28.3

UL25b	49350	GAGCTCATCAACCGCACCCGCGCGAACACGACCCCGGTCGGCGAGGTCATC	84.1	68.6	35.8
UL26.5a	51518	GACGTTGACGTCGGCCGGGCGGCAGACCTGTTTGTGTCGCAGATGATGGG	82	64	71.7
UL26.5b	51330	CCCACCCGCAGCCGTACGCGACTCCGGCCAACGCCCCCCTCGTTTCCGGG	87.2	76	56.7
UL26a	51518	GACGTTGACGTCGGCCGGGCGGCAGACCTGTTTGTGTCGCAGATGATGGG	82	64	**72.5**
UL26b	51330	CCCACCCGCAGCCGTACGCGACTCCGGCCAACGCCCCCCTCGTTTCCGGG	87.2	76	55.8
UL27a	52096	GACTTTGACGAGGCCAAGCTGGCGCAGGCGCGGGAGATGATCCGCTACATG	81.6	62.7	56.7
UL27b	51943	CCCGGCCCCGGTACTCTCCCCTCCGCGACACGGACGAGGAGGAACTGTAG	83	70	52.5
UL28a	54788	GGTCTCTACCGCTACCCGCCGGGCGTCTACCTCACCTACGACTCCGACTG	81	66	66.7
UL28b	54664	GACGTCTTCTCCATCCTCTACTCGGTCCTGCAGCACCTGGCCCCCAAGCTC	80.6	62.7	**73.3**
UL28c	54955	TTTCAAAAGTTTTACTCGTTTGACGGCATTTCCGGCATAACGCCCACCCAG	75.9	47.1	39.2
UL29a	57585	GTCGGGGAGGCGTTCGACTTCGGGGCGTTCGGGGCCGAGGACGACGACGC	85.8	74	53.3
UL29b	57722	GACGAGTACCTCGGCGAGGAGATGATGGCGTTCACCGCCCGCGCCCTGGAG	83.5	68.6	**67.5**
UL30a	65254	TCAAGGCCCTGTTTGGGAATAACGCCAAGATAACGGAAAATCTGTTGAAGC	74.5	45.1	**37.5**
UL30b	65126	CGAGCCCAAGCGGCGCAGGCTGCTGATCTCGGAGCTGGCCGAGGACCCGG	86.2	74	36.7
UL30c	65376	GCGGGCGCCGGCGTCACGGAGGCGGAAACTCGTCAAAGGCTGCGTACAGC	84.9	70	36.7
UL31a	65414	ATCAACTTCGACGGGGGGCTCATGCTAGAGTATCAAAGGCTGTACGCAGC	77.4	54	37.5
UL31b	65656	CTCTTCTACCCGGACCCCACCTACGGCGGCCACATGCTGTACGTCATCCTC	80.5	62.7	**53.3**
UL32a	66297	CACTCGATCTCGCTCCTGTCGCTCGAGCACACCCTCAGCACCTATGTATG	78	58	52.5
UL32b	66468	GAGGAGCTGCTCCTGTACAAGGCCCGGCTGGCCTGCGGCAACGAGTTCGG	83.6	68	30.0
UL33a	68252	GGTCTTTCCCACCACGGACGCCAAGCTCAACTACCTCTCTCGCACGCAGC	80.6	62	46.7
UL33b	68416	GTCGGTTCGCGGCGGTCATCAACCGCTTCCTCGACCTGCACCAGATCCTG	81.6	64	**80.0**
UL34a	69275	GCCTCGTCCTGGGGGCCGCGCTGTACTGGCTGGTCGGCCTCGCCCGATAA	85.8	72	40.0
UL34b	68950	GACCCGCGCCCGATGATGGCCAGCATGTGGATCCACTGCTTCGTCCGCATG	82.8	64.7	52.5
UL35a	69713	CCGCGTTCGGTCTCCGGCGAACCTACCTGCCGTTTCTGGCCGCCCCCGCC	86.4	74	56.7
UL35b	69453	GGTCCCGCAGTTCCACCGGCCCGACACCCTCCACCCCGAGCGCGTGCAGG	87.2	76	20.0
UL35Ra	69631	GACATGTGTCGCGCGCACGCGAACCACACGCTGGCGCTGCAGCCGATGTTC	84.3	66.7	40.0
UL35Rd	69700	GACTGGCTGCGCCCCGCGTTCGGTCTCCGGCGAACCTACCTGCCGTTTCTG	84.2	68.6	55.8
UL36a	69984	CACGCGCTCATCGATCGCAGCGAAGCGGTACTGACCGGCTTGTATCACGT	80.5	60	**53.3**
UL36b	70248	CCGGACGACACCGAGGGGTCGGACGACACCGAGGACTCGGACGAATCCTC	82.1	68	30.0
UL36c	70115	CACTTCGGACCCCCGCCCCTCTCGGCCAACGCGGCCCTCTCGCGGCGCTAC	87.5	76.5	28.3
UL37a	79388	CCCGGCCCCGGAACGACCCCGAGGAGGACGAGGAGGAGATGCTGCTGTAG	83.3	70	**35.8**

UL37b	79583	TCATGCTGGGGATGGACTGGCTGCTGATGAACGAGATGCTGCAGGTGACG	79.6	58	18.3
UL38a	84831	GTCGTGCTCCTGGAAGGCGTGGTCTGGAGACCGGGCGAGTGGCGGGTCTG	83.9	70	**26.7**
UL38b	84501	CTGGAGTCGGCCCTCGAGCGGGTGTTTGGGCGCCTGCGGACCACCAACAC	84.7	70	19.2
UL39a	88067	CATATAAGCGGGGCCTGAAGACGGGCATGTACTACTGCAAGGTTCGCAAG	77	54	56.7
UL39b	87907	CGGTTCAAGACGGCCTTCGACTACGACCAGGAGCTGCTGATCGACCTCTG	79	60	68.3
UL40a	89128	CTGAGCCTGATGTCCACCGAGAAGCACACCAACTTCTTTGAGTGCCGCAG	78.1	56	**80.0**
UL40b	88939	CTGTTCCGCGAGGCCGTCGAGATCGAGATCGGCTTCATCCGGTCGCAGGC	82.1	66	66.7
UL41a	89592	GAGCTCGTGGAGCTGAGGTACGTCAAGTACGTCATCTCCATCATCGCCCC	77.9	58	**40.0**
UL41b	89411	AGCTGTGGACGACCGTGCCCGCTCCTCCCCCGCCCTACCAGACCGTGCTG	85.6	72	23.3
UL41c	89738	GAGATCCTCACCCCTCCGGAGCTCGTGACCGTCCCGGACGCGCAGCTCGTC	84	70.6	27.5
UL42a	92477	GTACGCCCGATACTTCCCCGACCCGCCCATCCTCCGCGGGCCCCTGTTCTC	84.4	70.6	**63.3**
UL42b	92326	CCGAAGCGACCGACGCGCCCGCCCGCAAGAAGGCCCCGGAAGACCCCGTC	87.2	76	23.3
UL42Ra	92477	GTACGCCCGATACTTCCCCGACCCGCCCATCCTCCGCGGGCCCCTGTTCTC	84.4	70.6	62.5
UL42Rc	92192	CCCGAACTCGGTCTCGGCCGTGTTCCTGCTGGACCCCCGGAGCACCGACG	84.8	72	36.7
UL43a	93789	GTCCGGGGCCTCTTCCTCGCGGTGTACGTGTTGGGGTTGGGGGTCGGGAC	84	70	15.8
UL43b	93621	CACGCCTGCCTCCTGCTCGCGATCCACCTCGTCTGGTTCTTCGCCCCCGG	84.1	70	**34.2**
UL44a	95243	GAGCACAGCGAGTACACCTGCCGGCTCGTCGGCTACCCCGCGACCGTCCC	84.8	72	29.2
UL44b	95119	GGTCGAGCTGCGGTGGCTGCTGGGCGCCGACCCCGCGCCCGCGGAAAACG	88.8	78	36.7
UL45a	96058	GTCGGGCGCGTACTGGTGGGTCAACGGCGACGGCATCAAGGCCTGCCTGC	84.9	70	36.7
UL45b	96192	GGTTCGTGACCTTCGTCCTCGGCGCCCGCGACGCGCTGGGGCTGCCGTAG	86.2	74	**48.3**
UL46a	96429	GCGCCATGCTCACCAAGCTCAGGCACGACGGAGGACACCACAGCCGCTGA	83.6	66	56.7
UL46b	96721	GTCTACGAGGACCTCAGCGACTCGGAGGAGCCCGTCTACGAGGAGATCCC	79.6	64	**65.8**
UL47a	98814	CCGGGACCGTGGCCCCGCGGTTCCTGCTGGAGCAGTCCGAGCACTTCCCC	86.1	74	42.5
UL47b	98686	CTCCTCATGGGCGACGAGGAGGCGGCCGCGCTCCGGGCGCACGTGTCCGG	87.9	78	28.3
UL47c	98980	GAGCCCGCCGTGGGCGCCTACACCCCGCTGCGGTACGCGTGCGTGCTCCG	88.3	78	23.3
UL48a	101200	CTCTGGACATGGCGGACTTTGAGTTCGAGCAGATGTTTACCGACGCGCTG	78	56	**68.3**
UL48b	101352	GACCGTCCCGCGCACGGACGTGAGTCTGGGCGACGAGCTATGTCTCGACG	82.7	68	40.0
UL49a	103074	GGGCCACCAACATCCGCATCATCGTTTGCGAGGGCCTGAACCTGGTCCCG	82.1	64	**75.0**
UL49Aa	104214	CTGGTGCCTTTTTATCTCGGTATGGCGTCACTGGGCGTTTCTCTGGGGAC	77.8	56	35.0
UL49Ab	104348	ATGGGTGCGCGCGTGCTCTTCTTCTGGCTCCTGGCGCTCTGCGTCGTCGAC	84.1	66.7	45.0

UL49b	103264	CAGGAAACTGCACTTCAGCACCGCCCCCACCAGCCCGACGGCCCCCTGGAC	85.3	70.6	53.3
UL50a	105433	CGTCTGGTCTTCACCGCCGAGTTCGACGCCGAAGCCCCGCCCAGCGAACG	84.4	70	10.0
UL50b	105247	GTCCACAACATCACGGACGCCCCGGTCTTCCTGGAGCCCGGCTCGAAGGTG	82.9	66.7	**53.3**
UL50c	105126	CGTGCTGCTGGCGGCCGGGGGCGCGGAGGCCTACTACGTCCTGGGACGCT	87.7	76	15.8
UL51a	105885	CAGCATGGCGGAGAGCGACGTTGTCATGGAGGACGTGGCCATCGCCGAGC	82.7	66	35.0
UL51b	105678	CGTTTCGGGCCGGTCCAGGACCCCGGTTACCGATCGCCCGCTTCTTCTGG	83.2	68	**49.2**
UL52a	109449	CCAGCAAATGCGACAGCAACCGGCTCCGGACCCTGTTCACCGTTGACGCC	82.4	64	40.8
UL52b	109199	GTCGCCTGTCTGGAGACGCACTTTCCCGAGCACGCCCACGAGTACCAGTC	81.2	64	38.3
UL53a	110459	TGGTGGTGGCCTTCCGGTACGAGCAGGAGATCCAGCGGCGCATCTTTGAC	81.4	62	52.5
UL53b	110242	ATGTTCCTCAACATCATCACGTGGTGCTTCACCTCCGCCATCCTGCTCGC	78.9	56	50.0
UL54a	112273	CCCCGCTGTACGTGCACGGGAAATACTTTTATTGCAATTCCCTGTTTTGA	74.8	46	48.3
UL54b	112031	GCGCCGCCTGTCGGACATCAGGGAGATCGCGTCCTTCGTCTTCGTCACGC	82.3	66	**70.8**
UL55a	113010	GTACGCGCTGCACGGCGTCGTCTACCACTGCCACTGCCGCGACCCCTTCTC	84	68.6	**40.8**
UL55b	112888	CTGCCGCCGGTGCTGCGCGAGATGCGCGACAAGGCCAAGGGACGGGTGCTG	86.6	72.5	40.0
UL56a	113344	GGGTCCTGCTCCTCGTCGTGGTGTTTCTTCTGGTGTTCCTGTGGCGCTAG	79.4	60	15.0
UL56b	113552	GTCGTCTTTGAGGATGACCGCGACGGCCCCGATCCCCCGACCCAAGCCTC	83.1	68	**23.3**
UL56c	113675	GACTCCGGTCGTCGCGCCGAGGACTCCACCGAGTCGTCTTCGTCGTCGCC	83.4	70	20.0
US01a	131304	GAGAAACCGGCGGCTGTCGGAGGATGCGCTTTTCCACCACGTGTCCGTAC	80.9	62	**81.7**
US01b	131158	CTCCGTGGCCTCCAGGCTGCGGCGGGAGTTTGTCGCGTTCGACTGGACCC	84.4	70	77.5
US02a	131653	GAAGCGAGGCCGGAGTACGCGTGGGACTCGGACTTGGACTCGGACGAGCG	82.7	68	**33.3**
US02b	131914	GCCCACAACCACCCACTGGCGGCGCTGTTCGGGGCCTACGAGTACCGGTA	84	68	32.5
US03a	134075	CTGGACGCGCTACTACAAGCTGAACCTGGACGTGGAGTACCTCGTCTGCC	79	60	65.0
US03b	133954	GCATCATCCAGCAGGCGCAGGTGCACCCGGACGAGTTTCCGCAGCACGCG	84.3	68	52.5
US04a	135952	TTTCCGTTCCTCACCGCCTCGCCCGGCCTCGACTTCATCTTCCTGCTCAG	81.1	62	**73.3**
US04b	136079	CGCGCCAAGCGCGAGTTCGGTACGCGCGGCTGGCGACCTCCGAGGCGTGA	87.1	74	28.3
US04c	135733	GAGGAGGGGGAGGGGGCCGGGGACGGCGAGCTGCCTCCGGACGAGATTCC	86.3	76	36.7
US05a	136505	GTTTCTCCAGACGGACCGCGCCCTGCCCCTGGTGATCGGGGGCCTCTGCG	85.2	72	60.0
US05b	136339	CTCTCCCGCTCGTCGTCGGCGCCTGGGCGGCCGCGCTGGTGACCAACCTC	87.1	76	53.3
US06a	137921	CAGCTGCCGCCCAACTGGCACGTCCCGGAGGCCGACGACGTCACCATCCAG	85	70.6	40.8
US06b	137797	GACCACGAGCACCCTGCTCCCGCCCGAGGTGACGGAGACGGCCAACGGCAC	85.7	72.5	46.7

US07a	139485	CTGACCGCGATCTCGGAGGAGTCCGAGGCCCCCGCGATCCCGGAACTCTC	83.2	70	45.8
US07b	139204	CCTCCCGGTACGAGCTGACGGCCACGCGGATAGTGCAGATCGCCATCCCG	82.9	68	**59.2**
US08a	141338	GGCACGGACCGGCCTCCGAAGCGGGGCTCCGGGTTCCAGATCCTCTCCGG	85.7	74	67.5
US08Aa	141694	CGCTCGGGCTGGGCGCCGTGCTGTGTGGGGTCGCGTATTACATGACGTCG	83.6	68	70.0
US08Ab	141534	GACGTTCCCATCGGCGTCGCCGAAGACCTGCGGGCCCTCGCAATGACGTG	83.7	68	46.7
US08ARa	141711	GTGCTGTGTGGGGTCGCGTATTACATGACGTCGGCGCCCCGCGGCGCATAA	83.4	64.7	72.5
US08b	141214	AAAGCGGGACCCGGGAACGCAGACTTACATTCGCCTGGCGGACGACGAGC	82.4	64	46.7
US09a	141923	GCAGACCCAGAGCGCCGAGGCGTACTACACGGAGAGCGACGACGAGACGG	82.5	68	46.7
US09b	142077	CTCATCTCGGGCGGCCTGGGGGCCCTCGTGTGCTGGATGGCCCTCCGCTG	86.1	74	**73.3**
US10a	142624	CACGTGCTCGTGCGCCGGCATCTATTCCACCCACTGCCGCCACCTCTTTC	82.1	64	**90.0**
US10b	142794	GAAACTGGCCCATAACCAGTGGCTGCGCGCCAACAGAGACGCCGCCATCTG	82.1	62.7	83.3
US11a	143456	GGACATAAAGCCGCGCGAGGAGCTTCCTTGCCACAGAGCCGCGGGTATCG	81.8	64	46.7
US11b	143210	GCGGACCCCACGTCGTTACGTACAGCCCCCCGCGCCACGCGACCCCCGGG	87.9	78	60.0
US11c	143335	CGTGGAGATGCGCGTCTATCGCCGGCGCCCTCGCGCGCCCCTCGCGAGCC	88.1	78	35.0
US12a	143769	TCATCGTTATATTTGGCGACGGTCGACGCGTTCCTACGGAATCCGCACAC	77.7	54	**90.0**
US12b	143627	GACCGGCCCCTGCTCGCCCCTCCTTCCGCACCGCCCAACCACTCGCACCTC	86.8	74.5	74.2
US12cRa	143585	CTCGCACCTCGCGGCCCGCGAGACCGCGCCTCCTCCCGCCGCGACGCCATAA	88.8	76.9	28.3
US12cRc	146378	CATACGCCGACGAGGAGCGCCGAGAGGCCGCGAAGGCGATCGCCCATCCG	85	72	36.7

[a] Probe names in italic are designed for putative ORF sequences; Probe names in bold are "A" genes, as described in the text; Boxes indicate the multiple probes for a gene or a group of coterminal transcripts; [b] Probe location is described as the 5' end of the probe according to numbering on the sense strand of the SA8 genome, regardless of the direction of the probe. [c] Probe sequence is 5'-3', complementary to the appropriate strand of the SA8 viral DNA. [d] Calculation of the probe quality score is described in the text. Probes highlighted in grey are the 'best' probe for each ORF; Score values in bold indicate the best (or only) probe for each gene or coterminal transcript group.

9.2. Appendix B – qRT-PCR results of SA8 gene expression analysis. Gene expression was analysed at different times post-infection in the presence and absence of cycloheximide (CHX) and acyclovir (ACV).

A: Relative levels of SA8 gene expression compared to expression at 5hpi.

RL02	**Exp #5**	**Exp #6**	**Exp #1**	**Exp #2 / 4** [a]	**Mean**	**SD**
Mock	Undet. [c]	Undet.	Undet.	Undet.		
2 hr		0.10				
5 hr	1.00	1.00	1.00	1.00	1.00	0.00
5 hr+CHX	5.13	2.26	5.19	2.76	3.83	1.54
Late hr [b]	3.64	1.52	7.46	7.17	4.95	2.87
Late hr+ACV	0.72	0.31		0.80	0.61	0.26
RS01	**Exp #5**	**Exp #6**	**Exp #1**	**Exp #2 / 4**	**Mean**	**SD**
Mock	Undet.	Undet.	Undet.	Undet.		
2 hr		0.00				
5 hr	1.00	1.00	1.00	1.00	1.00	0.00
5 hr+CHX	0.01	0.05	0.10	0.08	0.06	0.04
Late hr	3193.68	8.43	Undet.	11.77	1071.30	1838.04
Late hr+ACV	359.89	0.13		1.97	120.66	207.18
US01	**Exp #5**	**Exp #6**	**Exp #1**	**Exp #2 / 4**	**Mean**	**SD**
Mock	0.18	0.01	Undet.	Undet.		
2 hr		0.01				
5 hr	1.00	1.00	1.00	1.00	1.00	0.00
5 hr+CHX	Undet.	0.06	0.03	0.05	0.05	0.01
Late hr	225.56	3.06	14.10	16.13	64.71	107.38
Late hr+ACV	31.45	0.38		2.67	11.50	17.31
US06	**Exp #5**	**Exp #6**	**Exp #1**	**Exp #2 / 4**	**Mean**	**SD**
Mock	Undet.	0.00	Undet.	Undet.		
2 hr		0.01				
5 hr	1.00	1.00	1.00	1.00	1.00	0.00
5 hr+CHX	0.03	0.00	Undet.	0.00	0.01	0.02
Late hr	126.77	37.54	59.18	5.95	57.36	51.17
Late hr+ACV	0.34	0.08		0.13	0.18	0.14
US12	**Exp #5**	**Exp #6**	**Exp #1**	**Exp #2 / 4**	**Mean**	**SD**
Mock	Undet.	0.00	Undet.	Undet.		
2 hr		0.08				
5 hr	1.00	1.00	1.00	1.00	1.00	0.00
5 hr+CHX	0.04	0.04	0.02	0.03	0.03	0.01
Late hr	3.11	1.42	2.40	4.20	2.78	1.17
Late hr+ACV	0.61	0.35		1.53	0.83	0.62

[a] Data from experiments (Exp) #2 and 4 are presented together since in the former, cells were harvested at an early time post-infection (5hpi), and in the latter, a late time (9hpi); [b] RNA samples were harvested a different "late" time points post-infection. In experiments #4 and 6, cells were harvested at 9hpi. In experiment #5, 27hpi was used. Treatment with ACV was not performed in experiment #1, but cells were harvested at the "late" time of 13hpi; [c] Undet. = value undetermined (below threshold of detection).

Appendix B (continued)

B: Quantitative levels of SA8 gene expression, expressed in gene copies per microgram of total RNA.

RL02	**Exp #5**	**Exp #6**	**Exp #1**	**Exp #2 / 4** [a]	**Mean**	**SD**
Mock	Undet. [c]	Undet.	Undet.	Undet.		
2 hr		137304				
5 hr	172389	888211	239406	300582	400147	329561
5 hr+CHX	430591	1389320	739232	570790	782483	423778
Late hr [b]	847816	1243115	1177410	959859	1057050	184690
Late hr+ACV	203416	338550		196183	246050	80189
RS01	**Exp #5**	**Exp #6**	**Exp #1**	**Exp #2 / 4**	**Mean**	**SD**
Mock	Undet.	Undet.	Undet.	Undet.		
2 hr		414				
5 hr	132932	101960	673450	704863	403301	330568
5 hr+CHX	5255	11694	133761	116421	66783	67750
Late hr	169404	480408	Undet.	3266502	1305438	1705435
Late hr+ACV	26458	20547		860007	302337	482965
US01	**Exp #5**	**Exp #6**	**Exp #1**	**Exp #2 / 4**	**Mean**	**SD**
Mock	11847	13728	Undet.	Undet.		
2 hr		14326				
5 hr	69142	877942	249504	171629	342054	364814
5 hr+CHX	Undet.	72858	15723	13622	34067	33610
Late hr	1740421	2336804	2814777	1334489	2056623	651859
Late hr+ACV	236781	324769		278114	279888	44021
US06	**Exp #5**	**Exp #6**	**Exp #1**	**Exp #2 / 4**	**Mean**	**SD**
Mock	Undet.	47209	Undet.	Undet.		
2 hr		103507				
5 hr	284788	835313	1315399	1284226	929931	482795
5 hr+CHX	55571	66151	Undet.	54865	58863	6322
Late hr	2811220	4192698	8903353	2469727	4594250	2967738
Late hr+ACV	172930	243874		444895	287233	141072
US12	**Exp #5**	**Exp #6**	**Exp #1**	**Exp #2 / 4**	**Mean**	**SD**
Mock	Undet.	89	Undet.	Undet.		
2 hr		264242				
5 hr	820744	4067525	1319006	1377139	1896103	1468994
5 hr+CHX	34889	192594	35565	49859	78227	76557
Late hr	4196218	5982287	3849429	3914538	4485618	1009065
Late hr+ACV	627567	1260304		1413918	1100597	416794

(See footnotes of appendix BA.)

Appendix B (continued)

C: Percentage change in number of SA8 gene copies per microgram of total RNA.

RL02	**Exp #5**	**Exp #6**	**Exp #1**	**Exp #2 / 4** [a]	**Mean**	**SD**
Mock	Undet. [c]	Undet.	Undet.	Undet.		
2 hr		85%				
5 hr	0%	0%	0%	0%	0%	0%
5 hr+CHX	150%	56%	209%	90%	126%	67%
Late hr [b]	222%	40%	392%	219%	218%	144%
Late hr+ACV	-23%	-62%		-35%	-40%	20%
RS01	**Exp #5**	**Exp #6**	**Exp #1**	**Exp #2 / 4**	**Mean**	**SD**
Mock	Undet.	Undet.	Undet.	Undet.		
2 hr		85%				
5 hr	0%	0%	0%	0%	0%	0%
5 hr+CHX	-96%	-89%	-80%	-83%	-87%	7%
Late hr	38654%	371%	Undet.	363%	13130%	22105%
Late hr+ACV	5953%	-80%		22%	1965%	3454%
US01	**Exp #5**	**Exp #6**	**Exp #1**	**Exp #2 / 4**	**Mean**	**SD**
Mock	-83%	-98%	Undet.	Undet.		
2 hr		85%				
5 hr	0%	0%	0%	0%	0%	0%
5 hr+CHX	Undet.	-92%	-94%	-92%	-92%	1%
Late hr	11885%	166%	1028%	678%	3439%	5642%
Late hr+ACV	1531%	-63%		62%	510%	886%
US06	**Exp #5**	**Exp #6**	**Exp #1**	**Exp #2 / 4**	**Mean**	**SD**
Mock	Undet.	-94%	Undet.	Undet.		
2 hr		85%				
5 hr	0%	0%	0%	0%	0%	0%
5 hr+CHX	-80%	-92%	Undet.	-96%	-89%	8%
Late hr	887%	402%	577%	92%	490%	332%
Late hr+ACV	-39%	-71%		-65%	-58%	17%
US12	**Exp #5**	**Exp #6**	**Exp #1**	**Exp #2 / 4**	**Mean**	**SD**
Mock	Undet.	-100%	Undet.	Undet.		
2 hr		85%				
5 hr	0%	0%	0%	0%	0%	0%
5 hr+CHX	-96%	-95%	-97%	-96%	-96%	1%
Late hr	305%	47%	192%	184%	182%	106%
Late hr+ACV	-39%	-69%		3%	-35%	36%

(See footnotes of appendix BA.)

9.3. Appendix C – Additional observations of similar kinetic characteristics among SA8 transcripts following treatment with cycloheximide (CHX).

Experiments #1, 2, and 3(A) provide comparison of signal intensities at 5hpi ± CHX. Data set B from experiment #3 reports the influence of CHX at 8hpi. Cluster characteristics are illustrated in the scatter plot of figures 13, 15, 16, and 17. Signal intensity and cluster membership data for each viral transcript is listed in appendix D.

Treatment with CHX identified a single "immediate-early" gene in SA8, RL02 (see Results, section II.). Additional observations of the RL1 putative transcripts (Groups A and B), and select late transcripts with cluster membership more characteristic of immediate-early genes (Group C) are outlined in Results section 5.4.4.1.

The following groups of SA8 transcripts were also identified as having similar kinetic characteristics following treatment with CHX by their shared or similar cluster membership across multiple experiments. Further analysis is required to assess whether the differences between clusters are statistically significant, but these observations provide preliminary data for further investigations. Table 6 summarizes the cluster membership of groups A through I for experiments #1 through 3.

Group D – UL08-09, UL30, US02

UL08-09 (helicase-primase, origin-binding protein), UL30 (DNA polymerase), US02 (tegument nuclear protein) share cluster membership in experiments #2 and 3A, characterized of relatively low CHX-treated signal intensity and significant inhibition by CHX. In experiment #1 these genes are split into two clusters with stronger patterns of CHX-inhibition, and they are not clustered together in experiment #3B.

Group E – UL15, UL37, UL38, UL43

These four single-ORF transcripts are often paired with (or in an adjacent cluster to) those of Group D. UL15 (capsid-associated protein), UL37 (tegument protein), UL38 (capsid protein VP19C), UL43 (membrane-associated protein) maintain cluster membership in all 5hpi ± CHX experiment (#1, 2, and 3A). They show weak but variable signal (mean 333), and approximately 50% reduction in abundance, in the presence of CHX.

Group F – UL23, UL29, UL36, UL41

UL23 (thymidine kinase), UL29 (ssDNA binding protein), UL36 (tegument protein), UL41 (tegument VHS) are found in the same or adjoining clusters in experiment #3A and 3B, demonstrating low CHX-treated signal intensity but considerable inhibition in the presence of CHX. However, it is interesting to note that in experiments #1 and 2, UL41 is upregulated following treatment with cycloheximide, sharing cluster membership with Group A. Confirmation of the kinetics of this transcript may require analysis by qRT-PCR.

Group G – UL18-20, UL46-47, UL10

UL18-20 (capsid proteins VP23 and VP5, and membrane protein), UL46-47 (tegument proteins VP11/12 and VP13/14), UL10 (glycoprotein M) are clustered together in experiment #1 and 3A, and two of the three transcripts share a cluster in experiments #2 and 3B. However, the cluster mean values for both CHX-treated signal intensity and percentage of signal change between CHX-treated and -untreated samples are quite different between experiments, suggesting that these genes may not have strong similarity in their kinetic behaviour.

Group H – UL03, UL04-05, UL06-07, UL22, UL33-35, UL39-40, UL44-45, UL49-49A, UL52-54, UL55, US01, US08-09

UL03 (membrane associated protein), UL04-05 (structural protein, helicase-primase), UL06-07 (both capsid proteins), UL22 (glycoprotein H), UL33-35 (two capsid/packaging proteins and one membrane associated protein), UL39-40 (small and large subunits of the ribonucleotide reductase), UL44-45 (glycoprotein C and a membrane protein), UL49-49A (tegument protein VP22, and a membrane-associated protein), UL52-54 (helicase-primase protein, glycoprotein K, and IE protein ICP27), UL55 (nuclear protein), US01 (IE protein ICP22), US08-09 (glycoprotein E, and a tegument protein) share a cluster (or pair of neighbouring clusters) in all three experiments conducted at 5hpi ± CHX. The transcript abundance of these genes is characterized by low signal in the presence of CHX, and strong inhibition (76 to 92% reduction) by CHX. Since many of the genes in these transcripts are predicted to have late kinetics, this is not an unexpected pattern of expression.

Group I – UL50 and UL51

The transcripts of Group H often also share cluster membership with UL50 (dUTPase) & UL51 (unknown function). These two transcripts have somewhat lower CHX-treated signal intensity compared to Group H but a similar level of inhibition of expression by CHX. Since the clusters to which Group H and Group I belong are probably not statistically significantly different, these groups could be amalgamated.

9.4. Appendix D –Microarray signal intensity and cluster membership of SA8 transcripts in the presence of cycloheximide (CHX).

A – Mean signal intensity, and gene and transcript cluster membership in experiments #1 and 2.

Gene Name	Probe Score	1 - 5h	1 - 5hCHX	Gene Cluster	Transcript Cluster	2 - 5h	2 - 5hCHX	Gene Cluster	Transcript Cluster
RL1a	5.8	701.91	782.78			735.13	1193.01		
RL1b	36.7	475.62	559.47			417.60	746.14	2	
RL1c	10.0	302.64	537.68	7		86.20	404.07	2	
RL1d	17.5	409.80	304.67	4 & 7(6%)		319.14	338.97	1	
RL1e	15.0	424.44	312.63	4 & 7(6%)		247.93	198.17	1	
RL1anti-a	29.2	474.13	789.48	7		456.39	1127.56	2	
RL1anti-b	20.0	345.95	250.07			188.00	173.46	1	
RL1anti-c	18.3	741.24	425.23	1		659.64	495.38	5	
RL1anti-d	21.7	393.62	170.22	6		274.44	104.93	3	
RL1anti-e	15.0	291.35	587.69	7		84.11	151.53	2	
RL02	**39.2**	746.07	1241.00	8	8	982.42	1803.68	4	4
SA8 LAT putv_CDS1	15.8	738.17	727.07	7		862.26	706.85	6	
SA8 LAT putv_CDS2	47.5	932.34	1881.53	8		1454.40	3176.27	4	
SA8 LAT putv_CDS3	16.7	427.81	352.77	4 & 7(6%)		261.56	281.76	1	
RS01	**26.7**	329.74	282.23	4 & 7(6%)	4 & 7(6%)	147.61	91.60	3	3
UL01	58.3	1774.23	335.33	6		2820.23	263.24	7	
UL02	63.3	2342.15	510.27	1	1	2597.27	572.80	5	5
UL03	**51.7**	1337.13	275.22	6	6	1512.44	269.51	7	7

UL04	46.7	819.63	269.02	6	6	1003.74	234.73	7	7
UL05	45.8	1016.22	264.45	6		1084.85	228.42	5	
UL06	44.2	837.04	198.09	6		904.67	195.62	8	
UL07	53.3	1543.85	301.42	6	6	1529.27	320.49	7	7
UL08	36.7	533.73	427.86	4	4	399.65	134.02	2	
UL09	36.7	446.55	212.63	5		400.19	142.93	3	3
UL10	**56.7**	1458.61	459.38	1	1	2116.39	795.94	6	6
UL11	66.7	2150.62	325.34	6		3346.02	352.81	7	
UL12	87.5	5449.11	861.30	3	3	7089.22	1304.91	4	4
UL13	60.0	1880.97	459.70	1		2493.76	651.29	5	
UL14	36.7	505.79	308.92	5		483.09	155.69	7	
UL15	**36.7**	445.26	243.08	5	5	392.90	158.80	3	3
UL16	41.7	926.12	289.07	6		971.97	335.98	7	
UL17	47.5	1094.14	230.53	6	6	1160.20	214.97	3	
UL18	70.8	2622.74	336.03	1	1	3691.42	143.92	8	8
UL19	65.8	2076.29	263.17	6		2597.77	234.82	7	
UL20	46.7	777.02	620.56	2		760.73	112.32	8	
UL21	**36.7**	477.55	597.09	7	7	507.35	117.61	8	8
UL22	**46.7**	918.80	242.50	6	6	989.78	223.04	7	7
UL23	**53.3**	1520.62	595.40	2	2	1883.28	166.05	8	8
UL24	38.3	530.23	219.75	5		488.28	165.40	3	
UL25	35.8	453.86	348.24	5		339.95	112.37	3	
UL26.5	71.7	2413.46	444.35	1		2948.35	168.47	8	
UL26	72.5	2498.93	293.50	6	6	3510.74	202.14	8	8
UL27	56.7	1716.25	266.39	6		2325.76	257.77	7	
UL28	73.3	3240.81	1106.46	3	3	3684.00	1668.94	4	4
UL29	**67.5**	2413.28	231.20	6	6	3897.95	160.75	8	8
UL30	**37.5**	479.18	416.55	4	4	372.03	158.40	3	3
UL31	53.3	1625.15	583.14	2	2	1765.62	740.50	6	6
UL32	52.5	1349.07	351.27	6		1245.80	291.01	7	

UL33	80.0	3468.64	321.96	6	6	4430.51	337.48	7	7
UL34	52.5	1278.22	193.08	6		1728.27	273.09	7	
UL35	56.7	1502.21	262.76	6		2155.96	292.86	7	
UL36	**53.3**	1301.90	604.48	2	2	2065.91	138.48	8	8
UL37	**35.8**	514.43	226.06	5	5	434.45	214.07	3	3
UL38	**26.7**	702.62	350.62	5	5	896.65	477.96	5	5
UL39	68.3	2092.02	239.43	6		3138.85	300.52	7	
UL40	80.0	3355.80	256.80	6	6	5722.51	287.89	7	7
UL41	**40.0**	509.46	858.33	7	7	486.58	380.92	2	2
UL42	**63.3**	2075.07	673.76	2	2	2929.58	1270.56	6	6
UL43	**34.2**	586.00	298.16	5	5	555.44	325.31	3 & 5(33%)	3 & 5(33%)
UL44	36.7	385.53	273.95	5		307.57	128.85	3	
UL45	48.3	978.74	263.12	6	6	1102.82	248.48	7	7
UL46	65.8	1609.28	417.15	1	1	2478.81	134.41	8	8
UL47	42.5	860.45	253.55	6		669.33	152.85	8	
UL48	**68.3**	2732.85	703.00	2	2	2242.54	628.74	5	5
UL49	75.0	2703.99	285.00	6	6	4074.25	291.63	7	7
UL49A	45.0	981.38	298.30	6		903.25	251.54	7	
UL50	**53.3**	1400.14	291.68	6	6	1898.07	129.00	8	8
UL51	**49.2**	1092.37	251.30	6	6	1218.45	135.44	8	8
UL52	40.8	731.25	234.54	6		925.23	307.83	7	
UL53	52.5	1394.75	285.00	6		1761.47	297.09	7	
UL53A	53.3	1293.73	443.93	1		1833.02	698.21		
UL54	70.8	2375.90	297.69	6	6	3948.22	427.40	8	7
UL55	**40.8**	722.64	248.21	6	6	862.26	245.03	7	7
UL56	**23.3**	771.15	408.75	1	1	1023.92	698.67	6	6
US01	**81.7**	3587.78	283.59	6	6	5258.21	328.25	7	7
US02	**33.3**	433.00	583.07	7	7	315.18	206.01	3	3
US03	65.0	1953.83	219.51	6		2962.39	161.47	8	

US04	73.3	2922.48	1626.98	3	3	4074.34	2078.95	4	4
US05	60.0	1473.12	299.87	6		1850.69	379.05	7	
US06	46.7	731.77	358.79	7		884.75	119.77	8	
US07	59.2	1467.83	231.07	6	6	2054.71	149.99	8	8
US08	67.5	1966.09	236.38	6		2913.92	207.64	8	
US08A	72.5	2781.39	330.73	6		3468.80	388.95	7	
US09	73.3	2351.38	289.60	6	6	3170.55	375.03	7	7
US10	90.0	6778.48	434.65	1	2	8028.38	489.76	5	6
US11	60.0	2191.04	578.49	2		2642.85	919.98	6	
US12	90.0	5696.79	1091.15	3		7582.08	1766.01	4	

Clusters were identified using signal intensity of "A" genes (bold) in the presence of 100 μg/ml (experiment #1) or 50 μg/ml (experiment #2) cycloheximide treatment at 5hpi compared to the difference between treated and untreated samples. Horizontal lines separate genes into coterminal transcript groups. "Transcript" cluster membership is not reported for probes located in regions of the SA8 genome where no ORF has been predicted.

Appendix D (continued)

B –Mean signal intensity, and gene and transcript cluster membership in experiment #3 data sets A and B.

Gene Name	Score	3A - 5h	3A - 5hCHX	Gene Cluster	Transcript Cluster	3B - 8h	3B - 8hCHX	Gene Cluster	Transcript Cluster
RL02	39.2	1706.04	2366.24	2	2	2560.56	4338.44	8	8
RL1anti-a	29.2	1063.09	1744.00	2		2285.18	2531.88	8	
RL1anti-c	18.3	855.55	518.40			1925.91	749.80		
RL1b	36.7	1610.70	1841.57			2774.34	3301.10	8	
RL1d	17.5	647.25	459.01			790.95	458.72	4 & 6(20%)	
RL1e	15.0	304.17	238.33	6		336.11	360.43	4 & 6(20%)	
RS01	25.0	2218.73	717.30	5	5	5676.76	1245.06	2	2
UL01	58.3	6371.61	719.16	1		8817.25	1173.13	3	
UL02	63.3	11151.89	890.66	5	5	13663.72	1074.64	3	3
UL03	51.7	6987.70	526.02	1	1	15385.30	794.52	3	3
UL04	46.7	2588.77	233.33	3	3	2526.79	467.63	6	6
UL05	45.8	2216.88	335.80	8		3080.54	459.75	6	
UL06	44.2	2110.81	253.24	3		2581.52	508.97	6	
UL07	53.3	3488.21	510.02	1	1	5117.89	538.92	5 & 7(2%)	5 & 7(2%)
UL08	36.7	1568.62	252.91	7	8	1514.10	608.38	4 & 6(20%)	
UL09	36.7	891.74	315.01	6		1011.56	410.47	4 & 6(20%)	4 & 6(20%)
UL10	56.7	3751.65	287.82	3	3	10174.20	602.65	5 & 7(2%)	5 & 7(2%)
UL11	66.7	4091.54	348.50	3		6990.40	359.31	5 & 7(2%)	
UL12	87.5	13766.06	1225.29	5	5	31526.46	1898.97	5 & 7(2%)	5 & 7(2%)
UL13	60.0	8235.81	807.04	5		11647.15	1201.47	2	
UL14	36.7	1763.11	253.96	8		3126.94	536.55	5 & 7(2%)	
UL15	36.7	930.85	273.03	8	8	1469.87	471.12	4 & 6(20%)	4 & 6(20%)
UL16	41.7	2864.77	344.13	3		7195.90	420.64	5 & 7(2%)	
UL17	47.5	3011.01	232.13	3	3	8862.15	560.14	5 & 7(2%)	5 & 7(2%)
UL18	70.8	12195.14	361.64	3	3	37034.11	654.32	5 & 7(2%)	

UL19	65.8	9449.64	854.67	5		19898.72	2016.99	8	8
UL20	46.7	3253.28	356.93	3		6162.57	612.75	5 & 7(2%)	
UL21	36.7	1015.48	321.65	6	6	2315.40	457.81	6	6
UL22	46.7	4196.42	280.72	3	3	7041.45	511.44	5 & 7(2%)	5 & 7(2%)
UL23	53.3	8629.64	414.28	1	1	7338.43	582.52	5 & 7(2%)	5 & 7(2%)
UL24	38.3	1629.40	253.71	8		3014.59	541.70	6	
UL25	35.8	693.42	291.81	4 & 8(13%)		2080.81	445.19	6	
UL26.5	71.7	6621.16	433.43	1		17539.68	990.50	3	
UL26	72.5	7821.05	616.34	1	1	32656.92	1390.48	2	2
UL27	56.7	3324.08	269.08	3		5457.55	507.47	5 & 7(2%)	
UL28	73.3	6475.35	847.45	5	5	15532.16	1491.60	2	2
UL29	67.5	5542.88	256.45	3	3	7208.38	267.64	5 & 7(2%)	5 & 7(2%)
UL30	37.5	1128.80	277.43	8	8	1668.20	255.97	7	7
UL31	53.3	3391.71	483.34	1	1	10201.46	956.73	3	3
UL32	52.5	2183.87	385.25	1 & 8(33%)		5210.92	632.84	5 & 7(2%)	
UL33	80.0	8772.07	601.74	1	1	14181.42	922.75	3	3
UL34	52.5	4207.73	397.84	3		5927.88	567.10	5 & 7(2%)	
UL35	56.7	6834.53	675.86	5		16028.04	1121.45	3	
UL36	53.3	2890.61	247.86	3	3	6721.28	423.01	5 & 7(2%)	5 & 7(2%)
UL37	35.8	833.79	297.73	4 & 8(13%)	4 & 8(13%)	2426.96	375.83	6	6
UL38	26.7	1291.09	272.03	8	8	2944.21	447.34	5 & 7(2%)	5 & 7(2%)
UL39	68.3	6000.40	365.26	3		7216.59	638.76	5 & 7(2%)	
UL40	80.0	16309.11	719.08	1	1	17888.67	834.02	3	3
UL41	40.0	1600.05	225.23	3	3	3109.67	406.39	5 & 7(2%)	5 & 7(2%)
UL42	63.3	6126.93	555.40	1	1	13733.12	978.62	3	3
UL43	34.2	2158.15	359.75	1 & 8(33%)	1 & 8(33%)	3492.95	641.78	4 & 6(20%)	4 & 6(20%)
UL44	36.7	1229.10	283.90	8		4307.14	365.71	5 & 7(2%)	
UL45	48.3	2629.61	253.38	3	3	10798.17	324.13	5 & 7(2%)	5 & 7(2%)
UL46	65.8	3671.28	231.03	3	3	11204.68	482.68	5 & 7(2%)	5 & 7(2%)
UL47	42.5	1934.50	297.32	8		7868.35	868.18	3	

UL48	68.3	7820.62	525.15	1	1	24940.03	1137.94	7	7
UL49	75.0	8935.80	1139.92	5	5	38510.67	2760.35	8	8
UL49A	45.0	2771.26	348.25	3		9580.58	699.25	3	
UL50	53.3	3708.54	230.37	3	3	3858.77	440.34	5 & 7(2%)	5 & 7(2%)
UL51	49.2	4098.97	369.41	3	3	10121.62	765.46	3	3
UL52	40.8	2008.19	267.75	3		3862.96	391.56	5 & 7(2%)	
UL53	52.5	3007.47	272.74	3		3778.18	463.51	5 & 7(2%)	
UL54	70.8	10452.06	510.14	1	1	11912.04	753.86	3	3
UL55	40.8	2084.62	301.21	8	8	2680.83	513.80	6	6
UL56	23.3	1428.69	362.36	8	8	1796.15	625.51	4 & 6(20%)	4 & 6(20%)
US01	81.7	12782.95	573.83	1	1	25084.05	1047.26	3	3
US02	33.3	684.80	165.32	8	8	1253.09	415.14	6	6
US03	65.0	6884.03	389.23	1		18493.95	858.75	3	
US04	73.3	12491.80	2409.18	5	5	24537.72	3984.51	8	8
US05	60.0	3879.59	361.27	3		13236.84	913.14	3	
US06	46.7	2776.86	286.17	3		13482.11	424.79	5 & 7(2%)	
US07	59.2	5330.53	462.84	1	1	13174.31	997.17	3	3
US08	67.5	4568.75	573.65	1		10591.83	1211.15	2	
US08A	70.0	7881.58	1793.01	5		15636.94	2611.71	8	
US09	73.3	10420.86	811.09	5	5	30658.72	1704.19	8	8
US10	90.0	27300.82	1014.31	5	5	48976.09	2358.40	8	8
US11	60.0	8520.02	952.05	5		17239.24	1461.28	2	
US12	90.0	15657.67	3424.34	5		31833.24	7257.67	8	

Clusters were identified using signal intensity of "A" genes (bold) in the presence of 100 μg/ml cycloheximide treatment at 5 or 8hpi compared to the difference between treated and untreated samples. Horizontal lines separate genes into coterminal transcript groups. "Transcript" cluster membership is not reported for probes located in regions of the SA8 genome where no ORF has been predicted.

9.5. Appendix E – Mean microarray signal intensity and cluster membership of SA8 transcripts in the presence of acyclovir (experiment #4).

Gene Name	Score	9h	9hACV	Gene	Transcript
RL1a	5.8	1993.51	1714.90	1	
RL1b	36.7	1209.48	337.57	6	
RL1c	10.0	273.40	67.42	6	
RL1d	17.5	948.63	392.60	2	
RL1e	15.0	179.40	153.78	1	
RL1anti-a	29.2	1371.56	381.31		
RL1anti-b	20.0	706.19	260.71	2	
RL1anti-c	18.3	3311.83	1599.18	2	
RL1anti-d	21.7	646.78	314.15	2	
RL1anti-e	15.0	535.47	93.21	5	
RL02	**39.2**	2250.84	688.18	7 & 2(2%)	7 & 2(2%)
SA8 LAT putv_CDS1	15.8	1656.40	589.79	7 & 2(2%)	
SA8 LAT putv_CDS2	47.5	5725.73	1226.53	3	
SA8 LAT putv_CDS3	16.7	2094.65	612.70	7 & 2(2%)	
RS01	**26.7**	430.76	116.55	6	6
UL01	58.3	8551.63	3831.21	4	
UL02	63.3	18566.57	5676.96	4	4
UL03	**51.7**	5834.07	1480.50	3	3
UL04	46.7	5732.82	1505.56	3	3
UL05	45.8	6461.21	1763.49	3	
UL06	44.2	4621.98	1192.87	3	
UL07	53.3	8810.93	2163.67	3	3
UL08	36.7	2006.73	566.98	7 & 2(2%)	6
UL09	36.7	1184.70	326.05	6	
UL10	**56.7**	10834.32	1717.57	5	5
UL11	66.7	12701.31	4277.05	4	
UL12	87.5	38677.11	13272.87	8	8
UL13	60.0	12956.07	4218.89	4	
UL14	36.7	3065.44	772.85	6	
UL15	**36.7**	1529.35	471.46	7 & 2(2%)	7 & 2(2%)
UL16	41.7	5578.80	902.06	5	
UL17	47.5	7208.27	924.11	5	5
UL18	70.8	20522.01	5185.71	4	4
UL19	65.8	18136.03	2806.54	5	
UL20	46.7	6578.33	1107.62	5	
UL21	**36.7**	3068.68	734.64	6	6
UL22	**46.7**	5611.14	1334.44	3	3

UL23	**53.3**	4886.05	2064.62	2	2
UL24	38.3	4367.45	702.43	5	
UL25	35.8	1617.62	191.63	5	
UL26.5	71.7	23005.33	4393.35	4	
UL26	72.5	25016.79	5089.98	4	4
UL27	56.7	11527.71	1644.51	5	
UL28	73.3	34117.83	4815.00	4	4
UL29	**67.5**	13726.08	3591.08	4	4
UL30	**37.5**	2467.63	730.41	7 & 2(2%)	7 & 2(2%)
UL31	53.3	14229.97	2598.89	5	5
UL32	52.5	11167.38	2451.04	3	
UL33	80.0	27075.59	6455.58	4	4
UL34	52.5	7337.25	1781.87	3	
UL35	56.7	12152.82	2381.78	3	
UL36	**53.3**	7186.63	1523.31	3	3
UL37	**35.8**	2326.97	322.13	5	5
UL38	**26.7**	3666.41	1124.22	7 & 2(2%)	7 & 2(2%)
UL39	68.3	16345.04	3966.44	4	
UL40	80.0	22929.26	8648.97	8	8
UL41	**40.0**	4732.81	1195.96	3	3
UL42	**63.3**	16694.38	5431.59	4	4
UL43	**34.2**	3849.03	684.48	5	5
UL44	36.7	1925.10	250.81	5	
UL45	48.3	8314.36	2055.10	3	3
UL46	65.8	24767.07	3007.07	5	5
UL47	42.5	5579.47	537.94	5	
UL48	**68.3**	23757.65	5256.25	4	4
UL49	75.0	28225.21	5767.25	4	4
UL49A	45.0	12348.06	1506.83	5	
UL50	**53.3**	6611.44	2129.85	2 & 3(44%) & 7(11%)	2 & 3(44%) & 7(11%)
UL51	**49.2**	9276.22	2500.37	3	3
UL52	40.8	4114.87	951.63	6	
UL53	52.5	8173.72	2240.17	5	
UL53A	53.3	9000.41	2395.14	3	
UL54	70.8	20551.04	4517.46	4	4
UL55	**40.8**	3579.62	1003.21	6	6
UL56	**23.3**	2982.78	1345.18	2	2
US01	**81.7**	24770.63	7997.24	8	8
US02	**33.3**	1691.72	371.11	6	6
US03	65.0	15188.42	2191.66	5	
US04	73.3	29354.08	7285.09	8	8
US05	60.0	16164.75	2714.83	5	
US06	46.7	6182.91	884.90	5	
US07	59.2	14660.73	2271.56	5	5

US08	67.5	15108.47	4248.17	4	
US08A	72.5	22930.66	6912.75	4	
US09	73.3	25181.17	7713.28	8	8
US10	90.0	67444.66	18042.03	8	8
US11	60.0	8794.64	3135.45	2 & 3(44%) & 7(11%)	
US12	90.0	46832.61	21147.63	8	

Clusters were identified using signal intensity of "A" genes (bold) in the presence of 10 μg/ml acyclovir (ACV) treatment at 9hpi compared to the difference between treated and untreated samples. Horizontal lines separate genes into coterminal transcript groups. "Transcript" cluster membership is not reported for probes located in regions of the SA8 genome where no ORF has been predicted.

REFERENCES

1. **Advani, S. J., R. Hagglund, R. R. Weichselbaum, and B. Roizman.** 2001. Posttranslational processing of infected cell proteins 0 and 4 of herpes simplex virus 1 is sequential and reflects the subcellular compartment in which the proteins localize. J. Virol. **75:**7904-7912.

2. **Advani, S. J., R. R. Weichselbaum, and B. Roizman.** 2003. Herpes simplex virus 1 activates cdc2 to recruit topoisomerase II alpha for post-DNA synthesis expression of late genes. Proc. Natl. Acad. Sci. U. S. A. **100:**4825-4830.

3. **Advani, S. J., R. R. Weichselbaum, and B. Roizman.** 2000. The role of cdc2 in the expression of herpes simplex virus genes. Proc. Natl. Acad. Sci. U. S. A. **97:**10996-11001.

4. **Aguilar, J. S., G. V. Devi-Rao, M. K. Rice, J. Sunabe, P. Ghazal, and E. K. Wagner.** 2006. Quantitative comparison of the HSV-1 and HSV-2 transcriptomes using DNA microarray analysis. Virology **348:**233-241.

5. **Ahn, J. W., K. L. Powell, P. Kellam, and D. G. Alber.** 2002. Gammaherpesvirus lytic gene expression as characterized by DNA array. Journal of Virology **76:**6244-6256.

6. **Andrade, M. R., J. Yee, P. Barry, A. Spinner, J. A. Roberts, P. H. Cabello, J. P. Leite, and N. W. Lerche.** 2003. Prevalence of antibodies to selected viruses in a long-term closed breeding colony of rhesus macaques (Macaca mulatta) in Brazil. Am. J. Primatol. **59:**123-128.

7. **apRhys, C. M., D. M. Ciufo, E. A. O'Neill, T. J. Kelly, and G. S. Hayward.** 1989. Overlapping octamer and TAATGARAT motifs in the VF65-response elements in herpes simplex virus immediate-early promoters represent independent binding sites for cellular nuclear factor III. J. Virol. **63:**2798-2812.

8. **Babb, R., C. C. Huang, D. J. Aufiero, and W. Herr.** 2001. DNA recognition by the herpes simplex virus transactivator VP16: a novel DNA-binding structure. Mol. Cell. Biol. **21:**4700-4712.

9. **Bailey, A. C. and R. Thompson.** 1992. A sequence-specific DNA-binding protein recognising a GA-rich element cooperates with Oct-1 at the herpes simplex virus type 1 IE3 promoter. Intervirology **34:**74-85.

10. **Beier, V., A. Bauer, M. Baum, and J. D. Hoheisel.** 2004. Fluorescent sample labeling for DNA microarray analyses. Methods Mol. Biol. **283:**127-135.

11. **Biron, C. A. and G. C. Sen.** 2001. Chapter 12: Interferons and other cytokines. In D. M. Knipe and P. M. Howley (ed.), Fields Virology, vol. 1. Lippincott, Williams & Wilkins, USA.

12. **Blaho, J. A., N. Michael, V. Kang, N. Aboul-Ela, M. E. Smulson, M. K. Jacobson, and B. Roizman.** 1992. Differences in the poly(ADP-ribosyl)ation patterns of ICP4, the herpes simplex virus major regulatory protein, in infected cells and in isolated nuclei. J. Virol. **66:**6398-6407.

13. **Bolovan, C. A., N. M. Sawtell, and R. L. Thompson.** 1994. ICP34.5 mutants of herpes simplex virus type 1 strain 17syn+ are attenuated for neurovirulence in mice and for replication in confluent primary mouse embryo cell cultures. J. Virol. **68:**48-55.

14. **Borchers, K. and M. Ozel.** 1993. Simian agent 8 (SA8): morphogenesis and ultrastructure. Zentralbl. Bakteriol. **279:**526-536.

15. **Bresnahan, W. A. and T. Shenk.** 2000. A subset of viral transcripts packaged within human cytomegalovirus particles. Science **288:**2373-2376.

16. **Brown, S. M., A. R. MacLean, E. A. McKie, and J. Harland.** 1997. The herpes simplex virus virulence factor ICP34.5 and the cellular protein MyD116 complex with proliferating cell nuclear antigen through the 63-amino-acid domain conserved in ICP34.5, MyD116, and GADD34. J. Virol. **71:**9442-9449.

17. **Bryant, P. A., D. Venter, R. Robins-Browne, and N. Curtis.** 2004. Chips with everything: DNA microarrays in infectious diseases. Lancet Infect. Dis. **4:**100-111.

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

19. **Carrier, F., P. T. Georgel, P. Pourquier, M. Blake, H. U. Kontny, M. J. Antinore, M. Gariboldi, T. G. Myers, J. N. Weinstein, Y. Pommier, and A. J. Fornace Jr.** 1999. Gadd45, a p53-responsive stress protein, modifies DNA accessibility on damaged chromatin. Mol. Cell. Biol. **19:**1673-1685.

20. **Carrozza, M. J. and N. A. DeLuca.** 1996. Interaction of the viral activator protein ICP4 with TFIID through TAF250. Mol. Cell. Biol. **16:**3085-3093.

21. **Cassady, K. A., M. Gross, and B. Roizman.** 1998. The herpes simplex virus US11 protein effectively compensates for the gamma1(34.5) gene if present before activation of protein kinase R by precluding its phosphorylation and that of the alpha subunit of eukaryotic translation initiation factor 2. J. Virol. **72:**8620-8626.

22. **Cassady, K. A., M. Gross, and B. Roizman.** 1998. The second-site mutation in the herpes simplex virus recombinants lacking the gamma134.5 genes precludes shutoff of protein synthesis by blocking the phosphorylation of eIF-2alpha. J. Virol. **72:**7005-7011.

23. **Chambers, J., A. Angulo, D. Amaratunga, H. Guo, Y. Jiang, J. S. Wan, A. Bittner, K. Frueh, M. R. Jackson, P. A. Peterson, M. G. Erlander, and P. Ghazal.** 1999. DNA microarrays of the complex human cytomegalovirus genome: profiling kinetic class with drug sensitivity of viral gene expression. J. Virol. **73:**5757-5766.

24. **Chee, A. V., P. Lopez, P. P. Pandolfi, and B. Roizman.** 2003. Promyelocytic leukemia protein mediates interferon-based anti-herpes simplex virus 1 effects. J. Virol. **77:**7101-7105.

25. **Cheng, G., Z. Feng, and B. He.** 2005. Herpes simplex virus 1 infection activates the endoplasmic reticulum resident kinase PERK and mediates eIF-2alpha dephosphorylation by the gamma(1)34.5 protein. J. Virol. **79:**1379-1388.

26. **Cheung, A. K.** 1989. DNA nucleotide sequence analysis of the immediate-early gene of pseudorabies virus. Nucleic Acids Res. **17:**4637-4646.

27. **Chipman, H., T. J. Hastie, and R. Tibshirani.** 2003. Clustering microarray data, p. 159. In T. Speed (ed.), Statisitcal analysis of gene expression microarray data, vol. 1. Chapman & Hall / CRC Press, USA.

28. **Chou, J., J. J. Chen, M. Gross, and B. Roizman.** 1995. Association of a M(r) 90,000 phosphoprotein with protein kinase PKR in cells exhibiting enhanced phosphorylation of translation initiation factor eIF-2 alpha and premature shutoff of protein synthesis after infection with gamma 134.5- mutants of herpes simplex virus 1. Proc. Natl. Acad. Sci. U. S. A. **92:**10516-10520.

29. **Chou, J., E. R. Kern, R. J. Whitley, and B. Roizman.** 1990. Mapping of herpes simplex virus-1 neurovirulence to gamma 134.5, a gene nonessential for growth in culture. Science **250:**1262-1266.

30. **Chou, J. and B. Roizman.** 1994. Herpes simplex virus 1 gamma(1)34.5 gene function, which blocks the host response to infection, maps in the homologous domain of the genes expressed during growth arrest and DNA damage. Proc. Natl. Acad. Sci. U. S. A. **91:**5247-5251.

31. **Chou, J. and B. Roizman.** 1992. The gamma 1(34.5) gene of herpes simplex virus 1 precludes neuroblastoma cells from triggering total shutoff of protein synthesis characteristic of programed cell death in neuronal cells. Proc. Natl. Acad. Sci. U. S. A. **89:**3266-3270.

32. **Churchill, G. A.** 2004. Using ANOVA to analyze microarray data. BioTechniques **37:**173-5, 177.

33. **Cook, W. J., B. Gu, N. A. DeLuca, E. B. Moynihan, and D. M. Coen.** 1995. Induction of transcription by a viral regulatory protein depends on the relative strengths of functional TATA boxes. Mol. Cell. Biol. **15:**4998-5006.

34. **Cui, X. and G. A. Churchill.** 2003. Statistical tests for differential expression in cDNA microarray experiments. Genome Biol. **4:**210.

35. **Cui, X., M. K. Kerr, and G. A. Churchill.** 2003. Transformations for cDNA microarray data. Stat. Appl. Genet. Mol. Biol. **2:**Article4.

36. **de Haro, C., R. Mendez, and J. Santoyo.** 1996. The eIF-2alpha kinases and the control of protein synthesis. FASEB J. **10:**1378-1387.

37. **Desai, P., R. Ramakrishnan, Z. W. Lin, B. Osak, J. C. Glorioso, and M. Levine.** 1993. The RR1 gene of herpes simplex virus type 1 is uniquely trans activated by ICP0 during infection. J. Virol. **67:**6125-6135.

38. **Douville, P., M. Hagmann, O. Georgiev, and W. Schaffner.** 1995. Positive and negative regulation at the herpes simplex virus ICP4 and ICP0 TAATGARAT motifs. Virology **207:**107-116.

39. **Dudoit, S. and J. Fridlyand.** 2003. Classification in microarray experiments, p. 93. In T. Speed (ed.), Statistical analysis of gene expression microarray data, vol. 1. Chapman & Hall / CRC Press, USA.

40. **Durand, L. O., S. J. Advani, A. P. Poon, and B. Roizman.** 2005. The carboxyl-terminal domain of RNA polymerase II is phosphorylated by a complex containing cdk9 and infected-cell protein 22 of herpes simplex virus 1. J. Virol. **79:**6757-6762.

41. **Eberle, R., D. H. Black, E. L. Blewett, and G. L. White.** 1997. Prevalence of Herpesvirus papio 2 in baboons and identification of immunogenic viral polypeptides. Lab. Anim. Sci. **47:**256-262.

42. **Eberle, R., D. H. Black, T. W. Lehenbauer, and G. L. White.** 1998. Shedding and transmission of baboon Herpesvirus papio 2 (HVP2) in a breeding colony. Lab. Anim. Sci. **48:**23-28.

43. **Eberle, R., D. H. Black, S. Lipper, and J. K. Hilliard.** 1995. Herpesvirus papio 2, an SA8-like alpha-herpesvirus of baboons. Arch. Virol. **140:**529-545.

44. **Eberle, R., M. Zhang, and D. H. Black.** 1993. Gene mapping and sequence analysis of the unique short region of the simian herpesvirus SA 8 genome. Arch. Virol. **130:**391-411.

45. **Elion, G. B., P. A. Furman, J. A. Fyfe, P. de Miranda, L. Beauchamp, and H. J. Schaeffer.** 1977. Selectivity of action of an antiherpetic agent, 9-(2-hydroxyethoxymethyl) guanine. Proc. Natl. Acad. Sci. U. S. A. **74:**5716-5720.

46. **Everett, R. D.** 2000. ICP0 induces the accumulation of colocalizing conjugated ubiquitin. J. Virol. **74:**9994-10005.

47. **Everett, R. D.** 1991. Construction and characterization of herpes simplex type 1 viruses without introns in immediate early gene 1. J. Gen. Virol. **72 (Pt 3):**651-659.

48. **Everett, R. D.** 1989. Construction and characterization of herpes simplex virus type 1 mutants with defined lesions in immediate early gene 1. J. Gen. Virol. **70 (Pt 5):**1185-1202.

49. **Everett, R. D., G. Sourvinos, C. Leiper, J. B. Clements, and A. Orr.** 2004. Formation of nuclear foci of the herpes simplex virus type 1 regulatory protein ICP4 at early times of infection: localization, dynamics, recruitment of ICP27, and evidence for the de novo induction of ND10-like complexes. J. Virol. **78:**1903-1917.

50. **Forster, T., D. Roy, and P. Ghazal.** 2003. Experiments using microarray technology: limitations and standard operating procedures. J. Endocrinol. **178:**195-204.

51. **Gaffney, D. F., J. McLauchlan, J. L. Whitton, and J. B. Clements.** 1985. A modular system for the assay of transcription regulatory signals: the sequence TAATGARAT is required for herpes simplex virus immediate early gene activation. Nucleic Acids Res. **13:**7847-7863.

52. **Gale, M.,Jr and M. G. Katze.** 1998. Molecular mechanisms of interferon resistance mediated by viral-directed inhibition of PKR, the interferon-induced protein kinase. Pharmacol. Ther. **78:**29-46.

53. **Gelman, I. H. and S. Silverstein.** 1987. Herpes simplex virus immediate-early promoters are responsive to virus and cell trans-acting factors. J. Virol. **61:**2286-2296.

54. **Godowski, P. J. and D. M. Knipe.** 1986. Transcriptional control of herpesvirus gene expression: gene functions required for positive and negative regulation. Proc. Natl. Acad. Sci. U. S. A. **83:**256-260.

55. **Gu, W., Q. Huang, and G. S. Hayward.** 1995. Multiple Tandemly Repeated Binding Sites for the YY1 Repressor and Transcription Factors AP-1 and SP-1 Are Clustered within Intron-1 of the Gene Encoding the IE110 Transactivator of Herpes simplex Virus Type 1. J. Biomed. Sci. **2:**203-226.

56. **Hagglund, R. and B. Roizman.** 2004. Role of ICP0 in the strategy of conquest of the host cell by herpes simplex virus 1. J. Virol. **78:**2169-2178.

57. **Hardy, W. R. and R. M. Sandri-Goldin.** 1994. Herpes simplex virus inhibits host cell splicing, and regulatory protein ICP27 is required for this effect. J. Virol. **68:**7790-7799.

58. **Harris-Hamilton, E. and S. L. Bachenheimer.** 1985. Accumulation of herpes simplex virus type 1 RNAs of different kinetic classes in the cytoplasm of infected cells. J. Virol. **53:**144-151.

59. **He, B., J. Chou, R. Brandimarti, I. Mohr, Y. Gluzman, and B. Roizman.** 1997. Suppression of the phenotype of gamma(1)34.5- herpes simplex virus 1: failure of activated RNA-dependent protein kinase to shut off protein synthesis is associated with a deletion in the domain of the alpha47 gene. J. Virol. **71:**6049-6054.

60. **He, B., J. Chou, D. A. Liebermann, B. Hoffman, and B. Roizman.** 1996. The carboxyl terminus of the murine MyD116 gene substitutes for the corresponding domain

of the gamma(1)34.5 gene of herpes simplex virus to preclude the premature shutoff of total protein synthesis in infected human cells. J. Virol. **70:**84-90.

61. **He, B., M. Gross, and B. Roizman.** 1998. The gamma134.5 protein of herpes simplex virus 1 has the structural and functional attributes of a protein phosphatase 1 regulatory subunit and is present in a high molecular weight complex with the enzyme in infected cells. J. Biol. Chem. **273:**20737-20743.

62. **He, B., M. Gross, and B. Roizman.** 1997. The gamma(1)34.5 protein of herpes simplex virus 1 complexes with protein phosphatase 1alpha to dephosphorylate the alpha subunit of the eukaryotic translation initiation factor 2 and preclude the shutoff of protein synthesis by double-stranded RNA-activated protein kinase. Proc. Natl. Acad. Sci. U. S. A. **94:**843-848.

63. **Herendeen, D. R., G. A. Kassavetis, J. Barry, B. M. Alberts, and E. P. Geiduschek.** 1989. Enhancement of bacteriophage T4 late transcription by components of the T4 DNA replication apparatus. Science **245:**952-958.

64. **Herrera, F. J. and S. J. Triezenberg.** 2004. VP16-dependent association of chromatin-modifying coactivators and underrepresentation of histones at immediate-early gene promoters during herpes simplex virus infection. J. Virol. **78:**9689-9696.

65. **Hill, A., P. Jugovic, I. York, G. Russ, J. Bennink, J. Yewdell, H. Ploegh, and D. Johnson.** 1995. Herpes simplex virus turns off the TAP to evade host immunity. Nature **375:**411-415.

66. **Hilliard, J. K., D. Black, and R. Eberle.** 1989. Simian alphaherpesviruses and their relation to the human herpes simplex viruses. Arch. Virol. **109:**83-102.

67. **Hilliard, J. K., R. Eberle, S. L. Lipper, R. M. Munoz, and S. A. Weiss.** 1987. Herpesvirus simiae (B virus): replication of the virus and identification of viral polypeptides in infected cells. Arch. Virol. **93:**185-198.

68. **Hobbs, W. E. and N. A. DeLuca.** 1999. Perturbation of cell cycle progression and cellular gene expression as a function of herpes simplex virus ICP0. J. Virol. **73:**8245-8255.

69. **Hoheisel, J. D.** 2006. Microarray technology: beyond transcript profiling and genotype analysis. Nat. Rev. Genet. **7:**200-210.

70. **Huang, C. C. and W. Herr.** 1996. Differential control of transcription by homologous homeodomain coregulators. Mol. Cell. Biol. **16:**2967-2976.

71. **Ihara, S., L. Feldman, S. Watanabe, and T. Ben-Porat.** 1983. Characterization of the immediate-early functions of pseudorabies virus. Virology **131:**437-454.

72. **Imbalzano, A. N., D. M. Coen, and N. A. DeLuca.** 1991. Herpes simplex virus transactivator ICP4 operationally substitutes for the cellular transcription factor Sp1 for efficient expression of the viral thymidine kinase gene. J. Virol. **65:**565-574.

73. **Johnson, P. A., C. MacLean, H. S. Marsden, R. G. Dalziel, and R. D. Everett.** 1986. The product of gene US11 of herpes simplex virus type 1 is expressed as a true late gene. J. Gen. Virol. **67 (Pt 5):**871-883.

74. **Johnson, P. A., A. Miyanohara, F. Levine, T. Cahill, and T. Friedmann.** 1992. Cytotoxicity of a replication-defective mutant of herpes simplex virus type 1. J. Virol. **66:**2952-2965.

75. **Jones, K. A. and R. Tjian.** 1985. Sp1 binds to promoter sequences and activates herpes simplex virus 'immediate-early' gene transcription in vitro. Nature **317:**179-182.

76. **Jones, P. C., G. S. Hayward, and B. Roizman.** 1977. Anatomy of herpes simplex virus DNA VII. alpha-RNA is homologous to noncontiguous sites in both the L and S components of viral DNA. J. Virol. **21:**268-276.

77. **Jones, P. C. and B. Roizman.** 1979. Regulation of herpesvirus macromolecular synthesis. VIII. The transcription program consists of three phases during which both extent of transcription and accumulation of RNA in the cytoplasm are regulated. J. Virol. **31:**299-314.

78. **Karaca, G., D. Hargett, T. I. McLean, J. S. Aguilar, P. Ghazal, E. K. Wagner, and S. L. Bachenheimer.** 2004. Inhibition of the stress-activated kinase, p38, does not affect the virus transcriptional program of herpes simplex virus type 1. Virology **329:**142-156.

79. **Katan, M., A. Haigh, C. P. Verrijzer, P. C. van der Vliet, and P. O'Hare.** 1990. Characterization of a cellular factor which interacts functionally with Oct-1 in the assembly of a multicomponent transcription complex. Nucleic Acids Res. **18:**6871-6880.

80. **Kawaguchi, Y., R. Bruni, and B. Roizman.** 1997. Interaction of herpes simplex virus 1 alpha regulatory protein ICP0 with elongation factor 1delta: ICP0 affects translational machinery. J. Virol. **71:**1019-1024.

81. **Kennedy, P. G., E. Grinfeld, M. Craigon, K. Vierlinger, D. Roy, T. Forster, and P. Ghazal.** 2005. Transcriptomal analysis of varicella-zoster virus infection using long oligonucleotide-based microarrays. J. Gen. Virol. **86:**2673-2684.

82. **Kent, J. R. and N. W. Fraser.** 2005. The cellular response to herpes simplex virus type 1 (HSV-1) during latency and reactivation. J. Neurovirol. **11:**376-383.

83. **Kerr, M. K.** 2003. Experimental design to make the most of microarray studies. Methods Mol. Biol. **224:**137-147.

84. **Kerr, M. K., M. Martin, and G. A. Churchill.** 2000. Analysis of variance for gene expression microarray data. J. Comput. Biol. **7**:819-837.

85. **Khadijah, S., S. Y. Neo, M. S. Hossain, L. D. Miller, S. Mathavan, and J. Kwang.** 2003. Identification of white spot syndrome virus latency-related genes in specific-pathogen-free shrimps by use of a microarray. J. Virol. **77**:10162-10167.

86. **Khoo, D., C. Perez, and I. Mohr.** 2002. Characterization of RNA determinants recognized by the arginine- and proline-rich region of Us11, a herpes simplex virus type 1-encoded double-stranded RNA binding protein that prevents PKR activation. J. Virol. **76**:11971-11981.

87. **Kim, H., B. Zhao, E. C. Snesrud, B. J. Haas, C. D. Town, and J. Quackenbush.** 2002. Use of RNA and genomic DNA references for inferred comparisons in DNA microarray analyses. BioTechniques **33**:924-930.

88. **Klemm, R. D., J. A. Goodrich, S. Zhou, and R. Tjian.** 1995. Molecular cloning and expression of the 32-kDa subunit of human TFIID reveals interactions with VP16 and TFIIB that mediate transcriptional activation. Proc. Natl. Acad. Sci. U. S. A. **92**:5788-5792.

89. **Kristie, T. M., J. H. LeBowitz, and P. A. Sharp.** 1989. The octamer-binding proteins form multi-protein--DNA complexes with the HSV alpha TIF regulatory protein. EMBO J. **8**:4229-4238.

90. **Kristie, T. M. and P. A. Sharp.** 1993. Purification of the cellular C1 factor required for the stable recognition of the Oct-1 homeodomain by the herpes simplex virus alpha-trans-induction factor (VP16). J. Biol. Chem. **268**:6525-6534.

91. **Lagunoff, M. and B. Roizman.** 1995. The regulation of synthesis and properties of the protein product of open reading frame P of the herpes simplex virus 1 genome. J. Virol. **69**:3615-3623.

92. **Lagunoff, M. and B. Roizman.** 1994. Expression of a herpes simplex virus 1 open reading frame antisense to the gamma(1)34.5 gene and transcribed by an RNA 3' coterminal with the unspliced latency-associated transcript. J. Virol. **68**:6021-6028.

93. **LaMarco, K. L. and S. L. McKnight.** 1989. Purification of a set of cellular polypeptides that bind to the purine-rich cis-regulatory element of herpes simplex virus immediate early genes. Genes Dev. **3**:1372-1383.

94. **Leib, D. A., T. E. Harrison, K. M. Laslo, M. A. Machalek, N. J. Moorman, and H. W. Virgin.** 1999. Interferons regulate the phenotype of wild-type and mutant herpes simplex viruses in vivo. J. Exp. Med. **189**:663-672.

95. **Leib, D. A., M. A. Machalek, B. R. Williams, R. H. Silverman, and H. W. Virgin.** 2000. Specific phenotypic restoration of an attenuated virus by knockout of a host resistance gene. Proc. Natl. Acad. Sci. U. S. A. **97**:6097-6101.

96. **Levin, J. L., J. K. Hilliard, S. L. Lipper, T. M. Butler, and W. J. Goodwin.** 1988. A naturally occurring epizootic of simian agent 8 in the baboon. Lab. Anim. Sci. **38:**394-397.

97. **Lieu, P. T. and E. K. Wagner.** 2000. The kinetics of VP5 mRNA expression is not critical for viral replication in cultured cells. J. Virol. **74:**2770-2776.

98. **Liu, H., S. Tarima, A. S. Borders, T. V. Getchell, M. L. Getchell, and A. J. Stromberg.** 2005. Quadratic regression analysis for gene discovery and pattern recognition for non-cyclic short time-course microarray experiments. BMC Bioinformatics **6:**106.

99. **Lua, D. T., M. Yasuike, I. Hirono, and T. Aoki.** 2005. Transcription program of red sea bream iridovirus as revealed by DNA microarrays. J. Virol. **79:**15151-15164.

100. **Luebcke, E., E. Dubovi, D. Black, K. Ohsawa, and R. Eberle.** 2006. Isolation and characterization of a chimpanzee alphaherpesvirus. J. Gen. Virol. **87:**11-19.

101. **Luke, K., A. Radek, X. Liu, J. Campbell, M. Uzan, R. Haselkorn, and Y. Kogan.** 2002. Microarray analysis of gene expression during bacteriophage T4 infection. Virology **299:**182-191.

102. **Mackem, S. and B. Roizman.** 1982. Structural features of the herpes simplex virus alpha gene 4, 0, and 27 promoter-regulatory sequences which confer alpha regulation on chimeric thymidine kinase genes. J. Virol. **44:**939-949.

103. **Mackem, S. and B. Roizman.** 1980. Regulation of herpesvirus macromolecular synthesis: transcription-initiation sites and domains of alpha genes. Proc. Natl. Acad. Sci. U. S. A. **77:**7122-7126.

104. **Martino, M. A., G. B. Hubbard, T. M. Butler, and J. K. Hilliard.** 1998. Clinical disease associated with simian agent 8 infection in the baboon. Lab. Anim. Sci. **48:**18-22.

105. **McGeoch, D. J., C. Cunningham, G. McIntyre, and A. Dolan.** 1991. Comparative sequence analysis of the long repeat regions and adjoining parts of the long unique regions in the genomes of herpes simplex viruses types 1 and 2. J. Gen. Virol. **72 (Pt 12):**3057-3075.

106. **McGeoch, D. J., M. A. Dalrymple, A. J. Davison, A. Dolan, M. C. Frame, D. McNab, L. J. Perry, J. E. Scott, and P. Taylor.** 1988. The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. J. Gen. Virol. **69 (Pt 7):**1531-1574.

107. **McNamee, E. E., T. J. Taylor, and D. M. Knipe.** 2000. A dominant-negative herpesvirus protein inhibits intranuclear targeting of viral proteins: effects on DNA replication and late gene expression. J. Virol. **74:**10122-10131.

108. **Memedula, S. and A. S. Belmont.** 2003. Sequential recruitment of HAT and SWI/SNF components to condensed chromatin by VP16. Curr. Biol. **13:**241-246.

109. **Minaker, R. L., K. L. Mossman, and J. R. Smiley.** 2005. Functional inaccessibility of quiescent herpes simplex virus genomes. Virol. J. **2:**85.

110. **Mohr, I. and Y. Gluzman.** 1996. A herpesvirus genetic element which affects translation in the absence of the viral GADD34 function. EMBO J. **15:**4759-4766.

111. **Mossman, K. L., P. F. Macgregor, J. J. Rozmus, A. B. Goryachev, A. M. Edwards, and J. R. Smiley.** 2001. Herpes simplex virus triggers and then disarms a host antiviral response. J. Virol. **75:**750-758.

112. **Mossman, K. L. and J. R. Smiley.** 2002. Herpes simplex virus ICP0 and ICP34.5 counteract distinct interferon-induced barriers to virus replication. J. Virol. **76:**1995-1998.

113. **Muller, M. T.** 1987. Binding of the herpes simplex virus immediate-early gene product ICP4 to its own transcription start site. J. Virol. **61:**858-865.

114. **Mulvey, M., J. Poppers, A. Ladd, and I. Mohr.** 1999. A herpesvirus ribosome-associated, RNA-binding protein confers a growth advantage upon mutants deficient in a GADD34-related function. J. Virol. **73:**3375-3385.

115. **Mulvey, M., J. Poppers, D. Sternberg, and I. Mohr.** 2003. Regulation of eIF2alpha phosphorylation by different functions that act during discrete phases in the herpes simplex virus type 1 life cycle. J. Virol. **77:**10917-10928.

116. **Murphy, D.** 2002. Gene expression studies using microarrays: principles, problems, and prospects. Adv. Physiol. Educ. **26:**256-270.

117. **Nakamura, H., M. Lu, Y. Gwack, J. Souvlis, S. L. Zeichner, and J. U. Jung.** 2003. Global changes in Kaposi's sarcoma-associated virus gene expression patterns following expression of a tetracycline-inducible Rta transactivator. J. Virol. **77:**4205-4220.

118. **Negorev, D. G., O. V. Vladimirova, A. Ivanov, F. Rauscher 3rd, and G. G. Maul.** 2006. Differential role of Sp100 isoforms in interferon-mediated repression of herpes simplex virus type 1 immediate-early protein expression. J. Virol. **80:**8019-8029.

119. **Nicholl, M. J. and C. M. Preston.** 1996. Inhibition of herpes simplex virus type 1 immediate-early gene expression by alpha interferon is not VP16 specific. J. Virol. **70:**6336-6339.

120. **Novoa, I., H. Zeng, H. P. Harding, and D. Ron.** 2001. Feedback inhibition of the unfolded protein response by GADD34-mediated dephosphorylation of eIF2alpha. J. Cell Biol. **153:**1011-1022.

121. **Nystrom, K., M. Biller, A. Grahn, M. Lindh, G. Larson, and S. Olofsson.** 2004. Real time PCR for monitoring regulation of host gene expression in herpes simplex virus type 1-infected human diploid cells. J. Virol. Methods **118:**83-94.

122. **Ogle, W. O., T. I. Ng, K. L. Carter, and B. Roizman.** 1997. The UL13 protein kinase and the infected cell type are determinants of posttranslational modification of ICP0. Virology **235:**406-413.

123. **Parkinson, J., S. P. Lees-Miller, and R. D. Everett.** 1999. Herpes simplex virus type 1 immediate-early protein vmw110 induces the proteasome-dependent degradation of the catalytic subunit of DNA-dependent protein kinase. J. Virol. **73:**650-657.

124. **Paulose-Murphy, M., N. K. Ha, C. Xiang, Y. Chen, L. Gillim, R. Yarchoan, P. Meltzer, M. Bittner, J. Trent, and S. Zeichner.** 2001. Transcription program of human herpesvirus 8 (kaposi's sarcoma-associated herpesvirus). J. Virol. **75:**4843-4853.

125. **Perelygina, L., L. Zhu, H. Zurkuhlen, R. Mills, M. Borodovsky, and J. K. Hilliard.** 2003. Complete sequence and comparative analysis of the genome of herpes B virus (Cercopithecine herpesvirus 1) from a rhesus monkey. J. Virol. **77:**6167-6177.

126. **Pfaffl, M. W.** 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. **29:**e45.

127. **Preston, C. M., M. C. Frame, and M. E. Campbell.** 1988. A complex formed between cell components and an HSV structural polypeptide binds to a viral immediate early gene regulatory DNA sequence. Cell **52:**425-434.

128. **Purves, F. C., W. O. Ogle, and B. Roizman.** 1993. Processing of the herpes simplex virus regulatory protein alpha 22 mediated by the UL13 protein kinase determines the accumulation of a subset of alpha and gamma mRNAs and proteins in infected cells. Proc. Natl. Acad. Sci. U. S. A. **90:**6701-6705.

129. **Quackenbush, J.** 2002. Microarray data normalization and transformation. Nat. Genet. **32 Suppl:**496-501.

130. **Quackenbush, J.** 2001. Computational analysis of microarray data. Nat. Rev. Genet. **2:**418-427.

131. **Rajcani, J., V. Andrea, and R. Ingeborg.** 2004. Peculiarities of herpes simplex virus (HSV) transcription: an overview. Virus Genes **28:**293-310.

132. **Randall, G., M. Lagunoff, and B. Roizman.** 1997. The product of ORF O located within the domain of herpes simplex virus 1 genome transcribed during latent infection binds to and inhibits in vitro binding of infected cell protein 4 to its cognate DNA site. Proc. Natl. Acad. Sci. U. S. A. **94:**10379-10384.

133. **Ravi, V., P. G. Kennedy, and A. R. MacLean.** 1998. Functional analysis of the herpes simplex virus type 2 strain HG52 RL1 gene: the intron plays no role in virulence. J. Gen. Virol. **79 (Pt 7):**1613-1617.

134. **Ray, N. and L. W. Enquist.** 2004. Transcriptional response of a common permissive cell type to infection by two diverse alphaherpesviruses. J. Virol. **78:**3489-3501.

135. **Regad, T. and M. K. Chelbi-Alix.** 2001. Role and fate of PML nuclear bodies in response to interferon and viral infections. Oncogene **20:**7274-7286.

136. **Revel, A. T., A. M. Talaat, and M. V. Norgard.** 2002. DNA microarray analysis of differential gene expression in Borrelia burgdorferi, the Lyme disease spirochete. Proc. Natl. Acad. Sci. U. S. A. **99:**1562-1567.

137. **Rice, S. A. and D. M. Knipe.** 1988. Gene-specific transactivation by herpes simplex virus type 1 alpha protein ICP27. J. Virol. **62:**3814-3823.

138. **Richmond, J. Y. and R. W. McKinney (eds.)** 1999. Biosafety in Microbiological and Biomedical Laboratories. U.S. Government Printing Office, Washington.

139. **Richter, A., C. Schwager, S. Hentze, W. Ansorge, M. W. Hentze, and M. Muckenthaler.** 2002. Comparison of fluorescent tag DNA labeling methods used for expression analysis by DNA microarrays. BioTechniques **33:**620-8, 630.

140. **Ritchey, J. W., K. A. Ealey, M. E. Payton, and R. Eberle.** 2002. Comparative pathology of infections with baboon and African green monkey alpha-herpesviruses in mice. J. Comp. Pathol. **127:**150-161.

141. **Rogers, K. M., K. A. Ealey, J. W. Ritchey, D. H. Black, and R. Eberle.** 2003. Pathogenicity of different baboon herpesvirus papio 2 isolates is characterized by either extreme neurovirulence or complete apathogenicity. J. Virol. **77:**10731-10739.

142. **Rogers, K. M., J. W. Ritchey, M. Payton, D. H. Black, and R. Eberle.** 2006. Neuropathogenesis of herpesvirus papio 2 in mice parallels infection with Cercopithecine herpesvirus 1 (B virus) in humans. J. Gen. Virol. **87:**267-276.

143. **Roizman, B. and D. M. Knipe.** 2001. Chapter 72: Herpes Simplex Viruses and Their Replication. In D. M. Knipe and P. M. Howley (ed.), Fields Virology, vol. 2. LIPPINCOTT WILLIAMS & WILKINS, 4th edition.

144. **Roizman, B. and P. E. Pellett.** 2001. Chapter 71: The family *herpesviridae:* a brief introduction. In D. M. Knipe and P. M. Howley (ed.), Fields Virology, vol. 2. LIPPINCOTT WILLIAMS & WILKINS, USA.

145. **Sacks, S. L., P. D. Griffiths, L. Corey, C. Cohen, A. Cunningham, G. M. Dusheiko, S. Self, S. Spruance, L. R. Stanberry, A. Wald, and R. J. Whitley.** 2004. HSV shedding. Antiviral Res. **63 Suppl 1:**S19-26.

146. **Sacks, W. R., C. C. Greene, D. P. Aschman, and P. A. Schaffer.** 1985. Herpes simplex virus type 1 ICP27 is an essential regulatory protein. J. Virol. **55:**796-805.

147. **Sacks, W. R. and P. A. Schaffer.** 1987. Deletion mutants in the gene encoding the herpes simplex virus type 1 immediate-early protein ICP0 exhibit impaired growth in cell culture. J. Virol. **61:**829-839.

148. **Samaniego, L. A., N. Wu, and N. A. DeLuca.** 1997. The herpes simplex virus immediate-early protein ICP0 affects transcription from the viral genome and infected-cell survival in the absence of ICP4 and ICP27. J. Virol. **71:**4614-4625.

149. **Sciortino, M. T., M. Suzuki, B. Taddeo, and B. Roizman.** 2001. RNAs extracted from herpes simplex virus 1 virions: apparent selectivity of viral but not cellular RNAs packaged in virions. J. Virol. **75:**8105-8116.

150. **Severini, A., A. R. Morgan, D. R. Tovell, and D. L. Tyrrell.** 1994. Study of the structure of replicative intermediates of HSV-1 DNA by pulsed-field gel electrophoresis. Virology **200:**428-435.

151. **Shepard, A. A. and N. A. DeLuca.** 1991. A second-site revertant of a defective herpes simplex virus ICP4 protein with restored regulatory activities and impaired DNA-binding properties. J. Virol. **65:**787-795.

152. **Shepard, A. A., A. N. Imbalzano, and N. A. DeLuca.** 1989. Separation of primary structural components conferring autoregulation, transactivation, and DNA-binding properties to the herpes simplex virus transcriptional regulatory protein ICP4. J. Virol. **63:**3714-3728.

153. **Singh, J. and E. K. Wagner.** 1993. Transcriptional analysis of the herpes simplex virus type 1 region containing the TRL/UL junction. Virology **196:**220-231.

154. **Smith, C. A., P. Bates, R. Rivera-Gonzalez, B. Gu, and N. A. DeLuca.** 1993. ICP4, the major transcriptional regulatory protein of herpes simplex virus type 1, forms a tripartite complex with TATA-binding protein and TFIIB. J. Virol. **67:**4676-4687.

155. **Smyth, G. K., J. Michaud, and H. S. Scott.** 2005. Use of within-array replicate spots for assessing differential expression in microarray experiments. Bioinformatics **21:**2067-2075.

156. **Spencer, C. A., M. E. Dahmus, and S. A. Rice.** 1997. Repression of host RNA polymerase II transcription by herpes simplex virus type 1. J. Virol. **71:**2031-2040.

157. **Stern, S. and W. Herr.** 1991. The herpes simplex virus trans-activator VP16 recognizes the Oct-1 homeo domain: evidence for a homeo domain recognition subdomain. Genes Dev. **5:**2555-2566.

158. **Stingley, S. W., J. J. Ramirez, S. A. Aguilar, K. Simmen, R. M. Sandri-Goldin, P. Ghazal, and E. K. Wagner.** 2000. Global analysis of herpes simplex virus type 1 transcription using an oligonucleotide-based DNA microarray. J. Virol. **74:**9916-9927.

159. **Taddeo, B., T. R. Luo, W. Zhang, and B. Roizman.** 2003. Activation of NF-kappaB in cells productively infected with HSV-1 depends on activated protein kinase R and plays no apparent role in blocking apoptosis. Proc. Natl. Acad. Sci. U. S. A. **100:**12408-12413.

160. **Talaat, A. M., S. T. Howard, W. Hale 4th, R. Lyons, H. Garner, and S. A. Johnston.** 2002. Genomic DNA standards for gene expression profiling in Mycobacterium tuberculosis. Nucleic Acids Res. **30:**e104.

161. **Tang, Q., L. Li, A. M. Ishov, V. Revol, A. L. Epstein, and G. G. Maul.** 2003. Determination of minimum herpes simplex virus type 1 components necessary to localize transcriptionally active DNA to ND10. J. Virol. **77:**5821-5828.

162. **Telford, E. A., M. S. Watson, K. McBride, and A. J. Davison.** 1992. The DNA sequence of equine herpesvirus-1. Virology **189:**304-316.

163. **Tinker, R. L., K. P. Williams, G. A. Kassavetis, and E. P. Geiduschek.** 1994. Transcriptional activation by a DNA-tracking protein: structural consequences of enhancement at the T4 late promoter. Cell **77:**225-237.

164. **Tran, R. K., P. T. Lieu, S. Aguilar, E. K. Wagner, and D. C. Bloom.** 2002. Altering the expression kinetics of VP5 results in altered virulence and pathogenesis of herpes simplex virus type 1 in mice. J. Virol. **76:**2199-2205.

165. **Triezenberg, S. J., K. L. LaMarco, and S. L. McKnight.** 1988. Evidence of DNA: protein interactions that mediate HSV-1 immediate early gene activation by VP16. Genes Dev. **2:**730-742.

166. **Tyler, S. D., G. A. Peters, and A. Severini.** 2005. Complete genome sequence of cercopithecine herpesvirus 2 (SA8) and comparison with other simplexviruses. Virology **331:**429-440.

167. **Tyler, S. D. and A. Severini.** 2006. The complete genome sequence of herpesvirus papio 2 (Cercopithecine herpesvirus 16) shows evidence of recombination events among various progenitor herpesviruses. J. Virol. **80:**1214-1221.

168. **Uprichard, S. L. and D. M. Knipe.** 1996. Herpes simplex ICP27 mutant viruses exhibit reduced expression of specific DNA replication genes. J. Virol. **70:**1969-1980.

169. **van Bakel, H. and F. C. Holstege.** 2004. In control: systematic assessment of microarray performance. EMBO Rep. **5:**964-969.

170. **van de Peppel, J., P. Kemmeren, H. van Bakel, M. Radonjic, D. van Leenen, and F. C. Holstege.** 2003. Monitoring global messenger RNA changes in externally controlled microarray experiments. EMBO Rep. **4:**387-393.

171. **Vogel, J. L. and T. M. Kristie.** 2000. The novel coactivator C1 (HCF) coordinates multiprotein enhancer formation and mediates transcription activation by GABP. EMBO J. **19:**683-690.

172. **Wagner, E. K.** 2003. Some Genetic Functions Encoded by Herpes Simplex Virus type 1. **2006:**http://www.dbc.uci.edu/~faculty/wagner/table.html.

173. **Wagner, E. K., J. J. Ramirez, S. W. Stingley, S. A. Aguilar, L. Buehler, G. B. Devi-Rao, and P. Ghazal.** 2002. Practical approaches to long oligonucleotide-based DNA microarray: lessons from herpesviruses. Prog. Nucleic Acid Res. Mol. Biol. **71:**445-491.

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

175. **Weigler, B. J.** 1992. Biology of B virus in macaque and human hosts: a review. Clin. Infect. Dis. **14:**555-567.

176. **Weigler, B. J., D. W. Hird, J. K. Hilliard, N. W. Lerche, J. A. Roberts, and L. M. Scott.** 1993. Epidemiology of cercopithecine herpesvirus 1 (B virus) infection and shedding in a large breeding cohort of rhesus macaques. J. Infect. Dis. **167:**257-263.

177. **Weil, M. R., T. Macatee, and H. R. Garner.** 2002. Toward a universal standard: comparing two methods for standardizing spotted microarray data. BioTechniques **32:**1310-1314.

178. **Weir, J. P.** 2001. Regulation of herpes simplex virus gene expression. Gene **271:**117-130.

179. **Whitley, R. J.** 2001. Chapter 73: Herpes Simplex Viruses. In D. M. Knipe and P. M. Howley (ed.), Fields Virology, vol. 2. LIPPINCOTT WILLIAMS & WILKINS, USA.

180. **Whitley, R. J. and J. K. Hilliard.** 2001. Chapter 83: Cercopithecine herpesvirus (B virus). In D. M. Knipe and P. M. Howley (ed.), Fields Virology, vol. 2. LIPPINCOTT WILLIAMS & WILKINS, USA.

181. **Williams, B. A., R. M. Gwirtz, and B. J. Wold.** 2004. Genomic DNA as a cohybridization standard for mammalian microarray measurements. Nucleic Acids Res. **32:**e81.

182. **Wirth, U. V., C. Fraefel, B. Vogt, C. Vlcek, V. Paces, and M. Schwyzer.** 1992. Immediate-early RNA 2.9 and early RNA 2.6 of bovine herpesvirus 1 are 3' coterminal and encode a putative zinc finger transactivator protein. J. Virol. **66:**2763-2772.

183. **Wu, N., S. C. Watkins, P. A. Schaffer, and N. A. DeLuca.** 1996. Prolonged gene expression and cell survival after infection by a herpes simplex virus mutant defective in the immediate-early genes encoding ICP4, ICP27, and ICP22. J. Virol. **70:**6358-6369.

184. **Xia, K., N. A. DeLuca, and D. M. Knipe.** 1996. Analysis of phosphorylation sites of herpes simplex virus type 1 ICP4. J. Virol. **70:**1061-1071.

185. **Xiao, H., A. Pearson, B. Coulombe, R. Truant, S. Zhang, J. L. Regier, S. J. Triezenberg, D. Reinberg, O. Flores, and C. J. Ingles.** 1994. Binding of basal transcription factor TFIIH to the acidic activation domains of VP16 and p53. Mol. Cell. Biol. **14:**7013-7024.

186. **Xu, F., M. R. Sternberg, B. J. Kottiri, G. M. McQuillan, F. K. Lee, A. J. Nahmias, S. M. Berman, and L. E. Markowitz.** 2006. Trends in herpes simplex virus type 1 and type 2 seroprevalence in the United States. JAMA **296:**964-973.

187. **Xu, Q., M. Dziejman, and J. J. Mekalanos.** 2003. Determination of the transcriptome of Vibrio cholerae during intraintestinal growth and midexponential phase in vitro. Proc. Natl. Acad. Sci. U. S. A. **100:**1286-1291.

188. **Yang, I. V., E. Chen, J. P. Hasseman, W. Liang, B. C. Frank, S. Wang, V. Sharov, A. I. Saeed, J. White, J. Li, N. H. Lee, T. J. Yeatman, and J. Quackenbush.** 2002. Within the fold: assessing differential expression measures and reproducibility in microarray assays. Genome Biol. **3:**research0062.

189. **Yang, W. C., G. V. Devi-Rao, P. Ghazal, E. K. Wagner, and S. J. Triezenberg.** 2002. General and specific alterations in programming of global viral gene expression during infection by VP16 activation-deficient mutants of herpes simplex virus type 1. J. Virol. **76:**12758-12774.

190. **Yao, F. and P. A. Schaffer.** 1994. Physical interaction between the herpes simplex virus type 1 immediate-early regulatory proteins ICP0 and ICP4. J. Virol. **68:**8158-8168.

191. **Yeh, L. and P. A. Schaffer.** 1993. A novel class of transcripts expressed with late kinetics in the absence of ICP4 spans the junction between the long and short segments of the herpes simplex virus type 1 genome. J. Virol. **67:**7373-7382.

192. **Zabierowski, S. and N. A. DeLuca.** 2004. Differential cellular requirements for activation of herpes simplex virus type 1 early (tk) and late (gC) promoters by ICP4. J. Virol. **78:**6162-6170.

193. **Zhu, H., J. P. Cong, G. Mamtora, T. Gingeras, and T. Shenk.** 1998. Cellular gene expression altered by human cytomegalovirus: global monitoring with oligonucleotide arrays. Proc. Natl. Acad. Sci. U. S. A. **95:**14470-14475.