TRANSCRIPTOME ACTIVITY OF HUMAN CYTOMEGALOVIRUS (STRAIN MERLIN) IN FIBROBLASTS, EPITHELIAL CELLS AND ASTROCYTES

BY

JAMES CHARLES TOWLER

THESIS SUBMITTED FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN THE INSTITUTE OF BIOMEDICAL AND LIFE SCIENCES AT THE UNIVERSITY OF GLASGOW

DECEMBER 2007

MRC VIROLOGY UNIT
INSTITUTE OF VIROLOGY
CHURCH STREET
GLASGOW
G11 5JR

SUMMARY

Global human cytomegalovirus (HCMV) gene expression was investigated during replication in three permissive human cell types in tissue culture; human foetal foreskin fibroblasts (HFFF-2), human retinal pigmented epithelial (RPE) cells, and human astrocytoma cells (U373Mg). A custom HCMV DNA microarray based on recent re-assessments of HCMV coding potential was designed. The Merlin strain of HCMV was used for these studies because it has been reported to contain the complete set of ORFs, only one of which (UL128) is mutated. The UL128 gene locus is invariably mutated in HCMV isolates propagated in fibroblasts, and the premature termination mutation of UL128 greatly enhanced infectious yields from Merlin passaged in HFFF-2 cells. The HCMV (Merlin) microarray consists of 60-mer oligonucleotide probes derived from both 3′- and 5′-proximal regions of each recognised ORF, and 3′-proximal probes for novel ORFs that have been proposed from *in silico* studies. Probes were also included for several previously reported ORFs that are now considered to non-protein coding. Positive and negative-sense bacterial sequence probes were included on the array, and were used in conjuction with spiked-in cognate RNAs in the cDNA synthesis reaction as controls for normalisation. The quality of the printed HCMV microarray was assessed and its specificity validated using Cy3-labelled cDNAs prepared from total RNA extracted from mock-infected and HCMV-infected cells at 96 h PI. Hybridisation conditions were then investigated to achieve optimal specificity and sensitivity of cDNA binding to cognate probes on the array.

Prior to commencing the microarray studies, the growth characteristics of HCMV (Merlin) were compared in each of the three cell lines. One-step virus growth curves revealed differential virus replication kinetics in the three cell types. Compared to infected HFFF-2 cultures, there was a 24 h delay in exit of virus from the viral eclipse phase in RPE and U373Mg cells and little or no release of infectivity to the extracellular medium. Differential growth kinetics in the three cell types were not due to differences in the ability of HCMV to enter cells or to induce virus gene expression, nor was it due to gross differences in the temporal expression kinetics of immediate-early, early and late protein synthesis in the three cell types. However, differences in the amount of protein made were evident, with viral protein expression lowest in RPE cells. Differences in virus growth kinetics were probably due to differences in the numbers of virus particles assembled, and/or their maturation and egress. It was therefore considered valid to

compare the temporal kinetics of HCMV (Merlin) transcript expression in the three cell types in order to identify genes that were regulated differently.

cDNAs were prepared from total RNA extracted from infected cells at 12, 24, 48 and 72 h PI, or at 72 h PI only from mock-infected cultures. After collection of the raw data, receiver operating characteristic (ROC) analysis was performed using the positive and negative control signals in order to determine cutoff points that discriminated between true-positive and true-negative hybridisation signals. The GeneSpring gene expression analysis software was used for statistical analysis of the microarray data. Human fibroblast cells have been extensively used in HCMV research, and so the HFFF-2 cell type was used as a reference cell type for the microarray work. Differential expression of a virus gene then relates to differences in amounts and/or expression kinetics between HFFF-2 and RPE cells, or between HFFF-2 and U373Mg cells. To identify differentially expressed virus genes, combined statistical tests were performed on the mean expression value for each gene from all data points over the time course, giving a single expression value for each gene in each cell type. The statistical tests then compared the expression values for individual HCMV genes in infected HFFF-2 cells against the corresponding expression values for individual HCMV genes in RPE or U373Mg cells.

Comparing the microarray data from HFFF-2 and RPE infected cells, 13 HCMV ORFs (UL4, UL16, UL45, UL148, IRS1, US11, US12, US13, US14, US15, US18, US19, US20), were found to be differentially expressed, and this was confirmed by examination of the expression kinetics of the individual genes. When the microarray data from infected HFFF and U373Mg cells were compared, 26 ORFs appeared to be differentially expressed. However, the microarray showed that late HCMV genes were expressed at unusually early times (24 h PI) in U373Mg cells; in contrast to the expression of their protein products. The data suggested that in U373Mg cells, either the HCMV transcriptome cascade was completed more rapidly, or that there was a breakdown in regulation of transcription control. Most of the 26 ORFs differentially expressed in HFFF-2 and U373Mg cells (identified by combined statistical testing) are expressed at a significantly higher level in U373Mg cells, but 7 were made in significantly reduced amounts (UL4, IRS1, US12, US14, US18, US19 and US20), and these were considered more likely candidates for differential expression, and were also differentially expressed in RPE cells. The disregulation of transit through the HCMV transcript cascade in U373Mg cells is thought to be due to the fact that p53 is mutated in this cell line. It has been reported that p53 mutants including the p53 mutation in U373Mg cells are capable of activating transcription from the HCMV MIEP and that the minimal promoter sequence is a TATA

box. It may be that mutant p53 activates HCMV early and late promoters resulting in an accelerated transit through the HCMV transcription cascade in infected U373Mg cells.

Of the 13 ORFs differentially expressed in HFFF-2 and RPE cells, UL4 is reported to be under both transcriptional and translational control. Cellular and viral transcription factors are involved in both positive and negative regulation of the UL4 promoter making it a good candidate for differential regulation in different cell types, although the function provided by UL4 is unknown. The immediate-early IRS1 gene is an important viral transactivator required throughout the virus replication cycle, and also self regulates its gene expression though an internal ORF. The functions of the US12 family genes (US12, US13, US14, US15, US18, US19 and US20) are unknown, but it has been suggested that they have a role in virus particle tegumentation, envelopment, and egress from infected cells. However, the identification of multiple members of the US12 gene family as differentially regulated should be interpreted with caution, since US18, US19 and US20 produce 3′ co-terminal transcripts, and it is probable that other family members also share polyadenylation sites. Down-regulation of US12 family genes with potential roles in virus maturation and egress are consistent with the impaired release of virus to the extracellular medium from RPE and U373Mg cells. Compared to HFFF-2 cells, the down-regulation of UL16 and US11 in RPE and U373Mg cells is interesting since the genes have immune evasion functions, and these cell types are located in immune-privileged organs, i.e. the brain (astrocytes) and eye (retinal epithelia). As yet, it remains unclear why the expression of two immune evasion genes from a total of ten, are down-regulated in these cell types. The down-regulation of UL45 in RPE cells may play a role in virus dissemination in the eye. While the HCMV UL45 gene product is a component of the virus tegument and might supply an important function early in the virus replication cycle, a UL45 mutant exhibited a growth defect in fibroblasts that changed cell-cell spread characteristics. Interestingly, M45, the MCMV homologue of HCMV UL45, is reported as a determinant of endothelial cell tropism.

With respect to proposed novel ORFs identified by *in silico* analysis, in most cases, we found no evidence for transcript expression. Of those that gave positive hybridisation signals, most might be explained by overlapping transcripts from genes in the same region and coding in the same direction. Other novel ORFs lie within regions of the genome now considered to be non-coding, but where transcripts have previously been reported, while the remainder may represent genuine coding ORFs. The lack of signal for previously described ORFs that are now considered non-protein coding confirms their status as discounted genes.

In order to test the microarray system, the temporal expression kinetics of selected virus genes were investigated by alternative methods including; real-time PCR, and northern blots to check the identity of specific transcripts, and where antibodies were available, western immunoblotting to confirm the expression kinetics of specific proteins. The expression kinetics obtained for specific genes both differentially and non-differentially expressed by these various methods were entirely consistent with those obtained for the same genes with the HCMV microarray. It was concluded that the Merlin microarray system was a valid and reliable research tool for the investigation of HCMV gene expression.

TABLE OF CONTENTS

SUMMARY 2
TABLE OF CONTENTS 6
LIST OF TABLES 9
LIST OF FIGURES 11
ACKNOWLEDGEMENTS 14
ABBREVIATIONS 15

1 INTRODUCTION 19
1.1 The clinical problem 19
1.2 Structural components of HCMV particles 21
1.2.1 The capsid 21
1.2.2 The envelope 22
1.2.3 The tegument 24
1.3 The HCMV genome and its coding potential 25
1.3.1 HCMV transcript mapping 26
1.4 The replication cycle 32
1.4.1 The major stages of the HCMV replication cycle 32
1.4.2 Kinetic class of gene expression 36
1.4.3 HCMV induced effects on intracellular signalling 36
1.5 Transcriptional transactivation in HCMV infected cells 40
1.5.1 Promoter structure 40
1.5.2 Switch from immediate-early, to early, to late gene expression 41
1.5.3 Regulatory activities of HCMV immediate-early proteins 45
1.5.3.1 Role of IE1 in the transactivation of gene expression 45
1.5.3.2 Role of IE2 in the transactivation of gene expression 46
1.6 Herpesvirus tropism 48
1.6.1 HCMV cell tropism 48
1.6.2 HCMV sequence hypervariability and its effects on cell tropism 51
1.6.3 HCMV genes involved in cell tropism 52
1.6.4 Role of viral determinants for cell tropism in non-human herpesviruses 54
1.7 Application of microarray technology to the study of viruses 55
1.7.1 Types of microarrays in current use 56
1.7.2 Microarrays used in transcriptome profiling of virus gene expression 58
1.7.3 Microarrays used in direct comparative studies of gene expression in different viruses 59
1.7.4 Microarrays used to determine the effect of herpesvirus infection on cellular gene expression 61
1.7.5 Microarrays used to compare the effect of regulatory virus genes on viral or cellular transcription 62
1.7.6 Microarrays used to compare of viral gene expression during lytic and latent phases of infection 63
1.7.7 Microarrays used to compare viral gene expression in different cell types infected in culture 64
1.7.8 Microarrays designed for the simultaneous comparison of viral and cellular gene expression 65
1.7.9 Viral diagnostic microarrays 65
1.8 Aims of the thesis 67
2 MATERIALS AND METHODS 68
2.1 Materials 68
2.1.1 Chemicals 68
2.1.2 Radiochemicals 68

2.1.3 Restriction Endonucleases 68
2.1.4 Antibodies 68
2.1.5 Composition of solutions and buffers 68
2.1.6 Miscellaneous reagents and commercial kits 70
2.1.7 Miscellaneous materials and plastics 71
2.1.8 Computer software and algorithms 72
2.2 Methods 73
2.2.1 Cell culture 73
2.2.1.1 Propagation of cell stocks 73
2.2.1.2 Storage of cells in liquid nitrogen 73
2.2.2 Preparation of virus stocks 74
2.2.2.1 Titration of virus stocks 74
2.2.3 Extraction of HCMV genomic DNA 74
2.2.4 One step virus growth curve 75
2.2.5 Preparation of mock-infected and HCMV infected total cellular RNA 76
2.2.6 Assessment of RNA integrity 76
2.2.7 DNase I treatment of total RNA 76
2.2.8 Synthesis of cDNA 77
2.2.8.1 Synthesis of cDNA for microarray hybridisation 77
2.2.8.2 Synthesis of cDNA for quantitative PCR 77
2.2.9 HCMV Microarray 77
2.2.9.1 Probe design and microarray fabrication 78
2.2.9.2 Microarray hybridisation 84
2.2.9.2.1 Preparation of cDNA 84
2.2.9.2.2 Pre-hybridisation of microarray slides 84
2.2.9.2.3 Hybridisation and stringency washes 84
2.2.9.2.4 Scanning and quantitation of microarrays 85
2.2.9.3 Data Processing 86
2.2.9.3.1 Determination of linear dynamic range and assessment of signal distribution 86
2.2.9.3.2 Receiver operating characteristic analysis 86
2.2.9.3.3 Normalisation 88
2.2.9.3.4 Data flagging and preparation for import into GeneSpring 88
2.2.9.4 Analysis of data in GeneSpring 89
2.2.9.4.1 Preparation of the GeneSpring software and import of microarray data 89
2.2.9.4.2 Examination of present and absent flags 90
2.2.9.4.3 Quality control based on clustering 90
2.2.9.4.4 Differential gene expression 90
2.2.10 PCR 91
2.2.11 Agarose gel electrophoresis 91
2.2.12 Recovery of DNA fragments 91
2.2.13 Restriction endonuclease digests 92
2.2.14 Preparation of FIX-BAC DNA 92
2.2.15 Real-time PCR 92
2.2.16 Northern Blotting 92
2.2.16.1 Formaldehyde-agarose gel electrophoresis 92
2.2.16.2 RNA transfer by capillary blotting 93
2.2.16.3 Preparation of DNA probes 93
2.2.16.4 Nucleic acid hybridisation 94
2.2.16.5 Phosphorimager analysis 94
2.2.17 Western Blotting 94
2.2.17.1 Preparation of mock-infected and HCMV infected cellular protein 94
2.2.17.2 SDS-PAGE electrophoresis 95
2.2.17.3 Western immunoblotting 95

2.2.18 Immunofluorescence 95
3 RESULTS I 97
3.1 Characterisation of HCMV strain Merlin replication in different cell types 97
3.2 One step virus growth curves 97
3.3 Efficiency of HCMV infection in HFFF-2, RPE and U373Mg cells 100
3.4 Expression kinetics of known HCMV immediate-early, early and late proteins 100
3.5 Discussion 107
4 RESULTS II 109
4.1 HCMV microarray validation 109
4.1.1 HCMV microarray print designs 109
4.1.2 Assessment of the specificity of viral oligonucleotide probes 19
4.1.3 The effect of hybridisation temperature on the specificity of the microarray 113
4.1.4 Comparison of cDNA synthesis methods on the specificity of the microarray 113
4.1.5 Specificity of the microarray for the detection of control features 116
4.2 Transcriptome profiling of HCMV in fibroblasts, epithelial cells and astrocytes 119
4.2.1 HCMV microarray hybridisations 119
4.2.2 Determining the optimum PMT settings 127
4.2.3 Assessment of signal distributions 131
4.2.4 Receiver operating characteristic analysis 131
4.2.5 Normalisation 137
4.2.6 Data flagging and preparation for import into GeneSpring 137
4.2.7 Microarry analysis using GeneSpring 137
4.2.7.1 Clustering analysis 137
4.2.7.2 Analysis based on present and absent flags 140
4.2.8 Expression profiles for representative non-differentially expressed immediate-early, early, and late regulated HCMV genes in HFFF-2 and RPE cells .144
4.2.9 Differential gene expression 144
4.2.10 Analysis of proposed novel ORFs and previously discounted ORFs 156
4.3 Discussion 160
5 RESULTS III 165
5.1 Validation of the data 165
5.2 HCMV 3' co-terminal transcripts, overlapping transcripts and spliced genes 165
5.3 Expression and validation of representative non-differentially regulated genes 169
5.4 Expression and validation of differentially regulated genes 181
5.5 Discussion 194
5.5.1 Genes differentially expressed in HFFF-2 and RPE cells 194
5.5.2 Gene differentially expressed in HFFF-2 and U373Mg cells 194
5.5.3 Conclusion 195
6 GENERAL DISCUSSION 196
6.1 Virus replication kinetics and the impact on gene expression 196
6.2 Viral genes specifying cell tropism factors 200
6.3 Future work on HCMV cell tropism 207
7 REFERENCES 208

LIST OF TABLES

1 INTRODUCTION

Table 1.1 HCMV ORFs in strain Merlin and associated homologues in MCMV and CCMV 28

Table 1.2 Additional ORFs contained in the HCMV microarray 31

Table 1.3 References for mapped HCMV transcripts.....33

2 METHODS

Table 2.1 List of antibodies 68

Table 2.2 List of HCMV microarray oligonucleotide sequences 79

Table 2.3 Bacterial spike control probe sequences.....83

4 RESULTS II

Table 4.1 Assessment of microarray signal intensities during hybridisation of HCMV-infected HFFF-2 RNA under different experimental conditions 115

Table 4.2 Assessment of median signal intensities for bacterial control probes 115

Table 4.3 Microarray data analysis determining selection cutoff points and correction factors.....121

Table 4.4 Extract of data used to generate the ROC curve for array number 27.....135

Table 4.5 Microarray dataset information for individual Merlin genes in each cell type..... 141

Table 4.6 Statistical analysis p-values for HFFF-2 and RPE or U373Mg 147

Table 4.7 HCMV genes that are differentially expressed in HFFF-2 and RPE cells 149

Table 4.8 HCMV genes that are differentially expressed in HFFF-2 and U373Mg 150

Table 4.9 Detection of transcripts in HFFF-2 and RPE cells hybridising to probes specific for the novel HCMV ORFs proposed by Murphy et al., 2003(a).....157

Table 4.10 Detection of transcripts in HFFF-2 and RPE cells hybridising to probes specific for the novel HCMV ORFs proposed by Murphy et al., 2003(b) 157

Table 4.11 Detection of transcripts in HFFF-2 and RPE cells hybridising to probes specific for the HCMV ORFs discounted by Dolan et al., 2002 159

5 RESULTS III

Table 5.1 List of RT-qPCR primers 166

Table 5.2 Northern blotting probe information 166

Table 5.3 HCMV polyadenylation signal (AATAAA) genome positions and predicted poly(A) usage by HCMV ORFs.....167

LIST OF FIGURES

1 INTRODUCTION

Figure 1.1 Gene map of the Merlin strain of HCMV....27

Figure 1.2 HCMV transcript map....34

Figure 1.3 Simplified diagram of the HCMV induced effects on intracellular signalling..38

Figure 1.4 Diagram depicting the mutual regulation of HCMV and NF-κB....39

Figure 1.5 Structure of the IE1/IE2 promoter-enhancer and the regulation of downstream gene expression....42

Figure 1.6 Promoter structure for early and late HCMV genes....43

2 MATERIALS AND METHODS

Figure 2.1 Analysis procedure for the HCMV microarray data....87

3 RESULTS I

Figure 3.1 One step HCMV growth curves obtained for HFFF-2, RPE and U373Mg cells infected at a m.o.i. of 1 p.f.u./cell....98

Figure 3.2 One step HCMV cell released virus (CRV) growth curves from HFFF-2, RPE and U373Mg cells infected at a m.o.i. of 6 p.f.u./cell....99

Figure 3.3 One step virus growth curve of HCMV cell associated virus (CAV) in HFFF-2, RPE and U373Mg cells at m.o.i. of 6 p.f.u./cell....101

Figure 3.4 (A) shows representative immunofluorescence images for HFFF-2, RPE and U373Mg cells infected with HCMV at m.o.i. of 3, 6 and 10 p.f.u./cell and expressing the HCMV early gene UL44 (FITC labelled), and (B) shows the numbers of UL44-FITC labelled cells in two representative micrscope fields of view for each cell type infected at each m.o.i. 103

Figure 3.5 Western blots for the immediate-early proteins IE1/IE2....105

Figure 3.6 Western blots for the tegument component pUS22....106

Figure 3.7 Western blots for the late gene UL99 (pp28) 106

4 RESULTS II

Figure 4.1 Design and validation procedure for the HCMV microarray....110

Figure 4.2 HCMV microarray print designs (P1) and (P2)....111

Figure 4.3 Assessment of probe specificity....112

Figure 4.4 Confirmation of different temperatures on specificity of the microarray....114

Figure 4.5 Hybridisation of cDNA that had been synthesised using (A) oligo-dT; or (B) random hexamer primers 117

Figure 4.6 Hybridisation of control 'spiked-in' RNA 118

Figure 4.7 HCMV microarray print design (P3) 120

Figure 4.8 Representative microarray images from one of three biological replicates for HCMV infected HFFF-2, RPE and U373Mg cells 122

Figure 4.9 Determination of linear dynamic range 129

Figure 4.10 Histograms of (A) linear microarray data and (B) log_2 transformed microarray data 130

Figure 4.11 Box-plots for (A) HFFF-2, (B) RPE and (C) U373Mg microarray datasets prior to scaling....132

Figure 4.11.1 Box-plots for (A) HFFF-2, (B) RPE and (C) U373Mg microarray datasets following scaling 133

Figure 4.11.2 Box-plots for (A) HFFF-2, (B) RPE and (C) U373Mg microarray datasets following normalisation in GeneSpring 134

Figure 4.12 ROC curve for array number 27 (true-positive vs false-positive rates) determined for the signal cutoff points shown in Table 4.2 136

Figure 4.13 Condition trees for the (A) HFFF-2, (B) RPE and (C) U373Mg microarray datasets 138

Figure 4.14 Expression profiles for representative non-differentially regulated genes in HFFF-2 and RPE cells 145

Figure 4.15 Microarray expression curves for genes listed in Table 4.7 as differentially expressed in HFFF-2 and RPE cells 152

Figure 4.16 Microarray expression curves for genes listed in Table 4.8 as differentially expressed in HFFF-2 and U373Mg cells....155

5 RESULTS III

Figure 5.1 Comparison of microarray expression data for UL123 (IE1) with associated RT-qPCR, northern blotting for UL123, and western blotting for IE1/IE2....170

Figure 5.1.1 Summary diagram of the IE1/IE2 gene locus, with the positions of the microarray, RT-qPCR and northern blot probe regions....171

Figure 5.2 Comparison of microarray data for UL55 (gB) with expression data detected by RT-qPCR, northern blot data....173

Figure 5.2.1 Summary diagram of the transcripts expressed in the UL54 to UL57 gene region....174

Figure 5.3 Comparison of microarray data for UL83 (pp65) with expression data detected by RT-qPCR, northern and western blotting 176

Figure 5.3.1 Summary diagram of the transcript expressed from the start of UL83 to a poly(A) signal downstream of UL82.....177

Figure 5.4 Comparison of microarray expression data for UL99 (pp28) with expression data from RT-qPCR, northern and western blotting 179

Figure 5.4.1 Summary diagram of the transcripts expressed from UL93 to UL99 180

Figure 5.5 Comparison of microarray expression data for UL43 with associated northern blot data.....182

Figure 5.6 Comparison of microarray expression data for UL4 with associated RT-qPCR and northern blot data.....183

Figure 5.7 Comparison of microarray expression data for UL16 with associated northern blot data.....185

Figure 5.8 Comparison of microarray expression data for IRS1 with associated RT-qPCR and northern blot data.....186

Figure 5.8.1 Summary diagram of IRS1 transcripts 187

Figure 5.9 Comparison of microarray expression data for US12 with associated RT-qPCR and northern blot data.....189

Figure 5.9.1 Summary diagram of the transcripts expressed from US12 to US15 190

Figure 5.10 Comparison of microarray expression data for US18 with associated RT-qPCR and northern blot data.....192

Figure 5.10.1 Summary diagram of transcripts expressed from US18 to US20.....193

ACKNOWLEDGMENTS

First and foremost, I would like to thank Dr Derrick Dargan for his excellent guidance, encouragement, patience and supervision during this project, and for the critical reading of this manuscript. I would also like to thank Fiona Jamieson and Elaine Douglas for their technical advice and emotional support over the last four years. I must also give many thanks to Dr Bahram Ebrahimi for sharing his expertise on microarray technology; specifically the training I received to perform the microarray experiments, his advice and guidance for the analysis of the data, and for the critical reading of certain sections of this thesis. I must also mention Dr Brian Lane who provided guidance for establishing a protocol for the analysis of microarray data, and for the critical reading of the microarray section within this document. I would also like to thank Dr Ewan Hunter for his advice and guidance on microarray data analysis.

Many other individuals within the Institute of Virology have provided advice and support throughout this project, but before I go on to name specific individuals, I must thank all members of the technical, clerical and computing support within the unit. I would like to mention Dr Andrew Davison, Dr Richard Adair, Dr Ania Owsianka, Dr Derek Gatherer, Prof Chris Preston and Mr Colin Loney, who have provided technical advice and/or discussion at some point during this project. I would also like to thank Prof Duncan McGeoch for giving me the opportunity to undertake this research, and the Medical Research Council for the funding I received for this project. I must go on to thank certain students of the virology intake of 2003: Tanya Chaudry, David Dalrymple, Amanda Sykes, Louise Yule, Jon Hubb and Sarah Mole; and also Amanda Bradley and Martin Higgs. I would like to mention and thank people who have given me support throughout the last four years and beyond: Matthew Sutton, Richard Lowthorpe, Jane McLeod, Danielle Newton, Louise Jackson, Daniel Saville, Alex Eslor, Kate Rhodes, Patrick O'Neill, David Wilson, Dean Backhouse, Scott Popham and Gary Hannar. I must also thank Dr Peter McParlin, who has been a good friend and mentor over the last five years. Last, but not least, I would like to thank my family who have given me unconditional encouragement and support, for which I am deeply indebted; my sisters, Jane and Karen, and my parents, Marilyn and Philip.

I declare that this thesis consists entirely of my own work, unless specifically indicated. This thesis has not been accepted in any previous application for a degree.

ABBREVIATIONS

A	AA	Amino Acid
	AC	Assembly Compartment
	AcMNPV	Autographa Californica Multicapsid Nucleopolyhedrovirus
	AIDS	Acquired Immune Deficiency Syndrome
	AP-1	Activator Protein
	APC	Antigen Presenting Cell
	ATF	Activating Transcription Factor
	ATP	Adenosine Triphosphate
B	BAC	Bacterial Artificial Chromosome
	BDGF	Bio-Dictionary Gene Finder
	BL	Burkitt's Lymphoma
	BLAST	Basic Local Alignment Search Tool
	bp	Base-Pair
	BSA	Bovine Serum Albumin
C	CAV	Cell Associated Virus
	CBP	CREB Binding Protein
	cDNA	Complementary Deoxyribonucleic Acid
	CMV	Cytomegalovirus
	CCMV	Chimpanzee Cytomegalovirus
	CHX	Cycloheximide
	CLL	Chronic Lymphocytic Leukaemia
	CNT	Count
	CPE	Cytopathic Effect
	CREB	cAMP-Response Element Binding Protein
	CRE	cAMP-Response Elements
	crs	Cis-Repsonse Signal
	CRV	Cell Released Virus
	CSF	Cerebrospinal Fluid
	CTL	Cytotoxic T-Lymphocyte
D	DAS	Downstream Activating Signal
	DC	Dendritic Cell
	DG	Diacylglycerol
	DMEM	Dulbecco's Modified Eagle's Medium
	DNA	Deoxyribonucleic Acid
E	E	Early
	EBV	Epstein-Barr Virus

	EGFR	Epidermal Growth Factor Receptor
	ERK	Extracellular-Signal Regulated Kinase
F	FCS	Foetal Calf Serum
	FITC	Fluorescein Isothiocyanate
G	GAL	Gene Array List
	GAPDH	Glyceraldehyde Phosphate Dehydrogenase
	GAPS	Gamma Amino Propyl Silane
	GCV	Gancyclovir
	GFAP	Glial Fibrallary Acidic Protein
	GPCMV	Guinea Pig Cytomegalovirus
H	HAART	Highly Active Anti-Retroviral Therapy
	HCMV	Human Cytomegalovirus
	HDAC	Histone Deactylase
	HELF	Human Embryonic Lung Fibroblasts
	HFF	Human Foreskin Fibroblasts
	HFFF-2	Human Foetal Foreskin Fibroblasts
	HHV	Human Herpesvirus
	HIV	Human Immunodeficiency Virus
	HL	Hodgkin's Lymphoma
	HMVEC	Human Microvascular Endothelial Cells
	HSV	Herpes Simplex Virus
	HUVEC	Human Umbilical Vein Endothelial Cells
I	IE	Immediate-Early
	IFN	Interferon
	IRF	Interferon Response Factor
	ISG	Interferon Stimulated Genes
	ISGF	Interferon Stimulated Gene Factor
	ISRE	Interferon Stimulated Response Element
J	JAK	Janus Kinase
	JPEG	Joint Photographic Experts Group
K	kb	Kilo-Base
	kbp	Kilo-Base Pair
	kDa	Kilo-Dalton
	KSHV	Kaposi's Sarcoma-Associated Herpesvirus
L	L	Late
	LDH	Lactate Dehyrogenase
	LPD	Lymphoproliferative Disorders
M	MAPK	Mitogen Activated Protein Kinase

	MCMV	Murine Cytomegalovirus
	MCP	Major Capsid Protein
	mCP	Minor Capsid Protein
	mC-BP	Minor Capsid Binding Protein
	MDM	Monocyte Derived Macrophage
	MEK	MAPK/ERK Kinase
	MHC	Major Hitocompatibility Complex
	MHV-68	Murine Gamma Herpesvirus-68
	MIE	Major Immediate-Early
	MIEP	Major Immediate-Early Promoter
	MKK	MAP Kinase Kinase
	m.o.i.	Multiplicity of Infection
	MOPS	3-(N-Morpholino) Propanesulfonic Acid
	mRNA	Messenger Ribonucleic Acid
	MTC	Multiple Testing Correction
N	NCBI	National Centre for Biotechnology Information
	NF-κB	Nuclear Factor-kappa B
	NHDF	Neonatal Human Dermal Fibroblasts
	NHL	Non-Hodgkin Lymphoma
	NIEP	Non-Infectious Enveloped Particles
	NK	Natural Killer
O	ORF	Open Reading Frame
P	PAA	Phosphonoacetic Acid
	PAL	Pyothorax-Associated Lymphoma
	PBMC	Peripheral Blood Mononuclear Lymphocytes
	PBS	Phosphate Buffered Saline
	PCR	Polymerase Chain Reaction
	PEL	Primary Effusion Lymphoma
	p.f.u.	Plaque Forming Units
	PI	Post Infection
	PI3-K	Phosphatidylinositol-3-Kinase
	PKR	Protein Kinase R
	PLC	Phospholipase C
	PML	Promyelocytic Leukaemia
	PMNL	Polymorphonuclear Leukocyte
	PMT	Photomultplier Tube
	POD	Promyelocytic Protein-Oncogenic Domain
	PRV	Pseudorabies Virus
R	Rb	Retinoblastoma

	REF	Rat Embryonic Fibroblasts
	RhCMV	Rhesus Cytomegalovirus
	RNA	Ribonucleic Acid
	rRNA	Ribosomal Ribonucleic Acid
	RNR	Ribonucleotide Reductase
	ROC	Receiver Operating Characteristic
	RPE	Retinal Pigmented Epithelia
	RT	Room Temperature
	RT-PCR	Real-Time Polymerase Chain Reaction
S	SCID	Severe Combined Immune Deficient
	SCP	Small Capsid Protein
	SDS	Sodium Dodecyl Sulphate
	SP-1	Stimulatory Protein
	SSC	Sodium Chloride/Sodium Citrate Buffer
	STAT	Signal Transducers and Activators of Transcription
	SUMO	Small Ubiquitin-Related Modifier
T	TBP	TATA Binding Protein
	TIFF	Tagged Image File Format
	TNF	Tumour Necrosis Factor
	TNFR	Tumour Necrosis Factor Receptor
	tRNA	Transfer Ribonucleic Acid
	TSP	Thrombospondin
U	UBF	Upstream Binding Factor
V	vICA	Viral Inhibitor of Caspase-8 Induced Apoptosis
	vMIA	Viral Mitochondria-Localised Inhibitor of Apoptosis
	VZV	Varicella-Zoster Virus

1 INTRODUCTION

1.1 The clinical problem

The human cytomegalovirus (HCMV) is a member of the *Herpesviridae*, and is the prototype virus representing the *Betaherpesvirinae*, which includes two other human viruses, HHV-6 and HHV-7 (Roizman and Pellett, 2001). In developed countries, the prevalence of HCMV infection is estimated at 40 % to 60 % of the population, but the incidence increases in the developing world and in the lower socioeconomic groups of industrialised societies (Pass, 2001). Transmission of HCMV requires direct contact with persons excreting virus in fluids such as saliva, urine, tears, semen and cervical secretions. It is estimated that 1 % to 4 % of babies born in the United States and Europe are infected with HCMV; symptomatic disease can result in mental retardation, blindness and sensorineural deafness (Fisher et al., 2000). Mother to foetus/neonate transmission of HCMV is an important factor for maintaining infection within the population, as virus can be transmitted via the placenta, during birth, and in human milk (Pass, 2001). Breast-feeding of neonates has been reported to be the most common route of mother-to-infant transmission (Dworsky et al., 1983). HCMV is also the leading opportunistic pathogen in immunosuppressed patients, specifically allograft recipients and patients in the advanced stages of AIDS. The severity of CMV infection in allograft recipients depends on whether HCMV infection was acquired from the donor, the type of organ transplanted and the degree of immunosuppression; bone marrow transplant recipients are particularly vulnerable to CMV-related pneumonia (Boeckh and Bowden, 1995). Clinical presentations of HCMV infection in AIDS patients include retinitis (blindness), oesophagitis and colitis (Cheung and Teich, 1999). However, the introduction of highly active anti-retroviral therapy (HAART) has resulted in a > 80 % reduction of cytomegalovirus disease in AIDS patients, due to the preservation of T-cells and associated cell-mediated immunity (Salmon-Ceron, 2001).

Development of a vaccine for the prevention of maternal and congenital cytomegalovirus infection has not been successful. When compared to clinical isolates, the live attenuated vaccine strain, Towne, displayed phenotypic changes such as altered growth characteristics *in vitro* and increased trypsin resistance (Plotkin et al., 1975). Immunisation of individuals with the Towne strain induced both humoral and cell-mediated immune responses to cytomegalovirus (Starr et al., 1981). However, when challenged with the low-passage Toledo strain (wild-type), volunteers inoculated with the Towne vaccine candidate exhibited only limited protective responses (Plotkin et al., 1989). In order to

develop a more potent immunogenic vaccine, recombinant Towne/Toledo chimeric vaccine strains have been generated. Phase I clinical trials with Towne/Toledo chimeric vaccine candidates were well-tolerated in HCMV-seropositive individuals but the vaccines failed to boost humoral or cell-mediated immune responses to HCMV, and the immunogenicity of the vaccine remains to be tested in HCMV-seronegative persons (Heineman et al., 2006). The development of subunit vaccines as an alternative to live virus candidates has focused on the viral glycoprotein, gB, a major antigen for the induction of neutralising antibodies (Speckner et al., 1999). Phase I clinical trials of gB subunit vaccines have been successful in eliciting neutralising antibody responses (Frey et al., 1999), but the need for a vaccine that can stimulate sufficient humoral and cell-mediated immunity remains (Gonczol and Plotkin, 2001).

Antiviral drugs currently licensed for the treatment of cytomegalovirus disease are significantly toxic. Ganciclovir, a nucleoside analogue and derivative of acyclovir, is recommended for prophylactic treatment of adult AIDS patients who are CMV seropositive, and have CD4+ T-lymphocyte counts of < 50/ml. However, treatment is not completely effective and ganciclovir is linked to myelotoxicity. In transplant recipients, foscarnet is administered as an alternative to ganciclovir, specifically in cases where bone marrow suppression is to be avoided; however, foscarnet has renal toxicity. Ganciclovir and foscarnet are reported to exhibit similar levels of effectiveness when used in preemptive therapy, where levels of viremia, detection of pp65 antigenemia and detection of CMV genomic DNA in the blood are used to monitor infection. Cidofovir is also licensed for the treatment of cytomegalovirus disease, but has renal toxicity and is linked to neutropenia (Pass, 2001).

The development of vaccines and more effective anti-HCMV drugs has proved difficult due to the lack of an animal model system resulting from the strict species specificity exhibited by cytomegaloviruses. Murine cytomegalovirus (MCMV), however, has been used as a surrogate system in order to draw parallels for HCMV transmission, infection and/or reactivation after immunosuppression and transfusion, but not for the study of foetal infection as MCMV, unlike HCMV, does not cross the placenta (Xiao et al., 2000). The Guinea pig cytomegalovirus (GPCMV) is reported to be a more accurate model for HCMV foetal infection, since GPCMV is able to cross the placental barrier, and shares biological characteristics with HCMV (McGregor and Schleiss, 2001). More recently, a rhesus macaque model of rhesus cytomegalovirus (RhCMV) infection *in utero* has been used to draw parallels with HCMV infection, as RhCMV crosses the placental barrier and

exhibits a similar pathogenesis to that of HCMV in immunocompetent and immunocompromised individuals (Lockbridge et al., 1999).

1.2 Structural components of HCMV particles

HCMV-infected cells in culture produce three different types of virus particles: virions, non-infectious enveloped particles (NIEP), and dense bodies. The electron-dense DNA core, containing the viral genome present in virions is absent in NIEPs, while dense bodies lack both the nucleocapsid and its DNA core, and are predominantly composed of a single tegument protein, pp65 (pUL83) (Irmiere and Gibson, 1983). It is not known whether the production of dense bodies is an artefact of the *in vitro* culture system, as it has been suggested that excess viral products might be stored in vacuoles and deposited into the extracellular medium (Severi et al., 1992). Neither is it known whether NIEPs are produced *in vivo*.

Several different types of nucleocapsid structure are produced in HCMV-infected cells (Type A, B and C). Type A capsids are dead-end products of abortive DNA packaging and lack the viral DNA genome, B capsids are precursors of fully mature particles and contain the scaffolding proteins, but lack the DNA genome, while C capsids are mature DNA containing nucleocapsids (Gibson, 1996). Fully mature herpesvirus particles contain a double-stranded DNA genome that is packaged into an icosahedral-shaped capsid that is surrounded by an amorphous protein layer known as the tegument, which is then enclosed in a lipid bilayer that is decorated with glycoprotein spikes (Roizman and Pellett, 2001). Previous studies estimated that the HCMV particle contained ~ 30 viral proteins, but a more recent assessment of HCMV particles identified 71 virus-encoded proteins (Varnum et al., 2004).

1.2.1 The capsid

All herpesvirus capsids are composed of 150 hexon capsomeres, 12 penton capsomeres and 320 triplex structures that bind the capsomeres together. The particle is organised on a $T = 16$ icosahedral lattice; however, aspects of capsid structure can vary between members of the herpesvirus family. HCMV B capsids (precursors of fully mature C capsids) have a diameter of 130 nm, compared to 125 nm in HSV-1, resulting in a 17 % increase in the overall capsid volume. Such differences in capsid size between HCMV and HSV-1 have been attributed to the spacing between hexamers and their relative tilt, the morphology of the hexon tips and the average diameter of the scaffold (Butcher et al.,

1998). The HCMV genome is ~ 50 % larger than HSV-1, thus a 17% increase in the capsid volume would not appear sufficient to allow packaging of the HCMV genome. However, the HCMV genome is packaged at a higher density, with an average inter-layer spacing of 23 Å compared to 26 Å in HSV-1 (Bhella et al., 2000).

The HCMV capsid contains four structural proteins: the major capsid protein (MCP; pUL86), the minor capsid protein (mCP; pUL85), the minor capsid binding protein (mC-BP; pUL46) and the small capsid protein (SCP; pUL48A) (Mocarski and Tan Courcelle, 2001). The MCP in HCMV (VP5 in HSV-1) forms the hexons (capsid faces) and pentons (capsid vertices), which are connected by the triplex proteins, themselves formed by the mCP (VP23 in HSV-1) and the mC-BP (VP19c in HSV-1). The SCPs share the least sequence similarity among the HCMV and HSV-1 capsid proteins. It has been reported that the SCP is essential for HCMV infection *in vivo* (Borst et al., 2001), but its homologue in HSV-1 (VP26) is dispensable for the formation of stable capsids in tissue culture (Desai et al., 1998). In HSV-1, VP26 is located at the hexon tips and was originally thought to have a role in directing tegument attachment to the capsid (Zhou et al., 1995), but it was later demonstrated that VP26 minus virions and wild-type virions are identical in tegument attachment patterns (Chen et al., 2001). Despite the sequence divergence between the HCMV SCP and HSV-1 VP26, it has been shown that SCPs also decorate the tips of the MCP hexon subunits (Yu et al., 2005). Tegument attachment to the capsid has not been fully resolved in either HCMV or HSV-1, but involves attachment to the capsid pentons, hexons and triplexes in HCMV (Chen et al., 1999), and the pentons in HSV-1 (Zhou et al., 1999). The capsid itself is constructed around a scaffold consisting of the HCMV assembly protein (pUL80A), and three other scaffold proteins that are all generated by auto-proteolytically cleaved pUL80, which contains protease activity in the amino-terminal half. Following construction of the capsid, the scaffold proteins are cleaved in order to allow them to vacate the nucleocapsid, so that the genomic DNA can then be packaged (Gibson, 1996; Welch et al., 1991).

1.2.2 The envelope

HCMV particles have a lipid bilayer envelope, which is derived from host intracellular membranes into which viral glycoproteins are embedded. The acquisition of the envelope is thought to occur in two separate stages, for which an envelopment-deenvelopment-reenvelopment model has been proposed (Mettenleiter, 2002). The first envelopment stage occurs by budding of DNA containing HCMV C capsids at the inner nuclear membrane into the lamellar space, followed by deenvelopment by fusion with the outer nuclear

membrane and then release of the naked capsids into the cytoplasm (Buser et al., 2007). Tegument proteins are thought to be added to the capsid in the cytoplasm, although a few may be added earlier when in the nucleus. Most tegument proteins are thought to aggregate on the cytoplasmic surface of membranes derived from the trans-golgi network, and which also contain viral envelope glycoproteins. Budding of the naked virus particle into the vacuole generates the tegumented and enveloped virus particle contained within a transport vesicle (Sanchez et al., 2000). However, this model is not universally accepted as two diverse pathways have been identified in HSV-1 envelopment. The first pathway is based on the envelopment-deenvelopment-reenvelopment model as mentioned above. The second pathway involves direct access of the capsids from the nucleus to the cytoplasm through impaired nuclear pores, and the capsids are then enveloped by budding from membranes derived from the rough endoplasmic reticulum (RER), trans-golgi apparatus, or large intracellular vacuoles (Leuzinger et al., 2005).

Approximately 50 glycoproteins are thought to be encoded by HCMV, but the majority have not yet been studied, and it is not clear which are expressed on the host cell surface and/or in the virion envelope (Mocarski and Tan Courcelle, 2001). Three major glycoprotein complexes are found in HCMV particles. Glycoprotein B (gpUL55) is a type 1 transmembrane protein that forms a homodimeric glycoprotein complex termed gcI. This complex binds to heparan sulphate proteoglycans to facilitate entry into host cells, and is also required for cell-to-cell spread of infection (Navarro et al., 1993; Boyle and Compton, 1998). Sequence analysis reveals that four gB subtypes exist, and that sequence variation is at its highest between codons 448 to 480 (Chou and Dennison, 1991). Such sequence variation is probably linked to immune evasion, as nearly all individuals develop anti-gB neutralising antibodies, reaching a neutralisation efficiency of 50 % at most (Speckner et al., 1999).

Glycoprotein complex gcII is composed of gM (gpUL100) and gN (gpUL73), which forms a gM/gN disulphide linked dimer that exhibits heparin-binding activity, thus having a role in virus entry and membrane fusion (Kari et al., 1992; Mach et al., 2000). These genes are essential for HCMV replication in fibroblasts as expression of one without the other leads to altered glycosylation and/or distribution phenotypes, resulting in replication deficient viruses (Hobom et al., 2000; Mach et al., 2000). Recently, it was demonstrated that the disulphide bond between gM and gN molecules was not required for complex formation or translocation of the gCII complex into the assembly compartments (Mach et al., 2005). However, disruption of the C-terminal domain of gN led to a reduction of

secondary envelopment of HCMV capsids, suggesting a critical role for gN in mediating this process (Mach et al., 2007).

Glycoprotein complex gcIII is composed of gH (gpUL75), gL (gpUL115) and gO (gpUL74), and is required for viral entry into host cells, since antibodies that mimic the gH/gL complex and bind to the same cell surface receptor as gCIII prevent HCMV from penetrating the cell membrane (Keay et al., 1991). Proteins that complex with gH/gL in EBV have cell tropism functions, and it is thought that HCMV gO might have a similar role (Mocarski and Tan Courcelle, 2001). Binding of gB or gH to cell surface receptors induces intracellular signalling that results in the expression of the cellular transcription factors SP1 and NF-κB (Yurochko et al., 1997).

1.2.3 The tegument

A recent assessment of the HCMV particle proteome identified 14 tegument proteins (Varnum et al., 2004). The tegument is a complex multifunctional structural component of the virus particle. Tegument proteins have a diverse range of functions that include: particle architecture, virus egress, trans-activation of immediate-early gene expression, and preparation of the cellular environment for lytic replication (see Table 1.1).

The most abundant tegument proteins are pp65 (pUL83) and pp150 (pUL32), comprising ~ 18 % and ~ 20 % of the total virion protein mass respectively (Irmiere and Gibson, 1983; Zipeto et al., 1993). Despite its abundance, pp65 is dispensable for growth in tissue culture (Schmolke et al., 1995). After infection, particle-delivered pp65 localises to the nucleus, but its function at this very early stage remains unknown. Expression of the UL83 (pp65) gene occurs at early-late times PI, and nascent pp65 is localised in the nucleus throughout the replication cycle, but at late times PI, is also accumulated in the cytoplasm (Hensel et al., 1995). It has been shown that pp65 has serine/threonine protein kinase activity capable of both autophosphorylation and phosphorylation of IE1 (Gilbert et al., 1996). It was also reported that pp65 prevents cleavage of IE1, therefore inhibiting the presentation of IE1 peptides via MHC class I (Gilbert et al., 1996). The association of pp65 with a cellular protein, Polo-like kinase (PIK1) might have effects on intracellular signalling, sub-cellular location and substrate specificity (Gallina et al., 1999).

The basic-phosphoprotein, pp150 is an immunodominant protein for which 85 % to 90 % of HCMV seropositive individuals have antibodies (Mocarski and Tan Courcelle, 2001). pp150 is modified by O-linked N-acetylglucosamine, a modification associated with

nuclear localisation. Initial investigations reported that like pp65, pp150 was localised in the nucleus at early times PI, but was found in the cytoplasm only at late times PI (Hensel et al., 1995); however, other workers only detected pp150 exclusively in the cytoplasm (Sanchez et al., 2000). As pp150 binds the MCP (Baxter and Gibson, 2001), it is possible that pp150 is involved in tegument attachment.

Other tegument proteins include the transactivators pp71 (pUL82), pUL69, pTRS1/pIRS1, which activate the virus immediate-early (IE) promoter/enhancer and other virus and cellular promoters. The tegument is also reported to contain multiple members of the US22 gene family: pUL23, pUL24, pUL43, pUS22 (Adair et al., 2002); pUL36, pTRS1/pIRS1 (Mocarski and Tan Courcelle, 2001); pUS23 and pUS24 (Varnum et al., 2004; Feng et al., 2006). pUL99 (pp28) is also a tegument protein and was reported to be essential for virus replication and envelopment of HCMV capsids (Silva et al., 2003). Together, these suggest a complex role for tegument proteins from initiating lytic infection, evading immune defences and/or modifying the cellular environment, to maturation and egress of virus particles.

1.3 The HCMV genome and its coding potential

HCMV is the largest member of the herpesvirus family with a linear double-stranded DNA genome of 235646 bp (strain Merlin; GenBank accession number AY446894). The genome is organised into two segments, designated UL and US, which are flanked by inverted repeat sequences (TRL/IRL and IRS/TRS) allowing the formation of four genome isomers: to date, two isomers have been shown to be viable (Mocarski and Tan Courcelle, 2001; Borst et al., 1999). HCMV gene predictions originally based on the AD169 laboratory strain predicted 208 ORFs of more than 100 amino acids (Chee et al., 1990) These ORFs were given prefixes of TRL, UL, IRL, IRS, US or TRS depending on the location of the ORF within the genome. However, the AD169 and Towne strain genomes lack 15 kbp and 12 kbp of sequence respectively, which are present in the UL/b' region of the genome in clinical isolates and the low passage Toledo strain. The 15 kb DNA sequence lacking in AD169 was reported to contain at least 19 ORFs. Moreover, the Toledo strain differs from clinical isolates, in that sequences from the Toledo UL/b' region are present in inverted orientation (Cha et al., 1996). Some stocks of AD169 (i.e. the sequenced AD169 genotype published by Chee et al., (1990)) also lack 929 bp of additional sequence from the UL42/UL43 region that are present in other laboratory stocks worldwide (Dargan et al., 1997). DNA sequence comparison between HCMV (AD169 plus the missing 15 kb fragment from the Toledo strain) and the chimpanzee cytomegalovirus

(CCMV) was undertaken to assess the number of HCMV protein coding ORFs. This study discounted 51 ORFs, predicted 10 new ORFs and modified the coding of a further 24 ORFs. It was therefore predicted that HCMV clinical isolates (i.e. wild-type), encodes 165 ORFs (Davison et al., 2003). It was later demonstrated that the low passage HCMV strain Merlin accurately reflected the wild-type complement of 165 genes, except for a point mutation in gene UL128 leading to a truncated protein (Dolan et al., 2004). Thirteen different gene families arising from gene duplication events during evolution have been identified: RL11, UL25, US1, GPCR (G-protein coupled receptor), UL14, UL82, US2, UL18, UL120, US6, US22, UL146 and US12. In addition, 40 herpesvirus-common 'core' genes are located centrally within UL, and have homologues in all members of the herpesviruses (Davison et al., 2003). A detailed map of the HCMV strain Merlin genome is shown in Fig. 1.1. Note that the genes are colour-coded in accordance with their respective gene family etc. A table of the Merlin ORFs, their known or predicted function and their homologues in CCMV and MCMV is shown in Table 1.1.

A different analysis of AD169 coding potential was undertaken using an *in silico* based approach. The Bio-Dictionary Gene Finder (BDGF) algorithm was used to interrogate the AD169 genome and the additional 15 kb sequence present in the Toledo strain. This analysis discounted 37 ORFs predicted by Chee et al. (1990) and identified a further 12 new ORFs (Table 1.2). It was therefore estimated that wild-type HCMV isolates encode 192 unique ORFs (Murphy et al., 2003a). Murphy et al. then set out to compare the coding potential of clinical and laboratory strains of HCMV. They sequenced the genomes of six HCMV isolates that had been cloned in the form of bacterial artificial chromosomes (BAC), two of which were from the laboratory strains (AD169 and Towne), one from the low passage laboratory strain (Toledo), while the others were prepared from clinical isolates (designated as PH, TR and FIX). A database was generated for the ORFs in the six HCMV genomes that encode polypeptides of ≥ 80 amino acids. The BLASTp algorithm was used to identify ORFs in all four clinical isolates that had not been previously recognised, and subjected them to additional filters: the presence of an AUG start codon, and an overlap of < 396 bp. This method identified 29 previously unrecognised ORFs (Murphy et al., 2003b) (Table 1.2). However, there is no evidence to suggest that any of the novel ORFs proposed by Murphy et al. are protein coding.

1.3.1 HCMV transcript mapping

Like other herpesviruses, HCMV has a genome that is densely packed with ORFs, but contains relatively few polyadenylation signals required for transcription termination and

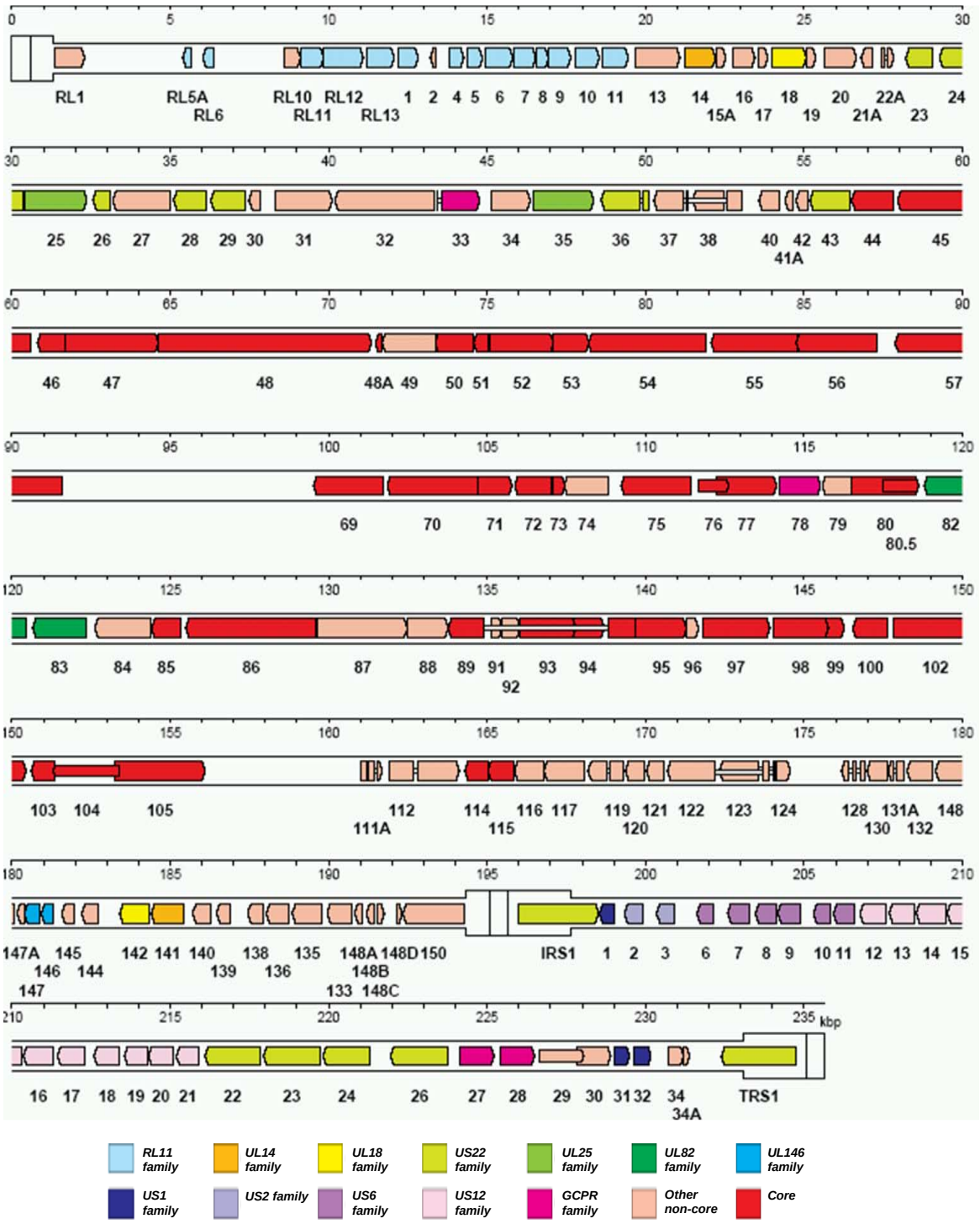


Figure 1.1. Gene map of the Merlin strain of HCMV (Modified from Dolan et al., 2004)

Gene families are grouped accordingly, those coloured red represent the herpesvirus core genes (located centrally within the genome) and code for proteins involved in DNA replication, and structural proteins. The terminal regions of the

HCMV	Function	MCMV	CCMV
RL1			RL1
RL5A	RL11 family		
RL6	RL11 family		
RL10	Virion envelope glycoprotein		RL10
RL11	RL11 family; IgG Fc-binding membrane glycoprotein		RL11
RL12	RL11 family; putative membrane glycoprotein		RL12
RL13	RL11 family; putative membrane glycoprotein		RL13
UL1	RL11 family; putative membrane glycoprotein		
UL2	Putative membrane protein		UL2
UL4	RL11 family; virion glycoprotein		UL4
UL5	RL11 family; putative membrane glycoprotein		UL5
UL6	RL11 family; putative membrane glycoprotein		UL6
UL7	RL11 family; putative membrane glycoprotein		UL7
UL9	RL11 family; putative membrane glycoprotein		UL9
UL10	RL11 family; putative membrane glycoprotein		UL10
UL11	RL11 family; membrane glycoprotein		UL11
UL13	Putative secreted protein		UL13
UL14	UL14 family; putative membrane glycoprotein		UL14
UL15A	Putative membrane protein		UL15A
UL16	Membrane glycoprotein; inhibits NK cell cytotoxicity		UL16
UL17			UL17
UL18	UL18 family; putative membrane glycoprotein; MHC-I homologue; possibly inhibits NK cell cytoxicity		UL18
UL19			UL19
UL20	Putative membrane glycoprotein		UL20
UL21A			UL21A
UL22A	Putative secreted glycoprotein		UL22A
UL23	US22 family; tegument protein	M23	UL23
UL24	US22 family; tegument protein	M24	UL24
UL25	UL25 family; tegument phosphoprotein	M25	UL25
UL26	US22 family; tegument protein; transcriptional activator of MIEP	M26	UL26
UL27		M27	UL27
UL28	US22 family (spliced)	M28	UL28
UL29	US22 family	M29	UL29
UL30		M30	UL30
UL31		M31	UL31
UL32	Major tegument phopshoglycoprotein (pp150)l; very immunogenic; binds to capsids	M32	UL32
UL33	GPCR family; membrane protein; putative chemokine receptor ; virion protein	M33	UL33
UL34	Represses US3 transcription	M34	UL34
UL35	UL25 family; tegument phosphoprotein; interacts with UL82 protein	M35	UL35
UL36	US22 family; immediate early tegument protein; inhibitor of caspase-8-induced apoptosis (vCIA)	M36	UL36
UL37	Immediate-early envelope glycoprotein; possible auxilliary role in DNA replication; exon 1 product is mitochondrial inhibitor of apoptosis (vMIA)	M37	UL37
UL38		M38	UL38
UL40	Putative membrane glycoprotein; inhibits NK cell cytotoxicity		UL40
UL41A	Putative membrane protein		UL41A
UL42	Putative membrane protein		UL42
UL43	US22 family; tegument protein	M43	UL43
UL44	Processivity subunit of DNA polymerase (ICP36)	M44	UL44
UL45	Large subunit of ribonucleotide reductase; tegument protein	M45	UL45
UL46	Minor capsid-binding protein (mC-BP)	M46	UL46
U47	Tegument protein; possible role in intracellular transport; binds to UL48 protein	M47	U47
UL48	High molecular weight tegument protein; binds to UL47	M48	UL48
UL48A	Small capsid protein (SCP)		UL48A
UL49		M49	UL49
UL50	Inner nuclear membrane protein; role in egress of capsids from nucleus	M50	UL50
UL51	Role in DNA packaging	M51	UL51
UL52	Role in DNA packaging	M52	UL52
UL53	Nuclear matrix protein; tegument protein; role in egress	M53	UL53
UL54	Catalytic subunit of DNA polymerase	M54	UL54
UL55	Virion glycoprotein gB; component of gCI	M55	UL55
UL56	Putative subunit of terminase; exhibits nuclease activity	M56	UL56
UL57	Single-stranded DNA-binding protein	M57	UL57
UL69	Regulatory protein; tegument	M69	UL69

HCMV	Function	MCMV	CCMV
UL70	Component of DNA helicase-primase complex	M70	UL70
UL71	Putative membrane protein	M71	UL71
UL72	Derived from deoxyuridine triphosphatase; enzymatically inactive	M72	UL72
UL73	Virion glycoprotein; gN; component of gCII	M73	UL73
UL74	Virion glycoprotein; gO,; component of gCIII		UL74
UL75	Virion glycoprotein; gH; component of gCIII	M75	UL75
UL76	Virion-associated regulatory protein	M76	UL76
UL77	Role in DNA packaging	M77	UL77
UL78	GPCR family; putative chemokine receptor	M78	UL78
UL79		M79	UL79
UL80	Protease (N terminus) and minor capsids scaffold protein (C terminus)	M80	UL80
UL80.5	Major capsids scaffold protein		UL80.5
UL82	UL82 family; (pp71; upper matrix protein); transcriptional activator	M82	UL82
UL83	UL82 family; tegument phosphoprotein (pp65; lower matrix protein) suppresses interferon response	M83	UL83
UL84	UL82 family; role in DNA replication; exhibits nucleocytoplasmic shuttling; transdominant inhibitor of IE transcription	M83	UL83
UL85	Minor capsid protein (mCP)	M85	UL85
UL86	Major capsid protein (MCP)	M86	UL86
UL87		M87	UL87
UL88	Tegument protein	M88	UL88
UL89	Putative ATPase subunit of terminase	M89	UL89
UL91		M91	UL91
UL92		M92	UL92
UL93	Role in DNA packaging; tegument protein	M93	UL93
UL94	Tegument protein; binds single-stranded DNA	M94	UL94
UL95		M95	UL95
UL96	Tegument protein	M96	UL96
UL97	Serine-threonine protein kinase; tegument protein; roles in DNA synthesis, DNA packaging and nuclear egress; phosphorylates UL44 protein	M97	UL97
UL98	Deoxyribonuclease	M98	UL98
UL99	Myristylated tegument phosphoprotein (pp28)	M99	UL99
UL100	Virion glycoprotein; gM; component of gCII	M100	UL100
UL102	Component of DNA helicase-primase complex	M102	UL102
UL103	Tegument protein	M103	UL103
UL104	Portal protein; possibly interacts with terminase	M104	UL104
UL105	Component of DNA helicase-primase complex; helicase	M105	UL105
UL111A	Viral interleukin 10 (vIL-10)		
UL112	Role in transcriptional activation or organising DNA replication proteins	M112	UL112
UL114	Uracil-DNA glycosylase; roles in excision of uracil from DNA and temporal regulation	M114	UL114
UL115	Virion envelope glycoprotein; gL; component of gCIII	M115	UL115
UL116	Putative membrane glycoprotein	M116	UL116
UL117	CC chemokine		UL117
UL119	IgG Fc-binding membrane glycoprotein related to COX-2	M119	UL119
UL120	UL120 family; putative membrane glycoprotein	M120	UL120
UL121	UL120 family; putative membrane glycoprotein	M121	UL121
UL122	Immediate-early transcriptional activator (IE2); interacts with basal transcription machinery and cellular transcription factors; specific DNA-binding protein	M122	UL122
UL123	Immediate-early transcriptional activator (IE1); enhances activation by IE2; interacts with basal transcriptional machinery; disrupts ND10	M123	UL123
UL124	Putative membrane glycoprotein		UL124
UL128	Putative secreted protein; putative CC chemokine		UL128
UL130	Putative secreted protein		UL130
UL131A	Putative secreted protein		UL131A
UL132	Putative membrane protein		UL132
UL148	Putative membrane protein		UL148
UL147A	Putative membrane protein		UL147A
UL147	UL146 family; putative secreted glycoprotein; putative CXC chemokine		UL147
UL146	UL146 family; secreted glycoprotein; CXC chemokine		UL146
UL145			UL145
UL144	Putative membrane glycoprotein; TNF receptor homologue		UL144
UL142	UL18 family; putative membrane glycoprotein; MHC-I homologue		UL142
UL141	UL14 family; membrane glycoprotein; inhibits NK cell cytoxicity by downregulating CD155		UL141
UL140	Putative membrane protein		UL140
UL139	Putative membrane glycoprotein		UL139
UL138	Putative membrane protein		UL138

HCMV	Function	MCMV	CCMV
UL136	Putative membrane protein		UL136
UL135	Putative secreted protein		UL135
UL133	Putative membrane protein		UL133
UL148A	Putative membrane protein		UL148A
UL148B	Putative membrane protein		UL148B
UL148C	Putative membrane protein		UL148C
UL148D	Putative membrane protein		UL148D
UL150	Putative secreted protein		UL150
IRS1	US22 family; immediate early transcriptional activator; tegument protein; involved in shutoff of host protein synthesis		IRS1
US1	US1 family		US1
US2	US2 family; membrane glycoprotein; selective degradation of MHC-I and MHC-II		US2
US3	US2 family; immediate-early gene; membrane glycoprotein; inhibits processing and transport of MHC-I and MHC-II		US3
US6	US6 family; putative membrane glycoprotein; inhibits TAP-mediated peptide transport		US6
US7	US6 family; membrane glycoprotein		US7
US8	US6 family; membrane glycoprotein; binds to MHC-I		US8
US9	US6 family; membrane glycoprotein		US9
US10	US6 family; membrane glycoprotein; delays trafficking of MHC-I		US10
US11	US6 family; membrane glycoprotein; selective degradation of MHC-I		US11
US12	US12 family; putative multiple transmembrane protein		US12
US13	US12 family; putative multiple transmembrane protein		US13
US14	US12 family; putative multiple transmembrane protein		US14
US15	US12 family; putative multiple transmembrane protein		US15
US16	US12 family; putative multiple transmembrane protein		US16
US17	US12 family; putative multiple transmembrane protein		US17
US18	US12 family; putative multiple transmembrane protein		US18
US19	US12 family; putative multiple transmembrane protein		US19
US20	US12 family; putative multiple transmembrane protein		US20
US21	US12 family; putative multiple transmembrane protein		US21
US22	US22 family; tegument protein; released from cells	m128Ex3 m139	US22
US23	US22 family	m140 m143	US23
US24	US22 family	m141	US24
US26	US22 family	m142	US26
US27	GPCR family; membrane protein		US27
US28	GPCR family; membrane protein; broad spectrum CC chemokine receptor; mediates cellular migration		US28
US29	Putative membrane glycoprotein		US29
US30	Putative membrane glycoprotein		US30
US31	US1 family		US31
US32	US1 family		US32
US34	Putative secreted protein		US34
US34A	Putative membrane protein		US34A
TRS1	US22 family; immediate-early transcriptional activator; tegument protein; involved in shutoff of host protein synthesis and capsids assembly		TRS1

Table 1.1. HCMV ORFs in strain Merlin, and associated homologues in MCMV and CCMV.

The table shows the ORFs in HCMV strain Merlin (Dolan et al., 2004), with the associated homologues in MCMV (Rawlinson et al., 1996) and in CCMV (Davison et al., 2003). The MCMV genome is not yet fully annotated.

ORF	Merlin (Genome Position)		Strand	Reference
ORF1	19446	19696	R	Murphy et al.2003a
ORF2	37178	37492	R	
ORF3	95417	95730	F	
ORF4	95470	95777	F	
ORF5	96843	97081	F	
ORF6	134719	135090	F	
ORF7	146369	146897	F	
ORF8	146369	146563	R	
ORF9	171496	172179	F	
ORF10	171872	172150	F	
ORF11	3729	3996	F	
ORF12	234784	234946	R	
C-ORF1	2103	2411	F	Murphy et al.2003b
C-ORF2	2505	2194	R	
C-ORF3	3656	3393	R	
C-ORF4	7647	7919	F	
C-ORF5	9143	8820	R	
C-ORF6	23893	23636	R	
C-ORF7	29239	29631	F	
C-ORF8	35007	35282	F	
C-ORF9	35626	36045	R	
C-ORF10	43175	43519	F	
C-ORF11	46367	46711	R	
C-ORF12	53935	54294	R	
C-ORF13	54728	54988	F	
C-ORF14	54901	55518	F	
C-ORF15	120492	120752	R	
C-ORF16	157314	157583	R	
C-ORF17	159230	159547	F	
C-ORF18	161724	161975	R	
C-ORF19				
C-ORF20	166444	166824	F	
C-ORF21	168066	168470	F	
C-ORF22	168531	168797	F	
C-ORF23	176123	176401	F	
C-ORF24	185685	185047	F	
C-ORF25	190448	190897	R	
C-ORF26	190553	191008	F	
C-ORF27				
C-ORF28	210088	210369	F	
C-ORF29	212116	212409	R	
RL3	3740	3933	F	Chee et al., 1990
RL4	793	472	R	
RL5	4772	5006	F	
RL8	7757	8092	R	
RL9	7920	8348	F	
UL21	26832	27359	R	
UL41	54294	54719	R	
UL60	93829	94130	R	
UL61	94555	95277	R	
UL101	147662	148009	F	
US5	201445	201825	F	
US25	221325	221864	F	
J1S	235125	235262	R	

Table 1.2. Additional ORFs that were contained in the HCMV microarray.
The table shows the Merlin genome coordinates for novel putative ORFs, and previously described ORFs that are now considered to be non-protein coding. Note that the coordinates for C-ORF19 and C-ORF27 could not be identified.

the subsequent addition of a polyadenylate tail on the free 3'-end of mRNAs. Consequently, many spliced and non-spliced HCMV genes share polyadenylation signals, leading to the generation of 3' co-terminal groups of genes within specific regions of the HCMV genome (Wing and Huang, 1995; Smuda et al., 1997; Guo and Huang, 1993). Some HCMV ORFs also overlap, e.g. UL146/UL147A (Lurain et al., 2006); while others are spliced, e.g. UL122/UL123 (Awasthi et al., 2004) and UL37 (Adair et al., 2003). To illustrate the point, a transcript map for HCMV ORFs is shown in Fig. 1.2. The data in the figure is partly obtained from published sources (green arrows), and partly speculative (black arrows), which does not take into account differential use of promoters and/or polyadenylation signals. A reference list for the mapped HCMV transcripts are shown in Table 1.3.

1.4 The replication cycle

1.4.1 The major stages of the HCMV replication cycle

HCMV is able to infect a wide range of cell types *in vivo* including fibroblasts, epithelial cells, smooth muscle cells and endothelial cells (Sinzger et al., 1995). The virus is disseminated within the host via leukocytes (monocytes and neutrophils) in the peripheral blood (Sinzger and Jahn, 1996). The following description of the HCMV replication cycle has been deduced from studies of infected fibroblast cells. Recent studies with epithelial and endothelial cells have revealed differences in the mechanism of entry and in maturation and release of infectious progeny; the HCMV replication cycle in fibroblasts then serves only as a model system.

Attachment of HCMV particles to heparan sulphate on the host cell surface precedes penetration via pH-independent fusion of the virion envelope and the host cell surface, releasing the viral nucleocapsid into the cytoplasm (Compton et al., 1992). HCMV capsids then associate with microtubules and are transported to the outer surface of the nucleus (Ogawa-Goto et al., 2003), with similar kinetics described for HSV-1 (Sodeik et al., 1997). Viral DNA is then deposited into the nucleus through a nuclear pore. The linear virus DNA circularises in the nucleus and replicates at intranuclear structures, which are associated with the periphery of promyelocytic protein (PML)-oncogenic domains (POD) (also known as ND10) (Penfold and Mocarski, 1997). The periphery of ND10 domains are also the site for the initiation of immediate-early gene expression. The particle delivered tegument protein pp71 (pUL82) and cellular hDaxx are reported to play an important role in HCMV genomic DNA deposition at the ND10 structures, and subsequent replication

HCMV Genes	Reference
UL34	Biegalke et al., 2004
UL35	Liu et al., 2002
UL36, UL37, UL38	Adair et al., 2003
UL44	Leach and Mocarski, 1989
UL47, UL48	Hyun et al., 1999
UL54, UL55, UL56, UL57	Smuda et al., 1997; Kiehl et al., 2003
UL84, UL85, UL86	He et al., 1992
UL93, UL94, UL95, UL96, UL97, UL98, UL99	Wing et al., 1995
UL102	Smith and Pari, 1995
UL103, UL104	Scalzo et al., 1995
UL105	Smith et al., 1996
UL115, UL116, UL117, UL119	Leatham et al., 1991
UL122/UL123	Rawlinson et al., 1993; Awasthi et al., 2004
IRS1, TRS1	Romanowski and Shenk, 1997
US3	Liu et al., 2002
US18, US19, US20	Guo and Huang, 1993
US34	Scott et al., 2002

Table 1.3. References for mapped HCMV transcripts.

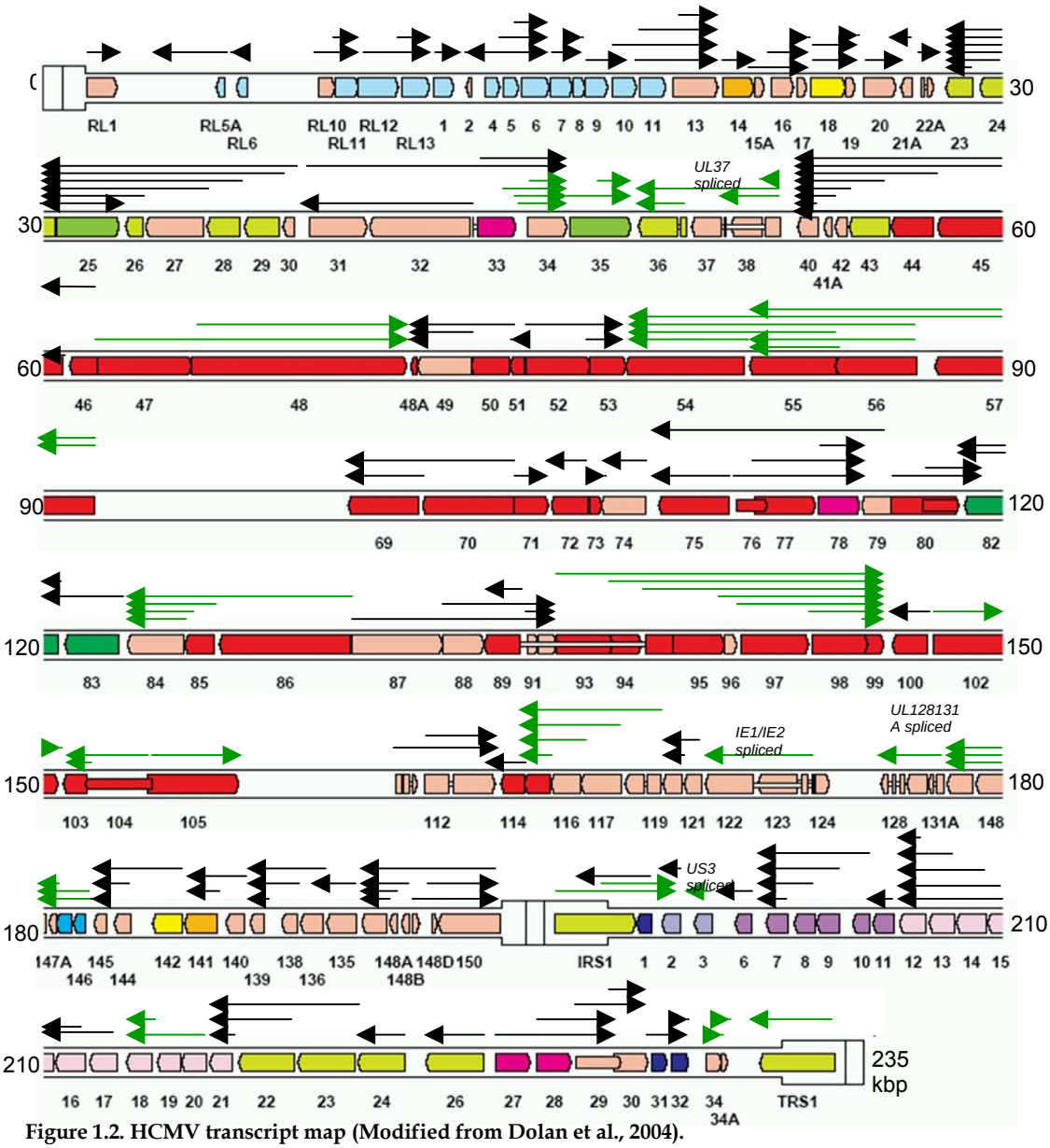


Figure 1.2. HCMV transcript map (Modified from Dolan et al., 2004).
The arrows shown in black are predicted transcripts that are drawn from the start of the ORF to the nearest proximal poly(A). The green arrows represent transcripts that have been mapped and reported in the literature, references are given to these in Table 1.3.

and onset of gene expression (Ishov et al., 1997; Ishov et al., 2002). However, the role of ND10 domains for HCMV DNA replication and the initiation of gene expression are controversial, since recent reports have suggested that ND10 domains form part of the cells intrinsic anti-viral defence mechanisms (Tavalai et al., 2006; Woodhall et al., 2006).

The circular DNA genome is thought to replicate by a 'rolling circle' mechanism leading to the formation of concatemeric DNA, which is cleaved into unit length during packaging (LaFemina and Hayward, 1983; McVoy et al., 2000). Genes required for the replication of HCMV DNA are UL36-UL38, UL44, UL54, UL57, UL70, UL84, UL102, UL105, UL112-113, UL122(IE2)/UL123(IE1), TRS1/IRS1. Initiation of DNA replication occurs at the lytic-phase replicator (ori-Lyt), located centrally within the UL region of the genome. Herpesvirus core genes required for the formation of the core replication machinery are UL54 (DNA polymerase), UL44 (DNA polymerase processivity factor), UL57 (single stranded DNA binding protein), and UL70-UL102-UL105 (helicase-primase complex). The remaining genes listed above are presumed regulatory proteins, as IE1/IE2 and TRS1/IRS1 are transactivators of gene expression. UL36-UL38 expression is also required for ori-Lyt dependent replication in human fibroblasts, but not in Vero cells, and UL36 and UL37 are known inhibitors of apoptosis (Goldmacher et al., 1999; Skaletskaya et al., 2001). UL84 is known to interact with IE2, and activate the ori-Lyt promoter in order initiate DNA replication, although the detailed mechanisms of this process at present remain unknown (Sarisky and Hayward, 1996; Xu et al., 2004a and b). pUL84 also functions as a transdominant inhibitor of IE2 transactivation (discussed in section 1.5.3.2). UL112-113 are thought to recruit replication-fork proteins (Ahn et al., 1999).

Concatemeric virus DNA is cleaved into unit length by the viral terminase activity and packaged into capsids through a portal located at one unique vertex. The portal protein is coded by UL104, and acts as a dock for the terminase enzyme, itself composed of two subunits coded for by UL56 and UL89 (Dittmer et al., 2005). Packaging of the viral DNA into the capsids is a dynamic process that requires the ATPase activity of the terminase enzyme in order to 'pump' the DNA into the capsid. The ATPase activity of the HCMV terminase enzyme is reported to be exclusive to pUL56 (Hwang and Bogner, 2002). However, recent evidence suggests that pUL89 contains an ATPase catalytic site, ATPase motor motifs, a zinc finger, DNA cutting sites and portal binding sites (Champier et al., 2007). Following packaging of the genomic DNA, capsids containing DNA are translocated from the nucleus into the cytoplasm. Tegumentation is thought to occur in the nucleus and in the cytoplasm; the capsids are then transported to the cell surface and released from the cell (see section 1.2).

1.4.2 Kinetic class of gene expression

Herpesvirus genes are transcribed by host cell RNA polymerase II in a regulated temporal cascade; virus genes are subsequently classified into three kinetic classes: immediate-early (IE), early (E), or late (L). The first viral gene products to be made during the replication cycle are expressed in the absence of *de novo* protein synthesis from ~ 1 h PI, and include the major immediate-early transactivators UL123/UL122 (IE1/IE2), together with accessory proteins TRS1/IRS1, UL36-UL38, US3 and UL115-UL119. The major transactivator of HCMV gene expression is IE2 (86 kDa), which acts co-operatively with IE1 (72 kDa) and the accessory proteins to regulate the expression of the early and late genes. Cellular transcription factors such as NF-κB, SP-1, Elk-1, NF-Y and CREB are also utilised to regulate both viral and cellular gene expression (Mocarski and Tan Courcelle, 2001).

Expression of the early genes starts from 4–12 h PI and is dependent on the presence of the immediate-early proteins. The early genes tend to code for non-structural proteins required for DNA replication, packaging and maturation of virus particles. Inhibitors of DNA replication such as phosphonoacetic acid (PAA) or ganciclovir (GCV) do not affect the transcription of early genes, and these drugs can be used to assign genes to the early or late kinetic classes. The late genes code for structural proteins and can be sub-divided into early-late (E-L) and true-late (L) genes. True-late (L) protein synthesis is dependent on DNA replication, whereas proteins for early-late (E-L) genes are expressed in small amounts in the presence, but larger amounts in the absence of DNA synthesis inhibitors (Mocarski and Tan Courcelle, 2001).

1.4.3 HCMV induced effects on intracellular signalling

To date, the HCMV induced effects of intracellular signalling have been studied in fibroblasts only, therefore, the specific effects of HCMV particle binding to other cell types and their effects on intracellular signalling may differ. However, the stimulation of intracellular signalling pathways by HCMV infection of fibroblasts serves as a model. HCMV particles form transient attachments to the surface of host cells via heparin sulphate proteoglycans (HSPG), followed by stable interactions of HCMV glycoproteins gB, gH, and others (see section 1.2.2), to surface receptors such as the epidermal growth factor receptor (EGFR) tyrosine kinase, and possibly other receptor tyrosine kinases (RTK), or the tumour necrosis factor receptor (TNFR), and Toll-like receptor (TLR). Stable attachment of viral glycoproteins to cell surface receptors leads to the activation of the

mitogen-activated protein kinase (MAPK) signalling pathway and the interferon-response pathway. Virus entry into the host cell is also mediated by the interaction of viral glycoproteins with the host cell surface receptors: pH independent membrane fusion (fibroblasts) or endocytosis (epithelial and endothelial cells) (Fortunato et al., 2000; Evers et al., 2004). A simplified diagram showing the HCMV induced effects on intracellular signalling in infected fibroblasts is shown in Fig 1.3.

Stimulation of the MAPK/ERK signalling pathway occurs by the activation of epidermal growth factor receptor (EGFR) tyrosine kinase via binding of HCMV gB, resulting in the recruitment of phosphatidylinositol-3-kinase (PI3-K), which phosphorylates and converts PIP to PIP_2. Activation of EGFR also stimulates phospholipase C (PLC-γ), which hydrolyses PIP_2 to diacylglycerol (DG) and inositol-triphosphate (IP_3). This in turn, leads to the activation of protein kinase C (PKC) via DG, and IP_3 induced release of Ca^{2+}, stimulating further second messenger activity (Evers et al., 2004). Protein kinase B (Akt) is also stimulated downstream of PI3-K signalling (Fig. 1.3), which in turn activates p70S6K (a serine/threonine kinase). Protein kinase B (Akt) and p70S6K both promote cell survival (anti-apoptosis) (Kandel and Hay, 1999). It has been reported that inhibition of PI3-K activity using drugs resulted in approximately 10,000 fold reduction in infectious virus progeny, suggesting an important role of PI3-K signalling for HCMV lytic infection (Johnson et al., 2001b).

The activation of NF-κB occurs by PI3-K phosphorylation of IκB via IκB kinase (IKK), which results in IκB ubiquitin-mediated proteasomal degradation. Since IκB is an inhibitor of NF-κB, this is then free to translocate to the nucleus, where it activates cellular genes and induces the HCMV MIEP, and subsequent expression of HCMV early genes (Evers et al., 2004). Note that induction of NF-κB also occurs downstream of TNFR and TLR signalling. NF-κB activation leads to the expression of cytokines such as TNF-α and Il-1β, thus inducing the innate anti-viral defence programme. However, it has been shown that NF-κB signalling is transient, and at late times PI, HCMV late proteins are thought to inhibit the NF-κB signalling pathways (Montag et al., 2006). This is supported by the ability of HCMV to inhibit TNFR surface expression (Baillie et al., 2003), and down-regulate interferon-response factor 3 (IRF3) activation of the interferon response (Abate et al., 2004). A diagram depicting the mutual regulation of HCMV and NF-κB is shown in Fig 1.4.

Activation of extracellular-signal regulated kinases (ERK) ERK1/2 via MEK1/2 also occurs downstream of PI3-K signalling (Fig. 1.3). ERKs have been shown to phosphorylate

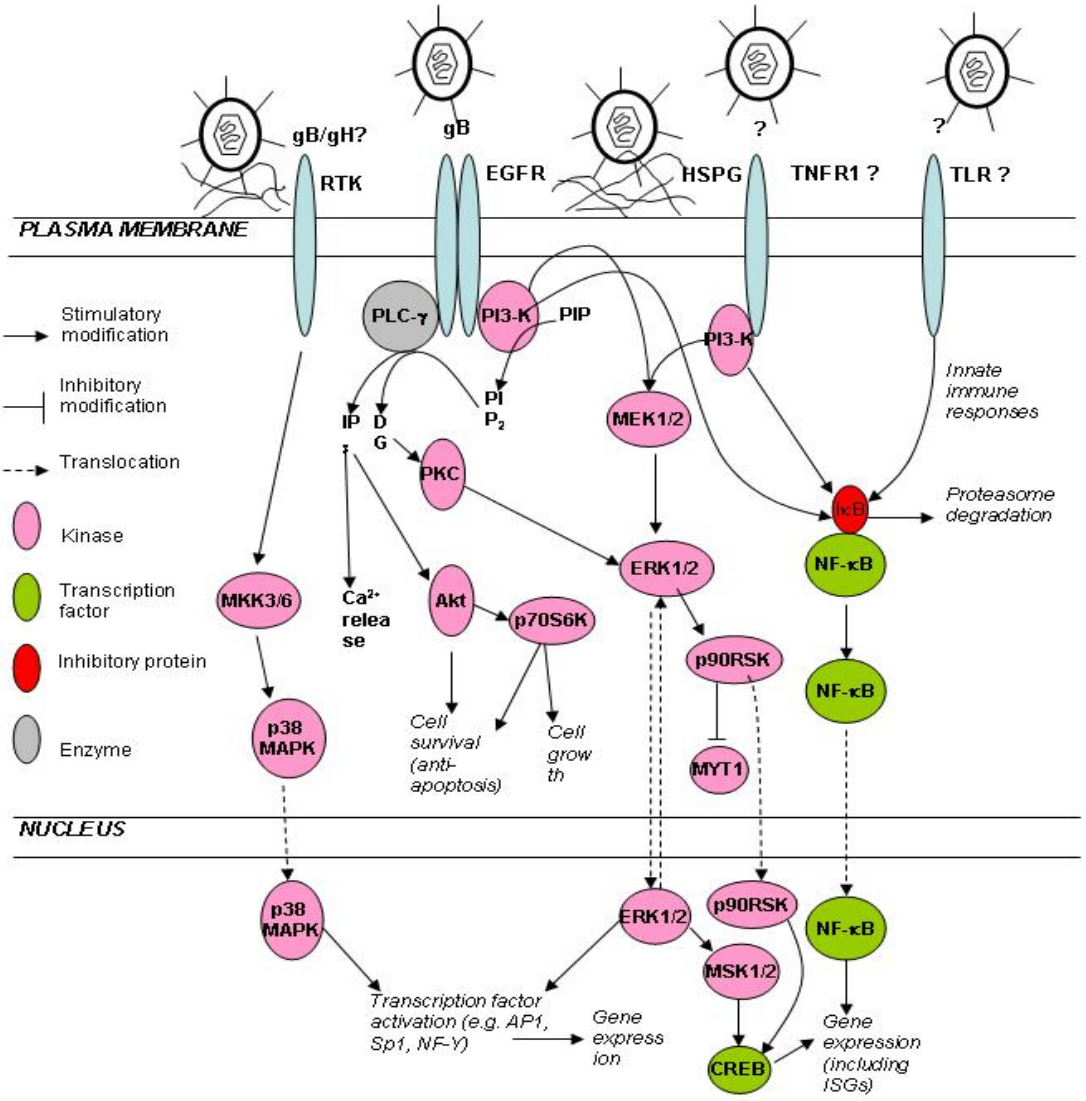


Figure 1.3. Simplified diagram of the HCMV induced effects on intracellular signalling.

HCMV particles associate with the HSPG, then HCMV particle glycoproteins gB and possible gH specifically interact with cell surface receptors, leading to the induction of different cellular signalling pathways. HCMV particles gain entry into the cell by fusion (fibroblasts) or endocytosis (epithelial and endothelial cells) (not shown on the diagram). RTK = receptor tyrosine kinase; EGFR = epidermal growth factor receptor; HSPG = heparin sulphate proteoglycan; TNFR = tumour necrosis factor receptor; TLR = toll-like receptor

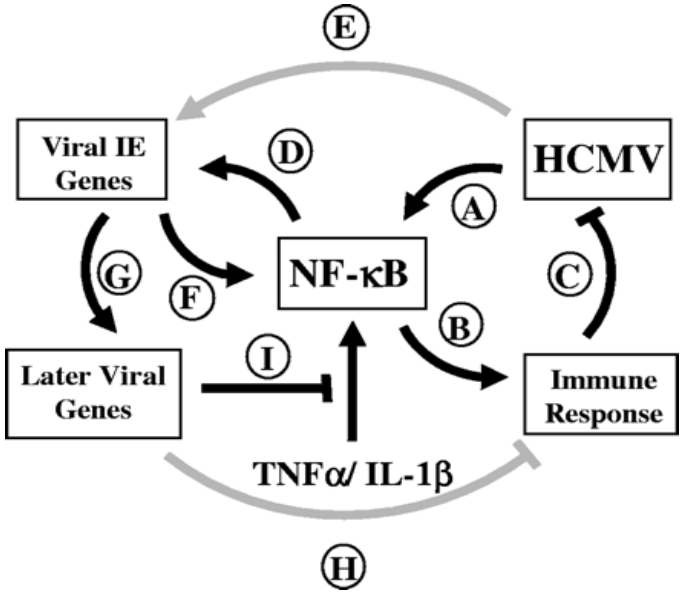


Figure 1.4. Diagram depicting the mutual regulation of HCMV and NF-κB (Modified from Montag et al., 2006).
(A) binding of HCMV activates NF-κB, which leads to the expression of immune response genes (B) to establish anti-viral defence mechanisms (C). HCMV IE gene expression is co-stimulated by NF-κB (D) along with viral accessory proteins and other cellular transcription factors, possibly induced by virus/cell interactions e.g. via EGFR signalling (E). In turn, it is suggested that HCMV IE proteins activate expression of NF-κB subunit genes (F). HCMV IE proteins activate the expression of early/late genes (G), which in turn recognise the cells anti-viral defence program (H), and turn it off (I).

p90RSK (p90 ribosomal kinase), which upregulates CREB (cAMP-response elememt binding protein), but also inhibits MYT1, which is a CDC2/cyclin B regulatory kinase. Since HCMV infection arrests the cell cycle at the G1/S boundary, and because CREB plays a role in activation of the MIEP, this suggests a possible role for ERK activation in preparing the cellular environment for HCMV lytic replication (Fortunato et al., 2000). HCMV infection of human fibroblasts also induces p38 MAPK signalling at early times PI (Fig. 1.3), where HCMV inhibits dephosphorylation of p38. At later times PI, MKK3/6 are upregulated, thus stimulating further p38 MAPK signalling (Johnson et al., 2001a). Inhibition of p38 signalling using the drug FHPI (4-(4-fluorophenyl)-2-(4-hydroxyphenyl)-5-(4-pyridyl)1H-imidazole), showed that HCMV DNA replication and late gene expression were inhibited, thus demonstrating a critical role for p38 MAPK signalling during HCMV infection of fibroblasts (Johnson et al., 1999; Johnson et al., 2000a).

Binding of HCMV gB to EGFR also activates a number of cellular interferon stimulated genes (ISGs), which are usually induced by IFN-α, which is a response normally associated with anti-viral defence (Navarro et al., 1998; Preston et al., 2001); however, different pathways are utilised by IFN-α and HCMV. IFN-α is a secreted cytokine that induces a signalling cascade via the JAK/STAT pathway, which leads to the nuclear localisation of ISGF3 (interferon stimulated growth factor-3), and activation of ISGs via binding of ISGF3 to the interferon-response element (ISRE). However, HCMV induced signalling via the interaction of gB/EGFR leads to the formation of a novel ISRE-binding protein complex, which contains interferon-regulatory factor 3 (IRF3) and CREB-binding protein, upregulating the expression of ISGs (Fortunato et al., 2000). As the infection progresses, cellular responses to IFN-α are diminished, thus disarming the anti-viral response. The HCMV tegument protein, pp65 (pUL83) has been shown to be an antagonist of ISG expression via inhibition of NF-κB, interferon-response factor-1 (IRF1) (Browne et al., 2003), and IRF3 (Abate et al., 2004). Interestingly, IE1 (72 kDa) has been reported to counteract STAT-mediated interferon signalling (Paulus et al., 2006).

1.5 Transcriptional transactivation in HCMV infected cells

1.5.1 Promoter structure

The major immediate-early promoter (MIEP) is composed of a promoter-enhancer region which spans from -500 to +1 bp (relative to the transcription start site) and is involved in the regulation of IE1/IE2 gene expression. A distal regulatory region spans from -1000 to -500 bp and is thought to repress the MIE promoter-enhancer. The entire region is densely

packed with cellular transcription factor binding sites in order to regulate the expression of the IE1/IE2 genes. Differentially spliced IE1/IE2 transcripts give rise to four proteins: IE1 (72 kDa), IE2 (86 kDa), IE2 (60 kDa), and a true late protein IE2 (40 kDa). Repression of IE1/IE2 gene expression occurs via binding of p53 and YY1 at sites on the distal-enhancer, or by the binding of IE2 (40 kDa) or IE2 (86 kDa) at the cis repression signal (crs) just upstream of the transcription start site (Mocarski and Tan Courcelle, 2001). A diagram showing the structure of the IE1/IE2 promoter-enhancer and its subsequent regulation by IE1/IE2 spliced variants is shown in Fig. 1.5.

Early genes have a promoter structure that consists of a TATA box and initiator elements just upstream of the transcription start site, and an upstream region spanning 100 to 200 bp that contains cis-acting regulatory elements for trans-acting viral and cellular proteins. Early promoters (and some early-late promoters) also have CAAT or CCAAT boxes in the upstream region located between -50 to -100 that can bind cellular transcription factors. Transcription does not usually start at the +1 site as RNA polymerase II can begin synthesis of mRNA from up to 150 bp upstream of the +1 site; this is referred to as the transcript leader (Mocarksi and Tan Courcelle, 2001). A diagram depicting the general structure of an early gene promoter is shown in Fig. 1.6A.

True-late genes have a very basic promoter structure that consists of a TATA box upstream of the transcription start site, and no upstream cis-acting regulatory signals. In HSV-1, true-late (γ-2) promoters may also have a downstream activation signal (DAS) at +20 to +44 (Costa et al., 1981). Expression of the true-late HCMV gene UL94 is reported to be fully activated from a promoter fragment from -45 to +1 in a transient expression assay; this is also the case for UL99 (pp28) where a fragment from -40 to +6 is sufficient for activation (Wing et al., 1998; Depto and Stenberg, 1992; Kohler et al., 1994). Because true-late promoters lack cis-acting regulatory signals, true-late gene expression is dependent on the replication of viral genomic DNA. A diagram depicting the general structure of a true-late gene promoter is shown in Fig. 1.6B.

1.5.2 Switch from immediate-early, to early, to late gene expression

A UL122 (IE2) mutant Towne-BAC transfected into HFF cells failed to generate viral plaques, and mRNA for early genes could not be detected by RT-PCR. This confirmed that UL122 (IE2) is an essential gene for the growth of HCMV and is required for the switch from E to L gene expression (Marchini et al., 2001). Therefore, the main regulatory protein for the switch from IE to E to L gene expression is IE2 (86 kDa); however, IE2

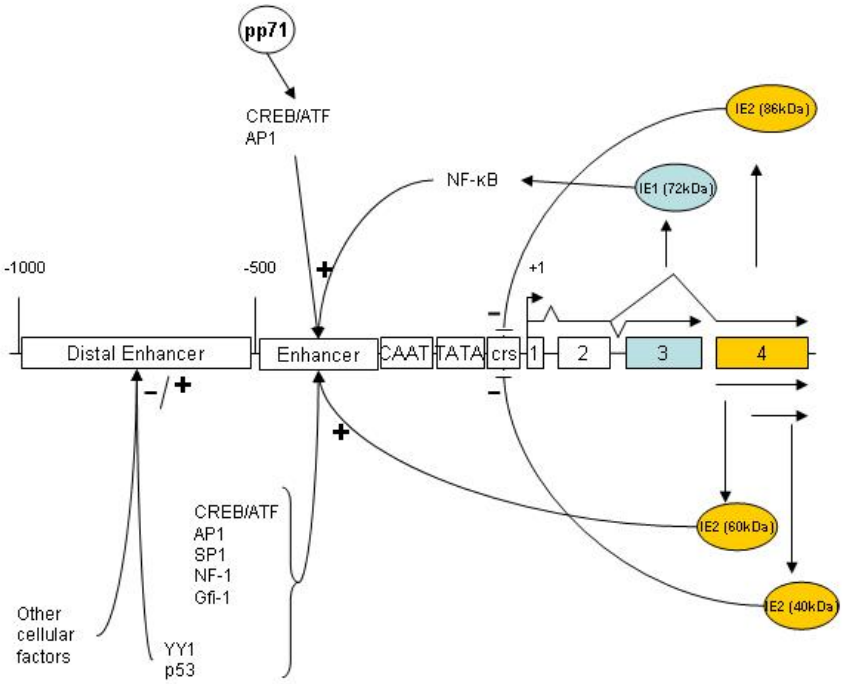


Figure 1.5. Structure of the IE1/IE2 promoter-enhancer and the regulation of downstream gene expression.

The main components are the promoter region, containing the TATA box, the cis-repression signal (crs) upstream of the transcription start site, and a CCAAT box. The enhancer and the distal-enhancer contain multiple cellular transcription factor binding sites for the regulation of the IE1/IE2 gene locus. Differential splicing of IE1/IE2 transcripts leads to the production of four immediate-early proteins: IE1 (72 kDa), IE2 (86 kDa), IE2 (60 kDa) and IE2 (40 kDa) - all of which play a role together with cellular transcription factors in the activation and/or repression of the downstream IE1/IE2 genes. pUL84 is also a transdominant inhibitor of IE2 (86 kDa), and can interact with IE2 (86 kDa) in order to bind to the crs and inhibit expression of the downstream exons.

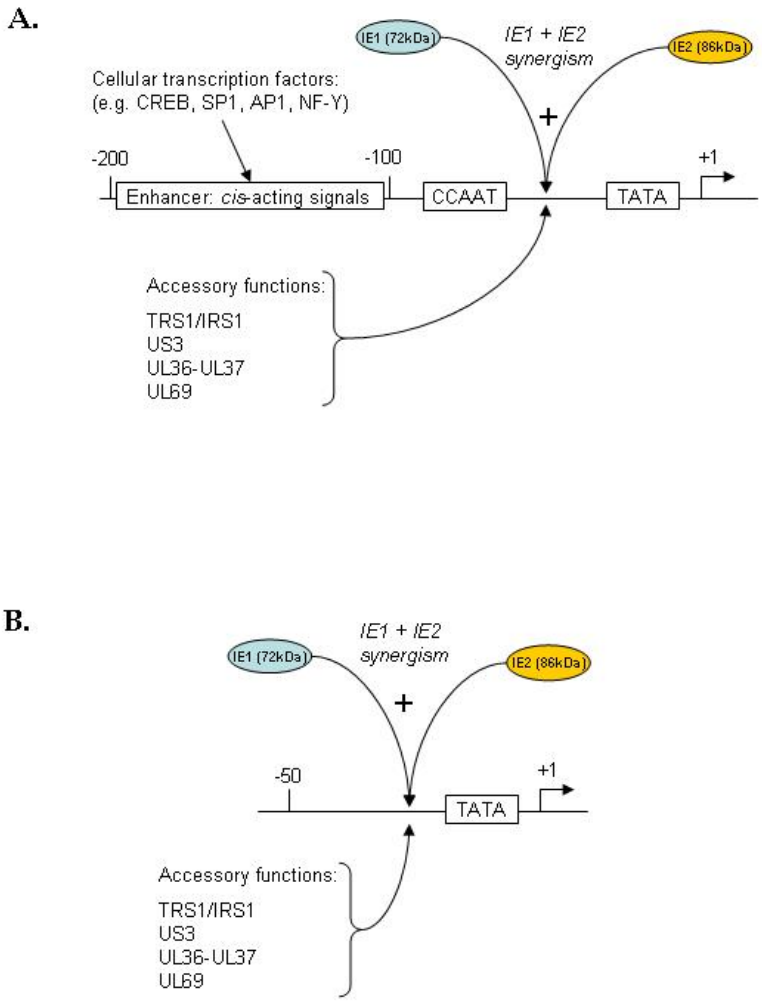


Figure 1.6. Promoter structure for early (E) and late (L) HCMV genes.

(A) Shows the general structure of a promoter for an early expressing gene, which contain a TATA box, CCAAT box, and upstream cis-acting regulatory elements. (B) Shows the general structure of a promoter for a late expressing gene, note that these promoters are very basic, and only contain a TATA box upstream of the transcription start site.

(86kDa) transactivation of early and late genes is more efficient in the presence of IE1 (72 kDa) (Mocarski et al., 1996). IE1 (72 kDa) also influences IE2 (86 kDa) transactivation in combination with TRS1/IRS1 ancillary regulators, and other viral IE proteins; US3, UL36, UL37 and UL69, which have accessory functions (Iskenderian et al., 1996). The expression of some genes is also influenced by the use of alternative promoters, which allows the production of both early and late transcripts from the same ORF. Transcripts are generated that have the same or related functions but differ in size because different promoters or polyadenylation signals are used, while their expression kinetics may also differ due to the operation of alternative posttranscriptional regulatory control mechanisms (Mocarski and Tan Courcelle, 2001).

UL4 is an early gene that has three differentially regulated transcription start sites; two are expressed at early times, and the third at late times (Chang et al., 1989a and b). UL4 expression is also influenced by the presence of cellular transcription factor binding sites for Elk-1 and NF-Y in an upstream element (Chen and Stinski, 2000). UL4 is one of the best characterised examples of posttranscriptional regulation in HCMV infected cells. UL4 mRNA contains three short ORFs in the 5′ leader, which affect downstream translation of the UL4 mRNA. During translation, the second short ORF (uORF2) prevents the release of tRNA from the ribosome, thus stalling the translation machinery on the mRNA (Alderete et al., 1999). Mutation of the uORF2 initiation codon results in abundant early expression of UL4 (Alderete et al., 2001).

UL44 (DNA polymerase processivity factor) also has three transcription start sites; the distal and proximal sites are active at early times, and the middle site is active at late times. It has been suggested that the early sites are responsive to the immediate-early regulatory proteins, while at late times, the middle site is induced by late-specific regulatory cis-acting sequences and TATA-binding transcription factors (Leach and Mocarski, 1989). HCMV viruses with mutated UL44 distal or middle TATA elements had lower levels of expression of UL44 and of late viral genes, which resulted in slower growth of the mutants compared to wild-type. Mutation of the UL44 distal site, but not the middle site, also resulted in a reduction in the level of viral DNA synthesis (Isomura et al., 2007).

UL99 (pp28) is a true-late gene for which the promoter sequences from -40 to +6 (relative to the transcription start site) are sufficient to induce maximal activation of gene expression, therefore, specific upstream elements are not required (Depto et al., 1992; Kohler et al., 1994). This has also been observed for the expression of the late UL94 gene,

where a promoter fragment from -45 to +1 achieved full activation of UL94 in a transient expression assay (Wing et al., 1998). It is therefore probable that HCMV true-late gene expression is induced solely from a basic TATA element just upstream of the transcription start site.

1.5.3 Regulatory activities of HCMV immediate-early proteins

1.5.3.1 Role of IE1 in the transactivation of gene expression

IE1 (72 kDa) is involved in regulating gene expression from the IE1/IE2 promoter-enhancer, and also has a role in the efficient expression of early and late genes. IE1 (72 kDa) is able to transactivate the MIEP indirectly by promoting the binding of NF-κB to the 18 bp repeats that are distributed along the MIE promoter-enhancer (Cherrington and Mocarksi, 1989). UL123-null mutants are able to replicate when cells are infected at a high m.o.i., but exhibit a growth defect at a low m.o.i. (Mocarski et al., 1996; Greaves et al., 1998). In transient transfection assays, the transactivation of viral promoters by IE2 (86 kDa) alone or in combination with the ancillary regulators, pTRS1/pIRS1 is enhanced by the presence of IE1 (72 kDa) (Iskenderian et al., 1996). IE1 (72kDa) alone is a poor transactivator of heterologous promoters, but is shown to increase transcription of target promoters when co-expressed with IE86 (Hagemeier et al., 1992a; Klucher et al., 1993).

Interaction of IE1 (72 kDa) with the CCAAT box binding protein, CTF1, induced expression from a TATA-less DNA polymerase α-promoter (Hayhurst et al., 1995), while IE1 (72 kDa) interaction with the E2F binding factor (E2F1) leads to the activation of the TATA-less dihydrofolate reductase (DHFR) receptor promoter (Margolis et al., 1995). Thus, the mechanism by which IE1 (72 kDa) influences gene expression from heterologous promoters occurs in a TATA box independent manner (Hagemeier et al., 1992b). IE1 (72 kDa) is also reported to affect the acetylation status of core histone and other nuclear proteins in order to promote the expression of viral genes. The acetylation of core histone proteins results in chromotin modifications, exposing the DNA sequences and activating transcription. HCMV genomic DNA is complexed with histone proteins, and it was reported that IE1 (72 kDa) was able to antagonise the action of histone deacetylases (HDACs), resulting in the activation of viral gene expression (Nevels et al., 2004).

HCMV infection rapidly arrests the progression of the cell cycle at multiple points; G1 phase, G2/M and G1/S boundaries (Lu and Shenk, 1996); a process that is mediated by

the actions of both IE1 (72 kDa) and IE2 (86 kDa). Direct interaction of IE1 (72 kDa) with the N-terminus of p107 (retinoblastoma family member protein), results in the displacement of p107 from cyclin E/cdk2 (Zhang et al., 2003), which allows cyclin E/cdk2 to function as a kinase, and leads to the activation of E2F (family of transcription factors involved in cell proliferation) responsive promoters (Poma et al., 1996). Further cell cycle regulation occurs by IE1 (72 kDa) induction of activator protein (AP-1) gene expression via an interaction with a cellular protein kinase, MEKK1, itself a component of the MAPK/ERK signalling pathway (see section 1.4.3). AP-1 is composed to two subunits, *c-fos* and *c-jun*, and is reported to be a cellular transcription factor with regulatory roles in cell proliferation and cell differentiation (Kim et al., 1999).

1.5.3.2 Role of IE2 in the transactivation of gene expression

It is known that IE2 (86 kDa) transactivates heterologous promoters by utilising cellular transcription factors to mediate the expression of viral and cellular genes (Lukac et al., 1994; Kim et al., 2000a). It has been reported that IE2 (86 kDa) transactivation properties are based on the ability of IE2 (86 kDa) to act as an adaptor protein to stabilise transcription factor II D (TFIID) associated factors on the promoter. This is evidenced by the fact that IE1 (72 kDa) can interact with TFIID associated factors, TFII 130 and TFII 110, while IE2 (86 kDa) simultaneously binds TFII 130 and the TATA box binding protein (TBP), thereby bridging interactions between proteins of the basal transcription machinery (Caswell et al., 1993; Lukac et al., 1997; Kim et al., 2000b). The ability of IE2 (86 kDa) to interact with target proteins is dependent on its phosphorylation status, and ERK2 (a MAPK family member), phosphorylates IE2 (86 kDa) at several different sites (Harel and Alwine, 1998). Sumoylation of IE2 (86 kDa) is also involved in regulation of its transactivation function, where interaction of IE2 (86 kDa) with SUMO-1 is suggested to provide IE86 with an additional cellular cofactor interaction motif (Hofmann and Stamminger, 2000; Ahn et al., 2001). The SUMO-1 modification of IE2 (86 kDa) is enhanced by an interaction with PIAS1 (protein inhibitor of STAT), which has SUMO E3 ligase activity, thus enhancing sumoylation of target proteins. This may modulate IE2 (86 kDa) transactivation of viral and cellular promoters (Lee et al., 2003).

At late times PI, IE2 (86 kDa) and IE2 (40 kDa) function directly to switch off immediate-early gene expression by binding to a cis-repression signal (crs) located between the TATA box and the +1 site at the start of the IE1/IE2 gene locus (Jenkins et al., 1994). It is suggested that the repression of US3 gene expression also occurs via binding of IE2 (86 kDa) and IE2 (40 kDa) at a crs located just upstream of the +1 site (Lashmit et al., 1998),

and crs-like sequences have also been identified in the early UL112 and UL4 promoters (Arlt et al., 1994; Huang et al., 1995). DNA binding of IE2 (86 kDa) and IE2 (40 kDa) to the crs occurs via the C-termini of the two proteins (Schwartz et al., 1994), and is thought to inhibit expression by preventing the assembly of the RNA polymerase II transcription machinery on the MIE promoter (Lee et al., 1996). The core domain in the IE2 (86 kDa) C-terminus required for interaction with the crs sequences consists of amino acids 450 to 544 (Asmar et al., 2004). However, IE2 (86 kDa) transactivation function requires both amino and carboxy domains. The domain comprising amino acids 86 to 542 was reported to be required for protein/protein interactions with cellular transcription factors: NF-κB, AP-1, Sp1, Tef-1, c-*jun*, *jun*B, ATF-2, protein kinase-A phosphorylated delta-CREB, p300, CBP, Nil-2A, CHD-1 and UBF (Spector, 1996).

As with IE1 (72 kDa), IE2 (86 kDa) interacts with proteins that regulate cell proliferation. IE2 (86 kDa) binds the retinoblastoma protein (Rb) preventing Rb inhibition of the E2F cellular transcription factor family proteins, leading to the expression of genes associated with the S phase of the cell cycle e.g. cyclin E, cdk-2, B-myb, E2F-1, ribonucleotide reductase 1 (RR1), ribonucleotide reductase 2 (RR2), thymidilate synthetase (TS), MCM3 and MCM7 (Hagemeier et al., 1994; Song and Stinski, 2002). The IE2 (86 kDa)/Rb complex also alleviates IE2 (86 kDa) repression of the MIEP, but suppresses IE2 (86 kDa) mediated transactivation of promoters. It was therefore suggested that Rb function may be a factor in the slow progression of the HCMV lytic cycle (Sommer et al., 1994; Choi et al., 1995).

HCMV infection of fibroblasts, astrocytes and human umbilical vein endothelial cells (HUVEC) results in elevated steady-state levels of the tumour suppressor protein, p53 (Muganda et al., 1994; Lokensgard et al., 1999; Kovacs et al., 1996). However, elevated levels of p53 are not associated with the activation of p53 responsive genes in HCMV infected cells, as evidenced by decreased levels of p21 (cyclin-dependent kinase inhibitor 1A), as the expression of p21 is regulated by p53 (Chen et al., 2001). p53 transactivation of both viral and cellular promoters is repressed by an interaction of the C-terminal domain of IE2 (86 kDa) with p53 (Subler et al., 1992; Tsai et al., 1996), while increased levels of p53 arise from an interaction of IE2 (86 kDa) with mdm2 (a negative regulator of p53 gene expression), leading to the proteasomal degradation of mdm2 by a mechanism not yet fully understood (Zhang et al., 2006). The affect of p53 on the replication of HCMV was examined in fibroblasts expressing wild-type p53, or in fibroblasts where p53 had been deleted. It was found that p53 is sequestered at viral replication centres in the nucleus, and that wild-type p53 enhanced the replication of HCMV. Furthermore, a delay in the release of infectious virus progeny was also observed in p53 negative fibroblasts

(Casavant et al., 2006). Similar delays in the accumulation and release of HCMV were also observed in the U373Mg cell line (although the data was not presented in the paper) (Casavant et al., 2006). The U373Mg astrocytoma cell line has a point mutation in p53 codon 273, resulting in an Arg → His substitution, which abates normal p53 function (Van Meir et al., 1994). It is known that wild-type p53, unlike mutant p53, can inhibit the expression of c-*fos*, and repress expression from several other cellular and viral promoters (Ginsberg et al., 1991). The mutant (R → H) p53 (present in U373Mg cells), unlike wild-type p53, can induce the expression from some cellular promoters, including the human multiple drug resistance gene, MDM1 (Chin et al., 1992). However, it remains unclear as to how mutant p53 functions, although interaction of mutant p53 with a constitutively expressed cellular heat shock protein (hsc70) was suggested to extend the half-life of mutant p53, enhancing the transformation of cells (Finlay et al., 1988). It has been shown that various p53 mutants (including the defective p53 protein present in U373Mg cells) activated viral and cellular promoters, requiring only the presence of a TATA box, while mutant p53 mediated expression was further enhanced by the presence of CREB binding sites in the promoter (Deb et al., 1992). It was also reported that HCMV early and late gene promoters are activated by IE2 (86 kDa) alone in transient transfection assays in U373Mg cells (Klucher et al., 1993; Wu et al., 1998; Wu et al., 2001). Therefore, the regulatory functions of mutant p53 are altered in U373Mg cells, leading to the activation of viral and cellular promoters, while the synergistic mechanisms of HCMV transactivation of early and late promoters by the viral immediate-early proteins may also be perturbed in this cell line.

Although IE2 (86 kDa) is essential for the transactivation of viral promoters, high levels of IE2 (86 kDa) are reported to be cytotoxic. As a result, the MIEP is autoregulated by IE82 (86 kDa) and IE2 (40 kDa) via a cis-repression sequence (crs) upstream of the transcription start site. pUL84 also binds to IE2 (86 kDa) in order to mediate the down-regulation of the MIEP. In addition, pUL84 is also responsible for the inhibition of IE2 (86 kDa) mediated transactivation of heterologous promoters (Gebert et al., 1997).

1.6 Herpesvirus tropism

1.6.1 HCMV cell tropism

During acute primary infection, HCMV infects a range of cell types and organs (Sinzger et al., 1995; Sinzger and Jahn, 1996); but the laboratory passage of the virus is generally undertaken in fibroblasts since virus yields are highest in this cell type. Unfortunately,

long term passage of HCMV in fibroblasts results in changes in the virus, both at the level of DNA sequence, and host-range phenotype. Mutations in virus tropism (host-range) genes limit the utility of fibroblast adapted HCMV as a research tool, and compromise its effectiveness as an *in vitro* model for viral infection in different cell types. This problem was highlighted following simultaneous parallel passage of a fresh HCMV isolate in human umbilical vein endothelial cells (HUVEC) or neonatal human dermal fibroblasts (NHDF); after the twentieth passage in NHDF, the isolate had lost its ability to induce cytopathic effects in HUVEC, and its ability to propagate in HUVEC had diminished greatly. However, the same isolate passaged in HUVEC retained natural endothelial and fibroblast cytopathogenicity and broad host-range cell tropism. This study demonstrated that HCMV encodes genes that confer cell type specific tropism, and that passage of HCMV in fibroblasts restricts the host cell range (Waldman et al., 1991).

During acute primary infection, fibroblasts, smooth muscle cells and epithelial cells produce the highest yields of infectious progeny. Peripheral blood mononuclear lymphocytes (PBML) and endothelial cells are also infected, and play an important role in the dissemination of HCMV. Assessment of PBML cell populations during *in vivo* or *in vitro* HCMV infections have identified two PBML cell types (neutrophils and monocytes) as containing infectious virus, the latter being the predominant infected cell type. Monocytes and neutrophils are not permissive for HCMV but simply serve as vehicles for the transport of HCMV, where PBML-mediated transfer of virus to uninfected cells at a site distal to the original focus of infection is accomplished by transient microfusion events (Grundy et al., 1998; Revello et al., 1998; Gerna et al., 2000). Subsequent differentiation of monocytes into macrophages results in fully permissivity for HCMV. Neutrophils are not capable of further differentiation and replication in neutrophils is blocked at the immediate-early (IE) stage of infection (Sinzger et al., 1995). Compared to fibroblasts, the growth kinetics of HCMV in monocyte-derived macrophages (MDM) was delayed, and the virus was compartmentalised in vacuoles in the cytoplasm. It was therefore inferred that the virus has evolved to replicate without causing lysis of MDM cells, thus allowing the virus to evade immune surveillance and spread infectivity by direct cell-to-cell contact (Fish et al., 1995).

Monocytes may also differentiate into dendritic cells (DC), and monocyte-derived dendritic cells are also fully permissive for HCMV. Fibroblast adapted, but not HUVEC adapted HCMV, loses ability to infect and/or replicate in monocyte-derived DCs. Dendritic cells are located in tissues and organs, and present antigens, delivering them to the lymph tissue to initiate the process of NK cell destruction. Mature dendritic cells are

more active than immature dendritic cells. Thus, the ability of the virus to replicate in immature dendritic cells might be expected to interfere with the immune response to infection. In addition, the migration of infected dendritic cells to lymph tissues might facilitate virus transmission to other body sites (Riegler et al., 2000).

Endothelial cell tropism has been genetically linked to the ability of HCMV isolates to transfer infectivity from infected fibroblasts or HUVEC cells to PMNL (monocytes and neutrophils) (Revello et al., 2001). Transfer of the HCMV isolates from HUVEC to PMNL is the function of a late viral gene product that appears to mediate microfusion events between the infected cell and PMNL cell plasma membranes, facilitating cell-to-cell transfer of cytoplasmic contents, including viral and cellular proteins, RNA, and virus particles. The infected PMNL cells serve as a transport vehicle for the haematological dissemination of virus. Virus transfer is mediated by subsequent microfusion events between the infected PMNL cells and uninfected cells, organs and tissues (Gerna et al., 2000). Clinical isolates serially passaged in human embryonic lung fibroblasts (HELF) cells simultaneously lose both endothelial cell tropism and the ability to transfer to PMNL cells indicating that the same gene(s) are responsible for both phenotypes (Gerna et al., 2000). Thus the ability of HCMV to transfer to PMNL can be considered to be a surrogate marker for its ability to replicate in endothelial cells, and hence pathogenicity of the virus. It follows then that, attenuation of the virus by long-term passage in fibroblast cell culture during vaccine production might be a consequence of the lack of transfer between these two cell types *in vivo* (Gerna et al., 2002).

Latency and persistence are important features of the HCMV life cycle, as they are responsible for lifelong infection. HCMV latently infects myeloid progenitor cells of the monocyte lineage (Maciejewski et al., 1992; Kondo et al., 1994). It was also suggested that HCMV persistently infects aortic endothelial cells, which then serve as a further reservoir of virus (Fish et al., 1998). However, it was reported that the virus does not undergo a replicative cycle, and that viral infection does not affect the normal growth of aortic endothelial cells (Kahl et al., 2000). It must be noted that these observations are yet to be confirmed by others. In general, the laboratory strains of HCMV (AD169 and Towne) and the low passage clinical isolate (Toledo) are unable to propagate in endothelial cells (Bolovin-Fritts and Weideman, 2001).

1.6.2 HCMV sequence hypervariability and its effects on cell tropism

Variation in the DNA sequences of some HCMV genes between different HCMV isolates has resulted in the recognition of different genotypes for the same gene e.g. there are four different genotypes of the gB gene (Chou and Dennison, 1991). Genotyping of such gene sequences can be utilised to identify strains of HCMV in the general population. There are a number of HCMV genes where the level of sequence variation is more extensive, referred to as variable or hypervariable genes (Dolan et al., 2004; see Fig. 2). The high amino acid variability is typically restricted to parts of the ORF, and is associated with ORFs that lie near the termini of the UL component of the HCMV genome, and is not a feature of the herpesvirus core genes. For the most part, hypervariable genes are thought to encode membrane proteins, and therefore, hypervariability might be expected to play a role in immune evasion. Although the number of genotypes coded by variable sequence is not yet known, it is clear from the high number of such genes that the potential number of different strains would be counted in thousands. It might be expected that differences in DNA sequence would result in HCMV strains with diffrerent pathogenicities, either through differences in immune evasion or cell tropism, though this hypothesis has yet to be studied.

Gene UL55 encodes the essential glycoprotein, gB, required for cell entry and cell-to-cell spread (Navarro et al., 1993; Boyle and Compton, 1998). Sequence analysis reveals that four major subtypes exist, and that sequence variation is at its highest between codons 448 to 480 (Chou and Dennison, 1991). Investigation of gB type 1 (gB1) reported this subtype to be generally associated with asymptomatic and non-fatal cases of HCMV infection, particularly in bone marrow recipients. Symptomatic HCMV infection among bone marrow recipients is thought to be associated with the gB type 2 (gB2) subtype (Fries et al., 1994). It was also reported that patients can be infected with two or more strains of HCMV, exhibiting different gB subtypes, and that the gB subtype did not correlate with prevalence in either the blood or urine from infected patients. However, it was reported that gB subtype may influence viral tropism *in vivo* since HCMV bearing either the gB2 or gB3 subtypes were able to infect T-lymphocytes, whereas virus with the gB1 subtype was not found to be associated with this cell type. It was concluded that gB subtype could influence the pathogenesis of HCMV *in vivo* and that it was linked to differences in viral tropism (Meyer-Konig et al., 1998).

The HCMV UL144 gene has a highly variable DNA sequence and encodes a homologue of the human tumour necrosis factor receptor. Three main subtypes, A, B and C, and two

minor UL144 subtypes have been reported (Arav-Boger et al., 2001), with variability recorded as approximately 20 % over 172 amino acids (Lurain et al., 1999). Hypervariability in DNA sequence typically results in very high rates of amino acid substitutions. Gene UL146 codes a functional CXC-chemokine, with a role in the dissemination of HCMV via monocytes (Penfold et al., 1999). Analysis of UL146 sequences revealed 14 subtypes, resulting from extensive sequence divergence, with amino acid substitutions at approximately 60 % (Dolan et al., 2004; Lurain et al., 2006). However, conserved regions include the core CXC motif, and also putative signal peptides and the spacing between the cysteine residues (Prichard et al., 2001). Such hypervariation and the presence of multiple UL146 subtypes could be a further determinant of HCMV virulence in the host.

1.6.3 HCMV genes involved in cell tropism

Viral tropism functions can operate at different stages in the replicative cycle of the virus. At the most basic level, the virus particle requires the appropriate receptor proteins to attach to, or to penetrate the plasma membranes of particular cell types. Binding of HCMV to cell surface receptors induces an anti-viral response e.g. interferon and/or cell apoptosis, and HCMV has evolved genes that interfere with these responses (Fortunato et al., 2000). If these viral functions operate more successfully in one cell type compared to another, these viral genes play a role in cell tropism.

Immediately after entry, herpesvirus particles become associated with the cytoskeleton network and are transported from the cytoplasm to a nuclear pore where the capsid releases its DNA molecule to enter the nucleus and initiate IE gene expression and viral replication (see section 1.4). In the case of non-endothelial-tropic strains of HCMV, it has been reported that transport of particles to the nucleus is blocked in endothelial cells (Sinzger et al., 2000). HSV-1 viral tegument proteins (VP16, VP22 and US11) are required for binding of the particle to the motor protein complex on microtubules (Diefenbach et al., 2002). With respect to non-endothelial-tropic HCMV strains, it may be that genetic changes involving a tegument protein(s) might affect cell tropism, since it has been reported that HCMV capsids associate with the microtubule network (Ogawa-Goto et al., 2003), and HCMV particles lacking the pUL47 and pUL48 tegument proteins have a post-entry defect, resulting in abortive infection (Bechtel and Shenk, 2002).

Apoptosis or programmed cell death is an innate property of many cell types, and can be induced when under viral attack. This suicide function operates to arrest viral replication

and prevent the spread of viral progeny in the host. The M45 gene of MCMV has strong anti-apoptotic function that confers endothelial cell tropism (Brune et al., 2001). The MCMV M45 homologue in HCMV is UL45; however, it was reported that pUL45 was not a determinant of HUVEC cell tropism, and although pUL45 exhibited a mild anti-apoptotic effect, it was of a much lower potency than that of M45 (Hahn et al., 2002). HCMV encodes other genes whose products inhibit apoptosis. HCMV gene UL37 is contained within a highly-differentially-spliced region of the virus genome within the gene locus UL36-UL38 (Adair et al., 2003). UL37 was found to have anti-apoptotic function, and was denoted vMIA (viral mitochondria-localised inhibitor of apoptosis). It was demonstrated that vMIA localised in the mitochondria and inhibited Fas-mediated apoptosis. No sequence similarity was detected between vMIA and the cellular equivalent, Bcl-2. It was therefore concluded that vMIA belongs to a new class of apoptosis inhibitors (Goldmacher et al., 1999). Further interrogation of UL37 identified two domains within exon 1, the first domain is required for trafficking of UL37 protein to the mitochondria, while the second domain prevents apoptosis by inhibiting the release of cytochrome c from the mitochondria (Hayajneh et al., 2001). It was later reported that a vMIA-Bax complex was formed, thus preventing Bax-induced mitochondrial membrane permeabilisation, and subsequent cell apoptosis (Arnoult et al., 2004).

The UL36 gene also codes for an anti-apoptotic factor, which was denoted the vICA (viral inhibitor of caspase-8 induced apoptosis). It was reported that the vICA inhibits apoptosis by directly binding to and inactivating caspase-8, a protein intermediate of the Fas-mediated apoptosis pathway. As with vMIA, no sequence similarity was detected between vICA and the cellular regulatory proteins of caspase-8 activity. It was concluded that vICA represents a novel class of viral apoptosis inhibitor (Skaletskaya et al., 2001). Murine cytomegalovirus (MCMV) gene M37 (a homologue of HCMV UL37) also localises in the mitochondria. A gene knockout of M37 severely diminishes the ability of MCMV to grow *in vivo* but not *in vitro*. Thus, M37 is a virulence factor and might be required for viral replication *in vivo* in certain cell types or tissues (Lee et al., 2000).

The UL128 gene locus provides an important cell tropism function. All HCMV isolates adapted to growth in fibroblasts invariably acquired mutations in one or more of three adjacently located genes UL128, UL130 and UL131A (termed the UL128 gene locus). This gene locus is clearly detrimental for virus growth in fibroblasts, but is indispensable for growth in endothelial and epithelial cells, and also infection of PMNL cells (monocytes and macrophages) (Akter et al., 2003; Hahn et al., 2004; Wang and Shenk, 2005a), and monocyte-derived dendritic-cells (Gerna et al., 2005). It was reported that two

glycoprotein gH complexes are present in HCMV infected cells: gH-gL-gO, and gH-gL-pUL128-pUL130. It was shown that the gH-gL-gO complex is sufficient for the entry of HCMV particles into fibroblasts, but the gH-gL-pUL128-pUL130 complex is also required for entry into epithelial and endothelial cells (Wang and Shenk, 2005b). Furthermore, the expression of pUL131A was reported to be required for the release of virus from endothelial cells; in contrast, wild-type pUL131A impaired the release of virus from infected fibroblasts (Adler et al., 2006).

1.6.4 Role of viral determinants for cell tropism in non-human herpesviruses

RhCMV gene Rh10 encodes an endothelial specific tropism factor, designated vCOX-2 (Rue et al., 2004), that is a homologue of cellular cyclooxygenase-2 (COX-2) – an enzyme required for the synthesis of eicosanoids which function to mediate innate immunity, homeostasis, and inflammatory responses (Smith et al., 2000). It was found that a recombinant vCOX-2 deletion mutant (Δ10RhCMV) was not able to replicate in endothelial cells (Rue et al., 2004). HCMV does not code a vCOX-2 homologue, but it has been shown that HCMV infection induces cellular COX-2 expression in infected cells, and that inhibition of cellular COX-2 resulted in decreased yield of HCMV, indicating an important role for COX-2 in HCMV replication, and suggesting that HCMV genes inducing up-regulation of COX-2 are involved in pathogenesis (Speir et al., 1998).

Many HCMV genes not required for growth in tissue culture, but retained in clinical isolates, are presumed to have roles in pathogenesis, immune evasion or cell/tissue tropism *in vivo* (Cha et al., 1996; Dargan et al., 1997). Examples of such genes have been identified in other herpesviruses. VZV ORF47 (HSV-1 homologue of UL13) is dispensable for growth *in vitro*, but essential for VZV replication in T cells and skin cells, identifying this gene as an essential virulence factor *in vivo* (Moffat et al., 1998). VZV ORF10 (homologous to HSV-1 VP16) encodes a tegument protein that is dispensable for growth *in vitro*, but is required for growth *in vivo* in epidermal and dermal cells (Che et al., 2006).

Murine cytomegalovirus (MCMV) infection of mice resembles that of HCMV infection in humans in many aspects, including the ability to replicate in different cell types and organs. The characterisation of MCMV tropism factors that allow MCMV to replicate within different cell types is an area of focused research. The generation of transposon insertion MCMV mutants and their ability to replicate *in vitro* in NIH 3T3 cells, and in the lungs, liver, spleen, kidneys and salivary glands of BALB/c and SCID mice has been investigated. It was shown that a mutant in gene M43 (RvM43) was able to replicate to

high titres in NIH 3T3 cells and in all the tissues examined, except for salivary glands where it exhibited a growth defect. It was therefore suggested that gene M43 was a viral determinant of salivary gland tropism (Xiao et al., 2000). Other MCMV viral determinants of cell tropism include M83 and M84, which are homologues of the HCMV gene, UL83 (pp65). It was shown that the MCMV M83 and M84 deletion mutants were able to replicate to wild-type levels in NIH 3T3 cells, but both mutants exhibited a growth defect in the spleen, liver and kidneys during *in vivo* infection. In addition, the M83 mutant was severely attenuated in the salivary glands (Morello et al., 1999).

Further MCMV determinants of cell tropism include the US22 family members (HCMV homologues are shown in backets): m142 (IRS1), m143 (TRS1), M36 (UL36), m139, m140, and m141. This study investigated the ability of MCMV transposon insertion mutants of the US22 gene family to replicate in fibroblasts, macrophages and endothelial cells. It was shown that M36, M43, m139, m140 and m141 were each viral determinants of macrophage tropism; m139, m141 and m141 proteins interact to provide the essential function. The tropism function of M36 was directly related to its anti-apoptotic function (Menard et al., 2003).

1.7 Application of microarray technology to the study of viruses

DNA microarrays have proved a powerful tool for the first-step, global analysis of gene expression. Typically, a DNA microarray consists of DNA oligonucleotide molecules (probes) representing hundreds or thousands of ORFs arrayed onto a glass slide. In gene expression studies, the arrayed probe sequences interrogate labelled cDNA molecules prepared from mRNAs, which then hybridise to their cognate probes. Microarray experiments might compare the expression of transcripts from two different samples e.g. drug-treated or non-treated cells. The cDNAs prepared from the two samples are labelled with either Cy3- or Cy5-conjugated nucleotides. The labelled cDNAs are then hybridised to the array in equal concentration to allow binding to the oligonucleotide probes. Analysis of the intensity of Cy3- and Cy5-labelled cDNAs prepared from the mRNA extracted from cells grown under different conditions allows for comparative gene expression studies (Stekel, 2003).

The applications of DNA microarrays include gene expression analysis, genotyping, diagnostics, and mapping of genomic libraries. Furthermore, tissue-based microarrays have also been designed in order to study gene expression from multiple tissues on one slide, and protein microarray chips have been developed to study nucleic acid-protein,

protein-protein, ligand-receptor, drug-protein targets, and also enzyme-substrate interactions. Like DNA microarrays, protein arrays are based on the attachment of capture molecules such as DNA, antibodies, enzyme substrates etc. to a glass slide, but with the hybridisation of fluorescently labelled protein molecules. This section provides an overview of microarray studies drawn mainly from herpesvirus literature, and is not a fully comprehensive review. However, microarrays for different herpesviruses that are novel and/or unique in their design have been selected for discussion in order to illustrate the application of this technology in virus research.

1.7.1 Types of microarrays in current use

Microarrays are generally printed on one of two solid supports; nylon membranes or glass slides. The solid support onto which the nucleic acid probes are printed influences the probe design and deposition, as well as the labelling strategy of the cDNA to be hybridised to the array. Nylon membrane arrays generally contain nucleic acid probes derived from PCR products of ~ 300 bp, and hybridised with cDNA that has been synthesised from RNA with direct incorporation of a radiolabelled dNTP. In contrast, arrays that are printed onto glass slides usually contain oligonucleotides (50– to 70-mers), which have been presynthesised prior to spotting, and hybridised with cDNA that has been synthesised and labelled directly or indirectly using fluorescent CyDye molecules (typically Cy3 or Cy5). On the other hand, the Affymetrix GeneChip is based on light directed *in situ* synthesis of 25-mers directly onto glass slides, and hybridised with biotinylated-labelled antisense cRNA. For custom DNA microarrays, the deposition of the nucleic acid probes onto the solid support is achieved using a robot that contains pins to spot the probes. However, inkjet arrays have been utilised for the *in situ* synthesis of oligonucleotide probes directly onto the surface of glass slides, however, unlike Affymetrix GeneChips, the inkjet system is inefficient at producing large batches of identical arrays (Stekel, 2003). Attempts to optimise microarray data capture have been reported, in which alternative methods to label cDNA and detect spot signals have been used. The use of nano-sized gold and silver particles for the labelling of cDNA and detection of signals is based on the resonance light scattering (RLS) properties of these metallic particles. The RLS system is much more sensitive than CyDyes, and the metallic particles do not photobleach upon exposure to light, as is the case with fluorophores (Yguerabide and Yguerabide, 2001). The RLS system was tested for the HCMV microarray described in this thesis. Despite the advantage of requiring a small input of total RNA for cDNA synthesis and labelling, the HCMV microarrays gave a poor and uneven

background, which was considered detrimental for the capture of high quality spot signals (data not shown).

While commercially available microarrays have been developed e.g. the Affymetrix GeneChip – these arrays are based on the cellular genomes of model organisms: Human, Arabidopsis, *E. coli*, *C. elegans*, mouse etc. To date, virus specific arrays are custom made, and have been developed using nylon membrane or glass slide based platforms, using a range of labelling and detection methods as discussed above. A HCMV microarray was first reported by Chambers et al., (1999) who designed an oligonucleotide array based on the AD169 mutant virus sequence, with additional sequences from the 15 kb fragment present in the Toledo strain (Chambers et al., 1999). However, this microarray has become dated due to recent assessments of the true coding potential of HCMV clinical isolates compared to AD169, and comparisons of the coding potential of HCMV with that of CCMV (Davison et al., 2003; Dolan et al., 2004; Murphy et al., 2003a and b). Since the development of the first HCMV microarray, virus-specific arrays have now been reported for other members of the herpesvirus family including: HSV-1 (Stingley et al., 2000), HSV-2 (Aguilar et al., 2005), VZV (Kennedy et al., 2005), EBV (Li et al., 2006), HHV-6 (Ohyashiki et al., 2005), KSHV (Jenner et al., 2001; Paulose-Murphy et al., 2001), MHV-68 (Ahn et al., 2002; Ebrahimi et al., 2003), and MCMV (Tang et al., 2006).

Microarrays have proven to be useful and robust tools for genomics studies, but each microarray system (i.e. Affymetrix or custom arrays based on nylon membranes or glass slides) is associated with specific advantages and disadvantages. The Affymetrix GeneChips can contain tens of thousands of different probes, which are standardised in design and manufacture; moreover, the system is also standardised for RNA preparation, cRNA synthesis/labelling, hybridisation, signal detection/data capture and normalisation. However, this system is not an open source, and there is an absolute requirement for the arrayed sequence to be known. This restricts the design of probes in regions of genomes that have yet to be fully sequenced and/or annotated. In contrast, custom made oligonucleotide based arrays are flexible in design, probes can be ORF- and strand-specific, while the system remains an open source. However, there are no standardised operating procedures with respect to the hybridisation of cDNA, or the subsequent data capture and analysis. These concerns have been addressed by microarray specialists through the introduction of MIAME (minimum information about a microarray experiment) and the development of public repositories for microarray data e.g. ArrayExpress hosted at the European Bioinformatics Institute (EBI) (Rocca-Serra et al.,

2003). The HCMV microarray data reported in thesis is yet to be deposited into ArrayExpress.

1.7.2 Microarrays used in transcriptome profiling of virus gene expression

Virus-specific custom DNA microarrays have been designed for the following alphaherpesviruses: HSV-1 (Stingley et al., 2000), HSV-2 (Aguilar et al., 2005) and VZV (Kennedy et al., 2005); beteherpesviruses: HCMV (Chambers et al., 1999), HHV-6 (Ohyashiki et al., 2005) and MCMV (Tang et al., 2006); gammaherpesviruses: EBV (Li et al., 2006), KSHV (Jenner et al., 2001; Paulose-Murphy et al., 2001), and MHV-68 (Ahn et al., 2002; Ebrahimi et al., 2003). For the HCMV, HSV-1, HHV-6, and the two MHV-68 arrays, the kinetic class of virus transcripts were examined using drugs to limit the expression to the IE, early and late gene classes. In contrast, the MCMV and the two KSHV arrays examined virus gene expression kinetics over a time course. For the KSHV arrays, hierarchical clustering was applied to the data in order to identify the major gene expression groups (Jenner et al., 2001; Paulose-Murphy et al., 2001). For the MCMV array, virus gene expression was examined in fibroblasts (NIH 3T3), and transcripts from ORFs apparently not expressed in fibroblasts were then examined in macrophages by real-time PCR, thus identifying potential viral determinants of macrophage cell tropism (Tang et al., 2006). The VZV array compared virus gene expression at a single time point in two different cell lines (MeWo and SVG) (Kennedy et al., 2005). An EBV custom DNA microarray was used to profile the transcriptome of EBV in cell lines derived from NK/T-cell lymphoproliferative disorders (LPD) in parallel with the examination of host cell gene expression using a human oligonucleotide array (Zhang et al., 2006).

The total percentage (%) expression from the various herpesvirus genomes detected by each microarray was as follows: HCMV (75 %), HSV-1 (100 %), VZV (97 %), HHV-6 (61 %), MCMV (84 %), MHV-68 (100 %; Ahn et al., 2002), MHV-68 (100 %; Ebrahimi et al., 2003), KSHV (100 %, Jenner et al., 2001) KSHV (100 %, Paulose-Murphy et al., 2001), EBV (detection limit not reported). For arrays that examined the kinetic class of transcripts expressed (HCMV, HSV-1, HHV-6, and both MHV-68 arrays), the results were broadly similar irrespective of the virus, with the percentage of genes falling within each class as approximately: IE (6 %), early (34 %), and late (60 %). However, this approximation does not include the data for the HHV-6 array, where it was shown that the percentage genes within each kinetic class were: IE (2 %), E (12 %), and L (86 %). It seems likely that the differences may be caused by the incomplete annotation of the HHV-6 virus genome (only 61 % of probes on the array detected a positive signal), and because the DNA synthesis

inhibitor used to define the early genes inhibited both viral and cellular DNA polymerases (Ohyashiki et al., 2005). Further difficulty in correlating data from different groups is caused by the introduction of a new E-L class. The HCMV (Chambers et al., 1999) and MHV-68 arrays (Ebrahimi et al., 2003) reported the herpesvirus kinetic classes as IE, E, E-L and L. However, for the HSV-1 and MHV-68 (Ahn et al., 2002) arrays, the percentage of genes that could be split into E-L and L are likely to be similar to that of HCMV and MHV-68 (Ebrahimi et al., 2003). The KSHV arrays investigated changes in gene expression in a time course experiment, covering the switch from latently infected cells through to lytic replication, revealing four major gene expression groups (as assessed by hierarchical clustering), but which roughly corresponded to latency associated genes, and the three kinetic classes: IE, E and L (Jenner et al., 2001; Paulose-Murphy et al., 2001). Scrutiny of the data published by the various groups showed that in general, viral transactivators and anti-apoptotic factors were the first groups of viral genes expressed, then enzymes and some tegument proteins, and finally the remaining tegument and structural proteins. However, a gene belonging to the early kinetic class in one herpesvirus might not be expressed with the same kinetics in another virus. For example, the herpesvirus core gene gB, is reported to be expressed with early kinetics in HCMV, HSV-1, and KSHV, but is considered a late gene in MHV-68 (supported by both Ahn et al., 2002, and Ebrahimi et al., 2003). These studies have shown that although the general transcription programme is conserved between different members of the herpesvirus family, the kinetic class of expression of some genes may vary between different herpesviruses.

1.7.3 Microarrays used in direct comparative studies of gene expression in different viruses

A microarray was developed to detect both HSV-1 transcripts and the equivalent HSV-2 homologues in order to compare the transcription profile of these two alphaherpesviruses during a lytic infection in a single cell line (NIH 3T3 cells). It was shown that the majority of transcripts from HSV-1 and HSV-2 belonged to the same kinetic class, although the level and accumulation of HSV-2 transcripts was generally lower than that of HSV-1 transcripts. Differential gene expression of HSV-1/HSV-2 in NIH 3T3 was examined using the correlation function within the Microsoft Excel package to identify significant differences in the expression kinetics of individual viral genes over the time course. It was shown that the expression of UL4, UL30, UL29 and UL31 differed between these two viruses. These genes are all involved in nuclear organisation and viral genome replication, and it was suggested that their differential regulation was probably due to different

properties of viral replication compartments during infection with HSV-1 compared to HSV-2 (Aguilar et al., 2006). This work followed on from an earlier report of a HSV-2 oligonucleotide microarray that had been applied to compare the global transcription programme of HSV-1/HSV-2, showing that the transcriptome profiles of these two viruses were very similar (Aguilar et al., 2005).

The cellular response to the alphaherpesvirus members, HSV-1 and pseudorabies virus (PRV), during infection of a common permissive cell type (rat embryonic fibroblasts; REF) was examined using Affymetrix RGU34A arrays. Despite PRV and HSV-1 displaying similarities in genome structure, gene conservation, virion structure, and replication cycle, the cellular response to these viruses was diverse. Approximately 1500 cellular genes were increased or decreased > 3 fold following infection with HSV-1 and PRV, falling into 24 different functional classes. However, only 500 of these cellular genes were increased or decreased in common, and the remaining cellular genes were expressed in a virus-specific manner. Classes of genes that were affected during infection with the alphaherpesviruses HSV-1 or PRV included anti-viral defence, cell signalling, apoptosis, and the heat shock and oxidative stress pathways (Ray and Enquist, 2004).

Similarly, the cellular response to the betaherpesvirus members, HCMV and RhCMV, was directly compared in a common permissive cell type (rhesus fibroblasts), in order to investigate the pathogenesis of these two viruses. As with the HSV-1/PRV study, it was found that there were similarities in the cellular response to HCMV/RhCMV, but virus-specific expression of some cellular genes was detected for genes involved in metabolic and physiological processes, or cell-cell communication. Surprisingly, the largest differences were observed for genes involved in the cellular innate immune response, and while HCMV induced expression of ISGs, proinflammatory cytokines etc., genes with these functions were not induced in cells infected with RhCMV. It was suggested that RhCMV particles have evolved to deter induction of these innate immune responses, in contrast to HCMV, which benefits from induction of these cellular responses during the early stages of infection (DeFilippis and Fruh, 2005).

These results suggest that while there are shared cellular responses to different viruses from the same herpesvirus subfamily, the majority of changes in cellular gene expression are virus-specific. In contrast, regulation of expression of virus genes from members of the same herpesvirus subfamily are similar. These comparative studies allow virus replication to be dissected with respect to the differential regulation of virulence factors, coded by

related viruses in the same cellular environment, and also highlight common or diverse cellular responses to infection with different virus types.

1.7.4 Microarrays used to determine the effect of herpesvirus infection on cellular gene expression

The effect of herpesvirus infection on cellular gene expression has been examined in alphaherpesviruses: HSV-1 (Mossman et al., 2001), VZV (Jones and Arvin, 2003); betaherpesviruses: HCMV (Zhu et al., 1998; Browne et al., 2001), HHV-6 (Mayne et al., 2001); and gammaherpesviruses: KSHV (Poole et al., 2002). These microarray studies have shown that the major classes of cellular genes that are modulated during a herpesvirus infection have roles in cell adhesion and structure, signal transduction/receptors, basal cell machinery (including synthesis, transport, and energy metabolism), immune and stress response, cell cycle, and apoptosis. Despite the commonality of functional classes of cellular genes affected, examples of virus-specific changes in the cellular transcription programme in response to infection with different herpesviruses have also been reported.

HCMV infection results in the induction of interferon stimulated genes (ISG), which occurs via intracellular signalling induced by the binding of HCMV particles to the cell surface. Increased ISG expression was also maintained in the presence of HCMV gene expression during the early stages of infection (Browne et al., 2001). Similarly, HHV-6 and VZV induce the expression of ISGs and proinflammatory cytokines (Mayne et al., 2006; Jones and Arvin, 2003). In contrast, while HSV-1 infection can also induce ISGs, the induction of these genes requires penetration of HSV-1 particles into the host cell, and ISG expression is blocked when HSV-1 gene expression is inhibited (Mossman et al., 2001). It is likely that the innate anti-viral defence mechanism induced by the interferon response is beneficial to HCMV, VZV, HHV-6 and KSHV in the early stages of a lytic infection, but is detrimental to HSV-1 infection. This difference in cellular response could highlight a major difference in the pathogenesis of HSV-1 infection compared to that of other herpesviruses. Interestingly, the induction of ISGs seen in HCMV infected cells was not observed during an infection with RhCMV (see section 1.7.3) (DeFilippis and Fruh, 2005).

Other examples of cellular genes that are modulated differently in response to virus infection include genes involved in apoptosis. It has been shown that VZV infection of T-cells leads to the down-regulation of caspase-8 (Jones and Arvin, 2003), in contrast, HHV-6 infection of T-cells leads to the up-regulation of caspase-8 (Mayne et al., 2001). It was suggested that this may reflect the need for VZV to inhibit apoptosis of T-cells to promote

the transfer of virus from T-cells to skin cells, and person-to-person spread of this virus (Jones and Arvin, 2003). However, it was also suggested that the different caspase-8 responses following infection of T-cells with VZV (Jones and Arvin, 2003) or HHV-6 (Mayne et al., 2001), could be due to the use of primary T-cells (VZV), versus immortalised T-cells (HHV-6). A response that was unique to HHV-6 infection of T-cells was the induction of IL-18 expression. This cytokine is expressed in demyelinating lesions of multiple sclerosis (MS) brains, linking HHV-6 infection with the progression of MS, and possibly other auto-immune disorders (Mayne et al., 2001).

It was reported that KSHV infection of human dermal microvascular endothelial cells (DMVEC) significantly induced the expression of tissue plasminogen activator (PLAT) and endothelial plasminogen activator inhibitor (PAI-1), and dysregulation of TSP1, angiopoietin 2, Cys-rich angiogenesis inducer (IGFBP10), and thrombomodulin, all of which may contribute to the process of angiogenesis and the formation of KS lesions. Down-regulation of the angiogenesis inhibitor, thrombospondin-1 (TSP1), was also observed in KSHV infected DMVECs (Poole et al., 2002). Interestingly, TSP1 is also down-regulated during HCMV infection in fibroblasts (Zhu et al., 1998) and human retinal glial cells (Cinatl et al., 2000), and in U373Mg cells stably expressing HCMV IE1 (72 kDa) (Lee et al., 2005). This may draw a common link between the pathogenesis of HCMV (in respect to retinitis and astrocyte malignancies), and KSHV with respect to the formation of KS lesions.

1.7.5 Microarrays used to compare the effect of regulatory virus genes on viral or cellular transcription

HSV-1 encodes an immediate-early protein, α27 (UL54), which is involved in the posttranscriptional regulation of viral RNA by inhibiting splicing and mediating the transport of viral transcripts from the nucleus into the cytoplasm. An HSV-1 mutant with a knockout mutation in UL54 was used to examine the regulatory activities of this protein with respect to viral gene expression, and also its effects on 57 cellular genes (Stingley et al., 2001; see section 1.7.2). The microarray was not able to fully resolve the effects of the UL54 knock-out mutant on the transcription programme of HSV-1, because some results were in conflict with previously reported observations e.g. the down-regulation of viral DNA polymerase (UL30) and helicase-primase (UL52) was not observed using the HSV-1 array. However, it was shown that immediate-early transcripts such as α4, α47 (US12), and α22 (US1) were increased, suggesting that α27 modulates their levels during infection. A HSV-1 gene that was down-regulated was that encoding ICP34.5, which has been reported to be a neurovirulence factor. The 57 cellular genes were generally

upregulated following infection with the HSV-1 UL54-KO mutant, suggesting that α27 has a role in the decline of cellular mRNA during a normal HSV-1 infection (Stingley et al., 2001).

It has been reported that HCMV IE2 (86 kDa) affects the cell cycle in infected cells by progressing quiescent (G0) cells to the G1/S boundary (Song and Stinski, 2005). The effect of IE2 (86 kDa) on the transcription programme of quiescent fibroblasts was examined in order to identify IE2 (86 kDa) induced cellular genes. This was achieved using transducing recombinant adenovirus vectors expressing IE2 (86 kDa) in quiescent fibroblasts, and examining the cellular transcriptome using an Affymetrix human genome GeneChip. This study showed that IE2 (86 kDa) induced the up-regulation of many E2F responsive genes including: myb, cyclin E, cdk-2, E2F-1, ribonucleotide reductase 1, ribonucleotide reductase 2, thymidylate synthetase, MCM3 and MCM7; all of which are involved in promoting G0/G1 phase transition to the S phase (Song and Stinski, 2002). In an analogous microarray study, the transcriptome of U373Mg cells stably expressing IE1 (72 kDa) was examined in order to gain insight into the potential role of HCMV on the pathogenesis of malignant glioma, since a high percentage of malignant gliomas are infected with HCMV (Cobbs et al., 2002). It was found that 14 cellular genes were modulated, including the down-regulation of glial fibrillary acidic protein (GFAP), TSP-1 and p53, and it was concluded that suppression of these anti-oncogenic genes contributed to the onset of glial tumourogenesis (Lee et al., 2005). These studies show that viral regulatory genes can modulate both viral and cellular transcription responses.

1.7.6 Microarrays used to compare of viral gene expression during lytic and latent phases of infection

Microarray studies into herpesvirus gene expression during latent and lytic phases of replication have been reported for KSHV (Jenner et al., 2001), and EBV (Li et al., 2006), while the changes in expression of cellular genes have been studied during HSV-1 latent and reactivation phases (Kent and Fraser, 2005). In the case of KSHV and EBV, a limited number of viral genes were expressed during latency (for KSHV: vFLIP, v-cyclin, and LANA-1; and EBV: EBNA1, EBNA2, EBER1 and EBER2). Reactivation of KSHV or EBV replication led to the expression of viral genes, subsequently defined in KSHV as primary, secondary and tertiary lytic genes. The expression of genes associated with an EBV lytic cycle were not identified according to the KSHV scheme (Jenner et al., 2001), and were reported only as expressed or not expressed (Li et al., 2006). However, both studies

revealed that the latency-associated genes were not expressed during the KSHV or EBV replication phase.

The effect on the cellular transcriptome was examined during reactivation of HSV-1 from latently infected mouse trigeminal ganglia. During latency, 18 cellular genes were up- or down-regulated > 2.5 fold, which included immune response genes such as those encoding interferon-inducible protein 1, STAT1, and cytokine inducible SH2-containing protein 7. Reactivation of HSV-1 resulted in the modulation of 48 cellular genes, including many more immune response genes, genes involved in cell signalling (e.g. TNF receptor) and cell-to-cell spread of the virus (e.g. alpha 2 catenin) (Kent and Fraser, 2005).

1.7.7 Microarrays used to compare viral gene expression in different cell types infected in culture

Transcriptome profiling of VZV gene expression was compared in the human skin melanoma cell line (MeWo) and the human astrocytoma cell line (SVG), because of the clinical presentations as varicella or herpes zoster, and the neurotropic nature of the virus respectively. The VZV transcriptome was compared at one time point (72 h PI), at which maximal CPE was observed in both cell lines. The relative abundance of VZV transcripts was then compared in the two cell lines. For the astrocytoma SVG cells, only 20 out of 71 VZV ORFs were significantly expressed, whereas 68 out of 71 VZV ORFs were significantly expressed in skin MeWo cells. Of the top six expressed ORFs in MeWo cells, only ORFs 49, 57 and 58 were also significantly expressed in SVG cells. Despite large differences in the transcription programme of VZV infection in skin (MeWo) and astrocyte (SVG) cells, analysis of differential expression of VZV genes was restricted because it was not possible to perform a synchronised high m.o.i. VZV infection of the two cell lines, and therefore the differential expression kinetics of viral genes could not be studied (Kennedy et al., 2005).

Differential viral gene expression in a time course experiment has been investigated for a baculovirus (autographa californica multicapsid nucleopolyhedrovirus; AcMNPV) in a time course experiment in two different cell types. The transcriptome profile of AcMNPV was determined and compared in insect *S. frugiperda* (Sf-9) and *T. ni* (TnHigh-Five) cells. Six viral genes displayed differential expression kinetics and the functions of two of these are known; one gene, p35, is involved in virus-origin specific DNA replication and exhibits anti-apoptotic activity, and the second, gene p10, has a role in stabilisation and release of infectious progeny virus from cells at the end of infection. The function of the

remaining four differentially expressed genes (lef-6, lef-3, PK-2 and ORF150), are as yet unknown (Yamagishi et al., 2003).

Together, the VZV and AcMNPV studies highlighted a number of points regarding the design of microarray experiments for differential analysis of viral gene expression in different cell types. Firstly, it is important to initiate a synchronised infection in each of the cell types studied, and to select time points that allow appropriate comparisons between virus gene expression profiles. Secondly, robust statistical analyses are required to extract significant and biologically relevant data.

1.7.8 Microarrays designed for the simultaneous comparison of viral and cellular gene expression

Microarrays that include both viral and cellular oligonucleotide probes have been reported for HSV-1 (Stingley et al., 2000) and MHV-68 (Ebrahimi et al., 2003), but the complement of probes representing cellular targets was limited (< 100). An EBV custom DNA mcroarray was used to profile the viral transcriptome in cell lines derived from NK/T-cell lymphomas in parallel to the examination of host cellular gene expression using a second human oligonucleotide array. Results indicated that EBV genes such as BARF-1 (a latency associated protein) was expressed in natural killer cells and T-cell lymphoproliferative disorders (NK/T-LPDs), thus supporting the data from earlier reports that BARF-1 is expressed in nasopharyngeal carcinoma (NPC) and gastric carcinomas. Several EBV lytic genes were also expressed including BKRF3, BDLF3, BFLF2, and BHRF1 (the homologue of human bcl-2 proto-oncogene). Cellular gene expression that was significantly upregulated included TNFRSF10D, which inhibits toll-like receptor (TRAIL) mediated apoptosis, CDK2, a key cell-cycle regulatory factor, and HSPCA, a heat shock protein associated with mantle-cell lymphoma (Zhang et al., 2006). While an attempt was made to define an EBV - NK/T-LPD molecular signature, the study was focused on the *in vitro* selection of EBV-infected NK or T-cells, and it was suggested that clinical samples would be required to clarify these findings.

1.7.9 Viral diagnostic microarrays

Microarrays have the potential to be ideal tools for viral diagnosis since they allow for the parallel screening of different human viruses; microarrays are fast, sensitive and specific, though expensive. However, in the case of HCMV at least, careful probe selection would be required because different strains of HCMV have different genotypes for some variable

sequence genes (see section 1.6.2). The development of viral diagnostic arrays is an interesting area of research, and the production of arrays that can detect human herpesviruses in human plasma and/or cell culture would be useful. An oligonucleotide based microarray that can differentiate between HSV-1, HSV-2, VZV, EBV, CMV and HHV-6 has already been developed (Foldes-Papp et al., 2004). The sensitivity (detection limit) of this diagnostic array was also enhanced by the use of fluorescently labelled dendrimers to optimise hybridisation signals (Striebal et al., 2004).

Another herpesvirus microarray has been developed for the detection of eight human herpesviruses in samples from cerebrospinal fluid (CSF), whole blood, plasma and serum. Oligonucleotide probes were designed for the viral DNA polymerase or the DNA polymerase processivity factors for each virus: HSV-1 (UL42), HSV-2 (UL30), VZV (ORF28), HCMV (UL44), EBV (BALF5), HHV-6A (U38), HHV-6B (U38) and HHV-7 (U38). PCR primers were also designed for each virus gene and used for RT-qPCR experiments for a direct comparison with the microarray data. It was shown the PCR and microarray data were in agreement for 94 % of the tests. However, some negative results for PCR were found to be positive in the microarray, and it was thought that these represented false-positives, that had not been eliminated during analysis of the microarray data. While the PCR and microarray data were largely in agreement, it was reported that each system had advantages and disadvantages. Diagnostic PCR tests were performed sequentially and therefore could be quite expensive, whereas with the microarray, several viruses were detected in a single hybridisation reaction. However, it was noted that optimisation of microarray sensitivity was required in order to avoid false-positive detection of viruses (Jaaskelainen et al., 2006).

1.8 Aims of the thesis

Human cytomegalovirus (HCMV) is the leading viral cause of neonatal abnormality and plays an important role in morbidity and mortality among immuno-compromised patients. During acute primary infection, HCMV is present in most organs, (heart, kidneys, lungs, spleen and liver), and replicates in many different cell types (fibroblasts, epithelial cells, endothelial cells, smooth muscle cells, macrophages and dendritic cells). The aims of this study were firstly to determine whether the pattern of HCMV gene expression was changed when the virus was grown in different cell types, and secondly to explore whether novel ORFs recently proposed by other workers are transcribed. A HCMV microarray platform was designed based on the Merlin strain of HCMV, since it contains a complete complement of intact ORFs, with the exception of UL128, which has a premature termination mutation. The microarray contains probes for all known and predicted ORFs. HCMV transcriptome activity was investigated in human foetal foreskin fibroblasts (HFFF), human retinal pigmented epithelial cells (RPE) and human astrocyte cells (U373Mg) at several time points post infection (PI). To summarise, the aims of this investigation were to:

1. Design and produce a custom DNA microarray based on the recent reassessment of HCMV coding potential.

2. Profile the expression of the low passage clinical isolate, Merlin, in permissive cell lines; human foetal foreskin fibroblasts (HFFF-2), human retinal pigmented epithelial cells (RPE), and human astrocytes (U373Mg).

3. Identify HCMV genes whose expression was differentially regulated in the different cell types.

4. Verify the microarray data using real-time PCR, northern and western blotting techniques.

5. To correlate virus growth characteristics in each cell type with the kinetics of viral gene expression.

2 MATERIALS AND METHODS

2.1 Materials

2.1.1 Chemicals

All chemicals were molecular biology grade, unless stated otherwise, and were purchased from SIGMA-ALDRICH (UK).

2.1.2 Radiochemicals

All radiochemicals were purchased from Amersham Biosciences (UK).

2.1.3 Restriction Endonucleases

All restriction endonucleases were purchased from New England Biolabs (UK) Ltd., and were supplied with their appropriate buffers.

2.1.4 Antibodies

Table 2.1. List of antibodies

Antibody	Species	Isotype	Company
Anti-HCMV US22	Mouse monoclonal	IgG2a	Advanced Biotechnologies Inc.
Anti-HCMV UL83 (pp65)	Mouse monoclonal	IgG3	Capricorn Products LLC
Anti-HCMV UL99 (pp28)	Mouse monoclonal	IgG3	Advanced Biotechnologies Inc.
Anti-HCMV UL44/FITC	Mouse monoclonal	IgG1	Dakocytomation
Anti-HCMV IE1/IE2	Rabbit polyclonal (recognises all protein species from the IE1/IE2 region)		Kind gift of Dr E. Mocarski

2.1.5 Composition of solutions and buffers

Bacterial culture*:*

L-broth 10 g/L NaCl, 10 g/L tryptone peptone (Becton Dickinson), 5 g/L yeast extract (Becton Dickinson), pH 7.5

L-broth agar L-broth medium with 15 g/L agar (Becton Dickinson)

Eukaryotic cell culture:

PBS	170 mM NaCl, 10 mM Na_2HPO_4, 3.4 mM KCl, 1.8 mM KH_2PO_4, 6.8 mM $CaCl_2$, 4.9 mM $MgCl_2$, pH 7.5
Trypsin	0.25 % trypsin in Tris saline pH 7.7 (140 mM NaCl, 0.7 mM Na_2HPO_4, 5.6 mM dextrose (D-glucose), 24.8 mM Tris-HCl, 2.5 mM KCl solution, containing 1 % (w/v) phenol red with 0.1 % (v/v) penicillin, 0.1 g/L streptomycin)
Versene	0.6 mM EDTA containing 0.002 % (v/v) phenol red in PBS

DNA extraction:

DNA extraction buffer	20 mM Tris-HCl pH7.5, 2 mM EDTA, containing 1.2 % (w/v) SDS
NTE buffer	140 mM NaCl, 10 mM Tris-HCl pH 7.5, 1 mM EDTA
RSB buffer	10 mM Tris-HCl pH 7.5, 10 mM KCl, 1.5 mM $MgCl_2$
TE buffer	10 mM Tris-HCl pH 7.4, 1 mM EDTA
TBE buffer (10X)	1.25 M Tris, 27 mM EDTA, 0.4 M boric acid

Microarray:

BSA blocking solution	5 X SSC/0.1 % SDS (v/v), containing 1 % (w/v) BSA

Northern blotting:

Orange G (10X)	5 mM Orange G, 1.2 M Sucrose
MOPS (10X)	0.4 M MOPS, pH 7.0, 0.1 M sodium acetate, 10 mM EDTA (SIGMA-ALDRICH)
SSC (20X)	3 M NaCl, 0.3 M tri-sodium citrate

SDS-PAGE and western immunoblotting

Resolving gel buffer	0.74 M Tris-HCl pH 8.9, containing 1 % (w/v) SDS
Stacking gel buffer	0.122 M Tris-HCl pH 6.7, containing 0.1 % (w/v) SDS
Gel running buffer	52 mM Tris, 53 mM glycine containing 0.1 % (w/v) SDS
Sample buffer (2X)	100 mM Tris-HCl pH 6.7, containing 20 % (v/v) glycerol, 4 % (w/v) SDS, 2 % (v/v) β-mercaptoethanol, 0.2 % (v/v) bromophenol blue
Towbins buffer	25 mM Tris, 192 mM glycine, containing 20 % (v/v) methanol

2.1.6 Miscellaneous reagents and commercial kits

DNA handling:

Platinum *Taq* polymerase	Invitrogen
PureLink Quick Gel Extraction Kit	Invitrogen
Rediprime II Random Prime Labelling System	Amersham Biosciences
SYBR green I	Invitrogen
Taq polymerase	QIAGEN
QIAGEN large construct kit	QIAGEN

RNA handling:

Agarose (nuclease-free)	SIGMA-ALDRICH
Amplification grade DNase I	Invitrogen
Formaldehyde (36.5-38 % in H_2O)	SIGMA-ALDRICH
NorthernMax Formaldehyde Load Dye	Ambion Inc.

Orange G	SIGMA-ALDRICH
RNA ladder (0.5 – 10 kb)	Invitrogen
Superscript II Reverse Transcriptase	Invitrogen
Tissue culture water (nuclease-free)	SIGMA-ALDRICH
QIAGEN RNeasy kit	QIAGEN

Protein handling:

ECL detection reagents	Amersham Biosciences
Rainbow markers RPN756	Amersham Biosciences

Microarray specific materials:

Centricon columns	Millipore
Hybridisation chamber (10 slide capacity)	GENETIX
LifterSlip™ (25 mm × 20 mm)	Eerie Scientific Company, USA

2.1.7 Miscellaneous materials and plastics

Chromatography paper 3MM	Whatman
ECL nitrocellulose western blotting membranes	Amersham Biosciences
Hybond-N+ nylon membrane	Amersham Biosciences
Photographic film	Kodak Ltd
Maximum recovery (nuclease-free) pipette tips	Axygen Scientific
MaxiClear (nuclease-free) 1.5m ml microtubes	Axygen Scientific

2.1.8 Computer software and algorithms

BLASTn	NCBI
GeneSpring version 7.2	Agilent Technologies
GRAIL version 1.0	Oak Ridge National Laboratories
Microsoft Excel	Microsoft
Oligo 6	Molecular Biology Insights
ScanArray Express version 3.0	PerkinElmer Inc.
SPSS version 13.0	SPSS Inc.
Quantity One 1-D Analysis Software	Bio-Rad Laboratories

2.2 Methods

2.2.1 Cell culture

Human Foetal Foreskin Fibroblasts (HFFF-2): Cat.# 86031405; European Collection of Cell Cultures (ECACC).

hTERT immortalised Retinal Pigmented Epithelial (hTERT-RPE) cells: Gift of Dr J. Shay and Dr W. Wright, University of Texas, Southwestern Medical Centre.

U373Mg Astrocytoma cells: Cat.# 89081403; ECACC.

2.2.1.1 Propagation of cell stocks

HFFF-2 and U373Mg cell cultures were grown in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10 % (v/v) FCS (DMEM/FCS) and 100 U of penicillin/streptomycin (GIBCO), and grown in 175 cm^2 flasks incubated at 37 °C in a humidified atmosphere comprising 95 % (v/v) air and 5 % (v/v) carbon dioxide. Cultures of RPE cells were grown in Dulbecco's Modified Eagle's Medium Ham's F-12 Mix 1:1 (DMEM F-12) supplemented with 10 % (v/v) FCS in 175 cm^2 flasks, and incubated as described for HFFF-2 and U373Mg cells. Confluent HFFF-2, RPE and U373Mg cell cultures were harvested by pouring off the medium, washing once with versene, and once with trypsin/versene solutions in a ratio 1:4. The cells were then resuspended in 10 ml of fresh medium and seeded into 175 cm^2 flasks with a split ratio of 1:2, with 50 ml of fresh medium added per flask.

2.2.1.2 Storage of cells in liquid nitrogen

Cells were harvested as outlined above and then pelleted at 500 x g for 10 min at RT. The cell pellets were resuspended in storage medium at a density of 1×10^6 cells/ml and 2 ml aliquots dispensed into plastic cryovials, which were placed in a polystyrene box and frozen overnight to -70 °C. After overnight incubation, the frozen cells were transferred to long-term liquid nitrogen storage. Recovery of frozen cells was accomplished by quick thawing of cell aliquots at 37 ^oC, resuspending in 25 ml of fresh medium and seeding of cells in 75 cm^2 flasks.

2.2.2 Preparation of virus stocks

HCMV strain Merlin was donated by Dr G. Wilkinson, University of Cardiff. HFFF-2 cells were seeded in 175 cm^2 flasks containing 50 ml of DMEM/FCS and grown until 90 % confluent; the cells were infected with HCMV strain Merlin at a m.o.i. of 0.01 p.f.u./cell in an inoculum of 30 ml DMEM/FCS. As the CPE developed, the medium was removed and replaced with 14 ml of fresh medium, and the infected cells were incubated at 37 ^oC until complete CPE was achieved at about 10-15 days PI. The cell supernate was clarified by pelleting detached cells at 500 x g for 10 min at RT. Cell debris was removed by centrifugation at 10000 x g for 20 min at 20 ^oC in a Sorvall RC-5B superspeed centrifuge using the SLA-1500 rotor. The cell-released virus (CRV) was then pelleted by centrifugation at 40000 x g for 1 h at 20 ^oC using the SS-34 rotor. The virus pellets were resuspended in fresh medium, pooled and dispensed, and then stored at -70 °C. Adherent infected cells were detached by tapping the culture flask and resuspension of the cells in either DMEM/FCS for cell-associated virus (CAV) stock preparations or resuspension in NTE buffer for DNA extraction. The cell suspension was sonicated to release the CAV and the preparation pelleted at low speed to clarify the virus preparation.

2.2.2.1 Titration of virus stocks

HCMV titres were determined by plaque-assay on HFFF-2 monolayers. Virus stocks were serially diluted in ten-fold steps, and 200 µl of each dilution plated on cell monolayers grown in 35 mm diameter tissue culture dishes. Virus was allowed 1 h at 37 ^oC to adsorb to cells, with gentle shaking every 15 min. Following virus adsorption, cell monolayers were overlaid with 2.5 ml of DMEM/FCS. Infected cell monolayers were incubated at 37 °C until visible plaques were observed (10-12 days PI). The medium was removed and 2 ml of Giemsa stain was added, and incubated at RT for 3 h. After staining, the fixed cell layers were washed with water and the plaques counted using a dissection microscope.

2.2.3 Extraction of HCMV genomic DNA

HCMV infected HFFF-2 cells were used to extract HCMV whole genomic DNA. Infected cells were harvested into NTE buffer and pelleted by centrifugation at 500 x g for 10 min at RT, the supernatant was discarded and the cell pellet resuspended in RSB buffer (containing NP40) and incubated on ice for 5 min to lyse the cells, and then pelleted by centrifugation as before. The supernatant was considered to be the cytoplasmic fraction and was carefully removed. The pellet was resuspended in RSB buffer, then sonicated to

release virus particles and to shear cellular and non-encapsidated DNA; this was considered to be the nuclear fraction. The cytoplasmic and nuclear sonicate were treated with DNase I (200 μg/ml) and RNase A (10 μg/ml) and incubated at 37 °C for 3 h. To extract encapsidated DNA, an equal volume of DNA extraction buffer was added, and the NaCl concentration adjusted to 0.3 M. The samples were then incubated with proteinase K (1 mg/ml) at 55 °C for 1 h. An equal volume of TE buffered phenol was added and mixed, and the sample centrifuged at 500 x g for 10 min at RT. The upper phenol layer was removed and mixed with an equal volume of TE buffered phenol:chloroform 1:1 (v/v) and centrifuged as before. The upper phase was removed and mixed with an equal volume of chloroform and centrifuged as before. Finally, the upper phase was removed, transferred to sterile dialysis tubing and dialysed against TE buffer overnight at RT. When required, DNA was concentrated by the addition of NaCl (to 100 mM), followed by ethanol precipitation and pelleting of the DNA (13000 x g for 10 min at RT), washing of the pellet in 70 % ethanol, air drying (5 min at RT), and resuspension of the DNA pellet in a small volume of sterile distilled water.

2.2.4 One step virus growth curve

35 mm tissue culture dishes seeded with 5×10^5 cells/dish of HFFF-2, RPE or U373Mg were incubated at 37 ^{o}C overnight. The cell monolayers were then infected with HCMV at a m.o.i. of 1 or 6 p.f.u./cell as previously described. After virus adsorption, the cell layers were washed twice with 1 ml of DMEM/FCS, and then overlaid with 2.5 ml of fresh medium, followed by incubation at 37 ^{o}C. At 0, 24, 48, 72, 96, 120, 144, 168 h PI, infected cell cultures were harvested by scraping the cells followed by sonication of the infected cell suspension and storage of samples at -70 ^{o}C, prior to titration of virus yield.

For separate assessment of infectivity contained in the cell-released (CRV) and cell-associated (CAV) virus fractions of the virus yield, the growth medium was pipetted off and pre-cleared by low speed centrifugation to provide the CRV fraction. Pelleted cell debris was resuspended in medium and this was used to harvest infected cells by scraping off the adherent cells into the medium, followed by sonication to provide the CAV fraction. The CAV and CRV virus preparations were then stored at -70 ^{o}C. Virus infectivity was titrated on HFFF-2 cells, as previously described.

2.2.5 Preparation of mock-infected and HCMV infected total cellular RNA

75 cm^2 flasks containing 6 × 10^6 cells/flask of HFFF-2, RPE or U373Mg were incubated at 37 ^oC overnight. The cell monolayers were either mock-infected or infected with HCMV at a m.o.i. of 6 p.f.u./cell as previously described. After virus adsorption, the cells layers were washed twice with 5 ml of DMEM/FCS, and then overlaid with 25 ml of fresh medium and incubated at 37 °C. Total RNA was harvested at 12, 24, 48, and 72 h PI, and at 72 h PI for MI cell cultures. The cell monolayers were washed and harvested as previously described. After centrifugation at 500 x g for 10 min at RT, the cell pellets were lysed using 1 ml of buffer RLT (supplied with the QIAGEN RNeasy kit), containing 1 % (v/v) β-mercaptoethanol. Cell lysates were passed through an 18-21 gauge needle 8-10 times in order to shear cellular and viral genomic DNA. Total cellular RNA was extracted using a RNeasy kit (QIAGEN), according to the manufacturer's instructions. RNA samples were quantitated by measuring the optical density at 260 nm, using an Eppendorf BioPhotometer.

2.2.6 Assessment of RNA integrity

The quality of RNA preparations were assessed by examining the integrity of the 28S and 18S ribosomal bands. 0.2-0.5 μg of RNA was diluted with water (SIGMA) to give a total volume of 7 μl, and incubated at 65 °C for 10 min, and then snap-cooled on ice. 3 μl of 5X OrangeG loading dye was added to the RNA sample, followed by electrophoresis on a 1 % agarose/TBE gel containing ethidium bromide. RNA was visualised using short-wave UV transillumination, and photographed using the BioRad Gel Doc system. High quality RNA exhibits a ribosomal band ratio for 28S and 18S rRNA of 2:1. If the 28S and 18S bands were smeared or deviated from the 2:1 ratio, the samples were discarded.

2.2.7 DNase I treatment of total RNA

RNA samples used for microarray or quantitative PCR analysis were treated with amplification grade DNase I (Invitrogen) to remove any contaminating cellular or viral genomic DNA. 80 μl of RNA sample was mixed with 10 μl of 10 X DNase I buffer, and 10 U of DNase I, then incubated at 25 °C for 15 min. The reaction was terminated by treatment with 10 μl of EDTA (25 mM), followed by incubation at 65 °C for 10 min. The DNase I treated RNA preparation was then cleaned using QIAGEN RNeasy spin-columns, according to the manufacturer's instructions.

2.2.8 Synthesis of cDNA

2.2.8.1 Synthesis of cDNA for microarray hybridisation

For microarray analysis, CyDye-3 conjugated nucleotides were used to directly label cDNA. 25 µg of total RNA from mock-infected or HCMV infected cells was resuspended in water (SIGMA) to give a final volume of 9 µl. The two control bacterial spike mRNAs (from *B.subtilis*, see section 2.2.9.1) each prepared at a concentration of 50 ng/µl, were mixed in a ratio of 1:1, and 2 µl of the pooled spike mRNA then added to the total RNA sample. 4 µl of random hexamer primers (250 ng/µl) were added to the RNA sample, which was then incubated at 70 °C for 10 min, and then snap-cooled on ice for 5 min. The following reaction mix was added to each sample: 6 µl of 5 X first strand synthesis buffer, 3 µl of DTT (10 mM), 0.6 µl of dNTP mix (dATP, dGTP, dTTP at 0.5 mM; dCTP at 0.3 mM) and 3 µl of Cy3-dCTP (0.1 mM) (Roche), and 400 U of Superscript II Reverse Transcriptase (Invitrogen). Samples were then incubated at 42 °C for 2 h, and the resulting labelled cDNA immediately stored at -20 °C.

2.2.8.2 Synthesis of cDNA for quantitative PCR

2 µg of total RNA from mock-infected or HCMV infected cells was mixed with 1 µl of random hexamer primers (250 ng/µl), 2 µl of dNTP mix (each at 5 mM), water (SIGMA) to a final volume of 13 µl, incubated at 65 °C for 5 min, and then snap-cooled on ice for 5 min. 6 µl of 5X first strand synthesis buffer, and 2 µl of DTT (0.1 M) was added to the sample and then incubated at 42 °C for 2 min. 200 U of Superscript II Reverse Transcriptase was added followed by incubation at 25 °C for 10 min, 42 °C for 2 h, and then 70 °C for 15 min, and cDNA samples then stored at -20 °C.

2.2.9 HCMV Microarray

The HCMV microarray work was performed under the guidance of our collaborator Dr. Bahram Ebrahimi (see Ebrahimi et al., 2003) of the Liverpool Microarray Facility. Protocols and the equipment for microarray hybridisations etc. were provided by Dr Ebrahimi during visits to the Liverpool Microarray Facility. Protocols for the analysis of microarray data were decided following discussions with Dr Bahram Ebrahimi, Dr Brian Lane (Bioinformatician at the North-West Institute for Bio-Health Informatics (NIBHI)) and Dr Ewan Hunter (Senior Scientist at Agilent Technologies), and consultation with my supervisor Dr Derrick Dargan.

2.2.9.1 Probe design and microarray fabrication

Microarray oligonucleotide probes were designed using Oligo6 software (Molecular Biology Insights). Viral ORFs were defined by 60-mer oligonucleotide probes based on the Merlin strain of HCMV (GenBank accession number AY446894), that shared the following properties: G+C content in the range of 40-60 %; melting temperature between 85-95 °C; lack of homo-oligomers or sequence repeats (Table 2.2). Probes were also selected either for regions of the genome that are now considered to be non-coding, or for regions containing proposed novel ORFs (Table 2.2). Normalisation control probes were also designed for two spiked *B.subtilis* genes; *SpoOB* (accession number M24537) and for the *lys* gene for diaminopimelate decarboxylase (accession number X17013) (Table 2.3).

Probe sequences were interrogated with BLASTn searches against the following databases: nr (non-redundant), EST (expressed sequence tag), and HTGS (high-throughput genomic sequences). Probes that scored an E-value < 0.1 were discarded, while those that showed little or no sequence similarity were selected for synthesis and microarray deposition. Viral sequences were not filtered out during the BLASTn searches, and sequence matches between the probe and the associated viral ORF were used to confirm the specifity of the 60-mer oligonucleotide sequence.

All Merlin ORFs were represented by 3'- and 5'- proximal probes, unless stated otherwise. To investigate transcription from the regions now considered to be non-coding (Dolan et al., 2003), probes were made for several previously proposed ORFs (Chee et al., 1990). Recently proposed novel ORFs (Murphy et al., 2003(a) and (b)) were represented by only a 3'- proximal probe. Oligonucleotide probes were synthesised by MWG-Biotech (U.S.A.).

Microarrays were printed at the Liverpool Microarray Facility by Dr Margaret Hughes. All viral oligonucleotide probes were printed in triplicate on Corning GAPS II microscope slides at a concentration of 40 pg /spot, using a BioRobotics Microgrid (Genomic Solutions). Control features were printed in forward (sense) and reverse (anti-sense) orientation (576 positive and 576 negative control features). A gene array list file (GAL) was generated during the printing process, to facilitate identification of probe features. After overnight incubation in a desiccator at RT, slides were briefly hydrated and snap dried on a 100 °C hot plate, and then UV cross-linked at 600 Jcm^{-2}. Slides were then stored in a desiccator at RT prior to hybridisation.

Table 2.2. List of HCMV microarray oligonucleotide sequences

HCMV ORF	5′ or 3′ Probe	Genome Position	Strand	Oligonucleotide Seqeunce (60-mer)
RL1	5′	1557	F	ACCTCTACGGTCAAACACAACCCCAACTACACGCATACCCCAACGCCAACCCACAGGAAA
RL5A	5′	5539	R	TGGGAAACATGGGATGACGGTGGAGAGTCTCTATACGATGTGACTAATAATGGTACAACG
RL6	5′	6148	R	ACGGGTATAGGTAATGCTACTGTGAATGGTAACGCAACTATTTGTGTGTCGAGCTGTGGT
RL10	5′	8686	F	CCACCAACTGTCTCGTGAAAGCAGAAAATACCCACCTGACATGTAAGTGCAATCCGAATA
RL11	5′	9162	F	CCTCACGCTTATCATAGTCACGTCGCTGTTTTTGTTCACAACTCAGGGAAGTTCATCGAA
RL12	5′	10708	F	TCAGGCAGATATAGTCGACGGTGTTACAAAGAGTACAACCACGATGGACCACACGAGGAT
RL13	5′	11426	F	GCAGCGTCGTCAACCATAATATCTCAAACCAACCGCAGTCATACGAGTAATGCTATCACA
UL1	5′	12368	F	CCCCTGCAACTTTTCTAGTTACTGGAATCACAATAACTGCGAACTTTGCGGTTGGACACC
UL2	5′	13283	R	TATCCGTCGTTCGGCAGCCTACCCGCCTCGCACGCTCAGTACGGCTTTCGACTGCTACGC
UL4	5′	13918	F	CCGTGGAATGGAAGTCTCCCGACCGTCAGATTCCTATGAATATTACTTGCGCTAATTACT
UL5	5′	14362	F	TTCTAGGCTACTCTGACTGTGTAGATCCCGGCCTTGCTGTGTATCGTGTATCTAGATCAC
UL6	5′	15514	F	GTAATGACACCCACGCTGGTTACAAACTCCACATTCAGTGTGTCACTTGTTGCGTTGAGA
UL7	5′	16197	F	CGGTCGTTACGGAGCAGAAGGTTTTACGAGAAGCGGCGAAAATGAAACGTTCCTGTGGTA
UL8	5′	16713	F	TGCAGTCAACCTCAGCAACGCGACTACGGTCATCCCACAACCTGTAGAATACCCGGCTGG
UL9	5′	17024	F	GACAATACCATGTACCCCTACAGTTGGATACGGCAGTCATAATATTAGCTTGCATCCGCT
UL10	5′	18005	F	CCTACGACCCTCCTAAACCTGGTAGACAAAAGACAATGAAAACTACTCCGTTGCCATCAC
UL11	5′	19076	F	CCCCACTACCACCACCAAAAAGCCCACTACGACTACGAGAACGACAACTACCACAACACA
UL13	5′	20778	F	GAGAAGGGAATGATGGGAGCCGAGCAAGTGGAAACGACGAGAGAAGGAATAATGCGGGAA
UL14	5′	21591	F	GCAGCGGCTGAGCCTCTCGTTTCGTCTCATCACCGAGACCGCGGGCACCTACACCTGCGT
UL15A	5′	22225	F	ATGAAGCGGATGATTCGCAGTCACGGCAGGAAAACGGAGTGTCAGATGACGGGCGCCGGC
UL16	5′	22774	F	GGTACCGCTGGGATGGGTGTTTTTTGTTCTTTGCTTATCTGCCTCTTCCCCGTGTGCTGT
UL17	5′	23593	F	CGTCACTGTCGGTTGGAAATGTTGATTCTGGACGAACAGGTGTCTAAGAGATCCTGGGAC
UL18	5′	24371	F	CCGGTTACGAGGGATTTGGGTGGGACGGGGAAACTTTAATGGAGCTCAAGGATAACCTGA
UL19	5′	25162	F	GCACTCGTGATGGGTTCTGCGGTCGGTTGACGTCCATCGCCACGAATTCCCACTACACTA
UL20	5′	26100	F	CGCTGGGAAATGAAACACGGTGCATGCTGCTCCCTAGACAGTATACTCTCAACGCCACGG
UL21A	5′	27112	R	GGAGGTAGCCCTGTTCCCCAGCTCACCACCGTCACTCAAGGACTCATGCCATCTGTGCGC
UL22A	5′	27710	F	GAGATGAAGATTACTCCGGGGGAGACTATGACGTTTTGATTACGGATACAGATGGAGGTA
UL23	5′	28700	R	TACCTATGGACCGACCACATCTACTCTGACTCGTTGACCTTTGTGGCCGAGAGCATCACT
UL24	5′	29451	R	CGCCCTTTCCTCACTTGCGCTGGCCCGTCGACCTTATCCCCATCGTCGTCGCCTATACCG
UL25	5′	30911	F	CGAGGGCGATGAGTTTTCCTTCTGCGACAGCGACATCGAAGACTTTGAGCGCGAATGTTA
UL26	5′	32915	R	ATACTGGAAGCTGCTGGTGGTGACCCAGGGTCAGTTGCGCGTCATCGGCACCATCGGACT
UL27	5′	34801	R	AGTTTTGTCACCAGTTCCTGCGCGCCTACCTGACGCCGATTCGTAATCGACAGGAGGCTG
UL28	5′	35873	R	CACGGCATCATGCTGGGCGACACTCAGTACTTTGGGGTGGTGCGCGATCACAAGACCTAC
UL29	5′	37151	R	CTTGGCCGACCTGTGCGTGCGACGCGACTATGAGGGCCTGCGGCGTTACTTGCGGCGCTT
UL30	5′	37605	R	ACCCGGAGCCCATCCAGAGCGAGACGGAGGGGGAGAATAAACAGTTTACGGAGCACACAC
UL31	5′	38530	F	CAGGAATGAGAGAAGAGAAAACAGGGAGGACGAAGGAGGGGAGACGAGGACGACGGGCAC
UL32	5′	42492	R	ACACGCTGCTGGAGCTGCTCATCGAGGACTTTGACATCTACGTGGACAGCTTCCCGCAGT
UL33	5′	43735	F	AACCCCCACCATCTACATGACCAACCTCTATTCCACCAACTTCCTCACGCTCACCGTACT
UL34	5′	45141	F	ATGAACTTCATCATCACCACCCGAGACTTCTCCAACGACGATTCAGTCCTGCGAGCCGCC
UL35	5′	47728	F	ACTTCCACGAAGGCAAGATCACCACCGAGACGTATCACCTCCAGCGCATCTATAGCATGA
UL36	5′	50014	R	GGGACTTCAACGGTCTCAACGACTTTCTGGAGCAGGAATGCGGCACCCGGCTGCACGTGG
UL37	5′	50871	R	GGCTGGTTTCCCCTGCGCATCCTGAACGAGAGCCGCGATATCAACGTCACCGAGGTCTAC
UL38	5′	52350	R	GCAGATGGACGTGGGGGGACTGATCCAGGCGTCGGCGTTGGGCAAGGTGGCGCTTCGCTA
UL41A	5′	NA	NA	3'-oligonucleotide only
UL40	5′	53980	R	CGTGTACCACCAAGGCGACATGGTCCTCATGACGCTCGACGTGTACTGCTGCCGCCAGAC
UL42	5′	54972	R	TTTCCACGCCACCGCCGCCACCACCCGACTGCAGCCCACCGCCCTATCGACCCCCGTACT
UL43	5′	56059	R	GCTGTCGAGCTGGTCAACGAGACCTTCCGCTGCTCTGTCACCTCCGACGCCCGCAAGGAC
UL44	5′	57553	R	ACAGCTCATGCCTCTACATCACTGACAAGTCGTTTCAGCCCAAGACCATTAACAATTCCA
UL45	5′	58563	R	TCATGCGGTTAGGGTTCACGTACTTTGCCTCTTGGGACTTGATCGAGCGCATCTTTGAGC
UL46	5′	61227	R	GCTGGGCTGCGTCAAGACCGTCTCGCTGGGCATCACCTCGTTGCTCACCTGCGTCATGCG
UL47	5′	62436	F	GCCGTGGAGTTTCAAAACTACGTCAAGAACAGCGTGCGGCACATGAGCTCTTTCGTCAGT
UL48	5′	65203	F	GCCCGCTTTCTACTATGAGGCCCTTTTTCTCTACATGCTGGATGTGGCGACCGTACCAGA
UL49	5′	73130	R	GCAGCAGCTGGTGCATCACGTGCCACGTTCTTGCGTCCTTCATCTCTTCGTGACGGATAA
UL50	5′	74395	R	TCTGTAAAAATCCCAATTACTCGGTGTGCGACGCCATGCTCAAGACAGACACGGTCTATT
UL51	5′	74848	R	GCGGTGGGCTCGAGGGGGGAGGTGGTGACGACGAGGACGGTGAAGACGGACACGCGCTAC
UL52	5′	75695	F	CGCCATCATCAGCATCTGCCTCAAACAGGACTGCGACCAGAGCTGGCTCCTCGAGTACAG
UL53	5′	77342	F	TCACGGGCAAAGAGTCCATCTGCTTACCCTTCAATTTCCACTCGCACCGGCAGCACACCT
UL54	5′	81424	R	GCGTCAACGTTTTCGGGCAGCGCAGCTACTTTTACTGTGAGTACAGCGACACCGATAGGC
UL55	5′	84697	R	GTTAACTTGTGTATCGTCTGTCTGGGTGCTGCGGTTTCCTCATCTTCTACTCGTGGAACT
UL56	5′	87024	R	GGAGATGCTGCTGAAGGCGCTGCACGAGGATACGGCTTTGCTGGATCGGGCGCTGATGGC
UL57	5′	91311	R	ACGGCGGGGTGCTCACGAAAGTCACGTCCTTCTGTCCTTTTGCGCTCTACTTTCATCATA
UL69	5′	101168	R	GCCGAGTTCCCAGTTGTCCTACGACTTTCGGTTCTTCTCACCCCTCCTCAGCCAATAATC
UL70	5′	103725	R	ACCGAGCTGCTGGACGTCATGCAAAAATACTTCTCGCTCGACAACTTTCTACACGATTAC
UL71	5′	105022	F	AATCGCAGCGTATCCTAGAGGCCCTGGACATTCTCATCCTCAAACTGGTGGTGGGCGAGT
UL72	5′	106772	R	ACATTTACGCTGCCCCAATCGACCGAGGAGAAATACGACAAAGAGCAGCATCCGGGAGAG
UL73	5′	107203	F	TACGGCGACAAGTACTTCATCGACGACTAGTGCCAAACCTGGTTCCACTACTCACGACCC
UL74	5′	108669	R	AGACGTTCCTGGCCGTATACCGTGCTATCTTATCGAGGTAAAGAGATTCTGAAGAAACAG
UL75	5′	111126	R	TTCCATATGCCTCGATGTCTTTTTGCGGGTCCTCTGGCGGAGCAGTTTCTGAACCAGGTA
UL76	5′	112187	F	TGTACCGATTCTGTTCTGGACTATCTGGGACGGCGTCAGGATGAGTCTGTTGCACACCTT
UL77	5′	112400	F	ACGGTATCTGCGTCAGGAGCTGAGGGATCTGGGTCACAGGGTACAGACTTACTGCGAGGA
UL78	5′	114530	F	GCGCTCTAAATCTCAGCCTCTGTCGCTTAGTGCTCTTTGTCGACGACGTGGGCCTATATT
UL79	5′	116168	R	GGACCACGCCGTGCTTAGTCGCAAGACGCCGCAACCCTACTGGCCTCATCTGTACCGCGA
UL80,UL80.5	5′	118165	F	GCGGGTTCCAATCAGCAGCAGCAACAACGTTACGATGAACTGCGGGATGCCATTCACGAG
UL82	5′	120179	R	AATCTCCGCACGACACCGTAGACCTGACCGACTTAAACATCAAGGGCCGCTGCGTGGTGG
UL83	5′	121477	R	GCAACGGCTTTACGGTGTTGTGTCCCAAAAATATGATAATCAAACCGGGCAAGATCTCGC
UL84	5′	123531	R	GCGCTGCATCTGCCAGAACTCACGTTCGAGCCGACGCTGGATATAAACAACGTAACGGAG
UL85	5′	125182	R	GCCGTTGTGCCCATTCCGCAGCGTCTACATCTCATCAAGCACTACCAGCTGGGCCTACAC
UL86	5′	128430	R	TTGCGGCGCGTCTACAAAAACACCGATACCAAAGATCCGCTAGAACGTAACGTGGACCTG
UL87	5′	123032	F	CGTAAAAGCGGCGGTAGATACGTAACGTGTGCGAGTTTTTCAGCGTCAATTCGTAAGGGA
UL88	5′	132631	F	CGTGTGTTACCGCCGCCGCCCGACGCCGAAGACGACGTGGTTTTTGCTTCCGAGCTGTGT
UL89	5′	139141	R	CATCTTTTTTATAGCCGTGACCAAGATCCCCGTGCTGGCCAATCGCGTGCTGCAGTACCT
UL91	5′	135171	F	ACTGGGGGTCGCGCACGCCACTACGGAGGATGTTTTTATCTTTGTCGACCGCCTCTTTCA
UL92	5′	135664	F	TGTTGGAACCCATCGAGGAGGCCGCACTGGACGACGTCAACATCATCAGCGCCGTGCTCA
UL93	5′	136290	F	GGCCTGTTTTCGTGTACTACCGCGCCCACTGGAGTTGCTCGATTACCTACGTCAAAGCGG
UL94	5′	137939	F	CGCATCTGTCATTTACCTACCTTGTACCAACTGAGCTTCGGAGGTCCTTTGGGTCCAGCC
UL95	5′	139707	F	GGCGGTGGTGCGAGCGGAGGTTAGGCGGCAGCGGCGAGAGGAGAGGAAAAAGATGGCGTC
UL96	5′	141323	F	GGACGTGATGCGCGTCGACCTTGAGCGACAGCAGCATCAGTTTCTGCGGCGTACCTACGG
UL97	5′	142736	F	CGTGCCCAAAGAGGACGATTTTTGCCACAAGATCTGTTATGCCGTGGACATGAGCGACGA
UL98	5′	144089	F	GCCCCTCAGTCTCTTTCTCATGAACACCTTTTTGCTGCACCAGGAGGGCTTCCGTAATCT
UL99	5′	145864	F	GGATGACGATAACGAGGAGCGGCAACAGAAGCTGCGGCTCTGCGGTAGTGGCTGCGGGGG
UL100	5′	147366	R	TTGGTGTGCTACGCCGTGTTCATGCAGCTCGTCTTTTTAGCCGTGACCATCTACTACCTG

HCMV ORF	5′ or 3′ Probe	Genome Position	Strand	Oligonucleotide Seqeunce (60-mer)
UL102	5′	149545	F	TGGTCTGGGGCGACGAGCGGTTGGTGGGTCCCTTCAACTTCTTCTACGGCAACGGCGGCG
UL103	5′	151283	R	GCGTGCTGGAGGTCCATACGGATTTCACTAGACAGAATGTGATGATCATGGAGCCGCAAG
UL104	5′	152874	R	TCCGCGACATTTTATATCAGTACGCCGATAACGACGACTATGGTCTCTACGTGGATTGGT
UL105	5′	154295	F	AACAAGCGTTGCACCGACCTGGACTTTGGCGACCTGCTCAAGTACATGGAGTTCGGTATC
UL111A	5′	161009	F	TCGGTGATGGTCTCTTCCTCTCTGGTCCTGATCGTCTTTTTTCTAGGCGCTTCCGAGGAG
UL112	5′	162293	F	AGAACGGTAATTTGCAAGTGACTTACGTGCGGCATTACCTGAAAAACCATGGCGAATCCT
UL114	5′	164777	R	ACCGCTGGAGCTATCTGTGTCCGCCCGAGCAGGTGCGCGTGGTGATCGTGGGCCAGGACC
UL115	5′	165661	R	TGCCCCGAACTAACGCGCCGATGCTTGTTGGGTGAGGTGTTTGAGGGTGACAAGTATGAG
UL116	5′	166428	R	ACCATCGGTACGAACGCTACTTCCCCCTCCTCTTCTGTATCCATACTTACCACCGTGACA
UL117	5′	168003	R	CTCCCAGGACCACGTCCAGATCGTCTACGGCTCCACACGCATCTGCAAGTCCCTGGCCCC
UL119	5′	169196	R	GGGAGTGAAAAGTAGCACCACAAGCGCCGTCACTTCCCCTAGTAATACCACCGTCACGTC
UL120	5′	169858	R	TGGGACTCGGTGACGGTGGCAACTACAATTAGGGTAGGGTGGTGGTATGAACCTCAAGTA
UL121	5′	170191	R	CACGCTACGGCTGCTGATACGAGATGGGATGTACGGGCGGGGGGAGAAAGAGTTGTGCAT
UL122	5′	173721	R	CTTTGAACAAGTGACCGAGGATTGCAACGAGAACCCCGAAAAAGATGTCCTGACAGAACT
UL123	5′	172481	R	ACAGTGATCAGGAAGAAAGTGAACAGAGTGATGAGGAACAGGAGGAGGGTGCTCAGGAGG
UL124	5′	174205	F	TGCCCTCCACCGTTAACAGCACCGCAACAGGAGTTACCTCTGACTCTCAACAGAACACAA
UL128	5′	176805	R	AGAAGAATGTTGCGAATTCATAAACGTCAACCACCCGCCGGAACGCTGTTACGATTTCAA
UL130	5′	177491	R	CCGTGGTCGACGCTAACAGCAAACCAGAATCCGTCCCCGCCATGGTCTAAACTGACGTAT
UL131A	5′	178086	R	TGCGGCTGTGTCGGGTGTGGCTGTCTGTTTGTCTGTGCGCCGTGGTGCTGGGTCAGTGCC
UL132	5′	178771	R	ATGAAAGTGCTGGCTATCCTCTTCTACATCGTGACAGGCACCTCCATTTTCAGCTTCATA
UL148	5′	179714	R	TTTCTAAGCTTTACGTGCCGCCTGCAGCTAGAACCCGTGGTGGAAAATGTTGGCCTCTAC
UL147A	5′	180268	R	ACTAGTATCCTCGCAGAGATTAACGAAAATTCCTGCTCCTCATCTTCTGTGGACCACGAA
UL147	5′	180722	R	TACTTGCCATATGTCCCATAGCCGTCCGTCTTTTAGAACTAGAAGATTACGACAAGCGGT
UL146	5′	181177	R	AGTGCATGGAACGGAATTACGCTGCAAATGTCTTGATGGTAAAAAACTGCCGCCCAAAAC
UL145	5′	181725	R	TGGCTACATCGCAAAGAAAGTCCCTAGTGTTACATCGATACAGTGCCGTGACAGCCGTGG
UL144	5′	182559	R	TGGGACAAAGAGTTACTAAAGTATGCACGGAGCGTACCAGCGTAACGTGTACCCCTTGCC
UL142	5′	184213	R	GGGAGCGAACGGTCATTATCCTATCGTTACCACCTGGAATCTAATTCATCTGCCAACGTG
UL141	5′	185001	R	TCTCGCAGTTGCATCTCGGCCAAATATTCTCGCTGACTTTCAACGTATCGACGGACACGG
UL140	5′	185957	R	GACGAGAGTCGTACTAACAGCGTGTCATCGTACGTTCTTTTATCACCCGCGTCCGATGGC
UL139	5′	186786	R	GCATCGGCGAGTGAAACCACTACAGGTACCAGCTCTAATTCCAGTCAATCTACTTCTGCT
UL138	5′	187855	R	ATCGGCGTGATGCTCGTGCTGATCGTGGCCATTCTCTGCTATCTAGCTTACCATTGGCAC
UL136	5′	188662	R	CGTGGGACTTGGACGTTGGAAATAAATGGCGGCGTCGAAAGGTCCTGAGTCGCATTCACC
UL135	5′	189443	R	TGTACGTCGCTCTACATAGGAGAGGATGGTCTGCCGATAGATAAACCCGAGTTTCCTCCG
UL133	5′	190595	R	GATAACATGCGCGGCTTTAGGAATTTGGTGTTTGGCGGGATCGTCGGCGGATGTCTCTTC
UL148A	5′	NA	NA	3'-oligonucleotide only
UL148B	5′	191358	R	CTAGCGGGGCTGGCTATCTGCGTGGGTCTAGTCATGGGGGTTACCGTGATCGCGTCGTGC
UL148C	5′	191545	F	GGCCGTGCTCTACCTCCTGGCGCTTGTCGTCTGGGTTGAGATGTTCTGCCTCGTCGCCGT
UL148D	5′	192228	F	GGATCAGCGTTGCCATCGTCATCTTCATCGGAGTCTGTCTGGTGGCCCTGATGTACTTTA
UL150	5′	193497	R	CGCCCCAACCTCACCTGCTGCGGCTACCACACATCGACTGTCGTTCCCTGAAGGATCGAG
IRS1	5′	197707	F	GGTCTCGACCTCACCACCGTCATGACAGAGCGTCAAAGTCAATTGCCCGAAAAGTATATC
US2	5′	199485	R	CTGGCTGCGTTGCGTCCCCGAACTGCGAGTGGATTACACGTCTAGCGCGTACATGTGGAA
US3	5′	200716	R	GCTGGTTCCATATGGGCATGCTGTACTTCAAGGGGAGGATGTCGGGCAACTTCACCAAGA
US6	5′	201862	R	GCCTAAAACGCACAAATCCCGTCCGAACGATAGGAACCTTGAGGGTAGGCTGACCTGTCA
US7	5′	202855	R	CCCCCGCCGACATCGCTAAAGTACTCATCTCACTGAAACCCGTGCAATTGTCTTCTGGAC
US8	5′	203652	R	CCCATGTGGTGCCAGCCTCGGTACCATATACGTTATTTCAGTTACGGGAATTCAGTGGAT
US9	5′	204839	R	TCCCCGTCCACCTGTTCCTTCTTCTGGCACTGGTGTCTGATCGCAGTAAGTGTACTCTCG
US10	5′	205722	R	GCCACGGAGGAGACGAGAGAACCTACCTACTTCACGTGCGGCTGTGTTATTCAGAACCAT
US11	5′	206480	R	CGGTCGCGGGTAGTATGCCTGAATTATCCTTGACTCTTTTCGATGAACCTCCGCCCTTGG
US12	5′	207507	R	TAAGAAAGAGAAACCCCCACCCGTGTCGCCGCCGTCACCACCGCCGATCCGTCGAGTCAC
US13	5′	208224	R	TGTCAACATGAATCTTCCCCTCTGCTGCTGGTTTTTGCCCCCTCCCTTCTCTGGTGTTTG
US14	5′	208814	R	ACTTACACCTGGTTGCATAAGACTCTGCTGTGTCTCTACACCGTGTTCGTGGGCTGCATC
US15	5′	209901	R	TGCTTCACGTGCCGCCGTTGATCGTAATATGTCTCTGCCTGGACGGTACGTTGGTCATCT
US16	5′	210769	R	GTCGTTTGTGCTCACGTTGGTCATCACCTTGCTCTGCGCGGAAGTAGTCTTCATCTGTCG
US17	5′	211880	R	GCGCCTGCTCCTCCTCTACGCTGGTGACCTTCTCGGGCCTCTTGGCTTGCGTGCTTTTTA
US18	5′	213053	R	TTTCGCCCACACACCACCAATTTTGCTCATCTCACAGTGGCGTGCTTGCTGATCAACACC
US19	5′	213660	R	CGACGACGCTGTTGACGGTGACGCTGATGCTAATTCACGATCTGTCCCTGATCACCTGTC
US20	5′	214680	R	AGTAGGTTCCTACGTGATGACCCTGGCGTTGTTTATCTCCTTTACGGGGCTGGCGTTTCT
US21	5′	215708	R	GTCTGACGCTGTCTTGCTTAATGATCTCCATCGTGTGCCTAGGTCTGTTGCGTTGGGCCG
US22	5′	217512	R	CGAACAGGCCGCTATCCCGCAGATTTACGCACGTTCTCTGGCCGCCGACTATCTCTGCTG
US23	5′	218843	R	CCGTTTCTGTCGTTTTTACGTGGTAATCGTCAAGTCGCACCTGGACCGCTCGCCGCCCCT
US24	5′	221160	R	CGACAAGTGTTGCGGGATTTTAAGGAACTGTTCTTTTGCCTGGAGCCAATGGAAATCACG
US26	5′	223618	R	GCAGGACATCCGCCACTTAGTTCGTAGCTATGCTGATATGAACATTAGCCTTCCCGTTTC
US27	5′	224293	F	TACCACCATCCTGTACTACCGTCGTAAGAAAAAATCTCCGAGCGATACCTACATCTGCAA
US28	5′	225543	F	TGTTTCTGTACGGCGTTGTCTTTCTCTTCGGTTCCATCGGCAACTTCTTGGTGATCTTCA
US29	5′	226666	F	GTGTTTCCGATGGTGGCTCTACAGTGGGTGGTGGTGGCTCACGTTTGGATGTGCTCGGAC
US30	5′	228324	F	GACGCCAAGTACGGTATTCGTAACCAGCATTTGAGTATACGGTTAATGTATCCTGGGGAA
US31	5′	229084	F	CGTGCGGTAACTGGCAAAGCCACGTTGAGATTCAGGACGAGGAGCCCAACTGCGAGCAGC
US32	5′	229711	F	AGCACCTCTATTGTGTGTACGACAGCCATTTTCAGCGACGACCCACGACCCGAGCCGAAC
US34	5′	230802	F	GGACTCGTTTATCGTGAGCTTCATGATTTCTACGGGTATCTGCAGCTGGACCTTCTGGGA
TRS1	5′	233004	R	ACCCTAGCCACCAAACATCCGGTCATCTGTGCAAATGTGGAAGACTATCTGCAAGACGCC
RL1	3′	2111	F	GCTCGGCCAGATCACACGCTCTTTCCTGTCCCTTCTACACCCTCAGCCACGGTTCACAAT
RL5A	3′	5472	R	ATACAACAGCCTGTGTTTCAAGTTGTTCGCATACGTCGCTTGTGCTTTGCAATATGACGC
RL6	3′	6025	R	TGGTTCGGTGAAAACATGGATGAATACAGTGGTGATGTGTGGCACTTGGAAGTCAGCTAA
RL10	3′	8969	F	ACGCAAAAAGCTGGAACAACACGCGGCTAAAAAGCAGAACATCTACGAACGGATTCCATA
RL11	3′	9786	F	GCTGCGGACGGACCTGGATACGGAACCTCTGTTGTTGACGGTGGACGGAGATTTGGAATA
RL12	3′	10801	F	GACACTGACAATCATACCACCCCACTATGTCCCAGATACGTAGGAACACAATCAGAAGAA
RL13	3′	11786	F	AGCCTCACGGTGATTTTTGCCAACACGACATCCATAGTCACCAACAGAACGTGTGATAGA
UL1	3′	12705	F	GTAGCCACTCACGTTGGTTGGACAGCCACCGTGGTGATAATTATCTGCGTTTTAACTTAC
UL2	3′	13211	R	ATCATGCTTGTCATTTGGACCGCAGTGTGGCTCAAGCTGCTTCGAGACGCTCTTTTGTAA
UL4	3′	14148	F	TGTACCTATGACCGCCTGACGTTGCTGAATTTGACGACGGAAAACAGCGGAAAGTATTAT
UL5	3′	14518	F	AGAGTAACAAGGCTATATCAAGGGACGAAGCAGCATTCACCTCCAGCGTGAGCACCCGTA
UL6	3′	15689	F	GCGTTGCGATGCTGCTGTTTGCCGCCGTGATGGTGCTCGTTGATTTGGGTTTGCCTCAAT
UL7	3′	16382	F	CACAACTACGTGCCGTCGCTGGATTCTCCCATCAGACGCCTTTGGAAAACAACACGCACC
UL8	3′	16854	F	CTGTCTACGAGCACCTCGAAAAATCTACCATCACTGGAAAGACAGTAAACAGTACGGACA
UL9	3′	17357	F	CCACCAAACACATTGCCATACAACTACATCATCCCTGTATCCACCTACATCTGTACACGA
UL10	3′	18109	F	GTAAACGTCTCAGATCCCAGAAACTATTGCAGGCGAAAATGTCCACCAAAGGGTAACTGT
UL11	3′	19158	F	ACCACCACCGCCAAGAAGACGACGATAAGCACTACCCATCATAAACACCCCAGTCCCAAA
UL13	3′	20901	F	TGAATAGGAGGTCACAGGAGGAGAACGAGGAAGGTGGACCGTGTGAATCGCCGCCAATGA
UL14	3′	21883	F	TGCGAGGAGGACGAGCGCGAACTGTGCGTGCCCTTCATCAGCCAGAGCATCGCGGACAAC
UL15A	3′	22455	F	GGCGGACATTGGGCTGTGGTTCATGTTTCTGGTTTTTGGACTGTGTTCGTGGTTGGCGAT
UL16	3′	23140	F	GGTCACTAATCTTACCGTGGGCCGCTATGACTGTTTACGCTGCGAGAACGGTACGATGAA
UL17	3′	23770	F	CACGATTTGTGCTGGCTGTTTCGACGACTTTTCTTTCCTCGGGAGGACTCAGAGCCACTG
UL18	3′	24915	F	TCACCCGACAGCTAACAAACGCTATAACACCATGACCATCAGCAGTGTCCTCCTAGCCCT

HCMV ORF	5′ or 3′ Probe	Genome Position	Strand	Oligonucleotide Seqeunce (60-mer)
UL19	3′	25249	F	GCATCGAGCGACCGAGTCAGTCAGAAGTGGATTGCGCCAGTTTAATGGAAACGCTGAAGC
UL20	3′	26562	F	CGTTGCCTCCACGACCTCCCCACTTAATAGAATTCCCGCCGTCTCCTCCGCCATCGCCTG
UL21A	3′	26883	R	ACCGACCTCCGCAGCAACCCCTACCCAATCCGCTGGTGCTACTGCTGGACGATGTTCCCC
UL22A	3′	27782	F	CACAAGAGAAGACCGACGAACACAAGGGAGAACACACCAAAGAAAATGAAAAGACCCAGT
UL23	3′	28620	R	GCGACGCTGCAACTTCATCACCGTCCCAGAGGAGCTGCCGCACACCGCCTCGCTTAGAGC
UL24	3′	29346	R	TGGGAGATAATCTCAACACGTTCATGTGTCTGGGACTTAACCTGCTGTTTGAAAACCGAC
UL25	3′	31720	F	AGAGCATCTCGGAGCTGTGTTACCTCATCTATGTACAACTGCCGTCGTTGCGCGAAGACT
UL26	3′	32574	R	TGACGCCGGGCAAGCAAGAAATCACCGACGCGATGTTCGAAGCCGGCAACGTGCCATCAG
UL27	3′	33301	R	CGGACGCCGAGCTGAGCAACCACGCCAACCGCTGCCGCCGCAAGGCCCCGCTGGAACTCG
UL28	3′	35549	R	CACACCTTTGCGGGCATGTACGAACTCTCCCAGATACTGCACGACCGCGCCAATCTGCTG
UL29	3′	36334	R	CGCGTGGGGCTGGCGCTGCTCATCGACGACTTTCGCTACGAGAGCATCGGCCCCGTGGAC
UL30	3′	37515	R	TCTTTGTATTTCCCTGTCTTGTGTTGCCCAACTGCTGTCAGGTCTCCGTAGATCGCTCCC
UL31	3′	39443	F	GCGTCCAAAAACGTACAGTACGAATTCATGGGTCTCATTTTCACCGTCAACGTGGATTCA
UL32	3′	41402	R	TCAGACGCCCGAGCAGTCGACGCCGTCCAGAATACGTAAAGCTAAGTTATCGTCTCCAAT
UL33	3′	44540	F	CAGAATCAGCGAGCTACAGCGGAGACAAATCTAGCGGCTGGCAACAATTCACAATCAGTT
UL34	3′	46295	F	GACCCCCTCACCTTCAGTCACCTCAGCCCTTACCCCCGTCACGTCCCCCATAACCCCGTT
UL35	3′	48064	F	CTTCTTCTACCACGGGTCTCTTACTCTCCTCTTCTTTGTCGGGGTCGCACGGCATTA
UL36	3′	49139	R	GGCCGAGGACGTGGTTATGTTCACCTGCGTCATGGGAAAGAAGGGACACCGAAACCACCG
UL37	3′	50500	R	TAACGTGACGTCGCTTCCCGAGTGGACACTACAACAATGCCAGGGGTGGAAATATTGGAC
UL38	3′	51483	R	GAAGGATGTGATGTACACGGCCGGGGAGGGCGACGTGGTACAGATGGTGGTCGTGGTCTA
UL41A	3′	54537	R	CGTCGACATGAACGTCTGCATCGGCATGATCGGCGTGGTTTGTTTCGTTTTCGGGGTCTT
UL40	3′	53889	R	CATCCTGCCGATAACACGTTGCTGATCGAAGTGGGTAACAACACGCGTCGCCACGTAGAC
UL42	3′	54743	R	CGTGGTGATTAACCGGGACAGTGCCAATATAACAACGGGGACGCAAGCATCATCGGGGTA
UL43	3′	55571	R	GTGCATGAAGCGCGAGGGTTCTATCTTCAGCTGGCGAGACGGTAACGAGGCGCTGACGAC
UL44	3′	56519	R	GCGAAAGAGGAGAGCGACAGCGAGGATTCTGTAACGTTCGAGTTCGTCCCTAACACCAAG
UL45	3′	57553	R	ACAGCTCATGCCTCTACATCACTGACAAGTCGTTTCAGCCCAAGACCATTAACAATTCCA
UL46	3′	60895	R	CTCTGTGTACCGGGGTCATACAAAAGCTGGGCTGGTGCCTGGCTGACGACATTCACACCT
UL47	3′	64006	F	GCGTACGTCAGATCAGCAGCGACACAGAGGAAGTAAGCCGACACGCCAAGGGTCACCGCA
UL48	3′	71227	F	GGTGCTCACGCAGAACGTCCTCAGTGCTCTAGAAATATTGCGACTAGTGCGATTGGACCT
UL49	3′	71720	R	CCACGTCGGCGGCCACATTGTGCAGCAGGCGCGCGGCGGTTTAGCAGCGACGTCCTCCTC
UL50	3′	73753	R	TGGCGGCGGGGTCGTGGCGGCTGTGTCGTTGTGAGGCCTGTATGGGACGATGCGGATGCG
UL51	3′	74601	R	GCAACTATTAACGGCGATCGTCAGCAAGCTCATGGACATTAACGGAATCCTGGAGGGAAA
UL52	3′	76526	F	CCGCGTCGTCTCGTACTGTCAGAACAATGTGAAGATGGTGGACCGCATTCAGCTGGTATT
UL53	3′	78132	F	CCACCAGTCTCAGTCTCAGCAGCATCATCACCGTCCCCAGTCACCACCGCCGCCGCTGTT
UL54	3′	78689	R	TCGTCGGTGCTTTCTAAGGACATCTCGTTGTACCGTCAATCTAACCTGCCGCACATTGCC
UL55	3′	82146	R	AGCAGAACGGTACAGATTCTTTGGACGGACGGACTGGCACGCAGGACAAGGGACAGAAGC
UL56	3′	84991	R	AGTCCAAAGATCTATATCTCCTCCTCTACAGGCATCTGTCCAGACCGGATGAGAGTGGCG
UL57	3′	88286	R	GGATTGCGACCCCGAAGTGATGGCCGTGTACGAGATTTTGAGCGTGCGTGAGGAGATCCC
UL69	3′	99671	R	CTTCTCACGGCGGCAGTCCACCCCAGGTTCCCCACAAGCAGCCCATCATCCCGGTGCAGT
UL70	3′	101881	R	CACCAAAAAGTGTCACTCCAACGCCAAAAACGTCCACATTTCCATCAAAATCAGACCGCC
UL71	3′	105147	F	CTGGTGGAGGTGCAGCGCGTTATGGGCCTGGTGGACATGGACTGTGAGAAAAGCGCGTAC
UL72	3′	105913	R	TCGACGAGACAGAAGAAGATGAAAAGAGCTCCGACGCAGAAAGCCCTGTCAACACGAGTG
UL73	3′	107345	F	TTTGCAGCCTGGTGGACTATGCTTAACGCTCTCATTCTGATGGGAGCTTTTTGTATCGTA
UL74	3′	107756	R	ATCGGAATGAACCCTTTTGTAAACCAGACCGTAACCGTACCGCCGTGTCAGAATTTATGA
UL75	3′	109364	R	TCGCCCTGGATCCCTACAACGAAGTGGTGGTCTCATCTCCGCGAACTCACTACCTCATGC
UL76	3′	112404	F	TATCTGCGTCAGGAGCTGAGGGATCTGGGTCACAGGGTACAGACTTACTGCGAGGATCTC
UL77	3′	113890	F	GCAAGAGCAACCCCGTGGCCGACTACATGTTCGCGCAGAGCTCCAAACAGTACGGCGATT
UL78	3′	115228	F	AAGCGGGCTATGTATAGCGTGGAGCTGGCCGTGTGTTACTTTTCTACGTCCGTCCGAGAC
UL79	3′	115956	R	TCCCCGAAGACCTGGCGCGCAACGGCAACATCCTCTTCTCCCTAGGCACGCTCTACGGAC
UL80,UL80.5	3′	118302	F	CCCAACTACTACTACCGTGTGTACTCCCACCAGCGAGCTGACGAGTGGCGGAGGAGAAAC
UL82	3′	118789	R	ACCCTCACCCATCAGCACCGCTTCCACTTCCAGCACCCCACGCAGTCGACCCCGCATCTA
UL83	3′	121207	R	GCGGGCCTCAGTACAGCGAGCACCCCACCTTCACCAGCCAGTATCGCATCCAGGGCAAGC
UL84	3′	123033	R	GTCCCTTACGAATTGACGCTGAAAAACTCGCACACGTTACGTATCTACCGCCGCTTTTAC
UL85	3′	124763	R	GCAGCAGCTGCAGACCATTACTTTTCGCGACGCCACCTTCACCATCCCCGATCCGGTCAT
UL86	3′	125722	R	CAACAAAACGCTCTTCAAAACCATCGACGAGTACCTGCTACGCGCCAAGGACTGCATCCG
UL87	3′	132033	F	CACAGCGGGGTCTCTTTGACTACAGCAAGAACCTCACGGCGCACACCAAGATCAAGCACA
UL88	3′	133332	F	TCTTTGTCAAGTCTGTCTTTACGAGTTGGACGAGGACGAGATGGGCGAGGAGATGCTGGG
UL89	3′	133945	R	CTACATCAAAGCCTCGCAAGAGCTCGTCTCCTACACCATCAAGCTGAGCCACGACCCCAT
UL91	3′	135363	F	CGTCGTGGAGACAAGCAACACCTGTCGTCCCTGCCCTTCTCCTGTTCCCTCCGCCCCCAA
UL92	3′	135910	F	AGCTTATCATCGGTATCTACTCAAAGCAGACCAAGTACGACGCGTGTGTCATCAAGGTTA
UL93	3′	137245	F	GCCTTGGCGCTCTGGATGGATCGGGCGGACGTGCGTAGCTGTATTATTAAGGCGCTAACG
UL94	3′	138527	F	GCGGCGCAGGTCACACTGGGAGACGGCCTGGATTATCACATCGGTGTCAAGGATTCTGAG
UL95	3′	140457	F	TACCTACGACGGCGAATTAATCTACGGCAGTTACCTGTTGTATCGCAAGGCTCACGTGGA
UL96	3′	141535	F	GACAGCTGCAATCCGGTAACGTGGACGACGCGCTGGACTCTTTAACCGAGCTGAAGGACA
UL97	3′	143108	F	CACGTTGGCCGACGCTATCAAATTTCTCAATCACCAGTGTCGTGTATGCCACTTTGATAT
UL98	3′	145602	F	ATCGTCACGCCCGTGGTCTTTGACCCTCAGTTTACGCGCCATGCCGTCTCTACCGTGTTA
UL99	3′	146183	F	CGGCCCGACACACCCAGGACTCCGCGCCAAAAGAAGATTTCACAACGTCCACCCACCCCC
UL100	3′	146706	R	ACCGGCATCAGCTGGTCGTTCGGCATGCTCTTTTTCATATGGGCCATGTTTACGACGTGT
UL102	3′	150120	F	TTGTCGTCCTCCACCGCTTCCACCACCACTTGTTCTTCTTCTTCCGTTCTCTCCTCCGCC
UL103	3′	150637	R	AGACGTGCTCAACGCCAACTACCCCAATTTGCCCTCACCCCCCAAGCTGCCACCGCGCTG
UL104	3′	151783	R	AACAGCTTCTTCTCGCAGTTCGTGCCCGGCACCGAGTCGCTGGAACGCTTCCTCACGCAG
UL105	3′	155655	F	GCTACACCAGCGACAACGTGCTCAGTCTGCCCAGTGACCGCCACCGCATCCACCCCGAGG
UL111A	3′	161574	F	TGATTAGTCGGCTGTCTCAGGAGGCGGAAAGGAAATCGGATAACGGCACGCGGAAAGGTC
UL112	3′	163711	F	TCCCCGGCCCGTGTGTGAAATCAAGCCCTACGTGGTAAACCCCGTTGTCGCCACCGCCGC
UL114	3′	164511	R	TCTCACCGTCACTTAGGCTGGCAGACGCTGAGTAACCATGTGATCCGGAGGCTGTCAGAG
UL115	3′	165511	R	AACTCCGTGCTGTTGGACGAGGCTTTCCTGGACACTCTGGCCCTGCTGTACAACAATCCG
UL116	3′	166195	R	AGACATTGCGGTGGATGAAGAAGAGCCAGAACTGGAGGACGACGACGAGTACGACGAACT
UL117	3′	167026	R	CCGTAGCCACGTTCAAATTCTTTCACCAGGATCCCAACCGCGTTCTGGACTGCATCCGTC
UL119	3′	168187	R	GAGACTGGAGGAGCCCGTTGAAGAAAAGAAACACCCGGTGCCCTACTTCAAGCAGTGGTA
UL120	3′	169480	R	AACACATTCATCCACTTTGTCGAAGATAGTGAGCTGGTGGAGAATCCCGCATACTTTAGG
UL121	3′	170090	R	TCGACGCCGACCTAGGACTCCTGTACGCCGTCTGTCTGATCCTGTCGTTCAGCATCGTCA
UL122	3′	170935	R	ACCCACCAATTGTGCCCCCGCTCTTCTGATTACCGCAACATGATTATCCACGCTGCCACC
UL123	3′	172428	R	CAGGAGGAGCGGGAGGACACTGTGTCTGTCAAGTCTGAGCCAGTGTCTGAGATAGAGGAA
UL124	3′	174306	F	TCTGCGTCAGCGGTGCAAGCCCATTCCCCGAGCTCATTTTCAGACACATACCCTACCGCC
UL128	3′	176542	R	GCGGGATCGTCACCACCATGACCCATTCATTGACACGCCAGGTCGTACACAACAAACTGA
UL130	3′	177039	R	GAGCTGGGCTCACGTCTTCCGGGACTACAGCGTGTCTTTTCAGGTGCGATTGACGTTCAC
UL131A	3′	177844	R	GGGAGCGAACGGTCATTATCCTATCGTTACCACCTGGAATCTAATTCATCTGCCAACGTG
UL132	3′	178464	R	AACATGATCCTGAGAACGTCATCTATTTCAGAAAGGATGGCAACTTGGACACGTCGTTCG
UL148	3′	179611	R	TTACACCGCAGGTAGACGTGGTACGCTTTGCTCTATATCTAGAAACGCTCTCCCGGATCG
UL147A	3′	180160	R	GCTCTGCTGGCCTTTTCCCTAGTGATTTGCGGTACGCTCCTCGTCACTTGTGTGATCTGA
UL147	3′	180391	R	AGCCTTTCTCTGGAAATACGCCAAGAAACTGAATTACCACTACTTTAGACTGCGCTGGTG
UL146	3′	181019	R	GACCTGGAGTATGTTTATCGCCCCATCACCCTTTTTCAAAATGGCTAGACAAACGCAACG

HCMV ORF	5′ or 3′ Probe	Genome Position	Strand	Oligonucleotide Seqeunce (60-mer)
UL144	3′	182386	R	AAGAACCATACGTACTTTTCCACTCCAGGCGTCCAACATCACAAGCAACGACAGCAAAAT
UL142	3′	184037	R	GTCATGGTTCACACCCTGGGTCCAAAATAAGAGTTACAGCAAACAACCCCTAAGTACCAC
UL141	3′	184486	R	TTCACCGTATGTTTCCTATGCTACCTGTGTTACCTGCAGTGCTGTGGACGCTGGTGCCCC
UL140	3′	185796	R	ACGCCGAGATTCGCAGCAACTGCTCATCGAGCTTCCGCCGGAGCCGCTCCCACCCGATGT
UL139	3′	186633	R	AGCGGCTGGACATTATCTGGACTCCTTCTCATATTTACTTGCTGCCTTTGCTGTTTTTGG
UL138	3′	187441	R	GGATCCGCGACGCAGTTCACCACCGTAGCCATGGTACATTATCATCAAGAATACACGTGA
UL136	3′	188097	R	GACGTCGTCGAACGCGCTGCTGCCAGAATGGATGGATGCGGTACATGTGGCGGTCCAAGC
UL135	3′	189131	R	AAGAACCTGAGCACGCCGCCCACCAAGAAAACACCGCCGCCCACGAAACCCAAGCCGGTC
UL133	3′	190133	R	ACCGAAGAAAGGTAGGGCGAAAGACAAACCGAAGGGTAGACCGAAGAACAAACCTCCGTG
UL148A	3′	190807	R	ACCGCCGGCCGAGCTGGAGATGGTGGAAGTGTCGGACGAGTGTTACTAGGAGATCGCCGC
UL148B	3′	191298	R	GCGCTGCTGGTGTTTTATTATTGCGATGAGAGGGAGGATGGCCGTCCGTCGAAGCTGTTG
UL148C	3′	NA	NA	5'-oligonucleotide only
UL148D	3′	NA	NA	5'-oligonucleotide only
UL150	3′	192511	R	CGGGGGGTGTGCGGGAGACGATTGGGGAGCGACAAGACAGGGACGTGGCGGACTCTGCGA
IRS1	3′	198090	F	CGACCAACGCTTCACTACCGAAGGCCACCAGGTCATCCCGTTGTGCGCGTCCGAACCCGA
US1	3′	198880	R	AGCCGACGACTCACACGACCTATGGTGTTGCTGTATGGACTGGAAGGCGCACGTTGAGTA
US2	3′	199333	R	TCGTGTATGTGACTGTGGACTGTAACTTGTCTATGATGTGGATGCGGTTTTTCGTGTGCT
US3	3′	200367	R	AACCCTGCTCGTATATCTGTTCTCTCTGGTTGTTCTTGTGCTCCTGACCGTAGGGGTGAG
US6	3′	201684	R	CGACACGGATTCTTCGCTGTGACTTTGTATTTATGTTGCGGGATTACCCTGCTGGTTGTT
US7	3′	202779	R	CCAATTACAACTACCCTGGGTTCCTCGACCGTCCTCATTCATGTACGACAGCTACAGGCT
US8	3′	203591	R	ACCTGCGACGACTACACTATGAATACAGACACTTAGAGCTCGGGGTCGTGATCGCCATAC
US9	3′	204425	R	CGACTCTCTTACGTGATGTTAACCGTTTACCCCTGCAGTGCCTGTAACCGGAGCGTGCTC
US10	3′	205347	R	CTAGGCGACTATGGGGCAATACTAAAAATTTATTTCGGACTGTTCTGCGGGGCCTGCGTC
US11	3′	205961	R	TGATGGTGGCAGTGATTCAAGTGTTTTGGGGGCTGTATGTGAAAGGTTGGCTGCACCGAC
US12	3′	206812	R	ACCAAGTCATTATGTTTCTGCTCGTGGTAGTACCCCTGACGGCCCCCATCTGGTATCCAA
US13	3′	208077	R	ACCTGGTGCTCCGACCTGCCCGCCATCCTCATCGACTACACACTGGTCCTCACGCTGTGG
US14	3′	208680	R	GGTACTTTACGCCACCGAGACACTCATCTATCACACCACACTCCTGATGCTCACGCCCGT
US15	3′	209699	R	GAGGCGTGGTTGGTCCTGGTGCTTTAAGATCGTCTTGTCCTTCAGCGTGCTCATTACCTG
US16	3′	210440	R	GAATCCGCAACTGTCAGGACCTGCTGAGCGATCTACGTCTCTTAGAACTACCGTCGTCCG
US17	3′	211595	R	AGGCCGACCTGCTCACCTTGTGTCTCTATGAGAACCTGGTGTACCTGTACCTGCTCATTC
US18	3′	212685	R	TAAAAACAGCGTCTGTGTCAGCGTCGTGTTGTACACCATCGTCATGAGTCTCCTAAACAT
US19	3′	213569	R	GCTACGCCTCTATGTCGAAAATGTGGCTTTATTCATCGGCATGTACCATCTTCTGAGGCT
US20	3′	214445	R	ATGGTCATCTTTTTCTGCCCACCTAACCAATGCATCCGTCACGCCGTCTGTCTCTACCTG
US21	3′	215591	R	TGGCTGTCACGGCGTCAGGCTTCCACTTTAGTCATCGTTCAGTCATTTATGCCATGGTGG
US22	3′	216300	R	CCGTATAGGCGATACTCCAGAAACCCCGTGTAATAGCGAAAACGAGGACGACACGACGGT
US23	3′	218100	R	AGGACACTGTCTTTGACCTCAAGGACGTAGACGAATGGTTTGAGCAACGGCGTTTGGCAG
US24	3′	219816	R	TACTCGGACCCTTTTCCCTTAAAGAGCCTGACGGGTGTGCGCGAGTACATCCAGATTTGA
US26	3′	221956	R	CCGCCCCACGCCCAAACCTACCATGCCTCATCCTGTTTCTTACTTGCCCTTTTGGATGTA
US27	3′	225158	F	ACCAAACGTACCCATTATGACAGAAAACATGCACCTATGGAGTCCGGGGAGGAGGAATTT
US28	3′	226268	F	CTCGCCTTTTGTCACTGTTGTCTCAATCCGCTGCTGTACGTCTTCGTGGGCACCAAGTTT
US29	3′	227597	F	TTGGGTCTGCCGCCGCCACCGTCAGTCGCACCACCTGGGGAAAAGAAGGAGCTACCGGCT
US30	3′	228729	F	GAAGGCAAACTGCAGGAAGAGAAGAAACGACAGTCTGCTCTGGTCTGGAACGTTGAGGCG
US31	3′	229397	F	GTCTAACTGTGGGGAAAGTAGCACTAGCAGCGGCCAGAGTAGCGGTGACGAGAGTAACTG
US32	3′	230112	F	CGCGTGGTGTTCCAGTTAGGGGGAGTACCTCCTCGCCGTCACCGGGAAACTTACGTGTAA
US34	3′	231108	F	GCTAATGCCAGTTTTTTATCTCGCTTCAATGTCGGTGATTTTCACGGAGCGTCATGGGAA
TRS1	3′	232450	R	TCCATCTGTGTACATGCCTCCCACCGTACCCTACCCCGACCCCGCTGCCAGATTGTGCCG
C-ORF1	3′	2337	F	TTTCGCCGCCACGCTGCTTCACCGATATCCAATAAACCCATCCCCTCGCCACGACGTCTC
C-ORF2	3′	2334	R	ACGTCGTGGCGAGGGGATGGGTTTATTGGATATCGGTGAAGCAGCGTGGCGGCGAAAGAC
C-ORF3	3′	3397	R	TCCGTCGCCGTTCTGCACCATGATTCGATCAATAACAACATCATCATCGGAGACCATCTT
C-ORF4	3′	7860	F	CTTCCAAACCATCTTGAGACCCGAGTAACGGTTTACAGGTCGCACGCCAGTCTCAGCTAA
C-ORF5	3′	8969	R	TATGGAATCCGTTCGTAGATGTTCTGCTTTTTAGCCGCGTGTTGTTCCAGCTTTTTGCGT
C-ORF6	3′	23771	R	TCAGTGGCTCTGAGTCCTCCCGAGGAAAGAAAAGTCGTCGAAACAGCCAGCACAAATCGT
C-ORF7	3′	29346	F	GTCGGTTTTCAAACAGCAGGTTAAGTCCCAGACACATGAACGTGTTGAGATTATCTCCCA
C-ORF8	3′	35184	F	GATGTATGGCTGCTCACCGGTTTCGCGGCAACGGTTGCGCTCGAGTCCAACGGCGAGAAG
C-ORF9	3′	NA		No suitable probe could be identified
C-ORF10	3′	43288	F	AAAGTTGACCAGGGCTACCACATCGCGCCGCTGTAGACCGATAAACTGCAAACTCATGCT
C-ORF11	3′	46552	R	GCACGCCAAAGTTGAGCAGCCCCACAAAGTTGGAGCAGCCCTGGGTCAACATCTCGAGCA
C-ORF12	3′	NA		No suitable probe could be identified
C-ORF13	3′	54728	F	ATGCTCTATCCCCGTTACCCCGATGATGCTTGCGTCCCCGTTGTTATATTGGCACTGTCC
C-ORF14	3′	55105	F	CGGCGTGGGCTCCATGTCGGTGGCAGTGACGGCGACGGTGGTAACTGTGGTGGAGACGGT
C-ORF15	3′	120571	R	ACGACTCTATAAAAACCCACGTCCACTCAGACACGCAACTTTTGGCCGCCACACCTGTCA
C-ORF16	3′	157339	R	GCGCTTTCTCTGGCTCGACACAATGATCACACCGCACGCTATAGACACGTCGTCATGGAC
C-ORF17	3′	159486	F	CATACGTCGTCACCGTCCTCTGGAGGAGGCACGGCGGCGCTGTTGTTGTTTGGATGCTT
C-ORF18	3′	161781	R	CGAACCCCGTCGCAGACTCTATTTATACATCATCTTTCCAGCCCGCCTAGCAACACCCAC
C-ORF19	3′	NA		No suitable probe could be identified
C-ORF20	3′	166588	F	GTGGTGGGAGTGGTGACGGTGTTACTCGTGGTTGAAGTGACGTTAGGGGAGGTAGTAGTG
C-ORF21	3′	168368	F	TCTTGATAAACACGTAGGTGGTGGTAAAACTTCGGCCCGCGATCTGGACGTGGAGACGCA
C-ORF22	3′	168687	F	GTTGTGATGGCGGCCCCAGCTGAAGAAAGAGCACGTGTACTCAGTGGTCTCCTGCGGCCT
C-ORF23	3′	176265	F	AGGCCCACGATCCGGGTTATCTTGTCGTATTCCAGATTGATCCATCAATAGGGAACGCTG
C-ORF24	3′	185172	F	CGCGAATGGAAACCATGGGCCAGGAGTGTGTCATGACCGTGCAGGGGATGGTAACTTGCT
C-ORF25	3′	190735	R	CCGGCGGACGTGACTCGGCAGCCGCTGTAGAGATAAATAGTGCGATGGCGTTTGTGGGAG
C-ORF26	3′	190605	F	CCGCCGACGATCCCGCCAAACACCAAATTCCTAAAGCCGCGCATGTTATCCAGGCCACAA
C-ORF27	3′			No suitable probe could be identified
C-ORF28	3′	210280	F	TACTTGAAACCCTTTTTCTCTTCTCATGGTGCGCTGCGTTCTCTGGAAACGGCTGCTCTG
C-ORF29	3′	212325	R	ACGGTGATTATATTACACTCTATAAACGGTTCTCATACGCGCCTTTTGATAGCCACCGCC
ORF1	3′	19594	R	TTGGTCTTTCTAGGGCCGGTCAACAGCTTTATGAAGGGAATCAGGGACGTCGGCTTTGGA
ORF2	3′	37357	R	GGTCGCCGAGACTACTACTGCTTCTGCTTTTTTGTCTCCTGTGGATCGTCGCGGACTGCC
ORF3	3′	NA		No suitable probe could be identified
ORF4	3′	95473		GGCCCGGCAGGTCACCCAACGTTGGTTCAGGCCCAGTCGAGTTTTTCCCCGGCACGAACG
ORF5	3′	96898	F	CGTCTGAATTTTTGCTTGATAGACACGTGTTTGGAACTCTGTCCCCCCACGTTTTCACTG
ORF6	3′	134825	F	CCAGAATGGTGTTGAAGGCCTCTTTCTTGATGAAGTGCGCCTCGTCCACCAGCAGCAAAT
ORF7	3′	146707	F	CACGTCGTAAACATGGCCCATATGAAAAAGAGCATGCCGAACGACCAGCTGATGCCGGTG
ORF8	3′	146381	R	TTATCTGTTCTTCCACCAGCTACTTTACGTGTTTCAAACAGCACATTAGACACACGGGGG
ORF9	3′	171789	F	TCTCGACTTCTTCACCCTGTTCTTCCTCGCTATCAGAGATCACGATACAGCCGGCGGTAT
ORF10	3′	171872	F	ATGGTAAAGTCGGGCTCGGGCTTGATGTCTTCCTGTTTGATGAGGGGCAGCATGATAGGC
ORF11	3′	3933	F	ATAAAAACTCATGGACGTGAAATCTGGCTTGGTTGTGGTGATTTCATTCTCATTATTGTT
ORF12	3′	234784	R	ACGGAGCCGTCCGCGTGTGTAAACGGCGTGGTCGCTGACGCGGGTTTGCTTCCTATATAG
RL3	3′	3831	F	TATACATCTAAGATTGGAGTACTAGTAGTCGTTTGTGGTTTCTATTTTTTTTTTATATTT
RL4	3′	304	R	CGGGGTGTGTCAGGGGTGTGTCGGGGTGTGTTGGCGGGCCGTGTCTGCGTGTGTCCTCGA
RL5	3′	4772	F	AGATTCGACCAGACAGAAGAGAAGGACCGGGGGCTTGGCGACCCTTCCACGACTGCTGTTG
RL8	3′	7870	R	TCCGCTGTTTTTAGCTGAGACTGGCGTGCGACCTGTAAACCGTTACTCGGGTCTCAAGAT
RL9	3′	8229	F	CGTCAGGGGCCGGGGTTCTTATTAGAGAAACAGCACGTAGGTCAGGATCCAGATGCTAAT
US5	3′	201540	F	CGCGACAGAAAAAATACCGTTCGTAGAGAATGCCGTGTTGAAGGAACGCGCTTTTATTGA

HCMV ORF	5′ or 3′ Probe	Genome Position	Strand	Oligonucleotide Seqeunce (60-mer)
US25	3′	221716	F	GGTGCCCGACCGTGAAGAGCCCTCATCCACCTGAACAGACCGCTAACCGAAGGACCCCGA
J1S	3′	235125	R	CCCACCGCAGCACACGCAACTAGTCGCCGTCGCCGTCCACACACGCAACTCCAAATTTCA
UL61	3′	95055	R	AGAGGGGGGAGGGGAGCCCAACCGGAGCCGCCGAGAGGGAGCCGCCGCGAGACCCGGAAG
UL21	3′	27215	R	AACATGACGCGTTCGGGACGAAAGACGACGTCAGGGGATTCACGGTATTTAGCCATGCAG
UL41	3′	54339	R	GAAGCCATAAAAATATGGGAATTCCCCTGCCTACGTCTTCATGACGGGCGCCTTTTTCTT
UL60	3′	93693	R	TCGTAGAACGTTTCGTTAGAAGAATTATGCTATATAGGGTATGCATCCTAGGGGTGGAAG
UL101	3′	147689	F	GGGGGCGCGCCGATGACGACAGGCTCGCGGGTCGTTAAATACTACGATGGGAGCCGCCGC

Table 2.2. List of HCMV microarray oligonucleotide probe sequences.
The table lists the 60-mer oligonucleotide sequences together with their genome position, and the strand that they represent, either the forward or reverse strand. Note that all viral probes sequences represented the sense strand of each ORF.

Bacterial Spike Control	Oligonucleotide Probe Sequence (60-mer)
X17013 (Sense)	GGATCGAACCGGGCCGTTCTCTCGTGGGAGACGCAGGCACAACTCTTTATACGGTTGGCT
X17013 (Anti-Sense)	CCTAGCTTGGCCCGGCAAGAGAGCACCCTCTGCGTCCGTGTTGAGAAATATGCCAACCGA
M24537 (Sense)	GAAAGCGTTTGATGATGTATTGATTCCAGGGGCCATGCAGGAGCTTGAAGCACTCGGCTG
M24537 (Anti-Sense)	CTTTCGCAAACTACTACATAACTAAGGTCCCCGGTACGTCCTCGAACTTCGTGAGCCGAC

Table 2.3. Bacterial spike control probe sequences.
Bacterial spike control probes were printed in forward (sense) and reverse (anti-sense) orientation, thus serving as positive and negative control elements.

2.2.9.2 Microarray hybridisation

2.2.9.2.1 Preparation of cDNA

cDNA was fragmented by adding 15 µl of 0.1 M NaOH to the 30 µl cDNA sample (see 2.2.8.1) and incubating at 70 °C for 10 min, and then neutralised by addition of 15 µl of 0.1 M HCl. 20 µg of COT I human DNA was added to the cDNA preparation, and the unincorporated nucleotides etc. removed with water (SIGMA) by centrifugation in a Centricon column at 13000 x g for 5 min at RT. The flow-through was discarded and ~ 0.5 ml of water (SIGMA) was added to the column. The column was then centrifuged at 13000 x g until the volume of sample in the column was reduced to < 8 µl. The column was transferred to a clean collection tube by inversion of the Centricon column, and centrifugation at 13000 x g for 1 min at RT. The sample volume was then increased to 12 µl by the addition of water (SIGMA). The cleaned labelled cDNA was prepared for hybridisation by adding 2.6 µl of 20 X SSC and 0.44 µl of 10 % SDS (3.5 X SSC/0.3 % SDS (v/v)), and incubating at 95 °C for 2 min followed by 37 °C for 30 min.

2.2.9.2.2 Pre-hybridisation of microarray slides

Microarray slides were re-hydrated prior to pre-hybridisation. The slides were placed array face down in a chamber in an atmosphere of 1 X SCC for 2 min at RT, and then placed array face up on a 70 °C hot plate for 20 sec. The slides were then pre-hybridised with BSA blocking solution at 42 °C for 45 min, then rinsed with water and dehydrated in isopropanol, followed by cleaning under a stream of filtered compressed air. The array area was carefully covered by a LifterSlip 25 mm × 20 mm (Eerie Scientific Company, U.S.A.), previously prepared by washing in absolute ethanol, and drying under a stream of filtered compressed air.

2.2.9.2.3 Hybridisation and stringency washes

The cDNA sample was carefully applied to the array area under the LifterSlip by capillary attraction, so as to avoid the formation of air bubbles. The array slides were then hybridised in a humidified GENETIX hybridisation chamber, which was incubated in a dry oven at 65 °C for 16-20 h. Following hybridisation, LifterSlips were removed without mechanical interference by briefly immersing the slides in 1 X SSC/0.2 % SDS (v/v) at 65 °C. The subsequent array washes described below were performed at RT in staining jars, in which solutions were continuously mixed using a magnetic stirrer. The slides were washed in 1 X SSC/0.2 % SDS (v/v) for 5 min, 0.1 X SSC/0.2 % SDS (v/v) for 5 min, then

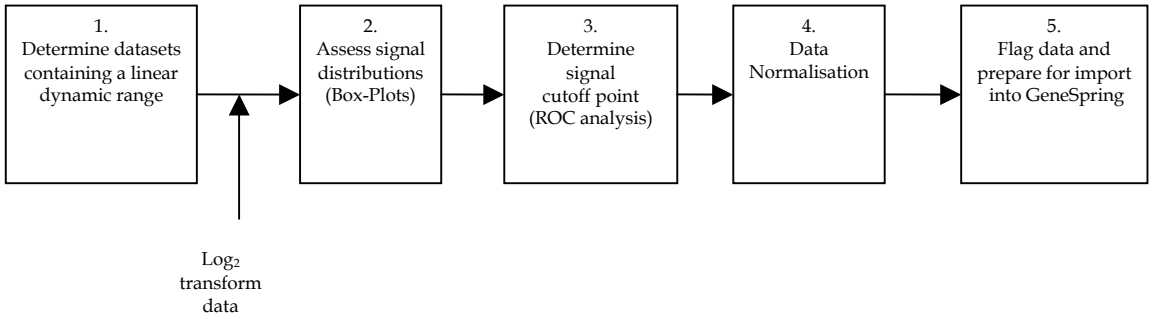
plunged 5 times in 0.1 X SSC in order to remove residual SDS. The slides were then dried by centrifugation (array face out) at 500 x g for 1 min at RT. The dry slides were then placed in a darkened box, and stored in a desiccator at RT.

2.2.9.2.4 Scanning and quantitation of microarrays

Microarrays were scanned and quantitated using Perkin-Elmer ScanArray Express hardware and associated software. Slides were scanned at a laser wavelength of 543 nm, with laser power set at 100 %, and a pixel resolution of 10 μm. The slides were each scanned 4 times at photomultiplier tube (PMT) gain settings of 40, 50, 60, and 70 %. Scanning of slides with increasing PMT gain was performed in order to identify array images whose hybridisation signals lay within a linear dynamic range, in the absence of signal saturation. Microarray images were saved as both greyscale TIFF and colour JPEG files before proceeding with filtration of poorly defined spots and quantitation of spot signal intensities. A flow-diagram depicting the stages of microarray data analysis is shown in Fig. 2.1.

Microarray hybridisation signals were quantitated using the 'ScanArray Express new fast spot-finding algorithm', on the TIFF image files. The spot-finding algorithm uses the GAL file to determine the position of each feature on the array. The spots were then quantitated by adaptive circle feature extraction; this places a circle over each spot and estimates the diameter for each spot individually (maximum diameter set at 100 μm), to optimise data capture. The spot-finding algorithm then defines background correction values and signal values (i.e. median and mean values), which are calculated for each spot. Visual inspection of the array grids was undertaken in order to confirm correct alignment, and to manually check the integrity of features (e.g. misshapen spots, smears or artefacts). The quantitation data was used to generate a spreadsheet listing signal intensities for each spot on the microarray (median minus background, and mean minus background), which was subsequently exported and saved as a Microsoft Excel file. This process was undertaken for each PMT gain setting (listed above), so that four Excel datasets were generated per microarray.

Stage 1



Stage 2

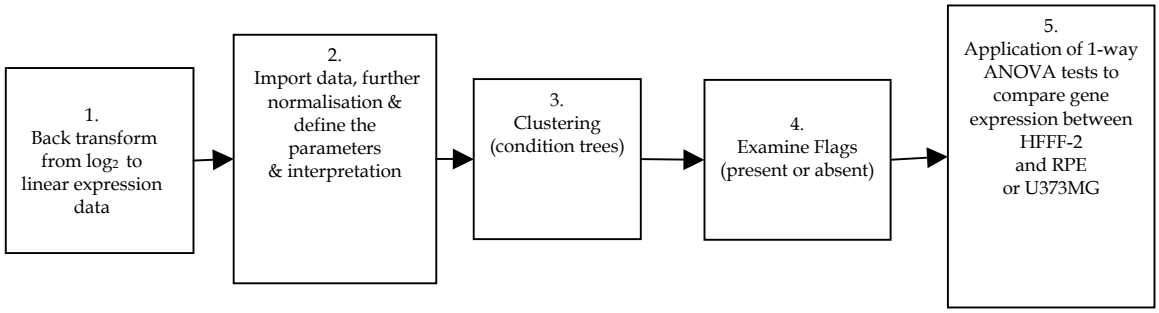


Figure 2.1:. Analysis procedure for the HCMV microarray data
The flow-diagram represents the steps taken to analyse the HCMV gene expression data. This procedure starts with the raw gene expression signal intensities, which are processed in a series of steps in order to determine linearality, signal cutoff points, and normalisation. The data is imported into GeneSpring and an experiment is created, followed by analysis such as filtering and clustering, and then identification of differential gene expression using 1-way ANOVA tests.

2.2.9.3 Data Processing

2.2.9.3.1 Determination of linear dynamic range and assessment of signal distribution

Each microarray slide was scanned four times at PMT gain settings of 40, 50, 60, and 70 %, generating four datasets per microarray. The signal intensity values of all probes on the microarray were extracted, and the 'mean minus background' signal values were selected for data processing and analysis. In order to determine the dataset with a linear dynamic range, scatterplots were drawn in Excel of PMT gain 40 vs 50, 50 vs 60, and 60 vs 70. The dataset giving signal values with optimum scatter in the absence of signal saturation and falling within a linear range (assessed by least-square fit analysis), were chosen for further analysis.

The signal distributions of the selected datasets were further assessed by drawing Box-Plots (SPSS statistical analysis software). Linear signal intensity data is 'right-skewed' when assessed by histograms. In order to transform the linear signal intensity data so that it follows a symmetrical distribution (a common assumption for many statistical tests, and a requirement for those employed here for analysing HCMV microarray data), the microarray data was $\log_2$ transformed. Box-plots were then drawn to ensure that the datasets were symmetrical. Datasets that fulfilled the requirements for both linear and symmetrical data were selected for Receiver Operating Characteristic (ROC) analysis.

2.2.9.3.2 Receiver operating characteristic analysis

An essential requirement in all microarray analyses is to determine a signal threshold cutoff point that discriminates between true-positive hybridisation signals and false-positive signals arising from non-specific binding. Unlike commercially available microarray technologies, where algorithms are designed to make such decisions with little user input, no such algorithms are applied to custom microarrays. Other workers have selected fairly arbitrary cutoff points, which do not take into account the sensitivity and specificity of such signal thresholds, and which can lead to elevated false-positive rates (Bilban et al., 2002). Receiver Operating Characteristic (ROC) analysis was performed to address this issue. ROC analysis determines signal thresholds that give optimum specificity (i.e. true-negative rate), without compromising the sensitivity (i.e. true-positive rate) of the microarray.

ROC analysis (contained in SPSS) was performed on the positive and negative control signal values for selected linear datasets (2.2.9.3.1). ROC analysis ranks the values for

positive and negative signals, then establishes a series of signal thresholds and measures the likelihood of type I (i.e. false-positive) and type II (i.e. false-negative) errors for such cutoff points. A ROC curve is drawn representing the relationship between false-positive and false-negative rates for every possible cutoff point; it is then possible to select a signal threshold for the whole array that gives optimum sensitivity and specificity. For all control signal datasets analysed, a signal threshold that gave a false positive rate of 5 % was selected. All signal values that fell below this cutoff point were assigned a signal value of 0.01.

2.2.9.3.3 Normalisation

Linear datasets were normalised against positive control signals generated from the *in vitro* transcribed spike mRNA of known concentration, in order to control for differences in the efficiencies of cDNA synthesis, hybridisation and scanning. The 75th percentile was calculated from the positive control signal values for each array dataset. An array chosen at random served as a reference against which all other array datasets, irrespective of cell type, were scaled. The following calculation was used to define a correction value for each dataset, allowing global normalisation against the reference.

$$\text{Correction Factor (CF)} = \frac{\text{75th Percentile of Positive Control Signals (Reference Array)}}{\text{75th Percentile of Positive Control Signals (Test Array)}}$$

A normalisation correction factor (CF) in the range 0.5 to 4 was deemed acceptable. Most of the HCMV array CF values fell within the range 0.75 to 2 (see table 4.2, results II). All probe signals in a particular dataset were then normalised by multiplying against the determined CF for that array. This allowed normalisation across all microarray datasets for all cell types.

2.2.9.3.4 Data flagging and preparation for import into GeneSpring

Following normalisation of the array datasets, signals were flagged as present (P), or absent (A). Signal values of 0.01 were flagged as absent (based on the ROC analysis, see 2.2.9.3.2). The normalised flagged data were transferred to new Excel files, which contained three data columns (Gene Name; Signal Value ($\log_2$); Flag). The new Excel files were then saved as 'tab delimited text files'; the format required for import into the expression analysis software, GeneSpring.

2.2.9.4 Analysis of data in GeneSpring

2.2.9.4.1 Preparation of the GeneSpring software and import of microarray data

In order to analyse HCMV gene expression data, the annotated HCMV (Merlin) genome file (acc. no. AY446894) was downloaded from GenBank (NCBI) into the GeneSpring software (Agilent Technologies). Annotations for regions of the genome now considered to be non-coding and the novel ORFs recently proposed by Murphy et al., (2003a and b) were added, creating a genome within GeneSpring that represented the full complement of probes on the microarray. The $\log_2$ datasets (tab delimited text files; see 2.2.9.3.4) were imported into GeneSpring using the software import wizard. This process had several stages:

1. Import data files and create an 'experiment'.

2. Describe the data contained within each data file (i.e. gene name; signal value; flag).

3. Transform $\log_2$ data to linear data (as GeneSpring uses the natural logarithm of linear data for its statistical analyses).

4. Further normalisation; per gene, normalise to the median. This normalisation step controls for differences in the detection efficiency between spots and was performed so that the relative change in gene expression could be compared for all genes between each cell type. This was calculated as follows:

$$\text{Per Gene : Normalise to the Median} = \frac{\text{Signal strength of gene A}}{\text{Median of all measurements for gene A from each cell type}}$$

5. Define the experiment parameters i.e. cell type and time.

6. Define the experiment interpretation and display options (i.e. calculate the arithmetic mean of the expression data for each gene).

The microarray data for each gene was displayed in the form of three separate "gene expression profile" graphs (one for each cell type), with time (h PI) on the x-axis, and normalised signal on the y-axis.

2.2.9.4.2 Examination of present and absent flags

All probes in each dataset had been flagged as present (P) or absent (A) prior to import into GeneSpring. Genes that failed to pass the cutoff points determined by the ROC analysis were assigned a value of 0.01, and flagged as absent (A). Examination of the number of datasets containing present/absent flags for each gene was undertaken in GeneSpring using the filtering tool. All genes from each cell type were assigned a value of from 0 to 12 (as there were 12 datasets per cell type), so genes that were not expressed could easily be identified, and to allow basic quality control analysis to be undertaken.

2.2.9.4.3 Quality control based on clustering

Condition trees were then assembled for each cell type, in order to confirm that the array datasets for each time point clustered together, and that there were no obvious anomalies (e.g. a 12 h dataset clustering with a 72 h dataset). The condition trees were drawn using the clustering tool within GeneSpring, and based on Spearman's correlation (rank analysis) in which the null hypothesis (no relationship between the array datasets) was tested in order to determine the clustering. This analysis was peformed for each cell type independently. Note that the 'per gene: normalise to the median' was removed prior to clustering so that comparisons could be made independently from one another. Mock-infected cell datasets were excluded from this analysis.

2.2.9.4.4 Differential gene expression

In order to identify genes whose expression were differentially regulated in different cell types, combined statistical tests (Student's t-test; Welch's t-test; and Wilcoxon-Mann Whitney test) were performed independently between HFFF-2 vs RPE, and HFFF-2 vs U373Mg. The statistical tests were performed on the mean expression value for each gene from all data points over the time course, giving a single expression value for each gene in each cell type. The statistical tests then compared the expression values for individual HCMV genes in HFFF-2 against the corresponding expression values for individual HCMV genes in RPE or U373Mg cells. The null hypothesis (no difference in the mean gene expression intensities in HFFF-2 and RPE or U373Mg cells) was tested. Significant differences in the mean expression of individual genes in HFFF-2 and RPE or U373Mg cells were identified at the 95 % significance level ($p < 0.05$).

In order to adjust the p-values derived from the statistical tests to correct for the occurrence of false positives, a multiple testing correction (MTC) was applied. We selected the Benjamini and Hochberg False Discovery Rate as this provides a good

balance between the limitation of false positive data and the discovery of statistically significant differences. The Benjamini and Hochberg MTC was based on an error rate of 5 %, therefore 5 % of genes considered significantly different would pass the MTC restriction by chance. The MTC ranks the p-values derived from the statistical tests from smallest to largest; the largest p-value remains as it is. The second largest p-value is corrected by:

Corrected p-value = p-value × (n/n-1)

n = the total number of genes tested in the gene list.

n-1 = the rank of the gene whose p-value is being adjusted.

The third largest p-value is then corrected by multiplying the p-value by (n/n-2), and so on. The test becomes more stringent as the rank of the p-value decreases. Genes whose p-values remain < 0.05 following the application of the MTC are considered significantly different, but note that 5 % of these genes will have passed this test by chance.

2.2.10 PCR

All PCR reactions were performed in a total volume of 50 µl and contained 100-500 ng of template DNA, using *Taq* polymerase (QIAGEN). Thermo-cycling conditions were; a hot start at 95 °C for 10 min; 35 cycles at 94 °C for 30 s; 55 °C for 20; 72 °C for 45 s; and finally a hold at 4 °C.

2.2.11 Agarose gel electrophoresis

All DNA fragments generated by PCR or restriction endonuclease digests were analysed on 1 % agarose/TBE gels containing ethidium bromide. DNA samples were diluted to give a final concentration of 1 X OrangeG (v/v), and electrophoresed at 100 V for 3 h. The gels were then visualised using short-wave or long-wave UV transillumination, as appropriate, and photographed using the BioRad Gel Doc system.

2.2.12 Recovery of DNA fragments

DNA fragments were separated on 1 % agarose/TBE gels at 100 V for 3 h. The DNA fragments were visualised using long-wave UV transillumination, and the bands excised.

DNA fragments were extracted from agarose gel slices using the PureLink Quick Gel Extraction Kit (Invitrogen), according to the manufacturer's instructions.

2.2.13 Restriction endonuclease digests

Digest reactions typically contained 10 U of enzyme, ~ 500 ng of DNA, prepared in a buffer solution compatible with the restriction enzyme. Total reaction volume was 20 µl, with incubation at 37 ^{o}C for 3 h.

2.2.14 Preparation of FIX-BAC DNA

The HCMV bacterial artificial chromosome FIX-BAC derived from the clinical VR1814 strain of HCMV was the kind gift of Dr G. Hahn (see Gerna et al., 2003). FIX-BAC DNA was isolated from *E.coli* DH12 bacteria using the QIAGEN large construct kit, according to the manufacturer's instructions.

2.2.15 Real-time PCR

Real-time PCR (RT-PCR) analyses were undertaken using an Applied Biosystems 7500 Fast Real-Time PCR machine and associated software. The PCR reaction mix contained 2 µl of 10 X PCR buffer (200 mM Tris.HCl, 500 mM KCl), 1 µl of 10 X $MgCl_2$ (50 mM), 1 µl each of forward and reverse primers (10 µM), 1 µl of 10 X SYBR green I, 2 µl of 10 X dNTP mix (each at 2 mM), 0.2 µl of Platinum *Taq* polymerase (5 U/µl) (Invitrogen), and 8 µl of template DNA (12.5 ng of cDNA; or 150 ng genomic DNA). Thermo-cycling conditions were; hot start at 95 °C for 10 min, then 35 cycles at 94 °C for 10 sec; 60 °C for 20 sec; 72 °C for 25 sec (with read plate); followed by dissociation analysis from 65-95 °C, with plate read every 0.2 °C increment; and hold at 4 ^{o}C.

2.2.16 Northern Blotting

2.2.16.1 Formaldehyde-agarose gel electrophoresis

In order to avoid RNase contamination, disposable plastic-ware was used whenever possible. Glassware was autoclaved and baked twice in the dry oven; gel tanks and other re-useable plastic-ware were washed with RNaseZap (SIGMA), rinsed in distilled water and allowed to drip-dry; nuclease-free and/or distilled water, nuclease-free eppendorf tubes and pipette tips were used throughout.

To prepare the gel, 1 g of agarose was dissolved by boiling in 85 ml of water (SIGMA), and cooling to 55 °C. 5.8 ml of formaldehyde and 10 ml of 10 X MOPS were pre-heated to 55 °C, added to the 1 % agarose (w/v), mixed and poured into a gel-casting tray and left to set at RT. The gel was pre-electrophoresed at 80 V for 10 min in 1 X MOPS buffer.

10 µg of total RNA in 5 µl of water (SIGMA), was mixed with 3 volumes of NorthernMax Formaldehyde Load Dye (Ambion Inc.), and incubated at 65 °C for 15 min. The RNA ladder (0.5-10 kb) (Invitrogen) was similarly prepared by adding 9 µl of Loading Dye to 3 µl of the RNA ladder and incubating at 65 °C for 15 min. The RNA samples and ladder were then snap cooled on ice for 5 min, and 0.25 µg of ethidium bromide added. RNA samples were electrophoresed at 50 V for 6 h, and the RNA bands briefly visualised and photographed using short-wave UV transillumination with the BioRad Gel Doc system. The intensities of the 28S and 18S rRNA bands across the gel tracks confirmed equal loading of the RNA samples.

2.2.16.2 RNA transfer by capillary blotting

RNA transfer from the gel to a nylon membrane (Hybond N+, Amersham Biosciences) was accomplished by capillary blotting (Southern, 1975). The gel was washed in distilled water for 10 min followed by washing in 10 X SSC for 15 min. The nylon membrane was washed in distilled water for 5 min followed by washing in 20 X SSC for 20 min. Whatman 3MM paper was washed in 2 X SSC for 20 min.

The gel was placed on Whatman 3MM paper that functioned as a wick descending from a raised platform into a reservoir of 20 X SSC. The nylon membrane was placed on top of the gel, followed by Whatman 3 MM paper soaked in 2 X SSC, and a stack of dry paper towels with a weight on top. After overnight transfer of RNA, the membrane was washed in 10 X SSC for 1 min (no shaking) and the RNA cross-linked to the membrane using a Stratagene UV cross-linker operating at 'auto-crosslink' (12000 Jcm^{-2}). The membranes were then wrapped in clingfilm and stored at 4 °C, prior to nucleic acid hybridisation.

2.2.16.3 Preparation of DNA probes

^{32}P-radiolabelled DNA probes were prepared from PCR fragments (250-400 bp) amplified using HCMV (Merlin) genomic DNA (see 2.2.3). Radiolabelled DNA probes for the cellular housekeeping gene, glyceraldehyde 3'-phosphate de-hydrogenase (GAPDH) were prepared from DECAtemplate GAPDH-mouse (905 bp) (Ambion Inc.). The Rediprime II

Random Prime Labelling System (Amersham Biosciences) was used to synthesise the probes (using 100 ng of template DNA and 50 μCi of α-^{32}P-dCTP), according to the manufacturer's instructions. The radiolabelled probes were then purified by acetate/ethanol precipitation using 150 μg/ml of GlycoBlue (Ambion Inc.) to facilitate DNA recovery. The DNA was precipitated with 6 μl of ammonium acetate (5 M) and 3 volumes of absolute ethanol -20 °C. The DNA probe was incubated at -20 °C for 1 h, then the DNA pelleted by centrifugation at 13000 x g for 10 min at 4 °C. The DNA was washed with 70 % ethanol 4 °C, followed by centrifugation as before. The DNA pellet was air dried for 5 min, before resuspension in 50 μl of water (SIGMA). The radiolabelled DNA probe was prepared for hybridisation by boiling for 10 min, cooled on ice for 5 min, and then added to the hybridisation reaction.

2.2.16.4 Nucleic acid hybridisation

Hybridisations were performed in glass hybridisation tubes (Hybaid) that were incubated in a Hybaid minioven MKII. Nylon membranes (Hybond N+) containing the cross-linked RNA fragments were prepared for hybridisation by washing in 20 ml of 1 X SSC/0.1 % SDS (v/v) at 68 °C for 1 h. The membranes were then pre-hybridised with 15 ml of Rapid-Hyb Buffer (Amersham Biosciences) at 68 °C for 2 h. Denatured radiolabelled DNA probe was then added and hybridisation continued at 68 °C overnight. Hybridised membranes were then washed with 20 ml of 2 X SSC/0.1 % SDS (v/v) at 68 °C for 20 min, and washed twice with 20 ml of 0.1 X SSC/0.1 % SDS (v/v) at 68 °C for 10 min. The radiolabelled membranes were then wrapped in clingfilm and exposed to a BioRad phosphorimager screen.

2.2.16.5 Phosphorimager analysis

The radiolabelled membranes were exposed to a phosphorimager screen for 3 h, after which the membrane was removed and the screen placed on the BioRad Personal FX phosphorimager. Radiolabelled bands on the northern blot were imaged using the Quantity One program.

2.2.17 Western Blotting

2.2.17.1 Preparation of mock-infected and HCMV infected cellular protein

35 mm tissue culture dishes containing 5 x 10^5 cells/dish of HFFF-2, RPE and U373Mg were incubated at 37 °C overnight. The cell monolayers were mock-infected or infected

with HCMV at a m.o.i. of 6 p.f.u./cell as previously described. After virus adsorption, the cell layers were washed twice with 1 ml of DMEM/FCS, and then overlaid with 2.5 ml of fresh medium followed by incubation at 37 ^{o}C. At 0, 12, 24, 48, and 72 h PI, and at 72 h PI for MI cell cultures, the medium was removed and the layers washed twice with 1 ml of PBS, and lysed with 200 µl of SDS-PAGE sample buffer, and stored at -20 ^{o}C prior to SDS-PAGE.

2.2.17.2 SDS-PAGE electrophoresis

The Bio-Rad mini-protean II cell apparatus was used for the preparation of SDS-PAGE gels. A 10 % SDS-PAGE gel was prepared; 5 ml of resolving gel buffer, 8 ml of distilled water, 7 ml of 30 % acrylamide/BIS (37.5:1), 20 µl of TEMED, and 80 µl of 25 % (w/v) ammonium persulphate. The stacking gel contained 3 ml of stacking gel buffer, 7 ml of distilled water, 2 ml of 30 % acrylamide/BIS, 20 µl of TEMED and 80 µl of 25 % (w/v) ammonium persulphate. 25 µl of each protein sample were loaded onto each gel slot, 3 µl of rainbow protein marker (Amersham Biosciences), and electrophoresed at 120 V.

2.2.17.3 Western immunoblotting

The Bio-Rad mini trans-blot apparatus was used to transfer proteins to ECL nitrocellulose membranes. SDS-PAGE gel and membrane were sandwiched between two layers of 3MM paper pre-soaked in Towbins buffer. The sandwich was immersed in a tank of Towbins buffer and transfer was accomplished by electrophoresis at 50 V for 2-3 h. The membranes were retrieved from the apparatus and incubated with blocking solution (PBS with 0.00005 % Tween 20 (PBS/T) and 5 % (w/v) 'Marvel' milk powder at 4 °C overnight, then washed 3 times with PBS/T for 15 min. Primary antibody (typically a 1:500 dilution in PBS/T containing 1 % BSA (w/v)) was incubated with the membrane at 37 °C for 2 h, followed by washes with PBS/T as previously described. Secondary antibody (typically 1:1000 dilution in PBS/T with 1 % BSA (w/v)) was then incubated with the membrane at 37 °C for 1 h, then washed 3 times with PBS for 15 min. The membrane was then treated with ECL reagents (Amersham Biosciences) for 1 min at RT, according to the manufacturer's instructions, and then exposed to photographic film prior to developing the image using the Konica SRX 101A film processor.

2.2.18 Immunofluorescence

13 mm sterile glass coverslips containing 1 × 10^4 cells/coverslip of HFFF-2, RPE and U373Mg were incubated at 37 °C overnight. Monolayers were infected with HCMV at

varying m.o.i. as previously described. The infected cultures were incubated at 37 °C for 48 h, after which the medium was removed and the cells fixed (5 % (v/v) formaldehyde, 2 % (w/v) sucrose in PBS) for 10 min at RT. The cells were then permeabilised by treatment with 0.5 % (v/v) NP40, 10 % (w/v) sucrose in PBS, for 5 min at RT, and the cells washed 3 times with PBS. Cells were then treated with primary antibody (anti-UL44 clone:CCH2 pre-conjugated to FITC; diluted 1:500) for 1 h at RT, after which the coverslips were washed 3 times with PBS. The coverlsips were set on glass microscope slides using mounting medium and sealed by the application of clear nail varnish around the sides of the coverslip. The coverslips were examined under UV illumination using the Zeiss Axioplan 2 microscope wavelength LSM512 imaging system, operating with the 488 nm wavelength laser. Fluorescent cells expressing the HCMV UL44 early gene product were counted from 2 representative fields of view for each coverslip.

3 RESULTS I

3.1 Characterisation of HCMV strain Merlin replication in different cell types

It is important to compare the replication kinetics of HCMV strain Merlin in HFFF-2, RPE and U373Mg cells prior to an interpretation of global HCMV gene expression data in these cell types. There are several stages in the virus lytic cycle that could be affected by growth in different cell types. Virus attachment and entry into host cells might be more efficient in one cell type compared to another. The kinetics of the viral transcription and/or protein expression cascade might vary in different cell types as a consequence of differences in the levels of specific cellular factors. The kinetics of viral DNA replication might also differ between cell types. Any one or more of these factors might affect the production and release of mature infectious particles.

3.2 One step virus growth curves

One step HCMV growth curves were obtained for HFFF-2, RPE and U373Mg cells infected at a m.o.i. of 1 p.f.u./cell. Samples were harvested every 24 h over a 7 day period for estimation of total infectious virus yield (Fig. 3.1). HCMV replication was most efficient in the HFFF-2 cell line. Exit from the viral eclipse phase of the growth curve occurred at 48 h PI in the case of HFFF-2 cells but was delayed by 24 h (i.e. at 72 h PI) in both RPE and U373Mg infected cultures. The rate of virus growth over the period 48 h to 96 h PI was similar in all of these cell types, although the final virus titres obtained varied widely. The curve obtained with HFFF-2 plateaus at 144 h PI with a maximum titre just over 10^5 p.f.u./ml. The maximum titre obtained from the RPE culture was 10 fold lower at 10^4 p.f.u./ml, although it should be noted that the curve has not reached a plateau; the maximum titre obtained from the U373Mg culture were 1000 fold lower than that obtained for HFFF-2, at 10^2 p.f.u./ml, and the curve plateaus after only 96 h PI.

The one step virus growth curves were then repeated at a m.o.i. of 6 p.f.u./cell, but in this experiment, the total yield was fractioned into cell-associated virus (CAV) and cell-released virus (CRV) yields. Fig. 3.2 shows the curves obtained for the CRV fraction. In HFFF-2 cells, released infectivity was first detected at 72 h PI, rising steadily to plateau at 144 h PI, with a yield of 10^6 p.f.u./ml. In RPE and U373Mg cell cultures, release of infectivity was first detected at 120 h PI (i.e. 48 h after HFFF-2), and the rate of accumulation of released infectivity was much slower, and failed to plateau over the 7 day period. Yield for RPE cultures was 100 fold less than that for HFFF-2 cultures, and for

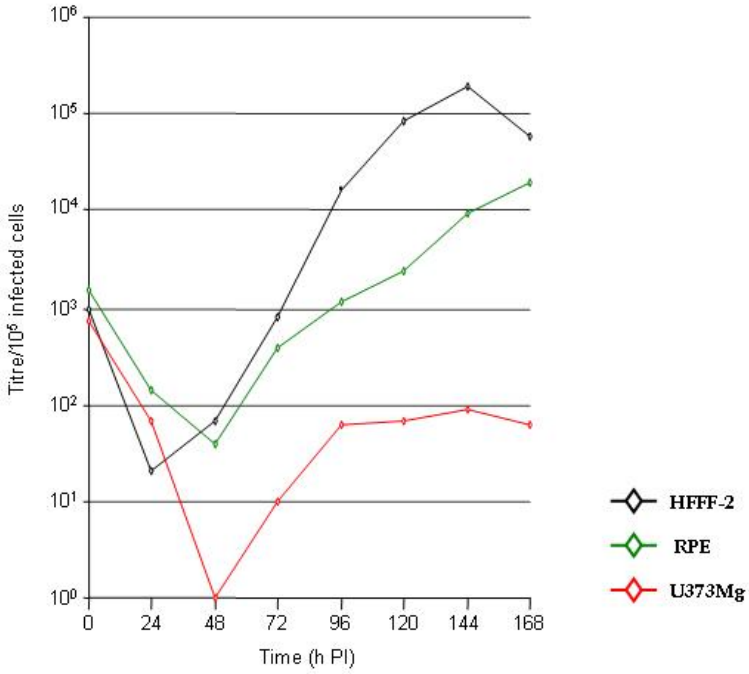


Figure 3.1. One step HCMV growth curves obtained for HFFF-2, RPE and U373Mg cells infected at a m.o.i. of 1 p.f.u./cell.

The growth curve shows the total virus yields over a 7 day period. Infected cells were harvested by scraping into the medium. Cell associated virus was released from infected cells by ultrasonic treatment and infectivity determined by titration on HFFF-2 cell layers.

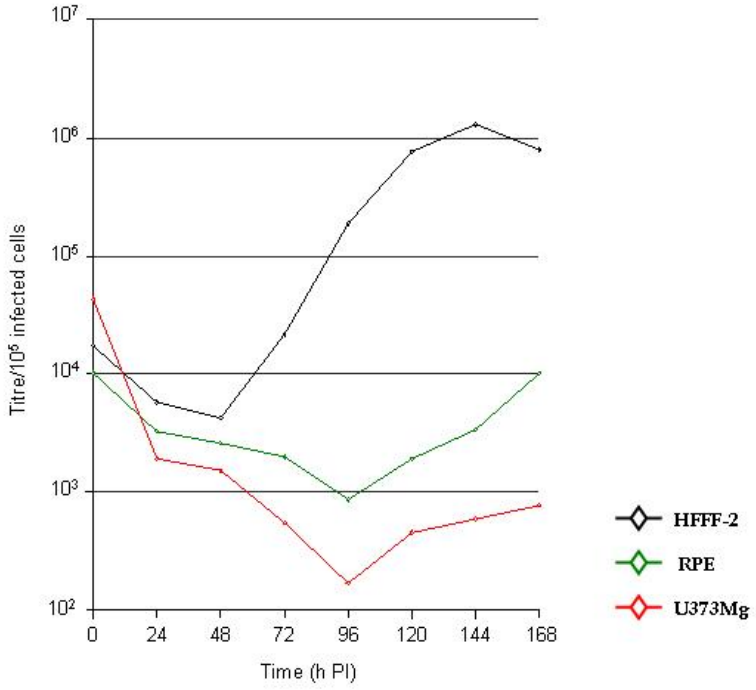


Figure 3.2. One step HCMV cell released virus (CRV) growth curves from HFFF-2, RPE and U373Mg cells infected at a m.o.i. of 6 p.f.u./cell.

The growth curve shows the yield of CRV every 24 hours over a period of 7 days.

U373Mg cultures, 1000 fold less. Fig. 3.3 shows the curves for the CAV fraction. In HFFF-2 cultures, exit from the viral eclipse phase occurred at 48 h PI, as before, RPE and U373Mg are relatively delayed by 24 h (i.e. at 72 h PI). The rate of accumulation of CAV is similar in HFFF-2 and RPE cultures, but slightly slower in U373Mg cultures. All three curves plateau at 96 h PI, with HFFF-2 and RPE having a similar titre at 10^4 p.f.u./ml, while the titre from U373Mg cultures is ~ 100 fold less at 10^2 p.f.u./ml.

Taken together, Fig. 3.1, 3.2 and 3.3 shows that accumulation of mature infectious virus is delayed by ~ 24 h in RPE and U373Mg infected cells compared to infected HFFF-2 cells. The rate of growth is similar in HFFF-2 and RPE infected cells and similar amounts of CAV are made. In contrast, replication of virus is limited both temporally and in infectious titre in U373Mg cells. Release of infectivity to the extracellular medium is both delayed (by 48 h) and accumulates more slowly from RPE and U373Mg cells compared to infected HFFF-2 cells.

3.3 Efficiency of HCMV infection in HFFF-2, RPE and U373Mg cells

In order to determine whether differences in the final virus yield were due to differences in the efficiency of HCMV infection of the different cell types, immunofluorescence experiments were performed at a m.o.i. of 3, 6 and 10 p.f.u./cell for each type. Infected cells were incubated for 48 h prior to fixing and staining with anti-UL44 antibody (CCH2 pre-conjugated to FITC). The HCMV UL44 is an early gene coding for the DNA polymerase processivity factor, and is used here as a marker for infection of the cells. Fluorescently labelled cells expressing UL44 were counted from 2 representative fields of view for each cell type (Fig. 3.4). No significant differences were seen in the numbers of cells infected for the different cell types. It was concluded that the observed differences in the final virus yields were not due to disproportionate numbers of cells infected for the different cell types.

3.4 Expression kinetics of known HCMV immediate-early, early and late proteins

Delays observed in the production of mature infectious particles and their subsequent release from RPE and U373Mg cells (Fig. 3.1, 3.2 and 3.3) could be due to delays in the onset of the gene expression cascade. The expression kinetics of representative known immediate-early (IE) (IE1/IE2), early (E) (US22; pUS22) and late (L) (UL99; pp28) proteins were examined in each cell type. Equal numbers of cells were seeded in 30 mm dishes and either mock-infected or infected with HCMV at a m.o.i. of 6 p.f.u./cell, and cell protein

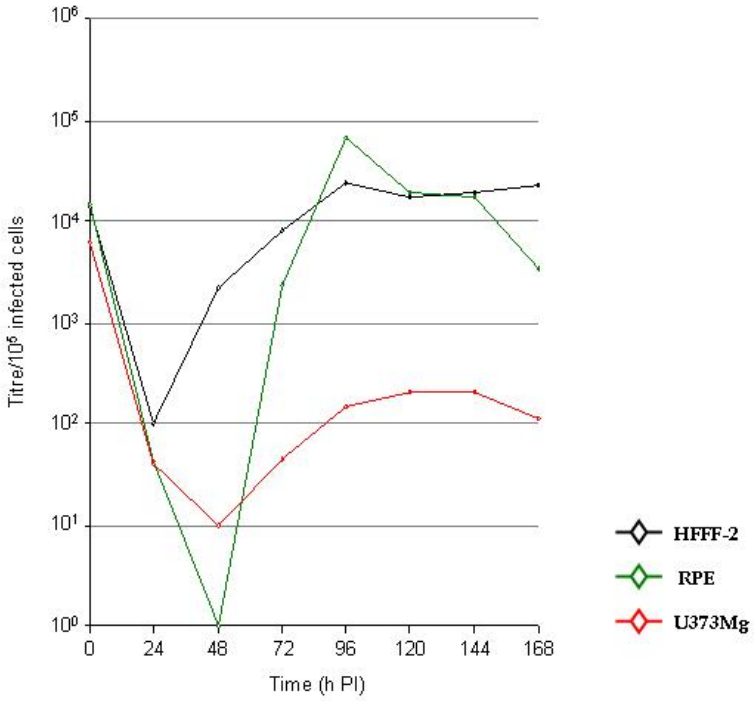
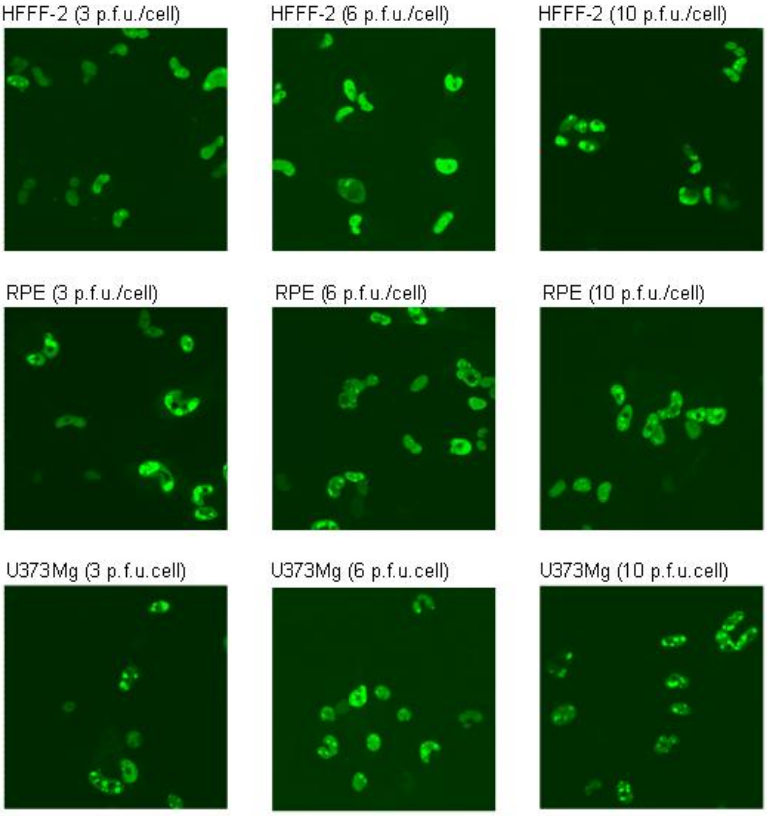


Figure 3.3. One step virus growth curve of HCMV cell associated virus (CAV) in HFFF-2, RPE and U373Mg cells at m.o.i. of 6 p.f.u./cell.

The growth curve shows the yield of cell-associated virus every 24 hours over a period of 7 days.

A.


HFFF-2 (3 p.f.u./cell)
HFFF-2 (6 p.f.u./cell)
HFFF-2 (10 p.f.u./cell)
RPE (3 p.f.u./cell)
RPE (6 p.f.u./cell)
RPE (10 p.f.u./cell)
U373Mg (3 p.f.u.cell)
U373Mg (6 p.f.u.cell)
U373Mg (10 p.f.u.cell)

B.

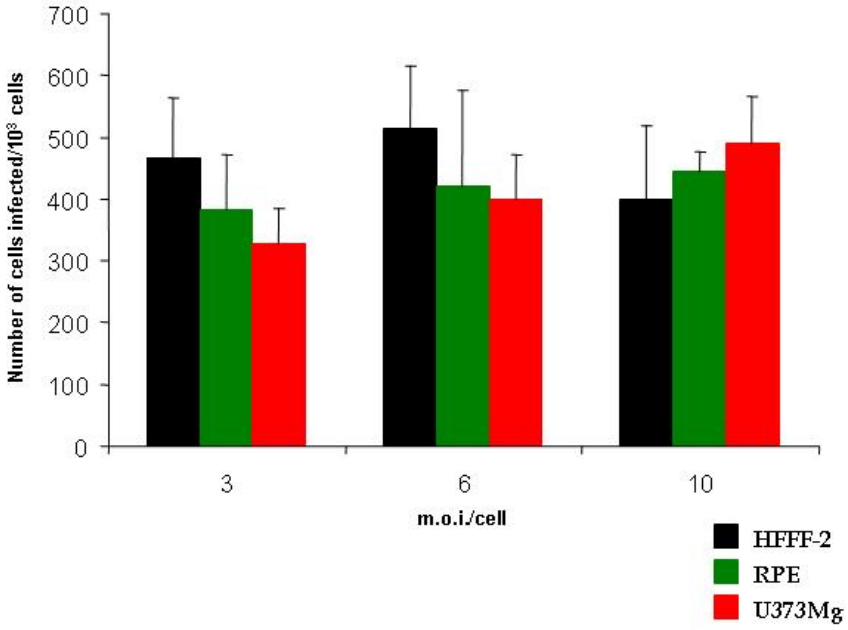


Figure 3.4. (A) shows representative immunofluorescence images for HFFF-2, RPE and U373Mg cells infected with HCMV at m.o.i. of 3, 6 and 10 p.f.u./cell and expressing the HCMV early gene UL44 (FITC labelled), and (B) shows the numbers of UL44-FITC labelled cells in two representative microscope fields of view for each cell type infected at each m.o.i.

The IF images in (A) were taken using the × 40 objective lens.

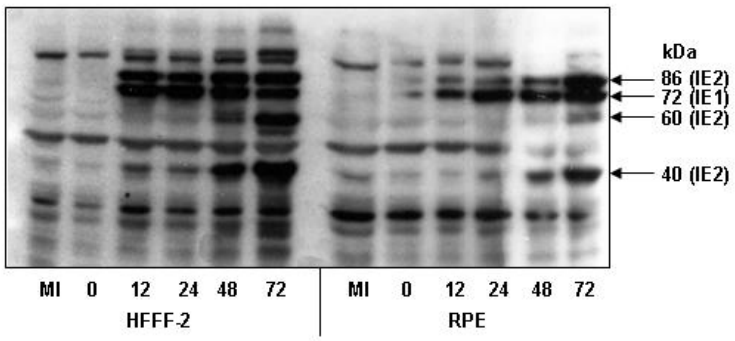
harvested in 200 µl/dish of SDS-PAGE sample buffer at 0, 12, 24, 48 and 72 h PI (and 72 h PI only for MI cell cultures). Equal volumes of protein samples were loaded into each gel slot and then separated by electrophoresis on SDS-polyacrylamide gels, transferred to a nylon membrane, then probed for IE1/IE2, pUS22 and pp28 using antibodies specifically directed against these target proteins.

The western blot for IE1/IE2 (Fig. 3.5) detected IE1 (72 kDa) expression from 12 h PI in each of the three cell types, reaching an apparent steady state level of expression between 12 and 24 h PI over the time course. IE2 (86 kDa) is similarly detected from 12 h PI in each cell type, but in contrast to IE1, IE2 gradually accumulates over the period 12 to 72 h PI. An early 60 kDa splice variant of IE2 is detected at 48 h PI in HFFF-2 cells and accumulates through to 72 h PI. This protein is also detected at 72 h PI in RPE infected cells, but not in U373Mg cells. The late IE2 (40 kDa) product is expressed from a transcript arising from an internal transcription start site within exon 4 of IE2 (Jenkins et al., 1994). Expression of IE2 (40 kDa) can be detected from 12 h PI in infected HFFF-2 cells but accumulates rapidly from 48 h PI through to 72 h PI in each cell type. No difference in the kinetics of expression of major IE transactivating proteins IE1 (72 kDa), IE2 (86 kDa) and the late IE2 (40 kDa) spliced product were detected in HFFF-2, RPE and U373Mg cells, although gross amounts differed between the cell types. In contrast, the IE2 (60 kDa) spliced product was not detected in U373Mg cells.

US22 is an early expressing gene, expressing pUS22 (76 kDa) a component of the virus tegument. The western blot for the pUS22 protein (Fig. 3.6) shows expression from 12 h PI in each of the three cell types, accumulating with time PI. As with the IE1 (72 kDa) and IE2 (86 kDa) proteins, there was no apparent delay in the onset of expression for this early protein in HFFF-2, RPE and U373Mg cells, although the amounts made, especially at 12 h and 24 h PI, was very much less in RPE and U373Mg cells.

Western blots for the late UL99 gene (Fig. 3.7) detected expression of pp28 (28 kDa) from 24 h PI in each cell type, that increases sharply at 48 h PI and 72 h PI. Again, no difference in the onset of pp28 expression is detected in HFFF-2, RPE and U373Mg cells, although amounts of the protein made at each time point varies between the cell types. It should be noted that both pUS22 and pp28 are components of the virus tegument, and bands for both proteins can be detected in the 0 h PI sample due to its presence in the input virus.

A.



B.

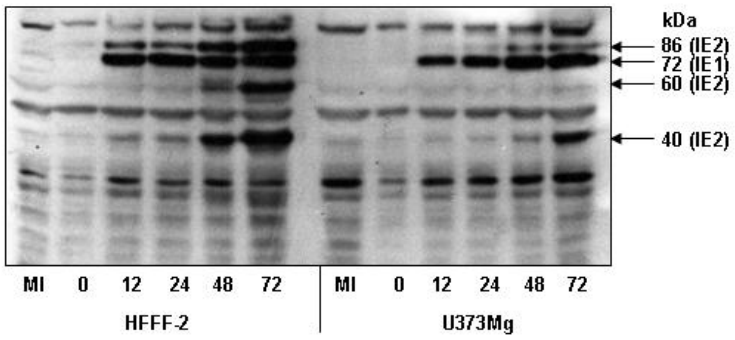


Figure 3.5. Western blots for the immediate-early proteins IE1/IE2.

Panel A shows the western blot for HCMV infected and mock-infected HFFF-2 and RPE cells, and panel B shows the western blot for HCMV infected and mock-infected HFFF-2 and U373Mg cells. In figures 3.5, 3.6 and 3.7, note that the HFFF-2 samples were identical and run in parallel in two SDS-PAGE gels. Apparent differences in accumulation of IE1/IE2 are not differences in amount but degree of separation in the two gels. Antibodies used were: anti-IE1/IE2 (gift from Dr E. Mocarski), which detects all spliced variants generated from the UL122 and UL123 genes. Rainbow markers (Amersham Biosciences) were run at either side of the of the gel in the first and last lanes.

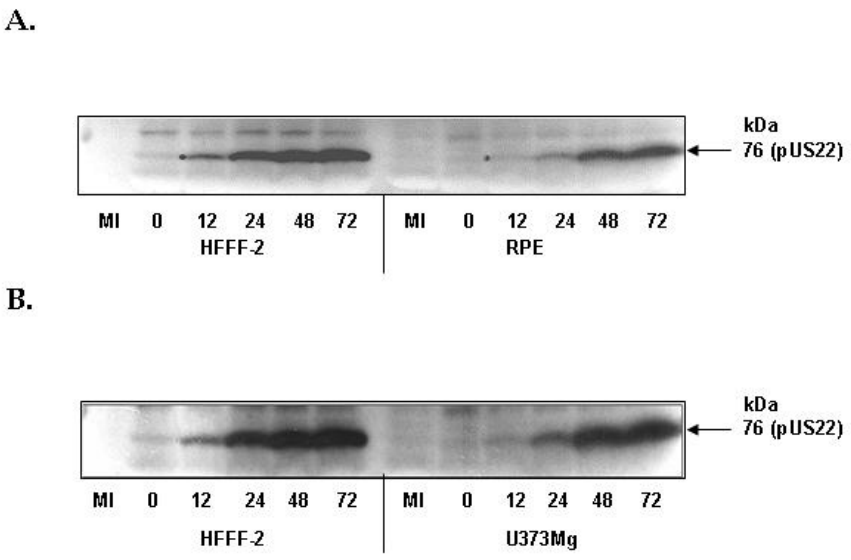


Figure 3.6. Western blots for the tegument component pUS22.

Panel A shows the western blot for HCMV-infected and mock-infected HFFF-2 and RPE cells, and panel B shows the western blot for infected and mock-infected HFFF-2 and U373Mg cells. Anti-pUS22 (Capricorn products). Rainbow markers were run at either side of the gel.

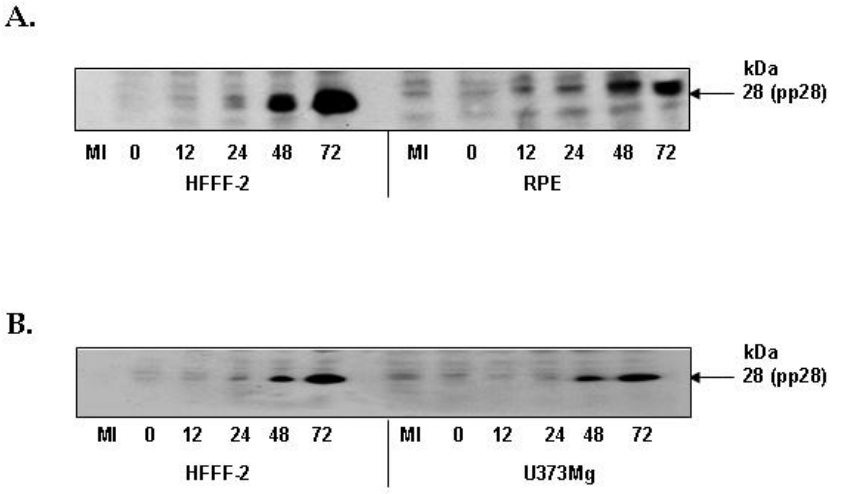


Figure 3.7. Western blot for the late gene UL99 (pp28).

Panel A a shows the western blot for HCMV-infected and mock-infected HFFF-2 and RPE cells, and panel B shows the western blot for infected and mock-infected HFFF-2 and U373Mg cells. Anti-pp28 (Capricorn Products). Rainbow markers were run at either side of the gel.

3.5 Discussion

Human foreskin fibroblast (HFF) cells have been used extensively in HCMV research because of the high yields of virus produced. Consequently, the HFFF-2 cell line was used here as a reference cell line for the microarray work. The biological studies have shown that there was no significant difference in the ability of HCMV to enter and initiate infection in the three cell types. However, the growth curves did reveal differential replication kinetics.

For the growth curve performed at a m.o.i. of 1 p.f.u./cell (Fig. 3.1), there was a delay of 24 h in the exit from the eclipse phase in RPE and U373Mg cells compared to HFFF-2 cells. The rate of replication from 48 to 96 h PI was similar in each cell type; however, the replication of HCMV plateaus at 96 h PI in U373Mg cells and at 144 h PI in HFFF-2 cells, while replication in RPE cells continues to increase up to 168 h PI. For the growth curves performed at a m.o.i. of 6 p.f.u./cell, the virus yield was fractionated into cell-released virus (CRV) (Fig. 3.3) and cell-associated virus (CAV) (Fig. 3.2) yields. As with the growth curve in Fig. 3.1, the CRV and CAV growth curves revealed differential replication kinetics with respect to the duration of the eclipse phase, delayed by a further 24 h in RPE and U373Mg cells compared to HFFF-2 cells, and also the release of CRV to the medium, and to a lesser extent, the accumulation of CAV. The accumulation of CAV was similar in HFFF-2 and RPE cells, but limited in amount and duration in U373Mg cells. When compared to HFFF-2 cells, the release of CRV to the medium was both delayed and occurred more slowly in RPE and U373Mg cells. The growth curves (Fig. 3.1 to 3.3) showed that the rate of replication in U373Mg cells was slower and shorter in duration compared to HFFF-2 and RPE cells, while the release of virus to the medium was delayed and slower in RPE and U373Mg cells when compared to HFFF-2 cells (Fig. 3.2). In the growth curve performed at a m.o.i of 1 p.f.u./cell (Fig. 3.1), the final virus yields were approximately 10 and 1000 fold lower in RPE and U373Mg cells respectively compared to HFFF-2 cells, while the CRV (Fig. 3.2) and CAV (Fig. 3.3) curves gave a final combined CRV/CAV virus yield that was approximately 25 and 1500 fold lower in RPE and U373Mg cells respectively when compared to HFFF-2 cells (data not shown). However, the growth curves were in agreement with respect to the duration of the eclipse phase in RPE and U373Mg cells, and the short duration of replication in U373Mg cells, when compared to HFFF-2 cells. The curves (Fig. 3.2 and 3.3) suggest impaired release of infectious virus from RPE cells and restricted production of infectious virus in U373Mg cells.

The differential replication kinetics were not due to a delay in the onset or progression through the virus gene expression cascade as evidenced by IE, E and L protein expression, though it may be explained by lower levels of viral protein made in RPE and U373Mg cells compared to HFFF-2 cells. Smaller amounts of protein feed through into lower numbers of infectious virions produced, especially during the early stages of infection (exit from the eclipse phase). The low yields of CAV obtained from U373Mg cells argues for some impairment of particle maturation into infectious progeny since protein synthesis appeared less affected than that in RPE cells (Fig. 3.5, 3.6 and 3.7).

The implications for the microarray work are that: 1. when cells are infected at a m.o.i. of 6 p.f.u./cell, equal numbers of cells are infected in all three cell types; 2. the temporal kinetics of transcription are the same in each of the three cells types; 3. but, based on protein expression data, there may be a general lower level of transcript abundance in RPE and U373Mg cells compared to HFFF-2 cells. It is concluded then that, despite some differences in viral growth kinetics and final virus yields, a valid comparison of the HCMV temporal kinetics of transcriptome expression can be made at the same time points using these three cell types.

4 RESULTS II

4.1 HCMV microarray validation

The main objectives of HCMV microarray validation were to assess the specificity of the viral oligonucleotide probes, and to optimise the conditions for cDNA synthesis and hybridisation to the array. A flow-diagram depicting the major steps in the validation procedure is shown in Fig. 4.1.

4.1.1 HCMV microarray print designs

Microarrays were printed onto Corning GAPSII microscope slides using a BioRobotics Microgrid (see methods, 2.2.9.1). Microarrays were printed in single batches, with 45 slides printed/batch. The first batch of arrays contained only HCMV probes, and were used in the validation hybridisations described in sections 4.1.2 (Fig. 4.3) and 0 (Fig. 4.4). The second batch contained both HCMV probes and bacterial control probes, which were used in the validation hybridisations described in sections 4.1.3 (Fig. 4.5) and 4.1.4 (Fig. 4.6). The subsequent batches of microarrays were used for transcriptome profiling of HCMV in different human cell types, section 4.2 (Fig. 4.8). Fig. 4.2 shows the print design of the first two batches of arrays used for microarray validation, batch P1 and P2 respectively, while Fig. 4.7 shows the print designs of the microarrays used for the profiling of HCMV gene expression in different cell types.

4.1.2 Assessment of the specificity of viral oligonucleotide probes

5′- and 3′- oligonucleotide probes were designed for HCMV ORFs or proposed ORFs as detailed in methods 2.8.1. The initial hybridisation experiments tested the specificity of the viral 60-mer oligonucleotide probes; this was achieved by hybridising mock-infected or HCMV-infected cell RNA to the array. At 96 h PI, total RNA was extracted from mock-infected or HCMV-infected HFFF-2 cells. Labelled cDNA was produced in a reaction containing 25 µg of DNase I treated total RNA, and labelled with Cy3-dCTP. Mock-infected or HCMV-infected labelled cDNAs were hybridised to separate arrays, each of which contained only HCMV probes, in order to test probe specificity (Fig. 4.3A and B). Hybridisation of HCMV-infected cell RNA showed that the spot morphology was uniform across the array, and triplicate spots gave consistent signal intensities, with low background (Fig. 4.3A). Hybridisation of mock-infected cell RNA confirmed the absence of cellular cDNA binding to the viral probes (Fig. 4.3B).

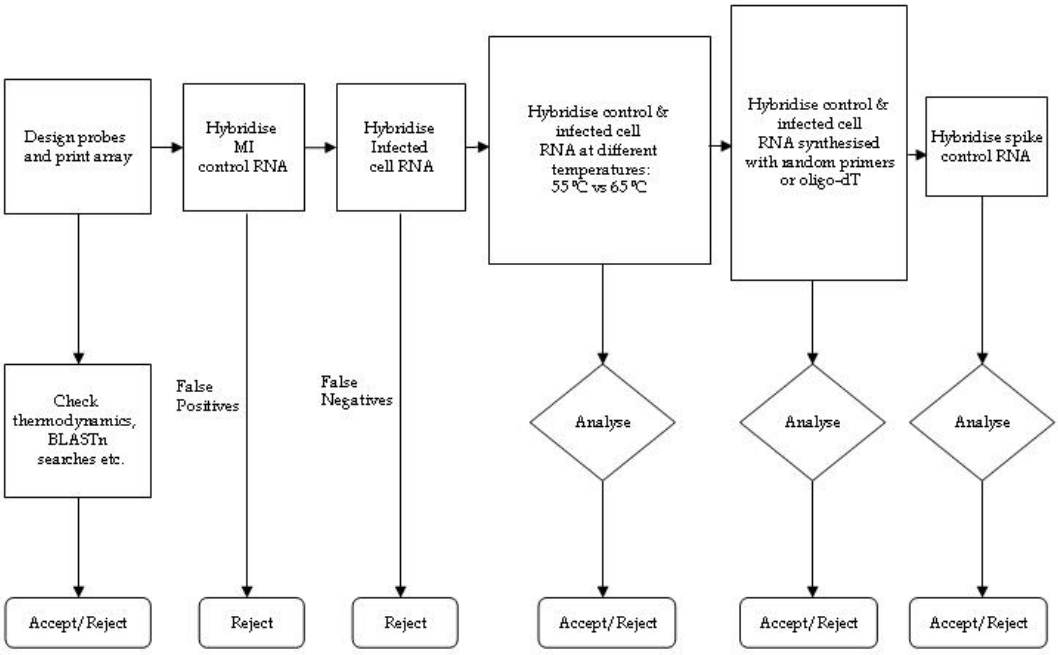


Figure 4.1: Design and validation procedure for the HCMV microarray

The flow-diagram shows the major steps in the design and validation of the HCMV microarray. Probes were designed and assessed using the Oligo6 software, followed by BLASTn searches; unsuitable probes were rejected. The probes were synthesised and printed onto microscope slides, following which, mock-infected or HCMV infected cellular RNA were hybridised in order to check the specificity of the viral probes. Hybridisations were then performed at different temperatures in order to optimise the array hybridisation temperature. cDNA synthesised using random primers or oligo-dT was then hybridised in order to determine which method gave the strongest hybridisation signals. Finally, spike control RNA was hybridised to the array to check the specificity of the control probes.

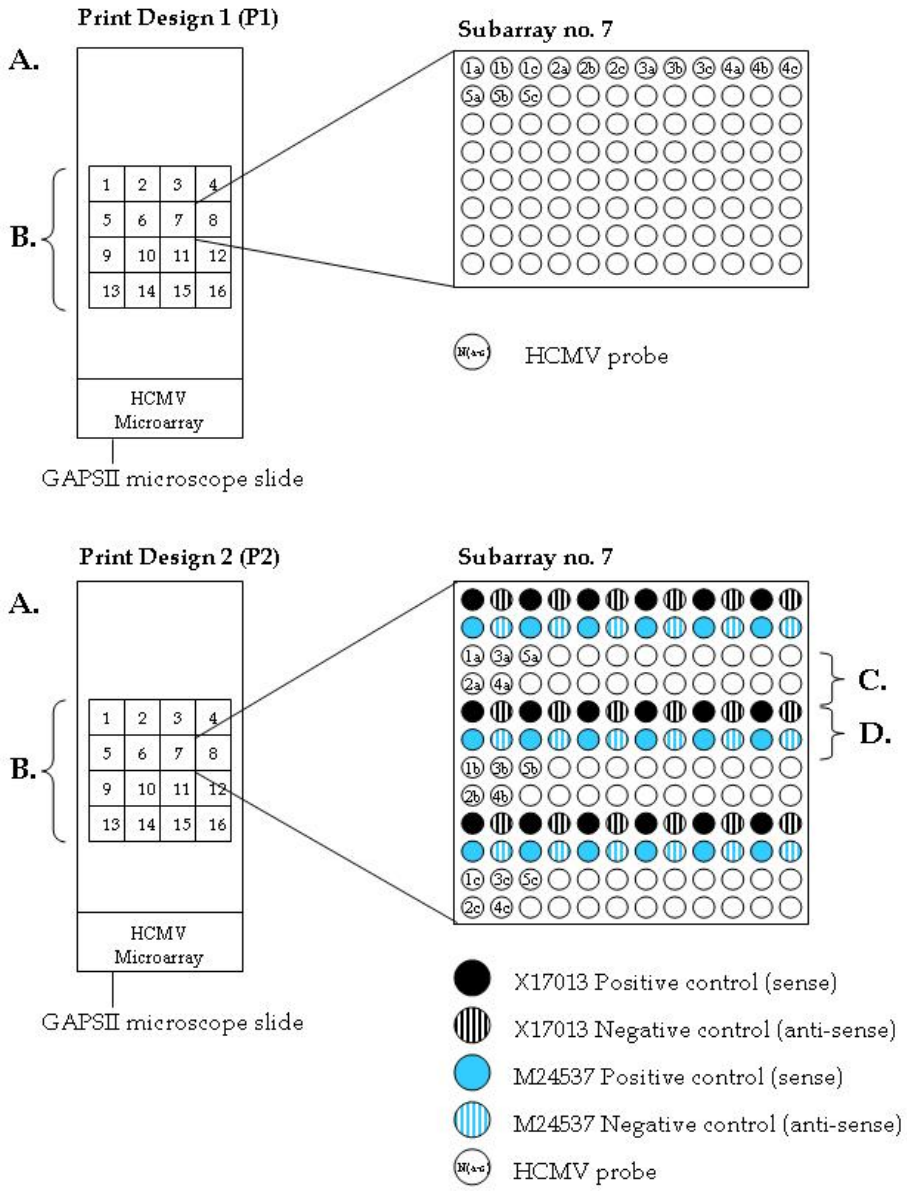
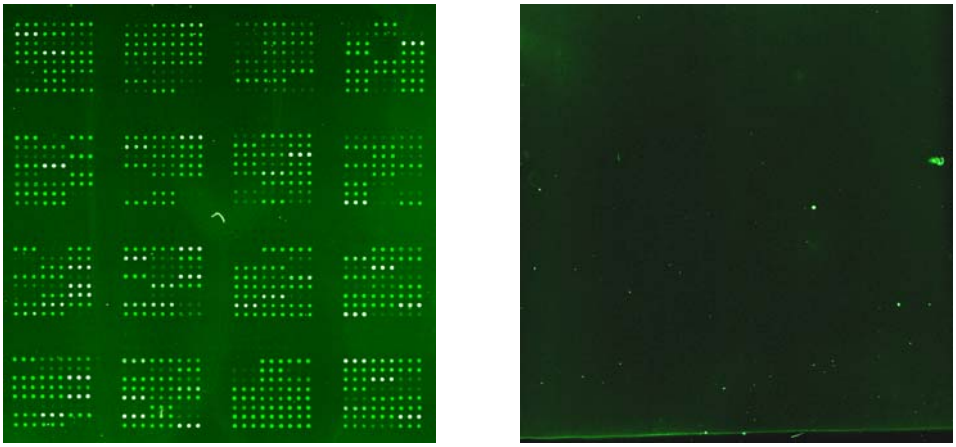


Figure 4.2. HCMV microarray print designs (P1) and (P2).

(P1) Shows the print design of the array for the first set of validation experiments; note that microarrays contained HCMV probes only. (P2) Shows the print design of the array for the second set of validaiton experiments; note that this batch of microarrays contained bacterial control probes and HCMV probes. (A) GAPSII microscope slide onto which the microarray was printed, as indicated by area (B); note that each microarray contains 16 subarrays. A magnified view of subarray number 7 shows the layout of the array probes. (C) In all arrays, HCMV probes are printed in triplicate, a, b and c; (D) bacterial control probes, printed in forward (sense) and reverse (anti-sense) orientation.



A. **B.**

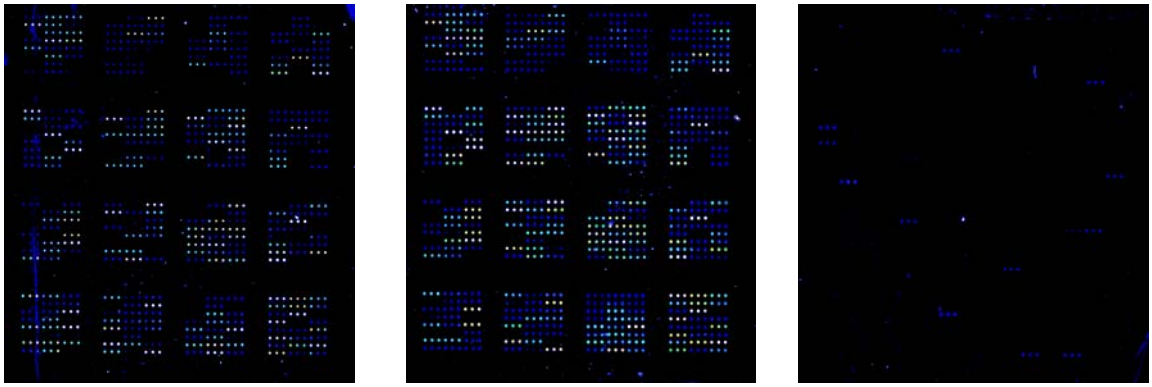
Figure 4.3. Assessment of probe specificity
(A) hybridisation of infected cell RNA to the array. (B) hybridisation of mock-infected cell RNA to the array.

4.1.3 The effect of hybridisation temperature on the specificity of the microarray

Varying the hybridisation temperature influences the binding efficiency and/or specificity of target cDNA molecules to cognate probes on the array. Optimising the temperature of hybridisation facilitates true-positive hybridisation and reduces non-specific (false-positive) hybridisation (Stekel, 2003). The hybridisations shown in Fig. 4.3 were performed at 65 ^{o}C, as described for the MHV-68 array (Ebrahimi et al., 2003). In order to investigate the effect of an alternative hybridisation temperature, two separate cDNA synthesis reactions were prepared using the same sample of HCMV-infected cell RNA. One cDNA preparation was hybridised to an array at 55 ^{o}C, and the second to another array at 65 ^{o}C. A mock-infected cDNA preparation was also hybridised to a third array at 55 ^{o}C. Following hybridisation, the three arrays were scanned at the same PMT gain setting (PMT50), so that a direct comparison could be made (Fig. 4.4A, B and C). The images clearly show that the hybridisation performed at 55 ^{o}C gave stronger and more abundant signals when compared with the hybridisation performed at 65 ^{o}C, the overall intensity of spots in (Fig. 4.4B) is higher than in (Fig. 4.4A), and some spots absent in (Fig. 4.4A) can be detected in (Fig. 4.4B). Assessment of the HCMV probe signal intensities of the two arrays hybridised at 55 °C (Fig. 4.4B) and 65 °C (Fig. 4.4A) showed that the median signal intensity increased by 1.3 fold for the array hybridised at 55 °C compared to the array hybridised at 65 °C (Table 4.1). The number of HCMV probes detecting a positive signal (an arbitrary signal cutoff point of 300 was selected) also increased when hybridised at 55 °C (Table 4.1). However, in contrast to hybridisations at 65 ^{o}C (Fig. 4.3B), hybridisation of mock-infected cDNA to the array at 55 ^{o}C resulted in a low level of non-specific binding to some viral probes on the array (Fig. 4.4C). To maintain probe specificity, all subsequent arrays were therefore hybridised at 65 ^{o}C.

4.1.3 Comparison of cDNA synthesis methods on the specificity of the microarray

cDNA was labelled by the direct incorporation of CyDye conjugated nucleotides during first strand synthesis of cDNA. We selected this strategy because labelled cDNA is generated easily and rapidly, and no further downstream processing is required other than cleaning the cDNA prior to hybridisation. However, there are some disadvantages in using this method: incorporation of large CyDye conjugated nucleotides can result in the premature termination of cDNA synthesis, yielding truncated cDNA copies of mRNAs (personal communication; Dr Tom Freeman, Sanger Institute). Because some regions of the HCMV genome code for 3' co-terminal transcripts spanning several kb, it is important that the cDNA generated represents the full length mRNA.



A. **B.** **C.**

Figure 4.4. Confirmation of different temperatures on specificity of the microarray

(A) HCMV infected cDNA was hybridised at 65 ºC. (B) HCMV infected cDNA hybridised at 55 ºC. (C) Mock-infected cDNA hybridised to the microarray at 55 ºC.

Experiment condition	Median signal intensity	Fold difference	Number of probes with a signal intensity > 300 Total no. 1152
65 °C	5883	1.3	925
55 °C	7652		1042
Random Priming	6579	1.25	1008
Oligo-dT	5247		903

Table 4.1. Assessment of microarray signal intensities during hybridisation of HCMV-infected HFFF-2 RNA under different experimental conditions.
Hybridisations were performed with HCMV-infected HFFF-2 RNA at 55 °C and 65 °C, for which the median signal intensity was calculated from the HCMV microarray probes. The table also shows the fold difference of the median signals between the conditions tested, and also the number of HCMV probes that gave signals > 300.

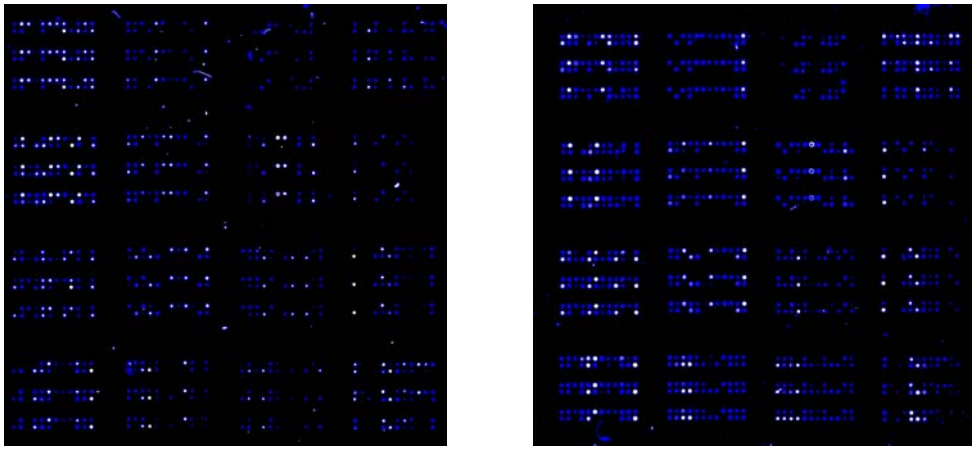
Bacterial control probes	Median signal intensity
X17013	5510
M24537	12382

Table 4.2. Assessment of median signal intensities for bacterial control probes.
The signal intensities were generated from bacterial control RNA (X17013 and M24537) each spiked into the cDNA synthesis reaction each at 50 ng. Note that control probes were printed on the microarray in a 2:1 ratio (M24537:X17013) which accounts for the doulbing of the median signal intensity of M24537 compared to X17013 when assessing the positive control signals (sense probes).

The cDNA synthesis and labelling strategy selected could skew or otherwise compromise the microarray data obtained. Priming cDNA synthesis reactions with oligo-dT (i.e. 3' orientated) might result in truncated cDNA copies of mRNAs, whereas priming the cDNA synthesis reactions with random hexamers should result in the synthesis of cDNAs that represent the full-length transcripts. To investigate the effect of cDNAs synthesised using random hexamers or oligo-dT on the efficiency and/or specificity of detection of array probes, two cDNA synthesis reactions were prepared from the same sample of HCMV-infected cell RNA: one reaction was primed using random hexamers, the other using oligo-dT primers. The two cDNA samples were then hybridised to separate arrays at 65 ^{o}C. Following hybridisation, the slides were scanned at the same PMT gain setting (PMT50), to allow direct comparison of the array images (Fig. 4.5A and B). While intra-array triplicate hybridisation signals appear evenly labelled in both (Fig. 4.5A) and (Fig. 4.5B), there is a greater sensitivity with cDNA synthesised using random hexamers (Fig 4.5B). Assessment of the HCMV probe signal intensities of the two arrays hybridised with cDNA synthesis primed with random primers (Fig. 4.5B) and oligo-dT (Fig. 4.5A) showed that the median signal intensity increased by 1.25 fold for the array hybridised with cDNA synthesised using random primers compared to the array hybridised with cDNA synthesised using oligo-dT (Table 4.1). The number of HCMV probes detecting positive signal intensities (> 300; see 1.1.1) also increased when hybridised with cDNA synthesis primed with random hexamer primers (Table 4.1). Consequently, all subsequent cDNA synthesis reactions were performed using the random hexamer priming method.

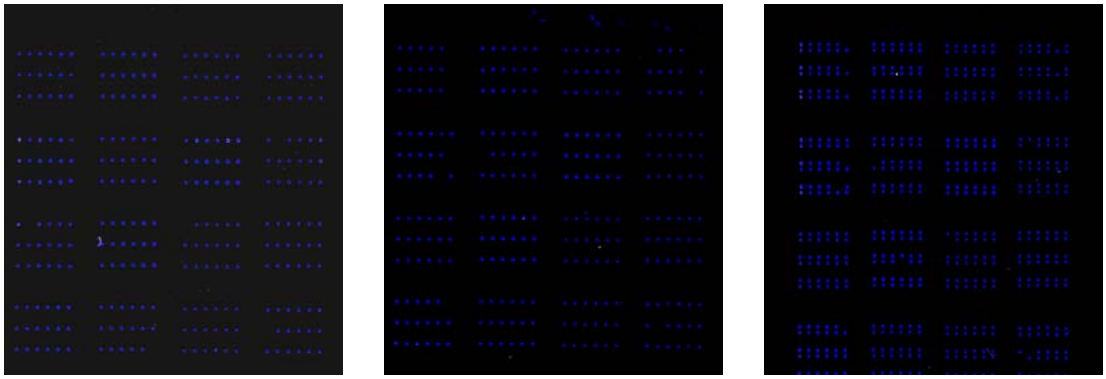
4.1.4 Specificity of the microarray for the detection of control features

In order to check the specificity of bacterial control probes (see 2.1.10.1), hybridisations were performed using the 'spiked in' control mRNA. cDNAs were synthesised and labelled using ~ 1 μg each of X17013 and M24537 mRNA in separate reaction tubes, then hybridised to separate arrays (that contained the full complement of viral and control probes). Scanning of the arrays (PMT50) showed that only the bacterial probes on the array were detected, thus confirming the specificity of the two control probes (Fig. 4.6A and B). Because the data in Fig. 4.6A and B was obtained with cDNA synthesised with control mRNA in excess (~ 1 μg), the experiment was repeated using 50 ng each of X17013 and M24537, with cDNA synthesis and labelling performed in the same reaction tube. The labelled cDNAs were then hybridised to the array and scanned as before (Fig. 4.6C). Since the array shows equal hybridisation levels, without evidence of signal saturation, poorly defined signals, or non-specific binding, 50 ng each of the bacterial control mRNAs was selected for subsequent spiked controls in cDNA synthesis reactions. Hybridisation of



A. B.

Figure 4.5. Hybridisation of cDNA that had been synthesised using (A) oligo-dT; or (B) random hexamer primers.
cDNA was synthesised using the same sample of HCMV infected HFFF-2 RNA (96 h PI) using (A) oligo-dT, or (B) random hexamer primers.



A. **B.** **C.**

Figure 4.6. Hybridisation of control 'spiked in' RNA.
(A) Hybridisation of spike mRNA 1 (X17013); (B) hybridisation of spike mRNA 2 (M24537); and (C) hybridisation of the spike RNA (X17013 and M23547) each at 50 ng.

mock-infected HFFF-2 RNA containing the bacterial spike control RNA (each spiked in at 50 ng) was also performed. The median signal intensities from the X17013 sense probes and M24537 sense probes generated from this hybridisation are shown in Table 4.2 (pg. 109).

4.2 Transcriptome profiling of HCMV in fibroblasts, epithelial cells and astrocytes

The HCMV microarray was employed to investigate the temporal pattern of global virus gene expression during replication in three different cell types (HFFF-2, human foetal foreskin fibroblasts; RPE, human retinal pigmented epithelial cells; and U373Mg, human astrocytes), with the specific aim of identifying viral genes that are differentially regulated. Because of the complexity of the experiment (i.e. a time course experiment in multiple cell types), single colour arrays were used. The microarray print design is shown in Fig. 4.7. The details of the experiment hybridisations and the associated results for the normalisation, ROC analysis etc. are shown in Table 4.3. A flow-diagram depicting the stages of microarray analysis is shown in Fig. 2.1 (see methods, 2.2.9).

4.2.1 HCMV microarray hybridisations

Total RNA samples were prepared from three biological replicates of mock-infected cells or HCMV-infected HFFF-2, RPE and U373Mg cell cultures. 25 µg each of total RNA sample (DNase I treated) was then spiked with 50 ng each of *B.subtilis* X17013 and M24537 mRNA ('spiked in' control RNA). cDNA was then synthesised using random primers and direct incorporation of Cy3-dCTP. The labelled cDNA preparations were then hybridised to separate HCMV microarrays at 65 ^{o}C and scanned using the Perkin-Elmer ScanArray Express scanner, at PMT gain settings of 40, 50, 60 and 70 %. To monitor progression through the HCMV global gene expression cascade, total RNA was prepared from three biological replicates at 12, 24, 48 and 72 h PI for infected cells, and at 72 h PI for MI control cell cultures. Time points prior to 12 h PI were excluded due to the low level of viral gene expression (personal communication; Dr Bahram Ebrahimi). Representative images (1 of 3 biological replicates for each cell type) of scanned arrays for each infected cell type from 12, 24, 48, and 72 h PI samples respectively are shown in Fig. 4.8. The microarrays from these experiments produced high quality data, with excellent spot morphology and low background noise.

Low numbers of virus specific signals were detected in infected HFFF-2 and RPE cells at 12 h PI (Fig. 4.8) but the intensity and abundance of signals continued to increase with time as the HCMV genome became more transcriptionally active. In the case of infected

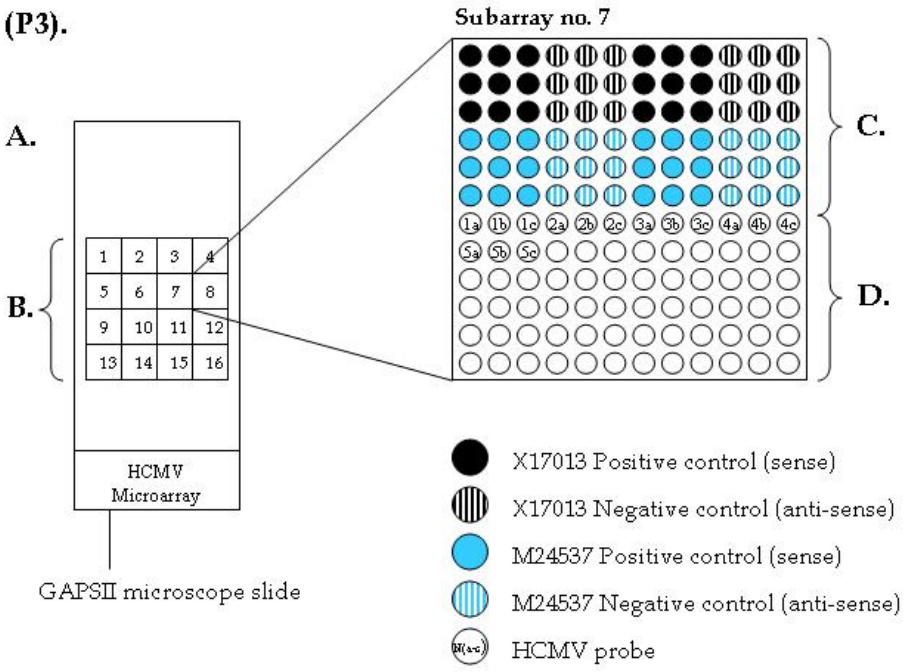


Figure 4.7. HCMV microarray print design (P3).

(P3) Shows the print design of the HCMV microarray that was used for transcriptome profiling in fibroblasts, epithelial cells and astrocytes. (A) GAPSII microscope slide onto which the probes were printed. (B) HCMV microarray area – contains 16 subarrays, each subarray contains 144 probes arranged in a 12 x 12 formation, and printed according to the layout displayed in the magnified view of subarray number 7. (C) Indicates the region of the subarray where the control probes were printed. (D) Indicates the region of the subarray where the virus probes were printed. Note that each virus probe was printed in triplicate (a, b and c).

Array Number	Cell Type	Time Point	R^2 Value	ROC Cutoff	Correction Factor (C.F.)
1	HFFF-2	12	0.988	10.423	3.744
2	HFFF-2	12	0.964	9.074	1.146
3	HFFF-2	12	0.941	7.185	3.584
4	HFFF-2	24	0.977	9.944	1.190
5	HFFF-2	24	0.911	9.350	1.014
6	HFFF-2	24	0.938	8.258	1.076
7	HFFF-2	48	0.997	8.384	1.111
8	HFFF-2	48	0.979	8.423	1.388
9	HFFF-2	48	0.998	9.343	1.069
10	HFFF-2	72	0.998	7.472	0.954
11	HFFF-2	72	0.998	7.833	1.262
12	HFFF-2	72	0.997	7.234	1.015
13	RPE	12	0.996	8.922	1.012
14	RPE	12	0.996	8.545	0.935
15	RPE	12	0.997	8.447	0.737
16	RPE	24	0.996	9.013	0.618
17	RPE	24	0.994	8.904	0.586
18	RPE	24	0.989	8.071	0.811
19	RPE	48	0.997	8.187	0.896
20	RPE	48	0.998	8.872	0.637
21	RPE	48	0.998	8.101	0.591
22	RPE	72	0.983	9.568	0.899
23	RPE	72	0.995	8.585	1.362
24	RPE	72	0.995	8.545	1.181
25	U373Mg	12	0.997	7.662	0.712
26	U373Mg	12	0.999	7.472	0.666
27	U373Mg	12	0.998	7.752	0.686
28	U373Mg	24	0.997	7.448	0.859
29	**U373Mg**	**24**	**0.998**	**6.889**	**1**
30	U373Mg	24	0.998	6.577	1.022
31	U373Mg	48	0.998	7.326	1.262
32	U373Mg	48	0.998	7.271	0.854
33	U373Mg	48	0.998	6.676	0.840
34	U373Mg	72	0.956	7.353	1.096
35	U373Mg	72	0.997	7.422	0.740
36	U373Mg	72	0.998	6.679	0.837

Table 4.3. Microarray data analysis determining selection cutoff points and correction factors.
The table lists the array number, cell type and time points, together with the associated data generated during assessment of the linear dynamic range, defining cutoff points using ROC analysis, and the calculated correction factors for data normalisation (scaling). The reference array (no. 29) values are in shown in red.

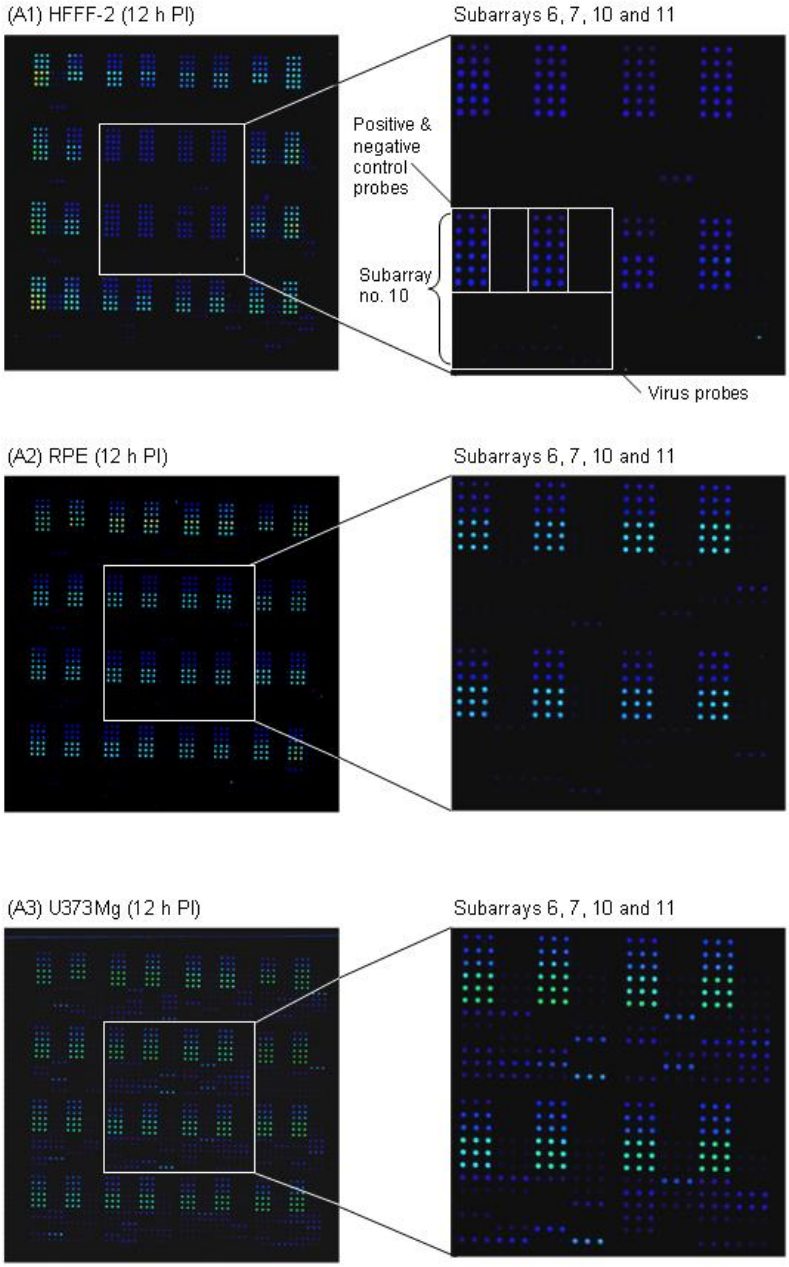
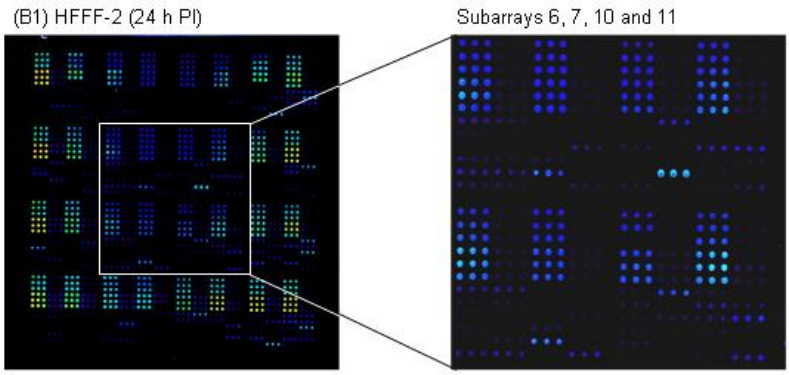
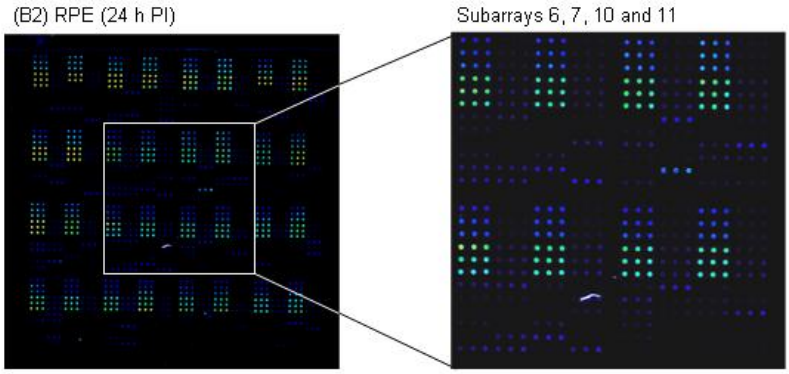


Figure 4.8. Representative microarray images from one of three biological replicates for HCMV infected HFFF-2, RPE and U373Mg cells.

cDNA was prepared from total RNA at 12, 24, 48 and 72 h PI from infected cells and from mock-infected cells after 72 h inoculation. The left panel images show the whole HCMV microarray area, while a magnified view of the central region with the micorarray is shown in the right panel of images, in order to highlight some of the control and virus probe spots within the central area. However, it must be noted that the surrounding subarrays (1,2,3,4,5,8,9,12,13,14,15,16) also contained both control and virus probe spots (see Fig. 4.7).





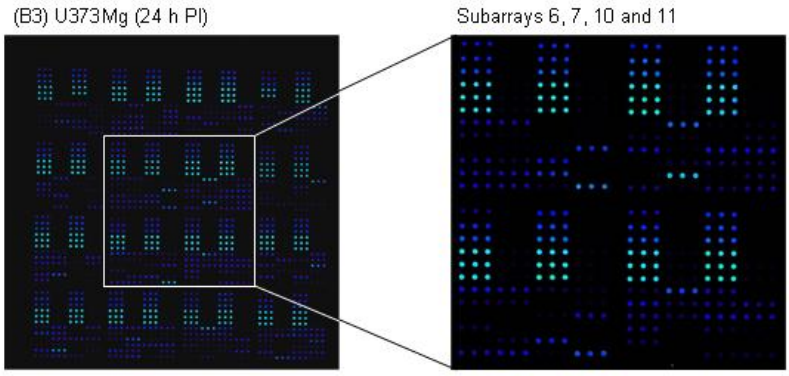


Figure 4.8 continued

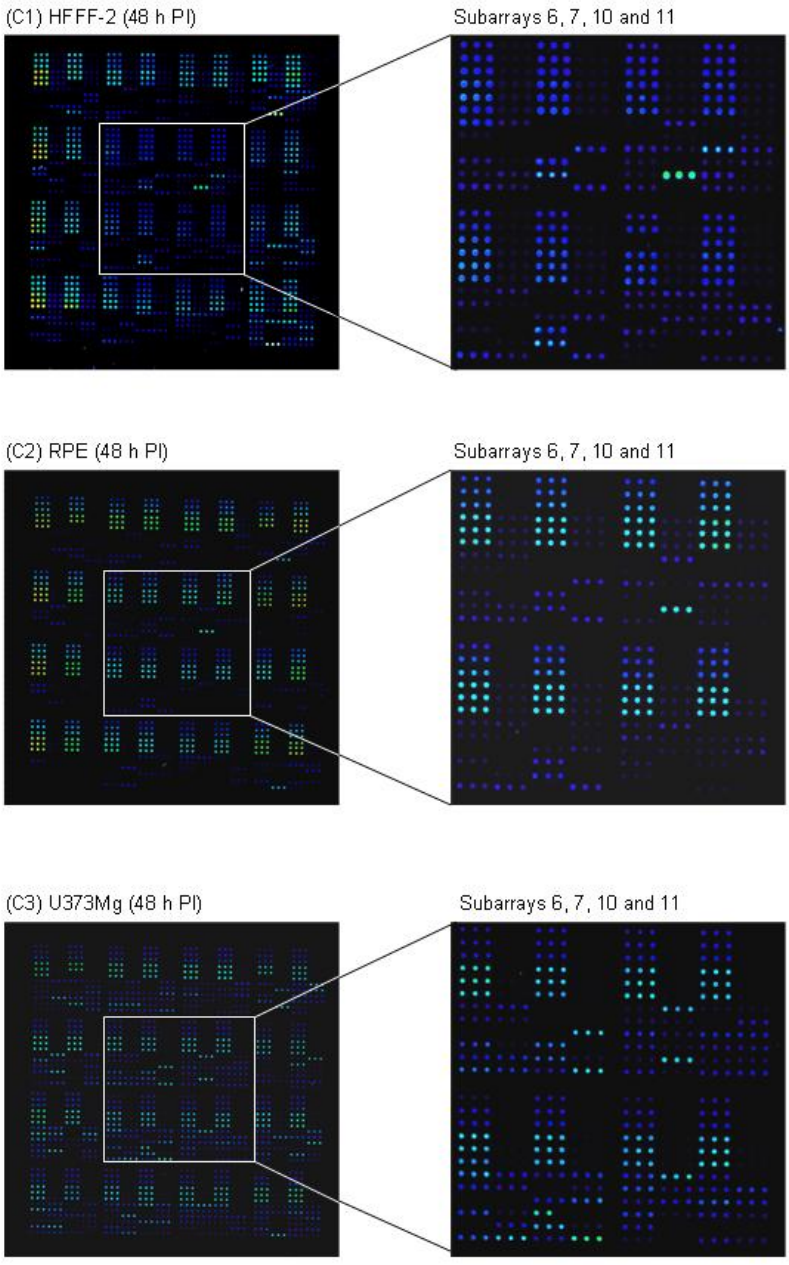
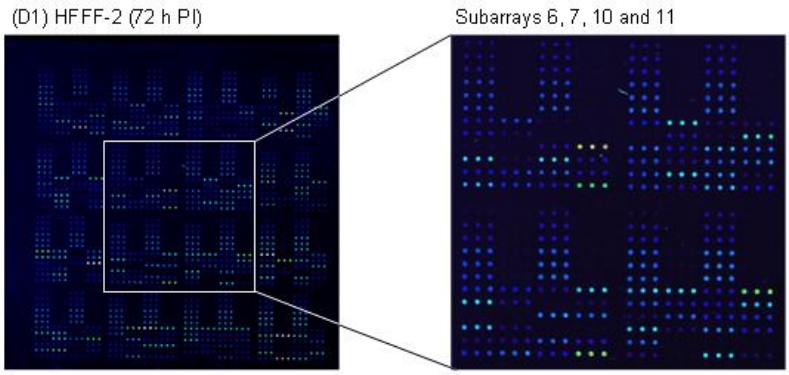
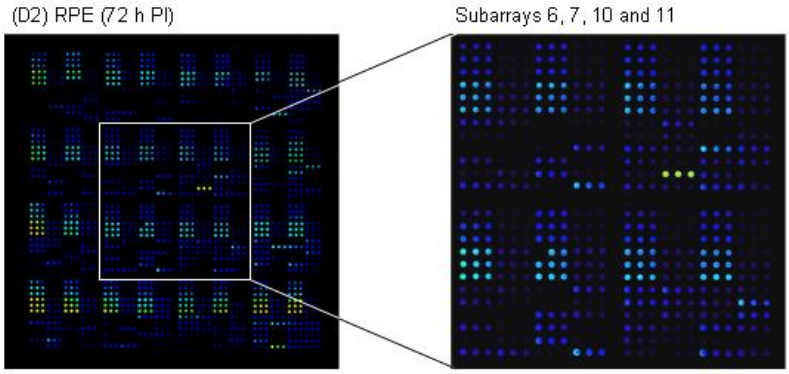


Figure 4.8 continued





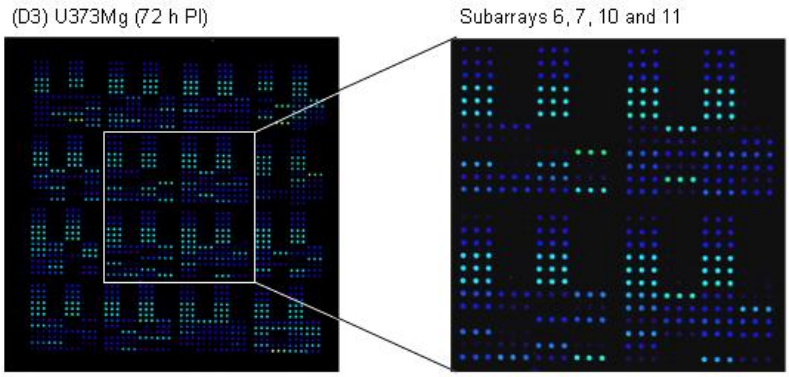


Figure 4.8 continued

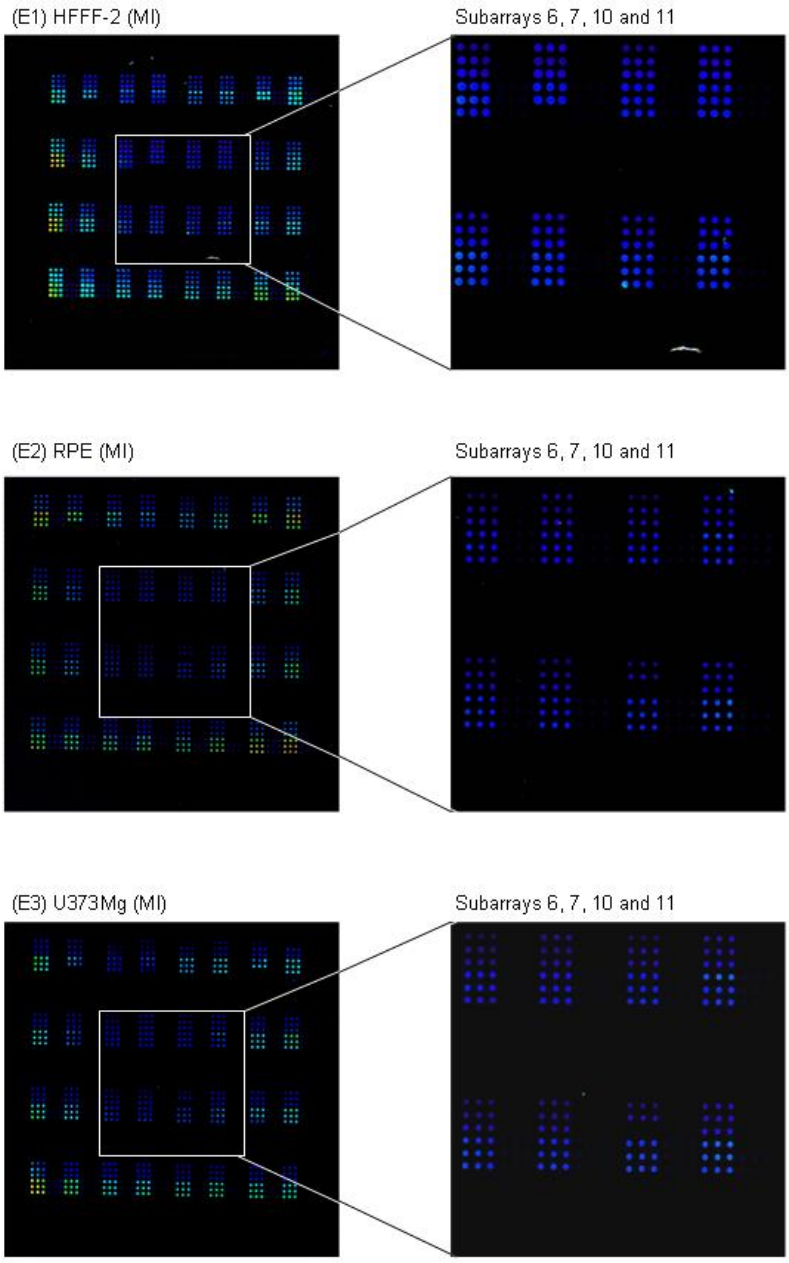


Figure 4.8 continued

U373Mg cells, it was apparent that a large proportion of viral genes were expressed as early as 12 h PI (Fig. 4.8). As expected, the strength of signals for U373Mg cells increased with time, but the incremental range over the period 12 h PI through to 72 h PI was not as large as seen with the HFFF-2 and RPE infected cells. The finding that the greater part of the HCMV transcriptome was active as early as 12 h PI in U373Mg cells was unexpected and contrasts with more strictly regulated activity seen with infected HFFF-2 and RPE cells. Consequently, the HCMV transcriptome profiling in U373Mg cells was repeated. Triplicate HCMV infections of fresh U373Mg cell cultures were performed and a second set of RNA samples were prepared at 12, 24, 48 and 72 h PI. Hybridisations were then repeated, and the expression profile analysed. Surprisingly, no difference could be found in the HCMV gene expression profiles obtained from the first and second set of U373Mg microarray data, confirming the unusually early transcription profile of HCMV in U373Mg cells (data not shown). Mock-infected cell RNA hybridisations for each cell type are also shown in Fig. 4.8, confirming the specificity of the virus probes.

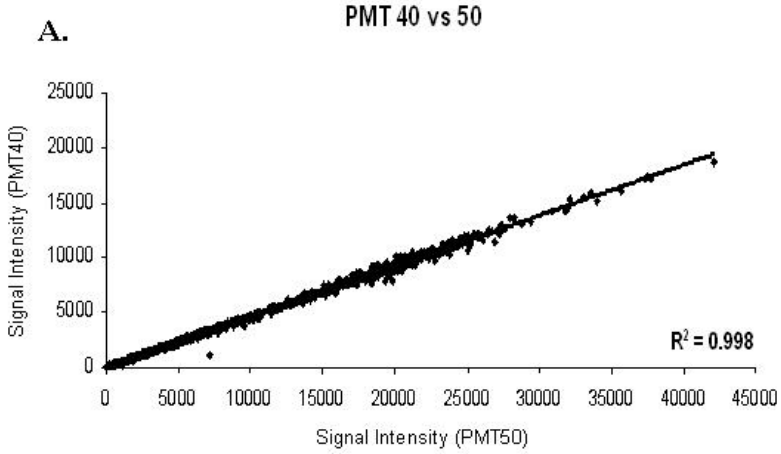
4.2.2 Determining the optimum PMT settings

(Fig. 2.1, stage 1, step 1)

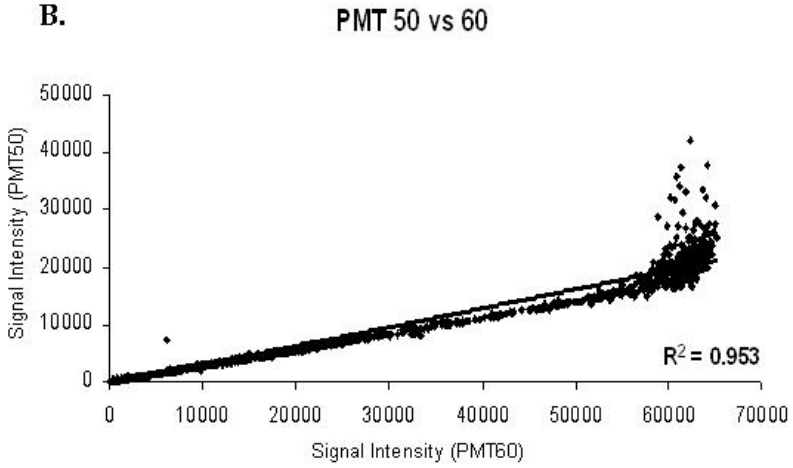
In order to determine the optimum gain settings, spot intensities were plotted at a series of PMT settings. Non-linearity indicates sub-optimal PMT since no linear increase in spot intensity is possible at that setting. The optimal PMT occurs at the highest PMT values where a linear range is obtained. The scatterplots were examined by R^2 least square fit analysis, and datasets exhibiting the optimum scatter were selected for further analysis. Fig. 4.9 shows the scatterplots of a representative microarray dataset. Scatterplot C (PMT 60 vs 70) shows deviation from the linear range and saturation of the microarray data at signal intensities > 60000. Scatterplot B (PMT 50 vs 60) also deviates from the linear range and also exhibits signal saturation, though at a lower level than scatterplot C. Finally, scatterplot A (PMT 40 vs 50) gives linear signal intensity data with a dynamic range between 0 to 45000 in the absence of saturation, and yielding a R^2 value of 0.9983. The dataset generated with PMT gain setting of 50 % was therefore selected for further analysis. The R^2 values for selected linear HCMV microarray datasets are shown in Table 4.3.

Raw signal intensity microarray data was right-skewed (when assessed by histograms, see Fig. 4.10), it is important therefore to $\log_2$ transform the data so that it assumes a normal distribution for subsequent statistical analysis. This is firstly because many statistical tests (used to analyse microarray data) are based on the assumption that the

A.


PMT 40 vs 50
Signal Intensity (PMT40)
Signal Intensity (PMT50)
R2 = 0.998
0
5000
10000
15000
20000
25000
30000
35000
40000
45000

B.


PMT 50 vs 60
Signal Intensity (PMT50)
Signal Intensity (PMT60)
R2 = 0.953
0
10000
20000
30000
40000
50000
60000
70000

C.

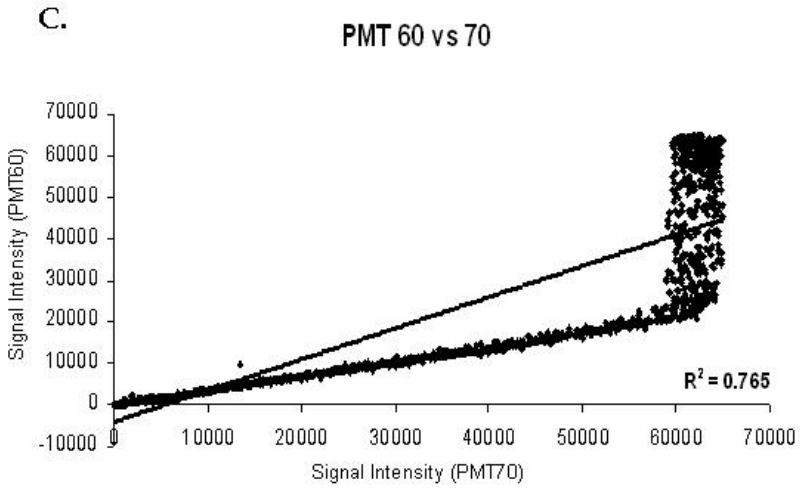

PMT 60 vs 70
Signal Intensity (PMT60)
Signal Intensity (PMT70)
R2 = 0.765
-10000
0
10000
20000
30000
40000
50000
60000
70000

Figure 4.9. Determination of linear dynamic range.

A, B and C show scatterplots for the array no. 36 (U373Mg at 72 h PI), with raw mean signal intensities plotted for PMT gain (%) settings (A) 40 vs 50; (B) 50 vs 60; (C) 60 vs 70. In graph C, there is evident signal saturation as the data curves at signal values > 60000; this is also the case in graph B, although to a lesser degree. Graph A shows the linear relationship achieved with PMT gain setting of 40 and 50, and a signal distribution from 0 to 45000. Least square fit values (R^2) are shown.

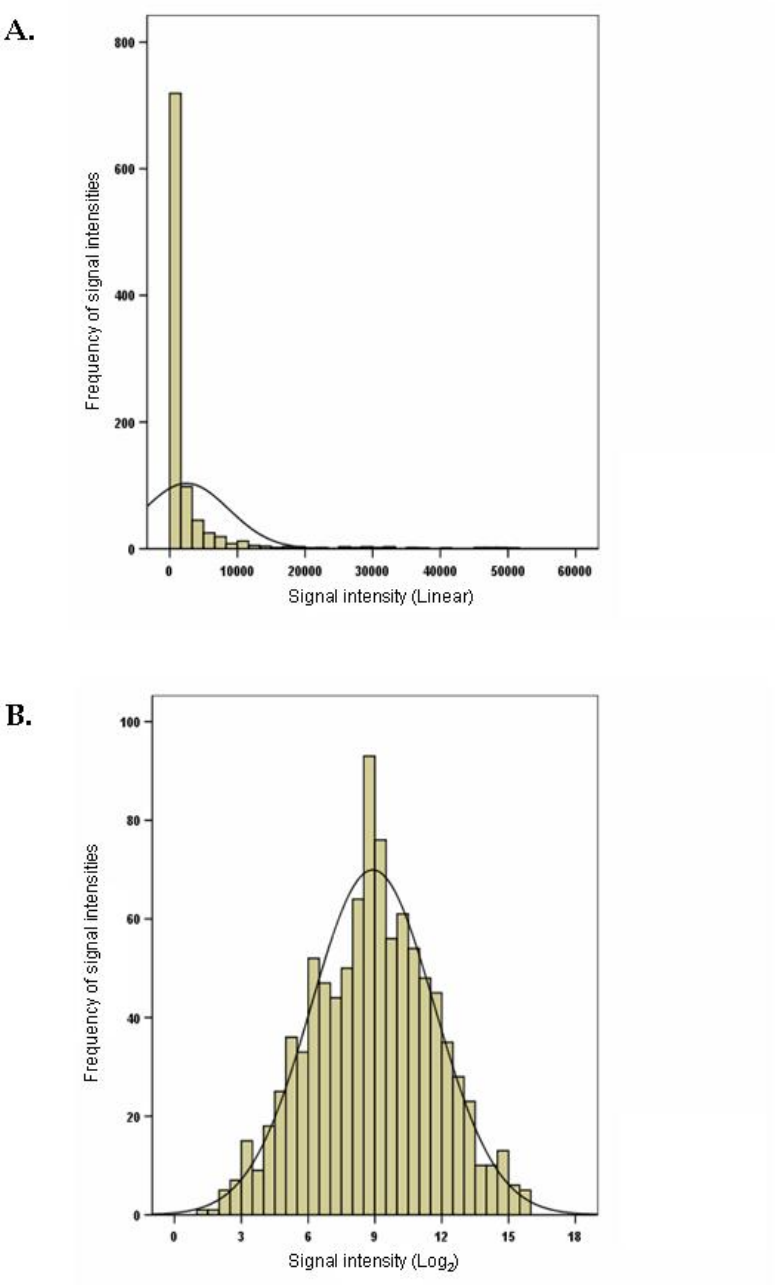


Figure 4.10. Histograms of (A) linear microarray data and (B) $\log_2$ transformed microarray data.

(A) Histogram of linear signal intensity data, demonstrating right-skewing of the data. (B) Histogram of the data following $\log_2$ transformation, demonstrating the approximation to a normal distribution.

data is symmetrically distributed, and secondly, because taking the logarithm of linear expression data limits the numerical range, and reduces the computational power required when performing statistical tests.

4.2.3 Assessment of signal distributions

(Fig. 2.1, stage 1, step 2)

Following base 2 logarithmic transformation of the raw signal intensity data, box-plots were drawn in order to ensure that the data followed a normal distribution before proceeding with further analysis. The box-plots for the entire HCMV microarray dataset for infected HFFF-2, RPE and U373Mg cells are shown in Fig. 4.11. The central line represents the median of the data, and the surrounding box represents the 25^{th} and 75^{th} percentiles, and the whiskers show the spread of the data. The box-plots confirm that following $\log_2$ transformation, the linear HCMV microarray datasets for each cell type are symmetrical. Box-plots were also drawn after the data was scaled (Fig. 4.11.1), and following normalisation in GeneSpring (Fig. 4.11.2) to ensure that the patterns of gene expression were not skewed following these analysis procedures.

4.2.4 Receiver operating characteristic analysis

(Fig. 2.1, stage 1, step 3)

Datasets that were linear and symmetrical were then subject to ROC analysis. Table 4.4 shows an example of ROC analysis for a representative microarray dataset. The table lists various cutoff points and their calculated false-positive and true-positive rates. The data can be plotted as a curve (Fig. 4.12A and B) of false-positive values against true-positive values for each cutoff point. A curve that lies toward the top left hand corner of the graph represents good discrimination between true-positive and true-negative signals, because the true-positive rate remains high while the false-positive rate is reduced with increasing signal cutoff thresholds. A curve that is linear represents data where no distinction can be made between true-positive and true-negative signals. Table 4.3 shows the ROC analysis cutoff points that were selected for the HCMV microarray datasets. This analysis was performed on the positive and negative control signals for all linear datasets, and cutoff points that gave a false-positive rate of 5 % were selected. Data that fell below the cutoff point was assigned a value of 0.01.

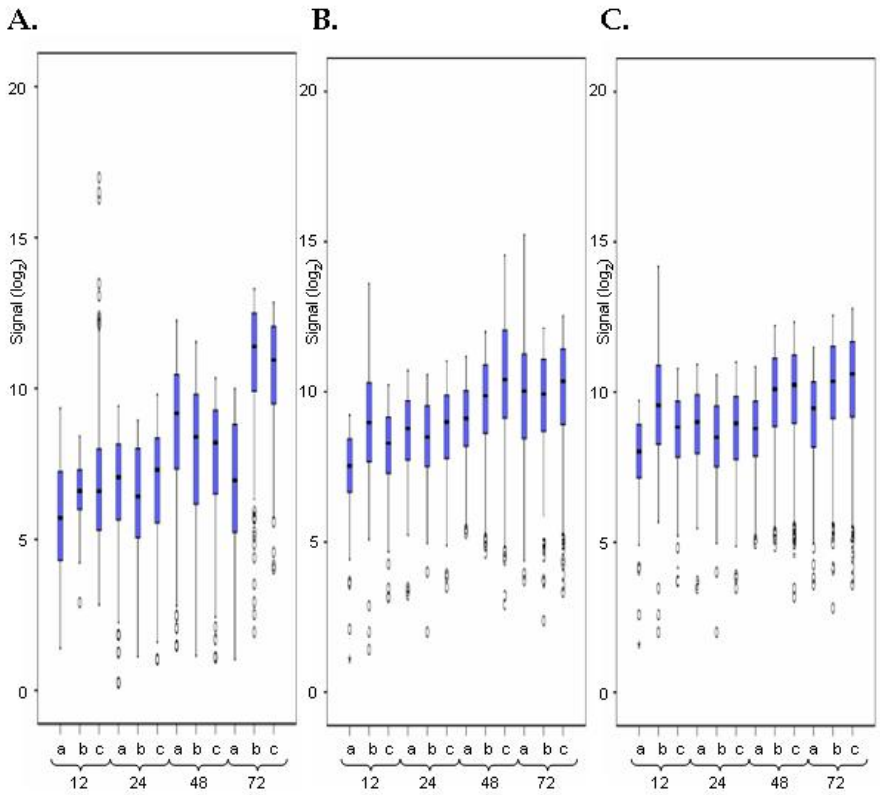


Figure 4.11. Box-plots for (A) HFFF-2, (B) RPE and (C) U373Mg microarray datasets prior to data scaling.

The time point and the biological replicate number is shown (a, b, or c). The box-plots represent the spread of the data following $\log_2$ transformation of the linear microarray data. The median of the data is represented by the thick black central line, and the box itself indicating the 25th and 75th percentiles. The whiskers represent the spread of the data. Signal values that are twice that of the median are represented by circles.

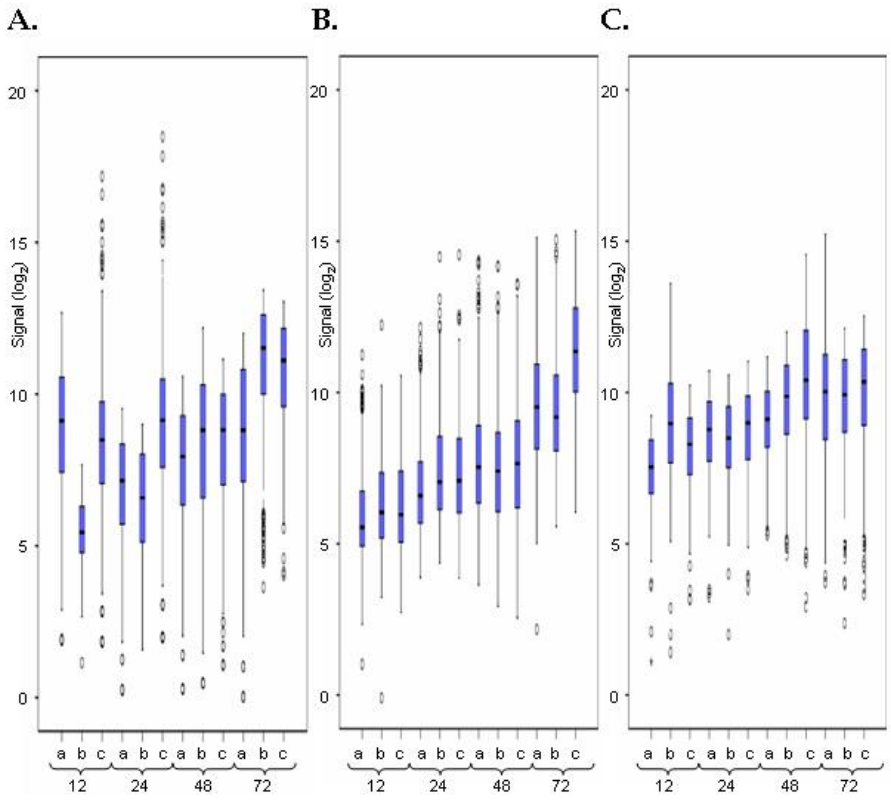


Figure 4.11.1. Box-plots for (A) HFFF-2, (B) RPE and (C) U373Mg microarray datasets following data scaling.

The microarray data was scaled against the 75th percentile of the positive control signals, in order to remove variation introduced during cDNA synthesis, hybridisation and scanning. Note that the datasets have been not centred, and that there is a general increase in the viral probe signals over the time course, from 12 h PI to 72 h PI.

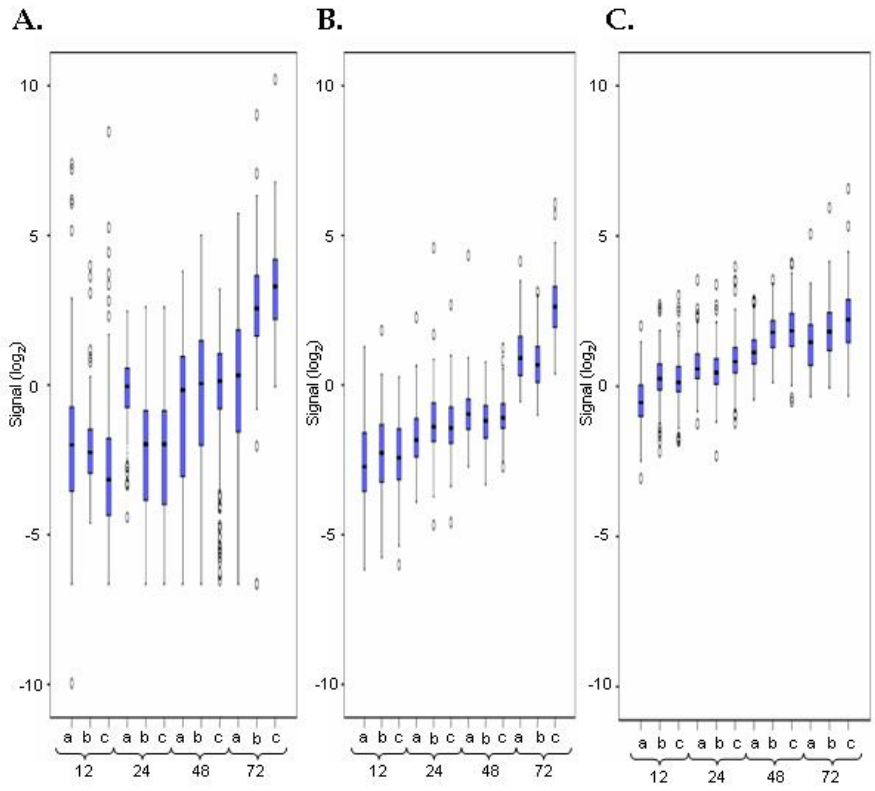


Figure 4.11.2. Box-plots for (A) HFFF-2, (B) RPE and (C) U373Mg microarray datasets following normalisation of the data in GeneSpring.

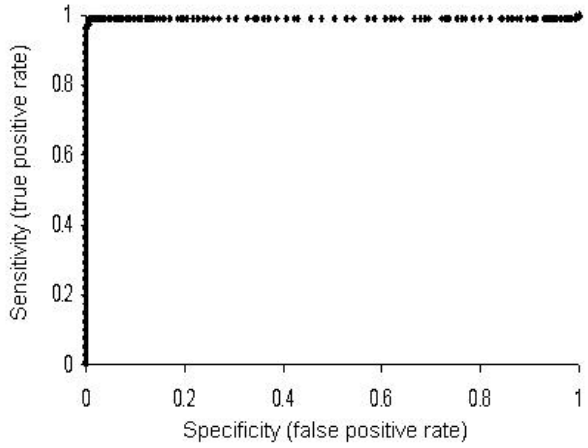
The normalisation procedure in GeneSpring 'per gene; normalise to the median' was performed so that the relative expression levels of each gene could be compared between each cell type (see methods, 2.2.9.4.1). The box-plots show that the data has not been 'centred' following the application of this normalisation step.

Signal Cutoff	Sensitivity (TPR)	Specificity (FPR)
7.604	0.989	0.051
7.658	0.989	0.048
7.697	0.989	0.044
7.714	0.989	0.037
7.823	0.989	0.034
8.175	0.989	0.031
8.523	0.989	0.027
8.629	0.989	0.024
8.667	0.989	0.020
8.716	0.989	0.017
8.768	0.989	0.014
8.873	0.989	0.010
9.539	0.989	0.007
10.171	0.982	0.007
10.275	0.979	0.007
10.351	0.975	0.007
10.417	0.972	0.007
10.477	0.972	0.003
10.509	0.968	0.003
10.537	0.965	0.003
10.575	0.965	0

Table 4.4. Extract of data used to generate the ROC curve for array number 27.

This is an extract from the table generated by ROC analysis on the positive and negative control signals for array no.27. The ROC curve for this data is shown in Fig. 4.12. Increasing the signal cutoff point leads to a reduction of the FPR, as the test for a positive signal becomes more stringent. As the signal cutoff rises, this eventually leads to a reduction of the TPR, because the test for a positive signal is more stringent. A cutoff point that gave a FPR of 5 % was selected for all microarray datasets; in this case, the cutoff point selected was 7.658.

A.



B.

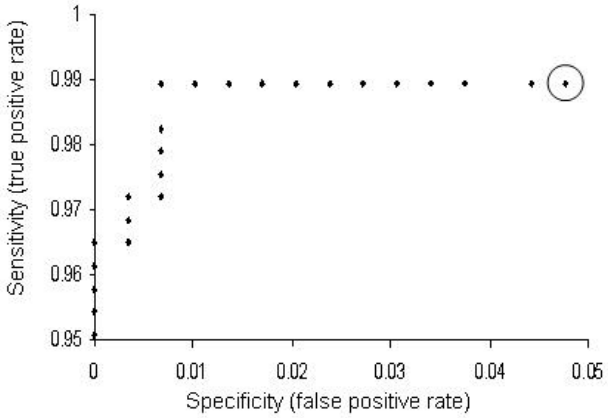


Figure 4.12. ROC curve for array number 27 (true-positive vs false-positive rates) determined for the signal cutoff points shown in Table 4.4.

(A) is a graphical representation of the data in Table 4.4. Increasing the signal cutoff (see Table 4.4) leads to a reduction of the FPR, but as the test becomes more stringent, the TPR also begins to fall. (B) shows a magnified view at the very top left hand corner of ROC curve (A). For the signal cutoff point 7.658 (Table 4.4), the TPR is 98.9 % and the FPR is 4.8 %, as shown by the circled point plotted at the top right of ROC curve (B).

4.2.5 Normalisation

(Fig. 2.1, stage 1, step 4)

Datasets were normalised against a reference array (see methods, 2.2.9.3.3), so that fluorescent signals were scaled against the 75th percentile of the positive control signals. This allowed normalisation across all microarrays irrespective of the cell type. The scaling correction factors (CF) are shown for all microarray datasets are shown in Table 4.3.

4.2.6 Data flagging and preparation for import into GeneSpring

(Fig. 2.1, stage 1, step 5)

The microarray data for individual probes were flagged as present (P) or absent (A), based on the cutoff points selected from the ROC analysis (Table 4.3). Data points falling below the cutoff point (assigned as 0.01) following ROC analysis were defined as absent (A). The data was then saved as tab delimited text files prior to import into GeneSpring.

4.2.7 Microarry analysis using GeneSpring

(Fig. 2.1, stage 2, step 1 to step 3)

The microarray datasets were imported into GeneSpring using the software import wizard, then 'back-transformed' from $\log_2$ to linear data, and normalised as previously described (see methods, 2.2.9.3.3). The principle aim of microarray analysis in GeneSpring was to identify HCMV genes that were differentially regulated in comparisons of infected cell lines over a time course. One-way ANOVA tests between HFFF-2 vs RPE, and HFFF-2 vs U373Mg (see methods, 2.2.9.4.4) were performed to identify statistically significant differences between individual HCMV genes for each cell type. Further quality control analysis was undertaken prior to the application of the combined statistical tests.

4.2.7.1 Clustering analysis

(Fig. 2.1, stage 2, step 3)

Condition trees generated by hierarchical clustering show how gene expression is related between microarray datasets. Comparisons made between individual datasets serves as a useful quality control step since datasets should cluster together according to time PI. Clustering was performed in GeneSpring generating distances based on Spearman's correlation (separation ratio 1, minimum distance 0.001), in which average linkage was applied in order to assemble the clusters into a hierarchical plot. Fig. 4.13 shows the

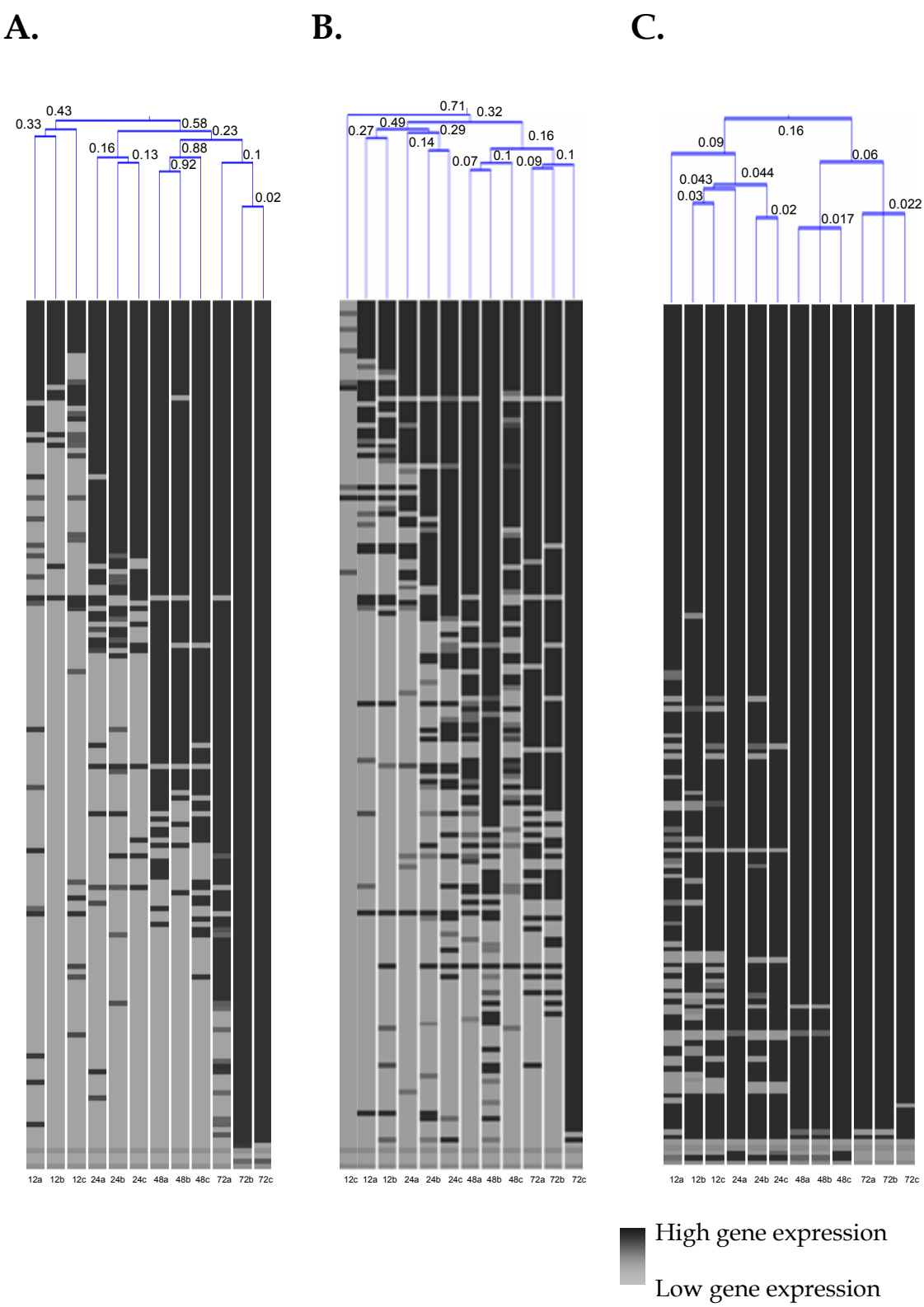


Figure 4.13. Condition trees for the (A) HFFF-2, (B) RPE and (C) U373Mg microarray datasets.

The condition trees show the relationships among the gene expression levels for each dataset, which are clustered accordingly. The time (h PI) is shown along with the biological replicate number (a, b, or c). The distances between the clusters are also displayed on the trees. Note that the distance value is calculated as: distance = 1 - correlation, hence values can range from 0 to 2, because Spearman's correlation reports values from -1 to +1.

condition trees for HFFF-2, RPE and U373Mg cells. The branches within each condition tree have been assigned distance values between the array datasets. Each row on the tree represents the expression level of an individual gene. It should be noted that the GeneSpring normalisation step (per gene: normalise to the median) was removed so that the array datasets could be compared independently. Mock-infected datasets were excluded from this analysis.

Fig. 4.13A shows the condition tree for the infected HFFF-2 cell line. There are 4 major clusters each representing a single time point, indicating good reproducibility between the triplicate biological samples. The 4 clusters are part of two branches, with the 24, 48 and 72 h PI clusters forming one branch, and the 12 h PI cluster forming the second. At 12 h, relatively few genes are expressed, but the number of expressing genes increases with time. At later time points, the triplicate dataset clusters are more conserved (confirmed by decreased distance values at later times PI between clusters), this could be due to either a greater uniformity in the regulation of gene expression with time, or reduced error given a larger P-flagged sample size at later time points.

Fig. 4.13B shows the condition tree for the infected RPE cell line. The branching pattern is more poorly conserved than that obtained for HFFF-2 (Fig. 4.13A). The 12 h dataset, replicate 12(c) forms an individual branch (displayed at the extreme left of the tree). The number of genes expressed in dataset 12(c) is much lower than in datasets 12(a) and 12(b) (which exhibit a similar banding pattern, representing good reproducibility between these two datasets). Dataset 12(c) exhibiting little positive expression data is therefore considered a cluster in its own right; although the expression data is more closely related to that of 12(a) and 12(b), than any other cluster. The reduced dataset in 12(c) may be due to experimental variation at the level of virus infection, microarray hybridisation, or the stringent cutoff points employed for these experiments. The remaining datasets representing the 24, 48 and 72 h PI time points cluster as expected. In the case of RPE cells, the two main arms of the condition tree group the 12 and 24 h PI clusters together on one branch while the 48 h and 72 h PI cluster together on the other branch.

Fig. 4.13C shows the condition tree for the infected U373Mg cell line. At the early time points, 12 and 24 h PI, the replicate datasets appear less conserved; one cluster is formed containing two 12 h and three 24 h PI datasets. As with the infected RPE cell line, increased error rates may be responsible for the branching pattern observed at these early time points. At 48 and 72 h PI, however, two discrete clusters are formed containing the

dataset for each time point, (note the almost perfect reproducibility for the triplicate datasets).

The condition trees generally demonstrate good reproducibility between the triplicate array datasets for each cell type. Where discrepancies appear, they are associated with the early time points (12 and 24 h PI), and may be due to increased error caused by either the limited number of genes expressed early in the virus replication cycle, or to a more coordinated regulation of gene expression at later time points (48 and 72 h PI). In general, the condition trees demonstrate a good and consistent clustering pattern.

4.2.7.2 Analysis based on present and absent flags

(Fig. 2.1, stage 2, step 4)

Following ROC analysis, present (P) or absent (A) flags were manually assigned to all probes in the array datasets (see methods, 2.8.5.2), with signals of 0.01 flagged as (A). As there were 12 datasets per microarray (excluding mock infected cell datasets), a single gene could have present flags in 1 to 12 datasets. Genes that had no present flags (i.e. only absent flags), were described as 'absent'. Table 4.5 lists the number of datasets containing present flags for all Merlin ORFs (AY446895) in each cell type.

For infected HFFF-2 and RPE cells, datasets contain a range of present signals for the genes, and are generally in agreement: this is not the case for the U373Mg data. Present flags are generally found in 10 or more datasets for each HCMV gene in U373Mg cells (each time point having the potential to score present flags in 3 datasets). This implies that the majority of genes are expressed at all time points, including genes that are reported to be expressed with late kinetics in infected HFF cells (Chambers et al., 1999). In HFFF-2 and RPE cells, HCMV genes that are reported to follow early expression kinetics (e.g. UL54, UL55, US18, US22) tend to have present flags in more datasets than those previously reported as late genes (UL99, UL128, UL130, UL144). This reflects the nature of gene expression in HCMV, which has been described as a regulated temporal cascade. It is not clear why the temporal expression of HCMV genes differs in U373Mg cells; it is possible that the temporal cascade is rapidly completed within the first 12 h of lytic infection, or alternatively, there may be a general breakdown in the cascaded regulation of HCMV gene expression U373Mg cells.

Absent genes are those whose expression cannot be detected at any time point in one or more cell types. Identification of such genes provides valuable information, particularly in

Table 4.5. Microarray dataset information for individual Merlin genes in each cell type.

Gene	HFFF-2	RPE	U373Mg
RL1	4	1	12
RL5A	6	6	11
RL6	2	1	4
RL10	5	6	9
RL11	8	9	12
RL12	8	9	12
RL13	9	8	12
UL1	5	3	10
UL2	2	1	7
UL4	12	11	12
UL5	10	10	12
UL6	5	2	10
UL7	4	1	10
UL8	4	1	11
UL9	5	1	11
UL10	5	3	11
UL11	7	5	12
UL13	6	6	12
UL14	2	1	12
UL15A	2	1	11
UL16	11	4	12
UL17	9	7	12
UL18	2	1	11
UL19	2	2	11
UL20	4	4	12
UL21A	6	7	12
UL22A	9	8	12
UL23	2	3	11
UL24	4	4	11
UL25	5	5	12
UL26	3	9	12
UL27	2	2	11
UL28	2	2	12
UL29	0	0	2
UL30	6	5	11
UL31	2	3	10
UL32	6	4	8
UL33	5	6	12
UL34	6	8	12
UL35	4	3	11
UL36	10	10	12
UL37	2	1	7
UL38	9	10	12
UL40	10	9	12
UL41A	7	8	11
UL42	8	7	12
UL43	9	9	12
UL44	10	11	12
UL45	11	9	12
UL46	9	5	12
UL47	3	1	10
UL48	6	6	12
UL48A	0	0	0
UL49	10	11	12
UL50	9	11	12

Gene	HFFF-2	RPE	U373Mg
UL51	2	2	12
UL52	6	6	9
UL53	8	8	12
UL54	11	10	12
UL55	9	9	12
UL56	9	9	12
UL57	5	4	12
UL69	10	7	12
UL70	6	4	12
UL71	3	4	12
UL72	6	3	12
UL73	5	4	12
UL74	3	1	7
UL75	6	5	11
UL76	2	1	10
UL77	3	1	10
UL78	9	8	12
UL79	2	1	7
UL80_UL80.5	6	7	12
UL82	9	8	12
UL83	8	7	12
UL84	6	8	12
UL85	7	6	12
UL86	6	4	9
UL87	2	1	9
UL88	4	8	12
UL89	6	6	12
UL91	7	7	11
UL92	4	5	12
UL93	6	5	12
UL94	6	8	12
UL95	4	5	12
UL96	7	8	12
UL97	9	11	12
UL98	11	11	12
UL99	10	11	12
UL100	6	5	12
UL102	2	1	11
UL103	6	5	12
UL104	6	2	9
UL105	2	3	12
UL111A	0	0	6
UL112	8	10	12
UL114	6	6	12
UL115	9	11	12
UL116	9	8	12
UL117	9	9	12
UL119	7	7	12
UL120	2	1	11
UL121	2	1	7
UL122	6	8	12
UL123	4	4	12
UL124	3	1	10
UL128	3	2	11
UL130	4	1	11

Gene	HFFF-2	RPE	U373Mg
UL131A	3	2	12
UL132	11	11	12
UL148	12	11	12
UL147A	12	11	12
UL147	10	11	12
UL146	3	1	9
UL145	3	2	12
UL144	8	9	12
UL142	4	2	6
UL141	8	7	12
UL140	5	4	12
UL139	5	2	12
UL138	11	10	12
UL136	9	12	12
UL135	11	12	12
UL133	9	9	12
UL148A	2	2	12
UL148B	2	3	11
UL148C	3	3	12
UL148D	6	7	12
UL150	2	1	12
IRS1	12	11	12
US1	0	0	0
US2	2	0	7
US3	2	1	12
US6	2	1	12
US7	3	1	8
US8	7	2	12
US9	4	5	12
US10	8	6	12
US11	11	10	12
US12	12	11	12
US13	12	11	12
US14	12	11	12
US15	11	9	12
US16	9	9	12
US17	3	2	12
US18	12	11	12
US19	12	11	12
US20	12	11	12
US21	4	3	10
US22	10	10	12
US23	9	6	12
US24	4	3	12
US26	2	1	12
US27	8	6	12
US28	9	8	12
US29	6	4	11
US30	7	5	12
US31	6	2	12
US32	7	4	12
US34	3	1	12
US34A	0	0	0
TRS1	4	6	12

Table 4.5. Microarray dataset information for individual Merlin genes in each cell type.
The table shows the number of Presence (P) flags for every Merlin gene in each cell type. As there were 12 datasets (representing 4 time points) for each cell type, the maximum number of Presence (P) flags for a gene in each cell type is 12. A gene with no presence flags is defined as absent.

cases where a gene could be present in one cell type, but absent in another, therefore providing evidence of cell-type specific differential regulation of gene expression. In the case of US2, present flags are found in two datasets in HFFF-2 cells, seven datasets in U373Mg cells, but US2 is absent in RPE cells; it should be noted that the difference in the expression of US2 in HFFF-2 and RPE cells is the only case where a gene is present in HFFF-2 cells, but absent in RPE cells. Gene US2 could be classified as a candidate for differential regulation. The two datasets containing present flags for US2 expression in HFFF-2 cells were obtained for the 72 h PI samples only, making it less likely that US2 is differentially expressed in RPE compared to HFFF-2 cells. It is probable that the US2 gene is expressed at low levels in both RPE and HFFF-2 cells. When comparing profiles from HFFF-2 and U373Mg infected cells, gene US2 is over-expressed in U373Mg compared to HFFF-2; however, this conclusion was not supported by the statistical tests (see Table 4.6). There are additional examples of genes which contain present flags in 1 dataset in RPE cells, but in 3 or 4 datasets in HFFF-2 cells (e.g. RL1, UL7, UL74). As with gene US2, examination of the expression profiles for RL1, UL7 and UL74 fails to provide strong evidence for differential regulation between HFFF-2 and RPE infected cells, but does appear to support differential expression between HFFF-2 and U373Mg cells; however, as before, these genes were not identified as differentially expressed by the statistical tests.

There are two examples where genes UL29 and UL111A are flagged as absent in both HFFF-2 and RPE cells, but present in U373Mg cells. In the case of UL29, there are two datasets containing present flags in U373Mg cells. However, comparison of the expression profiles for UL29 in HFFF-2 and U373Mg cells does not provide good evidence of differential regulation of UL29 expression; for the same reasons previously described for US2. There are six datasets containing present flags for gene UL111A in U373Mg cells, representing expression from at least 2 time points. Indeed, the statistical tests between HFFF-2 and U373Mg cells confirms expression of this gene to be significantly different (Table 4.6). Expression of genes such as UL111A in two or more time points (representing ≥ 4 datasets with present flags) against one or no time points provides better evidence for differential regulation.

Finally, there were three genes (UL48A, US1 and US34A) where no expression was detected at all in any of the cell types. No present flags are assigned to any dataset in any cell type. Possibly these genes are simply not expressed in any of the cell types tested here, or are expressed at very low levels and fail the statistical tests used in this analysis. Alternatively, there could be a problem with the oligonucleotide probes for these genes (e.g. poor thermodynamic performance), resulting in failure to bind under the

experimental conditions used. In order to resolve which of the above explanations is correct would require further application of techniques such as real-time PCR and/or northern blotting studies.

4.2.8 Expression profiles for representative non-differentially expressed immediate-early, early, and late regulated HCMV genes in HFFF-2 and RPE cells

Before dealing with HCMV genes that are differentially regulated, it is appropriate to first examine expression kinetic profiles obtained for non-differentially regulated genes representing IE, E and L classes, as determined by the combined statistical tests (Table 4.6). The expression profiles for TRS1 (IE), US22 (E), UL119 (E-L), UL25 (L), UL32 (L), UL54 (E), UL75 (E-L) and US10 (E) are shown in Fig. 4.14. No significant difference in the expression levels in HFFF-2 and RPE cells were found for these genes in the combined statistical tests, and examination of the expression profiles (Fig. 4.14) confirmed the statistical test p-values (Table 4.6). Biphasic expression kinetics were obtained for TRS1 (Fig. 4.14A), US22 (Fig. 4.14B) and US10 (Fig. 4.14D). Expression kinetics of TRS1 andUS22 increased between 12 to 24 h PI, plateaued between 24 and 48 h PI, then increased again from 48 to 72 h PI; however, there was an overall decrease in the expression kinetics of US10 between 12 to 48 h PI, followed by an increase from 48 to 72 h PI. Expression of HCMV gene UL119 (Fig. 4.14C) and UL53 (Fig. 4.14F) accumulated continuously from 12 through to 72 h PI, whereas expression of HCMV genes UL25, UL32 and UL75 all increased from 24 to 72 h PI in both HFFF-2 and RPE cells. The microarray data are in accord with the protein expression data (see results I, section 3.3) that showed that the kinetics of protein expression were the same in infected HFFF-2 and RPE cells. Correlations between reduced transcript and protein levels in RPE cells compared with HFFF-2 cells are observed between these two cell types.

4.2.9 Differential gene expression

(Fig. 2.1, stage 2, step 5)

In order to identify HCMV genes whose expression was differentially regulated, mean expression levels were compared using combined statistical tests performed independently between HFFF-2 vs RPE, and HFFF-2 vs U373Mg (see methods, 2.8.5.4). The null hypothesis (no difference in the gene expression intensities of individual HCMV genes in HFFF-2 and RPE or U373Mg cells) was tested, with significant differences identified at the 95 % significance level. The Benjamini and Hochberg false discovery rate (see methods, 2.2.9.4.4) was applied in order to reduce the error incurred by multiple

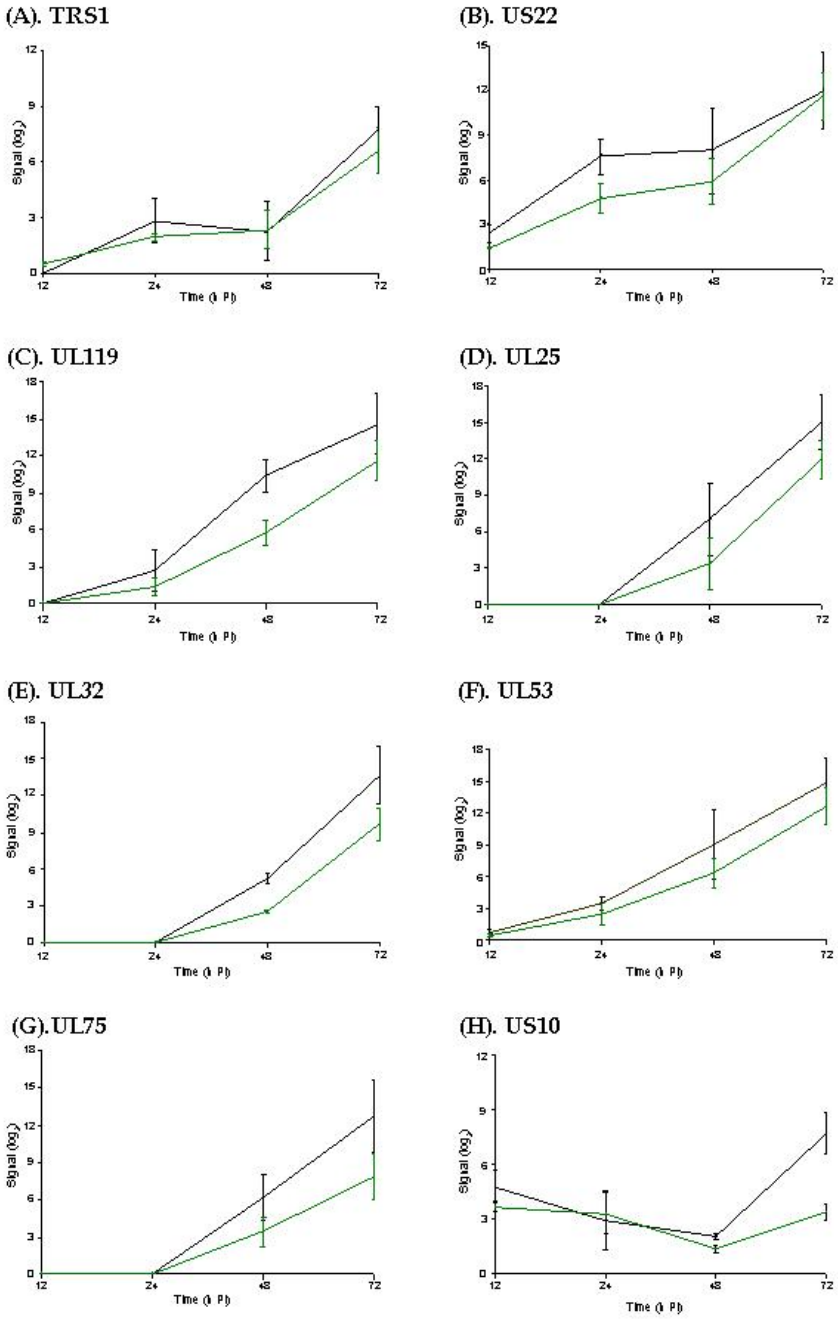


Figure 4.14. Expression profiles for representative non-differentially regulated genes in HFFF-2 and RPE cells.

The profiles show normalised signal (log_2) vs time h PI.

testing; note that % of genes reported as differentially expressed will have passed this restriction by chance. The statistical tests compared the overall level of transcript abundance across all time points without specific reference to the transcript expression kinetics. Because of this, there were genes whose overall expression was not determined to be significantly different when assessed by the statistical tests between HFFF-2 vs RPE, and HFFF-2 vs U373Mg, but exhibited expression kinetics that were different. Conversely, there were genes whose overall expression was determined to significantly different by the statistical tests, but did not exhibit differential expression kinetics (see discussion, section 4.3). Subsequently, differential gene expression was based on both the combined statistical test results and the assessment of the expression kinetics. All gene expression profiles were examined following the application of the combined statistical tests in order to compare the p-values with the associated gene expression kinetics. Note that these comparisons were also made for genes not determined to be differentially expressed according to the statistical tests.

Table 4.6 lists the p-values for all Merlin ORFs determined by the statistical tests between HFFF-2 and RPE, or HFFF-2 and U373Mg. Genes that are differentially expressed in HFFF-2 and RPE or U373Mg are shown in Tables 4.7 and 4.8 respectively. Genes arranged in genome order are listed with the corresponding p-value, the temporal kinetic class of gene expression (Chambers et al., 1999) and the gene function where known. The corresponding expression profiles for each gene in Table 4.7 (HFFF-2 vs RPE) are shown in Fig. 4.15. Note that significant differences in the expression of HCMV genes UL4, IRS1, US12, US14, US18, US19 and US20 were identified in the comparison between both HFFF-2 and RPE (Table 4.7) and the comparison between HFFF-2 and U373Mg (Table 4.8).

Expression of UL4 was detected from 12 h PI in all three cell types. The magnitude of expression of UL4 in HFFF-2 cells was greater than in U373Mg (from 24 to 72 h PI), and significantly higher than in RPE cells (from 12 to 48 h PI); though there is a steep incline in the expression of UL4 between 48 and 72 h PI in RPE cells. The combined statistical tests showed UL4 expression to be significantly different in each of the three cell types (Table 4.6). The UL4 expression profile for HFFF-2 and RPE was in accord with this p-value, as there is a clear difference in the expression profile in magnitude and the temporal kinetics of expression. When UL4 expression in HFFF-2 and U373Mg were compared, the overall level of expression was lower in U373Mg, but the expression profile appeared similar. Comparison of the corrected p-values for UL4 shown in Table 4.6, reported p-values of 0.00555 (HFFF-2/RPE) and 0.0126 (HFFF-2/U373Mg), which reflects the larger difference in the overall expression kinetics of UL4 in HFFF-2 and RPE cells, compared to HFFF-2

Table 4.6. Student's t-test p-values for HFFF-2 and RPE or U373Mg

Gene	Kinetic class	HFFF-2 vs RPE (p-value)	HFFF-2 vs U373Mg (p-value)	Gene	Kinetic class	HFFF-2 vs RPE (p-value)	HFFF-2 vs U373Mg (p-value)
RL1		0.342	0.109	UL52	L	0.61	0.987
RL5A		0.569	0.856	UL53	E	0.371	0.746
RL6	L	1	1	UL54	E	0.0342	0.0559
RL10	E-L	0.717	0.746	UL55	E	0.253	0.626
RL11	L	0.347	0.907	UL56	E	0.359	0.869
RL12	E-L	0.347	0.985	UL57	E	0.556	0.201
RL13	E-L	0.213	0.681	UL69	E-L	0.112	0.437
UL1	E-L	0.468	0.505	UL70	E-L	0.468	0.582
UL2	L	0.617	0.286	UL71		0.931	0.0734
UL4	E	0.00555	0.0126	UL72	E-L	0.4	0.626
UL5	E	0.187	0.303	UL73	E-L	0.569	0.238
UL6		0.45	0.296	UL74		0.47	0.626
UL7	L	0.347	0.312	UL75	E-L	0.527	0.857
UL8		0.4	0.136	UL76		0.639	0.0666
UL9	L	0.253	0.312	UL77	E	0.495	0.173
UL10		0.468	0.221	UL78	E	0.253	0.892
UL11	E	0.341	0.745	UL79		0.57	0.461
UL13	E	0.556	0.312	UL80_UL80.5	L	0.686	0.684
UL14	L	0.624	0.0126	UL82	L	0.257	0.519
UL15A		0.569	0.0126	UL83	L	0.347	0.746
UL16	E	0.00257	0.27	UL84	E-L	0.703	0.665
UL17	E	0.213	0.815	UL85	E-L	0.451	0.977
UL18	L	0.61	0.0269	UL86	E-L	0.47	0.977
UL19		0.746	0.0126	UL87		0.61	0.156
UL20		0.818	0.0594	UL88		0.717	0.175
UL21A		0.746	0.175	UL89	E-L	0.569	0.672
UL22A		0.255	0.523	UL91	L	0.556	0.907
UL23		0.974	0.0217	UL92	L	0.987	0.137
UL24		0.737	0.248	UL93	L	0.548	0.505
UL25	L	0.687	0.32	UL94	L	0.717	0.665
UL26	E	0.468	0.0594	UL95	E-L	0.971	0.136
UL27	E	0.82	0.0159	UL96	E-L	0.624	0.746
UL28		0.81	0.0126	UL97	E-L	0.468	0.752
UL29	L	1	0.29	UL98	E-L	0.0874	0.1
UL30		0.564	0.815	UL99	E-L	0.193	0.201
UL31	L	0.819	0.0126	UL100	E-L	0.556	0.643
UL32	L	0.459	0.896	UL102	L	0.671	0.0184
UL33	E	0.818	0.286	UL103	L	0.52	0.626
UL34	E-L	0.718	0.505	UL104	E	0.253	0.952
UL35	E	0.569	0.148	UL105	E	0.987	0.0126
UL36	E (IE)	0.253	0.907	UL111A	E-L	1	0.0326
UL37	IE	0.569	0.505	UL112	E	0.569	0.746
UL38	IE	0.296	0.76	UL114	E	0.569	0.643
UL40	E-L	0.15	0.263	UL115	L	0.385	0.746
UL41A		0.556	0.76	UL116	E-L	0.284	0.922
UL42		0.4	0.922	UL117	L	0.257	0.935
UL43	L	0.347	0.733	UL119	E	0.4	0.815
UL44	E-L	0.213	0.286	UL120	L	0.687	0.0126
UL45		0.0414	0.116	UL121	L	0.569	0.505
UL46	E-L	0.213	0.62	UL122	IE, L	0.671	0.684
UL47	E-L	0.48	0.224	UL123	IE	0.150	0.443
UL48	L	0.569	0.746	UL124	E	0.468	0.204
UL48A		1	1	UL128	E	0.737	0.092
UL49	E-L	0.213	0.29	UL130	E-L	0.347	0.211
UL50		0.569	0.94	UL131A	L	0.717	0.0621
UL51		0.987	0.0126	UL132	E-L	0.103	0.227

Gene	Kinetic class	HFFF-2 vs RPE (p-value)	HFFF-2 vs U373Mg (p-value)
UL148		0.0342	0.0615
UL147A		0.0178	0.0594
UL147	E-L	0.341	0.461
UL146		0.434	0.746
UL145		0.61	0.0376
UL144		0.411	0.952
UL142		0.468	0.952
UL141		0.45	0.952
UL140		0.556	0.397
UL139		0.254	0.417
UL138		0.0342	0.116
UL136		0.556	0.907
UL135		0.193	0.29
UL133		0.4	0.705
UL148A		0.569	0.0195
UL148B		1	0.0445
UL148C		0.974	0.0594
UL148D		0.671	0.505
UL150		0.573	0.109
IRS1	IE	0.00555	0.0126
US1		1	1
US2		0.347	0.189
US3	IE	0.556	0.0126
US6	E-L	0.61	0.0126
US7	E-L	0.501	0.355
US8	E	0.213	0.519
US9	E	0.987	0.0594
US10	E	0.253	0.907
US11	E	0.0342	0.0594
US12	E	0.00988	0.0126
US13	E	0.0342	0.0666
US14	E	0.0322	0.037
US15	E	0.037	0.204
US16	E	0.342	0.856
US17	E	0.556	0.125
US18	E	0.00555	0.0126
US19	E	0.00446	0.0126
US20	E	0.00988	0.0457
US21	E	0.527	0.296
US22	E	0.213	0.505
US23	E	0.213	0.746
US24	E	0.61	0.175
US26	E	0.569	0.0126
US27	E	0.342	0.94
US28	E	0.197	0.439
US29	E-L	0.433	0.746
US30	E	0.434	0.907
US31		0.213	0.443
US32	L	0.347	0.856
US34	E	0.434	0.0594
US34A		1	1
TRS1	IE	0.931	0.15

Table 4.6. Student's t-test p-values for HFFF-2 and RPE or U373Mg.
The table shows the corrected p-values for the Student's t-test between HFFF-2 and RPE, and HFFF-2 and U373Mg cells. Significant differences were considered at the 95 % significance level ($p < 0.05$).

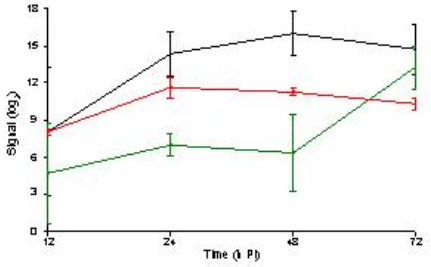
Gene	HFFF-2 vs RPE (p-value)	Kinetic Class	Function
UL4	0.00555	E	Glycoprotein, function unkown; Upstream cis acting element for NF-Y and Elk1. Upstream short ORF controls translation of the downstream mRNA.
UL16	0.00342	E	Evasion of immune surveillance. Expressed at early times and peaks at late times.
UL45	0.0414	L	Tegument protein; reported anti-apoptotic activity and role in cell-to-cell spread of infection.
UL148	0.0342		
IRS1	0.00555	IE	Important transactivator of gene expression. Efficient assembly of virus particles. Binds double stranded RNA and inhibits cellular interferon response. Internal transcription start site, produces a 263 aa protein that negatively regulates IRS1 expression. US22 family
US11	0.0342	E	Evasion of immune surveillance. Glycoprotein that targets MHC Class I heavy chains for degradation.
US12	0.00988	E	Putative transmembrane protein
US13	0.0342	E	Putative transmembrane protein
US14	0.0322	E	Putative transmembrane protein
US15	0.037	E	Putative transmembrane protein
US18	0.00555	E	Putative transmembrane protein
US19	0.00446	E	Putative transmembrane protein
US20	0.00988	E	Putative transmembrane protein

Table 4.7. HCMV genes that are differentially expressed in HFFF-2 and RPE cells.

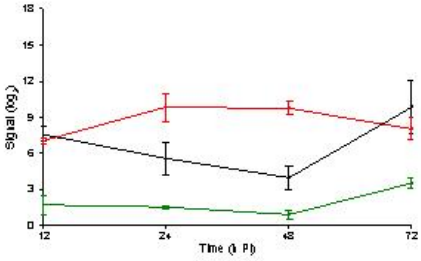
Gene	HFFF-2 vs U373Mg (p-value)	Kinetic Class (Chambers et al., 1999)	Function
UL4	0.0126	E	Glycoprotein, function unkown; Upstream cis acting element for NF-Y and Elk1. Upstream short ORF controls translation of the downstream mRNA
UL14	0.0126	L	UL14 family; putative membrane glycoprotein
UL15A	0.0126		Putative membrane protein
UL18	0.0269	L	gpUL18, similar to MHC class I; immune evasion
UL19	0.0126		
UL23	0.0217		US22 family member
UL27	0.0159	E	
UL28	0.0126		US22 family member
UL31	0.0126	L	
UL51	0.0126		DNA packaging protein
UL102	0.0184	L	Helicase-primase
UL105	0.0126	E	Helicase-primase
UL111A	0.0326	E-L	vIL-10
UL120	0.0126	L	Glycoprotein
UL145	0.0376		
UL148A	0.0195		Membrane protein
UL148B	0.0445		Membrane protein
IRS1	0.0126	IE	Important transactivator of gene expression. Efficient assembly of virus particles. Binds double stranded RNA and inhibits cellular interferon response. Internal transcription start site, produces a 263 aa protein that negatively regulates IRS1 expression
US3	0.0126	IE	Glycoprotein, inhibits processing of MHC class I and II; immune evasion
US6	0.0126	E-L	Glycoprotein, inhibits TAP mediated peptide transport; immune evasion
US12	0.0126	E	Putative transmembrane protein
US14	0.037	E	Putative transmembrane protein
US18	0.0126	E	Putative transmembrane protein
US19	0.0126	E	Putative transmembrane protein
US20	0.0457	E	Putative transmembrane protein
US26	0.0126	E	Putative transmembrane protein

Table 4.8. HCMV genes that are differentially expressed in HFFF-2 and U373Mg.

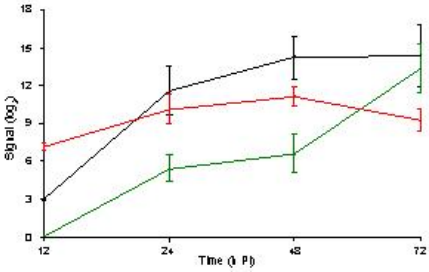
(A). UL4



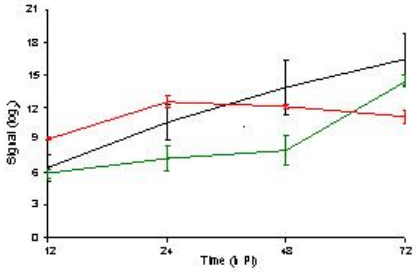
(B). UL16



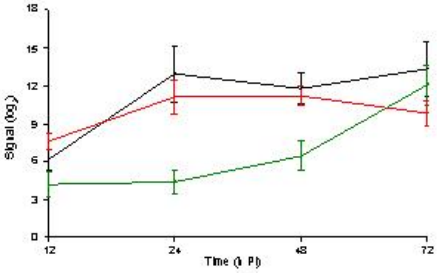
(C). UL45



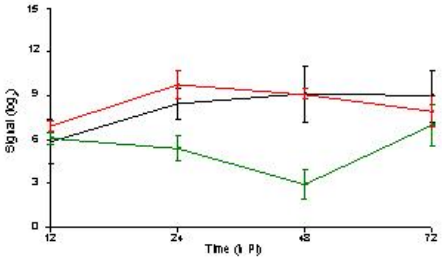
(D). UL148



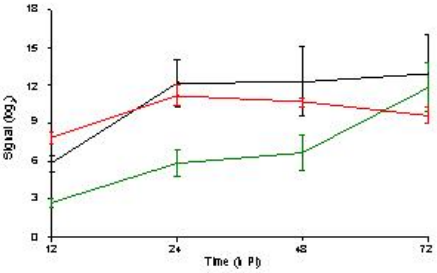
(E). IRS1



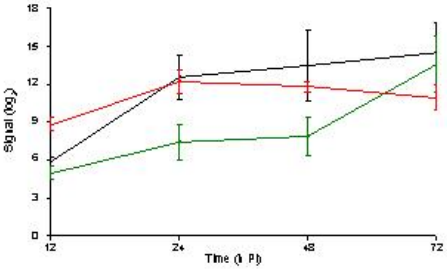
(F). US11



(G). US12



(H). US13



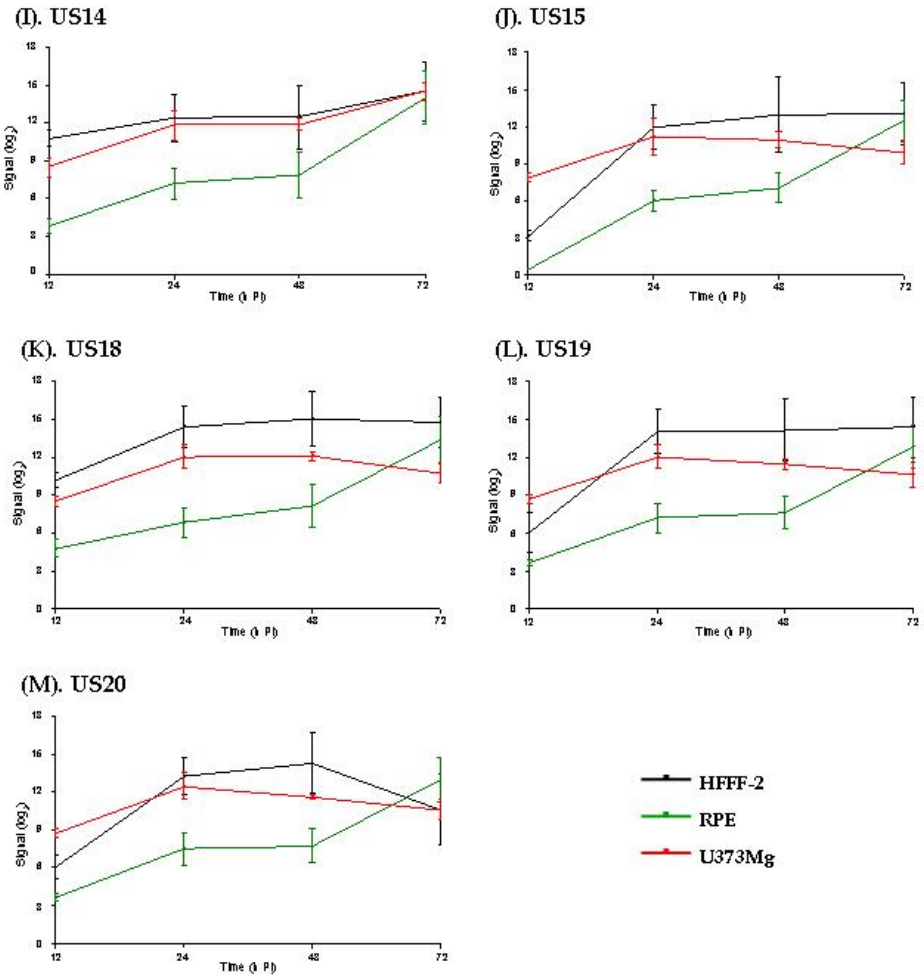


Figure 4.15. Microarray expression curves for genes listed in table 4.7 as differentially expressed in HFFF-2 and RPE cells.

The curves shows normalised signal (log_2) vs time (h PI).

and U373Mg cells. Indeed, this difference is approximately 2 fold between HFFF-2 and RPE, and 1.5 fold for U373Mg. It seems probable that expression of gene UL4 is up-regulated in HFFF-2 compared to RPE and U373Mg cells, as there was very high expression throughout the time course, and statistical analysis indicates that the difference in the expression for UL4 in HFFF-2 differs significantly from that in both RPE and U373Mg cells. Moreover, magnitude and expression kinetics obtained for UL4 in U373Mg cells is very similar to all other HCMV genes expressed in this cell type (Fig. 4.15, 4.16).

The expression kinetics for UL16 in HFFF-2 and RPE yields a concave profile albeit the profile is much flatter in RPE cells; in contrast, the expression profile in U373Mg was convex. UL16 is expressed at all time points in each cell type, with medium levels of expression in HFFF-2 and U373Mg cells, but much lower expression in RPE cells. The statistical test results show a significant difference in mean expression between HFFF-2 and RPE (p-value 0.00257), but no significant difference in expression between HFFF-2 and U373Mg. In RPE cells, expression of UL16 was extremely low (3-fold lower than in HFFF-2), suggesting possible down-regulation of gene UL16 in this cell type. While no statistically significant difference could be found in UL16 expression between HFFF-2 and U373Mg cells, the kinetics of expression were different, although the significance of that finding is not clear.

The expression kinetics of genes UL45, UL148 and IRS1 were similar in all three cell types. In HFFF-2, UL45 and UL148 transcripts increased steadily with time, while IRS1 transcript levels reached a plateau at 24 h PI. In RPE cells, UL148 and IRS1 expression was detected at 12 h PI with little change in transcript levels until after 48 h PI. UL45 expression was not detected in RPE cells until 24 h PI, thereafter transcript levels were of a similar magnitude and shared similar kinetics to those of IRS1. The expression profiles for UL45, UL148 and IRS1 in U373Mg were similar in kinetics and magnitude over the time course. No significant difference was found for UL45 and UL148 in the combined statistical tests comparing HFFF-2 and U373Mg cells.

US11 was detected in HFFF-2 and RPE cells from 12 h PI, after which the expression increased in HFFF-2 cells and plateaued at 48 h PI. However, the expression of US11 decreased in RPE cells until 48 h PI, and increased at 72 h PI. The US11 expression profile in U373Mg cells is characteristic of all other genes in that cell type (medium expression levels accumulating at a similar rate over the time course). The statistical tests

demonstrated a significant difference in US11 expression between HFFF-2 and RPE cells, but no significant difference between HFFF-2 and U373Mg cells.

US12, US13, US14, US15, US18, US19 and US20 are all members of the US12 gene family and all exhibited similar expression profiles in each of the three cell types. Since US12, US13, US14 and US15 are contained in one 3′ co-terminal group, and genes US18, US19 and US20 in a second 3′ co-terminal group, this may account for the shared expression kinetics among the family members US12 to US15, and US18 to US20. Apparent expression from all of these genes was detected throughout the time course. In HFFF-2, expression of all of these genes was fairly high and plateaued at about 24 h PI. In contrast, in RPE cells, the expression kinetics for these genes yielded a biphasic profile; an increase from 12 to 24 h PI, followed by a plateau between 24 and 48 h PI, and a sharp rise in expression between 48 and 72 h PI. The expression kinetics for each gene was different in HFFF-2 and RPE cells, and this correlated with a finding of significant difference in mean expression in the combined statistical tests. However, since these genes belong to 3′ co-terminal groups, further analysis such as northern blotting is required to confirm such conclusions.

In infected U373Mg cells, the profiles of HCMV genes US12, US13, US14, US15, US18, US19 exhibit the same magnitude and kinetics that characterise other HCMV genes. Expression levels plateaued after 24 h PI and generally declined at 72 h PI. Interestingly, the statistical tests determined that expression of genes US12, US14, US18, US19 and US20 as significantly different in HFFF-2 and U373Mg cells. In the case of US18, US19 and US20, the expression profiles obtained in HFFF-2 and U373Mg were similar, albeit slightly higher in HFFF-2 cells. According to the t-test p-values, there was a significant difference in the overall mean expression for US18, US19 and US20 between HFFF-2 and U373Mg cells; however, examination of the expression kinetics for these genes in these two cell types does not appear to show strong evidence in support of differential expression. Further analysis would be required to clarify such findings.

The expression curves for representative HCMV genes from Table 4.8 whose expression levels differed between infected HFFF-2 and U373Mg cells, but not between infected HFFF-2 and RPE cells are shown in Fig. 4.16. In infected HFFF-2 and RPE cells, the expression of HCMV genes UL18 (Fig. 4.16A), UL23 (Fig. 4.16B), US6 (Fig. 4.16C), UL102 (Fig. 4.16D), UL27 (Fig. 4.16E), UL31 (Fig. 4.16F), UL51 (Fig. 4.16G) and UL120 (Fig. 4.16H) are similar as they were generally only detected between 48 and 72 h PI, and no significant differences were reported by the combined statistical tests (Table 4.6).

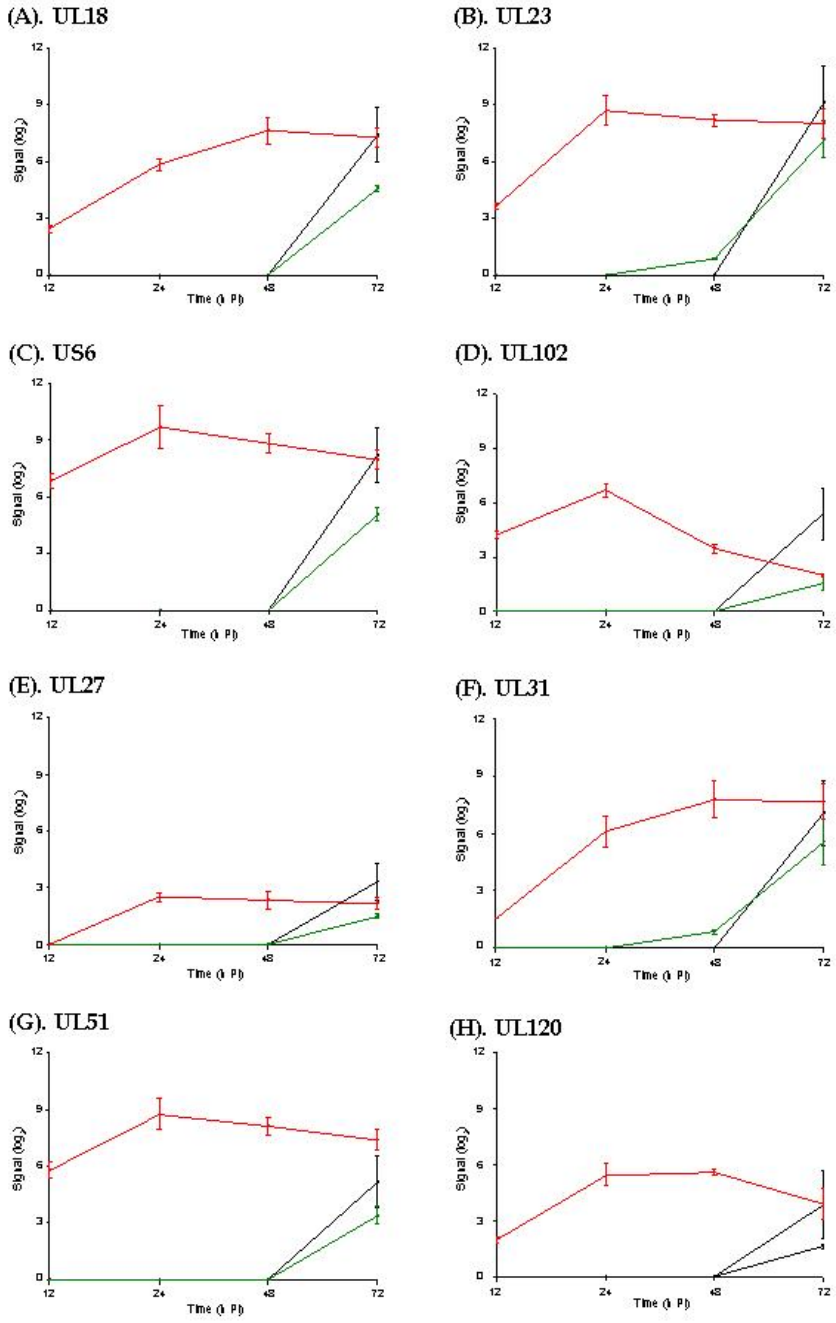


Figure 4.16. Microarray expression curves for representative genes listed in table 4.8 as differentially expressed in HFFF-2 and U373Mg cells.

The curves show normalised signal (log_2) vs time (h PI).

However, the differences in the expression kinetics of these genes in infected HFFF-2 and U373Mg cells confirms the result obtained by the combined statistical tests (Table 4.8).

4.2.10 Analysis of proposed novel ORFs and previously discounted ORFs

The microarray was designed to contain probes for 36 novel ORFs proposed by Murphy et al., 2003(a) and (b), in order to investigate whether the putative ORFs expressed detectable levels of transcripts. Tables 4.9 and 4.10 shows the microarray results obtained for these novel ORFs. It should be noted that only infected HFFF-2 and RPE cells were investigated to determine whether these ORFs were present (P) or absent (A) (disagreement between the cell types are indicated by an asterix). The proposed novel ORFs were defined as present (P) if both the HFFF-2 and RPE cells contained present flags in 3 or more datasets; this being the minimum present data required for expression at one time point. ORFs having present flags in only 1 or 2 datasets were considered unreliable.

Table 4.9 shows the data for the putative ORFs reported by Murphy et al., (2003a). No transcripts were detected for the majority of ORFs. However, signals were detected in 3 or more datasets for ORF4 and ORF5 in both HFFF-2 and RPE cells. Both ORF4 and ORF5 lie between UL54 and UL57, though coded on the opposite DNA strand. It is therefore possible that ORF4 and ORF5 code novel HCMV genes. ORF2 is defined as present in HFFF-2 cells, but absent in RPE cells. ORF2 codes in the same direction as, and overlaps with genes UL29 and UL30. The present flags for this gene in HFFF-2 cells could be due to a single transcript running through the two established HCMV genes. Supporting this conclusion is the observation that the only poly(A) signal in this region is located downstream of UL30.

Table 4.10 shows the data for the putative ORFs reported by Murphy et al., (2003b). C-ORF1, C-ORF3 and C-ORF4 are located within the RL region of the genome. This region of the genome was previously highly annotated (Chee et al., 1990) but more recently, most ORFs in the RL region were discounted as protein coding ORFs by Dolan et al. (2002). The three ORFs (C-ORF1, C-ORF3 and C-ORF4) all code in the same direction as adjacent established genuine ORFs, and long transcripts specifying the RL11 gene family members in this region of the genome could account for these positive signals. The novel C-ORF10, C-ORF15, C-ORF25 and C-ORF29 also code in the same direction as established genes (overlapping in some cases). Again, it is not possible to determine whether the positive signals hybridising to probes for these novel ORFs provide evidence for specific transcription or whether they are artefacts resulting from transcription of established genes within these regions. In contrast, C-ORF16, C-ORF18 and C-ORF26 all code in the

Table 4.9

ORF	P/A
ORF1	A
ORF2	P*
ORF4	P
ORF5	P
ORF6	A
ORF7	A
ORF8	A
ORF9	A
ORF10	A
ORF11	A
ORF12	P

Table 4.10

ORF	P/A
C-ORF1	P*
C-ORF2	A
C-ORF3	P
C-ORF4	P
C-ORF5	A
C-ORF6	A
C-ORF7	P*
C-ORF8	A
C-ORF10	P*
C-ORF11	A
C-ORF13	A
C-ORF14	A
C-ORF15	P
C-ORF16	P
C-ORF17	A
C-ORF18	P
C-ORF20	P*
C-ORF21	A
C-ORF22	A
C-ORF23	A
C-ORF24	A
C-ORF25	P
C-ORF26	P*
C-ORF28	A
C-ORF29	P*

Table 4.9. Detection of transcripts in HFFF-2 and RPE cells hybridising to probes specific for the novel HCMV ORFs proposed by Murphy et al., 2003(a).

Table 4.10. Detection of transcripts in HFFF-2 and RPE cells hybridising to probes specific for the novel HCMV ORFs proposed by Murphy et al., 2003(b).

(P) Present, (A) Absent, and (*) shows disagreement between the two cell types.

opposite direction to established HCMV genes in the locality and it is possible that the present signals obtained for these recently novel ORFs are the result of transcription from the proposed genes. Further work is necessary to confirm the findings reported here.

Probes for a subset of ORFs previously annotated by Chee et al. (1990) but discounted as protein coding genes by Dolan et al. (2002) were also included in the microarray (Table 4.11). Many of these ORFs were originally considered to belong to the RL11 gene family, while others are located within regions of the genome now thought to be non-protein coding. The microarray data confirmed the expectations of Dolan et al. (2002) that the majority of these ORFs do not express transcripts. For these discounted ORFs, that are flagged as present (e.g. UL41), the microarray data by itself is insufficient to determine whether the gene is protein-coding gene. As with all the present-flagged novel ORFs, additional techniques such as northern blotting are required to confirm their status as genes.

Gene	P/A
RL3	A
RL4	P*
RL5	A
RL8	P
RL9	A
UL21	A
UL41	P
UL60	A
UL61	P*
UL101	A
US5	A
US25	A
J1S	A

Table 4.11. Detection of transcripts in HFFF-2 and RPE cells hybridising to probes specific for the HCMV ORFs discounted by Dolan et al., 2002.
ORFs are reported as present (P) or absent (A); disagreement between HFFF-2 and RPE are marked with an asterisk.

4.3 Discussion

The aim of the microarray analysis was to investigate HCMV global gene expression in HFFF-2, RPE and U373Mg cells in order to identify genes whose expression differed from that in infected HFFF-2 cells, which was used as the reference cell line. In addition, transcription from recently proposed novel genes and from regions of the genome now discounted as protein coding was also assessed in HFFF-2 and RPE cells. In HFFF-2 and RPE cells, the microarray detected expression from approximately 96 % of the HCMV strain Merlin ORFs (AY446894), and this increased to approximately 98 % in U373Mg cells. No transcripts binding probes for the HCMV genes US1, US34A or UL48A could be detected at any time point in any cell type, and since the array contained both 5'- and 3'-probes, these were subsequently defined as absent genes. For UL111A and UL29, no transcripts were detected in HFFF-2 or RPE cells although expression of both genes was detected in U373Mg cells. There were also many examples of genes expressed throughout the time course in U373Mg cells, but expressed only at 1 time point (generally at 72 h PI) in HFFF-2 and RPE cells (e.g. UL146, UL150, US3, US6, US7, US26, UL131A). For these genes, and others in Table 4.5, present flags could only be detected in 1 or 2 datasets in HFFF-2 and RPE cells, but 11 or 12 datasets in U373Mg cells.

There are several possible explanations for the above findings. It is possible that genes defined as absent (US1, US34A and UL48A) were not expressed at the particular time points examined, or are expressed at levels below the detection limit, or the microarray probes do not bind their targets efficiently. For genes UL29 and UL111A, whose transcripts were not detected in HFFF-2 or RPE cells, but were detected in U373Mg cells, it is possible that transcript levels were low in HFFF-2 and RPE cells, falling below the detection limit of the microarray, but that their levels were greater in U373Mg cells; present flags were detected for UL29 in two datasets, and for UL111A in six datasets. While the statistical tests determined UL111A to be differentially expressed between HFFF-2 and U373Mg cells, the data for UL29 were insufficient to determine presence or absence by the statistical tests. In some cases, genes UL14, UL18, UL102, UL120, UL131A, where expression was detected in only one or two datasets in HFFF-2 and RPE cells, the genes belong to the late kinetic class (Chambers et al., 1999), and their expression was not detected until 72 h PI. The statistical tests between infected HFFF-2 and U373Mg cells identified these five genes as differentially expressed. The determination of differential gene expression is based firstly upon statistical analysis and secondly by examination of the kinetics of candidate differentially expressed genes identified by the combined statistical tests. The t-tests did not identify genes UL146, UL150, US3, US6, US7 and US26

as differentially expressed in infected HFFF-2 and U373Mg cells, but examination of the expression kinetics showed that these genes were expressed throughout the time course in U373Mg cells, but only at 72 h PI in HFFF-2 cells (data not shown). In contrast, genes UL54, UL147A and UL138 were determined to be significantly different ($p < 0.05$) in HFFF-2 and RPE cells, but examination of the expression profiles revealed that these genes shared the same kinetics in these two cell types (data not shown). Further studies are required to confirm or reject the statistical tests.

Statistically significant differences were identified with the various statistical tests for a total number of 13 genes in HFFF-2 and RPE cells, and 26 genes in HFFF-2 and U373Mg cells. Microarray experiments are generally based on a null hypothesis that there is no difference in the expression levels of individual genes under two (or more) conditions. For the HCMV microarray, the null hypothesis was: no difference in HCMV gene expression kinetics in HFFF-2, and RPE, or HFFF-2 and U373Mg cells. While this is true for infected HFFF-2 and RPE cells, the situation was less clear when comparing infected HFFF-2 and U373Mg cells since HCMV genes belonging to the late kinetic class were expressed at much earlier times in U373Mg cells than in HFFF-2. Consequently, the conclusions drawn from the comparison of HFFF-2 and U373Mg cells are less stringent than those drawn from a comparison of the data from HFFF-2 and RPE cells, where the temporal kinetics of viral gene expression are the same. It would be expected then that many more genes would be classified as differentially expressed when comparing HFFF-2 and U373Mg cells. Because of the differences in temporal expression of some HCMV genes and because individual HCMV genes were expressed at a fairly consistent level irrespective of sampling time PI, the combined statistical test data obtained for the HFFF-2/U373Mg experiment are less reliable than the data obtained for the HFFF-2/RPE comparisons. The data for gene expression in U373Mg cells then needs confirmation by other techniques.

The combined statistical tests between HFFF-2 and RPE cells provided evidence for 13 genes as differentially expressed. The expression kinetics were the same for each of the remaining genes in both cell types, although transcript and proteins levels were consistently lower in RPE cells compared to HFFF-2 cells. For each of the 13 differentially expressed genes, expression was down-regulated in RPE cells compared to HFFF-2 cells over the first 48 h PI, but increased thereafter, so that levels of UL4, UL45, UL148, IRS1, US11, US12, US13, US14, US15, US18, US19 and US20 approximated levels attained in infected HFFF-2 cells at 72 h PI. In contrast, the 72 h PI levels of UL16 never approached levels seen in HFFF-2 cells. The UL4 expression profile exhibits high expression throughout the time course in HFFF-2 cells (Fig. 4.15A). In addition to the reduction in

overall transcript level, expression kinetics suggests a delay in the induction of UL4, UL45, UL148, IRS1, US11, US12, US13, US14, US15, US18, US19 and US20 genes in infected RPE cells compared to HFFF-2 cells. The delay is gene specific and not related to progress through the temporal transcription cascade since other genes belonging to the early and late kinetic classes are expressed with the appropriate temporal kinetics and their overall expression level was not identified as significantly different by the statistical tests in HFFF-2 and RPE cells (Fig. 4.13). Expression of UL16 is down-regulated to the greatest degree in RPE cells, giving a 3 fold reduction compared to HFFF-2 cells. However, undue emphasis should not be placed on fold change in general as HCMV gene expression of non-differentially expressed genes in RPE cells was up to 2-fold lower than in HFFF-2 cells. In contrast, the statistical tests provide a more sensitive determination since they take into account the spread of the data rather than a single averaged point.

HCMV infected U373Mg gene expression kinetics were fairly consistent over the time course; however, examination of the profiles identified three types of expression kinetics within this cell type. (1) HCMV genes whose expression increased from 12 to 24 h PI, then plateaued or slowly declined to 72 h PI (e.g. Fig. 4.15A, B, D, E, F, G, H, J, K and L; Fig. 4.16B, C, D, E, G and H). (2) HCMV genes whose expression increased from 12 to 48 h PI then plateaued from 48 to 72 h PI (e.g. Fig 4.15C; Fig. 4.16A and F). (3) Biphasic response, where gene expression increased from 12 to 24 h PI, plateaued from 24 to 48 h PI, and increased again from 48 to 72 h PI (Fig. 4.15I). The gene expression kinetics described in (1) and (3) correlates with genes that belong to the early kinetic class, and in (2) the late kinetic class.

The data suggest that in infected U373Mg cells, HCMV early gene expression peaks at 24 h PI and late gene expression at 48 h PI. In infected HFFF-2 cells, early and late gene expression peaks at 48 h and72 h PI, as is the case with infected RPE cells (Fig. 4.16). These data suggest that HCMV gene expression progresses through the transcription cascade more rapidly in U373Mg cells than in HFFF-2 or RPE cells, and this may account for the general high level of viral transcripts in this cell type. Transcript levels of early HCMV genes in U373Mg cells reach saturating levels before equivalent transcripts had either appeared or reached significant levels in HFFF-2 cells. Consequently, overall transcript level averages are generally higher in U373Mg cells. In the case of those HCMV genes whose transcript levels were lower in infected U373Mg cells compared to infected HFFF-2 cells, these are likely specifically down-regulated in U373Mg cells and so considered differentially expressed. However, in the case of the majority of HCMV genes whose transcript level was higher in infected U373Mg cells compared to infected HFFF-2 cells,

the situation is less clear. Because HCMV gene expression appears to be progressing through the transcription cascade at a faster rate in infected U373Mg cells compared to infected HFFF-2 cells, the rapid transit progression through the transcription cascade in infected U373Mg cells would allow transcripts to reach maximal gene expression before they could be reached in infected HFFF-2 cells. In these cases, the statistical tests cannot discriminate between differential gene expression and different rates of transit through the gene expression cascade.

The combined statistical tests identified a total of 26 HCMV genes as differentially expressed in HFFF-2 and U373Mg cells (Table 4.8). However, separating these genes that were under- or over-expressed in infected U373Mg cells compared to infected HFFF-2 cells identifies 7 HCMV genes (UL4, IRS1, US12, US14, US18, US19 and US20) under-expressed and so good candidates to be truly differentially expressed in these two cell types. The remaining genes were all over-expressed in infected U373Mg cells compared to infected HFFF-2 cells (Fig. 4.15), and are regarded as less good candidates for differential expression. The reduction of the differentially expressed gene list for HFFF-2 and U373Mg cells from 26 to the 7 listed above highlights the fact that UL4, IRS1, US12, US14, US18, US19 and US20 are also identified as differentially expressed in infected HFFF-2 and RPE cells (Table 4.7). As with all microarray data, RT-qPCR, northern and western blotting would be required to confirm their status as differentially expressed HCMV genes. Interestingly, translation of transcripts into protein product (Fig. 3.5, 3.6 and 3.7) occurred with similar kinetics in HFFF-2, RPE and U373Mg cells as assessed by western blots. Rapid transit through the transcription cascade did not result in high yields of infectious progeny virus in U373Mg cells, conversely, virus yields were the lowest attained from any of the cell types (Fig. 3.1, 3.2 and 3.3).

With regard to putative novel ORFs in the HCMV genome, the microarray analysis did not find sufficient evidence of transcription from the majority of these to support their status as functional ORFs. It was considered that present flags in three datasets were sufficient to define an ORF as present, since this represented expression at a single time point. However, many of the ORFs that passed this filter and were flagged as present could still not be unambiguously defined as genuine ORFs since they were located between known genes whose transcripts are likely to include the region containing the proposed novel ORFs. This argument does not hold for C-ORF16, C-ORF18 and C-ORF26 which all code in the opposite direction to established HCMV genes in the locality, and these ORFs may represent novel ORFs. The ORFs previously annotated by Chee et al.,

(1990) but more recently discounted, provided no evidence for expression confirming the expectations of Dolan et al., (2002).

5 RESULTS III

5.1 Validation of the data

While the HCMV microarray facilitated the rapid screening of the HCMV transcriptome for genes that are differentially expressed in different cell types, the microarray data by itself is not conclusive and alternative techniques such as RT-qPCR and northern blotting are required to confirm the microarray expression data. The data obtained from RT-qPCR and northern blotting techniques allows direct comparison with microarray expression data since these techniques also measure transcript expression and abundance, albeit with considerable differences in the level of the sensitivity. Gene expression profiles obtained by RT-qPCR and northern blotting were compared with the microarray expression data reported in chapter 4 with the view to comparing expression trends for individual genes using different detection systems. However, in contrast to the microarray, it must be noted that northen blot and RT-qPCR assays were capable of detecting both sense and anti-sense transcripts as probes for the northern blots were generated using double-stranded DNA PCR products (see methods, 2.2.16.3), and RT-qPCR was based on the amplification of double-stranded DNA. This is particularly relevant as recent evidence suggests that anti-sense transcripts can be detected for approximately 46 % of the HCMV genome, and many of these anti-sense transcripts are overlapping with established HCMV genes (Zhang et al., 2007). This could lead to additional anti-sense transcript bands detected by the northern blots, while also influencing the signal detected during RT-qPCR. Based on the assumption that the kinetics and abundance of transcript expression are directly related to the kinetics and abundance of the associated protein, western blotting has also been employed to further test the microarray expression data. However, it must be noted that the availability of antibodies against HCMV proteins is limited, and no antibodies were available for those genes identified as differentially regulated in HFFF-2 and RPE cells. Primers that were used during RT-qPCR are shown in Table 5.1, while northern blotting probe information is shown in Table 5.2.

5.2 HCMV 3′ co-terminal transcripts, overlapping transcripts and spliced genes

The microarray probes will hybridise to all transcripts running through to a common shared poly(A) site, and so the signal detected will derive from multiple rather than a specific transcript. To assess which gene probes might be affected by binding multiple transcripts, an evaluation of poly(A) sites proximal to HCMV ORFs is shown in Table 5.3 and Fig. 1.2 (see introduction, 1.3.1). HCMV, like other herpesviruses, has a genome that

HCMV Primer	Genome Position	Sequence (5′-3′)
UL16(F)	23062	CGACATCACCACTAACATCT
UL16(R)	23200	AGAGGCGCTCGATTATT
UL55(F)	84316	GGTGTGGATGTAAGCGTA
UL55(R)	84401	GGCATCATGGTGGTCTACAA
UL83(F)	121209	TTGCCCTGGATGCGATACTG
UL83(R)	121285	TGCGCTCTTCTTTTTCGATA
UL99(F)	145715	GAACTCTGCAAACGAATATG
UL99(R)	145801	GGGATGTTGTCGTAGGAG
UL123(F)	172421	AGGCAACTTCCTCTATCTCA
UL123(R)	172550	CCCTCTGTCCTCAGTAATTG
US12(F)	206920	AATTTGACGGTGAGCGATA
US12(R)	207021	GTGCAGTCCTGGGAACCATA
US18(F)	212754	CCACGCTGGTAGATGAGA
US18(R)	212867	CCGTCATCGTCTTTTACCTA

Table 5.1. List of HCMV primers used for RT-qPCR

HCMV ORF	Probe Length (bases)	Genome Position
UL4	313	13936 - 14249
UL16	282	22861 - 23143
UL43	288	55891 - 56179
UL55	349	82460 - 82809
UL83	235	121565 - 121800
UL99	289	145796 - 146085
UL123	472	172777 - 173249
IRS1	286	198271 - 197985
US12	339	206906 - 207245
US18	328	212673 - 213001

Table 5.2. Northern blotting probe information

Table 5.3. HCMV polyadenylation signal (AATAAA) genome positions and predicted poly(A) usage by HCMV ORFs.

Genome Position	Strand	HCMV ORFs proximal to downstream poly(A) sites	Genome Position	Strand	HCMV ORFs proximal to downstream poly(A) sites
132	F		2510	R	RL5A
2367	F	RL1	5917	R	
3093	F		5980	R	RL6
3932	F		6782	R	
4309	F		11609	R	
4505	F		12126	R	
4517	F		12564	R	
6969	F		12856	R	
7502	F		12964	R	UL2
9842	F	RL10, RL11	14241	R	
11014	F		15387	R	
12081	F	RL12, RL13	16848	R	
12935	F	UL1	16990	R	
13152	F		17702	R	
14100	F		21103	R	
15088	F	UL4	22575	R	
15810	F	UL5, UL6	26803	R	UL21A
15857	F	UL6	27895	R	
15934	F		28158	R	UL23, UL24, UL26, UL27, UL28, UL29, UL30
17601	F	UL7, UL8	37629	R	UL32
18265	F	UL9	44792	R	
21144	F	UL10, UL11, UL13	47905	R	
22529	F	UL14	48471	R	UL36, UL37, UL38
23966	F	UL15A, UL16, UL17	51227	R	UL37, UL38
25557	F	UL18, UL19	51375	R	UL38
26787	F	UL20	53522	R	UL40, UL41A, UL42, UL43, UL44, UL45, UL46
28138	F	UL22A	68788	R	
28288	F		70441	R	
32458	F	UL25	70735	R	
37582	F		71441	R	UL48A, UL49, UL50
46379	F	UL31, UL33, UL34	74580	R	UL51
48423	F	UL35	77551	R	
54133	F		77620	R	UL54, UL55, UL56, UL57
54533	F		92481	R	
59257	F		92651	R	UL69, UL70
59934	F		105865	R	UL72
60704	F		106669	R	UL74
70073	F	UL47	109112	R	UL75, UL79
70091	F		117619	R	
71372	F	UL48	118726	R	UL82, UL83
92436	F	UL52, UL53	122486	R	UL84, UL85, UL86
97111	F		129353	R	
98195	F		130820	R	
99193	F		132697	R	UL89
104558	F		134534	R	
108127	F	UL71, UL73	135908	R	
109052	F		148606	R	UL100
115768	F	UL76, UL77, UL78	149261	R	UL103, UL104
118682	F	UL80, UL80.5 Overlapping genes	154291	R	
118734	F	UL80 and UL80.5	156136	R	
122432	F		160460	R	
129407	F		160676	R	
139210	F	UL87, UL88, UL91, UL92	161210	R	UL114
140263	F		165011	R	UL115, UL116, UL117, UL119
146447	F	UL93, UL94, UL95, UL96, UL97, UL98, UL99	168831	R	UL120, UL121
156281	F	UL102, UL105	170607	R	UL122 (IE2)/UL123 (IE1) Spliced transcripts
161739	F	UL111A	172236	R	
157595	F		173987	R	UL128
158069	F		176183	R	UL130, UL131A

Genome Position	Strand	HCMV ORFs proximal to downstream poly(A) sites	Genome Position	Strand	HCMV ORFs proximal to downstream poly(A) sites
161739	F		178174	R	UL132, UL148, UL147A, UL146, UL147
164240	F	UL112	181304	R	
168146	F		181499	R	UL144, UL145, UL142
174574	F	UL124	184383	R	UL141, UL140
174628	F		187122	R	UL139, UL138, UL136
174703	F		188701	R	UL135
182127	F		189884	R	UL133, UL148A, UL148B, UL150
182752	F		195536	R	US1
182822	F		199305	R	US2
184127	F		200129	R	US3
186856	F		200133	R	
187083	F		201597	R	US6
191339	F		202590	R	US7, US8, US9, US10
194306	F	UL148C, UL148D	205300	R	US11
199253	F	IRS1	206766	R	US12, US13, US14, US15, US16, US17
200114	F		212449	R	US18, US19,US20
200348	F		216033	R	US21, US22, US23
203416	F		219795	R	US24
205373	F		221909	R	US26
210531	F		227466	R	
212454	F		231967	R	
213596	F		232393	R	TRS1
216070	F				
217037	F				
226497	F	US27, US28			
228904	F	US29, US30			
230186	F	US31, US32			
231505	F	US34			
235198	F	US34A			

Table 5.3. HCMV polyadenylation signal (AATAAA) genome positions and predicted poly(A) usage by HCMV ORFs.

The table lists the positions of the most common polyadenylation signal (AATAAA) in the HCMV genome. The table also lists the strand on which the signals are located (F, forward; R, reverse), and the HCMV ORFs proximal to the polyadenylation signal. This table does not predict actual poly(A) usage by HCMV ORFs, but serves as an indication of the complexity of the elements of genetic control in the HCMV genome.

is densely packed with ORFs, but contains relatively few polyadenylation signals, required for transcription termination and the subsequent addition of a polyadenylate tail on the free 3'-end of mRNAs. Consequently, many spliced and non-spliced HCMV genes share polyadenylation signals, leading to the generation of 3' co-terminal groups of genes within specific regions of the HCMV genome (Wing and Huang, 1995; Smuda et al., 1997; Guo and Huang, 1998). Some HCMV ORFs also overlap, e.g. UL146/UL147A (Lurain et al., 2006), while others are spliced, e.g. UL122/UL123 (Awasthi et al., 2004) and UL37 (Adair et al., 2004). The table lists the most common polyadenylation signal (AATAAA) within the HCMV genome, together with location on the forward (F) or reverse strand (R), and also the HCMV ORFs that are 3'-proximal to the signal. While this table suggests poly(A) site usage by specific HCMV genes, it should be noted that HCMV poly(A) signals are not utilised equally or efficiently during transcription. For example, a poly(A) site is located downstream of UL94; however, UL93 and UL94 specific transcripts utilise a poly(A) site further downstream of UL99 (Wing and Huang, 1995).

In the case of 3' co-terminal genes using the same poly(A) site, it is possible that cDNA fragments from multiple transcripts will bind to the probes thus skewing the microarray expression data for individual members of the co-terminal gene family. For comparisons between microarray expression data and that obtained from northern blots, it is necessary to take into account whether individual genes are part of a 3' co-terminal family and/or spliced etc. The data calculated from band intensities and the band intensities of all transcripts detected by the microarray probe are added together to compare expression data in the two systems. Examples presented here include 3' co-terminal genes US12 to US17, US18 to US20, UL93 to UL99 and UL54 to UL57, and spliced genes UL122/UL123 (IE2/IE1). Summary diagrams of the relevant transcription units detected during northern blotting are presented for non-differentially expressed genes; UL123 (Fig. 5.1.1; note that microarray probe positions, the northern blot probe region, and the region amplified during RT-qPCR are also included in this figure), UL55 (Fig. 5.2.1), UL83 (Fig. 5.3.1), and UL99 (Fig. 5.4.1), and differentially expressed genes; IRS1 (Fig. 5.8.1), US12 (Fig. 5.9.1) and US18 (Fig. 5.10.1). Note that northern blot band sizes were determined using two RNA ladders (Invitrogen) that were run at either side of the gel, in the first and last lanes.

5.3 Expression and validation of representative non-differentially regulated genes

UL123 (IE1) transcripts were detected by the microarray (Fig. 5.1A) from 12 h PI in each cell type. Expression of UL123 in HFFF-2 and RPE cells differs at the early time points as transcript levels were stable from 12 h PI to 24 h PI in HFFF-2 cells, but decreased in RPE

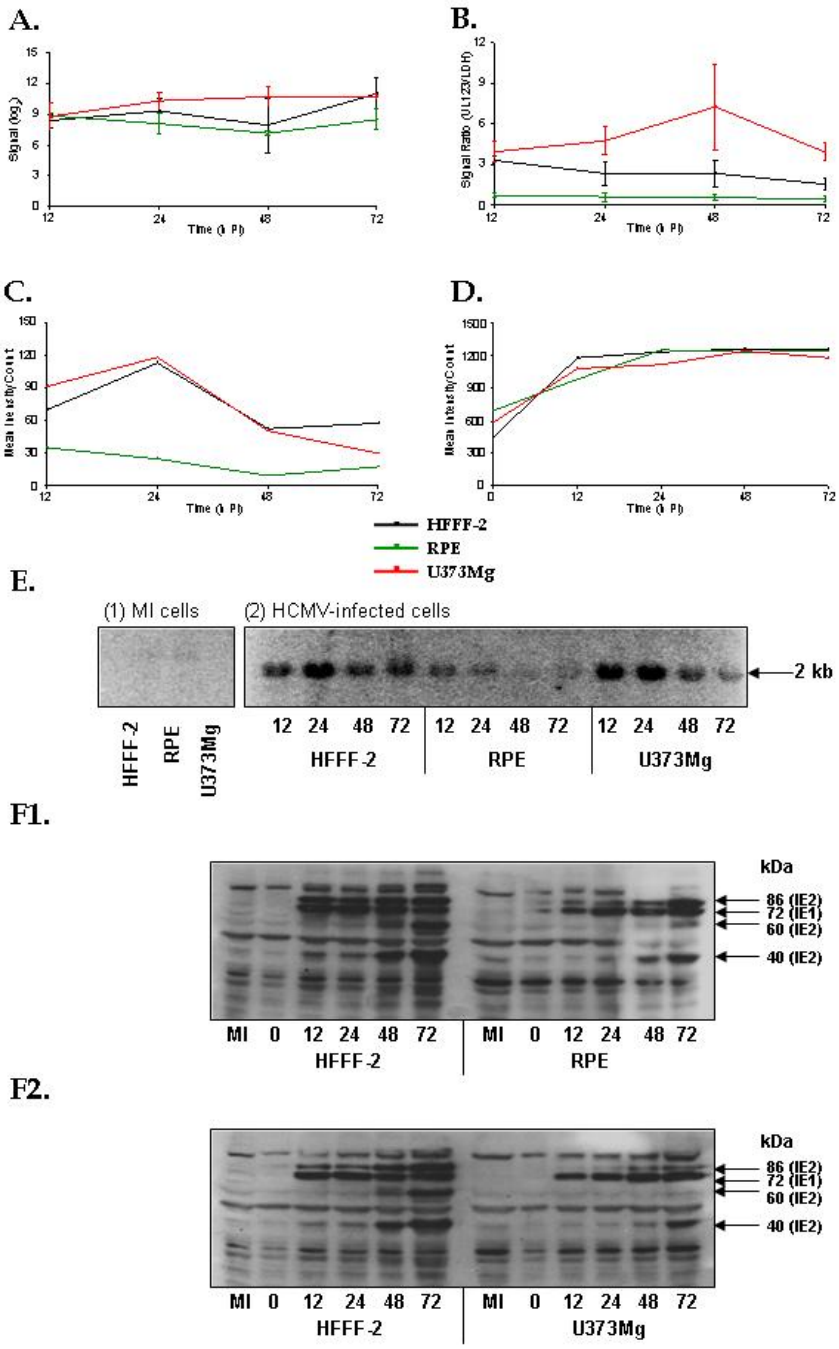


Figure 5.1. Comparison of microarray expression data for UL123 (IE1) with associated RT-qPCR, northern blotting for UL123, and western blotting for IE1/IE2.

(A), (B), (C) and (D) show the expression profiles for UL123 (IE1) for microarray, RT-qPCR, northern and western blotting respectively. (E) shows the northern blot for UL123 and (F1) and (F2) show the western blot for IE1/IE2 in HFFF-2 and RPE, and HFFF-2 and U373Mg cells respectively. Note that for microarray and RT-qPCR analysis, triplicate biological replicates were used, whereas only a single infected sample was used for northern and western blotting. Note that panels F1 and F2 are the same as those shown in Fig. 3.4.

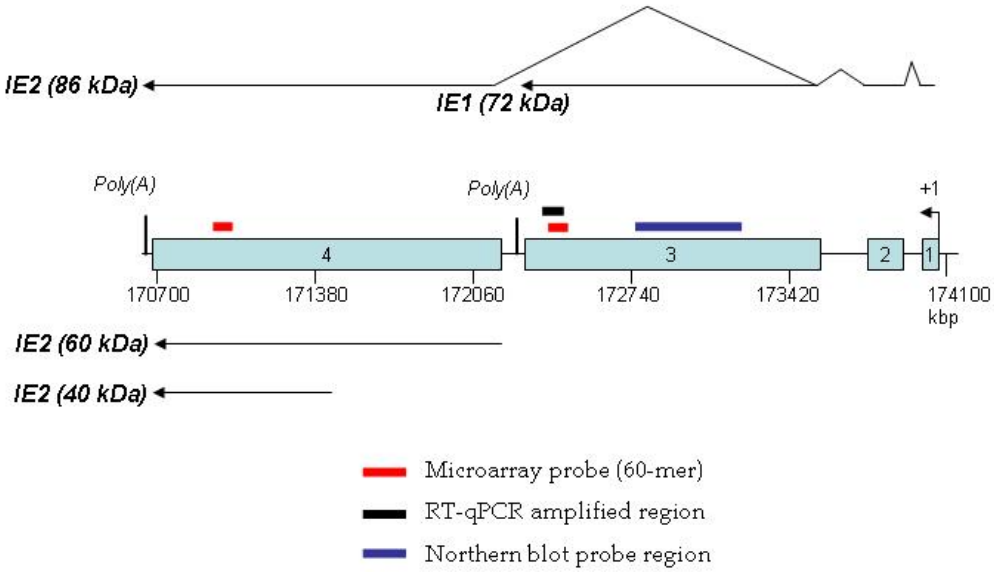


Figure 5.1.1. Summary diagram of the IE1/IE2 gene locus, with the positions of the microarray, RT-qPCR and northern blot probe regions.

The diagam shows the organisation of the IE1/IE2 gene locus, with distances shown in kbp (reverse strand), together with information regarding the expression of the spliced transcripts and their protein products. Microarray probes, RT-qPCR amplification, and northern blot probe regions are also indicated on the diagram. As indicated on the diagram, a Poly(A) signal is located at the end of exon 3, and another at the end of exon 4.

cells. However, from 24 h PI the expression kinetics were similar in these two cell types. The expression of UL123 in U373Mg cells increased from 12 to 24 h PI, and plateaued thereafter. The RT-qPCR expression data for UL123 (IE1) (Fig. 5.1B) confirms that the expression kinetics of UL123 in HFFF-2 and RPE cells are similar. The RT-qPCR data confirms that the levels of UL123 transcript in U373Mg cells (Fig. 5.1B) are significantly higher than in HFFF-2 and RPE cells, and that the kinetics are broadly similar to those obtained with the microarray (Fig. 5.1A). The apparent over-representation of UL123 in U373Mg cells (Fig. 5.1B) is a result of the normalisation procedure employed for RT-qPCR, where the cellular housekeeping gene, Lactate Dehydrogenase (LDH) was selected as an internal control for RNA input and the efficiency of cDNA synthesis. Examination of the LDH profile with RT-qPCR shows constant expression levels in all cell types over the 72 h time course. However, expression levels of LDH were consistently lower in U373Mg cells compared to HFFF-2 and RPE cells. Subsequently, normalising RT-qPCR gene expression data with LDH resulted in an artificially elevated gene expression profile for U373Mg cells. It should be noted that UL123 copy numbers in each cell type were of a similar order of magnitude prior to normalisation.

Northern blot analysis of the UL123 transcript (Fig. 5.1E) detected a single band of 2 kb, with expression from 12 h PI in each cell type. The band intensities were determined using the Quantity One program, where the mean intensity count (CNT) for each band was extracted with a global background correction. The band intensity data is displayed as a graph (Fig. 5.1C). The expression of UL123 as determined by the northern blot (Fig. 5.1C) follows the same kinetics as the microarray profile (Fig. 5.1A). However, the expression of UL123 in U373Mg cells from 24 h PI in (Fig. 5.1C) declines more quickly compared to the profile in (Fig. 5.1A).

Western blot data for IE1 and IE2 spliced variants are shown in panels (Fig. 5.1F1) and (Fig. 5.1F2) (HFFF-2 and RPE; HFFF-2 and U373Mg cells respectively). Expression of IE1 (72 kDa) and IE2 (86 kDa) were detected at all time points. Within each cell type the levels of IE1 protein were fairly constant at each time point, while levels of IE2 (86 kDa, 60 kDa and 40 kDa species) appeared to increase gradually over the time course. Band intensities for IE1 were extracted with a global background correction, and the data displayed as a graph in (Fig. 5.1D). IE1 expression was very similar in each cell type, with a rise in expression from 0 to 12 h PI, which plateaued from 24 to 48 h PI in each cell type.

Microarray expression data for the early-late gene UL55 (gB) (Fig. 5.2A) indicated that transcripts accumulated to 24 h PI and plateaued to 72 h PI in U373Mg cells; however,

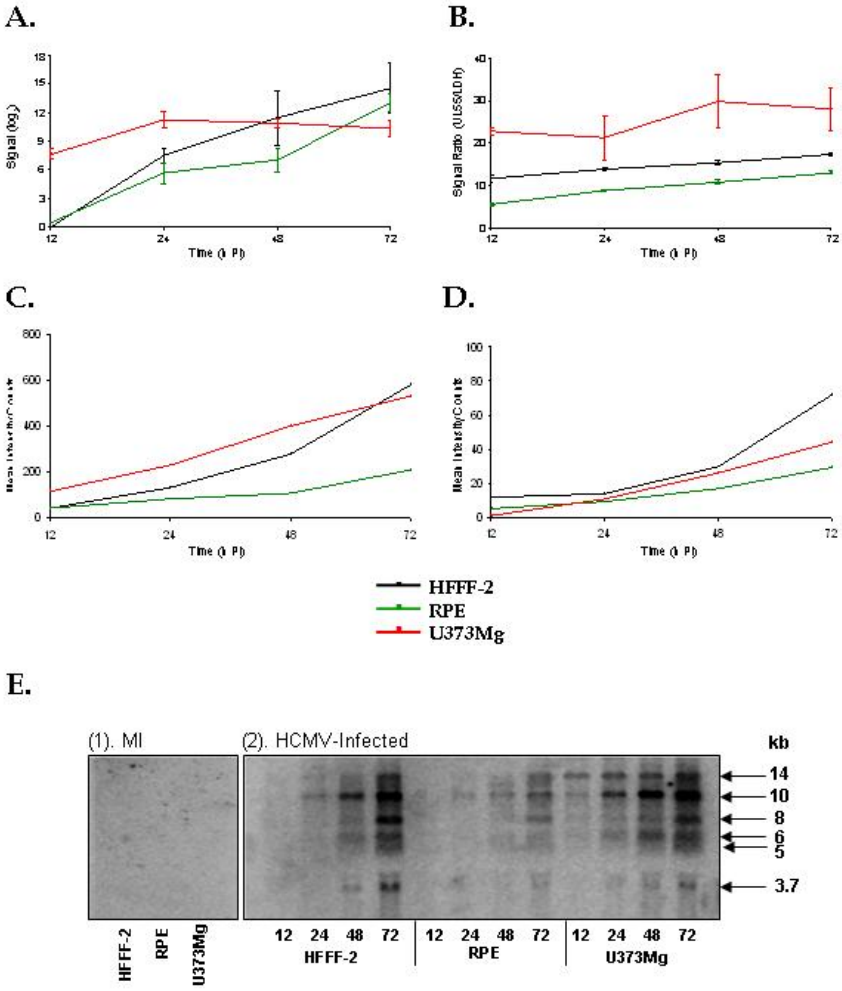


Figure 5.2 Comparison of microarray data for UL55 (gB) with expression data detected by RT-qPCR, northern blot data.

(A) shows the microarray expression profile for UL55, (B) shows the RT-qPCR expression profile. (C) shows the summed band intensity data in the northern blot (E), and (D) shows a graph for the UL55 specific bands on the northern blot.

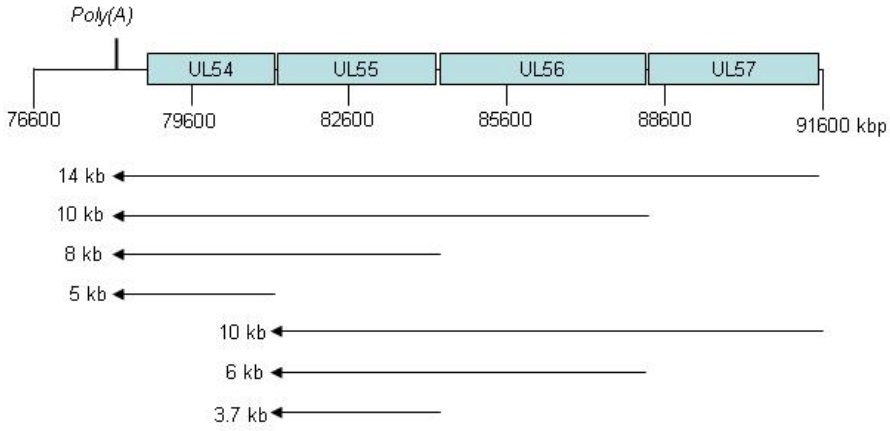


Figure 5.2.1. Summary diagram of the transcripts expressed in the UL54 to UL57 gene region.

The diagram shows the expression of transcripts, which begin at a TATA box at the start of each ORF and terminate at a poly(A) signal downstream of UL54; however, Smuda et al. (1997) also identified transcripts ending at a poly(A) signal downstream of UL55. In the case of HCMV strain Merlin, it is a non-canonical poly(A) signal.

UL55 transcripts were not detected until 24 h PI in HFFF-2 and RPE cells, but accumulated continuously thereafter with similar kinetics. In contrast, the RT-qPCR data (Fig. 5.2B) detected expression of UL55 in all cell types from 12 h PI. Expression levels of UL55 appear to be significantly higher in U373Mg cells compared to HFFF-2 and RPE cells due to the artefact of normalising as discussed previously. The kinetics of UL55 expression in HFFF-2 and RPE cells are similar in (Fig. 5.2A and B), with a steady increase in expression over the time course.

The northern blot for UL55 expression (Fig. 5.2E) detected multiple transcripts of 14 kb, 10 kb, 8 kb, 6 kb, 5 kb and 3.7 kb coded by the 3' co-terminal genes UL54, UL55, UL56 and UL57. The 5 kb and 3.7 kb transcripts are reported to be monocistronic, representing UL54 and UL55 respectively. The remaining overlapping large transcripts (14 kb, 10 kb, 8 kb and 6 kb) can be specifically mapped from a TATA box within this gene region, and extending to one or other of two polyadenylation signals located downstream of either UL55 or UL54 (Smuda et al., 1997). Because UL55 is the terminal gene, the UL55-specific probes on the microarray would detect all the transcripts expressed within this region. The 14 kb, 10 kb, 8 kb, and 6 kb transcripts were detected throughout the time course in U373Mg cells, but accumulated from 24 h PI onwards in HFFF-2 and RPE cells (Fig. 5.2E). In order to compare the microarray data (Fig. 5.2A) and the RT-qPCR (Fig. 5.2B) with the northern blot data, it is necessary to sum the individual northern band intensities for the 14 kb, 10 kb, 8 kb, 6 kb, 5 kb, and 3.7 kb transcripts at each time point for each cell type (Fig. 5.2C). The data in (Fig. 5.2C) for each cell type shows a continuous increase in transcripts binding to the UL55 probe over the time course, although the expression kinetics vary compared to (Fig. 5.2A). While it is clear that the microarray, RT-qPCR and northern data obtained for the UL55 probes are in agreement, the actual UL55 transcript (3.7 kb) is only a small part of the overall signal. Figure 5.2D was based on the band intensities of the 3.7 kb band as this shows the UL55 specific signal. The kinetics of UL55 expression were similar in each cell type as the curves showed a sustained increase from 24 h PI through to 72 h PI.

Figure 5.3A shows the microarray data for the late gene UL83 (pp65). Expression was detected in each cell type from 12 h PI, and increased rapidly from 24 h PI in HFFF-2 and RPE cells. UL83 expression levels plateaued in U373Mg cells after 24 h PI. The RT-qPCR profile (Fig. 5.3B) confirmed that UL83 transcript expression in HFFF-2 and RPE cells shared the same kinetics, although in contrast to Fig. 5.3A, the late increase in UL83 transcript levels was not detected until 72 h PI. The expression profile for UL83 obtained in U373Mg cells (Fig. 5.3B) indicated a constant level of expression throughout the time

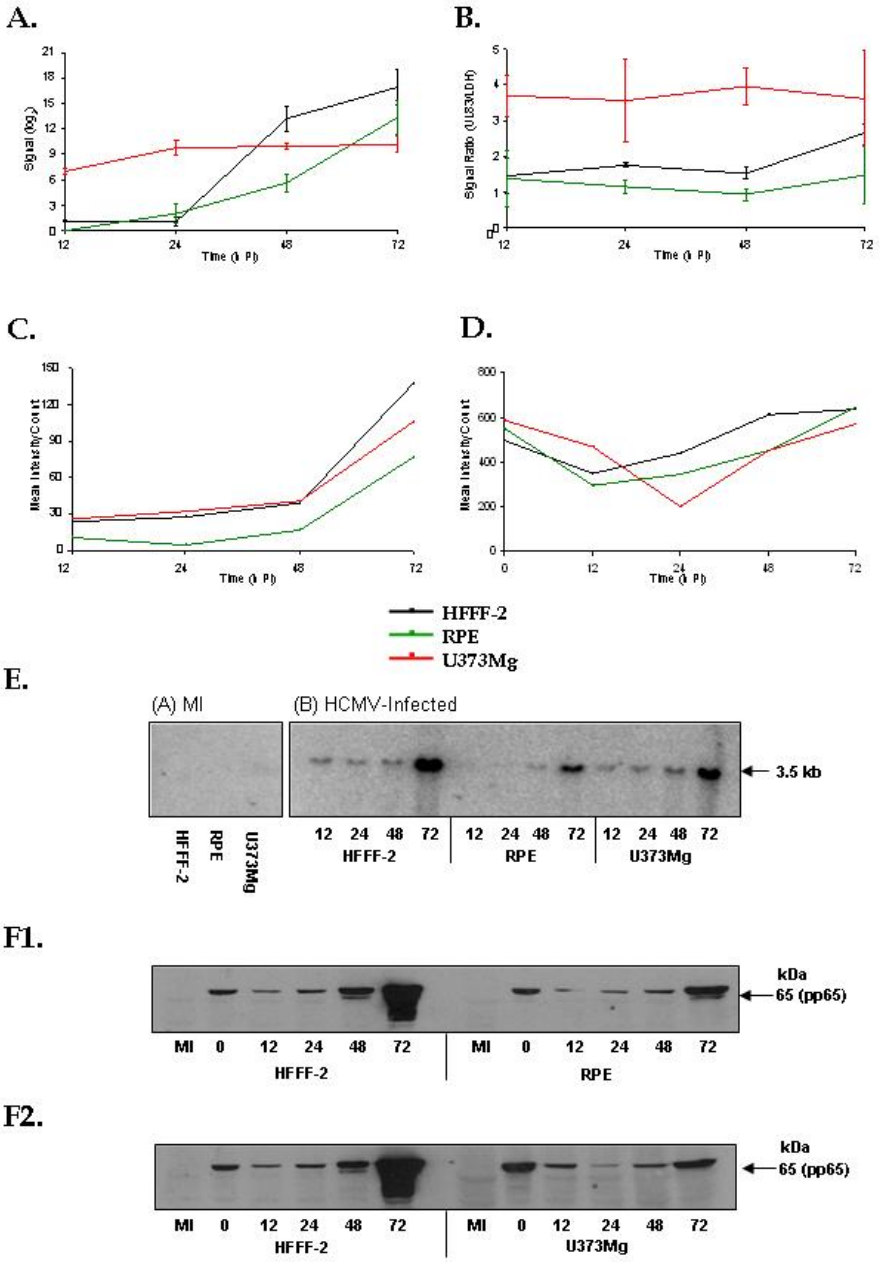


Figure 5.3. Comparison of microarray data for UL83 (pp65) with expression data detected by RT-qPCR, northern and western blotting.

(A) shows the microarray expression profile for UL83, (B) shows the RT-qPCR expression profile. (C) shows the band intensity data in the northern blot (E), and (D) shows a graph of the band intensity data for western blots for pp65 (F1 and F2).

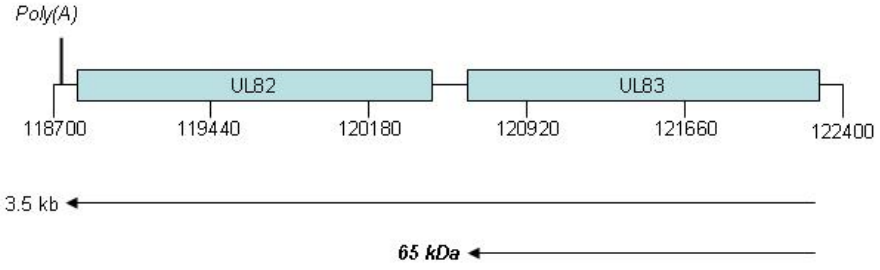
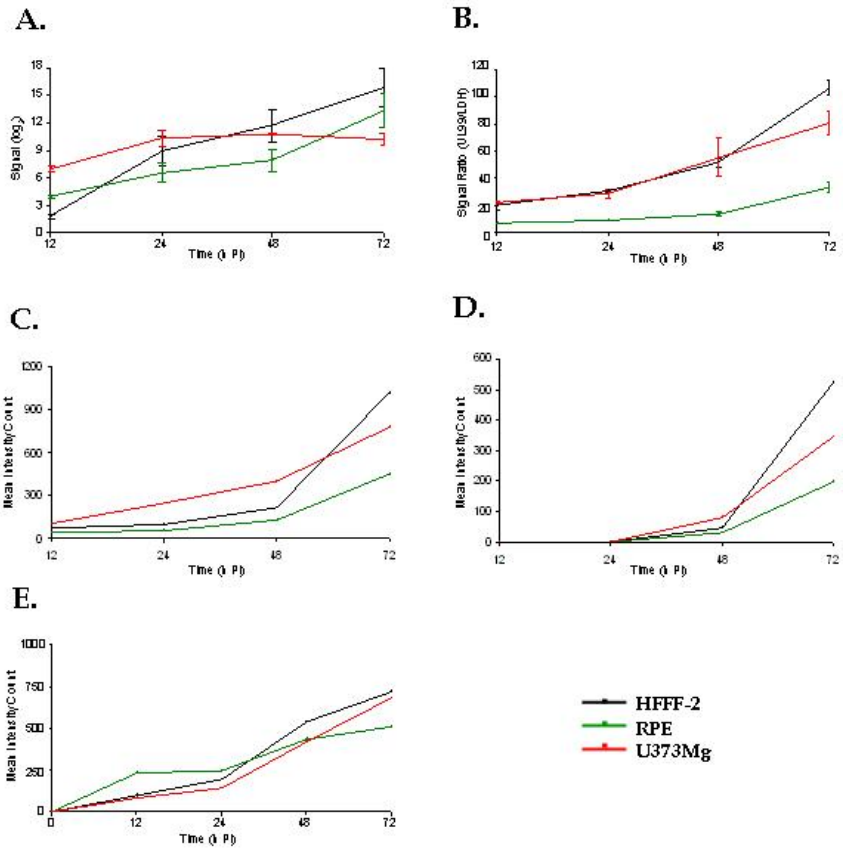


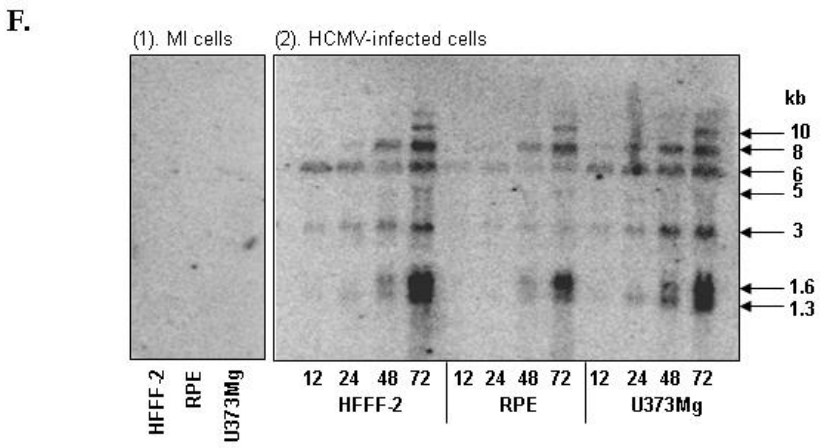
Figure 5.3.1. Summary diagram of the transcript expressed from the start of UL83 to a poly(A) signal downstream of UL82.

The diagram also indicates the protein that is translated from this transcript, pp65 (pUL83).

course and was consistent with the microarray data (Fig. 5.3A). The northern blot for UL83 (Fig. 5.3E) detected a single band of 3.5 kb, band intensity data is displayed in (Fig. 5.3C). This transcript is likely to contain both the UL82 (pp71) and UL83 (pp65) ORFs, as these genes share a single polyadenylation signal downstream of UL82. The expression profile for UL83 in HFFF-2 and RPE cells (Fig. 5.3C) exhibited similar kinetics and was consistent with the data obtained by both RT-qPCR (Fig. 5.3B) and microarray (Fig. 5.3A), albeit that UL83 expression in HFFF-2 is much greater at 48 h PI in (Fig. 5.3A) than in (Fig5.3B) and (Fig. 5.3C). Because pp65 (UL83) protein is a component of the virus tegument, it was detected at 0 h PI in western blots (Fig. 5.3F1 and F2) in each cell type. The levels of pp65 remained steady in both HFFF-2 and RPE cells, rising slowly from 24 h PI. The pp65 protein appeared to be less stable in U373Mg cells since levels decreased over the first 24 h, but increased again between 48 h PI and 72 h PI to give final amounts that were similar to those achieved in HFFF-2 cells.

The late UL99 gene codes for a tegument phosphoprotein, pp28, which is expressed with late kinetics (Chambers et al., 1999). The microarray data for UL99 (Fig. 5.4A) showed a continuous increase in expression from 12 h PI to 72 h PI in both HFFF-2 and RPE cells. In U373Mg cells, expression increased until 24 h PI, then plateaued thereafter. The RT-qPCR expression data for UL99 in HFFF-2 and RPE cells (Fig. 5.4B) shared similar kinetic trends with the microarray data (Fig. 5.4A). However, the incremental rise in transcript levels in HFFF-2 and RPE cells was slower in (Fig. 5.4B), and the plateau in UL99 levels in U373Mg cells seen in the microarray data (Fig. 5.4A) was not observed in the RT-qPCR data (Fig. 5.4B). The northern blot and associated band intensity data (Fig. 5.4E and C) showed that multiple transcripts of 10 kb, 8 kb, 6 kb, 5 kb, 3 kb, 1.6 kb and 1.3 kb are coded by the 3′ co-terminal genes UL93, UL94, UL95, UL96, UL97, UL98 and UL99. The UL99 specific transcripts are reported to be 1.3 kb and 1.6 kb (Wing et al., 1995). In the northern blots (Fig. 5.4F), the 1.3 kb transcript was detected from 24 h PI in each cell type, and the 1.6 kb transcript detected at 48 h PI and 72 h PI in each cell type. Note that approximate sizes of some transcripts differ slightly from those reported by Wing et al., who used the Towne strain of HCMV for transcript mapping. Because the UL99-specific probes on the microarray could detect all transcripts expressed in this region – UL99 being the terminal gene of this family of 3′ co-terminal transcripts, band intensities were aggregated according to time (h PI) for each cell type, and displayed in (Fig. 5.4C) for comparison with data in graphs (Fig. 5.4A) and (Fig. 5.4B). The expression kinetics for multiple transcripts binding to the UL99 microarray probe are the same in HFFF-2 and RPE cells, and the profile obtained from the northern blots (Fig. 5.4C) was consistent with the data obtained from RT-qPCR (Fig. 5.4B). In U373Mg cells, there was a continuous increase over


A.
B.
C.
D.
E.
Signal (log2)
Signal Ratio (U373/LDH)
Mean Intensity/Count
Time (h Pi)
HFFF-2
RPE
U373Mg


F.
(1). MI cells
(2). HCMV-infected cells
kb
10
8
6
5
3
1.6
1.3
HFFF-2
RPE
U373Mg
12 24 48 72
HFFF-2
12 24 48 72
RPE
12 24 48 72
U373Mg

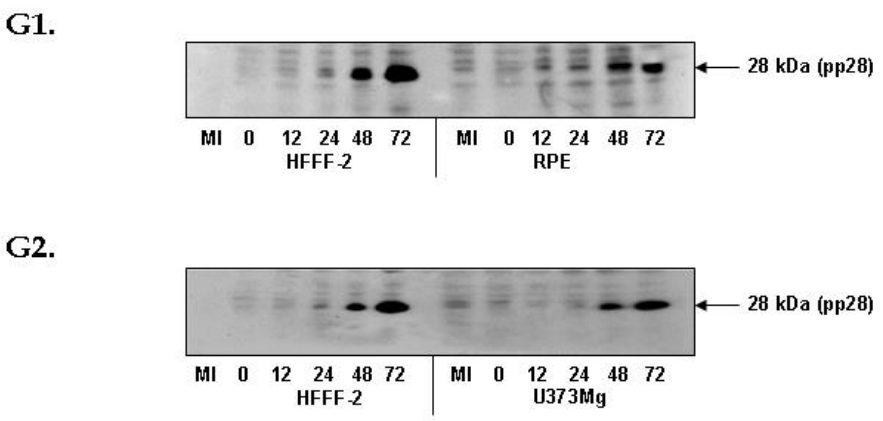


Figure 5.4. Comparison of microarray expression data for UL99 (pp28) with expression data from RT-qPCR, northern and western blotting.

(A) shows the microarray expression data for UL99, (B) shows the RT-qPCR expression data, (C) shows a graph of the intensity of the summed bands on the northern blot (F), and (D) shows a graph of the intensity of the UL99-specific bands. (E) shows a graph of the band intensities on the western blots (G1) and (G2).

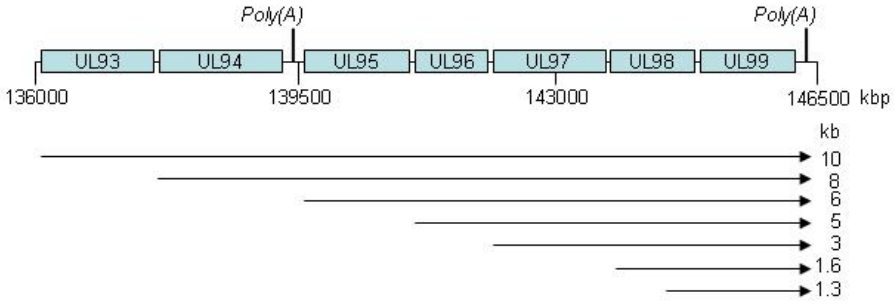


Figure 5.4.1. Summary diagram of the transcripts expressed from UL93 to UL99.

Note that there is a poly(A) signal downstream of UL94 that is not utilised during transcription by either UL93 or UL94. Transcripts for UL93 and UL94 both terminate at the poly(A) signal downstream of UL99 (Wing et al., 1995).

the time course, consistent with the RT-qPCR data (Fig. 5.4B), but not seen in the microarray data (Fig. 5.4A). Differences in UL99 expression kinetics in HFFF-2, RPE and U373Mg cells in (Fig. 5.4C) and (Fig. 5.4B) compared to (Fig. 5.4A) could be due to non-biological variation. Overall, the microarray data was in agreement with the data obtained for UL99 by RT-qPCR and northern blotting. To assess the actual UL99 expression levels and kinetics in each cell type, the levels of specific UL99 1.3 kb and 1.6 kb transcripts were summed and compared (Fig. 5.4D) with the levels of pp28 (UL99) protein detected in western blots (Fig. 5.4E). Western blot data for UL99 (Fig. 5.4G1 and G2) shows low levels of pp28 protein at 24 h PI in each cell type, which increased rapidly between 48 h PI and 72 h PI. The band intensity data is shown in (Fig. 5.4E). No difference is observed in the relative abundance or kinetics of expression of pp28 in each cell type, and this finding is in accord with the expression kinetics for transcripts binding the UL99 probes in (Fig. 5.4A, B, C and D).

Figure 5.5 shows the microarray (Fig. 5.5A) and northern blot (Fig. 5.5C) data for the late gene UL43, and the associated northern blot band intensity graph (Fig. 5.5B). The microarray detected UL43 from 24 h PI in HFFF-2 and RPE cells, and from 12 h PI in U373Mg cells (Fig. 5.5A). UL43 levels increased continuously over the 72 h period in HFFF-2 and RPE cells but reached plateau levels after 24 h PI in U373Mg cells (Fig. 5.5A). The northern blot (Fig. 5.5C) detected a single band of 5.4 kb, which is expressed from 24 h PI in HFFF-2 and RPE cells, but from 12 h PI in U373Mg cells. The band intensity data (Fig. 5.5B) confirms the increase in UL43 expression in HFFF-2 and RPE cells over the course of the experiment with the exception that in HFFF-2 cells levels were dramatically increased between 48 and 72 h PI. A steady increase in expression is observed in U373Mg over the time course. The microarray data (Fig. 5.5A) and the northern blot data (Fig. 5.5B) obtained for UL43 gene expression exhibit similar trends and are in agreement with each other.

5.4 Expression and validation of differentially regulated genes

Microarray expression data for the early gene UL4 showed that transcripts were detected from 12 h PI in all cell types (Fig. 5.6A). UL4 expression levels were very high throughout the time course in HFFF-2 cells, and combined statistical tests determined significant differences in expression in both comparisons of HFFF-2 and RPE, and HFFF-2 and U373Mg cells. Expression in RPE cells was much lower during the first 48 h PI, but by 72 h PI, levels were similar to those in HFFF-2 cells. The UL4 expression kinetics in U373Mg cells was similar to those for other early genes with maximum expression at 24 h PI.

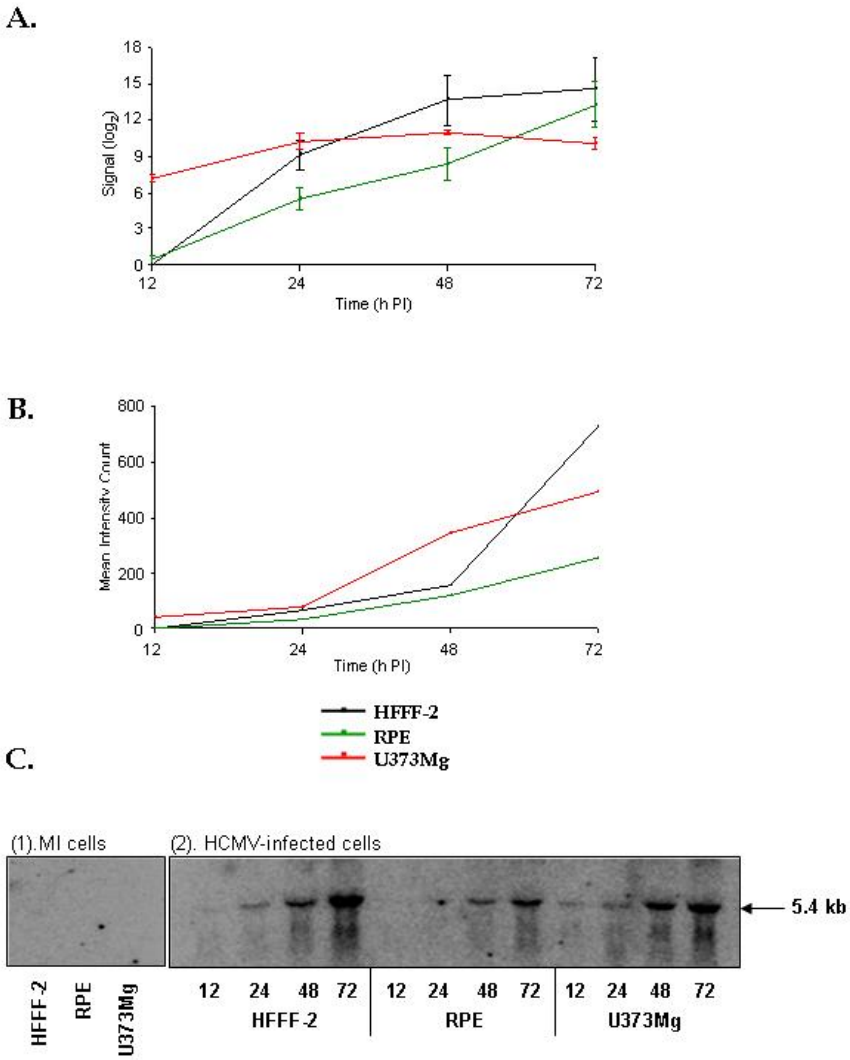


Figure 5.5. Comparison of microarray expression data for UL43 and the associated northern blot.

(A) shows the microarray expression profile for UL43, (B) shows the band intensity graph of the northern blot (C). It is not possible to predict which poly(A) signal is utilised because there are several alternatives. Detailed mapping of the transcript would be required to decipher the 5′ and 3′ ends of the transcript. UL43 is part of a 3′ co-terminal family (UL40 – UL46), however, only one band is shown in the northern blot (C).

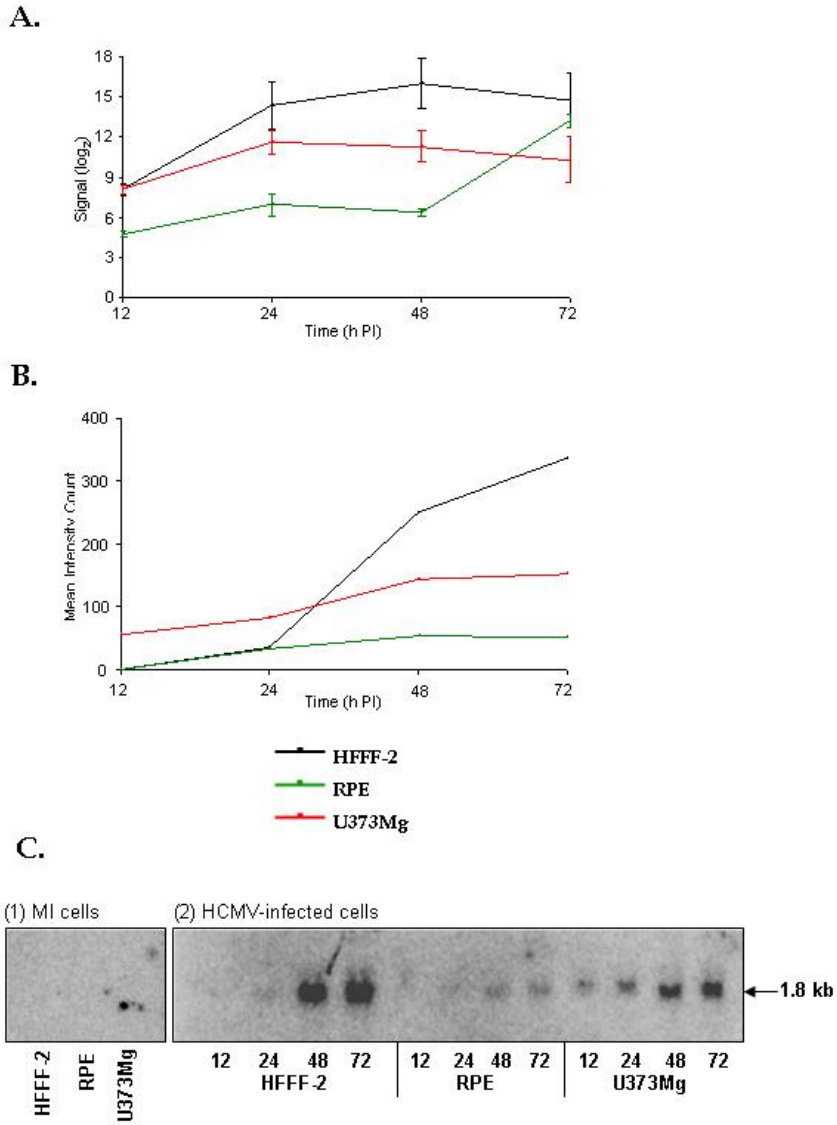


Figure 5.6. Comparison of microarray expression data for UL4 with associated RT-qPCR and northern blot data.

(A) shows the microarray expression profile, (B) shows the graph of the band intensites of the northern blot (C). A single UL4 transcript is expressed to a poly(A) signal at genome position 15088 kbp (see Table 5.1).

Northern blots (Fig. 5.6C) detected a single UL4 transcript of 1.8 kb from 12 h PI in U373Mg cells (Fig. 5.6C), but in HFFF-2 and RPE cells, UL4 was not detected until 24 h PI. Band intensity data for the northern blot (Fig. 5.6B) showed that UL4 expression in U373Mg and RPE cells remained fairly constant, with a gradual increase in expression over the time course. In contrast, UL4 expression levels in HFFF-2 and RPE cells were the same from 12 h PI to 24 h PI, but in HFFF-2 cells, UL4 expression increased sharply from 24 to 72 h PI. Despite the differences in temporal kinetics of UL4 expression between the microarray data (Fig. 5.6A) and the northern blot data (Fig. 5.6B and C), it is clear that the differential expression of UL4 in infected HFFF-2 and RPE, and HFFF-2 and U373Mg cells was confirmed by the northern blots.

Expression of the HCMV early gene UL16 is in RPE cells is considerably lower than in HFFF-2 cells, although expression kinetics were similar (Fig. 5.7A). In contrast, the expression profile of UL16 in U373Mg cells conforms to that of other early genes in this cell type with maximal expression levels at 24 h PI. The RT-qPCR expression profiles obtained for UL16 (Fig. 5.7B) in HFFF-2 and RPE cells were similar to those obtained with the microarray (Fig. 5.7A). UL16 expression in U373Mg cells assessed by microarray (Fig. 5.7A) and RT-qPCR (Fig. 5.7B) were also similar, differing in the time of maximal expression, although the large error bars in (Fig. 5.7B) may mask more similar kinetics. The northern blot for UL16 (Fig. 5.7D) showed a single UL16 transcript of 1.3 kb that was present throughout the time course in both HFFF-2 and U373Mg cells, but which appeared to be absent from infected RPE cells. It is clear from the microarray (Fig. 5.7A) and RT-qPCR (Fig. 5.7B) data that UL16 expression was very low in RPE cells; too low to permit detection by northern blotting. Northern blot band intensity data for UL16 is displayed in (Fig. 5.7C). Expression of UL16 increased gradually in HFFF-2 cells and plateaued at 48 h PI; this result was in conflict with the kinetic data obtained from both the microarray (Fig. 5.7A) and RT-qPCR (Fig. 5.7B) data. A similar disparity in UL16 expression kinetics in U373Mg cells is also apparent. Nevertheless, all three techniques support the conclusion that UL16 gene expression is down-regulated in RPE cells, compared to that in HFFF-2 and U373Mg cells.

The microarray expression kinetics obtained for the HCMV IE gene IRS1 (Fig. 5.8A) were reminiscent of those obtained for the UL4 gene (Fig. 5.6A). The expression kinetics for IRS1 in infected HFFF-2 and U373Mg cells correspond, increasing over the period 12 to 24 h PI, and plateauing thereafter. In contrast, the profile obtained for IRS1 expression in RPE cells exhibited a constant level to 24 h PI, increasing gradually thereafter to levels comparable to those found in the other cell types at 72 h PI. The northern blot (Fig. 5.8D)

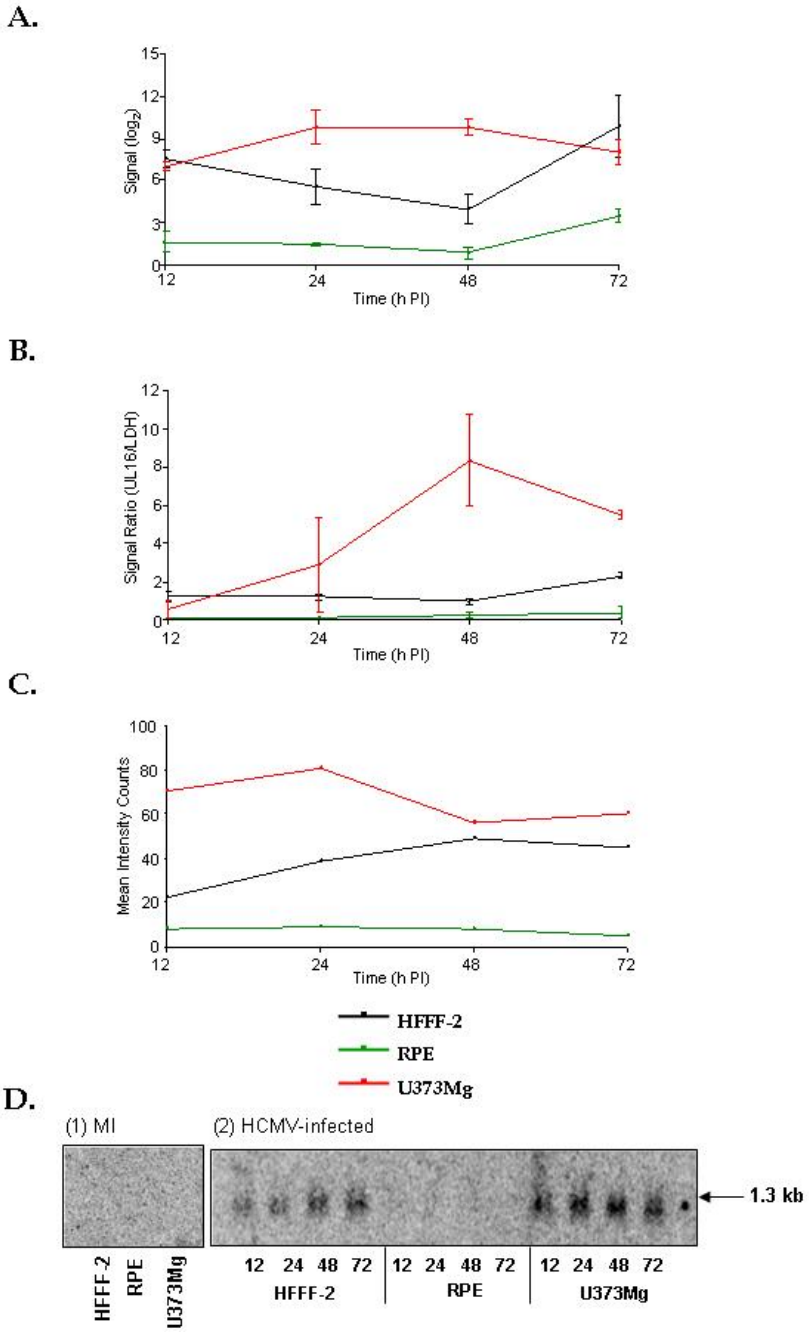


Figure 5.7. Comparison of microarray data with associated RT-qPCR and northern blot data for UL16.
(A) shows the microarray expression profile, (B) shows the RT-qPCR expression data, (C) shows the band intensity data in the northern blot (D). The single UL16 transcript is predicted to terminate at a poly(A) signal at genome position 23966 kbp (see Table 5.1). UL16 forms part of a 3′ co-terminal family, however, only one band was detected as UL16 is the leading gene in this 3′ co-terminal family.

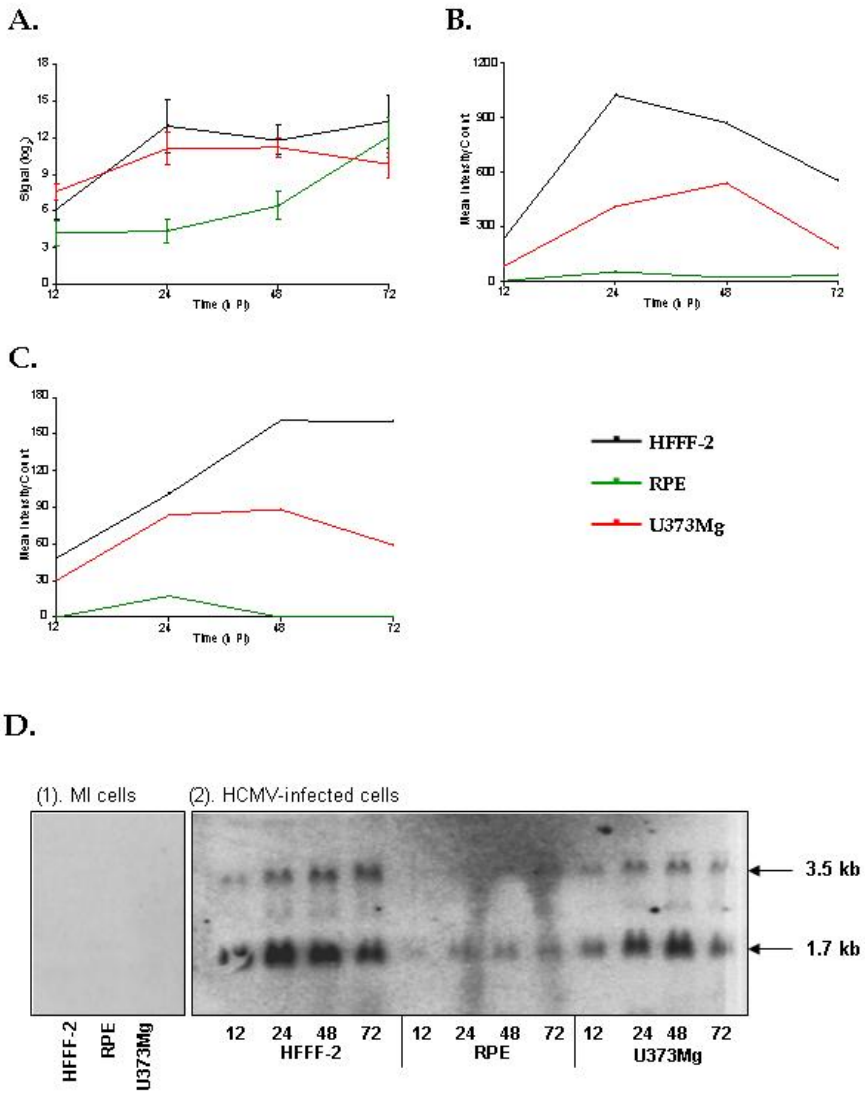


Figure 5.8. Comparison of microarray expression data for IRS1 with associated northern blot data.

(A) shows the microarray expression profile for IRS1, (B) shows the quantification of the two IRS1 bands in the northern blot (D), while (C) shows the quantification of the 3.5 kb IRS1 band only in the northern blot (D).

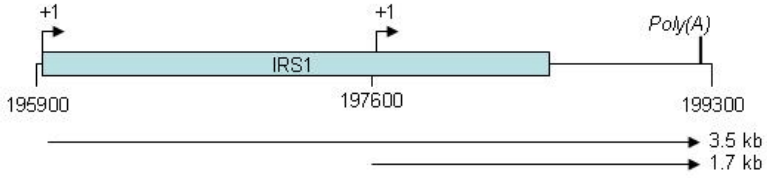


Figure 5.8.1. Summary diagram of IRS1 transcripts.

A ~ 3.5 kb transcript is expressed from the start of the gene to the poly(A) signal at genome position 199253. An internal transcription start site at genome position 197730 is utilised to produce a ~ 1.7 kb transcript.

identified two IRS1 specific bands of 1.7 kb and 3.5 kb. IRS1 has an internal transcription start site, which accounts for the small 1.7 kb transcript (Romanowski and Shenk, 1997). As the 3' proximal microarray probe would be able to detect both the 1.7 kb and 3.5 kb transcripts, the intensities of the 1.7 kb and 3.5 kb bands were summed at each time point in order to compare northern data (Fig. 5.8B) with the microarray data (Fig. 5.8A). The northern blot and microarray data show similar trends for IRS1 expression in HFFF-2 and RPE cells, except that in HFFF-2 cells, the northern showed a decline in IRS1 transcripts at 72 h PI, and in RPE cells, expression of IRS1 did not rise to levels seen in HFFF-2 cells. When IRS1 expression in HFFF-2 and U373Mg cells was compared, the microarray data (Fig. 5.8A) showed a fairly constant level of expression over the time course, but the northern blot data showed a sharp decline at 72 h PI (Fig. 5.8B), mirroring the kinetics seen in HFFF-2 cells. Examination of the IRS1 specific full length transcript (3.5 kb) in HFFF-2 cells shows that expression increases rapidly from 12 h PI, plateauing at 48 h PI (Fig. 5.8C). The magnitude and kinetics of expression of the 3.5 kb IRS1 transcript are different in RPE and U373Mg cells, where IRS1 is down-regulated in both cell types compared to HFFF-2 cells (Fig. 5.8C). Thus the data were in agreement with the results of the statistical tests, confirming that gene IRS1 was differentially expressed in infected HFFF-2 and RPE cells, and infected HFFF-2 and U373Mg cells.

HCMV US12 gene family members can be separated into two 3' co-terminal groups: US18, US19 and US20 forming one group using a polyadenylation signal downstream of US18 (Guo and Huang, 1993), and US12, US13, US14, US15, US16 and US17, forming a second group that are thought to use the polyadenylation signal downstream of US12. Possibly because the US12 gene family members form 3' co-terminal groups, they are over-represented in the 'differentially expressed' gene lists generated by the combined statistical tests.

The microarray expression profiles for transcripts binding the US12 probes (Fig. 5.9A) were similar for infected HFFF-2 and U373Mg cells, while expression levels in RPE cells were significantly lower from up to 48 h PI, and increased thereafter to levels similar to those in HFFF-2 cells at 72 h PI. The US12 expression data obtained by RT-qPCR (Fig. 5.9B) for infected HFFF-2 and RPE cells showed a steady-state level of expression over the time course, this deviates markedly from the kinetics obtained from the microarray data (Fig. 5.9A), although both techniques agree that US12 expression levels are lower in RPE cells. The apparently higher levels of US12 expression in U373Mg cells (Fig. 5.9B) is due to a normalisation artefact as previously discussed. The northern blot for US12 (Fig. 5.9D) shows two major bands of 4 kb and 1.1 kb. Since US12 is the terminal gene within the

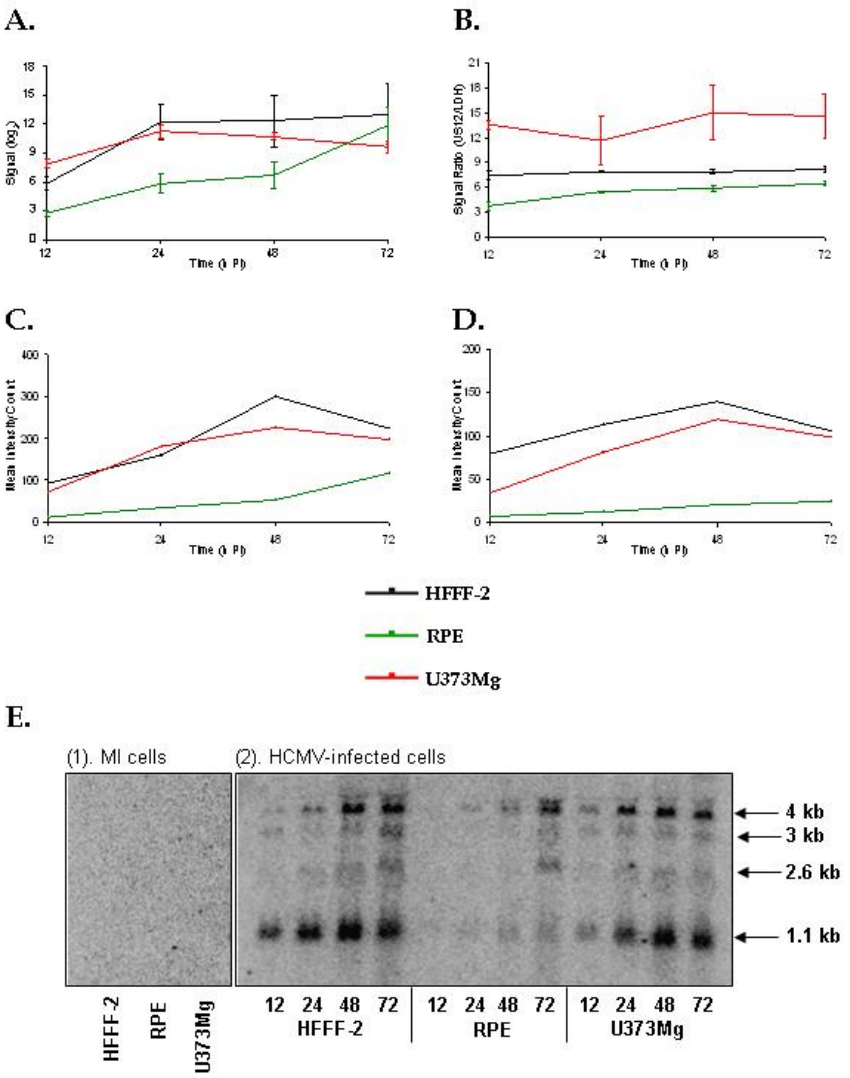


Figure 5.9. Comparison of microarray expression data for US12 with associated RT-qPCR and northern blotting.

(A) shows the microarray expression profile for US12, (B) shows the RT-qPCR profile, (C) shows the graph of the aggregated bands (4 kb plus 1.1 kb) in the northern blot (E), and (D) shows the graph for the likely US12 specific 1.1 kb band in the northern blot (E). The minor 2.6 kb and 3 kb bands are likely to represent transcript expression from the start of US13 and US14 respectively.

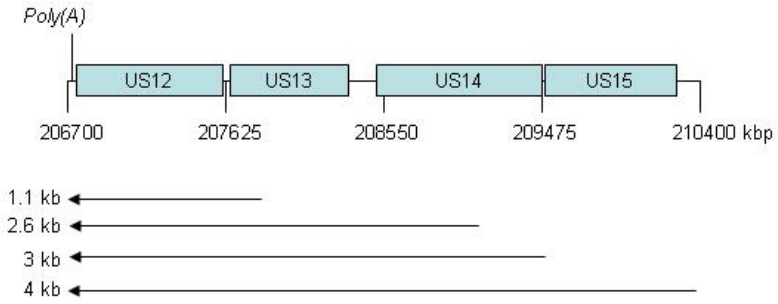


Figure 5.9.1. Summary diagram of the transcripts expressed from US12 to US15.

Based on northern blot data, it is suggested that transcripts are expressed from the start of US12 and from the start of US15, and both terminate at a poly(A) signal downstream of US12.

US12 gene family, the 4 kb and 1.1 kb band intensities were summed for each time point (h PI) plotted in the graph (Fig. 5.9C). It is clear from the northern blot (Fig. 5.9D) that the expression of both the 4 kb and 1.1 transcripts are much lower in RPE cells, and this is reflected in (Fig. 5.9C). Clearly, the microarray expression data (Fig. 5.9A) and the northern blot data (Fig. 5.9C) are in close agreement, confirming the down-regulation of transcripts binding the US12 probe in RPE cells. To further investigate which of the US12 gene family members whose transcripts bind the US12 gene probe is down-regulated will require detailed mapping of the 3′ and 5′ ends of the transcripts. The 4 kb transcript is likely to start at US15, therefore not taking into account genes US16 and US17; however, from its size, it is likely that the 1.1 kb transcript is specific for the US12 gene itself.

The microarray expression data for the second group of US12 gene family members i.e. those that bind the US18 probe (US18, US19 and US20) is shown in (Fig. 5.10A). All three cell types exhibited the same expression kinetics as those of the US12 group (US12, US13, US14, US15) (Fig. 5.9A). The expression kinetics for US18 obtained from the RT-qPCR data (Fig. 5.10B) yielded similar kinetic profiles to those obtained from the microarray for each cell type (Fig. 5.10A), except that in infected HFFF-2 cells, US18 expression decreased between 48 and 72 h PI. The northern blot (Fig. 5.10D) shows two major bands of 3.5 kb and 1.4 kb; the 3.5 kb band represents expression from the start of US20 through to a polyadenylation signal downstream of US18, and the 1.4 kb transcript represents expression from the start of US18 (Guo and Huang, 1993). The summed 3.5 and 1.4 kb band intensities at each time point are displayed in (Fig. 5.10C). The expression kinetics assessed by the northern blot (Fig. 5.10C) were similar to those obtained from the microarray (Fig. 5.10A) and RT-qPCR data (Fig. 5.10B). All three techniques agree that expression of US18 is lower in RPE cells then in HFFF-2 or U373Mg cells confirming the combined statistical tests that identified US18 as differentially expressed in RPE cells.

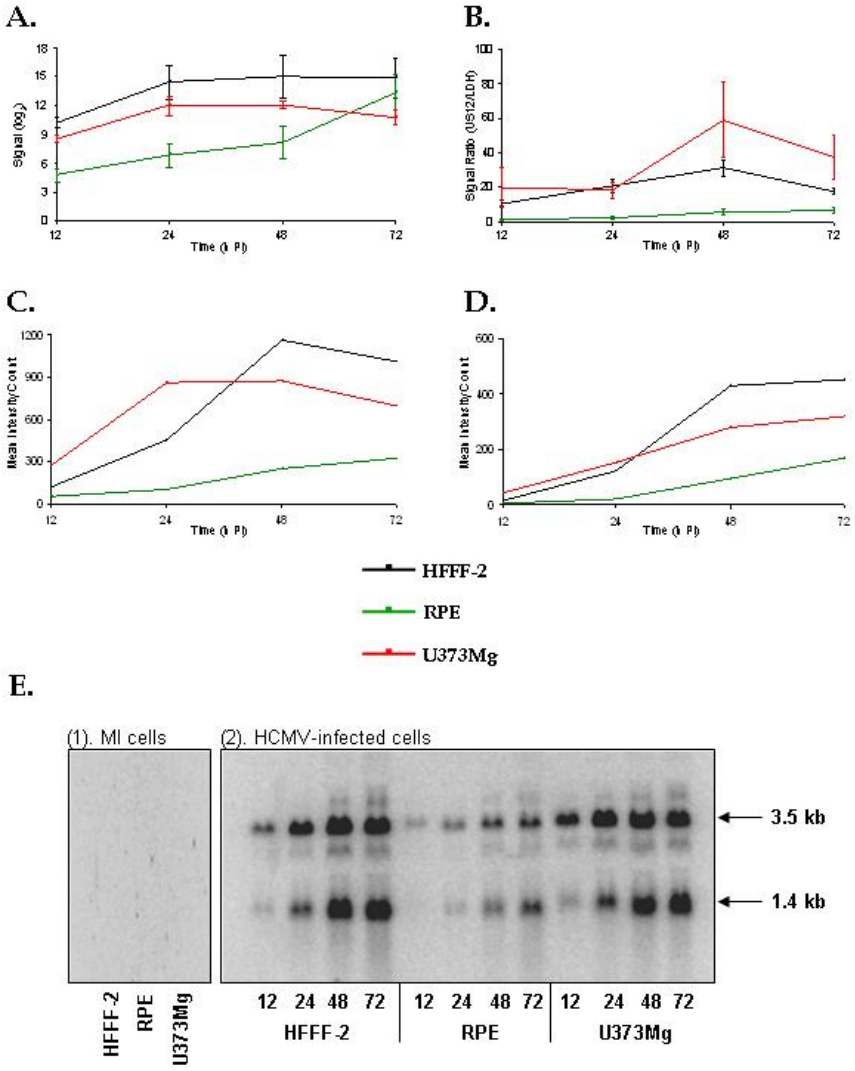


Figure 5.10. Comparison of microarray expression data with associated RT-qPCR and northern blot analysis.

(A) shows the microarray expression profile for US18, (B) shows the RT-qPCR expression profile, (C) shows the combined quantitated bands for a northern blot for US18, shown in (E), and (D) shows a graph of the US18 specific 1.4 kb bands in (E).

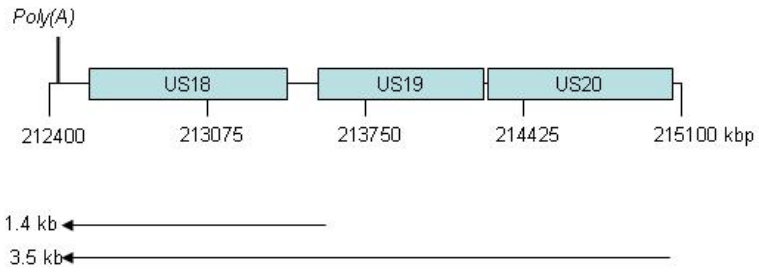


Figure 5.10.1. Summary diagram of transcripts expressed from US18 to US20.

Two major transcripts are expressed to a shared poly(A) signal downstream of US18. These transcripts have been also been mapped by Guo and Huang (1993).

5.5 Discussion

5.5.1 Genes differentially expressed in HFFF-2 and RPE cells

The majority of HCMV genes were not differentially expressed in HFFF-2 and RPE cells as determined by the combined statistical tests, and this was confirmed for representative IE, E and L kinetic class genes (UL123 (IE1), UL55 (gB), UL83 (pp65), UL99 (pp28), and UL43) by comparing the microarray expression profile data with RT-qPCR, northern and western blot data. All techniques supported the conclusions that the selected HCMV genes were non-differentially expressed in HFFF-2 and RPE cells.

Thirteen HCMV genes (UL4, UL16, UL45, IRS1, US11, US12, US13, US14, US18, US19 and US20) were differentially expressed in HFFF-2 and RPE cells (as assessed by the combined statistical tests at the 95 % significance level), and exhibited different expression levels and kinetics in the two cell types. Differentially expressed HCMV genes in HFFF-2 and RPE cells are considered specifically down-regulated in RPE cells; this was confirmed by RT-qPCR and/or northern blotting for UL4, UL16, IRS1, US12 and US18. Multiple transcripts were detected in the northern blot for HCMV genes US12 and US18 due to 3′ co-terminally expressed groups of the US12 family members. Assessment of US12 and US18 specific bands in the northern blots confirmed that these genes were down-regulated in RPE cells. Due to an internal transcription start site in HCMV gene IRS1 (Romanowski and Shenk, 1997), two distinct bands were obtained in northern blots. Both bands were significantly reduced in intensity in RPE cells confirming the microarray data indicating that IRS1 gene expression is down-regulated in RPE cells compared to HFFF-2 cells. When northern blot and/or RT-qPCR data were compared to HCMV infected HFFF-2 cells, the down-regulation of UL16 and UL4 was also confirmed in RPE cells. Thus, all techniques supported the conclusion that UL4, UL16, IRS1, US12 and US18 were differentially expressed in HFFF-2 and RPE cells.

5.5.2 Gene differentially expressed in HFFF-2 and U373Mg cells

Interpretation of the microarray gene expression data for HCMV infected U373Mg cells was made more difficult due to the unusual transcription profile of many HCMV genes, and because the dynamic range of expression over the time course is relatively small for most virus genes in this cell type. The microarray data for U373Mg cells found the majority of genes to be expressed from 12 h PI, irrespective of their kinetic class previously reported for infected HFF cells by Chambers et al. (1999). It is not known if the

rapid transit through the gene expression cascade results from a regulated acceleration in U373Mg cells, or whether there is a complete breakdown in the regulated cascade. It should be noted that according to the combined statistical tests, representative IE, E and L genes determined as differentially expressed in HFFF-2 and RPE cells were also found not to be differentially expressed in HFFF-2 and U373Mg cells. The status of UL123 (IE1), UL55 (gB), UL83 (pp65), UL99 (pp28) and UL43 as non-differentially expressed genes in HFFF-2 and U373Mg cells was confirmed by northern and western blotting; RT-qPCR data was inconclusive due to the elevated HCMV gene expression profiles obtained in U373Mg cells caused by the normalisation artefact (discussed previously). Importantly, expression of the late UL43 gene was detected by the microarray from 24 h PI in HFFF-2 and RPE cells, but from 12 h PI in U373Mg cells. This finding was confirmed by northern blotting, where UL43 transcripts were clearly expressed earlier (12 h PI) than in HFFF-2 and RPE cells (24 h PI). The microarray detected expression from most HCMV genes in U373Mg cells from 12 h PI, irrespective of the kinetic class of genes previously described in HFF cells by Chambers et al., 1999. This independent confirmation that the true-late UL43 gene is expressed as early as 12 h PI in U373Mg cells supports the conclusion that the HCMV transcriptome cascade is completed more rapidly in this cell line.

The statistical tests identified 26 genes as differentially expressed in HFFF-2 and U373Mg cells; however, removing genes in this list that are over-expressed and identification of genes that were down-regulated in U373Mg cells compared to HFFF-2 cells results in a more robust list of differentially expressed HCMV genes: UL4, IRS1, US12, US14, U18, US19 and US20. Comparisons of the microarray expression data for UL4, IRS1, US12 and US18 with northern blot data confirms that these genes are down-regulated in U373Mg cells compared to HFFF-2 cells. Interestingly, each of the genes down-regulated in U373Mg cells (UL4, IRS1, US12, US14, US18, US19 and US20) was also found to be down-regulated in RPE cells compared with HFFF-2 cells.

5.5.3 Conclusion

HCMV gene expression data generated by RT-qPCR and northern blotting were generally in accord with the expression kinetics identified by the microarray. The combined statistical tests, together with examination of the microarray gene expression profiles allows discrimination between non-differential and differentially regulated HCMV genes in different cell types. The HCMV (Merlin) microarray is therefore a robust tool for HCMV transcriptome profiling.

6 GENERAL DISCUSSION

6.1 Virus replication kinetics and the impact on gene expression

A custom DNA microarray was designed in order to profile the transcriptome activity of HCMV genes with the specific aim of identifying genes that were differentially expressed. The microarray design was based on the most up-to-date assessment of HCMV coding potential using a virus strain (Merlin) as close as possible to that of wild-type virus. The microarray was also utilised to assess transcription from regions of the HCMV genome recently proposed to code putative novel ORFs (Murphy et al., 2003a and b), and also regions of the genome now considered to be non-protein coding (Dolan et al., 2002). This thesis reports the first study to assess and compare the kinetics of HCMV gene expression in a time course experiment in multiple cell types in order to identify HCMV cell-type specific gene expression kinetics.

Previous HCMV microarray studies investigated the classification of HCMV genes according to their temporal (or kinetic class) of expression in HFF cells, using cycloheximide prior to and during viral infection to limit expression to the immediate-early (IE) class of genes, and gancyclovir from the time of viral infection to inhibit viral DNA synthesis to limit expression to the early (E) genes, while the absence of both drugs allowed expression through to the late (L) genes (Chambers et al., 1999). In contrast, this study investigated HCMV gene expression over a time course in multiple cell types (HFFF-2, RPE and U373Mg) in the absence of drugs. The use of three different cell types to investigate HCMV gene expression presented several challenges with respect to the differential replication kinetics of HCMV observed in these cell types. Several factors probably contribute to differential replication kinetics of HCMV in different cellular environments; different concentrations of specific cellular transcription factors and/or inhibitors of transcription within different cell types, and varying selective pressures resulting in the differential regulation of viral tropism factors. Ultimately, the degree of permissiveness of different cell types to HCMV infection can affect the interpretation of the virus gene expression data, if the virus transcription programme progresses at different rates. Transcriptome profiling of VZV (strain Dumas) was compared in two different cell types (human melanoma cell line, MeWo cells; and human atrocytoma cell line, SVG cells) at a single time point (72 h PI), when maximal CPE was observed in both cell lines, then compared the relative expression of VZV transcripts (see introduction, section 1.7.7). It was found that VZV gene expression was markedly reduced in infected SVG cells compare to infected MeWo cells and of the top six expressed VZV genes in

MeWo cells, only three (ORFs 49, 57 and 58), were also significantly expressed in infected SVG cells. It was concluded that the cellular environment did influence the expression of the viral transcriptome, but it was not possible to elaborate on cell-type specific expression kinetics as only one time point was examined (Kennedy et al., 2005). In contrast, Yamagishi et al., (2003) examined AcMNPV (autographa californica multicapsid nucleopolyhedrovirus) gene expression at 4 time points in two insect cell lines (Sf9; *S. frugiperda*; and TnHigh-Five; *T. ni*) (see introduction, section 1.7.7). They were able to identify 6 differentially expressed viral genes, one of which (AcMNPV gene p35) has anti-apoptotic function and is required for virus origin specific DNA replication, and a second (AcMNPV gene p10) is reported to be required for the release of virus from infected cells, the remaining genes have functions that are unknown. Data was not provided for AcMNPV growth kinetics in Sf9 or TnHigh-Five cells (Yamagishi et al., 2003).

This study compared HCMV differential gene expression in different cell types. In order to characterise the replication kinetics of HCMV in HFFF-2, RPE and U373Mg cells, infectious virus yields were measured in each cell type over a period of 7 days. The one step HCMV growth curves in RPE and U373Mg cells showed that there was a delay in infectious virus yield from the cells, resulting from effects on maturation and release of virus to the extracellular medium (Fig. 3.1 to 3.3). However, assessment of the protein expression kinetics in a time course experiment revealed that IE, E and L proteins were expressed at similar times although lower amounts of protein were made in RPE cells compared to HFFF-2 cells. This could be accounted for by an apparent delay in the exit from the eclipse phase in RPE cells, but the infection rises at a similar intracellular rate in HFFF-2 and RPE cells (Fig. 3.2), suggesting that the release of HCMV to the extracellular medium is impaired in RPE cells. The situation is similar in U373Mg cells compared to RPE cells, where there is an equivalent delay in the exit of virus from the eclipse phase (Fig. 3.1 and 3.3); however, unlike replication in RPE cells, the increase in infectious virus is slower and shorter in duration in U373Mg cells resulting in an approximately 1000-1500 fold lower infectious yield than in HFFF-2 cells. As is the case in RPE cells, IE, E and L protein expression kinetics are similar to those in HFFF-2 cells, but the release of infectious virus to the extracellular medium is also impaired in U373Mg cells (Fig. 3.2). In both RPE and U373Mg cells, the delay in the exit of the eclipse phase could be due to different mechanisms of HCMV entry into these two cell types and the subsequent delivery of nucleocapsids to the cell nucleus. It is known that HCMV enters RPE and HUVEC cells by endocytosis, whereas entry into fibroblasts occurs via fusion of the host and viral membranes (Bodaghi et al., 1999b). The entry mechanism of HCMV in U373Mg cells at present remains unknown. Immunofluorescence experiments showed that there

was no difference between the cell types in the numbers of cells infected and expressing UL44 (DNA polymerase processivity factor) at 48 h PI, but this does not exclude the possibility of delayed entry and/or delivery of nucleocapsids to the nucleus.

The microarray data showed that the temporal regulation of HCMV gene expression in U373Mg cells was perturbed compared to that in HFFF-2 or RPE cells. Moreover, infectious virus yields from U373Mg cells were the lowest of the cell types tested. These observations may be caused by rapid transit through, or breakdown in the regulation of the normal HCMV transcription cascade. Perturbed gene expression might be expected to impact on virus production and maturation, and could account for the short duration and low levels of infectious progeny made in U373Mg cells. One possible explanation for rapid transit through the HCMV gene expression cascade may be that the requirements for regulated HCMV gene expression are reduced in U373Mg cells. In fibroblasts, progression from immediate-early to early gene expression involves a complex interaction between the viral IE1 (72 kDa) and IE2 (86 kDa) proteins along with other viral and cellular gene products that provide gene specific accessory functions. In HCMV infected U373Mg cells, however, it has been reported that the IE2 (86 kDa) protein is the only viral gene product required to efficiently transactivate early and late viral gene promoters (Klucher et al., 1993; Wu et al., 1998; Wu et al., 2001), and that IE2 (86 kDa) can modulate the binding of the cellular transcription factor Sp1 in a cell type specific manner (Wu et al., 1998).

The finding that the cellular p53 gene is mutated in the U373Mg cell line (Van Meir et al., 1994) provides another explanation for the perturbation of the HCMV gene expression cascade. p53 plays a central role in several important cellular processes such as cell cycle regulation, DNA repair and apoptosis, and naturally occurring p53 mutations are associated with glioma oncogenesis (Louis et al., 1993). The effect of commonly encountered p53 mutations on viral gene promoter activity has been studied in transient transfection assays, and it was reported that the HCMV MIEP and HSV-1 UL9 promoters were activated by mutated p53 constructs. Viral gene activation induced by mutant p53 was enhanced in the presence of CREB, a cellular transcription factor induced by HCMV infection and which binds to the HCMV MIE promoter (Deb et al., 1992). It was also shown that the minimal promoter sequence required for activation by mutant p53 products was a functional TATA box (Deb et al., 1992). It may be then that, the mutant p53 in U373Mg cells binds all or most HCMV promoter sequences inducing early or deregulated expression from HCMV genes.

HCMV infection of fibroblasts, astrocytes and human umbilical vein endothelial cells results in elevated steady-state levels of (wild-type) p53 (Muganda et al., 1994; Lokensgard et al., 1999; Kovacs et al., 1996), but this is not associated with the activation of p53 responsive genes, or p53 activation of viral and cellular promoters (Subler et al., 1992; Tsai et al., 1996; Chen et al., 2001). Recently, an investigation into the ability of HCMV (strain Towne) to replicate in p53 negative fibroblasts and p53 positive fibroblasts revealed interesting observations that accord with the growth kinetics of HCMV in U373Mg cells presented here. It was shown that HCMV replication in p53 negative fibroblasts was delayed, with decreased infectious virus production and slow accumulation of viral DNA. In p53 negative fibroblasts, the onset of UL44 and pp28 protein expression was delayed compared to p53 positive fibroblasts (a phenomenon not observed in our U373Mg experiments). In p53 positive fibroblasts, p53 was sequestered at replication centres and enhanced the ability of HCMV to grow to high titres (Casavant et al., 2006).

The microarray data for U373Mg cells showed that HCMV early transcripts peaked at 24 h PI and late transcripts between 24 and 48 h PI, while in infected HFFF-2 and RPE cells, early transcripts peaked between 24 and 48 h PI and late transcripts between 48 and 72 h PI. In contrast, the temporal expression kinetics of viral proteins IE1 (72 kDa), and IE2 (86 kDa), pUS22, pp65, pp28 (representative of IE, IE, early, early/late and true late kinetic classes respectively) did not differ in the three cell types despite differences in transcript kinetics in U373Mg cells. This suggests some degree of post-transcriptional control over translation of transcripts in U373Mg cells, possibly operating at the level of nuclear export of mRNA or mRNA stability, or its binding to ribosomes and translation.

In summary then, it is speculated that in HCMV infected U373Mg cells, mutant p53 and IE2 (86 kDa) activate early promoters in the presence of specific cellular transcription factors e.g. CREB and Sp1, followed by late promoters that contain only TATA element. This leads to accelerated transition through the transcription cascade. Protein expression, however, is regulated, but the formation of viral replication centres is compromised by the absence of wild-type p53 (Casavant et al., 2006), resulting in a reduction of DNA replication and impairment of virus maturation and egress.

6.2 Viral genes specifying cell tropism factors

The HCMV Towne-BAC construct in which the US1 to US12 region of the genome have been replaced by the BAC sequence was used in a study to identify HCMV genes involved in cell tropism (Dunn et al., 2003). Transposon insertion mutagenesis provided a library of HCMV mutants covering most of the virus ORFs, which were then used to assess the replication efficiency of each mutant in HFF, RPE or HMVEC. Knockout mutations in genes UL10 or UL16 enhanced replication in RPE cells 100 fold, while a UL64 knockout (now considered to be non-protein coding) and US29 knockout mutations each decreased virus replication by 100 fold in RPE cells. While disruption of either the US16 and US29 ORFs enhanced replication by 100 fold in HMVEC, a UL24 knockout decreased replication by 100 fold (Dunn et al., 2003). In another study using the Towne-BAC, a US18 deletion mutant exhibited a 100 fold decrease in growth in cultured human gingival tissue compared to HFF cells (Hai et al., 2006). The finding that many HCMV gene products function to suppress virus replication is interesting and indicates the sophisticated level of control that HCMV exerts over its own replication, presumably to avoid immune surveillance.

The transposon mutant studies showed that the UL10, UL16, US29 ORFs, and sequences around the discounted UL64 ORF region may play a role in RPE cell tropism, while ORFs US16, US19 and UL24 may be involved in HMVEC tropism. US18, identified as differentially regulated gene in the microarray study, is a determinant of gingival tissue tropism (Hai et al., 2006). Since disruption of most of these ORFs resulted in a specific enhancement of virus replication in a single cell type, it might have been expected that during wild-type virus infection, the expression of such cell specific tropism factors would be down-regulated at the level of the promoter and would have been identified as differentially controlled genes in the microarray experiments. However, only one HCMV gene, UL16 identified as a determinant of RPE cell tropism, was found to be differentially expressed in the microarray experiments. In this case, there was a correlation between the findings that UL16 was down-regulated in infected RPE cells and that the knockout mutant exhibited a enhanced growth in RPE cells. The general lack of correlation between the list of HCMV genes identified as differentially expressed and the list of potential tropism factors is puzzling. In part, the discrepancy may be due to the use of the Towne-BAC virus to generate the mutants, since the Towne virus is itself a multiple mutant. The use of a transposon as a gene disruption tool may also have contributed. Because HCMV has many overlapping genes, genes that share a common poly(A) site or genes that are spliced, it may be that the transposon insertion has effects on the expression of upstream

or downstream genes that are real cell tropism factors. Another factor may be that the fact that the cell lines used in the two studies are separately sourced and derived from different lineages. Similar arguments regarding overlapping 3' co-terminal and spliced genes apply equally to the microarray study for those genes identified as differentially expressed, but whose true identity has not yet been confirmed by northern blotting.

A well-reported cell tropism factor that was not identified in the transposon based study or found to be differentially expressed in the microarray study is the function provided by the HCMV UL128 gene locus. This gene locus comprises three adjacently located ORFs UL128, UL130 and UL131A, whose products function co-operatively. Since the Towne strain has a premature termination mutation in gene UL130, and the Merlin strain has a termination mutation in UL128, the gene locus was non-functional in the viruses used in both studies (Akter et al., 2003; Dolan et al., 2004). While the UL128 mutation in Merlin would not be expected to impact on expression of the UL128 transcript *per se*, lack of function from the gene locus might eliminate the selective pressure to down-regulate expression from the gene locus in fibroblast cells. For reasons that are as yet unclear, expression of a fully functional UL128 gene locus is detrimental to growth in fibroblast cells and one or more of the UL128, UL130 and UL131A genes are invariably mutated in HCMV strains during passage in fibroblast cell cultures. In contrast, a fully functional UL128 gene locus is essential for propagation of HCMV in endothelial (HUVEC) cells (Hahn et al., 2004). How expression of the UL128 gene locus affects the replication of HCMV in different cell types *in vivo* at present remains unknown.

Transcriptome profiling of HCMV global gene expression in HFFF-2 and RPE cells revealed 13 HCMV genes that were differentially expressed during the time course (UL4, UL16, UL45, UL148, IRS1, US11, US12, US13, US14, US15, US18, US19 and US20). The application of combined statistical tests for HFFF-2 and U373Mg cells showed that 26 genes were expressed at significantly different levels ($p < 0.05$), although removing genes that were over-expressed in U373Mg cells compared to HFFF-2 cells reduced this list to 7 candidates (UL4, IRS1, US12, US14, US18, US19 and US20). That UL4 expression was down-regulated in both RPE and U373Mg cells compared to HFFF-2 cells was confirmed by comparison of UL4 northern blot expression data with the associated microarray data. UL4 is a glycoprotein that has an unusual mechanism of control over the translation of its protein product. The transcript leader of UL4 contains three transcription start sites, of which the second (designated as uORF2) is critical for the inhibition of the downstream cistron (Degnin et al., 1993). The uORF2 peptide is synthesised, but then retained on the ribosome blocking translation termination, stalling the ribosome on the transcript leader,

and restricting access to the downstream UL4 cistron (Cao and Geballe, 1995). With regard to UL4 as a candidate for differential gene expression, it has been shown that the UL4 promoter contains cis-acting elements for the cellular transcription factors NF-Y (that binds a CCAAT box) and Elk-1 (that binds to site 2), and activate UL4 gene expression along with IE2 (86 kDa) (Huang et al., 1994; Huang and Stinski, 1995; Chen and Stinski, 2000). It is known that HCMV infection of HFF cells activates ERK1/2 (MAPK) which has many substrates including AP-1, CREB, STAT proteins, SAP-1a and Elk-1. UL4 expression is influenced by cellular transcription factors (Elk-1) that are activated by the MAPK/ERK pathway (Chen and Stinski, 2000). As the function of UL4 is unknown, the significance of UL4 differential expression in these three cell types is unclear. The mechanism by which UL4 expression is controlled must involve the ability of Elk-1 and NF-Y to promote expression of this gene via upstream cis-acting regulatory elements. It is possible that these transcription factors are present in different concentrations in the three cell types, or there is a delay in their activation in RPE cells and/or U373Mg cells, that subsequently affects the expression from the UL4 promoter. Because UL4 appeared to be down-regulated in both RPE and U373Mg cells, it is equally viable to suggest that UL4 expression is up-regulated in HFFF-2 cells.

UL16 and US11 were both down-regulated in RPE cells compared to HFFF-2 cells, and both provide immune evasion functions. The UL16 gene codes for a glycoprotein dispensable for growth in cell culture (Kaye et al., 1992), but which binds to and retains the NK cell receptor ligands ULBP1, ULBP2 and MIC-B in the ER preventing their expression on the cell surface, thus avoiding NK cell activation (Cosman, 2001; Welte et al., 2003). Compared to HFFF-2 cells, UL16 gene expression was significantly reduced in RPE cells throughout the time course. It has been shown that a Towne-BAC ΔUL16 knockout mutant specifically enhanced replication in RPE cells by 100 fold compared to the parental HCMV Towne-BAC virus (Dunn et al., 2003). This suggests that high levels of UL16 expression are detrimental to growth and replication in RPE cells supporting the finding that UL16 is down-regulated in RPE cells. The expression of the US11 gene was also down-regulated in RPE cells compared to HFFF-2 cells (this observation was confirmed by northern blotting: data not shown). Interestingly, US2 was also a possible candidate for down-regulation in RPE cells compared to HFFF-2, although the data was not conclusive (see section 4.2.7.2). Both US2 and US11 are responsible for the targeting of MHC Class I heavy chains (HC) for ubiquitin-mediated proteasome degradation, thus preventing the activation of CD8+ CTLs (Wiertz, 1996a and b). Despite the down-regulation of US11, UL16, and possibly US2, there are many more HCMV genes whose products are involved in immune evasion, e.g. HCMV UL18; homologue for HLA-I

(Cosman, 1997), gpUL40; up-regulation of HLA-E (Tomasec et al., 2000), UL141; down-regulation of CD155 (Tomasec, 2005), UL142 inhibition of NK cell activation (Wills et al., 2005), and US3, US6, US9 and US10; inhibition of MHC Class I surface expression in HCMV infected cells (Loureiro and Ploegh, 2006), whose expression was not affected by growth in different cell types. The expression of HCMV gene UL18, a homologue of the human leukocyte antigen (HLA class I) protects against NK lysis of HCMV infected fibroblasts (Cosman, 1997), while the induction of CD8+ CTLs is avoided by the down-regulation of MHC class I molecules conferred by US2, US3, US6, US9, US10 and US11 in fibroblasts (Loureiro and Ploegh, 2006). However, it was reported that NK cytotoxicity of endothelial cells and macrophages is independent of the expression of UL18 or cellular HLA class I A, B and C antigens, because a ΔUL18 HCMV mutant, and a mutant lacking HCMV genes US1 to US9 and US11 (HCMV strain RV670) were both less susceptible to NK lysis in these two cell types than in fibroblasts (Odeberg et al., 2002). Based on this evidence, it is possible that the immune evasion functions provided by UL16 and US11 in HFFF-2 cells are not essential in RPE cells, and that resistance to NK and/or CD8+ CTL activity in RPE cells is independent of the expression of UL16 and US11.

A possible explanation for the down-regulation of some HCMV genes involved in immune evasion might be due to the fact that RPE cells located in the eye are an immune-privileged site *in vivo*, and this might alter their characteristics as antigen presenting cells (APCs) (Gabrielian et al., 1994). In contrast to HCMV infection in fibroblasts, infection of RPE cells circumvents the stimulation of NF-κB in order to avoid cell-mediated inflammatory mechanisms induced by this transcription factor. Clearly, the down-regulation of US11 and UL16 (and possibly US2) in RPE cells is part of a much more complicated network of immune effector mechanisms, for which cross-talk between these pathways and the possibility of antagonistic interactions and/or hierarchy within immune-privileged RPE cells remains unclear. Furthermore, as UL16 function is not required in RPE cells, this suggests that UL16 might have another function other than immune evasion, especially as Towne-BAC ΔUL16 replication was dramatically enhanced compared to parent Towne-BAC in RPE cells (Cosman, 2001; Welte et al., 2003; Dunn et al., 2003). Similarly, the US11 gene might also have an additional function that is detrimental to growth in RPE cells. Overexpression of these genes (US11 and UL16) in RPE cells would be required to gain understanding of their effects on the growth of HCMV in this cell type.

The expression of UL45 was down-regulated in RPE cells compared to HFFF-2 cells, as evidenced by examination of UL45 expression kinetics and the combined statistical test p-

values ($p < 0.05$). This gene was not found to be differentially expressed in HFFF-2 and U373Mg cells. UL45 is a tegument protein that is related to the cellular ribonucleotide reductase (RNR) large subunit (R1), and is reported to be dispensable for growth in HFF and endothelial cells (Hahn et al., 2002). However, the MCMV homologue (M45) of HCMV UL45 has been reported to be a determinant of MCMV endothelial cell tropism, and indispensable for virus growth and pathogenesis *in vivo* (Brune et al., 2001; Lembo et al., 2000; Lembo et al., 2004). The MCMV M45 product exhibits strong anti-apoptotic activity (Brune et al., 2001). Assessment of the growth of a HCMV UL45-KO mutant in HFF cells displayed a growth defect at low m.o.i. and its ability for cell-to-cell spread was diminished. The UL45 gene product displayed mild anti-apoptotic activity (Patrone et al., 2003). It is possible that HCMV UL45 plays a role in RPE cell tropism; however, in contrast to the studies of M45 in MCMV where it is reported to be essential for virus growth, UL45 was down-regulated in RPE cells. As it has been reported that UL45 is involved in fibroblast cell-to-cell spread at low m.o.i. (Patrone et al., 2003), it is possible that UL45 down-regulation in RPE cells might play in role in limiting the spread of HCMV within the retinal epithelium. This hypothesis is supported by the identification of several HCMV genes that appear to moderate the replication and/or spread of HCMV in a cell type specific manner (Dunn et al., 2003).

The US12 gene family were over-represented in the gene lists generated by the combined statistical tests between HFFF-2 and RPE cells, and between HFFF-2 and U373Mg cells. As discussed previously, these 10 genes (US12 to US20) form two distinct 3′ co-terminal groups. However, expression kinetics of HCMV genes US16 and US17 were the same in HFFF-2 and RPE cells, and overall expression was not considered significantly different in combined statistical tests for HFFF-2 and RPE, or HFFF-2 and U373Mg cells. For the remaining genes (US12 to US15; US18 to US20), the microarray data was supported by RT-qPCR and northern blot data (specifically for US12 and US18). The fact that these genes are 3′ co-terminal explains why so many family members appear to be differentially expressed. The US12 family are putative multiple transmembrane proteins (Rigoutsos et al., 2003). The localisations of US14, US17 and US18 have been studied in HFFs with the following observations: 1. US14 is distributed throughout the cytoplasm, but is occasionally found concentrated at virus assembly compartments (AC) in the cytoplasm; 2. US17 is expressed in a segmented manner (cleaved post-transcriptionally) with the N-terminus localising at the periphery of AC, and the C-terminus localising in the nucleus and cytoplasm; 3. US18 is distributed throughout the cytoplasm but localised at ACs at late stages of infection (Das et al., 2006; Das and Pellett, 2007). It has therefore been suggested that these gene products have roles in virus maturation and egress (Das and

Pellett, 2007). The down-regulation of ORFs US12 to US15 and US18 to US20 in RPE cells, and the down-regulation of US12, US14, and US18-US20 in U373Mg cells may impact on virus maturation and egress because it is apparent from the cell-released virus one step growth curves that virus egress is strictly limited in these cell types (Fig. 3.2). The down-regulation of these genes could be important in controlling cell-to-cell spread of infection in these cell types, promoting direct membrane-to-membrane route of egress rather than release of cell-free virus to the bloodstream or extracellular spaces. It is possible that different mechanisms of virus dissemination are appropriate in different organs and tissues.

IRS1 is a tegument protein that is dispensable for growth in fibroblast cell culture, but which cooperates with IE1/IE2 in the transctivation of early and late viral promoters (Jones and Muzithras, 1992; Pari et al., 2000). IRS1 was down-regulated in RPE and U373Mg cells compared to HFFF-2 cells, and this was confirmed by northern blotting. IRS1 and TRS1 are closely related genes that work synergistically with other transactivators (IE1/IE2) and immediate-early gene regulators (UL36-38 and UL69) for activation of early and late gene expression in infected HFF cells. It has also been shown that IRS1 or TRS1 are independently capable of initiating ori-Lyt dependent HCMV DNA replication together with UL36-38 and UL112-113 genes (Iskenderian et al., 1996). Recent studies have shown that IRS1 and TRS1 are protein kinase R (PKR) evasion genes. TRS1 sequesters PKR in the nucleus of infected cells, preventing interaction of PKR with cytoplasmic ds-RNA. This prevents the activation of PKR and the subsequent phosphorylation of elongation initiation factor-2-alpha (eIF2α), therefore preventing host protein synthesis shutoff (Hakki et al., 2005; Hakki et al., 2006). These data also suggest that IRS1 expression is not essential in the presence of TRS1. While TRS1 is not differentially expressed in HFFF-2 and RPE, or U373Mg cells, IRS1 was confirmed as differentially expressed in HFFF-2, RPE and U373Mg cells. Although related IRS1 and TRS-1 have different sequences at their 3'-ends, IRS1 gene expression is autoregulated by a protein that is expressed from an internal transcription start site within IRS1, and this protein negatively regulates the expression of the full length IRS1 transcript (Romanowski and Shenk, 1997). However, TRS1 does not exhibit a similar mechanism of control. The down-regulation of IRS1 in RPE cells and U373Mg cells suggests that IRS1 might have a unique function within these two cell types. Alternatively, since IRS1 was under-expressed in both RPE and U373Mg cells, it may be that IRS1 was up-regulated in HFFF-2 cells.

To summarise the above findings, HCMV genes with immune evasion or transactivation functions were down-regulated in RPE cells (UL16, US11, UL45 and IRS1), and U373Mg cells (IRS1), while HCMV genes that might have roles in virus maturation and egress were also down-regulated in RPE cells (US12-US15, US18-US20) and U373Mg cells (US12, US14 and US18-US20). The finding that the same viral genes are differentially expressed in both RPE and U373Mg cells compared to HFFF-2 cells is particularly relevant as both RPE and U373Mg cells are derived from immune privileged sites. HCMV infection of these two cell types may be subject to a more controlled replication strategy resulting in the dissemination of infection in the surrounding tissue via cell-to-cell junctions. This would contribute to immune evasion as it would prevent a large burst of virus to the extracellular space, leading to a delay the stimulation of the immune system. Moreover, genes that are essential for HCMV replication and envelopment are not affected in HFFF-2, RPE or U373Mg cells e.g. IE1/IE2 (Mocarski et al., 1996), pUL44 (Leach and Mocarski, 1989), and pUL99 (pp28) (Silva et al., 2003). In contrast, fibroblasts have a basic function in the formation of connective tissue within the body. HCMV replication *in vivo* in fibroblasts results in the quick release of virus into the blood stream that is disseminated throughout the body either as cell-free virus or by transient transfer to circulating neutrophils/monocytes. It is known that HCMV infection of the kidneys results in the release of HCMV into the urine, which ultimately contributes to the maintenance of HCMV within the human population. As fibroblasts are not major components of immune privileged organs, the virus will express all genes that function as immune evasins, to protect the infected cells from the host's immune response. High yields of cell-free infectious progeny from fibroblasts cells during a primary HCMV infection may be required to ensure establishment of persistence in the host, and to promote person-to-person spread of the virus via host secretions.

In order to fully understand the roles of individual differentially expressed HCMV genes in cell tropism or immune evasion requires further investigation. However, as the data stands, it can be concluded that certain HCMV genes are differentially regulated in HFFF-2, RPE and U373Mg cells. The mechanism by which differential expression of HCMV genes are affected is unclear, but may operate at the level of the promoter, polyadenylation and nuclear export of mRNA, or the stability of mRNAs. It is clear from this and other studies that the outcome of HCMV infection in a given cell type is a function of both the cell type and the virus itself. The HCMV microarray platform we have employed has provided valuable insight regarding the expression kinetics of HCMV genes in different cellular environments.

6.3 Future work on HCMV cell tropism

This is the first investigation to examine the temporal kinetics of HCMV global gene expression in different human cell types: fibroblasts, epithelial cells and astrocytes. The main conclusions from this study were that the majority of HCMV genes were expressed with similar kinetics in infected HFFF-2 and RPE cells, however, 13 genes (UL4, UL45, UL148, IRS1, US11, US12, US13, US14, US15, US18, US19 and US20) were expressed with cell-type specific kinetics. In HCMV infected U373Mg cells, the viral gene expression cascade appeared to be accelerated compared to that observed in HFFF-2 and RPE cells, probably as a consequence of a mutation in p53 in U373Mg cells. However, it was concluded that 7 HCMV genes (UL4, IRS1, US12, US14, US18, US19 and US20) were differentially expressed in U373Mg cells compared to HFFF-2 cells, and that each of the 7 genes differentially expressed in U373Mg cells were also differentially expressed in RPE cells.

Future studies might investigate the effects of specifically up- or down-regulating the expression of individual differentially expressed genes in order to examine their effects on the replication of HCMV in each cell type. This might be accomplished using interfering RNA (RNAi) technology to silence the expression of specific genes in order to examine the effects on the growth of the virus. Similarly, cells stably expressing specific HCMV genes under the control of an inducible promoter could also be engineered to examine their effects on the growth of the appropriate null mutant virus. Furthermore, antibodies could be raised against gene products that have been identified as differentially expressed in order to examine the intracellular distribution and/or potential interactions of the differentially expressed gene products with other viral or cellular proteins. It would also be interesting to examine the rate of HCMV DNA replication, and to examine virus maturation and egress in HFFF-2, RPE and U373Mg cells in order to gain further insight into the growth kinetics of HCMV in these cell types. Quantitative PCR could be used to measure the accumulation of HCMV genomic DNA in each cell type, while electron transmission microscopy could be used to examine the intracellular accumulation of virus particles. These techniques could highlight different strategies or mechanisms employed by HCMV when replicating in different cellular environments. Together, these studies will allow greater understanding of the complex nature of HCMV cell tropism.

7 REFERENCES

Abate, D. A., S. Watanabe, and E. S. Mocarski. 2004. Major human cytomegalovirus structural protein pp65 (ppUL83) prevents interferon response factor 3 activation in the interferon response. J Virol 78:10995-1006.

Adair, R., E. R. Douglas, J. B. Maclean, S. Y. Graham, J. D. Aitken, F. E. Jamieson, and D. J. Dargan. 2002. The products of human cytomegalovirus genes UL23, UL24, UL43 and US22 are tegument components. J Gen Virol 83:1315-1324.

Adair, R., G. W. Liebisch, and A. M. Colberg-Poley. 2003. Complex alternative processing of human cytomegalovirus UL37 pre-mRNA. Journal of General Virology 84:3353-3358.

Adler, B., L. Scrivano, Z. Ruzcics, B. Rupp, C. Sinzger, and U. Koszinowski. 2006. Role of human cytomegalovirus UL131A in cell type-specific virus entry and release. J Gen Virol 87:2451-60.

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 348:233-41.

Aguilar, J. S., P. Ghazal, and E. K. Wagner. 2005. Design of a herpes simplex virus type 2 long oligonucleotide-based microarray: global analysis of HSV-2 transcript abundance during productive infection. Methods Mol Biol 292:423-48.

Ahn, J. H., W. J. Jang, and G. S. Hayward. 1999. The human cytomegalovirus IE2 and UL112-113 proteins accumulate in viral DNA replication compartments that initiate from the periphery of promyelocytic leukemia protein-associated nuclear bodies (PODs or ND10). J Virol 73:10458-71.

Ahn, J. H., Y. Xu, W. J. Jang, M. J. Matunis, and G. S. Hayward. 2001. Evaluation of interactions of human cytomegalovirus immediate-early IE2 regulatory protein with small ubiquitin-like modifiers and their conjugation enzyme Ubc9. J Virol 75:3859-72.

Ahn, J. W., K. L. Powell, P. Kellam, and D. G. Alber. 2002. Gammaherpesvirus lytic gene expression as characterized by DNA array. J Virol 76:6244-56.

Akter, P., C. Cunningham, B. P. McSharry, A. Dolan, C. Addison, D. J. Dargan, A. F. Hassan-Walker, V. C. Emery, P. D. Griffiths, G. W. Wilkinson, and A. J. Davison. 2003. Two novel spliced genes in human cytomegalovirus. J Gen Virol 84:1117-22.

Alderete, J. P., S. J. Child, and A. P. Geballe. 2001. Abundant early expression of gpUL4 from a human cytomegalovirus mutant lacking a repressive upstream open reading frame. J Virol 75:7188-92.

Alderete, J. P., S. Jarrahian, and A. P. Geballe. 1999. Translational effects of mutations and polymorphisms in a repressive upstream open reading frame of the human cytomegalovirus UL4 gene. J Virol 73:8330-7.

Arav-Boger, R., R. E. Willoughby, R. F. Pass, J. C. Zong, W. J. Jang, D. Alcendor, and G. S. Hayward. 2002. Polymorphisms of the cytomegalovirus (CMV)-encoded tumor necrosis factor-alpha and beta-chemokine receptors in congenital CMV disease. J Infect Dis 186:1057-64.

Arlt, H., D. Lang, S. Gebert, and T. Stamminger. 1994. Identification of binding sites for the 86-kilodalton IE2 protein of human cytomegalovirus within an IE2-responsive viral early promoter. J Virol 68:4117-25.

Arnoult, D., L. M. Bartle, A. Skaletskaya, D. Poncet, N. Zamzami, P. U. Park, J. Sharpe, R. J. Youle, and V. S. Goldmacher. 2004. Cytomegalovirus cell death suppressor vMIA blocks Bax- but not Bak-mediated apoptosis by binding and sequestering Bax at mitochondria. Proc Natl Acad Sci U S A 101:7988-93.

Asmar, J., L. Wiebusch, M. Truss, and C. Hagemeier. 2004. The putative zinc finger of the human cytomegalovirus IE2 86-kilodalton protein is dispensable for DNA binding and autorepression, thereby demarcating a concise core domain in the C terminus of the protein. J Virol 78:11853-64.

Awasthi, S., J. A. Isler, and J. C. Alwine. 2004. Analysis of splice variants of the immediate-early 1 region of human cytomegalovirus. J Virol 78:8191-200.

Badie, B., C. S. Goh, J. Klaver, H. Herweijer, and D. A. Boothman. 1999. Combined radiation and p53 gene therapy of malignant glioma cells. Cancer Gene Ther 6:155-62.

Baillie, J., D. A. Sahlender, and J. H. Sinclair. 2003. Human cytomegalovirus infection inhibits tumor necrosis factor alpha (TNF-alpha) signaling by targeting the 55-kilodalton TNF-alpha receptor. J Virol 77:7007-16.

Baxter, M. K., and W. Gibson. 2001. Cytomegalovirus basic phosphoprotein (pUL32) binds to capsids in vitro through its amino one-third. J Virol 75:6865-73.

Bechtel, J. T., and T. Shenk. 2002. Human cytomegalovirus UL47 tegument protein functions after entry and before immediate-early gene expression. J Virol 76:1043-50.

Bhella, D., F. J. Rixon, and D. J. Dargan. 2000. Cryomicroscopy of human cytomegalovirus virions reveals more densely packed genomic DNA than in herpes simplex virus type 1. J Mol Biol 295:155-61.

Bilban, M., L. K. Buehler, S. Head, G. Desoye, and V. Quaranta. 2002. Defining signal thresholds in DNA microarrays: exemplary application for invasive cancer. BMC Genomics 3:19.

Bodaghi, B., M. E. Slobbe-van Drunen, A. Topilko, E. Perret, R. C. Vossen, M. C. van Dam-Mieras, D. Zipeto, J. L. Virelizier, P. LeHoang, C. A. Bruggeman, and S. Michelson. 1999. Entry of human cytomegalovirus into retinal pigment epithelial and endothelial cells by endocytosis. Invest Ophthalmol Vis Sci 40:2598-607.

Boeckh, M., and R. Bowden. 1995. Cytomegalovirus infection in marrow transplantation. Cancer Treat Res 76:97-136.

Bolovan-Fritts, C., and J. A. Wiedeman. 2001. Human cytomegalovirus strain Toledo lacks a virus-encoded tropism factor required for infection of aortic endothelial cells. J Infect Dis 184:1252-61.

Borst, E. M., G. Hahn, U. H. Koszinowski, and M. Messerle. 1999. Cloning of the human cytomegalovirus (HCMV) genome as an infectious bacterial artificial chromosome in Escherichia coli: a new approach for construction of HCMV mutants. J Virol 73:8320-9.

Borst, E. M., S. Mathys, M. Wagner, W. Muranyi, and M. Messerle. 2001. Genetic evidence of an essential role for cytomegalovirus small capsid protein in viral growth. J Virol 75:1450-8.

Boyle, K. A., and T. Compton. 1998. Receptor-binding properties of a soluble form of human cytomegalovirus glycoprotein B. J Virol 72:1826-33.

Browne, E. P., and T. Shenk. 2003. Human cytomegalovirus UL83-coded pp65 virion protein inhibits antiviral gene expression in infected cells. Proc Natl Acad Sci U S A 100:11439-44.

Browne, E. P., B. Wing, D. Coleman, and T. Shenk. 2001. Altered cellular mRNA levels in human cytomegalovirus-infected fibroblasts: viral block to the accumulation of antiviral mRNAs. J Virol 75:12319-30.

Brune, W., C. Menard, J. Heesemann, and U. H. Koszinowski. 2001. A ribonucleotide reductase homolog of cytomegalovirus and endothelial cell tropism. Science 291:303-5.

Buser, C., P. Walther, T. Mertens, and D. Michel. 2007. Cytomegalovirus Primary Envelopment Occurs at Large Infoldings of the Inner Nuclear Membrane. J. Virol. 81:3042-3048.

Butcher, S. J., J. Aitken, J. Mitchell, B. Gowen, and D. J. Dargan. 1998. Structure of the Human Cytomegalovirus B Capsid by Electron Cryomicroscopy and Image Reconstruction. Journal of Structural Biology 124:70-76.

Cao, J. H., and A. P. Geballe. 1995. Translational Inhibition by a Human Cytomegalovirus Upstream Open Reading Frame Despite Inefficient Utilization of Its Aug Codon. Journal of Virology 69:1030-1036.

Casavant, N. C., M. H. Luo, K. Rosenke, T. Winegardner, A. Zurawska, and E. A. Fortunato. 2006. Potential role for p53 in the permissive life cycle of human cytomegalovirus. J Virol 80:8390-401.

Cassady, K. A. 2005. Human cytomegalovirus TRS1 and IRS1 gene products block the double-stranded-RNA-activated host protein shutoff response induced by herpes simplex virus type 1 infection. J Virol 79:8707-15.

Caswell, R., C. Hagemeier, C. J. Chiou, G. Hayward, T. Kouzarides, and J. Sinclair. 1993. The human cytomegalovirus 86K immediate early (IE) 2 protein requires the basic region of the TATA-box binding protein (TBP) for binding, and interacts with TBP and transcription factor TFIIB via regions of IE2 required for transcriptional regulation. J Gen Virol 74 (Pt 12):2691-8.

Cha, T. A., E. Tom, G. W. Kemble, G. M. Duke, E. S. Mocarski, and R. R. Spaete. 1996. Human cytomegalovirus clinical isolates carry at least 19 genes not found in laboratory strains. J Virol 70:78-83.

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 73:5757-66.

Champier, G., S. Hantz, A. Couvreux, S. Stuppfler, M. C. Mazeron, S. Bouaziz, F. Denis, and S. Alain. 2007. New functional domains of human cytomegalovirus pUL89 predicted

by sequence analysis and three-dimensional modelling of the catalytic site DEXDc. Antivir Ther 12:217-32.

Chang, C. P., C. L. Malone, and M. F. Stinski. 1989a. A human cytomegalovirus early gene has three inducible promoters that are regulated differentially at various times after infection. J Virol 63:281-90.

Chang, C. P., D. H. Vesole, J. Nelson, M. B. Oldstone, and M. F. Stinski. 1989b. Identification and expression of a human cytomegalovirus early glycoprotein. J Virol 63:3330-7.

Che, X., L. Zerboni, M. H. Sommer, and A. M. Arvin. 2006. Varicella-zoster virus open reading frame 10 is a virulence determinant in skin cells but not in T cells in vivo. J Virol 80:3238-48.

Chee, M. S., A. T. Bankier, S. Beck, R. Bohni, C. M. Brown, R. Cerny, T. Horsnell, C. A. Hutchison, 3rd, T. Kouzarides, J. A. Martignetti, and et al. 1990. Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169. Curr Top Microbiol Immunol 154:125-69.

Chen, D.-H., J. Jakana, D. McNab, J. Mitchell, Z. H. Zhou, M. Dougherty, W. Chiu, and F. J. Rixon. 2001. The Pattern of Tegument-Capsid Interaction in the Herpes Simplex Virus Type 1 Virion Is Not Influenced by the Small Hexon-Associated Protein VP26 10.1128/JVI.75.23.11863-11867.2001. J. Virol. 75:11863-11867.

Chen, D. H., H. Jiang, M. Lee, F. Liu, and Z. H. Zhou. 1999. Three-Dimensional Visualization of Tegument/Capsid Interactions in the Intact Human Cytomegalovirus. Virology 260:10-16.

Chen, J., and M. F. Stinski. 2000. Activation of transcription of the human cytomegalovirus early UL4 promoter by the Ets transcription factor binding element. J Virol 74:9845-57.

Chen, Z., E. Knutson, A. Kurosky, and T. Albrecht. 2001. Degradation of p21cip1 in cells productively infected with human cytomegalovirus. J Virol 75:3613-25.

Cherrington, J. M., and E. S. Mocarski. 1989. Human cytomegalovirus ie1 transactivates the alpha promoter-enhancer via an 18-base-pair repeat element. J Virol 63:1435-40.

Cheung, T. W., and S. A. Teich. 1999. Cytomegalovirus infection in patients with HIV infection. Mt Sinai J Med 66:113-24.

Chin, K. V., K. Ueda, I. Pastan, and M. M. Gottesman. 1992. Modulation of activity of the promoter of the human MDR1 gene by Ras and p53. Science 255:459-62.

Choi, K. S., S. J. Kim, and S. Kim. 1995. The retinoblastoma gene product negatively regulates transcriptional activation mediated by the human cytomegalovirus IE2 protein. Virology 208:450-6.

Chou, S. W., and K. M. Dennison. 1991. Analysis of interstrain variation in cytomegalovirus glycoprotein B sequences encoding neutralization-related epitopes. J Infect Dis 163:1229-34.

Cobbs, C. S., L. Harkins, M. Samanta, G. Y. Gillespie, S. Bharara, P. H. King, L. B. Nabors, C. G. Cobbs, and W. J. Britt. 2002. Human cytomegalovirus infection and expression in human malignant glioma. Cancer Res 62:3347-50.

Compton, T., R. R. Nepomuceno, and D. M. Nowlin. 1992. Human cytomegalovirus penetrates host cells by pH-independent fusion at the cell surface. Virology 191:387-95.

Cosman, D., N. Fanger, L. Borges, M. Kubin, W. Chin, L. Peterson, and M. L. Hsu. 1997. A novel immunoglobulin superfamily receptor for cellular and viral MHC class I molecules. Immunity 7:273-82.

Cosman, D., J. Mullberg, C. L. Sutherland, W. Chin, R. Armitage, W. Fanslow, M. Kubin, and N. J. Chalupny. 2001. ULBPs, novel MHC class I-related molecules, bind to CMV glycoprotein UL16 and stimulate NK cytotoxicity through the NKG2D receptor. Immunity 14:123-33.

Costa, R. H., B. G. Devi, K. P. Anderson, B. H. Gaylord, and E. K. Wagner. 1981. Characterization of a major late herpes simplex virus type 1 mRNA. J Virol 38:483-96.

Dargan, D. J., F. E. Jamieson, J. MacLean, A. Dolan, C. Addison, and D. J. McGeoch. 1997. The published DNA sequence of human cytomegalovirus strain AD169 lacks 929 base pairs affecting genes UL42 and UL43. J Virol 71:9833-6.

Das, S., Y. Skomorovska-Prokvolit, F.-Z. Wang, and P. E. Pellett. 2006. Infection-Dependent Nuclear Localization of US17, a Member of the US12 Family of Human Cytomegalovirus-Encoded Seven-Transmembrane Proteins. J. Virol. 80:1191-1203.

Das, S. and P. E. Pellett. 2007. Members of the HCMV US12 family of predicted heptaspanning membrane proteins have unique intracellular distributions, including association with the cytoplasmic virion assembly complex. Virology. 361:263-73.

Davison, A. J., A. Dolan, P. Akter, C. Addison, D. J. Dargan, D. J. Alcendor, D. J. McGeoch, and G. S. Hayward. 2003. The human cytomegalovirus genome revisited: comparison with the chimpanzee cytomegalovirus genome. J Gen Virol 84:17-28.

Deb, S., C. T. Jackson, M. A. Subler, and D. W. Martin. 1992. Modulation of cellular and viral promoters by mutant human p53 proteins found in tumor cells. J Virol 66:6164-70.

DeFilippis, V., and K. Fruh. 2005. Rhesus cytomegalovirus particles prevent activation of interferon regulatory factor 3. J Virol 79:6419-31.

Degnin, C. R., M. R. Schleiss, J. H. Cao, and A. P. Geballe. 1993. Translational Inhibition Mediated by a Short Upstream Open Reading Frame in the Human Cytomegalovirus Gpul4 (Gp48) Transcript. Journal of Virology 67:5514-5521.

Depto, A. S., and R. M. Stenberg. 1992. Functional analysis of the true late human cytomegalovirus pp28 upstream promoter: cis-acting elements and viral trans-acting proteins necessary for promoter activation. J Virol 66:3241-6.

Desai, P., N. A. DeLuca, and S. Person. 1998. Herpes simplex virus type 1 VP26 is not essential for replication in cell culture but influences production of infectious virus in the nervous system of infected mice. Virology 247:115-24.

Diefenbach, R. J., M. Miranda-Saksena, E. Diefenbach, D. J. Holland, R. A. Boadle, P. J. Armati, and A. L. Cunningham. 2002. Herpes simplex virus tegument protein US11 interacts with conventional kinesin heavy chain. J Virol 76:3282-91.

Dittmer, A., J. C. Drach, L. B. Townsend, A. Fischer, and E. Bogner. 2005. Interaction of the putative human cytomegalovirus portal protein pUL104 with the large terminase subunit pUL56 and its inhibition by benzimidazole-D-ribonucleosides. J Virol 79:14660-7.

Dolan, A., C. Cunningham, R. D. Hector, A. F. Hassan-Walker, L. Lee, C. Addison, D. J. Dargan, D. J. McGeoch, D. Gatherer, V. C. Emery, P. D. Griffiths, C. Sinzger, B. P. McSharry, G. W. Wilkinson, and A. J. Davison. 2004. Genetic content of wild-type human cytomegalovirus. J Gen Virol 85:1301-12.

Dunn, W., C. Chou, H. Li, R. Hai, D. Patterson, V. Stolc, H. Zhu, and F. Liu. 2003. Functional profiling of a human cytomegalovirus genome. Proc Natl Acad Sci U S A 100:14223-8.

Durbin, B. P., J. S. Hardin, D. M. Hawkins, and D. M. Rocke. 2002. A variance-stabilizing transformation for gene-expression microarray data 10.1093/bioinformatics/18.suppl_1.S105. Bioinformatics 18:S105-110.

Dworsky, M., M. Yow, S. Stagno, R. F. Pass, and C. Alford. 1983. Cytomegalovirus infection of breast milk and transmission in infancy. Pediatrics 72:295-9.

Ebrahimi, B., B. M. Dutia, K. L. Roberts, J. J. Garcia-Ramirez, P. Dickinson, J. P. Stewart, P. Ghazal, D. J. Roy, and A. A. Nash. 2003. Transcriptome profile of murine gammaherpesvirus-68 lytic infection. J Gen Virol 84:99-109.

Evers, D. L., X. Wang, and E. S. Huang. 2004. Cellular stress and signal transduction responses to human cytomegalovirus infection. Microbes Infect 6:1084-93.

Fanger, N. A., L. Borges, and D. Cosman. 1999. The leukocyte immunoglobulin-like receptors (LIRs): a new family of immune regulators. J Leukoc Biol 66:231-6.

Feng, X., J. Schroer, D. Yu, and T. Shenk. 2006. Human cytomegalovirus pUS24 is a virion protein that functions very early in the replication cycle. J Virol 80:8371-8.

Finlay, C. A., P. W. Hinds, T. H. Tan, D. Eliyahu, M. Oren, and A. J. Levine. 1988. Activating mutations for transformation by p53 produce a gene product that forms an hsc70-p53 complex with an altered half-life. Mol Cell Biol 8:531-9.

Fish, K. N., A. S. Depto, A. V. Moses, W. Britt, and J. A. Nelson. 1995. Growth kinetics of human cytomegalovirus are altered in monocyte-derived macrophages. J Virol 69:3737-43.

Fish, K. N., C. Soderberg-Naucler, L. K. Mills, S. Stenglein, and J. A. Nelson. 1998. Human cytomegalovirus persistently infects aortic endothelial cells. J Virol 72:5661-8.

Foldes-Papp, Z., R. Egerer, E. Birch-Hirschfeld, H. M. Striebel, U. Demel, G. P. Tilz, and P. Wutzler. 2004. Detection of multiple human herpes viruses by DNA microarray technology. Mol Diagn 8:1-9.

Fortunato, E. A., A. K. McElroy, I. Sanchez, and D. H. Spector. 2000. Exploitation of cellular signaling and regulatory pathways by human cytomegalovirus. Trends Microbiol 8:111-9.

Frey, S. E., C. Harrison, R. F. Pass, E. Yang, D. Boken, R. E. Sekulovich, S. Percell, A. E. Izu, S. Hirabayashi, R. L. Burke, and A. M. Duliege. 1999. Effects of antigen dose and immunization regimens on antibody responses to a cytomegalovirus glycoprotein B subunit vaccine. J Infect Dis 180:1700-3.

Fries, B. C., S. Chou, M. Boeckh, and B. Torok-Storb. 1994. Frequency distribution of cytomegalovirus envelope glycoprotein genotypes in bone marrow transplant recipients. J Infect Dis 169:769-74.

Gabrielian, K., R. Osusky, B. D. Sippy, S. J. Ryan, and D. R. Hinton. 1994. Effect of TGF-beta on interferon-gamma-induced HLA-DR expression in human retinal pigment epithelial cells. Invest Ophthalmol Vis Sci 35:4253-9.

Gallina, A., L. Simoncini, S. Garbelli, E. Percivalle, G. Pedrali-Noy, K. S. Lee, R. L. Erikson, B. Plachter, G. Gerna, and G. Milanesi. 1999. Polo-like kinase 1 as a target for human cytomegalovirus pp65 lower matrix protein. J Virol 73:1468-78.

Gebert, S., S. Schmolke, G. Sorg, S. Floss, B. Plachter, and T. Stamminger. 1997. The UL84 protein of human cytomegalovirus acts as a transdominant inhibitor of immediate-early-mediated transactivation that is able to prevent viral replication. J Virol 71:7048-60.

Gerna, G., E. Percivalle, F. Baldanti, S. Sozzani, P. Lanzarini, E. Genini, D. Lilleri, and M. G. Revello. 2000. Human cytomegalovirus replicates abortively in polymorphonuclear leukocytes after transfer from infected endothelial cells via transient microfusion events. J Virol 74:5629-38.

Gerna, G., E. Percivalle, D. Lilleri, L. Lozza, C. Fornara, G. Hahn, F. Baldanti, and M. G. Revello. 2005. Dendritic-cell infection by human cytomegalovirus is restricted to strains carrying functional UL131-128 genes and mediates efficient viral antigen presentation to CD8+ T cells. J Gen Virol 86:275-84.

Gerna, G., E. Percivalle, A. Sarasini, F. Baldanti, and M. G. Revello. 2002. The attenuated Towne strain of human cytomegalovirus may revert to both endothelial cell tropism and leuko- (neutrophil- and monocyte-) tropism in vitro. J Gen Virol 83:1993-2000.

Gibson, W. 1996. Structure and assembly of the virion. Intervirology 39:389-400.

Gibson, W., and A. Irmiere. 1984. Selection of particles and proteins for use as human cytomegalovirus subunit vaccines. Birth Defects Orig Artic Ser 20:305-24.

Gilbert, M. J., S. R. Riddell, B. Plachter, and P. D. Greenberg. 1996. Cytomegalovirus selectively blocks antigen processing and presentation of its immediate-early gene product. Nature 383:720-2.

Ginsberg, D., F. Mechta, M. Yaniv, and M. Oren. 1991. Wild-type p53 can down-modulate the activity of various promoters. Proc Natl Acad Sci U S A 88:9979-83.

Goldmacher, V. S., L. M. Bartle, A. Skaletskaya, C. A. Dionne, N. L. Kedersha, C. A. Vater, J.-w. Han, R. J. Lutz, S. Watanabe, E. D. Cahir McFarland, E. D. Kieff, E. S. Mocarski, and T. Chittenden. 1999. A cytomegalovirus-encoded mitochondria-localized inhibitor of apoptosis structurally unrelated to Bcl-2. PNAS 96:12536-12541.

Gonczol, E., J. Ianacone, G. Furlini, W. Ho, and S. A. Plotkin. 1989. Humoral immune response to cytomegalovirus Towne vaccine strain and to Toledo low-passage strain. J Infect Dis 159:851-9.

Gonczol, E., and S. Plotkin. 2001. Development of a cytomegalovirus vaccine: lessons from recent clinical trials. Expert Opin Biol Ther 1:401-12.

Greaves, R. F., and E. S. Mocarski. 1998. Defective growth correlates with reduced accumulation of a viral DNA replication protein after low-multiplicity infection by a human cytomegalovirus ie1 mutant. J Virol 72:366-79.

Grundy, J. E., K. M. Lawson, L. P. MacCormac, J. M. Fletcher, and K. L. Yong. 1998. Cytomegalovirus-infected endothelial cells recruit neutrophils by the secretion of C-X-C chemokines and transmit virus by direct neutrophil-endothelial cell contact and during neutrophil transendothelial migration. J Infect Dis 177:1465-74.

Guo, Y. W., and E. S. Huang. 1993. Characterization of a structurally tricistronic gene of human cytomegalovirus composed of U(s)18, U(s)19, and U(s)20. J Virol 67:2043-54.

Hagemeier, C., R. Caswell, G. Hayhurst, J. Sinclair, and T. Kouzarides. 1994. Functional interaction between the HCMV IE2 transactivator and the retinoblastoma protein. Embo J 13:2897-903.

Hagemeier, C., S. Walker, R. Caswell, T. Kouzarides, and J. Sinclair. 1992a. The human cytomegalovirus 80-kilodalton but not the 72-kilodalton immediate-early protein transactivates heterologous promoters in a TATA box-dependent mechanism and interacts directly with TFIID. J Virol 66:4452-6.

Hagemeier, C., S. M. Walker, P. J. Sissons, and J. H. Sinclair. 1992b. The 72K IE1 and 80K IE2 proteins of human cytomegalovirus independently trans-activate the c-fos, c-myc and hsp70 promoters via basal promoter elements. J Gen Virol 73 (Pt 9):2385-93.

Hahn, G., H. Khan, F. Baldanti, U. H. Koszinowski, M. G. Revello, and G. Gerna. 2002. The human cytomegalovirus ribonucleotide reductase homolog UL45 is dispensable for growth in endothelial cells, as determined by a BAC-cloned clinical isolate of human cytomegalovirus with preserved wild-type characteristics. J Virol 76:9551-5.

Hahn, G., M. G. Revello, M. Patrone, E. Percivalle, G. Campanini, A. Sarasini, M. Wagner, A. Gallina, G. Milanesi, U. Koszinowski, F. Baldanti, and G. Gerna. 2004. Human cytomegalovirus UL131-128 genes are indispensable for virus growth in endothelial cells and virus transfer to leukocytes. J Virol 78:10023-33.

Hai, R., A. Chu, H. Li, S. Umamoto, P. Rider, and F. Liu. 2006. Infection of human cytomegalovirus in cultured human gingival tissue. Virol J 3:84.

Hakki, M., and A. P. Geballe. 2005. Double-stranded RNA binding by human cytomegalovirus pTRS1. J Virol 79:7311-8.

Hakki, M., E. E. Marshall, K. L. De Niro, and A. P. Geballe. 2006. Binding and nuclear relocalization of protein kinase R by human cytomegalovirus TRS1. J Virol 80:11817-26.

Hanfler, J., K. A. Kreuzer, K. Laurisch, S. Tomic, H. Oettle, and C. A. Schmidt. 2002. Quantitation of the spliced late gene of human cytomegalovirus and its kinetics during experimental infection. Medical Microbiology and Immunology 190:161-165.

Hayajneh, W. A., D. G. Contopoulos-Ioannidis, M. M. Lesperance, A. M. Venegas, and A. M. Colberg-Poley. 2001. The carboxyl terminus of the human cytomegalovirus UL37 immediate-early glycoprotein is conserved in primary strains and is important for transactivation. J Gen Virol 82:1569-79.

Hayhurst, G. P., L. A. Bryant, R. C. Caswell, S. M. Walker, and J. H. Sinclair. 1995. CCAAT box-dependent activation of the TATA-less human DNA polymerase alpha promoter by the human cytomegalovirus 72-kilodalton major immediate-early protein. J Virol 69:182-8.

Heineman, T. C., M. Schleiss, D. I. Bernstein, R. R. Spaete, L. Yan, G. Duke, M. Prichard, Z. Wang, Q. Yan, M. A. Sharp, N. Klein, A. M. Arvin, and G. Kemble. 2006. A phase 1 study of 4 live, recombinant human cytomegalovirus Towne/Toledo chimeric vaccines. J Infect Dis 193:1350-60.

Hensel, G., H. Meyer, S. Gartner, G. Brand, and H. F. Kern. 1995. Nuclear localization of the human cytomegalovirus tegument protein pp150 (ppUL32). J Gen Virol 76 (Pt 7):1591-601.

Hobom, U., W. Brune, M. Messerle, G. Hahn, and U. H. Koszinowski. 2000. Fast screening procedures for random transposon libraries of cloned herpesvirus genomes: mutational analysis of human cytomegalovirus envelope glycoprotein genes. J Virol 74:7720-9.

Hofmann, H., S. Floss, and T. Stamminger. 2000. Covalent modification of the transactivator protein IE2-p86 of human cytomegalovirus by conjugation to the ubiquitin-homologous proteins SUMO-1 and hSMT3b. J Virol 74:2510-24.

Huang, L., C. L. Malone, and M. F. Stinski. 1994. A human cytomegalovirus early promoter with upstream negative and positive cis-acting elements: IE2 negates the effect of the negative element, and NF-Y binds to the positive element. J Virol 68:2108-17.

Huang, L., and M. F. Stinski. 1995. Binding of cellular repressor protein or the IE2 protein to a cis-acting negative regulatory element upstream of a human cytomegalovirus early promoter. J Virol 69:7612-21.

Hwang, J. S., and E. Bogner. 2002. ATPase activity of the terminase subunit pUL56 of human cytomegalovirus. J Biol Chem 277:6943-8.

Hyun, J. J., H. S. Park, K. H. Kim, and H. J. Kim. 1999. Analysis of transcripts expressed from the UL47 gene of human cytomegalovirus. Archives of Pharmacal Research 22:542-548.

Irmiere, A., and W. Gibson. 1983. Isolation and characterization of a noninfectious virion-like particle released from cells infected with human strains of cytomegalovirus. Virology 130:118-33.

Ishov, A. M., R. M. Stenberg, and G. G. Maul. 1997. Human cytomegalovirus immediate early interaction with host nuclear structures: Definition of an immediate transcript environment. Journal of Cell Biology 138:5-16.

Ishov, A. M., O. V. Vladimirova, and G. G. Maul. 2002. Daxx-mediated accumulation of human cytomegalovirus tegument protein pp71 at ND10 facilitates initiation of viral infection at these nuclear domains. J Virol 76:7705-12.

Iskenderian, A. C., L. Huang, A. Reilly, R. M. Stenberg, and D. G. Anders. 1996. Four of eleven loci required for transient complementation of human cytomegalovirus DNA replication cooperate to activate expression of replication genes. J Virol 70:383-92.

Isomura, H., M. F. Stinski, A. Kudoh, S. Nakayama, S. Iwahori, Y. Sato, and T. Tsurumi. 2007. The late promoter of the human cytomegalovirus viral DNA polymerase

processivity factor has an impact on delayed early and late viral gene products but not on viral DNA synthesis. J Virol 81:6197-206.

Jaaskelainen, A. J., H. Piiparinen, M. Lappalainen, M. Koskiniemi, and A. Vaheri. 2006. Multiplex-PCR and oligonucleotide microarray for detection of eight different herpesviruses from clinical specimens. J Clin Virol 37:83-90.

Jenkins, D. E., C. L. Martens, and E. S. Mocarski. 1994. Human cytomegalovirus late protein encoded by ie2: a trans-activator as well as a repressor of gene expression. J Gen Virol 75 (Pt 9):2337-48.

Jenner, R. G., M. M. Alba, C. Boshoff, and P. Kellam. 2001. Kaposi's Sarcoma-Associated Herpesvirus Latent and Lytic Gene Expression as Revealed by DNA Arrays 10.1128/JVI.75.2.891-902.2001. J. Virol. 75:891-902.

Johnson, R. A., S. M. Huong, and E. S. Huang. 2000. Activation of the mitogen-activated protein kinase p38 by human cytomegalovirus infection through two distinct pathways: a novel mechanism for activation of p38. J Virol 74:1158-67.

Johnson, R. A., S. M. Huong, and E. S. Huang. 1999. Inhibitory effect of 4-(4-fluorophenyl)-2-(4-hydroxyphenyl)-5-(4-pyridyl)1H - imidazole on HCMV DNA replication and permissive infection. Antiviral Res 41:101-11.

Johnson, R. A., X. L. Ma, A. D. Yurochko, and E. S. Huang. 2001a. The role of MKK1/2 kinase activity in human cytomegalovirus infection. J Gen Virol 82:493-7.

Johnson, R. A., X. Wang, X. L. Ma, S. M. Huong, and E. S. Huang. 2001b. Human cytomegalovirus up-regulates the phosphatidylinositol 3-kinase (PI3-K) pathway: inhibition of PI3-K activity inhibits viral replication and virus-induced signaling. J Virol 75:6022-32.

Jones, J. O., and A. M. Arvin. 2003. Microarray analysis of host cell gene transcription in response to varicella-zoster virus infection of human T cells and fibroblasts in vitro and SCIDhu skin xenografts in vivo. J Virol 77:1268-80.

Jones, T. R., and S.-W. Lee. 2004. An Acidic Cluster of Human Cytomegalovirus UL99 Tegument Protein Is Required for Trafficking and Function 10.1128/JVI.78.3.1488-1502.2004. J. Virol. 78:1488-1502.

Jones, T. R., and V. P. Muzithras. 1992. A cluster of dispensable genes within the human cytomegalovirus genome short component: IRS1, US1 through US5, and the US6 family. J Virol 66:2541-6.

Kahl, M., D. Siegel-Axel, S. Stenglein, G. Jahn, and C. Sinzger. 2000. Efficient lytic infection of human arterial endothelial cells by human cytomegalovirus strains. J Virol 74:7628-35.

Kandel, E. S., and N. Hay. 1999. The regulation and activities of the multifunctional serine/threonine kinase Akt/PKB. Exp Cell Res 253:210-29.

Kari, B., and R. Gehrz. 1992. A human cytomegalovirus glycoprotein complex designated gC-II is a major heparin-binding component of the envelope. J. Virol. 66:1761-1764.

Kaye, J., H. Browne, M. Stoffel, and T. Minson. 1992. The UL16 gene of human cytomegalovirus encodes a glycoprotein that is dispensable for growth in vitro. J Virol 66:6609-15.

Keay, S., and B. Baldwin. 1991. Anti-idiotype antibodies that mimic gp86 of human cytomegalovirus inhibit viral fusion but not attachment. J Virol 65:5124-8.

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 86:2673-84.

Kennedy, P. G., E. O. Major, R. K. Williams, and S. E. Straus. 1994. Down-regulation of glial fibrillary acidic protein expression during acute lytic varicella-zoster virus infection of cultured human astrocytes. Virology 205:558-62.

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 11:376-83.

Kiehl, A., L. L. Huang, D. Franchi, and D. G. Anders. 2003. Multiple 5 ' ends of human cytomegalovirus UL57 transcripts identify a complex, cycloheximide-resistant promoter region that activates oriLyt. Virology 314:410-422.

Kim, J. M., Y. Hong, K. T. Jeang, and S. Kim. 2000a. Transactivation activity of the human cytomegalovirus IE2 protein occurs at steps subsequent to TATA box-binding protein recruitment. J Gen Virol 81:37-46.

Kim, J. M., Y. Hong, and S. Kim. 2000b. Artificial recruitment of Sp1 or TBP can replace the role of IE1 in the synergistic transactivation by IE1 and IE2. Biochem Biophys Res Commun 269:302-8.

Kim, S., S. S. Yu, I. S. Lee, S. Ohno, J. Yim, S. Kim, and H. S. Kang. 1999. Human cytomegalovirus IE1 protein activates AP-1 through a cellular protein kinase(s). J Gen Virol 80 (Pt 4):961-9.

Klucher, K. M., M. Sommer, J. T. Kadonaga, and D. H. Spector. 1993. In vivo and in vitro analysis of transcriptional activation mediated by the human cytomegalovirus major immediate-early proteins. Mol Cell Biol 13:1238-50.

Kohler, C. P., J. A. Kerry, M. Carter, V. P. Muzithras, T. R. Jones, and R. M. Stenberg. 1994. Use of recombinant virus to assess human cytomegalovirus early and late promoters in the context of the viral genome. J Virol 68:6589-97.

Kondo, K., H. Kaneshima, and E. S. Mocarski. 1994. Human cytomegalovirus latent infection of granulocyte-macrophage progenitors. Proc Natl Acad Sci U S A 91:11879-83.

Kovacs, A., M. L. Weber, L. J. Burns, H. S. Jacob, and G. M. Vercellotti. 1996. Cytoplasmic sequestration of p53 in cytomegalovirus-infected human endothelial cells. Am J Pathol 149:1531-9.

LaFemina, R. L., and G. S. Hayward. 1983. Replicative forms of human cytomegalovirus DNA with joined termini are found in permissively infected human cells but not in non-permissive Balb/c-3T3 mouse cells. J Gen Virol 64 (Pt 2):373-89.

Lashmit, P. E., M. F. Stinski, E. A. Murphy, and G. C. Bullock. 1998. A cis repression sequence adjacent to the transcription start site of the human cytomegalovirus US3 gene is required to down regulate gene expression at early and late times after infection. J Virol 72:9575-84.

Leach, F. S., and E. S. Mocarski. 1989. Regulation of cytomegalovirus late-gene expression: differential use of three start sites in the transcriptional activation of ICP36 gene expression. J Virol 63:1783-91.

Leatham, M. P., P. R. Witte, and M. F. Stinski. 1991. Alternate promoter selection within a human cytomegalovirus immediate-early and early transcription unit (UL119-115) defines true late transcripts containing open reading frames for putative viral glycoproteins. J Virol 65:6144-53.

Lee, G., J. Wu, P. Luu, P. Ghazal, and O. Flores. 1996. Inhibition of the association of RNA polymerase II with the preinitiation complex by a viral transcriptional repressor. Proc Natl Acad Sci U S A 93:2570-5.

Lee, J. M., H. J. Kang, H. R. Lee, C. Y. Choi, W. J. Jang, and J. H. Ahn. 2003. PIAS1 enhances SUMO-1 modification and the transactivation activity of the major immediate-early IE2 protein of human cytomegalovirus. FEBS Lett 555:322-8.

Lee, K., K. Jeon, J. M. Kim, V. N. Kim, D. H. Choi, S. U. Kim, and S. Kim. 2005. Downregulation of GFAP, TSP-1, and p53 in human glioblastoma cell line, U373MG, by IE1 protein from human cytomegalovirus. Glia 51:1-12.

Lee, M., J. Xiao, E. Haghjoo, X. Zhan, G. Abenes, T. Tuong, W. Dunn, and F. Liu. 2000. Murine cytomegalovirus containing a mutation at open reading frame M37 is severely attenuated in growth and virulence in vivo. J Virol 74:11099-107.

Lembo, D., M. Donalisio, A. Hofer, M. Cornaglia, W. Brune, U. Koszinowski, L. Thelander, and S. Landolfo. 2004. The ribonucleotide reductase R1 homolog of murine cytomegalovirus is not a functional enzyme subunit but is required for pathogenesis. J Virol 78:4278-88.

Lembo, D., G. Gribaudo, A. Hofer, L. Riera, M. Cornaglia, A. Mondo, A. Angeretti, M. Gariglio, L. Thelander, and S. Landolfo. 2000. Expression of an altered ribonucleotide reductase activity associated with the replication of murine cytomegalovirus in quiescent fibroblasts. J Virol 74:11557-65.

Leuzinger, H., U. Ziegler, E. M. Schraner, C. Fraefel, D. L. Glauser, I. Heid, M. Ackermann, M. Mueller, and P. Wild. 2005. Herpes simplex virus 1 envelopment follows two diverse pathways. J Virol 79:13047-59.

Li, C., R. S. Chen, S. K. Hung, Y. T. Lee, C. Y. Yen, Y. W. Lai, R. H. Teng, J. Y. Huang, Y. C. Tang, C. P. Tung, T. T. Wei, B. Shieh, and S. T. Liu. 2006. Detection of Epstein-Barr virus infection and gene expression in human tumors by microarray analysis. J Virol Methods 133:158-66.

Lockridge, K. M., G. Sequar, S. S. Zhou, Y. Yue, C. P. Mandell, and P. A. Barry. 1999. Pathogenesis of experimental rhesus cytomegalovirus infection. J Virol 73:9576-83.

Lokensgard, J. R., M. C. Cheeran, G. Gekker, S. Hu, C. C. Chao, and P. K. Peterson. 1999. Human cytomegalovirus replication and modulation of apoptosis in astrocytes. J Hum Virol 2:91-101.

Louis, D. N., A. von Deimling, R. Y. Chung, M. P. Rubio, J. M. Whaley, R. H. Eibl, H. Ohgaki, O. D. Wiestler, A. D. Thor, and B. R. Seizinger. 1993. Comparative study of p53 gene and protein alterations in human astrocytic tumors. J Neuropathol Exp Neurol 52:31-8.

Loureiro, J., and H. L. Ploegh. 2006. Antigen presentation and the ubiquitin-proteasome system in host-pathogen interactions. Adv Immunol 92:225-305.

Lu, M., and T. Shenk. 1996. Human cytomegalovirus infection inhibits cell cycle progression at multiple points, including the transition from G1 to S. J Virol 70:8850-7.

Lukac, D. M., N. Y. Harel, N. Tanese, and J. C. Alwine. 1997. TAF-like functions of human cytomegalovirus immediate-early proteins. J Virol 71:7227-39.

Lukac, D. M., J. R. Manuppello, and J. C. Alwine. 1994. Transcriptional activation by the human cytomegalovirus immediate-early proteins: requirements for simple promoter structures and interactions with multiple components of the transcription complex. J Virol 68:5184-93.

Lurain, N. S., A. M. Fox, H. M. Lichy, S. M. Bhorade, C. F. Ware, D. D. Huang, S. P. Kwan, E. R. Garrity, and S. Chou. 2006. Analysis of the human cytomegalovirus genomic region from UL146 through UL147A reveals sequence hypervariability, genotypic stability, and overlapping transcripts. Virol J 3:4.

Lurain, N. S., K. S. Kapell, D. D. Huang, J. A. Short, J. Paintsil, E. Winkfield, C. A. Benedict, C. F. Ware, and J. W. Bremer. 1999. Human cytomegalovirus UL144 open reading frame: sequence hypervariability in low-passage clinical isolates. J Virol 73:10040-50.

Mach, M., B. Kropff, P. Dal Monte, and W. Britt. 2000. Complex formation by human cytomegalovirus glycoproteins M (gpUL100) and N (gpUL73). J Virol 74:11881-92.

Mach, M., B. Kropff, M. Kryzaniak, and W. Britt. 2005. Complex formation by glycoproteins M and N of human cytomegalovirus: structural and functional aspects. J Virol 79:2160-70.

Mach, M., K. Osinski, B. Kropff, U. Schloetzer-Schrehardt, M. Krzyzaniak, and W. Britt. 2007. The carboxyterminal domain of glycoprotein N of human cytomegalovirus is required for virion morphogenesis. J Virol.

Maciejewski, J. P., E. E. Bruening, R. E. Donahue, E. S. Mocarski, N. S. Young, and S. C. St Jeor. 1992. Infection of hematopoietic progenitor cells by human cytomegalovirus. Blood 80:170-8.

Marchini, A., H. Liu, and H. Zhu. 2001. Human cytomegalovirus with IE-2 (UL122) deleted fails to express early lytic genes. J Virol 75:1870-8.

Margolis, M. J., S. Pajovic, E. L. Wong, M. Wade, R. Jupp, J. A. Nelson, and J. C. Azizkhan. 1995. Interaction of the 72-kilodalton human cytomegalovirus IE1 gene product with E2F1 coincides with E2F-dependent activation of dihydrofolate reductase transcription. J Virol 69:7759-67.

Mayne, M., C. Cheadle, S. S. Soldan, C. Cermelli, Y. Yamano, N. Akhyani, J. E. Nagel, D. D. Taub, K. G. Becker, and S. Jacobson. 2001. Gene expression profile of herpesvirus-infected T cells obtained using immunomicroarrays: induction of proinflammatory mechanisms. J Virol 75:11641-50.

McGregor, A., and M. R. Schleiss. 2001. Molecular cloning of the guinea pig cytomegalovirus (GPCMV) genome as an infectious bacterial artificial chromosome (BAC) in Escherichia coli. Mol Genet Metab 72:15-26.

McVoy, M. A., D. E. Nixon, J. K. Hur, and S. P. Adler. 2000. The ends on herpesvirus DNA replicative concatemers contain pac2 cis cleavage/packaging elements and their formation is controlled by terminal cis sequences. J Virol 74:1587-92.

Menard, C., M. Wagner, Z. Ruzsics, K. Holak, W. Brune, A. E. Campbell, and U. H. Koszinowski. 2003. Role of murine cytomegalovirus US22 gene family members in replication in macrophages. J Virol 77:5557-70.

Mettenleiter, T. C. 2002. Herpesvirus Assembly and Egress. 2002. J. Virol. 76:1537-1547.

Meyer-Konig, U., C. Vogelberg, A. Bongarts, D. Kampa, R. Delbruck, G. Wolff-Vorbeck, G. Kirste, M. Haberland, F. T. Hufert, and D. von Laer. 1998. Glycoprotein B genotype correlates with cell tropism in vivo of human cytomegalovirus infection. J Med Virol 55:75-81.

Mocarski, E. S., and C. T. Courcelle. 2001. Cytomegaloviruses and their replication, 4th ed, vol. 2. Lippincott Williams and Williams, Philadelphia, Pa.

Mocarski, E. S., G. W. Kemble, J. M. Lyle, and R. F. Greaves. 1996. A deletion mutant in the human cytomegalovirus gene encoding IE1(491aa) is replication defective due to a failure in autoregulation. Proc Natl Acad Sci U S A 93:11321-6.

Moffat, J. F., L. Zerboni, M. H. Sommer, T. C. Heineman, J. I. Cohen, H. Kaneshima, and A. M. Arvin. 1998. The ORF47 and ORF66 putative protein kinases of varicella-zoster virus determine tropism for human T cells and skin in the SCID-hu mouse. Proc Natl Acad Sci U S A 95:11969-74.

Montag, C., J. Wagner, I. Gruska, and C. Hagemeier. 2006. Human cytomegalovirus blocks tumor necrosis factor alpha- and interleukin-1beta-mediated NF-kappaB signaling. J Virol 80:11686-98.

Morello, C. S., L. D. Cranmer, and D. H. Spector. 1999. In vivo replication, latency, and immunogenicity of murine cytomegalovirus mutants with deletions in the M83 and M84 genes, the putative homologs of human cytomegalovirus pp65 (UL83). J Virol 73:7678-93.

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 75:750-8.

Muganda, P., O. Mendoza, J. Hernandez, and Q. Qian. 1994. Human cytomegalovirus elevates levels of the cellular protein p53 in infected fibroblasts. J Virol 68:8028-34.

Murphy, E., I. Rigoutsos, T. Shibuya, and T. E. Shenk. 2003a. Reevaluation of human cytomegalovirus coding potential. Proc Natl Acad Sci U S A 100:13585-90.

Murphy, E., D. Yu, J. Grimwood, J. Schmutz, M. Dickson, M. A. Jarvis, G. Hahn, J. A. Nelson, R. M. Myers, and T. E. Shenk. 2003b. Coding potential of laboratory and clinical strains of human cytomegalovirus. Proc Natl Acad Sci U S A 100:14976-81.

Navarro, D., P. Paz, S. Tugizov, K. Topp, J. La Vail, and L. Pereira. 1993. Glycoprotein B of human cytomegalovirus promotes virion penetration into cells, transmission of infection from cell to cell, and fusion of infected cells. Virology 197:143-58.

Navarro, L., K. Mowen, S. Rodems, B. Weaver, N. Reich, D. Spector, and M. David. 1998. Cytomegalovirus activates interferon immediate-early response gene expression and an interferon regulatory factor 3-containing interferon-stimulated response element-binding complex. Mol Cell Biol 18:3796-802.

Nevels, M., C. Paulus, and T. Shenk. 2004. Human cytomegalovirus immediate-early 1 protein facilitates viral replication by antagonizing histone deacetylation. Proc Natl Acad Sci U S A 101:17234-9.

Odeberg, J., C. Cerboni, H. Browne, K. Karre, E. Moller, E. Carbone, and C. Soderberg-Naucler. 2002. Human cytomegalovirus (HCMV)-infected endothelial cells and macrophages are less susceptible to natural killer lysis independent of the downregulation of classical HLA class I molecules or expression of the HCMV class I homologue, UL18. Scand J Immunol 55:149-61.

Ogawa-Goto, K., K. Tanaka, W. Gibson, E. Moriishi, Y. Miura, T. Kurata, S. Irie, and T. Sata. 2003. Microtubule network facilitates nuclear targeting of human cytomegalovirus capsid. J Virol 77:8541-7.

Ohyashiki, J. H., T. Takaku, T. Ojima, K. Abe, K. Yamamoto, Y. Zhang, and K. Ohyashiki. 2005. Transcriptional profiling of human herpesvirus type B (HHV-6B) in an adult T cell leukemia cell line as in vitro model for persistent infection. Biochem Biophys Res Commun 329:11-7.

Pari, G. S., M. A. Kacica, and D. G. Anders. 1993. Open reading frames UL44, IRS1/TRS1, and UL36-38 are required for transient complementation of human cytomegalovirus oriLyt-dependent DNA synthesis. J Virol 67:2575-82.

Patrone, M., E. Percivalle, M. Secchi, L. Fiorina, G. Pedrali-Noy, M. Zoppe, F. Baldanti, G. Hahn, U. H. Koszinowski, G. Milanesi, and A. Gallina. 2003. The human cytomegalovirus UL45 gene product is a late, virion-associated protein and influences virus growth at low multiplicities of infection. J Gen Virol 84:3359-70.

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 75:4843-53.

Paulus, C., S. Krauss, and M. Nevels. 2006. A human cytomegalovirus antagonist of type I IFN-dependent signal transducer and activator of transcription signaling. Proc Natl Acad Sci U S A 103:3840-5.

Penfold, M. E., D. J. Dairaghi, G. M. Duke, N. Saederup, E. S. Mocarski, G. W. Kemble, and T. J. Schall. 1999. Cytomegalovirus encodes a potent alpha chemokine. Proc Natl Acad Sci U S A 96:9839-44.

Penfold, M. E., and E. S. Mocarski. 1997. Formation of cytomegalovirus DNA replication compartments defined by localization of viral proteins and DNA synthesis. Virology 239:46-61.

Plotkin, S. A., T. Furukawa, N. Zygraich, and C. Huygelen. 1975. Candidate cytomegalovirus strain for human vaccination. Infect. Immun. 12:521-527.

Plotkin, S. A., S. E. Starr, H. M. Friedman, E. Gonczol, and R. E. Weibel. 1989. Protective effects of Towne cytomegalovirus vaccine against low-passage cytomegalovirus administered as a challenge. J Infect Dis 159:860-5.

Poma, E. E., T. F. Kowalik, L. Zhu, J. H. Sinclair, and E. S. Huang. 1996. The human cytomegalovirus IE1-72 protein interacts with the cellular p107 protein and relieves p107-mediated transcriptional repression of an E2F-responsive promoter. J Virol 70:7867-77.

Poole, L. J., Y. Yu, P. S. Kim, Q. Z. Zheng, J. Pevsner, and G. S. Hayward. 2002. Altered patterns of cellular gene expression in dermal microvascular endothelial cells infected with Kaposi's sarcoma-associated herpesvirus. J Virol 76:3395-420.

Preston, C. M., A. N. Harman, and M. J. Nicholl. 2001. Activation of interferon response factor-3 in human cells infected with herpes simplex virus type 1 or human cytomegalovirus. J Virol 75:8909-16.

Prichard, M. N., M. E. Penfold, G. M. Duke, R. R. Spaete, and G. W. Kemble. 2001. A review of genetic differences between limited and extensively passaged human cytomegalovirus strains. Rev Med Virol 11:191-200.

Rawlinson, W. D., and B. G. Barrell. 1993. Spliced Transcripts of Human Cytomegalovirus. Journal of Virology 67:5502-5513.

Revello, M. G., E. Percivalle, E. Arbustini, R. Pardi, S. Sozzani, and G. Gerna. 1998. In vitro generation of human cytomegalovirus pp65 antigenemia, viremia, and leukoDNAemia. J Clin Invest 101:2686-92.

Revello, M. G., E. Percivalle, A. Di Matteo, F. Morini, and G. Gerna. 1992. Nuclear expression of the lower matrix protein of human cytomegalovirus in peripheral blood leukocytes of immunocompromised viraemic patients. J Gen Virol 73 (Pt 2):437-42.

Riegler, S., H. Hebart, H. Einsele, P. Brossart, G. Jahn, and C. Sinzger. 2000. Monocyte-derived dendritic cells are permissive to the complete replicative cycle of human cytomegalovirus. J Gen Virol 81:393-9.

Rigoutsos, I., J. Novotny, T. Huynh, S. T. Chin-Bow, L. Parida, D. Platt, D. Coleman, and T. Shenk. 2003. In silico pattern-based analysis of the human cytomegalovirus genome. Journal of Virology 77:4326-4344.

Rocca-Serra, P., A. Brazma, H. Parkinson, U. Sarkans, M. Shojatalab, S. Contrino, J. Vilo, N. Abeygunawardena, G. Mukherjee, E. Holloway, M. Kapushesky, P. Kemmeren, G. G. Lara, A. Oezcimen, and S. A. Sansone. 2003. ArrayExpress: a public database of gene expression data at EBI. C R Biol 326:1075-8.

Roizman, B. A., and P. E. Pellet. 2001. The family herpesviridae: a brief introduction, 4th edn, vol. 2, Philadelphia: Lippincott-Williams & Wilkins, Pa.

Romanowski, M. J., and T. Shenk. 1997. Characterization of the human cytomegalovirus irs1 and trs1 genes: a second immediate-early transcription unit within irs1 whose product antagonizes transcriptional activation. J Virol 71:1485-96.

Rue, C. A., M. A. Jarvis, A. J. Knoche, H. L. Meyers, V. R. DeFilippis, S. G. Hansen, M. Wagner, K. Fruh, D. G. Anders, S. W. Wong, P. A. Barry, and J. A. Nelson. 2004. A

cyclooxygenase-2 homologue encoded by rhesus cytomegalovirus is a determinant for endothelial cell tropism. J Virol 78:12529-36.

Salmon-Ceron, D. 2001. Cytomegalovirus infection: the point in 2001. HIV Med 2:255-9.

Sanchez, V., K. D. Greis, E. Sztul, and W. J. Britt. 2000. Accumulation of virion tegument and envelope proteins in a stable cytoplasmic compartment during human cytomegalovirus replication: characterization of a potential site of virus assembly. J Virol 74:975-86.

Schmolke, S., H. F. Kern, P. Drescher, G. Jahn, and B. Plachter. 1995. The dominant phosphoprotein pp65 (UL83) of human cytomegalovirus is dispensable for growth in cell culture. J Virol 69:5959-68.

Schwartz, R., M. H. Sommer, A. Scully, and D. H. Spector. 1994. Site-specific binding of the human cytomegalovirus IE2 86-kilodalton protein to an early gene promoter. J Virol 68:5613-22.

Scott, G. M., B. G. Barrell, J. Oram, and W. D. Rawlinson. 2002. Characterisation of transcripts from the human cytomegalovirus genes TRL7, UL20a, UL36, UL65, UL94, US3 and US34. Virus Genes 24:39-48.

Severi, B., M. P. Landini, G. Cenacchi, N. Zini, and N. M. Maraldi. 1992. Human cytomegalovirus nuclear and cytoplasmic dense bodies. Arch Virol 123:193-207.

Silva, M. C., Q.-C. Yu, L. Enquist, and T. Shenk. 2003. Human Cytomegalovirus UL99-Encoded pp28 Is Required for the Cytoplasmic Envelopment of Tegument-Associated Capsids
10.1128/JVI.77.19.10594-10605.2003. J. Virol. 77:10594-10605.

Sinzger, C., A. Grefte, B. Plachter, A. S. Gouw, T. H. The, and G. Jahn. 1995. Fibroblasts, epithelial cells, endothelial cells and smooth muscle cells are major targets of human cytomegalovirus infection in lung and gastrointestinal tissues. J Gen Virol 76 (Pt 4):741-50.

Sinzger, C., and G. Jahn. 1996. Human cytomegalovirus cell tropism and pathogenesis. Intervirology 39:302-19.

Sinzger, C., M. Kahl, K. Laib, K. Klingel, P. Rieger, B. Plachter, and G. Jahn. 2000. Tropism of human cytomegalovirus for endothelial cells is determined by a post-entry step dependent on efficient translocation to the nucleus. J Gen Virol 81:3021-35.

Skaletskaya, A., L. M. Bartle, T. Chittenden, A. L. McCormick, E. S. Mocarski, and V. S. Goldmacher. 2001. A cytomegalovirus-encoded inhibitor of apoptosis that suppresses caspase-8 activation. PNAS 98:7829-7834.

Smith, J. A., S. Jairath, J. J. Crute, and G. S. Pari. 1996. Characterization of the human cytomegalovirus UL105 gene and identification of the putative helicase protein. Virology 220:251-255.

Smith, J. A., and G. S. Pari. 1995. Human Cytomegalovirus Ul102 Gene. Journal of Virology 69:1734-1740.

Smith, W. L., D. L. DeWitt, and R. M. Garavito. 2000. Cyclooxygenases: structural, cellular, and molecular biology. Annu Rev Biochem 69:145-82.

Smuda, C., E. Bogner, and K. Radsak. 1997. The human cytomegalovirus glycoprotein B gene (ORF UL55) is expressed early in the infectious cycle. Journal of General Virology 78:1981-1992.

Sodeik, B., M. W. Ebersold, and A. Helenius. 1997. Microtubule-mediated transport of incoming herpes simplex virus 1 capsids to the nucleus. J Cell Biol 136:1007-21.

Soderberg-Naucler, C., D. N. Streblow, K. N. Fish, J. Allan-Yorke, P. P. Smith, and J. A. Nelson. 2001. Reactivation of latent human cytomegalovirus in CD14(+) monocytes is differentiation dependent. J Virol 75:7543-54.

Sommer, M. H., A. L. Scully, and D. H. Spector. 1994. Transactivation by the human cytomegalovirus IE2 86-kilodalton protein requires a domain that binds to both the TATA box-binding protein and the retinoblastoma protein. J Virol 68:6223-31.

Song, Y. J., and M. F. Stinski. 2002. Effect of the human cytomegalovirus IE86 protein on expression of E2F-responsive genes: a DNA microarray analysis. Proc Natl Acad Sci U S A 99:2836-41.

Song, Y. J., and M. F. Stinski. 2005. Inhibition of cell division by the human cytomegalovirus IE86 protein: role of the p53 pathway or cyclin-dependent kinase 1/cyclin B1. J Virol 79:2597-603.

Speckner, A., D. Glykofrydes, M. Ohlin, and M. Mach. 1999. Antigenic domain 1 of human cytomegalovirus glycoprotein B induces a multitude of different antibodies which, when combined, results in incomplete virus neutralization. J Gen Virol 80 (Pt 8):2183-91.

Spector, D. H. 1996. Activation and regulation of human cytomegalovirus early genes. Intervirology 39:361-377.

Speir, E., Z. X. Yu, V. J. Ferrans, E. S. Huang, and S. E. Epstein. 1998. Aspirin attenuates cytomegalovirus infectivity and gene expression mediated by cyclooxygenase-2 in coronary artery smooth muscle cells. Circ Res 83:210-6.

Stamminger, T., E. Puchtler, and B. Fleckenstein. 1991. Discordant Expression of the Immediate-Early-1 and Immediate-Early-2 Gene Regions of Human Cytomegalovirus at Early Times after Infection Involves Posttranscriptional Processing Events. Journal of Virology 65:2273-2282.

Starr, S. E., J. P. Glazer, H. M. Friedman, J. D. Farquhar, and S. A. Plotkin. 1981. Specific cellular and humoral immunity after immunization with live Towne strain cytomegalovirus vaccine. J Infect Dis 143:585-9.

Stekel, D. 2003. Microarray bioinformatics. 1st edn, Cambridge University Press.

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 74:9916-27.

Striebel, H. M., E. Birch-Hirschfeld, R. Egerer, Z. Foldes-Papp, G. P. Tilz, and A. Stelzner. 2004. Enhancing sensitivity of human herpes virus diagnosis with DNA microarrays using dendrimers. Exp Mol Pathol 77:89-97.

Subler, M. A., D. W. Martin, and S. Deb. 1992. Inhibition of viral and cellular promoters by human wild-type p53. J Virol 66:4757-62.

Tal-Singer, R., R. Pichyangkura, E. Chung, T. M. Lasner, B. P. Randazzo, J. Q. Trojanowski, N. W. Fraser, and S. J. Triezenberg. 1999. The transcriptional activation domain of VP16 is required for efficient infection and establishment of latency by HSV-1 in the murine peripheral and central nervous systems. Virology 259:20-33.

Tang, Q., E. A. Murphy, and G. G. Maul. 2006. Experimental confirmation of global murine cytomegalovirus open reading frames by transcriptional detection and partial characterization of newly described gene products. J Virol 80:6873-82.

Tavalai, N., P. Papior, S. Rechter, M. Leis, and T. Stamminger. 2006. Evidence for a role of the cellular ND10 protein PML in mediating intrinsic immunity against human cytomegalovirus infections. J Virol 80:8006-18.

Tomasec, P., V. M. Braud, C. Rickards, M. B. Powell, B. P. McSharry, S. Gadola, V. Cerundolo, L. K. Borysiewicz, A. J. McMichael, and G. W. Wilkinson. 2000. Surface expression of HLA-E, an inhibitor of natural killer cells, enhanced by human cytomegalovirus gpUL40. Science 287:1031.

Tomasec, P., E. C. Wang, A. J. Davison, B. Vojtesek, M. Armstrong, C. Griffin, B. P. McSharry, R. J. Morris, S. Llewellyn-Lacey, C. Rickards, A. Nomoto, C. Sinzger, and G. W. Wilkinson. 2005. Downregulation of natural killer cell-activating ligand CD155 by human cytomegalovirus UL141. Nat Immunol 6:181-8.

Tsai, H. L., G. H. Kou, S. C. Chen, C. W. Wu, and Y. S. Lin. 1996. Human cytomegalovirus immediate-early protein IE2 tethers a transcriptional repression domain to p53. J Biol Chem 271:3534-40.

Van Meir, E. G., T. Kikuchi, M. Tada, H. Li, A. C. Diserens, B. E. Wojcik, H. J. Huang, T. Friedmann, N. de Tribolet, and W. K. Cavenee. 1994. Analysis of the p53 gene and its expression in human glioblastoma cells. Cancer Res 54:649-52.

Varnum, S. M., D. N. Streblow, M. E. Monroe, P. Smith, K. J. Auberry, L. Pasa-Tolic, D. Wang, D. G. Camp, II, K. Rodland, S. Wiley, W. Britt, T. Shenk, R. D. Smith, and J. A. Nelson. 2004. Identification of Proteins in Human Cytomegalovirus (HCMV) Particles: the HCMV Proteome. 2004. J. Virol. 78:10960-10966.

Waldman, W. J., W. H. Roberts, D. H. Davis, M. V. Williams, D. D. Sedmak, and R. E. Stephens. 1991. Preservation of natural endothelial cytopathogenicity of cytomegalovirus by propagation in endothelial cells. Arch Virol 117:143-64.

Wang, D., and T. Shenk. 2005a. Human cytomegalovirus UL131 open reading frame is required for epithelial cell tropism. J Virol 79:10330-8.

Wang, D., and T. Shenk. 2005b. Human cytomegalovirus virion protein complex required for epithelial and endothelial cell tropism. Proc Natl Acad Sci USA 102:18153-8.

Welch, A. R., A. S. Woods, L. M. McNally, R. J. Cotter, and W. Gibson. 1991. A herpesvirus maturational proteinase, assemblin: identification of its gene, putative active site domain, and cleavage site. Proc Natl Acad Sci U S A 88:10792-6.

Welte, S. A., C. Sinzger, S. Z. Lutz, H. Singh-Jasuja, K. L. Sampaio, U. Eknigk, H. G. Rammensee, and A. Steinle. 2003. Selective intracellular retention of virally induced NKG2D ligands by the human cytomegalovirus UL16 glycoprotein. Eur J Immunol 33:194-203.

Wiertz, E. J., T. R. Jones, L. Sun, M. Bogyo, H. J. Geuze, and H. L. Ploegh. 1996a. The human cytomegalovirus US11 gene product dislocates MHC class I heavy chains from the endoplasmic reticulum to the cytosol. Cell 84:769-79.

Wiertz, E. J., D. Tortorella, M. Bogyo, J. Yu, W. Mothes, T. R. Jones, T. A. Rapoport, and H. L. Ploegh. 1996b. Sec61-mediated transfer of a membrane protein from the endoplasmic reticulum to the proteasome for destruction. Nature 384:432-8.

Wing, B. A., and E. S. Huang. 1995. Analysis and mapping of a family of 3'-coterminal transcripts containing coding sequences for human cytomegalovirus open reading frames UL93 through UL99. J Virol 69:1521-31.

Wing, B. A., R. A. Johnson, and E. S. Huang. 1998. Identification of positive and negative regulatory regions involved in regulating expression of the human cytomegalovirus UL94 late promoter: role of IE2-86 and cellular p53 in mediating negative regulatory function. J Virol 72:1814-25.

Woodhall, D. L., I. J. Groves, M. B. Reeves, G. Wilkinson, and J. H. Sinclair. 2006. Human Daxx-mediated repression of human cytomegalovirus gene expression correlates with a repressive chromatin structure around the major immediate early promoter. J Biol Chem 281:37652-60.

Wu, J., J. O'Neill, and M. S. Barbosa. 2001. Late temporal gene expression from the human cytomegalovirus pp28US (UL99) promoter when integrated into the host cell chromosome. J Gen Virol 82:1147-55.

Wu, J., J. O'Neill, and M. S. Barbosa. 1998. Transcription factor Sp1 mediates cell-specific trans-activation of the human cytomegalovirus DNA polymerase gene promoter by immediate-early protein IE86 in glioblastoma U373MG cells. J Virol 72:236-44.

Xiao, J., T. Tong, X. Zhan, E. Haghjoo, and F. Liu. 2000. In Vitro and In Vivo Characterization of a Murine Cytomegalovirus with a Transposon Insertional Mutation at Open Reading Frame M43. 2000. J. Virol. 74:9488-9497.

Xu, Y., S. A. Cei, A. R. Huete, and G. S. Pari. 2004a. Human cytomegalovirus UL84 insertion mutant defective for viral DNA synthesis and growth. J Virol 78:10360-9.

Xu, Y., S. A. Cei, A. Rodriguez Huete, K. S. Colletti, and G. S. Pari. 2004b. Human cytomegalovirus DNA replication requires transcriptional activation via an IE2- and UL84-responsive bidirectional promoter element within oriLyt. J Virol 78:11664-77.

Yamagishi, J., R. Isobe, T. Takebuchi, and H. Bando. 2003. DNA microarrays of baculovirus genomes: differential expression of viral genes in two susceptible insect cell lines. Arch Virol 148:587-97.

Yguerabide, J., and E. E. Yguerabide. 2001. Resonance light scattering particles as ultrasensitive labels for detection of analytes in a wide range of applications. J Cell Biochem Suppl 37:71-81.

Yu, X., S. Shah, I. Atanasov, P. Lo, F. Liu, W. J. Britt, and Z. H. Zhou. 2005. Three-Dimensional Localization of the Smallest Capsid Protein in the Human Cytomegalovirus Capsid. 2005. J. Virol. 79:1327-1332.

Yurochko, A. D., E. S. Hwang, L. Rasmussen, S. Keay, L. Pereira, and E. S. Huang. 1997. The human cytomegalovirus UL55 (gB) and UL75 (gH) glycoprotein ligands initiate the rapid activation of Sp1 and NF-kappaB during infection. J Virol 71:5051-9.

Zhang, G., B. Raghavan., M. Kotur, J. Cheatham, J. Sedmak, C. Cook, J. Waldman, and J. Trgovcich. 2007. Antisense transcription in the human cytomegalovirus transcriptome. J Virol 81:11267-81.

Zhang, Y., J. H. Ohyashiki, T. Takaku, N. Shimizu, and K. Ohyashiki. 2006. Transcriptional profiling of Epstein-Barr virus (EBV) genes and host cellular genes in nasal NK/T-cell lymphoma and chronic active EBV infection. Br J Cancer 94:599-608.

Zhang, Z., S. M. Huong, X. Wang, D. Y. Huang, and E. S. Huang. 2003. Interactions between human cytomegalovirus IE1-72 and cellular p107: functional domains and mechanisms of up-regulation of cyclin E/cdk2 kinase activity. J Virol 77:12660-70.

Zhou, Z. H., D. H. Chen, J. Jakana, F. J. Rixon, and W. Chiu. 1999. Visualization of tegument-capsid interactions and DNA in intact herpes simplex virus type 1 virions. J Virol 73:3210-8.

Zhou, Z. H., J. He, J. Jakana, J. D. Tatman, F. J. Rixon, and W. Chiu. 1995. Assembly of VP26 in herpes simplex virus-1 inferred from structures of wild-type and recombinant capsids. Nat Struct Biol 2:1026-30.

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. Proceedings of the National Academy of Sciences of the United States of America 95:14470-14475.

Zhu, H., Y. Q. Shen, and T. Shenk. 1995. Human Cytomegalovirus Ie1 and Ie2 Proteins Block Apoptosis. Journal of Virology 69:7960-7970.

Zipeto, D., F. Baldanti, E. Percivalle, G. Gerna, and G. Milanesi. 1993. Identification of a human cytomegalovirus mutant in the pp150 matrix phosphoprotein gene with a growth-defective phenotype. J Gen Virol 74 (Pt 8):1645-8.