

***‘Developing a Human
Cytomegalovirus strain for
better in vitro research’***

A thesis submitted in candidature for the degree of

DOCTOR OF PHILOSOPHY (PhD)

By

Isa Imanial Murrell

October 2014

HCMV and Adenovirus Research Group,
Institute of Infection and Immunity,
School of Medicine,
Cardiff University,
CF14 4XN, UK

ACKNOWLEDGEMENTS AND DEDICATIONS

My thanks and gratitude extend to all members of the HCMV and Adenovirus Research Group in the laboratory of Professor Gavin Wilkinson, as well as to all colleagues throughout the entire School of Medicine, Cardiff University. The support, advice, and encouragement that I have received from these colleagues has been essential not only to the success of this work, but also for my own professional and personal development.

I am particularly indebted to my supervisor Dr. Richard Stanton for the opportunity to undertake this project, and perhaps more importantly, for the immeasurable guidance and patience that he has provided throughout this work – at times, I must have driven Rich as crazy as the work has driven me!

I am grateful to the National Institute for Social Care and Health Research (NISCHR) body of the Welsh Government body that funded this work. I would also like acknowledge the contribution of collaborators in the laboratory of Dr Andrew Davison (MRC-University of Glasgow Centre for Virus Research, Glasgow, UK) who performed high-throughput whole-genome analyses of the viruses investigated in this work, and also collaborators in the laboratory of Dr Paul Lehner (Cambridge Institute for Medical Research, University of Cambridge, UK) who performed mass-spectrometry analysis of virions from the Merlin variants investigated.

I dedicate this thesis to my family and also many of my life-long friends who have been ever-present during this journey – whilst these people may not appreciate the finer details of this work and the causes of my sporadic insanity, they have suffered my continual tiredness and moaning without question. Worthy of special mention are my mother Caryle and my grandmother Mary, both of whom have kept me fed, clothed, sheltered, and in pocket until this very day. I would also like to mention my aunty Cynthia whose counsel provided the stimulus for me to pursue a career in academic science.

‘bono animo esto...’

SUMMARY

Investigations into Human Cytomegalovirus (HCMV) pathogenesis should be based on strains that closely reflect the causative agent of disease, however HCMV invariably mutates *in vitro*, generating phenotypically distinct laboratory-adapted strains. In particular, mutations are selected in the HCMV genome UL128 locus (UL128L) that encodes sub-units of a virion envelope pentameric complex that impedes virus propagation in fibroblasts (the cell type most commonly used *in vitro*), but is required for infection of several naturally targeted cell-types (e.g. epithelial, endothelial, and myeloid cells). Addressing this issue, the genome of wildtype HCMV strain Merlin was cloned as a stable bacterial artificial chromosome (BAC), however similarly to clinical isolates, viruses reconstituted from the Merlin-BAC grow poorly and are prone to *de novo* mutation in cell culture.

Direct comparison to viruses from the Merlin-BAC revealed that viruses from the BAC-cloned versions of strains TR (TR-BAC), TB40E (TB40-BAC4) and VR1814 (FIX-BAC) could be propagated more efficiently in fibroblasts, despite containing intact UL128L ORFs. Unique nucleotide variations identified in TB40-BAC4 and FIX-BAC UL128L ORFs were transferred into the Merlin-BAC, generating variants that produced greater titres of cell-free virus following reconstitution. Virions from these novel Merlin variants displayed reduced pentameric complex content, but remained able to infect epithelial cells, albeit with slightly compromised efficiency. The greater fitness of viruses from these novel Merlin-BAC variants alleviated the selective pressures for the selection of *de novo* UL128L mutations in fibroblasts. The Merlin virion proteome was determined, with up to 30 novel components identified to provide a more comprehensive picture of wildtype HCMV virion composition. Comparison of virions from several Merlin variants demonstrated that varying pentameric complex content impacted the incorporation of other components, however virions from the novel Merlin variants produced in this work closely matched those from the parental Merlin variant containing wildtype UL128L ORFs. Thus, the novel Merlin-BAC variants produced in this work represent valuable reagents for future HCMV research.

CONTENTS

DECLARATION.....	I
ACKNOWLEDGEMENTS AND DEDICATIONS.....	III
SUMMARY.....	V
CONTENTS.....	VII
SYMBOLS/ACRONYMS/ABBREVIATIONS.....	XV
1 Introduction	1
1.1 Human Cytomegalovirus (HCMV) - Human Herpesvirus 5 (HHV5)	3
1.1.1 Taxonomy	3
1.1.2 Evolution, phylogeny and related species.....	3
1.2 Discovery, isolation, and propagation of HCMV <i>in vitro</i>.....	5
1.3 Classification of HCMV as a herpesvirus.....	7
1.3.1 HCMV Virions.....	7
1.3.2 Defective particles: Non-infectious enveloped particles (NIEPs)	8
1.3.3 Defective particles: Dense bodies (DBs)	8
1.3.4 HCMV lytic replication cycle – overview.....	10
1.4 HCMV infection <i>in vivo</i>	12
1.4.1 Immune response to HCMV.....	12
1.4.2 Tropism of HCMV <i>in vivo</i>	13
1.5 Clinical manifestations of HCMV infection.....	15
1.5.1 Congenital and perinatal infection.....	16
1.5.2 Infection in the immuno-compromised	16
1.6 Current anti-HCMV Therapies: successes and short-comings	17
1.6.1 Anti-viral chemotherapeutics.....	17
1.6.2 Immunoprophylaxis	17
1.7 Historic limitations in HCMV research.....	20
1.7.1 Adaptation of HCMV to cell culture.....	20
1.7.2 Definition of wildtype and ‘laboratory-adapted’ strains, and strain variants	20
1.7.3 Genetic heterogeneity between ‘laboratory-adapted’ strains.....	21
1.8 HCMV genetics	22
1.8.1 Genome organisation	22
1.8.2 HCMV coding potential.....	23
1.8.3 Natural intra-strain genetic variation.....	27
1.9 BAC-mid cloning of HCMV genomes	28

1.10 HCMV genes that are essential, dispensable, and/or inhibitory, for virus growth in cell culture	29
1.10.1 Cell-type specific functions	30
1.10.2 Mutations acquired during adaptation of HCMV to cell culture	30
1.11 Biological differences between wildtype and 'laboratory-adapted' HCMV strains.....	32
1.11.1 Restricted tropism	32
1.11.2 A genetic determinant of broad-tropism: the UL128L genome region	33
1.11.3 Restricted virulence	36
1.12 HCMV virus particle composition and morphogenesis.....	39
1.12.1 HCMV particle nucleocapsid compartment	39
1.12.2 HCMV particle tegument compartment	40
1.12.3 HCMV particle envelope.....	42
1.13 HCMV particle envelope glycoproteins in binding, entry and maturation/egress.....	43
1.13.1 Glycoprotein complex family gCI	44
1.13.2 Glycoprotein complex family gCII.....	44
1.13.3 Glycoprotein complex family gCIII	45
1.13.4 Entry of HCMV particles into cells	46
1.13.5 Progeny virion maturation and egress	49
1.14 Variant gH/gL glycoprotein complexes encoded by HCMV	49
1.14.1 gCIII; gH/gL/gO.....	49
1.14.2 gH/gL/gpUL128/gpUL130/gpUL131A.....	50
1.14.3 Roles of gH/gL complex variants in entry of cell-free HCMV into cells.....	52
1.14.4 Roles of gH/gL complex variants in cell-to-cell spread of HCMV.....	54
1.14.5 Roles of gH/gL complex variants in cell-free HCMV production.....	54
1.14.6 The role of HCMV gH/gL complex variants during virus replication: summary	55
1.15 Biologically relevant HCMV strains currently available for research <i>in vitro</i>	55
1.15.1 'Broad-tropism' BAC-cloned strains	55
1.15.2 Cloning of the wildtype Merlin genome as a BAC-construct	56
1.15.3 Limitations of the wildtype Merlin-BAC construct.....	56
1.16 Aims of current study	57
2 Methods and Materials	59
2.1 Solutions, buffers and media	61
2.2 Cells and Viruses.....	63

2.2.1	Cells.....	63
2.2.2	HCMV strains	63
2.2.3	Adenovirus vectors	64
2.3	Tissue and virus culture	64
2.3.1	Human cell master stocks	64
2.3.2	Tissue culture reagent preparation and manipulations.....	65
2.3.3	Growth of human cells.....	65
2.3.4	Infections.....	66
2.3.5	Monitoring infections.....	67
2.4	Molecular Biology Techniques.....	69
2.4.1	<i>In silico</i> analysis of nucleic acid and amino acid sequences	69
2.4.2	Molecular biology reagent storage and preparation.....	69
2.4.3	Nucleic acid modification and amplification.....	69
2.4.4	Primer Design.....	69
2.4.5	DNA gel electrophoresis	69
2.4.6	Purification of PCR-amplified DNA	70
2.4.7	Determination of DNA concentration	70
2.5	Recombineering of BAC-cloned HCMV and Adenovirus genomes	70
2.5.1	Phage Lambda Red-mediated Homologous Recombination.....	70
2.5.2	Design and preparation of DNA sequences for transformation	71
2.5.3	Preparation and transformation of electro-competent <i>E.coli</i> SW102s	72
2.5.4	Screening for transformants	72
2.5.5	Analysis of transformant genome integrity and coding potential.....	73
2.5.6	Maintenance of transformed <i>E.coli</i> SW102s as glycerol stocks.....	75
2.5.7	Extraction and purification of BAC DNA from <i>E.coli</i> SW102s	75
2.6	Reconstitution, growth and preparation of BAC-derived virus.....	76
2.6.1	Reconstitution of virus from BAC-cloned HCMV genomes.....	76
2.6.2	Growth of HCMV stocks	77
2.6.3	Preparation of infected cell culture fractions and sub-fractions	77
2.6.4	Preparation of concentrated HCMV working stocks	77
2.6.5	HCMV infectivity quantification	78
2.6.6	Growth and preparation of Adenovirus vector stocks	79
2.6.7	Titration of Adenovirus stocks	79
2.7	HCMV-derived nucleic acid extraction.....	80
2.7.1	Extraction of viral nucleic acids from infected cells or cell-free virions.....	80
2.8	Nucleic acid sequence analysis.....	81
2.8.1	PCR amplification of sequences from HCMV genome regions	81

2.8.2	Cloning of PCR products	84
2.8.3	RT-PCR of RNA transcripts and production of cDNA libraries	84
2.8.4	DNA sequencing	84
2.9	Biochemical analysis of HCMV particles	85
2.9.1	Purification of virus derived particles	85
2.9.2	Immunoblot detection of proteins	87
2.9.3	Mass spectrometry analysis	88
3	Growth Characteristics of Viruses Reconstituted from ‘Broad-Tropism’ BAC-cloned Strains.....	91
3.1	Growth phenotypes of ‘broad-tropism’ HCMV strains in fibroblasts.....	94
3.1.1	Cell-to-cell spread.....	94
3.1.2	Dissemination kinetics and cell-free virus production	94
3.1.3	Validation of BAC-cloned strain growth phenotypes in fibroblasts	96
3.1.4	Conclusion.....	96
3.2	Growth phenotype of ‘broad-tropism’ HCMV strains in epithelial cells	98
3.2.1	Cell-to-cell spread.....	98
3.2.2	Dissemination kinetics and cell-free virus production	98
3.2.3	Validation of BAC-cloned strain growth phenotypes in epithelial cells.....	100
3.2.4	Conclusion.....	100
3.3	Role of variant UL128L in the differential growth phenotypes of each strain	102
3.4	Comparative nucleotide sequence analysis for the identification of <i>in vitro</i> acquired mutations, or other variations that conveyed distinct growth phenotype of each strain.....	103
3.5	Genome regions contributing to the growth phenotype of virus from the TR-BAC	106
3.5.1	Impact of UL131A 5’-UL132 5’ genome region on RV-TR growth phenotype	106
3.5.2	Conclusion.....	108
3.6	Nucleotides contributing to the growth phenotype of viruses from TB40-BAC4110	
3.6.1	Impact of TB40-BAC4 UL128 G>T variation in fibroblasts	112
3.6.2	Impact of TB40-BAC4 UL128 intron 1 G>T in epithelial cells	114
3.6.3	Validation of TB40-BAC4 UL128 intron 1 G>T growth phenotype impact ...	116
3.6.4	Conclusion.....	116
3.7	Nucleotides contributing to the growth phenotype of viruses from FIX-BAC	118

3.7.1	Impact of FIX UL130 A>G In Fibroblasts	120
3.7.2	Impact of FIX UL130 A>G in epithelial cells.....	122
3.7.3	Validation of FIX UL130 A>G growth phenotype impact	124
3.7.4	Conclusion.....	124
3.8	Summary	126
4	Impact of UL128L Sequence Variations on Pentameric Complex Functionality	127
4.1	Variant UL128L ORF expression efficiency	129
4.1.1	TB40-BAC4 intron G>T reduces UL128 mRNA and gpUL128L production..	132
4.1.2	UL130 sequence variations, including FIX UL130 A>G, do not impact gpUL130 production.....	133
4.1.3	Each variant UL131A is expressed with similar efficiency.....	134
4.1.4	Conclusion.....	135
4.2	Impact of UL128L coding potential on pentameri complex assembly and/or incorporation into virions	135
4.2.1	Purification of virions from virus working stocks.....	135
4.2.2	Pentameric complex content of virions	138
4.2.3	Conclusion.....	138
4.3	Impact of virion gH/gL/gpUL128/gpUL130/gpUL131A content on epitheliotropism.....	140
4.3.1	Conclusion.....	140
4.4	Impact of UL128L coding potential on progeny virion production and/or release from epithelial cells	142
4.4.1	Relative CRV and CAV infectivity produced during infection of epithelial cells	142
4.4.2	Relative epitheliotropism of CRV and CAV populations produced during infection of epithelial cells	144
4.4.3	Conclusion.....	144
4.5	Impact of UL128L coding potential on progeny virion production and/or release from fibroblast cells.....	145
4.5.1	Relative CRV and CAV infectivity produced by viruses containing non-intact or wildtype UL128L ORFs	145
4.5.2	Relative CRV and CAV produced by viruses containing TB40-BAC4 UL128L sequence features	147
4.5.3	Relative CRV and CAV produced by viruses containing FIX-BAC UL128L sequence features	147
4.5.4	Tropism of CRV and CAV sub-populations produced in fibroblasts.....	147

4.5.5 Conclusion.....	148
4.6 Summary	150
5 Genetic Stability of Virus from BAC-cloned Strains and Novel Merlin-BAC	
Variants <i>in vitro</i>	151
5.1 Adaptation of TR-BAC-derived viruses during previous growth kinetics assays	153
5.1.1 Virus reconstituted from the TR-BAC adapted during a single passage in HFFFs	153
5.1.2 Virus reconstituted from the TR-BAC adapted during a single passage in RPE-1s	154
5.1.3 Conclusion.....	155
5.2 Genetic stability of viruses in epithelial cell culture	155
5.2.1 Viruses containing stably-incorporated BAC-vector sequences rapidly adapt during propagation in epithelial cell culture, involving mutation at the site of BAC-vector insertion.....	156
5.2.2 Viruses from the Merlin-BAC variants rapidly adapt during propagation in epithelial cell culture, involving mutation in the U_L/b' genome region.....	157
5.2.3 Conclusion.....	158
5.3 Genetic stability of viruses in fibroblast cells.....	161
5.3.1 Viruses containing stably-incorporated BAC-vector sequences rapidly adapt during sequential passage in fibroblasts, involving mutation at the site of BAC-vector insertion	161
5.3.2 Virus from the Merlin-BAC variant lacking intact UL128L is stable in fibroblasts; virus from the Merlin-BAC variants containing wildtype UL128L are prone to rapid mutation in this genome region, and others, during sequential passage in fibroblasts.....	162
5.3.3 TB40-BAC4 UL128L is not stable when expressed in the Merlin genome background in fibroblasts, yet the unique UL128 intron 1 G>T variation does stabilise Merlin UL128L ORFs in fibroblasts	164
5.3.4 FIX-BAC UL128L is stable when expressed in the Merlin genome background in fibroblasts, and the unique UL130 A>G variation also stabilises Merlin UL128L ORFs in fibroblasts.....	165
5.3.5 Conclusion.....	165
5.4 Stabilisation of genomes by repression of ORFs most inhibitory to virus propagation <i>in vitro</i>.....	167
5.4.1 Genetic stability of virus produced in tetracycline repression culture system	167

5.4.2	Growth phenotype of virus produced during subsequent infection with virus stocks produced in tetracycline repression culture system.....	168
5.4.3	Conclusion.....	169
5.5	Summary	170
6	The HCMV Virion Proteome	171
6.1	Optimising recovery of virions purified over gradients	173
6.1.1	Concentration of virus by high-speed centrifugation (HSC) or Ultrafiltration (UF)	174
6.1.2	Recovery and purification of virions following ultra-centrifugation of HSC and dUF concentrated viruses over gradients	174
6.1.3	Conclusion.....	176
6.2	Production and preparation of virions for mass spectrometry analysis	176
6.3	Merlin virion proteome.....	177
6.3.1	Virion proteins identified with high-confidence.....	177
6.3.2	Proteins identified with low confidence.....	178
6.3.3	Proteins previously identified in HCMV virions, or homologous proteins identified virions of other CMVs, but not on MS analysis of Merlin virions	179
6.4	Relative abundance of proteins in virions from Merlin variants.....	189
6.4.1	Inter-assay variation	189
6.4.2	Impact of variations in gpRL13 content	189
6.4.3	Impact of variations in pentameric complex content	191
6.5	Summary	201
7	Discussion	203
7.1	The need to develop the Merlin BAC – why not use other ‘broad tropism’ strains?	205
7.2	Wildtype UL128L is inhibitory to growth of virus – but is Merlin UL128L unusual?	207
7.3	The alternative gH/gL complex content of virions is a major factor dictating the virus growth phenotype	208
7.3.1	The role of the trimer and its impact on titration data	208
7.3.2	Impact of pentamer/trimer ratios on cell-free virus entry	209
7.3.3	Impact of pentamer/trimer on cell-free virus production	211
7.3.4	Impact of pentamer/trimer ratio on cell-to-cell spread.....	212
7.3.5	Impact of pentamer/trimer ratio on wildtype virus during natural infection	213
7.4	Novel Merlin BAC reagents: assessment, propagation and utility.....	214
7.4.1	Merlin viruses containing TB40-BAC4 UL128L sequences.....	214

7.4.2	Merlin viruses containing FIX UL128L sequences	215
7.4.3	Propagation of novel Merlin virus stocks.....	216
7.4.4	Utility of the novel Merlin reagents	218
7.5	Scope for further Merlin BAC development.....	219
7.6	Composition of Merlin virions – a closer reflection of wildtype virus <i>in vivo</i>	
	221	
8	References.....	225
9	Appendix I: Recombineering experiment design	227
9	228	
10	Appendix II: Impact of variant UL128L on virus growth phenotype.....	235
	A.II.1 Impact of variant UL128L in fibroblasts	238
	A.II.1.1 Cell-to-cell spread.....	238
	A.II.1.2 Dissemination kinetics and cell-free virus production	238
	A.II.1.3. Validation of variant UL128L impact in fibroblasts	240
	A.II.2.1 Impact of variant UL128L in epithelial cells	242
	A.II.2.2 Cell-to-cell spread.....	242
	A.II.2.3 Dissemination kinetics and cell-free virus production	242
	A.II.2.4. Validation of variant UL128L impact in epithelial cells.....	244
11	Appendix III: Plaque size comparison data	247
12	Appendix IV: Bioinformatics analysis of TB40-BAC4 UL128 intron 1 G>T	257
13	Appendix V: Whole genome sequencing protocol.....	263
14	Appendix VI: MS- analysis of Merlin virions.....	267
15	Appendix VII: SILAC-MS Raw Data	271
16	Appendix VIII: Publications	277

SYMBOLS/ACRONYMS/ABBREVIATIONS

°C – degrees Celsius

Δ – deletion

μM – micromolar

μL – microlitre

μm – micrometre

A

aa – amino acid

Ab – Antibody

AdZ-5 – Adenovirus vector

AIDS – Acquired Immunodeficiency Syndrome

ARPE – Aotaki-Keen Retinal Pigment Epithelial cells

ATCC – American Tissue Culture Collection

B

BAC – Bacterial Artificial Chromosome

BLAST – Basic Local Alignment Search Tool

BM – Bone Marrow

bp – base pairs

BSA – Bovine Serum Albumin

C

CAV – Cell-Associated Virus

CD – Cluster of Differentiation

CCMV – Chimpanzee Cytomegalovirus

CID – Cytomegalic Inclusion Disease

cm – centimetre

CMV – Cytomegalovirus

CNS – Central Nervous System

CPE – Cytopathic Effect

CRV – Cell-free Released Virus

cUF – centrifugal Ultra-Filtration

D

Da – Dalton

DAXX – Death-Associated protein 6

DB – Dense Body

DE/ β – Delayed Early

DMEM – Dulbecco's Modified Eagle's Medium

DMSO – Dimethyl Sulfoxide

DNA – Deoxyribonucleic Acid

DNase – Deoxyribonuclease

dNTP – Deoxynucleotide Triphosphate

DTT – Dithiothreitol

dUF – dialysis Ultra-Filtration

DC – Dendritic Cell

DURP – Deoxyuridine triphosphatase-Related Protein

E

EBV – Epstein-Barr Virus

EC – Endothelial Cell

EDTA – EthyleneDiamineTetraacetic Acid

eGFP – enhance Green Fluorescing Protein

EGFR – Epidermal Growth Factor Receptor

EM – Electron Microscopy

ER – Endoplasmic Reticulum

ERGIC – Endoplasmic Reticulum-Golgi Intermediate Compartment

F

FBS – Fetal Bovine serum

FS – Frameshift

FS-ins – Frameshifting insertion

G

g - Gravity

gC – glycoprotein Complex

GCV – Ganciclovir

GM-CSF – Granulocyte Macrophage Colony-Stimulating Factor

gp – glycoprotein

GPCR – G-Protein Coupled Receptor

H

hCAR – human Coxsackievirus-Adenovirus-Receptor

HCMV – Human Cytomegalovirus

HF – Human Fibroblast

HFFF – Human Foetal Foreskin Fibroblast

HHV – Human Herpesvirus

HIG – Hyperimmune Globulin

HIV – Human Immunodeficiency Virus

HRP – Horseradish Peroxidase

hr(s) – hour(s)

HSC – High-Speed Centrifugation

HSCT – Hematopoietic Stem Cell Transplantation

HSPG – Heparan Sulphate Proteoglycan

HSV-1 – Herpes Simplex Virus 1

HSV-2 – Herpes Simplex Virus 2

hTERT – human Telomerase Reverse Transcriptase

I

ID – Identity

IF – Immuno-Fluorescence

IFN- β – Interferon Beta

IFN- γ – Interferon Gamma

IHW – International Herpesvirus Workshop

Ig – Immunoglobulin

IL2 – Interleukin 2

IL6 – Interleukin 6

IE/ α – Immediate Early

IMS – Industrially Methylated Spirits

IPTG – Isopropyl β -D-1-thiogalactopyranoside

IRES – Internal Ribosome Entry Site

IRL/IR_L – Internal Repeat Long

IRS/IR_S – Internal Repeat Short

K

Kb – Kilobase

Kbp – Kilobase pairs

KDa – Kilodalton

KSHV – Kaposi's Sarcoma-associated Herpesvirus

KV – Kilovolt

L

L – Litre

.L⁻¹ – per Litre

L/ γ - Late

LB – Luria's Broth

LMP – Lower Matrix Protein

LTP – Largest Tegument Protein

M

M – Molar (moles per litre)

mAb – monoclonal Antibody

MCMV – Murine Cytomegalovirus

MCP – Major Capsid Protein

mCP – Minor Capsid Protein

mC-BP – Minor Capsid Protein Binding Protein

MDa – Megadalton

MEM – Minimal Eagle's Medium

mg – milligram

MHC – Major Histocompatibility Complex

MICA – MHC class I polypeptide-related sequence A

MICB – MHC class I polypeptide-related sequence A

MIE – Major Immediate Early

MIEP – Major Immediate Early Protein

min(s) – minutes

miRNA – micro RNA

mL – milliliter

mM – millimolar

M/M – Macrophage/Monocyte

MOI – Multiplicity of Infection

MOPS - 3-(N-morpholino)propanesulfonic acid

MRC – Medical Research Council

mRNA – messenger RNA

MRP – Multiple Regulatory Protein

MS – Mass Spectrometry

mut – mutated

MVP – Mixed Virion Preparation

N

NIEP – Non-infectious Enveloped Particle

NK cell – Natural Killer cell

NLS – Nuclear Localisation Sequences

nm – nanometer

NSP – Nucleocapsid-proximal Stabilization Protein

NUC – deoxyribonuclease

nt – nucleotide

O

OD – Optical Density

OHSU – Oregon Health & Science University

ORF – Open Reading Frame

P

PAGE – Polyacrylamide Gel Electrophoresis

PBS – Phosphate Buffered Saline

PBS-T – Phosphate Buffer Saline + tween 20

PCR – Polymerase Chain Reaction

PDGFR – Platelet-Derived Growth Factor Receptor

PEG – Polyethylene glycol

PES – Polyethersulfone

PFA – Paraformaldehyde

PFU – Plaque forming units

p.i. – post-infection

PKR – Protein Kinase R

PMNL – Polymorphonuclear Leukocyte

pmol – picomolar

POL – Polymerase

PORT – Portal protein

pp – phosphoprotein

PPS – Polymerase Processivity Sub-unit

p.t. – post-transfection

PVDF - Polyvinylidene fluoride

Q

qPCR – quantitative PCR

R

RCMV – Rat Cytomegalovirus

RhCMV – Rhesus Cytomegalovirus

RFLP – Restriction Fragment Length Polymorphism

RAdZ – Recombinant Adebovirus Vector

RNA – Ribonucleic acid

RNAse – Ribonuclease

RPE – Retinal Pigment Epithelial Cell

RPM – Revolutions per minute

RT-PCR – Reverse Transcriptase PCR

RV – Reconstituted Virus

S

s – seconds

SCMV – Simian Cytomegalovirus

SCP – Smallest Capsid Protein

SD – Standard Deviation

SDS - Sodium dodecyl sulfate

SEM – Standard Error Mean

SILAC – Stable Isotope Labeling by/with Amino acids in Cell culture

siRNA – small inhibitory RNA

SNP - Single Nucleotide Polymorphism

SSB – Single-stranded DNA Binding Protein

SU – Surface Sub-unit

T

TAE – Tris-acetate-EDTA

tet – tetracycline

THV – Tupaia Herpesvirus

TM – Transmembrane

TNF – Tumour Necrosis Factor

TNFR – Tumour Necrosis Factor Receptor

TRL/TR_L – Terminal Repeat Long

TRS/TR_S – Terminal Repeat Short

U

UF – Ultra-Filtration

UL/U_L – Unique Long

UMP – Upper Matrix Protein

UNG – Uracil-DNA Glycosylase

US/U_S – Unique Short

UV – Ultraviolet

V

V – Volts

VAC – Virus Assembly Compartment

VEP – Virion Egress Protein

VPK – Viral Protein Kinase

v/v – volume/volume

VZV – Varicella Zoster Virus

W

wt – wildtype

wv – wavelength

w/v – weight/volume

X

X-gal – 5-bromo-4-chloro-3-indolyl-beta-D-galacto-pyranoside.

1 Introduction

1.1 Human Cytomegalovirus (HCMV) - Human Herpesvirus 5 (HHV5)

1.1.1 Taxonomy

Human Cytomegalovirus (HCMV) is one of the nine currently recognised herpesviruses that infect humans. The herpesviruses are a large group of DNA viruses with diverse animal hosts, collectively assigned to the order *Herpesvirales* (McGeoch, Rixon et al. 2006, Davison, Eberle et al. 2009). Herpesviruses of mammals, reptiles, and birds are assigned to the family *Herpesviridae* that contains 3 major sub-families distinguished by nucleotide sequence and biological characteristics: the *Alphaherpesvirinae*, *Betaherpesvirinae* and *Gammapherpesvirinae* (Roizmann, Desrosiers et al. 1992, Davison 2010). The human herpesviruses (HHV1-8; including HHV6A and HHV6B) include representatives of each sub-family, with HCMV (HHV5) the prototype species of the betaherpesvirinae and the genus Cytomegalovirus.

1.1.2 Evolution, phylogeny and related species

Studies into the evolutionary origins and phylogenetic relationship between the *herpesviridae* members have been based on the alignment of amino acid sequences encoded by genes conserved amongst the herpesviruses (Figure 1.1.). Data from these studies suggest that the *alpha*-, *beta*- and *gammapherpesvirinae* lineages evolved from a common ancestor that existed 350-400 million years ago, with genera within each sub-family having arisen prior to the mammalian radiation event that occurred 60-80 million years ago (McGeoch, Cook et al. 1995, McGeoch and Gatherer 2005). The closest relatives of HCMV are the other members of the genus Cytomegalovirus that comprises viruses of apes and other primates. More distantly related viruses with biological similarities to HCMV are often called ‘cytomegaloviruses’, yet are not members of the Cytomegalovirus genus. These include murine CMV (MCMV) and rat CMV (RCMV) from the genus Muromegalovirus, and guinea pig CMV (GPCMV) yet assigned to any genus (Davison 2002, Davison 2013).

Since they are derived from a common ancestor the herpesviruses share many biological qualities, and several aspects of HCMV biology have been elucidated by comparison to other human herpesviruses and betaherpesviruses of other animals. One common biological feature is the establishment of life-long persistent infections, contributing to the survival of HCMV in host populations. However, herpesviruses have also undergone divergent speciation during co-evolution with their respective hosts; this was suggested by the observation that the branching pattern of the mammalian herpesvirus phylogenetic trees closely overlapped the phylogenetic trees for the corresponding mammalian host lineages (McGeoch, Cook et al. 1995, McGeoch, Dolan et al. 2000). As a result, humans are the exclusive host of HCMV, with the evolutionary success of this virus in part due to the acquisition of mechanisms to evade or subvert the human host immune response; reviewed in (Vossen, Westerhout et al. 2002).

1.2 Discovery, isolation, and propagation of HCMV *in vitro*

The earliest documented accounts of HCMV pathogenesis date back to the late 19th century when HCMV-infected cells were observed during histological examination of the kidney tissues of a stillborn, and the parotid glands of infants suffering unknown diseases. However, these observations were reported some years later (Ribbert 1904) following independent reports of similar clinical findings (Jesionek 1904). Each report described abnormal cellular structures containing a large central nuclear body surrounded by a clear halo, together termed ‘nuclear inclusions’ or ‘Owl’s eye’ inclusions, as well as cytoplasmic inclusions at juxtanuclear sites (Figure 1.2.). These cellular structures were initially considered ‘protozoan-like’, however investigation into their occurrence and intra-host distribution led to their identification as human cells, albeit with an unusual morphology that was termed ‘cytomegalia’ (*cyto* – cell, *megalia* – large) (Goodpasture and Talbot 1921). A viral aetiology was postulated (Von Glahn 1925) on comparison to reports of similar inclusion-containing cells from individuals with then known herpesvirus infections (Lipscheutz 1921). The presentation of cytomegalia was soon correlated with a lethal congenital infection in humans and other animals, for which the term ‘cytomegalic inclusion disease’ (CID) was applied, and it was also recognised that similar phenomena occurred in other animals (Wyatt, Saxton et al. 1950). The term ‘cytomegalic inclusion disease virus’ (CIDV) was proposed, as was ‘salivary gland virus’ (SGV) as cytomegalic cells were consistently identified in the salivary glands of CID patients.

Early attempts to isolate and propagate human CIDV/SGV were generally unsuccessful; this may be attributed to the use of non-human tissues together with the high degree of host-specificity displayed by this virus, whilst various human tissue types targeted *in vivo* also did not support propagation *in vitro*. However successful isolation and serial propagation of CIDV/SGV *in vitro* was achieved following the development of techniques for the routine culture of diploid fibroblast cells, reported almost simultaneously by three independent groups (Rowe, Hartley et al. 1956, Smith 1956, Weller, Macauley et al. 1957). During each of these studies, propagation of the same clinical isolates in non-human tissues and non-fibroblast human tissues was conclusively shown to be inefficient, indicating that clinical HCMV isolates displayed specificity for fibroblasts *in vitro*. Serological testing demonstrated antigenic heterogeneity amongst these isolates to identify distinct strains of a single viral species that was subsequently named Cytomegalovirus (Weller 1960, Weller 1970). The prototype laboratory strain AD169 was derived from these earliest isolations, and has since been the virus reagent of choice for historic HCMV research (Rowe, Huebner et al. 1953, Rowe, Hartley et al. 1956). However during these works, AD169 and the other clinical isolates were subjected to extensive passaging to produce virus stocks, and it is now widely appreciated that this occurred with concomitant selection of mutations during adaptation to cell culture (see Section 1.7.).

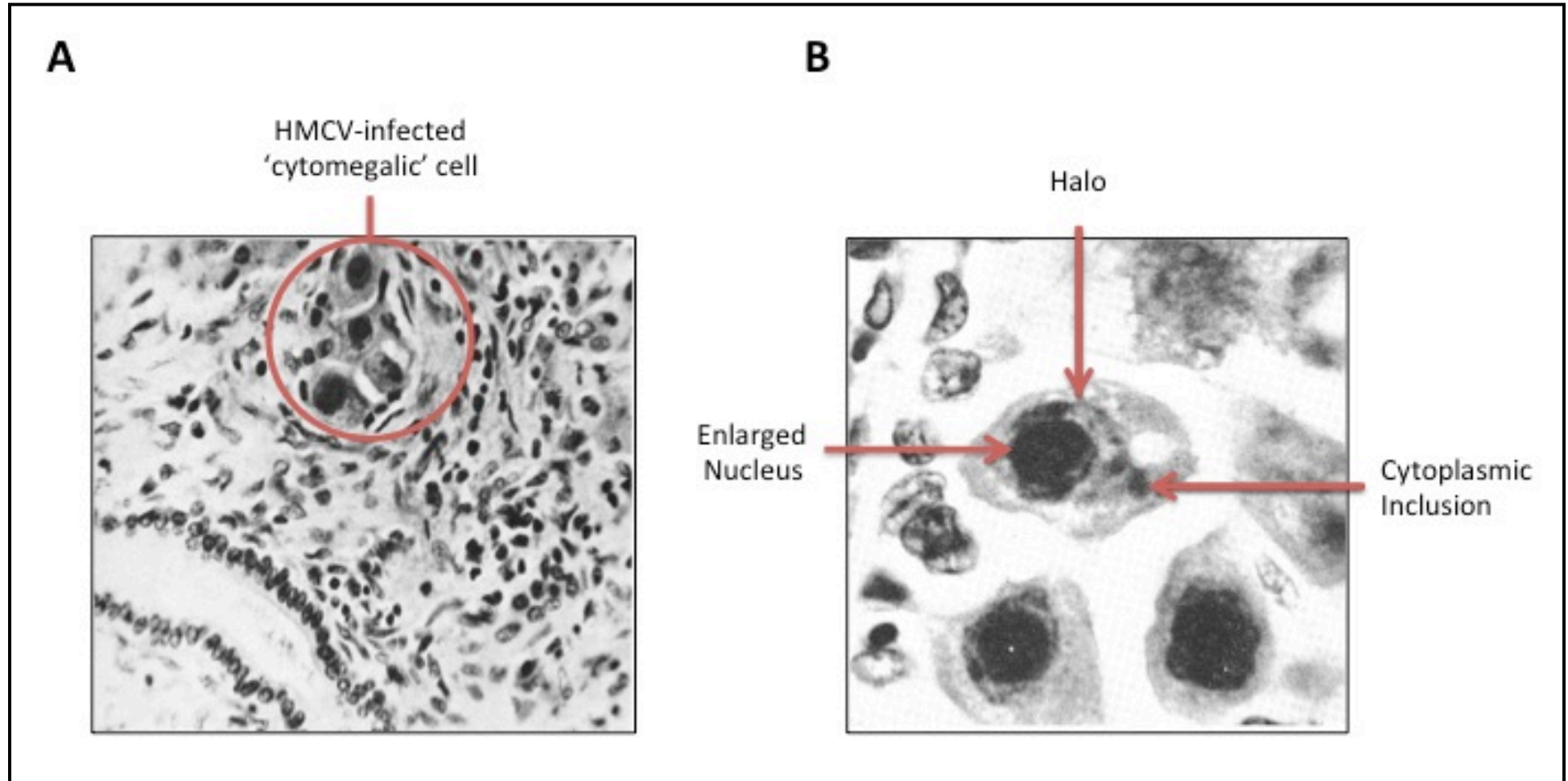


Figure 1.2. Cytomegalic cells. Haematoxylin and eosin stained tissue biopsies derived from HCMV infected individuals with CID. (A) Infection foci (encircled) of liver bile duct cells (at 400X magnification). (B) HCMV infected renal tubule epithelial cells (at 960X magnification), displaying enlarged nucleus and surrounding clear halo (nuclear, or 'Owl's eye' inclusion), as well cytoplasmic inclusions. Images reproduced from (Wyatt, Saxton et al. 1950), with permission of Elsevier publishers.

1.3 Classification of HCMV as a herpesvirus

Prior to the availability of nucleic acid sequencing technologies, virus species classification was based on observation of common biological properties (i.e. pathologies, replication characteristics, host range etc.), as well as the biochemical and architectural features of virus particles. Under these criteria, HCMV was tentatively classified based on biological similarities with then known herpesviruses (Andrews 1961), and more conclusively when electron microscopy (EM) analysis of HCMV particles revealed structural similarities with herpes simplex virus particles (Smith and Rasmussen 1963). An interesting observation made during the earliest analysis of HCMV particles was a high particle: infectivity ratio (Smith and Rasmussen 1963). Later investigations, including those aided by gravity-flow sedimentation techniques for the purification of viral particles (Figure 1.3.A), identified three distinct HCMV particle variants (Craighead, Kanich et al. 1972, Talbot and Almeida 1977, Irmiere and Gibson 1983). EM analysis of purified infectious HCMV particles (virions) further demonstrated the four-layered architecture definitive of the herpesviruses (see section 1.3.1). Characterisation of the two alternative particle variants demonstrated that they were morphologically and/or biochemically distinct from virions, and represented defective non-infectious particles (see sections 1.3.2 and 1.3.3). However, these investigations were based on virus that had been propagated in cell-culture, and neither defective particle variant has been identified in clinical samples. Hence the significance of these defective particle variants during natural infection remains speculative.

1.3.1 HCMV Virions

HCMV virions (Figure 1.3.B) are the largest and most complex of all studied human herpesviruses, ~200-230 nm in diameter, with pleomorphic quasi-spherical morphology (Irmiere and Gibson 1983, Chen, Jiang et al. 1999). The four-layered herpesvirion architecture includes, from the innermost regions outwards: a dsDNA genome; a highly ordered protein structure, termed the capsid; a matrix predominantly composed of viral proteins, termed the tegument; and a limiting membrane envelope (Roizmann, Desrosiers et al. 1992, Pellet 2013). As with other herpesviruses, the linear dsDNA genome of HCMV virions is arranged in a torus of concentric hexagonal shells within the capsid (Furlong, Swift et al. 1972); these two layers form the core structure termed the nucleocapsid. HCMV contains the largest genome of any human herpesvirus (~236 Kbp in length), which is in turn packaged with considerable density in the largest capsid of any human herpesvirus (~130 nm in diameter) (Chen, Jiang et al. 1999, Bhella, Rixon et al. 2000, Dolan, Cunningham et al. 2004). Like all herpesviruses, HCMV capsids are assembled from 162 capsomeres (capsid sub-units) and 320 hetero-trimeric protein triplexes, and these are arranged into icosahedral lattices with T=16 symmetry to display 20 triangular faces and 12 vertices (Chen, Jiang et al. 1999, Brown and Newcomb 2011, Pellet

2013). 161 capsomeres are multimers of a single protein; 150 hexamers form the capsid faces and walls, whilst 11 pentamers are located at the vertices. One specialised capsomere at the final vertex forms a channel through which the genome is inserted during virion assembly, and released following *de novo* infection. The HCMV virion tegument is ~50 nm thick and arranged into two distinct layers: the innermost nucleocapsid-proximal layer contains densely packaged materials with a highly ordered arrangement; the envelope proximal layer contains less densely materials with amorphous arrangement (Irmiere and Gibson 1983, Chen, Jiang et al. 1999, Yu, Shah et al. 2011). The HCMV virion envelope is a ~10 nm thick host-cell derived lipid bilayer, and contains an array of viral encoded glycoproteins as a further hallmark of herpesvirion morphology (Irmiere and Gibson 1983, Pellet 2013).

1.3.2 Defective particles: Non-infectious enveloped particles (NIEPs)

Non-infectious enveloped particles (NIEPs) (Figure 1.3.C) are biochemically and structurally similar to virions, however they lack a copy of the viral genome and are therefore unable to support viral replication (Irmiere and Gibson 1983). Compared to virions, NIEPs are less densely packaged and are separated from other viral particles by gravity-flow sedimentation as slower migrating particles (Figure 1.3.A). In place of the fully matured nucleocapsid, NIEPs contain a smaller immature capsid variant (Irmiere and Gibson 1985), and are therefore distinguishable on EM analysis by their less electron dense core.

1.3.3 Defective particles: Dense bodies (DBs)

Dense bodies (DBs) (Figure 1.3.D) are composed of enveloped tegument materials, and lack both the viral genome and capsid (Craighead, Kanich et al. 1972). In contrast to NIEPs, DBs are more densely packaged than virions, and therefore migrate at a greater rate during gravity-low sedimentation (Figure 1.3.A). These defective virus particles are distinguishable from both virions and NIEPs on EM analysis by the presence of electron dense material throughout the entirety of their structure. DBs are also the most heterogeneous particle variant, existing in a range of sizes that are often up to twice the diameter of virions and NIEPs (Craighead, Kanich et al. 1972, Talbot and Almeida 1977).

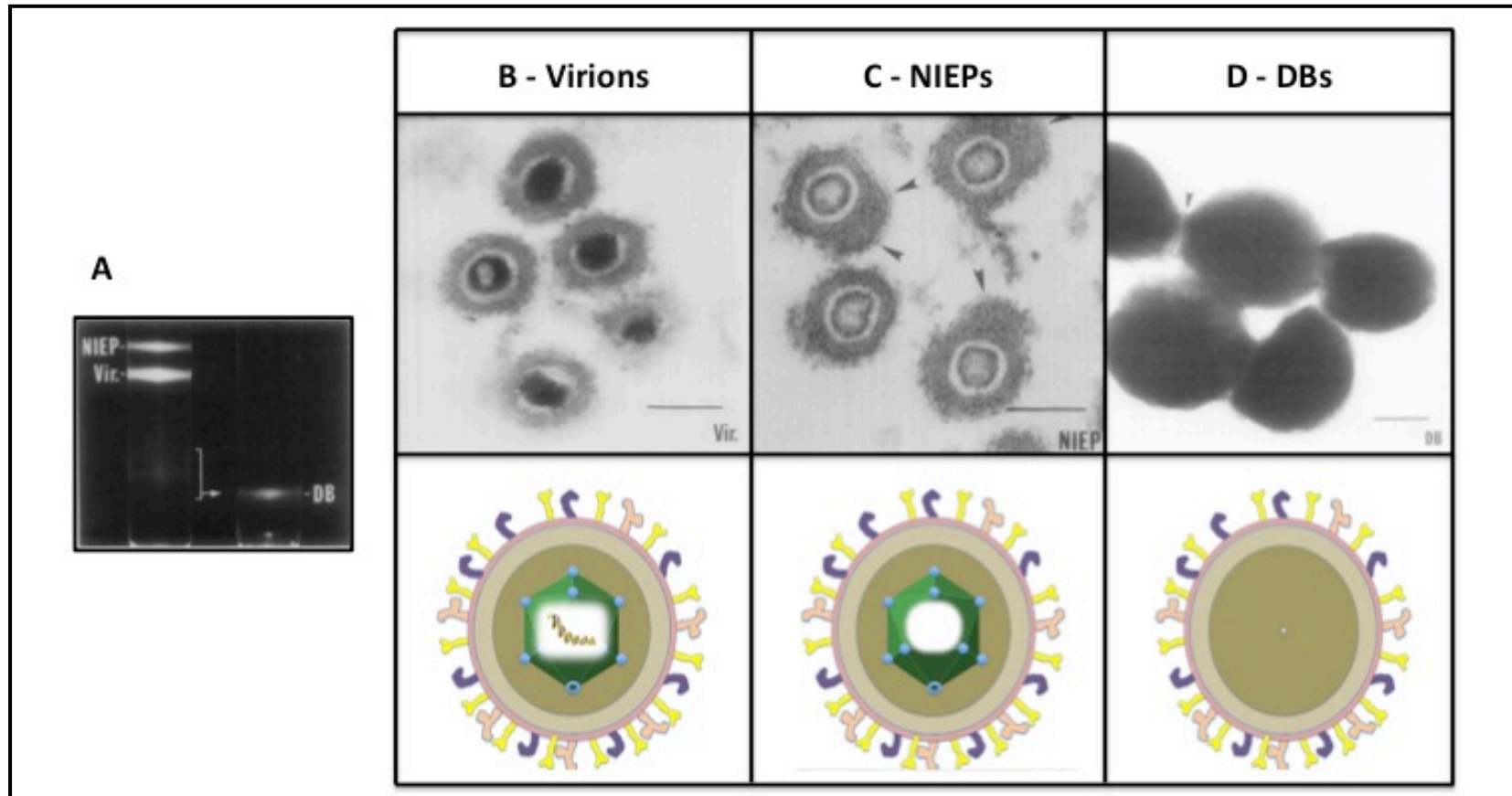


Figure 1.3. HCMV particle variants observed *in vitro*. (A) Separation of virus particle variants by gravity-flow sedimentation over gradients. (B) EM image and schematic of infectious virion architecture. (C) EM image and schematic of defective NIEPs. (D) EM image and schematic of defective DBs. In virions, the linear dsDNA genome is encased in the capsid, together forming the nucleocapsid. The capsid is surrounded by tegument arranged in two distinct layers: a densely packed and highly ordered capsid-proximal layer (dark brown), and the amorphous envelope proximal layer (light brown). The envelope (pink) contains an array of embedded viral encoded glycoproteins (pink, yellow, purple). NIEPs differ from virions in containing an immature capsid devoid of an inserted genome. DBs lack both a viral genome and also the capsid structure. Gradient purified virus particles and EM images adapted from (Irmiere and Gibson 1983) with the permission of Elsevier publishers.

1.3.4 HCMV lytic replication cycle – overview

The core mechanisms employed at each phase in the HCMV lytic (productive) replication cycle (Figure 1.4.) are common amongst the herpesviruses. Infection with cell-free HCMV commences with virions binding onto cells, then subsequently penetration occurs following fusion between the virion envelope and cellular membranes; either by direct fusion at the plasma membrane, or with endosomal membranes after internalisation by receptor-mediated endocytosis, depending on cell type (Compton 2004, Vanarsdall and Johnson 2012). The binding and membrane fusion mechanisms are orchestrated by virion envelope glycoproteins via interactions with receptor and entry mediator molecules on the host cell membranes. HCMV can also spread by the direct cell-to-cell route. This is believed to involve the same viral envelope glycoproteins that facilitate the entry of cell-free virus, though when they are presented at the infected cell plasma membrane to interact with closely opposed receptor and entry-mediator molecules on an adjacent cell (Kinzler and Compton 2005).

Following the deposition of virion contents into cells, the outermost tegument materials diffuse to specific sites where they perform a wide range of activities (e.g. priming the cellular environment for viral replication, and also subverting both intrinsic and innate cellular responses to viral pathogens) (Kalejta 2008, Kalejta 2013). The nucleocapsid is then transported to the nucleus where herpesvirus genome transcription and replication occur. Viral lytic genes are transcribed in a temporal cascade: the first genes transcribed encode transcriptional trans-activators of other viral genes; next are genes that encode functions for replication of the viral genome; and last are genes that encode virion structural components (Stinski 1978, Chambers, Angulo et al. 1999). There is also considerable crossover between genes with different transcription kinetics that encode functions to modulate the cellular environment and also the host immune response. The viral genome circularises following delivery into the nucleus and is replicated by the rolling circle mode; a concatamer of genomes is produced, with cleavage into single genome-length copies occurring concurrently with insertion into pre-formed capsids shells (Mcvoy and Adler 1994).

Newly formed nucleocapsids then enter the virion maturation pathway, acquiring tegument materials at the nucleus, on transit through the cytoplasm, and at the juxtannuclear virion assembly compartment (VAC) (Tandon and Mocarski 2012). Herpesvirions undergo sequential envelopment processes during maturation (Skepper, Whiteley et al. 2001). On nuclear egress, nascent HCMV nucleocapsids undergo primary envelopment at the inner nuclear membrane during deposition into the peri-nuclear space, and de-envelopment at the outer nuclear membrane during deposition into the cytoplasm (Buser, Walther et al. 2007). Nucleocapsids undergo final envelopment at the VAC where they bud into secretory vesicles destined for release from the cell by exocytosis (Cepeda, Esteban et al. 2010). Alternatively, HCMV disseminates via the direct cell-to-cell route by unknown mechanisms.

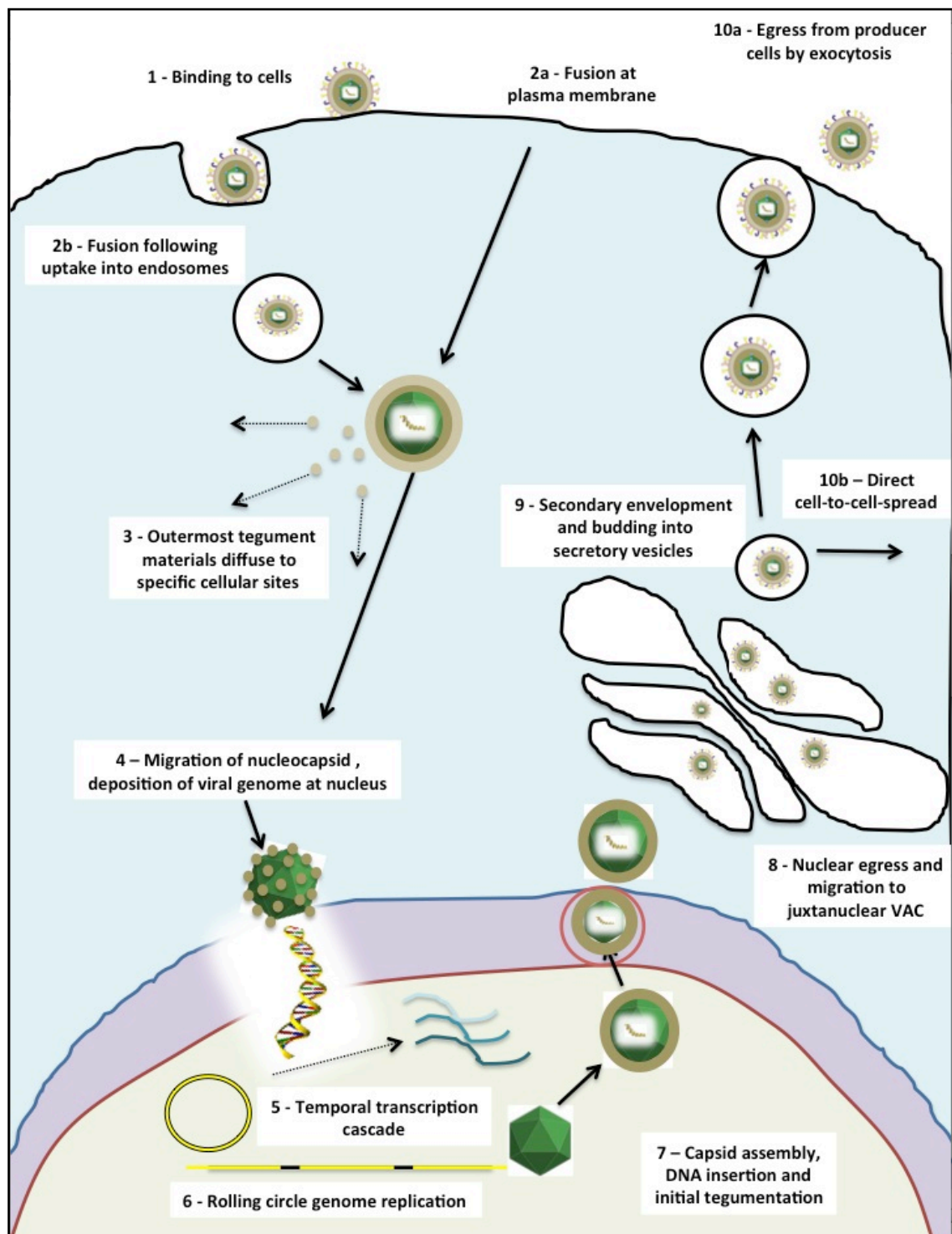


Figure 1.4. HCMV lytic replication cycle – overview. 1 – virions bind at plasma membrane (black line). 2 – virion penetration following fusion at plasma membrane (black line) (2a), or following internalisation by vesicle membranes (black circle) following endosomal uptake (2b), and deposition of virion contents in cytoplasm (light blue). 3 – diffusion of outermost tegument materials. 4 – migration of nucleocapsid and deposition of viral genome into nucleus. 5 – temporal genome transcription. 6 – rolling circle genome replication. 7 – initiation of progeny virion morphogenesis. 8 – nuclear egress following primary envelopment at inner nuclear membrane (red), deposition into peri-nuclear space (purple), and de-envelopment at outer nuclear membrane, and migration of nascent nucleocapsid to juxtannuclear VAC. 9 – secondary envelopment following budding into secretory vesicles. 10 – egress of progeny virions from producer cells by exocytosis into extracellular space (10a), or by direct cell-to-cell spread (10b).

1.4 HCMV infection *in vivo*

Following initial acquisition of HCMV (primary infection), an acute phase of productive viral replication occurs where the host sheds high titres of virus (Zanghellini, Boppana et al. 1999). This acute infection typically lasts weeks to a few months. The mounting of an immune response results in markedly restricted viral replication and shedding during the ensuing convalescent phase. However, HCMV is never completely eliminated and remains throughout the lifetime of the host, requiring constant immune surveillance (Goodrum, Caviness et al. 2012). In an immuno-competent host two potential outcomes arise, though are not necessarily mutually exclusive: chronic persistent infection, with lytic replication and dissemination limited to discrete infection foci, and continued shedding of low titres of virus (Britt 2008); or latent infection, with viral genomes maintained within particular cellular subsets in a state of quiescence by epigenetic repression of lytic genes (Ioudinkova, Arcangeletti et al. 2006, Reeves 2011). Acute infection can again be encountered following re-activation of chronic persistent or latent infections, whilst carriers also remain susceptible to re-infection with new strains.

1.4.1 Immune response to HCMV

Both innate and adaptive immunity are critical for the control of HCMV infection (Loewendorf and Benedict 2010) (Figure 1.5.). The importance of Natural Killer (NK) cells is demonstrated by NK-cell deficiency being correlated with severe outcomes of infection (Biron, Byron et al. 1989). NK cells have several modes of activity for the control of HCMV infection: activated NK cells induce IFN- β production from infected cells, which in turn stimulates various pathways that limit virus replication in an autocrine fashion (Iversen, Norris et al. 2005); activated NK cells also release IFN- γ that activates other effector cells of the innate immune system (e.g. macrophages), and also up regulates MHC class II molecule expression in infected cells, thereby facilitating viral antigen presentation for recognition by cytotoxic T-cells; NK cells directly eliminate HCMV infected cells directly via cytolytic mechanisms (Wilkinson 2013). A defining immunological hallmark of HCMV infection includes CMV specific CD4⁺ and CD8⁺ T-cell sub-sets representing a significant proportion of the total memory T-cell population in an immuno-competent individual (Sylwester, Mitchell et al. 2005, Wills 2013). Furthermore, T-cell depletion correlates with the progression of more severe disease (see section 1.5.2). These observations suggest a critical role for the cellular branch of adaptive immunity for the control of infection. The contribution of the humoral branch of adaptive immunity is evident with titres of anti-HCMV antibodies in the sera of infected individuals increasing throughout the lifetime of the host (Wang, Li et al. 2011, Aberle and Puchhammer-Stöckl 2012). The earliest investigations into the composition of virions and the immune response elicited during natural infection identified virus-encoded envelope glycoproteins as

major targets for neutralising antibodies that limit virus dissemination by alternative activities; i) targeting of glycoproteins incorporated into virions to prevent the binding and entry of cell-free virus into cells, ii) targeting glycoproteins presented at the infected cell surface to prevent cell-to-cell spread (Stinski 1976, Pereira, Hoffman et al. 1982, Farrar and Oram 1984, Nowak, Sullivan et al. 1984).

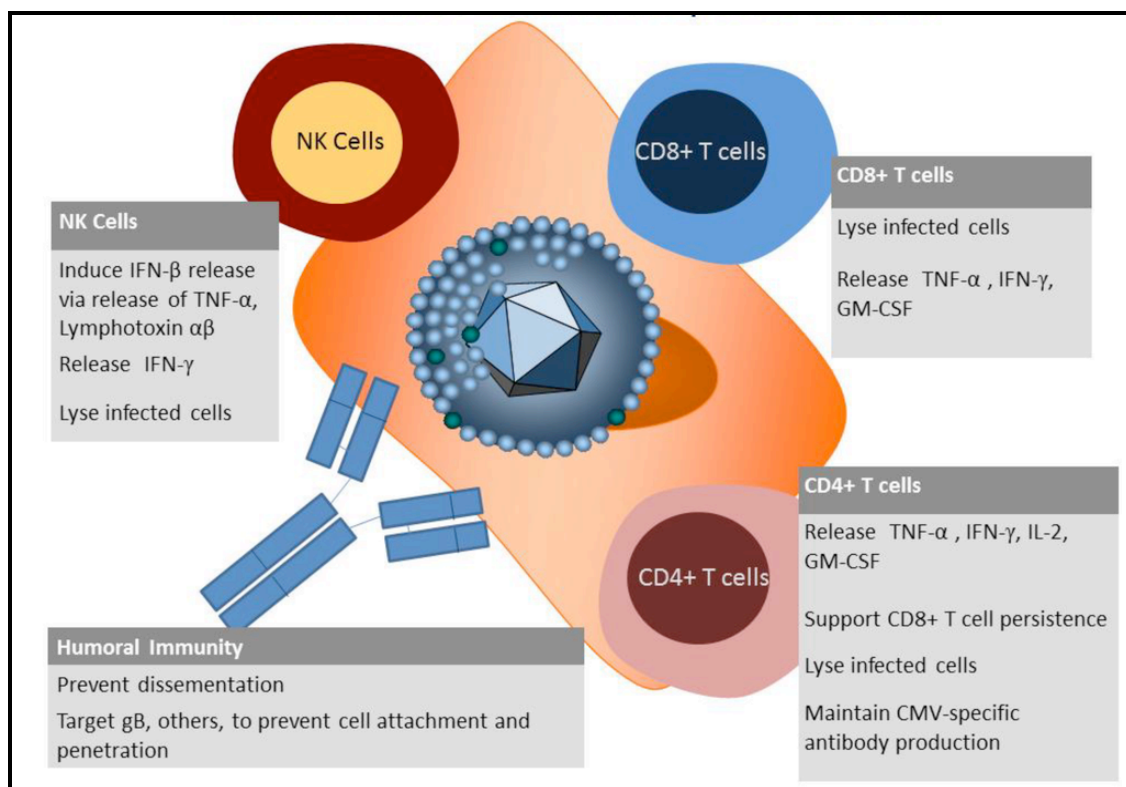


Figure 1.5. Immune response to HCMV infection. Figure reproduced from (Hanley and Bollard 2014), with permission of MDPI publishers.

1.4.2 Tropism of HCMV *in vivo*

During phases of overt lytic replication, virus is capable of disseminating throughout the host and establishing infection in most major visceral organs (Myerson, Hackman et al. 1984, Bissinger, Sinzger et al. 2002). Histological examination of tissues from immuno-compromised hosts, in which HCMV replication is not efficiently suppressed and infections are broadly disseminated, has highlighted the broad range of cell types that are targeted *in vivo* (Sinzger, Muntefering et al. 1993, Sinzger, Grefte et al. 1995, Sinzger, Plachter et al. 1996, Sinzger, Bissinger et al. 1999, Bissinger, Sinzger et al. 2002). However, the major cell types infected vary amongst different organs, as can the organs implicated in different HCMV-related pathologies (Emery 2001).

1.4.2.1 Sites of primary infection and shedding

The principal mode of horizontal transmission is believed to occur by direct contact with infectious bodily secretions e.g. saliva, breast milk or semen (Britt 2007). Following transmission via the oral route or via sexual contact, the first tissues encountered are likely mucosal epithelial tissues of the rhinopharynx or the genital tract, respectively (Revello and Gerna 2010). When transmission occurs via haematogeneous transfer, e.g. via blood transfusions, solid organ transplants, or during intra-uterine transmission, the first tissues encountered are likely vascular endothelial tissues (Revello and Gerna 2010). Glandular epithelial tissues are recognised as sites where virus is secreted into bodily fluids including saliva, urine, breast milk, semen and cervical secretions (Britt 2007).

1.4.2.2 Intra-host dissemination

Intra-host dissemination of HCMV relies on direct cell-to-cell transfer. This includes localised radial dissemination within any single or adjacent solid tissues, and also dissemination to distal tissues that occurs via the haematogeneous route. Epithelial and endothelial tissues other than those at sites of primary infection are also major targets in visceral organs (Sinzger, Muntefering et al. 1993, Sinzger, Grefte et al. 1995, Bissinger, Sinzger et al. 2002). Vascular endothelial tissues located at the interface between solid tissues and circulating blood are crucial in the haematogeneous dissemination of HCMV (Adler and Sinzger 2009, Revello and Gerna 2010). Virus in the blood tends to be cell-associated, with detached and circulating infected cytomegalic vascular endothelial cells (ECs) representing one source of viraemia (Grefte, van der Giessen et al. 1993, Percivalle, Revello et al. 1993). The major sites of viraemia are polymorphonuclear leukocytes (PMNLs) (Gerna, Zipeto et al. 1992, Gerna, Furione et al. 1994) and cells of the monocyte/macrophage (M/M) lineage (Taylor-Wiedeman, Sissons et al. 1991, Gerna, Zipeto et al. 1992, Gerna, Furione et al. 1994). Each of these cell types are recruited to infected vascular endothelial tissues, and acquire virus following transient membrane micro-fusion interactions with infected ECs (Grundy, Lawson et al. 1998, Gerna, Percivalle et al. 2000). Transfer of virus between these cell types is bi-directional (Waldman, Knight et al. 1995), with PMNLs and M/Ms ferrying virus through the blood to initiate infection at distal vascular endothelial tissues. PMNLs are non-permissive to lytic replication or the maintenance of viral genomes as a site of latent infection, and HCMV establishes abortive replication in this cell type (Gerna, Percivalle et al. 2000). Naïve monocytes are also non-permissive to lytic replication, though they do maintain viral genomes as a site of latent infection (see section 1.4.2.4) (Ibanez, Schrier et al. 1991, Taylor-Wiedeman, Sissons et al. 1991). Thus, PMNLs and naïve monocytes ferry pre-formed virions between distal vascular endothelial tissues. Productive replication may be initiated from within infected naïve monocytes following infiltration into solid tissues and transition to more terminally differentiated monocytes, macrophages and dendritic cells (DCs) (Ibanez, Schrier et al. 1991, Sinzger, Plachter et al.

1996, Riegler, Hebart et al. 2000, Smith, Bentz et al. 2004, Bentz, Jarquin-Pardo et al. 2006). PMNLs are also implicated in the direct delivery of virus to solid tissues, and can acquire virus following recruitment and during extravasion through infected EC tissues (Sinzger, Grefte et al. 1995, Grundy, Lawson et al. 1998). These modes of dissemination, together with their proximity to epithelial and endothelial tissues, potentially underlie various stromal tissue cells (fibroblast, smooth muscle cells) and parenchyma tissue cells (hepatocytes, neurons, glial cells) as targets of natural infection (Sinzger, Grefte et al. 1995, Sinzger, Bissinger et al. 1999).

1.4.2.3 *Sites of chronic persistent infections*

Epithelial and endothelial tissues are proposed sites of persistence (Fish, Soderberg-Naucler et al. 1998, Jarvis and Nelson 2002), as suggested by the prolonged shedding of virus from these tissues during chronic infection (Britt 2008). Terminally differentiated monocytes, macrophages and DCs are also proposed as sites of persistent infection (Taylor-Wiedeman, Sissons et al. 1991, Taylor-Wiedeman, Sissons et al. 1994, Jarvis and Nelson 2002, Goodrum, Caviness et al. 2012, Gredmark-Russ and Soderberg-Naucler 2012).

1.4.2.4 *Reservoirs of latency*

Contentiously, bone marrow derived endothelial cells have been proposed as sites of latency (Reeves, Coleman et al. 2004, Jarvis and Nelson 2007). However, more widely accepted sites of latency are bone-marrow derived CD34⁺ precursors of monocytes, macrophages and DCs, and also CD14⁺ monocytes (Taylor-Wiedeman, Sissons et al. 1991, Minton, Tysoe et al. 1994, Mendelson, Monard et al. 1996, Hahn, Jores et al. 1998, Khaiboullina, Maciejewski et al. 2004, Sinclair and Sissons 2006, O'Connor and Murphy 2012). The reactivation of lytic replication in these cells occurs following differentiation into macrophages and DCs (Taylor-Wiedeman, Sissons et al. 1994, Reeves and Sinclair 2013).

1.5 Clinical manifestations of HCMV infection

Infection in an otherwise healthy host is most commonly asymptomatic, yet in rare cases, mild febrile illness, infectious mononucleosis, prolonged fever, and mild hepatitis can be encountered (Sissons and Carmichael 2002, Gandhi and Khanna 2004). More recently, HCMV has been associated with a range of chronic inflammatory and other progressive degenerative disorders in the immuno-competent host (Rafailidis, Mourtzoukou et al. 2008). Examples include malignancies of the CNS (Soderberg-Naucler 2006, Soroceanu and Cobbs 2011), accelerated immune senescence (Pawelec and Derhovanessian 2011), and exasperation of pathologies during critical illness (Cook and Trgovcich 2011). However, HCMV pathogenesis has classically been appreciated in the immune-naïve and immuno-compromised host. In these settings, the lytic viral replication that occurs unabated can result in significant tissue destruction, and ultimately end-organ failure and fatality.

1.5.1 Congenital and perinatal infection

HCMV is the leading viral cause of congenital malformation, affecting ~0.5-1% of live births in the developed countries, with the prevalence of congenital CMV greater still in developing countries (Sissons and Carmichael 2002, Kenneson and Cannon 2007). Vertical transmission of HCMV *in utero* to the immuno-naïve host frequently results in disseminated infection affecting multiple organ systems. HCMV's 'predilection' for the central nervous system (CNS) can result in severe neurological sequelae, with microcephaly, mental retardation, sensory neural hearing loss (SNHL) or visual impairment, amongst the most frequently encountered outcomes of congenital infection. Clinical presentations of Cytomegalic Inclusion Disease (CID) may also include petechiae, hepatosplenomegaly, jaundice, as well as growth retardation (Sissons and Carmichael 2002, Malm and Engman 2007). Symptomatic congenital infection (as defined by the presentation of clinical abnormalities in neonates) accounts for ~10-15% of cases, whilst ~85-90% of congenital infections are asymptomatic (Dollard, Grosse et al. 2007). The prognosis is generally poorer in symptomatic infection, and the risk of developing a permanent disability correlates with increased viral load (Boppana, Fowler et al. 2005). Symptomatic infection is also more frequent when HCMV is transmitted to the developing foetus from a mother experiencing a primary infection, as opposed to a pregnant mother experiencing re-activation of an existing infection or re-infection with new strains. (Reynolds, Stagno et al. 1978, Fowler, Stagno et al. 1992). Therefore, whilst maternal antibodies do not prevent the acquisition of CMV by the developing foetus, they do offer some protection. Approximately 17-20% of all congenitally infected neonates will develop permanent neurological sequelae, whilst mortality is predicted in ~0.5% of all congenital HCMV infections (Dollard, Grosse et al. 2007). Perinatal and post-natal infection can also occur following transmission to new-borns via the ingestion of cervical secretions during birth, or breast milk during lactation (Reynolds, Stagno et al. 1973, Stagno, Reynolds et al. 1980, Dworsky, Yow et al. 1983). Long-term outcome studies indicate that post-natal HCMV infection may have a negative impact on cognitive outcome, particularly when acquired by pre-term infants (Bevot, Hamprecht et al. 2012, Goelz, Meisner et al. 2013).

1.5.2 Infection in the immuno-compromised

HCMV is an important opportunistic pathogen in immuno-compromised individuals, including allograft recipients undergoing immunosuppressive therapy, as well as those who are immuno-deficient due to infection e.g. HIV-AIDS (Emery 2001). The major commonality amongst these patients is compromised adaptive cellular immunity, and therefore a reduced ability to respond to primary infection. Acute productive infection can also arise in the immuno-compromised host due to re-activation of existing persistent/latent infections, or re-infection with different strains. A common clinical presentation in these groups is the development of a HCMV syndrome that presents as a mononucleosis-like syndrome similar to that encountered

in the immuno-competent host, but generally more severe (Emery 2001, Fishman 2007). Various tissue invasive inflammatory diseases can manifest in allograft recipients, most frequently during phases of acute infection. These outcomes of infection predominantly, though not exclusively, affect the particular transplanted tissues. Common clinical presentations include, pneumonitis, gastroenteritis, hepatitis, retinitis (Emery 2001, Boeckh, Nichols et al. 2003, Boeckh 2011), as well as myelo-suppression and pancytopenia bone marrow hyperplasia particularly in bone marrow (BM) transplant recipients (Emery 2001, Castagnola, Cappelli et al. 2004). Chronic persistent infection can result in prolonged stimulation of a pro-inflammatory environment that can contribute to accelerated allograft destruction and rejection (Streblow, Orloff et al. 2007, Arthurs, Eid et al. 2008). HCMV co-infection is also frequent in HIV-infected patients, with high level HCMV viraemia a marker of poor prognosis and correlated with accelerated progression to AIDS defining pathologies (Deayton, Prof Sabin et al. 2004, Griffiths 2006). The most frequently encountered HCMV-related clinical manifestation in AIDS patients is retinitis, whilst gastroenteritis, encephalopathy and poly-radiculopathy are also common (Emery 2001, Fishman 2007).

1.6 Current anti-HCMV Therapies: successes and short-comings

1.6.1 Anti-viral chemotherapeutics

Anti-viral chemotherapeutic drugs currently available for the management of HCMV infection target viral DNA replication. However, these therapies are only modestly effective and also convey dose-limiting toxicity to the host (Lurain and Chou 2010, Prichard and Kern 2011). Moreover, prolonged antiviral therapy is associated with the selection of resistant strains (Emery and Griffiths 2000, Lurain and Chou 2010, Prichard and Kern 2011). The most commonly used antiviral drug is ganciclovir (GCV), a synthetic analogue of 2'-deoxy-guanosine (Matthews and Boehme 1988, Sullivan, Talarico et al. 1992). GCV is administered in an inactive form that is sequentially phosphorylated, firstly by viral encoded protein kinase enzyme (VPK), and subsequently by cellular kinase enzymes, ultimately forming GCV-triphosphate that is a competitive inhibitor of deoxy-guanosine triphosphate (dGTP) (Matthews and Boehme 1988, Sullivan, Talarico et al. 1992). GCV-triphosphate is incorporated into DNA chains during viral genome replication, preferentially inhibiting viral DNA polymerase and arresting chain elongation. Resistance to GCV arises following the acquisition of mutations in either the viral protein kinase (VPK) or viral DNA polymerase-encoding genes (Smith, Cherrington et al. 1997, Chou 2008).

1.6.2 Immunoprophylaxis

A vaccine with high efficacy for the prevention and control of HCMV infection has long been sought, though trials to date have been met with limited success (Krause, Bialek et al.

2013). In one early trial the vaccination potential of live attenuated virus derived from the classical laboratory strain Towne (Plotkin, Furukawa et al. 1975) was investigated in renal transplant recipients (Plotkin, Smiley et al. 1984). Vaccination did not prevent the acquisition of virus or development of acute infection, as seronegative vaccinees remained susceptible to infection when receiving transplants from seropositive donors, and seropositive vaccinees remained prone to reactivation of existing persistent/latent infection, or re-infection with new strains. However, some degree of protection was afforded in vaccinees, and pathologies encountered in these people were often less severe compared to those that received placebo. To further investigate the protective immunity bestowed by vaccination with this live attenuated virus, vaccinees were exposed to pathogenic challenge virus Toledo, and the occurrence and severity of clinical outcomes was compared to those in seronegative and naturally seropositive people exposed to the challenge virus (Quinnan, Delery et al. 1984, Plotkin, Starr et al. 1989). Whilst acute infection was less readily induced in vaccinees compared to seronegative persons, it was more readily induced than in naturally seropositive persons. Furthermore, severe symptomatic infection occurred in vaccinees. Thus, vaccination with this live attenuated virus could not recapitulate the degree of immunity as seen in natural infection. More recent strategies have focused on recombinant sub-unit vaccines for the delivery of specific HCMV antigens. These have frequently been based on virion envelope glycoproteins that are major targets for neutralizing antibodies in natural infection. However, like the previous live-attenuated virus vaccines, these approaches have been only moderately effective for the prevention of later virus acquisition (Pass, Duliege et al. 1999, Pass, Zhang et al. 2009). There was some success in controlling infection in allograft recipients where cellular mediated immunity is compromised, and also for the boosting of existing antibody and T-cell responses in people with chronic persistent infection (Griffiths, Stanton et al. 2011, Sabbaj, Pass et al. 2011). Encouragement for the development of better antibody-mediated immunoprophylaxis strategies comes from the observable benefits of administering hyper-immune globulin (HIG) preparations to those in high-risk groups (e.g. pregnant women with primary infection, transplant recipients); this preparation is derived from the sera of naturally infected individuals and the anti-HCMV antibodies therein bestow passive immunity to recipients (Snydman, Werner et al. 1993, Nigro, Adler et al. 2005, Alexander, Hladnik et al. 2010). However, the degree of protection afforded by administration of HIG in certain groups has been questioned, particularly in HSCT transplant recipients (Cordonnier, Chevret et al. 2003, Raanani, Gafer-Gvili et al. 2009).

The sera of naturally infected individuals and the derived HIG preparations has been shown to display potent neutralising activity in cell types implicated in primary infection and dissemination (e.g. epithelial and endothelial cells), yet this was not re-capitulated in the sera of individuals who received previous vaccines (Figure 1.6.) (Cui, Meza et al. 2008, Gerna, Sarasini et al. 2008). Since the trials of the above-described vaccines, it has become clear that HCMV employs different envelope glycoproteins to gain entry into different cell types, and it was later

demonstrated that the early vaccines did not include important envelope glycoproteins that virus requires for entry into epithelial and endothelial cells. Antibodies against a recently identified virion envelope glycoprotein complex (gH/gL/gpUL128/gpUL130/gpUL131A; see section 1.4.2) are those with greatest potency for the inhibition of epithelial and endothelial cell infection, and these have been detected at high titres in the sera of naturally infected individuals (Cui, Meza et al. 2008, Genini, Percivalle et al. 2011, Cui, Lee et al. 2013). Thus, it is envisioned that there is scope for the development of novel vaccines/immunotherapies with superior efficacy where the pentameric complex is included (Schleiss 2010, Lilja and Mason 2012).

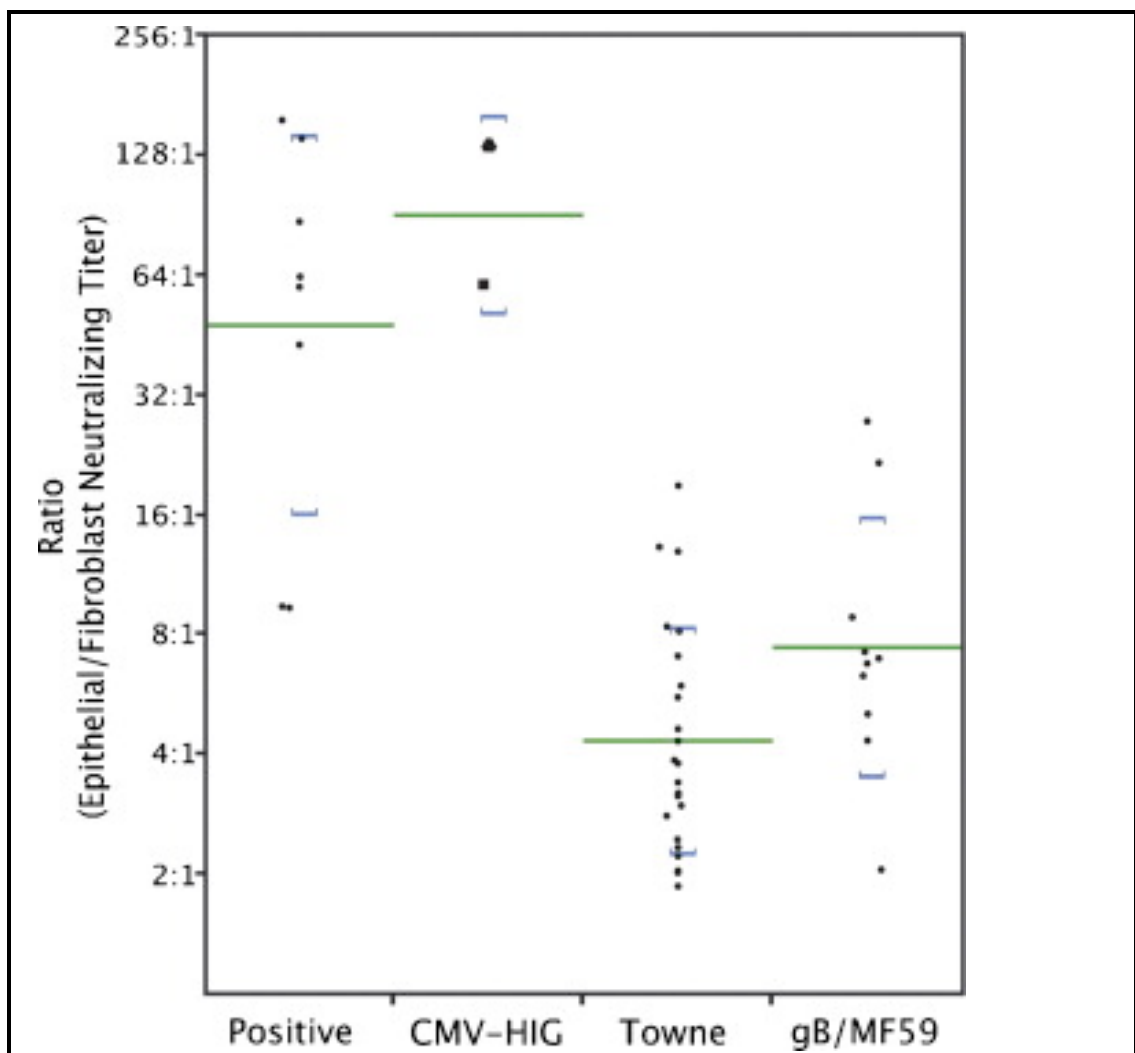


Figure 1.6. Relative titres of anti-bodies that neutralise Epithelial Cell/Fibroblast infection of in sera of individuals with natural infection, HIG preparations, and sera of individuals vaccinated with live-attenuated virus, or recombinant sub-unit vaccines. Ratios of Epithelial Cell/Fibroblast Neutralising titres in human sera following natural exposure to wildtype virus (sero-positive), in CMV-HIG preparation, in sera of subjects vaccinated with live attenuated virus (derived from strain Towne), or subjects vaccinated with glycoprotein sub-unit vaccine (gB/MF59). Lines (green) represent logarithmic means; brackets (blue) represent one standard deviation above/below means. Figure reproduced from (Cui, Meza et al. 2008), with permission of Elsevier publishers.

1.7 Historic limitations in HCMV research

In efforts to develop better therapeutics, the complete biology of HCMV encompassing interactions with all naturally targeted cell-types warrants investigation. However, HCMV research has to date been hindered by the limited availability of virus strains *in vitro* that accurately reflect the biology of wildtype virus *in vivo*. Over the last 3 decades it has become clear that HCMV undergoes rapid adaptation to cell culture, resulting in the generation of laboratory-adapted strains that are genetically and phenotypically distinct from wildtype strains.

1.7.1 Adaptation of HCMV to cell culture

The isolation and propagation of wildtype virus from clinical samples has historically been performed on cultured fibroblast cells. This is due to the relative ease by which this cell type can be infected with clinical isolates as compared to other cultured cell types (Weller 1970, Weller 1971, Weller 1971), and possibly relates to clinical samples containing neutralising antibodies that prevent infection of non-fibroblasts (e.g. epithelial cells, endothelial cells). Initially, the development of cytopathic effects (CPE) by new or recent clinical isolates following inoculation of fibroblasts is timely. Furthermore, infectivity remains mostly cell-associated with low-titres of cell-free virus released into the infected cell culture supernatant, and dissemination is predominantly driven by the cell-to-cell route to form discrete infectious foci plaques (Weller 1971, Weller 1971, Sinzger, Knapp et al. 1997). As a result, inoculated fibroblasts require extensive sub-culturing to encourage the dissemination of clinical HCMV isolates. In contrast to this, virus that has already undergone extensive *in vitro* passage disseminates faster, and produces greater titres of cell-free virus (Weller 1971, Weller 1971, Sinzger, Knapp et al. 1997). Hence, efficient propagation of HCMV *in vitro* has historically been achieved following adaptation to fibroblast cell culture.

1.7.2 Definition of wildtype and ‘laboratory-adapted’ strains, and strain variants

HCMV strains used *in vitro* can be defined by their cell culture passage history and adaptation status (Cha, Tom et al. 1996). Viruses/strains that have been passaged extensively and adapted to fibroblasts are often described as ‘laboratory’ or ‘high-passage’ strains. The more ambiguous term ‘clinical strain’ is often used to describe virus undergone limited fibroblast passage, or indeed no passage at all. For clarity, virus undergone no passage at all shall hereon be referred to as ‘wild-type’, and virus undergone relatively few passages shall be referred to as ‘low-passage’. The generation of genetically and phenotypically distinct variants of a given strain is also correlated with passage history, and these variants can arise from both ‘wildtype’ and ‘laboratory’ strains; however variants differ from the parental virus by a lesser degree than distinct wildtype and laboratory strains.

1.7.3 Genetic heterogeneity between ‘laboratory-adapted’ strains

Separate virus stocks used in laboratories worldwide often contain phenotypically distinct variants derived from a common ‘laboratory-strain’ progenitor (Brown, Kaneshima et al. 1995, MacCormac and Grundy 1999, Hahn, Rose et al. 2003). Retrospective analysis of variants from the most frequently used ‘laboratory-strain’ AD169 (Rowe, Hartley et al. 1956) has correlated their divergence with continued propagation following distribution to different laboratories; this work concerned whole-genome sequencing analysis of viruses derived from the same lineage, though differing in passage history (Bradley, Lurain et al. 2009). Following initial isolation, strain AD169 was passaged up to 14 times in fibroblast in order to produce the virus stock NIH76559 (Rowe, Hartley et al. 1956). The two AD169 variants most commonly used in HCMV research were derived from the NIH76559 lineage. Following distribution to laboratories at St George’s Hospital Medical School in London, NIH76559 was initially subjected to ~40 passages in fibroblasts, then another 16-24 additional passages in attempts to generate an attenuated vaccine strain (Elek and Stern 1974). Virus from these latter passages was plaque purified, and genome restriction fragments were cloned into plasmids that were subsequently used to determine the genome of AD169 variant AD169varUK (Oram, Downing et al. 1982, Chee, Bankier et al. 1990). The passage history of NIH76559 following distribution to laboratories in America is not as well documented, but gave rise to the variant AD169varATCC. Whilst all *in vitro* mutations identified NIH76559 stocks were also identified in AD169varUK and AD169varATCC, each variant also contained further mutations. In particular, AD169varATCCC contained the greatest number of unique mutations, potentially reflecting that this virus had been subject to more extensive *in vitro* passage history. A less commonly used AD169 variant (AD169varUC) was also assessed; this variant was predicted to have arisen from a progenitor that was subjected to a minimum of 50 passages in fibroblasts following distribution to laboratories in the University of Chicago. However, unlike the most commonly used variants, AD169varUC did not contain all of the mutations identified in NIH76559. Thus, AD169varUC contained mutations that were selected during the earliest of the first 14 passages of the AD169 clinical isolate, but was derived from a progenitor earlier than NIH76559. Similar variation in coding potential of variants from ‘laboratory-strain’ Towne (Plotkin, Furukawa et al. 1975) (TownevarL and TownevarS) were also noted during this work, however each of these variants are contained in a mixed population virus, and no correlation between divergence and *in vitro* passage history was made.

The lack of genetic equivalence amongst the commonly used ‘laboratory-strains’ compounded the difficulties in early HCMV research, with many of the earliest reported phenotypes of HCMV not demonstrated in the context of the full wildtype HCMV gene complement. Thus, since the growing appreciation that the classically used ‘laboratory-strains’ are genetically heterogeneous, better descriptions of wildtype HCMV coding potential have since continuously been sought.

1.8 HCMV genetics

The first complete HCMV genome sequence to be determined was that of the prototype laboratory strain AD169 variant AD169varUK (Chee, Bankier et al. 1990). However subsequent studies soon revealed this ‘high-passage’ strain to be a multiple mutant with lesions at several genome loci (Cha, Tom et al. 1996, Skaletskaya, Bartle et al. 2001). Sequencing of several ‘low-passage’ strains contributed to the incremental identification of further HCMV genes that are absent in AD169 (Murphy, Yu et al. 2003, Dolan, Cunningham et al. 2004, Cunningham, Gatherer et al. 2010). However, by far the most comprehensive picture of the HCMV gene complement was gained by sequencing the un-passaged wildtype strain Merlin directly from the clinical sample from which it was isolated (urine of a congenitally infected infant) (A. Davison – Personal Communication). In light of this, strain Merlin been designated the reference HCMV genome sequence by the National Centre for Biotechnology Information (GenBank AY446894; RefSeq NC_006273) (Dolan, Cunningham et al. 2004).

1.8.1 Genome organisation

The HCMV genome is organised in class E herpesvirus genome configuration (Figure 1.7.); ORFs are dispersed over distinct ‘long’ and ‘short’ genome segments, each containing central regions of unique sequence flanked by shorter sequences that are repeated in inverted orientation at each end (Kilpatrick and Huang 1977, Roizman 1980, Weststrate, Geelen et al. 1980, Spector, Hock et al. 1982, Pellet 2013). Annotations for genome regions include a prefix to indicate their location in unique or repeated sequence genome regions, as well as a suffix to indicate their location in either the long or short genome segments: U_L and U_S in the long and short genome segments respectively, with repeated sequence regions denoted TR_L/TR_S at the termini and IR_L/IR_S at the U_L/U_S junction. Under this scheme, the physical HCMV genome map is represented as $TR_L-U_L-IR_L-IR_S-U_S-TR_S$. At the termini of each genome segment as direct repeats, and also in the IR_L-IR_S junction as an inverted repeat, are short sequences (~300-600 bp) termed the *a* sequence (Spaete and Mocarski 1985). The genome configuration can therefore be alternatively denoted as $ab-U_L-b'a'c'-U_S-ca$, with *b* and *c* sequences specific to either the ‘long’ and ‘short’ genome segment terminal repeats, respectively (‘ indicates inverted orientation). The *a* sequences contain signals for genome packaging and cleavage (Kemble and Mocarski 1989). Since these sequences are repeated in inverse orientation at the junction between the U_L and U_S genome segments, they can facilitate genome segment inversion by homologous recombination during replication. This can give rise to four natural isomeric genome arrangements that contain U_L and U_S segments in alternative orientations with respect to one another (Weststrate, Geelen et al. 1980) (Figure 1.7.B).

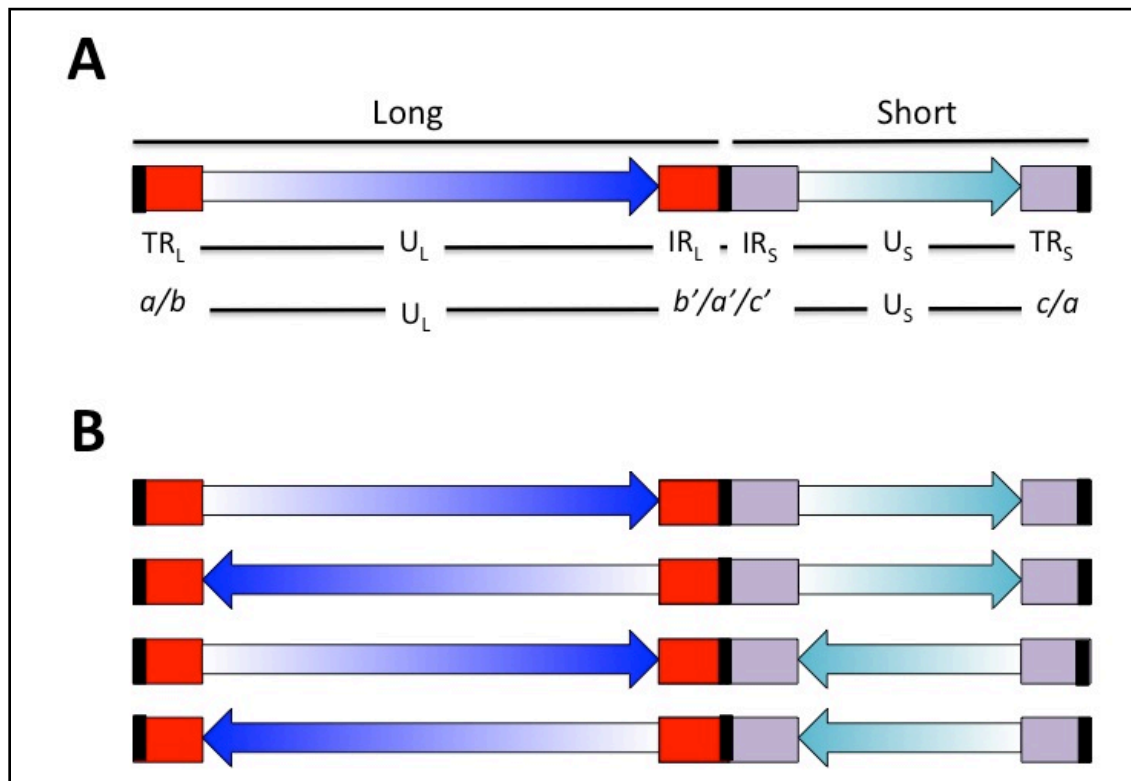


Figure 1.7. HCMV genome organisation and natural isomeric variants. **(A)** Layout and alternative nomenclature for genome regions. In wide boxes are repeat sequences that flank the long (TR_L and IR_L) and short (TR_S and IR_S) genome segments: *a* sequences (black); *b* sequences (red) that; *c* sequences (purple). Unique, single-copy sequences in the long segment (dark blue) and short segment (light blue) are depicted by arrows showing the progressive order of ORFs contained therein. **(B)** Alternative isomeric genome configurations, with U_L and U_S genome segments arranged in various orientations to one another.

1.8.2 HCMV coding potential

Preliminary HCMV genome annotations based on the nucleotide sequence data from strain AD169varUK included 189 putative protein-coding genes (Chee, Bankier et al. 1990). The criteria applied for the identification of protein coding genes focused on sizeable ORFs (≥ 100 contiguous amino acid coding nucleotides) with no significant overlap ($< 60\%$), with gene defining features (e.g. transcriptional elements) used to identify smaller ORFs. This preliminary annotation has since been subjected to multiple revisions, aided by nucleotide sequence data from the wildtype strain Merlin, as well as interpretation by several alternative and complimentary approaches. These include: comparative genomics analysis between distinct HCMV strains and CMVs of other animals to identify conserved ORFs, and also for the re-interpretation of previously described ORFs (Cha, Tom et al. 1996, Davison, Dolan et al. 2003); *in silico* pattern-based approaches to validate putative protein coding genes by identification of structural motifs and functional domains with homology to those of known proteins (Novotny, Rigoutsos et al. 2001, Murphy, Rigoutsos et al. 2003, Rigoutsos, Novotny et al. 2003); and transcriptional analysis to investigate RNA encoding genome regions and deduce intricate gene

splicing patterns (Akter, Cunningham et al. 2003, Mitchell, Savaryn et al. 2009, Gatherer, Seirafian et al. 2011). According to current estimates, the HCMV genome contains ~170 canonical protein-coding genes (See Figure 1.8.). More recently, analysis of HCMV strain Merlin by a ribosome profiling approach has predicted as many as 751 translated ORFs (Stern-Ginossar, Weisburd et al. 2012). This complexity is due to the use of alternative transcription start sites for numerous overlapping ORFs. However the functional significance of many of these proteins remains to be demonstrated.

The alphanumeric nomenclature for the naming of HCMV genes includes a prefix to indicate their location in the genome (UL or US in unique regions; TRL or IRL in repeat regions flanking the long genome segment; TRS or IRS in repeat regions flanking the short genome segment), as well as a numeric suffix to indicate their positional order. To maintain this nomenclature, revised and newly discovered genes are often afforded an additional suffix to distinguish them from existing genes, or to reflect the existence of alternative splice variants of any given gene. It should be noted that the prefixes TRL or IRL are not used in the annotation of all HCMV genomes, with ORFs located in the TRL genome region often simply denoted with the prefix RL. This alternative nomenclature arose since strain AD169varUK that was the basis for the preliminary HCMV genome annotation has undergone duplication of sequences in the TRL genome region which were inserted into the U_L/b' genome region (Cha, Tom et al. 1996). As a result, the 'laboratory-strain' AD169varUK is therefore diploid for ORFs that are haploid in wildtype strains, which normally contain no recognised ORFs in the IRL region.

1.8.2.1 Protein-coding gene products

A striking feature of the HCMV genome is the large number of genes predicted to encode integral membrane proteins, as well as proteins that are post-translationally modified by glycosylation or phosphorylation (Novotny, Rigoutsos et al. 2001, Rigoutsos, Novotny et al. 2003). To reflect the post-translational modification of particular gene products, annotation of encoded proteins often includes a prefix to indicate their existence in un-modified form (p-), or alternatively glycosylated (gp-) and phosphorylated states (pp-). Alternative naming schemes for HCMV gene products reflect the biochemical properties and function of proteins. HCMV genes are also frequently described by their transcriptional kinetics class, classically defined as immediate early (IE/ α), delayed early (DE/ β) and late (L/ γ) genes (Chambers, Angulo et al. 1999). However, similar to recent re-interpretations of the HCMV coding potential, these transcriptional kinetic classes have also been subject to recent re-assessment (Weekes, Tomasec et al. 2014).

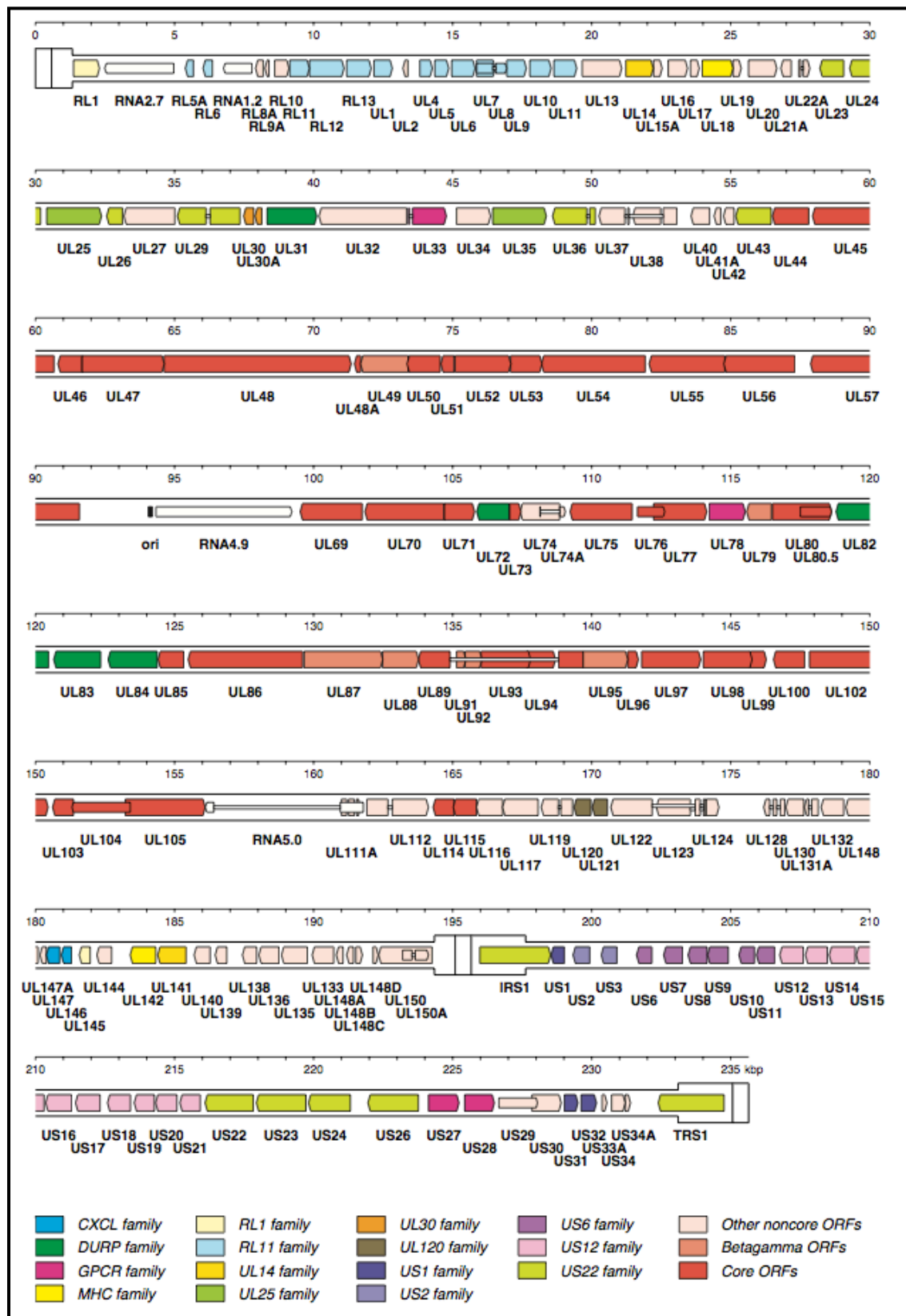


Figure 1.8. Wildtype HCMV strain Merlin genome. Genome organised in $TR_L-U_L-IR_L-IR_S-U_S-TR_S$ orientation. Repeat regions are depicted in broad lines, unique regions in thinner lines. Canonical ORF blocks from, left to right, include: RL1-RL13 and UL1-UL150A in the U_L genome segment; IRS1, US1-US34A, and TRS1 in the U_S genome segment. Herpesvirus conserved (core) genes are in red. Gene families are colour coded. Figure kindly supplied by Dr Andrew Davison (MRC-University of Glasgow Centre for Virus Research, Glasgow, UK).

1.8.2.2 *Non-coding RNAs*

The HCMV genome contains vast expanses of sequence that contain no protein-coding genes, and instead specify non-coding poly-adenylated RNAs. These include a 2.7 Kb RNA (β 2.7) described to have anti-apoptotic activity (McSharry, Tomasec et al. 2003, Reeves, Davies et al. 2007), and a 1.1 Kb spliced RNA and its associated 5 Kb stable intron (RNA 5.0, which by homology to a similar transcript encoded by MCMV, is predicted to be a virulence factor (Kulesza and Shenk 2004). HCMV also specifies 1.2 Kb (RNA 1.2) and 4.9 Kb RNAs (RNA 4.9), though neither has a function assigned (Hutchinson and Tocci 1986, Gatherer, Seirafian et al. 2011). Further non-coding RNAs include those that are transcribed in anti-sense orientation to the ORFs they over-lap (Zhang, Raghavan et al. 2007), as well as various micro RNAs (miRNAs) derived from longer transcripts and implicated in the regulation of genome expression and replication (Pfeffer, Sewer et al. 2005). It should be noted that several of the non-canonical ORFs described by ribosome-profiling over-lap these regions (Stern-Ginossar, Weisburd et al. 2012).

1.8.2.3 *Unique and conserved functions encoded by HCMV*

The substantial divergence of HCMV during co-speciation with the human host is reflected the fact that as few as 40 of the ~170 canonical protein-coding genes (~24%) are conserved across all herpesviruses, whilst only 75 canonical protein-coding genes (~45%) are conserved amongst the beta-herpesviruses (Davison and Bhella 2007, Davison 2007). Genes encoding conserved herpesvirus core functions are dispersed amongst 7 genome blocks (Gompels, Nicholas et al. 1995) located centrally within the U_L genome segment (UL44-UL115) (Figure 1.8.), and likely represent the minimal genetic content derived from the ancestor of all herpesviruses (Dolan, Cunningham et al. 2004). Genes located progressively closer to the termini of each genome segment are those more recently acquired. In general, genes flanking the blocks of herpesvirus core gene are specific to the betaherpesvirinae (UL124-UL115, and several ORFs in the U_S genome region) (Gompels, Nicholas et al. 1995), whilst genes located in the outermost terminal regions are more unique to CMVs, or indeed HCMV alone.

1.8.2.4 *Tropism and immuno-modulatory/immune evasion functions*

Genes specific to HCMV include those that encode tropism factors and immuno-modulatory/immune evasion functions, and these are important determinants of HCMV virulence (Davison 2002, Davison 2013). Intriguingly, several of the proteins encoded by these genes share significant amino acid sequence and structural homology with host cell gene products (Holzerlandt, Orengo et al. 2002). In particular, immuno-modulatory/immune evasion strategies employed by HCMV are commonly centred on molecular mimicry to subvert cellular signalling pathways, and include homologues of cellular cytokines and chemokines, as well as

their receptors (McSharry, Avdic et al. 2012). These observations indicate that HCMV has acquired genetic material from the human host, with evolution of immuno-modulatory/immune evasion mechanisms occurring in parallel with the evolution of the host immune response (Michelson 2004). HCMV also appears to encode proteins with significant homology to those encoded by other viruses, potentially reflecting the acquisition of genetic material from other infections co-existing in the human host (Davison, Akter et al. 2003). Several distinct HCMV genes of this class encode proteins with conserved amino acid sequence and structural motifs, and are assumed to have arisen by divergence following duplication of a single progenitor gene. Such genes are assigned to the 15 currently described gene families (Figure 1.8.), most frequently existing in clusters at the genome segment termini (Chee, Bankier et al. 1990, Davison 2013).

1.8.3 Natural intra-strain genetic variation

The genomes of wildtype HCMV strains are co-linear with ORFs in each genome segment presented in the same relative order (Murphy, Yu et al. 2003, Dolan, Cunningham et al. 2004, Cunningham, Gatherer et al. 2010). Inter-strain variability is most apparent with genetic polymorphisms at several discrete genome loci, particularly involving genes that encode membrane associated or secreted proteins, such as virion envelope glycoproteins and immune evasion functions (Murphy, Yu et al. 2003, Cunningham, Gatherer et al. 2010). These genes often exist as divergent alleles that are believed to have arisen due to immune selection processes. Investigations into the significance of these polymorphisms have focussed on inter-strain differences in tropism and virulence *in vivo*, though to date, no conclusive correlation has been demonstrated; these investigations are reviewed in (Pignatelli, Monte et al. 2004, Puchhammer-Stöckl and Görzer 2011). These genetic signatures are frequently used in the differentiation and monitoring of multiple-strain infections, however this is made difficult given that the recombination that can occur during viral genome replication can result in generation of chimeric strains with diverse genotype pairings (Rasmussen, Geissler et al. 2003). However, analysis of multiple clinical isolates derived from several geographic locations has demonstrated appreciable genotype linkage for virion envelope glycoproteins encoded by adjacent ORFs UL73 (gN) and UL74 (gO) in eight groups (Mattick, Dewin et al. 2004, Bates, Monze et al. 2008, Yan, Koyan et al. 2008, Gorzer, Guelly et al. 2010). Each of these glycoproteins are incorporated into distinct multimeric complexes implicated in the entry of cell-free virus, the cell-to-cell dissemination of virus, and also the formation and/or release of nascent progeny virions (see sections 1.13.2, 1.14.1, and 1.14.3). However, any direct interaction between glycoproteins gN and gO in the virus replication cycle has yet to be demonstrated, and the linkage between these genotypes may simply reflect the fact that adjacent ORFs are less likely to be separated by recombination. In contrast to this, the adjacent ORFs UL146 and UL147 that each encodes viral homologues of α CXC chemokines are also

polymorphic (see section 1.11.3.1), but do not display any observable genotype linkage in distinct strains (Arav-Boger, Zong et al. 2005, Bradley, Kovacs et al. 2008).

Analysis of viruses over time had previously led to the conclusion that HCMV genotypes are relatively stable *in vivo* and also in cell culture, indicating that the selective pressures that resulted in their divergence were no longer as stringent (Mattick, Dewin et al. 2004, Stanton, Westmoreland et al. 2005, Lurain, Fox et al. 2006, Bradley, Kovacs et al. 2008, Gorzer, Guelly et al. 2010). However, with the development of more sensitive high-throughput whole-genome sequencing technologies to assess the genetic content of clinical HCMV isolates, a new picture regarding the stability of these polymorphic alleles is unfolding. It has recently been reported that polymorphic and also well-conserved HCMV ORFs do in fact diverge and give rise to novel variants within a given host (Renzette, Bhattacharjee et al. 2011). As a result, clinical samples from an infected individual often contain variants derived from a single strain. Despite this, the consensus genetic content of wildtype strains is relatively consistent. A potential explanation for this includes distinct quasi-species arising within a host following acquisition and dissemination to different tissues, with a bottleneck process occurring following transmission to a new host (Renzette, Gibson et al. 2013). With regards to hosts experiencing multi-strain infections, deep sequencing analysis over an extended time-course has demonstrated complex dynamics of strain/genotype dominance (Gorzer, Guelly et al. 2010, Ross, Novak et al. 2011).

1.9 BAC-mid cloning of HCMV genomes

The capture of HCMV genomes as bacterial artificial chromosome (BAC) clones represented a major turning point in HCMV research. These clones provided for the first time a genetically stable source of clonal, single-strain genomes, from which virus of known coding potential can be reconstituted. BAC-cloning was first demonstrated to be an effective strategy for the stable capture and maintenance of large DNA fragments (Shizuya, Birren et al. 1992), and subsequently adapted for the capture of entire CMV genomes; firstly that of MCMV strain Smith (Messerle, Crnkovic et al. 1997), and later HCMV strain AD169 (Borst, Hahn et al. 1999). Since this, the genomes of several 'low-passage' strains have also been captured as BAC constructs.

BAC-targeting vectors are based on the sizeable single copy fertility plasmid (F-plasmid) from *E.coli*, and contain several features to ensure BAC constructs are replicated and retained in bacterial cells with exquisite stability (Borst, Hahn et al. 1999, Stanton, Baluchova et al. 2010). To capture HCMV genomes in BAC constructs, linearised BAC targeting vectors are first engineered to contain flanking arms with homology to sequences at the site of BAC insertion in the viral genome. These BAC-targeting vectors are then delivered into HCMV infected cells where they are integrated into the HCMV genome by homologous recombination.

Insertion of the BAC targeting vector increases the total genome size, and given the constraints for genome packaging, this can in turn reduce the efficiency of virus replication (Yu, Smith et al. 2002, Cui, McGregor et al. 2009). To limit the likelihood of virus acquiring large compensating deletions following reconstitution from the BAC, genome regions encoding functions known to be dispensable for growth in cell culture are often replaced by stably incorporated BAC vectors; this commonly includes the sacrifice of ORFs in the U_S genome region that have been shown to be dispensable (Jones and Muzithras 1992). In an alternative approach, self-excising BAC sequences can be inserted. BAC vectors can be engineered to encode mediators of excision (Cre-recombinase), and the signals they recognise (lox P) flanking the BAC vector sequence (Messerle, Crnkovic et al. 1997). The insertion of an intron to disrupt the Cre gene ensures Cre recombinase is not produced in the bacterial cells where BACs are maintained, and the BAC vector sequence is only excised following reconstitution of virus in eukaryotic cells (Yu, Smith et al. 2002). This approach is advantageous as no compromise in coding potential is required to accommodate stably incorporated BAC vector sequences.

1.10 HCMV genes that are essential, dispensable, and/or inhibitory, for virus growth in cell culture

The contribution of single genes in the context of productive infection has been investigated by global mutational analysis of BAC-cloned HCMV genomes and phenotype analysis of the viruses they produce. Applying a reverse genetics approach to BAC-cloned versions of the ‘laboratory-strains’ AD169 or Towne, the viability of single gene knockout mutants has been used to distinguish genes that are essential or non-essential for virus propagation *in vitro* (Dunn, Chou et al. 2003, Yu, Silva et al. 2003). Whilst mutants lacking non-essential genes remained viable, they often displayed a distinct growth phenotype compared to parental virus. Mutants often grew less efficiently than parental virus, demonstrating that whilst the corresponding deleted gene was not essential, it enhanced/augmented virus replication. Conversely, particular mutants grew more efficiently than parental virus, demonstrating that the corresponding deleted genes are inhibitory to the growth of virus in cell culture; genes encoding functions inhibitory to virus growth in cell culture are commonly termed ‘temperance factors’ (Dunn, Chou et al. 2003, Yu, Silva et al. 2003, Stanton, Baluchova et al. 2010). However, since each of the laboratory-adapted HCMV strains used in these investigations lack the full complement of wildtype genes, these analyses cannot be considered comprehensive. It should be appreciated that the true impact of any given gene may only be demonstrable in the presence of functional partners, competitive inhibitors, or regulatory components that modulate the expression and functionality of the products they encode. Whilst the ‘laboratory-strains’ used in these studies contain similar genetic lesions, they also contain unique mutations. This likely explains the few discrepancies in the data from these previous studies where the impact of particular genes was reported differently. However, the

impact of the vast majority of HCMV genes were reported similarly, and these works provided a reliable picture of the of the HCMV genome complexity, and also highlighted the redundancy of numerous encoded functions for the growth of virus in cell culture.

1.10.1 Cell-type specific functions

As little as 25-30% of the genes encoded by HCMV are essential for growth in cultured fibroblasts, mostly involving those that are conserved amongst the herpesviruses and encode core functions in the viral replication cycle (Figure 1.9) (Dunn, Chou et al. 2003, Yu, Silva et al. 2003). There is considerable overlap within this sub-set of genes that are essential for the growth of virus in other cultured cell types, as well as with the genes that are dispensable for the growth of virus in different cell types – the latter commonly involve those genes that encode immuno-modulatory functions. Furthermore, some genes are essential or enhance the growth of virus in one cell type, yet are dispensable or even temperance factors in others.

1.10.2 Mutations acquired during adaptation of HCMV to cell culture

The identification of temperance genes offered some explanation to the selective pressures imposed for the adaptation of HCMV *in vitro*. Dispensable genes encoding ‘temperance-factors’ are distributed throughout the genome, with lesions identified at several loci in the genomes of extensively passaged laboratory viruses (Cha, Tom et al. 1996, Davison, Dolan et al. 2003). The number of mutations contained in lab-adapted strains correlates with increasing passage history, yet consistent themes are observed amongst all passaged viruses. This sequential adaptation of HCMV has been demonstrated experimentally, and also shown to occur during repeated passage in cultured epithelial cells and endothelial cells, as well as fibroblasts (Dargan, Douglas et al. 2010). Gene ablating mutations are found as frame-shifting insertions and deletions (indels); substitutions that introduce premature stop coding codons; sequence inversions interrupting ORFs; and large deletions of sequence that encompass entire genes. Non-synonymous amino acid substitutions are also frequently observed.

Genes that encode integral membrane proteins performing non-essential functions are generally those most prone to ablating mutation. Genes encoding core functions in the viral replication cycle, e.g. genome replication and packaging machinery, have been shown to acquire non-synonymous mutations *in vitro* (Dargan, Douglas et al. 2010), however the essential functions these genes encode suggest that such mutations will not be totally ablating (Dunn, Chou et al. 2003, Yu, Silva et al. 2003). Of the genes encoding virion structural components, those that encode capsid proteins rarely mutate. Tegument protein encoding genes frequently mutate, but genes most prone to mutation are those encoding virion envelope components.

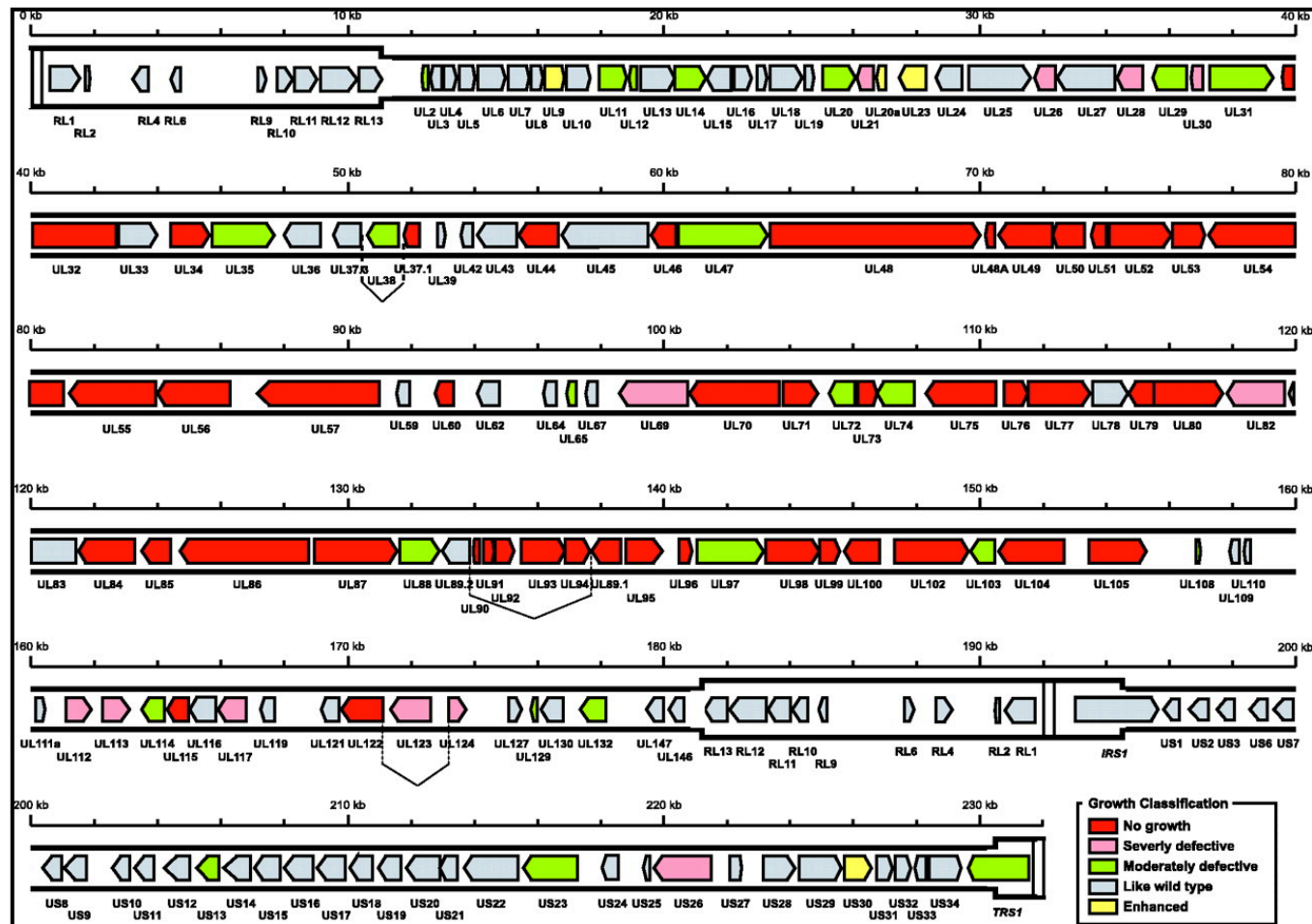


Figure 1.9. Impact of single ORFs on virus growth in fibroblasts. Data based on strain Towne BAC-derived virus. ORFs are colour-coded to reflect growth phenotype of mutants; no growth (red) - the corresponding ORF was classified as essential; severely or moderately defect (pink/green) - corresponding ORF enhanced/augmented replication; no differential phenotype (grey) – corresponding ORF dispensable, but neither enhances or tempers virus growth; enhanced (yellow) – corresponding ORFs were inhibitory and tempered virus growth. Figure reproduced from (Dunn, Chou et al. 2003), with permission of PNAS publishers

1.11 Biological differences between wildtype and ‘laboratory-adapted’ HCMV strains

1.11.1 Restricted tropism

‘Laboratory-strains’ display restricted tropism as a hallmark of their adaptation to cell culture, such that increased fitness in fibroblasts occurs with a reduced fitness in other naturally targeted cell types (Figure 1.10.). Efficient infection of cultured epithelial cells (Miceli, Newsome et al. 1989), endothelial cells (Kahl, Siegel-Axel et al. 2000), monocytes, macrophages (Ibanez, Schrier et al. 1991, Minton, Tysoe et al. 1994) and PMNLs (Gerna, Percivalle et al. 2000), has to date been a particular challenge in HCMV research. Thus, investigations in these naturally targeted cell types have only been possible with ‘low-passage’ viruses that grow slowly and to limiting titres.

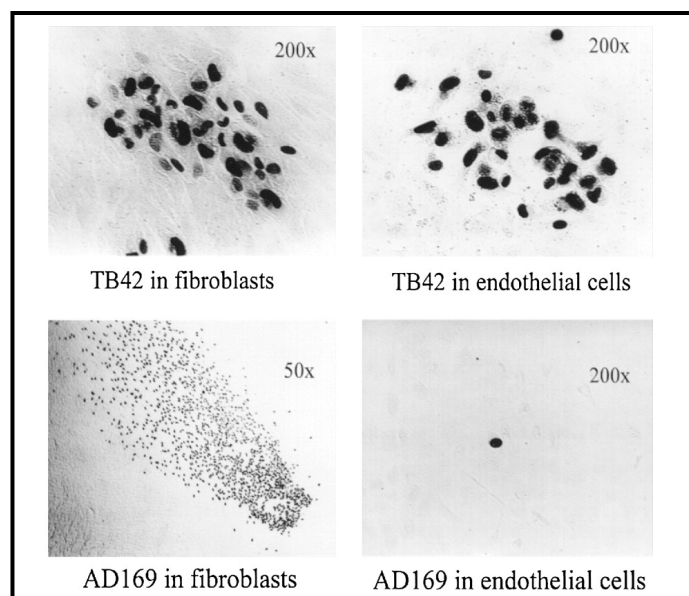


Figure 1.10. Differential tropism of HCMV strains with contrasting cell-culture passage history. As demonstrated by infectious focus expansion assay (Sinzger, Knapp et al. 1997), ‘laboratory-strain’ AD169 displays greater fibroblast tropism, yet drastically restricted endotheliotropism, compared to the low-passage strain TB42. Figure reproduced from (Sinzger, Schmidt, 1999), with the permission of Society of General Microbiology publishers.

The association between fibroblast passage history and tropism was verified by the successful re-capitulation of lab-adaptation under controlled experimental conditions (Sinzger, Schmidt et al. 1999, Revello, Baldanti et al. 2001). Sequential passage of broad tropism ‘low-passage’ strains in fibroblasts resulted in the generation of tropism-restricted ‘laboratory-adapted’ progeny, and this occurred concurrently with changes in the viral genome. The hereditary nature of HCMV tropism determining factors was demonstrated in these investigations, and the restricted tropism of extensively passaged strains was shown not to be

due to epigenetic phenomena. Firstly, fibroblasts transfected with DNA extracted from either 'broad-tropism' or 'restricted-tropism' strain variants produced progeny virus with tropism identical to that of the respective parental viruses. Secondly, fibroblasts simultaneously infected with two distinct 'restricted-tropism' lab-adapted strain variants produced 'broad-tropism' progeny. This latter observation indicated that the restricted tropism of each parental strain was due to loss of distinct genetic functions, and together these strains provided the full range of genes encoding all necessary tropism determining factors, as contained in the recombinant broad-tropism progeny virus that was produced.

Interestingly, similar growth kinetics and productivity improving adaptations have also been observed following prolonged propagation of low-passage virus in endothelial cell culture; viruses with an extensive passage history in this cell type grew more efficiently in endothelial cells compared to fibroblasts (James Waldman, Sneddon et al. 1989). Thus, adaptation of HCMV to cell culture occurs in a cell-type specific manner, and is not restricted to fibroblasts. This observation has been exploited for the propagation of virus that retains broad tropism, although the titres produced following adaptation to ECs were still lower than that of virus adapted to fibroblasts (Waldman, Roberts et al. 1991, Sinzger, Schmidt et al. 1999). Furthermore, broad-tropism progeny have been recovered from within stocks of laboratory-adapted strains following sequential passage in cell types other than fibroblasts (Gerna, Percivalle et al. 2002, Gerna, Percivalle et al. 2003).

1.11.2 A genetic determinant of broad-tropism: the UL128L genome region

Assessment of the genetic differences between broad-tropism and restricted-tropism strains identified three adjacent ORFs in the U_L/b' genome region (Figure 1.11.) as determinants of tropism: UL128, UL130 and UL13A, together comprising the UL128 locus (UL128L) (Hahn, Revello et al. 2004). The proposed role of these ORFs as tropism determinants was supported by observation that all sequenced restricted-tropism strains contained mutations in this genome region (Revello, Baldanti et al. 2001, Akter, Cunningham et al. 2003, Davison, Dolan et al. 2003, Dolan, Cunningham et al. 2004, Cunningham, Gatherer et al. 2010, Stanton, Baluchova et al. 2010). The observation that 'laboratory-adapted' strains were often mutated in just one UL128L gene suggested that the functions they encode acted in concert. The specific contribution of each of UL128, UL130 and UL131A was investigated by a reverse genetics approach for the repair or ablation of each gene in 'restricted-tropism' and 'broad-tropism' strains, respectively, as well as complementation of 'restricted-tropism' strains with wildtype copies of UL128L genes for the rescue of broad tropism. These investigations demonstrated that each of the UL128L genes was required for infection of epithelial cells (Wang and Shenk 2005), endothelial cells (Patrone, Secchi et al. 2005, Adler 2006), monocytes (Straschewski, Patrone et al. 2011), and dendritic cells (Riegler, Hebart et al. 2000, Gerna, Percivalle et al. 2005).

1.11.2.1 *UL128L gene structure*

The UL128L genome region in AD169 was originally described to contain four genes (UL128, UL129, UL130, UL131) (Chee, Bankier et al. 1990). Later interpretations identified novel splice patterns for UL128 and UL131, and ultimately the identification of 3 genes (Akter, Cunningham et al. 2003, Davison, Dolan et al. 2003). The newly interpreted UL128 over-lapped and included UL129, and thus retained its original name. UL131 was frame-shifted in AD169, with the proteins encoded by the original and re-interpreted ORFs sharing no significant amino acid sequence homology, resulting in renaming of UL131 as UL131A. UL128 contains 516 coding nucleotides dispersed over three exons (exon 1 – 164 nt, exon 2 – 135 nt, and exon 3 – 217 nt); UL130 contains 645 coding nucleotides with no introns; and UL131A contains 390 coding nucleotides dispersed over two exons (exon 1 – 236 nt, and exon 2 – 154 nt) (Akter, Cunningham et al. 2003, Baldanti, Paolucci et al. 2006, Sun, Ji et al. 2009).

1.11.2.2 *UL128L gene diversity*

Nucleotide sequence analysis of distinct wildtype strains from separate geographic locations has demonstrated that each of the three UL128L genes are highly conserved *in vivo*: UL128, UL130 and UL131A amongst clinical strains display 94.3-96.0%, 92.3-95.8% and 95.9-96.2% nucleotide sequence identity, respectively (Baldanti, Paolucci et al. 2006, Sun, Ji et al. 2009). Each gene shares no significant sequence homology with genes in other herpesviruses, though positional and functional homologues are found in CMVs of other animals (Davison, Dolan et al. 2003, Gnanandarajah, Gillis et al. 2014, Malouli, Hansen et al. 2014).

1.11.2.3 *UL128L gene transcription*

A number of transcripts have been identified in UL128L (Akter, Cunningham et al. 2003, Hahn, Revello et al. 2004, Sun, Ren et al. 2010). Read-through transcripts corresponding to the coding regions of all 3 genes and originating from a start site situated upstream of UL131A have consistently been identified. UL128-specific transcripts are also produced, most likely derived from an initiation site located within the UL130 coding region. All observed UL128L transcripts are co-terminal, determined by the same poly-adenylation site downstream of UL128.

There are conflicting reports of the temporal kinetics with which the UL128L genes are expressed. Two groups report the detection of transcripts exclusively with late (L/ γ) kinetics (Akter, Cunningham et al. 2003, Hahn, Revello et al. 2004), whilst another group detected UL128L transcripts with immediate early (IE/ α) and late (L/ γ) kinetics (Sun, Ren et al. 2010). This latter observation opens up the possibility that the UL128L genome region, particularly the UL128 gene that appears to be transcribed independently of UL130 and UL131A, encodes multiple protein variants that potentially perform different functions at different phases in the lytic replication cycle.

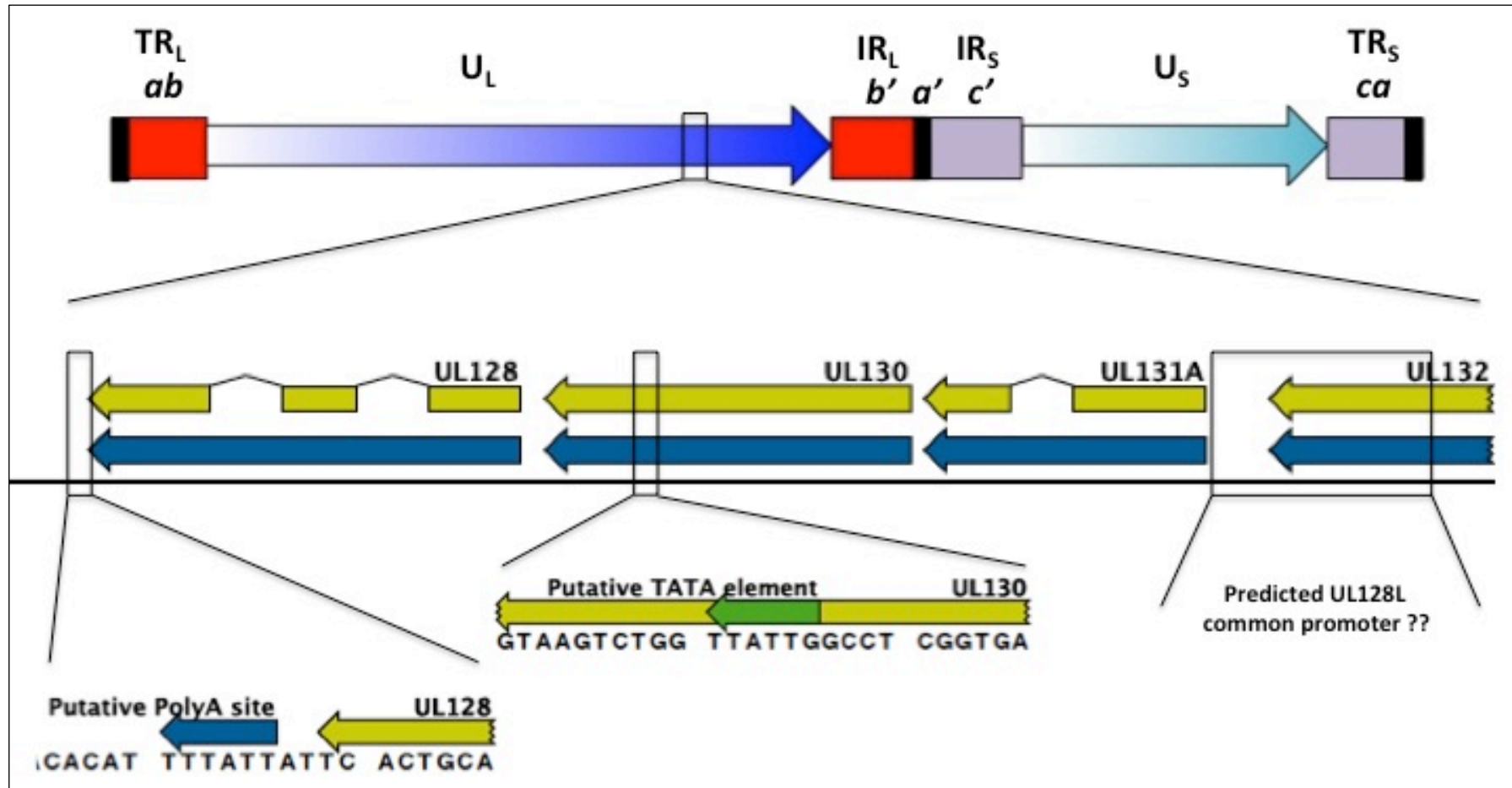


Figure 1.11. UL128 locus (UL128L): location and ORF structure. (A) Position of the UL128 locus (UL128L) in HCMV U_L/b' genome region. (B) In focus are UL128, UL130 and UL131A ORFs, and part of UL132: entire ORFs (blue); coding domain sequences (CDS) (yellow); introns (arches linking CDS regions). Arrow directions reflect 5'-3' direction of UL128L ORF transcription, in complementary orientation to the conventional genome annotation. Highlighted are (C) the putative polyadenylation common to all UL128L transcripts downstream of UL128 (dark blue); (D) the putative TATA element for UL128-specific transcripts initiated from within the UL130 coding domain (green); and (E) the region upstream of UL131A predicted to contain the UL128L common promoter.

1.11.2.4 Impact of UL128L expression on virus propagation in cell culture

The UL128L genome region encodes functions required for efficient infection of a broad range of naturally targeted cells types, yet are dispensable for infection of fibroblasts. In fact, UL128L encoded functions are temperance factors for the growth of virus in fibroblasts, as demonstrated when the repair of mutated UL128L genome regions, or alternatively complementation of viruses with intact UL128L ORFs, each resulted in reduced dissemination kinetics and cell-free virus production of infections compared to parental virus (Patrone, Secchi et al. 2005, Wang and Shenk 2005, Adler 2006, Dargan, Douglas et al. 2010, Stanton, Baluchova et al. 2010). These observations explained the pressure for the rapid selection of virus containing ablating mutations in this genome region when passaged in fibroblasts, and the subsequent alterations in tropism when virus undergoes adaptation to cell culture. Furthermore, the selection of UL128L mutant progeny occurs only slightly less rapidly than that of RL13 mutant progeny (see section 1.11.3.2), and within as few as 1-5 passages (Dargan, Douglas et al. 2010, Stanton, Baluchova et al. 2010).

1.11.3 Restricted virulence

Laboratory strains are generally avirulent, as can be seen from previous vaccine clinical trials based on live-attenuated viruses from strains with varying *in vitro* passage history. The classical laboratory strains AD169 and Towne were each subjected to extensive *in vitro* passage during vaccine development, passaged more than 50 and 128 times in fibroblasts respectively (Rowe, Hartley et al. 1956, Elek and Stern 1974, Plotkin, Furukawa et al. 1975). In contrast to this, the challenge virus strain Toledo developed during clinical trials of Towne vaccine was subjected to less extensive *in vitro* passage (5 fibroblast passages) (Quinnan, Delery et al. 1984, Plotkin, Starr et al. 1989). As previously mentioned, exposure to strain Toledo resulted in the development of acute infection and often severe clinical symptoms, whilst vaccination with AD169 and Towne most frequently resulted in no clinical outcome at all (Just, Buergin-Wolff et al. 1975, Fleisher, Starr et al. 1982, Quinnan, Delery et al. 1984, Plotkin, Starr et al. 1989). The reduced virulence of the more extensively passaged strains may be in part due to restricted tropism and the inability to establish infection in a wide range of normally targeted tissues. However, genome regions that encode virulence factors have since been shown to be prone to mutation during the passage of virus *in vitro*.

1.11.3.1 U_L/b' encoded functions

The U_L/b' genome region (UL148-UL150 in Merlin genome figure above – Figure 1.8.) is a major site of heterogeneity amongst laboratory strains, first evidenced by RFLP analysis of related laboratory strain variants (Tamashiro and Spector 1986). These observations, together with the phenotypic differences between ‘laboratory-adapted’ and ‘low-passage’ strains, inspired the comparative analysis of the U_L/b' genome regions of strains with differing passage

history (Cha, Tom et al. 1996, Davison, Dolan et al. 2003, Dolan, Cunningham et al. 2004). These investigations identified up to 15 Kb of sequence in the ‘low-passage’ strain Toledo that was absent in extensively passaged strains AD169varATCC and Towne. Furthermore, Towne contained sequence absent in the U_L/b' region of both AD169 and Toledo, whilst Toledo also contained inverted sequence in this genome region. Ultimately, 22 putative U_L/b' ORFs were identified compared to that in annotations based on strain AD169varUK, although subsequent analysis now indicates there are 21 ORFs in this region. However, the selective pressures for mutations in this genome region are not well understood like that for mutation in the UL128L genome region. In particular, the inhibitory impact of specific ORFs in this genome region has not been demonstrated experimentally, and therefore many may have been lost during the passage of strains AD169 and Towne co-incidentally along with ORFs that do indeed encode temperance factors.

A general commonality amongst the ORFs in the U_L/b' genome region is the encoding of integral membrane proteins with immuno-modulatory/immune evasion functions. These include (in order of presentation in the Merlin genome): the polymorphic gene UL144, a member of the RL1 gene family that encodes a tumour necrosis factor receptor (TNFR) homologue that subverts lymphocyte responses (Lurain, Kapell et al. 1999, Arav-Boger, Willoughby et al. 2002, Poole, King et al. 2006); UL142 of the MHC homologue gene family, and UL141 of the UL14 gene family, that both encode integral membrane proteins that subvert natural killer (NK) detection and elimination of virus infected cells (Tomasec, Wang et al. 2005, Wills, Ashiru et al. 2005, Prod'homme, Sugrue et al. 2010); the poly-cistronic UL138-UL133 locus (ORFs UL138, UL136, UL135 and UL133) that encodes four integral membrane proteins which act in synergy to repress viral replication in CD34⁺ myeloid cells during latency (Umashankar, Petrucelli et al. 2011). Each of UL144, UL142, and UL141 has been shown to mutate independently during the passage of virus in cultured fibroblasts and endothelial cells, whilst UL141 has also been shown to mutate in epithelial cells (Dargan, Douglas et al. 2010). Hence the functions encoded by these ORFs may indeed be inhibitory to the growth of virus in cell culture. Interestingly, whilst the UL138-UL133 locus is dispensable for growth in fibroblasts and epithelial cells, the functions they encode may enhance virion morphogenesis in endothelial cells (Umashankar, Petrucelli et al. 2011, Bughio, Elliott et al. 2013). Other than ORFs encoding integral membrane proteins, the polymorphic genes UL147 and UL146 each encode secreted viral homologues of α CXC chemokines (vCXC-2 and vCXC-1 respectively), and together comprise the UL146 CXC chemokine homologue gene family (Arav-Boger, Foster et al. 2006, Lurain, Fox et al. 2006, McSharry, Avdic et al. 2012).

Other genes within the U_L/b' region encode putative membrane proteins of unknown function. This includes UL140, as well as the polymorphic genes UL139, UL148A, UL148B, UL148C and UL148D (Ji, Rong Sun et al. 2006). UL150 and UL50A each encode putative

secreted proteins with unknown function {Mocraski, 2013 #777}. Many of these genes have also been shown to acquire mutations independently (Dargan, Douglas et al. 2010). For example, although UL140, UL145 and UL150 were identified in Toledo, each contained frame-shifting mutations that likely ablated the functionality of proteins encoded (Dolan, Cunningham et al. 2004). Thus, as well as being dispensable, the functions encoded by these genes may be inhibitory to the growth of virus in cell culture.

1.11.3.2 *RL11 gene family*

Encompassing the TR_L/U_L junction are 14 polymorphic genes that comprise the RL11 gene family, each predicted to encode integral membrane glycoproteins, including several that are incorporated into the virion envelope (Davison, Akter et al. 2003, Davison, Dolan et al. 2003). Several members encode immuno-modulatory functions and potentially tropism factors, although the function of others remains to be determined. Many of these genes are non-essential for growth in fibroblasts, and have been observed to undergo mutation independent of other RL11 genes following passage in this cell type (Dolan, Cunningham et al. 2004, Spaderna, Hahn et al. 2004, Cunningham, Gatherer et al. 2010, Dargan, Douglas et al. 2010). Interestingly, RL5A mutations have also been identified in wildtype virus undergone no cell culture passage at all (Cunningham, Gatherer et al. 2010). RL13 is a ‘temperance factor’ for the growth of virus in both fibroblasts and epithelial cells, and is the most commonly reported RL11 family member to acquire mutations *in vitro* (Dolan, Cunningham et al. 2004, Stanton, Baluchova et al. 2010). Indeed, it is the most rapidly mutated gene within the entire genome (Dargan, Douglas et al. 2010, Stanton, Baluchova et al. 2010). UL1 is dispensable for growth in fibroblasts, though loss of this gene causes a growth defect in epithelial cells (Shikhagaie, Mercé-Maldonado et al. 2012). Like RL5A, naturally occurring UL1 mutants have been identified in wildtype virus (Sekulin, Gorzer et al. 2007). Conversely, UL10 has been demonstrated to be inhibitory and dispensable for the growth of virus in epithelial cells, though not fibroblasts (Dunn, Chou et al. 2003).

1.12 HCMV virus particle composition and morphogenesis

Descriptions of the HCMV particle composition have been subject to incremental revision in parallel to that of the genome coding potential. The identification of virion components has historically relied on the analysis of purified virions by biochemical and immunological methods, with the assignment of components to particular virion compartments deduced by a combination of direct and indirect experimental approaches (Sarov and Abady 1975, Fiala, Honess et al. 1976, Baldick Jr and Shenk 1996). The identification of glycoproteins as components of the virion envelope relied on their detection by immuno-blot analysis of purified virions, coupled with infectivity neutralization assays to demonstrate their exposure to the external environment and involvement in virus entry and binding (Stinski 1976, Farrar and Oram 1984). Identification and compartmentalization of other components has been aided by indirect observations that include: confocal microscopy analysis to identify proteins that co-localise with known virion components at the virion assembly compartment (VAC); yeast-two-hybrid and other protein-protein binding assays to determine interactions between known components already assigned to a virion compartment; as well as *in silico* analysis of amino acid sequence and structural motifs to identify features indicative of virion envelope insertion (e.g. trans-membrane domains and signal peptide sequences). Direct approaches have included proteolytic profiling, based on a combination of biochemical fractionation, proteolytic digestion, and detection of recovered proteins from both intact and detergent solubilized virions. More recently, mass spectrometry analysis of HCMV virions has been applied in a global proteomics approach using the AD169 strain. This identified 71 viral encoded proteins and more than 70 host cell derived proteins that co-purified with virions (Varnum, Streblow et al. 2004).

1.12.1 HCMV particle nucleocapsid compartment

The HCMV nucleocapsid is structurally constrained and thus the virion compartment with greatest homogeneity amongst strains. There are currently 5 known capsid protein components, all of which are encoded by herpesviruses core genes. These include: the major capsid protein (MCP; pUL86) (Chee, Rudolph et al. 1989) present at 955 copies per virion, and forming the previously described pentameric and hexameric capsomeres at the capsid faces, walls and vertices; portal protein (PORT; pUL104) (Dittmer, Drach et al. 2005), present at 12 copies per virion, and forming the single specialised capsomere through which the genome is inserted and released; 320 copies the minor capsid protein (mCP; pUL85) (Baldick Jr and Shenk 1996), and 640 copies of the minor capsid binding protein (mC-BP; pUL46) (Gibson, Baxter et al. 1996), together forming the 320 hetero-trimeric triplexes that link three adjacent capsomeres, in a 1:2 ratio respectively; and 900 copies of the smallest capsid protein (SCP; pUL48A) that decorate the MCP hexons in 6-member ring-like structures (Gibson, Clopper et al. 1996, Yu, Shah et al. 2005).

Descriptions of the HCMV capsid assembly pathway (Figure 1.12.) have been heavily dependent on observations made for other herpesviruses (Homa and Brown 1997, Brown and Newcomb 2011, Cardone, Heymann et al. 2012). In HCMV, two proteins orchestrate capsid assembly: the assembly protein (AP/pUL80.5) and protease assemblin (PR/pUL80), each initially contained as precursors within a fusion protein (pPR-pAP). Assembly begins with the formation of capsid sub-unit protomers in the cytoplasm, where pUL86/MCP associates with pAP and pPR, and then the nuclear localisation sequences (NLS) of pAP facilitate the delivery of protomers into the nucleus (Wood, Baxter et al. 1997, Plafker and Gibson 1998). Similarly, the capsid hetero-triplexes also form in the cytoplasm, and the single copy pUL46/mC-BP is also thought to contain NLS sequences (Gibson 1996). Following migration into the nucleus, capsid protomers coalesce to form a spherical pre-capsid shell via protein-protein interactions between different AP molecules, as well as between the hetero-trimers and capsomeres. Auto-activation of PR results in excision of AP and PR itself, occurring concurrently with genome insertion (Yu, Trang et al. 2005). Nucleocapsid maturation culminates in a transitional change to the final icosahedral structure. Three capsid isoforms are produced, two of which are defective and commonly observed in NIEPs: *A* capsids, lacking AP and a copy of the viral genome, and *B* capsids, containing the AP but no genome. *C* capsids contain a copy of the viral genome and represent the fully matured capsid incorporated into virions (Irmieri and Gibson 1985).

1.12.2 HCMV particle tegument compartment

~32 viral proteins have been identified as tegument components, the most abundant by mass being phosphoproteins. The tegument materials are incorporated to the capsid-proximal and envelope-proximal layers (Figure 1.12.) by sequential protein-protein interactions (Phillips and Bresnahan 2011, To, Bai et al. 2011, Tandon and Mocarski 2012).

Initial tegumentation occurs in the nucleus where tegument protein pp150 (ppUL32) associates with pro-capsid shells through interactions with SCP (Sampaio, Cavignac et al. 2005, Dai, Yu et al. 2013). Following nuclear egress, pp150 recruits pUL96, both of which are critical in offering structural support to nucleocapsids in the cytoplasm (Tandon and Mocarski 2008, Tandon and Mocarski 2011). Further capsid-proximal materials include the high molecular weight protein (HMWP; pUL48), and also with the HMWP-binding (HMWP-BP; pUL47) (Yu, Shah et al. 2011). These components are predicted to support intracellular transport of the nucleocapsids along microtubule filaments during *de novo* infection, and also control un-coating and release of genome at the infected cell nucleus (Bechtel and Shenk 2002, Brock, Kruger et al. 2013). The major site of virion tegumentation is the VAC; formed from re-modelled cellular secretory apparatus (endosome/exosomes) at the ER-Golgi intermediate compartment (ERGIC) (Tooze, Hollinshead et al. 1993, Sanchez, Greis et al. 2000, Das, Vasanji et al. 2007). Materials incorporated into the virion from this site likely include those that reside in the envelope-

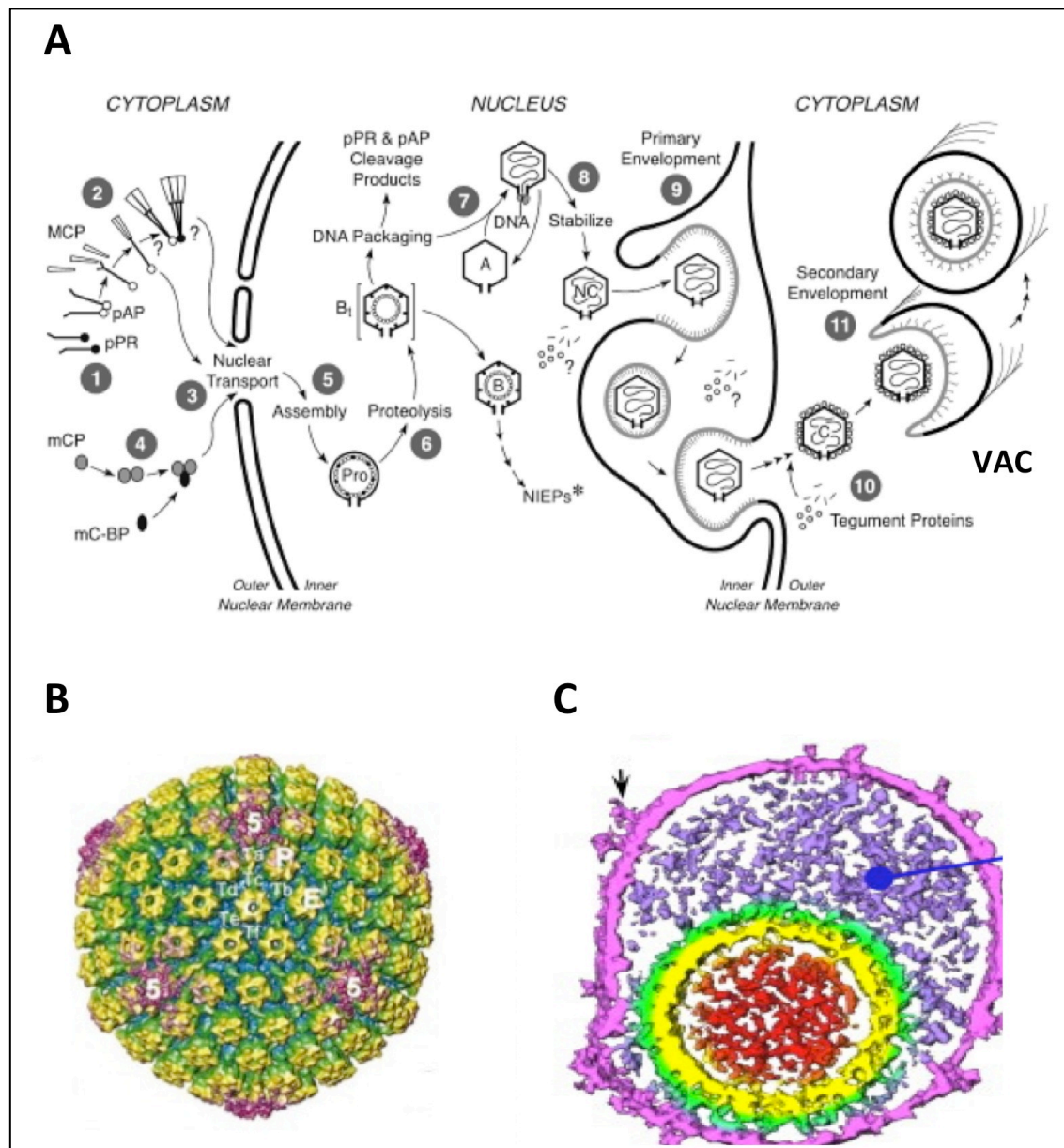


Figure 1.12. HCMV virion maturation pathway. (A) HCMV virion maturation pathway: 1 – pAP and pPR oligomerise in the cytoplasm; 2 – pAP oligomers bind MCP; 3 – nuclear translocation of pAP-MCP aggregates; 4 – mCP and mC-BP associate to form hetero-trimers that move in to the nucleus; 5 – capsid protomers (pUL86/pUL80.5, pUL85/pUL46, and pUL104) coalesce to form spherical pro-capsid shell; 6 + 7 – auto-activation of pUL80 results in cleavage of pUL80 itself and pUL80.5, forming transition B-capsids. Nucleocapsids then: i) undergo defective maturation, lose all pUL80 and pUL80.5, yet gain no genome and become A capsids; ii) retain pUL80 and pUL85, yet gain no genome and become B capsids; iii) lose all pUL80 and pUL80.5, gain a genome and become fully matured C capsids; 8 – initial tegumentation occurring in the nucleus stabilises nascent nucleocapsids; 9 + 10 – nuclear egress of nascent nucleocapsids, by primary envelopment and deposition into the peri-nuclear space, and de-envelopment at outer nuclear membrane and deposition into cytoplasm; 10 + 11 – further tegumentation at cytoplasmic sites and secondary envelopment during budding into secretory vesicles at the VAC. (B) Fully matured icosahedral C capsids: MCP hexons at surfaces and walls are (yellow); MCP pentamers at vertices (purple); triplex sub-units at capsid floor (green). (C) Density map showing distinct protein layers in the HCMV virion, including: genome containing core region (red); capsid (yellow); capsid-proximal tegument (green); envelope proximal tegument (purple); envelope (pink). Figure A reproduced from (Gibson, Bogner et al. 2013); Figure B adapted from (Chen, Jiang et al. 1999), with permission of Elsevier publishers; Figure C adapted from (Yu, Shah et al. 2011), with permission of Elsevier publishers.

proximal tegument layer. The phosphoprotein pp65 (ppUL83) is the most abundant protein in the virion, with a broad range of immuno-modulatory activities to subvert both innate and adaptive immune response pathways (Odeberg, Plachter et al. 2003, Arnon, Achdout et al. 2005). This highly abundant protein performs a role during the assembly of the virion, providing structural support as a scaffold to which other materials are incorporated (Chevillotte, Landwehr et al. 2009). However, whilst pp65 has been demonstrated to augment/enhance virus growth (Becke, Fabre-Mersseman et al. 2010), it is not absolutely required (Schmolke, Kern et al. 1995). The phosphoprotein pp71 (ppUL82) is a trans-activator of IE gene transcription, and virus that does not produce pp71 displays a considerable growth defect in fibroblasts (Cantrell and Bresnahan 2005, Nicholson, Sutherland et al. 2009). Tegument protein pp28 (ppUL99) localizes at the VAC where it orchestrates nascent virion secondary envelopment via interactions with the ER resident chaperone cellular protein Bip/GRP78 (Sanchez, Sztul et al. 2000, Silva, Yu et al. 2003, Seo and Britt 2008, Buchkovich, Maguire et al. 2009). Several cellular proteins, as well as both cellular and virus encoded RNAs, are contained in the virion tegument compartment (Terhune, Schroer et al. 2004, Varnum, Streblow et al. 2004). Whilst virions and NIEPs have similar tegument profiles, DBs contain dis-proportionate levels pp65 (Varnum, Streblow et al. 2004).

1.12.3 HCMV particle envelope

21 viral proteins have been identified as envelope components, 19 of which are predicted or known to be glycosylated {Mocraski, 2013 #777}. One further putative envelope glycoprotein encoding ORF (UL74A) is predicted though, yet to be detected in virions. As few as 5 of the viral encoded virion envelope components are essential for growth in fibroblasts, and of the 16 proteins which are dispensable, only one enhances virus growth in this cell type (Hobom, Brune et al. 2000, Dunn, Chou et al. 2003, Yu, Silva et al. 2003, Spaderna, Kropff et al. 2005). Consequently, several envelope glycoproteins are absent in laboratory strains. In fact, this virion compartment is currently recognised as the major site of heterogeneity amongst laboratory and clinical virus, as demonstrated by the detection of as few as 13 envelope glycoproteins identified in the virions derived from strain AD169 (Varnum, Streblow et al. 2004).

Virion envelope glycoproteins often contain membrane anchors (transmembrane domains or signal peptide sequences) that dictate their transition through the cellular secretory pathway, and also their ultimate retention in vesicle membranes at the VAC where virions acquire undergo secondary envelopment (Sanchez, Greis et al. 2000, Rigoutsos, Novotny et al. 2003). Exosome trafficking can result in glycoproteins being presented at the infected cell surface, with endosomal re-cycling often employed for retrieval and delivery of glycoproteins back to the VAC. Many glycoproteins sequester cellular membrane receptors and their ligands

during endosomal recycling, providing a mechanism by which HCMV modulates cell-signalling pathways to subvert an anti-viral response (Mocarski Jr 2002). Cell-surface presentation of glycoproteins is also believed to induce micro-fusion events between the plasma membranes of adjacent cells, forming pores through which virus can spread by the cell-to-cell route (Gerna, Percivalle et al. 2000). This model is supported by the observation of syncytia formation by 'broad-tropism' strains in dendritic cell and epithelial cell cultures (Gerna, Percivalle et al. 2005, Wang and Shenk 2005), and also by the fact that cytoplasmic material can be transferred between adjacent infected and non-infected cells (Digel, Sampaio et al. 2006).

The most widely studied glycoproteins are those implicated in the binding and entry of virus into cells, as well as the secondary envelopment of nucleocapsids during virion budding into secretory vesicles (see section 1.13.). Besides these, the virion envelope contains several RL11 gene family encoded glycoproteins, with gpRL11 the best described functionally (Lilley, Ploegh et al. 2001, Cortese, Calo et al. 2012). This type I membrane glycoprotein binds the IgG Fc domain and has antibody sequestering activity based on endosomal retrieval. A similar function has been suggested for both gpRL12 and gpRL13, however this has not been demonstrated in the context of infection (Lilley, Ploegh et al. 2001, Atalay, Zimmermann et al. 2002, Cortese, Calo et al. 2012, Cortese 2013). Furthermore, gpRL12 has yet to be identified as a virion envelope component. Another virion envelope glycoprotein, gpUL118-119, has IgG Fc binding properties, and thus may perform similar functions (Mocarski 2004). The GPCR homologue gene family products (pUL33x1, gpUL78, gpUS27 and gpUS28) are best known for their activity in the interference of cell-signalling pathways by sequestration of cellular GPCR ligands (chemokines) (Vischer, Leurs et al. 2006). Interestingly, gpUS28 is capable of binding cell-associated chemokine CX3CL1, potentially contributing to the initial binding of virus to cells during infection (Casarosa, Waldhoer et al. 2005). The function of gpUL132 has yet to be determined, but it has been reported to enhance virus replication in fibroblasts (Hobom, Brune et al. 2000, Spaderna, Kropff et al. 2005).

Two further proteins are frequently co-purified with virions, though are not considered *bona fide* envelope components. This includes the non-glycosylated pUL76 with unknown function, as well as gpUL22A, a secreted glycoprotein with reported chemokine binding and subversion activity (Wang, Bresnahan et al. 2004, Wang, Duh et al. 2004).

1.13 HCMV particle envelope glycoproteins in binding, entry and maturation/egress

Glycoproteins that orchestrate essential mechanisms in the HCMV replication cycle are incorporated into the virion envelope in the form of multimeric glycoprotein complexes. Three distinct glycoprotein complex families have been described: gCI, gCII, gCIII (gC; glycoprotein complex) (Gretch, Kari et al. 1988). Most members of these families are conserved amongst the herpesviruses, and were named (gB, gH, gL, gM and gN) to reflect their functional homology

with glycoproteins in other herpesviruses. The assembly, processing, and composition of these complexes have been investigated by a combination of pulse-chase assays coupled with SDS-PAGE and immuno-blot detection. The stoichiometry of the glycoprotein members contained within each complex has been inferred by comparison of the molecular weight of complete intact complexes, to that of sub-units released under reducing conditions. However, multiple variant complexes have been identified within each family, and it remains possible that glycoprotein complexes exist as a range of variants that may perform distinct functions.

1.13.1 Glycoprotein complex family gCI

Glycoprotein complex family gCI comprises homo-oligomers of the type I glycoprotein gB that is encoded by HCMV ORF UL55 (Cranage, Kouzarides et al. 1986, Mach, Utz et al. 1986, Gretch, Gehrz et al. 1988, Kari, Liu et al. 1990). The primary UL55 translation product is a contiguous polypeptide of ~906-907 amino acids in length that is co-translationally glycosylated to a 150 KDa precursor then modified to 160 KDa form (Britt and Auger 1986, Spaete, Thayer et al. 1988, Britt and Vugler 1989). The 160 KDa intermediate subsequently undergoes proteolytic cleavage by cellular furin-like enzymes to generate a ~55 KDa C-terminal sub-unit (denoted gp55) and a ~116 KDa N-terminal sub-unit (denoted gp116) (Spaete, Thayer et al. 1988, Vey, Schäfer et al. 1995). The mature gB is a homodimer of both gp55 and gp116 sub-units, each held together by di-sulphide bonding that occurs during oligomerisation prior to protein folding in the ER (Britt and Auger 1986, Spaete, Thayer et al. 1988, Britt and Vugler 1989, Britt and Vugler 1992, Billstrom and Britt 1995). The gp55 sub-unit contains a signal sequence and forms the gB transmembrane (TM) domain, and the gp116 sub-unit forms the surface (SU) domain presented at the exterior of the virion, and also externally to cell when presented at the infected cell plasma membrane (Rigoutsos, Novotny et al. 2003).

The HCMV ORF UL55 is polymorphic, with four main genotypes described (gB1, gB2, gB3 and gB4) (Chou and Dennison 1991, Meyer-Konig, Haberland et al. 1998), though further less common genotypes (gB5, gB6, and gB7) have also been described (Rasmussen, Hong et al. 1997, Shepp, Match et al. 1998, Trincado, Scott et al. 2000). Amino acid sequence variation is observed in both the gp55/TM and gp116/SU sub-units (Chou 1992).

1.13.2 Glycoprotein complex family gCII

Glycoprotein complexes assigned to the family gCII exist as hetero-oligomers of glycoproteins gM and gN, and are heterogeneous in terms of the composition and size (Kari, Goertz et al. 1990, Kari and Gehrz 1993, Mach, Kropff et al. 2000, Dal Monte, Pignatelli et al. 2001, Mach, Kropff et al. 2005). Glycoprotein gM, encoded by HCMV ORF UL100 (Lehner, Meyer et al. 1989, Kari, Li et al. 1994), is a ~372 amino acid type III transmembrane protein with seven membrane spanning regions (Mach, Kropff et al. 2000, Rigoutsos, Novotny et al. 2003). gM contains one N-linked carbohydrate modification that is acquired in the ER, and

exists as a 42 KDa glycoprotein in its fully mature state (Mach, Kropff et al. 2000). Glycoprotein gN, encoded by HCMV ORF UL73, is a type I transmembrane glycoprotein or ~138 amino acids, predicted to contain an N-terminal signal peptide and C-terminal TM domain (Mach, Kropff et al. 2000, Rigoutsos, Novotny et al. 2003, Mach, Osinski et al. 2007). This gCII family member is translated as an 18 KDa glycoprotein pre-cursor and is subsequently modified to a 50-60 KDa glycoprotein following the addition of O-linked sugars (Kari and Gehrz 1993, Mach, Kropff et al. 2000). Oligomerisation of gM and gN is required for the progression of either glycoprotein through the secretory pathway, yet complex formation by covalent bonding is not a strict requisite as non-covalently bonded gM/gN complexes have been detected in purified virions (Mach, Kropff et al. 2000, Mach, Kropff et al. 2005).

HCMV ORF UL100 (gM) is highly conserved amongst different strains. In contrast to this, UL73 (gN) that is one of the most polymorphic HCMV ORFS and exists in four main genotypes, two of which contain further divergent variants (gN1, gN2, gN3a, gN3b, gN4a, gN4b, gN4c, and gN4d), with amino acid variation is most commonly isolated to the N-terminal domain that is exposed to the external, environment (Lehner, Stamminger et al. 1991, Dal Monte, Pignatelli et al. 2001, Pignatelli, Dal Monte et al. 2003).

1.13.3 Glycoprotein complex family gCIII

Glycoprotein complexes of the family gCIII contain glycoproteins gH and gL as a disulphide bonded heterodimer (Kaye, Gompels et al. 1992, Huber and Compton 1999), forming a scaffold to which alternative auxiliary glycoproteins associate. Glycoprotein gH, encoded by the HCMV ORF UL75, is a ~742-744 amino acid type I transmembrane protein containing a C-terminal TM domain, and an N-terminal signal sequence (Rigoutsos, Novotny et al. 2003). gH is first translated to a 75 KDa precursor, then modified to an 86 KDa mature glycoprotein (Cranage, Smith et al. 1988, Kinzler, Theiler et al. 2002). Glycoprotein gL, encoded by HCMV ORF UL115 (Kaye, Gompels et al. 1992), is a ~278 amino acid in length, and whilst no TM domain is predicted, an N-terminal signal peptide is predicted (Rigoutsos, Novotny et al. 2003). gL contains just a single glycosylation modification and in its mature state is 34 KDa (Kinzler, Theiler et al. 2002). The gH/gL heterodimer is formed shortly following translation at the ER, and this interaction is sufficient for the progression through the secretory pathway and presentation at the cell surface (Cranage, Smith et al. 1988, Kaye, Gompels et al. 1992, Spaete, Perot et al. 1993).

HCMV ORFS UL75, existing in two genotypes (gH1 and gH2) (Chou 1992, Rasmussen, Geissler et al. 2002), and UL115, existing in four distinct genotypes (gL1, gL2, gL3 and gL4) (Rasmussen, Geissler et al. 2002), are moderately polymorphic. Furthermore, amino acid variation is most commonly observed in the N-terminal regions in each of these gCIII family members.

1.13.4 Entry of HCMV particles into cells

The binding and entry of HCMV into cells is a multi-phase process, engaging sequential glycoprotein complex-host cell receptor interactions (Figure 1.13.). Envelope glycoprotein-host cell receptor interactions also stimulate host-cell signalling cascades that may contribute to entry, cell migration, and/or prime the cell for virus replication (Compton 2004, Isaacson, Juckem et al. 2008). Whilst NIEPs and DBs contain envelopes comparable to that in virions and can also achieve penetration of cells, this section will focus on the entry of virions.

1.13.4.1 *Binding and docking of virions at cell surface*

The first step in binding is the tethering of virions to the cell surface, facilitated by low-affinity interactions with virion envelope glycoproteins and heparan sulfate proteoglycans (HSPGs); these molecules that are ubiquitous on all cell types permissive to infection. Both gB and gM/gN have heparan sulphate binding abilities, underlying the fact that virus-cell binding can be reduced using anti-bodies against these glycoprotein complexes (Compton, Nowlin et al. 1993, Kari and Gehrz 1993). The next phase includes the docking of virions at the cell surface via more stable glycoprotein-receptor interactions. gB is also implicated in the docking of virions, and has been shown to bind cells in a bi-phasic manner, firstly through low affinity interaction with HSPGs, and subsequently through higher affinity interaction with other receptors (Boyle and Compton 1998). Several candidate host-cell receptor/entry mediator molecules have been proposed, though their significance in cell-signalling, virion docking, or mediating entry is not clearly defined. Epidermal growth factor receptor (EGFR) is one proposed participant, as suggested when cells expressing EGFR were permissive to infection, whilst otherwise identical EGFR knock out cells were refractory to infection (Wang, Huong et al. 2003). Furthermore, in this work, gB was been shown to co-precipitate with EGFR from infected cell lysates, and also contain an epitope with homology to EGF ligand. However, contrasting reports claim no such reduction in virus infection efficiency in fibroblasts, epithelial and endothelial cells when EGFR was blocked using antibodies, implicating the use of other receptors in the absence of available EGFR (Isaacson, Feire et al. 2007). One group has demonstrated the blocking of infection using anti-bodies and siRNAs to PDGFR (Soroceanu, Akhavan et al. 2008), whilst another group reported no such effect, instead suggesting this receptor invokes entry by an abnormal pathway (Vanarsdall, Wisner et al. 2012). Integrins are also proposed as entry mediators, and gB contains a disintegrin like domain that has been shown to interact with $\beta 1$ integrin (Feire, Koss et al. 2004). The proposed involvement of this interaction was supported by the blocking of infection in fibroblasts using synthetic gB disintegrin peptides, as well as integrin blocking antibodies (Feire, Roy et al. 2010). Alternatively, gH has been shown to bind $\alpha V\beta 3$ integrins, offering further support to the proposed role of this class of cell surface molecule in the early stages of infection (Wang, Huang et al. 2005).

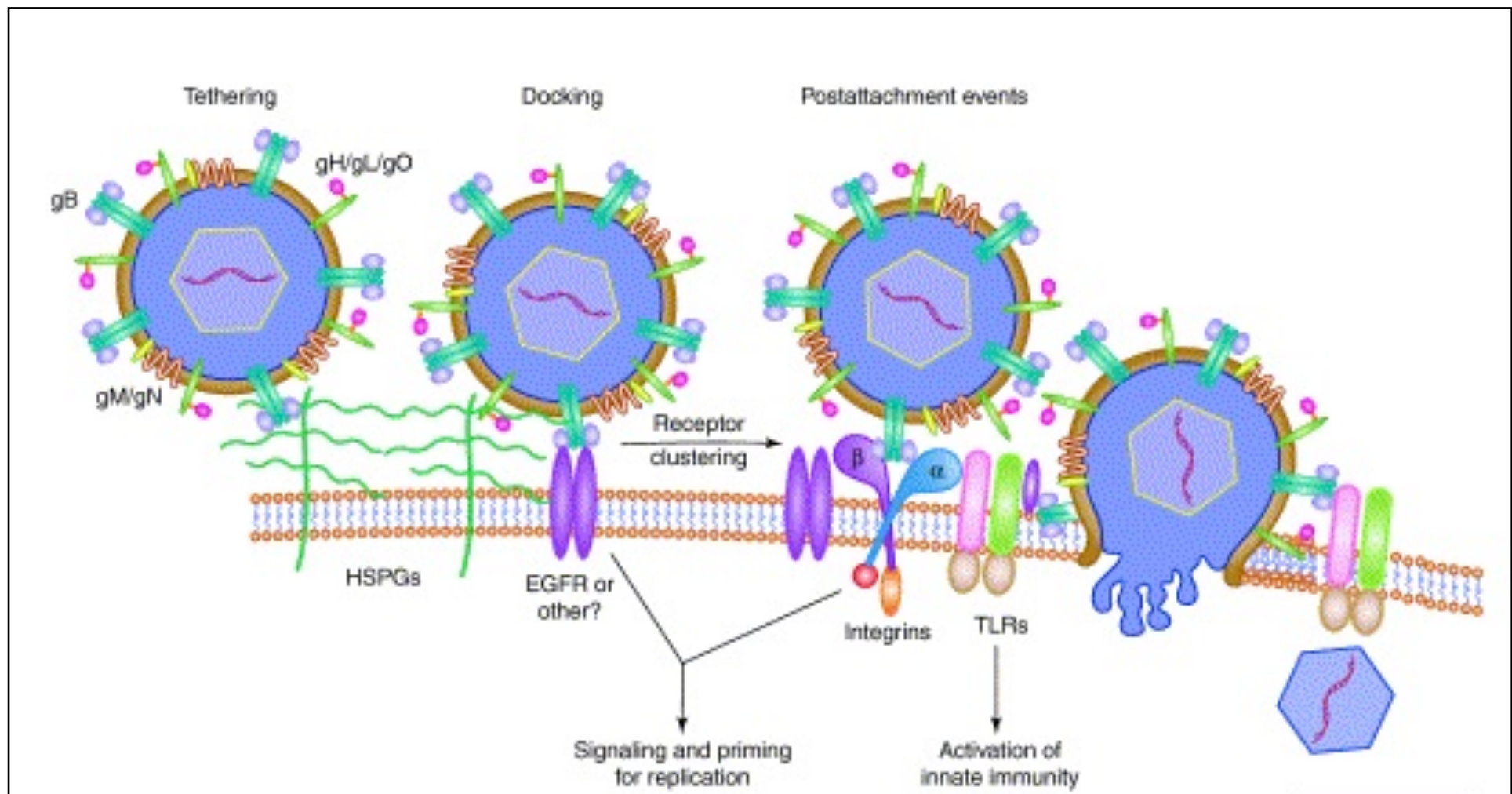


Figure 1.13. Sequential phases of HCMV virion binding and entry into cells. Model summarises entry into fibroblasts, with fusion of the virion envelope and cell plasma membrane prior to penetration. Figure reproduced from (Compton 2004), with the permission of Elsevier publishers.

1.13.4.2 *Virion envelope-cell membrane fusion – penetration of the cell and internalisation of virion contents*

Fusion between the virion envelope and cellular membranes is orchestrated by glycoproteins gB and gH/gL complexes; together, gB, gH, and gL form the core fusion machinery conserved amongst the herpesviruses (different HCMV gH/gL complexes are required for infection of different cell-types; see section 1.14). The observation that virus retains the ability to bind, but not enter cells when these glycoproteins are absent or blocked by antibodies, together with the use of the chemical fusogen polyethylene glycol (PEG) to overcome these entry defects, confirmed the activity of gB and gH/gL complexes in membrane fusion (Keay and Baldwin 1991, Vanarsdall, Ryckman et al. 2008, Isaacson and Compton 2009, Wille, Wisner et al. 2013). The contribution of gH/gL and gB complexes to membrane fusion has been further investigated in cells transduced to endogenously express each glycoprotein complex singly or in combination, with fusogenic potential displayed by the formation of syncytia. Cell surface presentation of gH/gL or gB alone was insufficient to induce cell fusion, while co-presentation of all three glycoproteins, whether in *cis* (in the same cells) or in *trans* (on opposing cells), was necessary and sufficient (Milne, Paterson et al. 1998).

The currently accepted model of HCMV glycoprotein induced fusion includes gB as the major fusogenic component, and gH/gL complexes as a receptor-binding activator of gB, and is supported by several experimental observations. Firstly, gB expressing cells are permissive to infection by gB-null mutant virus, whilst gH/gL-expressing cells could not support infection of gH-null mutants (Vanarsdall, Ryckman et al. 2008, Wille, Wisner et al. 2013). This was interpreted to suggest that presentation of gB in any orientation was sufficient for fusion; in this case gB was inserted into the virion envelope and fusion was led from within the transduced cell plasma membrane, whilst gH/gL was required to be presented in an orientation emanating from the virion, suggestive of receptor binding activity. Furthermore, this interpretation was in agreement with results from interference assays where endogenous expression of gH/gL complexes, but not gB, could block infection by cell-free virus into cells (Ryckman, Chase et al. 2008, Vanarsdall, Chase et al. 2011). This finding was again interpreted to indicate that gB did not bind any saturable cellular receptor, whilst gH/gL did.

Thus, whilst gB contributes to the initial binding of virions to the cell and potentially stimulates cellular signalling pathways with various outcomes (Simmen, Singh et al. 2001, Isaacson and Compton 2009), the major activity of this glycoprotein is in the post-attachment fusion of cellular and virion membranes. Further evidence for the specific contribution of each of these components has come from investigation of the conformational and structural features of each glycoprotein(s). The crystal structure of the HSV-2 gH/gL ectodomain has been determined, and contains no architectural features that resemble those of known viral fusogen proteins (Chowdary, Cairns et al. 2010). However, determination of the HCMV gB crystal structure identified two internal hydrophobic fusion loops that are a hallmark of class III viral

fusion proteins (Sharma, Wisner et al. 2013). In the virion envelope, the ectodomain of three adjacent gB molecules form trimers to bury the hydrophobic regions of this glycoprotein complex when in the pre-fusion conformation. Conformational changes in fusion proteins result in the exposure of the hydrophobic fusion loops to the external environment, then the triggered fusogen protein is rapidly inserted into the adjacent membrane (host cell). This disrupts the outer membrane layer and ultimately culminates in fusion between the opposed membranes (Baquero, Albertini et al. 2013). Furthermore, each of gB and gH have been shown to contain alpha helical coiled-coil domains, in the gp55 TM domain gB and near the N-terminal region in gH; interaction between these amino acid structural motifs are believed to necessary for the gB conformational change required to trigger fusion (Lopper and Compton 2004).

1.13.5 Progeny virion maturation and egress

Glycoprotein complexes of the gCII (gM/gN), and gCIII (gH/gL) families are also implicated in the maturation and egress of progeny virions. Site-directed mutagenesis experiments demonstrated the presence of motifs that orchestrated secondary envelopment at the VAC in cytoplasmic tails of both gM and gN (Krzyzaniak, Mach et al. 2007, Mach, Osinski et al. 2007). The contribution of gH/gL complexes to virion maturation and egress is less clearly established (see 1.14.3).

1.14 Variant gH/gL glycoprotein complexes encoded by HCMV

More recently, distinct glycoprotein complex variants in the gCIII family have been identified and each has been shown to be required for efficient virus replication in different cell types (Figure 1.14.).

1.14.1 gCIII; gH/gL/gO

The classically described gCIII variant is a 240 KDa hetero-trimeric complex that includes gH/gL, along with glycoprotein gO encoded by HCMV ORF UL74 (Huber and Compton 1997, Li, Nelson et al. 1997, Huber and Compton 1998, Huber and Compton 1999) (Figure 1.14.A). This glycoprotein is variable in length, ranging from 464-472 amino acids. Like gL, gO is not predicted to contain a transmembrane domain, though is predicted to contain a N-terminal signal peptide sequence (Rigoutsos, Novotny et al. 2003). Efficient transport of gO from the ER is achieved only when incorporated into gH/gL/gO (Huber and Compton 1999, Theiler and Compton 2001, Kinzler, Theiler et al. 2002). The assembly of the hetero-trimeric gH/gL complex variant commences with the addition of a 100 KDa gO pre-cursor that is subsequently modified at the Golgi to contain both N-linked and O-linked carbohydrates in its fully matured 125 KDa isoform (Huber and Compton 1999). The inclusion of gO increases the efficiency of gH/gL transport through the secretory pathway, and thus has gO been assigned chaperone-like activities (Ryckman, Chase et al. 2010, Wille, Knoche et al. 2010).

HCMV ORF UL74 (gO) is one of the most polymorphic HCMV genes that exists as 4 distinct genotypes, two of which contain further distinguishable variants (gO-1a, gO-1b, gO-1c, gO-2a, gO-2b, gO-3, gO-4), with amino acid variation commonly centred around the first 100 N-terminal residues (Huber and Compton 1998, Paterson, Dyer et al. 2002, Rasmussen, Geissler et al. 2002, Mattick, Dewin et al. 2004).

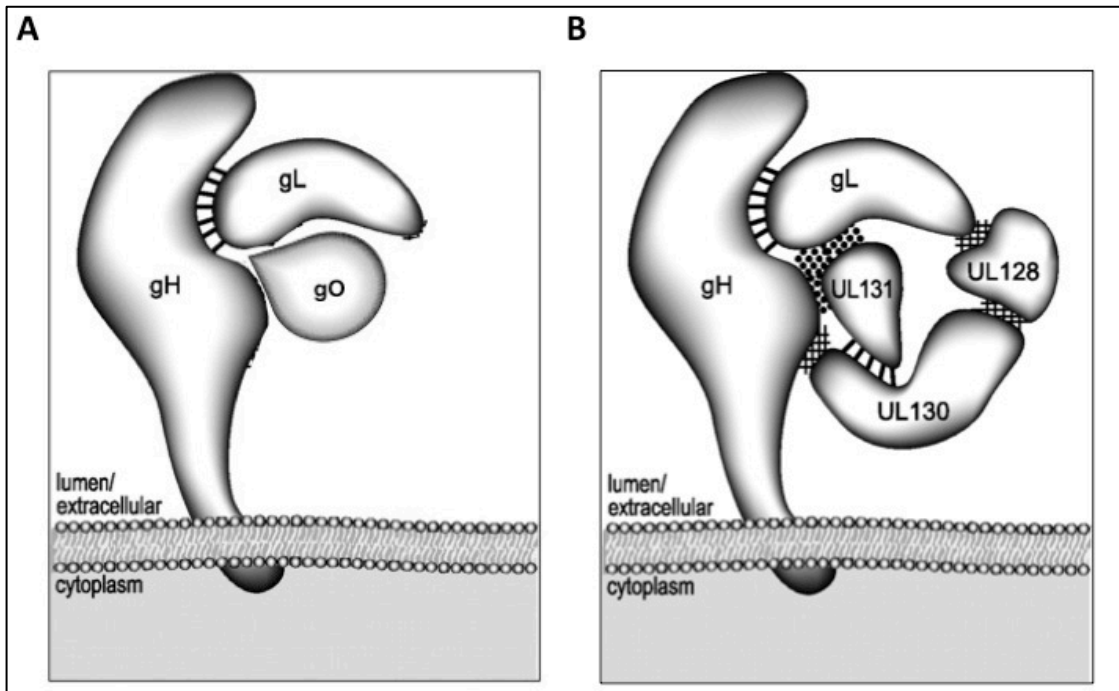


Figure 1.14. Predicted conformations of alternative gH/gL glycoprotein complexes. (A) The classical gCIII glycoprotein complex gH/gL/gO. (B) The more recently described gCIII family member containing UL128L ORF-encoded glycoprotein sub-units: gH/gL/UL128/gpUL130/gpUL131A. Lines depict covalent bonds; cross-hatching depicts non-covalent bonds. The surface created by gH/gL to which gpUL131A bind is depicted by dark dots. Image 1.14. (B) reproduced from (Ryckman, Rainish et al. 2008) and adapted for 1.14.A, with permission of American Society for Microbiology Publishers.

1.14.2 gH/gL/gpUL128/gpUL130/gpUL131A

The most recently described gH/gL variant is a hetero-pentameric complex containing UL128L-encoded glycoproteins: gH/gL/gpUL128/gpUL130/gpUL131A (Wang and Shenk 2005, Ryckman, Rainish et al. 2008) (Figure 1.14.B).

gpUL128 contains 171 amino acids with a signal peptide sequence predicted to span the first 27 N-terminal residues (Akter, Cunningham et al. 2003, Sun, Ji et al. 2009). A single site for N-linked glycosylation is predicted at amino acid position 55 (Rigoutsos, Novotny et al. 2003), and gpUL128 can be detected by immunoblot analysis as multiple bands that include a

16 KDa entity likely representing the fully mature glycoprotein (Wang and Shenk 2005). A putative beta (cc) chemokine-like domain is observed in the N-terminal half of the gpUL128, consisting of conserved cysteine residues at positions 30, 31, 49 and 64 (Akter, Cunningham et al. 2003, Rigoutsos, Novotny et al. 2003, Hahn, Revello et al. 2004). Amino acid diversity is low in gpUL128 (95.4-96.9% aa identity in wildtype virus), and most frequently observed in the N-terminal region (Baldanti, Paolucci et al. 2006).

gpUL130 is the largest of the UL128L encoded glycoproteins, 214 amino acids long, with a signal peptide sequence that spans the first 25 N-terminal residues (Rigoutsos, Novotny et al. 2003, Sun, Ji et al. 2009). gpUL130 is a triglyceride containing sites for N-linked glycosylation modifications at residues 85, 118 and 201, as well as a putative methylation site at residue 145 (Rigoutsos, Novotny et al. 2003, Patrone, Secchi et al. 2005). Thus, gpUL130 identified by multiple bands on immunoblot analysis, with a 35 KDa band representing the fully mature glycoprotein (Patrone, Secchi et al. 2005, Wang and Shenk 2005). Like gpUL128, gpUL130 also contains a putative chemokine domain; the amino acid tract 46-120 resembles the structure of CXC chemokines, although only two of the four cysteine residues definitive of this chemokine class are present (aa residues 52 and 83) (Akter, Cunningham et al. 2003, Rigoutsos, Novotny et al. 2003, Hahn, Revello et al. 2004). gpUL130 displays the lowest amino acid sequence identity of the UL128L-encoded glycoproteins (91.1-96.6% aa identity in wildtype virus), with variation more diffuse than in UL128 and UL131A, though most commonly observed within residue tracts 8-78 and 128-151 (Baldanti, Paolucci et al. 2006).

gpUL131A is the smallest of the UL128L encoded glycoproteins, 129 amino acids in length, and predicted to contain a single N-linked glycosylation site located at amino acid residue 70 (Rigoutsos, Novotny et al. 2003, Sun, Ji et al. 2009). This UL128L-encoded glycoprotein is detected as a band of 14-15 KDa on immunoblot analysis (Adler 2006). gpUL131A is the most conserved UL128L-encoded glycoprotein (97.0-97.2% aa identity in wildtype virus), with variations localized to the amino acid tract covering residues 22-51 in the first exon that also contains a peptide signal spanning the first 18 N-terminal amino acids (Rigoutsos, Novotny et al. 2003, Adler 2006, Baldanti, Paolucci et al. 2006, Sun, Ji et al. 2009).

The UL128L encoded sub-units are stabilized in the pentameric complex by both disulphide and non-covalent bonds (Ryckman, Rainish et al. 2008). gpUL130 and gpUL131A form associations stabilized by disulphide bonds, whilst gpUL130 forms further associations with both gH and gpUL128 via non-covalent bonds. gpUL128 in turn associates with gL via non-covalent interactions, and gpUL131A forms non-covalent association with a surface created by both gH and gL (Ryckman, Rainish et al. 2008). The inclusion of each UL128L encoded sub-unit promotes the incorporation other sub-units, with gpUL128 incorporation in turn promoting incorporation of gpUL130 and gpUL13A, whilst the absence of gpUL131A

leads to incorporation of an immature gpUL130 variant. Each of the UL128L encoded sub-units is absolutely required for complex assembly, with loss of any one component resulting in the retention of the remaining complex sub-units in the ER (Adler 2006, Ryckman, Rainish et al. 2008). Furthermore, the incorporation of UL128L proteins increases the efficiency of gH/gL transport via the secretory pathway, in a similar manner to gO. Peptides within the UL128L encoded sub-units that contribute to the pentameric complex functionality have been investigated by charge-cluster-to-alanine mutagenesis. Functionally significant peptides have been proposed to exist within the central region of gpUL128 (aa 72-106), in particular, residue tracts 72-74, 82-86, and 104-108 (Schuessler, Sampaio et al. 2008); in the C-terminal region (aa 142-214) of gpUL130, particularly involving an amino acid sequence motif (PNLIV) that closely resembles a cell-penetrating penta-peptide (CPP) motif implicated in internalization of proteins by endocytosis (Rhee and Davis 2006, Schuessler, Sampaio et al. 2010); as well as in numerous regions dispersed throughout gpUL131A (Schuessler, Sampaio et al. 2012).

1.14.3 Roles of gH/gL complex variants in entry of cell-free HCMV into cells

HCMV enters different cell-types by different pathways. In fibroblasts, virus penetration is achieved by direct pH-independent fusion of the virion envelope and host cell plasma membrane (Compton, Nepomuceno et al. 1992), whilst in endothelial cells and epithelial cells, virions are first internalised by endocytosis prior to fusion of the virion envelope and endosomal membranes (Bodaghi, Slobbe-van Drunen et al. 1999, Ryckman, Jarvis et al. 2006, Sinzger 2008). There are conflicting reports as to the requirement of endosomal acidification for the induction of fusion between the virion envelope and endosomal membranes (Ryckman, Jarvis et al. 2006, Sinzger 2008). Internalisation into dendritic cells has been reported to occur by a macropinocytosis-like pathway, with subsequent fusion occurring via a pH-independent mechanism (Haspot, Lavault et al. 2012). The alternative gH/gL glycoprotein complexes facilitate the internalisation of cell-free virus in different cell types. This model is supported by data from interference assays where endogenous expression of gH/gL/gO rendered fibroblasts but not epithelial cells refractory to infection, whilst endogenous expression of gH/gL/gpUL128/gpUL130/gpUL131A blocked infection in epithelial cells but not fibroblasts (Ryckman, Chase et al. 2008, Vanarsdall, Chase et al. 2011). Hence each distinct gH/gL complex potentially engages alternative saturable receptor molecules present on different target cell types (Figure 1.14.B).

Whilst the gH/gL/gO complex facilitates the entry of cell-free virus into fibroblasts, the specific contribution of gO to this activity is unclear. A direct contribution of glycoprotein gO to fusion has been suggested since anti-gO antibodies can prevented virus-mediated cell fusion (fusion-from-without) (Paterson, Dyer et al. 2002). However, syncytia formation in cells endogenously expressing gH/gL and gB was not enhanced by co-expression of gO (fusion-from-within) (Vanarsdall, Chase et al. 2011). Glycoprotein gO has also been reported to reduce

the inhibitory impact of anti-gH and anti-gB antibodies on the entry of virus into cells, suggesting that gO may play an indirect role by preventing exposure of the HCMV virion core fusion machinery to components of the humoral immune system (Jiang, Sampaio et al. 2011); a similar role has been postulated for glycoprotein gN (Kropff, Burkhardt et al. 2012). A role for gH/gL/gO during the entry of non-fibroblasts has also been suggested (Wille, Knoche et al. 2010). In this work, virions from a gO-null mutant remained at the surface of inoculated epithelial and endothelial cells, and were only internalised following treatment with fusogenic agents, suggestive of an entry defect.

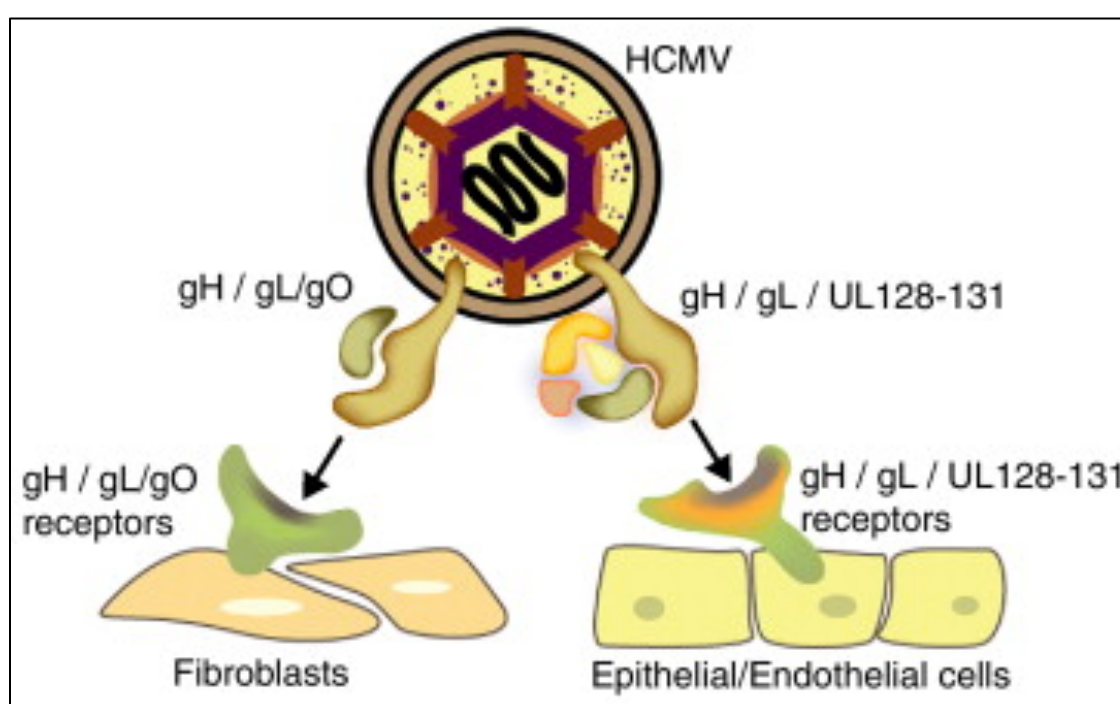


Figure 1.15. Requirement of alternative gH/gL complexes for entry into different cell types. gH/gL/gO, required for infection of fibroblasts, and gH/gL/gpUL128/gpUL130/gpUL131A, required for entry into epithelial and endothelial cells, as well as others. Figure reproduced from (Vanarsdall and Johnson 2012), with permission of Elsevier publishers.

No reports detailing any specific role of the UL128L-encoded glycoproteins during the entry of cell-free virus into fibroblasts are currently available. However the requirement of the pentameric complex for entry into non-fibroblasts is strict, with UL128L-null mutants markedly restricted in the ability to infect epithelial and endothelial cells (Sinzger, Knapp et al. 1997, Hahn, Revello et al. 2004, Adler 2006, Straszewski, Patrone et al. 2011). Furthermore, antibodies specific for UL128L encoded proteins block HCMV infection of epithelial and endothelial cells, but not fibroblasts (Wang and Shenk 2005, Adler 2006). The pentameric

complex is active in a post-attachment step, though prior to penetration of virus following fusion between the viral envelope and endosomal membranes; virions lacking the pentameric gH/gL complex variant are internalised into endothelial cells by endocytosis, but are unable to deliver its genome to the nucleus (Slobbe-van Drunen, Hendrickx et al. 1998, Sinzger, Kahl et al. 2000). The UL128L proteins have also been suggested to mediate interactions between gH and gB, which in turn trigger the conformational changes that occur in gB to facilitate membrane fusion and the entry of virus into endothelial cells (Patrone, Secchi et al. 2007).

1.14.4 Roles of gH/gL complex variants in cell-to-cell spread of HCMV

Assembly of the trimeric gH/gL complex variant (gH/gL/gO) is not absolutely required for HCMV replication and dissemination in fibroblasts; gO-null mutants reconstituted from BAC-cloned HCMV genomes display a severe growth defect compared to parental virus, but do remain viable (Hobom, Brune et al. 2000, Dunn, Chou et al. 2003). Thus, gO enhances the growth of HCMV in fibroblasts. Whilst gO-null mutant virus has been described to display a small plaque phenotype in fibroblasts, cell-to-cell spread in epithelial or endothelial cells has been observed to be unaffected, if not enhanced (Hobom, Brune et al. 2000, Paterson, Dyer et al. 2002, Jiang, Adler et al. 2008, Wille, Knoche et al. 2010). Therefore, assembly of the trimeric gH/gL complex is potentially inhibitory to the cell-to-cell spread of virus in non-fibroblasts.

A hallmark of ‘laboratory-strains’ that contain ablating UL128L mutations is greater plaque formation in fibroblast cultures; hence assembly of the pentameric complex is correlated with reduced cell-to cell spread in this cell type. Interestingly, a UL74-null (gO-null) and UL128L-null double mutant virus has been shown to be completely non-viable in fibroblasts (Jiang, Adler et al. 2008). Thus, whilst UL128L is non-essential and impedes the cell-to-cell spread of virus in fibroblasts, the pentameric gH/gL complex does apparently contribute to this activity in the absence of the trimer. As with the dissemination of cell-free virus, the cell-to-cell spread of a UL128L-null mutant is also heavily restricted in non-fibroblasts (Sinzger, Knapp et al. 1997, Hahn, Revello et al. 2004, Adler 2006, Straschewski, Patrone et al. 2011).

1.14.5 Roles of gH/gL complex variants in cell-free HCMV production

The trimer has been implicated in progeny virion morphogenesis and/or release, and is potentially involved in the secondary envelopment of nascent nucleocapsids similar to gM/gN complexes (Jiang, Adler et al. 2008). This was based on the observation that reconstitution of a gO-null mutant in fibroblasts resulted in the accumulation of non-enveloped nucleocapsids in the cytoplasm, whilst infections with the gO-null mutant in both fibroblasts and endothelial cells were predominantly cell-associated and produced low titres of cell-free virus. However these results contradicted those from the previously mentioned study (Wille, Knoche et al. 2010), where a gO-null mutant displayed no difference in cell-free virus production compared to

parental virus. Expression of intact UL128L ORFs has been clearly demonstrated to result in a vast reduction in cell-free virus production, but the mechanism by which pentameric complex assembly contributes to this growth phenotype is not established. Interestingly, a gO-null mutant has also been reported to produce virions containing increased quantities of gH/gL/gpUL128/gpUL130/gpUL131A, suggesting competition for available gH/gL for the formation of either complex variant (Wille, Knoche et al. 2010).

1.14.6 The role of HCMV gH/gL complex variants during virus replication: summary

The requirements for each the distinct gH/gL complexes in different cell types are widely contrasting, however each complex variant may not be totally redundant in cell types where they are not strictly required. In summary, the trimer complex (gH/gL/gO) is necessary for the entry of cell-free virus in to fibroblasts, and is a major participant in the cell-to-cell dissemination of virus in this cell-type. Furthermore, gH/gL/gO may also be involved in the entry of cell-free virus in non-fibroblasts (e.g. epithelial and endothelial cells). The pentameric complex (gH/gL/gpUL128/gpUL130/gpUL131A) is necessary for the entry and cell-to-cell dissemination of HCMV in a range of naturally targeted cell types other than fibroblasts. The specific contribution of either gH/gL variant to the production of cell-free virus during infection of fibroblasts and non-fibroblasts remains unclear.

1.15 Biologically relevant HCMV strains currently available for research *in vitro*

There is a clear need for translational HCMV research to be based on strains that best reflect the causative agent of disease, including those that assemble the pentameric complex and are suitable for investigation in a broad range of naturally infected cell types. The availability of low-passage BAC-cloned strains partially addressed the problem where recent clinical isolates undergo rapid adaptation to cell culture, and provided a stable source from which the same ‘broad-tropism’ virus could be reconstituted.

1.15.1 ‘Broad-tropism’ BAC-cloned strains

Several of the currently used BAC-cloned HCMV strains produce virus with broad tropism resembling that of wild type virus, and by extension, contain intact UL128L genome regions. The most widely used ‘broad-tropism’ BAC-cloned strains include: TR-BAC, captured from isolate TR derived from the ocular swab of an AIDS patient with retinitis, with unknown passage history (Smith, Taskintuna et al. 1998, Murphy, Yu et al. 2003); TB40-BAC4, cloned from the mixed virus population TB40/E derived from the throat wash of a bone marrow transplant recipient, following 5 passages in fibroblasts and 22 passages in endothelial cells (Sinzger, Schmidt et al. 1999, Sinzger, Hahn et al. 2008); and FIX-BAC, cloned from the clinical strain VR1814 derived from the cervical swab of a pregnant lady with primary infection, following 46 passages in fibroblasts (Revello, Baldanti et al. 2001, Hahn, Khan et al.

2002). However, none of these BAC-cloned strains contain the full complement of wildtype genes contained by virus *in vivo*: each lack sequences in U_S genome region where the BAC vector sequences were stably inserted; and each has been demonstrated to contain lesions affecting ORFs throughout the genome. Importantly, sequencing data from the un-passaged wildtype progenitors viruses from which these strains were cloned are not available, and it is not possible to confirm whether the ORFs in these strains contain wildtype sequence.

1.15.2 Cloning of the wildtype Merlin genome as a BAC-construct

In light of the uncertainties regarding the coding potential of the above-described ‘broad-tropism’ BAC-cloned strains, the genome from the wildtype strain Merlin was captured as a BAC construct (Stanton, Baluchova et al. 2010). The Merlin isolate was subjected to limited passaging in fibroblasts prior to BAC cloning, in which time virus containing mutations at 2 genetic loci were selected: RL13, and UL128. The premature stop codon in UL128 abrogated gH/gL/gpUL128/gpUL130/gpUL131A functionality, and virus displayed a growth phenotype and tropism range similar to that of ‘laboratory-adapted’ strains (i.e. was propagated efficiently and produced high titres of cell-free virus, though displayed restricted tropism for non-fibroblasts). This mutation was repaired in the BAC by reference to sequencing data derived directly from the un-passaged wildtype virus in the clinical sample. Furthermore, unlike many other BAC-cloned strains, the Merlin-BAC contains a self-excising BAC vector sequence and no genes were sacrificed during the clonal of this strain. As a result, the Merlin-BAC represents the only cloned HCMV genome with coding potential identical to that of wildtype virus *in vivo*.

1.15.3 Limitations of the wildtype Merlin-BAC construct

Following the repair of ORF UL128 in the BAC-cloned Merlin genome, the inhibitory impact of wildtype UL128L (and by extension the pentameric gH/gL complex variant) on the production of cell-free virus was clearly demonstrated; infections with virus from the Merlin-BAC variant containing wildtype UL128L produced peak titres of cell-free virus that were ~1000-fold lower than that produced by otherwise identical virus containing the ablating UL128 mutation (UL128L^{mut}) (Stanton, Baluchova et al. 2010). Furthermore, *de novo* UL128L mutations were repeatedly selected following attempts to propagate of virus from the Merlin-BAC containing wildtype UL128L sequence. Thus, similar to low-passage clinical isolates, this newly generated reagent is limited in its utility. Interestingly, indirect comparisons between growth kinetics data for viruses from the wildtype Merlin-BAC to that published for viruses from the TR-BAC, TB40-BAC4 and FIX-BAC clones, suggested that infections with these latter strains produced greater titres of cell-free progeny virus. Furthermore, unlike Merlin, these strains have not been reported to lose their broad tropism when passaged on fibroblasts. Thus, these strains potentially contain features that allow virus to be propagated more efficiently without involvement of mutations in the UL128L genome region.

1.16 Aims of current study

This project aimed to investigate ways by which viruses derived from the wildtype Merlin BAC-cloned strain could be propagated more efficiently, and more stably *in vitro*, as well as to assess the biochemical composition of virions from this strain. There were four avenues of exploration:

- Investigate whether viruses reconstituted from the TR-BAC, TB40-BAC4 and FIX-BAC cloned genomes grow more efficiently compared to virus from the Merlin-BAC, and also whether each contained features that could be transferred to the Merlin-BAC genome to produce viruses that grow more efficiently.
- Determine the impact of any such features on specific activities in the viral replication cycle.
- Assess the impact of any altered growth phenotype displayed by viruses from the Merlin-BAC variants produced on the genetic stability of virus in cell culture, and develop methods of growing Merlin virus without the risk of mutations in UL128L being selected.
- Determine the proteome composition of wildtype Merlin virions to inform studies of tropism, entry of virus into cells, and vaccine development.

2 Methods and Materials

2.1 Solutions, buffers and media

Unless otherwise stated, the tissue culture reagents used were from the Gibco product line of Invitrogen/Life Technologies, and the analytical grade chemicals and bacterial culture reagents used were from Sigma and/or Fisher. Double-distilled ultra-pure water (ddH₂O) used for the preparation of solutions, buffers, and media was delivered from a Purelab Ultra water system (Elgin).

2 x DMEM:	50% (v/v) ddH ₂ O, 20% (v/v) 10X MEM, 20% (v/v) FBS, 1000 U.L ⁻¹ penicillin, 1000 µg.L ⁻¹ streptomycin, 4 mM L ⁻¹ glutamine, 0.45% (w/v) sodium bicarbonate.
TE buffer (pH 8.0):	10 mM Tris, 1 mM EDTA in ddH ₂ O.
DNA loading buffer:	ddH ₂ O, 10 mM Tris-HCl, 0.03% bromophenol blue, 60% glycerol, 60 mM EDTA.
DNA running buffer:	1x TAE (0.04M tris acetate, 2 mM Na ₂ EDTA), in ddH ₂ O
Luria-Bertani (LB) broth:	ddH ₂ O, 1% (w/v) tryptone (Fluka), 0.5% (w/v) yeast extract, 0.85M NaCl.
All formulated bacterial culture media was sterilized by autoclaving. Prior to addition of supplements, media was tempered to 50°C.	
LB Agar:	LB broth, 15% (w/v) agar (Oxoid).
LB Agar + Sucrose:	ddH ₂ O, 1% (w/v) tryptone (Fluka), 0.5% (w/v) yeast extract, 15% (w/v) agar (Oxoid) 5% (w/v) sucrose.
Ampicillin solution:	50 mg/mL Ampicillin (Melford) in ddH ₂ O, sterilised by passage through 0.22 µm filters.
Chloramphenicol solution:	12.5 mg/mL chloramphenicol (Melford), in 100% ethanol
IPTG solution:	100 mM isopropyl β-D-1-thiogalactopyranoside (Melford), in ddH ₂ O.
X-gal solution:	40 mg/mL 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (Melford) in N-dimethyl formamide.

PBS (pH 7.3):	0.8% (w/v) sodium chloride, 0.02% (w/v) potassium chloride, 0.115 % (v/v) di-sodium hydrogen phosphate, 0.002% potassium di-hydrogen phosphate (Oxoid) in ddH ₂ O.
4% PFA solution	PBS, 4% (w/v) paraformaldehyde.
IF wash/block buffer:	PBS, 1% (w/v) BSA. Depending on use, IF buffers differed on the inclusion or exclusion of sodium azide. For IE1 immuno-staining, 0.01 % (w/v) sodium azide was added to IF buffer.
SILAC 'L' Medium:	<p>SILAC DMEM (Thermo) supplemented with 10% (v/v) dialysed FBS, 500 U.L⁻¹ penicillin, 500 µg.L⁻¹ streptomycin, 250 µg.L⁻¹ 'light' L-Proline, 50 µg.L⁻¹ 'light' L-Lysine dihydrochloride (CK Isotopes), 50 µg.L⁻¹</p> <p>Dialysed FBS was cleared of large debris by centrifuging at x g for 10 mins at room temperature, and passed through a 0.45 filter prior to addition. Isobaric-labeled amino acids were dissolved in excess SILAC DMEM and passed through a 0.22 filter prior to addition.</p>
SILAC 'M' Medium:	<p>SILAC DMEM (Thermo) supplemented with 10% (v/v) dialysed FBS, 500 U.L⁻¹ penicillin, 500µg.L⁻¹ streptomycin, 250 µg.L⁻¹ 'light' L-Proline, 50 µg.L⁻¹ 'medium' L-Lysine di-hydrochloride (CK Isotopes), 50 µg.L⁻¹ 'medium' L-Arginine di-hydrochloride (CK Isotopes).</p>
SILAC 'H' Medium:	<p>SILAC DMEM (Thermo) supplemented with 10% (v/v) dialysed FBS, 500U.L⁻¹ penicillin, 500 µg.L⁻¹ streptomycin, 250 µg.L⁻¹ 'light' L-Proline, 50 µg.L⁻¹ 'heavy' L-Lysine di-hydrochloride (CK Isotopes), 50 µg.L⁻¹ 'heavy' L-Arginine di-hydrochloride (CK Isotopes).</p>
Dexamethasone solution:	25 mM Dexamethasone, in ethanol
NaPh buffer:	0.0076 M sodium di-hydrogen phosphate, 0.0324 M di-sodium hydrogen phosphate, in ddH ₂ O
Gradient 'Light' solution:	NaPh buffer, 15% (w/w) sodium tartrate, 30% (w/w) glycerol

Gradient 'Heavy' solution:	NaPh buffer, 35% (w/w) sodium tartrate
Sorbitol cushion:	PBS, 20% (w/v) sorbitol
MOPs SDS running buffer:	2.5 mM MOPS, 2.5 mM Tris Base, 0.005% (w/v) SDS, 0.05 mM EDTA
Carbonate transfer buffer:	3 mM sodium carbonate, 10 mM sodium bicarbonate, 10% (v/v) methanol
PBS-T:	PBS, 0.1% (v/v) Tween 20
Immunoblot block/stain buffer:	PBST, 5% (w/v) fat-free milk proteins (Marvel)

2.2 Cells and Viruses

2.2.1 Cells

Human cell lines used include: two variant hTERT-immortalised epithelial cell lines, RPE-1 (ATCC #: CRL-4000) and ARPE-19 (ATCC #: CRL-2302); and Human Foetal Foreskin Fibroblast (HFFF) cells, kindly supplied by Dr. Graham Farrar (CAMR, Salisbury, UK). Further cell lines used were derived from hTERT-immortalised HFFFs (HFFF-hTERTs) (McSharry, Jones et al. 2001). The HFFF-tet cell line was generated by transducing HFFF-hTERTs to constitutively express tetracycline (Stanton, Baluchova et al. 2010); these were used for the selective repression of viral genome regions engineered to contain tetracycline-binding operators. The HF-hCAR cell line was generated from HFFF-hTERT cells engineered to express the human Coxsackie Adenovirus Receptor (hCAR) (McSharry, Burgert et al. 2008); these were used for the expression of HCMV transgenes from RAdZ-5 vectors. The 293-TREx helper cell-line (Invitrogen/Life Technologies), based on HEK293 cells transduced to express the Adenovirus E1 genome region (Graham, Smiley et al. 1977), were used for the propagation of AdZ-5 vectors.

2.2.2 HCMV strains

HCMV strains used are detailed in Table 2.1. Several Merlin-BAC variants were readily available prior to the onset of this work. BAC-cloned versions of strains TR (TR-BAC), TB40/E (TB40-BAC4), and VR1814 (FIX-BAC) were kindly supplied by Jay Nelson (OHSU, Portland Oregon, USA), Christian Sinzger (University of Tübingen, Germany) and Gabi Hahn (Universitätsklinikum Carl Gustav Carus, Dresden, Germany), respectively. Each was received in the form of purified cell-free DNA. To provide an observable marker of infection, all BAC-

cloned genomes were engineered to express enhanced green fluorescent protein (eGFP) from the same transcriptional unit as the immediate early genes IE1/IE2 (UL122/UL123), as previously described (Stanton, Baluchova et al. 2010). Purified cell-free genomic DNA from the un-passaged HCMV strain 3301 was kindly donated by Andrew Davison.

Table 2.1. HCMV Strains investigated

Strain	BAC-cloned version	GenBank no.	References
Merlin ^a	Merlin-BAC	GU179001.1	(Stanton, Baluchova et al. 2010)
TR	TR-BAC	AC146906.1	(Murphy, Yu et al. 2003)
TB40E	TB40-BAC4	EF999921.1	(Sinzger, Hahn et al. 2008)
VR1814	FIX-BAC	AC146907.1	(Hahn, Khan et al. 2002)
3301	n/a	GQ466044.1	(Cunningham, Gatherer et al. 2010)

^a – several Merlin-BAC variants were used.

2.2.3 Adenovirus vectors

AdZ-5 vector constructs (Stanton, McSharry et al. 2008) contained within *E.coli* SW102 cells were readily available prior to the onset of this work. The AdZ-5 vectors lacked the essential E1 and E3 genome regions, and thus were replication deficient. In place of the deleted genome regions, the AdZ vector contained features that allowed the efficient insertion and expression of HCMV-derived transgenes. These included, in tandem: the HCMV major immediate early (MIE) promoter, a selectable cassette (see section 2.5) occupying the transgene insertion site, and an HCMV MIE polyadenylation sequence (polyA tail).

2.3 Tissue and virus culture

2.3.1 Human cell master stocks

Master stocks of each cell type were stored in cell banks at -196°C in liquid nitrogen. Cells were suspended in ‘freezing medium’ [90% (v/v) FBS, 10% (v/v) DMSO], and stored in Cryo.S 1mL cryovials (Grenier Bio-One). To ensure cells remained viable during freezing, cells were initially cooled to -70°C at a rate of -1°C min⁻¹ using a Cryo 1°C Freezing Container (Nalgene), before being transferred to the cell bank. All cell-lines were screened for Mycoplasma infection using the VenorGeM Mycoplasma PCR detection kit (Sigma-Aldrich), and deemed suitable for use only when found to be negative.

2.3.2 Tissue culture reagent preparation and manipulations

Tissue culture reagents were defrosted and warmed to 37°C in a water bath prior to use. All manipulations of human cells and viruses were conducted within a Class II biological safety cabinet (BSC), and performed with aseptic technique. Internal surfaces of the BSC and external surfaces of sample and reagent containers were de-contaminated using a solution of 70% (v/v) IMS before performing any manipulations.

2.3.3 Growth of human cells

Standard tissue culture conditions for the growth of human cells were: 37°C, in an environment of 5% CO₂. Unless otherwise stated, cells were propagated in standard tissue culture media [DMEM supplemented with 10% (v/v) FBS, 100 U.mL⁻¹ penicillin, and 100 µg.mL⁻¹ streptomycin]. Cells were fed with fresh media every 2-3 days. All cell types utilised were adherent and grown in monolayers attached to the surface of flat-bottomed tissue culture flasks and plates (Corning). 293TREx cells adhered less sturdily than other cell types and were grown in culture vessels of the 'Cell Bind' product line (Corning). When grown to confluency, cells were harvested for use in assays, or for continued growth by 'splitting'.

2.3.3.1 *Harvesting adherent cells*

Cell culture supernatants were removed before cells were washed in excess PBS. Cells were then dissociated from the culture flask surface by the addition of 0.05% trypsin/EDTA and incubation at 37°C for 3-5 mins. Flasks were then gently tapped to disrupt the cell monolayer, with detached cells aspirated in excess culture medium to quench residual trypsin, and generate a homogenous single-cell suspension.

2.3.3.2 *Counting harvested cells*

Harvested cells were counted using a glass haemocytometer. The counting chamber was loaded with 10 µL of homogenous single-cell suspensions, and the number of cells within grids that held 0.1 µL (1 x 10⁻⁴ mL) was counted. The averaged cell count from several grids was multiplied by a factor of 10,000 (10⁴) to determine the number of cells in 1 mL of the original suspension.

2.3.3.3 *Seeding cell culture vessels*

To maintain cell cultures at optimal sub-confluency by 'splitting', absolute cell counts were not determined and harvested cells were instead diluted by nominal factors. For use in infections and assays, cell culture flasks and plates were seeded with a known number of cells predicted to grow overnight to a suitable degree of confluency (40-80%). Media volumes, nominal dilution factors for splitting, and cell numbers used to seed all culture vessels are depicted in Table 2.1.

Table 2.2. Dilution factors and seed numbers for planned growth of cell

Cell Type			
	Fibroblasts (HFFF, HF-hCAR, HF-tet)	Epithelial cells (RPE-1, ARPE-19)	293TREx
Dilution factor for cell stock ‘splitting’			
Culture vessels ^a	1:3 – 1:5	1:6 – 1:8	1:8 – 1:10
Cell number for seeding cell culture flasks			
T150 (15-20 mL)	4 x 10 ⁶	4 x 10 ⁶	
T75 (10-15 mL)	3 x 10 ⁶	2 x 10 ⁶	
T25 (5-7 mL)	6 x 10 ⁵	7 x 10 ⁵	2 x 10 ⁶
Cell number for seeding cell culture cluster plates			
6-well (2 mL)	2.5 x 10 ⁵	3.0 x 10 ⁵	
12-well (1 mL)	1.25 x 10 ⁵	1.5 x 10 ⁵	5 x 10 ⁵
96-well (100 µL)	1.5 x 10 ⁴	2.0 x 10 ⁴	

^a – culture vessel and volume of media used for growth of cells therein.

2.3.4 Infections

Unless otherwise stated, all virus infections were performed under standard culture conditions. To maximise virus-cell interactions, inocula were prepared in the minimal volume sufficient to cover all cells. Inoculated cell cultures were initially incubated on a rocking platform (at 5-10 RPM) for two hours during which time virus was adsorbed onto cells. Following adsorption, the inoculum was removed and replaced with fresh culture medium. To limit virus dissemination to the cell-to-cell route only, infected cell cultures were incubated under a semi-solid overlay composed of a mixture (1:1) of 2X DMEM and Avicel; a mixture of microcrystalline cellulose and sodium carboxymethylcellulose (FMC Biopolymer), as previously described (Matrosovich, Matrosovich et al. 2006).

2.3.5 Monitoring infections

2.3.5.1 Microscopy and photography

Infections were monitored by phase-contrast microscopy for observable cytopathic effects (CPE), and by fluorescence microscopy for eGFP expression using DM:IL or DM:IRBE microscopes (Leica) (Figure 2.1.). Photographic imaging of infected cell cultures was performed using an ORCA-ER camera attached to the Leica DM:IRBE microscope. Image analysis was performed using Image J 64 and OpenLab 3 software (Perkin Elmer), with measurements of plaque dimensions made using tools within each software package following calibration using a 100-micron graticule. Estimation of the sizes of approximately elliptical plaques formed in HCMV-infected cell cultures, in μm^2 , was performed using the following calculation:

$$\text{Plaque length } (\mu\text{m}) \times \text{Plaque width } (\mu\text{m}) \times 0.8$$

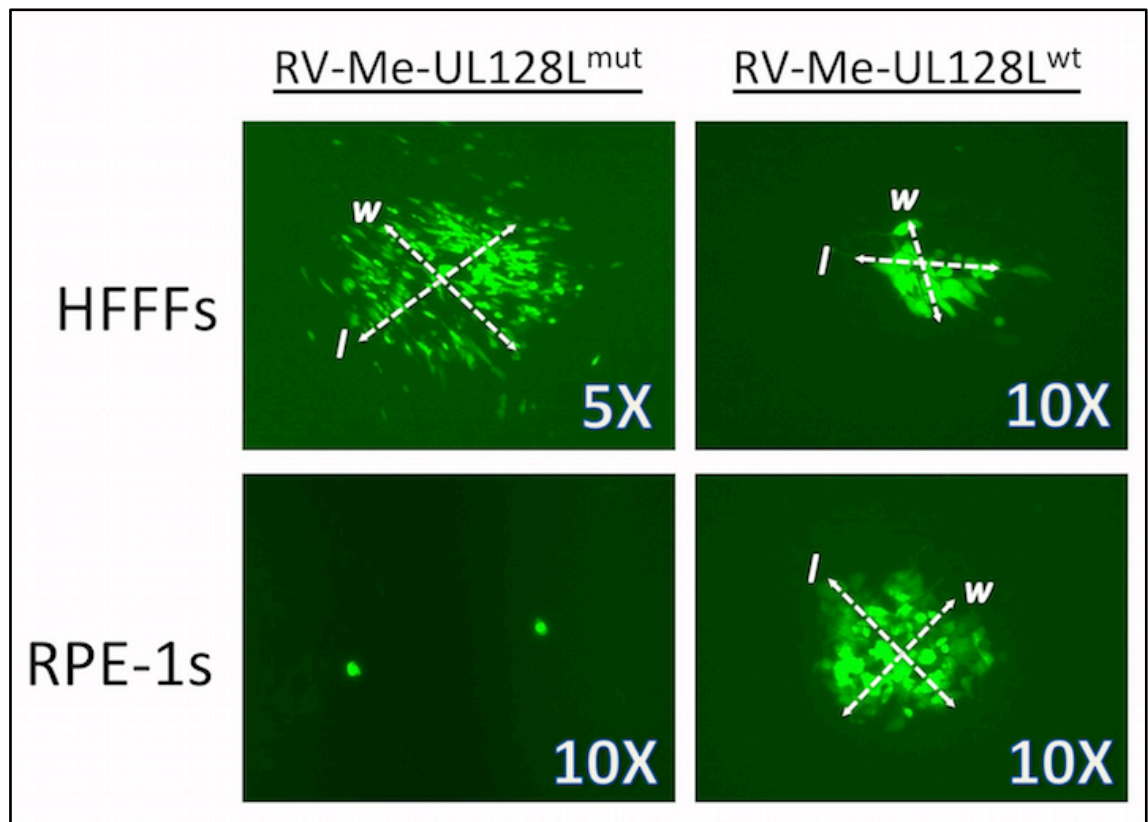


Figure 2.1. HFFF and RPE-1 plaque imaging and size measurement. Representative HFFF and RPE-1 plaques formed Merlin BAC-derived viruses either lacking intact UL128L, or containing wildtype UL128L. Magnification of images is shown. *w* - indicates plaque width; *l* - indicates plaque length

2.3.5.2 Flow cytometry

Infections in epithelial cell cultures were monitored by flow-cytometry using an Accuri C6 (BD Bioscience) (Figure 2.2.). Samples were prepared as follows: total cells (infected and uninfected) were harvested as previously described, then collected by centrifuging the resultant suspensions at 470 in x g for 3 mins at room temperature. Following removal of the supernatant, cells were washed by re-suspending the pellet in excess PBS and repeat of the centrifuging. Washed cells were re-suspended in PBS and then fixed in 2% (w/v) PFA for 15 mins at 4°C. Samples were analysed on the Accuri C6 flow cytometer with 20,000 events recorded under parameters defined using the CFlow software. Gating of live cells was based on two criteria: size and granularity. Non-infected and infected cells were distinguished by fluorescence signal, indicative of eGFP expression. To account for auto-fluorescence, and also any fluorescence from the originally transfected cells and not disseminated infectivity, the proportion of eGFP+ epithelial cells in cultures transfected with Merlin virus lacking intact UL128L (and unable to disseminate in epithelial cells) was used as a threshold (~ 1%) to which all other infections were compared.

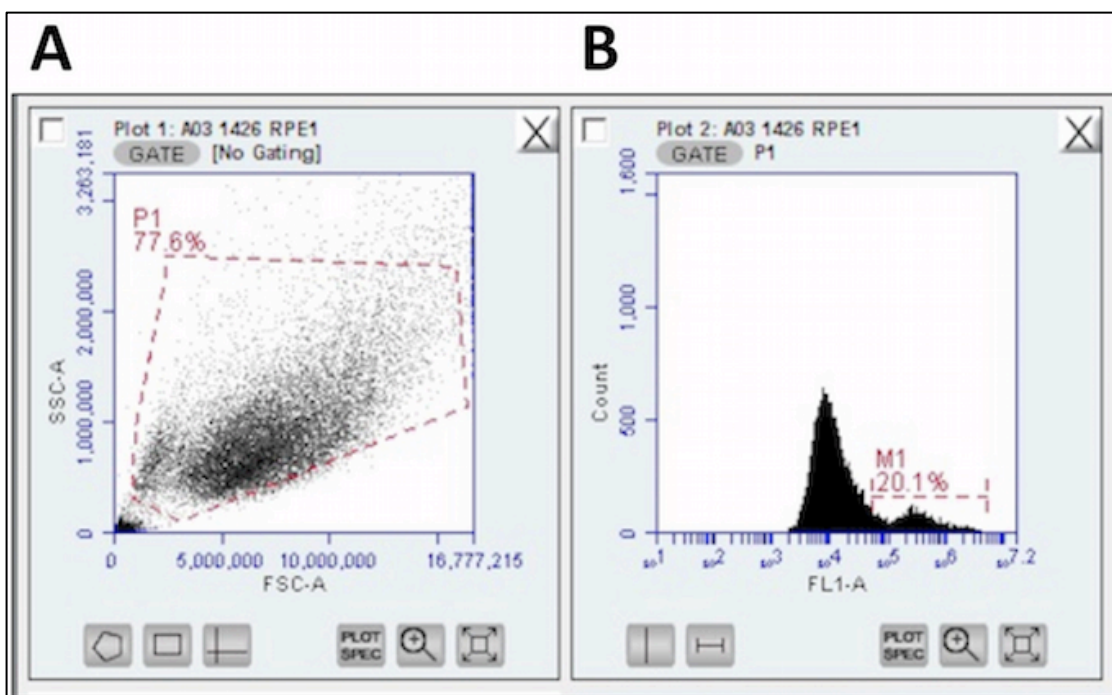


Figure 2.2. Monitoring infection of epithelial cell culture by flow cytometry. Example of data analysis. (A) The recorded events considered to represent live cells were gated for (red dash line), based on granularity (SSC-A), and size (FSC-A). (B) eGFP expression was detected in channel FL1-A.

2.4 Molecular Biology Techniques

2.4.1 *In silico* analysis of nucleic acid and amino acid sequences

Nucleic acid and amino acid sequence comparisons, recombineering experiment design, primers and oligonucleotide design, and sequencing data analysis, was performed using CLC MAIN 6 software (CLC Bio). Primers and oligonucleotide attributes were verified using Oligo Explorer 1.4 Beta (GeneLink) and Oligo Analyser (GeneLink) software packages.

2.4.2 Molecular biology reagent storage and preparation

Unless otherwise stated, all thermo-labile reagents and reaction products were stored at -20°C. To increase thermo-stability, nucleic acids were dissolved and stored in TE buffer, unless otherwise stated. All thermo-labile reagents were defrosted gradually and stored on ice during use, whilst thermo-stable reagents were defrosted in a water bath and stored at room temperature.

2.4.3 Nucleic acid modification and amplification

Modification and amplification of nucleic acids was performed with various commercial kits (indicated below), with all reactions performed within thin-walled DNase and RNase-free tubes (ELKay), under thermal cycling conditions as specified for each kit using a T3000 Thermocycler (Biometra). When not included in a given kit, dNTPs (New England BioLabs) and PCR-grade ddH₂O supplied by the user were used. PCR reaction mixtures were supplemented with 3% (v/v) DMSO to limit template secondary structure formation and increase amplification efficiency.

2.4.4 Primer Design

Where possible, primers were designed to have a melting temperature (T_m) 60-65°C, as well as display limited self/pair annealing and secondary structure formation. All primers and oligonucleotides were acquired from Sigma-Aldrich as Reverse Phase Cartridge purified lyophilised DNA. Primers were reconstituted by the addition of an appropriate volume of ddH₂O to achieve a final 100 µM concentration.

2.4.5 DNA gel electrophoresis

DNA fragments were resolved and isolated from mixed samples by agarose gel electrophoresis. Gels were formed by dissolving agarose in tris-acetate-EDTA (TAE) buffer, and were supplemented with ethidium bromide to stain captured DNA. Gels were cast using trays and accompanying well-forming combs, and allowed to set prior to submersion in excess TAE running buffer. Samples were mixed with 0.2 volumes of 6X DNA loading buffer, then loaded

into individual wells. In each electrophoresis performed, a sample of DNA Smart ladder (Eurogentec) was loaded to a separate well and electrophoresed in parallel, allowing estimation of the sizes of resolved DNA fragments by extrapolation to the migration of fragments of known size. Resolved amplicons were visualised by UV light activated ethidium bromide fluorescence using an Autochemi Bioimaging system (UVP) and LabWorks software (Perkin Elmer), or a bench-top transilluminator (Spectroline).

2.4.6 Purification of PCR-amplified DNA

DNA was purified from PCR reactions using the Illustra GFX DNA and Gel Band purification Kit (GE Healthcare) according to manufacturer's instructions. In brief, nucleic acids and other macromolecules were first denatured in the presence of a chaotropic buffer (supplied), and then non-specifically captured onto silica membranes by centrifugation. Residual salts and other impurities were cleared from captured DNA by the addition of wash buffer (supplied) supplemented with ethanol, then further centrifugation. Captured and purified DNA was then eluted and recovered in an appropriate buffer (supplied) during one further spin. Where target amplicons were required to be separated from other non-specific amplicons, they were first isolated by agarose gel electrophoresis and then excised directly from gels. Excised gel segments were then transferred to sterile 1mL sample tubes, together with an appropriate volume of capture buffer, and incubated at 60°C for ~5 mins to melt the gel and release DNA. Alternatively, where purity was of little consequences for downstream applications, DNA was purified directly from PCR reaction mixtures diluted in capture buffer.

2.4.7 Determination of DNA concentration

The concentration of DNA samples was determined using a Nano-drop ND-1000 spectrophotometer (ThermoScientific). The spectrophotometer was initialised with ddH₂O according to standard operating procedures, then the loading pedestal was wiped using fibre-less paper. The instrument was then blanked by loading 2 µL of the relevant solvent buffer and performance of a measurement. The loading pedestal was again wiped before loading 2 µL of analyte sample to be measured.

2.5 Recombineering of BAC-cloned HCMV and Adenovirus genomes

2.5.1 Phage Lambda Red-mediated Homologous Recombination

Recombination mediated genetic engineering ('Recombineering') (Warming, Costantino et al. 2005) of BAC-cloned HCMV genomes and for the insertion of transgenes into the AdZ-5 vector was performed as previously described (Stanton, McSharry et al. 2008, Stanton, Baluchova et al. 2010). In brief, recombineering was performed in *E. coli* SW102 cells transduced to express phage lambda red recombinant proteins that facilitate recombination between homologous

nucleotide sequences as short as 50 bp. Cells containing the ‘recipient’ BAC-cloned genomes to be modified were subjected to two successive rounds of transformation. In the first round, a DNA construct that encoded selectable characteristics (*amp^r/lacZ* cassette) was inserted into the region to be modified. In the second round of transformation, ‘donor’ sequences were inserted in place of the selectable cassette. Details of each recombineering experiment, coordinates to the regions to be modified in the ‘recipient’ genome and sequences/regions to be transferred from the ‘donor’ genome, as well as all primers and oligonucleotide used, are detailed in Appendix I.

2.5.2 Design and preparation of DNA sequences for transformation

To facilitate incorporation by homologous recombination, sequences to be inserted were engineered to contain sequence extensions with homology to regions flanking the intended site of insertion. The selectable *amp^r/lacZ* cassette and ‘donor’ sequences were PCR amplified and gel-purified prior to use, with primer pairs used designed to contain ~20-nucleotides at their 3’ end identical to sequences at the termini of the sequence to be inserted (cassette or ‘donor’ sequences), and ~80-nucleotide at their 5’ end identical to sequences flanking the insertion site in the ‘recipient’ genome. To transfer single ‘donor’ nucleotides, commercially produced oligonucleotides of ~100-nucleotide length were used in the second transformation step.

2.5.2.1 *amp^r/lacZ* cassette amplification

The 4.217 Kb *amp^r/lacZ* cassette was PCR-amplified using the Expand HiFi PCR system (Roche), which employs the robust 5’-3’ reading *Taq* DNA polymerase in combination with a 3’-5’ proofreading exonuclease for increased amplification fidelity. Amplification was performed by a 2-step PCR reaction, under the following thermal-cycling conditions:

- 95°C - 2 mins;
- 10 cycles: 95°C - 30 s, 55°C - 30 s and 68°C - 4.5 mins;
- 25 cycles: 95°C - 30 s, 55°C - 30 s and 68°C - 4.5 mins + 20 s per additional cycle;
- 68°C - 15 mins.

2.5.2.2 ‘Donor’ sequence amplification

‘Donor’ sequences were PCR-amplified using the Phusion High Fidelity kit (New England BioLabs) that employs an enzyme that is less robust, though amplifies with greater fidelity, compared to *Taq* DNA polymerase. Amplification was performed as following:

- 98°C - 1 mins;
- 35 cycles: 98°C - 30 s, 55°C - 30 s and 72°C - 15-30 s per Kb;
- 72°C - 10 mins

2.5.3 Preparation and transformation of electro-competent *E.coli* SW102s

Phage-lambda red homologous recombineering protein expression is temperature-controlled, occurring between 37°C and 45°C (optimal 42°C). To limit recombineering protein expression prior to transformation, *E.coli* SW102s were grown at 32°C in a shaking incubator (at 200 RPM). All bacterial cultures were grown in media based on low-salt LB broth supplemented with the anti-biotic chloramphenicol (12.5 µg/mL); this enriched for cells that contained the BAC vector encoding a chloramphenicol resistance gene. To generate an abundance of viable electro-competent cells, cultures were grown to upper cellular concentration limits whilst remaining in 'log-phase' of growth. The cellular concentration of cultures was monitored during growth by spectrophotometry using an UltraSpec 3000 (Pharmacia Biotech), with bacterial culture growth halted when cultures reached a density of 0.6 OD units as detected at 600 *nm*. Phage lambda red homologous recombination protein expression was induced by incubating cultures at 42°C for 15 mins, prior to incubation on ice for a further 15 mins to arrest any further growth and ensure bacterial cells remained transformation competent. Cells were then washed by to clear residual salts by three successive rounds of centrifuging at 3345 x *g* at 0°C for 5 mins and re-suspension in ice-cold ddH₂O. Cells in the final washed pellet were then re-suspended in ~400-500 µL ice-cold ddH₂O, with 25 µL aliquots transferred to pre-cooled 0.5 mL sample tubes and mixed with 4 µL of DNA to be transfected. Cell-DNA mixes were incubated on ice for 5 mins prior to being transferred into a chilled Geneflow 2mm electroporation cuvette (Cell Projects), then electroporated using a Micropulser electroporation unit (BioRad) set on program EC2 to deliver a voltage of 2.5 KV. Following delivery of the selectable *amp^r/sacB/lacZ* cassette, electroporated cells were diluted in 1 mL LB broth, then incubated under culture conditions for 1 hr to recover and allow induction of selectable transgene expression. Following replacement of the selectable cassette in the second transformation, extra time was afforded for the recovery of electroporated cells to ensure that the SacB protein was cleared; transformed cells were recovered in 5 mL LB broth during incubation for 4 hrs.

2.5.4 Screening for transformants

Following each transformation step, electroporated cells were spread onto LB agar plates supplemented with agents to select for, or against, *amp^r/sacB/lacZ* cassette encoded functions, and grown at 32°C overnight.

2.5.4.1 Positive selection for *amp^r/sacB/lacZ* cassette following first transformation

LB agar was supplemented with ampicillin (50 ng/mL) to select for *amp^r* expression, as well as X-gal (80 mg/mL) and IPTG (2 mM) to select for *lacZ* expression as indicated by the development of blue colonies.

2.5.4.2 Negative selection for *amp^r/lacZ* cassette following second transformation

LB agar was supplemented with 5% (w/v) sucrose, which is metabolised to a toxic product in the presence of SacB protein, as well as X-gal (80 mg/mL) and IPTG (2 mM) to identify transformants that lack *lacZ* expression by development of white colonies.

2.5.5 Analysis of transformant genome integrity and coding potential

BAC DNA was extracted and purified from screened transformants (see 2.5.6.1.). The integrity of 'recipient' genomes in selected recombinants was assessed by restriction digest analysis to indicate whether recombination had occurred in non-targeted viral genome regions. To ensure insertion of DNA into the region to be modified occurred with high fidelity, the nucleotide sequence at the site of cassette or 'donor' sequence insertion was analysed.

2.5.5.1 Restriction digest

Restriction digest reaction mixtures in a final volume of 10 µL were composed of purified BAC-DNA mixed with HindIII (10 Units) and associated buffer at working concentration (NEB). Digests were performed by incubating reaction mixtures at 37°C for 4 hours, with generated fragments resolved over 0.7% (w/v) agarose gels.

2.5.5.2 Sequence analysis

Where possible, sequencing reactions were performed using the BigDye Terminator v3.1 cycle sequencing kit (Invitrogen/Life Technologies). Reaction mixtures in a final volume of 10 µL were composed of BAC DNA (0.5-1.0 µg), primers (3.2 mM) and pre-formed Big Dye mix. Reactions were performed under the following thermal cycling conditions:

- 95°C for 5 mins;
- 100 cycles: 95°C - 30 s; 55°C - 10 s; - 60°C - 4 mins.

Reaction products were cleared of excess dye-terminators and other contaminants using Performa DTR gel filtration columns (Edge Bio), according to manufacturer's instructions. Initially, columns were centrifuged to dry the gel, with BigDye reaction mixtures loaded directly into the gel before repeated centrifuging. The recovered eluate was then analysed using an ABI sequencer (Central Biotechnology Services, School of Medicine, Cardiff University). Following insertion of the cassette, regions corresponding to the ~80 bp engineered arms of homology were analysed using primers that bind at the 3' termini of the *amp^r/lacZ* cassette and read outwards into the regions flanking the site of insertion: ACG GAA ATG TTA TAC TCA TAC TCT and CAC CGT TTT CAT CTG TGC ATA T. Following insertion of 'donor' sequences, regions corresponding to the engineered arms of homology, as well as the transferred 'donor' sequence itself, were analysed using primers specific for the genome region modified (Table 2.2.).

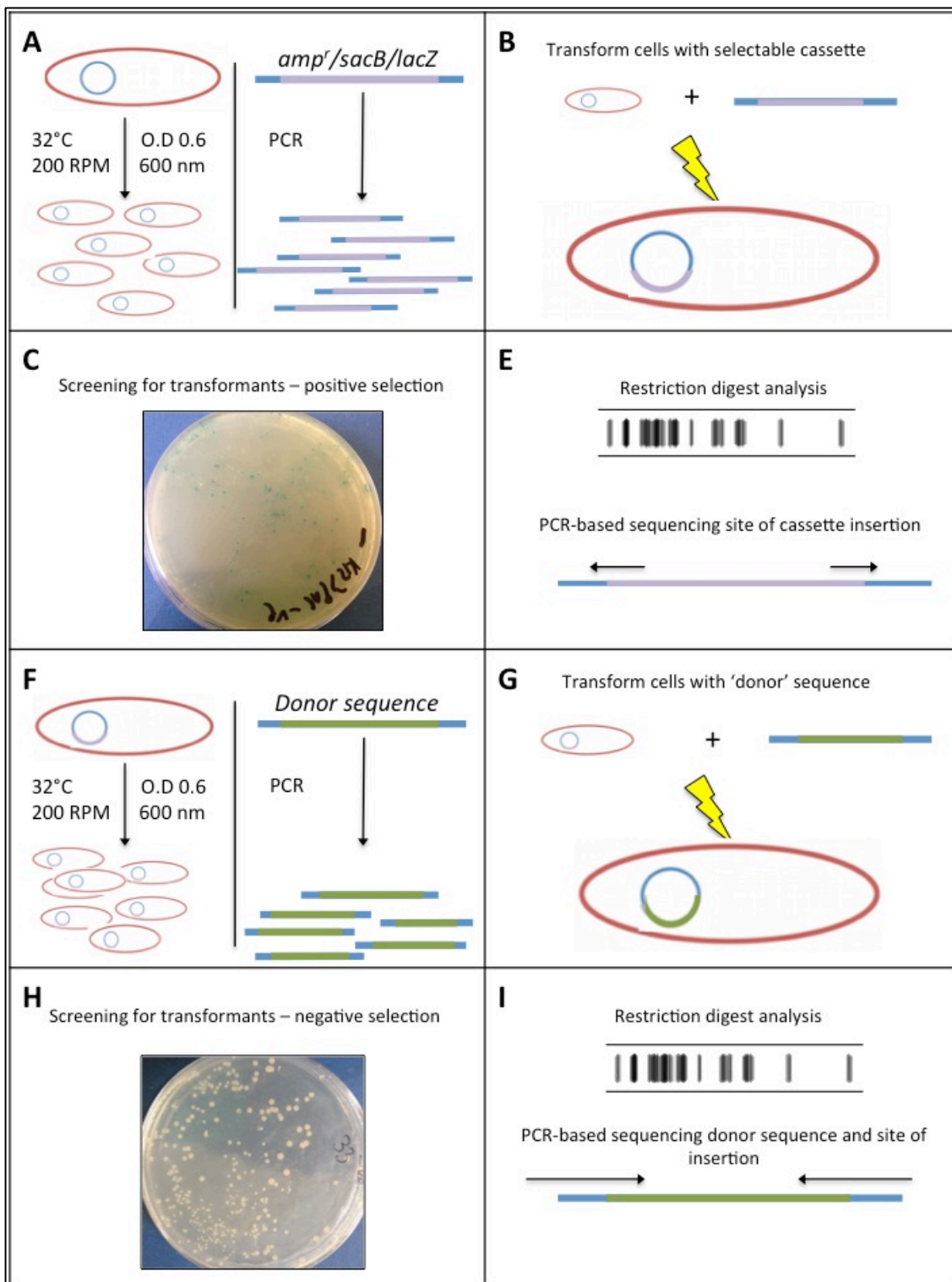


Figure 2.3. BAC-cloned viral genome recombineering - overview. (A) *E. coli* SW102 cells were grown to optimum density and the selectable cassette was PCR-amplified; (B) cells were transformed with the selectable cassette by electroporation; (C) recovered transformant were screened by growth on LB agar containing ampicillin, X-gal, and IPTG for successfully transformed (blue) colonies; (D) recombinant BACs were assessed by restriction digest analysis, and PCR-based sequencing at site of cassette insertion; (E – I) the same process as in A-D was followed for the replacement of the inserted selectable cassette with 'donor' sequences, with screening by growth on LB agar supplement with sucrose, X-gal, and IPTG for successfully transformed (white) colonies, and PCR-based sequencing to analyse the inserted sequence and also site of insertion.

2.5.6 Maintenance of transformed *E.coli* SW102s as glycerol stocks

E.coli SW102s containing BAC-DNA constructs were stored as plasmid library stocks at -70°C. Plasmid library stocks were prepared from bacterial cultures grown over night, with 425 µL of bacterial cultures thoroughly mixed with 75 µL of the cryo-preservent glycerol (Fisher), then transferred to suitable cryovials before freezing.

2.5.7 Extraction and purification of BAC DNA from *E.coli* SW102s

Extraction and purification of BAC-DNA from *E.coli* SW102 cells was performed using various kits, each based on the release and precipitation of total DNA by alkaline lysis, followed by the selective re-naturing of BAC DNA by the addition of acetate in lysis neutralisation buffers (Bimboim and Doly 1979). *E.coli* SW102s grown in LB broth cultures over night were harvested by centrifuging under conditions specified by each commercial kit.

2.5.7.1 Small-scale preparation of BAC-DNA

Small-scale preparations of purified BAC-DNA were prepared using a Mini-prep Kit (QIAGEN). Total DNA was extracted by alkaline lysis as previously described. Lysates were cleared of precipitated cellular debris by centrifuging at 16,000 x g for 10 mins at room temperature. The resulting supernatant was transferred to a fresh tube, with BAC-DNA precipitated and pelleted by the addition of room temperature isopropanol and centrifuging at 16,000 x g for 10 mins at 4°C. Recovered BAC-DNA was retained and washed by the addition of 70% (v/v) ethanol and repeated centrifuging. Following drying, the DNA pellet was re-suspended in an appropriate volume of TE buffer.

2.5.7.2 Large-scale preparation of transfection grade BAC-DNA

Large-scale working stocks of transfection-grade cell-free BAC-DNA constructs were prepared by Maxi-prep, using a Nucleobond BAC 100 plasmid DNA purification kit (Machery Nagel), according to manufacturer's instructions for the purification of low-copy number plasmids. In this method, silica-based anion exchange resin technology utilising latex beads of exquisite size uniformity allowed nucleic acids of different chemistry (e.g. RNA, DNA), structure (single or double-stranded), and size, to be resolved within strictly defined elution profiles. In brief, total DNA was extracted from *E.coli* SW102 cells by alkaline lysis, and the resultant flocculate of precipitated cellular debris was then cleared by passage of the lysate across a fine filter (supplied). The recovered flow-through was then passed across a BAC 100 column by gravity-flow filtration. Columns containing captured BAC DNA were repeatedly washed by the addition of wash buffer (supplied), and all flow-through from these steps was discarded. Captured BAC DNA was recovered by addition of pre-warmed (50°C) elution buffer (supplied) to the column, with the flow-through from this step recovered. Eluted DNA was precipitated by the addition of room temperature isopropanol and pelleted by centrifuging at 15,000 x g for 30

mins at 4°C. The resulting DNA pellet was retained and washed by the addition of room temperature 70% (v/v) ethanol, and further centrifugation at 15,000 x g for 10 mins at room temperature. The DNA pellet was retained and allowed to air-dry, before being re-suspended in appropriate TAE buffer and stored 4°C.

2.6 Reconstitution, growth and preparation of BAC-derived virus

2.6.1 Reconstitution of virus from BAC-cloned HCMV genomes

HCMV viruses were reconstituted from BACs and grown in various fibroblast and epithelial cell types. Productive infection in each of these cell types was initiated by transfection of cells with BAC DNA. All grown viruses and infected cell-culture supernatant samples were stored at -70°C. To minimise the likelihood for the selection of mutants during the production of virus stock in cell-culture, no virus was plaque-purified.

2.6.1.1 *Transfections of fibroblasts*

HFFF and HF-tet cells were transfected by electroporation using a Nucleofector II (AMAXA), and a basic nucleofector kit (LONZA). In brief, 10⁶ harvested fibroblast cells were collected by centrifuging homogenised cellular suspensions at 75 x g for 10 mins at room temperature. The pelleted cells were then re-suspended in 100 µL of transfection reagent (supplied), and mixed with 2 µg of purified BAC-DNA. The entire transfection reaction mixture was then transferred to cuvettes (supplied) and electroporated using Nucleofector II program T-16. Electroporated cells were diluted in pre-warmed culture medium and transferred to suitable cell culture vessels, then allowed to adhere and recover during overnight incubation. On occasions when cultures were <70% confluent after overnight recovery, additional non-transfected cells were added.

2.6.1.2 *Transfection of epithelial cells*

RPE-1 and ARPE-19 cells were transfected by chemical transfection using the Effectene kit (QIAGEN), according to instructions for stable transfection of adherent cells. Transfection mixtures were generated in two steps. Firstly, 0.4 µg purified BAC DNA was condensed by the addition of 3.2 µL enhancer reagent and buffer (supplied) to a final volume of 100 µL, with this mixture incubated at room temperature for 2-5 mins. Secondly, transfection complexes were formed by the addition of the 10 µL effectene reagent (supplied), and further incubation at room temperature for 5-10 mins. The transfection complex mixture was diluted by the addition of pre-warmed culture medium, and then added to epithelial cell cultures grown to 40-80% confluency in the well of a 6-well plate.

2.6.2 Growth of HCMV stocks

Productive infections were initiated in cell cultures grown to near confluency (>90%). Cells were either inoculated by the addition of existing infected cells, initially derived from transfection, or by the addition of cell-free virus at an approximate MOI of 0.1. To facilitate the dissemination of slow growing strains that displayed a tendency towards focal spread and formation of discrete plaques, infected cell cultures were disrupted as in 'splitting', with the entire resultant cellular suspension returned to the flask to allow the monolayer to re-form with greater distribution of infectivity. When 70-100% of all cells were infected, infectious supernatant containing cell-free virus were collected and replaced with fresh culture medium. Infectious supernatants were collected every two days until all cells in infected cell cultures were lysed, and stored at -70°C. No virus was plaque purified.

2.6.3 Preparation of infected cell culture fractions and sub-fractions

Infected cell-culture supernatant was removed and centrifuged at 470 x g for 5 mins at room temperature, with the cleared supernatant recovered to form the cell-released virus (CRV) fraction. The pelleted cellular materials were collected and mixed with infected cells recovered from culture vessels by trypsinisation. The total infected cell material was re-suspended in fresh standard culture medium, forming the cell-associated virus (CAV) fraction.

2.6.3.1 *Liberation of CAV infectivity, preparation of Cell-extract and Cell-debris sub-fractions*

Infected cells were lysed using a VibraCell sonicator (Sonics), with infected cells subjected to 30-second pulses of sonication at 100% amplitude. To prevent overheating, samples were incubated on ice during sonication. To generate cellular debris and cell-extract sub-fractions, lysates were centrifuged at 345 x g for 10 mins at room temperature. The cleared supernatant was recovered as the cell-extract CAV sub-fraction, and the pellet was re-suspended in fresh standard culture medium to form the Cell-debris CAV sub-fraction.

2.6.4 Preparation of concentrated HCMV working stocks

Working stocks of virus were prepared by concentrating cell-free virus contained within infected cell culture supernatant 'bleeds'. Bleeds were initially defrosted and warmed to 37°C in a water bath, then cleared of high-mass cellular debris by centrifugation at 470 x g for 5 mins at room temperature. Following concentration by various methods, virus was aliquoted to suitable cryo-preserved tubes and stored at -70°C.

2.6.4.1 *Concentration by high-speed centrifugation*

Cleared supernatants were pooled and centrifuged at 30,000 x g for 2 hr at 20°C. The pelleted infectious virions were then re-suspended in a reduced volume of fresh culture medium. To

ensure dissociation of any virion aggregates and generate a homogenous single virion suspension, the re-suspended virus was subjected to limited passes (3-5) through a 19 G needle.

2.6.4.2 Concentration by ultra-filtration

Virus samples were concentrated by the gradual removal of water using ultrafiltration technology. In these methods, virus samples were applied to a polyethersulfone (PES) membrane with pores that excluded the passage of particles larger than 1 MDa.

2.6.4.2.1 Dialysis ultra-filtration

VivaFlow50 ultra-filtration cartridges (Sartorius) were used according to manufacturer's instruction. In brief, a reservoir of infectious supernatant was stored in a suitable open-necked vessel, drawn along an inlet tube connected to the cartridge, circulated through the cartridge compartments lined with the PES filtration membrane, then returned to the reservoir through an efflux tube. Circulation of supernatant through the efflux tube was restricted by a flow-restrictor, driving the negative-pressure dialysis of excess media across the PES membrane. Circulation of the supernatant was driven by use of a peristaltic pump and occurred at a flow rate of 400 mL.min⁻¹. When the final volume was less than 20 mL, the circulation of supernatant was stopped. The inlet tube was then transferred to a reservoir containing ~10 mL ddH₂O which was circulated throughout the cassette and collected to increase recovery of concentrated virus.

2.6.4.2.2 Centrifugal ultra-filtration

Cleared supernatants were loaded into the upper chamber of a VivaSpin 20 column (Sartorius), with passage of excess water and solutes through the filtration membrane achieved by centrifugation at 3000 x g at 25°C.

2.6.5 HCMV infectivity quantification

The infectivity in virus samples (not plaque purified) was enumerated by minimal dilution assay. Samples were serially diluted to generate a 10-fold dilution series, with cells inoculated with 100 µL from successive dilutions predicated to contain 1-1000 infectious units. Following incubation detectable markers of infection were counted, and titres of original samples, expressed as total PFU in 1 mL (PFU.mL⁻¹), were determined by the following calculation:

$$\text{plaques per 100 } \mu\text{L inoculum} \times 10 \times \text{inoculum dilution factor}$$

2.6.5.1 Plaque titration

Cells were grown to 40-80% confluency in either 6-well, 12-well, or 24-well culture plates. On performance of infections, the cell culture media was removed and immediately replaced with fresh culture medium (500 µL in 6 well plates; 250 µL in 12-well plates; and 150 µL in 24 well plates. Inocula (100 µL) were then added to wells, with infections performed as described. To ensure only plaques that were derived from a single infection event were formed, cells were

incubated under a semi-solid over-layer. Following sufficient incubation (2-3 weeks, depending on cell type), the semi-solid overlay was removed before the infected cultures were washed twice in excess PBS and plaques were counted.

2.6.5.2 *IE1 immuno-fluorescence titration*

Cells grown in clear-bottomed 96-well plates were infected under standard conditions. On infection, culture media was removed from wells, and 100 μ L inocula applied directly to cells. Following 12-24 hrs incubation, cells were prepared for staining of nuclear-localised pIE1. Infected cell cultures were fixed in 4% (w/v) PFA, permeabilised using immunofluorescence buffer (IF buffer) supplemented with 0.1% (w/v) Triton X100, then blocked in excess IF buffer. All treatments of cells were performed for 10 mins at room temperature, and followed by three washes in excess PBS. The nuclei of infected cells were stained with monoclonal mouse anti-IE1 antibody (ThermoScientific, MA1-7596) diluted (1:1000) in IF buffer, during incubation for 1 hour at 37°C on a rocking platform. Un-bound primary antibody was removed by washing in excess PBS, and AlexaFluor 594 conjugated goat anti-mouse antibody F(ab')₂ fragment (Invitrogen, A11020) diluted (1:500) in IF buffer was applied, with cells further incubated for 30 mins at 37°C on a rocking platform. Stained cells were then washed in excess PBS, and fluorescently stained nuclei were counted.

2.6.6 Growth and preparation of Adenovirus vector stocks

Adenovirus vectors were grown in 293-TREx helper cells, with infections initiated by transfection using the Effectene chemical transfection kit as described above. Following the dissemination of infection throughout the entire 293TREx cell culture, cells were harvested and cleared of supernatant by centrifuging at 470 x g for 3 mins at room temperature. Pelleted cells were then washed by re-suspension in excess PBS and repeated centrifuging. Cells in the resulting pellet were again re-suspended in PBS, then mixed with an equal volume of the detergent tetrachloroethylene, before being shaken vigorously to lyse infected cells and liberate cell-associated Adenovirus virions. The resulting emulsion was centrifuged at 836 x g for 20 mins at room temperature, generating three fractions: a low density fraction from which extracted Adenovirus virions in PBS were recovered (top layer); a fraction of cellular debris (middle layer); and a high-density fraction of tetrachloroethylene (bottom layer). Adenoviruses were recovered from the top layer and stored in cryo-preserved tubes at -70°C.

2.6.7 Titration of Adenovirus stocks

The infectivity of RAdZ stocks was quantified by titration on 293TREx cells. To prevent disturbance of the loosely adhered 293TREx cells, the cell culture supernatant was not removed prior to inoculation, with inocula added directly to the cultures. No preliminary incubation on the rocker platform was performed. At 48 hours post infection, cell culture supernatant was

removed and cells were allowed to air-dry, before being fixed and permeabilised in a cocktail of ice-cold acetone and methanol (1:1) at -20°C for 10 mins. Cells were then washed in excess blocking buffer composed of PBS supplemented with 0.2% (w/v) BSA, before infected cells were stained with polyclonal goat anti-adenovirus primary antibody (Abcam, AB1056) diluted (1:5,000) in blocking buffer, during incubation at 37°C for 1 hour on a rocking platform. Unbound primary antibody was removed by washing in excess blocking buffer, before HRP-conjugated polyclonal donkey anti-goat antibody (SantaCruz, 2056) diluted (1:1000) in blocking buffer was added and cells, which were further incubated. Cells were again washed in excess blocking buffer, before the addition of a colorigenic HRP substrate. DAB substrate (Vector Labs, SK100) was used according to manufacturer's instructions, and prepared by the addition of DAB substrate (3, 3'-diaminobenzidine), as well as associated buffers and other components (all supplied) to ddH₂O. Substrate was applied to cells that were then incubated for sufficient time for signal development before counting stained cells.

2.7 HCMV-derived nucleic acid extraction

2.7.1 Extraction of viral nucleic acids from infected cells or cell-free virions

Viral nucleic acids were extracted from cells and cell-free virus using various kits of the QIAgen product line, according to manufacturer's instructions. All protocols were based on a combination of silica resin anionic binding and microspin technology. In brief, cells and virions were disrupted by protease enzymes and lysis buffers to release nucleic acids; cellular and virion derived contaminants in lysates were then neutralised by various buffers, or captured within spin columns; then nucleic acids were selectively captured on silica membranes by centrifugation under binding conditions optimised by the inclusion of salts and alcohols; captured nucleic acids were washed by the addition of buffers and/or alcohols and further centrifugation, then were recovered by in elution buffers.

2.7.1.1 *Extraction of viral RNA from infected cells*

Total RNA, including virus derived RNA, was extracted from infected cell cultures using an RNeasy Plus Universal Kit. Infected cells were harvested and disrupted in lysis buffer, with the resulting lysate homogenised by centrifuging through a QIAshredder spin column. DNA was removed from the homogenised lysate by centrifuging through a gDNA Eliminator spin column. Cleared lysate was mixed with 70% (v/v) ethanol, with RNA then captured onto the membrane of an RNeasy spin column, washed and eluted in RNase-free water.

2.7.1.2 *Extraction of viral DNA from infected cells*

Viral DNA was extracted from infected cells using the DNeasy Blood and Tissue Kit. Harvested infected cells were cleared of cell-culture supernatant and re-suspended in PBS, before being disrupted by the addition of protease and accompanying lysis buffers and

incubation at 56°C for 10 mins. The lysate was mixed with 100% ethanol and loaded into a DNeasy mini prep column, with DNA captured onto the membrane then subjected to two rounds of washing. Finally, DNA was recovered in an appropriate volume of elution buffer.

2.7.1.3 *Extraction of viral DNA from cell-free virions*

DNA was extracted from cell-free virions using a MinElute Virus spin Kit. Initially, cell-free virions were disrupted by the activity of protease and accompanying lysis buffers, during incubation at 56°C for 15 mins. The resultant lysate was mixed with 100% ethanol and incubated at room temperature for 5 mins, before viral DNA was captured by centrifugation through a spin column. Bound DNA was subjected to three washes, firstly by the addition of wash buffers, then by addition of 100% ethanol in the final wash. Washed membranes were allowed to air dry, prior to recovery of viral DNA in elution buffer.

2.8 Nucleic acid sequence analysis

Details of all primers used for the PCR, RT-PCR, and sequencing reactions, are in Table 2.2.

2.8.1 PCR amplification of sequences from HCMV genome regions

2.8.1.1 *Amplification from virion incorporated genomic DNA*

When HCMV genome sequences were PCR-amplified directly from virion incorporated genomic DNA, the Advantage II PCR system (Clontech) employing the robust Titanium *Taq* dDdP with hot-start PCR technology was used. Cycling conditions included:

- 95°C - 5 mins (to disrupt virions and release virus genomic DNA)
- 10 cycles: 95°C - 30 s, 68°C - 60-90 s per Kb
- 68°C - 4.5 mins, time depending on template size
- 4°C - indefinite

2.8.1.2 *Amplification from extracted and purified viral genomic DNA*

Amplification of genome regions from extracted and purified HCMV genomic DNA was performed using the Expand HiFi PCR kit (Roche) (described above). Amplification was performed by one-step PCR under the following thermal cycle conditions:

- 94°C - 2 mins (to disrupt virions and release virus genomic DNA);
- 35 cycles: 94°C - 30 s; 65°C - 30 s; 72°C 1 min per Kb
- 72°C, time depending on template size
- 4°C – indefinite

Table 2.3. Primers used for PCR amplification of genome regions, RT-PCR of transcripts and PCR amplification of cDNAs, and sequencing reactions.

Genome Region	Primer	Merlin	TR	TB40-BAC4	FIX	3301
UL128L Genome Region amplification primers and co-ordinates in BAC-cloned strain genomes						
Narrow	PCR Forward: CAG AAA CTC ACA TCG GCG ACA	175,970 – 175,990	211,709 – 211,729	174,763 – 174,783	52,168 – 52,188	
	PCR Reverse: CCA TCA CCT CGC CTA TAC TAT GTG	178,201 – 179,013	213,876 – 213,899	176,926 – 176,949	49,996 – 50,019	
Wide	PCR Forward: GGC TAA TGG CCA ATA TTG ATT CAA TGT A	175,715 – 175,742	211,456 – 211,483	174,510 – 174,537	52,414 – 52,441	
	PCR Reverse: GAG CCC TTA TCA GCG GTT GGA	178,514 – 178,515	214,265 – 214,286	177,317 – 177,338	49,608 – 49,629	
UL128L genome region sequencing primers						
Internal	Seq 1: TCT TCC AAT ATC GCC ATC TC	176,022 – 176,041	211,761 – 211,780	174,815 – 174,834	52,117 – 52,136	
	Seq 2: CGG ATT GTA GTT GCA GCT CG	176,471 – 176,490	212,210 – 212,229	175,260 – 175,279	51,666 – 51,685	176,621 – 176,640
	Seq 3: TCT GGT TAT TGG CCT CGG TG	176,971 – 176,990	212,710 – 212,729	175,760 – 175,779	51,166 – 51,185	177,121 – 177,140

Seq 4: GCG CAC AGA AGC AGG CAG		177,528 – 177,545	213,267 – 213,284	176,317 – 176,334	50,611 – 50,628	177,678 – 177,695
UL131A 5' – UL132 5' Region amplification and sequencing primers						
UL131A 5' - UL132 5'	PCR Forward + seq: TGT AAC GGG TTT GGT CGG	177,945 – 177,962	213,684 – 213,701			
	PCR Reverse + seq: TGG GCT CGC GTT TTG ATA	179,117 – 179,134				
	Seq: GAA GTC CAT GTG AAG CAG	178,936 – 178,953	214,675 – 214,692			
UL128 and UL131A transcript RT-PCR and cDNA amplification primers						
UL128	PCR Forward + seq: CAT AAA CGT CAA CCA CCC	176,778 – 176,795	212,517 – 212,534	175,567 – 175,584	51,361 – 51,378	
	PCR Reverse + seq: CAC TGC AGC ATA TAG CCC	176,136 – 176,153	211,875 – 211,892	174,929 – 174,946	52,005 – 52,022	
UL131A	PCR Forward + seq: GTG TGG CTG TCT GTT TGT	178,063 – 178,080	213,802 – 213,819	176,852 – 176,869	50,076 – 50,093	
	PCR Reverse + seq: CGT GGT CCT TTT GTT GGT	177,670 – 177,687	213,409 – 213,426	176,459 – 176,476	50,469 – 50,486	

^a - Primer pairs used to amplify the UL128L genome region of viruses (narrow), and also the UL128L genome plus flanking sequences (wide).

^b - Primers that bound regions internally within the UL128L genome region to generate sufficient contigs for sequencing.

^c - Primer pair used for amplification of virus UL131A 5' – UL132 5' genome region and sequencing, and internal primer used for sequencing reactions.

^d - Primer pairs used for the RT-PCR of UL128 and UL131A specified transcripts, PCR-amplification of cDNAs, and also for sequencing reactions.

Nucleotide positions are given relative to sequences of BAC-cloned strains deposited in GenBank: Merlin (GU179001.1); TR (AC146906.1); TB40-BAC4 (EF2999921.1); FIX (AC146907.1); strain 3301 (GQ466044.1).

2.8.2 Cloning of PCR products

PCR products were cloned into plasmid vectors using the TOPO TA Cloning Kit (Invitrogen/Life technologies). In this kit, a linearised plasmid vector contains 3' terminal deoxythymine overhangs that are bound with *Vaccinia* virus Topoisomerase I, and this enzyme facilitates the incorporation of PCR products that display 3' deoxyadenosine overhangs. To incorporate the necessary 3' deoxyadenosine overhangs, gel purified PCR products were first mixed with dNTPs (New England BioSciences) and *Taq* polymerase (Invitrogen/Life Technologies) that has terminal transferase activity, and incubated for 10 mins at 72°C. Ligation reaction mixtures were then prepared from modified PCR products, the linearised plasmid vector and salt water (200 mM NaCl, 10 mM MgCl₂) (both supplied), and ddH₂O. Ligation reactions were incubated at room temperature for 30 mins prior to rapid cooling on ice. Electro-competent One Shot TOP10 *E.coli* cells (supplied) were grown to competency, transformed by electroporation, and recovered as described above in 'recombineering'. To select for cells containing the plasmid that encodes selectable ampicillin resistance (Amp^r), recovered cells were grown on LB agar plates supplemented with ampicillin (50 µg/mL).

2.8.3 RT-PCR of RNA transcripts and production of cDNA libraries

Libraries of cDNAs complimentary to virus-derived mRNAs were generated using a NanoScript Precision reverse transcriptase (RT) kit (Primer Design). In the initial annealing reaction, total infected cell RNA extracts were mixed with target-specific primers and ddH₂O, then incubated at 65°C for 5 mins before being rapidly cooled by transfer to ice. An extension reaction mixture was then prepared from primed mRNAs mixed with the nanoscript RT enzyme and associated reaction buffer, as well as dithiotreitol (DTT), dNTPs and RNase/DNase free water (all supplied). RT extension reactions for the generation of cDNA amplicons was performed during incubation at 55°C for 20 mins.

2.8.4 DNA sequencing

2.8.4.1 Sanger sequencing

Sequencing analysis of short genome regions or PCR products was performed by the Sanger sequencing services offered by Eurofins (MWG). Genome regions of interest were PCR amplified, with DNAs purified as described above. The concentration of DNAs was also determined as previously described. All DNAs and primers were sent in volumes and at concentrations as required. Primers used for sequencing reactions are listed in Table 2.2.

2.8.4.2 Illumina platform whole-genome sequencing

Whole genome sequence analysis, including the detection of mutant populations within grown HCMV working stocks, was performed by Illumina platform high-throughput sequencing,

conducted by collaborators in the laboratory of Dr Andrew Davison (MRC-Centre for Virology, University of Glasgow). Viral genomes were extracted either directly from cell-free virus preparations or from within infected cell cultures by methods described above, and delivered as frozen samples on dry ice.

2.9 Biochemical analysis of HCMV particles

2.9.1 Purification of virus derived particles

Virus particles were purified from infectious supernatants and concentrated virus preparations by ultra-centrifugation over solutions of defined density and viscosity. Ultra-centrifuging was performed using an Optima XPN-80 Ultra-centrifuge (Beckman Coulter), using the SW41 rotor and compatible thin-walled Ultra-Clear centrifuge tubes (Beckman Coulter). All spins were performed at $90465.7 \times g$ for varying times at 20°C.

2.9.1.1 *Purification of total virus particles*

Total virus-derived particles (virions, NIEPs and DBs) were purified from viruses by ultra-centrifugation over 4 mL ‘cushions’ of various solutions, including PBS containing 20% (w/v) sorbitol, and also sodium phosphate buffer containing 15% (w/w) sodium tartrate and 30% (w/w) glycerol (gradient ‘light’ solution). Virus derived particles were purified by ultra-centrifugation over each solution for 1 hr.

2.9.1.2 *Purification of virions*

Virion purification was performed by gravity-flow rate sedimentation of virus-derived particles during ultracentrifugation through glycerol-tartrate positive density negative-viscosity continuous gradients, as previously described (Talbot and Almeida 1977, Irmieri and Gibson 1983). In this method, virions are separated from other virus particles and cellular debris as they migrate at different rates dictated by differences in particle buoyant density.

2.9.1.2.1 *Forming continuous gradients*

Solutions used to form gradients were made in a sodium-phosphate buffer (pH 7.4): ‘heavy’ solution, containing 35% (w/w) sodium tartrate; and ‘light’ solution, containing 15% (w/w) sodium tartrate and 30% (w/w) glycerol. Gradients were formed using a SG50 gradient maker (Hoefer): the ‘heavy’ and ‘light’ solutions were loaded to the two separate chambers of the gradient maker and allowed to gradually mix as they were released via a single exit tube. Gradients were poured into Ultra-clear centrifuge tubes, with the flow of both solutions driven by using a peristaltic pump (Pump P1, Pharmacia Fine Chemicals). Gradients were poured ‘from-the-top’, with ‘heavy’ solution forming the lowest gradient regions, and ‘light’ solution forming the upper regions of the gradient.

2.9.1.2.2 Isolation of virions

Virus preparations were gently loaded on top of gradients such that they did not mix, and formed a clearly defined bi-layer. Gradients were balanced prior ultra-centrifugation by the addition of PBS loaded on top of virus samples. Loaded gradients were ultra-centrifuged for 45 mins at conditions described above. Banded virus particles were recovered using a syringe and a 20 G needle. To remove gradient derived salts and other contaminants from the purified virus particles, recovered bands were washed by gradual dilution in excess NaPh buffer, and pelleted by ultra-centrifugation for 1 hr. The final purified virion pellet was re-suspended in a solvent suitable for downstream applications.

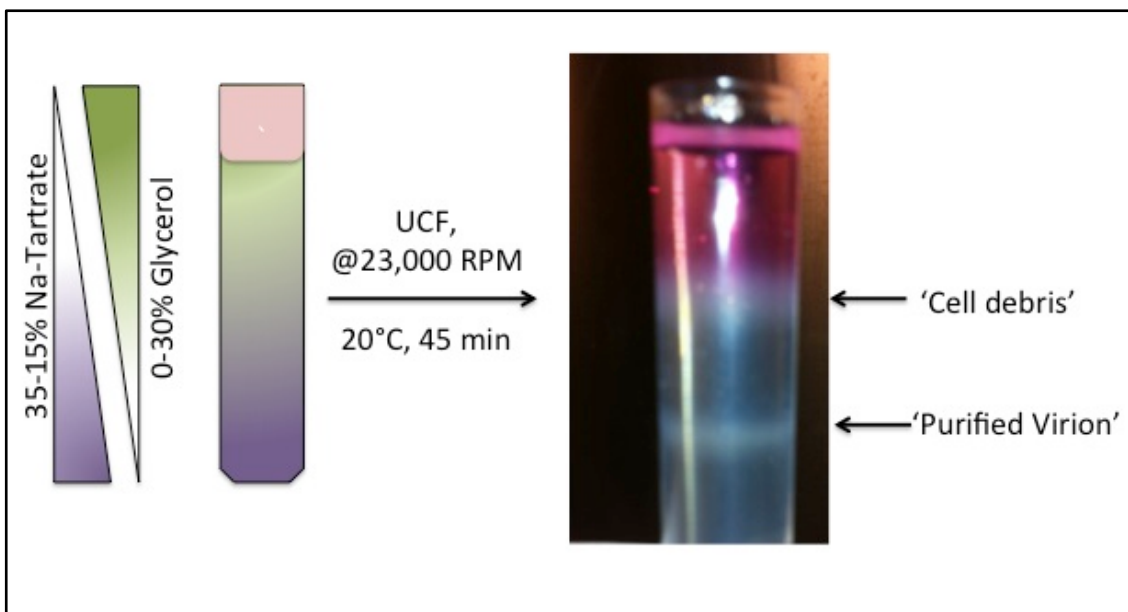


Figure 2.4. Purification of virions over glycerol-tartrate gradients. Glycerol-tartrate gradients were poured from the ‘top-up’, to include lower most regions that were composed of gradient heavy solution (purple; 35% sodium tartrate in NaPh buffer), and uppermost regions composed of gradient light solution (green; 15% sodium tartrate, 30% glycerol, in NaPh buffer). Concentrated virus samples (pink) were loaded on top of gradients, with virus particles purified by gravity-flow rate sedimentation during ultra-centrifugation at conditions described. Bands formed included, ‘cell debris’, ‘purified virions, and ‘DBs’ (not shown).

2.9.2 Immunoblot detection of proteins

Semi-quantitative detection of proteins was performed by immuno-blot techniques, with targeted proteins first resolved and isolated from mixed samples using the NuPage Novex protein isolation kit (Invitrogen/Life technologies).

2.9.2.1 Sample preparation

Samples were prepared in NuPage LDS sample buffer (supplied), supplemented with 10 mM DTT for the reduction of proteins and polypeptides. Where infected cells or virions were re-suspended directly following pelleting, sample buffer based on NuPage LDS (4x) was first diluted to working concentration by the addition of ddH₂O. When applied to infected cells that remained adhered to cell culture vessels, cells were first washed in excess PBS before lysis buffer was added, then cells were scraped before lysates were recovered. When infected cells or virions were suspended in cell culture supernatant or other media, appropriate volumes of NuPage LDS (4x) and DTT were added to achieve final working concentration. Proteins in all samples were denatured during incubation at 100°C for 10 mins using a heating block.

2.9.2.2 Resolution and isolation of polypeptides by electrophoresis

Proteins and polypeptides present in mixed samples were resolved and isolated based on molecular weight by one dimensional polyacrylamide gel electrophoresis, using the XCell 4 SureLock electrophoresis system and compatible pre-poured gels containing 10% (w/v) polyacrylamide in Bis-Tris buffer (supplied). Samples were loaded to individual wells, with a separate well also loaded with Novex Sharp Pre-stained Protein Standard for the estimation of isolated protein size by extrapolation to the migration of proteins of known size. Gels were submersed in MOPs SDS running buffer (supplied) diluted to working concentration, and electrophoresed at 150-200 V for 1-2 hrs until sufficient migration had occurred.

2.9.2.3 Western transfer

Resolved and isolated proteins/polypeptides were electrophoretically transferred to PVDF (GE) or Nitrocellulose (GE) membranes by semi-dry transfer using a TransBlot SD Model 200/2.0 (BioRad). Initially, gels were gently released from the casings in which they were contained, and washed in western transfer buffer. Blotting paper and capture membrane were pre-soaked in the relevant western transfer buffers for 5-10 mins, which included NuPage transfer buffer (supplied), or carbonate transfer buffer prepared by the user. PVDF was initially soaked in methanol for 5 mins. Transfer stacks comprised, from the top down, soaked blotting paper, electrophoresed gels, capture membranes and soaked blotting paper. Proteins were transferred during electrophoresis at 10-20 V for 1-2 hrs. Following transfer, membranes were rinsed in ddH₂O, then stored in excess PBS-T at 4°C.

2.9.2.4 *Immuno-staining of captured proteins and polypeptides*

Two successive rounds of immuno-labelling detected captured proteins, firstly with antibody reactive for targeted proteins (Table 2.4), secondly with HRP-conjugated antibody reactive with the primary antibody used. All treatments were performed on a rocking platform (10-20 RPM), and were followed by washes for 3 mins in excess PBS-T. Membranes were ‘blocked’ in PBS-T supplemented with 5% (w/v) milk proteins (Marvel). Blocking was performed either for 1 hr at room temperature, or overnight at 4°C. Target specific primary antibodies were diluted in blocking buffer and added to membranes that were again incubated at 4°C overnight, or at room temperature for 1 hr. Secondary HRP-conjugated anti-bodies were added to membranes and allowed to bind primary antibodies during incubation at room temperature for 1 hr. Primary anti-bodies used are detailed in HRP-conjugated secondary anti-bodies used were goat anti-mouse monoclonal IgG (Bio-Rad; Cat #170-6515) and goat anti-rabbit monoclonal IgG (Bio-Rad; Cat #170-6516), each used at a 1:2000 dilution.

2.9.2.5 *Detection of captured protein*

As an HRP substrate for the generation of a detectable signal, Super Signal West Pico Chemiluminescent Substrate was used (Pierce). Equal volumes of each supplied component were added to stained membranes that were then incubated for 3-5 mins at room temperature. Chemiluminescent signal was visualised using an Autochemi Bioimaging system (UVP), with images analysed using LabWorks software (UVP).

2.9.3 Mass spectrometry analysis

The proteome composition of virions derived from several Merlin variants was determined by mass spectrometry analysis performed by collaborators in the laboratory of Professor Paul Lehner (Cambridge institute for Medical research, University of Cambridge).

2.9.3.1 *SILAC labelling virions*

Proteins incorporated into the virions analysed were differentially labelled by the stable isotope labelling by amino acids in cell culture (SILAC) method (Ong, Blagoev et al. 2002). In brief, viruses were grown in SILAC medium supplemented with isotopically labelled amino acids arginine, lysine and proline. To ensure only labelled amino acids were available for incorporation into virion proteins, cells (HFFFs) grown to >90% confluency were subjected to a minimum of 2 passages in SILAC medium prior to infection. To increase the efficiency of virus infection, cells were treated over night with the addition of dexamethasone to cell culture medium to a final concentration of 10 µM (Tanaka, Ogura et al. 1984). The following day, cells were infected at an MOI of 2-4 as previously described, then washed in PBS to remove any non-internalised virions. Infections were incubated at standard conditions, with labelled progeny virions released into the supernatant recovered every two days until monolayers were destroyed.

Table 2.4. Details of primary anti-bodies used in detection of proteins by immuno-blot.

Target protein	Source	Product code	Host Animal	Isotype	Clonality	Dilution
Cellular Targets						
Calnexin	Millipore	MAB3126	Mouse	IgG	Monoclonal	1:2,000
S6RP	Cell Signaling	5G10	Rabbit	IgG	Monoclonal	1:2,000
TGN46	Abcam	Ab50595	Rabbit	IgG	Polyclonal	1:2,000
Giantin	Abcam	Ab24586	Rabbit	IgG	Polyclonal	1:5,000
Actin	Sigma	A-2066	Rabbit		Polyclonal	1:10,000
Viral Targets						
pp28	Abcam	Ab6502	Mouse	IgG	Monoclonal	1:10,000
gB	Abcam	Ab6499	Mouse	IgG	Monoclonal	1:4,000
gpUL128 ^a	G. Gerna	SURN			Monoclonal	1:100
Anti-V5	Serotec	MCA1360	Mouse	IgG	Monoclonal	1:20,000

^a – anti-gpUL128 antibody was kindly donated from the laboratory of Giuseppe Gerna (Servizio di Virologia, Università di Pavia, Italy).

2.9.3.2 *Virion concentration estimations by total protein concentration determination*

Total protein concentration of purified virion samples was estimated using a Micro BCA Protein Assay Kit (Thermo Scientific), according to manufacturer's instructions. In brief, total protein concentration was estimated by indirect colorimetric reaction, with protein abundance resulting in the gradual reduction of Cu^{2+} to Cu^{1+} , which in turn chelates molecules of bicinchoninic acid to generate a purple colour. Working reagent was prepared as instructed, and mixed with appropriate volumes of each unknown sample in wells of a 96-well plate, before incubation at 37°C for 2 hrs for the colorimetric reaction to occur. The plate was cooled, and the intensity of colours formed was read at 562 nm using an Omega Fluostar plate reader (BMG Labtech). Protein concentration was determined by extrapolation of the colour intensity readings generated by proteins of known concentration in standard samples.

2.9.3.3 *Preparation of samples*

Purified virion preparations were prepared by suspension in NuPage loading buffer supplemented with 10 mM DTT, then loaded to the wells of pre-poured continuous gradient gels containing 4-12% (w/v) polyacrylamide in Bis-Tris buffer and electrophoresed such that proteins migrated approximately 1-2 cm into the gel. To visualise resolved proteins, the Colloidal coomassie blue staining kit (Invitrogen) was used according to manufacturer's instructions. Captured proteins were fixed by submersion in freshly generated fixing solution, which contained 50% (v/v) methanol, 40% (v/v) ddH_2O and 10% (v/v) acetic acid. Submersed gels were incubated on a rocking platform (10-20 RPM) at room temperature for 10 mins. A staining solution was generated in two steps, with a final formulation composed of 55% (v/v) ddH_2O , 20% (v/v) methanol, 20% (v/v) Stainer A (supplied) and 5% (v/v) Stainer B (supplied). Captured and fixed proteins were submersed in staining solution minus Stainer B and incubated on a rocker (10-20 RPM) from 10 mins. Finally, Stainer B was added and gels were then incubated for sufficient time until maximum staining was achieved (3-12 hrs). Using a sterile blade, segments containing migrated proteins were excised from electrophoresed gels and sectioned into 6 pieces that were frozen and sent to collaborators for mass spectrometry analysis.

3 Growth Characteristics of **Viruses Reconstituted from** **‘Broad-Tropism’ BAC-cloned** **Strains**

Initial investigations aimed to directly compare the growth characteristics of viruses reconstituted from the TR-BAC, TB40-BAC4 and FIX-BAC clones, to that of viruses from two Merlin-BAC variants: Merlin virus that contains wildtype UL128L ORFs and grows to low titres; and Merlin virus that contains non-intact UL128L ORFs and grows to higher titres (Table 3.1). The Merlin-BACs differed at only a single nucleotide position (in UL128 exon 1), yet each contained an otherwise full complement of wildtype genes, with the exception of RL13; the frame-shifting RL13 mutation selected during the *in vitro* propagation of the Merlin clinical isolate prior to BAC cloning was present in each Merlin-BAC variant used. The RL13 ORFs in TB40-BAC4 and FIX also contain lesions, though TR RL13 contained no obvious disruptions. However, given that RL13 mutants are consistently selected during the passaging of HCMV isolates *in vitro*, it is likely that the TR RL13 ORF is not expressed, or the encoded gpRL13 is non-functional. Thus, no attempts were made to repair RL13 in any strain investigated. To provide a detectable marker to monitor infections, all viruses were engineered to express eGFP from an IRES inserted downstream from UL122 (encoding the immediate early protein IE2). The growth characteristics of viruses were monitored immediately following transfection of cells with infectious BAC DNAs: cell-to-cell spread was assessed when transfected cells were incubated under semi-solid overlay, and the sizes of plaques formed were compared (plaque size comparisons and statistical analyses presented in Appendix III); dissemination kinetics and cell-free virus production were assessed from transfected cells incubated without overlay, with the proportion of infected cells (i.e. eGFP+) scored at weekly time-points until monolayers were destroyed, and supernatant samples collected for enumeration of cell-free virus produced by plaque titration on fibroblasts.

Table 3.1. BAC-cloned HCMV strains investigated.

BAC-clone	UL128L status	Designation in text: BAC /Reconstituted Virus
Merlin-BAC ^a	Mutated: G>A (R>stop) in UL128 exon 1 (nt 176,260)	Merlin-UL128L ^{mut} /RV-Me-UL128L ^{mut}
Merlin-BAC ^b	Wildtype	Merlin-UL128L ^{wt} /RV-Me-UL128L ^{wt}
TR-BAC	Intact	TR/RV-TR
TB40-BAC4	Intact	TB40-BAC4/RV-TB40-BAC4
FIX-BAC	Intact	FIX/RV-FIX

^a – Merlin variant containing UL128 mutation acquired during passage in fibroblast prior to BAC cloning

^b – Merlin variant containing wildtype UL128 following repair of mutation acquired during passage in fibroblast prior to BAC cloning

3.1 Growth phenotypes of ‘broad-tropism’ HCMV strains in fibroblasts

The growth characteristics of viruses from each BAC-cloned strain were first investigated in fibroblast cells (HFFFs) that are the cell-type most commonly used for the propagation of HCMV *in vitro*.

3.1.1 Cell-to-cell spread

RV-Me-UL128L^{mut} formed HFFF plaques that were significantly larger than those formed by RV-Me-UL128L^{wt} ($p < 0.0001$), demonstrating the inhibitory impact of intact UL128L ORFs on the cell-to-cell spread of HCMV in fibroblasts (Figure 3.1.A). RV-TR formed the smallest HFFF plaques of any virus investigated, even smaller than those formed by RV-Me-UL128L^{wt} ($p < 0.05$), and therefore displayed the least efficient cell-to-cell spread in fibroblasts. RV-TB40-BAC4 and RV-FIX BACs each formed significantly larger HFFF plaques than RV-Me-UL128L^{wt} ($p < 0.0001$ and $p < 0.01$, respectively), and therefore displayed a greater capacity for cell-to-cell spread in fibroblasts. In fact, HFFF plaques RV-TB40-BAC4 formed by were almost as large as those formed by RV-Me-UL128L^{mut}, whilst HFFF plaques formed by RV-FIX were of intermediate size compared to those from the Merlin viruses.

3.1.2 Dissemination kinetics and cell-free virus production

Transfection of fibroblasts by electroporation is relatively inefficient and results in productive infection in only a very low proportion of cells. Furthermore, the dissemination of viruses containing intact UL128L ORFs is severely restricted until mutants arise (Stanton, Baluchova et al. 2010). Thus, to encourage the early dissemination of each infection, transfected HFFF cultures were disrupted by trypsinisation, with all cells (infected and non-infected) re-seeded to fresh culture vessels where monolayers re-formed with greater distribution of infectivity. Infections were treated in this way at weekly intervals until ~10% of all HFFFs were infected.

Consistent with previous work (Stanton, Baluchova et al. 2010), RV-Me-UL128L^{mut} disseminated rapidly (Figure 3.1.B) and produced high peak titres of cell-free virus (Figure 3.1.C), whereas RV-Me-UL128L^{wt} disseminated considerably slower and produced peak titres of cell free virus that were ~1000-fold lower. Thus, intact UL128L ORFs clearly impeded HCMV dissemination and cell-free virus production in fibroblasts. RV-TR initially spread with delayed kinetics compared to both Merlin viruses and produced the lowest titres of cell-free virus. However this changed dramatically from week 5 post-transfection such that the RV-TR infection was completed earlier and produced a peak titre of cell-free virus >100-fold greater than the RV-Me-UL128L^{wt} infection. RV-TB40-BAC4 and RV-FIX each disseminated with intermediate kinetics compared to the Merlin viruses, and compared to RV-Me-UL128L^{wt}, also produced greater titres of cell-free virus (~1000-fold, and ~35-fold, respectively).

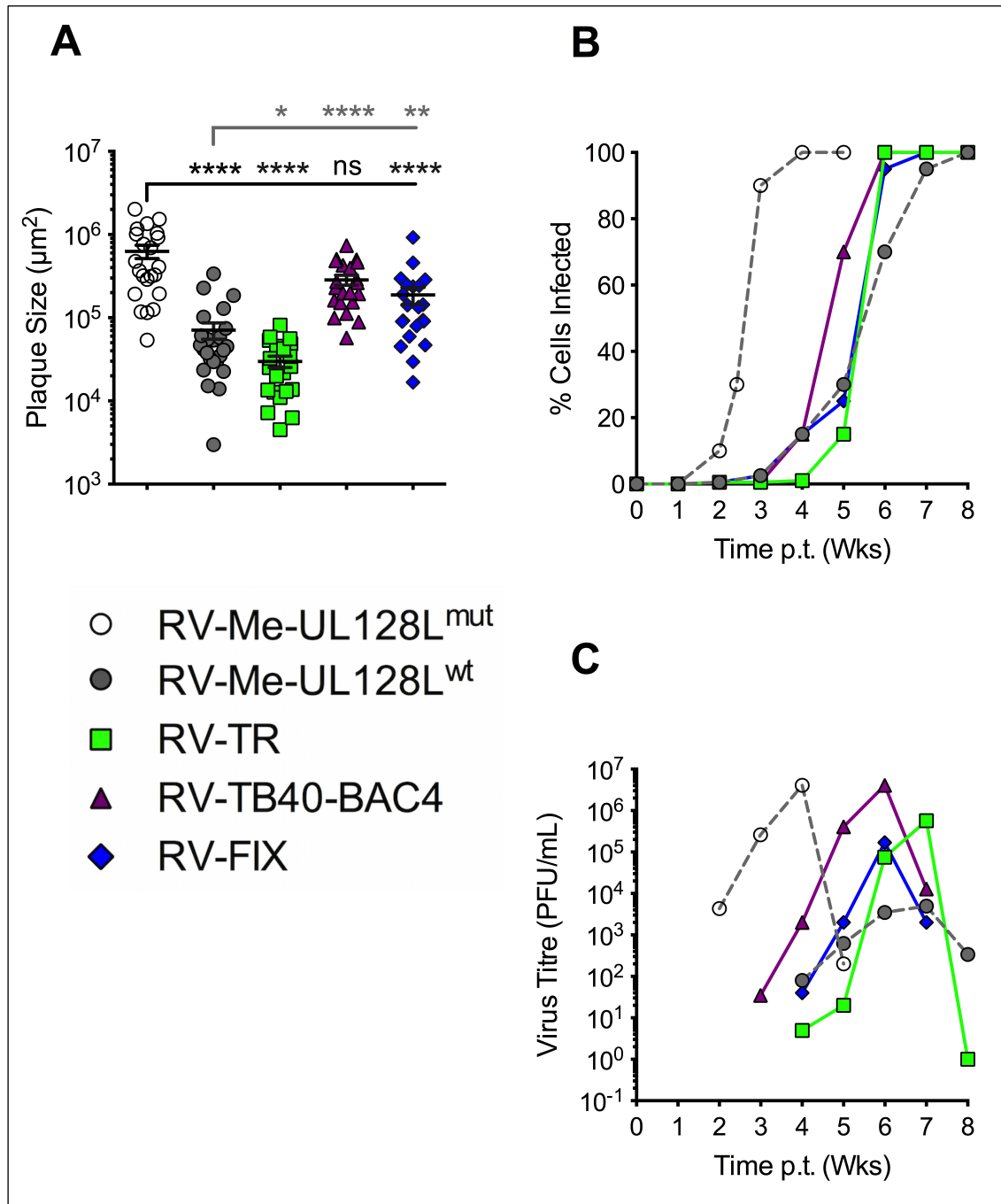


Figure 3.1. Growth characteristics of virus derived from BAC-cloned strains Merlin (UL128L^{mut} or ^{wt}), TR, TB40-BAC4 and FIX in fibroblasts. (A) HFFF cells were transfected with BAC DNA for the indicated viruses and incubated under semisolid overlay, with plaque sizes measured 2 weeks post transfection. A minimum of 20 plaques per infection was measured. All data from 1 experiment (n = 1). Mean plaque sizes and \pm SEM are shown. (Where indicated, the mean plaque sizes formed by RV-Me-UL128L^{wt}, RV-TR, RV-TB40-BAC4 and RV-FIX were compared to those formed by RV-Me-UL128L^{mut} (black line) and/or RV-Me-UL128L^{wt} (grey line): * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$). (B) HFFF cells were transfected with BAC DNA for the indicated viruses, and the dissemination of infection was allowed to progress until the monolayer was destroyed. At weekly time points post transfection (p.t.), the level of infection was estimated by fluorescence microscopy analysis of eGFP-expressing cells. (C) Supernatant samples collected at weekly time-points from the infections shown in panel B were titrated on HFFFs to provide a measure of cell-free virus release throughout the time-course of each infection.

3.1.3 Validation of BAC-cloned strain growth phenotypes in fibroblasts

Since HCMV can undergo adaptation during a single passage in fibroblasts, we sought to determine whether any mutations had been selected during the above-described growth kinetics assays. The relative order of HFFF plaque sizes formed during titration of supernatant samples collected at final time-point in each infection (HFFF-p1) was compared to that formed immediately following transfection of HFFFs (p0). The relative order of plaque sizes at p0 was closely re-capitulated by all HFFFp1 progeny (Figure 3.2. A+B). However, whereas RV-TR-p0 formed plaques significantly smaller than RV-Me-UL128L^{wt}-p0, RV-TR-HFFFp1 formed HFFF plaques with more similar size to those produced by RV-Me-UL128L^{wt}-HFFFp1. This implied adaptation was apparently not due to mutations in RV-TR-HFFFp1 UL128L, since PCR-based sequencing showed that UL128L in all viruses was equivalent to the parental BAC clone.

3.1.4 Conclusion

Compared to the Merlin BAC-derived virus containing wildtype UL128L ORFs, viruses derived from the TB40-BAC4 and the FIX-BAC clones spread by the direct cell-to-cell route in fibroblasts more efficiently, whilst together with virus from the TR-BAC, they also disseminated throughout fibroblast cultures more rapidly and produced greater titres of cell-free virus. However, in comparison to virus from the Merlin-BAC variant containing non-intact UL128L, none spread in fibroblasts by the cell-to-cell route as efficiently, disseminated as rapidly, nor produced greater quantities of cell-free virus.

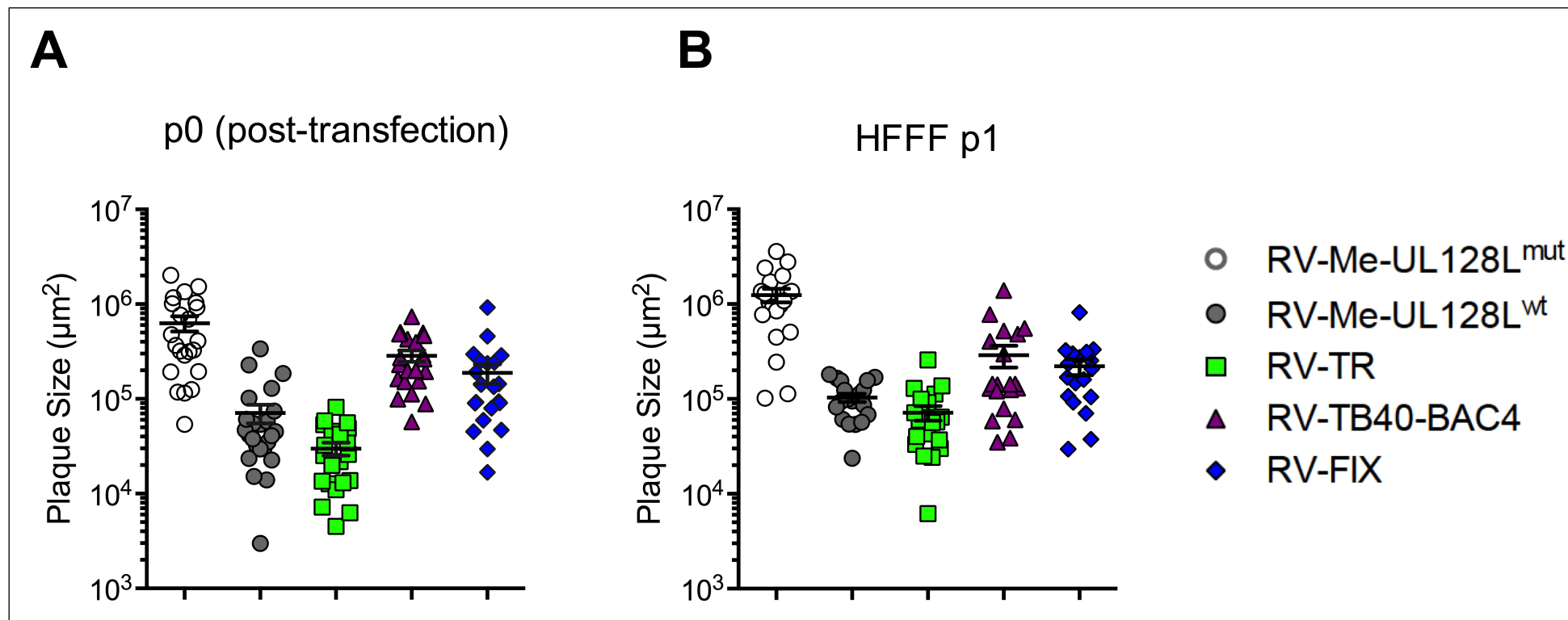


Figure 3.2. Cell-to-cell spread of viruses derived from each BAC-cloned strain in fibroblasts, before and after passage in fibroblasts cells (A) Plaque sizes \pm SEM formed 2 weeks post-transfection (p.t.) of HFFFs with BAC DNA for each indicated virus and incubation under overlay (from Figure 1.A). (B) HFFF plaque sizes and \pm SEM of formed by HFFF-p1 viruses, at 2 weeks post-infection (p.i.) and incubation under overlay.

3.2 Growth phenotype of ‘broad-tropism’ HCMV strains in epithelial cells

The growth properties of viruses reconstituted from each BAC-clone were also assessed in epithelial cells (RPE-1s), where virus requires an intact UL128L genome region and virion-incorporated pentameric complex for efficient infection.

3.2.1 Cell-to-cell spread

RV-Me-UL128L^{wt} formed the largest RPE-1 plaques of any virus investigated, and thus displayed the greatest cell-to-cell spread in epithelial cells (Figure 3.3.A). In contrast, RPE-1 plaques formed by RV-Me-UL128L^{mut} were significantly smaller than those formed by RV-Me-UL128L^{wt} ($p < 0.0001$); in the vast majority of cases these plaques were represented by single cells that were likely infected directly following transfection. This demonstrated the severely restricted cell-to-cell spread of virus lacking intact UL128L in epithelial cell culture. Unlike RV-Me-UL128L^{mut}, RV-TR and RV-TB40-BAC4 each consistently formed multi-cellular RPE-1 plaques, however they were significantly smaller than those formed by RV-Me-UL128L^{wt} (both $p < 0.0001$). RV-FIX displayed similar epithelial cell-to-cell spread properties as RV-Me-UL128L^{mut} and produced RPE-1 plaques significantly smaller than those formed by RV-Me-UL128L^{wt} ($p < 0.0001$). In fact, numerous RPE-1 plaques produced by RV-FIX represented single infected cells, yet this occurred less frequently than in the RV-Me-UL128L^{mut} infection.

3.2.2 Dissemination kinetics and cell-free virus production

The dissemination of infections in epithelial cell culture was monitored by flow cytometry; infected RPE-1 monolayers were trypsinised at weekly time-points post-transfection and ~1/7th of the total cells were collected for analysis, with the remaining cells seeded into fresh flasks for monolayers to reform and for infections to continue.

In a contrast to results seen in fibroblasts, RV-Me-UL128L^{wt} disseminated throughout an RPE-1 monolayer with the greatest kinetics of any virus investigated, whilst RV-Me-UL128L^{mut} did not spread by any noteworthy degree (Figure 3.3.B). Thus, not only was virus lacking intact UL128L unable to spread by the cell-to-cell route in epithelial cells, spread by the cell-free supernatant driven route was also severely restricted. RV-TR and RV-TB40-BAC4 did spread throughout RPE-1 monolayers, albeit slower than RV-Me-UL128L^{wt}. Like RV-Me-UL128L^{mut}, RV-FIX did not spread throughout an RPE-1 monolayer. In contrast to the direct relationship between dissemination kinetics and cell-free virus production seen in fibroblasts, the slower spreading RV-TR and RV-TB40-BAC4 infections produced peak titres of cell-free virus that were ~85-fold, and ~140-fold greater than that produced by RV-Me-UL128L^{wt}, respectively (Figure 3.3.C). As RV-Me-UL128L^{mut} and RV-FIX did not establish productive infection in RPE-1s, supernatants from these transfected cultures were not analysed.

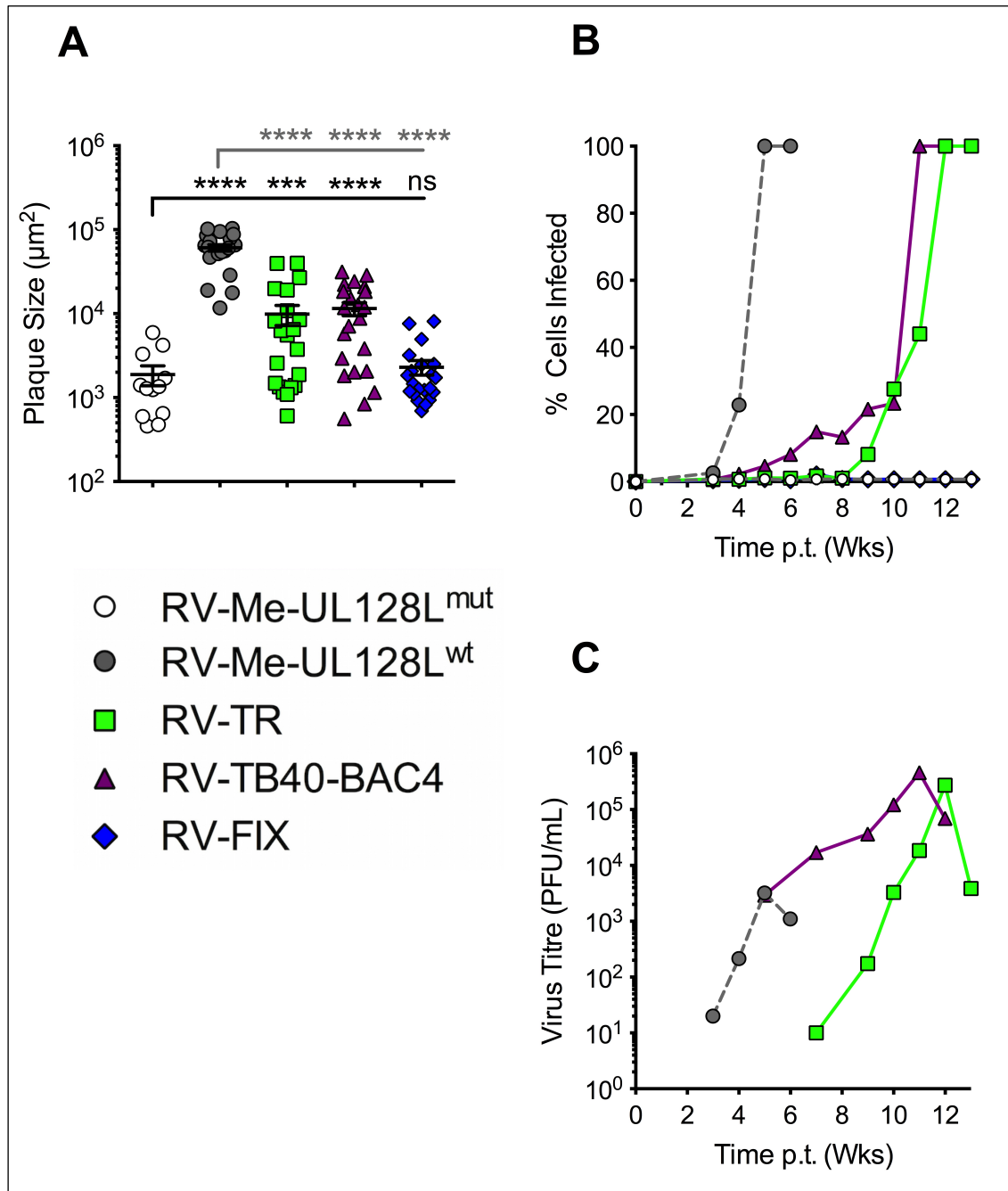


Figure 3.3. Growth characteristics of virus derived from BAC-cloned strains Merlin (UL128L^{mut} or UL128L^{wt}), TR, TB40-BAC4 and FIX in epithelial cells. (A) RPE-1 cells were transfected with BAC DNA for the indicated viruses and incubated under semi-solid overlay, with plaque sizes measured 3 weeks post transfection. Where possible, a minimum of 20 plaques from each infection was measured. All data from 1 experiment (n = 1). Means and \pm SEM are shown. (Where indicated, the mean plaque sizes formed by RV-Me-UL128L^{mut}, RV-TR, RV-TB40-BAC4 and RV-FIX were compared to those formed by RV-Me-UL128L^{mut} (black line) and RV-Me-UL128L^{wt} (grey line) by 1-way analysis of variance [ANOVA] followed by Dunnett's posttest: ns – non-significant, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$). (B) RPE-1 cells were transfected with BAC DNA for the indicated viruses, and the dissemination of infection was allowed to progress until the monolayer was destroyed. At weekly time points post transfection (p.t.), the level of infection was estimated by flow cytometry analysis of eGFP-expressing cells. (C) Supernatant samples collected at weekly time-points from the infections shown in panel B were titrated on HFFFs to provide a measure of cell-free virus release throughout the time-course of each infection.

3.2.3 Validation of BAC-cloned strain growth phenotypes in epithelial cells

As in fibroblasts, HCMV can also undergo rapid adaptation to epithelial cell culture, and this could involve mutations in genome regions that are unstable in both cell types (Dargan, Douglas et al. 2010). Hence, to assess whether mutations had been selected in the growth kinetics assays performed in epithelial cells, the relative HFFF plaque sizes formed during titration of progeny virus from each RPE-1 infection (RPE-1p1) were compared to those formed immediately following transfection of HFFFs (p0) (Figure 3.4.A+B). The relative HFFF plaque sizes formed by TB40-BAC4-RPE-1p1 and RV-Me-UL128L^{wt}-RPE-1p1 were comparable to those formed immediately following transfection, indicating that the TB40-BAC4 BAC-derived virus had not undergone any major adaptation. However, whereas RV-TR-p0 formed smaller HFFF plaques than RV-Me-UL128L^{wt}-p0, RV-TR-RPE-1p1 formed slightly larger HFFF plaques than RV-Me-UL128L^{wt}-RPE-1p1. Thus, RV-TR had apparently adapted in RPE-1s. Again, PCR-based sequencing indicated that UL128L in all RPE-1-p1 viruses was equivalent to that in each parental BAC.

3.2.4 Conclusion

Compared to Merlin virus containing wildtype UL128L, virus derived from TR-BAC and TB40-BAC4 spread cell-to-cell and disseminated throughout epithelial cultures less efficiently, yet produced greater quantities of cell-free virus. Surprisingly, virus reconstituted from the FIX-BAC was unable to establish productive infection in RPE-1s and displayed a growth phenotype similar to the Merlin virus lacking intact UL128L ORFs.

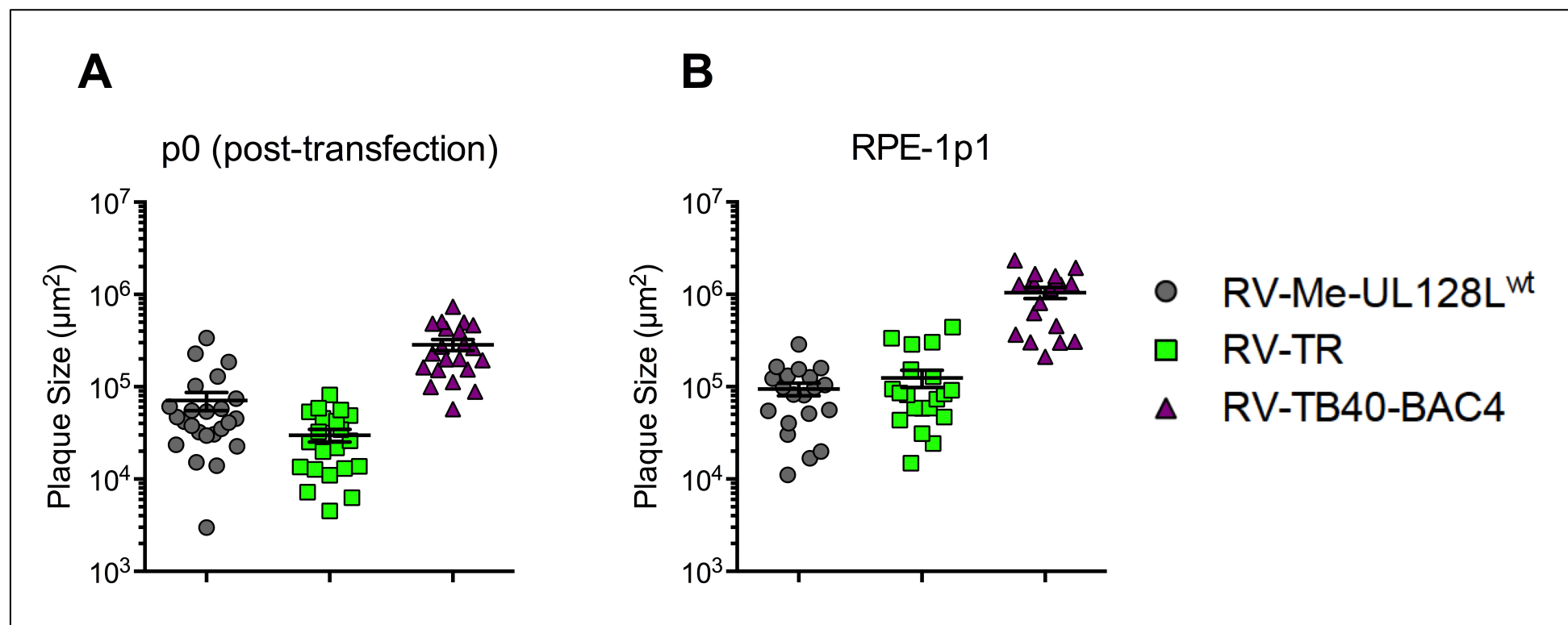


Figure 3.4. Cell-cell spread of viruses derived from each BAC-cloned strain in fibroblasts before and after passage in epithelial cells. (A) Plaque sizes \pm SEM formed 2 weeks post transfection of HFFFs with BAC DNA and incubation under overlay. (B) HFFF plaque sizes and \pm SEM formed by RPE-1-p1 viruses, at 2 weeks post-infection and incubation under overlay.

3.3 Role of variant UL128L in the differential growth phenotypes of each strain

Data from the growth kinetics assays suggested that whilst UL128L was apparently intact in each of the TR-BAC and TB40-BAC4 genomes, the inhibitory impact of this genome region on the growth of virus in fibroblasts was less severe than in wildtype Merlin virus infections. The same consideration applied to FIX, however the assumption that FIX-BAC UL128L was intact was based on the reported use of this strain for infections in cell types where virus requires UL128L (i.e. endothelial cells). We hypothesised that the distinct growth phenotypes of these strains was due to one, or a combination, of three potential circumstances: i) each contains natural UL128L variants that are less inhibitory to the growth of virus in fibroblasts cells; ii) each contain subtle UL128L mutations, acquired during the fibroblast passage of the corresponding parental viruses prior to BAC-cloning, that suppress but not ablate pentameric complex functionality; iii) other genome regions in each strain compensate for the inhibitory effects of UL128L. Sequencing data for the un-passaged clinical isolates from which the TR, TB40-BAC4 and FIX BAC-cloned genomes were captured was not available for comparison. We therefore sought to determine the specific impact of each variant UL128L on the growth phenotypes of virus. To this end, Merlin BAC variants containing each variant UL128L in place of the native UL128L were constructed, and the derived recombinant viruses were assessed to see whether the growth characteristics of each donor strain were transferred - performed by Dr Richard Stanton (School of Medicine, Cardiff University) (Appendix II).

In brief, the Merlin virus expressing TR UL128L displayed an almost identical growth phenotype as the parental Merlin virus containing native wildtype UL128L. Thus, UL128L was not a major contributing factor to the distinct growth phenotype of TR BAC-derived virus. Interestingly, RV-Me-UL128L^{TR} did produce greater titres of cell-free virus at late time-points in fibroblast culture infections compared to RV-Me-UL128L^{wt}. However, this was later shown to have occurred following the acquisition of a large deletion encompassing the UL128L ORFs (UL128 and UL130), and not due to features inherent to TR UL128L. Transfer of TB40-BAC4 and FIX UL128L did alter the growth phenotype of Merlin virus. Compared to RV-Me-UL128L^{wt}, RV-Me-UL128L^{TB40} and RV-Me-UL128L^{FIX} each displayed greater cell-to-cell spread, more efficient dissemination kinetics, and increased cell-free virus production in fibroblasts. Furthermore, both also displayed reduced cell-to-cell spread, yet greater cell-free virus production in epithelial cells. The dissemination kinetics of RV-Me-UL128L^{FIX} in epithelial cell culture, but not RV-Me-UL128L^{TB40}, also differed to that of RV-Me-UL128L^{wt}. Thus, UL128L contained within the TB40-BAC4 and FIX BACs did contribute, at least partially, to the distinct growth phenotypes of virus produced by each corresponding BAC-cloned strain. Moreover, FIX UL128L was confirmed to be functionally intact when expressed in the Merlin genome background.

3.4 Comparative nucleotide sequence analysis for the identification of *in vitro* acquired mutations, or other variations that conveyed distinct growth phenotype of each strain

The UL128L genome region of TB40-BAC4 and FIX contained features that apparently altered pentameric complex functionality, ultimately impacting the growth phenotype of virus. In contrast, the pentameric complex incorporated into Merlin virions yet composed of TR UL128L encoded sub-units was functionally equivalent to the pentameric complex composed of wildtype Merlin UL128L encoded sub-units. Therefore, genome regions other than UL128L appeared to be responsible for the distinct growth phenotype of TR BAC-derived viruses. To identify nucleotides that defined the different growth phenotype of each strain, comparative nucleotide sequence analysis of each variant UL128L ORF, and/or other genome regions, was performed.

Initially, the sequences of TR, TB40-BAC4 and FIX UL128L were compared to Merlin UL128L (Table 3.2.). Similarly to descriptions of UL128L coding potential in distinct wildtype strains (Baldanti, Paolucci et al. 2006, Sun, Ji et al. 2009), each variant UL128L ORF contained protein coding nucleotide sequences that were highly homologous: compared to the corresponding ORF in Merlin UL128L, each variant UL128 shared between 98.16-98.95% nucleotide sequence identity, with variations mostly commonly observed in exon 1; each variant UL130, the most variable of the UL128L ORFs, shared between 97.52-98.29% nucleotide sequence identity with variations distributed throughout the entire ORF; and each variant UL131A shared 98.19-98.39% nucleotide sequence identity, with variations most frequently observed in exon 1.

Of interest were coding nucleotide variations that potentially represented *in vitro* acquired mutations. Given the high degree of homology between UL128L from each strain, the primary criteria applied for candidates of '*in vitro* acquired mutations' were nucleotides that were uncommon, or indeed unique to any given strain, and thus unlikely to represent natural variations. The common or unique nature of nucleotide variations were then assessed by performing BLAST searches against all other HCMV UL128L sequences (~75 either partial or complete UL128L sequences) deposited in GenBank, including those derived from both highly-passaged laboratory-adapted strains, and clinical isolates undergone no passage at all.

Table 3.2. Nucleotide variations in UL128L ORFs from TR-BAC, TB40-BAC4, and FIX-BAC compared to the wildtype Merlin-BAC.

BAC clone	ORF	Nucleotide Identity to Merlin ORF^a	<i>Nucleotide Identity/Variations in coding nucleotides^{b,c}</i>	Amino Acid Identities/Variation
TR-BAC	UL128	748/759 (98.55%)	Exon 1 159/164 (96.95%): A15G, C34T¹ , G43A, A49G, C75T Exon 2 134/135 (99.26%): C325T² Exon 3 215/217 (99.08%): T587C, A647G	168/171 (98.24%): A12T ¹ , V77I ²
	UL130	633/645 (98.14%)	633/642 (98.86%): G25A¹ , G79A² , T96C, T99C, T198C, A206G³ , C305T⁴ , G399A, C351G, T357C, C525T, A540G	210/214 (98.13%): H9Y ¹ , D27S ² , L69S ³ , R102K ⁴
	UL131A	490/498 (98.39%)	Exon 1 233/236 (98.73%): C11T¹ , T69C, C75T Exon 2 153/154 (99.35%): T453C	128/129 (99.22%): C4Y ¹
TB40-BAC4	UL128	745/759 (98.16%)	Exon 1 159/164 (96.95 %): C13T¹ , A15G, G46A, C53T² , C75T Exon 2 135/135 (100%) Exon 3 215/217 (99.08 %): T587C, A647G	169/171 (98.83%): D5N ¹ , G18D ²
	UL130	634/645 (98.29%)	634/642 (98.75%): C6T, C87T, T96C, T99C, G119A¹ , G178T² , T198C, G233A³ , C351G T357C, C390T	211/214 (98.60%): L40I ¹ , L60I ² , S78L ³

	UL131A	489/498 (98.19%)	Exon 1 234/236 (99.15%): T69C, C75T Exon 2 153/154 (99.35%): T453C	129/129 (100%)
FIX-BAC	UL128	751/759 (98.95%)	Exon 1 161/164 (98.78%): C13T ¹ , A15G, C75T Exon 2 135/135 (100%) Exon 3 215/217 (99.08%): T587C, A647G	170/171 (99.42%): D5N ¹
	UL130	629/645 (97.52%)	629/642 (97.98%): C6T, G86A ¹ , C87G, T96C, T99C, A139G ² , T198C, A206G ³ , A214G ⁴ , G399A, C351G, G354A T357C, G480A, C525T, A540G	210/214 (98.13%): S29F ¹ , S47P ² , L69S ³ , S72P ⁴
	UL131A	490/498 (98.39%)	Exon 1 233/236 (98.73%): T69C, C75T, G135A Exon 2 153/154 (99.35%): T453A	129/129 (100%)

^a - nucleotide identities through entire corresponding ORF, including coding nucleotides in UL128 and UL131A exons, as well as non-coding nucleotide in the introns of these ORFs.

^b - nucleotide co-ordinates denoted reflect position in Merlin UL128, UL130 and UL131A ORF sequences deposited in GenBank (GU179001).

^c - Bold typeface indicates non-synonymous nucleotide variations; ^{number} matches amino acid variation in corresponding glycoprotein product.

^d - amino acid co-ordinates denoted reflect position in Merlin-encoded gpUL128, gpUL130 and gpUL131A glycoproteins.

3.5 Genome regions contributing to the growth phenotype of virus from the TR-BAC

Compared to Merlin UL128L, each TR UL128L ORF contained non-synonymous nucleotide variations. As potential candidates of *in vitro* acquired mutations, the R102K amino acid substitution encoded by TR UL130 was identified only in one further HCMV sequence deposited in GenBank, derived from ‘Isolate 39J’ (DQ229947.1), whilst the C4Y amino acid substitution encoded by TR UL131A was identified only in ‘Strain AL’ (GQ22015.2) and ‘Isolate Burns H’ (GU574781.1). However, with few reports regarding the origins and passage history of these strains, their relationship to the strain TR viral lineage could not be assessed in order to predict whether these features represent uncommon, yet natural sequence variations. Furthermore, given the inability of TR UL128L to convey any distinct growth characteristics when expressed by Merlin virus, these features were not investigated.

3.5.1 Impact of UL131A 5’-UL132 5’ genome region on RV-TR growth phenotype

TR-BAC derived virus displayed distinct phenotypes in pentameric complex influenced activities, yet the TR UL128L ORFs were not responsible for this. We therefore considered whether transcriptional regulatory components that govern TR UL128L expression might be responsible. Comparative nucleotide sequence analysis was applied to genome regions flanking TR UL128L: upstream of UL131A, the promoter from which a polycistronic single read-through transcript encoding all UL128L glycoproteins is predicted to reside; downstream of UL128, the polyadenylation signal common to all UL128L derived transcripts is located. A region containing multiple uncommon and unique nucleotide variations was identified immediately upstream of UL131A and overlapping UL132 (Figure 3.5.) due to several of these nucleotides laying outside of described ORFs, nucleotide co-ordinates are denoted in relation to the corresponding position in the Merlin BAC genome sequence (GU179001). Within the TR UL131A-UL132 intragenic region were the uncommon nucleotide residues G at nt 178,115 and A at nt 178,156, each absent in the vast majority of other HCMV sequences in GenBank. TR UL132 also contained uncommon nucleotides, with a G>T at nt 178,618 identified only in strains ‘HAN 33’ (GQ221998.1) and ‘HAN 21’ (GQ221987.1), and a G>A at nt 178,861 identified only in strains ‘NT’ (GQ222016.1) and strain ‘HAN 8’ (GQ221979.1). Unique nucleotide variations identified in TR UL132 included a G>A at nt 178,233 and a G>A at nt 178,750. The variations at nt 178,233 and nt 178,861 were non-synonymous, specifying P261S and A51T amino acid substitutions, respectively. Whilst these features possibly impacted the UL128L ORF common promoter, they also potentially impacted UL132 that encodes a virion envelope glycoprotein (gpUL132) reported to enhance the growth of virus in fibroblasts (Spaderna, Kropff et al. 2005).

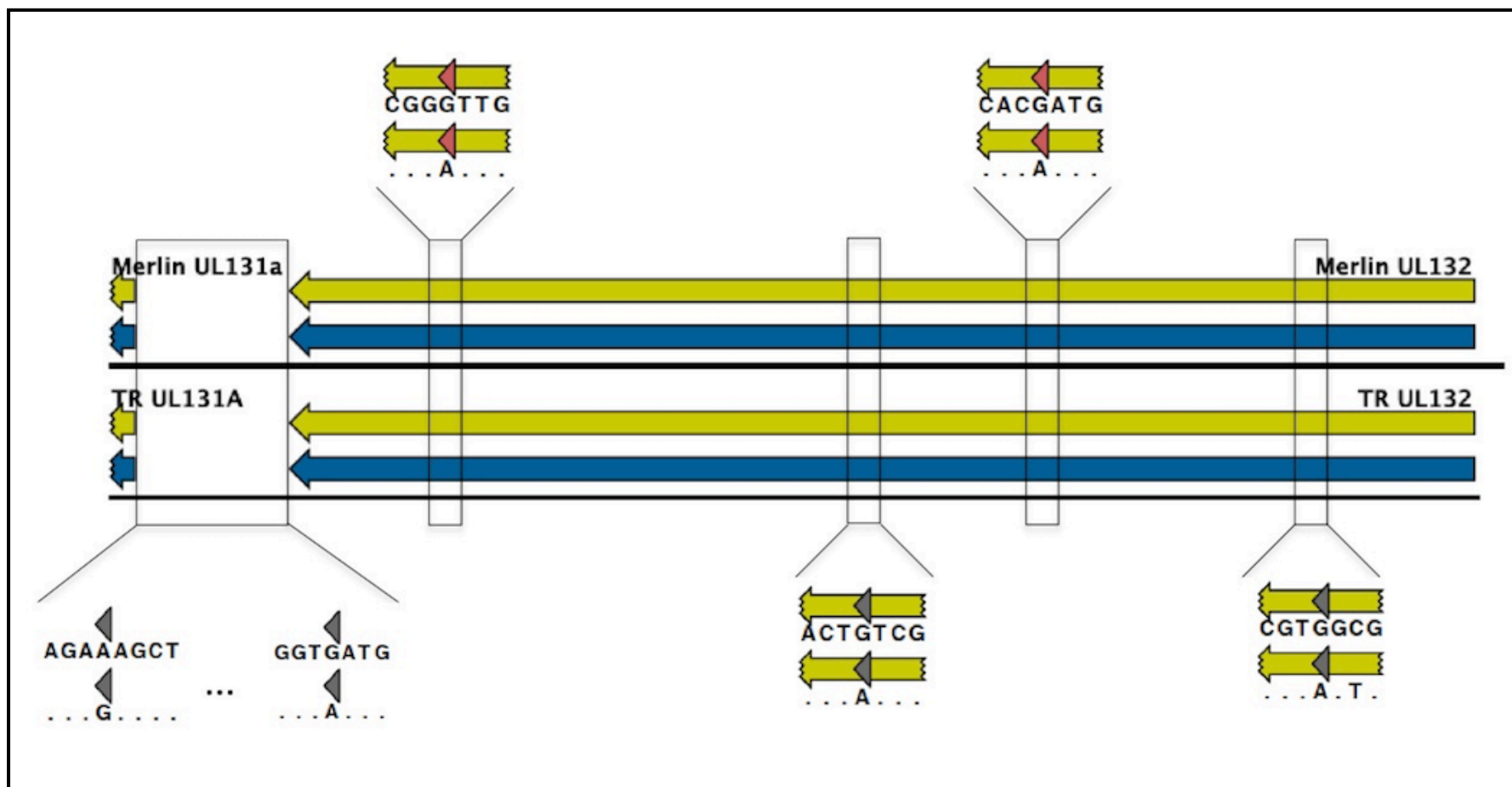


Figure 3.5. Nucleotide sequence identified in TR UL131A 5' – UL132 5' genome region. The UL131A 5' – UL132 5' genome region of the Merlin and TR BAC-cloned genomes is indicated by a single black line; ORFs are depicted in yellow. Nucleotide positions are given relative to the Merlin BAC sequence (GU179001). Dots indicate identical nucleotides in comparison sequence. Nucleotide variations highlighted in grey are uncommon amongst all sequences deposited in GenBank, though not unique to TR. In the UL131A – UL132 intragenic region, A>G variation at nt 178,115 and G>A at nt 178,156, and in the UL132 ORF, G>A at nt 178,618 and G>A at nt 178,861. Variations highlighted in pink are unique to TR, and represent G>A substitutions at positions nt 178,233 and nt 178,750.

The TR UL131A 5' – UL131 5' genome region was inserted into the Merlin-UL128L^{wt} BAC genome background to generate the novel BAC variant Merlin-UL132^{TR}, and the growth characteristics of the derived recombinant Merlin virus were compared to those of the parental viruses containing native UL131A 5' – UL132 5' genome regions (Table 3.3).

Table 3.3. Viruses containing UL131A 5' – UL132 5' from TR or Merlin.

Genome Background	UL131A 5'- UL132 5' Origin	Designation in text: BAC /Reconstituted Virus
Merlin	Native	Merlin-UL128L ^{wt} /RV-Me-UL128L ^{wt}
Merlin	TR	Merlin-UL132 ^{TR} /RV-Me-UL132 ^{TR}
TR	Native	TR /RV-TR

3.5.1.1 Impact of TR UL131A 5' – UL132 5' in fibroblasts

The cell-to-cell spread of each virus was investigated in transfected fibroblasts (Figure 3.6.A). Consistent with previous results, RV-Me-UL128L^{wt} formed HFFF plaques that were larger than those formed by RV-TR ($p < 0.05$). RV-Me-UL132^{TR} formed plaques of comparable size to those formed by RV-Me-UL128L^{wt}, and larger than those of RV-TR ($p < 0.01$).

3.5.1.2 Impact of TR UL131A 5' – UL132 5' in epithelial cells

The cell-to-cell spread of RV-Me-UL132^{TR} was also assessed in epithelial cells (3.6.B). Again, consistent with previous findings, RV-Me-UL128L^{wt} produced larger RPE-1 plaques compared to RV-TR ($p < 0.001$). Similar to what was observed in fibroblasts, RV-Me-UL132^{TR} displayed cell-to-cell spread in epithelial cells that closely reflected that of the parental Merlin virus RV-Me-UL128L^{wt}, and importantly, formed RPE-1 plaques that were significantly larger than those formed by RV-TR ($p < 0.0001$). The dissemination kinetics of RV-Me-UL132^{TR} in RPE-1s were also identical that of RV-Me-UL128L^{wt}, and dramatically faster than that of RV-TR (Figure 3.6.C).

3.5.2 Conclusion

The growth characteristics of the recombinant Merlin virus containing the TR UL131A 5' – UL132 5' genome region closely reflected that of the parental Merlin virus and not that of the TR BAC-derived virus. Hence, the nucleotide sequence features identified in TR UL131A 5' – UL132 5' were not the cause of the distinct growth phenotype of virus reconstituted from the TR-BAC compared to virus from the Merlin-UL128L^{wt} BAC.

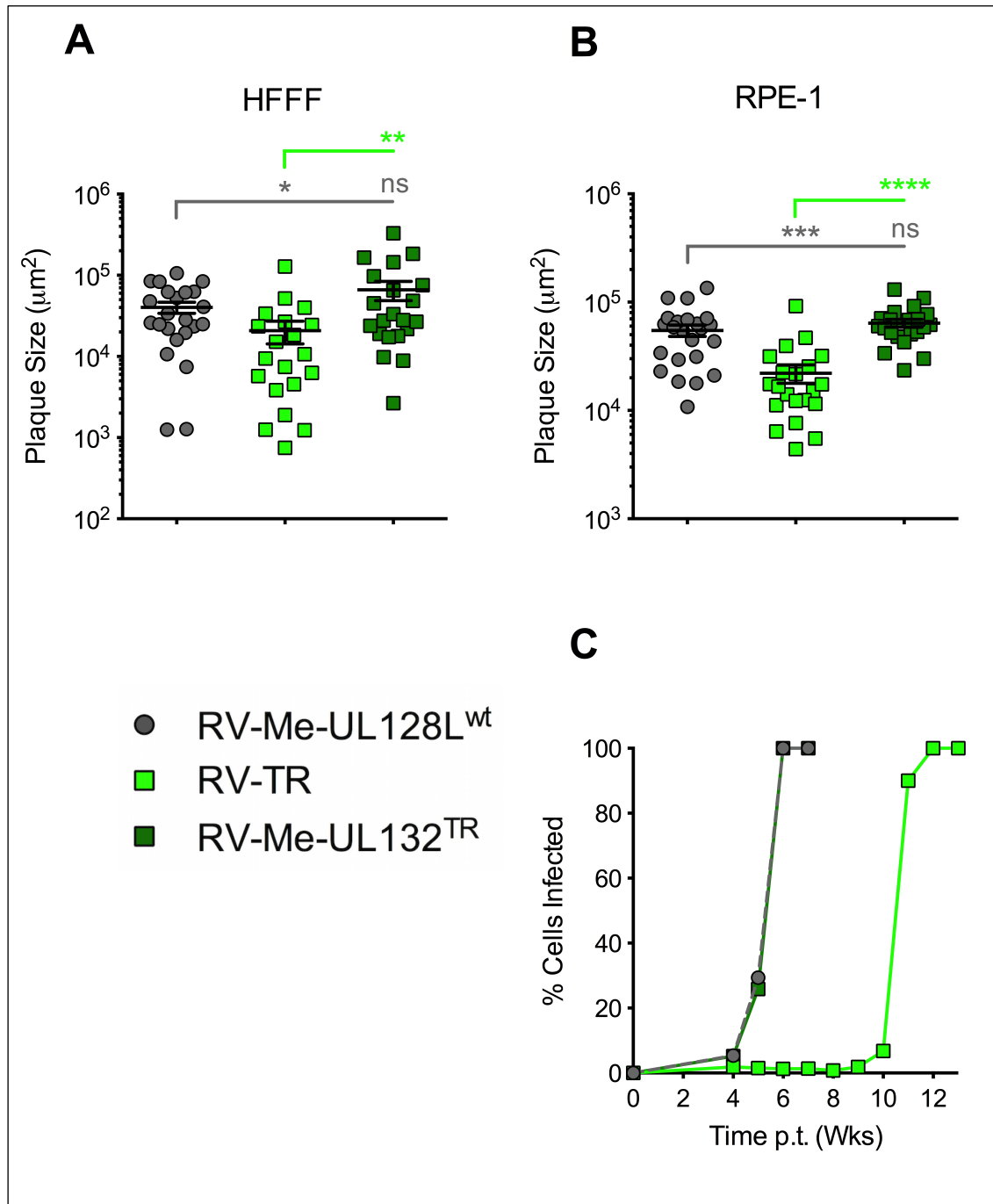


Figure 3.6. Growth characteristics of virus containing TR UL131A 5' – UL132 5' genome region in fibroblasts and epithelial cells. (A) HFFF cells were transfected with BAC DNA for the indicated viruses and incubated under semisolid overlay, with plaque sizes measured 2 weeks post-transfection (p.t). (B) RPE-1 cells were transfected with BAC DNA for the indicated viruses and incubated under semisolid overlay, with plaque sizes measured 3 weeks post transfection. All data from 1 experiment ($n = 1$). A minimum of 20 plaques from each infection was measured, where possible. Means and \pm SEM are shown. (Where indicated, the mean plaque sizes formed by RV-Me-UL128L^{wt}, RV-TR, and RV-Me-UL132^{TR} were compared to those formed by RV-Me-UL128L^{wt} (grey line) and/or RV-TR (green line) by 1-way analysis of variance [ANOVA] followed by Dunnett's posttest: ns – non-significant, * $p < 0.05$, ** $p < 0.01$, *** $p < .001$, **** $p < 0.0001$). (C) RPE-1 cells were transfected with BAC DNA for the indicated viruses. At weekly time-points post-transfection (p.t.), the level of infection was estimated by fluorescence microscopy analysis of eGFP-expressing cells.

3.6 Nucleotides contributing to the growth phenotype of viruses from TB40-BAC4

TB40-BAC4 UL128 and UL130 each contained non-synonymous nucleotide variations compared to the corresponding Merlin ORFs, yet none were unique to this strain. TB40-BAC4 UL131A contained 3 coding nucleotide variations, yet each was synonymous. Thus, TB40-BAC4 UL128L encoded no amino acid variations that were strong candidates for *in vitro* acquired mutations. The TB40-BAC4 UL128L sequence is highly homologous to that in strain ‘3301’ (GQ466044.1); this strain has not been passaged *in vitro* and can be considered to contain wildtype UL128L sequence (Cunningham, Gatherer et al. 2010). TB40-BAC4 and strain 3301 UL128 ORFs differed at three nucleotide positions, only 1 of which was a coding nucleotide but was synonymous. UL131A in these strains differ at only one position in the intron. The UL130 ORFs differed at 5 nucleotide positions, 2 of which encoded amino acid variations; in comparison to Merlin UL130, TB40-BAC4 encodes an S78L amino acid substitution, whilst strain 3301 encodes an A20G amino acid substitution. The fact that TB40-BAC4 and strain 3301 UL128L ORFs differed to Merlin UL128L ORFs in a near identical way provided an opportunity to determine whether natural variations were responsible for the distinct growth phenotype impact of TB40-BAC4 UL128L. Thus, the UL128L genome region from strain 3301 was inserted into the Merlin-BAC genome to generate Merlin-UL128L³³⁰¹. Interestingly, unique nucleotide variation was identified in TB40-BAC4 UL128 intron 1 (G>T; nt 282), 6 nucleotides upstream from the intron 1 splice acceptor site (Figure 3.7). This feature was absent in all HCMV sequences deposited in GenBank, including strain 3301. Moreover, this variation was also absent in the consensus sequence from the mixed population virus ‘TB40/E’ from which TB40-BAC4 was cloned, and also sibling clone ‘TB40/E clone Lisa’ (KF297339.1). The UL128 intron 1 G>T variation was therefore a candidate for an *in vitro* acquired mutation and inserted into the Merlin-UL128L^{wt} BAC, generating Merlin-UL128L^{G>T}. The growth phenotypes of Merlin viruses expressing TB40-BAC4 and strain 3301 UL128L sequences were compared to viruses from the original Merlin-BACs variants (Table 3.4).

Table 3.4. Merlin Viruses containing TB40-BAC4, or strain 3301 UL128L sequences.

Genome Background	UL128L Origin / Nucleotide features	Designation in text: BAC /Reconstituted Virus
Merlin	Native, mutated	Merlin-UL128L ^{mut} /RV-Me-UL128L ^{mut}
	Native; wildtype	Merlin-UL128L ^{wt} /RV-Me-UL128L ^{wt}
	Strain 3301	Merlin-UL128L ³³⁰¹ /RV-Me-UL128L ³³⁰¹
	TB40-BAC4	Merlin-UL128L ^{TB40} /RV-Me-UL128L ^{TB40}
	Native + UL128 intron 1 G>T	Merlin-UL128L ^{G>T} /RV-Me-UL128L ^{G>T}

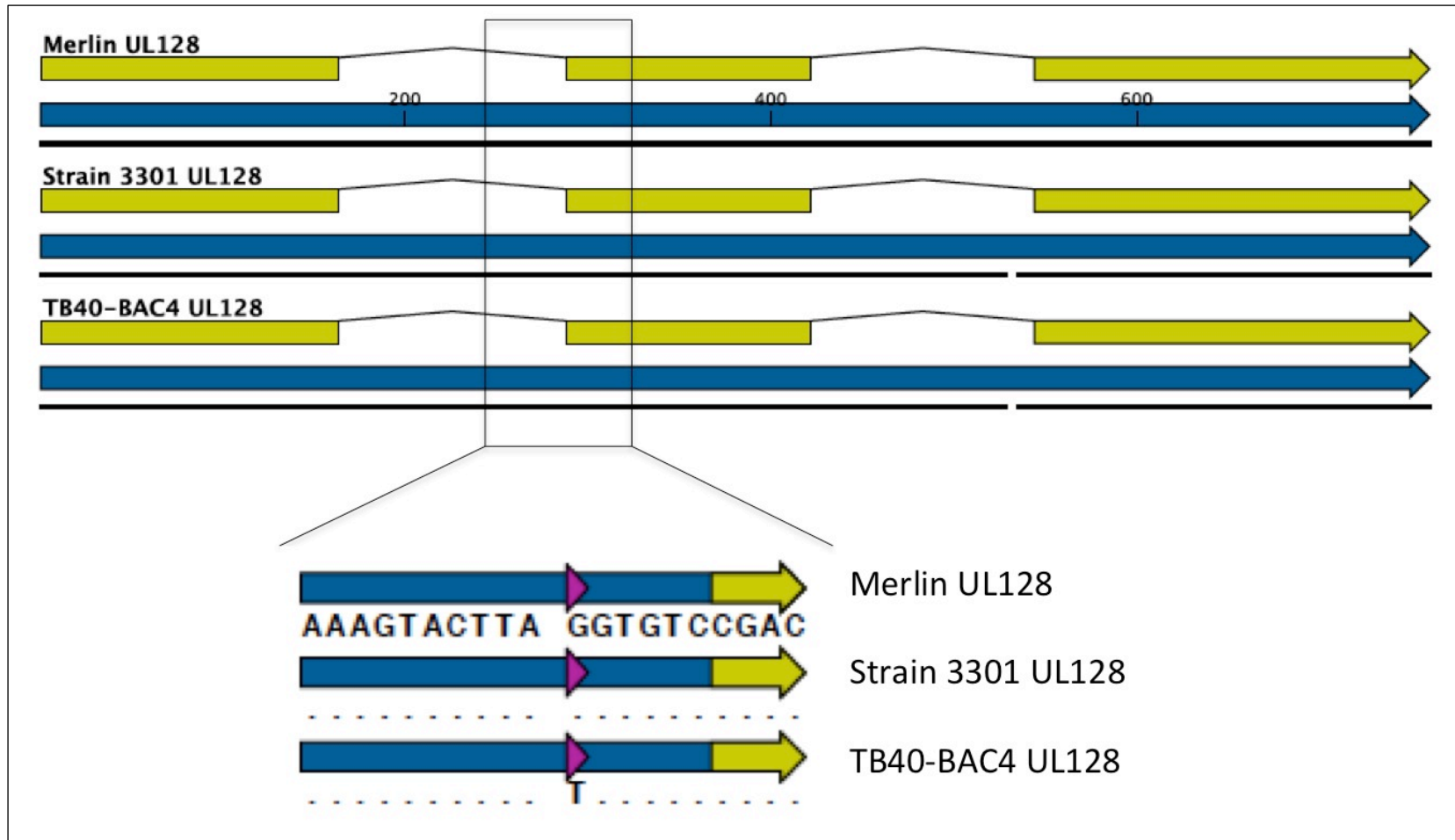


Figure 3.7. Unique nucleotide identified in TB40-BAC4 UL128 intron 1. Alignment of UL128 ORFs from Merlin, strain 3301, and TB40-BAC4. ORFs depicted in 5'-' orientation, inverted to presentation in the HCMV genome. A solid black line depicts the continuous nucleotide sequence. Entire ORFs (blue); coding region (yellow); introns (black arch). Nucleotide positions are given relative to the Merlin BAC sequence (GU179001). Dots indicate identical residues in comparison sequences. Purple arrow highlights unique TB40-BAC4 UL128 G>T variation (nt 282 in Merlin UL128).

3.6.1 Impact of TB40-BAC4 UL128 G>T variation in fibroblasts

3.6.1.1 Cell-to-cell spread

RV-Me-UL128L³³⁰¹ and RV-Me-UL128L^{wt} produced HFFF plaques with comparable size (Figure 3.8.A), demonstrating that wildtype UL128L ORFs of strains 3301 and Merlin impeded the cell-to-cell spread of HCMV in fibroblasts by the same degree, regardless of natural strain-strain variations. The larger HFFF plaques formed by RV-Me-UL128L^{TB40} compared to RV-Me-UL128L³³⁰¹ ($p < 0.01$) confirmed that the UL128L nucleotide variations common to TB40-BAC4 and strain 3301 were not responsible for the more efficient cell-to-cell spread of Merlin virus following transfer of TB40-BAC4 UL128L ORFs. Intriguingly, RV-Me-UL128L^{G>T} produced HFFF plaques with comparable size to those formed by RV-Me-UL128L^{TB40}, hence the increased cell-to-cell spread in fibroblasts displayed by Merlin virus following transfer of TB40-BAC4 UL128L was recapitulated by transfer of the UL128 G>T variation alone.

3.6.1.2 Dissemination kinetics and cell-free virus production

RV-Me-UL128L³³⁰¹ disseminated throughout HFFF monolayers with identical kinetics compared to RV-Me-UL128L^{wt}, both being dramatically slower than RV-Me-UL128L^{mut} (Figure 3.8.B). Thus, the wildtype UL128L ORFs of strain 3301 also impeded the dissemination of HCMV in fibroblast by a similar degree as Merlin wildtype UL128L ORFs. RV-Me-UL128L^{TB40} again spread throughout an HFFF monolayer with kinetics intermediate to that of the Merlin viruses containing wildtype or non-intact UL128L ORFs, and crucially, faster than RV-Me-UL128L³³⁰¹. Furthermore, the dissemination kinetics of RV-Me-UL128L^{TB40} were entirely re-capitulated by RV-Me-UL128L^{G>T}.

The peak titre of cell-free virus produced in the RV-Me-UL128L^{wt} infection (Figure 3.8.C) was greater than seen in previous assays – only 10-fold lower than those produced by RV-Me-UL128L^{mut}, whereas previously the difference had been ~1000-fold (see Figure 3.1.C). The RV-Me-UL128L³³⁰¹ infection also produced surprisingly high titres of cell-free virus. In later investigations (see section 5.3), RV-Me-UL128L^{wt} and RV-Me-UL128L³³⁰¹ viruses derived from these infections were shown to have acquired ablating UL128L genome mutations. Based on this, the cell-free virus production displayed by RV-Me-UL128L^{wt} and RV-Me-UL128L³³⁰¹ infections were deemed un-suitable for comparison. However, both RV-Me-UL128L^{TB40} and RV-Me-UL128L^{G>T} displayed comparable cell-free virus production in fibroblasts, each producing peak titres of cell-free virus that were ~30-fold and ~20-fold lower than produced by RV-Me-UL128L^{mut}, respectively.

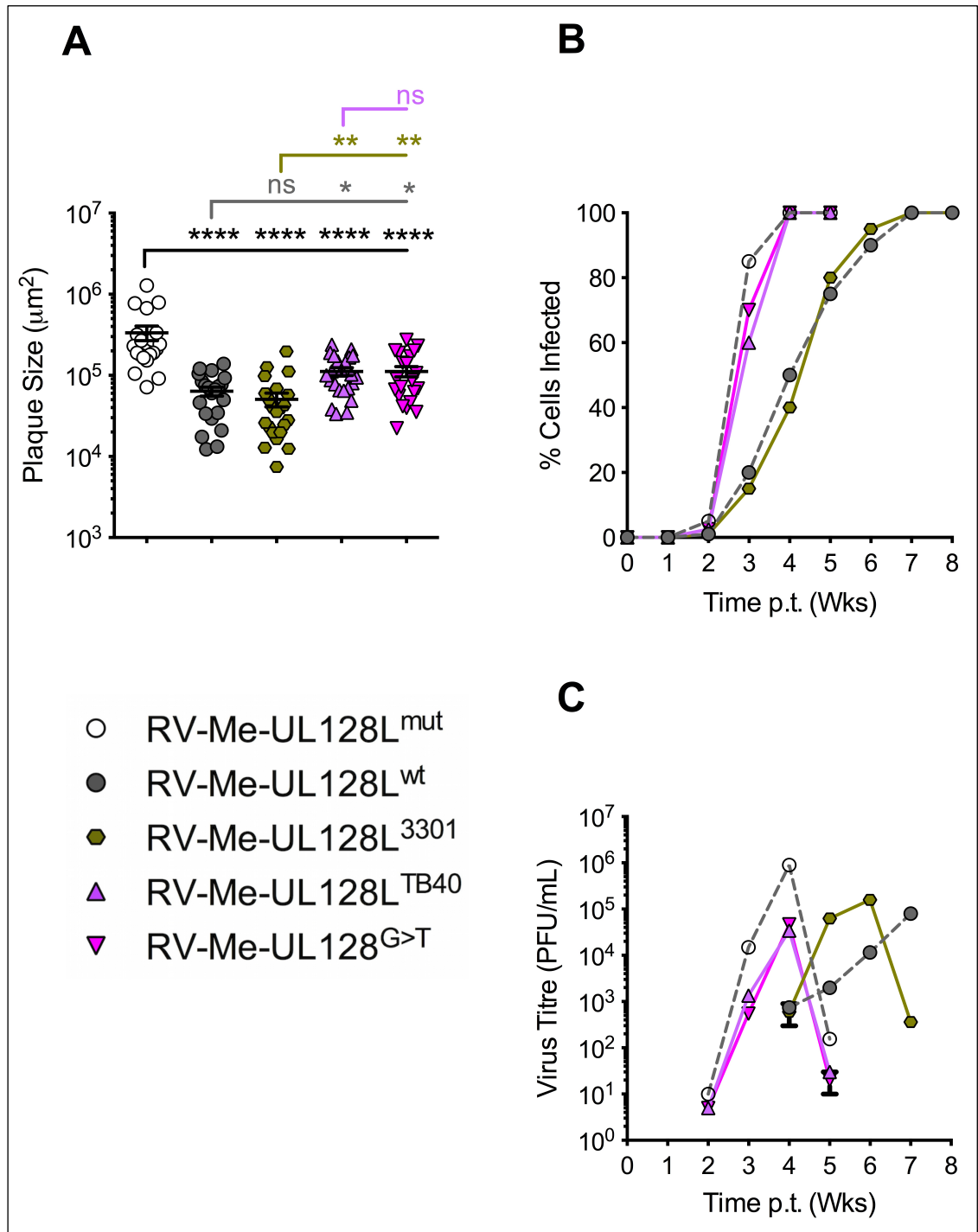


Figure 3.8. Growth characteristics of virus containing TB40-BAC4 UL128L or strain 3301 UL128L nucleotides in fibroblasts. (A) HFFF cells were transfected with BAC DNA for the indicated viruses and incubated under semisolid overlay, with plaque sizes measured 2 weeks post-transfection (p.t.). Where possible, a minimum of 20 plaques from each infection was measured. All data from 1 experiment ($n = 1$). Means and \pm SEM are shown. (Where indicated, the mean plaque sizes formed by RV-Me-UL128L^{wt}, RV-Me-UL128L³³⁰¹, RV-Me-UL128L^{TB40}, RV-Me-UL128L^{G>T} were compared to those formed by RV-Me-UL128L^{mut} (black line), RV-Me-UL128L^{wt} (grey line), RV-Me-UL128L³³⁰¹ (gold line), and/or RV-Me-UL128L^{TB40} (purple line) by 1-way analysis of variance [ANOVA] followed by Dunnett's posttest: ns – non-significant, * $p < 0.05$, ** $p < 0.01$, *** $p < .001$, **** $p < 0.0001$). (B) HFFF cells were transfected with BAC DNA for the indicated viruses, and the dissemination of infection was allowed to progress until the monolayer was destroyed. At weekly time-points post-transfection (p.t.), the level of infection was estimated by fluorescence microscopy analysis of eGFP-expressing cells. (C) Supernatant samples collected at weekly time-points from the infections shown in panel B were titrated on HFFFs to provide a measure of cell-free virus release throughout the time-course of each infection.

3.6.2 Impact of TB40-BAC4 UL128 intron 1 G>T in epithelial cells

3.6.2.1 Cell-to-cell spread

A similar, but reciprocal, pattern of findings as seen in fibroblasts was observed following comparison of RPE-1 plaque sizes formed by the Merlin variants containing UL128L sequences from strain 3301 and TB40-BAC4 (Figure 3.9.A). Importantly, RV-Me-UL128L³³⁰¹ formed RPE-1 plaques with comparable size to those formed by RV-Me-UL128L^{wt}, further demonstrating that natural strain-strain variations had little impact on the growth phenotype influence of wildtype UL128L ORFs. Similar to previous findings, RV-Me-UL128L^{TB40} again formed multi-cellular RPE-1 plaques that were significantly larger than those formed by RV-Me-UL128L^{mut} ($p < 0.0001$), but smaller than those formed by RV-Me-UL128L^{wt} ($p < 0.05$). Furthermore, RPE-1 plaques formed by RV-Me-UL128L^{TB40} were also significantly smaller than those formed by RV-Me-UL128L³³⁰¹ ($p < 0.05$). RV-Me-UL128L^{G>T} formed RPE-1 plaques with similar size to those formed by RV-Me-UL128L^{TB40}, and these too were smaller than those formed by RV-Me-UL128L^{wt} or RV-Me-UL128L³³⁰¹ (both $p < 0.05$). This pattern of results indicated that natural variations common to the UL128L genome regions of TB40-BAC4 and strain 3301 did not account for the reduced capacity for cell-to-cell spread in epithelial cells displayed by RV-Me-UL128L^{TB40}, whilst the UL128 G>T variation unique to TB40-BAC4 was responsible.

3.6.2.2 Dissemination kinetics and cell-free virus production

Each of RV-Me-UL128L^{wt}, RV-Me-UL128L³³⁰¹, RV-Me-UL128L^{TB40} and RV-Me-UL128L^{G>T} achieved complete infection of epithelial cell cultures within identical time frames (Figure 3.9.B). Thus sequence variations amongst the UL128L ORFs in the Merlin-BAC and TB40-BAC4 clones or in the UL128L ORFs in strain 3301, whether natural or representative of in vitro acquired mutations, had no apparent differential impact on the dissemination of Merlin virus in epithelial cell culture.

Like RV-Me-UL128L^{wt}, RV-Me-UL128L³³⁰¹ produced limited titres of cell-free virus; the peak titres produced by these infections differed by only ~ 2-fold (Figure 3.9.C). However, RV-Me-UL128L^{TB40} produced peak titres of cell-free virus that were ~45-fold and ~20-fold greater than that produced by RV-Me-UL128L^{wt} and RV-Me-UL128L³³⁰¹, respectively, and this was entirely re-capitulated by RV-Me-UL128L^{G>T}.

Thus, similar to the impact on the cell-to-cell spread of HCMV in epithelial cell culture, the TB40-BAC4 UL128L genome region and the single UL128 G>T variation also had an identical impact on the dissemination and cell-free virus production of HCMV in this cell-type.

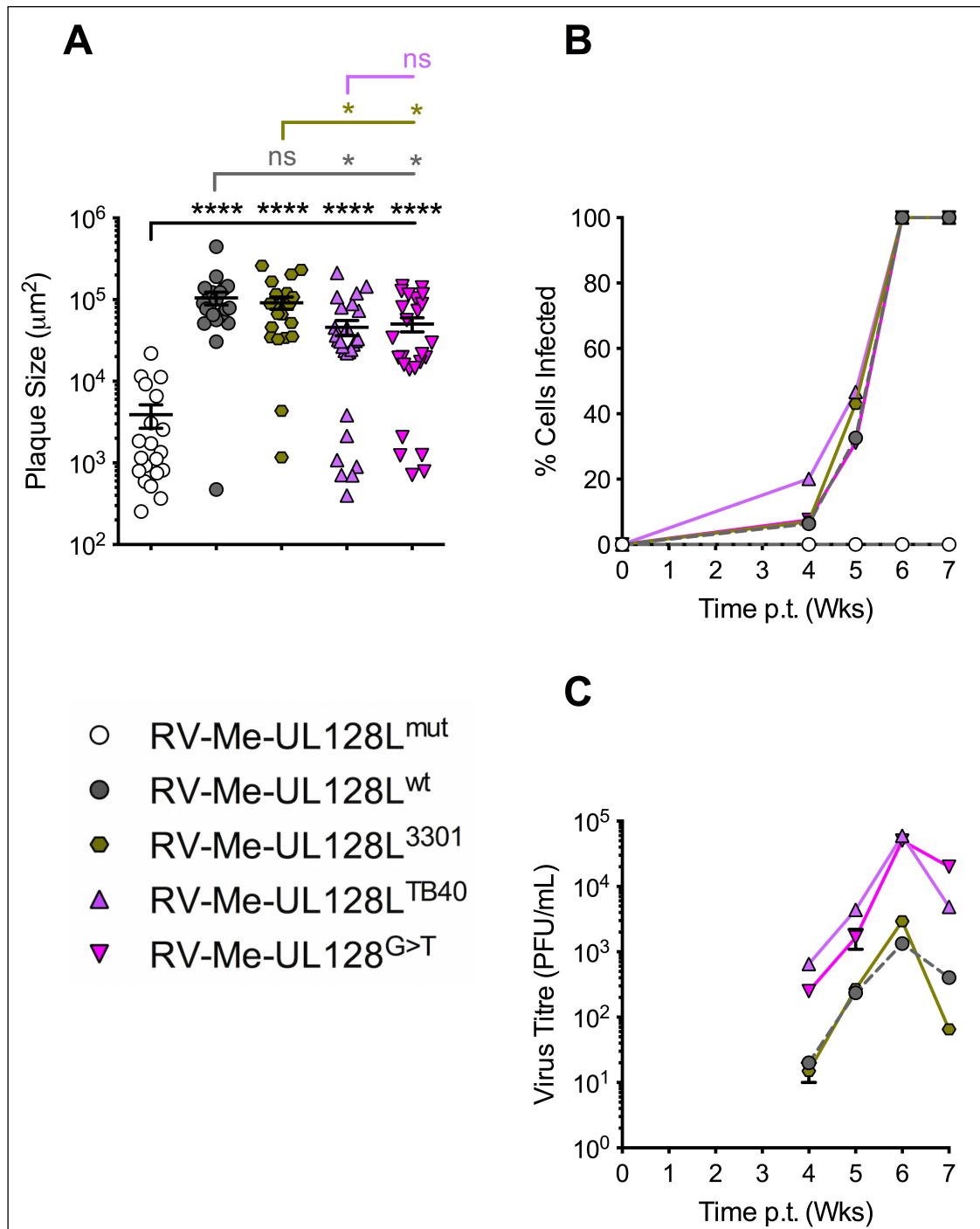


Figure 3.9. Growth characteristics of virus containing TB40-BAC4 UL128 intron 1 G>T unique nucleotide in epithelial cells. (A) RPE-1 cells were transfected with BAC DNA for the indicated viruses and incubated under semisolid overlay, with plaque sizes measured 3 weeks post-transfection (p.t.). Means and \pm SEM are shown. All data from 1 experiment ($n = 1$). Means and \pm SEM are shown. (Where indicated, the mean plaque sizes formed by RV-Me-UL128^{wt}, RV-Me-UL128³³⁰¹, RV-Me-UL128^{TB40}, RV-Me-UL128^{G>T} were compared to those formed by RV-Me-UL128^{mut} (black line), RV-Me-UL128^{wt} (grey line), RV-Me-UL128³³⁰¹ (gold line), and/or RV-Me-UL128^{TB40} (purple line) by 1-way analysis of variance [ANOVA] followed by Dunnett's posttest: ns – non-significant, * $p < 0.05$, ** $p < 0.01$, *** $p < .001$, **** $p < 0.0001$). (B) RPE-1 cells were transfected with BAC DNA for the indicated viruses, and the dissemination of infection was allowed to progress until the monolayer was destroyed. At weekly time points post transfection (p.t.), the level of infection was estimated by fluorescence microscopy analysis of eGFP-expressing cells. (C) Supernatant samples collected at weekly time-points from the infections shown in panel B were titrated on HFFs to provide a measure of cell-free virus release throughout the time-course of each infection.

3.6.3 Validation of TB40-BAC4 UL128 intron 1 G>T growth phenotype impact

Again, the relative HFFF plaque sizes formed immediately following transfection with BAC DNA for each virus (p0) (Figure 3.10.A) were compared to those formed during titration of progeny virus produced in fibroblasts (HFFFp1) (Figure 3.10.B), or in epithelial cell infections (RPE-1p1) (Figure 3.10.C). As mentioned, the titres of cell-free virus produced during the RV-Me-UL128L^{wt} and RV-Me-UL128L³³⁰¹ infections in HFFFs suggested that these viruses may have undergone adaptation early following reconstitution, and the sizes of HFFF plaques formed by RV-Me-UL128L^{wt}-HFFFp1 and RV-Me-UL128L³³⁰¹-HFFFp1 viruses reflected this; HFFF plaques formed immediately following transfection with Merlin-UL128L^{wt} and Merlin-UL128L³³⁰¹ BAC DNA were smaller than those formed immediately following transfection with Merlin-UL128L^{TB40} and Merlin-UL128L^{G>T} BAC DNA, whilst the HFFFp1 progeny from the RV-Me-UL128L^{wt} and RV-Me-UL128L³³⁰¹ infections formed larger HFF plaques than those formed HFFFp1 progeny from the RV-Me-UL128L^{TB40} and RV-Me-UL128L^{G>T} infections. However, comparison to the HFFF plaque sizes formed by RV-Me-UL128L^{mut}-HFFFp1 gave no indications that either of the Merlin variants containing TB40-BAC4 UL128L sequence had undergone adaptation in fibroblasts. In addition, there was no evidence that any virus had undergone adaptation during passage in epithelial cells.

3.6.4 Conclusion

The growth phenotype impact of UL128L ORFs with alternative, but wildtype sequences was identical, and apparently not subject to natural strain-strain variation. This was apparent in both fibroblasts, where UL128L is inhibitory to virus growth, and also in epithelial cells, where virus requires UL128L for efficient infection. The different growth characteristics of RV-Me-UL128L^{TB40} compared to RV-Me-UL128L³³⁰¹, together with the recapitulation of the RV-Me-UL128L^{TB40} growth phenotype by RV-Me-UL128L^{G>T}, indicated that the G>T nucleotide variation identified in intron 1 of the TB40-BAC4 UL128 ORF alone, and not natural strain-strain variations, was the defining difference between TB40-BAC4 UL128L and Merlin wildtype UL128L genome regions.

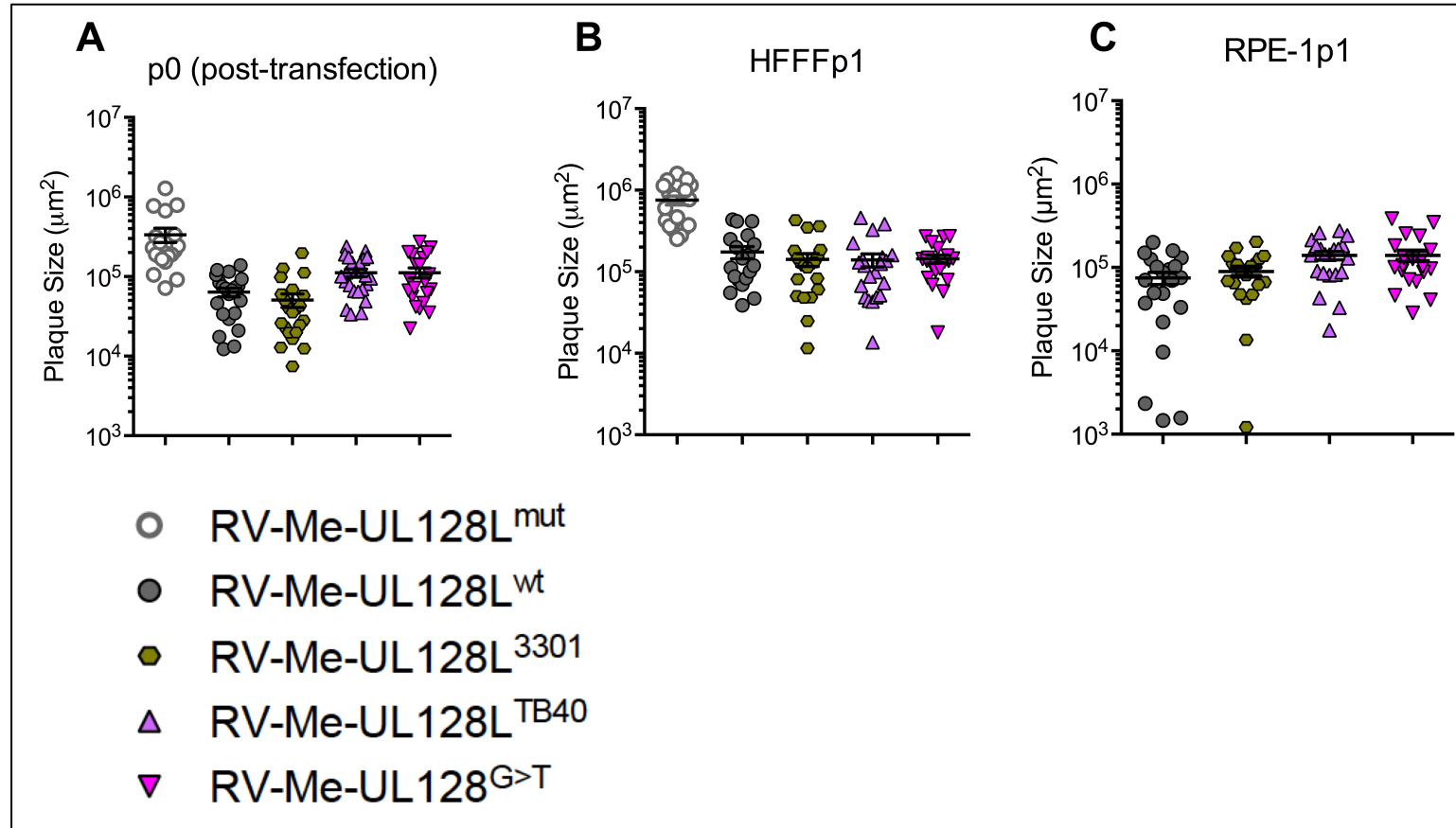


Figure 3.10. Fibroblast cell-to-cell spread of recombinant Merlin viruses containing TB40-BAC4 or strain 3301 UL128L nucleotides, before and after passage in fibroblast and epithelial cells. (A) Plaque sizes \pm SEM formed 2 weeks post-transfection (p.t) of HFFFs with BAC DNA for each indicated virus and incubation under overlay. (B) HFFF plaque sizes \pm SEM formed during titration of virus in supernatant samples collected at final time-point in growth kinetics assay performed in HFFFs, at 2 weeks post-infection (p.i.) and incubation under overlay. (C) HFFF plaque sizes \pm SEM of formed during titration of virus in supernatant samples collected at final time-point in growth kinetics assay performed in HFFFs, at 2 weeks post-infection (p.i.) and incubation under overlay.

3.7 Nucleotides contributing to the growth phenotype of viruses from FIX-BAC

FIX UL128 contained a single non-synonymous nucleotide variation compared to Merlin UL128, however the encoded D5N amino acid substitution was also identified in the sequences of several other distinct strains deposited in GenBank. FIX UL131A differed from that in Merlin at 4 nucleotide positions, each being synonymous variations and not encoding for alternative amino acids. FIX UL130 contained 11 nucleotide variations, four of which were non-synonymous. One amino acid variation (L69S) was common to several other strains, and potentially represented natural strain-strain variation. However, three amino acid variations were less common. The S29F amino acid substitution was identified in sequences from strain VR1814 (GU179289.1) (FIX-BAC parental strain), and also from ‘Isolate 20M’ (DQ22943.1). The 2 more exclusive nucleotide sequence features were A>G nucleotide variations at nt 140 and nt 214 in FIX UL130, coding for S47P and S72P amino acid substitutions respectively. The S47P amino acid substitution is also present in FIX parental strain VR1814, whilst the S72P amino acid substitution was entirely unique to FIX (Figure 3.11). Thus, the S72P variation was selected for further investigation and transferred into the Merlin genome to generate Merlin-UL130^{A>G}. The contribution of the unique A>G (S72P) nucleotide variation was investigated by determining the growth phenotype of Merlin virus engineered to contain this nucleotide variation in direct comparison to Merlin virus expressing the entire FIX UL128L, as well as to Merlin viruses containing native UL128L sequences (wildtype or non-intact) (Table 3.5).

Table 3.5. Merlin BACs/Viruses containing native or FIX UL128L sequence features.

Genome Background	UL128L Origin / Nucleotide feature	Designation in text: BAC /Reconstituted Virus
Merlin	Native, mutated	Merlin-UL128L ^{mut} /RV-Me-UL128L ^{mut}
Merlin	Native; wildtype	Merlin-UL128L ^{wt} /RV-Me-UL128L ^{wt}
Merlin	FIX	Merlin-UL128L ^{FIX} /RV-Me-UL128L ^{FIX}
Merlin	Native; + FIX UL130 A>G	Merlin-UL130 ^{A>G} /RV-Me-UL130 ^{A>G}

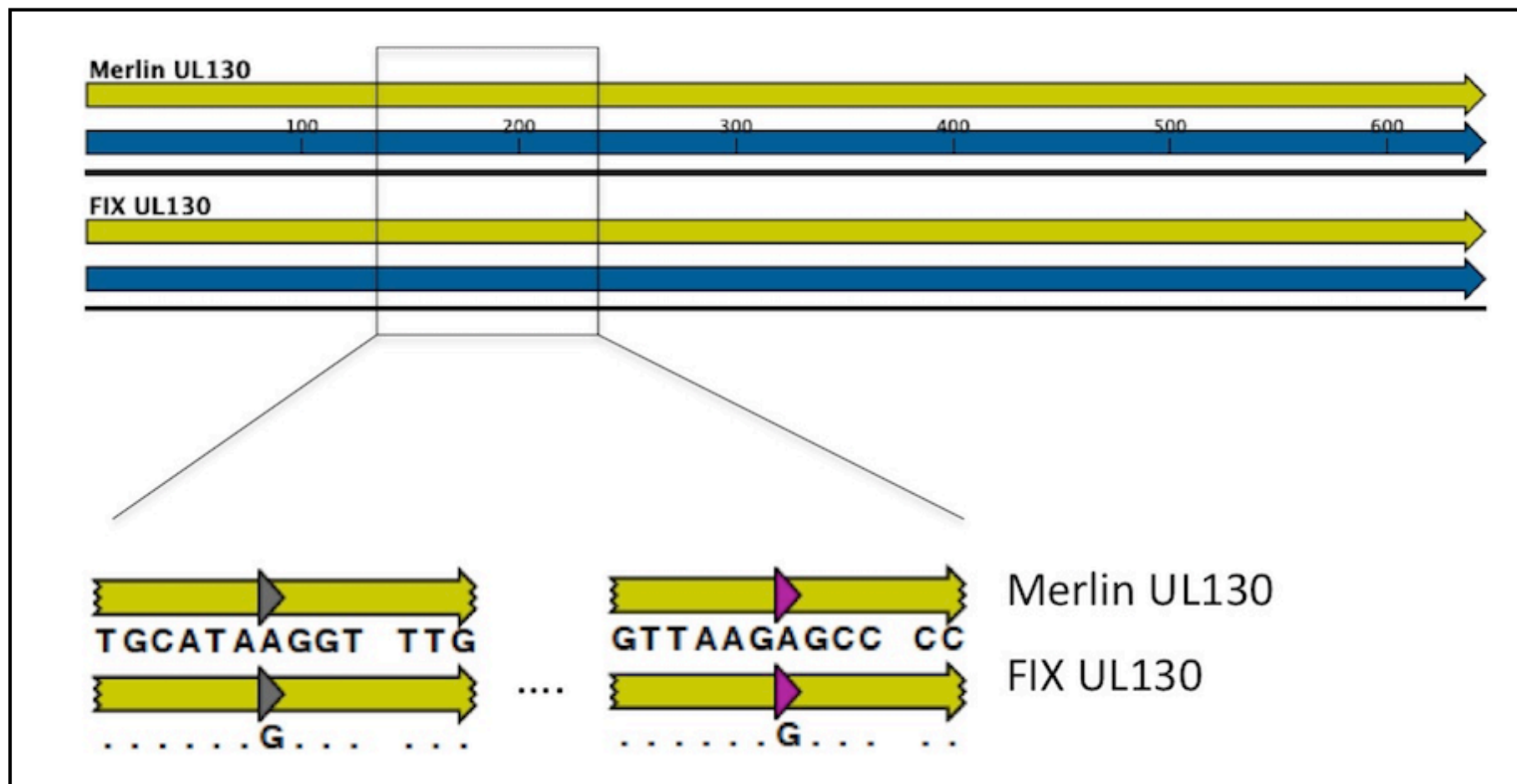


Figure 3.11. Unique nucleotide sequence features identified in FIX UL130. UL130 ORFs from Merlin and FIX, in inverse orientation as to how presented in the HCMV genome. A solid black line depicts the continuous nucleotide sequence. Entire ORFs (blue); coding region (yellow). Nucleotide positions are given relative to Merlin UL130 sequences. Dots indicate identical residues in comparison sequences. Grey arrow highlights A>G (S47P) variation identified only in FIX and VR1814 at nt 139. Pink arrow highlights A>G (S72P) nucleotide variation unique to FIX, at nt 214 (pink).

3.7.1 Impact of FIX UL130 A>G In Fibroblasts

3.7.1.1 Cell-to-cell spread

As previously demonstrated, RV-Me-UL128L^{FIX} again formed HFFF plaques of intermediate size compare to those from RV-Me-UL128L^{mut} ($p < 0.05$) and RV-Me-UL128L^{wt} (Figure 3.12.A); the fact this latter difference was deemed non-significant by statistical analysis likely reflects that fact that relatively few RV-Me-UL128L^{FIX} plaques were measured. The UL130 A>G nucleotide variation was conclusively demonstrated to be a contributing feature of the less inhibitory UL128L genome region in the FIX-BAC clone, as RV-Me-UL130^{A>G} formed HFFF plaques with comparable size to those produced by RV-Me-UL128L^{FIX}, and significantly larger than those formed by RV-Me-UL128L^{wt} ($p < 0.001$).

3.7.1.2 Dissemination kinetics and cell-free virus production

RV-Me-UL128L^{FIX} also displayed more rapid dissemination kinetics in fibroblast cultures compared to RV-Me-UL128L^{wt} as had previously been observed, and completed infection of an HFFF monolayer was achieved within an a time-frame intermediate to the Merlin viruses containing non-intact and wildtype UL128L ORFs (Figure 3.12.B). These dissemination kinetics were closely matched by RV-Me-UL130^{A>G} that achieved complete infection of the HFFF monolayer 1 week later than RV-Me-UL128L^{FIX}. Thus transfer of the UL130 A>G nucleotide variation also alleviated the inhibitory impact of the UL128L genome for the dissemination of HCMV in fibroblast culture.

As previously stated, RV-Me-UL128L^{wt} potentially mutated during these assays and produced titres of cell-free virus that were deemed unsuitable for comparison. However, comparison to the peak titres produced by RV-Me-UL128L^{mut} provided evidence to suggest that the FIX UL128L genome region did not impede cell-free virus production by the same degree as previously demonstrated by Merlin wildtype UL128L ORFs, and the UL130 A>G variation was a contributing factor (Figure 3.12.C). However, the peak titre of cell-free virus produced by RV-Me-UL128L^{FIX} was ~13-fold lower than that from RV-Me-UL128L^{mut}, whilst the peak titre of cell-free virus produced by RV-Me-UL130^{A>G} was ~60-fold lower. Hence, RV-Me-UL130^{A>G} produced peak titres ~6-fold lower than those produced by RV-Me-UL128L^{FIX}, suggesting that the unique UL130 A>G nucleotide variation alone was not responsible for the increased cell-free virus production achieved by expression of FIX UL128L ORFs in the Merlin virus.

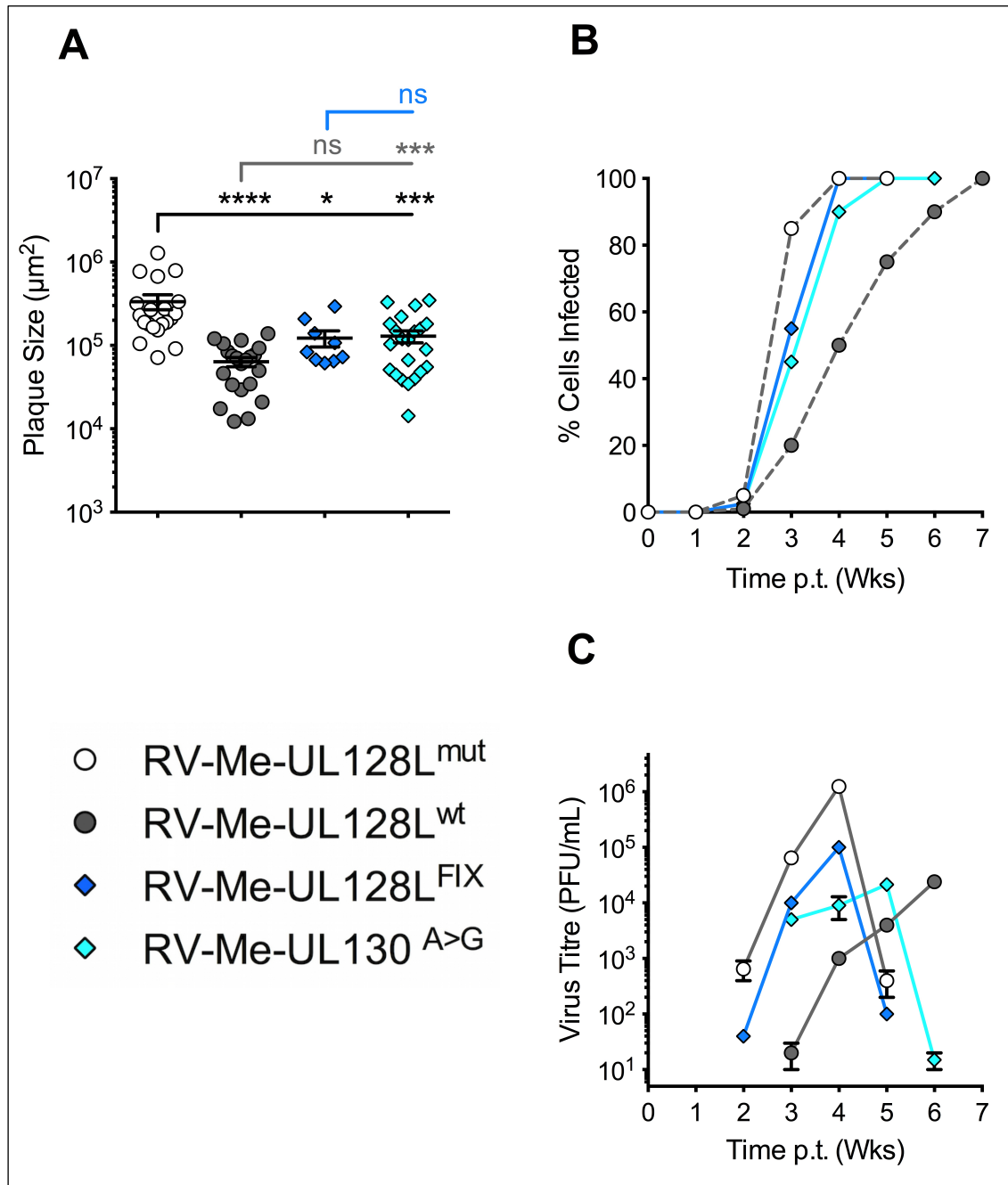


Figure 3.12. Growth characteristics of recombinant Merlin viruses containing FIX UL128L nucleotides in fibroblasts. (A) HFFF cells were transfected with BAC DNA for the indicated viruses and incubated under semisolid overlay, with plaque sizes measured 2 weeks post transfection. Where possible, a minimum of 20 plaques from each infection was counted; an exception to this was the RV-Me-UL128L^{FIX} infection where only 10 plaques could be measured accurately. All data from 1 experiment (n = 1). Means and \pm SEM are shown. (Where indicated, the mean plaque sizes formed by RV-Me-UL128L^{wt}, RV-Me-UL128L^{FIX} and RV-Me-UL130^{A>G} were compared to those formed by RV-Me-UL128L^{mut} (black line), RV-Me-UL128L^{wt} (grey line) and/or RV-Me-UL130^{A>G} by 1-way analysis of variance [ANOVA] followed by Dunnett's posttest: ns – non-significant, * $p < 0.05$, ** $p < 0.01$, *** $p < .001$, **** $p < 0.0001$). (B) HFFF cells were transfected with BAC DNA for the indicated viruses, and the dissemination of infection was allowed to progress until the monolayer was destroyed. At weekly time points post transfection (p.t.), the level of infection was estimated by fluorescence microscopy analysis of eGFP-expressing cells. (C) Supernatant samples collected at weekly time points from the infections shown in panel B were titrated on HFFFs to provide a measure of cell-free virus release throughout the time-course of each infection.

3.7.2 Impact of FIX UL130 A>G in epithelial cells

3.7.2.1 Cell-to-cell spread

The specific impact of FIX UL128L sequences on the cell-to-cell spread of virus in epithelial cells was also determined (Figure 3.13.A). Unlike previous findings, RV-Me-UL128L^{FIX} formed RPE-1 plaques that were significantly larger than those formed by RV-Me-UL128L^{mut} ($p < 0.0001$); previously, RV-Me-UL128L^{FIX} had formed RPE-1 plaques that were comparable in size to those formed by RV-Me-UL128L^{mut} (Appendix II). Nonetheless, the RPE-1 plaques formed by RV-Me-UL128L^{FIX} in this assay were again significantly smaller than those formed by RV-Me-UL128L^{wt} ($p < 0.0001$), and thus expression of FIX UL128L ORFs did reduce the capacity of Merlin virus for cell-to-cell spread in epithelial cell culture. RV-Merlin-UL130^{A>G} also formed RPE-1 plaques that were significantly larger than those formed by RV-Me-UL128L^{mut} ($p < 0.0001$), yet significantly smaller than those formed by RV-Me-UL128L^{wt} ($p < 0.0001$). In fact, RV-Merlin-UL130^{A>G} spread by the cell-to-cell route in epithelial cells with similarly reduced efficiency as did the Merlin virus containing the entire FIX UL128L genome region.

3.7.2.2 Dissemination kinetics and cell-free virus production

Monitoring the dissemination and cell-free virus production of Merlin variants containing either native UL128L or FIX UL128L nucleotides in epithelial cells further revealed the impact of the FIX UL130 A>G variation to the growth properties of virus in this cell type (3.13.B & C). RV-Me-UL128L^{FIX} again displayed reduced dissemination kinetics in RPE-1s compared to RV-Me-UL128L^{wt}, but the delayed dissemination was more prominent than in previous assays; complete infection of an RPE-1 monolayer was achieved by 12 weeks post-transfection in this assay, whereas in previous assays complete infection had been achieved within 9 weeks post-transfection. RV-Me-UL130^{A>G} also spread throughout the RPE-1 monolayer slower than RV-Me-UL128L^{wt}, however this infection disseminated considerably faster than RV-Me-UL128L^{FIX}.

The RV-Me-UL128L^{FIX} infection produced peak titres of cell-free virus that were >150-fold greater than that produced by RV-Me-UL128L^{wt}, whereas the peak titre of cell-free virus produced by the RV-Me-UL130^{A>G} infection were only ~40-fold greater than that produced by RV-Me-UL128L^{wt}.

Ultimately, transfer of the UL130 A>G variation alone impeded the dissemination of Merlin virus yet increased the cell-free virus production in epithelial cell culture, though each of these effects were less dramatic than transfer of the entire FIX-BAC UL128L genome region.

3.7.3 Validation of FIX UL130 A>G growth phenotype impact

Comparison of the relative plaques sizes formed following transfection of HFFFs with BAC DNA for each virus (Figure 3.14.A), to those formed during titration of progeny virus produced in fibroblasts (HFFFp1) (Figure 3.14.B), or in epithelial cell infections (RPE-1p1) (Figure 3.14.C), gave no indication to suggest recombinant Merlin viruses containing FIX UL128L sequence features had undergone adaptation in the growth kinetics assays. The apparent adaptation of RV-Me-UL128L^{wt} in fibroblasts was reflected by the HFFF plaques formed by the RV-Me-UL128L^{wt}-HFFFp1 progeny virus (as described in section 3.6.3.).

3.7.4 Conclusion

The unique A>G nucleotide variation identified in FIX UL130 contributed to the differential impact of FIX UL128L compared to wildtype UL128L, in both fibroblasts and epithelial cell culture. However, the growth characteristics of the Merlin virus expressing the entire FIX UL128L genome region (RV-Me-UL128L^{FIX}) were not entirely re-capitulated by the Merlin virus engineered to contain the UL130 A>G variation alone (RV-Me-UL130^{A>G}) in either cell-type, most notably the increased cell-free virus production. Therefore, other nucleotide features must also contribute to distinct growth phenotype influence of the FIX-BAC UL128L genome region.

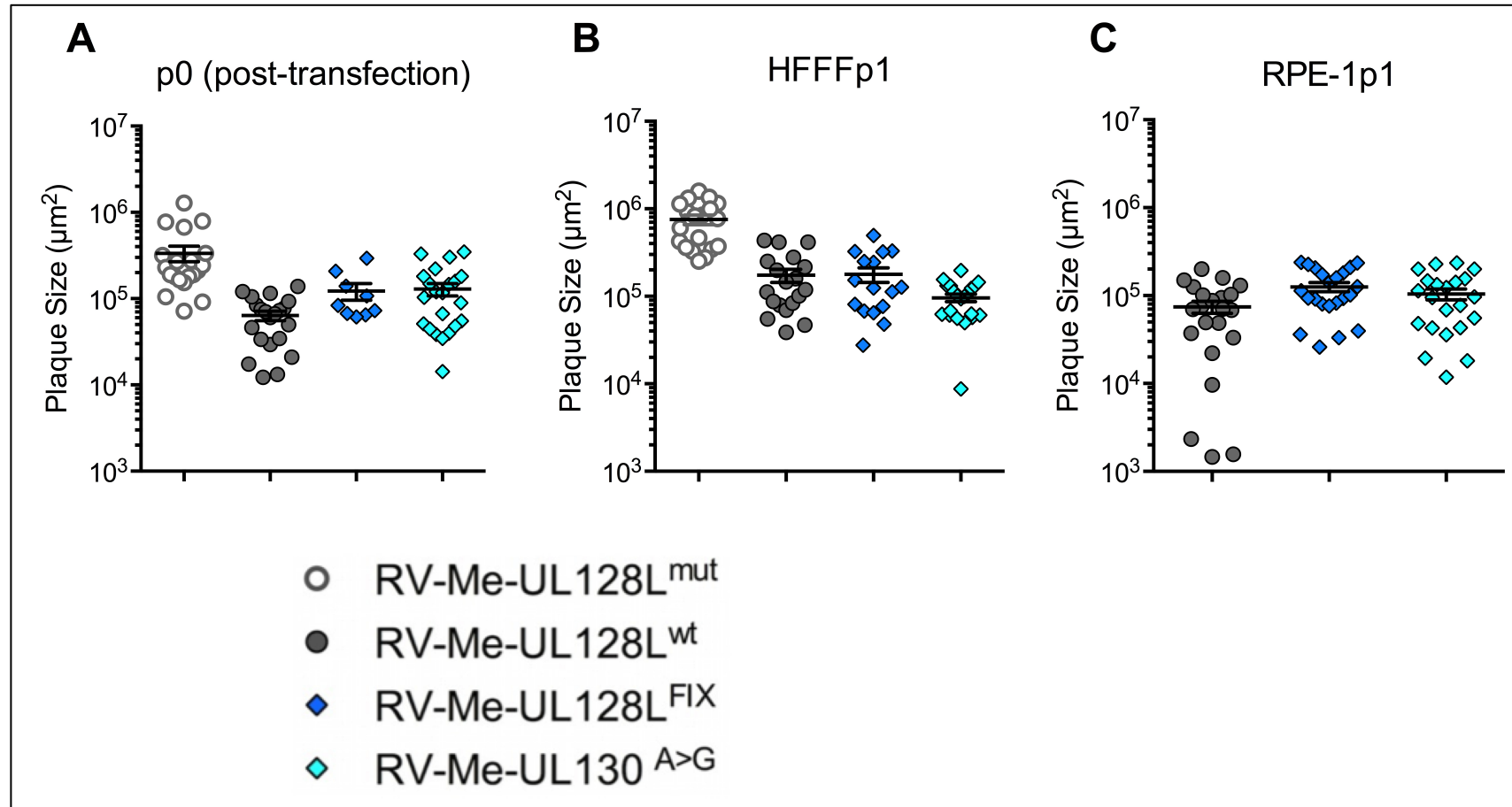


Figure 3.14. Fibroblast cell-to-cell spread of recombinant Merlin viruses containing TB40-BAC4 or strain 3301 UL128L nucleotides, before and after passage in fibroblast and epithelial cells. (A) Plaque sizes \pm SEM formed 2 weeks post-transfection (p.t) of HFFFs with BAC DNA for each indicated virus and incubation under overlay. (B) HFFF plaque sizes \pm SEM formed during titration of virus in supernatant samples collected at final time-point in growth kinetics assay performed in HFFFs, at 2 weeks post-infection (p.i.) and incubation under overlay. (C) HFFF plaque sizes \pm SEM of formed during titration of virus in supernatant samples collected at final time-point in growth kinetics assay performed in HFFFs, at 2 weeks post-infection (p.i.) and incubation under overlay.

3.8 Summary

The inhibitory impact of Merlin wildtype UL128L ORFs on the growth of virus in fibroblasts was repeatedly demonstrated in 4/4 repeats of each assay: virus from Merlin-UL128L^{mut} BAC consistently formed significantly larger HFFF plaques than virus from the Merlin-UL128L^{wt} BAC; virus from the Merlin-UL128L^{mut} BAC achieved complete infection of the HFFF monolayer around twice as fast as virus from the Merlin-UL128L^{wt} BAC; and virus from the Merlin-UL128L^{mut} consistently produced titres that were ~1000-fold greater than that from infections with virus from the Merlin-UL128L^{wt} BAC (except in assays where this virus had apparently undergone adaptation to fibroblast cell culture). Similar consistency was observed for the growth phenotype of viruses from the Merlin-UL128L^{wt} and Merlin-UL128L^{mut} BAC variants in epithelial cell cultures.

The growth phenotypes of viruses derived from the BAC-cloned versions of strains TR (TR-BAC), TB40E (TB40-BAC4), and VR1814 (FIX-BAC) were distinct from that of virus from the Merlin BAC containing wildtype UL128L. The UL128L genome regions in the TB40-BAC4 and FIX-BAC clones contributed to the differential growth phenotype of virus from each corresponding strain, and unique single nucleotide variations that were responsible for this difference were identified. Importantly, Merlin viruses engineered to contain these unique nucleotides (TB40-BAC4 UL128 intron 1 G>T and FIX UL130 A>G [S72P]) remained viable in epithelial cells, albeit with slightly compromised fitness in this cell type compared to virus from the parental Merlin-BAC variant containing wildtype UL128L ORFs. Thus, these nucleotides suppressed, but did not ablate virion pentameric complex functionality. In contrast to this, the distinct growth phenotype of TR BAC-derived virus was independent of differences in UL128L; TR UL128L conveyed a growth phenotype similar to that conveyed by UL128L with wildtype sequence, and TR UL128L mutated within a time frame comparable to that seen during the passage of wildtype virus and clinical isolates in fibroblasts, yet this was only apparent when expressed from the Merlin genome background. The distinct growth phenotypes of TB40-BAC4 and FIX BAC-derived viruses were only partially recapitulated by Merlin viruses containing TB40-BAC4 and FIX UL128L sequences. Hence, similar to TR BAC-derived viruses, other genome regions of TB40-BAC4 and FIX also contributed to the distinct growth phenotypes of these viruses.

Ultimately, the novel Merlin viruses containing these nucleotides from TB40-BAC4 UL128 and FIX UL130 could be grown to greater titres than the parental Merlin virus containing wildtype UL128L, and this was apparent following infection of fibroblasts, and interestingly, also in epithelial cell infections.

4 Impact of UL128L Sequence

Variations on Pentameric

Complex Functionality

The reduced fitness in epithelial cell culture displayed by Merlin viruses engineered to contain TB40-BAC4 and FIX UL128L sequences (less efficient cell-to-cell spread and dissemination kinetics) suggested that the transferred nucleotides impacted the functionality of the pentameric complex. Thus, we next sought to determine the mechanisms, and degree, by which the nucleotide sequence variations amongst UL128L genome regions of each strain affected the pentameric complex, and also determine the specific impact of these changes on activities in the viral replication cycle that are directly or indirectly influenced.

4.1 Variant UL128L ORF expression efficiency

Bioinformatics-based predictions as to the impact of the nucleotide variation identified in TB40-BAC4 UL128 intron 1 were made by Dr Gavin Wilkie (MRC-University of Glasgow Centre for Virus Research, Glasgow, UK) (Appendix IV). This mutation was predicted to disrupt the splicing of UL128 transcripts for the generation of translatable UL128 mRNA, and this could ultimately limit the production and availability of gpUL128 for pentameric complex assembly. Alternatively, novel UL128 mRNA splice variants could be translated to produce distinct gpUL128 isoforms incorporated into functionally deficient complexes. Further nucleotide variations were also observed in intron 2 amongst each variant UL128, and also in the intron of each variant UL131A (Figure 4.1), however bioinformatics-based predictions as to the impact of these sequence variations were not made.

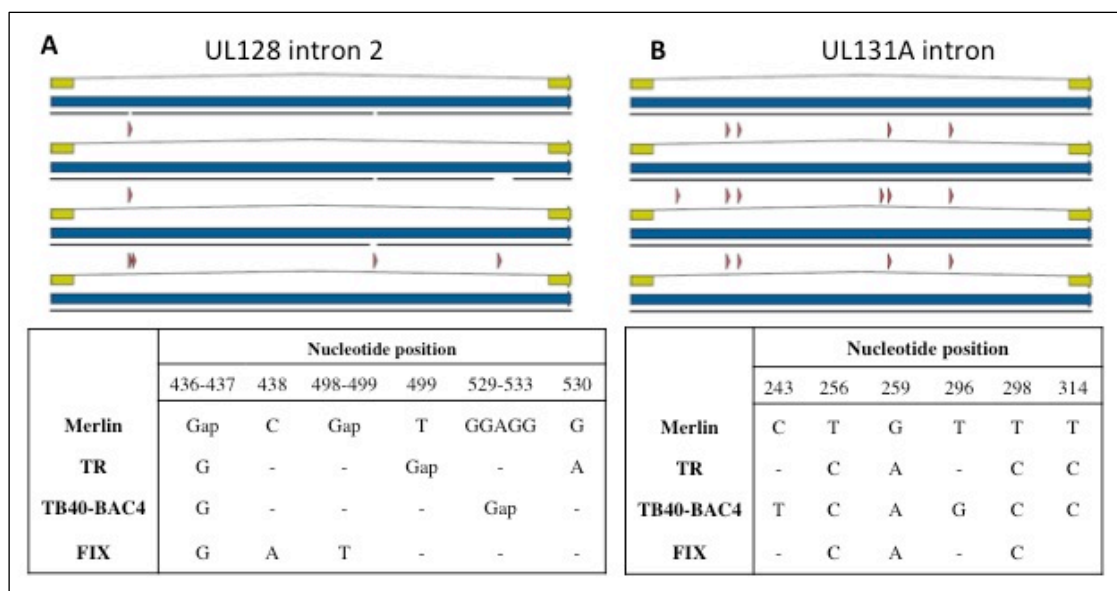


Figure 4.1. Nucleotide variations in UL128 intron 2 and UL131A intron. (A) Aligned variant UL128 intron 2. ORF (blue) and protein coding region (yellow) annotations are depicted in solid lines. The solid black line under ORF and coding domain annotations depicts the continuous nucleotide sequence. The branching line depicts intron sequence regions. Nucleotide variations are depicted in pink arrows; gaps are indicated by breaks in solid black lines. (B) Aligned variant UL131A intron. Tables - Locations of nucleotide variations are given relative nucleotide positions in Merlin UL128 and UL131A; - indicates identical residue.

To assess the impact of the UL128 G>T mutation and other UL128L ORF intronic nucleotide variations, the splice patterns of transcripts derived from each variant UL128 and UL131A ORF during infection was assessed by RT-PCR. Fibroblasts were the only cell type that supported productive infection of each virus, therefore the splice pattern of transcripts were analysed following infection of HFFFs. Details of each of virus stocks used for the initiation of productive infections are given in Table 4.1. Where possible, virus stocks used were those produced in epithelial cell culture; these stocks were chosen, as UL128L mutants were unlikely to have been selected during the growth of virus in epithelial cells. Transcripts were analysed at 72-hrs post-infection when the UL128L ORFs are transcriptionally active (Akter, Cunningham et al. 2003, Hahn, Revello et al. 2004, Sun, Ren et al. 2010).

Table 4.1. HCMV virus stocks.

BAC Origin/ Reconstituted Virus	UL128L origin/sequence features ^{a,b}	Cell type ^c
Merlin-UL128L ^{mut} / RV-Me-UL128L ^{mut}	Native; mutated: UL128 G>A at nt 176260 (R> stop)	HFFF
Merlin-UL128L ^{wt} / RV-Me-UL128L ^{wt}	Native wildtype	RPE-1
TR-BC/ RV-TR	Native	RPE-1
TB40-BAC4/ RV-TB40-BAC4	Native	RPE-1
Merlin-UL128L ^{TB40} / RV-Me-UL128L ^{TB40}	TB40-BAC4	RPE-1
Merlin-UL128 ^{G>T} / RV-Me-UL128 ^{G>T}	Merlin; wildtype + UL128 G>T at nt 176663 (intron 1)	RPE-1
FIX-BAC/ RV-FIX	Native	HFFF
Merlin-UL128L ^{FIX} / RV-Me-UL128L ^{FIX}	FIX	RPE-1
Merlin-UL130 ^{A>G} / RV-Me-UL130 ^{A>G}	Merlin; wildtype + UL130 A>G at nt 177364 (S72P)	RPE-1

^a – UL128L genome region sequence features.

^b – Nucleotide positions are given relative to sequence of BAC-cloned strain Merlin (GU179001).

^c – Cell type in which each virus stock was grown.

Variations in coding nucleotide positions, including the FIX A>G (S72P) mutation, could also impact the expression of each variant UL128L ORF; these potentially introduced codons that reduced the efficiency of ORF transcription, or reduced the efficiency by which mRNAs were translated. The expression of each variant UL128L ORF was therefore analysed at a translational level following expression from replication deficient recombinant adenovirus (RAdZ) vectors (Stanton, McSharry et al. 2008), with the quantities of the glycoproteins produced compared. To provide a standard marker for detection, all cloned UL128L ORFs were engineered to encode glycoproteins containing a C-terminal V5 epitope tag (~ 3 KDa). Details of each RAdZ vector are given in Table 4.2.

Table 4.2. RAdZ vectors containing variant UL128L-derived transgenes.

Transgene	BAC Origin	Sequence features ^a	Designation in Text/Figures ^b	Translation product
UL128	Merlin-UL128L ^{wt}	Wildtype	Me-UL128 ^{wt}	gpUL128-V5
	TR		TR-UL128	
	TB40-BAC4		TB-UL128	
	FIX		FIX-UL128	
	Merlin-UL128 ^{G>T}	G>T at nt	Me-UL128 ^{G>T}	
UL130	Merlin-UL128L ^{wt}	Wildtype	Me-UL130 ^{wt}	gpUL130-V5
	TR		TR-UL130	
	TB40-BAC4		TB-UL130	
	FIX		FIX-UL130	
	Merlin-UL130 ^{A>G}	A>G at nt 214	Me-UL130 ^{A>G}	
UL131A	Merlin-UL128L ^{wt}		Me-UL131A	gpUL131A-V5
	TR		TR-UL131A	
	TB40-BAC4		TB-UL131A	
	FIX		FIX-UL131A	

^a - Defining sequence features of variant UL128L ORF

^b - Designation of recombinant AdZ vector (RAdZ) containing variant UL128L ORF

4.1.1 TB40-BAC4 intron G>T reduces UL128 mRNA and gpUL128L production

Following RT-PCR, three amplicons corresponding to differently spliced UL128 transcripts were detected across all infections. Sequencing demonstrated that these transcripts corresponded to: i) non-spliced UL128 pre-mRNA (656-662 bp, depending on strain origin); ii) partially spliced UL128 pre-mRNA (540 bp transcript; lacking intron 2) pre-mRNAs; iii) fully spliced UL128 mRNA (417 bp transcript, lacking introns 1 and 2) (Figure 4.2.A). The quantities of non-spliced and partially spliced UL128 pre-mRNAs from each infection were equivalent, however all viruses that contained the UL128 intron G>T nucleotide variation (RV-TB40-BAC4, RV-Me-UL128^{TB40}, and RV-Me-UL128^{G>T}) produced significantly reduced amounts of fully spliced transcripts. Thus, as predicted, the intron 1 G>T nucleotide variation disrupted the splicing of exons 1 and 2. Sequencing identified additional amplicons detected in each infection as hetero-duplexes of PCR products, and not representative of splice variants.

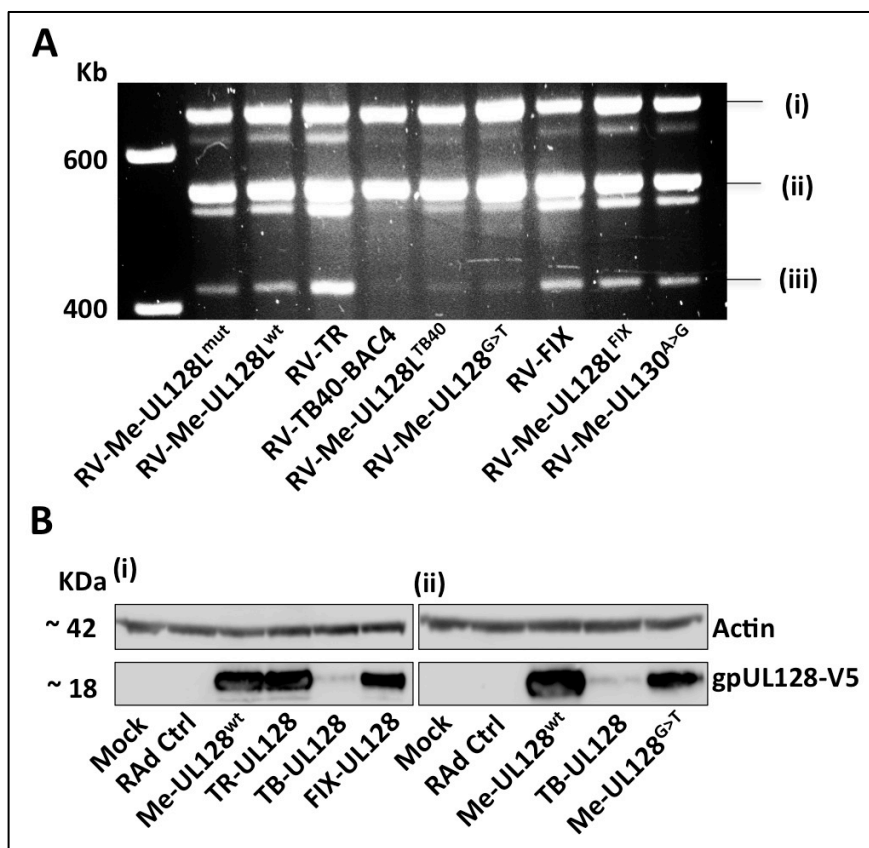


Figure 4.2. Variant UL128 Expression. (A) UL128 transcript splice patterns. HFFFs were infected with each indicated virus (MOI 3), and total infected cell RNA was extracted at 72-hrs post-infection. RT-PCR was performed using primers that bound conserved regions of exons 1 and 3 of all variant UL128 ORFs. cDNAs were resolved over 4% agarose gels. Three PCR products corresponding to the differentially spliced UL128 transcripts were detected: (i) non-spliced, (ii) partially spliced (lacking intron), and (iii) fully spliced (lacking introns 1 and 2). (B) gpUL128 production. HF-hCARs were infected with RADZ vectors (MOI 3) containing each variant UL128 ORF. At 72 hrs post-infection cell lysates were prepared and analysed by western blot for actin (upper band) and gpUL128-V5 (lower band). (i) gpUL128-V5 produced by expression of UL128 from each BAC cloned strain. (ii) gpUL128-V5 produced by expression of Merlin UL128 containing intron 1 G>T.

Expression of each variant UL128 from RAdZ vectors demonstrated a pattern of gpUL128 production similar to that of fully spliced UL128 transcript production (Figure 4.2.B (i)). Expression of Merlin wildtype UL128 and TR UL128 produced equivalent levels of gpUL128-V5, whilst expression of FIX produced only slightly reduced levels of gpUL128-V5. Expression of TB40-BAC4 UL128 produced dramatically reduced levels of gpUL128-V5. However, unlike the production of fully spliced mRNA, the reduced levels of gpUL128-V5 produced from TB40-BAC4 UL128 were not completely re-capitulated by Merlin UL128 containing the G>T mutation (Figure 4.2.B (ii)).

4.1.2 UL130 sequence variations, including FIX UL130 A>G, do not impact gpUL130 production

Since the UL130 ORF contains no introns, investigations into the expression of the variant UL130 ORFs at a transcriptional level were not pursued. However, no major differences in gpUL130-V5 production were observed following expression of each variant UL130 ORF from the RAd vector background (Figure 4.3.A). Therefore, nucleotide variations amongst each UL130, the most variable of the UL128L ORFs, had no apparent influence on expression. Importantly, the A>G (S72P) mutation also had no impact of the expression of UL130 (Figure 4.3.B).

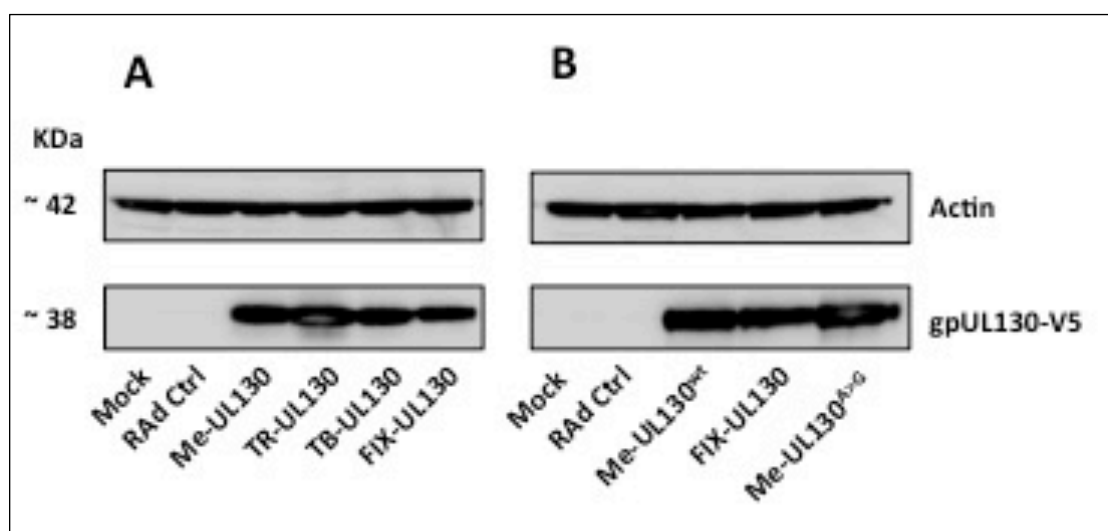


Figure 4.3. Variant UL130 expression. HF-hCARs were infected with RAd vectors (MOI 3) expressing each variant UL130, and at 72 hrs post-infection, cell lysates were analysed by western blot for actin (upper band) and gpUL130-V5 (lower band). (A) gpUL130-V5 produced by expression of variant UL130 contained in the each BAC cloned strain. (B) gpUL130-V5 produced by expression of Merlin UL130 ORF containing FIX UL130 A>G (S72P).

4.1.3 Each variant UL131A is expressed with similar efficiency

Analysis of UL131A cDNA PCR products identified two amplicons corresponding to the differentially spliced UL131A transcripts: (i) non-spliced UL131A pre-mRNA (411 bp); (ii) fully spliced UL131A mRNA (303 bp) (See Figure 4.4.A). The levels of UL131A pre-mRNA transcripts varied only marginally across infections, whilst the fully spliced UL131 mRNA produced by viruses derived from each original BAC-cloned strain were comparable. The only exception to this pattern was the reduced levels of fully spliced mRNA produced from by RV-Me-UL128L^{FIX}, suggesting that FIX UL131A pre-mRNAs are spliced less efficiently when contained within transcripts derived from Merlin genome background.

Expression of each variant UL131A from the RAd vector background produced comparable amounts of gpUL131A-V5, further indicating that each variant UL131A ORF contained no nucleotide features that impacted expression (see Figure 4.4.B).

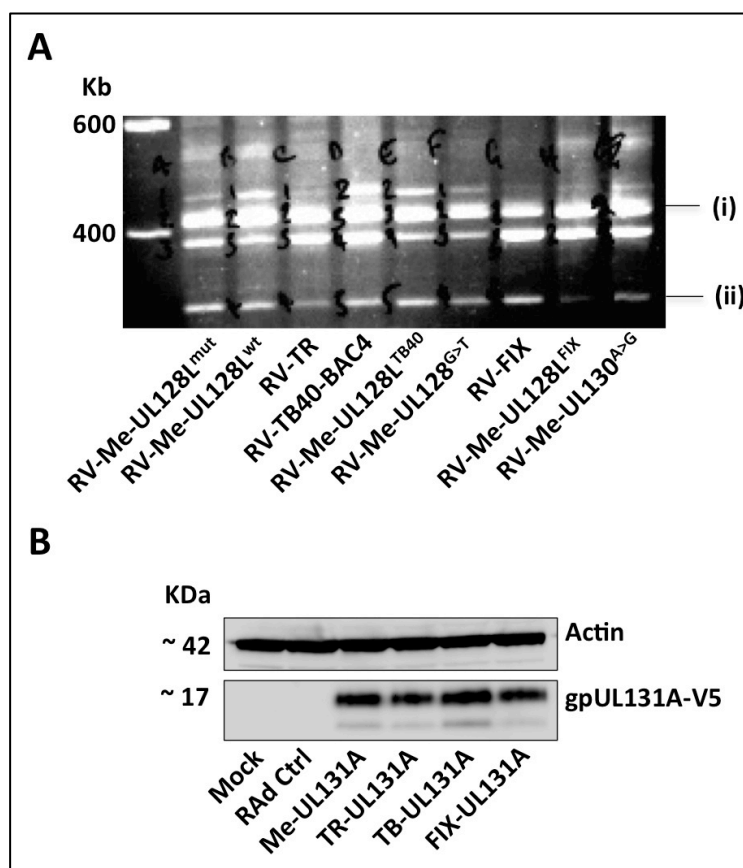


Figure 4.4. Variant UL131A Expression. (A) UL131A transcript splice patterns. HFFs were infected with each indicated virus (MOI 3), and total RNA was extracted from at 72-hrs post-infection. RT-PCR was performed using primers that bound conserved regions in the two exons. cDNAs were resolved over 4% agarose gels. Two amplicons corresponding to the differentially spliced UL131A transcripts were detected: (i) non-spliced, and (ii) fully spliced (lacking intron). (B) gpUL131A production. HF-hCARs were infected with RAd vectors (MOI 3) containing each variant UL131A. At 72 hrs post-infection, cell lysates were prepared and analysed by western blot for cellular actin (upper band) and gpUL131A-V5 (lower band).

4.1.4 Conclusion

Nucleotide variations identified in TR and FIX UL128, as well as those amongst UL130 and UL131A from each of TR, TB40-BAC4 and FIX, whether natural variations or *in vitro* acquired mutations, did not limit the production of monomeric sub-units for pentameric complex assembly. However, the G>T mutation in TB40-BAC4 UL128 affected pre-mRNA splicing, and ultimately gpUL128 production. Importantly, no UL128 mRNA splice variant or novel gpUL128 isoforms that potentially arose due to the G>T variation was detected.

4.2 Impact of UL128L coding potential on pentameri complex assembly and/or incorporation into virions

The fact that the UL128 G>T limited gpUL128 production strongly suggested that virions from virus encoding this mutation could contain reduced quantities of pentameric complex. Whilst the FIX UL130 A>G (S72P) mutation had no effect on gpUL130 production, protein modelling using the Phyre 2 bioinformatics tool (Kelley and Sternberg 2009) predicted that this amino acid transition disrupted a beta sheet, and therefore the folding of gpUL130 into native conformation. This could impact binding interactions between gpUL130 and other pentameric complex sub-units, ultimately impeding the assembly and/or incorporation of the pentameric complex into virions. To explore these possibilities, the pentameric complex content of virions from each virus was determined. To directly relate pentameric complex content to data from investigations into the expression of each variant UL128L, virions analysed were those from the same concentrated virus working stocks previously used (Table 4.1).

4.2.1 Purification of virions from virus working stocks

Purification of virus over sorbitol cushions is common practice in HCMV research, however does not allow separation of virions from NIEPs or DBs. Greater purification can be achieved by rate-velocity sedimentation of virus particles during ultra-centrifugation over gradients (Talbot and Almeida 1977). These two approaches are often used in tandem, including initial purification of virus over sorbitol cushions prior to gradient purification (Varnum, Streblow et al. 2004, Patrone, Secchi et al. 2005, Ryckman, Rainish et al. 2008). We sought to determine the most suitable approach (or combination of approaches) for the purification of virions from concentrated virus stocks prepared by high-speed centrifugation. To assess the degree of purification achieved, recovered viral particles were analysed for the presence of contaminating cellular proteins. Markers of contamination were proteins contained in cellular structures intimately exposed to virion structural components, or those that localise in close proximity with the virion assembly compartment (VAC) (see Table 4.3). Of the markers used, only TGN46 has previously been reported to be incorporated into virions (Cepeda, Esteban et al. 2010).

Table 4.3. Cellular protein markers of virion sample purity.

Protein/ Marker	Cellular location	Comments
S6RP	40S ribosomal sub-unit	Exposed to virion structural components during translation
Giantin	ER resident	Localises in ring-like structure surrounding VAC (Cepeda, Esteban et al. 2010)
Calnexin	ER resident	Chaperone specifically implicated in virion envelope glycoprotein folding (Yamashita, Shimokata et al. 1996)
TGN46	Trans-Golgi network resident	Co-localises with virion materials in structures proximal to the VAC (Sanchez, Greis et al. 2000), and also with gB during endosome retrieval from the plasma membrane (Crump, Hung et al. 2003)

The high-titre RV-Me-UL128L^{mut} (produced in HFFFs) virus stock was purified over a sorbitol cushion, with the levels of co-purified cellular contaminants compared to that in the original virus stock, and also infected cell culture supernatant which had undergone no concentration at all (Figure 4.5.A). Giantin was detected only in the infected cell culture supernatant, indicating that preparation of virus stocks by high-speed centrifugation achieved some degree of purification. However, all other markers of contamination detected in the original virus stock were also present in the sorbitol cushion purified virus. Thus, ultra-centrifugation over sorbitol cushions alone did not sufficiently purify virions.

To determine the utility of gradients for the purification of virions, the recovery and purity of virions isolated from the low-titre RV-Me-UL128L^{wt} stock (produced in RPE-1 cells) was assessed. Virus ultra-centrifuged over glycerol: tartrate gradients produced a distinct banding pattern: a band present at the gradient: sample interface, predicted to contain cellular and virus derived contaminants; a band 2-3 cm into the gradient predicted to contain purified virions; and bands within lower gradient regions predicted to contain dense bodies - this lower 'DB' band was more diffuse less clearly defined. The 'NIEPs' band described to localise between the gradient: sample interface and the major virion-containing band (Irmiere and Gibson 1983) was not observed. Each band was recovered and analysed for markers of contamination, and also glycoprotein gB as a marker of virions (Figure 4.5.B). Samples of the original virus stock, and virus from this stock purified over a sorbitol cushion were also analysed for comparison. As previous, calnexin and TGN46 were each detected in the original RV-Me-UL128L^{wt} virus preparation, and also co-purified with virus over sorbitol cushions. From the gradient purified virus, both calnexin and TGN46 were detected in the band derived from the gradient: sample interface, whilst calnexin was also detected in the diffuse 'DB' band recovered from lower regions in the gradient. Importantly, no markers of contamination were

observed in the band predicted to contain purified virions. gB was readily detected in all gradient samples analysed; each was subsequently shown to contain infectious virus following inoculation of HFFFs, however the highest titres were found in the ‘purified virion’ band. Thus infectious virions were retained at the gradient: sample interface, and also in the lower band that migrated faster than purified virions. However the ‘purified virion’ band contained the greatest titres and was free of contamination.

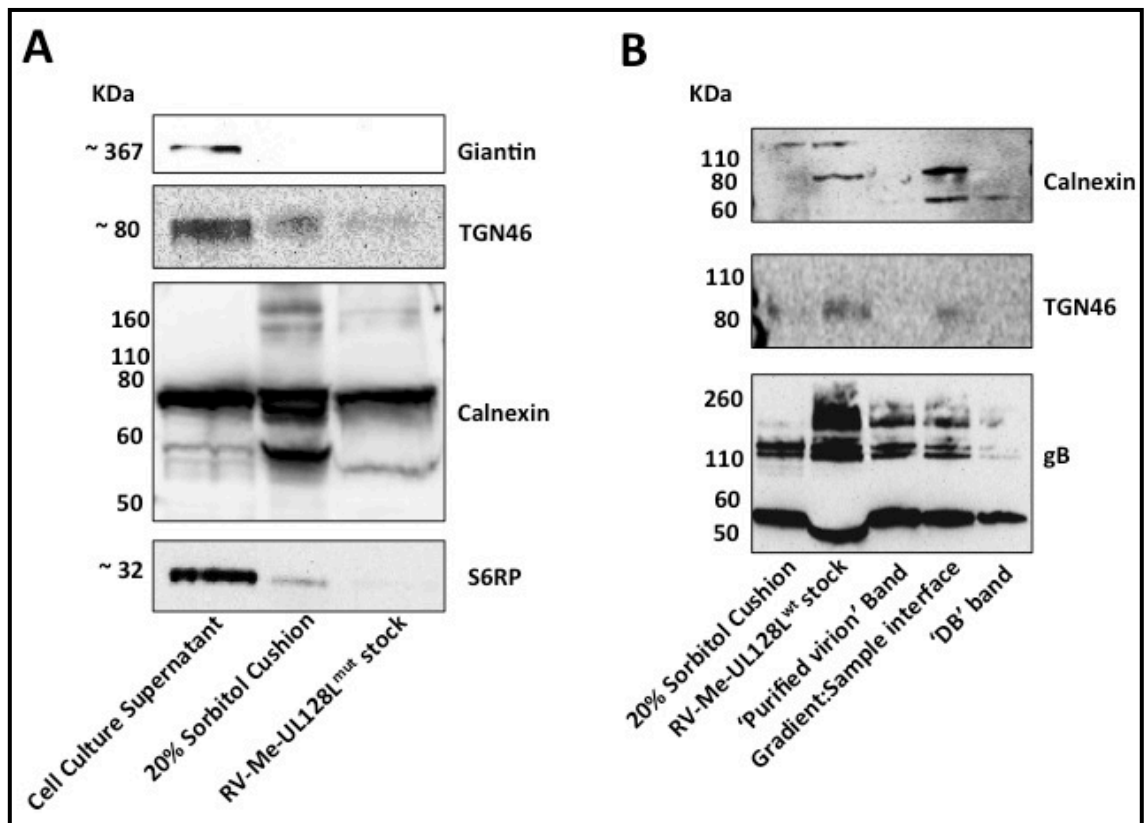


Figure 4.5. Purification of virus over sorbitol cushions and Glycerol: Tartrate gradients. (A) Virus within the high titre RV-Me-UL128L^{mut} stocks was ultra-centrifuged over a cushion of 20% solution in PBS. Recovered virus was analysed by Western blot for the presence of cellular contamination markers: Giantin, TGN46, Calnexin, S6RP. Samples from the original RV-Me-UL128L^{mut} stock, as well as infected cell-culture supernatant undergone no concentration were analysed in parallel. (B) Virus within the low titre RV-Me-UL128L^{wt} stocks were ultra-centrifuged over a glycerol: tartrate gradients, and produced a distinct banding pattern. Each band was recovered and materials therein were diluted in NaPh buffer, and collected by further ultra-centrifugation. Pelleted materials were analysed by western blot for markers of contamination, and also gB as a marker of virions. Samples of the same virus purified over sorbitol cushions and a sample undergone no purification at all were analysed in parallel.

4.2.2 Pentameric complex content of virions

The pentameric gH/gL complex is believed to contain equimolar quantities of each glycoprotein sub-unit, and each UL128L sub-unit is absolutely required for assembly and incorporation into virions (Wang and Shenk 2005, Ryckman, Rainish et al. 2008). We therefore used detection of gpUL128 as surrogate a marker for the entire pentameric complex (Figure 4.6.A). Samples of virions purified from each virus stock over gradients were normalised according to the quantities of envelope glycoprotein gB.

Virions purified from virus lacking intact UL128L (RV-Me-UL128L^{mut}) did not contain detectable gpUL128. In contrast, virions purified from virus encoding wildtype UL128L (RV-Me-UL128L^{wt} and RV-Me-UL128L³³⁰¹) contained readily detectable gpUL128, the most abundant of any virus analysed. Virions purified from RV-TR stocks contained comparable levels of gpUL128 to those detected in RV-Me-UL128L^{wt}. In contrast, RV-TB40-BAC4 virions contained reduced levels of gpUL128, as did virions purified from Merlin variants containing the entire TB40-BAC4 UL128L (RV-Me-UL128L^{TB40}), or the single G>T mutation identified in TB40-BAC4 UL128 (RV-Me-UL128^{G>T}). RV-FIX virions contained vastly reduced levels of gpUL128 that were initially un-detectable, and this too was closely repeated in RV-Me-UL128L^{FIX} virions. However, whilst RV-Me-UL130^{A>G} virions also contained reduced levels of gpUL128, they contained more than virions from viruses expressing the entire FIX UL128L.

In a second blot (Figure 4.6.B), the maximum volume of purified virion samples was analysed (i.e. without normalisation), facilitating the detection of virion incorporated gpUL128 even where levels were drastically reduced. This analysis conclusively demonstrated that gpUL128 was present in virions from RV-TB40-BAC4, RV-Me-UL128L^{TB40}, RV-Me-UL128^{G>T}, RV-FIX and RV-Me-UL128L^{FIX}.

4.2.3 Conclusion

The pentameric complex content of virions from Merlin viruses containing alternative wildtype UL128L ORFs (native, or from strain 3301) was comparable. Thus, pentameric complex assembly and/or incorporation did not vary due to natural inter-strain differences in UL128L. TR UL128L ORFs, which are expressed with similar efficiency compared to Merlin wildtype UL128L, also contain no features that convey different complex assembly and/or incorporation into virions. Virions from the Merlin viruses expressing TB40-BAC4 UL128L or the single UL128 G>T variation alone displayed similar pentameric complex content, thus the reduced gpUL128 production due to the UL128 G>T variation appeared entirely responsible for this effect. In contrast to this, observation that virions from Merlin virus containing FIX UL128L displayed lower pentameric complex content compared to those from the Merlin virus modified by the UL130 A>G variation alone indicated that FIX UL128L contains other features that impede complex assembly and/or incorporation into virions.

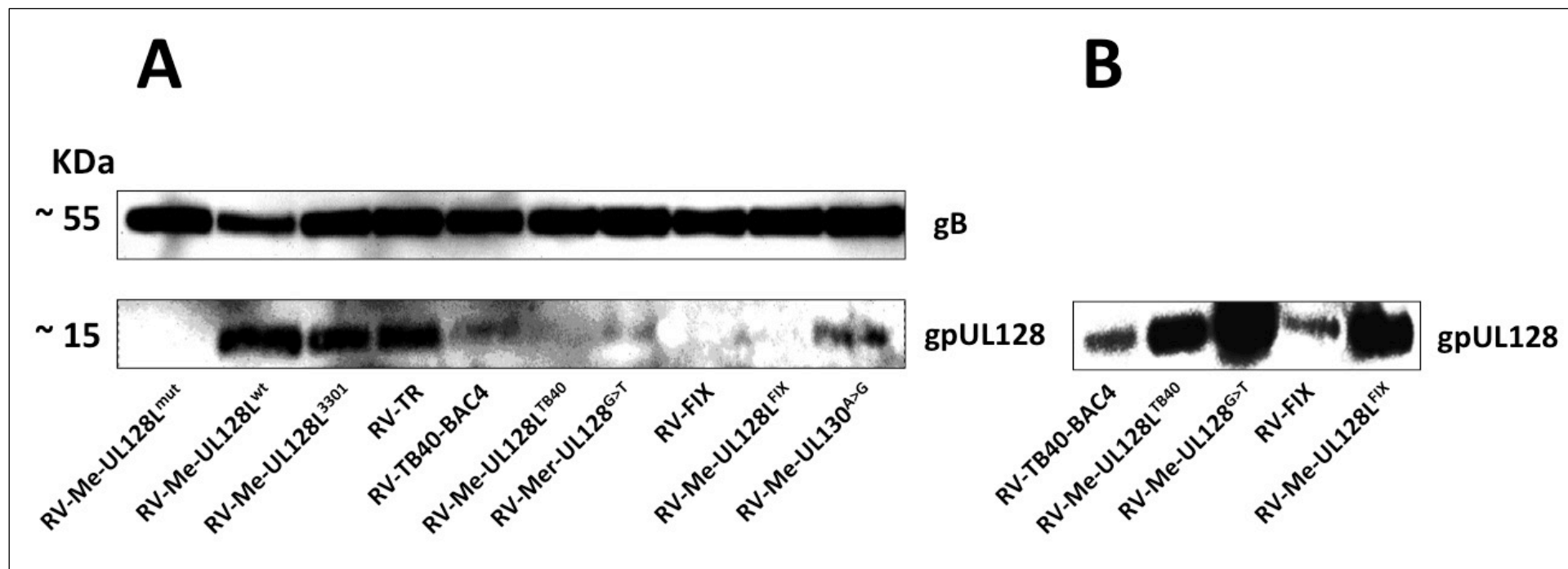


Figure 4.6. gpUL128 content of virions. (A) Virions were purified from working stocks of each virus indicated on glycerol-tartrate gradients, and analysed by western blot. Samples loads were normalised to gB content (upper panel) and stained for gpUL128 (lower panel). (B) The gpUL128 content in the maximised loads of samples analysed in A that were not non-normalised to gB content.

4.3 Impact of virion gH/gL/gpUL128/gpUL130/gpUL131A content on epitheliotropism

The pentameric complex content of virions had the potential to impact the efficiency by which the virus infects cell types where the complex is required. Therefore, the epitheliotropism of the different viruses investigated was assessed: the same virus stocks in which pentameric complex content was analysed (Table 4.1.) were titrated onto fibroblasts (HFFFs) and epithelial cells (RPE-1s and ARPE-19s) in parallel, and titres reported on the different epithelial cell types were normalised to that reported on fibroblasts. Thus, the ability of each virus to infect epithelial cells was quantified relative to its ability to infect fibroblasts.

RV-Me-UL128L^{mut} infected ARPE-19s with ~5,000-fold reduced efficiency, and RPE-1s with ~100-fold reduced efficiency, than in HFFFs (Figure 4.7.A). Hence the lack of pentameric complex in virions from this virus resulted in significantly reduced tropism for both epithelial cell types (both $p < 0.0001$). In contrast to this, viruses that produce virions with the greatest pentameric complex content (RV-Me-UL128L^{wt}, RV-Me-UL128L³³⁰¹, and RV-TR) displayed similar tropism for fibroblasts and epithelial cells.

RV-TB40-BAC4 infected ARPE-19s with comparable efficiency to HFFFs, but infected RPE-1s with ~9-fold lower efficiency ($p < 0.05$) (Figure 4.7.B). RV-Me-UL128L^{TB40} and RV-Me-UL128L^{G>T} each infected ARPE-19s with ~4-fold reduced efficiency than HFFFs ($p < 0.001$ and $p < 0.0001$, respectively), and RPE-1s 8-10-fold less efficiently (both $p < 0.0001$). Thus, the similarly reduced pentameric complex content of virions from the Merlin variants resulted in a near identical reduction in relative epitheliotropism.

RV-FIX infected ARPE-19s and RPE-1s with ~75-fold and >180-fold reduced efficiency compared to HFFFs, respectively (both $p < 0.0001$) (Figure 4.7.C). RV-Me-UL128L^{FIX}, the Merlin variant containing FIX UL128L that produced virions with vastly reduced pentamer content, infected ARPE-19s with ~25-fold reduced efficiency and RPE-1s with ~50-fold reduced efficiency (both $p < 0.0001$), thus displaying severely reduced epitheliotropism. This was only partially recapitulated by RV-Me-UL130^{A>G} that infected ARPE-19s and RPE-1s with ~2-fold and ~6.5-fold reduced efficiency, respectively ($p < 0.05$).

4.3.1 Conclusion

The relative epitheliotropism of viruses, particularly the Merlin variants, correlated with the pentameric complex content of virions: as pentameric complex content pentamer decreased, so did epitheliotropism. However, whilst the epitheliotropism of the novel Merlin viruses containing TB40-BAC4 and FIX single unique nucleotides was reduced, these Merlin variants displayed at most 10-fold reduced relative epitheliotropism, as compared to the 100-5000-fold reduced relative epitheliotropism of Merlin virus completely lacking the pentameric complex.

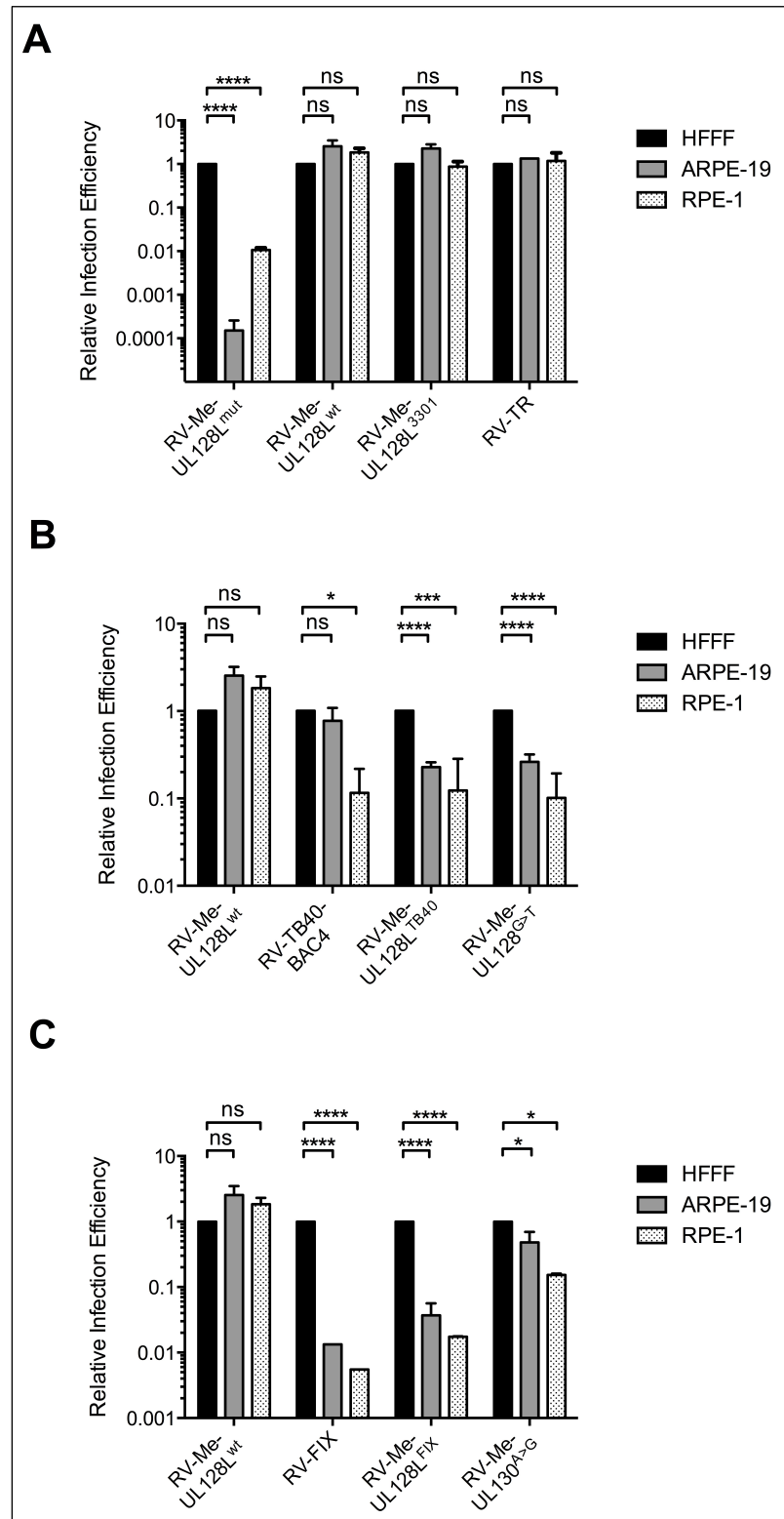


Figure 4.7. Impact of virion pentameric complex content on epitheliotropism. Virus working stocks were plaque titrated onto HFFFs, ARPE-19s and RPE-1s in parallel, with all titres normalised to the titres of the same virus on HFFFs. Data are based on 2-3 independent experiments ($n = 2-3$). \pm SD is shown. **(A)** Relative epitheliotropism of viruses containing undetectable or the greatest levels of pentameric complex. **(B)** Relative epitheliotropism of viruses containing reduced pentameric complex due to expression of TB40-BAC4 UL128L sequence features. **(C)** Relative epitheliotropism of viruses containing reduced pentameric complex levels due to expression of FIX UL128L sequence features. (Where indicated, the fibroblast infection efficiency of each virus was compared to the epithelial cell infection efficiency by 1-way analysis of variance [ANOVA] followed by Dunnett's posttest: ns – non-significant, * $p < 0.05$, ** $p < 0.01$, *** $p < .001$, **** $p < 0.0001$).

4.4 Impact of UL128L coding potential on progeny virion production and/or release from epithelial cells

The pentameric complex content of virions could be inversely correlated with cell-free virus production displayed during the growth kinetics assays. However it was unclear whether the expression of UL128L limited the formation, or alternatively the release, of progeny virions. Furthermore, it has been reported that TB40-BAC4 BAC-derived virus produced two distinct progeny sub-populations in endothelial cells, each varying in virion-incorporated pentameric complex content and tropism: progeny released into the supernatant contain lower levels of pentameric complex, and in turn display reduced endotheliotropism, whilst progeny retained associated with cells contain greater levels of pentameric complex, and display the greatest endotheliotropism (Scrivano, Sinzger et al. 2011). Thus, it was considered possible that a major source ‘broad-tropism’ virus could be retained within producer cells. To explore this possibility, the production of cell-free released virus (CRV) or cell-association virus (CAV) sub-populations during infection of epithelial cells was assessed, as was the tropism of any recovered sub-populations. RPE-1s were infected at low MOI with RV-Me-UL128L^{wt}, and CRV and CAV fractions prepared 4 days after all cells were productively infected (i.e. GFP+): the CRV fraction was prepared from the infectious supernatants that was cleared of cellular materials by centrifugation; the CAV fraction was prepared from cells that remained intact, and also the pelleted cellular material cleared from the CRV fraction.

4.4.1 Relative CRV and CAV infectivity produced during infection of epithelial cells

Sonication was explored as a method to disrupt cells and liberate CAV infectivity. To determine whether this treatment had an impact on virion integrity, infectious supernatant was sonicated under conditions required to achieve total cell lysis, with the infectivity that could be recovered quantified by plaque titration in parallel to supernatants that had undergone no treatment at all (Figure 4.8.A). The near complete recovery of infectivity from the sonicated supernatant sample indicated that this treatment did not disrupt virions.

Following sonication of the infected cellular materials (intact cells and cellular debris), lysates were centrifuged to generate two further CAV sub-fractions: the cell-extract compartment (cleared supernatant), and the cell-debris compartment (pelleted cellular material). For comparison of the relative quantities of progeny virus located to either the CRV or CAV fractions, the ‘CAV total’ was estimated as the sum of infectivity reported in the cell-extract and cell-debris sub-fractions. The CAV fraction was a considerable site of infectivity, and contained similar levels of total infectivity to the CRV fraction (Figure 4.8.B). However the relative virus titres from the CRV and CAV (total) sub-populations were of different magnitudes on the different cell types used. Interestingly, the CAV ‘total’ infectivity was ~ 4-fold higher than the CRV fraction on HFFs ($p < 0.05$), but ~3.5-fold lower than in the CRV fraction on ARPE-19s

($p < 0.01$). The infectivity within each infected cell-culture fraction was reported with similar titres on RPE-1s.

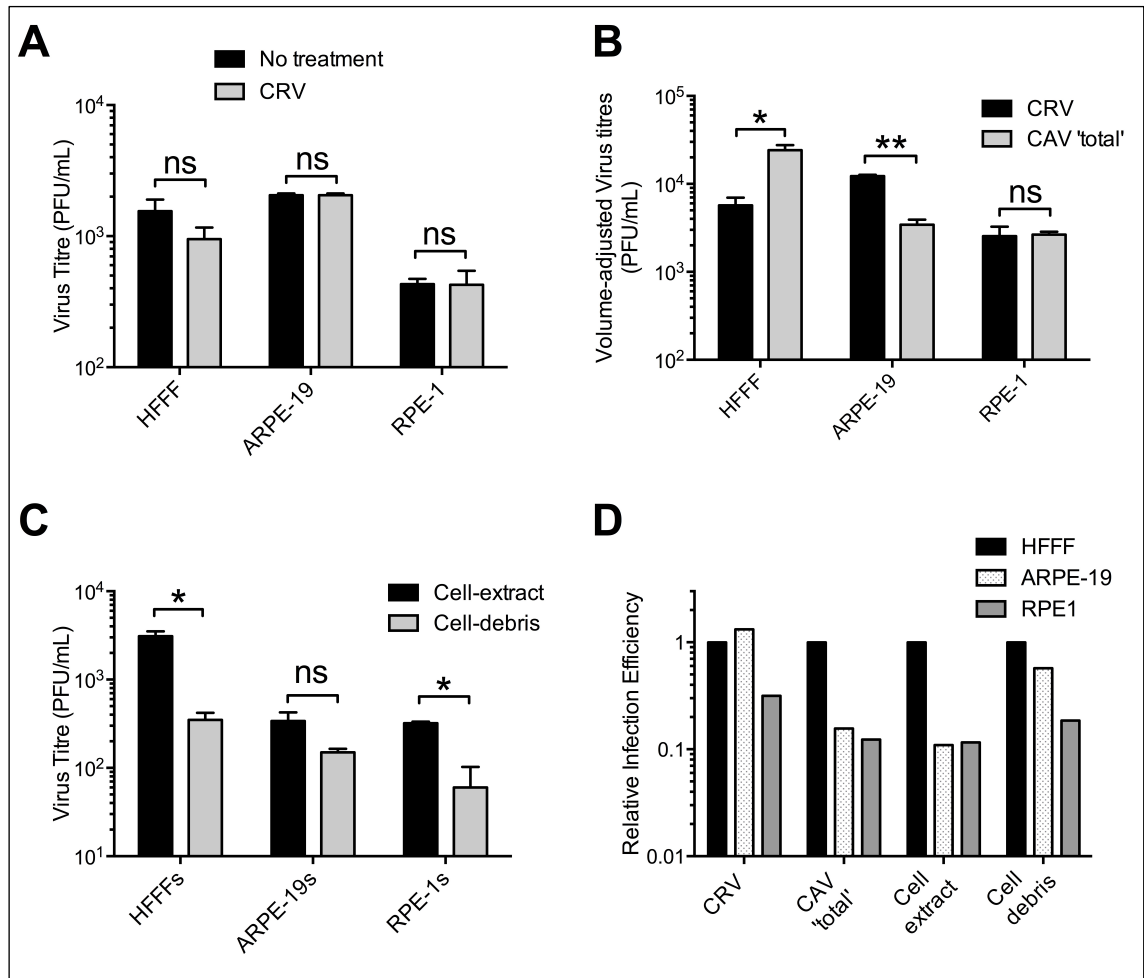


Figure 4.8. CRV and CAV sub-populations produced in epithelial cells. RPE-1 cells were infected at low MOI with Merlin virus containing wildtype UL128L ORFs (RV-Me-UL128L^{wt}). 4 days after all cells were productively infected, cell-released virus (CRV) and cell-associated virus (CAV) populations were analysed. All data is based on 1 experiment (n = 1). \pm SEM are shown. **(A)** To assess the effect of sonication on virions, treated (sonicated) or untreated infectious supernatants were analysed in parallel by plaque titration onto HFFFs, ARPE-19s, and RPE-1s in duplicate. **(B)** Infectivity recovered from the CRV sub-population and CAV fractions was quantified by plaque titration onto HFFFs, ARPE-19s, and RPE-1 in parallel. The 'CAV total' infectivity was estimated as the sum of infectivity recovered from both the Cell-extract and Cell-debris fractions. To account for the preparation of CRV and CAV sub-populations in different volumes (6 mL and 7 mL, respectively), the total infectivity in each sample was calculated from the determined titres. **(C)** Titres of virus recovered from Cell-extract and Cell-debris fractions, reported on each different cell type. **(D)** Relative epitheliotropism of virus recovered from the CRV and CAV sub-fractions was determined by normalising infectivity reported in all infected-cell culture fractions (and sub-fractions) on ARPE-19s and RPE-1s to titres of the same samples reported on HFFFs. (Where indicated, titres of virus progeny recovered from the stated infected cell culture sub-population and sub-fraction were compared by paired 1-tailed t-test: ns – non-significant, * $p < 0.05$, ** $p < 0.01$, *** $p < .001$, **** $p < 0.0001$).

Within the CAV sub-population, the Cell-extract sub-fraction was the major site of infectivity and contained greater titres of virus than that in the cell debris sub-fraction. As reported on HFFs, the cell-extract fraction contained ~9-fold greater titres compared to the cell debris fraction ($p < 0.05$), whilst on RPE-1s the cell-extract fraction was also reported to contain ~5-fold greater titres ($p < 0.05$). Furthermore, the titres of the cell-extract fraction were also greater than in the cell-debris fraction on ARPE-19s, but this was not significant. Thus the majority of the virus progeny that remained associated with producer cells could be separated from cellular materials.

4.4.2 Relative epitheliotropism of CRV and CAV populations produced during infection of epithelial cells

The relative epitheliotropism of progeny contained within the CRV and CAV sub-fractions was quantitated by normalising titres of each sample reported on epithelial cells to those of the same sample reported on fibroblasts (see Figure 4.8.C). Virus within the CRV fraction infected ARPE-19s slightly more efficiently than HFFs, though infected RPE-1s with ~5-fold reduced efficiency compared to HFFs. Interestingly, infectivity in the 'CAV total' fraction displayed reduced epitheliotropism compared to the CRV infectivity. Whilst the epithelial cell infection efficiency of virus in the cell-debris compartment was also reduced compared to that of the CRV, the infectivity contained within the cell-debris compartment displayed greater epitheliotropism than that in the cell-extract compartment.

4.4.3 Conclusion

The CAV fraction produced during infection of epithelial cell cultures was a major source of infectious virus. However the total infectivity recovered from both CRV and CAV fractions of epithelial cells infected with virus containing wildtype UL128L (RV-Me-UL128L^{wt}) was still lower than CRV infectivity produced during previous infections with viruses engineered to contain UL128L adaptations (i.e. RV-Me-UL128L^{TB40}, RV-Me-UL128L^{G>T}, RV-Me-UL128L^{FIX}, RV-Me-UL130^{A>G}). Therefore, expression of wildtype UL128L apparently inhibits the production of infectious virus during infection in epithelial cells. Furthermore, virus localised to the CAV fraction of infected epithelial cell cultures displayed reduced epitheliotropism compared to that of the CRV fraction, and presumably included virions with reduced pentameric complex content.

4.5 Impact of UL128L coding potential on progeny virion production and/or release from fibroblast cells

We next sought to determine whether pentameric complex impeded the morphogenesis or cell-free release of progeny virions produced in fibroblasts. In the previously mentioned report, it was suggested that CRV and CAV progeny sub-populations are produced in fibroblasts, with the CAV displaying the more restricted tropism compared to the CRV sub-population (Scrivano, Sinzger et al. 2011). Thus, the tropism of any sub-population recovered from fibroblasts was also assessed. One-step growth curve assays were performed, with the CRV and CAV produced over the time-course of each infection compared. Viruses used in these assays are those described in Table 4.1, except for an alternative version of the Merlin virus containing wildtype UL128L; the titre of the previous RV-Me-UL128L^{wt} stock grown in RPE-1s was not sufficient for the performance of high MOI infections. Instead, an alternative version of this virus was grown to high titres by alleviating the inhibitory impact of UL128L in a tetracycline repression culture system (Stanton, McSharry et al. 2008, Stanton, Baluchova et al. 2010). This virus, RV-Me-tet-UL128L^{wt}, differed at only two genetic loci compared to RV-Me-UL128L^{wt}: at UL122, no eGFP expression cassette was inserted; in the UL131A-UL132 inter-genic region, a binding sequence for the tetracycline operator was inserted.

HFFF monolayers infected with all Merlin viruses were almost entirely obliterated by the final time-point, (day 10 p.i.) however RV-TR, RV-TB40-BAC4 and RV-FIX infected HFFFs displayed delayed degradation. CRV samples were collected every two days post infection, and cellular materials cleared by low-speed centrifugation were returned to each infection. Total CAV samples were prepared at the final time-point of each infection as previously mentioned, with CAV infectivity released by sonication.

4.5.1 Relative CRV and CAV infectivity produced by viruses containing non-intact or wildtype UL128L ORFs

The inhibitory impact of Merlin UL128L on cell-free virus production was again demonstrated, with the total estimated CRV produced by RV-Me-UL128L^{mut} ~900-fold greater than that produced by RV-Me-tet-UL128L^{wt} (Figure 4.9.A). The CAV sub-population represented <1% of the total infectivity produced by RV-Me-UL128L^{mut}, thus the vast majority of progeny virus produced in this infection had egressed producer cells. Surprisingly, no CAV infectivity was recovered from the RV-Me-tet-UL128L^{wt} infection, providing strong evidence that the pentameric complex was inhibitory to the production of progeny virions, and not simply cell-free release. Surprisingly, RV-TR produced the least CRV infectivity of any infection. In fact, cell-free virus was un-detectable other than at day 2 or day 10 post-infection. The vast majority of the RV-TR infected HFFFs (>80%) remained intact at day 10 post-infection, with ~98% of the total infectivity recovered located in the CAV fraction.

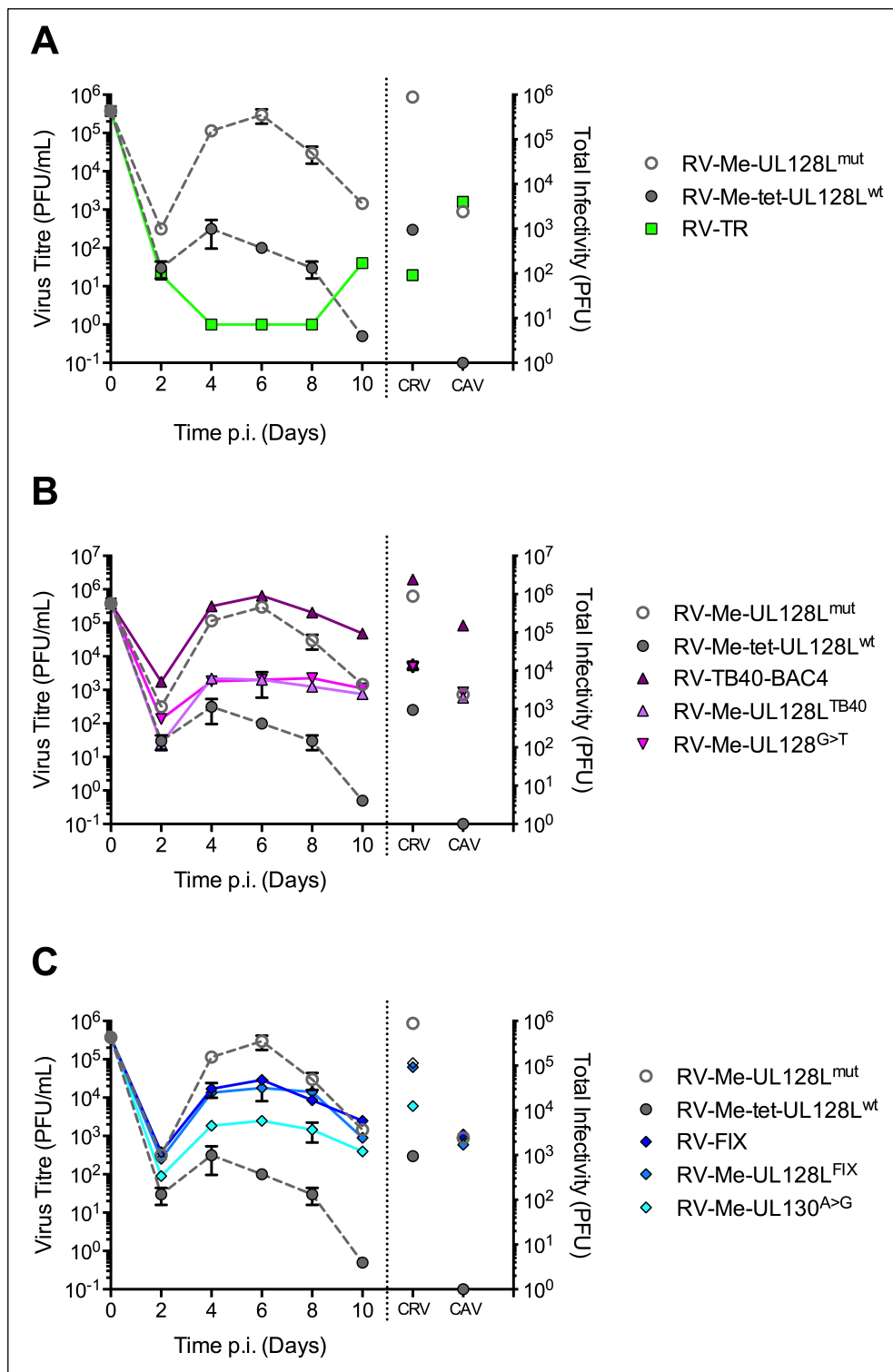


Figure 4.9. CRV and CAV sub-populations produced in fibroblast infections. HFFFs were infected (MOI 3) with each indicated virus, and supernatant samples were collected every two days and cleared of cellular materials by centrifugation. Pelleted cellular materials were returned to each original infection. At day 10 post-infection, the total cellular material was sonicated to release CAV infectivity. All CRV and CAV samples were analysed by plaque titration onto HFFFs. Plotted on the left Y-axis, infectivity contained within each CRV samples collected over time-course of infections. On the right Y-axis, total CRV and CAV infectivity. Total CRV infectivity was estimated as the sum of infectivity produced at each time-point over the time-course of infections. All data based on 1 experiment (n = 1). \pm SEM are shown. Plotted (A) CRV and CAV subpopulations produced by virus containing greatest/least pentamer. (B) CRV and CAV subpopulations produced by virus containing pentamer levels dictated by TB40-BAC4 UL128L nucleotides. (C) CRV and CAV subpopulations produced by virus containing pentamer levels dictated FIX UL128L nucleotides.

4.5.2 Relative CRV and CAV produced by viruses containing TB40-BAC4 UL128L sequence features

The RV-TB40-BAC4 infection produced the most cell-free virus, with total CRV infectivity >2500-fold more than produced by RV-Me-tet-UL128L^{wt} and also more than twice that produced by RV-Merlin-UL128L^{mut} (Figure 4.9.B). However, this infection also produced the greatest quantities of CAV infectivity; this may reflect the fact that more cells remained intact at the final infection time point (day 10 p.i.). Merlin viruses expressing TB40-BAC4 UL128L or the single UL128 intron 1 G>T variation (RV-Me-UL128L^{TB40} and RV-Me-UL128L^{G>T}) again displayed almost identical cell-free virus production that was intermediate to the Merlin control virus infections. They also produced similar levels of CAV infectivity, which compared to RV-Me-UL128L^{mut} infection, represented a greater proportion of the total infectivity produced (~14% and ~16%, respectively). However, the CAV produced in each infection could not account for the deficit in CRV production compared to RV-Merlin-UL128L^{mut}. Thus, TB40-BAC4 UL128L was also inhibitory to infectious virus production, although less so than Merlin wildtype UL128L. Furthermore, the UL128 intron 1 G>T nucleotide variation was solely responsible for this difference.

4.5.3 Relative CRV and CAV produced by viruses containing FIX-BAC UL128L sequence features

The total CRV produced by the RV-FIX infection was >120-fold greater than that produced by RV-Me-tet-UL128L^{wt}, though ~7.5-fold less than that produced by RV-Me-UL128L^{mut} (Figure 4.9.C). This was almost re-capitulated by the Merlin virus containing FIX UL128L, which produced total CRV that was ~100-fold greater than that produced by RV-Me-UL128L^{wt}. The CAV from each of these infections was broadly equivalent, representing 2-3% of the total infectivity produced. Thus, as in the RV-Me-UL128L^{mut} infection, the vast majority of RV-FIX and RV-Me-UL128L^{FIX} progeny virions that also display dramatically reduced pentameric complex content had successfully egressed producer cells. The total CRV produced by RV-Me-UL130^{A>G} was >13-fold greater than that produced by RV-Me-UL128L^{wt}, and thus lower than that produced by each of RV-Me-UL128L^{mut}, RV-FIX, or RV-Me-UL128L^{FIX}. The titres of CAV produced by RV-Me-UL130^{A>G} were similar to that produced by RV-FIX or RV-Me-UL128L^{FIX}; this represented a greater source of the total progeny virus produced (~17%), however could not account for the lower CRV produced.

4.5.4 Tropism of CRV and CAV sub-populations produced in fibroblasts

The tropism of CRV and CAV sub-populations produced in fibroblasts were assessed from separate infections. In the previous one-step growth curve assays, peak titres of cell-free virus were consistently produced at day 6 post-infection, other than in the RV-TR infection.

Furthermore, HFFFs infected with Merlin viruses were almost completely degraded by day 10 post-infection, whereas more HFFFs infected with viruses from TR-BAC, TB40-BAC4, and to a lesser extent FIX-BAC, remained intact. Thus, to assess progeny derived from a comparable number of intact cells across all infections, CRV and CAV fractions were produced from infected HFFFs at day 8 post-infection. CRV and CAV samples were then titrated onto fibroblasts (HFFFs) and epithelial cells (RPE-1s) in parallel, with titres of each sample reported on epithelial cells normalised to that for the same sample reported on fibroblasts.

CRV from the RV-Me-UL128L^{mut} infection displayed ~100-fold reduced infection efficiency in RPE-1s compared to HFFFs, similar to that previously demonstrated for stocks of RV-Me-UL128L^{mut} grown in HFFFs (Figure 4.10.A). Interestingly, RV-Me-tet-UL128L^{wt} CRV infected RPE-1s with >10-fold reduced efficiency than HFFFs, whilst RV-TR CRV infected RPE-1s with >100-fold reduced efficiency. This was in contrast to the comparable fibroblast and epithelial cell infection efficiency by counterpart viruses that had been produced in epithelial cells. Similarly, CRV from infections with viruses expressing TB40-BAC4 UL128L sequence features infected RPE-1s 15-30-fold less efficiently than they did HFFFs, whereas versions of each corresponding virus grown in epithelial cells had previously displayed 8-10-fold reduced RPE-1 cell infection efficiency. CRV from the RV-FIX infection again displayed ~100-fold reduced epitheliotropism, similarly to the RV-FIX stock produced in HFFFs. CRV from infections with Merlin viruses expressing FIX UL128L sequence features infected epithelial cells in a similar pattern to that seen previously: RV-Me-UL128L^{FIX} displayed the most reduced (~40-fold) epitheliotropism, and RV-Me-UL130^{A>G} displayed the least reduced (~14-fold) epitheliotropism. The epithelial infection efficiency of each virus was however slightly lower than previously demonstrated for counterpart virus stocks produced in epithelial cell culture. A similar pattern of results was seen in the reported infectivity in the CAV fractions of each infection (Figure 4.10.B).

4.5.5 Conclusion

Increasing levels of pentameric complex in virions inversely correlated with progeny virion production, and not simply the cell-free release, from producer fibroblast cells. Whilst increasing pentameric complex levels could be correlated with greater relative proportions of CAV produced several infections in this cell type, this compartment did not contain any source of infectious virions that could account for differences in CRV production. The tropism of viruses containing intact UL128L and produced in fibroblast was different to that of the same virus produced in epithelial cells; viruses produced in fibroblasts consistently displayed reduced epitheliotropism compared to counterparts grown in epithelial cells. Furthermore, unlike progeny sub-populations produced in epithelial cell infections, the CRV and CAV sub-populations displayed no observable differences in tropism.

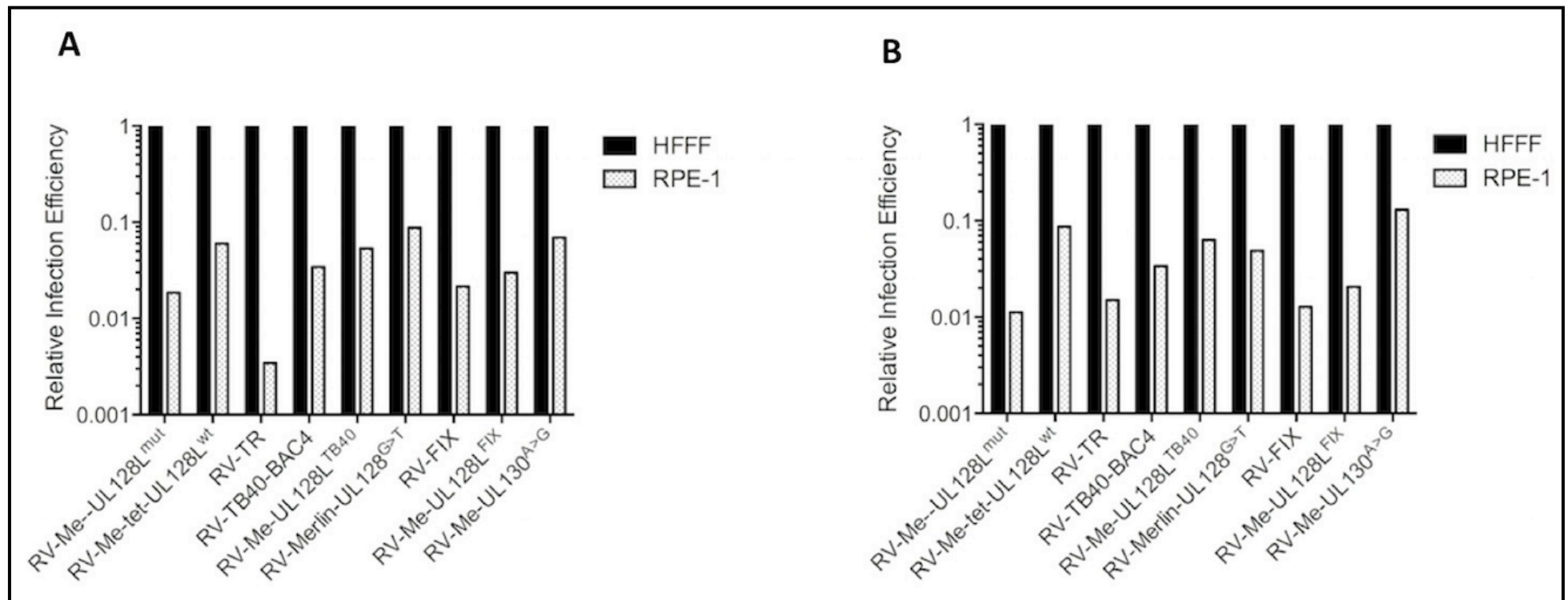


Figure 4.10. Tropism of CRV and CAV Sub-populations produced in Fibroblasts. HFFFs were infected with each stated virus (MOI 3), and at day 8 post-infection, CRV and CAV fractions were prepared as previously described. CRV and CAV fractions were prepared in identical final volumes, and analysed by titration onto HFFFs and RPE-1s in parallel. Titres reported on RPE-1s were normalized to those reported on HFFFs, giving the relative epithelial cell infection efficiency (epitheliotropism). All data based on one experiment (n=1). **(A)** The relative epithelial cell infection efficiency of CRV fraction infectivity from each infection. **(B)** The relative epithelial cell infection efficiency of CAV fraction infectivity from each infection.

4.6 Summary

An investigation as to how the TB40-BAC4 UL128 G>T and FIX UL130 A>G (S72P) mutations impacted pentameric complex functionality revealed that they each reduced the assembly and/or incorporation of the pentameric complex into virions. However, these effects were brought about by different mechanisms: the UL128 intron 1 G>T nucleotide variation reduced the efficiency by which UL128 pre-mRNAs were spliced, in turn reducing the production and availability of gpUL128 for pentameric complex assembly; the UL130 A>G (S72P) nucleotide variation did not impact UL130 expression, and it is most likely that the gpUL130 structural changes that were introduced impacted peptides that facilitate binding interactions with other pentameric complex sub-units.

Except for the TB40-BAC4 UL128 G>T mutation, the UL128L ORFs of each strain investigated contained no further nucleotide sequence variations that impacted expression at either the transcriptional or translational levels. The similar expression efficiency of TR UL128L to that of Merlin UL128L resulted in RV-TR virions containing pentameric complex levels that were equivalent to that of the Merlin virus containing wildtype UL128L sequences. The reduced levels of pentameric complex in TB40-BAC4 virions appeared to be heavily contributed to by the UL128 G>T mutation, whilst the reduced pentameric complex content of FIX virions was only partially due to the UL130 A>G (S72P) mutation; virions from the recombinant Merlin virus containing the A>G (A72P) mutation contained greater pentameric complex levels than the recombinant Merlin virus containing the entire FIX UL128L genome region.

The pentameric complex content impacted various activities in the viral replication cycle, offering some explanation to the previously determined distinct growth phenotypes of the viruses assessed. Reduced pentameric complex content in virions in turn reduced the ability of virus to gain entry into epithelial cells relative to fibroblasts, however the relative epitheliotropism of Merlin viruses containing the single UL128 G>T and UL130 A>G mutations was only slightly compromised (≤ 10 -fold) in comparison to virus completely lacking the pentameric complex (100-5000-fold). Greater cell-free virus production of strains also correlated with reduced pentameric complex content. Data from infections performed in epithelial cells and fibroblasts suggested that whilst the pentameric complex may promote cell-association and impede the release of progeny virions, its main effect was to inhibit infectious virion morphogenesis. Furthermore, any progeny virus recovered from cellular sites displayed reduced epitheliotropism, and thus likely contained less pentameric complex. Importantly, viruses derived from the same BAC-clone though produced in different cell-types display differences in tropism; cell-free released virus produced during infection of epithelial cell culture displayed greater epitheliotropism than counterparts produced in fibroblasts.

5 Genetic Stability of Virus from
BAC-cloned Strains and Novel
Merlin-BAC Variants *in vitro*

As previously mentioned, viruses from the TR, TB40-BAC4 and FIX BACs appear to retain broad tropism when propagated in fibroblast cell culture. We hypothesised that the apparent greater stability of UL128L in these strains, compared to that contained within clinical isolates, was directly related to the more efficient growth phenotype of the viruses they produced. Encouragingly, the transfer of TB40-BAC4 UL128 and FIX UL128L sequences also conveyed a greater fitness to Merlin viruses in fibroblasts, and we this potentially alleviated the selective pressures normally encountered during passage in cell culture. However, the adaptation of TR BAC-derived viruses in the previous growth kinetics assays had apparently occurred without mutation in UL128L, suggesting that other genome regions may have also contributed to the greater stability of TR UL128L. This potentially applied to the apparent greater stability of TB40-BAC4 and FIX UL128L genome regions too.

5.1 Adaptation of TR-BAC-derived viruses during previous growth kinetics assays

5.1.1 Virus reconstituted from the TR-BAC adapted during a single passage in HFFFs

Initially, the RV-TR-HFFFp1 virus (section 3.1) was further passaged in fibroblasts to enrich any UL128L mutant population that may not have been detected by PCR-based sequencing. The resultant RV-TR-HFFFp2 virus displayed further increased HFFF plaque formation (Figure 5.1.), yet PCR-based sequencing again suggested that UL128L appeared unchanged.

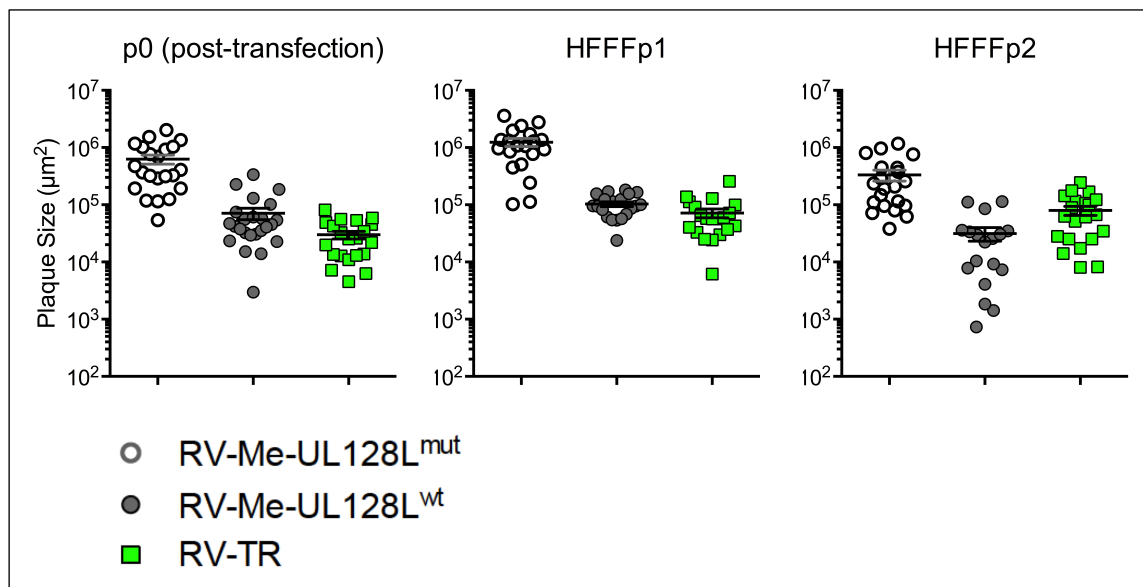


Figure 5.1. Adaptation of RV-TR to fibroblasts. The cell-to-cell spread of TR BAC-derived virus in fibroblast following sequential passage in the same cell type. Where possible, a minimum of 20 plaques following 2 weeks incubation under overlay was measured. Mean and \pm SEM are shown. (A) HFFF plaques formed following transfection of HFFFs with BAC DNA for each virus (p0). (B) HFFF plaques sizes following 1 passage in HFFFs. (C) HFFF plaques formed by RV-TR-HFFFp2, compared to that of RV-Me-UL128L^{mut}-HFFFp1 and RV-Me-UL128L^{wt}-HFFFp1.

The RV-TR-HFFFp2 virus was assessed by whole genome sequencing, performed by Dr Andrew Davison and Dr Gavin Wilkie (MRC-University of Glasgow Centre for Virus Research, Glasgow, UK) (Appendix V). To generate sufficient quantities of viral genomic DNA for analysis, HFFFs were infected at high MOI with RV-TR-HFFFp2, and viral genomic DNA was extracted at 72-hr post-infection. Whole genome sequencing confirmed that the RV-TR-HFFFp2 virus contained UL128L genome regions that were unchanged. Furthermore, RL13 that is apparently intact in TR also remained unchanged. However, a significant population (~30%) in the RV-TR-HFFFp2 virus contained a large deletion encompassing sequence originating from the *a'* genome region terminus, and ending in the BAC vector sequence (Figure 5.2.). Thus, TR BAC-derived virus had lost ORFs IRS1 and US1, along with the remainder of US2 and part of the BAC vector, within as few as 2 passages in fibroblasts.

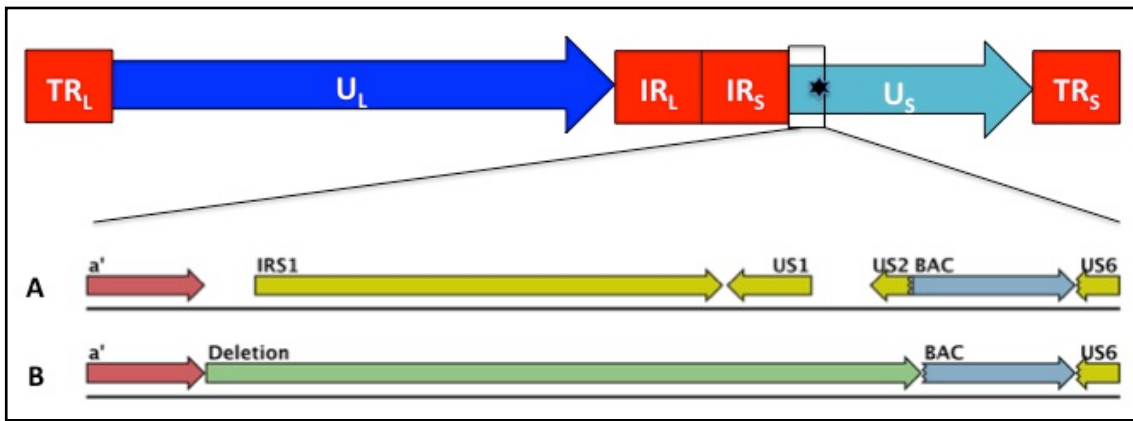


Figure 5.2. TR-BAC derived virus acquired a large deletion at the site of BAC vector insertion during passage in fibroblasts. Schematic of the BAC-cloned HCMV strain TR genome; black star indicates the site of BAC vector insertion. (A) TR-BAC *a'* – US6 genome region. (B) TR-HFFFp2 *a'* – US6 genome region. Sequence deleted (green) originated from the *a'* region termini (red) and spanned ORFs (yellow) up to and including part of the BAC vector (grey). Average coverage depth = 478 nt.

5.1.2 Virus reconstituted from the TR-BAC adapted during a single passage in RPE-1s

We next examined the TR BAC-derived virus from RPE-1 cells (RV-TR-RPE-1p1) that had also apparently undergone adaptation during the growth kinetics assays (section 3.2). Genomic DNA was again prepared from HFFFs infected at high MOI with RV-TR-RPE-1p1, and analysed by whole genome sequencing (Figure 5.3.). The UL128L genome region in this TR-BAC-derived virus contained identical sequence to that of the parental BAC, as did RL13. The majority population was that of virus containing mutations at two genomic loci: a frame-shifting mutation in UL84, following the insertion of an additional T into a tract of 8 T residues; and a deletion from within US9 and encompassing ORFs up to and including part of US16.

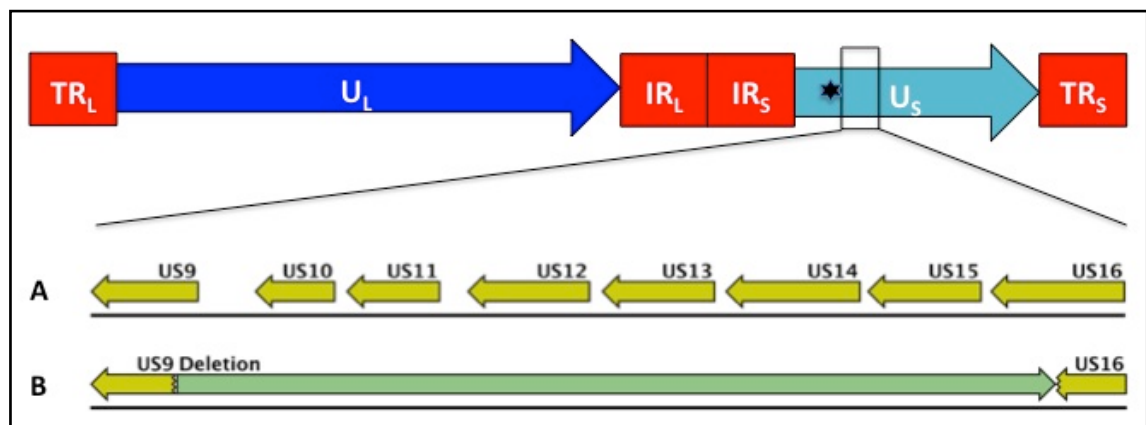


Figure 5.3. TR-BAC derived virus acquired a large deletion proximal to the site of BAC vector insertion during passage in epithelial cells. Schematic of the TR-BAC genome as previously described; the black star indicates the site of BAC vector insertion. (A) TR-BAC US9-US16 genome region. (B) RV-TR-RPE-1p1 US9-US16 genome region. Sequence deleted (green) originated from within US9 and spanned all ORFs (yellow) up to and including part of US16. Bases with coverage = 99.945%. Average coverage depth = 207.117 nt.

5.1.3 Conclusion

TR BAC-derived virus had mutated within a single passage following reconstitution in both fibroblasts (HFFF) and epithelial cells (RPE-1s), and in both instances, this involved the deletion of sequences that included, or were proximal to, the site of the stably incorporated BAC vector sequence insertion.

5.2 Genetic stability of viruses in epithelial cell culture

The demonstration that TR BAC-derived virus underwent adaptation as early as the first passage following reconstitution in RPE-1s invoked the question as to whether other viruses were also prone to rapid mutation in this cell type. To answer this question, we examined the virus stocks grown in RPE-1 epithelial cells and used in previous investigations (Table 4.1). An exact assessment of the number of passages each virus had undergone was not possible, however each was subjected to the minimal passage history required for the growth of a stock from a BAC; productive infections from which each stock was prepared were initiated by inoculating RPE-1s in T25 culture flasks with RPE-1p1 viruses produced in the growth kinetics assays. When all cells were infected (i.e. eGFP+), harvested cells were diluted ~1:5 with fresh RPE-1s before being seeded to new culture flasks (firstly to T75 culture flasks, then to single T150 culture flasks, and finally to 5 T150 culture flasks).

Genomic DNA from each virus stock was again prepared from high MOI infection of HFFFs, other than that from the RV-Me-UL128L^{wt} and RV-Me-UL128L³³⁰¹ stocks. The low titre of these stocks were sufficient only for the performance of low MOI HFFF infections, and the multiple rounds of lytic virus replication that would be required to generate a sufficient

abundance of genomes could result in the selection of mutated progeny at population levels that were not representative of the virus stock. Thus, we attempted sequencing of viral genomic DNA extracted directly from these virus stocks.

5.2.1 Viruses containing stably-incorporated BAC-vector sequences rapidly adapt during propagation in epithelial cell culture, involving mutation at the site of BAC-vector insertion

Stocks of RV-TR were prepared from the same virus lineage that contained the UL84 mutation and deletion of US9-US16 described above; these mutations were identified in ~80% of all genomes derived from the RV-TR-RPE-1 stock (Table 5.1). However, no further genetic changes had occurred during the additional passaging in RPE-1s prior to stock preparation. Like strain TR, TB40-BAC4 also contains a stably incorporated BAC vector in U_S genome region: the left hand end of the BAC interrupts US2 that is fused to the a' sequence; the right hand end of the BAC replaces sequence within the US6-US7 intergenic region (Sinzger, Hahn et al. 2008). RV-TB40-BAC4-RPE-1 stocks contained similar *de novo* mutant populations to those identified in the RV-TR-RPE-1 stocks. The majority population (~80%) had undergone a large deletion resulting in the loss of the entire BAC vector sequence, as well as ORFs in flanking genome regions (to the left of the BAC vector, the loss of the US2 fragment; to the right of BAC vector, loss of US7, US8 and part of US9) (Figure 5.4.). Also similar to the RV-TR-RPE-1 stock, ~1/2 of all genomes in the RV-TB40-BAC4-RPE-1 stock contained a premature stop codon in UL84.

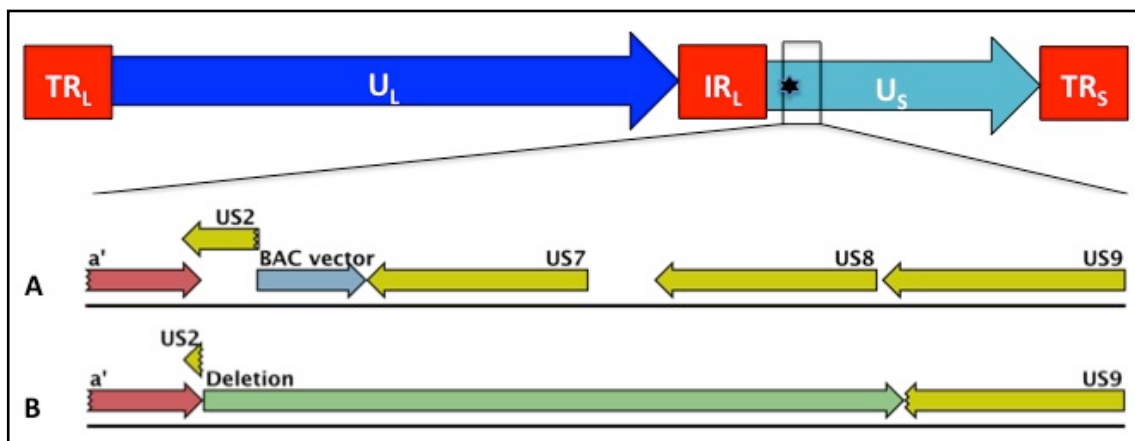


Figure 5.4. TB40-BAC4 BAC derived virus acquired a large deletion at the site of BAC vector insertion during passage in epithelial cells. Schematic of the TB40-BAC4 genome; repeated sequence regions (red box); unique sequence regions depicted by arrows (U_L , dark blue; U_S light blue) in relative orientation of ORF blocks; site of stably incorporated BAC at US2 to US6-US7 intergenic region (black star). (A) TB40-BAC4 a' -US9 genome region. (B) RV-TB40-BAC4-RPE-1p5 genome region. Sequence deleted (green) originated from the a' (red) terminus that overlapped the remainder of the ORF US2, and spanned the entire BAC vector (grey) as well as ORFs (yellow) up to and including part of US9. Bases with coverage = 99.907%. Average coverage depth = 383.475 nt.

5.2.2 Viruses from the Merlin-BAC variants rapidly adapt during propagation in epithelial cell culture, involving mutation in the U_L/b' genome region

Unlike the TR and TB40-BAC4 BAC clones, the BAC vector inserted into the Merlin genome (at the US28-US29 intergenic region) is excised following reconstitution of virus in eukaryotic cells, with the only remnant being a single 40 bp LoxP binding site. In turn, the genome region in which the BAC vector sequence was inserted remained intact in all Merlin viruses propagated in epithelial cells. However, stocks of all but one of the Merlin viruses (RV-Me-UL130^{A>G}-RPE-1) grown in RPE-1s contained *de novo* mutant populations, predominantly affecting the U_L/b' genome region.

In the stock of the Merlin virus containing native wildtype UL128L (RV-Me-UL128L^{wt}-RPE-1), a minority population (<3%) contained a deletion from within UL128 and extending into UL136 (Figure 5.5.). This deletion occurred with a concomitant insertion of an 11895 bp fragment of DNA that contained sequence from the HCMV ORF US20 flanked by sequences from *E.coli*.

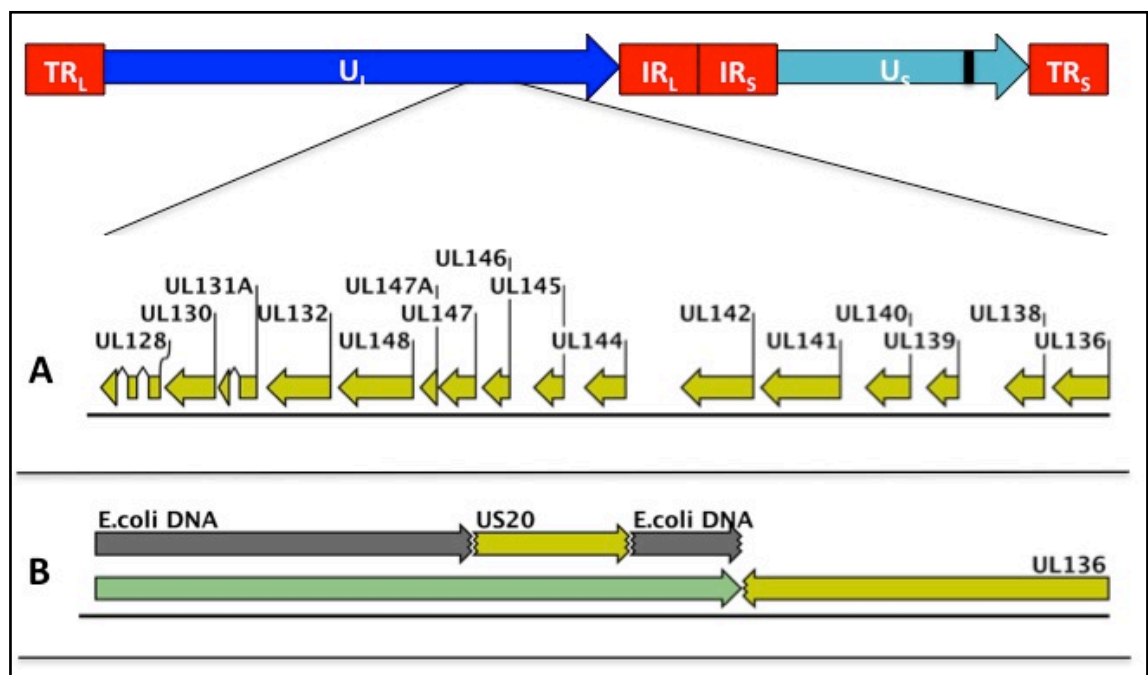


Figure 5.5. Minority mutant population present in RV-Me-UL128L^{wt}-RPE-1 stocks. Schematic of the BAC-cloned wildtype Merlin genome; repeated sequence regions (red box); unique sequence regions depicted by arrows (U_L, dark blue; U_S light blue) in relative orientation of ORF blocks; site of self-excising BAC vector at US28-US29 intergenic region (black line). (A) ORFs (yellow) UL128-UL136 in the Merlin BAC U_L/b' genome region. (B) Sequence deleted (green) in RV-Me-UL128L^{wt}, spanning UL128-UL138, and part of UL136, occurring with concomitant insertion of *E.coli* DNA (dark grey) and duplication of part of US20. Bases with coverage = 99.994%. Average coverage depth = 138.897 nt.

Stocks of recombinant Merlin viruses containing UL128L sequence from other strains each contained more significant mutant populations (Figure 5.6.). The UL141 ORF was consistently compromised across all Merlin viruses grown in epithelial cells, either due to large deletions encompassing multiple ORFs in the U_L/b' genome, or by the acquisition of mutations affecting this ORF alone. In RV-Me-UL128L³³⁰¹-RPE-1 stocks, deletion of ORFs UL147–UL150 occurred with the concomitant insertion of duplicated sequence from U_L/b genome region (TR_{L/b} sequence-RL12) (Figure 5.6.B). RV-Me-UL128L^{TB40}-RPE-1 stocks contained two mutant populations: the major population contained a deletion of sequence from within UL145 and spanning all ORFs up to, and including UL148C, with concomitant insertion of *E.coli* DNA (Figure 5.6.C.i); the minor population had acquired a deletion of sequence commencing from within UL142 and spanning all ORFs up to, and including, part of UL139 (Figure 5.5.6.ii). Each of RV-Me-UL128L^{G>T} and RV-Me-UL128L^{FIX} contained majority populations in which UL141 alone was compromised (Figure 5.6.D & E).

Mutations outside of the U_L/b' genome region were identified in the Merlin virus containing the UL128 G>T mutation (RV-Me-UL128L^{G>T}-RPE-1), with amino acid substitutions affecting UL26 and also UL75 in ~1/2 of all genomes. A population containing US8 compromised by an amino acid substitution was also identified in this virus, though this represented relatively few of the genomes analysed.

5.2.3 Conclusion

When passaged in epithelial cells, viruses that contained stably incorporated BAC vector sequences consistently acquired large deletions in genome regions that included, or were proximal too, the site of BAC vector insertion. Mutations affecting UL84 were also identified in these viruses. In contrast, the passage of various Merlin viruses resulted in the rapid selection of mutations affecting ORFs in the U_L/b' genome region, as well as occasional mutations at other genomic loci.

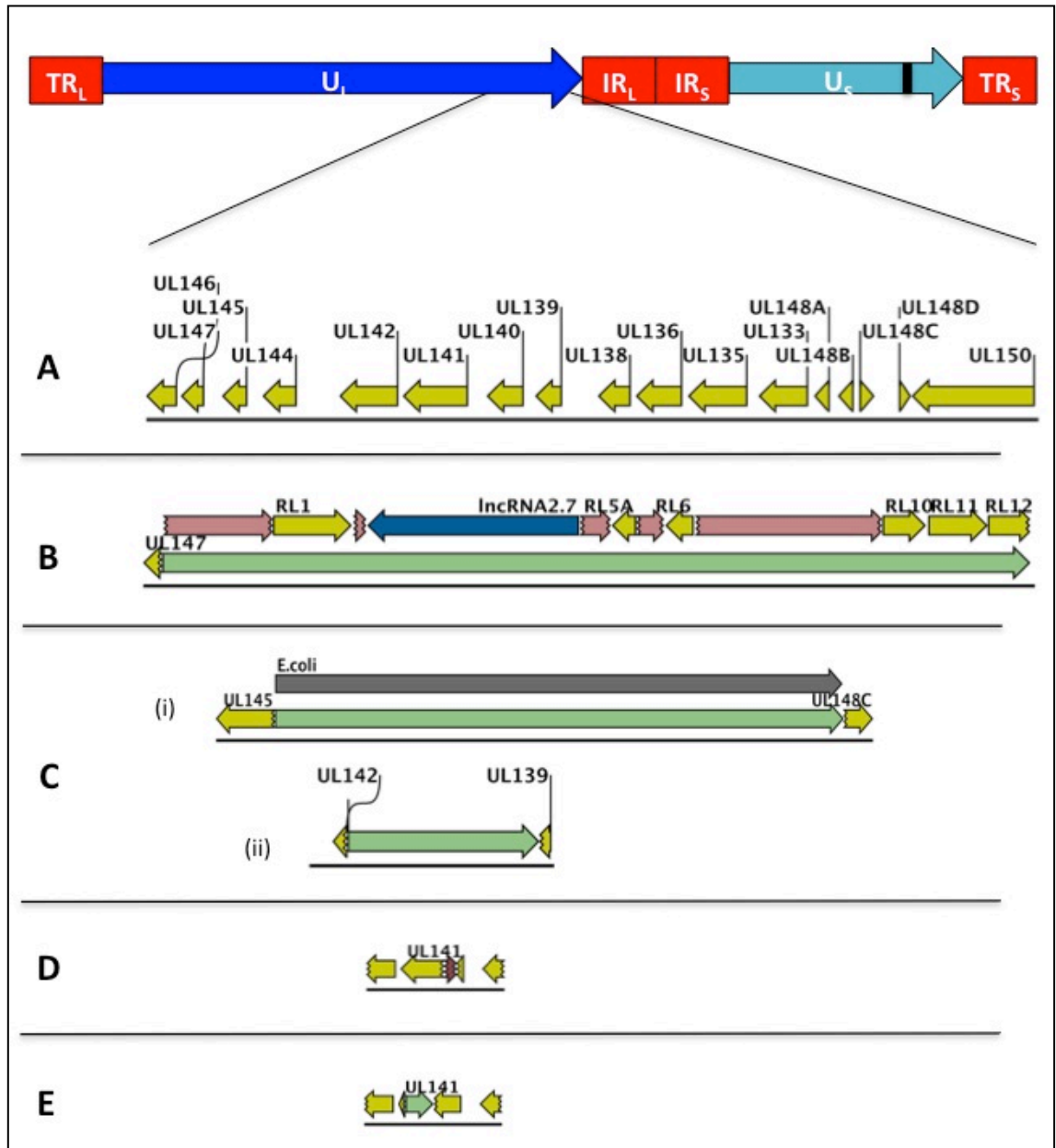


Figure 5.6. Mutation in U_L/b' genome region of Merlin viruses grown in epithelial cells. Schematic of the BAC-cloned wildtype Merlin genome, as previous. **(A)** ORFs UL147-UL150 in the Merlin-BAC variants in the U_L/b' genome region. Mutated regions in each passaged virus are aligned to corresponding region in U_L/b' region, though not to scale. Deleted regions are depicted in green. **(B)** Deletion of sequence from within RV-Me-UL128L³³⁰¹ UL147 and spanning ORFs up to and including UL150, occurring with concomitant insertion of duplicated sequence (pink) (UL to RL12). Bases with coverage = 99.956%. Average coverage depth = 179.172 nt. **(C)** Deletion of U_L/b' ORFs in RV-Me-UL128L^{TB40} (i) Major population – loss of part of UL145 and ORFs up to and including UL148B, with concomitant insertion of *E.coli* DNA; (ii) minor population – loss of part of UL142 and ORFs up to and including part of UL139. Bases with coverage = 99.855%. Average coverage depth = 769.865 nt. **(D)** Acquisition of amino acid substitution encoding SNP (purple) in RV-Me-UL128L^{G>T} UL141. Bases with coverage = 99.975%. Average coverage depth = 396.27 nt. **(E)** In frame deletion of RV-Me-UL128L^{FIX} UL141. Bases with coverage = 99.919%. Average coverage depth = 546.484 nt.

Table 5.1. Mutant populations in virus stocks prepared in RPE-1 epithelial cells.

Virus Stock	Location (nt) ^a	ORFs affected	Mutation/(Coding Effect) ^b	Population (%)
RV-TR-RPE-1-RPE-1	3,687-9,738	US9-US16	Δ 6052 bp	80
	159,574-159,581	UL84	Frame-shifting insertion: TTTTTTTT>TTTTTTTTT	80
RV-TB40-BAC4-RPE-1	228,878 – 1973	<i>α</i> '-US9	Δ 173 bp of US2, BAC vector, and 1,973 bp over-lapping US7, US8 and part US9	80
	155,174	UL84	G>A (Q>Stop)	46
RV-Me-UL128L ^{wt} -RPE-1	176,129-188,023	UL128-UL136	Δ 11,895 bp; insertion 1,217 bp (<i>E.coli</i> DNA-UL20 [25,726-26,006]- <i>E.coli</i> DNA)	< 3
RV-Me-UL128L ³³⁰¹ -RPE-1	180,563-194,029	UL147-UL150	Δ ~13,750 bp; insertion 10,300 bp from UL-RL12 [1-10,300]	45
RV-Me-UL128L ^{TB40} -RPE-1	183,591-186,632	UL145-UL148C	Δ 3,042 bp	90
	181,910-191,511	UL142-UL139	Δ 9,602 bp; insertion 3733 bp <i>E.coli</i> DNA	9
RV-Me-UL128 ^{G>T} -RPE-1	32,652	UL26	G>T (R>S)	50
	110,792	UL75	G>C (L>M)	50
	185,141	UL141	C>T (W>Stop)	65
	204,004	US8	A>C (V>G)	low
RV-Me-UL128L ^{FIX} -RPE-1	184,459-184,914	UL141	Δ 456 bp (In-frame Deletion)	100

^a Location/Nucleotide co-ordinates in TR BAC (AC146906), TB40-BAC4 BAC (EF29999), and Merlin BAC (GU179001) genome sequences deposited in GenBank. The sequences of strains TR and TB40-BAC4 deposited in GenBank commence immediately following the site of BAC vector insertion. The positions of deletions in viruses derived from these clones are described to include sequence at each end of the deposited sequences.

^b Mutation and coding effect: Δ indicates deletion; bp – base pairs. Numbers enclosed in [] detail the origin of HCMV sequences duplicated in mutant virus genomes.

5.3 Genetic stability of viruses in fibroblast cells

In contrast to viruses grown in epithelial cells, whole genome sequencing of the RV-Me-UL128L^{mut} and RV-FIX stocks grown in fibroblasts identified no populations that had acquired *de novo* mutations. These stocks were prepared by a similar method as stocks produced in RPE-1s, though were subjected to 1 fewer passage; productive infections were initiated by transfection of HFFFs with BAC DNA, and not infection of cells with virus derived from the growth kinetics assays. Despite this, the lack of any mutants in either of these stocks suggested that either the selection of mutants occurred less rapidly in fibroblasts, or the lack (or reduced levels) of pentameric complex and increased fitness of virus in this cell type (greater cell-to-cell spread, cell-free virus production, and dissemination kinetics) reduced the chance of mutation.

To investigate these possibilities, the genetic stability of all viruses following successive passages in HFFFs was assessed. Where possible, the passage series of each virus was a continuation of infections from the previously described growth kinetics assays. Each successive infection was initiated by inoculating fresh HFFFs with cell-free virus collected when the previous passage was completed (i.e. when mono-layers were destroyed). Infectivity in supernatants was quantified by immunofluorescence titration, and all new infections were performed at an approximate MOI of 0.05-0.1. Viral genomic DNA was extracted directly from virions released into the supernatant at the end of passage 5, and assessed by whole genome sequencing.

5.3.1 Viruses containing stably-incorporated BAC-vector sequences rapidly adapt during sequential passage in fibroblasts, involving mutation at the site of BAC-vector insertion

TR BAC-derived virus had previously undergone adaptation within a single passage in fibroblasts, involving the deletion of sequences at the site of BAC vector insertion. The RV-TR-HFFFp5 assessed in this investigation was derived from a separate transfection of HFFFs, yet contained populations with near identical mutations to that of the previous RV-TR-HFFFp2 virus. Two distinct populations were detected, each containing deletions from the *a'* genome region terminus to within the BAC vector, and differing only as to the size of the deletion: the major population lost 6087 bp, whilst the minor population lost 4306 bp (Table 5.2). The RV-TB40-BAC4-HFFFp5 virus also contained a major population that lacked sequence originating from within the *a'* genome region, spanning the entire BAC vector sequence, and also part of US7.

The FIX BAC-clone also contains a stably incorporated BAC vector in the U_S genome region (Hahn, Khan et al. 2002), with the BAC vector interrupting the US2 ORF that is fused to

the end of the *a* sequence at the outermost U_S genome segment terminus. RV-FIX-HFFFp5 virus contained a mutant population in which a large deletion had removed sequence commencing immediately outside of the terminal *a* sequence, spanning the entire BAC vector sequence, and removing part of the remaining US2 ORF (Figure 5.7.). Thus like viruses produced from TR and TB40-BAC4, FIX BAC-derived virus is prone to mutation involving the site of the stably incorporated BAC insertion.

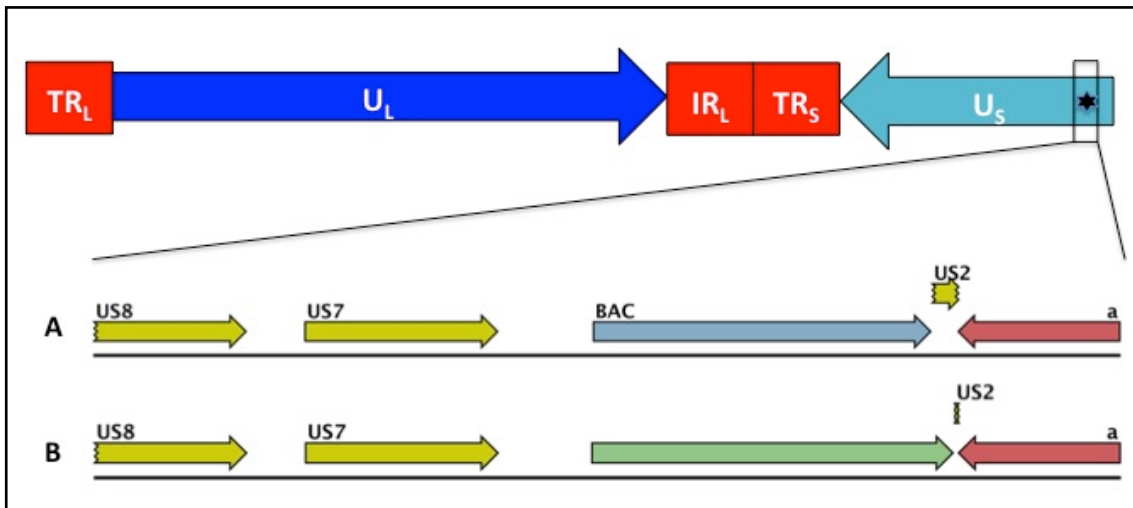


Figure 5.7. Deletion of BAC vector in RV-FIX-HFFFp5. Schematic of the FIX-BAC genome; repeated sequence regions (red box); unique sequence regions depicted by arrows (U_L , dark blue; U_S light blue) in relative orientation of ORF blocks; site of stably incorporated BAC vector at US2 to US6-US7 intergenic region (black star). ORFs in the U_S segment in reverse orientation relative to that of the U_L genome segment. The FIX BAC clone lacks any IRS genome region at the U_L/U_S junction, which instead contains the single inverted copy of the TRS sequence. (A) FIX-BAC genome region spanning from US8 to the terminal *a* sequence. (B) RV-FIX-HFFFp5 US8 to terminal *a* sequence region. Sequence deleted (green) originated from the *a* (red) terminus fused with the remainder of the ORF US2, and spanned the entire BAC vector (grey), as well as 7 bp in the US6-US7 intergenic region. Bases with coverage = 99.481%. Average coverage depth = 562.065 nt.

5.3.2 Virus from the Merlin-BAC variant lacking intact UL128L is stable in fibroblasts; virus from the Merlin-BAC variants containing wildtype UL128L are prone to rapid mutation in this genome region, and others, during sequential passage in fibroblasts

No mutant populations were detected in the sequentially passaged Merlin virus lacking UL128L (RV-Me-UL128L^{mut}-HFFFp5). In contrast, Merlin viruses containing wildtype UL128L each contained mutations in the encompassing this genome region (Figure 5.8.). The vast majority of genomes (~92%) in the RV-Me-UL128L^{wt}-HFFFp5 virus contained a deletion of sequence originating from within UL128 up to, and including UL139, as well as sequence from within the U139-UL138 intergenic region (Figure 5.8.Bi). A smaller population (~6%)

also contained a deletion of 4 bp within UL131A (Figure 5.8.Bii). RV-Me-UL128L³³⁰¹-HFFFp5 virus contained a mutant population that had acquired a frame-shifting single nucleotide insertion adjacent to a nucleotide substitution in UL128 exon 1 (Figure 5.8.C). Each of these stocks was derived from HFFFp1 progenitor viruses produced during growth kinetics assays. In these assays, each of the Merlin viruses containing wildtype UL128L had apparently undergone adaptation during the first passage in fibroblasts, as evidenced by the production of unexpectedly high titres of cell-free virus, and unusually large HFFF plaques formed during titration of supernatants. The RV-Me-UL128L³³⁰¹-HFFFp5 virus also contained a G>A nucleotide substitution located in the intergenic region between UL54 and UL55.

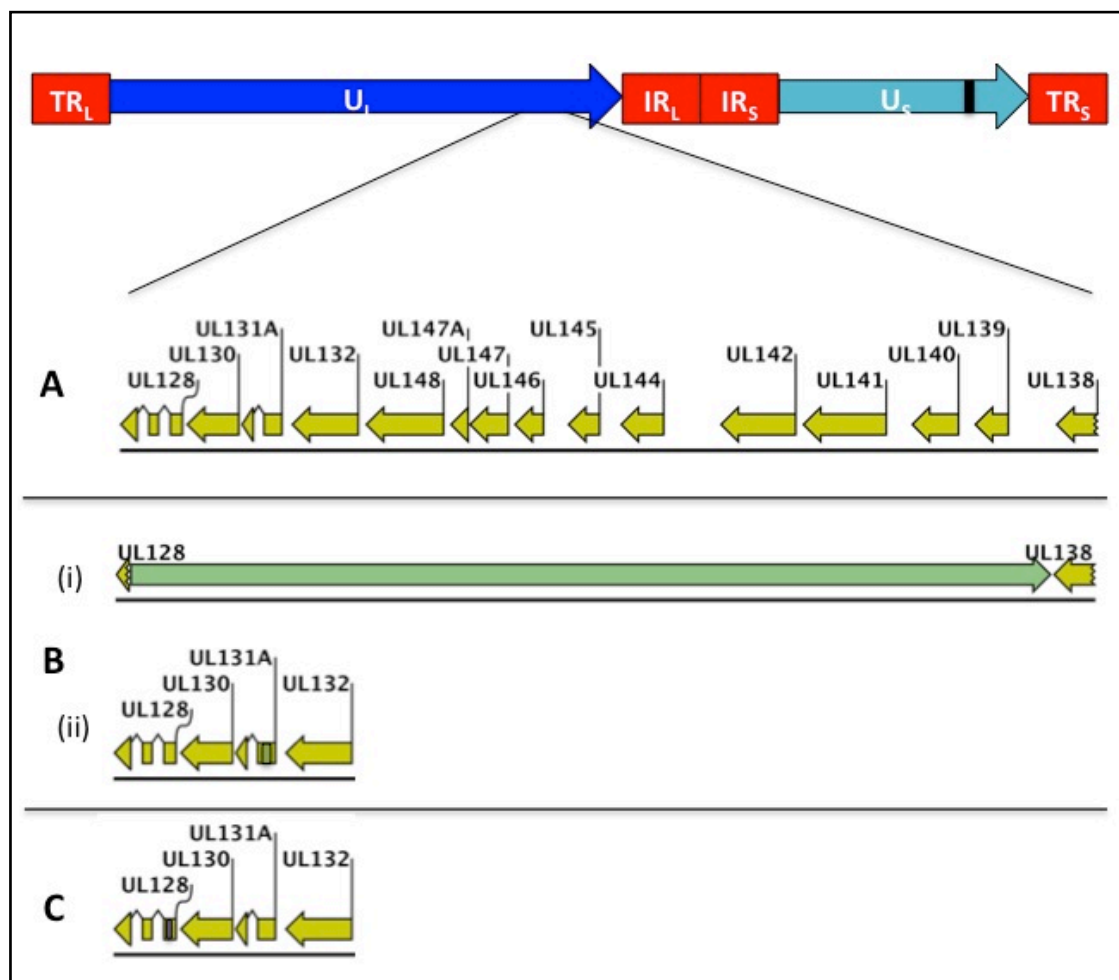


Figure 5.8. Mutation in U_L/b' genome region of Merlin containing wildtype UL128L when grown in fibroblasts. Schematic of the Merlin genome, as previous. (A) ORFs UL128-UL138 in the Merlin-BAC variants in the U_L/b' genome region. Mutated regions in each passaged virus are aligned to corresponding region in U_L/b' region, though not to scale. Deleted regions are depicted in green. (B) Deletion of sequence from within RV-Me-UL128L^{wt} (i) major population – loss of sequence commencing from within UL128 and spanning ORFs up to and including UL139, and ending in the UL139-UL138 intergenic region; (ii) small deletion of 4 bp within UL131A exon 1. Bases with coverage = 100%. Average coverage depth = 594.686 nt. (C) Single nucleotide substitution and frame-sifting nucleotide insertion in RV-Me-UL128L³³⁰¹ UL128 exon 1. Bases with coverage = 100%. Average coverage depth = 533.51 nt.

5.3.3 TB40-BAC4 UL128L is not stable when expressed in the Merlin genome background in fibroblasts, yet the unique UL128 intron 1 G>T variation does stabilise Merlin UL128L ORFs in fibroblasts

Whole genome sequencing of the sequentially passaged Merlin virus containing TB40-BAC4 UL128L (RV-Me-UL128L^{TB40}-HFFFp5) identified a population (~60%) in which a deletion had occurred in the U_L/b' genome region, resulting in the loss of sequence from within UL130 up to and including part of UL148 (Figure 5.9.). In contrast, Merlin virus containing the G>T mutation identified in TB40-BAC4 UL128 contained no mutations in the U_L/b' genome region. In fact, the RV-Me-UL128L^{G>T}-HFFFp5 virus contained relatively few mutations at all, with amino acid substitution encoding SNPs in just ORFs UL47 and UL55.

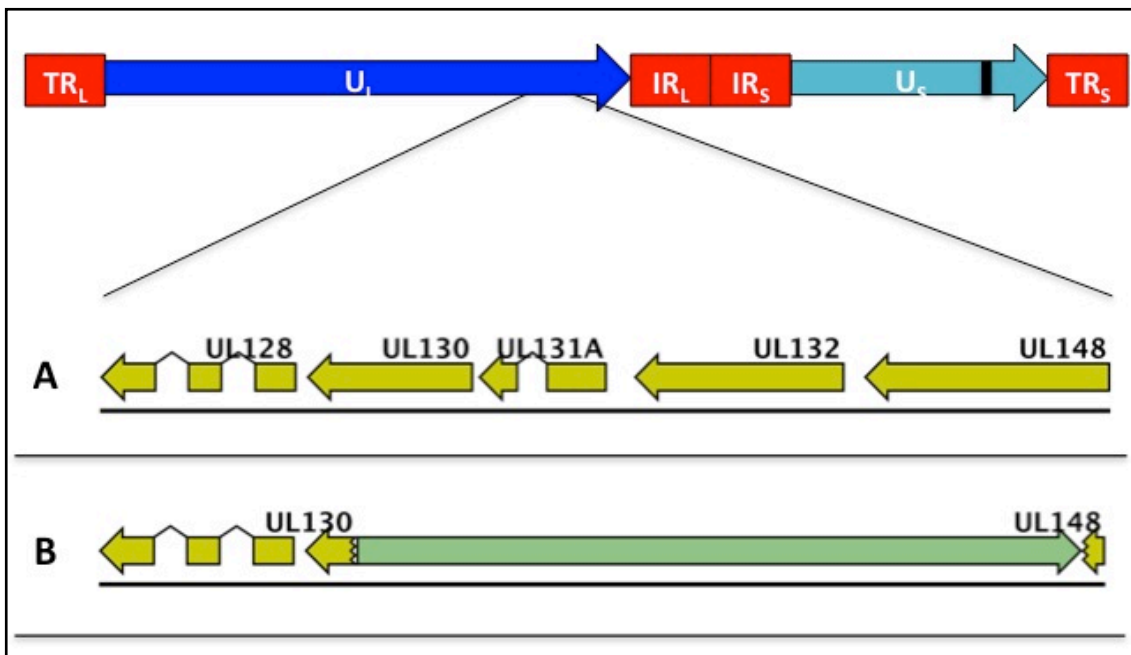


Figure 5.9. Mutation in U_L/b' genome region of Merlin containing TB40-BAC4 UL128L when grown in fibroblasts. Schematic of the Merlin genome, as previous. (A) ORFs UL128-UL148 in the Merlin-BAC variants in the U_L/b' genome region. Mutated regions in each passaged virus are aligned to corresponding region in U_L/b' region, though not to scale. Deleted regions are depicted in green. (B) Deletion of sequence from within RV-Me-UL128L^{TB40} originating from within UL130, and spanning ORFs up to, and partially including UL148. Bases with coverage = 99.966%. Average coverage depth = 463.882 nt.

5.3.4 FIX-BAC UL128L is stable when expressed in the Merlin genome background in fibroblasts, and the unique UL130 A>G variation also stabilises Merlin UL128L ORFs in fibroblasts

All genomes detected in the RV-Me-UL128L^{FIX}-HFFFp5 virus contained sequence that was identical to the parental BAC, and therefore *de novo* mutations had not been acquired during the passage of this virus. This result was closely matched by that of the Merlin virus containing the UL130 A>G mutation identified in FIX UL128L, with the only mutation identified in RV-Me-UL130^{A>G}-HFFFp5 being a deletion within the synthetic eGFP cassette inserted upstream of UL122.

5.3.5 Conclusion

As in epithelial cell culture, viruses containing stably incorporated BAC vector sequences were prone to the acquisition of large deletions surrounding the site of BAC vector insertion. Wildtype UL128L was rapidly mutated following growth of virus in fibroblasts, regardless of natural strain-strain sequence variation. Like TR UL128L (see Appendix II), TB40-BAC4 UL128L was stable in the TB40-BAC4 background, but was prone to mutation in when expressed in the Merlin genome. However, the FIX UL128L genome region that is more severely compromised than UL128L from other strains was not prone to mutation, whether expressed in the native FIX genome background, or in the Merlin genome background. The UL128L genome region of Merlin viruses engineered to contain mutations from TB40-BAC4 UL128 or FIX UL130 were more stable than wildtype UL128L in the parental Merlin virus in fibroblasts, although point mutations were occasionally acquired elsewhere in these genomes of these viruses.

Table 5.2. Mutant populations in viruses produced following by sequential passage in fibroblasts (HFFFp5).

Virus Stock	Location (nt) ^a	ORFs affected	Mutation/(Coding Effect) ^b	Population (%)
RV-TR-HFFFp5	194,202-BAC	<i>a'</i> -US3	Δ 6,087 bp: 3,899 bp from <i>a'</i> - to BAC vector, + 2,188 BAC nt	>
			Δ 4,306 bp: 3,899 bp from <i>a'</i> - to BAC vector, + 407 BAC nt	<
RV-TB40-BAC4-HFFFp5	228,878-446	US7	Δ173 bp: <i>a'</i> - BAC vector, entire BAC vector, 446 bp into US7	
RV-FIX-HFFFp5	228,002-60		Δ 7,650 bp: Sequence from <i>a'</i> to BAC vector + 7 bp	
RV-Me-UL128L ^{wt} -HFFFp5	176,320-187,358	UL128 – UL139	Δ 11,039 bp	92
	178,063-178,066	UL131A	Δ 4 bp; Frame-shift	6
RV-Me-UL128L ³³⁰¹ -HFFFp5	176,827-176,828	UL128 exon 1	Insertion: G; Frame-shift	
	176,827		Substitution: C>T	
RV-Me-UL128L ^{TB40} -HFFFp5	177,130-179,950	UL130 – UL148	Δ 2,807 bp	60
	n/a	eGFP	Substitution: T>A	
RV-Me-UL128 ^{G>T} -HFFFp5	61,647	UL47	Substitution: G>C (D>H)	
	83,915	UL55	Substitution: T>A (D>V)	
RV-Me-UL130 ^{A>G} -HFFFp5	n/a	eGFP	Δ 720 bp	99

^a Location/Nucleotide co-ordinates in TR BAC (AC146906), TB40-BAC4 BAC (EF29999), and Merlin BAC (GU179001) genome sequences deposited in GenBank. The sequences of strains TR and TB40-BAC4 deposited in GenBank commence immediately following the site of BAC vector insertion. The positions of deletions in viruses derived from these clones are described to include sequence at each end of the deposited sequences.

^b Mutation and coding effect: Δ indicates deletion; bp – base pairs. Numbers enclosed in [] detail the origin of HCMV sequences duplicated in mutant virus genomes.

5.4 Stabilisation of genomes by repression of ORFs most inhibitory to virus propagation *in vitro*

Previous work demonstrated the utility of the tetracycline repressor culture system for the growth of virus containing wildtype RL13 and UL128L ORFs to high titres, comparable to that of virus containing ablation mutations in these genome regions (Stanton, McSharry et al. 2008). We speculated that tet-repression might stabilise these ORFs that are normally rapidly mutated when virus is passaged in fibroblasts, and as suggested by the exquisite stability of the Merlin virus lacking intact RL13 and UL128L (RV-Me-UL128L^{mut}) in fibroblasts, other genome regions that are prone to mutation in this cell type.

5.4.1 Genetic stability of virus produced in tetracycline repression culture system

Alternative versions of Merlin-UL128^{G>T} and Merlin-UL130^{A>G} BACs were constructed, and similarly to the alternative version of Merlin-UL128L^{mut} that was used in previous 1-step growth curve assays (see section 4.5), each differed from previously used counterparts at only two genomic loci: at UL122, no eGFP expression cassette was inserted; in the UL131A-UL132 inter-genic region, binding sequences for the tetracycline operator were inserted. One further virus was analysed, containing wildtype RL13 as well as wildtype UL128L (RV-Me-RL13^{wt}/UL128L^{wt}). This virus contained no eGFP cassette, and binding sequences for tetracycline operators inserted were inserted both upstream of RL13, and in the UL131A-UL132 inter-genic region. Details of all viruses investigated are listed in Table 5.3. Stocks of each virus were grown in cells expressing tetracycline (HFFF-Tet), and virus genomic DNA was extracted directly from each stock and analysed by whole genome sequencing. No *de novo* mutant populations were identified in any stock, with all viruses displaying coding potential equivalent to the parental BACs.

Table 5.3. HCMV stocks grown in HF-tets.

Recombinant Merlin Virus	RL13	UL128L
RV-Me-tet-UL128L ^{wt}	Mutated; FS-ins ^a - A at nt 226 ^b	Wildtype
RV-Me-tet-UL128 ^{G>T}	Mutated; FS-ins ^a - A at nt 226 ^b	Wildtype + UL128 G>T
RV-Me-tet-UL130 ^{A>G}	Mutated; FS-ins ^a - A at nt 226 ^b	Wildtype + UL130 A>G (S72P)
RV-Me-tet-RL13 ^{wt} -tet-UL128L ^{wt}	Wildtype	Wildtype

^a – FS-ins – frame-shifting insertion.

^b – Nucleotide positions are given relative to sequence of BAC-cloned strain Merlin (GU179001).

5.4.2 Growth phenotype of virus produced during subsequent infection with virus stocks produced in tetracycline repression culture system

It was important to determine whether viruses produced in the tetracycline repression system could efficiently express RL13 and UL128L in subsequent infection of cell types where these ORFs were not repressed, ultimately producing progeny that contained gpRL13 and the pentameric complex. To explore this, one-step growth curve assays were performed to assess the growth phenotype of progeny produced following infection of 'normal' fibroblasts with viruses produced in the tetracycline repression culture system (Table 5.3.). The growth properties of the Merlin virus containing existing mutations in RL13 and UL128L (RV-Me-UL128L^{mut}) were also determined for comparison.

In previous assays, investigation into the relative CRV and CAV sub-populations relied on the collection of supernatant samples throughout the time-course of infection, but total cellular materials at the final time-point of infections. Surprisingly, no RV-Me-tet-UL128L^{wt} progeny sub-population was detected. This was potentially due to: i) the vast majority of cells being lysed at the time of CAV fraction preparation, and any progeny that was originally retained within cells may have ultimately been released; ii) virions are thermo-labile at standard culture conditions with a half-life of 24-48hrs (Christian Sinzger – Pers Comm.), and CAV infectivity that was retained at these conditions for extended times compared to CRV infectivity may have been subjected to greater decay. To more accurately reflect the relative CRV and CAV infectivity produced by infections in this repeat analysis, cellular materials cleared from supernatant samples collected every two days were stored for later analysis, rather than being returned to the original infection.

The Merlin virus containing wildtype UL128L ORFs again produced the lowest titres of cell-free virus, with the total CRV infectivity produced by RV-Me-tet-UL128L^{wt} 750-800-fold lower than that produced by RV-Me-UL128L^{mut} (Figure 5.10. – left axis). Merlin viruses encoding the single nucleotide variations identified in TB40-BAC4 UL128 intron 1 (RV-Me-tet-UL128^{G>T}) and FIX UL130 (RV-Me-tet-UL130^{A>G}) again produced greater total CRV titres compared to RV-Me-tet-UL128L^{wt}, yet less than produced by RV-Me-UL128L^{mut}. In comparison to previous assays, greater CAV infectivity was recovered from the infection with virus completely lacking the pentameric complex (RV-Me-UL128L^{mut}), representing ~55% of the total infectivity produced (Figure 5.10. – right axis); this difference likely reflects the different approach to CAV sample preparation between the repeat analyses. Furthermore, CAV infectivity recovered from the RV-Me-tet-UL128L^{wt} infection represented ~70% of the total infectivity produced, whereas the counterpart to this virus used in previous assays produced no detectable CAV. Importantly, the CAV produced by RV-Me-UL128L^{mut} was >380-fold greater than that produced by RV-Me-tet-UL128L^{wt}, providing further evidence to suggest that the difference in CRV production by these infections was not simply due to the pentameric complex

impeding the cell-free release of progeny. Merlin viruses encoding TB40-BAC4 and FIX UL128L mutational features, and containing intermediate quantities of virion-incorporated pentameric complex, each produced CAV infectivity that represented ~60% and 75% of the total infectivity produced in each corresponding infection respectively.

Consistent with a prior report (Stanton, Baluchova et al. 2010), the Merlin virus containing wildtype RL13 and UL128L ORFs (RV-Me-tet-UL128L^{wt}-tet-RL13^{wt}) produced the least cell-free virus; total CRV from this infection was >2500-fold lower than that produced by RV-Me-UL128L^{mut}, and also ~3.5-fold lower than produced RV-Me-tet-UL128L^{wt}. The CAV produced by the Merlin virus containing wildtype UL128L and RL13 was also the lowest of any virus assessed, however this represented ~80% of the total infectivity produced in this infection.

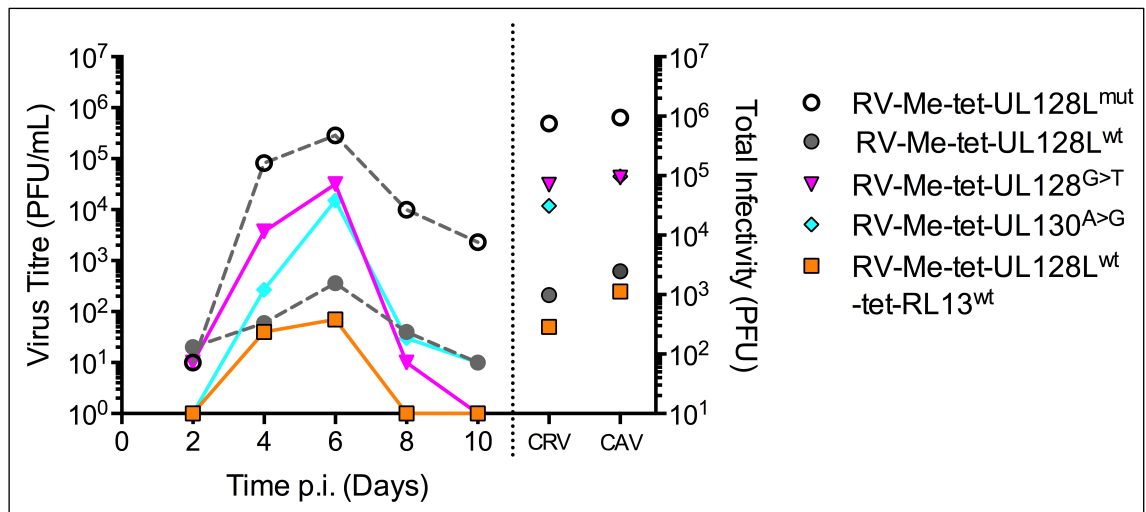


Figure 5.10. Growth phenotype of progeny produced by viruses grown in tetracycline repression culture system. HFFs were infected (MOI 5) with each indicated virus. Supernatants were collected every two days and centrifuged at low speed to collect cellular materials that were stored for later analysis. Infectivity within each sample was enumerated by IE1 immuno-fluorescence titration on HFFs. Cellular Left Y-axis: Infectivity within CRV supernatant samples collected over the time course of infection. Right Y-axis: CAV infectivity was investigated from pooled and sonicated cellular material samples, in parallel to pooled and sonicated CRV samples. To account for differences in the volumes of these reported the CRV and CAV sub-populations, total CRV infectivity was calculated from the titres of each sample.

5.4.3 Conclusion

Repression of genome regions most inhibitory to the growth of virus in fibroblasts allowed the stable propagation of virus without the selection of *de novo* mutations in any genome region, i.e. the normally rapidly mutated RL13 and UL128L ORFs were stabilised, as were other genome regions prone to mutation in fibroblasts when virus contains intact UL128L. Furthermore, as judged by growth phenotypes, virus produced in the tetracycline repression culture system still retained the ability to express RL13 and UL128L in subsequent infection of cell types where these ORFs were not repressed.

5.5 Summary

Whilst phenotype analysis provided indications that viruses had undergone adaptation to cell culture, the genome regions in which mutations were selected was not easily predicted. High-throughput whole-genome sequencing analysis provided a reliable assessment of the coding potential of viruses studied. Assuming random sampling of genomes, the average nucleotide coverage depth (at lowest, 138.897-fold for sequencing data directly the low-titre stock of RV-Me-UL128L^{wt} produced in RPE-1s) was sufficient to reliably depict the genetic diversity of each virus assessed. The incomplete genome sequences reported (ie nucleotides with coverage <100%) reflect the lack of data in high-G+C or repetitive regions (eg in UL57, OriLyt, or *b/b'*), or regions deleted from major genome populations.

Viruses derived from each of the TR-BAC, TB40-BAC4 and FIX-BAC clones repeatedly underwent rapid adaptation following reconstitution in both fibroblasts and epithelial cells, and this consistently involved the loss of sequence at, or proximal to, the site of BAC vector insertion. Moreover, this occurred within the minimum passaging required to generate virus stocks of each strain.

Merlin viruses were prone to adaptation when passaged in epithelial cells, most commonly involving the selection of mutants in the *U_L/b'* genome region, as well as occasionally at other genetic loci. In fibroblasts, Merlin viruses containing wildtype UL128L (whether native or from strain 3301) rapidly acquired ablating mutations in this genome region. In contrast, Merlin viruses containing TB40-BAC4 and FIX-BAC UL128L sequences were generally more stable in fibroblasts, compared to the parental Merlin viruses containing wildtype UL128L sequences. Importantly, incorporation of the TB40-BAC4 UL128 G>T or FIX UL130 A>G (S72P) single nucleotide variations stabilised the Merlin UL128L genome region. Thus, the more efficient growth properties that correlated with reduced virion pentameric complex content apparently alleviated the selective pressures for the acquisition of further UL128L mutations. However, these viruses often acquired mutations affecting other genome regions. Taken together, these data suggest that the stable passage of HCMV in cell culture may be best achieved in fibroblasts when UL128L is modified such that the growth inhibition of this genome region is less severe.

The Merlin virus containing existing ablating mutations in RL13 and UL128L remained stable in all genome regions following successive passaging in fibroblasts – this has been recapitulated in seven independent virus stocks derived from the Merlin-UL128L^{mut} BAC variant. Similarly, tet-repression of wildtype RL13 and UL128L ORFs permitted the propagation of viruses from the Merlin-BAC without selection of any *de novo* mutations. Growth phenotype analysis of progeny from viruses produced in this repression culture system indicated that RL13 and UL128L were unchanged and again expressed in normal fibroblasts.

6 The HCMV Virion Proteome

Conditional repression of wildtype RL13 and UL128L ORFs allowed the growth of genetically intact virus from the Merlin-BAC- to high titres. The growth phenotype of viruses following subsequent infection of fibroblasts where RL13 and UL128L are not repressed suggested that progeny virions that were produced contained gpRL13 and the pentameric complex. Importantly, the limited number of lytic replication cycles virus was subjected to following high MOI infection in 'regular' fibroblasts reduced the likelihood of the selection of *de novo* mutations. Ultimately, virions containing the full array of viral proteins encoded by the wildtype Merlin-BAC could be produced, and this provided an opportunity to gain valuable insights into the composition of wildtype HCMV virions. To investigate the full range of virus-derived proteins that are incorporated into virions, virions derived from Merlin viruses produced by this two-step propagation method were analysed by mass spectrometry.

The reduced pentameric complex content in the novel Merlin viruses containing UL128L sequences from TB40-BAC4 and FIX potentially impacted the composition of virions. Based on previous demonstration that a gO-null mutant that does not assemble the trimeric gH/gL complex (gH/gL/gO) produces virions with increased pentameric complex content (Wille, Knoche et al. 2010), I hypothesised that reduced pentameric complex content may occur with concomitant increase in trimeric complex (gH/gL/gO). Similarly, the presence or absence of gpRL13 in virions potentially affected the incorporation of other components, and it was also important to determine how our most commonly used variants (RL13^{mut}) differ from fully wildtype virus. Thus, as well as determining the absolute composition of virions, the relative abundances of viral proteins incorporated into virions different Merlin viruses. To this end, quantitative stable isotope labelling by amino acid mass spectrometry (SILAC-MS) was employed (Ong, Blagoev et al. 2002).

6.1 Optimising recovery of virions purified over gradients

Mass spectrometry analysis required a suitable abundance of purified virions. However, previous investigations demonstrated that gradient purification of concentrated virus stocks prepared by high-speed ultra-centrifugation resulted in limited recovery of purified virions (Section 4.5). This was at least partially due to virions localising with cellular debris at the sample: gradient interface, and also in lower bands predicted to contain DBs. The detection of cellular contaminants in the 'DB' gradient bands suggested that viral particles and contaminating cellular proteins were incorporated into aggregates that may have segregated virions from the 'purified virion' band, to lower regions of the gradient. One potential reason for this was the method of virus concentration; the compaction of cellular contaminants along with virus particles into a pellet during high-speed ultra-centrifugation of supernatants may have promoted aggregate formation. Therefore, alternative methods for the preparation of concentrated virus stocks were explored.

6.1.1 Concentration of virus by high-speed centrifugation (HSC) or Ultrafiltration (UF)

Ultra-filtration of infected cell-culture supernatants has been described as a suitable method for the preparation of concentrated herpesvirus stocks (Smith and De Harven 1973). In this approach, excess media is removed from infectious supernatants via membranes with pores that allow the passage of water and other small molecules (<1 MDa), but not macromolecular structures (e.g. HCMV particles). The elimination of small molecules via membranes can be driven by negative pressure dialysis (for larger volumes), or low-speed centrifugation (for smaller volumes). Initially, the recovery of infectivity in virus stocks concentrated by virus stocks prepared by high-speed centrifugation (as performed previously; HSC) or by dialysis ultra-filtration (dUF) was compared. Neither method resulted in any significant loss in infectivity (Figure 6.1.A).

6.1.2 Recovery and purification of virions following ultra-centrifugation of HSC and dUF concentrated viruses over gradients

HSC and dUF concentrated viruses were then purified over gradients, and the infectivity that localised to the 'purified virion' band and the lower 'DB' band of the different gradients compared (Figure 6.1.B). The 'purified virion' band from the HSC virus gradient contained as little as 3.5% of the total infectivity loaded, with ~1.75% localised to the 'DB' band. In contrast, the 'purified virion' band from the dUF virus gradient contained ~20% of the total infectivity loaded, whilst the 'DB' band contained <1% of the infectivity.

The purity of virions recovered following ultra-centrifugation of HSC and dUF concentrated viruses over gradients was also assessed (Figure 6.2.B). As a marker of virions and contaminating proteins, 'purified virion' and 'DB' band samples were analysed for the virion tegument protein pp28 and calnexin, respectively. Interestingly, whilst concentrated virus stocks prepared by either method each contained calnexin (Figure 6.2.C lanes 1 & 6), the levels were markedly reduced in the dUF concentrated virus. Thus, ultra-filtration of supernatants resulted in elimination of some contaminating proteins. The 'purified virion' band from both HSC and dUF concentrated virus gradients contained pp28, but not calnexin (Figure 6.2.C lanes 2 & 7), indicating that virions had been successfully purified from both concentrated stocks. Significantly, whilst the 'DB' band from the HSC virus contained readily detectable levels of pp28, the 'DB' band from dUF virus was completely devoid of pp28 (Figure 6.2.C lanes 3 & 8).

HSC and dUF viruses were also purified over sorbitol cushions (Figure 6.2.C lanes 4 & 9), or the 'light solution' used to form the uppermost region of gradients (Figure 6.2.C lanes 5 & 10); it was unclear whether the previously observed impurity of virus particles ultra-centrifuged over sorbitol was due to aggregate formation that occurred during concentration by high-speed centrifugation. However, calnexin co-purified with viral particles following ultra-centrifugation of both concentrated virus stocks over these solutions.

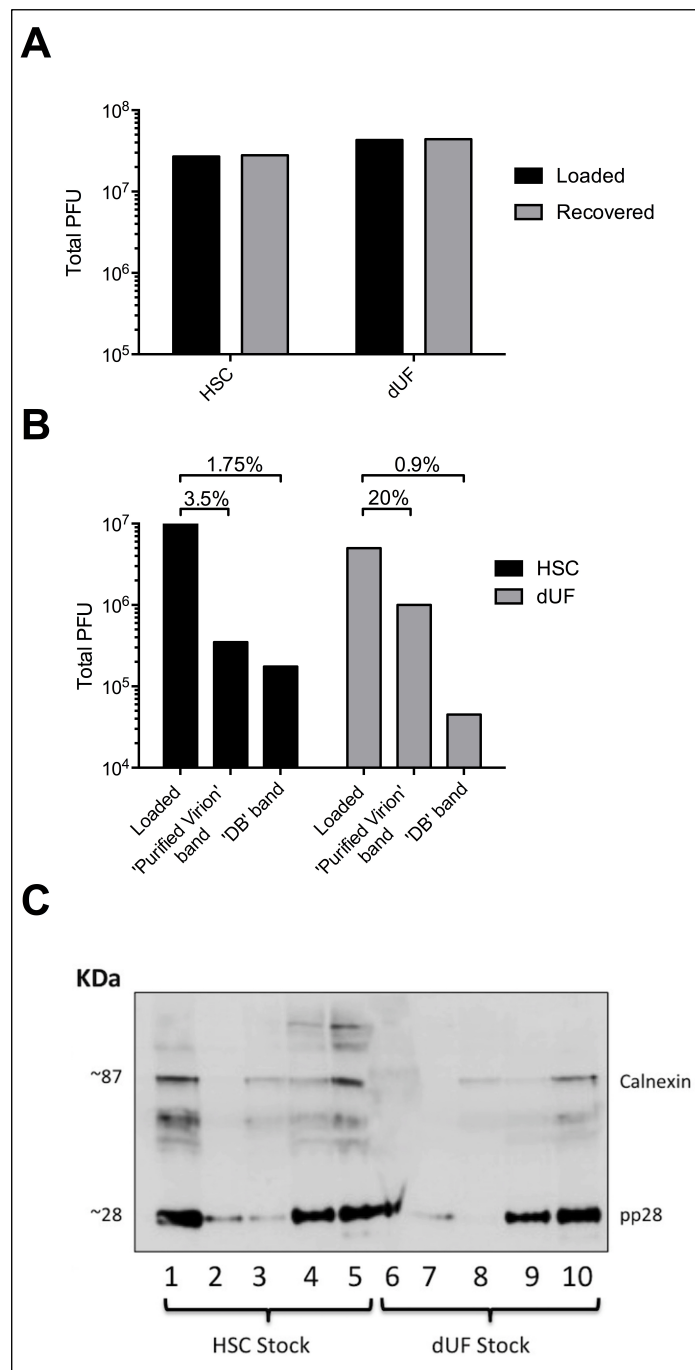


Figure 6.1. Recovery of infectivity following concentration of virus by high-speed centrifugation (HSC) or dialysis ultra-filtration (dUF), and subsequent purification over glycerol-tartrate gradients. RV-Me-UL128L^{mut} was grown in HFFs according to standard virus propagation procedures. Supernatants collected over the time-course of infection were initially cleared of large cellular debris by low-speed centrifugation. Aliquots were then concentrated ~40-fold by high-speed centrifugation (HSC), or ~20-fold by dialysis ultra-filtration (dUF), before being ultra-centrifuged over glycerol tartrate gradients. **(A)** The original (un-concentrated) supernatant, as well as HSC and dUF concentrated viruses were titrated onto HFFs in singlicate. To reflect differences in the volume of infectious supernatants concentrated by either method, and also the reduced volume of the concentrated virus stocks that were prepared, the total infectivity was estimated from the titres of each sample as reported on HFFs. **(B)** Following recovery from gradients, infectivity in 'purified virion' and 'DB' bands was enumerated by plaque titration onto HFFs. Total infectivity loaded to gradients and recovered from each band are reported. **(C)** Immunoblot analysis of gradient purified HSC or dUF concentrated viruses: 1 – HSC virus stock; 2 – HSC virus 'purified virion' band; 3 – HSC virus 'DB' band; 4 – HSC virus over gradient light solution; 5 – HSC virus over sorbitol cushion; 6 – dUF virus stock; 7 – dUF virus 'purified virion' band; 8 – dUF virus 'DB' band; 9 – dUF virus over gradient light solution; 10 – dUF virus over sorbitol cushion.

6.1.3 Conclusion

Whilst neither method of virus concentration resulted in significant loss of infectivity, 5-6-fold greater infectious-virion recovery was achieved following gradient purification of virus stocks concentrated by dialysis ultra-filtration compared to high-speed centrifugation. Taken together with the banding patterns of virions and contaminants from HSC and dUF concentrated viruses over gradients, the improved recovery of virions from gradient purified dUF virus was likely due to reduced aggregate formation during concentration by ultra-filtration.

6.2 Production and preparation of virions for mass spectrometry analysis

The panel of Merlin viruses analysed included 2 variants that contained ablating mutations in both RL13 and UL128L: RV-Me-UL128L^{mut}, and a derived virus that was also deleted for the ORF UL141 (RV-Me-UL128L^{mut}/ΔUL141). Stocks of each of these viruses were produced in ‘regular’ HFFFs. The remaining 4 viruses each contained intact, but tet-regulated UL128L genome regions, whilst one also contained wildtype tet-regulated RL13 (RV-Me-tet-RL13^{wt}/UL128L^{wt}). Stocks of these viruses were produced in the tetracycline repression culture system (HFFF-tet). All virus stocks used to initiate productive infections were assessed by whole genome sequencing and shown to be equivalent to the parental BACs. Quantitative SILAC-MS is performed by the simultaneous analysis of proteins labelled with different isotopes. Virions containing differently labelled proteins were produced following infection of HFFFs at high MOI (2-4), with each virus grown in SILAC media supplemented with ‘light’, ‘medium’ or ‘heavy’ amino acids (See Table 6.1.).

Table 6.1. Merlin virions analysed by mass spectrometry.

BAC Origin	Reconstituted Virus	SILAC label
Merlin-UL128L ^{mut}	RV-Me-UL128L ^{mut}	Medium
Merlin-tet-UL128L ^{wt}	RV-Me-tet-UL128L ^{wt}	Light
Merlin-tet-RL13 ^{wt} /tet-UL128L ^{wt}	RV-Me-tet-RL13 ^{wt} /tet-UL128L ^{wt}	Heavy
Merlin-tet-UL128 ^{G>T}	RV-Me-tet-UL128 ^{G>T}	Heavy
Merlin-tet-UL130 ^{A>G}	RV-Me-tet-UL130 ^{A>G}	Heavy
Merlin-UL128L ^{mut} /Δ UL141 ^a	RV-Me-UL128L ^{mut} /Δ UL141	Light

^a – Δ denotes deleted gene

Supernatants collected over the time-course of each productive infection were first cleared of large cellular debris by low-speed centrifugation, then concentrated by ultra-filtration: firstly by dialysis ultra-filtration (dUF), and subsequently by low-speed centrifugal ultra-filtration (cUF). Preliminary experiments demonstrated that viruses in SILAC media concentrated >15-fold were too dense to be layered onto glycerol: tartrate gradients, therefore all samples were concentrated 10-fold. Virions were purified over gradients and re-suspended in lysis buffer lacking DTT, and the total protein concentrations of samples determined by Micro BCA assay.

Two mixed virion preparations (MVP) containing equal protein loads from 3 different purified virion samples were analysed. To assess the impact of gpRL13 and pentameric complex incorporation on the composition of virions, the mixed virion preparation MVP-1 contained ‘light’ virions from RV-Me-tet-UL128L^{wt}, ‘medium’ virions from RV-Me-UL128L^{mut}, and ‘heavy’ virions from RV-Me-tet-RL13^{wt}/tet-UL128L^{wt}; to assess the specific impact of variations in pentameric complex content, the mixed virion preparation MVP-2 contained the same ‘light’ and ‘medium’ virions, but with ‘heavy’ virions from RV-Me-tet-UL128^{G>T}. Virions purified from each virus were also analysed separately. Samples were prepared as in Methods and Materials, then analysed by mass spectrometry analysis at the laboratory of Professor Paul Lehner (Cambridge institute for Medical research, University of Cambridge) under procedures as previously reported (Weekes et al, 2014), according to the protocol described in Appendix VI.

6.3 Merlin virion proteome

Mass spectrometry analysis is limited in its ability to identify all proteins in complex mixtures on each independent run. Whilst the Merlin viruses analysed differed in coding potential, they did so only in a small sub-set of ORFs (i.e. RL13, UL128, UL130, UL131A, UL141). Analysis of each virus individually, along with the two mixed samples, provided a total of 8 independent analyses of near identical Merlin virions. The absolute Merlin virion proteome was determined as the consensus of proteins identified across all virion preparations. Proteins were designated as being identified with ‘high confidence’ according to the following criteria: when a minimum of 2 peptides unique to that protein were detected, and when the peptides were present in at least half of all virus preparations analysed. An exception was made for gpRL13, gpUL128, gpUL130, gpUL131A and gpUL141, where ‘high confidence’ identification was based on identification of these proteins in at least half of the preparations containing virions from viruses in which the relevant ORFs were intact.

6.3.1 Virion proteins identified with high-confidence

70 HCMV proteins were identified with high confidence (Table 6.2.A). Of these, 51 were previously identified by mass spectrometry analysis of AD169 virions, including all 5

recognised capsid components, as well as the capsid maturational protease and capsid assembly protein pre-cursor (pUL80/PR-AP) fusion protein; 25 known tegument proteins; 8 proteins yet to be assigned to any virion compartment; and 12 proteins either known, or predicted to be incorporated into the envelope due to the presence of membrane anchors. A further 6 proteins not identified by mass spectrometry analysis of AD169 virions, but recognised as HCMV virion components following assessment of virions from other strains by alternative approaches, were also detected with high confidence (UL36x1, gpRL13, gpUL128, gpUL130, gpUL131A and gpUS28).

A further 13 proteins identified with high confidence have not been previously described as virion components. Of these, 5 are implicated in viral genome transcription/replication and virion assembly (pUL52, pUL34, pUL70, pUL98, and ppUL114). The viral CXC chemokine homologue encoded by UL146/vCXC-1 was also identified with high confidence. 7 integral membrane bound proteins (gpRL12, gpUL116, gpUL141, gpUL148, pUS12, pUS14 and pUS20), likely incorporated into the virion envelope were also identified. In support of these novel proteins being virion components, homologous proteins to pUL52 (Rh83), gpUL116 (Rh148), gpUL141 (Rh164), pUS12 (Rh190) have been identified in RhCMV virions, whilst homologous proteins of pUL70 (M70) and pUL116 (M116) have also been identified in MCMV virions.

6.3.2 Proteins identified with low confidence

Criteria applied for the identification of proteins with low confidence included the detection of a minimum of 2 peptides unique to that protein in at least one virion preparation. A further 22 proteins were identified across the virion preparations analysed that met this criteria (Table 6.2.B). Of these, two were previously identified during mass spectrometry analysis of AD169 virions (pUL54, and pUL56). Homologues for each of these proteins were identified in MCMV virions (M54 and M56, respectively), suggesting that they may represent *bona fide* virion components. Two further proteins identified with low confidence have previously been described as virion components, though were not detected in AD169 virions (pUL28/UL29, and gpUL78). Homologous proteins to pUL28/pUL29 were also previously identified in MCMV virions (M28).


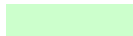
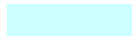

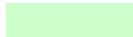
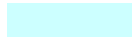

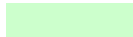
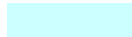

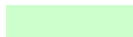


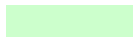


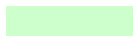


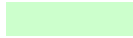

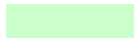
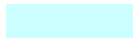
No reports exist of the remaining 17 proteins as being HCMV virion components. Proteins homologous to pUL102 and pUL105 were identified in RhCMV virions (Rh138), and also in MCMV virions (M102 and M105). The major immediate early protein pUL123 and the viral chemokine homologue pUL147/vCXC-2 were both also detected. Novel integral membrane proteins identified include gpUL16, gpUL74A, gpUL139 and pUL140, gpUS9 (US6 family members), and pUS13 and pUS18 (US12 family members). A further 6 proteins were detected with low confidence, none of which are characterised biochemically or functionally:


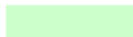

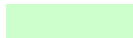
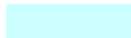

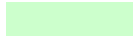




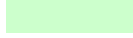
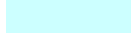

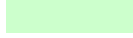
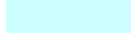

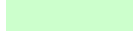


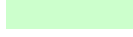
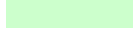

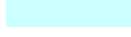

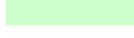

pRL1, pUL31 (a DURP family member), pUL145, pUL150A, pUS1 (US1 family member) and pUS26 (US22 family member). Homologues of pUL31 (M31) have been identified in MCMV, whilst homologues of pUS1 (Rh181) and pUS26 (Rh211) have been identified in RhCMV.


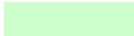


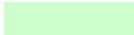


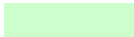
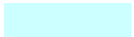

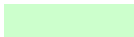
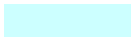

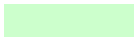

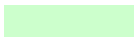


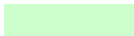
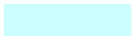

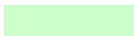


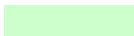



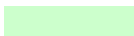
6.3.3 Proteins previously identified in HCMV virions, or homologous proteins identified in virions of other CMVs, but not on MS analysis of Merlin virions


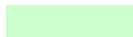

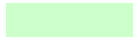
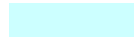


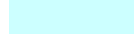

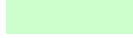
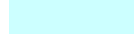

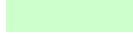


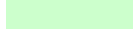
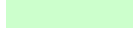
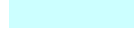

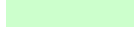
Several proteins that have previously been described as HCMV virion components were not detected by MS analysis of Merlin virions. This includes 6 proteins previously identified by MS analysis of AD169 virions (although these were only identified by a second, more sensitive methodology in this previous study), including the tegument proteins pUL38, pUL72; the envelope proteins gpRL14 and gpUL5; the late gene transactivator protein pUL79; and also the virion associated glycoprotein gpUL22A (Varnum, Streblow et al. 2004). Of these proteins, only homologues of pUL72 have also been identified in other CMV virions (Kattenhorn, Mills et al. 2004). Unidentified proteins previously detected in HCMV virions, though by alternative approaches to mass spectrometry analysis include the tegument protein pUL23 (Dunn, Chalupny et al. 2003), virion envelope components gpUL1 (Shikhagaie, Mercé-Maldonado et al. 2012) and gpUL4 (Bhushan, Meyer et al. 2010), as well as the virion associated protein gpUL76 (Wang, Duh et al. 2004). Further HCMV protein homologues were also identified in virions other CMVs, though have yet to be identified in HCMV by any analytical approach: in RhCMV, homologues of UL13 (Rh31), UL111a (Rh143) and US30 (Rh223) have been identified (Malouli, Nakayasu et al. 2012), whilst in MCMV, homologues of proteins UL51 (M51), UL87 (M87), UL95 (M95), UL121 and (M121) have been identified (Kattenhorn, Mills et al. 2004). In consideration of these undetected components, single peptides corresponding to pUL13 (1/8 preparations), gpUL22A (3/8 preparations), pUL23 (4/8 preparations) and gpUL76 (2/8 preparations), and pUL95 (3/8 preparations) were detected across the Merlin virion preparations.


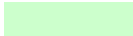

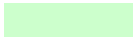

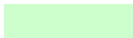

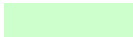

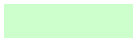
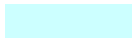

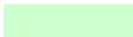

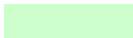
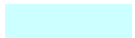
Table 6.2. Virus encoded proteins identified by MS-analysis of Merlin virions.


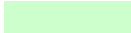
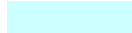

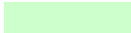


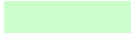
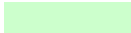

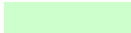




ORF ^a	ID ^b	Protein/Function ^c	AD169 ^d	RhCMV ^e	MCVM ^f
<i>A- Proteins identified with high confidence</i>					
<i>Capsid:</i>					
UL46	8/8	pUL46/mCBP (minor capsid protein binding protein); interacts with mCP to form capsid triplex			
UL48A	8/8	pUL48A/SCP (smallest capsid protein); decorates capsid hexons			
UL80	7/8	pUL80/PR-AP (assembly protein) precursor; scaffold for capsid assembly			
UL85	8/8	pUL85/mCP (minor capsid protein); interacts with mCBP to form capsid triplex			
UL86	8/8	pUL86/MCP (major capsid protein); forms capsid walls and faces			
UL104	8/8	pUL104/PORT; forms specialized portal capsomere for genome insertion/release			
<i>Known Tegument:</i>					
UL24	8/8	pUL24; US22 family member; enhances replication in endothelial cells			
UL25	8/8	pUL25; UL25 family member			

UL26	7/8	pUL26; US22 family member, MIEP transcription transactivator, tegument material phosphorylation, stabilises virion			
UL32	8/8	pp150 / NSP (nucleocapsid-proximal stabilization protein); virion maturation			
UL35	8/8	pUL35; UL25 family member, virion morphogenesis			
UL36x1	6/8	pUL36x1; US22 family member, vICA (viral inhibitor of caspase-8-induced apoptosis)			
UL43	8/8	pUL43; US22 family member			
UL45	8/8	pUL45/RR1 (Ribonucleotide reductase homologue sub-unit)			
UL47	8/8	pUL47/LTPbp (Largest tegument protein binding protein); interacts with ppUL48 intra-cellular transport of nucleocapsid			
UL48	8/8	pUL48/LTP (Largest tegument protein); interacts with ppUL47 intra-cellular transport of nucleocapsid			
UL50	8/8	pUL50/NEC1 (nuclear egress complex membrane anchoring component 1); interacts with pUL53 (NEC2) to orchestrate nuclear egress of nucleocapsids			
UL52	8/8	pUL52; DNA encapsidation, nucleocapsid formation			
UL69	8/8	ppUL69; MRP (multiple regulatory protein)			
UL71	8/8	pUL71; Secondary envelopment, virion egress			

ORF ^a	ID ^b	Protein/Function ^c	AD169 ^d	RhCMV ^e	MCMV ^f
UL82	7/8	ppUL82/pp71/UMP (upper matrix protein); secondary envelopment, relieves DAXX-mediated repression			
UL83	8/8	ppUL83/pp65/LMP (lower matrix protein); phosphorylase, innate and adaptive immune response evasion			
UL88	8/8	pUL88; Putative cytoplasmic egress function			
UL94	8/8	pUL94; interacts with pp28 to facilitate secondary envelopment			
UL96	6/8	pUL96; interacts with pp150 to stabilise nucleocapsid during cytoplasmic egress			
UL97	8/8	ppUL97/VPK (viral protein kinase); phosphorylates viral and cellular proteins			
UL99	7/8	ppUL99/pp28; myristylated protein, secondary envelopment			
UL103	7/8	pUL103; VEP (Virion and DB egress protein); orchestrates release from producer cells			
IRS1	8/8	pIRS1; US22 family member, PKR (Protein kinase R) inhibitor, Transcriptional activator			
US22	8/8	pUS22; US22 family member			
US23	6/8	pUS23; US22 family member			
US24	5/8	pUS24; US22 family member; enhances early infection			

TRS1	8/8	pTRS1; US22 family member, PKR (Protein kinase R) inhibitor, Transcriptional activator, capsid assembly			
<i>Transcription/replication machinery, yet to be assigned to virion compartment:</i>					
UL34	7/8	pUL34; Represses US3 transcription			
UL44	8/8	ppUL44/PPS (Viral DNA polymerase processivity sub-unit); DNA synthesis			
UL57	8/8	ppUL57/SSB (single-stranded DNA binding protein); DNA synthesis			
UL70	5/8	pUL70/HP2 (DNA helicase-primase sub-unit 2); DNA synthesis			
UL77	8/8	pUL77/CVC1 (putative capsid vertex-specific component 1); DNA encapsidation			
UL84	8/8	ppUL84/Viral DNA replication accessory; DURP family member, DNA synthesis			
UL89	6/8	pUL89/TER1 (terminase sub-unit 1); DNA encapsidation			
UL93	7//8	pUL93/CVC2 (putative capsid vertex-specific component 2); DNA encapsidation			
UL98	7/8	pUL98/NUC (deoxyribonuclease)			
UL112-UL113	8/8	pUL112/UL113; Orchestrates DNA synthesis, transcriptional activator			

ORF ^a	ID ^b	Protein/Function ^c	AD169 ^d	RhCMV ^e	MCMV ^f
UL114	4/8	ppUL114; DNA synthesis UNG (Uracil-DNA glycosylase), control DNA synthesis			
UL122	8/8	pUL122; IE2; transactivator or host-cell transcription machinery			
UL146	7/8	gpUL146; vCXC chemokine homologue family, putative chemokine			
<i>Envelope:</i>					
RL10	7/8	gpRL10			
RL12	7/8	gpRL12/gp95; RL11 family member, IgG Fc binding			
RL13	½	gpRL13; RL11 family member, putative IgG Fc binding, impedes replication in fibroblasts and epithelial cells			
UL33x1	8/8	pUL33x1; GPCR homologue family member, constitutive signalling GPCR			
UL41A	6/8				
UL55	8/8	gB; gCI subunit, virion binding and entry			
UL73	8/8	gN; gCII sub-unit, virion binding, progeny virion secondary envelopment			
UL74	8/8	gO; gCIII sub-unit, fibroblast cell entry, release of progeny virions from producer cells			

UL75	8/8	gH; gCIII/pentameric complex sub-unit, entry into cells			
UL100	8/8	gM; gCII sub-unit, virion binding, progeny virion secondary envelopment			
UL115	7/8	gL; gCIII/pentameric complex sub-unit, entry into cells			
UL116	7/8	gpUL116; putative membrane glycoprotein			
UL118-UL119	8/8	gpUL118-UL119/gp68; IgG Fc binding			
UL128	4/6	gpUL128,; pentameric complex, entry into non-fibroblasts			
UL130	4/6	gpUL130; pentameric complex, entry into non-fibroblasts			
UL131A	4/6	gpUL131A; pentameric complex, entry into non-fibroblasts			
UL132	8/8	gpUL132; enhances replication in fibroblasts			
UL141	7/7	gpUL141; UL14 family member, NK evasion, down regulates CD155 and CD112			
UL148	8/8	gpUL148, putative membrane glycoprotein			
US12	7/8	pUS12; US12 family member, 7TM protein			

ORF ^a	ID ^b	Protein/Function ^c	AD169 ^d	RhCMV ^e	MCMV ^f
US14	8/8	pUS14; US12 family member, 7TM protein			
US20	6/8	pUS20; US12 family member, 7TM protein			
US27	8/8	gpUS27; GPCR homologue family member, 7TM protein, enhances virion release			
US28	8/8	gpUS28; GPCR homologue family member, 7TM protein, CC and CXC3 chemokine receptor			

B- Proteins identified with low confidence

<i>Tegument</i>					
UL28/UL29	1/8	pUL28/UL29; US22 family member, Stimulates IE gene expression			
<i>Transcription/replication machinery, immune evasions, yet assigned to virion compartment:</i>					
UL54	1/8	pUL54; POL; catalytic DNA polymerase sub-unit			
UL56	3/8	pUL56; TER2 (terminase sub-unit 2); DNA encapsidation			
UL102	1/8	pUL102/HP3 (helicase primase sub-unit 3); DNA synthesis			

UL105	2/8	pUL105; HP1 (helicase-primase sub-unit 1), DNA synthesis	
UL123	2/8	ppUL123/IE1 (major IE protein); promotes transcription cascade, enhances IE2 activation	
UL147	3/8	pUL147; vCXC-2 chemokine homologue family, putative chemokine	
<hr/>			
<i>Envelope:</i>			
<hr/>			
RL11	3/8	gpRL11/gp34; RL11 family member, IgG Fc binding	
UL16	3/8	gpUL16; membrane glycoprotein, NK cell evasion, blocks NKG2D ligands MICB, ULBP1 and ULBP2	
UL74A	1/8	gpUL74A; putative membrane glycoprotein	
UL78	2/8	gpUL78; GPCR homologue family member; putative chemokine receptor	
UL139	1/8	gpUL139; putative membrane glycoprotein	
UL140	3/8	pUL140; putative membrane protein	
US9	2/8	gpUS9; US6 family member, membrane glycoprotein	
US13	3/8	pUS13; US12 family member, 7TM protein	

ORF ^a	ID ^b	Protein/Function ^c	AD169 ^d	RhCMV ^e	MCMV ^f
US18	1/8	pUS18; US12 family member, 7TM protein			
<i>Uncharacterised:</i>					
RL1	1/8	pRL1; RL1 family member			
UL31	1/8	pUL31; DURP family member			
UL145	3/8				
UL150A	3/8	Putative secreted protein			
US1	3/8	US1 family member			
US26	3/8	pUS26; US22 family member			

^a ORFs encoding proteins identified in Merlin virions. Highlighted in purple are proteins identified in HCMV virions by approaches other than mass spectrometry, in yellow are novel HCMV virions components;

^b identities depict the number of virion preparations in which proteins were detected;

^c - comments as to the biochemistry and function of proteins identified, as described in Field's Virology {Mocarski, 2007 #511}

^d - proteins previously identified during mass spectrometry analysis of AD169 virions (pink) {Varnum, 2004 #65};

^e - homologous proteins identified during mass spectrometry analysis of RhCMV virions {Malouli, 2012 #991};

^f - homologous proteins identified during mass spectrometry analysis of MCMV virions {Kattenhorn, 2004 #992}.

(A) Proteins identified with high confidence.

(B) Proteins identified with low confidence.

6.4 Relative abundance of proteins in virions from Merlin variants

In the quantitative SILAC-MS analysis approach, the relative abundance of a given protein incorporated into the different virions is inferred from the ratio of differently labelled versions of that protein identified in the mixed virion preparation. Since the capsid is structurally constrained and the virion compartment with greatest homogeneity between strains, accurate comparison of protein levels in an equal number of different virions will include capsid protein ratios approximating the value 1. Based on this, SILAC-MS raw data were corrected by normalising the ratios of all identified proteins within a sample to the ratios of the major capsid protein (pUL86/MCP) in that same sample.

6.4.1 Inter-assay variation

To compare the relative abundance of components incorporated into the different virions in each mixed virion preparation, it was important to assess the precision of the MS analysis. The composition of RV-Me-tet-UL128L^{wt} and RV-Me-UL128L^{mut} virions were compared during analysis of both the MVP-1 and MVP-2 preparations, and the reproducibility of the SILAC-MS data was assessed by comparing the relative abundances of proteins as reported in these duplicate analyses (See figure 6.3.A). 68 HCMV virion components were identified in both MVP-1 and MVP-2, and the ratios of 61 of these proteins were reported with <1.3-fold variation in the duplicate analyses to demonstrate the high precision of the SILAC-MS technique. The reported ratios of 7 proteins were more variable (See Figure 6.3.B). The ratios of pUL34, pUL50 and pUL89 showed 1.5-2.2-fold variation on repeat analyses. The reported ratios of pUS24 were widely contrasting in the two assays: on analysis of the MVP-1 preparation, RV-Me-tet-UL128L^{wt} virions contained ~8.4-fold more pUS24 than RV-Me-UL128L^{mut} virions, whilst on analysis of the MVP-2 preparation, RV-Me-tet-UL128L^{wt} virions contained ~2.5-fold less pUS24 than RV-Me-UL128L^{mut} virions. The remaining virion proteins reported with great variation were the UL128L encoded glycoproteins (see section .6.4.3.2)

6.4.2 Impact of variations in gpRL13 content

In the MVP-1 preparation, RV-Me-tet-RL13^{wt}/tet-UL128L^{wt} virions contained significantly greater levels of the vast majority of proteins compared to virions from both RV-Me-tet-UL128L^{wt} (68/73 proteins) and RV-Me-UL128L^{mut} (56/73 proteins), suggesting a major impact of gpRL13 on virion composition. However, pUL86/MCP ratios in the raw data indicated that approximately ~10-fold fewer RV-Me-tet-RL13^{wt}/UL128L^{wt} virions were loaded compared to those of the other viruses, potentially affecting the accuracy of quantitation. Thus, a repeat of this experiment will be required before firm conclusions can be drawn.

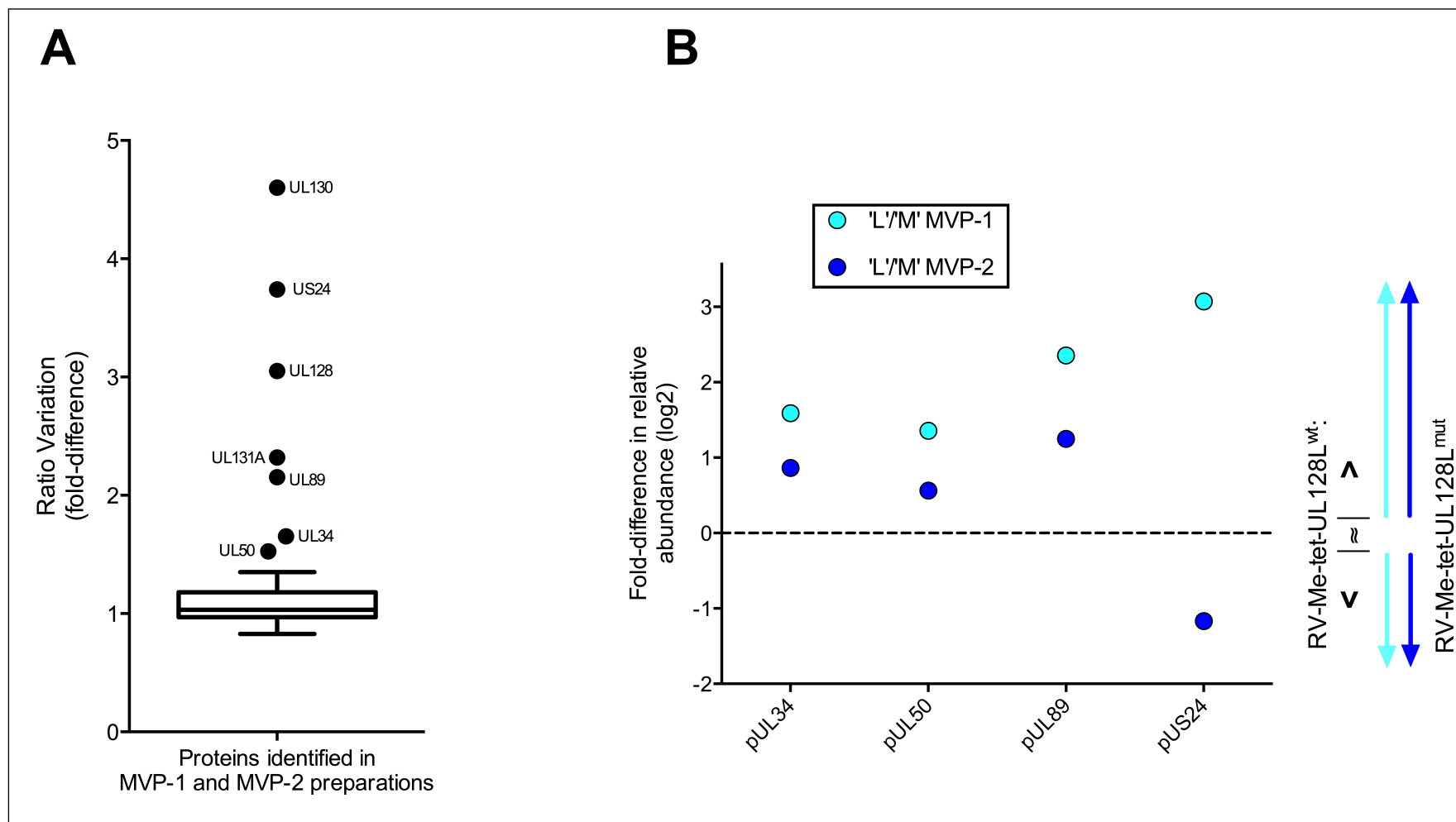


Figure 6.3. SILAC-MS inter-assay variation and outlier proteins. (A) Tukey Box-plot displaying variation in 'light'/'medium' ratios of all 68 proteins identified during analysis of MVP-1 and MVP-2 preparations. pUL34, pUL50, pUL89, gpUL128, gpUL130, gpUL131A and pUS24 showed the greatest variation between runs. Of the remaining 61 proteins, median variance = 1.013. (B) 'Light' / 'Medium' ratios of pUL34, pUL50, pUL89, and pUS24 as detected in the MVP-1 and MVP-2 duplicate analyses.

6.4.3 Impact of variations in pentameric complex content

6.4.3.1 Heterogeneity amongst capsid proteins as a measure of the accuracy of the assay

Following normalisation to pUL86/MCP ratios, the ratios of all other known capsid proteins closely approximated the value 1 (Figure 6.4.) to indicate a high level of accuracy in the SILAC-MS assay. One protein (pUL80/PR-AP) tentatively assigned to the capsid compartment was identified with greater abundance in RV-Me-tet-UL128L^{wt} and RV-Me-tet-UL128L^{G>T} virions compared to those of RV-Me-UL128L^{mut}. However the current model of herpesvirion assembly includes homologues of this protein being eliminated from nascent nucleocapsids during maturation (Yu, Trang et al. 2005). It is therefore possible that residual pUL86 was detected at differing levels that did not reflect the heterogeneity of the virion capsid compartment. The most variable *bona fide* capsid protein between the different virions was pUL48A/SCP, varying at most by 1.1665-fold; given that the capsid proteins should be present in different virions at equimolar amounts, ± 1.1665 -fold variation was assumed to represent the accuracy range of the SILAC-MS technique. A threshold of ± 1.5 -fold variation was applied to distinguish proteins that varied between virions with confidence; this figure was calculated as 2 standard deviations (2 S.D.) from the ‘accuracy range’ of the assay [**accuracy range** (± 1.1665 -fold) + **2 S.D.** (0.122) = **confidence threshold** (± 1.4108 -fold; rounded to 1.5-fold)].

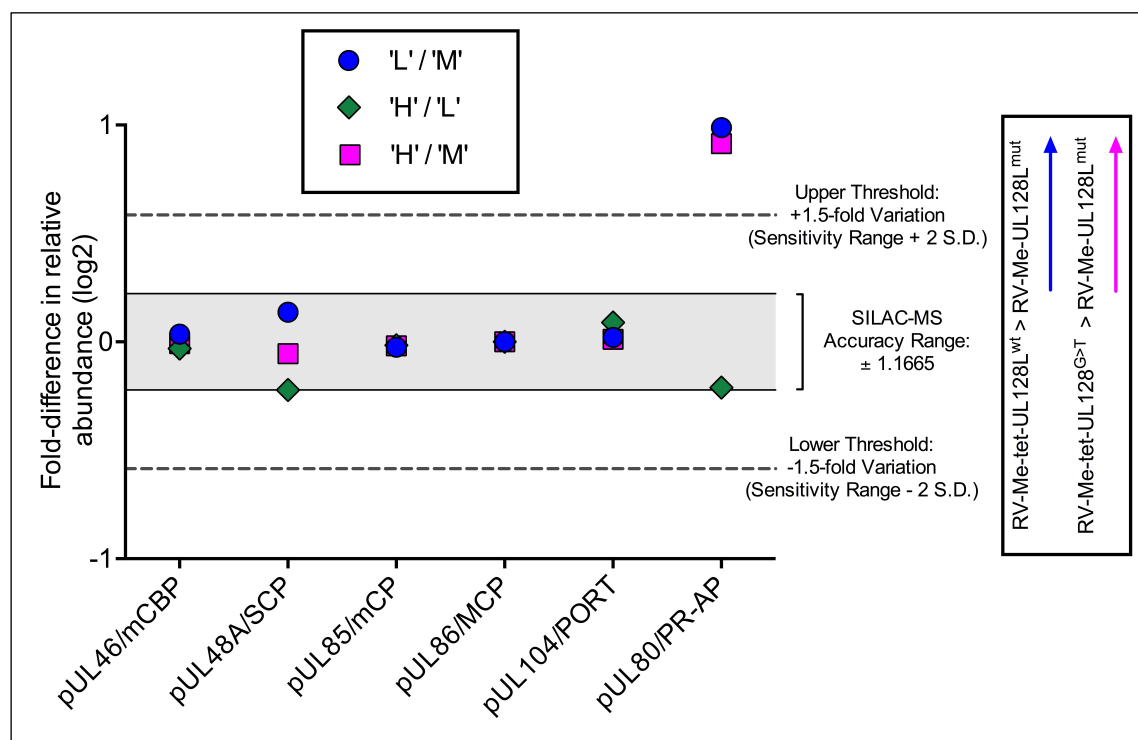


Figure 6.4. Relative abundance of capsid proteins in RV-Me-tet-UL128L^{wt} (L) RV-Me-UL128L^{mut} (M), RV-Me-tet-UL128L^{G>T} (H) virions. Highlighted (grey) is the degree of acceptable variation within the accuracy range of the SILAC-MS technique. Broken lines (dark grey) shows the variance threshold applied (± 1.5 -fold) to highlight proteins that differed between virions with confidence.

6.4.3.2 *Relative abundance of gH/L complexes (pentameric complex and (gH/gL/gO))*

SILAC-MS analysis of both MVP-1 and MVP-2 preparations detected the UL128L-encoded glycoproteins with greater abundance in virions from RV-Me-tet-UL128L^{wt} compared to RV-Me-UL128L^{mut} (Figure 6.5.A). As previously mentioned, the relative abundance of these glycoproteins in RV-Me-tet-UL128L^{wt} and RV-Me-UL128L^{mut} virions was reported with different magnitudes between each independent analysis. Furthermore, relative abundances of each of gpUL128, gpUL130, and gpUL131A also differed within each independent analysis; this was surprising given that the UL128L-encoded glycoproteins are believed to be incorporated into the pentameric complex (and therefore virions) at equimolar quantities. These discrepancies were likely due to the complete absence of these components in RV-Me-UL128L^{mut} virions, with the relative abundance compared to RV-Me-tet-UL128L^{wt} virions reported against background noise of the SILAC-MS analysis. Conversely, glycoprotein gO was detected with reduced abundance in virions from RV-Me-tet-UL128L^{wt} compared to those from RV-Me-UL128L^{mut}, yet this variation was reported with similar magnitude on duplicate analyses. Importantly, glycoproteins gH and gL (that are common to the pentameric complex and gH/gL/gO) were detected with comparable abundance in virions from RV-Me-tet-UL128L^{wt} and RV-Me-UL128L^{mut}. Thus, the inverse relative abundances of the UL128L-encoded glycoproteins and glycoprotein gO likely reflect the contrasting pentameric and trimeric complex content of RV-Me-tet-UL128L^{wt} and RV-Me-UL128L^{mut} virions.

Virions from the Merlin virus engineered to express the UL128 intron 1 G>T variation (RV-Me-tet-UL128^{G>T}) contained intermediate quantities of the UL128L-encoded glycoprotein compared to virions from RV-Me-tet-UL128L^{wt} and RV-Me-UL128L^{mut} (Figure 6.5.B). The greater abundance of these glycoproteins in virions from RV-Me-tet-UL128^{G>T} compared to those from RV-Me-UL128L^{mut} was clear, although the reduced abundance of gpUL128 and gpUL131A in comparison to RV-Me-tet-UL128L^{wt} virions did not exceed the confidence threshold. Each of gpUL128, gpUL130, and gpUL131A were however detected with similar relative abundance between RV-Me-tet-UL128L^{wt} and RV-Me-tet-UL128^{G>T} virions, consistent with being incorporated into pentameric complexes at equimolar quantities. Glycoprotein gO was also detected in RV-Me-tet-UL128^{G>T} virions with intermediate abundance compared to virions from RV-Me-tet-UL128L^{wt} and RV-Me-tet-UL128L^{mut}, and these variations did exceed the threshold for confidence. Furthermore, glycoproteins gH and gL were detected in RV-Me-tet-UL128^{G>T} virions at similar abundance as in virions from both RV-Me-tet-UL128L^{wt} and RV-Me-tet-UL128L^{mut}. These latter results provide further evidence for an inverse linear correlation between pentameric and trimeric complex content in virions, and support the hypothesis that reduce pentameric complex content occurs with a concomitant increase trimer complex content.

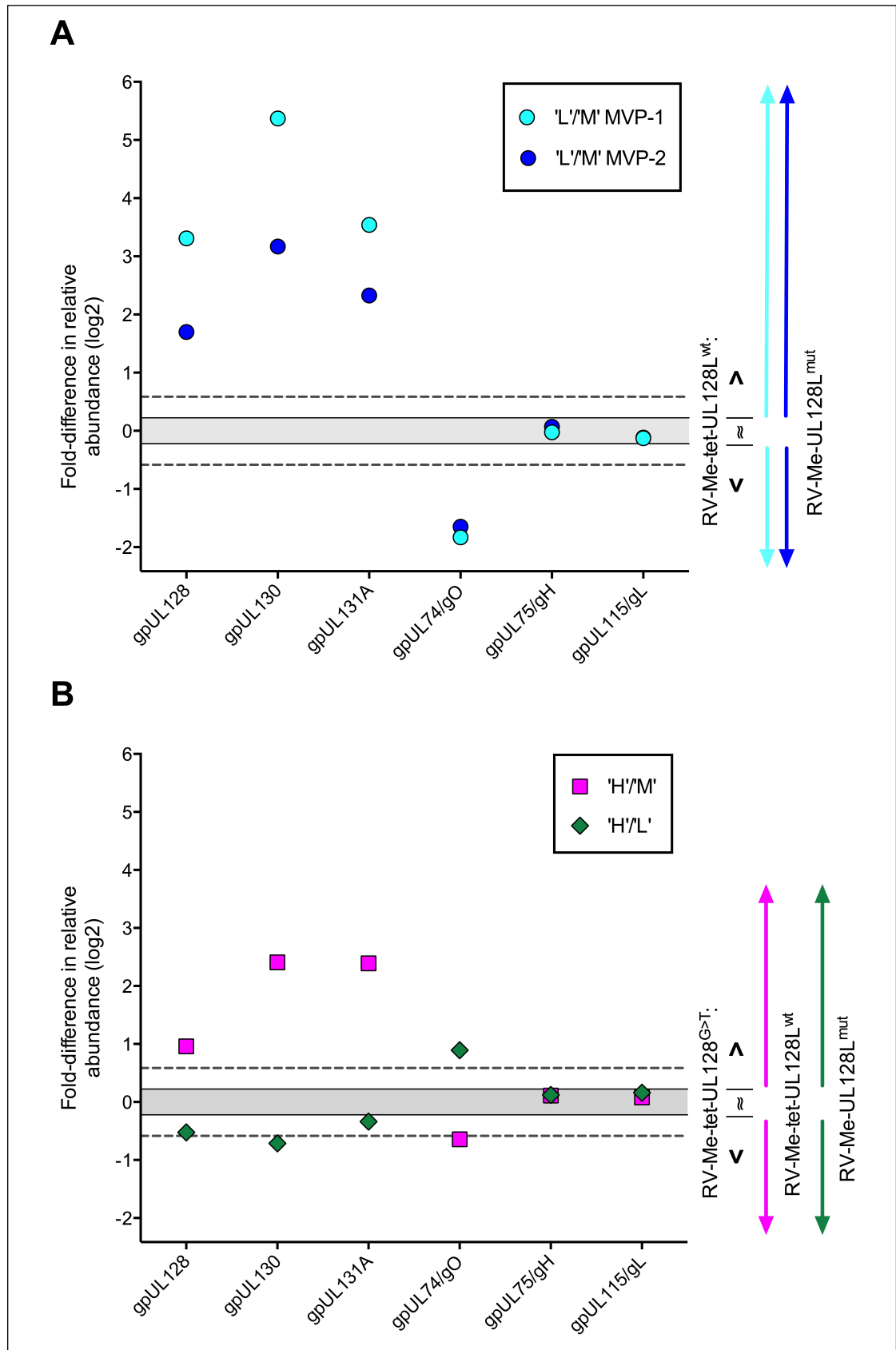


Figure 6.5. Pentameric complex and gCIII (gH/gL/gO) content in virions from Merlin viruses differing in UL128L coding potential. Highlighted (grey) is the acceptable degree of intra-assay variation ($\pm 1.1.665$ -fold). Broken lines (dark grey) show the threshold (± 1.5 -fold) to highlight proteins that differed between virions with confidence. **(A)** Relative abundance of proteins in RV-Me-tet-UL128L^{wt} (L) compared to RV-Me-UL128L^{mut} (M) virions. **(B)** Relative abundance of proteins in RV-Me-tet-UL128^{G>T} (H) compared to RV-Me-tet-UL128L^{wt} (L) and RV-Me-UL128L^{mut} (M).

6.4.3.3 Variations in tegument compartment materials

All tegument proteins with variance exceeding the threshold for confidence were detected at greater abundance in RV-Me-tet-UL128L^{wt} virions compared to those from RV-Me-UL128L^{mut} (Figure 6.6.A). One protein (pUL36) was dramatically more abundant in RV-Me-tet-UL128L^{wt} virions, identified at ~3.7-fold greater levels than in RV-Me-UL128L^{mut} virions. A further 9 tegument proteins (pUL52, ppUL69, ppUL82, pUL96, ppUL97, pIRS1, pUS22, pUS2, pTRS1) were moderately more abundant. Only 6 tegument proteins (ppUL32, pUL45, pUL47, pUL83, ppUL99, pUL103) were detected with comparable abundance in RV-Me-tet-UL128L^{wt} and RV-Me-UL128L^{mut} virions, whilst 9 proteins (pUL24, pUL25, pUL26, pUL35, pUL43, pUL48, pUL71, pUL88, pUL94) were slightly more abundant in RV-Me-tet-UL128L^{wt} virions, but did not exceed the threshold of variance applied for confidence.

Virions from RV-Me-tet-UL128L^{G>T} that contained intermediate pentamer/trimer ratios displayed a tegument protein profile with similarities to virions from both RV-Me-tet-UL128L^{wt} and RV-Me-UL128L^{mut} (Figure 6.6.B). The most variable tegument protein, pUL36x1, was dramatically more abundant (> 3-fold) in RV-Me-tet-UL128L^{G>T} virions compared to those from RV-Me-UL128L^{mut}, but also slightly more abundant than in RV-Me-tet-UL128L^{wt} virions. Hence, the incorporation of this virion component did not correlate with increasing pentameric complex content, but rather the presence of the pentamer at any level. 10 tegument proteins varied moderately (< 3-fold) between RV-Me-tet-UL128L^{G>T} virions and those of the comparison viruses. Three tegument proteins (pUL25, pUL26 and ppUL97) more closely matched that in RV-Me-UL128L^{mut} virions and were less abundant than in RV-Me-tet-UL128L^{wt} virions, whilst 6 proteins (pUL52, ppUL69, ppUL82, pIRS1, pUS22 and pTRS1) closely matched RV-Me-tet-UL128L^{wt} virions with greater abundance than in RV-Me-UL128L^{mut} virions. Incorporation of pUL96 into virions potentially correlated with increasing pentameric complex levels, with RV-Me-tet-UL128L^{G>T} containing more than virions from RV-Me-UL128L^{mut}, but less than in RV-Me-tet-UL128L^{mut} virions. However the ratios of this protein between RV-Me-tet-UL128L^{G>T} and RV-Me-tet-UL128L^{wt} virions did not exceed the applied threshold for confidence.

A further 9 tegument proteins varied between virions from RV-Me-tet-UL128L^{G>T} and the comparison viruses, and whilst none exceeded the threshold for confidence, a trend suggestive of a correlation between their incorporation and the pentamer/trimer ratios of virions was apparent. These included 4 proteins (pUL24, pUL43, ppUL83, and pUL94) that closely matched RV-Me-UL128L^{mut} virions (low pentamer, high trimer), and 5 proteins (e.g. pUL47, pUL48, pUL71, pUL88 and pUS23) more closely matched those in RV-Me-tet-UL128L^{wt} virions (high pentamer, low trimer). Only 5 tegument proteins (ppUL32, pUL35, pUL45, ppUL99, pUL103) were detected in RV-Me-tet-UL128L^{G>T} virions at comparable levels to those from both comparison viruses (RV-Me-UL128L^{mut} and RV-Me-tet-UL128L^{wt}).

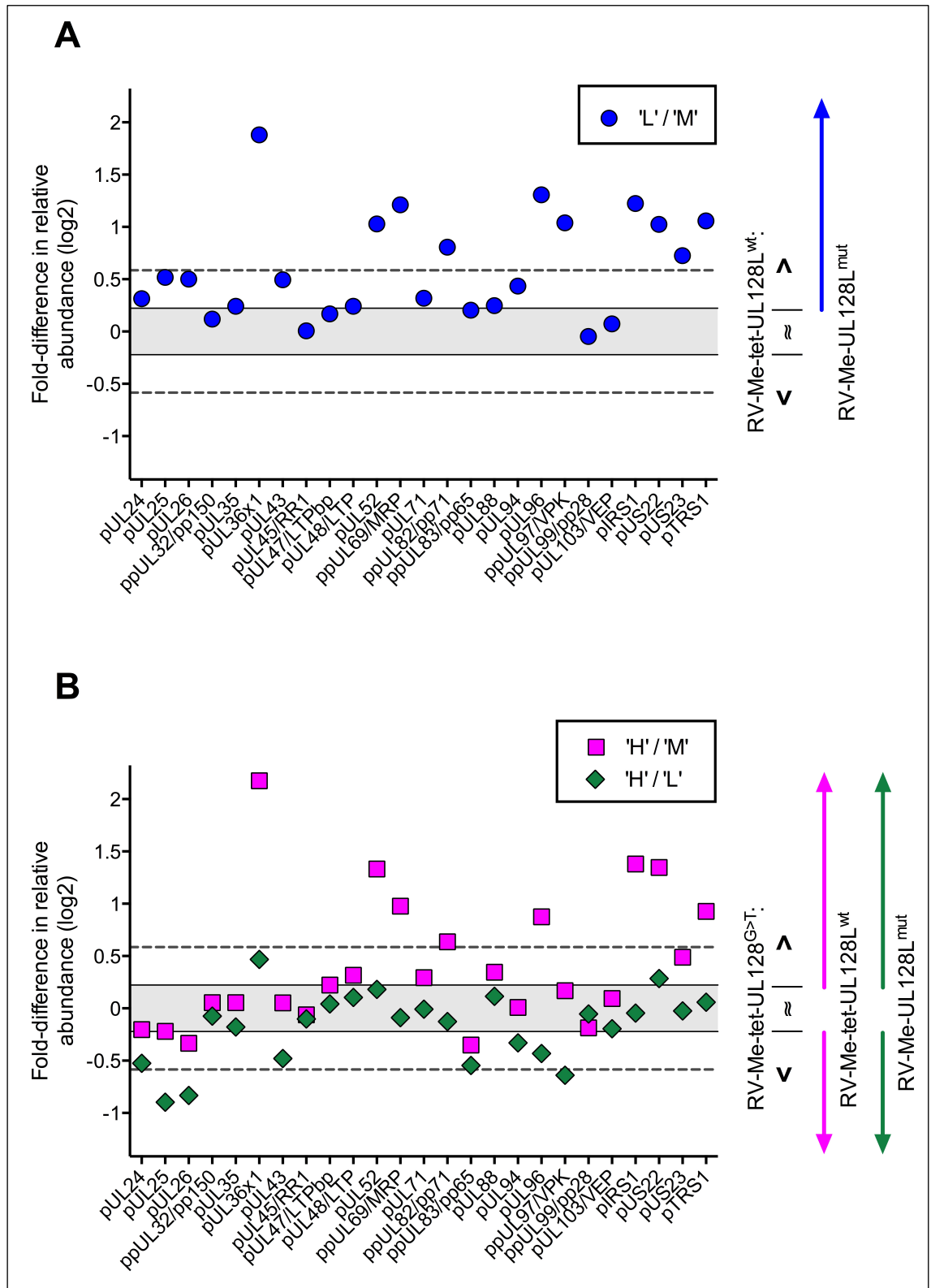


Figure 6.6. Tegument protein content of virions with contrasting pentameric/trimeric complex ratios. Highlighted (grey) is the degree of acceptable intra-assay variation ($\pm 1.1.665$ -fold). Broken lines (dark grey) show the threshold (± 1.5 -fold) to highlight proteins that differed between virions with confidence. (A) Relative abundance of proteins in RV-Me-tet-UL128L^{wt} (L) compared to RV-Me-UL128L^{mut} (M) virions. (B) Relative abundance of proteins in RV-Me-tet-UL128^{G>T} (H) compared to RV-Me-tet-UL128L^{wt} (L) and RV-Me-UL128L^{mut} (M).

6.4.3.4 *Variation in proteins yet to be assigned to any particular virion compartment*

Similarly to the tegument proteins, the majority of the variable viral proteins that are yet assigned to any virion compartment were more abundant in virions containing the pentameric complex (RV-Me-tet-UL128L^{wt}) than those lacking the pentameric complex (RV-Me-tet-UL128L^{mut}) (Figure 6.7.A). Amongst the most variable proteins in this category were the novel virion components pUL70 and pUL147, detected in RV-Me-tet-UL128L^{wt} virions with dramatically greater abundance (~6.4-fold and ~6.7-fold greater, respectively) than in RV-Me-UL128L^{mut} virions. A further 6 proteins (ppUL44, ppUL57, ppUL84, pUL98, pUL112-113, pUL122) were moderately more abundant in RV-Me-tet-UL128L^{wt} virions. In contrast, the novel virion component gpUL146 was identified with dramatically reduced abundance (~4.25-fold) in virions from RV-Me-tet-UL128L^{wt} compared to those from RV-Me-tet-UL128L^{mut}. The remaining proteins in this category (pUL77 and pUL93) were detected with comparable abundance in RV-Me-tet-UL128L^{wt} and RV-Me-tet-UL128L^{mut} virions; each of these proteins are sub-units of a capsid vertex complex, and similar to the known capsid proteins, will likely be heterogeneous amongst different strains and strain variants. Proteins pUL102 and pUL114 were not detected on SILAC-MS analysis of the MVP-2 preparation. However in the MVP-1 preparation, pUL102 was present at greater levels in RV-Me-tet-UL128L^{wt} virions but not by a degree that exceeded the threshold of confidence, whereas pUL114 was confidently detected at greater abundance than in RV-Me-tet-UL128L^{mut} virions.

In RV-Me-tet-UL128L^{G>T} virions, the abundance of these viral proteins (those yet assigned to any virion compartment) more closely resembled that in RV-Me-tet-UL128L^{wt} virions (Figure 6.7.B). This included the proteins that varied the most between virions containing or lacking the pentameric complex, with RV-Me-tet-UL128L^{G>T} virions containing comparable levels of pUL70 and gpUL147 to virions from RV-Me-tet-UL128L^{wt}, and therefore dramatically more than in virions from RV-Me-UL128L^{mut}. Furthermore, RV-Me-tet-UL128L^{G>T} virions also contained comparable levels of gpUL146 as in those from RV-Me-tet-UL128L^{wt}, and therefore dramatically less than in RV-Me-UL128L^{mut}. Interestingly, RV-Me-tet-UL128L^{G>T} virions contained ppUL57 at greater levels than in RV-Me-UL128L^{mut} and RV-Me-UL128L^{wt} virions. This was also apparent for proteins pUL98 and pUL122, but the variance in these components between RV-Me-tet-UL128L^{G>T} and RV-Me-tet-UL128L^{wt} virions did not exceed the threshold for confidence. Thus, incorporation of ppUL57, pUL98, and pUL122 into virions did not correlate with increasing pentameric complex content, but simply the presence of the pentameric complex at any level. The 5 remaining variable proteins (pUL44, pUL84, pUL98, pUL112-113, pUL122) were present at similar levels to RV-Me-tet-UL128L^{wt} virions, yet were greater than in RV-Me-UL128L^{mut} virions. The capsid-vertex complex proteins pUL77 and pUL93 were present at levels similar to that in both RV-Me-UL128L^{mut} and RV-Me-tet-UL128L^{mut} virions.

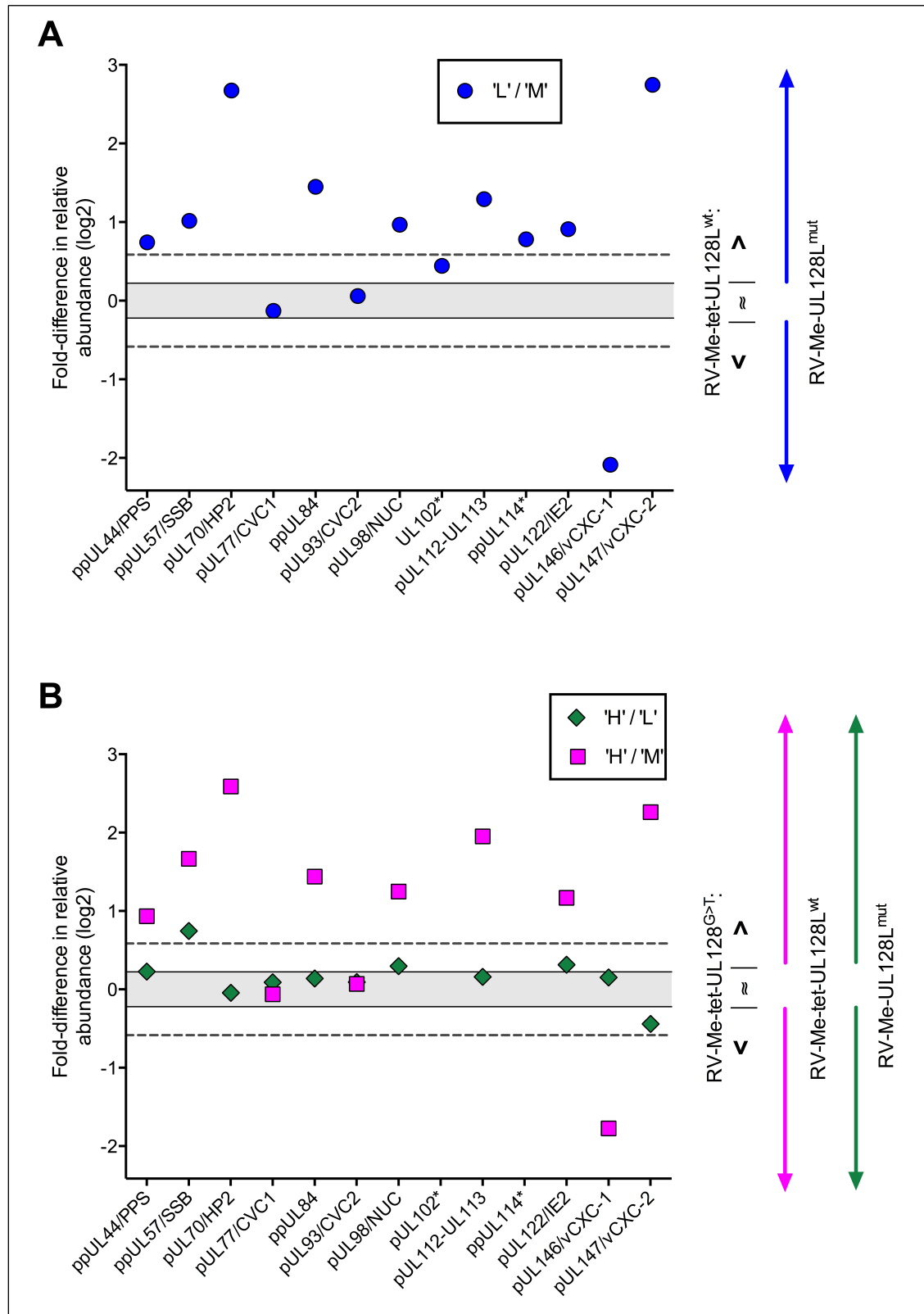


Figure 6.7. Relative abundance of proteins ‘yet to be assigned’ to any compartment in virions with contrasting pentameric/trimeric complex ratios. Highlighted (grey) is the degree of acceptable intra-assay variation ($\pm 1.1.665$ -fold). Broken lines (dark grey) show the threshold (± 1.5 -fold) to highlight proteins that differed between virions with confidence. **(A)** Relative abundance of proteins in RV-Me-tet-UL128L^{wt} (L) compared to RV-Me-UL128L^{mut} (M) virions. **(B)** Relative abundance of proteins in RV-Me-tet-UL128^{G>T} (H) compared to RV-Me-tet-UL128L^{wt} (L) and RV-Me-UL128L^{mut} (M). * denotes protein detected in MVP-1 preparation only.

6.4.3.5 *Variation in envelope incorporated glycoproteins/proteins*

Incorporation of several virion envelope components was apparently influenced by incorporation of the pentameric complex (Figure 6.8.A). The envelope proteins that varied in relative abundance most dramatically were the novel virion components gpUL16 and gpUL148, each identified with ~3.3-, and ~6-fold greater abundance in RV-Me-tet-UL128L^{wt} virions compared to those from RV-Me-UL128L^{mut}. A further 8 envelope proteins (gpRL12, pUL33x1, gpUL78, gpUL132, pUS12, pU13, pUS14 and pUS20) were also more abundant in RV-Me-tet-UL128L^{wt} virions. Conversely, RV-Me-tet-UL128L^{wt} virions contained less gpUL118-UL119. Only 4 (gpRL10, gpUL55, gpUL116, and gpUS28) of the 21 identified envelope components were detected at comparable levels in the virions from RV-Me-tet-UL128L^{wt} and RV-Me-UL128L^{mut}. The relative abundance of the remaining 6 envelope components that were identified in the MVP-2 preparation varied between virions with contrasting pentamer/trimer ratios, but did not exceed the threshold for confidence; RV-Me-tet-UL128L^{wt} virions contained marginally greater levels of gpUL141 and gpUS27, yet less pUL41A, gpUL73, gpUL100, and pUL140. Of particular interest were gpUL73 (gN) and gpUL100 (gM) that are components of glycoprotein complex family gCII members. The identification of these envelope components with similar relative abundance between RV-Me-tet-UL128L^{wt} and RV-Me-UL128L^{mut} virions likely reflects the inclusion gM and gN in the different virions at similar stoichiometry. Furthermore, the fact that each was less abundant in RV-Me-tet-UL128L^{wt} virions compared to and RV-Me-UL128L^{mut} virions may indicate that pentameric complex incorporation may impede the incorporation of gM/gN complexes.

Ten of the 21 identified envelope proteins were present at comparable levels in virions from RV-Me-tet-UL128^{G>T} and RV-Me-tet-UL128L^{wt} virions, whilst the ratios of a further 9 envelope proteins were below the threshold of confidence (i.e. < ±1.5-fold) (Figure 6.8.B). Interestingly, RV-Me-tet-UL128^{G>T} virions contained the greatest abundance of 8 envelope proteins (gpRL12, gpUL16, gpUL141, pUS12, pUS13, pUS20, gpUS28, gpUL148). Since these proteins were more abundant in RV-Me-tet-UL128^{G>T} compared to RV-Me-tet-UL128L^{wt} virions, their incorporation was not correlated with increasing pentameric complex content of virions. In fact, incorporation of only 2 of the 21 identified envelope components correlated with increasing pentameric complex content, with RV-Me-tet-UL128^{G>T} virions potentially containing less gpUL78 and pUS14 than RV-Me-tet-UL128L^{wt} virions, yet more than in RV-Me-UL128^{mut} virions. Furthermore, whilst variation in the levels gpUL73 (gN) and gpUL100 (gM) between virions from RV-Me-tet-UL128^{G>T} and the comparison viruses did not exceed the threshold of confidence, a trend to further suggest that pentameric complex assembly may impede the incorporation of gM/gN complexes was observed.

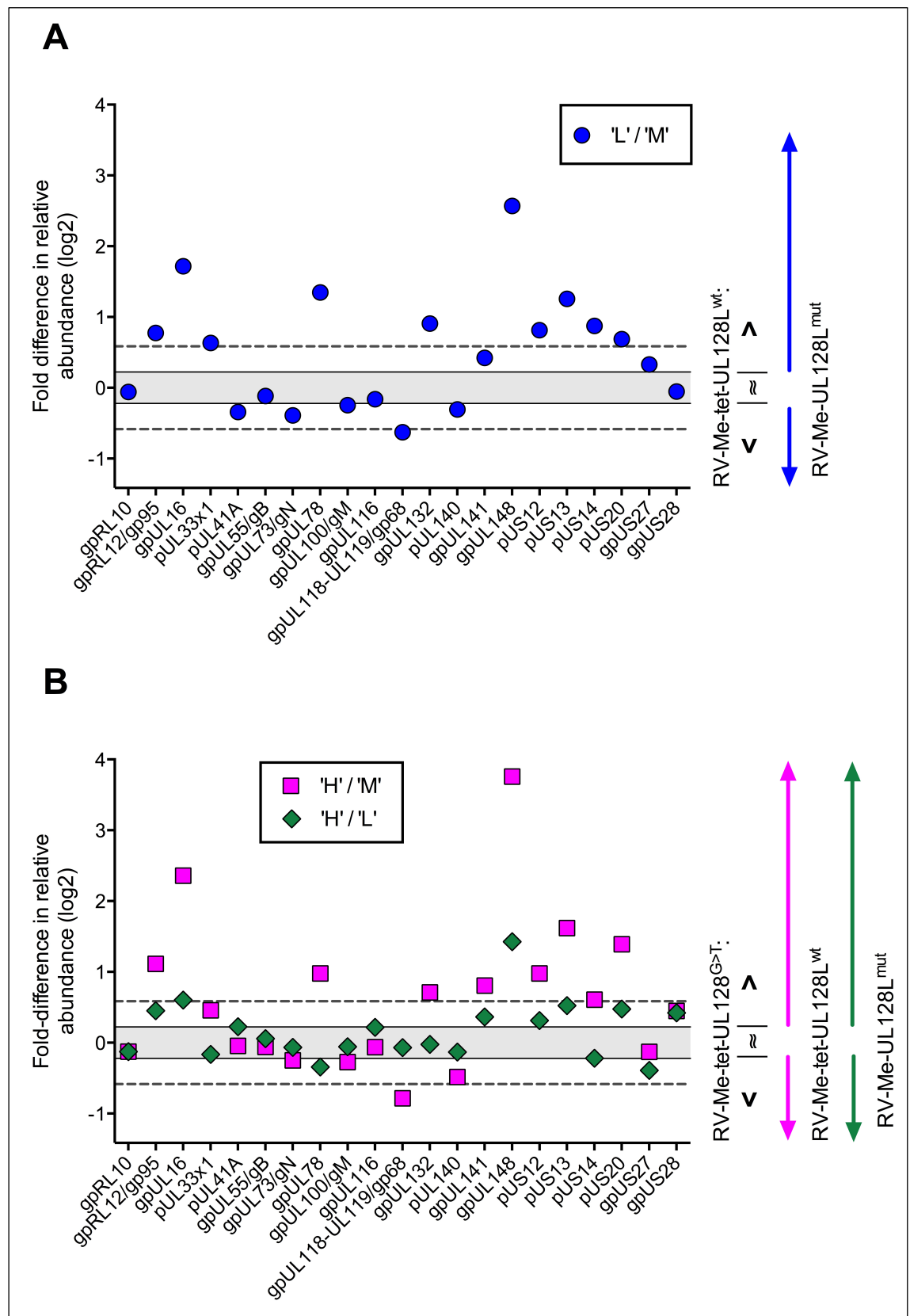


Figure 6.8. Relative abundance of membrane proteins in virions with contrasting pentameric/trimeric complex ratios. Highlighted (grey) is the degree of normal intra-assay variation ($\pm 1.1.665$ -fold). Broken lines (dark grey) show the threshold (± 1.5 -fold) to highlight proteins that differed between virions with confidence. (A) Relative abundance of proteins in RV-Me-tet-UL128L^{wt} (L) compared to RV-Me-UL128L^{mut} (M) virions. (B) Relative abundance of proteins in RV-Me-tet-UL128^{G>T} (H) compared to RV-Me-tet-UL128L^{wt} (L) and RV-Me-UL128L^{mut} (M).

6.4.3.6 Variation in uncharacterised proteins

Of the viral proteins yet to be characterised biochemically or functionally, only two were detected during SILAC-MS analysis of the MVP-2 preparation. Compared to RV-Me-UL128L^{mut} virions, RV-Me-tet-UL128L^{wt} virions contained moderately greater abundance of pUL31, yet a dramatically greater abundance of pUL145 (Figure 6.9.). Interestingly, RV-Me-tet-UL128^{G>T} virions contained the greatest abundance of these proteins. However, only pUL31 was detected with different abundance between virions from RV-Me-tet-UL128^{G>T} and RV-Me-tet-UL128L^{wt} virions above the confidence threshold. Thus, whilst virus containing non-intact UL128L ORFs and producing virions that lacked the pentameric complex contained the lowest abundance of these proteins, their incorporation could not be correlated with the pentameric/trimeric complex ratio of virions. One uncharacterised viral protein was detected only during SILAC-MS analysis of the MVP-1 preparation, with pUL150A identified with moderately greater abundance in RV-Me-tet-UL128L^{wt} virions compared to RV-Me-UL128L^{mut} virions.

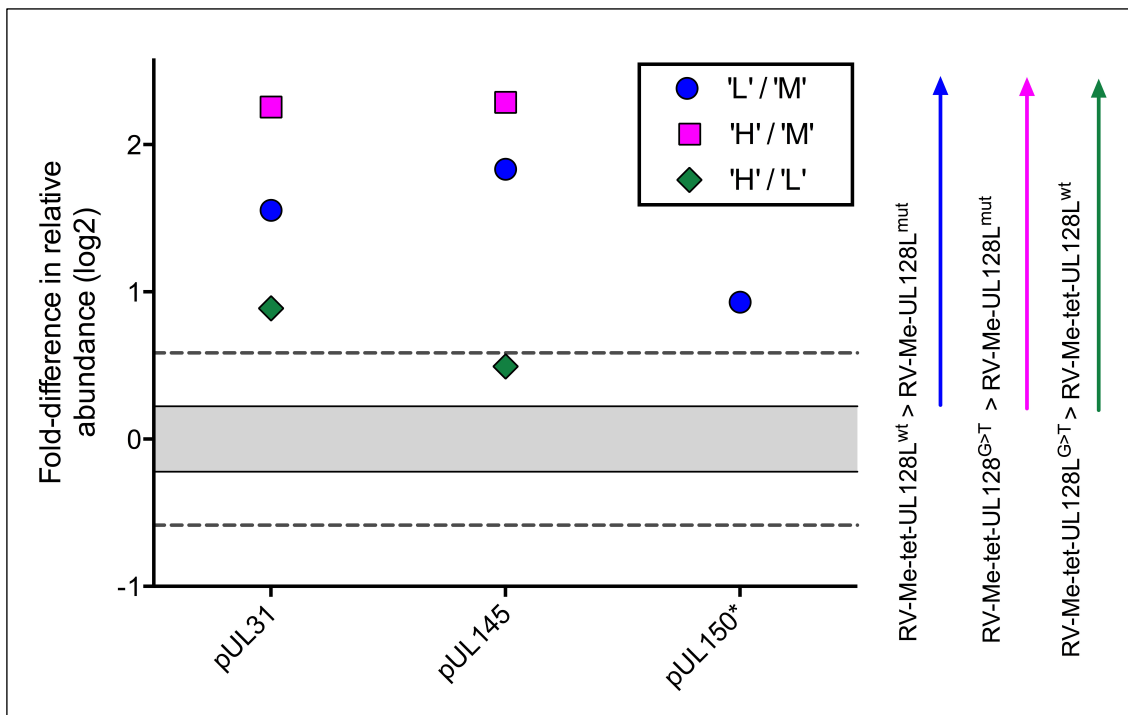


Figure 6.9. Relative abundance of uncharacterised viral proteins in virions with contrasting pentameric/trimeric complex ratios. Highlighted (grey) is the degree of acceptable intra-assay variation (± 1.1665 -fold). Broken lines (dark grey) show the threshold (± 1.5 -fold) to highlight proteins that differed between virions with confidence. **(A)** Relative abundance of proteins in RV-Me-tet-UL128L^{wt} (L) compared to RV-Me-UL128L^{mut} (M) virions. **(B)** Relative abundance of proteins in RV-Me-tet-UL128^{G>T} (H) compared to RV-Me-tet-UL128L^{wt} (L) and RV-Me-UL128L^{mut} (M). * denotes protein detected in MVP-1 preparation only.

6.5 Summary

Preliminary investigations demonstrated that the method of virus concentration prior to purification over gradients is an important consideration for any investigation of viral particles. Concentration of virus by high-speed centrifugation likely promotes aggregate formation between virus particles and contaminating proteins. This limits the recovery virus particles over gradients, and also potentially the degree of purification that can be achieved. In contrast, the reduced aggregate formation in virus stocks concentrated by dialysis ultra-filtration improves virus particle segregation and recovery over gradients, and may also limit the co-purification of contaminating proteins. Thus proteins detected in virions purified from viruses concentrated by ultra-filtration can be more confidently assumed to represent *bona fide* components.

Mass spectrometry analysis of purified Merlin virions identified a total of 92 virus-encoded proteins. This included 39 proteins that were not identified during previous by MS analysis of AD169 virions; 19 identified with high confidence, and 20 identified with low confidence. Furthermore, 13 of these proteins identified with high confidence were novel and have not been identified as HCMV virion components by any previous approach, whilst 17 of the proteins identified with low confidence were also novel components. 10 proteins previously reported to be components of HCMV virions were not identified by MS analysis of Merlin virions.

SILAC-MS analysis was validated to be a reliable approach for the comparison of virions from distinct Merlin variants; a high degree of precision in this technique was demonstrated by the reproducibility of data generated on repeat analyses, and the high degree of accuracy was demonstrated by the fact that components that should be heterogeneous amongst distinct HCMV strains and strain variants (capsid proteins) were detected with near identical abundance in the different virions. In agreement with the original hypothesis, reduced pentameric complex content of virions occurred with a concomitant increase in the abundance of the alternative gH/gL complex (gH/gL/gO). The pentamer/trimer ratio of virions also correlated with incorporation of various other virion components, and in general, virions with high pentamer/low trimer content contained increased quantities of numerous envelope components, tegument materials, and also proteins that have yet to be assigned to any particular virion compartment. Thus, viruses that have undergone adaptation to cell culture involving the selection of ablating UL128L mutations produce virions that are biochemically dissimilar to wildtype HCMV virions beyond differences in pentameric complex content.

Importantly, virions from the novel Merlin variant containing the UL128 G>T mutation, that can be propagated more efficiently and with greater stability, displayed a proteome profile that most closely matched that of Merlin virus containing wildtype UL128L, and not that of the Merlin variant lacking intact UL128L.

7 Discussion

Data described in this thesis demonstrate significant advances in addressing a major challenge in HCMV research – the efficient *in vitro* propagation of virus that closely represents the causative agent of disease. The isolation of HCMV from clinical samples has by necessity been performed on cultured fibroblasts, yet a number of functions encoded by wildtype strains inhibit the propagation of virus in this cell type. The ORFs that are most inhibitory are RL13 with unknown function, and the UL128L ORFs that encode tropism factors required for efficient infection of a wide range of other naturally targeted cell types. The passage of HCMV in cell culture (fibroblasts) invariably results in the selection of ablating mutations in these genome regions, giving rise to viruses that can be propagated more efficiently, but display restricted tropism. The cloning of the HCMV strain Merlin genome as a BAC construct and repair of all *in vitro* acquired mutations represented an important milestone in addressing these problems, providing for the first time a stable source of virus with coding potential equivalent to that of wildtype virus *in vivo* (Stanton, Baluchova et al. 2010). However, prior to this work, efficient propagation of ‘broad-tropism’ viruses (UL128L^{wt}) from the Merlin-BAC remained a technical challenge; similar to recent clinical isolates, the wildtype UL128L ORFs of Merlin impeded the dissemination and cell-free virus production of infections, and propagation in fibroblasts resulted in the selection of *de novo* UL128L mutations. Addressing these issues, the work detailed in this thesis concerns the production of novel Merlin-BAC variants engineered to contain subtle non-ablating UL128L mutations, with the ‘broad-tropism’ viruses produced growing more efficiently in fibroblasts, thereby alleviating the selective pressure imposed against the UL128L ORFs. In addition, the composition of virions from a range of Merlin variants was determined, highlighting an array of previously un-recognised virus encoded protein components that are likely to be incorporated into virions from wildtype virus *in vivo*.

7.1 The need to develop the Merlin BAC – why not use other ‘broad tropism’ strains?

The fact that the BAC-cloned Merlin genome displays coding potential equivalent to that of wildtype virus *in vivo* makes it an ideal strain for HCMV research. However, the difficulties associated with the *in vitro* propagation of the Merlin-BAC derived viruses have to date prevented the use of this strain by many laboratories. Data presented in thesis offer some explanation for the preference of the most commonly used ‘broad tropism’ BAC-cloned strains (TR-BAC, TB40-BAC4, and FIX-BAC). These strains do not suffer the above-described problems: each contains intact UL128L genome regions and produce viruses that display broad tropism, yet these viruses can be grown to high titres without the selection of *de novo* UL128L mutations. However, whilst these qualities are clearly desirable, none of these strains contain the full complement of wildtype HCMV genes. TR was cloned from a multi-drug resistant isolate and contains GCV-resistance associated mutations in UL97, and cifodovir-resistance associated mutations in UL54 (Smith, Taskintuna et al. 1998). TB40-BAC4 contains mutations

in RL5A, RL6, UL141, and UL40 (Sinzger, Hahn et al. 2008, Magri, Muntasell et al. 2011). TR-BAC, TB40-BAC4 and FIX-BAC may also contain further genetic lesions that are yet to be identified; this is particularly evident for FIX-BAC, as data in this thesis and also previous reports (Dargan, Douglas et al. 2010) indicates that this strain lacks intact copies of ORFs required for efficient infection of epithelial cells.

A major site of heterogeneity between the genomes of the TR, TB40-BAC4 and FIX BACs and wildtype virus is in the U_S region. The capture of the TB40-BAC4 and FIX genomes involved the sacrifice of ORFs US2-US5 in order to accommodate stably incorporated BAC vector sequences, however recombination events that occurred during cloning resulted in the loss of additional sequences flanking the intended site of BAC insertion (Hahn, Khan et al. 2002, Sinzger, Hahn et al. 2008). No report detailing the construction of the TR BAC is available, yet it is believed to have involved the replacement of ORFs US2-US5, similar to the construction of TB40-BAC4 and FIX (Murphy, Yu et al. 2003). Nucleotide sequence analysis of TR-BAC derived viruses produced in this work revealed discrepancies at the intended site of BAC vector insertion, prompting the re-analysis of the TR BAC coding potential; the BAC vector sequence was identified as starting from within the 5' region of US2, and ending within the 3' region in US6 (performed by Andrew Davison). Ultimately, each of TR-BAC, TB40-BAC4 and FIX-BAC lacks intact copies of US2, US3 and US6 (US4 and US5 have since been deemed non-coding; Andrew Davison – Pers. Comm.). These ORFs encode immunomodulatory functions that impede MHC complex presentation of viral antigens at the infected cell surface, and are therefore important virulence factors in natural infection (Jones, Wiertz et al. 1996, Ahn, Gruhler et al. 1997, Jones and Sun 1997, Lehner, Karttunen et al. 1997). Furthermore, both TB40-BAC4 and FIX have lost sequence close to the BAC vector insertion site, such that the a'/a sequences present at the U_S genome segment termini are fused with a portion of US2. As a result, TB40-BAC4 and FIX also lack copies of IRS1 that encodes a multi-functional protein involved in viral gene transcription transactivation (Romanowski and Shenk 1997), viral genome synthesis (Pari, Kacica et al. 1993), and subversion of intrinsic anti-viral responses (Child, Hakki et al. 2004), as well as US1 with unknown function.

As demonstrated in this work, viruses from each of the BAC-cloned versions of strains TR (TR-BAC), TB40E (TB40-BAC4) and VR1814 (FIX-BAC) are prone to rapid mutation following reconstitution in both fibroblasts and epithelial cells. This repeatedly involved the deletion of sequence at, or close to, the site of BAC vector insertion in the U_S genome region. Importantly, this occurred within the very first passage of viruses from the TR-BAC, but also in the minimum number of passages required to grow stocks of virus derived from TB40-BAC4 and FIX-BAC. The selective pressure for these mutations is likely due to the increase in genome size that occurred following insertion of BAC vector sequences, coupled with genome size restrictions during virion morphogenesis (Yu, Smith et al. 2002, Cui, McGregor et al. 2009).

However, if genome size constraints alone were responsible, compensating deletion mutations may be expected to occur in any genome region that contains multiple dispensable ORFs in tandem. Whilst the U_S ORFs involved in these mutations are clearly dispensable, there are no reports to suggest that the functions they encode are inhibitory to the growth of virus. Therefore, it seems that the BAC vector sequences in these strains itself may contain features that impede virus replication, making the U_S genome region prone to mutation.

Hence, whilst viruses from the TR, TB40-BAC4 and FIX BAC do not suffer the same growth restriction and UL128L instability seen with viruses from the wildtype Merlin-BAC, they are genetically and phenotypically dissimilar to wildtype virus, and are also prone to rapid mutation when propagated in cell culture. These limitations with the most commonly used 'broad-tropism' strains underpin the need for the development of the wildtype BAC-cloned strain Merlin.

7.2 Wildtype UL128L is inhibitory to growth of virus – but is Merlin UL128L unusual?

Data presented in this work recapitulate that of previous findings to clearly demonstrate the inhibitory impact of the wildtype UL128L of BAC-cloned strain Merlin (Stanton, Baluchova et al. 2010). However, no investigations directly comparing the UL128L genomes regions from distinct wildtype strains have previously been reported, and it remained to be determined whether natural strain-strain variations made Merlin UL128L more inhibitory to the growth of virus in cell culture. The high nucleotide sequence homology of this genome region amongst distinct wildtype strains (Baldanti, Paolucci et al. 2006, Sun, Ji et al. 2009) suggested that the growth phenotype impact would be similar, and this is now confirmed experimentally; the UL128L genome region of the un-passaged strain 3301 contained natural sequence variations compared to Merlin UL128L, yet conveyed an identical growth phenotype. The growth phenotype influence of TR UL128L (when expressed from the Merlin genome background) was also identical to that of Merlin UL128L; together with the absence of unique nucleotide variations, these observations make a strong case that TR UL128L contains fully wildtype sequence.

The inhibitory impact of TB40-BAC4 and FIX UL128L genome regions was less severe compared to Merlin wildtype UL128L due to nucleotide sequence variations in these genome regions. Whilst it is not possible to entirely discount that the nucleotides concerned are natural strain-strain variations, several lines of evidence indicate that they represent mutations, most likely acquired during the passage of each virus in cell-culture prior to BAC-cloning: they conveyed a strikingly altered growth phenotype when expressed by Merlin viruses, whereas natural variations had little to no differential impact; they reduced the fitness of virus in epithelial cells, an important cell type in natural infection; and they were unique to each strain,

and absent even in closely related or even parental strains. These subtle ‘mutations’ suppress, but do not ablate pentameric complex functionality. As a result, TB40-BAC4 and FIX BAC-derived viruses display a broad tropism phenotype, and these strains are classified as ‘clinical strains’ that are considered a close representation of wildtype virus. However, the identification of these ‘mutations’ highlight important considerations when using these, or other BAC cloned strains: i) the phenotypes of BAC-cloned strains cannot be assumed to accurately represent that of wildtype virus; ii) it is essential to compare BAC-captured genomes to that of viruses undergone no passage at all in order to determine whether genome regions of interest are wildtype in sequence.

7.3 The alternative gH/gL complex content of virions is a major factor dictating the virus growth phenotype

The effect of the non-ablating UL128L ‘mutations’ was to impede the assembly and/or incorporation of pentameric complex (gH/gL/gpUL128/gpUL130/gpUL131A) into Merlin virions, and particularly for the UL128 G>T mutation, this occurred with a concomitant increase in trimeric complex (gH/gL/gO) content. Based on this, there appears to be competition between the UL128L-encoded glycoproteins and gO for available gH/gL and the assembly of either glycoprotein complex variant. This model is supported by a similar but reciprocal observation, where a gO-null mutant that does not assemble the trimeric complex produced virions with increased pentameric complex content (Wille, Knoche et al. 2010). The relative abundance of the pentamer or trimer incorporated into virions is a major factor that dictates the growth phenotype of virus in cell culture. Both complexes are implicated in core activities during the viral replication cycle (e.g. entry of cell-free virus into cells, cell-to-cell spread, and cell-free virus production), yet there is a different requirement for each complex in different cell types. In fibroblasts, the trimeric complex is necessary for all of these activities, whilst the pentameric complex is inhibitory to cell-to-cell spread and cell-free release; hence the selective pressure imposed against UL128L. In epithelial cells, the pentameric complex is required for cell entry and cell-to-cell spread, while the role of the trimeric complex is unclear. Whilst an inverse relationship between the pentamer/trimer content of virions was shown for only a select panel of viruses grown in fibroblasts, the growth phenotypes of all Merlin variants correlated with the projected pentamer/trimer ratios of virions. In summary, high pentamer/low trimer viruses displayed reduced cell-to-cell spread, dissemination kinetics, and cell-free virus production in fibroblasts. In epithelial cells, high pentamer/low trimer viruses displayed greater cell-to-cell spread and dissemination kinetics, but reduced cell-free virus production.

7.3.1 The role of the trimer and its impact on titration data

Analyses of the virion production and tropism of the different Merlin variants were based on plaque titration data to enumerate the virion load of various samples. However,

special consideration should be given to the proposed roles of the trimer when interpreting this data. The role of trimer in cell entry is widely accepted (Vanarsdall and Johnson 2012), yet there are conflicting reports regarding a role in the production and/or cell-free release of progeny virions (Jiang, Adler et al. 2008, Wille, Knoche et al. 2010). If the trimer promotes cell-free virus production, then the ranging titres reported on analysis of the different samples correlate well with the projected pentamer/trimer content of virions, i.e. titres of low pentamer/high trimer virus samples were consistently higher than in high pentamer/low trimer virus samples. Alternatively, it is possible that the trimer content of virions influences the efficiency by which virus infects cells (see section 7.3.2). Thus, the plaque titration assay may detect virions with different pentamer/trimer ratios with different degrees of sensitivity. Ultimately this may result in the virion loads in samples from the different Merlin viruses being reported with varying degrees of accuracy.

Several lines of indirect evidence from this work support the model where the trimer contributes to the production of cell-free virus. Firstly, gradient purification of UL128L-null Merlin viruses containing the highest trimer content produced noticeably stronger ‘purified virion’ bands, and this occurred without any noticeable difference in ‘DB’ band strength that may indicate the pentamer contributes to the formation of aggregates that would limit the recovery of virions. These observations indicate a greater number of virions were released into the supernatant from infections with high trimer virus. Secondly, SILAC-MS analysis revealed that high pentamer/low trimer Merlin virions contained greater quantities of the vast majority of all other incorporated HCMV proteins. There is no evidence to suggest that UL128L or the encoded glycoproteins regulate the expression of other HCMV ORFs, and it is hard to imagine that the pentamer directly facilitates the incorporation of all other virion components by direct protein-protein interactions. Thus, the differing composition of virions most likely represents equivalent quantities of proteins being produced by the Merlin variants that are distributed between different numbers of progeny virions. Whilst this does not exclude the possibility that the virion trimer content impacts the efficiency by which virus enters cells, it supports the interpretation of titration data as reflecting different virion loads in samples.

7.3.2 Impact of pentamer/trimer ratios on cell-free virus entry

The Merlin variants clearly displayed differences in tropism. However it is not easy to determine whether this was solely due to variations in the pentamer content of virions alone, whether variations in trimer content also contributed, or whether variations in both complexes played a role but possibly with contrasting effects. It should be noted that the virions with the lowest trimer content were those produced by the Merlin variant containing wildtype UL128L; this variant can be considered a close representative of wildtype virus *in vivo*, and likely produces virions with a similar pentamer/trimer content. It would be seemingly counterintuitive

for wildtype virus to produce virions with a trimer content that did not allow efficient infection of naturally targeted cell types, particularly in epithelial cells that are first encountered following intra-host transmission of virus; the requirement for the pentameric complex for efficient entry into this cell type would mean that by default virions would contain a relatively low trimer content. Based on this, the different epitheliotropism displayed by the Merlin variants was likely dictated by the pentameric complex content of virions, and not trimeric complex content. If the virion trimer content does impact the efficiency of cell entry, then this may be most apparent in fibroblasts where the trimer but not the pentamer is required. To accurately determine the impact of varying pentamer/trimer ratios on the absolute tropism of virus for each cell type, the viruses assessed in this thesis will be need to be quantified by methods independent of the biological characteristics (e.g. EM analysis and HCMV particle counts, qPCR genome copy estimation) in parallel to plaque titration assays.

Aside from these uncertainties, the projected pentamer/trimer content of virions correlated with the relative abilities of virus to infect epithelial cells compared to fibroblasts. This was best demonstrated by comparison of viruses produced in epithelial cells: Merlin viruses containing wildtype UL128L (greatest quantities of pentameric complex, and lowest quantities of trimer) infected epithelial cells with equivalent, if not greater efficiency, than fibroblasts, whilst the novel Merlin viruses containing subtle UL128L mutations (medium pentamer, medium trimer) infected epithelial cells with reduced efficiency than fibroblasts. The efficiency by which viruses infected different epithelial cell types followed an interesting, and consistent pattern: virus lacking the pentameric complex was able to infect RPE-1s much more efficiently than ARPE19's, whilst viruses containing the complex, at any abundance, infected ARPE-19s with greater efficiency than RPE-1s. One explanation for this includes ARPE-19s being more permissive to HCMV infection than RPE-1s, but with a stricter requirement for the pentameric complex. This observation underpins the need for further investigations into the tropism of viruses in a greater range of cell types, particularly primary cells such as endothelial cells and DCs.

Interestingly, cell-free viruses produced in fibroblasts did not show the same relative epitheliotropism as their counterparts produced in epithelial cells; differences were most striking for Merlin viruses predicted to contain the greatest levels of pentameric complex (UL128L^{wt}, UL128L^{TR}). This may reflect the fact that the levels of pentameric complex incorporated into virions produced in either epithelial cells or fibroblasts differs. Indeed, analysis of the gpUL128 content in virions from the Merlin-UL128L^{wt} and Merlin-UL128^{G>T} Merlin-BAC variants and produced in epithelial cells was dramatic (by immunoblot), whilst the difference in the gpUL128 content of virions from the same BACs but produced in fibroblasts was <1.5-fold (shown by SILAC-MS). Clearly further characterisation of the novel Merlin

viruses produced in this work will require an assessment of tropism when produced in different cell types.

7.3.3 Impact of pentamer/trimer on cell-free virus production

Important clues as to how the pentameric complex impedes cell-free virus production came from investigation into the cell-free released virus (CRV) and cell-associated virus (CAV) sub-populations produced in fibroblast infections. For viruses with increasing predicted virion pentamer content, the CAV sub-population represented a greater proportion of the total progeny produced, suggesting that the pentamer may promote cell-association and/or impede the release of progeny virions. However, the CAV titres across all infections were broadly equivalent, and also considerably lower than the CRV titres from the same infections. Importantly, the CAV produced by high pentamer virus infections was unable to account for the differences in CRV infectivity produced by infections with low pentamer virus. Thus, whilst it cannot be conclusively discounted that the pentameric complex actively promotes the cell-association and/or impedes the cell-free release of virus, the major effect appears to be reducing the production of infectious virions. Furthermore, the CAV sub-populations displayed reduced epitheliotropism compared to the CRV sub-populations produced in infections, suggesting that virions retained within the cell contained reduced pentamer content. These data corroborate previous reports based on strain TB40-BAC4 where the CAV sub-populations produced in fibroblasts showed reduced endotheliotropism compared to the CRV sub-population (Scrivano, Sinzger et al. 2011), and further support a role of the trimer in progeny virion morphogenesis (Jiang, Adler et al. 2008).

Similar findings were evident on analysis of the CRV and CAV sub-populations produced in epithelial cell infections, with the progeny that remained cell-associated displaying reduced epitheliotropism compared to the cell-free released progeny. However this investigation was based only on the Merlin variant containing wildtype UL128L, and therefore firm conclusions regarding the impact of varying pentamer/trimer ratios cannot be drawn. The fact that the CAV sub-population from this infection represented a more significant site for the total progeny produced compared to that in fibroblasts infected with the same Merlin variant likely reflects the analysis of CRV and CAV sub-populations at a different time point post infection. The demonstration that UL128L inhibits cell-free virus production during infection of epithelial cells was unexpected, and conflicts with previous findings; in an early investigation, complementation of a 'broad tropism' strain with additional copies of UL128L ORFs (UL131A) did not reduce cell-free virus production during infection of endothelial cells (Adler 2006). However the 'broad-tropism' strain used was VR1814, the parental virus of the FIX BAC. VR1814 also contains a unique UL128L nucleotide (UL130 A>G that encodes an S42P amino acid substitution (discussed in section 7.4.2). This nucleotide likely represents an *in vitro*

acquired mutation that could impact the pentameric complex that may not have been overcome by complementation with UL131A. However, data presented in this thesis conclusively demonstrate that, like in fibroblasts, pentameric complex assembly impedes the production of cell-free virus in cell types where it is required by the virus for efficient infection.

7.3.4 Impact of pentamer/trimer ratio on cell-to-cell spread

The mechanisms of direct virus cell-to-cell spread are not well described. One proposed mechanism includes the same virion envelope glycoproteins that orchestrate the entry of cell-free herpesviruses being presented at the infected cell surface (Cocchi, Menotti et al. 2000, Kinzler and Compton 2005), and these induce microfusion events between the plasma membranes of adjacent cells to form pores through which virus migrates (Gerna, Percivalle et al. 2000). Recent work has confirmed the presence of all three trimer sub-units (gH, gL, gO) on the surface of fibroblasts infected with HCMV strain Merlin (Weekes, Tomasec et al. 2014). Indications that the pentameric complex is presented at the cell-surface in HCMV-infected fibroblasts, and also the pentamer and trimer are presented at the cell-surface in HCMV-infected epithelial cells, have come from experiments using recombinant expression systems (Ryckman, Rainish et al. 2008, Vanarsdall, Chase et al. 2011). However there may be different glycoprotein requirements for cell-to-cell and cell-free spread, as the trimer has been reported to be required for entry of cell-free virus into epithelial cells, yet a gO-null virus displayed no compromise in cell-to-cell spread in this cell type (Jiang, Adler et al. 2008, Wille, Knoche et al. 2010). It should also be considered that HCMV might spread cell-to-cell by alternative routes than via intracellular pores. One further possibility includes the dissemination of cell-free virus between closely opposed cells, such that infectious virions encounter a new cell immediately following egress from a producer cell.

None-the-less, data in this thesis demonstrate a direct correlation between the projected pentamer/trimer ratios of the viruses assessed and their abilities to spread by the cell-to-cell route in both fibroblasts and epithelial cells. Again, data generated in this thesis complement the reciprocal observations made from work on the gO-null mutant. This offers some explanation to the apparently conflicting ideas that UL128L expression inhibits the cell-to-cell spread of viruses in fibroblasts (Stanton, Baluchova et al. 2010), yet the pentameric complex has been reported to participate in the cell-to-cell spread of HCMV in fibroblasts; a gO-null mutant remained capable of cell-to-cell in fibroblasts, albeit displaying a small plaque phenotype, whereas a gO-null/UL128L-null double mutant was completely non-viable (Jiang, Adler et al. 2008, Wille, Knoche et al. 2010). The inhibitory impact of UL128L expression on the cell-to-cell spread of virus in fibroblasts may therefore be simply due to the reduced trimer assembly that occurs with pentamer assembly. Conversely, the previously described greater cell-to-cell

spread of gO-null viruses in epithelial cells is likely due to an increase in pentameric complex assembly (Wille, Knoche et al. 2010).

7.3.5 Impact of pentamer/trimer ratio on wildtype virus during natural infection

It is possible that the pentamer/trimer content of virions also impacts HCMV-host interactions during natural infection. Intra-host dissemination of HCMV is heavily dependent on the direct cell-to-cell transfer of virus; an obvious benefit for this mode of dissemination includes avoiding components of the humoral immune system. Data in this thesis make the case that the reduced cell-free virus production that occurs with increasing pentamer assembly is not simply due to infectious virions remaining cell-associated, but is instead due to lower infectious progeny virion production. This invokes the question as to how such a virus could disseminate throughout an infected host. Direct cell-to-cell transfer can be significantly more efficient than cell-free virus dissemination, with relatively few virions required to infect neighbouring cells (Sattentau 2008). Furthermore cell-to-cell spread may not require fully formed virions to be produced; a UL99-null mutant does not produce the tegument protein ppUL28 that orchestrates secondary envelopment for the production of enveloped virions, but is still capable of cell-to-cell spread in fibroblasts (Silva, Schroer et al. 2005). Thus, immature non-enveloped virus particles may represent one form of infectivity that disseminates by direct cell-to-cell transfer.

Based on the above described possibilities, the pentamer could have two effects to promote the cell-to-cell spread of virus: i) when expressed at the surface, it could contribute to the formation of intra-cellular pores through which virus migrates; ii) the limited trimer assembly could impede morphogenesis of progeny virions destined to egress the cell by exocytosis, with the accumulated sub-viral particles instead driven to disseminate by the direct cell-to-cell route. In this model, pentamer assembly would have different effects on the cell-to-cell spread of virus in fibroblasts: reduced progeny virion morphogenesis could drive cell-to-cell dissemination of immature viral particles, yet reduced trimer assembly could impede intracellular pore formation. It would be interesting to determine whether UL128L-null mutants arise in infected fibroblast tissues *in vivo*. The structure of fibroblast tissues in the host includes densely packed cells that are arranged in close proximity, and dissemination of HCMV throughout these tissues will most likely be heavily dependent on direct cell-to-cell transfer. The increased cell-free release of any UL128L-null mutants that arose during natural infection would result in exposure to humoral components of the immune system, selecting against these mutants. However, cell-associated UL128L-null progeny would not be exposed to the humoral immune system, and the more efficient cell-to-cell spread in fibroblasts could allow these mutants to persist.

Despite these comments concerning intra-host spread, cell-free virus must be produced for inter-host transmission. It is becoming increasingly recognised that clinical samples from an

infected individual often contain variants derived from a single strain. Following acquisition and intra-host dissemination, distinct quasi-species may arise within a host due to the selective pressures imposed during infection of different tissues (Renzette, Gibson et al. 2013). Following inter-host transmission (which occurs at epithelial tissues, and would thus require UL128L), a bottleneck process occurs, presumably with the selection of virus containing the full set of wildtype ORFs required to initiate new infection. Both wildtype UL128L and mutated UL128L populations could arise in the host, and these distinct variants may simultaneously infect glandular epithelial tissues from where virus is secreted. Genomes containing intact UL128L ORFs could complement genomes in which the UL128L genome region is mutated, and the net effect of this could be the production of progeny virions with reduced pentamer content, yet increased trimer content. These hypothetical virions would be biochemically similar to those of the novel Merlin viruses produced in this work, with greater production and release, yet still displaying epitheliotropism. Such a process was observed experimentally with the generation of UL128L mutant populations when passaging wildtype Merlin virus in epithelial cells (described below). However, arguing against this possibility, when the genomes of un-passaged virus are sequenced directly from clinical samples, they have been invariably reported to contain intact UL128L genome regions. This could indicate that a more complex process of active gene regulation may be occurring in secretory cells to suppress, but not ablate, UL128L ORF expression, or alternatively increase UL74 (gO) expression.

7.4 Novel Merlin BAC reagents: assessment, propagation and utility

The preference for cell-to-cell dissemination displayed by wildtype virus in natural infection is an important consideration, and future investigations into this method of spread are essential. However, the requirement for infectious cell-free virus preparations for experimental applications, including studies on inter-host transmission, remains equally important.

7.4.1 Merlin viruses containing TB40-BAC4 UL128L sequences

The UL128 G>T ‘mutation’ impacted pentameric complex assembly by impeding UL128 pre-mRNA splicing for the formation of translatable mRNA, and ultimately the production of gpUL128. Importantly, there was no evidence to suggest that a distinct gpUL128 isoform was produced. Several important assumptions regarding the pentameric complex assembled in the Merlin viruses containing the G>T mutation can be made: pentameric complexes are composed of UL128L encoded sub-units with wildtype amino acid sequence; thus they are arranged with normal conformation; and they are functionally equivalent to complexes composed of Merlin wildtype UL128L encoded sub-units. This also holds true for the Merlin virus containing the entire TB40-BAC4 UL128L genome region - no unique coding nucleotide variations were identified in TB40-BAC4 UL128L and all but one of the acid variations identified were also identified in the un-passaged strain 3301, and therefore likely

represented natural variations. There was discrepancy in the levels of gpUL128 production following expression of UL128 in isolation from the recombinant adenovirus vectors, with Merlin UL128 containing the G>T mutation producing more gpUL128 than TB40-BAC4 UL128. The reasons for this are unclear, but may include TB40-BAC4 UL128 mRNAs containing codons that are translated less efficiently, or have reduced thermo-stability, or produce a protein with reduced thermo-stability. If this is the case, these limitations must only be apparent in isolation from other pentamer sub-units, as virions purified from these viruses contained similar levels of gpUL128. Each of the novel Merlin viruses containing either the entire TB40-BAC4 UL128L genome region or the G>T mutation alone grew to significantly greater titres than the Merlin virus containing wildtype UL128L (in both fibroblasts and epithelial cells), yet displayed only moderately reduced relative epitheliotropism (≤ 10 -fold) compared to the severely reduced relative epitheliotropism of virus lacking intact UL128L (100-5000-fold). Therefore, each of these Merlin variants can be considered valuable reagents for future research into the biology of HCMV in cell-types where the pentameric complex is for efficient infection.

7.4.2 Merlin viruses containing FIX UL128L sequences

The FIX UL130 A>G (S72P) mutation did not impact the expression of UL130, but potentially disrupted the folding of gpUL130 into native conformation. All three members of UL128L are required for egress of the complex from the ER, explaining why modification in one component would affect the entire complex. In support of the prediction that this impacted pentameric complex sub-unit binding interactions, gpUL130 has been demonstrated to play a central role in complex assembly via interactions with both gH and gpUL131A (Ryckman, Rainish et al. 2008). In addition to affecting the amounts of UL128L components in the virion, this may also have affected the functionality of the pentameric complex that was assembled; the FIX S72P mutation resulted in virions containing significantly more pentameric complex than those containing the TB40-BAC G>A mutation, yet the relative epitheliotropism of the Merlin virus containing the S72P mutation was only marginally higher.

Unlike the TB40-BAC4 mutation, the FIX UL130 A>G (S72P) mutation contributed only partially to the distinct growth phenotypes influenced by the FIX UL128L genome region. A candidate for an additional contributing feature is the FIX UL130 A>G (S42P) variation that was present only in FIX and the parental strain VR1814. Being unique to the FIX virus lineage, this variation also likely represents an *in vitro* acquired mutation. The S42P mutation may only have a modest impact on pentameric complex functionality; it did not apparently alleviate the selective pressures of fibroblast cell culture passage since the S72P mutation was acquired later. However, given that Merlin virus containing both mutations (in the entire FIX UL128L) was

severely restricted in epithelial cells, a Merlin double mutant containing both of these features would most likely be of limited utility.

7.4.3 Propagation of novel Merlin virus stocks

The choice of cell culture system used for the propagation of viruses is an important consideration. Several observations from this work suggest that viruses produced in different cell types are phenotypically distinct. As previously mentioned, viruses produced in epithelial cells displayed greater relative epitheliotropism than counterparts from the same Merlin BAC variant that were produced in fibroblasts. Furthermore, the pentameric complex content of virions from Merlin viruses containing wildtype UL128L compared to virus containing the UL128 G>T mutation differed by different degrees when produced from epithelial cells or fibroblasts. These producer cell-type specific variations may occur due to the different capacities of cell types to translate, modify, and process the pentameric complex sub-units, as has been reported when comparing the processing of viral encoded glycoproteins in fibroblasts and astrocytoma cells (Kari, Radeke et al. 1992). An assessment of the composition of virions produced in different cell types would directly answer these questions. Nonetheless, the choice of cell type used for the production of virus stocks should likely be guided by the intended experimental use.

7.4.3.1 In epithelial cells

In experiments where virus containing the greatest pentameric complex content possible is desirable, it should be advantageous to grow viruses in epithelial cell culture. Furthermore, the UL128L genome region of viruses should be more stable in this cell type as opposed to fibroblasts. However, similar to observations made in previous studies (Dargan, Douglas et al. 2010), viruses were prone to rapid mutation during propagation in this cell type. Importantly, this occurred within the minimal passage history required to grow virus stocks, and occurred in similar timeframes across the panel of Merlin viruses. Most strikingly, the wildtype UL128L genome region of Merlin was observed to mutate on one occasion, and this occurred with concomitant insertion of *E.coli* DNA. The source of contaminating DNA was most likely derived from the bacterial cells from which BAC-DNA was prepared prior to transfection – deep sequencing of maxipreps identified bacterial DNA in some preparations, indicating that recombination could have occurred following transfection into mammalian cells. Alternatively recombination could have occurred within the bacterial cells. The expansion of this UL128L mutant population that occurred on further propagation in epithelial cells was likely due to complementation of mutated genomes with pentameric complexes derived from UL128L intact genomes present in the same cell. Whilst this virus was not used for further experimentation, it demonstrated a growth phenotype similar to viruses containing UL128L sequences from TB40-BAC4 i.e. it grew to higher titres than the parental virus, yet showed slightly reduced relative

epitheliotropism. This observation underpins the crucial importance of assessing the genetic integrity of any virus stock prior to use in experiments.

The U_L/b' genome region consistently mutated when Merlin viruses were propagated in epithelial cells, and this always involved the UL141 ORF. Given that this ORF often acquired mutations independently from the surrounding ORFs in this genome region, UL141 may be inhibitory to the growth of virus in epithelial cell culture. UL141 encodes an NK-cell evasion function that sequesters multiple NK activating ligands in the ER (Tomasec, Wang et al. 2005, Prod'homme, Sugrue et al. 2010), and is therefore an important virulence factor in natural infection. ORFs flanking UL141 that were frequently also lost in the Merlin viruses passaged in epithelial cells, including UL145, UL144, UL142, and UL140, have previously been shown to mutate during passage of virus in epithelial cells (Dargan, Douglas et al. 2010). In one Merlin virus (containing wildtype UL128L from strain 3301), all ORFs in the U_L/b' genome region (UL148-UL150) were lost during growth in epithelial cells. This resembled the mutation of strains AD169 and Towne following extensive passage in fibroblasts (Cha, Tom et al. 1996), although data presented in this work suggest that loss of U_L/b' occurs more rapidly during propagation of virus in epithelial cell culture. It remains to be determined whether all ORFs in this genome region encode temperance factors, or whether the deletion of some ORFs occurred co-incidentally with the deletion of UL141. In light of this data, it may be worth identifying the ORFs in this region that encode functions which are inhibitory to virus growth, as Merlin viruses deleted (or inhibited) for these ORFs (e.g. UL141) could be significantly more stable during passage in epithelial cell culture.

7.4.3.2 *In fibroblasts*

In contrast to the growth of viruses in epithelial cells, Merlin viruses were generally more stable during propagation in fibroblasts. Thus, virus stocks produced in this cell type may be desirable where the inclusion of the pentameric complex at any level is sufficient, yet genome wide integrity is essential. Unlike in epithelial cells where all Merlin variants mutated genetic stability in fibroblasts correlated with the fitness of virus in fibroblasts; the reduced fitness conveyed by expression of wildtype UL128L ORFs resulted in their rapid mutation as has previously been reported (Dolan, Cunningham et al. 2004, Dargan, Douglas et al. 2010, Stanton, Baluchova et al. 2010), whilst Merlin virus containing an ablating mutation in UL128L was exquisitely stable. Recapitulation of this genome-wide stability by repression of UL128L and RL13 demonstrated that without the inhibitory impact of these genome regions, the selective pressure against other genome regions in fibroblast culture is less severe. Thus, repression of the genome regions most inhibitory to the growth of virus in cell culture is an attractive option for the production of virus retaining the coding potential of the wildtype Merlin-BAC. An obvious limitation to this approach includes any virus produced in this repression culture system being restricted in tropism and suited only for infection in fibroblasts.

‘Broad-tropism’ progeny that contain all encoded virion components (gpRL13 and the pentameric complex) can be produced in subsequent infection of fibroblasts that do not endogenously express tetracycline, however expression of RL13 and UL128L would severely impede cell-free virus production. Thus, the titres of genetically stable, ‘broad-tropism’ virus stocks produced in this two-step culture system may be limiting for infections in large numbers of endothelial/epithelial cells.

Crucially, the reduced pentameric complex content of the novel Merlin viruses also conveyed a growth phenotype that alleviated the selective pressures for *de novo* UL128L mutations, and relatively few mutations were acquired at other loci in the genomes of these viruses. Mutations acquired by the Merlin virus containing the G>T mutation include non-synonymous substitutions acquired in UL47 that encoded the capsid-proximal tegument protein HMWPbp, and UL55 that encodes glycoprotein gB. Whilst the impact of these mutations was not assessed, they are unlikely to be ablating since these ORFs either enhance virus growth (UL47), or are essential to virus growth (UL55), in cell culture (Dunn, Chou et al. 2003). Surprisingly, TB40-BAC4 UL128L mutated when expressed in the Merlin genome; this mutant population accounted ~60% of all genomes present. It remains to be determined at what passage this mutation was acquired and became dominant, and also whether this was an unusual occurrence. Neither of the Merlin viruses containing FIX UL128L nucleotides acquired mutation in any HCMV ORF, although Merlin virus containing the UL130 A>G mutation did acquire an ablating deletion in the eGFP cassette, demonstrating that mutations can be acquired even in the absence of obvious selective pressure.

7.4.4 Utility of the novel Merlin reagents

With the exception of an ablating mutation in RL13 and modified UL128L ORFs, the novel viruses produced in this work display coding potential equivalent to wildtype virus *in vivo*. Hence, the composition of virions from these viruses can be considered to more accurately represent that of the causative agent of disease than other strains previously used. It can be envisaged that these qualities will aid more reliable HCMV research, including investigations into HCMV pathogenesis; HCMV tissue tropism *in vivo*; and the mechanisms of binding and entry, as well as cell-to-cell spread. A greater understanding in these aspects in HCMV biology will likely herald the development of novel therapeutic strategies.

An example of the utility of these novel reagents may include the development of vaccine strategies based on the pentameric complex sub-units. This is currently an area of great interest, and preliminary studies to assess the immunogenicity of the pentameric complex or UL128L encoded glycoproteins, as well as the homologues encoded by other animal CMVs, has been encouraging (Saccoccio, Sauer et al. 2011, Auerbach, Yan et al. 2013, Freed, Tang et al. 2013, Gnanandarajah, Gillis et al. 2014, Wen, Monroe et al. 2014). Since the novel Merlin

viruses containing TB40-BAC4 and FIX nucleotides can be propagated to higher titres without loss of UL128L genome region integrity, they may be of utility in vaccine development. The Merlin virus containing the entire TB40-BAC4 UL128L genome region may also be valuable here, providing a virus that assembles pentameric complexes complex containing natural epitope variations. Merlin viruses containing wildtype UL128L (native or derived from strain 3301) and producing virions with the greatest levels of pentameric complex may be of use as challenge strains to assess the efficacy of any vaccine produced. Further utility of these novel variants includes investigations into the alternative roles of the UL128L-encoded glycoproteins when in monomeric form, particularly gpUL128. It has recently been reported that this glycoprotein may have an additional function as a soluble chemokine homologue, and facilitates the attraction of PBMCs, and stimulates expression of TNF-alpha and IL-6 (Zheng, Tao et al. 2012, Tao, Xu et al. 2014). The novel Merlin viruses containing the UL130 A>G mutation may be of particular use here; since UL128 in this virus is expressed efficiently, but is not as efficiently incorporated into pentameric complexes, it may be secreted at greater levels. The array of novel virus-encoded components identified in Merlin virions may also include important virulence factors that are either ideal targets for novel therapies, or be responsible for the limited efficacy of existing therapies, and warrant further investigation.

7.5 Scope for further Merlin BAC development

Genome regions other than UL128L clearly contributed to the distinct growth phenotype of TR, TB40-BAC4 and FIX BAC-derived viruses. These may include features that could be exploited to further increase the efficiency by which Merlin virus is propagated in cell culture. The most obvious candidates for this include the other glycoproteins that are present in the virion envelope and contribute to the mechanisms employed during the binding and entry of virions in cells, and also the morphogenesis of progeny virions. Glycoproteins gB, gM and gN are major contributors to these processes, and each exists as multiple genotypes. Alignment of ORFs encoding these virion envelope glycoproteins to those of other strains deposited in GenBank identified no unique nucleotide features that could represent *in vitro* acquired mutations. However, the different gB genotypes have been correlated with different tropism and tissue invasiveness (Meyer-König, Vogelberg et al. 1998). Thus it could be worthwhile investigating the impact of natural variations in these glycoproteins.

Asides from the UL128L encoded glycoproteins, variations in the gH, gL, and gO could also influence assembly of either the pentamer or trimer, or alternatively the functionality of these complexes. The glycoproteins common to the pentamer and trimer, gH and gL, exist in 2 and 4 major genotypes, respectively (Chou 1992, Rasmussen, Geissler et al. 2002). Comparative nucleotide sequence analysis revealed sequence variations in the UL75 (gH) and UL115 (gL) ORFs amongst each BAC-cloned strain, indicating that each of the alternative

genotypes was represented. Glycoprotein gO displays extraordinary amino acid sequence diversity amongst strains, and exists as 8 genotypes (Rasmussen, Geissler et al. 2002). Analysis of UL74 (gO) sequences amongst the strains investigated revealed that each was of a distinct genotype; this could explain the distinct growth phenotype of some of the viruses, particularly TR BAC-derived virus that did not fit the pentamer/trimer ratio patterns displayed by all other viruses. During the time of this work, independent investigations were performed by another group to assess and compare levels of pentameric and trimeric gH/gL complexes incorporated into virions from different HCMV strains, including those derived from the Merlin and TR BACs (Zhou, Yu et al. 2013). The findings of this study potentially contradicted data presented in this thesis. Unlike in this work, virions from Merlin-BAC-derived virus contained greater levels of the pentameric complex than virions from TR-BAC-derived virus, whilst conversely, virions from the TR-BAC-derived virus contained greater levels of trimer compared to those from Merlin. This discrepancy may be explained by the fact that virions compared in this thesis were produced from epithelial cells, whilst virions compared in this alternative report were produced from fibroblasts, underlining that the producer cell can influence the relative virion composition. Data in this thesis suggest that the differences between TR and Merlin were clearly not due to UL128L, and alternatively ORFs UL75 (gH), UL115 (gL) or UL74 (gO) may be responsible. An interesting observation is the apparent chimeric gH encoded by the TR BAC, which contains sequence motifs of gH group 1 in N-terminal most domain, and sequence motifs of the gH group 2 within the internal region. It is therefore possible that the chimeric gH of TR displays greater affinity for gO than the UL128L-encoded glycoproteins. These observations open up exciting avenues of exploration and suggest that other components of the gH/gL complexes may be manipulated to the optimise the growth phenotype of HCMV *in vitro*.

Assessment of the TB40-BAC4 and TR BAC-derived viruses propagated in epithelial cells suggested that the UL84 ORF might be manipulated to increase the Merlin virus progeny turnover. This ORF encodes a multi-functional protein involved in DNA synthesis initiation at the origin of lytic replication (Ori_{Lyt}) via interactions with the viral DNA polymerase (UL44), and also as a transdominant inhibitor of IE2 (ppUL123) that is the major transactivator of early gene expression (Gebert, Schmolke et al. 1997). Interactions between pUL84 and ppUL123 increase the autologous repression of ppUL123, and suppress the cascade of viral gene transcription that precedes genome replication. Thus, UL84 has conflicting activities to both promote and impede virus replication. UL84 was initially deemed essential for virus growth (Sarisky and Hayward 1996), yet it has since been reported that a UL84-null mutant of TB40-BAC4 remained viable in HFFs (Spector and Yetming 2010). Domains that facilitate the interactions with pUL44 required for genome replication and nucleocapsid translocation are located within N-terminal residues 1-68 (Strang, Bender et al. 2012), whilst domains that are involved in the transcriptional repression interaction with ppUL123 are found in both the N-terminal and C-terminal regions (Gebert, Schmolke et al. 1997). The UL84 mutations acquired

by TR and TB40-BAC4 derived viruses in epithelial cells each resulted in the production of truncated pUL84 lacking C-terminal residues, and were dominant in each of these viruses (~80% population size). Thus it appears that domains specifically implicated in the IE2-mediated repression activity of UL84 are dispensable, and may result in a growth advantage *in vitro*.

7.6 Composition of Merlin virions – a closer reflection of wildtype virus *in vivo*

By far the most comprehensive description of the HCMV virion composition to date was that based on mass-spectrometry analysis of AD169 virions (Varnum, Streblow et al. 2004). However, given that this strain is a multiple mutant, it is not surprising that as many as 40 novel virus-encoded components (21 high confidence, 19 low confidence) were identified in wildtype Merlin virions. The MS technique is limited in its ability to detect proteins that may present at low levels within a complex mixture, as well as those that are small and release few signature peptides following trypsinisation. As previously mentioned, increasing virion production appears to result in virions containing lower levels of almost all incorporated components. AD169 displays virion production that is greater even than that of the Merlin virus containing non-intact UL128L, and the abundance of virus encoded proteins incorporated into AD169 virion may be further reduced. Thus, several of the novel proteins identified in Merlin virions could in fact be incorporated into AD169 virions, though at levels undetectable by MS. Recent advances in mass spectrometers may also account for the detection of a greater number of proteins in the present study. Furthermore, on MS analysis, protein identification is achieved by comparing peptide detection events to a database of predicted peptide signatures. Interpretation of ORFs within the HCMV genome have been subject to many revisions since MS analysis of AD169 virions, and the more comprehensive database used in this thesis may have facilitated the more reliable identification of proteins. For example gpRL14 was detected in AD169 virions, though not in Merlin virions, since RL14 has been deemed non-coding in more recent interpretations. On this note it should be appreciated that the database used to identify proteins included in the Merlin virions did not account for the 700+ distinct ORFs as are predicted by ribosome profiling (Stern-Ginossar, Weisburd et al. 2012). Thus it is likely that HCMV virions contain further components that were not identified during this work. An alternative reason for the differences between Merlin and AD169 virions may also include the different preparation protocols used. Purification of AD169 virions involved an initial centrifugation of virus over glycerol prior to gradients. However, data presented in this thesis demonstrated the centrifugation of virus over glycerol could result in significant aggregate formation, potentially explaining detection of gpUL22A in AD169 but not Merlin virions.

Of the novel HCMV virion components identified, several are implicated in viral genome transcription, replication and encapsidation. This includes, pUL70, the helicase-primase

sub-unit 2 that was identified with high confidence, as well as pUL102 and pUL105, helicase-primase sub-units 3 and 1 respectively that were identified with low confidence (McMahon and Anders 2002). Further participants in viral genome replication are pUL98, an alkaline deoxyribonuclease (Kuchta, Parikh et al. 2012), and ppUL114, a Uracil-DNA glycosylase (Courcelle, Courcelle et al. 2001), each implicated in the repair of viral genomic DNA during synthesis and identified with high confidence. Others are involved in the transcription of the viral genome, including pUL34 that suppresses expression of US3 (LaPierre and Biegelke 2001), and pUL123/IE1 protein that enhances activation of IE2 (the major transactivator of viral gene expression) (Malone, Vesole et al. 1990). pUL52 orchestrates DNA encapsidation and nucleocapsid maturation/formation (Borst, Wagner et al. 2008). The significance of these novel virion components is hard to predict, however their identification may simply reflect the sensitive detection of components that remain in association with the genome or capsid at residual levels of following virion morphogenesis. Alternatively, the detection of these proteins may also reflect limitation in the purity that can be achieved when isolating virions, together with the highly sensitive detection by MS-analysis. It would be interesting to determine the virion compartment to which the viral chemokine homologues pUL146/vCXC-1 and pUL147/vCXC-2 localise. pUL146/vCXC-1 has been shown to be an agonistic ligand for cellular chemokine receptors CXCR1 and CXCR2 to attract neutrophils and DCs to sites of infected endothelial cells (Luttichau 2010). However, this has been demonstrated for soluble pUL146/vCXC-1, and the role of this protein when contained in virions is unclear. Most interestingly, up to 14 novel envelope components (7 with high confidence, and 7 with low confidence) were also identified. These may be involved in the binding, docking and entry of virus in permissive cells, and potentially represent targets to block intra-host virus dissemination and inter-host transmission. Like the RL11 family members gpRL1 and gpRL13, gpRL12 has possesses anti-body Fc binding and sequestration activities, though this has only been demonstrated when expressed from a recombinant expression system (Cortese, Calo et al. 2012, Cortese 2013). None of pUS9, pUS12, pUS13, pUS14, are described functionally, however by homology to US16 (another US12 family member) (Bronzini, Luganini et al. 2012), these novel proteins could be tropism factors for epithelial and endothelial cells. US18 and US20 are also members of the US12 family and have recently been described to encode NK-cell evasion functions by promoting the lysosomal degradation of MIC A, an activating ligand for the NK cell receptor NKG2D (Fielding, Aicheler et al. 2014). However it is hard to see how this activity would be significant in the virion. Similarly, none of gpUL148, pUL140, or gpUL139, are well described functionally, but the ORFs encoding these envelope components are situated in the U_L/b' region where multiple proteins are immuno-modulatory, or virulence factors. The novel component gpUL141 is another example of an immuno-modulatory protein encoded in this genome region. This glycoprotein binds CD155 that is also found at the junctions of adjacent epithelial cells (Ogita and Takai 2006), and has been described as a receptor for

poliovirus entry (Mendelsohn, Wimmer et al. 1989). Thus, when presented in the virion, gpUL141 may be implicated in the entry of virus in this cell type, or perhaps cell-to-cell spread. Conversely, gpUL16 also impairs NK activity, but has been reported to impair the replication of virus in epithelial cells (Mocarski, Shenk et al. 2007). The roles of gpUL74A and gpUL116 in the virion are unknown. Peptides from >2000 cellular proteins were detected by MS analysis of Merlin virions, though the number that met the criteria for high confidence identification would likely be much lower. However the significance of these proteins is not easily predicted, as the cellular proteins incorporated into virions produced in cell culture may be significantly different to those incorporated *in vivo*. To validate proteins identified in Merlin virions as *bona fide* HCMV virion components, and also identify cellular proteins that are incorporated into wildtype virions during natural infection, analysis of virions directly from clinical samples will need to be performed.

As previously mentioned, virions containing greater levels of pentamer and lower levels of trimer, or low production of virions, generally contained greater levels of all other components. In fact, only 5 virus-encoded proteins were present at reduced levels as pentamer levels increased: pUL146/vCXC-1, and envelope components gp41A, gpUL118-gpUL119, gN/gpUL73 and pUL140. Of these, the levels of only pUL146/vCXC-1 and gpUL118-gpUL119 varied by a significant degree, potentially indicating direct interaction with the trimer. In contrast to this, proteins that were significantly more abundant with increasing pentamer levels included tegument protein pUL36x1, pUL147/vCXC-1, as well as the envelope glycoproteins gpUL16 and gpUL148. The biological significance of these variations cannot simply be inferred from the difference in relative abundance, but it should be appreciated that experiments performed with UL128L-null virus may not accurately reflect the true magnitude of observed biological effects. It is of interest that whilst the majority viral proteins did vary between the different virions assessed by SILAC-MS analysis, others were present at comparable abundance. Since the capsid vertex complex sub-units pUL77 and pUL93 will likely be incorporated into different virions with at equimolar quantities, the equivalent relative abundance of these proteins amongst the different virions potentially offers a further demonstration of the high degree of accuracy in these assays. The other proteins that did not apparently vary in abundance between virions may be expressed with efficiency such that they are not limiting for the assembly of virions.

In summary, the Merlin virion proteome data presented in this thesis provides a more comprehensive picture of the composition of wildtype HCMV virions *in vivo*, and highlights a range of components that may be important as tropism determinants, or even immune evasion functions. Moreover, virions from the newly produced Merlin variants were found to contain a proteome that closely reflected those of the parental wildtype Merlin-BAC, further demonstrating the suitability of these reagents for future research.

8 References

- Aberle, J. H. and E. Puchhammer-Stöckl (2012). "Age-dependent increase of memory B cell response to cytomegalovirus in healthy adults." Experimental gerontology **47**(8): 654-657.
- Adams, M. J. and E. B. Carstens (2012). "Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses (2012)." Arch Virol **157**(7): 1411-1422.
- Adler, B. and C. Sinzger (2009). "Endothelial cells in human cytomegalovirus infection: One host cell out of many or a crucial target for virus spread?" Thrombosis and Haemostasis **102**(6): 1057-1063.
- Adler, B. S., L.; Ruzcics, Z.; Rupp, B.; Sinzger, C.; Koszinowski, U. (2006). "Role of human cytomegalovirus UL131A in cell type-specific virus entry and release." Journal of General Virology **87**: 2451-2460.
- Ahn, K., A. Gruhler, B. Galocha, T. R. Jones, E. J. Wiertz, H. L. Ploegh, P. A. Peterson, Y. Yang and K. Fruh (1997). "The ER-luminal domain of the HCMV glycoprotein US6 inhibits peptide translocation by TAP." Immunity **6**(5): 613-621.
- 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**(Pt 5): 1117-1122.
- Alexander, B. T., L. M. Hladnik, K. M. Augustin, E. Casabar, P. S. McKinnon, R. M. Reichley, D. J. Ritchie, P. Westervelt and E. R. Dubberke (2010). "Use of cytomegalovirus intravenous immune globulin for the adjunctive treatment of cytomegalovirus in hematopoietic stem cell transplant recipients." Pharmacotherapy **30**(6): 554-561.
- Andrews, C. N. B., F.M.; Enders, G.K.; Hirst, M.M.; Zhdanov, V.M. (1961). "Taxonomy of viruses infecting vertebrates: present knowledge and ignorance." Virology **15**: 52-55.
- Arav-Boger, R., C. B. Foster, J. C. Zong and R. F. Pass (2006). "Human cytomegalovirus-encoded alpha -chemokines exhibit high sequence variability in congenitally infected newborns." J Infect Dis **193**(6): 788-791.
- 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**(8): 1057-1064.
- Arav-Boger, R., J. C. Zong and C. B. Foster (2005). "Loss of linkage disequilibrium and accelerated protein divergence in duplicated cytomegalovirus chemokine genes." Virus Genes **31**(1): 65-72.
- Arnon, T. I., H. Achdout, O. Levi, G. Markel, N. Saleh, G. Katz, R. Gazit, T. Gonen-Gross, J. Hanna, E. Nahari, A. Porgador, A. Honigman, B. Plachter, D. Mevorach, D. G. Wolf and O. Mandelboim (2005). "Inhibition of the NKp30 activating receptor by pp65 of human cytomegalovirus." Nat Immunol **6**(5): 515-523.
- Arthurs, S. K., A. J. Eid, R. A. Pedersen, W. K. Kremers, F. G. Cosio, R. Patel and R. R. Razonable (2008). "Delayed-onset primary cytomegalovirus disease and the risk of allograft failure and mortality after kidney transplantation." Clin Infect Dis **46**(6): 840-846.

Atalay, R., A. Zimmermann, M. Wagner, E. Borst, C. Benz, M. Messerle and H. Hengel (2002). "Identification and expression of human cytomegalovirus transcription units coding for two distinct Fcγ receptor homologs." *J Virol* **76**(17): 8596-8608.

Auerbach, M., D. Yan, A. Fouts, M. Xu, A. Estevez, C. D. Austin, F. Bazan and B. Feuerbach (2013). "Characterization of the guinea pig CMV gH/gL/GP129/GP131/GP133 complex in infection and spread." *Virology*.

Baldanti, F., S. Paolucci, G. Campanini, A. Sarasini, E. Percivalle, M. G. Revello and G. Gerna (2006). "Human cytomegalovirus UL131A, UL130 and UL128 genes are highly conserved among field isolates." *Arch Virol* **151**(6): 1225-1233.

Baldick Jr, C. J. and T. Shenk (1996). "Proteins associated with purified human cytomegalovirus particles." *Journal of virology* **70**(9): 6097-6105.

Baquero, E., A. A. Albertini, P. Vachette, J. Lepault, S. Bressanelli and Y. Gaudin (2013). "Intermediate conformations during viral fusion glycoprotein structural transition." *Curr Opin Virol* **3**(2): 143-150.

Bates, M., M. Monze, H. Bima, M. Kapambwe, F. C. Kasolo, U. A. Gompels and C. s. group (2008). "High human cytomegalovirus loads and diverse linked variable genotypes in both HIV-1 infected and exposed, but uninfected, children in Africa." *Virology* **382**(1): 28-36.

Bechtel, J. T. and T. Shenk (2002). "Human cytomegalovirus UL47 tegument protein functions after entry and before immediate-early gene expression." *Journal of virology* **76**(3): 1043-1050.

Becke, S., V. Fabre-Mersseman, S. Aue, S. Auerochs, T. Sedmak, U. Wolfrum, D. Strand, M. Marschall, B. Plachter and S. Reyda (2010). "Modification of the major tegument protein pp65 of human cytomegalovirus inhibits virus growth and leads to the enhancement of a protein complex with pUL69 and pUL97 in infected cells." *J Gen Virol* **91**(Pt 10): 2531-2541.

Bentz, G. L., M. Jarquin-Pardo, G. Chan, M. S. Smith, C. Sinzger and A. D. Yurochko (2006). "Human cytomegalovirus (HCMV) infection of endothelial cells promotes naive monocyte extravasation and transfer of productive virus to enhance hematogenous dissemination of HCMV." *J Virol* **80**(23): 11539-11555.

Bevot, A., K. Hamprecht, I. Krageloh-Mann, S. Brosch, R. Goelz and B. Vollmer (2012). "Long-term outcome in preterm children with human cytomegalovirus infection transmitted via breast milk." *Acta Paediatr* **101**(4): e167-172.

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." *Journal of molecular biology* **295**(2): 155-161.

Bhushan, S., H. Meyer, A. L. Starosta, T. Becker, T. Mielke, O. Berninghausen, M. Sattler, D. N. Wilson and R. Beckmann (2010). "Structural basis for translational stalling by human cytomegalovirus and fungal arginine attenuator peptide." *Mol Cell* **40**(1): 138-146.

Billstrom, M. A. and W. J. Britt (1995). "Postoligomerization folding of human cytomegalovirus glycoprotein B: identification of folding intermediates and importance of disulfide bonding." *J Virol* **69**(11): 7015-7022.

Bimboim, H. and J. Doly (1979). "A rapid alkaline extraction procedure for screening recombinant plasmid DNA." Nucleic acids research **7**(6): 1513-1523.

Biron, C. A., K. S. Byron and J. L. Sullivan (1989). "Severe herpesvirus infections in an adolescent without natural killer cells." N Engl J Med **320**(26): 1731-1735.

Bissinger, A. L., C. Sinzger, E. Kaiserling and G. Jahn (2002). "Human cytomegalovirus as a direct pathogen: Correlation of multiorgan involvement and cell distribution with clinical and pathological findings in a case of congenital inclusion disease." Journal of Medical Virology **67**(2): 200-206.

Bodaghi, B., M. E. P. Slobbe-van Drunen, A. Topilko, E. Perret, R. C. R. M. Vossen, M. C. E. 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." Investigative Ophthalmology & Visual Science **40**(11): 2598-2607.

Boeckh, M. (2011). "Complications, diagnosis, management, and prevention of CMV infections: current and future." Hematology Am Soc Hematol Educ Program **2011**(1): 305-309.

Boeckh, M., W. G. Nichols, G. Papanicolaou, R. Rubin, J. R. Wingard and J. Zaia (2003). "Cytomegalovirus in hematopoietic stem cell transplant recipients: Current status, known challenges, and future strategies." Biol Blood Marrow Transplant **9**(9): 543-558.

Boppana, S. B., K. B. Fowler, R. F. Pass, L. B. Rivera, R. D. Bradford, F. D. Lakeman and W. J. Britt (2005). "Congenital cytomegalovirus infection: association between virus burden in infancy and hearing loss." J Pediatr **146**(6): 817-823.

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." Journal of virology **73**(10): 8320-8329.

Borst, E. M., K. Wagner, A. Binz, B. Sodeik and M. Messerle (2008). "The essential human cytomegalovirus gene UL52 is required for cleavage-packaging of the viral genome." Journal of Virology **82**(5): 2065-2078.

Boyle, K. A. and T. Compton (1998). "Receptor-binding properties of a soluble form of human cytomegalovirus glycoprotein B." Journal of virology **72**(3): 1826-1833.

Bradley, A. J., I. J. Kovacs, D. Gatherer, D. J. Dargan, K. R. Alkharsah, P. K. Chan, W. F. Carman, M. Dedicoat, V. C. Emery, C. C. Geddes, G. Gerna, B. Ben-Ismaeil, S. Kaye, A. McGregor, P. A. Moss, R. Puztai, W. D. Rawlinson, G. M. Scott, G. W. Wilkinson, T. F. Schulz and A. J. Davison (2008). "Genotypic analysis of two hypervariable human cytomegalovirus genes." J Med Virol **80**(9): 1615-1623.

Bradley, A. J., N. S. Lurain, P. Ghazal, U. Trivedi, C. Cunningham, K. Baluchova, D. Gatherer, G. W. Wilkinson, D. J. Dargan and A. J. Davison (2009). "High-throughput sequence analysis of variants of human cytomegalovirus strains Towne and AD169." J Gen Virol **90**(Pt 10): 2375-2380.

Britt, W. (2007). Virus entry into host, establishment of infection, spread in host, mechanisms of tissue damage. Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis. A. Arvin, G. Campadelli-Fiume, E. Mocarski et al., Cambridge university Press.

Britt, W. (2008). Manifestations of human cytomegalovirus infection: Proposed mechanisms of acute and chronic disease. T. E. Shenk and M. F. Stinski. **325**: 417-470.

Britt, W. (2008). "Manifestations of human cytomegalovirus infection: proposed mechanisms of acute and chronic disease." Curr Top Microbiol Immunol **325**: 417-470.

Britt, W. J. and D. Auger (1986). "Synthesis and processing of the envelope gp55-116 complex of human cytomegalovirus." J Virol **58**(1): 185-191.

Britt, W. J. and L. G. Vugler (1989). "Processing of the gp55-116 envelope glycoprotein complex (gB) of human cytomegalovirus." J Virol **63**(1): 403-410.

Britt, W. J. and L. G. Vugler (1992). "Oligomerization of the human cytomegalovirus major envelope glycoprotein complex gB (gp55-116)." J Virol **66**(11): 6747-6754.

Brock, I., M. Kruger, T. Mertens and J. von Einem (2013). "Nuclear Targeting of Human Cytomegalovirus Large Tegument Protein pUL48 Is Essential for Viral Growth." Journal of Virology **87**(10): 6005-6019.

Bronzini, M., A. Luganini, V. Dell'Oste, M. De Andrea, S. Landolfo and G. Griboaldo (2012). "The US16 Gene of Human Cytomegalovirus Is Required for Efficient Viral Infection of Endothelial and Epithelial Cells." Journal of virology **86**(12): 6875-6888.

Brown, J. C. and W. W. Newcomb (2011). "Herpesvirus capsid assembly: insights from structural analysis." Curr Opin Virol **1**(2): 142-149.

Brown, J. M., H. Kaneshima and E. S. Mocarski (1995). "Dramatic interstrain differences in the replication of human cytomegalovirus in SCID-hu mice." J Infect Dis **171**(6): 1599-1603.

Brunak, S., J. Engelbrecht and S. Knudsen (1991). "Prediction of human mRNA donor and acceptor sites from the DNA sequence." J Mol Biol **220**(1): 49-65.

Buchkovich, N. J., T. G. Maguire, A. W. Paton, J. C. Paton and J. C. Alwine (2009). "The endoplasmic reticulum chaperone BiP/GRP78 is important in the structure and function of the human cytomegalovirus assembly compartment." J Virol **83**(22): 11421-11428.

Bughio, F., D. A. Elliott and F. Goodrum (2013). "An endothelial cell-specific requirement for the UL133-UL138 locus of human cytomegalovirus for efficient virus maturation." J Virol **87**(6): 3062-3075.

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**(6): 3042-3048.

Cantrell, S. R. and W. A. Bresnahan (2005). "Interaction between the human cytomegalovirus UL82 gene product (pp71) and hDaxx regulates immediate-early gene expression and viral replication." Journal of Virology **79**(12): 7792-7802.

Cardone, G., J. B. Heymann, N. Cheng, B. L. Trus and A. C. Steven (2012). "Procapsid assembly, maturation, nuclear exit: dynamic steps in the production of infectious herpesvirions." Adv Exp Med Biol **726**: 423-439.

Casarosa, P., M. Waldhoer, P. J. LiWang, H. F. Vischer, T. Kledal, H. Timmerman, T. W. Schwartz, M. J. Smit and R. Leurs (2005). "CC and CX3C chemokines differentially interact with the N terminus of the human cytomegalovirus-encoded US28 receptor." J Biol Chem **280**(5): 3275-3285.

Castagnola, E., B. Cappelli, D. Erba, A. Rabagliati, E. Lanino and G. Dini (2004). "Cytomegalovirus infection after bone marrow transplantation in children." Hum Immunol **65**(5): 416-422.

Cepeda, V., M. Esteban and A. Fraile - Ramos (2010). "Human cytomegalovirus final envelopment on membranes containing both trans - Golgi network and endosomal markers." Cellular microbiology **12**(3): 386-404.

Cha, T., 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." Journal of virology **70**(1): 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**(7): 5757-5766.

Chee, M., A. Bankier, S. Beck, R. Bohni, C. Brown, R. Cerny, T. Horsnell, C. Hutchison III, T. Kouzarides and J. Martignetti (1990). Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169. Cytomegaloviruses, Springer: 125-169.

Chee, M., S. A. Rudolph, B. Plachter, B. Barrell and G. Jahn (1989). "Identification of the major capsid protein gene of human cytomegalovirus." J Virol **63**(3): 1345-1353.

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**(1): 10-16.

Chevillotte, M., S. Landwehr, L. Linta, G. Frascaroli, A. Luske, C. Buser, T. Mertens and J. von Einem (2009). "Major tegument protein pp65 of human cytomegalovirus is required for the incorporation of pUL69 and pUL97 into the virus particle and for viral growth in macrophages." J Virol **83**(6): 2480-2490.

Child, S. J., M. Hakki, K. L. De Niro and A. P. Geballe (2004). "Evasion of cellular antiviral responses by human cytomegalovirus TRS1 and IRS1." J Virol **78**(1): 197-205.

Chou, S. (1992). "Comparative analysis of sequence variation in gp116 and gp55 components of glycoprotein B of human cytomegalovirus." Virology **188**(1): 388-390.

Chou, S. (1992). "Molecular epidemiology of envelope glycoprotein H of human cytomegalovirus." J Infect Dis **166**(3): 604-607.

Chou, S. (2008). "Cytomegalovirus UL97 mutations in the era of ganciclovir and maribavir." Rev Med Virol **18**(4): 233-246.

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**(6): 1229-1234.

Chowdary, T. K., T. M. Cairns, D. Atanasiu, G. H. Cohen, R. J. Eisenberg and E. E. Heldwein (2010). "Crystal structure of the conserved herpesvirus fusion regulator complex gH-gL." Nature structural & molecular biology **17**(7): 882-888.

Cocchi, F., L. Menotti, P. Dubreuil, M. Lopez and G. Campadelli-Fiume (2000). "Cell-to-cell spread of wild-type herpes simplex virus type 1, but not of syncytial strains, is mediated by the immunoglobulin-like receptors that mediate virion entry, nectin1 (PRR1/HveC/HlgR) and nectin2 (PRR2/HveB)." J Virol **74**(8): 3909-3917.

Compton, T. (2004). "Receptors and immune sensors: the complex entry path of human cytomegalovirus." Trends Cell Biol **14**(1): 5-8.

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**(1): 387-395.

Compton, T., D. M. Nowlin and N. R. Cooper (1993). "Initiation of human cytomegalovirus infection requires initial interaction with cell surface heparan sulfate." Virology **193**(2): 834-841.

Cook, C. H. and J. Trgovcich (2011). "Cytomegalovirus reactivation in critically ill immunocompetent hosts: a decade of progress and remaining challenges." Antiviral Res **90**(3): 151-159.

Cordonnier, C., S. Chevret, M. Legrand, H. Rafi, N. Dhedin, B. Lehmann, F. Bassompierre, E. Gluckman and G. S. Group (2003). "Should immunoglobulin therapy be used in allogeneic stem-cell transplantation? A randomized, double-blind, dose effect, placebo-controlled, multicenter trial." Ann Intern Med **139**(1): 8-18.

Cortese, M. (2013). "Identification and functional characterization of human cytomegalovirus Fc binding proteins."

Cortese, M., S. Calo, R. D'Aurizio, A. Lilja, N. Pacchiani and M. Merola (2012). "Recombinant human cytomegalovirus (HCMV) RL13 binds human immunoglobulin G Fc." PLoS One **7**(11): e50166.

Courcelle, C. T., J. Courcelle, M. N. Prichard and E. S. Mocarski (2001). "Requirement for uracil-DNA glycosylase during the transition to late-phase cytomegalovirus DNA replication." J Virol **75**(16): 7592-7601.

Craighead, J. E., R. E. Kanich and J. D. Almeida (1972). "Nonviral microbodies with viral antigenicity produced in cytomegalovirus-infected cells." J Virol **10**(4): 766-775.

Cranage, M., T. Kouzarides, A. Bankier, S. Satchwell, K. Weston, P. Tomlinson, B. Barrell, H. Hart, S. Bell and A. Minson (1986). "Identification of the human cytomegalovirus glycoprotein

B gene and induction of neutralizing antibodies via its expression in recombinant vaccinia virus." The EMBO journal **5**(11): 3057.

Cranage, M. P., G. Smith, S. Bell, H. Hart, C. Brown, A. Bankier, P. Tomlinson, B. Barrell and T. Minson (1988). "Identification and expression of a human cytomegalovirus glycoprotein with homology to the Epstein-Barr virus BXLF2 product, varicella-zoster virus gpIII, and herpes simplex virus type 1 glycoprotein H." Journal of virology **62**(4): 1416-1422.

Crump, C. M., C.-H. Hung, L. Thomas, L. Wan and G. Thomas (2003). "Role of PACS-1 in trafficking of human cytomegalovirus glycoprotein B and virus production." Journal of virology **77**(20): 11105-11113.

Cui, X., R. Lee, S. P. Adler and M. A. McVoy (2013). "Antibody inhibition of human cytomegalovirus spread in epithelial cell cultures." J Virol Methods **192**(1-2): 44-50.

Cui, X., A. McGregor, M. R. Schleiss and M. A. McVoy (2009). "The impact of genome length on replication and genome stability of the herpesvirus guinea pig cytomegalovirus." Virology **386**(1): 132-138.

Cui, X., B. P. Meza, S. P. Adler and M. A. McVoy (2008). "Cytomegalovirus vaccines fail to induce epithelial entry neutralizing antibodies comparable to natural infection." Vaccine **26**(45): 5760-5766.

Cunningham, C., D. Gatherer, B. Hilfrich, K. Baluchova, D. J. Dargan, M. Thomson, P. D. Griffiths, G. W. Wilkinson, T. F. Schulz and A. J. Davison (2010). "Sequences of complete human cytomegalovirus genomes from infected cell cultures and clinical specimens." J Gen Virol **91**(Pt 3): 605-615.

Dai, X., X. Yu, H. Gong, X. Jiang, G. Abenes, H. Liu, S. Shivakoti, W. J. Britt, H. Zhu, F. Liu and Z. H. Zhou (2013). "The smallest capsid protein mediates binding of the essential tegument protein pp150 to stabilize DNA-containing capsids in human cytomegalovirus." PLoS Pathog **9**(8): e1003525.

Dal Monte, P., S. Pignatelli, M. Mach and M. P. Landini (2001). "The product of human cytomegalovirus UL73 is a new polymorphic structural glycoprotein (gpUL73)." Journal of Human Virology **4**(1): 26-34.

Dargan, D. J., E. Douglas, C. Cunningham, F. Jamieson, R. J. Stanton, K. Baluchova, B. P. McSharry, P. Tomasec, V. C. Emery, E. Percivalle, A. Sarasini, G. Gerna, G. W. G. Wilkinson and A. J. Davison (2010). "Sequential mutations associated with adaptation of human cytomegalovirus to growth in cell culture." Journal of General Virology **91**(6): 1535-1546.

Das, S., A. Vasanji and P. E. Pellett (2007). "Three-dimensional structure of the human cytomegalovirus cytoplasmic virion assembly complex includes a reoriented secretory apparatus." J Virol **81**(21): 11861-11869.

Davison, A. and D. Bhella (2007). Basic virology and viral gene effects on host cell functions: betaherpesviruses: comparative genome and virion structure. . Human Herpesviruses: Biology, Therapy and Immunoprophylaxis. A. C.-F. Arvin, G.; Mocarski, E.; Moore, P.S.; Roizman, B.; Whitley, R.; Yamanishi, K., Cambridge University Press: 177-203.

Davison, A. J. (2002). "Evolution of the herpesviruses." Veterinary microbiology **86**(1): 69-88.

Davison, A. J. (2007). Introduction: definition and classification of the human herpesviruses: comparative analysis of the genomes. . Human Herpesviruses: Biology, Therapy and Immunoprophylaxis. A. C.-F. Arvin, G.; Mocarski, E.; Moore, P.S.; Roizman, B.; Whitley, R.; Yamanishi, K., Cambridge University Press: 10-26.

Davison, A. J. (2010). "Herpesvirus systematics." Vet Microbiol **143**(1): 52-69.

Davison, A. J., P. Akter, C. Cunningham, A. Dolan, C. Addison, D. J. Dargan, A. F. Hassan-Walker, V. C. Emery, P. D. Griffiths and G. W. G. Wilkinson (2003). "Homology between the human cytomegalovirus RL11 gene family and human adenovirus E3 genes." Journal of General Virology **84**(3): 657-663.

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**(Pt 1): 17-28.

Davison, A. J., R. Eberle, B. Ehlers, G. S. Hayward, D. J. McGeoch, A. C. Minson, P. E. Pellett, B. Roizman, M. J. Studdert and E. Thiry (2009). "The order Herpesvirales." Arch Virol **154**(1): 171-177.

Davison, A. J. H., M.; Dolan, A.; Dragen, D.K.; Gatherer, D.; Hayward, G.S. (2013). Comparative Genomics of Primate Cytomegaloviruses. Cytomegaloviruses: From Molecular Pathogenesis to Intervention. M. J. L. Reddehase, N.A.W., Caister Academic Press. **I**: 1-22.

Deayton, J. R., C. A. Prof Sabin, M. A. Johnson, V. C. Emery, P. Wilson and P. D. Griffiths (2004). "Importance of cytomegalovirus viraemia in risk of disease progression and death in HIV-infected patients receiving highly active antiretroviral therapy." Lancet **363**(9427): 2116-2121.

Digel, M., K. L. Sampaio, G. Jahn and C. Sinzger (2006). "Evidence for direct transfer of cytoplasmic material from infected to uninfected cells during cell-associated spread of human cytomegalovirus." J Clin Virol **37**(1): 10-20.

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**(23): 14660-14667.

Dogan, R. I., L. Getoor, W. J. Wilbur and S. M. Mount (2007). "SplicePort--an interactive splice-site analysis tool." Nucleic Acids Res **35**(Web Server issue): W285-291.

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. G. Wilkinson and A. J. Davison (2004). "Genetic content of wild-type human cytomegalovirus." Journal of General Virology **85**(5): 1301-1312.

Dollard, S. C., S. D. Grosse and D. S. Ross (2007). "New estimates of the prevalence of neurological and sensory sequelae and mortality associated with congenital cytomegalovirus infection." Rev Med Virol **17**(5): 355-363.

Dunn, C., N. J. Chalupny, C. L. Sutherland, S. Dosch, P. V. Sivakumar, D. C. Johnson and D. Cosman (2003). "Human cytomegalovirus glycoprotein UL16 causes intracellular sequestration

of NKG2D ligands, protecting against natural killer cell cytotoxicity." J Exp Med **197**(11): 1427-1439.

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**(24): 14223-14228.

Dworsky, M., M. Yow, S. Stagno, R. F. Pass and C. Alford (1983). "Cytomegalovirus infection of breast milk and transmission in infancy." Pediatrics **72**(3): 295-299.

Elek, S. D. and H. Stern (1974). "Development of a vaccine against mental retardation caused by cytomegalovirus infection in utero." Lancet **1**(7845): 1-5.

Emery, V. C. (2001). "Investigation of CMV disease in immunocompromised patients." J Clin Pathol **54**(2): 84-88.

Emery, V. C. and P. D. Griffiths (2000). "Prediction of cytomegalovirus load and resistance patterns after antiviral chemotherapy." Proc Natl Acad Sci U S A **97**(14): 8039-8044.

Farrar, G. H. and J. D. Oram (1984). "Characterization of the human cytomegalovirus envelope glycoproteins." J Gen Virol **65** (Pt 11)(11): 1991-2001.

Feire, A. L., H. Koss and T. Compton (2004). "Cellular integrins function as entry receptors for human cytomegalovirus via a highly conserved disintegrin-like domain." Proc Natl Acad Sci U S A **101**(43): 15470-15475.

Feire, A. L., R. M. Roy, K. Manley and T. Compton (2010). "The glycoprotein B disintegrin-like domain binds beta 1 integrin to mediate cytomegalovirus entry." J Virol **84**(19): 10026-10037.

Fiala, M., R. W. Honess, D. C. Heiner, J. W. Heine, Jr., J. Murnane, R. Wallace and L. B. Guze (1976). "Cytomegalovirus proteins. I. Polypeptides of virions and dense bodies." J Virol **19**(1): 243-254.

Fielding, C. A., R. Aicheler, R. J. Stanton, E. C. Wang, S. Han, S. Seirafian, J. Davies, B. P. McSharry, M. P. Weekes, P. R. Antrobus, V. Prod'homme, F. P. Blanchet, D. Sugrue, S. Cuff, D. Roberts, A. J. Davison, P. J. Lehner, G. W. Wilkinson and P. Tomasec (2014). "Two novel human cytomegalovirus NK cell evasion functions target MICA for lysosomal degradation." PLoS Pathog **10**(5): e1004058.

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**(7): 5661-5668.

Fishman, J. A. (2007). "Infection in solid-organ transplant recipients." N Engl J Med **357**(25): 2601-2614.

Fleisher, G. R., S. E. Starr, H. M. Friedman and S. A. Plotkin (1982). "Vaccination of pediatric nurses with live attenuated cytomegalovirus." Am J Dis Child **136**(4): 294-296.

Fowler, K. B., S. Stagno, R. F. Pass, W. J. Britt, T. J. Boll and C. A. Alford (1992). "The outcome of congenital cytomegalovirus infection in relation to maternal antibody status." New England Journal of Medicine **326**(10): 663-667.

Freed, D. C., Q. Tang, A. M. Tang, F. S. Li, X. He, Z. Huang, W. X. Meng, L. Xia, A. C. Finnefrock, E. Durr, A. S. Espeseth, D. R. Casimiro, N. Y. Zhang, J. W. Shiver, D. Wang, Z. Q. An and T. M. Fu (2013). "Pentameric complex of viral glycoprotein H is the primary target for potent neutralization by a human cytomegalovirus vaccine." Proceedings of the National Academy of Sciences of the United States of America **110**(51): E4997-E5005.

Furlong, D., H. Swift and B. Roizman (1972). "Arrangement of herpesvirus deoxyribonucleic acid in the core." J Virol **10**(5): 1071-1074.

Gandhi, M. K. and R. Khanna (2004). "Human cytomegalovirus: clinical aspects, immune regulation, and emerging treatments." The Lancet infectious diseases **4**(12): 725-738.

Gatherer, D., S. Seirafian, C. Cunningham, M. Holton, D. J. Dargan, K. Baluchova, R. D. Hector, J. Galbraith, P. Herzyk, G. W. Wilkinson and A. J. Davison (2011). "High-resolution human cytomegalovirus transcriptome." Proc Natl Acad Sci U S A **108**(49): 19755-19760.

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**(9): 7048-7060.

Genini, E., E. Percivalle, A. Sarasini, M. G. Revello, F. Baldanti and G. Gerna (2011). "Serum antibody response to the gH/gL/pUL128-131 five-protein complex of human cytomegalovirus (HCMV) in primary and reactivated HCMV infections." J Clin Virol **52**(2): 113-118.

Gerna, G., M. Furione, F. Baldanti and A. Sarasini (1994). "Comparative quantitation of human cytomegalovirus DNA in blood leukocytes and plasma of transplant and AIDS patients." J Clin Microbiol **32**(11): 2709-2717.

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**(12): 5629-5638.

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**(Pt 2): 275-284.

Gerna, G., E. Percivalle, A. Sarasini, F. Baldanti, G. Campanini and M. G. Revello (2003). "Rescue of human cytomegalovirus strain AD169 tropism for both leukocytes and human endothelial cells." Journal of general virology **84**(6): 1431-1436.

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." Journal of General Virology **83**: 1993-2000.

Gerna, G., A. Sarasini, M. Patrone, E. Percivalle, L. Fiorina, G. Campanini, A. Gallina, F. Baldanti and M. G. Revello (2008). "Human cytomegalovirus serum neutralizing antibodies

block virus infection of endothelial/epithelial cells, but not fibroblasts, early during primary infection." J Gen Virol **89**(Pt 4): 853-865.

Gerna, G., D. Zipeto, E. Percivalle, M. Parea, M. G. Revello, R. Maccario, G. Peri and G. Milanesi (1992). "Human cytomegalovirus infection of the major leukocyte subpopulations and evidence for initial viral replication in polymorphonuclear leukocytes from viremic patients." J Infect Dis **166**(6): 1236-1244.

Gibson, W. (1996). "Structure and assembly of the virion." Intervirology **39**(5-6): 389-400.

Gibson, W., M. K. Baxter and K. S. Clopper (1996). "Cytomegalovirus "missing" capsid protein identified as heat-aggregable product of human cytomegalovirus UL46." Journal of Virology **70**(11): 7454-7461.

Gibson, W., E. Bogner and M. Reddehase (2013). "Morphogenesis of the cytomegalovirus virion and subviral particles." Cytomegaloviruses: From Molecular Pathogenesis to Intervention **1**: 230.

Gibson, W., K. S. Clopper, W. J. Britt and M. K. Baxter (1996). "Human cytomegalovirus (HCMV) smallest capsid protein identified as product of short open reading frame located between HCMV UL48 and UL49." Journal of Virology **70**(8): 5680-5683.

Gnanandarajah, J. S., P. A. Gillis, N. Hernandez-Alvarado, L. Higgins, T. W. Markowski, H. Sung, S. Lumley and M. R. Schleiss (2014). "Identification by mass spectrometry and immune response analysis of guinea pig cytomegalovirus (GPCMV) pentameric complex proteins GP129, 131 and 133." Viruses **6**(2): 727-751.

Goelz, R., C. Meisner, A. Bevot, K. Hamprecht, I. Kraegeloh-Mann and C. F. Poets (2013). "Long-term cognitive and neurological outcome of preterm infants with postnatally acquired CMV infection through breast milk." Arch Dis Child Fetal Neonatal Ed **98**(5): F430-433.

Gompels, U. A., J. Nicholas, G. Lawrence, M. Jones, B. J. Thomson, M. E. Martin, S. Efsthathiou, M. Craxton and H. A. Macaulay (1995). "The DNA sequence of human herpesvirus-6: structure, coding content, and genome evolution." Virology **209**(1): 29-51.

Goodpasture, E. W. and F. B. Talbot (1921). "Concerning the nature of "protozoan-like" cells in certain lesions of infancy." American Journal of Diseases of Children **21**(5): 415-425.

Goodrum, F., K. Caviness and P. Zagallo (2012). "Human cytomegalovirus persistence." Cellular Microbiology **14**(5): 644-655.

Gorzer, I., C. Guelly, S. Trajanoski and E. Puchhammer-Stockl (2010). "Deep sequencing reveals highly complex dynamics of human cytomegalovirus genotypes in transplant patients over time." J Virol **84**(14): 7195-7203.

Graham, F., J. Smiley, W. Russell and R. Nairn (1977). "Characteristics of a human cell line transformed by DNA from human adenovirus type 5." Journal of General Virology **36**(1): 59-72.

Gredmark-Russ, S. and C. Soderberg-Naucler (2012). "Dendritic cell biology in human cytomegalovirus infection and the clinical consequences for host immunity and pathology." Virulence **3**(7): 621-634.

Grefte, A., M. van der Giessen and W. van Son (1993). "Circulating cytomegalovirus (CMV)-infected endothelial cells in patients with an active CMV infection." Journal of Infectious Diseases **167**(2): 270-277.

Gretch, D. R., R. C. Gehrz and M. F. Stinski (1988). "Characterization of a human cytomegalovirus glycoprotein complex (gcl)." J. Gen. Virol **69**: 1205-1215.

Gretch, D. R., B. Kari, L. Rasmussen, R. C. Gehrz and M. F. Stinski (1988). "Identification and characterization of three distinct families of glycoprotein complexes in the envelopes of human cytomegalovirus." J Virol **62**(3): 875-881.

Griffiths, P. D., A. Stanton, E. McCarrell, C. Smith, M. Osman, M. Harber, A. Davenport, G. Jones, D. C. Wheeler, J. O'Beirne, D. Thorburn, D. Patch, C. E. Atkinson, S. Pichon, P. Sweny, M. Lanzman, E. Woodford, E. Rothwell, N. Old, R. Kinyanjui, T. Haque, S. Atabani, S. Luck, S. Prideaux, R. S. Milne, V. C. Emery and A. K. Burroughs (2011). "Cytomegalovirus glycoprotein-B vaccine with MF59 adjuvant in transplant recipients: a phase 2 randomised placebo-controlled trial." Lancet **377**(9773): 1256-1263.

Griffiths, P. P. (2006). "CMV as a cofactor enhancing progression of AIDS." Journal of Clinical Virology **35**(4): 489-492.

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 CXC chemokines and transmit virus by direct neutrophil-endothelial cell contact and during neutrophil transendothelial migration." Journal of Infectious Diseases **177**(6): 1465-1474.

Hahn, G., R. Jores and E. S. Mocarski (1998). "Cytomegalovirus remains latent in a common precursor of dendritic and myeloid cells." Proc Natl Acad Sci U S A **95**(7): 3937-3942.

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." Journal of Virology **76**(18): 9551-9555.

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**(18): 10023-10033.

Hahn, G., D. Rose, M. Wagner, S. Rhiel and M. A. McVoy (2003). "Cloning of the genomes of human cytomegalovirus strains Toledo, TownevarRIT3, and Towne< sub> long</sub> as BACs and site-directed mutagenesis using a PCR-based technique." Virology **307**(1): 164-177.

Hanley, P. J. and C. M. Bollard (2014). "Controlling Cytomegalovirus: Helping the Immune System Take the Lead." Viruses-Basel **6**(6): 2242-2258.

Haspot, F., A. Lavault, C. Sinzger, K. Laib Sampaio, Y. D. Stierhof, P. Pilet, C. Bressolette-Bodin and F. Halary (2012). "Human cytomegalovirus entry into dendritic cells occurs via a macropinocytosis-like pathway in a pH-independent and cholesterol-dependent manner." PLoS One **7**(4): e34795.

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**(17): 7720-7729.

Holzerlandt, R., C. Orengo, P. Kellam and M. M. Alba (2002). "Identification of new herpesvirus gene homologs in the human genome." Genome Res **12**(11): 1739-1748.

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

Huber, M. T. and T. Compton (1997). "Characterization of a novel third member of the human cytomegalovirus glycoprotein H-glycoprotein L complex." J Virol **71**(7): 5391-5398.

Huber, M. T. and T. Compton (1998). "The human cytomegalovirus UL74 gene encodes the third component of the glycoprotein H-glycoprotein L-containing envelope complex." J Virol **72**(10): 8191-8197.

Huber, M. T. and T. Compton (1999). "Intracellular formation and processing of the heterotrimeric gH-gL-gO (gCIII) glycoprotein envelope complex of human cytomegalovirus." J Virol **73**(5): 3886-3892.

Hutchinson, N. I. and M. J. Tocci (1986). "Characterization of a major early gene from the human cytomegalovirus long inverted repeat; predicted amino acid sequence of a 30-kDa protein encoded by the 1.2-kb mRNA." Virology **155**(1): 172-182.

Ibanez, C. E., R. Schrier, P. Ghazal, C. Wiley and J. A. Nelson (1991). "Human cytomegalovirus productively infects primary differentiated macrophages." J Virol **65**(12): 6581-6588.

Ioudinkova, E., M. C. Arcangeletti, A. Rynditch, F. De Conto, F. Motta, S. Covan, F. Pinardi, S. V. Razin and C. Chezzi (2006). "Control of human cytomegalovirus gene expression by differential histone modifications during lytic and latent infection of a monocytic cell line." Gene **384**: 120-128.

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**(1): 118-133.

Irmiere, A. and W. Gibson (1985). "Isolation of human cytomegalovirus intranuclear capsids, characterization of their protein constituents, and demonstration that the B-capsid assembly protein is also abundant in noninfectious enveloped particles." Journal of virology **56**(1): 277-283.

Isaacson, M., L. Juckem and T. Compton (2008). Virus entry and innate immune activation. Human Cytomegalovirus, Springer: 85-100.

Isaacson, M. K. and T. Compton (2009). "Human cytomegalovirus glycoprotein B is required for virus entry and cell-to-cell spread but not for virion attachment, assembly, or egress." J Virol **83**(8): 3891-3903.

Isaacson, M. K., A. L. Feire and T. Compton (2007). "Epidermal growth factor receptor is not required for human cytomegalovirus entry or signaling." J Virol **81**(12): 6241-6247.

Iversen, A. C., P. S. Norris, C. F. Ware and C. A. Benedict (2005). "Human NK cells inhibit cytomegalovirus replication through a noncytolytic mechanism involving lymphotoxin-dependent induction of IFN-beta." *J Immunol* **175**(11): 7568-7574.

James Waldman, W., J. M. Sneddon, R. E. Stephens and W. H. Roberts (1989). "Enhanced endothelial cytopathogenicity induced by a cytomegalovirus strain propagated in endothelial cells." *Journal of medical virology* **28**(4): 223-230.

Jarvis, M. A. and J. A. Nelson (2002). "Human cytomegalovirus persistence and latency in endothelial cells and macrophages." *Current opinion in microbiology* **5**(4): 403-407.

Jarvis, M. A. and J. A. Nelson (2007). "Human cytomegalovirus tropism for endothelial cells: not all endothelial cells are created equal." *Journal of virology* **81**(5): 2095-2101.

Jesionek, A. K., B. (1904). "Über einen Befund von Protozoenartigen gebilden in den Organen eines hereditärluetischen Fetus." *Munch Med Wochenschr* **51**: 1905-1907.

Ji, Y. H., Z. Rong Sun, Q. Ruan, J. J. Guo, R. He, Y. Qi, Y. P. Ma, Z. Q. Mao and Y. J. Huang (2006). "Polymorphisms of human cytomegalovirus UL148A, UL148B, UL148C, UL148D genes in clinical strains." *J Clin Virol* **37**(4): 252-257.

Jiang, X. J., B. Adler, K. L. Sampaio, M. Digel, G. Jahn, N. Ettischer, Y. D. Stierhof, L. Scrivano, U. Koszinowski, M. Mach and C. Sinzger (2008). "UL74 of human cytomegalovirus contributes to virus release by promoting secondary envelopment of virions." *J Virol* **82**(6): 2802-2812.

Jiang, X. J., K. L. Sampaio, N. Ettischer, Y. D. Stierhof, G. Jahn, B. Kropff, M. Mach and C. Sinzger (2011). "UL74 of human cytomegalovirus reduces the inhibitory effect of gH-specific and gB-specific antibodies." *Arch Virol* **156**(12): 2145-2155.

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**(4): 2541-2546.

Jones, T. R. and L. Sun (1997). "Human cytomegalovirus US2 destabilizes major histocompatibility complex class I heavy chains." *J Virol* **71**(4): 2970-2979.

Jones, T. R., E. J. Wiertz, L. Sun, K. N. Fish, J. A. Nelson and H. L. Ploegh (1996). "Human cytomegalovirus US3 impairs transport and maturation of major histocompatibility complex class I heavy chains." *Proc Natl Acad Sci U S A* **93**(21): 11327-11333.

Just, M., A. Buerger-Wolff, G. Emoedi and R. Hernandez (1975). "Immunisation trials with live attenuated cytomegalovirus TOWNE 125." *Infection* **3**(2): 111-114.

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**(16): 7628-7635.

Kalejta, R. F. (2008). "Tegument proteins of human cytomegalovirus." *Microbiol Mol Biol Rev* **72**(2): 249-265, table of contents.

- Kalejta, R. F. (2013). "Pre-Immediate Early Tegument Protein Functions." Cytomegaloviruses: From Molecular Pathogenesis to Intervention **1**: 141.
- Kari, B. and R. Gehrz (1993). "Structure, composition and heparin binding properties of a human cytomegalovirus glycoprotein complex designated gC-II." The Journal of general virology **74**: 255-264.
- Kari, B., R. Goertz and R. Gehrz (1990). "Characterization of cytomegalovirus glycoproteins in a family of complexes designated gC-II with murine monoclonal antibodies." Arch Virol **112**(1-2): 55-65.
- Kari, B., W. Li, J. Cooper, R. Goertz and B. Radeke (1994). "The human cytomegalovirus UL100 gene encodes the gC-II glycoproteins recognized by group 2 monoclonal antibodies." The Journal of general virology **75**: 3081.
- Kari, B., Y. N. Liu, R. Goertz, N. Lussenhop, M. F. Stinski and R. Gehrz (1990). "Structure and composition of a family of human cytomegalovirus glycoprotein complexes designated gC-I (gB)." J Gen Virol **71** (Pt 11)(11): 2673-2680.
- Kari, B., R. Radeke and R. Gehrz (1992). "Processing of human cytomegalovirus envelope glycoproteins in and egress of cytomegalovirus from human astrocytoma cells." The Journal of general virology **73**: 253.
- Kattenhorn, L. M., R. Mills, M. Wagner, A. Lomsadze, V. Makeev, M. Borodovsky, H. L. Ploegh and B. M. Kessler (2004). "Identification of proteins associated with murine cytomegalovirus virions." J Virol **78**(20): 11187-11197.
- Kaye, J. F., U. A. Gompels and A. C. Minson (1992). "Glycoprotein H of human cytomegalovirus (HCMV) forms a stable complex with the HCMV UL115 gene product." J Gen Virol **73** (Pt 10): 2693-2698.
- Keay, S. and B. Baldwin (1991). "Anti-idiotypic antibodies that mimic gp86 of human cytomegalovirus inhibit viral fusion but not attachment." Journal of virology **65**(9): 5124-5128.
- Kelley, L. A. and M. J. Sternberg (2009). "Protein structure prediction on the Web: a case study using the Phyre server." Nat Protoc **4**(3): 363-371.
- Kemble, G. W. and E. S. Mocarski (1989). "A host cell protein binds to a highly conserved sequence element (pac-2) within the cytomegalovirus a sequence." J Virol **63**(11): 4715-4728.
- Kenneson, A. and M. J. Cannon (2007). "Review and meta-analysis of the epidemiology of congenital cytomegalovirus (CMV) infection." Rev Med Virol **17**(4): 253-276.
- Khaiboullina, S. F., J. P. Maciejewski, K. Crapnell, P. A. Spallone, A. Dean Stock, G. S. Pari, E. D. Zanjani and S. St Jeor (2004). "Human cytomegalovirus persists in myeloid progenitors and is passed to the myeloid progeny in a latent form." British Journal of Haematology **126**(3): 410-417.
- Kilpatrick, B. A. and E. S. Huang (1977). "Human cytomegalovirus genome: partial denaturation map and organization of genome sequences." J Virol **24**(1): 261-276.

- Kinzler, E. R. and T. Compton (2005). "Characterization of human cytomegalovirus glycoprotein-induced cell-cell fusion." *J Virol* **79**(12): 7827-7837.
- Kinzler, E. R., R. N. Theiler and T. Compton (2002). "Expression and reconstitution of the gH/gL/gO complex of human cytomegalovirus." *J Clin Virol* **25 Suppl 2**: S87-95.
- Krause, P. R., S. R. Bialek, S. B. Boppana, P. D. Griffiths, C. A. Laughlin, P. Ljungman, E. S. Mocarski, R. F. Pass, J. S. Read, M. R. Schleiss and S. A. Plotkin (2013). "Priorities for CMV vaccine development." *Vaccine* **32**(1): 4-10.
- Kropff, B., C. Burkhardt, J. Schott, J. Nentwich, T. Fisch, W. Britt and M. Mach (2012). "Glycoprotein N of human cytomegalovirus protects the virus from neutralizing antibodies." *PLoS Pathog* **8**(10): e1002999.
- Krzyzaniak, M., M. Mach and W. J. Britt (2007). "The cytoplasmic tail of glycoprotein M (gpUL100) expresses trafficking signals required for human cytomegalovirus assembly and replication." *Journal of virology* **81**(19): 10316-10328.
- Kuchta, A. L., H. Parikh, Y. Zhu, G. E. Kellogg, D. S. Parris and M. A. McVoy (2012). "Structural modelling and mutagenesis of human cytomegalovirus alkaline nuclease UL98." *J Gen Virol* **93**(Pt 1): 130-138.
- Kulesza, C. A. and T. Shenk (2004). "Human cytomegalovirus 5-kilobase immediate-early RNA is a stable intron." *Journal of Virology* **78**(23): 13182-13189.
- LaPierre, L. A. and B. J. Biegalke (2001). "Identification of a novel transcriptional repressor encoded by human cytomegalovirus." *J Virol* **75**(13): 6062-6069.
- Lehner, P. J., J. T. Karttunen, G. W. Wilkinson and P. Cresswell (1997). "The human cytomegalovirus US6 glycoprotein inhibits transporter associated with antigen processing-dependent peptide translocation." *Proc Natl Acad Sci U S A* **94**(13): 6904-6909.
- Lehner, R., H. Meyer and M. Mach (1989). "Identification and characterization of a human cytomegalovirus gene coding for a membrane protein that is conserved among human herpesviruses." *Journal of virology* **63**(9): 3792-3800.
- Lehner, R., T. Stamminger and M. Mach (1991). "Comparative sequence analysis of human cytomegalovirus strains." *J Clin Microbiol* **29**(11): 2494-2502.
- Li, H. and R. Durbin (2009). "Fast and accurate short read alignment with Burrows-Wheeler transform." *Bioinformatics* **25**(14): 1754-1760.
- Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin and S. Genome Project Data Processing (2009). "The Sequence Alignment/Map format and SAMtools." *Bioinformatics* **25**(16): 2078-2079.
- Li, L., J. A. Nelson and W. J. Britt (1997). "Glycoprotein H-related complexes of human cytomegalovirus: identification of a third protein in the gCIII complex." *J Virol* **71**(4): 3090-3097.

- Lilja, A. E. and P. W. Mason (2012). "The next generation recombinant human cytomegalovirus vaccine candidates-beyond gB." Vaccine **30**(49): 6980-6990.
- Lilley, B. N., H. L. Ploegh and R. S. Tirabassi (2001). "Human cytomegalovirus open reading frame TRL11/IRL11 encodes an immunoglobulin G Fc-binding protein." J Virol **75**(22): 11218-11221.
- Lipscheutz, B. (1921). "Untersuchungen ueber die aetiologie der krankeiten der Herpes genitalis." Arch Dermatol Syph **136**: 428-482.
- Loewendorf, A. and C. A. Benedict (2010). "Modulation of host innate and adaptive immune defenses by cytomegalovirus: timing is everything." J Intern Med **267**(5): 483-501.
- Lopper, M. and T. Compton (2004). "Coiled-coil domains in glycoproteins B and H are involved in human cytomegalovirus membrane fusion." J Virol **78**(15): 8333-8341.
- Lurain, N. S. and S. Chou (2010). "Antiviral drug resistance of human cytomegalovirus." Clin Microbiol Rev **23**(4): 689-712.
- 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. W. Chou (2006). "Analysis of the human cytomegalovirus genomic region from UL146 through UL147A reveals sequence hypervariability, genotypic stability, and overlapping transcripts." Virology Journal **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**(12): 10040-10050.
- Luttichau, H. R. (2010). "The cytomegalovirus UL146 gene product vCXCL1 targets both CXCR1 and CXCR2 as an agonist." J Biol Chem **285**(12): 9137-9146.
- MacCormac, L. P. and J. E. Grundy (1999). "Two clinical isolates and the toledo strain of cytomegalovirus contain endothelial cell tropic variants that are not present in the AD169, Towne, or Davis strains." Journal of Medical Virology **57**(3): 298-307.
- 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**(24): 11881-11892.
- 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**(4): 2160-2170.
- Mach, M., K. Osinski, B. Kropff, U. Schloetzer-Schrehardt, M. Krzyzaniak and W. Britt (2007). "The carboxy-terminal domain of glycoprotein N of human cytomegalovirus is required for virion morphogenesis." J Virol **81**(10): 5212-5224.
- Mach, M., U. Utz and B. Fleckenstein (1986). "Mapping of the major glycoprotein gene of human cytomegalovirus." J Gen Virol **67** (Pt 7): 1461-1467.
- Magri, G., A. Muntasell, N. Romo, A. Saez-Borderias, D. Pende, D. E. Geraghty, H. Hengel, A. Angulo, A. Moretta and M. Lopez-Botet (2011). "NKP46 and DNAM-1 NK-cell receptors drive

the response to human cytomegalovirus-infected myeloid dendritic cells overcoming viral immune evasion strategies." Blood **117**(3): 848-856.

Malm, G. and M.-L. Engman (2007). Congenital cytomegalovirus infections. Seminars in Fetal and Neonatal Medicine, Elsevier.

Malone, C. L., D. H. Vesole and M. F. Stinski (1990). "Transactivation of a human cytomegalovirus early promoter by gene products from the immediate-early gene IE2 and augmentation by IE1: mutational analysis of the viral proteins." J Virol **64**(4): 1498-1506.

Malouli, D., S. G. Hansen, E. S. Nakayasu, E. E. Marshall, C. M. Hughes, A. B. Ventura, R. M. Gilbride, M. S. Lewis, G. Xu, C. Kreklywich, N. Whizin, M. Fischer, A. W. Legasse, K. Viswanathan, D. Siess, D. G. Camp, 2nd, M. K. Axthelm, C. Kahl, V. R. DeFilippis, R. D. Smith, D. N. Streblow, L. J. Picker and K. Fruh (2014). "Cytomegalovirus pp65 limits dissemination but is dispensable for persistence." J Clin Invest **124**(5): 1928-1944.

Malouli, D., E. S. Nakayasu, K. Viswanathan, D. G. Camp, 2nd, W. L. Chang, P. A. Barry, R. D. Smith and K. Fruh (2012). "Reevaluation of the coding potential and proteomic analysis of the BAC-derived rhesus cytomegalovirus strain 68-1." J Virol **86**(17): 8959-8973.

Matrosovich, M., T. Matrosovich, W. Garten and H. D. Klenk (2006). "New low-viscosity overlay medium for viral plaque assays." Virology **3**: 63.

Matthews, T. and R. Boehme (1988). "Antiviral activity and mechanism of action of ganciclovir." Rev Infect Dis **10 Suppl 3**: S490-494.

Mattick, C., D. Dewin, S. Polley, E. Sevilla-Reyes, S. Pignatelli, W. Rawlinson, G. Wilkinson, P. Dal Monte and U. A. Gompels (2004). "Linkage of human cytomegalovirus glycoprotein gO variant groups identified from worldwide clinical isolates with gN genotypes, implications for disease associations and evidence for N-terminal sites of positive selection." Virology **318**(2): 582-597.

McGeoch, D. J., S. Cook, A. Dolan, F. E. Jamieson and E. A. Telford (1995). "Molecular phylogeny and evolutionary timescale for the family of mammalian herpesviruses." Journal of molecular biology **247**(3): 443-458.

McGeoch, D. J., A. Dolan and A. C. Ralph (2000). "Toward a comprehensive phylogeny for mammalian and avian herpesviruses." Journal of virology **74**(22): 10401-10406.

McGeoch, D. J. and D. Gatherer (2005). "Integrating reptilian herpesviruses into the family Herpesviridae." Journal of virology **79**(2): 725-731.

McGeoch, D. J., F. J. Rixon and A. J. Davison (2006). "Topics in herpesvirus genomics and evolution." Virus Res **117**(1): 90-104.

McMahon, T. P. and D. G. Anders (2002). "Interactions between human cytomegalovirus helicase-primase proteins." Virus Research **86**(1-2): 39-52.

McSharry, B., C. Jones, J. Skinner, D. Kipling and G. Wilkinson (2001). "Human telomerase reverse transcriptase-immortalized MRC-5 and HCA2 human fibroblasts are fully permissive for human cytomegalovirus." Journal of General Virology **82**(4): 855-863.

McSharry, B. P., S. Avdic and B. Slobedman (2012). "Human Cytomegalovirus Encoded Homologs of Cytokines, Chemokines and their Receptors: Roles in Immunomodulation." Viruses **4**(11): 2448-2470.

McSharry, B. P., H.-G. Burgert, D. P. Owen, R. J. Stanton, V. Prod'homme, M. Sester, K. Koebernick, V. Groh, T. Spies and S. Cox (2008). "Adenovirus E3/19K promotes evasion of NK cell recognition by intracellular sequestration of the NKG2D ligands major histocompatibility complex class I chain-related proteins A and B." Journal of virology **82**(9): 4585-4594.

McSharry, B. P., P. Tomasec, M. L. Neale and G. W. Wilkinson (2003). "The most abundantly transcribed human cytomegalovirus gene (β 2. 7) is non-essential for growth in vitro." Journal of general virology **84**(9): 2511-2516.

Mcvoy, M. A. and S. P. Adler (1994). "Human Cytomegalovirus DNA Replicates after Early Circularization by Concatemer Formation, and Inversion Occurs within the Concatemer." Journal of Virology **68**(2): 1040-1051.

Mendelsohn, C. L., E. Wimmer and V. R. Racaniello (1989). "Cellular receptor for poliovirus: molecular cloning, nucleotide sequence, and expression of a new member of the immunoglobulin superfamily." Cell **56**(5): 855-865.

Mendelson, M., S. Monard, P. Sissons and J. Sinclair (1996). "Detection of endogenous human cytomegalovirus in CD34+ bone marrow progenitors." J Gen Virol **77 (Pt 12)**(12): 3099-3102.

Messerle, M., I. Crnkovic, W. Hammerschmidt, H. Ziegler and U. H. Koszinowski (1997). "Cloning and mutagenesis of a herpesvirus genome as an infectious bacterial artificial chromosome." Proceedings of the National Academy of Sciences of the United States of America **94**(26): 14759-14763.

Meyer-König, U., M. Haberland, D. von Laer, O. Haller and F. T. Hufert (1998). "Intragenic variability of human cytomegalovirus glycoprotein B in clinical strains." J Infect Dis **177**(5): 1162-1169.

Meyer - König, U., C. Vogelberg, A. Bongarts, D. Kampa, R. Delbrück, 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." Journal of medical virology **55**(1): 75-81.

Miceli, M. V., D. A. Newsome, L. C. Novak and R. W. Beuerman (1989). "Cytomegalovirus replication in cultured human retinal pigment epithelial cells." Curr Eye Res **8**(8): 835-839.

Michelson, S. (2004). "Consequences of human cytomegalovirus mimicry." Human immunology **65**(5): 465-475.

Milne, I., M. Bayer, L. Cardle, P. Shaw, G. Stephen, F. Wright and D. Marshall (2010). "Tablet—next generation sequence assembly visualization." Bioinformatics **26**(3): 401-402.

Milne, R., D. A. Paterson and J. C. Booth (1998). "Human cytomegalovirus glycoprotein H/glycoprotein L complex modulates fusion-from-without." Journal of general virology **79**(4): 855-865.

- Minton, E. J., C. Tysoe, J. H. Sinclair and J. G. Sissons (1994). "Human cytomegalovirus infection of the monocyte/macrophage lineage in bone marrow." J Virol **68**(6): 4017-4021.
- Mitchell, D. P., J. P. Savaryn, N. J. Moorman, T. Shenk and S. S. Terhune (2009). "Human cytomegalovirus UL28 and UL29 open reading frames encode a spliced mRNA and stimulate accumulation of immediate-early RNAs." J Virol **83**(19): 10187-10197.
- Mocarski, E., T. Shenk, R. Pass, D. Cytomegaloviruses and P. H. Knipe (2007). *Fields virology*, Lippincott, Philadelphia.
- Mocarski, E. S., Jr. (2004). "Immune escape and exploitation strategies of cytomegaloviruses: impact on and imitation of the major histocompatibility system." Cell Microbiol **6**(8): 707-717.
- Mocarski Jr, E. S. (2002). "Immunomodulation by cytomegaloviruses: manipulative strategies beyond evasion." Trends in microbiology **10**(7): 332-339.
- Murphy, E., I. Rigoutsos, T. Shibuya and T. E. Shenk (2003). "Reevaluation of human cytomegalovirus coding potential." Proc Natl Acad Sci U S A **100**(23): 13585-13590.
- 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 (2003). "Coding potential of laboratory and clinical strains of human cytomegalovirus." Proc Natl Acad Sci U S A **100**(25): 14976-14981.
- Myerson, D., R. C. Hackman, J. A. Nelson, D. C. Ward and J. K. McDougall (1984). "Widespread presence of histologically occult cytomegalovirus." Hum Pathol **15**(5): 430-439.
- Nicholson, I. P., J. S. Sutherland, T. N. Chaudry, E. L. Blewett, P. A. Barry, M. J. Nicholl and C. M. Preston (2009). "Properties of virion transactivator proteins encoded by primate cytomegaloviruses." Virology **6**: 65.
- Nigro, G., S. P. Adler, R. La Torre and A. M. Best (2005). "Passive immunization during pregnancy for congenital cytomegalovirus infection." New England Journal of Medicine **353**(13): 1350-1362.
- Novotny, J., I. Rigoutsos, D. Coleman and T. Shenk (2001). "< i> In silico</i> structural and functional analysis of the human cytomegalovirus (HHV5) genome." Journal of molecular biology **310**(5): 1151-1166.
- Nowak, B., C. Sullivan, P. Sarnow, R. Thomas, F. Bricout, J. C. Nicolas, B. Fleckenstein and A. J. Levine (1984). "Characterization of monoclonal antibodies and polyclonal immune sera directed against human cytomegalovirus virion proteins." Virology **132**(2): 325-338.
- O'Connor, C. M. and E. A. Murphy (2012). "A myeloid progenitor cell line capable of supporting human cytomegalovirus latency and reactivation, resulting in infectious progeny." J Virol **86**(18): 9854-9865.
- Odeberg, J., B. Plachter, L. Branden and C. Soderberg-Naucler (2003). "Human cytomegalovirus protein pp65 mediates accumulation of HLA-DR in lysosomes and destruction of the HLA-DR alpha-chain." Blood **101**(12): 4870-4877.

Ogita, H. and Y. Takai (2006). "Nectins and nectin-like molecules: roles in cell adhesion, polarization, movement, and proliferation." IUBMB Life **58**(5-6): 334-343.

Ong, S. E., B. Blagoev, I. Kratchmarova, D. B. Kristensen, H. Steen, A. Pandey and M. Mann (2002). "Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics." Molecular & Cellular Proteomics **1**(5): 376-386.

Oram, J. D., R. G. Downing, A. Akrigg, A. A. Dollery, C. J. Duggleby, G. W. G. Wilkinson and P. J. Greenaway (1982). "Use of Recombinant Plasmids to Investigate the Structure of the Human Cytomegalovirus Genome." Journal of General Virology **59**(Mar): 111-129.

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**(5): 2575-2582.

Pass, R. F., A. M. Duliege, S. Boppana, R. Sekulovich, S. Percell, W. Britt and R. L. Burke (1999). "A subunit cytomegalovirus vaccine based on recombinant envelope glycoprotein B and a new adjuvant." J Infect Dis **180**(4): 970-975.

Pass, R. F., C. Zhang, A. Evans, T. Simpson, W. Andrews, M. L. Huang, L. Corey, J. Hill, E. Davis, C. Flanigan and G. Cloud (2009). "Vaccine prevention of maternal cytomegalovirus infection." N Engl J Med **360**(12): 1191-1199.

Paterson, D. A., A. P. Dyer, R. S. Milne, E. Sevilla-Reyes and U. A. Gompels (2002). "A role for human cytomegalovirus glycoprotein O (gO) in cell fusion and a new hypervariable locus." Virology **293**(2): 281-294.

Patrone, M., M. Secchi, E. Bonaparte, G. Milanesi and A. Gallina (2007). "Cytomegalovirus UL131-128 products promote gB conformational transition and gB-gH interaction during entry into endothelial cells." Journal of Virology **81**(20): 11479-11488.

Patrone, M., M. Secchi, L. Fiorina, M. Ierardi, G. Milanesi and A. Gallina (2005). "Human cytomegalovirus UL130 protein promotes endothelial cell infection through a producer cell modification of the virion." Journal of Virology **79**(13): 8361-8373.

Pawelec, G. and E. Derhovanessian (2011). "Role of CMV in immune senescence." Virus Research **157**(2): 175-179.

Pellet, P. E. R., B. (2013). Herpesviridae. Field's Virology. D. M. H. Knipe, P.M.; Cohen, J.I.; Griffin, D.E.; Lamb, R.A.; Martin, M.A.; Racaniello, V.R.; Roizman, B., LIPPINCOTT WILLIAMS & WILKINS, a WOLTERS KLUWER business. **II**: 2770-3809.

Percivalle, E., M. G. Revello, L. Vago, F. Morini and G. Gerna (1993). "Circulating endothelial giant cells permissive for human cytomegalovirus (HCMV) are detected in disseminated HCMV infections with organ involvement." J Clin Invest **92**(2): 663-670.

Pereira, L., M. Hoffman, D. Gallo and N. Cremer (1982). "Monoclonal antibodies to human cytomegalovirus: three surface membrane proteins with unique immunological and electrophoretic properties specify cross-reactive determinants." Infect Immun **36**(3): 924-932.

Pfeffer, S., A. Sewer, M. Lagos-Quintana, R. Sheridan, C. Sander, F. A. Grassler, L. F. van Dyk, C. K. Ho, S. Shuman, M. Chien, J. J. Russo, J. Ju, G. Randall, B. D. Lindenbach, C. M. Rice, V.

- Simon, D. D. Ho, M. Zavolan and T. Tuschl (2005). "Identification of microRNAs of the herpesvirus family." Nat Methods **2**(4): 269-276.
- Phillips, S. L. and W. A. Bresnahan (2011). "Identification of Binary Interactions between Human Cytomegalovirus Virion Proteins." Journal of Virology **85**(1): 440-447.
- Pignatelli, S., P. Dal Monte, G. Rossini, S. Chou, T. Gojobori, K. Hanada, J. Guo, W. Rawlinson, W. Britt and M. Mach (2003). "Human cytomegalovirus glycoprotein N (gpUL73-gN) genomic variants: identification of a novel subgroup, geographical distribution and evidence of positive selective pressure." Journal of General Virology **84**(3): 647-655.
- Pignatelli, S., P. D. Monte, G. Rossini and M. P. Landini (2004). "Genetic polymorphisms among human cytomegalovirus (HCMV) wild - type strains." Reviews in medical virology **14**(6): 383-410.
- Plafker, S. M. and W. Gibson (1998). "Cytomegalovirus assembly protein precursor and proteinase precursor contain two nuclear localization signals that mediate their own nuclear translocation and that of the major capsid protein." J Virol **72**(10): 7722-7732.
- Plotkin, S. A., T. Furukawa, N. Zygraich and C. Huygelen (1975). "Candidate cytomegalovirus strain for human vaccination." Infect Immun **12**(3): 521-527.
- Plotkin, S. A., M. L. Smiley, H. M. Friedman, S. E. Starr, G. R. Fleisher, C. Wlodaver, D. C. Dafoe, A. D. Friedman, R. A. Grossman and C. F. Barker (1984). "Towne-vaccine-induced prevention of cytomegalovirus disease after renal transplants." Lancet **1**(8376): 528-530.
- 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**(5): 860-865.
- Poole, E., C. A. King, J. H. Sinclair and A. Alcami (2006). "The UL144 gene product of human cytomegalovirus activates NFkappaB via a TRAF6-dependent mechanism." EMBO J **25**(18): 4390-4399.
- Prichard, M. N. and E. R. Kern (2011). "The search for new therapies for human cytomegalovirus infections." Virus Res **157**(2): 212-221.
- Prod'homme, V., D. M. Sugrue, R. J. Stanton, A. Nomoto, J. Davies, C. R. Rickards, D. Cochrane, M. Moore, G. W. Wilkinson and P. Tomasec (2010). "Human cytomegalovirus UL141 promotes efficient downregulation of the natural killer cell activating ligand CD112." J Gen Virol **91**(Pt 8): 2034-2039.
- Puchhammer-Stöckl, E. and I. Görzer (2011). "Human cytomegalovirus: an enormous variety of strains and their possible clinical significance in the human host." Future Virology **6**(2): 259-271.
- Quinnan, G. V., Jr., M. Delery, A. H. Rook, W. R. Frederick, J. S. Epstein, J. F. Manischewitz, L. Jackson, K. M. Ramsey, K. Mittal, S. A. Plotkin and et al. (1984). "Comparative virulence and immunogenicity of the Towne strain and a nonattenuated strain of cytomegalovirus." Ann Intern Med **101**(4): 478-483.

- Raanani, P., A. Gaftor-Gvili, M. Paul, I. Ben-Bassat, L. Leibovici and O. Shpilberg (2009). "Immunoglobulin prophylaxis in hematopoietic stem cell transplantation: systematic review and meta-analysis." J Clin Oncol **27**(5): 770-781.
- Rafailidis, P. I., E. G. Mourtzoukou, I. C. Varbobitis and M. E. Falagas (2008). "Severe cytomegalovirus infection in apparently immunocompetent patients: a systematic review." Viol J **5**(1): 47.
- Rasmussen, L., A. Geissler, C. Cowan, A. Chase and M. Winters (2002). "The genes encoding the gCIII complex of human cytomegalovirus exist in highly diverse combinations in clinical isolates." J Virol **76**(21): 10841-10848.
- Rasmussen, L., A. Geissler and M. Winters (2003). "Inter- and intragenic variations complicate the molecular epidemiology of human cytomegalovirus." J Infect Dis **187**(5): 809-819.
- Rasmussen, L., C. Hong, D. Zipeto, S. Morris, D. Sherman, S. Chou, R. Miner, W. L. Drew, R. Wolitz, A. Dowling, A. Warford and T. C. Merigan (1997). "Cytomegalovirus gB genotype distribution differs in human immunodeficiency virus-infected patients and immunocompromised allograft recipients." J Infect Dis **175**(1): 179-184.
- Reese, M. G., F. H. Eeckman, D. Kulp and D. Haussler (1997). "Improved splice site detection in Genie." J Comput Biol **4**(3): 311-323.
- Reeves, M. B. (2011). "Chromatin-mediated regulation of cytomegalovirus gene expression." Virus Res **157**(2): 134-143.
- Reeves, M. B., H. Coleman, J. Chadderton, M. Goddard, J. G. P. Sissons and J. H. Sinclair (2004). "Vascular endothelial and smooth muscle cells are unlikely to be major sites of latency of human cytomegalovirus in Vivo." Journal of General Virology **85**(11): 3337-3341.
- Reeves, M. B., A. A. Davies, B. P. McSharry, G. W. Wilkinson and J. H. Sinclair (2007). "Complex I binding by a virally encoded RNA regulates mitochondria-induced cell death." Science **316**(5829): 1345-1348.
- Reeves, M. B. and J. H. Sinclair (2013). "Circulating dendritic cells isolated from healthy seropositive donors are sites of human cytomegalovirus reactivation in vivo." J Virol **87**(19): 10660-10667.
- Renzette, N., B. Bhattacharjee, J. D. Jensen, L. Gibson and T. F. Kowalik (2011). "Extensive genome-wide variability of human cytomegalovirus in congenitally infected infants." PLoS Pathog **7**(5): e1001344.
- Renzette, N., L. Gibson, B. Bhattacharjee, D. Fisher, M. R. Schleiss, J. D. Jensen and T. F. Kowalik (2013). "Rapid Intrahost Evolution of Human Cytomegalovirus Is Shaped by Demography and Positive Selection." Plos Genetics **9**(9): e1003735.
- Revello, M. G., F. Baldanti, E. Percivalle, A. Sarasini, L. De-Giuli, E. Genini, D. Lilleri, N. Labo and G. Gerna (2001). "In vitro selection of human cytomegalovirus variants unable to transfer virus and virus products from infected cells to polymorphonuclear leukocytes and to grow in endothelial cells." Journal of General Virology **82**(6): 1429-1438.

- Revello, M. G. and G. Gerna (2010). "Human cytomegalovirus tropism for endothelial/epithelial cells: scientific background and clinical implications." Rev Med Virol **20**(3): 136-155.
- Reynolds, D. W., S. Stagno, T. S. Hosty, M. Tiller and C. A. Alford, Jr. (1973). "Maternal cytomegalovirus excretion and perinatal infection." N Engl J Med **289**(1): 1-5.
- Reynolds, D. W., S. Stagno, R. Reynolds and C. A. Alford, Jr. (1978). "Perinatal cytomegalovirus infection: influence of placentally transferred maternal antibody." J Infect Dis **137**(5): 564-567.
- Rhee, M. and P. Davis (2006). "Mechanism of uptake of C105Y, a novel cell-penetrating peptide." J Biol Chem **281**(2): 1233-1240.
- Ribbert, H. (1904). "Ueber protozoenartige Zellen in der Niere eines syphilitischen Neugeborenen und in der Parotis von Kindern." Zentralbl Allg Pathol **15**: 945-948.
- 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**(Pt 2): 393-399.
- 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." J Virol **77**(7): 4326-4344.
- Roizman, B. (1980). "Genome variation and evolution among herpes viruses." Ann N Y Acad Sci **354**: 472-483.
- Roizmann, B., R. Desrosiers, B. Fleckenstein, C. Lopez, A. Minson and M. Studdert (1992). "The family Herpesviridae: an update." Archives of virology **123**(3-4): 425-449.
- Romanowski, M. J. and T. Shenk (1997). "Characterization of the human cytomegalovirus *irsl* and *trsl* genes: a second immediate-early transcription unit within *irsl* whose product antagonizes transcriptional activation." J Virol **71**(2): 1485-1496.
- Ross, S. A., Z. Novak, S. Pati, R. K. Patro, J. Blumenthal, V. R. Danthuluri, A. Ahmed, M. G. Michaels, P. J. Sanchez, D. I. Bernstein, R. W. Tolan, A. L. Palmer, W. J. Britt, K. B. Fowler and S. B. Boppana (2011). "Mixed infection and strain diversity in congenital cytomegalovirus infection." J Infect Dis **204**(7): 1003-1007.
- Rowe, W. P., J. W. Hartley, S. Waterman, H. C. Turner and R. J. Huebner (1956). "Cytopathogenic agent resembling human salivary gland virus recovered from tissue cultures of human adenoids." Proc Soc Exp Biol Med **92**(2): 418-424.
- Rowe, W. P., R. J. Huebner, L. K. Gilmore, R. H. Parrott and T. G. Ward (1953). Isolation of a cytopathogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture. Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, NY), Royal Society of Medicine.
- Ryckman, B. J., M. C. Chase and D. C. Johnson (2008). "HCMV gH/gL/UL128-131 interferes with virus entry into epithelial cells: evidence for cell type-specific receptors." Proc Natl Acad Sci U S A **105**(37): 14118-14123.

Ryckman, B. J., M. C. Chase and D. C. Johnson (2010). "Human cytomegalovirus TR strain glycoprotein O acts as a chaperone promoting gH/gL incorporation into virions but is not present in virions." Journal of Virology **84**(5): 2597-2609.

Ryckman, B. J., M. A. Jarvis, D. D. Drummond, J. A. Nelson and D. C. Johnson (2006). "Human cytomegalovirus entry into epithelial and endothelial cells depends on genes UL128 to UL150 and occurs by endocytosis and low-pH fusion." J Virol **80**(2): 710-722.

Ryckman, B. J., B. L. Rainish, M. C. Chase, J. A. Borton, J. A. Nelson, M. A. Jarvis and D. C. Johnson (2008). "Characterization of the human cytomegalovirus gH/gL/UL128-131 complex that mediates entry into epithelial and endothelial cells." J Virol **82**(1): 60-70.

Sabbaj, S., R. F. Pass, P. A. Goepfert and S. Pichon (2011). "Glycoprotein B vaccine is capable of boosting both antibody and CD4 T-cell responses to cytomegalovirus in chronically infected women." J Infect Dis **203**(11): 1534-1541.

Saccoccio, F. M., A. L. Sauer, X. Cui, A. E. Armstrong, E. L. S. E. Habib, D. C. Johnson, B. J. Ryckman, A. J. Klingelutz, S. P. Adler and M. A. McVoy (2011). "Peptides from cytomegalovirus UL130 and UL131 proteins induce high titer antibodies that block viral entry into mucosal epithelial cells." Vaccine **29**(15): 2705-2711.

Sampaio, K. L., Y. Cavignac, Y. D. Stierhof and C. Sinzger (2005). "Human cytomegalovirus labeled with green fluorescent protein for live analysis of intracellular particle movements." J Virol **79**(5): 2754-2767.

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**(2): 975-986.

Sanchez, V., E. Sztul and W. J. Britt (2000). "Human cytomegalovirus pp28 (UL99) localizes to a cytoplasmic compartment which overlaps the endoplasmic reticulum-golgi-intermediate compartment." J Virol **74**(8): 3842-3851.

Sarisky, R. T. and G. S. Hayward (1996). "Evidence that the UL84 gene product of human cytomegalovirus is essential for promoting oriLyt-dependent DNA replication and formation of replication compartments in cotransfection assays." J Virol **70**(11): 7398-7413.

Sarov, I. and I. Abady (1975). "The morphogenesis of human cytomegalovirus: isolation and polypeptide characterization of cytomegalovirions and dense bodies." Virology **66**(2): 464-473.

Sattentau, Q. (2008). "Avoiding the void: cell-to-cell spread of human viruses." Nat Rev Microbiol **6**(11): 815-826.

Schleiss, M. R. (2010). "Cytomegalovirus Vaccines in the Pipeline." Drugs of the Future **35**(12): 999-1006.

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." Journal of Virology **69**(10): 5959-5968.

- Schuessler, A., K. L. Sampaio, L. Scrivano and C. Sinzger (2010). "Mutational mapping of UL130 of human cytomegalovirus defines peptide motifs within the C-terminal third as essential for endothelial cell infection." Journal of virology **84**(18): 9019-9026.
- Schuessler, A., K. L. Sampaio and C. Sinzger (2008). "Charge cluster-to-alanine scanning of UL128 for fine tuning of the endothelial cell tropism of human cytomegalovirus." J Virol **82**(22): 11239-11246.
- Schuessler, A., K. L. Sampaio, S. Straschewski and C. Sinzger (2012). "Mutational mapping of pUL131A of human cytomegalovirus emphasizes its central role for endothelial cell tropism." J Virol **86**(1): 504-512.
- Scrivano, L., C. Sinzger, H. Nitschko, U. H. Koszinowski and B. Adler (2011). "HCMV spread and cell tropism are determined by distinct virus populations." PLoS Pathog **7**(1): e1001256.
- Sekulin, K., I. Gorzer, D. Heiss-Czedik and E. Puchhammer-Stockl (2007). "Analysis of the variability of CMV strains in the RL11D domain of the RL11 multigene family." Virus Genes **35**(3): 577-583.
- Seo, J. Y. and W. J. Britt (2008). "Multimerization of tegument protein pp28 within the assembly compartment is required for cytoplasmic envelopment of human cytomegalovirus." Journal of Virology **82**(13): 6272-6287.
- Sharma, S., T. W. Wisner, D. C. Johnson and E. E. Heldwein (2013). "HCMV gB shares structural and functional properties with gB proteins from other herpesviruses." Virology **435**(2): 239-249.
- Shepp, D. H., M. E. Match, S. M. Lipson and R. G. Pergolizzi (1998). "A fifth human cytomegalovirus glycoprotein B genotype." Res Virol **149**(2): 109-114.
- Shikhagaie, M., E. Mercé-Maldonado, E. Isern, A. Muntasell, M. M. Albà, M. López-Botet, H. Hengel and A. Angulo (2012). "The Human Cytomegalovirus-Specific UL1 Gene Encodes a Late-Phase Glycoprotein Incorporated in the Virion Envelope." Journal of virology **86**(8): 4091-4101.
- Shizuya, H., B. Birren, U. J. Kim, V. Mancino, T. Slepak, Y. Tachiiri and M. Simon (1992). "Cloning and Stable Maintenance of 300-Kilobase-Pair Fragments of Human DNA in Escherichia-Coli Using an F-Factor-Based Vector." Proceedings of the National Academy of Sciences of the United States of America **89**(18): 8794-8797.
- Silva, M. C., J. Schroer and T. Shenk (2005). "Human cytomegalovirus cell-to-cell spread in the absence of an essential assembly protein." Proc Natl Acad Sci U S A **102**(6): 2081-2086.
- 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." Journal of Virology **77**(19): 10594-10605.
- Simmen, K. A., J. Singh, B. G. M. Luukkonen, M. Lopper, A. Bittner, N. E. Miller, M. R. Jackson, T. Compton and K. Fruh (2001). "Global modulation of cellular transcription by human cytomegalovirus is initiated by viral glycoprotein B." Proceedings of the National Academy of Sciences of the United States of America **98**(13): 7140-7145.

Sinclair, J. and P. Sissons (2006). "Latency and reactivation of human cytomegalovirus." J Gen Virol **87**(Pt 7): 1763-1779.

Sinzger, C. (2008). "Entry route of HCMV into endothelial cells." J Clin Virol **41**(3): 174-179.

Sinzger, C., A. L. Bissinger, R. Viebahn, H. Oettle, C. Radke, C. A. Schmidt and G. Jahn (1999). "Hepatocytes are permissive for human cytomegalovirus infection in human liver cell culture and In vivo." J Infect Dis **180**(4): 976-986.

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)(4): 741-750.

Sinzger, C., G. Hahn, M. Digel, R. Katona, K. L. Sampaio, M. Messerle, H. Hengel, U. Koszinowski, W. Brune and B. Adler (2008). "Cloning and sequencing of a highly productive, endotheliotropic virus strain derived from human cytomegalovirus TB40/E." J Gen Virol **89**(Pt 2): 359-368.

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." Journal of General Virology **81**(12): 3021-3035.

Sinzger, C., J. Knapp, B. Plachter, K. Schmidt and G. Jahn (1997). "Quantification of replication of clinical cytomegalovirus isolates in cultured endothelial cells and fibroblasts by a focus expansion assay." Journal of virological methods **63**(1): 103-112.

Sinzger, C., H. Muntefering, T. Loning, H. Stoss, B. Plachter and G. Jahn (1993). "Cell types infected in human cytomegalovirus placentitis identified by immunohistochemical double staining." Virchows Arch A Pathol Anat Histopathol **423**(4): 249-256.

Sinzger, C., B. Plachter, A. Grefte and G. Jahn (1996). "Tissue macrophages are infected by human cytomegalovirus in vivo." Journal of Infectious Diseases **173**(1): 240-245.

Sinzger, C., K. Schmidt, J. Knapp, M. Kahl, R. Beck, J. Waldman, H. Hebart, H. Einsele and G. Jahn (1999). "Modification of human cytomegalovirus tropism through propagation in vitro is associated with changes in the viral genome." Journal of General Virology **80**(11): 2867-2877.

Sissons, J. G. and A. J. Carmichael (2002). "Clinical aspects and management of cytomegalovirus infection." J Infect **44**(2): 78-83.

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." Proc Natl Acad Sci U S A **98**(14): 7829-7834.

Skepper, J. N., A. Whiteley, H. Browne and A. Minson (2001). "Herpes simplex virus nucleocapsids mature to progeny virions by an envelopment --> deenvelopment --> reenvelopment pathway." J Virol **75**(12): 5697-5702.

Slobbe-van Drunen, M., A. Hendrickx, R. Vossen, E. Speel, M. van Dam-Mieras and C. Bruggeman (1998). "Nuclear import as a barrier to infection of human umbilical vein endothelial cells by human cytomegalovirus strain AD169." Virus research **56**(2): 149-156.

Smith, I. L., J. M. Cherrington, R. E. Jiles, M. D. Fuller, W. R. Freeman and S. A. Spector (1997). "High-level resistance of cytomegalovirus to ganciclovir is associated with alterations in both the UL97 and DNA polymerase genes." Journal of Infectious Diseases **176**(1): 69-77.

Smith, I. L., I. Taskintuna, F. M. Rahhal, H. C. Powell, E. Ai, A. J. Mueller, S. A. Spector and W. R. Freeman (1998). "Clinical failure of CMV retinitis with intravitreal cidofovir is associated with antiviral resistance." Arch Ophthalmol **116**(2): 178-185.

Smith, J. D. and E. De Harven (1973). "Concentration of herpesviruses." J Virol **11**(2): 325-328.

Smith, K. O. and L. Rasmussen (1963). "Morphology of Cytomegalovirus (Salivary Gland Virus)." J Bacteriol **85**(6): 1319-1325.

Smith, M. G. (1956). Propagation in tissue cultures of a cytopathogenic virus from human salivary gland virus (SGV) disease. Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, NY), Royal Society of Medicine.

Smith, M. S., G. L. Bentz, J. S. Alexander and A. D. Yurochko (2004). "Human cytomegalovirus induces monocyte differentiation and migration as a strategy for dissemination and persistence." Journal of Virology **78**(9): 4444-4453.

Snydman, D. R., B. G. Werner, N. N. Dougherty, J. Griffith, R. H. Rubin, J. L. Dienstag, R. H. Rohrer, R. Freeman, R. Jenkins, W. D. Lewis, S. Hammer, E. O'Rourke, G. F. Grady, K. Fawaz, M. M. Kaplan, M. A. Hoffman, A. T. Katz, M. Doran and C. S. G. Boston Center for Liver Transplantation (1993). "Cytomegalovirus immune globulin prophylaxis in liver transplantation. A randomized, double-blind, placebo-controlled trial." Ann Intern Med **119**(10): 984-991.

Soderberg - Naucler, C. (2006). "Does cytomegalovirus play a causative role in the development of various inflammatory diseases and cancer?" Journal of internal medicine **259**(3): 219-246.

Soroceanu, L., A. Akhavan and C. S. Cobbs (2008). "Platelet-derived growth factor-alpha receptor activation is required for human cytomegalovirus infection." Nature **455**(7211): 391-395.

Soroceanu, L. and C. S. Cobbs (2011). "Is HCMV a tumor promoter?" Virus research **157**(2): 193-203.

Spaderna, S., G. Hahn and M. Mach (2004). "Glycoprotein gpTRL10 of human cytomegalovirus is dispensable for virus replication in human fibroblasts." Archives of Virology **149**(3): 495-506.

Spaderna, S., B. Kropff, Y. Kodel, S. Shen, S. Coley, S. Lu, W. Britt and M. Mach (2005). "Deletion of gpUL132, a structural component of human cytomegalovirus, results in impaired virus replication in fibroblasts." J Virol **79**(18): 11837-11847.

Spaete, R. R. and E. S. Mocarski (1985). "The alpha sequence of the cytomegalovirus genome functions as a cleavage/packaging signal for herpes simplex virus defective genomes." Journal of virology **54**(3): 817-824.

Spaete, R. R., K. Perot, P. I. Scott, J. A. Nelson, M. F. Stinski and C. Pachl (1993). "Coexpression of truncated human cytomegalovirus gH with the UL115 gene product or the truncated human fibroblast growth factor receptor results in transport of gH to the cell surface." Virology **193**(2): 853-861.

Spaete, R. R., R. M. Thayer, W. S. Probert, F. R. Masiarz, S. H. Chamberlain, L. Rasmussen, T. C. Merigan and C. Pachl (1988). "Human cytomegalovirus strain Towne glycoprotein B is processed by proteolytic cleavage." Virology **167**(1): 207-225.

Spector, D. H., L. Hock and J. C. Tamashiro (1982). "Cleavage maps for human cytomegalovirus DNA strain AD169 for restriction endonucleases EcoRI, BglII, and HindIII." J Virol **42**(2): 558-582.

Spector, D. J. and K. Yetming (2010). "UL84-independent replication of human cytomegalovirus strain TB40/E." Virology **407**(2): 171-177.

Stagno, S., D. W. Reynolds, R. F. Pass and C. A. Alford (1980). "Breast milk and the risk of cytomegalovirus infection." N Engl J Med **302**(19): 1073-1076.

Stanton, R., B. McSharry, M. Armstrong, P. Tomasec and G. Wilkinson (2008). "Re-engineering adenovirus vector systems to enable high-throughput analyses of gene function." BioTechniques **45**(6): 659-668.

Stanton, R., D. Westmoreland, J. D. Fox, A. J. Davison and G. W. Wilkinson (2005). "Stability of human cytomegalovirus genotypes in persistently infected renal transplant recipients." J Med Virol **75**(1): 42-46.

Stanton, R. J., K. Baluchova, D. J. Dargan, C. Cunningham, O. Sheehy, S. Seirafian, B. P. McSharry, M. L. Neale, J. A. Davies, P. Tomasec, A. J. Davison and G. W. G. Wilkinson (2010). "Reconstruction of the complete human cytomegalovirus genome in a BAC reveals RL13 to be a potent inhibitor of replication." Journal of Clinical Investigation **120**(9): 3191-3208.

Stanton, R. J., B. P. McSharry, M. Armstrong, P. Tomasec and G. W. Wilkinson (2008). "Re-engineering adenovirus vector systems to enable high-throughput analyses of gene function." Biotechniques **45**(6): 659-662, 664-658.

Stern-Ginossar, N., B. Weisburd, A. Michalski, V. T. Le, M. Y. Hein, S. X. Huang, M. Ma, B. Shen, S. B. Qian, H. Hengel, M. Mann, N. T. Ingolia and J. S. Weissman (2012). "Decoding human cytomegalovirus." Science **338**(6110): 1088-1093.

Stinski, M. F. (1976). "Human cytomegalovirus: glycoproteins associated with virions and dense bodies." Journal of virology **19**(2): 594-609.

Stinski, M. F. (1978). "Sequence of protein synthesis in cells infected by human cytomegalovirus: early and late virus-induced polypeptides." J Virol **26**(3): 686-701.

Strang, B. L., B. J. Bender, M. Sharma, J. M. Pesola, R. L. Sanders, D. H. Spector and D. M. Coen (2012). "A mutation deleting sequences encoding the amino terminus of human cytomegalovirus UL84 impairs interaction with UL44 and capsid localization." J Virol **86**(20): 11066-11077.

Straschewski, S., M. Patrone, P. Walther, A. Gallina, T. Mertens and G. Frascaroli (2011). "Protein pUL128 of human cytomegalovirus is necessary for monocyte infection and blocking of migration." *J Virol* **85**(10): 5150-5158.

Streblow, D. N., S. L. Orloff and J. A. Nelson (2007). "Acceleration of allograft failure by cytomegalovirus." *Curr Opin Immunol* **19**(5): 577-582.

Sullivan, V., C. L. Talarico, S. C. Stanat, M. Davis, D. M. Coen and K. K. Biron (1992). "A protein kinase homologue controls phosphorylation of ganciclovir in human cytomegalovirus-infected cells." *Nature* **358**(6382): 162-164.

Sun, Z., Y. Ji, Q. Ruan, R. He, Y. Ma, Y. Qi, Z. Mao and Y. Huang (2009). "Structure characterization of human cytomegalovirus UL131A, UL130 and UL128 genes in clinical strains in China." *Genet. Mol. Res* **8**(3): 1191-1201.

Sun, Z., G. Ren, Y. Ma, N. Wang, Y. Ji, Y. Qi, M. Li, R. He and Q. Ruan (2010). "Transcription pattern of UL131A-128 mRNA in clinical strains of human cytomegalovirus." *J Biosci* **35**(3): 365-370.

Sylwester, A. W., B. L. Mitchell, J. B. Edgar, C. Taormina, C. Pelte, F. Ruchti, P. R. Sleath, K. H. Grabstein, N. A. Hosken, F. Kern, J. A. Nelson and L. J. Picker (2005). "Broadly targeted human cytomegalovirus-specific CD4⁺ and CD8⁺ T cells dominate the memory compartments of exposed subjects." *J Exp Med* **202**(5): 673-685.

Talbot, P. and J. D. Almeida (1977). "Human cytomegalovirus: purification of enveloped virions and dense bodies." *J Gen Virol* **36**(2): 345-349.

Tamashiro, J. C. and D. H. Spector (1986). "Terminal structure and heterogeneity in human cytomegalovirus strain AD169." *J Virol* **59**(3): 591-604.

Tanaka, J., T. Ogura, S. Kamiya, H. Sato, T. Yoshie, H. Ogura and M. Hatano (1984). "Enhanced replication of human cytomegalovirus in human fibroblasts treated with dexamethasone." *Journal of general virology* **65**(10): 1759-1767.

Tandon, R. and E. S. Mocarski (2008). "Control of cytoplasmic maturation events by cytomegalovirus tegument protein pp150." *J Virol* **82**(19): 9433-9444.

Tandon, R. and E. S. Mocarski (2011). "Cytomegalovirus pUL96 Is Critical for the Stability of pp150-Associated Nucleocapsids." *Journal of Virology* **85**(14): 7129-7141.

Tandon, R. and E. S. Mocarski (2012). "Viral and host control of cytomegalovirus maturation." *Trends Microbiol* **20**(8): 392-401.

Tao, R., J. Xu, H. Gao, W. Zhao and S. Shang (2014). "Characteristics and functions of human cytomegalovirus UL128 gene/protein." *Acta Virol* **58**(2): 103-107.

Taylor-Wiedeman, J., J. G. Sissons, L. K. Borysiewicz and J. H. Sinclair (1991). "Monocytes are a major site of persistence of human cytomegalovirus in peripheral blood mononuclear cells." *J Gen Virol* **72** (Pt 9): 2059-2064.

Taylor-Wiedeman, J., P. Sissons and J. Sinclair (1994). "Induction of endogenous human cytomegalovirus gene expression after differentiation of monocytes from healthy carriers." J Virol **68**(3): 1597-1604.

Terhune, S. S., J. Schroer and T. Shenk (2004). "RNAs are packaged into human cytomegalovirus virions in proportion to their intracellular concentration." Journal of Virology **78**(19): 10390-10398.

Theiler, R. N. and T. Compton (2001). "Characterization of the signal peptide processing and membrane association of human cytomegalovirus glycoprotein O." J Biol Chem **276**(42): 39226-39231.

To, A., Y. Bai, A. Shen, H. Gong, S. Umamoto, S. Lu and F. Liu (2011). "Yeast two hybrid analyses reveal novel binary interactions between human cytomegalovirus-encoded virion proteins." PLoS One **6**(4): e17796.

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**(2): 181-188.

Tooze, J., M. Hollinshead, B. Reis, K. Radsak and H. Kern (1993). "Progeny Vaccinia and Human Cytomegalovirus Particles Utilize Early Endosomal Cisternae for Their Envelopes." European Journal of Cell Biology **60**(1): 163-178.

Trincado, D. E., G. M. Scott, P. A. White, C. Hunt, L. Rasmussen and W. D. Rawlinson (2000). "Human cytomegalovirus strains associated with congenital and perinatal infections." J Med Virol **61**(4): 481-487.

Umashankar, M., A. Petrucelli, L. Cicchini, P. Caposio, C. N. Kreklywich, M. Rak, F. Bughio, D. C. Goldman, K. L. Hamlin and J. A. Nelson (2011). "A Novel Human Cytomegalovirus Locus Modulates Cell Type-Specific Outcomes of Infection." PLoS Pathogens **7**(12): e1002444.

Vanarsdall, A. L., M. C. Chase and D. C. Johnson (2011). "Human cytomegalovirus glycoprotein gO complexes with gH/gL, promoting interference with viral entry into human fibroblasts but not entry into epithelial cells." J Virol **85**(22): 11638-11645.

Vanarsdall, A. L. and D. C. Johnson (2012). "Human cytomegalovirus entry into cells." Curr Opin Virol **2**(1): 37-42.

Vanarsdall, A. L., B. J. Ryckman, M. C. Chase and D. C. Johnson (2008). "Human cytomegalovirus glycoproteins gB and gH/gL mediate epithelial cell-cell fusion when expressed either in cis or in trans." J Virol **82**(23): 11837-11850.

Vanarsdall, A. L., T. W. Wisner, H. Lei, A. Kazlauskas and D. C. Johnson (2012). "PDGF Receptor-alpha Does Not Promote HCMV Entry into Epithelial and Endothelial Cells but Increased Quantities Stimulate Entry by an Abnormal Pathway." PLoS Pathog **8**(9): e1002905.

Varnum, S. M., D. N. Streblow, M. E. Monroe, P. Smith, K. J. Auberry, L. Pasa-Tolic, D. Wang, D. G. Camp, 2nd, 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." J Virol **78**(20): 10960-10966.

- Vey, M., W. Schäfer, B. Reis, R. Ohuchi, W. Britt, W. Garten, H.-D. Klenk and K. Radsak (1995). "Proteolytic processing of human cytomegalovirus glycoprotein B (gpUL55) is mediated by the human endoprotease furin." Virology **206**(1): 746-749.
- Vischer, H. F., R. Leurs and M. J. Smit (2006). "HCMV-encoded G-protein-coupled receptors as constitutively active modulators of cellular signaling networks." Trends Pharmacol Sci **27**(1): 56-63.
- Von Glahn, W. P., AM. (1925). "Intranuclear inclusions in visceral disease." Am J Pathol **1**: 445-446.
- Vossen, M. T., E. M. Westerhout, C. Soderberg-Naucler and E. J. Wiertz (2002). "Viral immune evasion: a masterpiece of evolution." Immunogenetics **54**(8): 527-542.
- Waldman, W. J., D. A. Knight, E. H. Huang and D. D. Sedmak (1995). "Bidirectional Transmission of Infectious Cytomegalovirus between Monocytes and Vascular Endothelial-Cells - an in-Vitro Model." Journal of Infectious Diseases **171**(2): 263-272.
- 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**(3-4): 143-164.
- Wang, D., W. Bresnahan and T. Shenk (2004). "Human cytomegalovirus encodes a highly specific RANTES decoy receptor." Proc Natl Acad Sci U S A **101**(47): 16642-16647.
- Wang, D., F. Li, D. C. Freed, A. C. Finnefrock, A. Tang, S. N. Grimes, D. R. Casimiro and T. M. Fu (2011). "Quantitative analysis of neutralizing antibody response to human cytomegalovirus in natural infection." Vaccine **29**(48): 9075-9080.
- Wang, D. and T. Shenk (2005). "Human cytomegalovirus UL131 open reading frame is required for epithelial cell tropism." J Virol **79**(16): 10330-10338.
- Wang, D. and T. Shenk (2005). "Human cytomegalovirus virion protein complex required for epithelial and endothelial cell tropism." Proc Natl Acad Sci U S A **102**(50): 18153-18158.
- Wang, S. K., C. Y. Duh and C. W. Wu (2004). "Human cytomegalovirus UL76 encodes a novel virion-associated protein that is able to inhibit viral replication." J Virol **78**(18): 9750-9762.
- Wang, X., D. Y. Huang, S.-M. Huong and E.-S. Huang (2005). "Integrin $\alpha v \beta 3$ is a coreceptor for human cytomegalovirus." Nature medicine **11**(5): 515-521.
- Wang, X., S. M. Huong, M. L. Chiu, N. Raab-Traub and E. S. Huang (2003). "Epidermal growth factor receptor is a cellular receptor for human cytomegalovirus." Nature **424**(6947): 456-461.
- Warming, S., N. Costantino, D. L. Court, N. A. Jenkins and N. G. Copeland (2005). "Simple and highly efficient BAC recombineering using galK selection." Nucleic Acids Res **33**(4): e36.
- Weekes, M. P., P. Tomasec, E. L. Huttlin, C. A. Fielding, D. Nusinow, R. J. Stanton, E. C. Wang, R. Aicheler, I. Murrell, G. W. Wilkinson, P. J. Lehner and S. P. Gygi (2014).

"Quantitative temporal viromics: an approach to investigate host-pathogen interaction." Cell **157**(6): 1460-1472.

Weller, T. H. (1970). "Cytomegaloviruses: the difficult years." Journal of Infectious Diseases **122**(6): 532-539.

Weller, T. H. (1971). "The cytomegaloviruses: ubiquitous agents with protean clinical manifestations. I." The New England journal of medicine **285**(4): 203-214.

Weller, T. H. (1971). "The cytomegaloviruses: ubiquitous agents with protean clinical manifestations. II." N Engl J Med **285**(5): 267-274.

Weller, T. H., J. Macauley, J. Craig and P. Wirth (1957). Isolation of intranuclear inclusion producing agents from infants with illnesses resembling cytomegalic inclusion disease. Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, NY), Royal Society of Medicine.

Weller, T. H. H., J.B.; Scott, D.E. (1960). "Serological differentiation of viruses responsible for cytomegalic inclusion disease." Virology **108**: 843-846.

Wen, Y., J. Monroe, C. Linton, J. Archer, C. W. Beard, S. W. Barnett, G. Palladino, P. W. Mason, A. Carfi and A. E. Lilja (2014). "Human cytomegalovirus gH/gL/UL128/UL130/UL131A complex elicits potently neutralizing antibodies in mice." Vaccine **32**(30): 3796-3804.

Weststrate, M. W., J. L. Geelen and J. van der Noordaa (1980). "Human cytomegalovirus DNA: physical maps for restriction endonucleases BglII, hindIII and XbaI." J Gen Virol **49**(1): 1-21.

Wilkinson, G. W. A., R.J.; Wang, E.C. (2013). Natural Killer Cells and Human Cytomegalovirus. Cytomegaloviruses: From Molecular Pathogenesis to Intervention. M. J. L. Reddehase, N.A.W., Caister Academic Press. **I**: 173-191.

Wille, P. T., A. J. Knoche, J. A. Nelson, M. A. Jarvis and D. C. Johnson (2010). "A human cytomegalovirus gO-null mutant fails to incorporate gH/gL into the virion envelope and is unable to enter fibroblasts and epithelial and endothelial cells." J Virol **84**(5): 2585-2596.

Wille, P. T., T. W. Wisner, B. Ryckman and D. C. Johnson (2013). "Human Cytomegalovirus (HCMV) Glycoprotein gB Promotes Virus Entry In Trans Acting as the Viral Fusion Protein Rather than as a Receptor-Binding Protein." mBio **4**(3).

Wills, M. R., O. Ashiru, M. B. Reeves, G. Okecha, J. Trowsdale, P. Tomasec, G. W. G. Wilkinson, J. Sinclair and J. G. P. Sissons (2005). "Human cytomegalovirus encodes an MHC class I-like molecule (UL142) that functions to inhibit NK cell lysis." Journal of Immunology **175**(11): 7457-7465.

Wills, M. R. M., G.M.; Sissons, J.G.P. (2013). Adaptive Cellular Immunity to HCMV. Cytomegaloviruses: Molecular Pathogenesis to Intervention. M. J. L. Reddehase, N.A.W., Caister Academic Press: 142-172.

Wood, L. J., M. K. Baxter, S. M. Plafker and W. Gibson (1997). "Human cytomegalovirus capsid assembly protein precursor (pUL80.5) interacts with itself and with the major capsid protein (pUL86) through two different domains." J Virol **71**(1): 179-190.

- Wyatt, J. P., J. Saxton and et al. (1950). "Generalized cytomegalic inclusion disease." J Pediatr **36**(3): 271-294, illust.
- Yamashita, Y., K. Shimokata, S. Mizuno, T. Daikoku, T. Tsurumi and Y. Nishiyama (1996). "Calnexin acts as a molecular chaperone during the folding of glycoprotein B of human cytomegalovirus." Journal of virology **70**(4): 2237-2246.
- Yan, H. N., S. Koyan, Y. Inami, Y. Yamamoto, T. Suzutani, M. Mizuguchi, H. Ushijima, I. Kurane and N. Inoue (2008). "Genetic linkage among human cytomegalovirus glycoprotein N (gN) and gO genes, with evidence for recombination from congenitally and post-natally infected Japanese infants." Journal of General Virology **89**: 2275-2279.
- Yu, D., M. C. Silva and T. Shenk (2003). "Functional map of human cytomegalovirus AD169 defined by global mutational analysis." Proc Natl Acad Sci U S A **100**(21): 12396-12401.
- Yu, D., G. A. Smith, L. W. Enquist and T. Shenk (2002). "Construction of a self-excisable bacterial artificial chromosome containing the human cytomegalovirus genome and mutagenesis of the diploid TRL/IRL13 gene." J Virol **76**(5): 2316-2328.
- 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." J Virol **79**(2): 1327-1332.
- Yu, X., S. Shah, M. Lee, W. Dai, P. Lo, W. Britt, H. Zhu, F. Liu and Z. H. Zhou (2011). "Biochemical and structural characterization of the capsid-bound tegument proteins of human cytomegalovirus." J Struct Biol **174**(3): 451-460.
- Yu, X., P. Trang, S. Shah, I. Atanasov, Y. H. Kim, Y. Bai, Z. H. Zhou and F. Liu (2005). "Dissecting human cytomegalovirus gene function and capsid maturation by ribozyme targeting and electron cryomicroscopy." Proc Natl Acad Sci U S A **102**(20): 7103-7108.
- Zanghellini, F., S. B. Boppana, V. C. Emery, P. D. Griffiths and R. F. Pass (1999). "Asymptomatic primary cytomegalovirus infection: virologic and immunologic features." J Infect Dis **180**(3): 702-707.
- Zhang, G., B. Raghavan, M. Kotur, J. Cheatham, D. Sedmak, C. Cook, J. Waldman and J. Trgovcich (2007). "Antisense transcription in the human cytomegalovirus transcriptome." J Virol **81**(20): 11267-11281.
- Zheng, Q., R. Tao, H. Gao, J. Xu, S. Shang and N. Zhao (2012). "HCMV-encoded UL128 enhances TNF-alpha and IL-6 expression and promotes PBMC proliferation through the MAPK/ERK pathway in vitro." Viral Immunol **25**(2): 98-105.
- Zhou, M., Q. Yu, A. Wechsler and B. J. Ryckman (2013). "Comparative analysis of gO isoforms reveals that strains of human cytomegalovirus differ in the ratio of gH/gL/gO and gH/gL/UL128-131 in the virion envelope." Journal of virology **87**(17): 9680-9690.

9 Appendix I: Recombineering experiment design

A.I.1. Recombineering of BAC-cloned HCMV genomes (Tables A.I.1-4)

- Designation of the novel BAC-cloned HCMV genome construct.
- Co-ordinates for the region in the 'recipient' genome to be modified.
- Primers used for the amplification of the selectable cassette that was inserted into the genome region to be modified during the first round of recombineering; primers designated SacBF or SacBR are the forward and reverse reading primers, respectively. Primers included ~80bp in the 5' region with sequence homologous to 'recipient' genome regions flanking the intended site of insertion (bold, underlined), and ~20bp in the 3' region with sequence homologous to the termini of the selectable cassette.
- Co-ordinates within the 'donor' genome for regions/sequences to be transferred into the 'recipient' genome in the second recombineering step.
- Primers used for the amplification of 'donor' genome regions/sequences, or oligonucleotides used for the delivery of short 'donor' sequences. Primers were designed to bind and amplify regions with identical sequence across all strains. For the amplification of each variant UL128L genome region, short ~20bp primers that bound ~80bp outside of UL128 and UL131A were used (Table A.I.1.). Primers used for the amplification of other 'donor' genome regions included: ~80bp in the 5' region with sequence homologous to 'recipient' genome regions flanking the intended site of insertion (bold, underlined); ~20bp in the 3' region with sequence homologous to the termini of the 'donor' genome region/sequence (*italic*). Oligonucleotides used for the transfer of single 'donor' nucleotides or small sequences (red, *italic*) into recipient genomes contained flanking ~50bp sequences homologous to the intended site of insertion in the 'recipient' genome (bold, underlined).

A.I.2. Recombineering of AdZ expression vectors (Table A.I.5)

- V5-tagged transgene inserted into AdZ expression vector construct.
- Transgene co-ordinates in 'donor' strain.
- Primers used for amplification of variant UL128L ORF transgenes: Primers were designed to bind and amplify regions with identical sequence across all strains. Primers were designed to contain ~80bp in the 5' region with sequence homologous to intended site of insertion in the AdZ vector construct (bold, underlined), and ~20bp in the 3' region with sequence homologous to the termini of each variant UL128L ORF (*italic*). Reverse reading primers were designed to exclude natural stop codons in each variant UL128L ORF, and include sequence encoding a 3' V5 tag.

All denoted nucleotide positions refer to sequences of each BAC-cloned strain deposited in GenBank: Merlin - GU179001; TR AC146906; TB40-BAC4 - EF999921; and FIX - AC146907 and 3301 - GQ466044.1.

Table A.I.1. Recombinant Merlin BACs containing UL128L from other strains in place of the native UL128L in Merlin.

Recombinant Merlin BAC	Merlin Genome Region Modified	Primers for cassette amplification	‘Donor’ genome Co-ordinates	Primers for amplification of sequence to be transferred
Merlin-UL128L ^{TR}	Merlin nt 176135-178095	SacBF into Merlin UL128L: <u>ATC CAG CCG TTT GTG TTT CTT</u> <u>AAC GCT CTC CAG GTA CTG ATC</u> <u>CAG GCC CAC GAT CCG GGT TAT</u> <u>CTT GTC GTA TTC CAG CCT GTG ACG</u> <i>GAA GAT CAC TTC G</i>	TR nt 211874-213834	UL128L F: GCG TAT TTC GGA CAA ACA CAC A UL128L R: CGC ATG TTG CAG ACT GAG AAA GA
Merlin-UL128L ^{TB40}			TB40-BAC4 nt 208578-210534	
Merlin-UL128L ^{FIX}			FIX nt 50061-52023	
Merlin-UL128L ³³⁰¹			3301 nt 176289-178245	

Table A.I.2. Recombinant Merlin BACs containing unique mutations identified in TB40-BAC4 UL128 (G>T) and FIX UL130 (A>G).

Recombinant Merlin BAC	Merlin Genome Region Modified	Primers for cassette amplification	'Donor' Sequence	Oligonucleotide for delivery of SNPs
Merlin-UL128 ^{G>T}	Merlin nt 176612	<p>SacBF into Merlin UL128 intron 1:</p> <p><u>GGT GGT GAC GAT CCC GCG AAT CTC AGC CGT TTT</u> <u>CTC GGG ACT GTA GCA GAC TTC GCC GTC CGG ACA</u> <u>CCG CAG CCT GTG CCT GTG ACG GAA GAT CAC TTC G</u></p> <p>SacBR into Merlin UL128 intron 1:</p> <p><u>CTG GAT CTG TCT CTC GAC GTT TCT GAT AGC CAT</u> <u>GTT CCA TCG ACG ATC CTC GGG AAT GCC AGA GTA</u> <u>GAT TTT CAT GAA TCT GAG GTT CTT ATG GCT CTT G</u></p>	TB40-BAC4 nt 209051	<p>UL128 G>T insert oligo:</p> <p><u>GTT TTC TCG GGA CTG TAG</u> <u>CAG ACT TCG CCG TCC GGA</u> <u>CAC CGC AGC CTG TGT ATT</u> <u>CAT GAA AAT CTA CTC TGG</u> <u>CAT TCC CGA GGA TCG TCG</u> <u>ATG GAA CAT G</u></p>
Merlin-UL130 ^{A>G}	Merlin nt 177372	<p>SacBF into Merlin UL130 CDS:</p> <p><u>GTC TGG CCT TCC CGG TTG TAC AGC AGA TAC AGG</u> <u>GTC TCG TTG CGA CAC TCG GGA CCC GTT GAT ACC</u> <u>CGC TGG AAC CCC CCT GTG ACG GAA GAT CAC TTC G</u></p> <p>SacBR into Merlin UL130 CDS:</p> <p><u>TAT TCC AAA CCG CAT GAC GCG GCG ACG TTT TAC TGT CCT</u> <u>TTT CTC TAT CCC TCG CCC CCA CGA TCC CCC TTG CAA TTC</u> <u>CCT GAG GTT CTT ATG GCT CTT G</u></p>	FIX nt 50792	<p>UL130 A>G insert oligo:</p> <p><u>CAG GGT CTC GTT GCG ACA</u> <u>CTC GGG ACC CGT TGA TAC</u> <u>CCG CTG GAA CCC CGG GAA</u> <u>TTG CAA GGG GGA TCG</u> <u>TGG GGG CGA GGG ATA</u> <u>GAG AAA AGG ACA GTA A</u></p>

Table A.I.3. Recombinant Merlin BAC containing TR UL131A 5' – UL132 5' genome region

Recombinant Merlin BAC	Merlin Genome Region Modified	Primers for cassette amplification	'Donor' genome Co-ordinates	Primers for amplification of sequence to be transferred
Merlin-UL132 ^{TR}	Merlin nt 178096-179013	<p>SacBF-into Merlin UL131A-132:</p> <p><u>TTC CGC GGT TTC CCG CTG GCA CTG</u> <u>ACC CAG CAC CAC GGC GCA CAG ACA</u> <u>AAC AGA CAG CCA CAC CCG ACA CAG</u> <u>CCG CAT</u> CCT GTG ACG GAA GAT CAC TTC G</p> <p>SacBR- into Merlin UL131A-132:</p> <p><u>TAG AGG CTT GCG GAA ACC ACG TCC</u> <u>TCG TCA CAC GTC GTT CGC GGA CAT</u> <u>AGC AAG AA ATC CAC GTC GCC ACG TCT</u> <u>CGA GAC</u> TGA GGT TCTT ATG GCT CTT G</p>	TR nt 213835-214752	<p>UL131A-132F:</p> <p><u>TTT TCCG CGG TTT CCC GCT</u> <u>GGC ACT GAC CCA GCA CCA CGG</u> <u>CGC ACA GAC AAA CAG ACA GCC</u> <u>ACA CCC GAC ACA GCC GCA</u> TGT TGC AGA CTG AGA AAG AGA</p> <p>UL131A-132R:</p> <p><u>GTA GAG GCT TGC GGA AAC CAC</u> <u>GTC CTC GTC ACA CGT CGT TCG</u> <u>CGG ACA TAG CAA GAA ATC CAC</u> <u>GTC GCC ACG TCT CGA GAA</u> TGC CGG CCC CGC GGG GT</p>

Table A.I.4. Recombinant Merlin BACs containing tetracycline-binding sequences upstream of UL131A.

Recombinant Merlin BAC	'Recipient' Genome Region Modified	Primers for amplification of selectable cassette	'Donor' Sequence	Oligonucleotide for delivery of tet- binding sequence
Merlin-tet-UL128L	Merlin nt: 178128	SacB F into Merlin UL131A region: <u>ACA GCC ACA CCC GAC ACA GCC</u> <u>GCA TGT TGC AGA CTG AGA AAG</u> <u>AAA GCT TTA TTA TGA GTC TCT ATC</u> <u>ACT GAT AGG GAG CCT GTG ACG GAA</u> <i>GAT CAC TTC G</i>	Tetracycline binding operator site	<u>GCA GAC TGA GAA AGA AAG</u> <u>CTT TAT TAT GAG GCA GAC TGA</u> <u>GAA AGA AAG CTT TAT TAT GAG</u> <i><u>TCT CTA TCA CTG ATA GGG AGA</u></i> <i><u>TCT CTA TCA CTG AGG GAC ATC</u></i> <u>ATA CAC ATA GTA TAG GCG</u> <u>AGG TGA T</u>
Merlin-tet-UL128 ^{G>T}		SacB R into Merlin UL131A region: <u>TCT TTC GGT TCC AAC TCT TTC CCC</u> <u>GCC CCA TCA CCT CGC CTA TAC</u> <u>TAT GTG TAT GAT GTT CCC TAT</u> <u>CAG TGA TAG AGA TCT GAG GTT CTT</u> <i>ATG GCT CTT G</i>		
Merlin-tet-UL130 ^{A>G}				

Table A.I.5. Recombinant AdZ vectors (RAdZ) for the expression of variant UL128L transgenes.

RAdZ Construct	Transgene Co-ordinates in ‘Donor’ genome	Primers for amplification of sequence to be transferred
RAdZ-Me-UL128 ^{wt}	Merlin-UL128L ^{wt} - nt: 176,138 – 176,901	UL128 insert F:
RAdZ-TR-UL128	TR - nt: 211,877 – 212,640	<u>AAC CGT CAG ATC GCC TGG AGA CGC CAT CCA CGC TGT TTT GAC CTC CAT AGA</u> <u>AGA CAC CGG GAC CGA TCC AGC CTG GAT CCC GCG CGT CAT GA G TCC CAA A</u>
RAdZ-TB-UL128	TB40-BAC4 - nt: 174,931 – 175,690	UL128 insert R:
RAdZ-FIX-UL128	FIX - nt: 51,263 – 52,020	<u>ATA GAG TAT ACA ATA GTG ACG TGG GAT CCT TAC GTA GAA TCA AGA CCT AGG</u> <u>AGC GGG TTA GGG ATT GGC TTA CCA GCG CTC TGC AGC ATA TAG CCC ATT T</u>
RAdZ-Me-UL128 ^{G>T}	Merlin-UL128 ^{G>T} - nt: 176,138 – 176,901	
RAdZ-Me-UL130 ^{wt}	Merlin- UL128L ^{wt} - nt: 176,936 – 177,598	UL130 insert F:
RAdZ-TR-UL130	TR nt: 212,675 – 213,337	<u>AAC CGT CAG ATC GCC TGG AGA CGC CAT CCA CGC TGT TTT GAC CTC CAT AGA</u> <u>AGA CAC CGG GAC CGA TCC AGC CTG GAT CCG CCT GCG TCA CGG GAA ATA A</u>
RAdZ-TB-UL130	TB40-BAC4 - nt: 175,25 – 176,387	UL130 insert R:
RAdZ-FIX-UL130	FIX - nt: 50,579 – 51,220	<u>ATA GAG TAT ACA ATA GTG ACG TGG GAT CCT TAC GTA GAA TCA AGA CCT AGG</u> <u>AGC GGG TTA GGG ATT GGC TTA CCA GCG CTA ACG ATG AGA TTG GGA TGG G</u>
RAdZ-Me-UL130 ^{A>G}	Merlin-UL130 ^{A>G} - nt: 176,936 – 177,598	
RAdZ-Me-UL131A	Merlin - nt 177,601 – 178,109	UL131A insert F:
RAdZ-TR-UL131A	TR - nt 213,340 – 213,848	<u>AAC CGT CAG ATC GCC TGG AGA CGC CAT CCA CGC TGT TTT GAC CTC CAT AGA</u> <u>AGA CAC CGG GAC CGA TCC AGC CTG GAT CCT CTC AGT CTG CAA CAT GCG G</u>
RAdZ-TB-UL131A	TB40-BAC4 - nt: 176,390 – 176,898	UL131A insert R:
RAdZ-FIX-UL131A	FIX - nt: 50,047– 50,536	<u>ATA GAG TAT ACA ATA GTG ACG TGG GAT CCT TAC GTA GAA TCA AGA CCT AGG</u> <u>AGC GGG TTA GGG ATT GGC TTA CCA GCG CTG TTG GCA AAG AGC CGC ACG C</u>

10 Appendix II: Impact of variant

UL128L on virus growth

phenotype

To determine the contribution of TR, TB40-BAC4 and FIX UL128L genome regions to the distinct growth phenotypes of virus, each variant UL128L was inserted into Merlin genome in place of the native UL128L genome region, to assess whether the growth phenotype of the donor virus was also transferred (Table A.II.1.).

Table A.II.1. Recombinant Merlin BAC variants containing UL128L from TR, TB40-BAC4 and FIX

Genome Background	UL128L origin	Designation in text: BAC /Reconstituted Virus
	Native; mutated: G>A in UL128 at nt 176260 (R>stop)	Merlin-UL128L ^{mut} /RV-Me-UL128L ^{mut}
	Native; wildtype	Merlin-UL128L ^{wt} /RV-Me-UL128L ^{wt}
Merlin	TR	Merlin-UL128L ^{FIX} /RV-Me-UL128L ^{TR}
	TB40-BAC4	Merlin-UL128L ^{TB40} /RV-Me-UL128L ^{TB40}
	FIX	Merlin-UL128L ^{FIX} /RV-Me-UL128L ^{FIX}

A.II.1 Impact of variant UL128L in fibroblasts

A.II.1.1 Cell-to-cell spread

Fibroblasts (HFFFs) were again transfected with BAC DNA for the recombinant Merlin viruses to assess the specific impact of each variant UL128L genome region on the cell-to-cell spread of virus in this cell-type (Figure A.II.1.A). As previously demonstrated, the Merlin virus lacking intact UL128L (RV-Me-UL128L^{mut}) displayed the greatest capacity for cell-to-cell spread in fibroblasts and formed HFFF plaques that were significantly larger than any other virus assessed ($p < 0.0001$). However, whereas RV-TR had formed HFFF plaques that were smaller than those produced by RV-Me-UL128L^{wt}, RV-Me-UL128L^{TR} formed HFFF plaques with comparable size to those formed by RV-Me-UL128L^{wt}. RV-Me-UL128L^{TB40} and RV-Merlin-UL128L^{FIX} each formed larger HFFFs plaques than RV-Me-UL128L^{wt} ($p < 0.001$ and $p < 0.01$, respectively), mirroring the previously observed greater cell-to-cell spread displayed by both RV-TB40-BAC4 and RV-FIX compared to RV-Me-UL128L^{wt}.

A.II.1.2 Dissemination kinetics and cell-free virus production

RV-Me-UL128L^{mut} again spread the most efficiently (Figure A.II.1.B) and produced the greatest peak titres of cell-free virus (Figure A.II.1.C), whilst dissemination of RV-Me-UL128L^{wt} was again delayed and produced the lowest peak titre of cell-free virus. However, in this assay, the peak titre of cell-free virus produced by RV-Me-UL128L^{wt} was only ~ 270-fold lower than that produced by RV-Me-UL128L^{mut}, whereas the difference had previously been ~1000-fold. The dissemination kinetics of RV-Me-UL128L^{TR} resembled that of RV-Me-UL128L^{wt} at all time points through the infection time-course, yet similar to what was demonstrated by RV-TR, cell-free virus production by RV-Me-UL128L^{TR} rapidly increased at later time points, and the peak titre of cell-free virus produced was ~100-fold greater than that produced by RV-Me-UL128L^{wt}. Merlin viruses expressing TB40-BAC4 and FIX UL128L each displayed dissemination kinetics and cell-free virus production that was intermediate compared to the Merlin control infections. RV-Me-UL128L^{TB40} produced peak titres of cell free virus ~4-fold greater than that produced by RV-Me-UL128L^{wt}, whilst Me-UL128L^{FIX} produced a peak titre of cell- free virus ~7.5-fold greater than that produced by RV-Me-UL128L^{wt}.

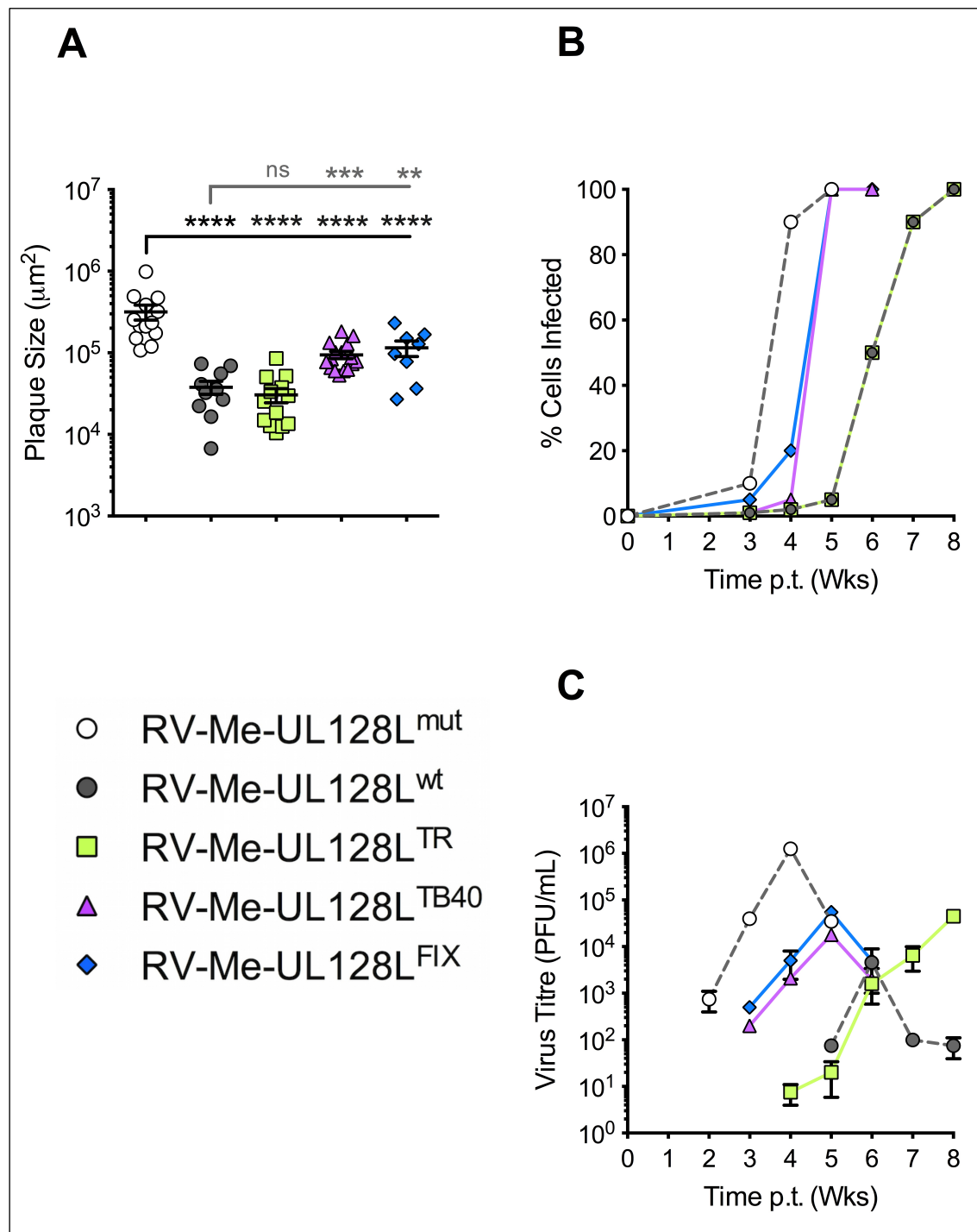


Figure A.II.1 Growth phenotype impact of variant UL128L in fibroblasts. (A) HFFF cells were transfected with BAC DNA for the indicated viruses and incubated under semisolid overlay, with plaque sizes measured 2 weeks post transfection. Means and \pm SEM are shown. (Where indicated, the mean plaque sizes formed by RV-Me-UL128L^{wt}, RV-Me-UL128L^{TR}, RV-Me-UL128L^{TB40}, RV-Me-UL128L^{FIX} were compared to those formed by RV-Me-UL128L^{mut} (black line) and/or RV-Me-UL128L^{wt} (grey line) by 1-way analysis of variance [ANOVA] followed by Dunnett's posttest: * $p < 0.05$, ** $p < 0.01$, *** $p < .001$, **** $p < 0.0001$). (B) HFFF cells were transfected with BAC DNA for the indicated viruses, and the dissemination of infection was allowed to progress until the monolayer was destroyed. At weekly time points post transfection (p.t.), the level of infection was estimated by fluorescence microscopy analysis of eGFP-expressing cells. (C) Supernatant samples collected at weekly time points from the infections shown in panel B were titrated on HFFFs to provide a measure of cell-free virus release throughout the time-course of each infection.

A.II.1.3. Validation of variant UL128L impact in fibroblasts

To determine whether the growth phenotypes of the recombinant Merlin viruses were an accurate reflection of the impact of each variant UL128L, we again sought to determine whether any virus had undergone adaptation during the growth kinetics assays in HFFFs. The relative plaque sizes formed following transfection of HFFFs with BAC DNA for each virus (p0 virus) (Figure A.II.2.A) were again compared to those formed during titration of progeny virus produced at the latest time-points in each infection (HFFFp1 virus) (Figure A.II.2.B). With the exception of RV-Me-UL128L^{TR}—HFFFp1, no other progeny displayed any signs of adaptation. Furthermore, PCR-based sequencing confirmed UL128L in these viruses to be equivalent to each parental BAC clone. Whereas transfection of HFFFs with Merlin-UL128L^{TR} BAC DNA resulted in the formation in smallest plaques of any virus investigated, HFFF plaques formed during titration of the progeny produced in the RV-Me-UL128L^{TR}-HFFFp1 of comparable size to those formed by progeny of the Merlin virus lacking intact UL128L (RV-Me-UL128L^{mut}-HFFFp1). PCR-based sequencing analysis revealed the RV-Me-UL128L^{TR}-HFFFp1 to contain a major mutant population that acquired a 507 bp deletion affecting ORFs UL128 and UL130 (nt 176,742-177,249, as well as a premature stop coding SNP (Q>Stop) in UL130 (nt 177,355). Thus, the dramatic increase in cell-free virus production by the RV-Me-UL128L^{TR} infection occurred secondarily to mutation in UL128L.

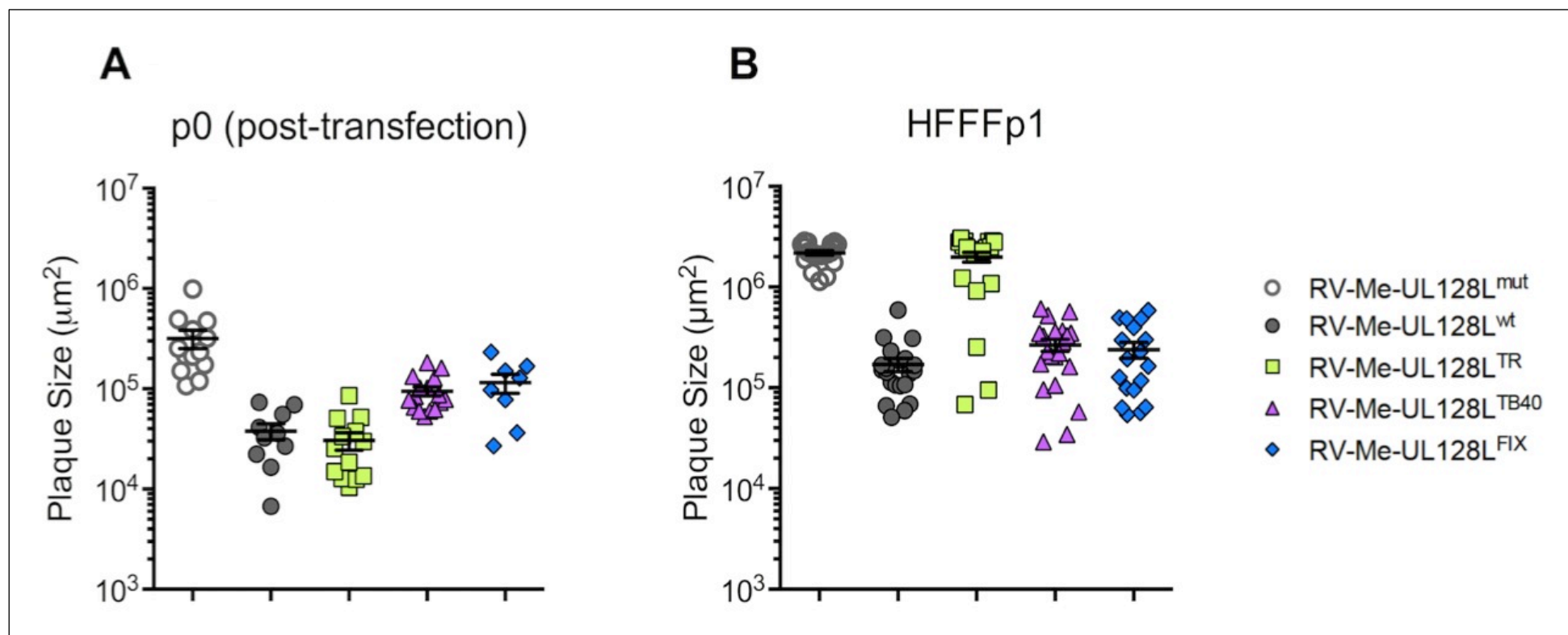


Figure A.II.2 Cell-to-cell spread of viruses derived from Merlin-BAC variants containing UL128L from other strains in fibroblasts, before and after passage in fibroblasts. (A) Plaque size (\pm SEM) formed 2 weeks post-transfection of HFFFs with BAC DNA (p0) for each indicated virus, and incubation under overlay. (B) HFFF plaque sizes (\pm SEM) formed 2 weeks post-infection of HFFFs with viruses in supernatants collected at final time-point of growth kinetics assays performed in fibroblasts (HFFF-p1), and incubation overlay.

A.II.2.1 Impact of variant UL128L in epithelial cells

A.II.2.2 Cell-to-cell spread

The specific contribution of each variant UL128L to cell-to-cell spread of virus in epithelial cells (RPE-1s) was assessed (Figure A.II.3.A). As previously demonstrated, RV-Me-UL128L^{wt} again formed the largest RPE-1 plaques, whilst RV-Me-UL128L^{mut} was unable to form multi-cellular RPE-1 plaques. RV-Me-UL128L^{TR} formed RPE-1 plaques with mean size that did not differ from those produced by RV-Me-UL128L^{wt} by any significant degree. In contrast to this, whilst RV-Me-UL128L^{TB40} formed multi-cellular RPE-1 plaques that were significantly larger than those formed by RV-Me-UL128L^{mut} ($p < 0.0001$), these were still smaller than those formed by RV-Me-UL128L^{wt} ($p < 0.01$). Similar to RV-FIX, RV-Me-UL128L^{FIX} formed RPE-1 plaques with comparable size to those formed by RV-Me-UL128L^{mut} and significantly smaller compared to those formed by RV-Me-UL128L^{wt} ($p < 0.0001$), the smallest of any virus investigated. However unlike the majority of RPE-1 plaques formed by RV-FIX, multi-cellular plaques formed by RV-Me-UL128L^{FIX} were consistently observed.

A.II.2.3 Dissemination kinetics and cell-free virus production

An assessment of the dissemination kinetics and cell-free virus production of each recombinant Merlin virus was also performed in epithelial cell (Figure A.II.3.B+C). RV-Me-UL128L^{wt} spread the fastest in RPE-1s, and similar to findings in fibroblasts, RV-Me-UL128L^{TR} spread with identical kinetics. Furthermore the peak titre of cell-free virus produced by RV-Me-UL128L^{TR} was also comparable to that produced by RV-Me-UL128L^{wt}. RV-Me-UL128L^{TB40} also displayed identical dissemination kinetics to RV-Me-UL128L^{wt}, unlike the parental virus RV-TB40-BAC4 that displayed delayed dissemination kinetics in this cell type. However, like RV-TB40-BAC4, RV-Me-UL128L^{TB40} did produce greater peak titres of cell-free virus that were ~18-fold greater than that produced by RV-Me-UL128L^{wt}. Interestingly, RV-Me-UL128L^{FIX} was able to establish productive infection in RPE-1s, whereas RV-FIX previously could not. Nonetheless, transfer of FIX UL128L did reduce the dissemination kinetics of Merlin virus in RPE-1s, and also increased the productivity of cell-free virus in this cell type such that the peak titre of cell-free virus produced was ~12-fold greater than that produced by RV-Me-UL128L^{wt}.

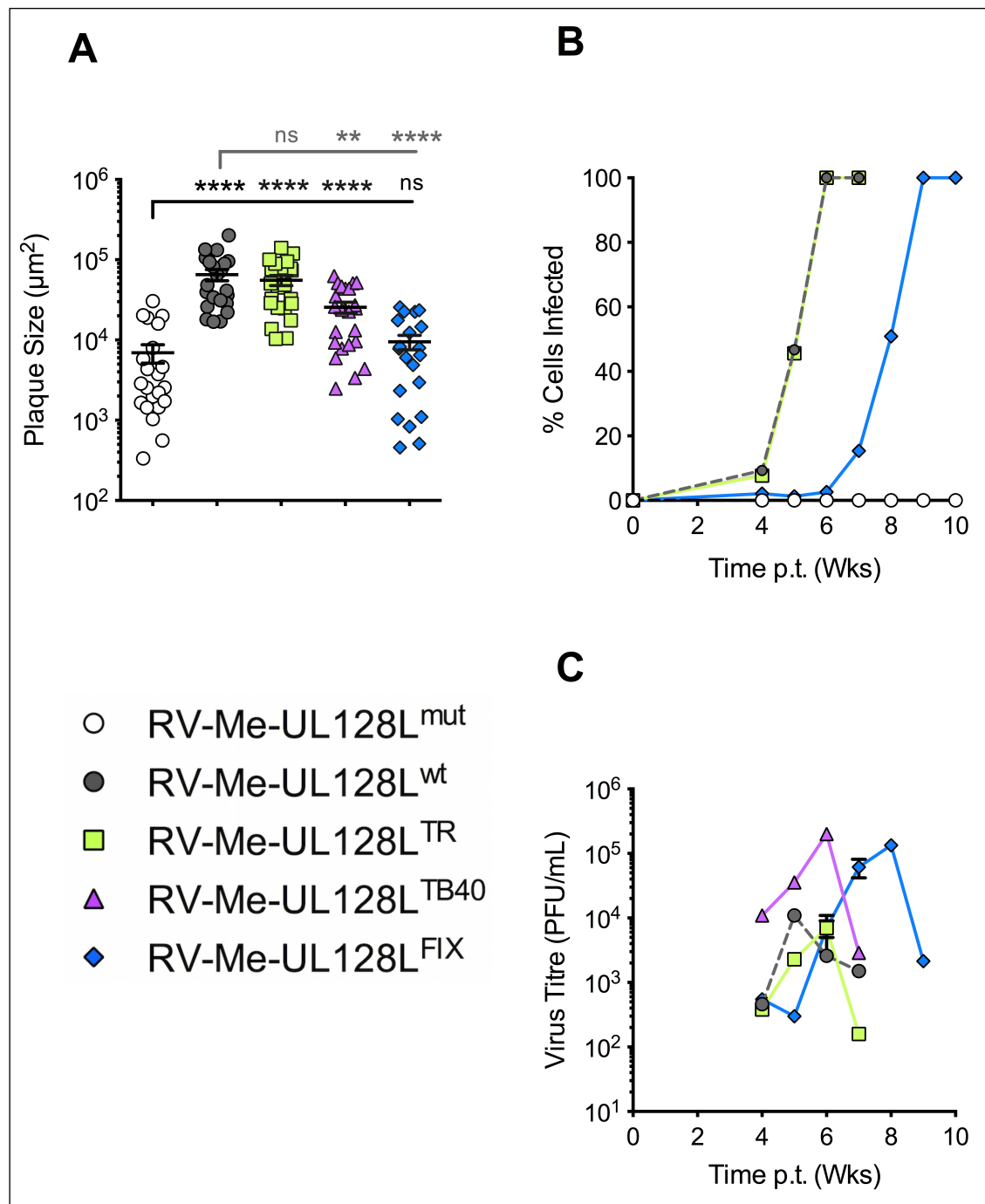


Figure A.II.2 Growth phenotype impact of variant UL128L in Epithelial cells. (A) RPE-1 cells were transfected with BAC DNA for the indicated viruses and incubated under semisolid overlay, with plaque sizes measured 3 weeks post transfection. Means and \pm SEM are shown. (Where indicated, the mean plaque sizes formed by RV-Me-UL128L^{wt}, RV-Me-UL128L^{TR}, RV-Me-UL128L^{TB40}, RV-Me-UL128L^{FIX} were compared to those formed by RV-Me-UL128L^{mut} (black line) and/or RV-Me-UL128L^{wt} (grey line) by 1-way analysis of variance [ANOVA] followed by Dunnett's posttest: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$). (B) RPE-1 cells were transfected with BAC DNA for the indicated viruses, and the dissemination of infection was allowed to progress until the monolayer was destroyed. At weekly time points post transfection (p.t.), the level of infection was estimated by fluorescence microscopy analysis of eGFP-expressing cells. (C) Supernatant samples collected at weekly time points from the infections shown in panel B were titrated on HFFFs to provide a measure of cell-free virus release throughout the time-course of each infection.

A.II.2.4. Validation of variant UL128L impact in epithelial cells

Comparison of HFFF plaque sizes formed immediately following transfection (Figure A.II.4.A), to those produced during titration of progeny virus produced in RPE-1 epithelial cells (Figure A.II.4.B), gave no indication that any virus had undergone adaptation in epithelial cell culture. PCR-based sequencing confirmed UL128L in these viruses to be equivalent to each parental BAC clone. Hence, the impact of each variant UL128L on the growth phenotype of virus in epithelial cells was accurately represented.

A.II.2.5. Conclusion

In summary, TR UL128L was not the major difference responsible for the distinct growth phenotype of RV-TR in epithelial cell culture compared to RV-Me-UL128L^{wt}. In fact, in all investigations performed, TR UL128L was functionally equivalent to that of wildtype UL128L. However, TB40-BAC4 and FIX UL128L were responsible, at least partially for the distinct growth phenotypes of RV-TB40-BAC4 and RV-FIX in epithelial cell culture. Interestingly, RV-FIX was non-viable in RPE-1s, yet RV-Me-UL128L^{FIX} was, albeit with dramatically reduced fitness compared to RV-Me-UL128L^{wt}. It was therefore confirmed that FIX UL128L is indeed functionally intact, and sequence features in other genome regions must dictate the incompatibility of FIX in RPE-1s.

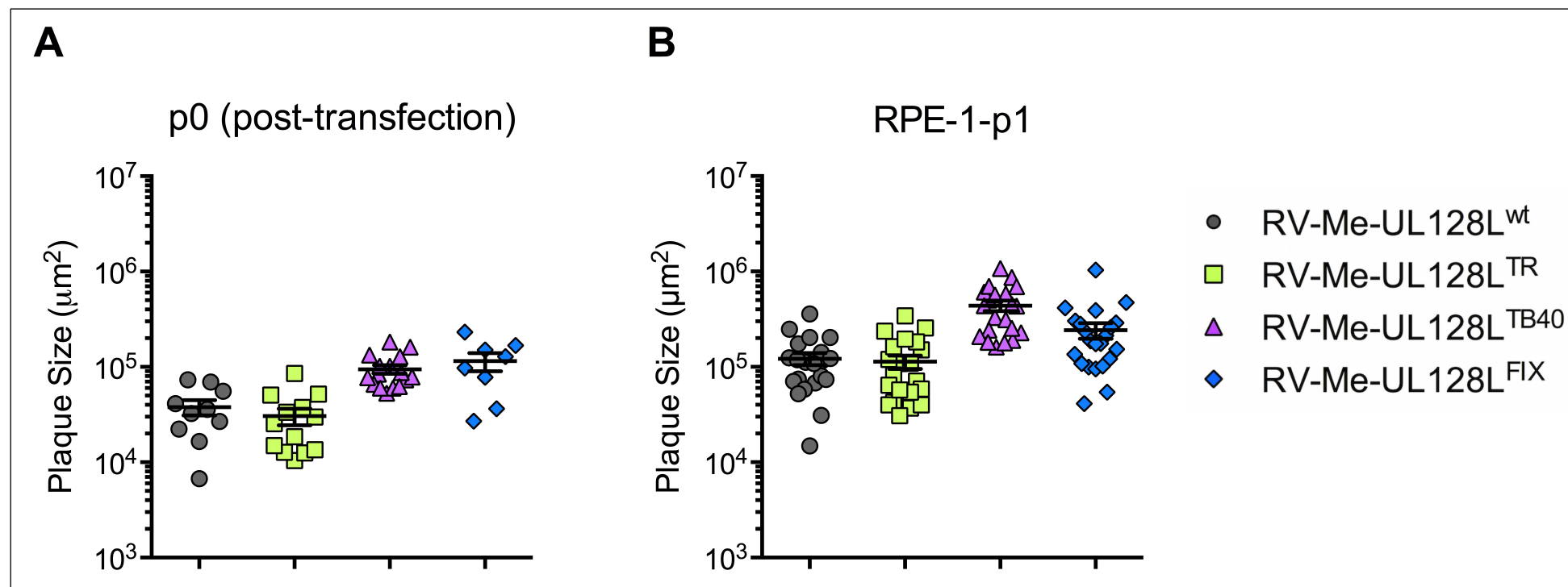


Figure A.II.4 Cell-to-cell spread of viruses derived from Merlin-BAC variants containing UL128L from other strains in fibroblasts, before and after passage in epithelial cells. (A) Plaque size (\pm SEM) formed 2 weeks post-transfection of HFFFs with BAC DNA (p0) for each indicated virus, and incubation under overlay. (B) HFFF plaque sizes (\pm SEM) formed 2 weeks post-infection of HFFFs with viruses in supernatants collected at final time-point of growth kinetics assays performed in epithelial cells (RPE-1-p1), and incubation overlay.

11 Appendix III: Plaque size comparison data

Plaque size comparisons were performed as follows:

- Plaque sizes formed by each virus in either fibroblast (HFFFs) or epithelial cells (RPE-1s and ARPE-19s) were first logged, thus normalising each data set.
- Normalized plaque size data sets were then analysed by one-way analysis of variance (ANOVA) and Dunnett's post-test.
- In the tables below, the mean plaque size formed by each infection is noted, as are the fold-differences in plaque sizes between infections, and the significance (p) of any difference observed (as determined by Dunnett's post-test).
- Mean plaque sizes formed by each infection that differed by a significant degree are highlighted: where the mean plaque sizes formed by any given virus were larger than those formed the comparison virus, in green; where smaller than those formed by the comparison virus, in pink.

Table A.III.1. Comparison of mean HFFF and RPE-1 plaque sizes formed immediately following transfection with BAC DNA for each strain investigated.

Virus	HFFFs			RPE-1s		
	Mean Plaque Size (μm^2)	Compared to RV-Me-UL128L ^{mut}	Compared to RV-Me-UL128L ^{wt}	Mean Plaque Size (μm^2)	Compared to RV-Me-UL128L ^{mut}	Compared to RV-Me-UL128L ^{wt}
RV-Me-UL128L ^{mut}	627681.875		+ 8.75x ($p < 0.0001$)	1886.625		- 32.25x ($p < 0.0001$)
RV-Me-UL128L ^{wt}	70942.625	- 8.75x ($p < 0.0001$)		60977.625	+ 32.25x ($p < 0.0001$)	
RV-TR	29882.25	- 21x ($p < 0.0001$)	- 2.5x ($p < 0.05$)	9897.625	+ 5.25x ($p < 0.001$)	- 6.25x ($p < 0.0001$)
RV-TB40-BAC4	285380.125	- 2.25x (ns)	+ 4x ($p < 0.0001$)	11497.25	+ 6x ($p < 0.0001$)	- 5.25x ($p < 0.0001$)
RV-FIX	188327.50	- 3.25x ($p < 0.0001$)	+ 2.75x ($p < 0.01$)	2306.00	+ 1.25x (ns)	- 26.5x ($p < 0.0001$)

Table A.III.2. Comparison of mean HFFF and RPE-1 plaque sizes formed immediately following transfection with BAC DNA for recombinant Merlin viruses containing variant UL128L genome regions from other strains.

Virus	HFFFs			RPE-1s		
	Mean Plaque Size (μm^2)	Compared to RV-Me-UL128L ^{mut}	Compared to RV-Me-UL128L ^{wt}	Mean Plaque Size (μm^2)	Compared to RV-Me-UL128L ^{mut}	Compared to RV-Me-UL128L ^{wt}
RV-Me-UL128L ^{mut}	317676.00		+ 8.5x ($p < 0.0001$)	6945.625		-9.5x ($p < 0.0001$)
RV-Me-UL128L ^{wt}	37989.50	- 8.25x ($p < 0.0001$)		65130.50	+ 9.5x ($p < 0.0001$)	
RV-TR	30536.50	- 10.5x ($p < 0.0001$)	- 1.25x (ns)	55584.125	+ 8x ($p < 0.0001$)	- 1.25x (ns)
RV-TB40-BAC4	94620.50	- 3.25x ($p < 0.0001$)	+ 2.5x ($p < 0.001$)	25579.25	+ 3.75x ($p < 0.0001$)	- 2.5x ($p < 0.01$)
RV-FIX	114900.125	- 2.75x ($p < 0.0001$)	+ 3x ($p < 0.01$)	9476.92	+ 1.25x (ns)	- 6.75x ($p < 0.0001$)

Table A.III.3. Comparison of HFFF and RPE-1 plaques formed immediately following transfection with BAC DNA for recombinant Merlin viruses containing TR UL131A 5' – UL132 5' genome region.

Virus	HFFFs			RPE-1s		
	Mean Plaque Size (μm^2)	Compared to RV-Me-UL128L ^{wt}	Compared to RV-TR	Mean Plaque Size (μm^2)	Compared to RV-Me-UL128L ^{wt}	Compared to RV-TR
RV-Me-UL128L ^{wt}	40302.375		+ 2x ($p < 0.05$)	54918.375		+ 2.5x ($p < 0.001$)
RV-TR	20798.25	- 2x ($p < 0.05$)		22072.125	- 2.5x ($p < 0.001$)	
RV-Me-UL132 ^{TR}	66365.50	+1.75x (ns)	+ 3.25x ($p < 0.01$)	63939.75	+ 1.25x (ns)	+ 3x ($p < 0.0001$)

Table A.III.4.a Comparison of mean HFFF plaque sizes formed immediately following transfection with BAC DNA for recombinant Merlin viruses containing UL128L genome regions from strain 3301 or TB40-BAC4, as well as the unique nucleotide identified in TB40-BAC4 UL128.

Virus	HFFFs				
	Mean Plaque Size (µm ²)	Compared to RV-Me-UL128L ^{mut}	Compared to RV-Me-UL128L ^{wt}	Compared to RV-Me-UL128L ³³⁰¹	Compared to RV-Me-UL128L ^{TB40}
RV-Me-UL128L ^{mut}	336104.75				
RV-Me-UL128L ^{wt}	63658.50	- 5.25x (<i>p</i> < 0.0001)			
RV-Me-UL128L ³³⁰¹	50859.50	- 6.5x (<i>p</i> < 0.0001)			
RV-Me-UL128L ^{TB40}	112035.875	- 3x (<i>p</i> < 0.0001)	+ 1.75x (<i>p</i> < 0.05)	+ 2.25x (<i>p</i> < 0.01)	
RV-Me-UL128 ^{G>T}	112027.625	- 3x (<i>p</i> < 0.0001)	+ 1.75x (<i>p</i> < 0.05)	+ 2.25x (<i>p</i> < 0.01)	+/- 1x ns

Table A.III.4.b Comparison of mean RPE-1 plaque sizes formed immediately following transfection with BAC DNA for recombinant Merlin viruses containing UL128L genome regions from strain 3301 or TB40-BAC4, as well as the unique nucleotide identified in TB40-BAC4 UL128.

Virus	RPE-1s				
	Mean Plaque Size (µm ²)	Compared to RV-Me-UL128L ^{mut}	Compared to RV-Me-UL128L ^{wt}	Compared to RV-Me-UL128L ³³⁰¹	Compared to RV-Me-UL128L ^{TB40}
RV-Me-UL128L ^{mut}	3841.625				
RV-Me-UL128L ^{wt}	109047.00	+ 28.5x (<i>p</i> < 0.0001)			
RV-Me-UL128L ³³⁰¹	46353.25	+ 24x (<i>p</i> < 0.0001)	- 1.25x (ns)		
RV-Me-UL128L ^{TB40}	50317.25	+ 12x (<i>p</i> < 0.0001)	- 2.25x (<i>p</i> < 0.05)	- 2x (<i>p</i> < 0.05)	
RV-Me-UL128 ^{G>T}	91813.125	+ 13x (<i>p</i> < 0.0001)	- 2.25x (<i>p</i> < 0.05)	- 2x (<i>p</i> < 0.05)	+/- 1x (ns)

Table A.III.5.a Comparison of mean HFFF plaque sizes formed immediately following transfection with BAC DNA for recombinant Merlin viruses containing UL128L genome regions from strain FIX, as well as the unique nucleotide identified in FIX UL130.

Virus	HFFFs			
	Mean Plaque Size (µm ²)	Compared to RV-Me-UL128L ^{mut}	Compared to RV-Me-UL128L ^{wt}	Compared to RV-Me-UL128L ^{FIX}
RV-Me-UL128L ^{mut}	336104.75			
RV-Me-UL128L ^{wt}	63658.50	- 5.25x (<i>p</i> < 0.0001)		
RV-Me-UL128L ^{FIX}	122350.375	- 2.75x (<i>p</i> < 0.05)	+ 2x (ns)	
RV-Me-UL130 ^{A>G}	128950.125	- 2.5x (<i>p</i> < 0.001)	+ 2x (<i>p</i> < 0.001)	+/- 1x (ns)

Table A.III.5.b Comparison of mean RPE-1 plaque sizes formed immediately following transfection with BAC DNA for recombinant Merlin viruses containing UL128L genome regions from strain FIX, as well as the unique nucleotide identified in FIX UL130.

Virus	RPE-1s			
	Mean Plaque Size (μm^2)	Compared to RV-Me-UL128L ^{mut}	Compared to RV-Me-UL128L ^{wt}	Compared to RV-Me-UL128L ^{FIX}
RV-Me-UL128L ^{mut}	3915.50			
RV-Me-UL128L ^{wt}	104883.625	+ 26.75x ($p < 0.0001$)		
RV-Me-UL128L ^{FIX}	26572.125	+ 6.75x ($p < 0.0001$)	-4x ($p < 0.0001$)	
RV-Me-UL130 ^{A>G}	35727.125	+ 9x ($p < 0.0001$)	-3x ($p < 0.0001$)	+ 1.25x (ns)

12 Appendix IV: Bioinformatics
analysis of TB40-BAC4 UL128
intron 1 G>T

The unique G>T identified in TB40-BAC4 UL128 lay 6 nucleotides in from the intron 1 splice acceptor site determined by mRNA mapping experiments (Akter, Cunningham et al. 2003). Given the location of this nucleotide variation, the splicing of UL128 pre-mRNA transcripts could be interrupted by this mutation. Thus, investigations into the impact of the G>T mutation commenced with the mapping of splice donor and acceptor sites in Merlin wildtype UL128, and also in Merlin UL128 containing the G>T mutation using three alternative predictive tools: SplicePort (Dogan, Getoor et al. 2007), NNSplice v.0.9 (Reese, Eeckman et al. 1997) and NetGene2 (Brunak, Engelbrecht et al. 1991). Predicted splice donor and acceptor sites are listed in Table A.IV.1.

The splice donor and acceptor sites were initially mapped in Merlin wildtype UL128L. In all analyses, the intron 1 splice donor site was predicted to reside at nucleotide position 164, corresponding to the splice donor site in established UL128 ORF interpretations. An alternative splice donor site was proposed at nucleotide position 196 using NNSplice v.0.9, however previous mRNA mapping experiments have generated no data to suggest the authenticity of this position. Various UL128 intron 1 splice acceptor sites were predicted, including nucleotide position 287 that corresponded to the splice acceptor site in current UL128 ORF interpretations. Alternative splice acceptor sites were predicted at nucleotide position 231, using SplicePort and NetGene2, at positions 320 and 327 using NetGene2 alone, and also at position 385 using NNSplice v.0.9 alone. Though like the alternative splice donor sites predicted by NNSplice 0.9, no data from UL128 mRNA mapping experiments support these proposed positions as authentic splice acceptor sites. Various candidate intron 2 splice donor and acceptor sites donor sites that were predicted by each analysis, including those determined to be authentic by mRNA mapping experiments at positions 422 and 542 respectively.

Splice donor/acceptor mapping of Merlin UL128 containing the unique G>T nucleotide substitution differed only in the prediction of intron 1 splice acceptor sites. The alternative splice acceptor sites were again proposed, whilst no analysis predicted the authentic splice acceptor site at nucleotide position 287. Therefore the unique G>T nucleotide feature identified in TB40-BAC4 UL128 intron 1 had the potential to disrupt the splice acceptor function at nucleotide position 287.

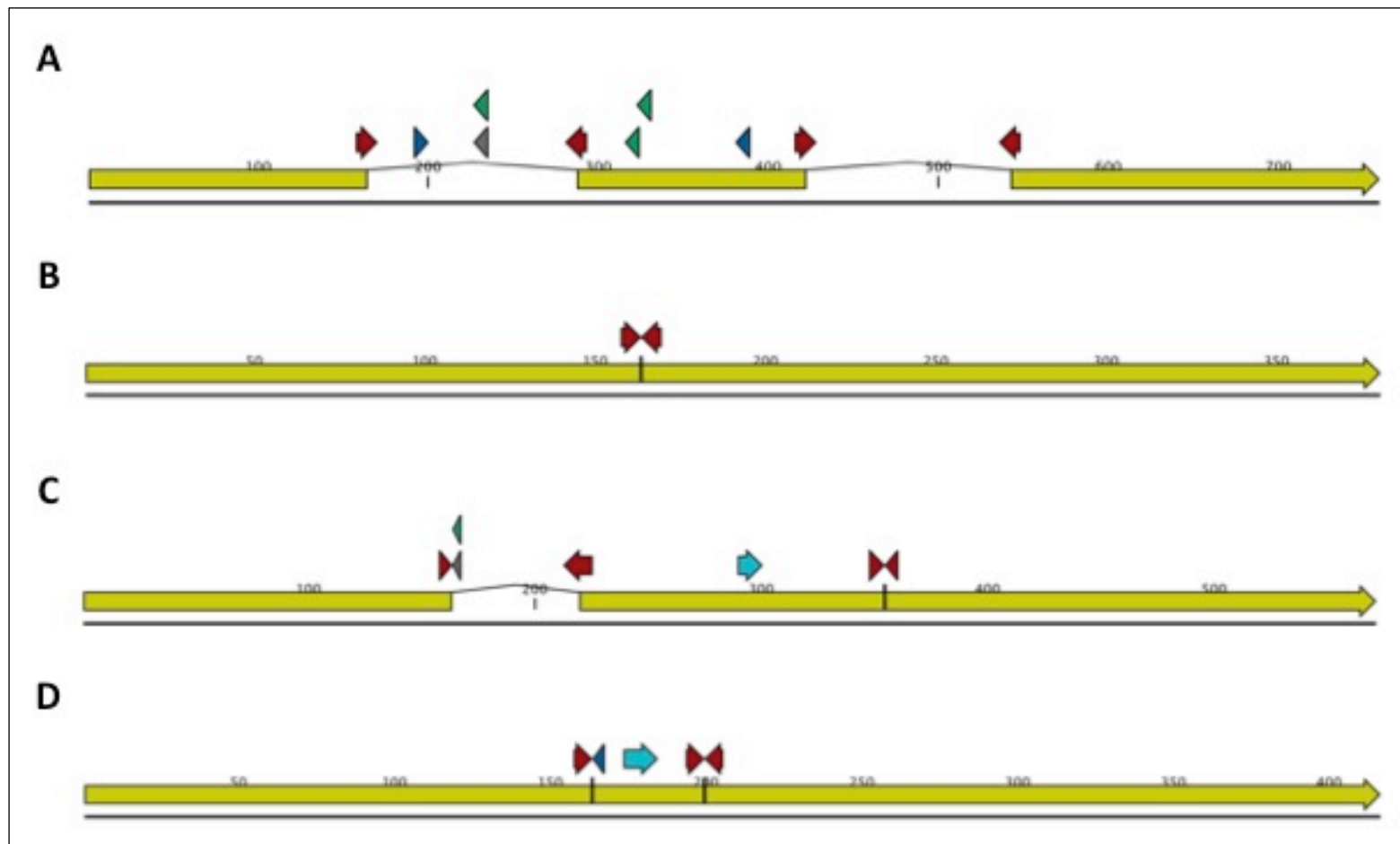


Figure A.IV.1. Potential UL128 intron 1 and intron 2 splice donor/acceptor sites (A) Splice donor/acceptor sites predicted for intron 1 and intron 2 of HCMV ORF UL128. Arrows pointing rightwards indicate splice donor sites; arrows pointing leftwards indicate splice acceptor sites. Positions highlighted in pink arrows are authentic donor/acceptor sites determined by mRNA mapping experiments; in dark blue are cryptic sites predicted by NNSplice 0.9; in grey are cryptic sites predicted by Spliceport; in green are cryptic sites predicted by NetGene. (B) Effect of exon 2 skipping with intron 1 splice donor site using intron 2 splice acceptor site. (C) Effect of activating cryptic splice site at position 21 – premature stop codon is highlighted by light blue arrow. (D) Effect of activating cryptic splice acceptor site at position 385 – premature stop codon is highlighted by light blue arrow.

Table A.IV.1. Predicted splice ‘donor’ and ‘acceptor’ sites for UL128 introns 1 and 2.

UL128 intron	Splice Donor/Acceptor Site	SplicePort	NNSplice 0.9	NetGene
Intron 1	Donor	164	164	164
			<i>196</i>	
	Acceptor	287*	287*	287*
		<i>231</i>	<i>385</i>	<i>231</i>
				<i>320</i>
				<i>327</i>
Intron 2	Donor	422	422	422
		542	542	542
	Acceptor	<i>556</i>		<i>556</i>
		<i>623</i>		<i>583</i>
				<i>594</i>
				<i>600</i>

Nucleotide positions are given relative to the sequence of the UL128 ORF in Merlin BAC-cloned genome deposited in GenBank (GU179001). In bold are splice ‘donor’ and ‘acceptor’ sites shown to be authentic by mRNA mapping experiments. In italics are novel ‘donor’ and ‘acceptor’ sites predicted in the corresponding predictive tool. * - denotes intron 1 acceptor site predicted in Merlin UL128 with wildtype sequence, though absent in Merlin UL128 containing the unique G>T nucleotide.

This loss of the intron 1 splice acceptor site at nucleotide position 287 would potentially manifest in a range of possible effects that included, but were not limited to:

- The splicing of intron 1 donor and intron 2 splice acceptor site at position 542, resulting in the skipping of exon 2. In this scenario, exon 3 was predicted to remain in frame, but the amino acids encoded by sequence in exon 2 would be absent from gpUL128.
- The use of cryptic splice acceptor site at position 231, resulting in the inclusion of 56 nucleotides from intron 1. This would cause a frame-shift in exons 2 and 3, resulting in the inclusion of 43 abnormal amino acids and introduce a premature stop codon that truncated gpUL128 after 98 amino acids.
- The use of cryptic splice acceptor site at position 385 within exon 2. This was also predicted that causes a frame shift in the remainder of exon 2 sequences, resulting in the inclusion of 5 abnormal amino acids and the truncation of gpUL128 after 60 residues.

13 Appendix V: Whole genome sequencing protocol

All virus stocks were sequenced following reconstitution from BAC DNA to ensure they had not undergone mutation during passage. All matched the original BAC sequence perfectly. Sequence data sets were generated using either a Genome Analyzer IIx or MiSeq instrument (Illumina, San Diego, CA, USA). 1000ng of DNA from each sample was sheared by sonication and DNA fragments were purified and size-selected using AMPure XP beads (Beckman-Coulter, High Wycombe, UK). Purified DNA fragments were end-repaired, 3'-adenylated, ligated to paired-end multiplexing adapters and amplified by PCR using standard methods (Illumina). Sequence read data sets were filtered to remove bases less than phred quality 30 and processed for adapter removal using Trim Galore v. 0.2.2, which is a wrapper for Cutadapt and FastQC (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore). Filtered reads were aligned against reference genomes using BWA v0.6.2-r126 (Li and Durbin 2009). Indexed Bam files were generated from the alignments using Samtools v. 0.1.18 (Li, Handsaker et al. 2009) and viewed using Tablet 1.12.12.05 (Milne, Bayer et al. 2010).

14 Appendix VI: MS- analysis of Merlin virions

Solutions

- MS solvent; 3% (V/V) MeCN, 0.1% (V/V) TFA
- In-gel destain solution; 1:1 (V/V) H₂O: acetonitrile
- In-gel reduction solution; 10 mM DTT, 100 mM NH₄HCO₃
- In-gel alkylation solution; 55 mM iodoacetamide, 100 mM NH₄HCO₃
- In-gel digestion solution; 12.5 ng/μL modified trypsin, 50 mM NH₄HCO₃
- In-gel peptide extraction solution 1; 1:1 (V/V) 50 mM NH₄HCO₃/acetonitrile
- In-gel peptide extraction solution 2; 1:1 (V/V) 5% (V/V) formic acid/acetonitrile

In-gel digestion:

- Gel pieces are destained at room temperature for 15 min with shaking at 800 rpm using In-gel destain solution. Multiple washes may be necessary to fully destain samples.
- Destained samples are equilibrated in 100 mM NH₄HCO₃, after 5 min equilibration an equal volume of acetonitrile is added and incubated at room temperature for 15 min with shaking at 800 rpm.
- Equilibration solution is removed and replaced with In-gel reduction solution. Samples are reduced at 56°C for 45 min with shaking at 800 rpm.
- In-gel reduction solution is removed and replaced with an equal volume of In-gel alkylation solution. Samples are alkylated for 30 min at room temperature in the dark with shaking at 800 rpm.
- Samples are washed twice with 100 mM NH₄HCO₃ followed by equal volumes of acetonitrile as in step 2. Following washes, gel pieces are dehydrated and dried using acetonitrile and a vacuum centrifuge.
- Sufficient In-gel digestion solution is added to cover the gel pieces and samples are incubated for 30 min on ice to allow gel pieces to rehydrate. Excess In-gel digestion solution is removed from the rehydrated gel pieces and replaced with fresh 50 mM NH₄HCO₃. Samples are digested overnight at room temperature.
- Digested peptides are extracted once with In-gel peptide extraction solution 1 and twice with In-gel peptide extraction solution 2. All extraction steps are performed at room temperature with shaking at 800 rpm. The three washes are pooled in a 0.6 mL Maxymum Recovery microtube (Axygen, #MCT-060-L-C) and dried almost to completion using a vacuum centrifuge.
- Re-suspend sample in 10 μL MS solvent.

LC-MSMS

LC-MSMS data were generated using a RSLC3000 (Dionex) coupled to a Q Exactive mass spectrometer with an EASY-spray source (Thermo). Peptides were trapped on a 5 mm PepMap 5 μ m C18 μ -Precolumn and eluted to a 50 cm, 2 μ m PepMap column (Dionex). Data were acquired in a Top 10 DDA fashion with MS scans in the Orbitrap at 70,000 (fwhm at 200 m/z), a target value of 1e6 and a maximum ion accumulation time of 250ms. The 10 most intense peaks with charges ≥ 2 were fragmented in the HCD collision cell with normalized collision energy of 28%. Tandem mass spectra were acquired in the Orbitrap at 17,500 (fwhm at 200 m/z), a target value of 5e4 and a maximum accumulation time of 200 ms. Raw files were processed using MaxQuant 1.3.0.5 and searched against a concatenated human/HCMV database. Data were searched allowing a MSMS tolerance of 20 ppm and a protein and peptide FDR of 0.01. CAM cysteine was defined as a fixed modification while oxidized methionine, acetyl (protein N-terminus), and deamidated N/Q were defined as potential variable modifications.

15 Appendix VII: SILAC-MS Raw

Data

Raw data from the SILAC-MS analysis of Merlin virion preparations described the ratios of ‘heavy’-‘medium’, ‘medium’-‘light’, and ‘heavy’-‘light’ versions of all proteins identified, according to a log2 scale. Data from the MV-P2 mixed virion preparation were corrected to accurately reflect the relative abundance of proteins in an equivalent number of different virions, by normalising the ratios of all proteins in each comparison by the reported ratios of differently labelled pUL86 (MCP). Corrected raw SILAC-MS data are included in Table A.VI.1.

Table A.VI.1 SILAC-MS raw data.

ORF	RV-Me-UL128L ^{mut} Vs. RV-Me-UL128L ^{wt}		RV-Me-UL128 ^{G>T} Vs. RV-Me-UL128L ^{wt}		RV-Me-UL128 ^{G>T} Vs. RV-Me-UL128L ^{mut}	
	<i>Ratio (Log2)</i>	<i>Fold - difference</i>	<i>Ratio (Log2)</i>	<i>Fold - difference</i>	<i>Ratio (Log2)</i>	<i>Fold - difference</i>
<i>Capsid:</i>						
UL46	- 0.0359	1.0252	- 0.0312	1.0219	- 0.0084	1.0058
UL48A	- 0.1374	1.1000	- 0.2222	1.1665	- 0.0547	1.0387
UL80	- 0.9882	1.9837	- 0.2104	1.1570	0.9148	1.8853
UL85	0.0240	1.0168	- 0.0152	1.0106	- 0.0187	1.0131
UL86	-	1.0000	-	1.0000	-	1.0000
UL104	- 0.0209	1.0146	0.0889	1.0636	0.0118	1.0082
<i>Tegument:</i>						
UL24	- 0.3143	1.2434	- 0.5243	1.4383	- 0.2025	1.1507
UL25	- 0.5177	1.4317	- 0.8973	1.8626	- 0.2202	1.1649
UL26	- 0.5018	1.4159	- 0.8334	1.7819	- 0.3335	1.2600
UL28/UL29	-	-	-	-	-	-
UL32	- 0.1200	1.0868	- 0.0748	1.0532	0.0536	1.0379
UL35	- 0.2423	1.1828	- 0.1785	1.1317	0.0544	1.0384
UL36x1	- 1.8800	3.6807	0.4680	1.3832	2.1753	4.5168
UL43	-0.4949	1.4092	-0.4798	1.3945	0.0523	1.3945
UL45	-0.0066	1.0046	-0.1019	1.0732	-0.0611	1.0432

UL47	-0.1680	1.1235	0.0432	1.0304	0.2235	1.1676
UL48	-0.2408	1.1817	0.1038	1.0746	0.3166	1.2454
UL50	-0.5637	1.4780	0.7703	1.7056	0.9588	1.9437
UL52	-1.0296	2.0414	0.1814	1.1340	1.3314	2.5165
UL69	-1.2121	2.3167	-0.0894	1.0640	0.9762	1.9673
UL71	-0.3189	1.2473	-0.0078	1.0054	0.2943	1.2263
UL82	-0.8067	1.7492	-0.1260	1.0912	0.6360	1.5540
UL83	-0.2037	1.1517	-0.5461	1.4601	-0.3500	1.2745
UL88	-0.2473	1.1870	0.1158	1.0836	0.3462	1.2712
UL94	-0.4347	1.3516	-0.3294	1.2565	0.0082	1.0057
UL96	-1.3074	2.4750	-0.4311	1.3482	0.8750	1.8340
UL97	-1.0396	2.0557	-0.6395	1.5578	0.1683	1.1237
UL99	0.0477	1.0336	-0.0535	1.0378	-0.1862	1.1378
UL103	-0.0740	1.0526	-0.1949	1.1447	0.0938	1.0672
IRS1	-1.2251	2.3377	-0.0455	1.0321	1.3804	2.6034
US22	-1.0256	2.0359	0.2853	1.2187	1.3462	2.5425
US23	-0.7245	1.6524	-0.0235	1.0164	0.4894	1.4039
US24	1.1675	2.2462	1.1049	2.1508	0.9019	1.8685
TRS1	-1.0583	2.0825	0.0589	1.0416	0.9272	1.9016

Proteins yet assigned to any virion compartment:

UL34	-0.8632	1.8190	0.5081	1.4222	1.3833	2.6087
UL44	-0.7416	1.6720	0.2268	1.1703	0.9343	1.9110
UL54	-	-	-	-	-	-
UL56	-	-	-	-	-	-
UL57	-1.0152	2.0212	0.7459	1.6770	1.6652	3.1717
UL70	-2.6732	6.3784	-0.0435	1.0306	2.5879	6.0123
UL77	0.1294	1.0938	0.0892	1.0638	-0.0613	1.0434
UL84	-1.4501	2.7323	0.1392	1.1013	1.4389	2.7111
UL89	-1.2479	2.3749	-0.4749	1.3898	0.7396	1.6698
UL93	-0.0597	1.0422	0.0922	1.0660	0.0704	1.0500
UL98	-0.9679	1.9560	0.2966	1.2282	1.2485	2.3760
UL102*	- 0.4435	1.3599	-	-	-	-
UL105	-	-	-	-	-	-

UL112-UL113	- 1.2906	2.4462	0.1602	1.1175	1.9507	3.8657
UL114*	- 0.7805	1.7178	-	-	-	-
UL122	- 0.9112	1.8806	0.3139	1.2430	1.1671	2.2456
<i>UL123</i>	-	-	-	-	-	-
UL146	2.0872	4.2493	0.1527	1.1117	- 1.7744	3.4210
UL147	- 2.7470	6.7131	- 0.4411	1.3576	2.2612	4.7938

Envelope:

RL10	0.0577	1.0408	-0.1261	1.0913	-0.1265	1.0916
<i>RL11</i>	-	-	-	-	-	-
RL12	- 0.7738	1.7098	0.4515	1.3675	1.1123	2.1619
UL16	- 1.7176	3.2888	0.6025	1.5184	2.3578	5.1259
UL33x1	- 0.6340	1.5519	-0.1656	1.1216	0.4572	1.3729
UL41A	0.3420	1.2675	0.2251	1.1688	-0.0434	1.0306
UL55	0.1167	1.0842	0.0583	1.0412	-0.0603	1.0427
UL73	0.3914	1.3117	-0.0651	1.0461	-0.2499	1.1892
UL74	1.6496	3.1374	0.8920	1.8558	-0.6439	1.5625
<i>UL74A</i>	-	-	-	-	-	-
UL75	- 0.0679	1.0482	0.1208	1.0873	0.1123	1.0810
UL78	- 1.3477	2.5451	-0.3420	1.2675	0.9781	1.9699
UL100	0.2464	1.1862	-0.0575	1.0406	-0.2729	1.2083
UL115	0.1165	1.0841	0.1609	1.1180	-0.0791	1.0564
UL116	0.1616	1.1185	0.2160	1.1615	-0.0614	1.0435
UL118/UL119	0.6278	1.5452	-0.0701	1.0498	-0.7847	1.7228
UL128	-1.7007	3.2505	-0.5242	1.4381	0.9592	1.9442
UL130	-3.1677	8.9860	-0.7144	1.6408	2.4085	5.3092
UL131A	-2.3270	5.0176	-0.3358	1.2621	2.3913	5.2463
UL132	-0.9072	1.8754	-0.0229	1.0160	0.7117	1.6377
<i>UL139</i>	-	-	-	-	-	-
UL140	0.3068	1.2370	-0.1319	1.0958	-0.4835	1.3981
UL141	-0.4235	1.3412	0.3657	1.2885	0.8060	1.7483
UL148	-2.5699	5.9377	1.4249	2.6850	3.7576	13.5251
<i>US9</i>						
US12	-0.8126	1.7564	0.3122	1.2416	0.9777	1.9694

US13	-1.2570	2.3900	0.5247	1.4387	1.6178	3.0690
US14	-0.8727	1.8311	-0.2201	1.1648	0.6077	1.5238
<i>US18</i>	-	-	-	-	-	-
US20	-0.6888	1.6120	0.4755	1.3904	1.3908	2.6223
US27	-0.3294	1.2565	-0.3922	1.3124	-0.1286	1.0932
US28	0.0539	1.0381	0.4218	1.3396	0.4487	1.3648

Uncharacterised:

<i>RL1</i>	-	-	-	-	-	-
UL31	-1.5535	2.9352	0.8880	1.8506	2.2549	4.7729
UL145	-1.8333	3.5636	0.4933	1.4077	2.2864	4.8785
UL150A*	- 0.9295	1.9046				
<i>US1</i>	-	-	-	-	-	-
<i>US26</i>	-	-	-	-	-	-

Where *log2* ratios are preceded by (-), virions contained lower levels of the corresponding protein than virions from the comparison virus.

In italics are ORFs encoding proteins identified as Merlin virion components, though not in the MVP-1 or MVP-2 mixed virion preparations.

* - Denotes ORFs whose proteins were detected in the MVP-1 mixed virion preparation, but not the MVP-2 preparation

16 Appendix VIII: Publications

Impact of Sequence Variation in the UL128 Locus on Production of Human Cytomegalovirus in Fibroblast and Epithelial Cells

Isa Murrell, Peter Tomasec, Gavin S. Wilkie, Derrick J. Dargan, Andrew J. Davison and Richard J. Stanton
J. Virol. 2013, 87(19):10489. DOI: 10.1128/JVI.01546-13.
Published Ahead of Print 24 July 2013.

Updated information and services can be found at:
<http://jvi.asm.org/content/87/19/10489>

REFERENCES

These include:

This article cites 81 articles, 51 of which can be accessed free at: <http://jvi.asm.org/content/87/19/10489#ref-list-1>

CONTENT ALERTS

Receive: RSS Feeds, eTOCs, free email alerts (when new articles cite this article), [more»](#)

Information about commercial reprint orders: <http://journals.asm.org/site/misc/reprints.xhtml>
To subscribe to to another ASM Journal go to: <http://journals.asm.org/site/subscriptions/>

Impact of Sequence Variation in the UL128 Locus on Production of Human Cytomegalovirus in Fibroblast and Epithelial Cells

Isa Murrell,^a Peter Tomasec,^a Gavin S. Wilkie,^b Derrick J. Dargan,^b Andrew J. Davison,^b Richard J. Stanton^a

Institute of Infection and Immunity, School of Medicine, Cardiff University, Cardiff, United Kingdom^a; MRC–University of Glasgow Centre for Virus Research, Glasgow, United Kingdom^b

The human cytomegalovirus (HCMV) virion envelope contains a complex consisting of glycoproteins gH and gL plus proteins encoded by the UL128 locus (UL128L): pUL128, pUL130, and pUL131A. UL128L is necessary for efficient infection of myeloid, epithelial, and endothelial cells but limits replication in fibroblasts. Consequently, disrupting mutations in UL128L are rapidly selected when clinical isolates are cultured in fibroblasts. In contrast, bacterial artificial chromosome (BAC)-cloned strains TB40-BAC4, FIX, and TR do not contain overt disruptions in UL128L, yet no virus reconstituted from them has been reported to acquire mutations in UL128L *in vitro*. We performed BAC mutagenesis and reconstitution experiments to test the hypothesis that these strains contain subtle mutations in UL128L that were acquired during passage prior to BAC cloning. Compared to strain Merlin containing wild-type UL128L, all three strains produced higher yields of cell-free virus. Moreover, TB40-BAC4 and FIX spread cell to cell more rapidly than wild-type Merlin in fibroblasts but more slowly in epithelial cells. The differential growth properties of TB40-BAC4 and FIX (but not TR) were mapped to single-nucleotide substitutions in UL128L. The substitution in TB40-BAC4 reduced the splicing efficiency of UL128, and that in FIX resulted in an amino acid substitution in UL130. Introduction of these substitutions into Merlin dramatically increased yields of cell-free virus and increased cell-to-cell spread in fibroblasts but reduced the abundance of pUL128 in the virion and the efficiency of epithelial cell infection. These substitutions appear to represent mutations in UL128L that permit virus to be propagated in fibroblasts while retaining epithelial cell tropism.

Human cytomegalovirus (HCMV) is ubiquitous throughout populations worldwide and represents a significant public health challenge in both developed and developing countries (1). Like other herpesviruses, HCMV establishes life-long persistent infections with periodic episodes of reactivation that require constant immunosurveillance. Productive infection is commonly asymptomatic in an immunocompetent host, but the virus remains a leading infectious cause of congenital malformation and is responsible for a broad spectrum of pathological consequences in immunocompromised (e.g., AIDS patients) or immunosuppressed individuals (e.g., transplant recipients). HCMV displays tropism for a broad range of cell types and tissues *in vivo*, with disease being associated with most major organs. For example, initial infection occurs in mucosal epithelial tissues, infection of endothelial tissues can result in transfer of virus to solid organs and leukocytes for dissemination, and CD34⁺ bone marrow progenitor cells and monocytes are sites of latency (1–3). Thus, studies of HCMV pathogenesis require the use of virus with the ability to infect a wide range of cell types *in vitro*. However, there are significant challenges associated with *in vitro* propagation of virus that exhibits the broad tropism characteristic of clinical virus.

Classically, three distinct virion envelope glycoprotein complexes, designated gCI, gCII, and gCIII, have been implicated in recognition and uptake of HCMV by the cell. gCI, composed of glycoprotein gB, and gCII, composed of glycoproteins gM and gN, mediate the initial attachment of virions. Both are capable of binding to heparin sulfate proteoglycans (4–7), while gB is also capable of binding to integrins, epidermal growth factor receptor (EGFR), and platelet-derived growth factor receptor α (PDGFR α). All of these molecules have been reported to be important for virus entry (8–11), although the roles of EGFR and PDGFR α have been disputed (12, 13). Following initial binding, fusion with cellular membranes is orchestrated by gB and gCIII, which is formed of

glycoproteins gH, gL, and gO (14–18). More recently it has become apparent that gH and gL also form a second glycoprotein complex, and that infection of different cell types occurs by different mechanisms involving these two different complexes.

Infection of fibroblasts occurs by direct fusion of the virion envelope with the plasma membrane, whereas in epithelial, endothelial, and myeloid cells, membrane fusion takes place in vesicles following internalization by endocytosis or micropinocytosis (19–22). gH/gL/gO is required for infection, virion maturation, egress, and cell-to-cell spread in fibroblasts, as well as for infection of epithelial and endothelial cells (6, 23, 24). A second complex, gH/gL/UL128L, is formed by gH/gL along with the products of the UL128 locus (UL128L), pUL128, pUL130, and pUL131A. gH/gL/UL128L is required for efficient infection and cell-to-cell spread in epithelial, endothelial, and myeloid cells (22, 25–35), either by binding to cell surface receptors (22, 27, 28) or by promoting nuclear translocation of virions (21, 32, 36). Infection of fibroblasts does not require gH/gL/UL128L; in fact, virus containing gH/gL/UL128L displays reduced cell-to-cell spread and cell-free release in fibroblasts *in vitro* (37, 38). As a result, there is considerable selection pressure against UL128L in this cell type. Thus, routine isolation of HCMV strains from clinical material in fibroblasts is associated with rapid acquisition of disabling mutations in UL128L, which are usually apparent as frameshifts caused by

Received 11 June 2013 Accepted 17 July 2013

Published ahead of print 24 July 2013

Address correspondence to Richard J. Stanton, StantonRJ@cf.ac.uk.

Copyright © 2013 Murrell et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 Unported license.

doi:10.1128/JVI.01546-13

insertion or deletion of one or more nucleotides, in-frame termination codons caused by single-nucleotide substitutions, or deletions (37–44). This results in the generation of laboratory-adapted viruses that display efficient growth in fibroblasts but limited growth in other cell types.

To provide a genetically stable source of HCMV, the genome can be cloned into a bacterial artificial chromosome (BAC) and virus recovered by transfection (38, 45–47). However, HCMV is invariably subjected to some degree of passaging *in vitro* prior to BAC cloning, and as a result, BAC-cloned strains exhibit various degrees of *in vitro* adaptation. We have previously described the cloning of the complete HCMV strain Merlin genome into a self-excising BAC following five passages in fibroblasts (38). *In vitro*-acquired mutations were identified by reference to the original clinical material and repaired, with the resulting BAC having the genetic competence of wild-type virus. However, as with clinical HCMV strains, the presence of wild-type UL128L in reconstituted virus results in the production of very low titers of cell-free virus *in vitro*, and the virus is prone to mutation when passaged in fibroblasts (38).

Several HCMV strains in addition to Merlin have been BAC cloned. These include TB40-BAC4, which was cloned from a mixed population of TB40/E following five passages in fibroblasts and 22 passages in endothelial cells (48), FIX, which was derived from strain VR1814 following 46 passages in fibroblasts (49, 50), and TR, which was isolated from an ocular swab from an AIDS patient (51, 52). Strains TB40-BAC4, FIX, and TR were cloned by replacing sequences in one region of the genome (at the left end of U_L) with a nonexcising BAC vector. UL128L in each BAC is apparently intact, and reconstituted virus is able to infect epithelial and endothelial cells. However, unlike strains isolated from clinical material, there are no reports of viruses reconstituted from these BACs acquiring mutations in UL128L during passage *in vitro*. This raised the hypothesis that these viruses contain subtle mutations in UL128L that were acquired during passage of the clinical isolates prior to BAC cloning. Indeed, our study identifies single-nucleotide substitutions in strains TB40-BAC4 and FIX that impact UL128 and UL130, respectively. Introduction of these substitutions into wild-type UL128L in strain Merlin dramatically increased yields of cell-free virus and increased cell-cell spread in fibroblasts but reduced both the abundance of pUL128 in the virion and the efficiency of infection in epithelial cells.

MATERIALS AND METHODS

Cells and viruses. Primary human fetal foreskin fibroblast (HFFF) cells and human telomerase reverse transcriptase (hTERT)-immortalized retinal pigmented epithelial (RPE-1) cells were grown in Dulbecco's modified Eagle medium (DMEM) (Life Technologies) supplemented with fetal bovine serum (10%, vol/vol), penicillin (500 U/ml), and streptomycin (500 µg/ml) at 37°C in 5% CO₂. Two variants of the Merlin BAC were used and have been described previously (38). Merlin-UL128L^{mut} (previously called pAL1158) contains a premature stop codon in UL128, while Merlin-UL128L^{wt} (previously called pAL1160) contains wild-type UL128. Both BACs contain a frameshift in RL13 and an internal ribosomal entry site (IRES) followed by enhanced green fluorescent protein (EGFP) after UL122. TB40-BAC4 was kindly donated by Christian Sinzger (48), FIX by Gabi Hahn (50), and TR by Jay Nelson (51). For these BACs, recombinering was used to insert an IRES followed by EGFP after UL122 as described previously (38). HCMV strain 3301 DNA had been extracted previously from the urine of a congenitally infected infant (40).

Infections were performed at 37°C for 2 h on a rocker, followed by

removal of the inoculum and addition of fresh medium (53). In titrations and assays to investigate cell-to-cell spread, supernatant-driven spread was limited by use of a 1% Avicel semisolid overlay (54). After incubation for 14 days (for fibroblasts) or 21 days (for epithelial cells), the overlay was removed and cells were washed in phosphate-buffered saline (PBS). Unless otherwise stated, quantitation of cell-free virus produced during infection of both fibroblast and epithelial cells was performed using HFFF, since this is the only cell type in which all viruses could infect and spread. Plaques were identified based on EGFP expression and imaged using an ORCA-ER camera and Leica DMIRBE microscope. Plaque sizes were determined using OpenLab 3 software.

Preparation of BACs. Stocks of each BAC-cloned genome were prepared using a Nucleobond plasmid purification kit (Macherey-Nagel) according to the manufacturer's instructions. The concentration of purified plasmid DNA was determined by use of an ND1000 spectrophotometer (Nanodrop).

Transfections. BACs were transfected into HFFF cells by electroporation using a Nucleofector (Amaxa) and basic fibroblast kit (Lonza) and program T-16 according to the manufacturer's instructions. On occasions when the monolayer formed was less than 70% confluent on the following day, additional cells were added. The number of plaques formed following electroporation varied only marginally, ranging from 20 to 35 per transfection. RPE-1 cells were transfected using Effectene (Qiagen) according to the manufacturer's instructions. The number of plaques formed varied only minimally, with 50 to 70 plaques generated per transfection. These variations were not observed to have a major impact on the relative rate of virus spread through the monolayer in experimental repeats.

FACS analysis. At weekly time points posttransfection, infected RPE-1 cultures were trypsinized and reseeded into fresh flasks. An aliquot of cells was kept for fluorescence-activated cell sorter (FACS) analysis using an Accuri C6 and CFlow software for the detection of EGFP⁺ cells.

Recombineering. All recombineering was performed as described previously (38, 55, 56), using *E. coli* SW102 cells containing the BAC to be modified. A selectable *amp/sacB/lacZ* cassette was PCR amplified and inserted into the region to be modified, followed by positive selection for expression of ampicillin resistance on medium supplemented with ampicillin (50 µg/ml). In a second round of recombineering, the selection cassette was swapped with the DNA sequence to be inserted, followed by negative selection on medium supplemented with sucrose (5%, wt/vol) to select against *sacB* expression and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) and isopropyl β-D-1-thiogalactopyranoside (IPTG) to identify white colonies lacking *lacZ* expression. Amplification of the selectable cassette was performed using the Expand HiFi system (Roche) under the following conditions: 95°C for 2 min; 10 cycles at 95°C for 30 s, 55°C for 30 s, and 68°C for 4.5 min; 25 cycles at 95°C for 30 s, 55°C for 30 s, and 68°C for 4.5 min; and 68°C for 15 min. Primer pairs were designed with approximately 20 bp of identity to the selectable cassette at each 3' end and approximately 80 bp of identity to sequences adjacent to the insertion site at the 5' end. In the primer sequences shown below, regions identical to sequences immediately up- and downstream from the insertion site are underlined. Primers were designed to cover regions with 100% identity in all strains.

Insertion of UL128L sequences into the Merlin genome. For insertion of the complete UL128L from strains TR, TB40-BAC4, FIX, and 3301 in place of the wild-type Merlin UL128L, the *amp/sacB/lacZ* cassette was amplified using primers SacBR-131A (CAG TCT GCA ACA TGC GGC TGT GTC GGG TGT GGC TGT CTG TTT GTC TGT GCG CCG TGG TGC TGG GTC AGT GCC AGC GGG ACT GAG GTT CTT ATG GCT CTT G) and SacBF-128 (ATC CAG CCG TTT GTG TTT CTT AAC GCT CTC CAG GTA CTG ATC CAG GCC CAC GAT CCG GGT TAT CTT GTC GTA TTC CAG CCT GTG ACG GAA GAT CAC TTC G). UL128L from each strain was amplified using a Phusion high-fidelity kit (New England BioLabs) and primers UL128LF (GCG TAT TTC GGA CAA ACA CAC A) and UL128LR (CGC ATG TTG CAG ACT GAG AAA GA). PCR

was performed under the following conditions: 98°C for 1 min; 35 cycles at 98°C for 30 s, 55°C for 30 s, and 72°C for 1 min; and 72°C for 10 min.

Insertion of the unique TB40-BAC4 nucleotide into the Merlin genome. The *amp/sacB/lacZ* selection cassette was amplified using primers SacBF (GGT GGT GAC GAT CCC GCG AAT CTC AGC CGT TTT CTC GGG ACT GTA GCA GAC TTC GCC GTC CGG ACA CCG CAG CCT GTG CCT GTG ACG GAA GAT CAC TTC G) and SacBR (CTG GAT CTG TCT CTC GAC GTT TCT GAT AGC CAT GTT CCA TCG ACG ATC CTC GGG AAT GCC AGA GTA GAT TTT CAT GAA TCT GAG GTT CTT ATG GCT CTT G). The following oligonucleotide was used to insert the TB40-BAC4 UL128 nucleotide (underlined): GTT TTC TCG GGA CTG TAG CAG ACT TCG CCG TCC GGA CAC CGC AGC CTG TTG ATT CAT GAA AAT CTA CTC TGG CAT TCC CGA GGA TCG TCG ATG GAA CAT G.

Insertion of unique FIX UL130 nucleotide into the Merlin genome. The *amp/sacB/lacZ* selection cassette was amplified using primers SacBF-FIX (GTC TGG CCT TCC CGG TTG TAC AGC AGA TAC AGG GTC TCG TTG CGA CAC TCG GGA CCC GTT GAT ACC CGC TGG AAC CCC CCT GTG ACG GAA GAT CAC TTC G) and SacBR-FIX (TAT TCC AAA CCG CAT GAC GCG GCG ACG TTT TAC TGT CCT TTT CTC TAT CCC TCG CCC CCA CGA TCC CCC TTG CAA TTC CCT GAG GTT CTT ATG GCT CTT G). The following oligonucleotide was used to insert the FIX UL130 nucleotide (underlined): AGG GTC TCG TTG CGA CAC TCG GGA CCC GTT GAT ACC CGC TGG AAC CCC GAT AAT TGC AAG GGG GAT CGT GGG GGC GAG GGA TAG AGA AAA GGA CAG TAA A.

Sanger DNA sequencing. UL128L was amplified by PCR using primers UL128LF (GCG TAT TTC GGA CAA ACA CAC A) and UL128LR (CGC ATG TTG CAG ACT GAG AAA GA). UL128L was amplified from all viruses during time courses (e.g., from infected cell culture supernatants) to determine whether it retained its original sequence. The Advantage II PCR system (Clontech) was used according to the manufacturer's instructions under the following conditions: 95°C for 10 min; 35 cycles at 95°C for 30 s, 55°C for 30 s, and 68°C for 3 min; and 68°C for 10 min. When UL128L was amplified from a BAC, the HiFi expand PCR system (Roche) was used according to the manufacturer's instructions under the following conditions: 94°C for 2 min; 30 cycles at 94°C for 15 s, 68°C for 30 s, and 72°C for 2 min; and 72°C for 7 min. PCR products were purified from agarose gels by using an Illustra GFX PCR DNA gel band purification kit (GE Healthcare) and then sequenced (EurofinsMWG) using the primers CGG ATT GTA GTT GCA GCT CG, TCT GGT TAT TGG CCT CGG TG, TCT TCC AAT ATC GCC ATC TC, and GCG CAC AGA AGC AGG CAG. Sequence alignments, BLAST searches, and other analyses were performed by using CLC MAIN 6 software.

Generation of cDNA libraries of UL128 transcripts. Total cell RNA samples from HFFF cultures infected at a multiplicity of infection (MOI) of 3 were extracted 72 h postinfection by using an RNeasy plus universal kit (Qiagen) according to the manufacturer's instructions. cDNAs were generated using a NanoScript Precision reverse transcriptase kit (Primer Design). cDNAs were then amplified using primers UL128F (CATAAAC GTCAACACCCC) and UL128R (CACTGCAGCATATAGCCC) with the HiFi expand PCR system (Roche) under the conditions stated above.

Bioinformatics analysis. Analysis of UL128 splice sites was carried out by using NNsplice v. 0.9 (57), NetGene2 (58), SplicePort (59), and Human Splicing Finder v. 2.4.1 (60). Protein structure prediction was performed using Phyre 2 (61).

Preparation of virus stocks. Where feasible, virus stocks were generated in RPE-1 cells to prevent the selection of mutations in UL128L. Merlin strains mutated in UL128L and FIX were grown in HFFF cells owing to their inability to spread productively in RPE-1 cells. Supernatant from infected cell cultures was collected and centrifuged at $470 \times g$ for 5 min at room temperature to pellet cellular debris. Cleared supernatants were centrifuged at $30,000 \times g$ for 2 h at 20°C. Pelleted virions were resuspended in DMEM containing 10% fetal bovine serum (FBS) and stored at -80°C .

Gradient purification of virions. Virions were purified from noninfectious enveloped particles, dense bodies, and cellular debris by ultracentrifugation through glycerol-tartrate positive-density, negative-viscosity gradients, as described previously (62, 63). The gradients were centrifuged at $90,465.7 \times g$ for 45 min at 20°C. Virions were recovered by piercing the tubes using a 20-gauge needle and syringe, diluted in 0.04 M sodium-phosphate buffer (pH 7.4), and pelleted by centrifugation at $90,465.7 \times g$ for 1 h at 20°C. The virion pellet was resuspended in a medium suitable for downstream applications (see below).

Western blot analysis. Purified virions were resuspended in Nu-Page LDS sample buffer, and proteins were separated by using a Nu-page Tris-acetate gel system (Invitrogen) according to the manufacturer's instructions. Electrophoresed proteins were subsequently transferred to Hybond-P polyvinylidene difluoride (PVDF) membranes (GE Healthcare) by semidry transfer at 10 V for 1 h using carbonate transfer buffer. The membranes were incubated for 1 h at room temperature in blocking buffer (PBS containing 0.1% Tween 20 [PBST] plus 5% [wt/vol] fat-free milk). They were then incubated with primary antibody for 1 h at room temperature, washed three times with PBST, and incubated with secondary antibody for 1 h at room temperature. Antibody was detected by SuperSignal West Pico (Thermo) using an AutoChem imaging system and Labworks software (UVP Bioimaging). Primary antibodies were mouse-anti pUL128 antibody SURN, provided by Giuseppe Gerna (1:100), and mouse anti-gB (1:4,000; Abcam). The secondary antibody was goat anti-mouse horseradish peroxidase (HRP; 1:1,000; GE Healthcare).

RESULTS

Growth properties of HCMV strains FIX, Merlin, TB40-BAC4, and TR in fibroblast cells. Clinical isolates of HCMV consistently replicate inefficiently in fibroblast cultures until mutations arise and are selected, first in RL13 and then in UL128L (37, 38). A defect in any of UL128, UL130, or UL131A suppresses formation of the pentameric glycoprotein complex (gH/gL/UL128L) in the virion (37). To establish the growth characteristics of virus lacking or containing gH/gL/UL128L, viruses derived from two Merlin BAC variants were used in this study. Merlin-UL128L^{mut} contains a premature stop codon mutation in UL128 that was selected during growth of Merlin in fibroblasts prior to BAC cloning. This mutation was repaired in Merlin-UL128L^{wt}. In addition, because the FIX and TB40-BAC4 clones contain mutations in RL13 (38), which invariably mutates upon passage in fibroblast or epithelial cells (37, 38), a preexisting RL13 frameshift mutation present in these Merlin BACs was not repaired. TR contains an intact RL13, which could contribute to the growth characteristics of this virus. However, given the speed with which mutants are consistently selected in RL13 when clinical virus is grown *in vitro* (37, 38), it seems likely that RL13 is either not expressed or is nonfunctional in virus derived from TR.

Interestingly, all three protein-coding regions in UL128L appear to be intact in the TB40-BAC4, FIX, and TR BACs, yet reconstituted viruses have not been reported to acquire obvious mutations during passage in fibroblasts. It is possible that these three BACs contain natural variants of UL128L that are stable in fibroblasts, or that subtle UL128L mutations have been selected that preserve the integrity of the protein-coding regions but suppress their functions. Unfortunately, the clinical samples from which these strains were derived are not available for comparison. Therefore, we compared the growth properties of TB40-BAC4, FIX, and TR to those of Merlin-UL128L^{wt} and Merlin-UL128L^{mut}. To monitor infection, all viruses were engineered to express EGFP from an IRES inserted downstream from UL122 (encoding IE2) (Table 1).

TABLE 1 BAC-cloned and recombinant HCMV strains used in this study

Strain	Reference	UL128L origin ^a	GenBank accession no.	Designation in text
BAC cloned				
Merlin ^b	38	Mutated: G>A in UL128 at nt 176260 (R>stop)	GU179001.1	Merlin-UL128L ^{mut}
Merlin ^b	38	Native	GU179001.1	Merlin-UL128L ^{wt}
TR ^b	51	Native	AC146906.1	TR
TB40-BAC4	48	Native	EF999921.1	TB40-BAC4
FIX ^b	50	Native	AC146907.1	FIX
Clinical sample (nonpassaged)				
3301	40	Native	GQ466044.1	3301
Recombinant Merlin containing variant UL128L				
Merlin		3301		Merlin-UL128L ³³⁰¹
Merlin		TR		Merlin-UL128L ^{TR}
Merlin		TB40-BAC4		Merlin-UL128L ^{TB40}
Merlin		FIX		Merlin-UL128L ^{FIX}
Recombinant Merlin strains containing unique substitutions in UL128L				
Merlin		TB40-BAC4: G>T in UL128 at nt 176663 (near splice acceptor site)		Merlin-UL128 ^{G>T}
Merlin		FIX: A>G in UL130 at nt 177364 (S72P)		Merlin-UL130 ^{A>G}

^a Nucleotide positions are relative to the sequence of BAC-cloned HCMV strain Merlin (GU179001).
^b GenBank accession numbers for parental viruses are NC_006273 (Merlin), GU179289 (VR1814; parental virus of FIX), and KF021605 (TR).

UL128L mutants can arise within a single passage of wild-type HCMV in fibroblasts (43). To ensure that the experiments were initiated using genetically homogenous virus preparations, HFFF cells were transfected with infectious BAC clones of each virus. The capacity of virus infection to progress by cell-to-cell spread was assessed by direct measurement of plaque sizes formed under a semisolid overlay (Fig. 1A). In this assay, Merlin-UL128L^{mut} consistently generated the largest plaques, whereas TB40-BAC4 and FIX plaques were smaller than those of Merlin-UL128L^{mut} but were 4- to 2.5-fold larger, respectively, than those of Merlin-UL128L^{wt}. TR produced the smallest plaques of all viruses tested.

Consistent with previous work (38), Merlin-UL128L^{mut} spread through the HFFF monolayer the fastest (Fig. 1B) and produced the greatest amounts of cell-free virus (Fig. 1C), whereas Merlin-UL128L^{wt} spread the slowest and produced the lowest cell-free titers (approximately 1,000-fold less than Merlin-UL128L^{mut}). Compared to Merlin-UL128L^{wt}, TB40-BAC4 and FIX each spread through the HFFF monolayer faster and produced much higher yields (1,000- and 50-fold, respectively) of cell-free virus. Indeed, the amount of cell-free TB40-BAC4 release was similar to that of Merlin-UL128L^{mut}. The TR infections spread very slowly at first, but the rate increased from week 5. Like TB40-BAC4 and FIX, TR also produced peak titers of cell-free virus that exceeded those of Merlin-UL128L^{wt} by more than 100-fold.

Thus, TB40-BAC4 and FIX displayed more efficient cell-to-cell spread in fibroblasts and produced greater yields of cell-free virus than Merlin containing wild-type UL128L. Cell-to-cell spread of TR in fibroblasts was reduced compared to that of Merlin containing wild-type UL128L, yet peak cell-free titers were significantly higher.

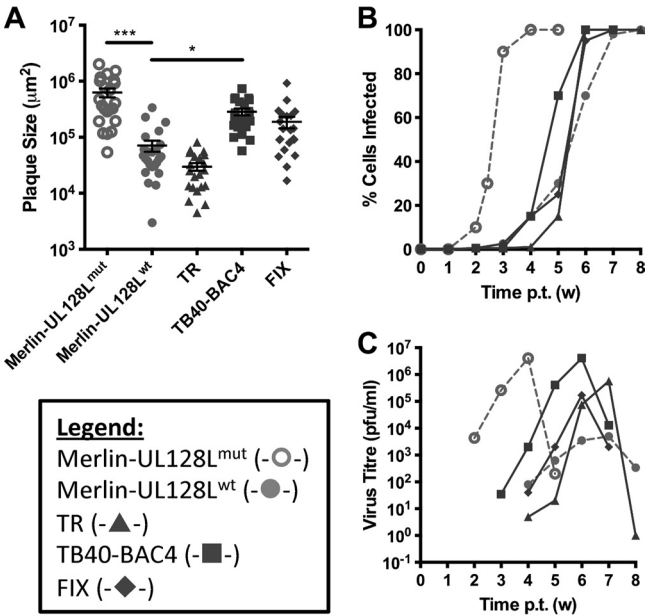


FIG 1 Growth characteristics of BAC-cloned strains in fibroblasts. (A) HFFF cells were transfected with BAC DNA for the indicated viruses and then placed under semisolid overlay. Plaque sizes were measured 2 weeks later. Means and standard deviations are shown. (B) HFFF cells were transfected with BAC DNA for the indicated viruses, and infection was allowed to progress until the monolayer was destroyed. At weekly time points posttransfection (p.t.), the level of infection was estimated by FACS analysis of EGFP-expressing cells. (C) Supernatants from the infections shown in panel B were retained at weekly intervals and titrated on HFFFs to provide a measure of cell-free virus release. (Where indicated, samples were compared by 1-way analysis of variance [ANOVA] followed by Dunnett's posttest to compare each sample to Merlin-UL128L^{wt}. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$).

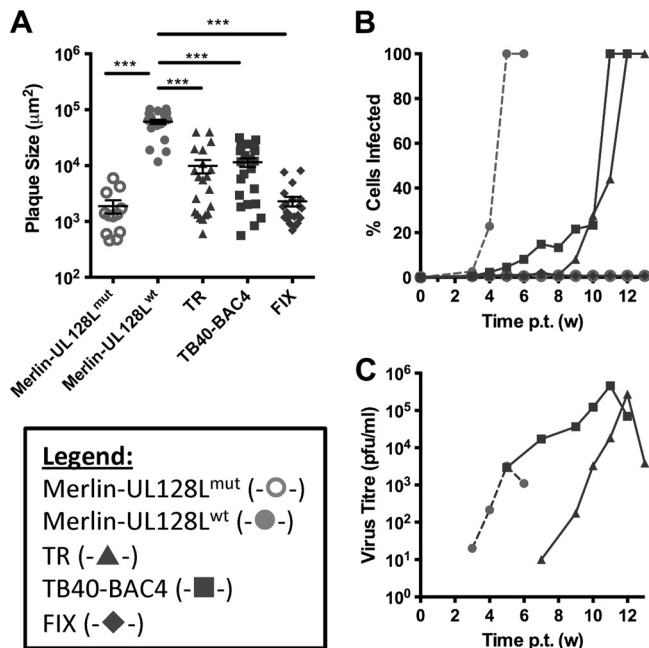


FIG 2 Growth characteristics of BAC-cloned strains in epithelial cells. (A) RPE-1 cells were transfected with BAC DNA for the indicated viruses and then placed under semisolid overlay. Plaque sizes were measured 3 weeks later. Means and standard deviations are shown. (B) RPE-1 cells were transfected with BAC DNA for the indicated viruses, and infection was allowed to progress until the monolayer was destroyed. At weekly time points, cells were trypsinized and the level of infection was measured by FACS analysis of EGFP-expressing cells. (C) Supernatants from the infections shown in panel B were retained at weekly intervals and titrated on HFFs to provide a measure of cell-free virus release. (Where indicated, samples were compared by 1-way ANOVA followed by Dunnett's posttest to compare each sample to Merlin-UL128L^{wt}. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$).

Growth properties of HCMV strains FIX, Merlin, TB40-BAC4, and TR in epithelial cells. In order to investigate the growth characteristics of FIX, Merlin, TB40-BAC4, and TR in an epithelial cell line (which requires UL128L for efficient infection), RPE-1 cells were transfected with the infectious BAC clones. In marked contrast to the findings for fibroblasts, Merlin-UL128L^{wt} displayed much more efficient plaque formation than Merlin-UL128L^{mut} (Fig. 2A). TB40-BAC4 and TR plaques were of an intermediate size (approximately 6 times smaller than Merlin-UL128L^{wt}), while FIX formed plaques that were 25 times smaller, similar to those produced by Merlin-UL128L^{mut}. TB40-BAC4 and TR each spread throughout the RPE-1 monolayer much more slowly than Merlin-UL128L^{wt} (Fig. 2B and C) yet produced 100- to 150-fold higher titers of cell-free virus. Neither FIX nor Merlin-UL128L^{mut} supported significant spread through RPE-1 monolayers.

Thus, the growth phenotypes displayed by TB40-BAC4, FIX, and TR in fibroblasts and epithelial cells were distinct from those of Merlin containing wild-type UL128L. Cell-to-cell spread of TB40-BAC4 and FIX was more efficient in fibroblasts but less efficient (along with TR) in epithelial cells. TB40-BAC4, TR, and FIX all produced significantly higher cell-free titers than Merlin-UL128L^{wt} in both cell types.

The differential growth characteristics of TB40-BAC4 and FIX are determined by UL128L. To assess whether the growth

characteristics of each strain could be attributed to UL128L, we replaced UL128L in Merlin with that from TB40-BAC4, FIX, or TR, thereby generating the recombinant viruses Merlin-UL128L^{TB40}, Merlin-UL128L^{FIX}, and Merlin-UL128L^{TR}, respectively (Table 1). Compared to Merlin-UL128L^{wt}, acquisition of TB40-BAC4 or FIX UL128L was consistently associated with a 2.5- to 3-fold increased plaque size in fibroblasts (Fig. 3A), an increased rate of cell-free spread (Fig. 3B), and approximately 10-fold increased yields of cell-free virus (Fig. 3C). In contrast, TR UL128L did not alter the rate of cell-to-cell spread of strain Merlin and produced titers of cell-free virus that were comparable to those of Merlin-UL128L^{wt}, at least until week 6 posttransfection. However, production of cell-free virus by Merlin-UL128L^{TR} increased dramatically by week 7. DNA sequencing of UL128L from cell-free virus during the final time point revealed that all viruses retained their original sequence in UL128L, with the exception of Merlin-UL128L^{TR}. Merlin-UL128L^{TR} harbored a deletion (nt 176742 to 177247) affecting both UL130 and UL128, as well as a G-to-A mutation at nt 177355 that resulted in a premature stop codon in UL130. Thus, the rapid increase in cell-free virus production by Merlin-UL128L^{TR} was correlated with mutation of UL128L.

Measurement of the size of plaques formed by these recombinant viruses in RPE-1 cells (Fig. 3D) revealed that expression of TB40-BAC4 and FIX UL128L reduced the epithelial cell-to-cell spread of Merlin, with plaques 2.5 to 7 times smaller, respectively. However, TR UL128L did not affect the growth characteristics of Merlin. Likewise, time course experiments in epithelial cells showed that transfer of FIX UL128L resulted in reduced speed of dissemination, and transfer of FIX and TB40-BAC4, but not TR, UL128L resulted in 10- to 20-fold increased production of cell-free virus (Fig. 3E and F). Thus, compared to Merlin-UL128L^{wt}, the ability of TB40-BAC4 and FIX to produce larger plaques in fibroblasts but smaller plaques in epithelial cells, as well as to produce higher titers of cell-free virus in both cell types, appears to be attributable at least in part to genetic differences in UL128L. However, the differential growth characteristics between TR and Merlin-UL128L^{wt} appear to be independent of UL128L.

A single-nucleotide difference in a UL128 intron increases cell-free virus production. TB40-BAC4 infection produced significantly higher titers of cell-free virus than Merlin-UL128L^{wt} in both epithelial and fibroblast cells. Since UL128L clearly makes a major contribution to the differential growth characteristics of TB40-BAC4, we hypothesized that UL128L had acquired a mutation during passage that is compatible with growth in fibroblasts yet permits the virus to retain a degree of epithelial cell tropism. If this were the case, any such mutation would be expected to be unique to TB40-BAC4. UL128L sequences from different strains exhibit a high level of sequence conservation (>92.3% at the nucleotide level [64, 65]); however, BLAST searches revealed that TB40-BAC4 had a unique G-to-T (G>T) substitution at nt 176612 (with reference to the equivalent location in the Merlin BAC genome). The G residue is conserved in 50 other HCMV strains and is located in the first of the two UL128 introns, 6 nt from the splice acceptor site. To determine whether this substitution affects the growth properties of TB40-BAC4, it was introduced into Merlin-UL128L^{wt} to generate Merlin-UL128L^{G>T} (Table 1). As an additional control, UL128L from strain 3301 was inserted into Merlin (generating Merlin-UL128L³³⁰¹). Strain 3301 had not been subjected to *in vitro* passage prior to sequencing and

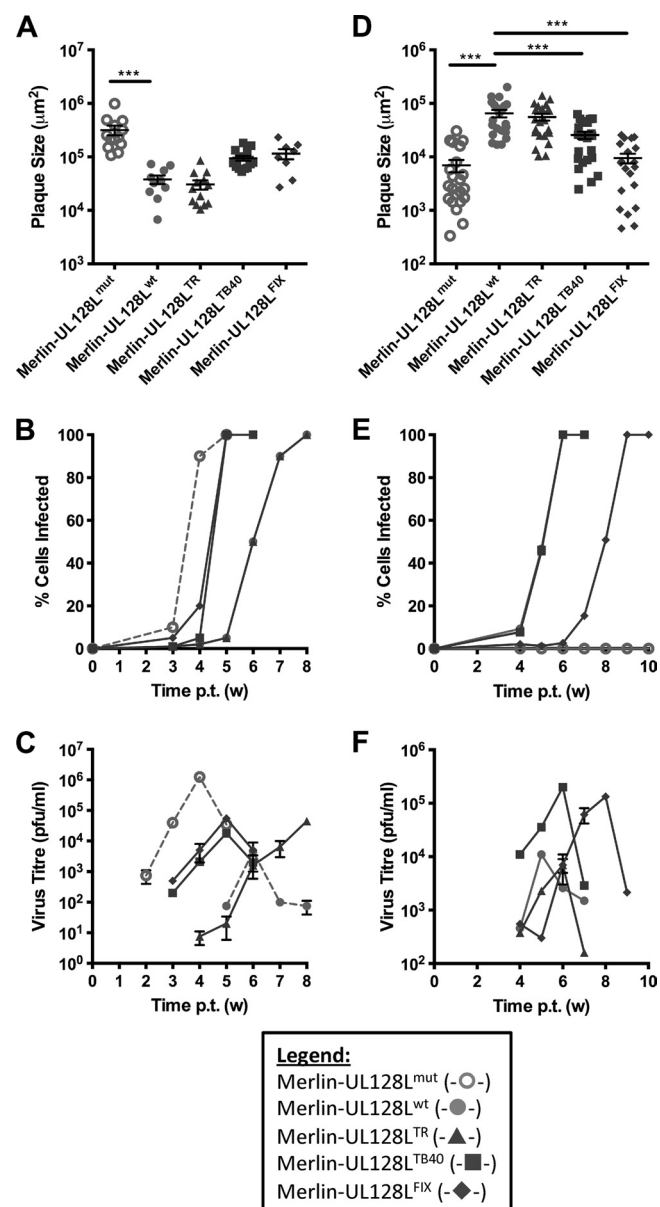


FIG 3 Growth characteristics of recombinant Merlin viruses containing UL128L from other strains. HFFF (A) or RPE-1 (D) cells were transfected with BAC DNA for the indicated viruses and then placed under semisolid overlay. Plaque sizes were measured 2 (A) or 3 (D) weeks later. Means and standard deviations are shown. HFFF (B) or RPE-1 (E) cells were transfected with BAC DNA for the indicated viruses, and infection was allowed to progress until the monolayer was destroyed. At weekly time points, the level of infection was estimated from EGFP expression (B), or cells were trypsinized and the level of infection was measured by FACS analysis of EGFP-expressing cells (E). (C and F) Supernatants from the infections shown in panels B and E were retained at weekly intervals and titrated on HFFFs to provide a measure of cell-free virus release. (Where indicated, samples were compared by 1-way ANOVA followed by Dunnett's posttest to compare each sample to Merlin-UL128L^{wt}. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.)

was thus representative of a UL128L sequence that was different from that of Merlin but that was known to be wild type. UL128L from 3301 was also highly homologous to UL128L from TB40-BAC4 (99.49% at the nucleotide level); however, it did not contain the G>T nucleotide difference identified in the intron of UL128L.

Therefore, it was useful to determine whether the G>T substitution, as opposed to natural strain variation elsewhere, contributed to the different characteristics of UL128L from TB40-BAC4 compared to that from Merlin.

Merlin-UL128L^{wt} and Merlin-UL128L³³⁰¹ displayed similar cell-cell spread properties in RPE-1 cells. Transfer of the substitution identified in TB40-BAC4 UL128L was sufficient to reduce epithelial cell-to-cell spread of Merlin to the same degree (approximately 2-fold) as transfer of the entire TB40-BAC4 UL128L (Fig. 4A). Investigations of infection kinetics (Fig. 4B) and production of cell-free virus in epithelial cell culture (Fig. 4C) also showed that Merlin-UL128L³³⁰¹ had growth characteristics similar to those of Merlin-UL128L^{wt}. Transfer of the G>T substitution to Merlin did not alter the rate of spread through the monolayer, but it did increase the production of cell-free virus by Merlin by the same degree as transfer of the entire TB40-BAC4 UL128L (approximately 50-fold). When plaque sizes were measured in fibroblasts (Fig. 4D), Merlin-UL128L³³⁰¹ again formed plaques of sizes comparable to those of Merlin-UL128L^{wt}, whereas the plaques formed by Merlin-UL128L^{G>T} were similar in size to those formed by Merlin-UL128L^{TB40} and were consistently 2-fold larger than those formed by Merlin-UL128L^{wt} and Merlin-UL128L³³⁰¹.

A single-nucleotide difference in a TB40 UL128 intron reduces splicing efficiency. The proximity of the G>T substitution in TB40-BAC4 UL128 to the splice acceptor site in intron 1 suggested that it had the potential to disrupt splicing of UL128 mRNA. Indeed, bioinformatics analysis predicted that this would be the case (Fig. 5A). Total infected cell RNA was extracted 72 h postinfection, and RT-PCR was performed across both introns of UL128 (Fig. 5B). Three differently spliced products were detected, and DNA sequencing demonstrated that they corresponded to (i) an unspliced transcript (656 to 662 bp, depending on the strain), (ii) a transcript with intron 2 excised (540 bp), and (iii) a transcript with both introns 1 and 2 excised (417 bp). All three UL128 products were detected in all strains tested; however, in the constructs containing the G>T substitution (TB40-BAC4, Merlin-UL128L^{TB40}, and Merlin-UL128L^{G>T}), the cDNAs corresponding to the fully spliced transcript were present at significantly lower abundance.

Thus, the G>T substitution identified in TB40-BAC4 reduced the efficiency of UL128 mRNA splicing and appeared to be entirely responsible for the different growth characteristics conferred by TB40-BAC4 UL128L compared to those of Merlin UL128L.

A single-nucleotide difference in FIX UL130 increases virus production. BLAST searches of UL128L sequences also revealed a unique A-to-G (A>G) substitution in FIX UL130 at nt 177364 (with reference to the Merlin BAC sequence). This difference manifests as a serine-to-proline amino acid change (S72P), and according to protein structure prediction, it disrupts a beta sheet within the protein. The alteration was introduced into Merlin-UL128L^{wt} (generating Merlin-UL130^{A>G}) (Table 1). Merlin-UL130^{A>G} produced plaques in RPE-1 cells approximately 3-fold smaller than those of Merlin-UL128L^{wt} but similar in size to those formed by Merlin-UL128L^{FIX} (Fig. 6A). The effects of this substitution on dissemination during infection (Fig. 6B) and cell-free virus production (Fig. 6C) in epithelial cell culture were also investigated. It reduced the rate of dissemination through the epithelial cell monolayer compared to that of Merlin-UL128L^{wt}, although not as much as transfer of the entire FIX UL128L. It also

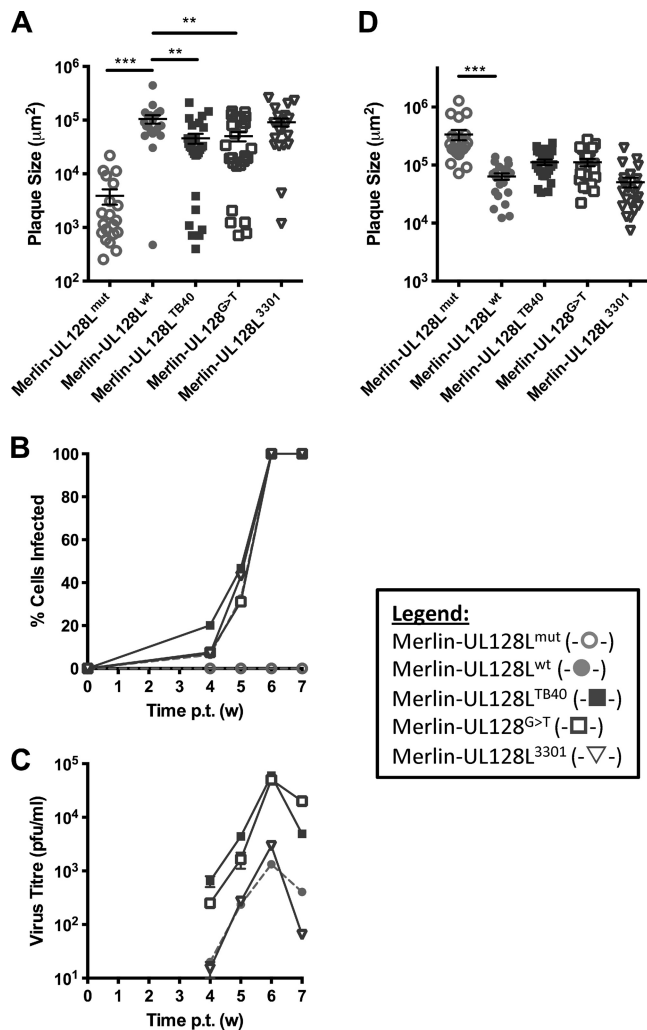


FIG 4 Characteristics of virus containing TB40-BAC4 UL128 G>T unique nucleotide. (A) RPE-1 cells were transfected with BAC DNA for the indicated viruses and then incubated under semisolid overlay. Three weeks later plaque sizes were measured. Means and standard deviations are shown. (B) RPE-1 cells were transfected with BAC DNA for the indicated viruses, and infection was allowed to progress until the monolayer was destroyed. At weekly time points, cells were trypsinized and the level of infection was measured using FACS analysis of EGFP-expressing cells. (C) Also at weekly time points, supernatants were kept and titrated on HFFFs to provide a measure of cell-free virus release. (D) HFFF cells were transfected with BAC DNA for the indicated viruses and then placed under semisolid overlay. Two weeks later, plaque sizes were measured. Means and standard deviations are shown. (Where indicated, samples were compared by 1-way ANOVA followed by Dunnett's posttest to compare each sample to Merlin-UL128^{wt}. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.)

increased the production of cell-free virus approximately 50-fold compared to that of Merlin-UL128L^{wt}. As with the rate of spread through the monolayer, the increase in cell-free release was not as dramatic as that achieved by transfer of the entire FIX UL128L (approximately 150-fold). When cell-to-cell spread ability was tested in HFFFs (Fig. 6D), Merlin-UL130^{A>G} and Merlin-UL128L^{FIX} formed plaques of comparable size, and the plaques were consistently twice as large as those formed by Merlin-UL128L^{wt}.

Thus, the A>G substitution in FIX UL130 contributed signif-

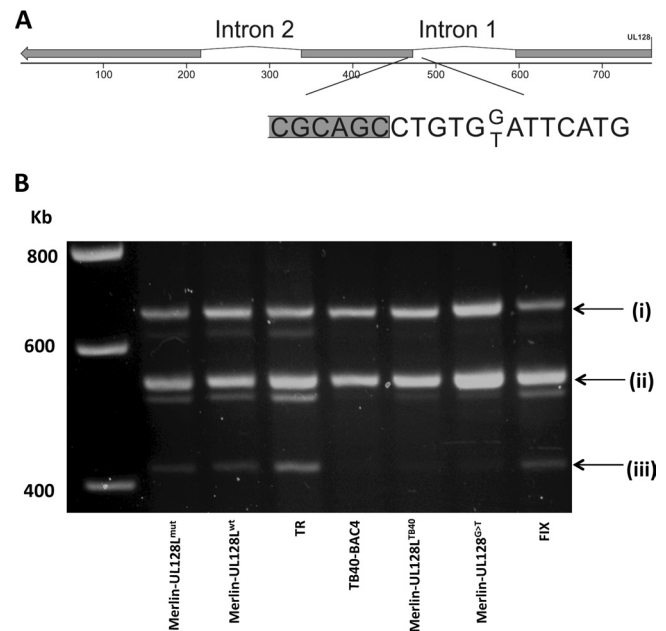


FIG 5 UL128 gene map and transcripts. (A) The position of the unique A residue identified in TB40-BAC4 UL128 intron 1, in place of the C residue present in all other strains, is indicated. Protein-coding exons are shaded. (B) HFFF cells were infected with the indicated virus, and total RNA was extracted at 72 h postinfection. RT-PCR was performed using primers binding in the first and third exons of UL128. The bands correspond to transcripts that are (i) unspliced, (ii) lacking intron 2, or (iii) lacking both intron 1 and intron 2.

icantly to the different growth characteristics of FIX UL128L compared to those of Merlin UL128L.

pUL128 is reduced in TB40-BAC4 and FIX virions. The substitutions in TB40-BAC4 and FIX UL128L identified above have the potential to influence the function of gH/gL/UL128L by various mechanisms. Impairing the efficiency of RNA splicing in TB40-BAC4 would be expected to limit expression of full-length pUL128, and the S72P substitution in pUL130 might reduce the stable incorporation of pUL130 into gH/gL/UL128L or affect a functional domain of the complex. Since all five subunits are required for stable incorporation of gH/gL/UL128L into the virion (28, 66), alteration of the amounts or structural attributes of any one component has the potential to affect the levels of the entire complex incorporated into the virion. To investigate this, levels of pUL128 were analyzed in HCMV virions purified on glycerol-tartrate gradients (Fig. 7A), with sample loading normalized to that of gB. Note that although the epitope recognized by the anti-UL128 antibody used in this assay is not mapped, any differences in detection are unlikely to be due to different antibody affinities for pUL128 from different strains; compared to Merlin, UL128 in the other strains differed by only one (FIX) or two (TB40-BAC4 and TR) amino acids, while in Merlin viruses containing the single-nucleotide substitutions from FIX or TB40-BAC4, pUL128 is identical to Merlin.

pUL128 was undetectable in Merlin-UL128L^{mut} but was readily detected in viruses having wild-type UL128L sequences (Merlin-UL128L^{wt} and Merlin-UL128L³³⁰¹). Consistent with TR UL128L having a similar influence on growth kinetics compared to Merlin UL128L, TR contained levels of pUL128 that were comparable to those in Merlin-UL128L^{wt} and Merlin-UL128L³³⁰¹.

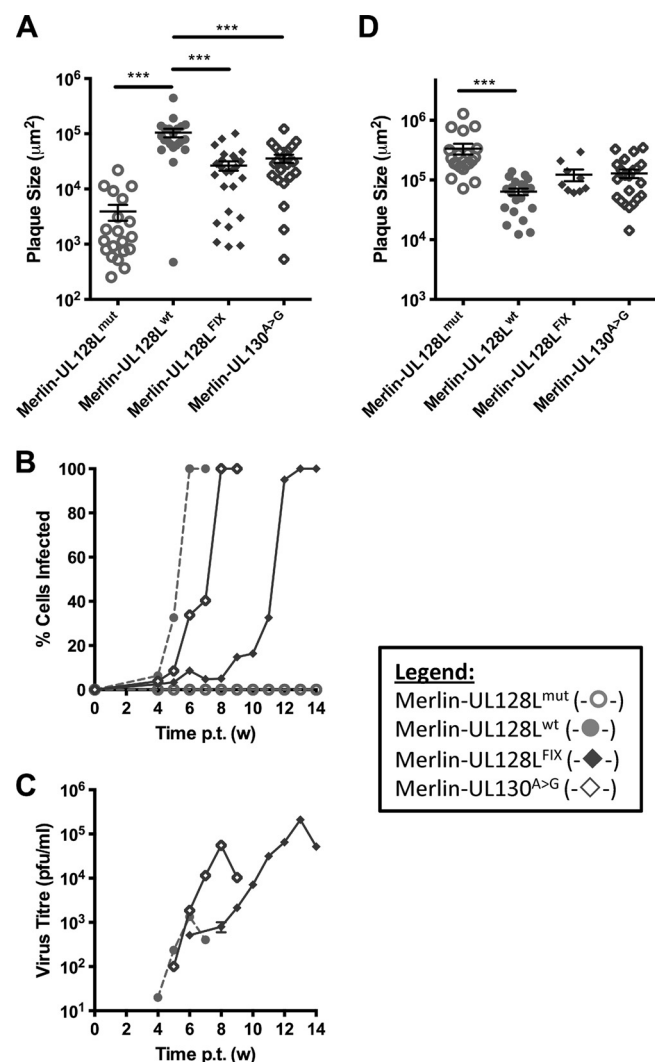


FIG 6 Growth characteristics of virus containing the FIX UL130 A>G substitution. (A) RPE-1 cells were transfected with BAC DNA for the indicated viruses and incubated under semisolid overlay. Plaque sizes were measured 3 weeks later. Means and standard deviations are shown. (B) RPE-1 cells were transfected with BAC DNA for the indicated viruses, and infection was allowed to progress until the monolayer was destroyed. At weekly time points, cells were trypsinized and the level of infection was measured by FACS analysis of EGFP-expressing cells. (C) Supernatants from the infections shown in panel B were retained at weekly intervals and titrated on HFFFs to provide a measure of cell-free virus release. (D) HFFF cells were transfected with BAC DNA for the indicated viruses and then placed under semisolid overlay. Plaque sizes were measured 2 weeks later. Means and standard deviations are shown. (Where indicated, samples were compared by 1-way ANOVA followed by Dunnett's posttest to compare each sample to Merlin-UL128^{wt}. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.)

However, TB40-BAC4 contained reduced levels of pUL128, as did Merlin containing either the entire TB40-BAC4 UL128L region (Merlin-UL128L^{TB40}) or only the G>T substitution (Merlin-UL128L^{G>T}). pUL128 was undetectable in FIX and Merlin-UL128L^{FIX}, while Merlin-UL130^{A>G} contained reduced, though detectable, levels of pUL128.

Based on the spread of these viruses in epithelial cells, it seemed likely that all except Merlin-UL128L^{mut} contained gH/gL/UL128L in the virion, and that lack of detection of pUL128 in some was due

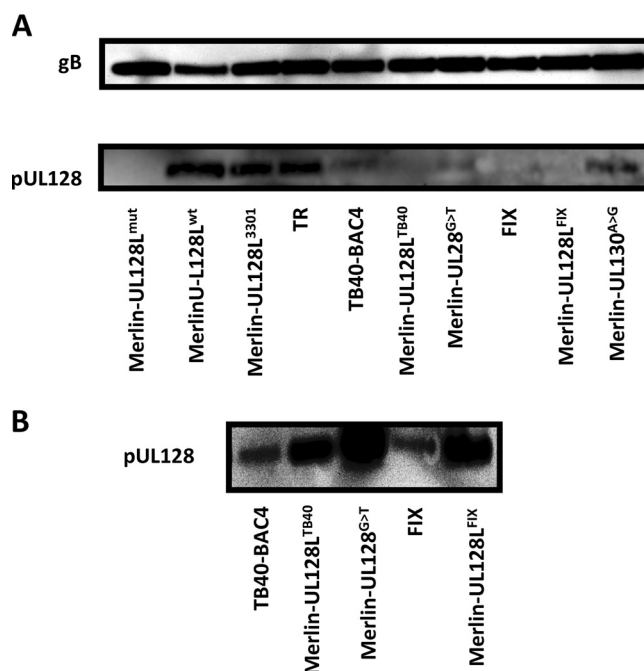


FIG 7 Virion pUL128 content of BAC-cloned strains and recombinant Merlin viruses. (A) Stocks of the indicated viruses were grown, and infectious virions were purified on glycerol-tartrate gradients before being analyzed by Western blotting. Loading was normalized to gB levels (upper panel) and stained for pUL128 (lower panel). (B) The same samples as those used for panel A were analyzed by Western blotting. However, rather than being normalized to gB content, virion loads were maximized to facilitate detection of pUL128 in each strain.

to the small amounts of virus loaded in order to keep sample loading comparable to that of the low-titer viruses Merlin-UL128L^{wt} and Merlin-UL128L³³⁰¹. In a separate blot where the virion load was not normalized, pUL128L was detectable in all virions except Merlin-UL128L^{mut} (Fig. 7B).

Thus, all virions except Merlin-UL128L^{mut} contain pUL128; however, the substitutions identified in UL128L of TB40-BAC4 and FIX resulted in reduced levels of gH/gL/UL128L being incorporated into the virion.

TB40-BAC4 and FIX UL128L bestow impaired epithelial cell tropism. Since gH/gL/UL128L is required for efficient infection of epithelial cells, its reduced incorporation into virions would be expected to restrict epithelial cell tropism. Viruses derived from each BAC-cloned strain were plaque titrated on HFFF and RPE-1 cells in parallel (Fig. 8A). For each virus, the titer reported in RPE-1 cells was normalized to that reported for HFFF cells for the same virus, thereby quantitating the ability of each virus to infect epithelial cells relative to its ability to infect fibroblasts. Merlin-UL128L^{mut} infected RPE-1 cells approximately 100-fold less efficiently than HFFFs; however, the majority of plaques were represented by single cells, indicating that cell-to-cell spread was strongly inhibited. Viruses containing wild-type UL128L (Merlin-UL128L^{wt} and Merlin-UL128L³³⁰¹) infected RPE-1 cells with comparable or slightly greater efficiency than they did HFFFs. Like Merlin-UL128L^{wt} and Merlin-UL128L³³⁰¹, TR infected RPE-1 and HFFF cells with similar efficiencies, consistent with these viruses containing comparable levels of gH/gL/UL128L in virions. However, infection of RPE-1 cells by TB40-BAC4 was approxi-

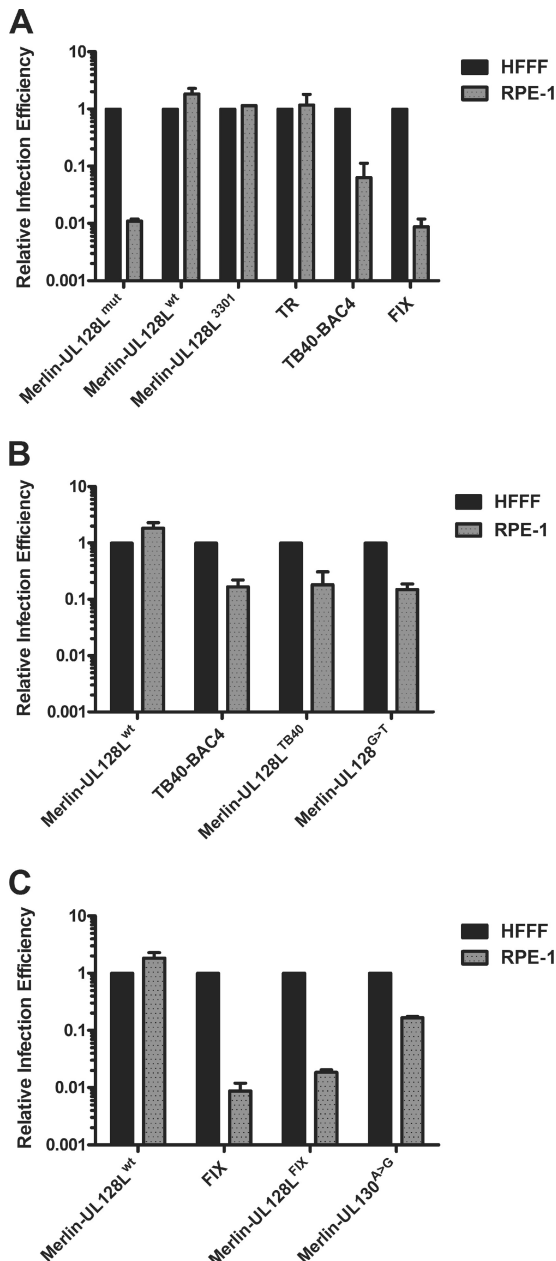


FIG 8 Relative infection efficiency of fibroblast and epithelial cells. Stocks of the indicated viruses were simultaneously titrated onto HFFF or RPE-1 cells. Titters for each virus on RPE-1 cells were then normalized to the titer of the same virus on HFFF cells, which was arbitrarily set to 1, giving the relative efficiency of infection of epithelial cells for each strain. The data are based on two independent experiments. (A) Relative infection efficiencies of TR, TB40-BAC4, and FIX compared to various Merlin viruses. (B) Relative infection efficiencies of Merlin containing either the entire UL128L or the substitution from TB40-BAC4 compared to TB40-BAC4 and Merlin. (C) Relative infection efficiencies of Merlin containing either the entire UL128L or the single nucleotide from FIX compared to FIX and Merlin.

mately 10-fold less efficient than infection of HFFFs, whereas the efficiency of infection of RPE-1 cells by FIX was approximately 100-fold less than that of HFFFs and comparable to a virus lacking gH/gL/UL128L (Merlin-UL128L^{mut}).

Like TB40-BAC4, Merlin-UL128L^{TB40} and Merlin-UL128L^{G>T}

infected RPE-1 cells approximately 10-fold less efficiently than HFFFs (Fig. 8B), displaying a lower infection efficiency than Merlin-UL128L^{wt}, Merlin-UL128L³³⁰¹, and TR.

FIX infected epithelial cells with approximately 100-fold less efficiency than fibroblasts, and this was closely recapitulated by Merlin containing FIX UL128L (Fig. 8C). The A>G substitution in FIX UL130 also reduced the ability of Merlin to infect RPE-1 cells, although only by approximately 10-fold.

In summary, the unique nucleotide differences identified in TB40-BAC4 UL128L and FIX UL130 resulted in a reduction in the amounts of gH/gL/UL128L incorporated into the virion and a concomitant reduction in the relative ability of virus to infect epithelial cells.

DISCUSSION

Previous attempts to propagate wild-type HCMV from clinical material by passage *in vitro* in fibroblasts have shown that mutations invariably occur in the viral genome. RL13 mutates first, followed by UL128L, and following these steps virus grows to much higher cell-free titers (37, 38). However, the adapted virus lacks gH/gL/UL128L in the virion envelope; therefore, it does not efficiently infect cells other than fibroblasts. The single-nucleotide substitutions identified in UL128L of TB40-BAC4 and FIX may explain the apparently conflicting fact that these strains retain the ability to infect endothelial and epithelial cells when passaged on fibroblasts. gH/gL/UL128L is expressed and is present in virions of these strains, but the level is reduced. This permits greater release of cell-free virus and potentially reduces the selective pressure for further mutations in UL128L.

It is important to determine whether these substitutions represent natural strain variation or mutations acquired *in vitro*. The clinical material from which the strains were derived is not available; therefore, it is not possible to answer this question directly. However, the S72P substitution in FIX UL130 is absent from the passaged parental strain VR1814 (37); thus, it probably occurred during further passage of VR1814 prior to BAC cloning. Similarly, the fact that the G>T substitution in intron 1 of TB40-BAC4 is unique to TB40-BAC4 strongly suggests that it was acquired during passage in fibroblasts prior to BAC cloning. This mutation is also absent from the consensus sequence of the parental strain (TB40/E), which instead contains a unique C207S variation in UL130. TB40/E is known to comprise a mixture of genomes (48), and the G>T mutation may be present at low levels in this mixture. Nevertheless, the identification of the substitutions in FIX and TB40-BAC4 demonstrates the importance of comparing viral genomes to unpassaged virus when determining whether a gene is wild type in sequence and whether (by extension) the phenotype of a virus can be assumed to be the same as the clinical virus.

The effect of these substitutions was a reduction in levels of gH/gL/UL128L in the virion. This manifested in impaired efficiency of epithelial cell entry, an increase in cell-to-cell spread in fibroblasts but a decrease in cell-to-cell spread in epithelial cells, and an increase in cell-free release in both cell types. Taken together, these observations indicate that the level of gH/gL/UL128L in the virion is directly related to its ability to inhibit replication in fibroblasts and to promote entry and cell-to-cell spread in epithelial cells. Surprisingly, it also suggests that wild-type UL128L inhibits the production of cell-free virus in both epithelial and fibroblast cells. The mechanism by which this occurs is unclear but may be related to the fact that gH and gL are common to two glycopro-

tein complexes (gH/gL/UL128L and gH/gL/gO). Greater amounts of gH/gL/UL128L may result in there being less gH/gL available for the formation of gH/gL/gO. gH/gL/gO may be important for secondary envelopment of progeny virions and egress into the supernatant (23) and for subsequent infection by cell-free released virus (24, 67), explaining the increased titers of cell-free virus as the level of gH/gL/UL128L is reduced. This model is supported by the observation that expression of wild-type UL128L reduces levels of both cell-free and cell-associated virus (37), and that loss of gO results in greater accumulation of gH/gL/UL128L in the virion with a concomitant increase in cell-to-cell spread in epithelial and endothelial cells (24).

Although the mutations identified in FIX and TB40-BAC4 strongly contributed to the greater production of cell-free virus and differences in cell-to-cell spread, it is clear that they are not the sole reason for the growth differences seen in TB40-BAC4 and FIX compared to Merlin-UL128L^{wt}. TB40-BAC4 grew to approximately 1,000-fold higher cell-free titers than Merlin-UL128L^{wt}, whereas transfer of the G>T mutation in TB40-BAC4 UL128 resulted in an increase of only 50-fold. Likewise, FIX was unable to spread cell to cell in RPE-1 epithelial cells, yet Merlin-UL128L^{FIX} was able to spread, albeit more slowly than Merlin-UL130^{A>G}, which itself spread more slowly than Merlin-UL128L^{wt}. Similar considerations apply to TR. Despite containing levels of gH/gL/UL128L similar to those in Merlin-UL128L^{wt}, TR generated higher cell-free titers yet displayed lower efficiency of cell-to-cell spread in both fibroblast and epithelial cells. The observation that Merlin-UL128L^{TR} behaved the same as Merlin-UL128L^{wt} and Merlin-UL128³³⁰¹, and that TR UL128L readily mutated when expressed in the Merlin genome during growth in fibroblasts but not when expressed within the TR genome, underlines the conclusion that the differential growth characteristics of TR were independent of UL128L yet are sufficient to enhance the stability of UL128L *in vitro*. These differences may be due to a number of other genome regions that can influence growth in a cell type-specific manner (37, 68–70). Sequence differences in other glycoproteins might also be responsible, with gO being a particularly prominent candidate because it exists in several highly divergent, yet stable, genotypes (43, 71–73). Alignments of the protein-coding regions of gH, gL, gM, gN, gO, and gB of TR, TB40-BAC4, and FIX with those of other strains did not reveal any unique substitutions that could represent *in vitro* adaptations, although all 3 viruses lack sequences in Us, where the BAC cassette was inserted, and TB40-BAC4 contains mutations in RL5A, RL6, UL141, and UL40 (48, 74), and TR contains a mutation in UL97 (C607Y) that conveys ganciclovir resistance (75). These mutations could also contribute to the growth characteristics observed.

Owing to its central role in determining tropism for a broad range of clinically significant cell types, it is essential that an intact, wild-type gH/gL/UL128L is present in any strain used to investigate HCMV pathogenesis. More recently, other important implications of UL128L function have underscored the need for research based on virus strains that are competent in this genome region. Specifically, gH/gL/UL128L elicits a potent neutralizing antibody response in a natural infection (76–81) and is considered ideal for inclusion in new vaccine strategies (34, 82). The differences identified between levels of gH/gL/UL128L in various strains analyzed have the potential to affect read-outs of the efficacy of anti-gH/gL/UL128L antibodies. However, they also offer significant advantages relating to the important issue of being able

to grow HCMV while minimizing the risk of mutation. For work performed in fibroblast cells, viruses based on Merlin-UL128L^{mut} offer an ideal defined, full-length genome that produces relatively high titers. For work requiring infection of other cell types, Merlin-UL128^{G>T} produces larger amounts of cell-free virus that is still able to infect epithelial and endothelial cells, potentially with reduced risk of mutation in UL128L.

ACKNOWLEDGMENTS

This work was funded by the National Institute for Social Care and Health (NISCHR), the Medical Research Council (MRC), and the Wellcome Trust.

We thank Gavin Wilkinson for critical reading of the manuscript.

REFERENCES

1. Mocarski ES. 2007. Cytomegalovirus, p 2447–2492. In Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA, Roizman B, Straus SE (ed), Fields virology, 5th ed. Lippincott Williams & Wilkins, Philadelphia, PA.
2. Adler B, Sinzger C. 2009. Endothelial cells in human cytomegalovirus infection: one host cell out of many or a crucial target for virus spread? *Thromb. Haemost.* 102:1057–1063.
3. Sinclair J, Sissons P. 2006. Latency and reactivation of human cytomegalovirus. *J. Gen. Virol.* 87:1763–1779.
4. Compton T, Nowlin DM, Cooper NR. 1993. Initiation of human cytomegalovirus infection requires initial interaction with cell surface heparan sulfate. *Virology* 193:834–841.
5. Boyle KA, Compton T. 1998. Receptor-binding properties of a soluble form of human cytomegalovirus glycoprotein B. *J. Virol.* 72:1826–1833.
6. Britt WJ, Boppana S. 2004. Human cytomegalovirus virion proteins. *Hum. Immunol.* 65:395–402.
7. Kari B, Gehrz R. 1992. A human cytomegalovirus glycoprotein complex designated gC-II is a major heparin-binding component of the envelope. *J. Virol.* 66:1761–1764.
8. Feire AL, Koss H, Compton T. 2004. Cellular integrins function as entry receptors for human cytomegalovirus via a highly conserved disintegrin-like domain. *Proc. Natl. Acad. Sci. U. S. A.* 101:15470–15475.
9. Wang X, Huang SM, Chiu ML, Raab-Traub N, Huang ES. 2003. Epidermal growth factor receptor is a cellular receptor for human cytomegalovirus. *Nature* 424:456–461.
10. Soroceanu L, Akhavan A, Cobbs CS. 2008. Platelet-derived growth factor- α receptor activation is required for human cytomegalovirus infection. *Nature* 455:391–395.
11. Chan G, Nogalski MT, Yurochko AD. 2009. Activation of EGFR on monocytes is required for human cytomegalovirus entry and mediates cellular motility. *Proc. Natl. Acad. Sci. U. S. A.* 106:22369–22374.
12. Isaacson MK, Feire AL, Compton T. 2007. Epidermal growth factor receptor is not required for human cytomegalovirus entry or signaling. *J. Virol.* 81:6241–6247.
13. Vanarsdall AL, Wisner TW, Lei H, Kazlauskas A, Johnson DC. 2012. PDGF receptor- α does not promote HCMV entry into epithelial and endothelial cells but increased quantities stimulate entry by an abnormal pathway. *PLoS Pathog.* 8:e1002905. doi:10.1371/journal.ppat.1002905.
14. Kinzler ER, Compton T. 2005. Characterization of human cytomegalovirus glycoprotein-induced cell-cell fusion. *J. Virol.* 79:7827–7837.
15. Chowdary TK, Cairns TM, Atanasiu D, Cohen GH, Eisenberg RJ, Heldwein EE. 2010. Crystal structure of the conserved herpesvirus fusion regulator complex gH-gL. *Nat. Struct. Mol. Biol.* 17:882–888.
16. Isaacson MK, Compton T. 2009. Human cytomegalovirus glycoprotein B is required for virus entry and cell-to-cell spread but not for virion attachment, assembly, or egress. *J. Virol.* 83:3891–3903.
17. Lopper M, Compton T. 2004. Coiled-coil domains in glycoproteins B and H are involved in human cytomegalovirus membrane fusion. *J. Virol.* 78:8333–8341.
18. Vanarsdall AL, Ryckman BJ, Chase MC, Johnson DC. 2008. Human cytomegalovirus glycoproteins gB and gH/gL mediate epithelial cell-cell fusion when expressed either in cis or in trans. *J. Virol.* 82:11837–11850.
19. Compton T, Nepomuceno RR, Nowlin DM. 1992. Human cytomegalovirus penetrates host cells by pH-independent fusion at the cell surface. *Virology* 191:387–395.
20. Haspot F, Lavault A, Sinzger C, Laib Sampaio K, Stierhof YD, Pilet P,

- Bressollette-Bodin C, Halary F. 2012. Human cytomegalovirus entry into dendritic cells occurs via a macropinocytosis-like pathway in a pH-independent and cholesterol-dependent manner. *PLoS One* 7:e34795. doi:10.1371/journal.pone.0034795.
21. Sinzger C. 2008. Entry route of HCMV into endothelial cells. *J. Clin. Virol.* 41:174–179.
 22. Ryckman BJ, Jarvis MA, Drummond DD, Nelson JA, Johnson DC. 2006. Human cytomegalovirus entry into epithelial and endothelial cells depends on genes UL128 to UL150 and occurs by endocytosis and low-pH fusion. *J. Virol.* 80:710–722.
 23. Jiang XJ, Adler B, Sampaio KL, Digel M, Jahn G, Ettischer N, Stierhof YD, Scrivano L, Koszinowski U, Mach M, Sinzger C. 2008. UL74 of human cytomegalovirus contributes to virus release by promoting secondary envelopment of virions. *J. Virol.* 82:2802–2812.
 24. Wille PT, Knoche AJ, Nelson JA, Jarvis MA, Johnson DC. 2009. An HCMV gO-null mutant fails to incorporate gH/gL into the virion envelope and is unable to enter fibroblasts, epithelial, and endothelial cells. *J. Virol.* 84:2585–2596.
 25. Adler B, Scrivano L, Ruzcics Z, Rupp B, Sinzger C, Koszinowski U. 2006. Role of human cytomegalovirus UL131A in cell type-specific virus entry and release. *J. Gen. Virol.* 87:2451–2460.
 26. Hahn G, Revello MG, Patrone M, Percivalle E, Campanini G, Sarasini A, Wagner M, Gallina A, Milanese G, Koszinowski U, Baldanti F, Gerna G. 2004. Human cytomegalovirus UL131-128 genes are indispensable for virus growth in endothelial cells and virus transfer to leukocytes. *J. Virol.* 78:10023–10033.
 27. Ryckman BJ, Chase MC, Johnson DC. 2008. HCMV gH/gL/UL128-131 interferes with virus entry into epithelial cells: evidence for cell type-specific receptors. *Proc. Natl. Acad. Sci. U. S. A.* 105:14118–14123.
 28. Ryckman BJ, Rainish BL, Chase MC, Borton JA, Nelson JA, Jarvis MA, Johnson DC. 2008. Characterization of the human cytomegalovirus gH/gL/UL128-131 complex that mediates entry into epithelial and endothelial cells. *J. Virol.* 82:60–70.
 29. Schuessler A, Sampaio KL, Scrivano L, Sinzger C. 2010. Mutational mapping of UL130 of human cytomegalovirus defines peptide motifs within the C-terminal third as essential for endothelial cell infection. *J. Virol.* 84:9019–9026.
 30. Schuessler A, Sampaio KL, Sinzger C. 2008. Charge cluster-to-alanine scanning of UL128 for fine tuning of the endothelial cell tropism of human cytomegalovirus. *J. Virol.* 82:11239–11246.
 31. Schuessler A, Sampaio KL, Straszewski S, Sinzger C. 2012. Mutational mapping of pUL131A of human cytomegalovirus emphasizes its central role for endothelial cell tropism. *J. Virol.* 86:504–512.
 32. Straszewski S, Patrone M, Walther P, Gallina A, Mertens T, Frascaroli G. 2011. Protein pUL128 of human cytomegalovirus is necessary for monocyte infection and blocking of migration. *J. Virol.* 85:5150–5158.
 33. Wang D, Shenk T. 2005. Human cytomegalovirus UL131 open reading frame is required for epithelial cell tropism. *J. Virol.* 79:10330–10338.
 34. Revello MG, Gerna G. 2010. Human cytomegalovirus tropism for endothelial/epithelial cells: scientific background and clinical implications. *Rev. Med. Virol.* 20:136–155.
 35. Gerna G, Percivalle E, Lilleri D, Lozza L, Fornara C, Hahn G, Baldanti F, Revello MG. 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.* 86:275–284.
 36. Sinzger C, Kahl M, Laib K, Klingel K, Rieger P, Plachter B, Jahn G. 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.* 81:3021–3035.
 37. Dargan DJ, Douglas E, Cunningham C, Jamieson F, Stanton RJ, Baluchova K, McSharry BP, Tomasec P, Emery VC, Percivalle E, Sarasini A, Gerna G, Wilkinson GW, Davison AJ. 2010. Sequential mutations associated with adaptation of human cytomegalovirus to growth in cell culture. *J. Gen. Virol.* 91:1535–1546.
 38. Stanton RJ, Baluchova K, Dargan DJ, Cunningham C, Sheehy O, Seirafian S, McSharry BP, Neale ML, Davies JA, Tomasec P, Davison AJ, Wilkinson GW. 2010. Reconstruction of the complete human cytomegalovirus genome in a BAC reveals RL13 to be a potent inhibitor of replication. *J. Clin. Invest.* 120:3191–3208.
 39. Akter P, Cunningham C, McSharry BP, Dolan A, Addison C, Dargan DJ, Hassan-Walker AF, Emery VC, Griffiths PD, Wilkinson GW, Davison AJ. 2003. Two novel spliced genes in human cytomegalovirus. *J. Gen. Virol.* 84:1117–1122.
 40. Cunningham C, Gatherer D, Hilfrich B, Baluchova K, Dargan DJ, Thomson M, Griffiths PD, Wilkinson GW, Schulz TF, Davison AJ. 2010. Sequences of complete human cytomegalovirus genomes from infected cell cultures and clinical specimens. *J. Gen. Virol.* 91:605–615.
 41. Sinzger C, Schmidt K, Knapp J, Kahl M, Beck R, Waldman J, Hebart H, Einsele H, Jahn G. 1999. Modification of human cytomegalovirus tropism through propagation in vitro is associated with changes in the viral genome. *J. Gen. Virol.* 80:2867–2877.
 42. Cha T, Tom E, Kemble GW, Duke GM, Mocarski ES, Spaete RR. 1996. Human cytomegalovirus clinical isolates carry at least 19 genes not found in laboratory strains. *J. Virol.* 70:78–83.
 43. Dolan A, Cunningham C, Hector RD, Hassan-Walker AF, Lee L, Addison C, Dargan DJ, McGeoch DJ, Gatherer D, Emery VC, Griffiths PD, Sinzger C, McSharry BP, Wilkinson GW, Davison AJ. 2004. Genetic content of wild-type human cytomegalovirus. *J. Gen. Virol.* 85:1301–1312.
 44. Prichard M, Penfold M, Duke G, Spaete R, Kemble G. 2001. A review of genetic differences between limited and extensively passaged human cytomegalovirus strains. *Rev. Med. Virol.* 11:191–200.
 45. Borst EM, Hahn G, Koszinowski UH, Messerle M. 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.* 73:8320–8329.
 46. Paredes AM, Yu D. 2012. Human cytomegalovirus: bacterial artificial chromosome (BAC) cloning and genetic manipulation. *Curr. Protoc. Microbiol.* Chapter 14:Unit14E.14. doi:10.1002/9780471729259.mc14e04s24.
 47. Yu D, Smith GA, Enquist LW, Shenk T. 2002. Construction of a self-excisable bacterial artificial chromosome containing the human cytomegalovirus genome and mutagenesis of the diploid TRL/IRL13 gene. *J. Virol.* 76:2316–2328.
 48. Sinzger C, Hahn G, Digel M, Katona R, Sampaio KL, Messerle M, Hengel H, Koszinowski U, Brune W, Adler B. 2008. Cloning and sequencing of a highly productive, endotheliotropic virus strain derived from human cytomegalovirus TB40/E. *J. Gen. Virol.* 89:359–368.
 49. Grazia Revello M, Baldanti F, Percivalle E, Sarasini A, De-Giuli L, Genini E, Lilleri D, Labò N, Gerna G. 2001. In vitro selection of human cytomegalovirus variants unable to transfer virus and virus products from infected cells to polymorphonuclear leukocytes and to grow in endothelial cells. *J. Gen. Virol.* 82:1429–1438.
 50. Hahn G, Khan H, Baldanti F, Koszinowski UH, Revello MG, Gerna G. 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.* 76:9551–9555.
 51. Murphy E, Yu D, Grimwood J, Schmutz J, Dickson M, Jarvis MA, Hahn G, Nelson JA, Myers RM, Shenk TE. 2003. Coding potential of laboratory and clinical strains of human cytomegalovirus. *Proc. Natl. Acad. Sci. U. S. A.* 100:14976–14981.
 52. Smith IL, Taskintuna I, Rahhal FM, Powell HC, Ai E, Mueller AJ, Spector SA, Freeman WR. 1998. Clinical failure of CMV retinitis with intravitreal cidofovir is associated with antiviral resistance. *Arch. Ophthalmol.* 116:178–185.
 53. Stanton RJ, McSharry BP, Rickards CR, Wang EC, Tomasec P, Wilkinson GW. 2007. Cytomegalovirus destruction of focal adhesions revealed in a high-throughput Western blot analysis of cellular protein expression. *J. Virol.* 81:7860–7872.
 54. Matrosovich M, Matrosovich T, Garten W, Klenk HD. 2006. New low-viscosity overlay medium for viral plaque assays. *Virol. J.* 3:63.
 55. Warming S, Costantino N, Court DL, Jenkins NA, Copeland NG. 2005. Simple and highly efficient BAC recombining using galK selection. *Nucleic Acids Res.* 33:e36. doi:10.1093/nar/gni035.
 56. Stanton RJ, McSharry BP, Armstrong M, Tomasec P, Wilkinson GW. 2008. Re-engineering adenovirus vector systems to enable high-throughput analyses of gene function. *Biotechniques* 45:659–664.
 57. Reese MG, Eeckman FH, Kulp D, Haussler D. 1997. Improved splice site detection in Genie. *J. Comput. Biol.* 4:311–323.
 58. Brunak S, Engelbrecht J, Knudsen S. 1991. Prediction of human mRNA donor and acceptor sites from the DNA sequence. *J. Mol. Biol.* 220:49–65.
 59. Dogan RI, Getoor L, Wilbur WJ, Mount SM. 2007. SplicePort—an interactive splice-site analysis tool. *Nucleic Acids Res.* 35:W285–W291.
 60. Desmet FO, Hamroun D, Lalande M, Collod-Beroud G, Claustres M,

- Beroud C. 2009. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* 37:e67. doi:10.1093/nar/gkp215.
61. Kelley LA, Sternberg MJ. 2009. Protein structure prediction on the Web: a case study using the Phyre server. *Nat. Protoc.* 4:363–371.
 62. Talbot P, Almeida JD. 1977. Human cytomegalovirus: purification of enveloped virions and dense bodies. *J. Gen. Virol.* 36:345–349.
 63. Irmieri A, Gibson W. 1983. Isolation and characterization of a noninfectious virion-like particle released from cells infected with human strains of cytomegalovirus. *Virology* 130:118–133.
 64. Baldanti F, Paolucci S, Campanini G, Sarasini A, Percivalle E, Revello MG, Gerna G. 2006. Human cytomegalovirus UL131A, UL130 and UL128 genes are highly conserved among field isolates. *Arch. Virol.* 151:1225–1233.
 65. Sun ZR, Ji YH, Ruan Q, He R, Ma YP, Qi Y, Mao ZQ, Huang YJ. 2009. Structure characterization of human cytomegalovirus UL131A, UL130 and UL128 genes in clinical strains in China. *Genet. Mol. Res.* 8:1191–1201.
 66. Wang D, Shenk T. 2005. Human cytomegalovirus virion protein complex required for epithelial and endothelial cell tropism. *Proc. Natl. Acad. Sci. U. S. A.* 102:18153–18158.
 67. Scrivano L, Esterlechner J, Muhlbach H, Ettischer N, Hagen C, Grunewald K, Mohr CA, Ruzsics Z, Koszinowski U, Adler B. 2010. The m74 gene product of murine cytomegalovirus (MCMV) is a functional homolog of human CMV gO and determines the entry pathway of MCMV. *J. Virol.* 84:4469–4480.
 68. Bronzini M, Lugini A, Dell'Oste V, De Andrea M, Landolfo S, Gribo G. 2012. The US16 gene of human cytomegalovirus is required for efficient viral infection of endothelial and epithelial cells. *J. Virol.* 86:6875–6888.
 69. Davison AJ, Akter P, Cunningham C, Dolan A, Addison C, Dargan DJ, Hassan-Walker AF, Emery VC, Griffiths PD, Wilkinson GWG. 2003. Homology between the human cytomegalovirus RL11 gene family and human adenovirus E3 genes. *J. Gen. Virol.* 84:657–663.
 70. O'Connor CM, Shenk T. 2011. Human cytomegalovirus pUS27 G protein-coupled receptor homologue is required for efficient spread by the extracellular route but not for direct cell-to-cell spread. *J. Virol.* 85:3700–3707.
 71. Paterson DA, Dyer AP, Milne RS, Sevilla-Reyes E, Gompels UA. 2002. A role for human cytomegalovirus glycoprotein O (gO) in cell fusion and a new hypervariable locus. *Virology* 293:281–294.
 72. Rasmussen L, Geissler A, Cowan C, Chase A, Winters M. 2002. The genes encoding the gCIII complex of human cytomegalovirus exist in highly diverse combinations in clinical isolates. *J. Virol.* 76:10841–10848.
 73. Stanton R, Westmoreland D, Fox JD, Davison AJ, Wilkinson GW. 2005. Stability of human cytomegalovirus genotypes in persistently infected renal transplant recipients. *J. Med. Virol.* 75:42–46.
 74. Magri G, Muntasell A, Romo N, Saez-Borderias A, Pende D, Geraghty DE, Hengel H, Angulo A, Moretta A, Lopez-Botet M. 2011. NKp46 and DNAM-1 NK-cell receptors drive the response to human cytomegalovirus-infected myeloid dendritic cells overcoming viral immune evasion strategies. *Blood* 117:848–856.
 75. Erice A. 1999. Resistance of human cytomegalovirus to antiviral drugs. *Clin. Microbiol. Rev.* 12:286–297.
 76. Fouts AE, Chan P, Stephan JP, Vandlen R, Feierbach B. 2012. Antibodies against the gH/gL/UL128/UL130/UL131 complex comprise the majority of the anti-CMV neutralizing antibody response in CMV-HIG. *J. Virol.* 86:7444–7447.
 77. Genini E, Percivalle E, Sarasini A, Revello MG, Baldanti F, Gerna G. 2011. Serum antibody response to the gH/gL/pUL128-131 five-protein complex of human cytomegalovirus (HCMV) in primary and reactivated HCMV infections. *J. Clin. Virol.* 52:113–118.
 78. Gerna G, Sarasini A, Patrone M, Percivalle E, Fiorina L, Campanini G, Gallina A, Baldanti F, Revello MG. 2008. Human cytomegalovirus serum neutralizing antibodies block virus infection of endothelial/epithelial cells, but not fibroblasts, early during primary infection. *J. Gen. Virol.* 89:853–865.
 79. Lilleri D, Kabanova A, Lanzavecchia A, Gerna G. 2012. Antibodies against neutralization epitopes of human cytomegalovirus gH/gL/pUL128-130-131 complex and virus spreading may correlate with virus control in vivo. *J. Clin. Immunol.* 32:1324–1331.
 80. Saccoccio FM, Sauer AL, Cui X, Armstrong AE, el Habib SE, Johnson DC, Ryckman BJ, Klingelutz AJ, Adler SP, McVoy MA. 2011. Peptides from cytomegalovirus UL130 and UL131 proteins induce high titer antibodies that block viral entry into mucosal epithelial cells. *Vaccine* 29:2705–2711.
 81. Macagno A, Bernasconi NL, Vanzetta F, Dander E, Sarasini A, Revello MG, Gerna G, Sallusto F, Lanzavecchia A. 2010. Isolation of human monoclonal antibodies that potently neutralize human cytomegalovirus infection by targeting different epitopes on the gH/gL/UL128-131A complex. *J. Virol.* 84:1005–1013.
 82. Schleiss M. 2010. Cytomegalovirus vaccines in the pipeline. *Drugs Future* 35:999.